FEDERATED LEARNING-BASED PRIVACY PROTECTION METHODS FOR INTERNET OF THINGS SYSTEMS

Mahmuda Akter and Nour Moustafa

The School of Systems and Computing, University of New South Wales
Canberra, ACT 2612, Australia

ABSTRACT

The Internet of Things (IoT) forms intelligent systems, such as smart cities and factories, to enhance productivity and provide revolutionary and automated services to end-users and organisations. An IoT ecosystem requires more dynamics and heterogeneity with advanced privacy preservation. Federated Learning (FL)addresses the challenge of maintaining data privacy using a privacy-preserving sharing mechanism instead of transmitting raw data. However, the latest cyber threats cause privacy and security breaches. This study systematically analyses federated learning-based privacy-preserving methods in IoT systems. A standard IoT architecture with possible privacy threats is illustrated. Also, Federated Learning schemes and their taxonomies are discussed in a privacy-preserving manner, with initial experiments proving the significance of FL-based privacy preservation in IoT environments. This finds acceptable noise addition in differential privacy by keeping higher testing accuracy in different settings to enhance privacy preservation of federated learning. Various Federated Learning schemes, challenges and future research directions are covered.

KEYWORDS

Federated Learning, Privacy-Preserving, Internet of Things (IoT),Privacy Threats

1. INTRODUCTION

The Internet of Things (IoT) represents an embedded giant network interconnected with physical objects and people, with its devices exchanging generated data. IoT is heavily integrated across many aspects of modern life, with approximately 7 billion IoT devices connected worldwide [1]. These devices have many uses, including as the foundation of smart cities, factories being integrated with smart manufacturing, the basis of smart healthcare, and increasing efficiency and safety in future airports. Across these application areas, Artificial Intelligence introduces a unique deployment for standardising and adding value to IoT.

Over 250 million vehicles become globally connected; moreover, according to Telefonica investigation, various vehicles 90% increase from the year 2013 to the year 2020 those are connected to the Internet [2]. Similarly, the International Data Corporation (IDC) will expand IoT around the US $1.7 trillion by 2024 [3]. However, i-SCOOP reported that global IoT security spending will amount to $3.1 billion by 2021. By 2025, the world's IoT devices are estimated to number more than 64 billion, and cost savings are the primary source of revenue for 54% of company IoT projects. Also, in every Gartner's IT Hype cycle [4], predicted that IoT might hold market assumptions for 5-10 years.

David C. Wyld et al. (Eds): SPTM, AIS, SOENG, AMLA, NLPA, IPPR, CSIT – 2024
pp. 31-44, 2024. CS & IT - CSCP 2024 DOI: 10.5121/csit.2024.141103

The number of connected devices to the network is rising day by day. Hence, the overall architecture of IoT must satisfy the rising demand. Researchers are addressing various challenging issues related to the effective distribution of IoT with the active data privacy protocol. Several application domains are considered to solve those issues in covering the larger area and self-adoption. These application domains in the IoT system can be Data-centric, Location-based, and Hierarchical-based [5]. With recent computing and communication technology advances, around 7 billion IoT devices are connected worldwide [1]. They perform diverse crowd sensing tasks. Within 2025, the IoT nodes might take place in a single object; as a result, connected device numbers will rise significantly to the Internet [6]. Cisco predicted that around 500 billion devices will be added to the network by 2030.

Managing a massive volume of IoT data from all of a network's knobs is a tedious operation. The data centres' computing efficiency and owner privacy should be considered. Artificial Intelligence (AI) solutions for automated self-decision-making and energy efficiency are used to address decision-making challenges [5]. Although clients' data must be protected from snooping and meddling, security and privacy are the most essential but open concerns in IoT design. Authenticity and the integrity of data should be maintained on the client side. Several cryptographic techniques are provided for data authentication. But it retains major privacy and security breaches, energy consumption and bandwidth issues [7]. Effective collaboration of IoT devices and data uploading to the data centre process faces significant data security and privacy concerns in IoT networks. Given this explosive growth and increasing number of uses, there is a need to design a secure IoT architecture that satisfies rising demand and protects users' data privacy. To solve the issues of larger areas and self-adaptation, several application domains, which can be data-centric, location-based or hierarchical-based, have been considered [5].

In contrast to centralised machine learning models, federated learning frameworks naturally encourage confidentiality and privacy because all data created on an end device does not leave that device. In the federated learning system, data owners are not required to make their data available to the central aggregator. Instead of sending raw data, this learning process uses a sharing model parameter that ensures data protection and privacy at a price that may be much higher than the accuracy loss [8]. Using Federated Learning in wireless IoT networks has several advantages [9], including local Machine Learning system settings can reduce power consumption and wireless bandwidth usage by not exchanging massive amounts of training data; local transmission delay can be significantly decreased by calibrating an ML model's parameters; and only the local learning model variables are transferred when Federated Learning is used, and training data remains on the edge devices themselves, helping to maintain the privacy of the data. However, Federated Learning meets the requirement for computing the IoT data [10].

1.1. Research Motivation

Multi-node machine learning systems are employed to enhance performance, both by allowing for scalable input data volumes and by reducing the number of errors. However, such systems are not restricted to using raw data for input. From a privacy perspective, federated learning is a strong choice when considering IoT design, as it guarantees raw data's privacy, trust, confidentiality and security. The data's authenticity and integrity should be maintained on the client's side. However, cloud-based applications provide Machine Learning-based predictive maintenance solutions to manage the high levels of heterogeneity and diversity in IoT systems. Centralised systems do present an architectural disadvantage from a security perspective. No matter how trusted a centralised server is, it still presents concerns for confidentiality, integrity and availability. Federated Learning paradigms might be a viable solution if direct data is not shared with a central server. In federated learning, devices can run the learning process while charging, connected with the network even if not in use and upload learnt model parameters for an update. Although

several cryptographic techniques, anonymisation, randomisation, perturbation, condensation etc., are available for privacy preservation [1][5].

Significant privacy breaches, energy consumption, and bandwidth are still issues. The effective collaboration of IoT devices for uploading data to a data centre for processing faces significant concerns regarding data security and privacy. IoT data is increasing exponentially due to its cost and open vulnerability. Therefore, monitoring the relevant network using IDS/IPS, SIEM tools, and other advanced security analytics is necessary to detect malicious activity in networks, apps and data. However, a Federated Learning-enabled privacy-preserving framework is capable of handling this challenge.

Recently, all interconnected IoT devices have various protocols and platforms. Data-driven machine learning has been widely applied to develop inference- and decision-making in wireless communications and IoT systems. These ever-broadening sectors require the raw data for processing to transmit to central machine learning. As user privacy and data confidentiality are significant concerns that are not always considered feasible by different parties and organisations, data can be exploited by privacy attacks. Clients' personal information would be used or abused for commercial or political goals without authorisation.

1.2. Key Contributions

Organisations that use IoT systems should incorporate data authentication, access control, attack resistance and client privacy in their business activities as added benefits. This necessitates addressing privacy issues during learning. This study presents Federated Learning-based privacy-preserving methods in IoT systems considering the difficulties and significance of designing them. The primary contributions are explained as follows.

- We present a standard IoT architecture, elucidating and analysing possible threats and providing extensive visualisations of the privacy risks of IoT networks.
- We discuss federated learning from an IoT perspective with its classifications based on different dimensions.
- We propose a federated learning-enabled privacy-preserving framework for IoT networks, with initial experiments demonstrating its significance.

2. RECENT STUDIES RELATED TO FEDERATED LEARNING AND IOT

There are several modern academic analyses evaluating IoT applications and network architecture. For example, the authors [11] reviewed the classification of Federated Learning and a cloud server design regarding security and privacy. This work explicitly discusses the robustness and optimisation schemes of Federated Learning and notes the research challenges and future directions of the field. The authors [12] noted that it categorises attacks against the privacy of ML paradigms. Moreover, possible reasons for privacy leakage were explained, and protective actions against different attacks were analysed. The authors [13] presented a novel questions-based taxonomy analysing privacy leakage in FL. Extensive analyses of data security and privacy preservation standardisations were studied in [14]. Classifications of the privacy issue regarding the life cycles of big data and comparative evaluations of security and privacy preservation were also surveyed.

2.1. IoT and Security Perspectives

The IoT is a paradigm with the express design of creating smart systems. It has significant implementations across several fields, including medical, smart energy, manufacturing, the commercial industry, and homes. IoT systems are comprised of sensors, actuators, networking, and cloud processing and storage. These systems can transfer data without requiring any human-to-human or human-to-computer interaction. IoT systems have multiple known privacy issues, including dependency on vendors, interoperability and transparency, and lack of consent. Each of these needs to be addressed as IoT is projected to expand.

Several IoT architectures have been developed over time. In Figure 1, the latest is compared by visualising their possible privacy threats. Due to technological advances, they are continually upgraded.

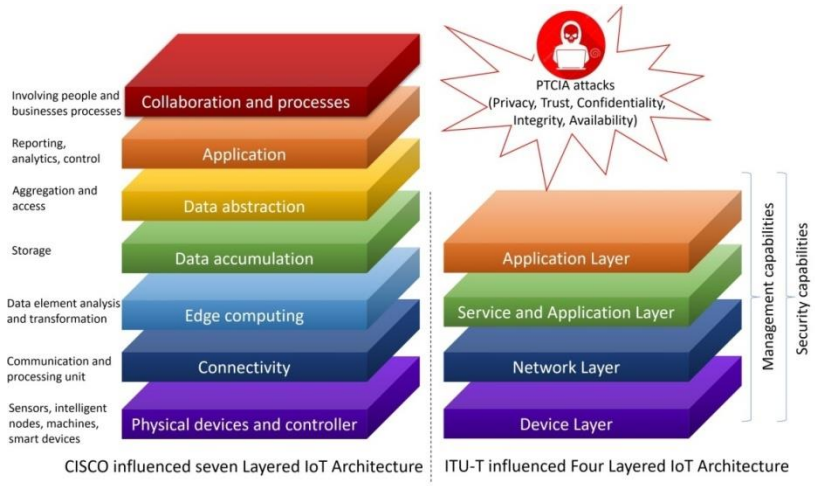

Fig.1. A comparative Architectures of IoT System with possible Privacy Threats

2.2.Layers of IoT Architecture

The basic IoT architecture includes the following layers:

– Physical layer: This layer comprises a wide range of physical IoT devices and controllers responsible for sending and receiving information. Key considerations for this layer include the assumption that the significant scale of large-scale heterogeneous devices accommodates multiple open and proprietary standards for communication.

– Connectivity Layer: This layer is responsible for the reliable and timely transmission of information from devices across a network. However, other referenced IoT architectures contain network layers that perform similar communications. One of the goals of connectivity in the IoT reference model is to enable a current network to handle communications and processing using a gateway.

– Edge Computing layer: the primary function of this layer is to convert a high volume of network data into information for higher-level processing and analysis that makes it suitable for storage. It performs evaluation, expansion, reduction, formation, assessment, etc.

– Data Accumulation layer: this layer carries data flows through it at the rate and in the order

dictated by the devices that generate them and switches to querybased processing. This is critical in linking the differences between real-time networking and non-real-time applications.

– Data Abstraction layer: this layer focuses on presenting data and its storage and is responsible for enabling the development of more straightforward, highperformance applications in IoT systems. Various IoT devices continue to produce data stored for an event which can increase query times.

– Application layer: this layer emphasizes monitoring and controlling models, programming patterns and the flexibility of IoT data. It is not required to run at network speeds. Moreover, perpendicular markets, the type of device data, and business demand influence it.

– Collaboration and processes layer: The essential characteristics of an IoT system are communication and collaboration that involve people and processes. People use programs and data to meet their specific requirements. Multiple users frequently use the same program for a variety of objectives. This layer empowers them to collaborate and process their work to improve it.

2.3. Security and Privacy on IoT Architecture

IoT architectural concepts are based on several factors, including geographic distribution of devices, scale, optimisations for speed, network aims and business constraints [15]. One of the most popular reference architectures is the Cisco-referenced IoT architecture, which incorporates an IoT network's privacy and security issues. In this model, the following layers are considered: identity management (software), authentication/authorisation (software), secure storage (hardware & software), tamper-resistant (software), secure communications (protocols and encryption), secure network access (hardware & protocols) and secure content (silicon).

Table1. Summary of referenced IoT Architecture

Referenced	Number of Layers	Protocol stacks
Basic architecture	3	Perception layer, Network Layer, Application Layer
ITU-T reference	4	Device Layer, Network Layer, service and Application Layer, Application Layer
WSO2	5	Device Layer, Communication Layer, Aggregation/Bus Layer, Event Processing and Analytics Layer, Client/External Communications Layer
FP7 reference	6	Data layer, end to end layer, ID layer, Network Layer, Link layer, physical layer
Cisco	7	Physical devices and controller Layer, Connectivity Layer, Edge computing Layer, Data accumulation Layer, Data abstraction Layer, Application Layer, Collaboration and processes Layer

Table 1 summarises various referenced IoT architectures with their layers and protocol stacks, including a four-layered IoT cybersecurity framework and classified 6 IoT cybersecurity attacks. IoT architecture is a vast concept. There is no proposed uniform architecture such as IoT Forum Architecture with three layers including the transportation layer, Qian Xiaocong, Zhang Jidong architecture, Kun Han, Shurong Liu, Dacheng Zhang and Ying Han's (2012) Architecture including tracking and position is discussed in [16].

3. SUBEDGE INTELLIGENCE

Future IoT networks will continue to develop into a converged Cloud-Edge-Terminal ecosystem that can support various essential artificial intelligence applications on edge computing devices, creating a pervasive Edge Intelligence paradigm to support the intelligent transformation of vertical industries and differentiated service innovations. However, privacy preservation, data security, and data transmission over IoT networks necessitate building artificial intelligence using conventional machine learning techniques in isolated edge units where the distributed data is assembled. In this regard, federated learning may take advantage of the computing capacity of edge servers and the data gathered by widely scattered edge devices, making it a potential solution for applications utilising edge computing and intelligence.

3.1. Federated Learning

The fast development of IoT applications has coincided with an increase in the need for securely and reliably learning data in dispersed systems. For efficient data processing, various IoT applications are currently choosing distributed ML for which there have been recent advancements, i.e., Federated Learning (FL) is preferred for accessing heterogeneous data, privacy, security and rights without sharing data. In FL, data is collected and handled locally for each user or note. Updated data parameters are transferred from clients to a central aggregator for final aggregation. Federated Learning has many advantages over centralised learning in IoT systems due to its distributed storage, processing and privacy protection.

TYPES OF FEDERATED LEARNING FL is a potential distributed framework adopted in many application settings. Also, it considers a client's privacy, data size, computation and energy. State-of-the-art Federated Learning techniques for optimising resources are roughly partitioned into two forms: opaque (or black)-box and transparent (or white)-box forms [17].

3.2. Opaque-box method

Opaque-box FL methods It has strategies for training tricks, hierarchical aggregation, client selection and data reimbursement.

3.3. Transparent-Box Method

It involves the concepts of model compression, feature blend, knowledge extraction and asynchronous updating.

However, federated learning can be classified into three types [8].These are Horizontal Federated Learning, Vertical Federated Learning, and Federated Transfer Learning, as explained below.

Horizontal Federated Learning In horizontal federated learning, client datasets with the same feature spaces across all devices are employed. In Sample-Partitioned Federated Learning, the overlapping features from data samples maintained by different participants are used to train a model collaboratively. This provides a simple yet more effective solution than the standard centralised learning paradigm for preventing private local data from being leaked.

Vertical Federated Learning In Vertical Federated Learning, different datasets with different feature spaces jointly train a global model. This feature-partitioned Federated Learning is a method for cooperatively training a model using data samples with non-overlapping or partially overlapping features maintained by many participants.

Federated transfer learning This improves statistical models in a data federation by sharing knowledge without violating user privacy and allowing complementary knowledge to be transported across the network. It is a predictive model that predicts labels for unlabelled samples using feature representations from aligned samples.

1. Centralised federated learning A central server organises all the participating nodes using an algorithm during this process. The server is responsible for client selection at the start of the training operation and collecting the model's aggregated updates. The server may become a system bottleneck because all the chosen nodes must submit updates to a single entity.
2. Decentralised federated learning In this process, the clients select themselves to receive the global model. This avoids the collapse of a particular point during a model exchange because the model's updates are transferred only between connecting nodes without the involvement of a central server. However, the network's topology may influence the performance of this learning process.
3. Heterogeneous federated learning In it, IoT devices act as heterogeneous clients. Researchers are now working on the HeteroFL framework for addressing heterogeneous clients with varying computational and communication capabilities. In this process, heterogeneous local models can be trained with ongoing computational difficulties while a global model is generated.

4. FEDERATED LEARNING IN IOT RELATED STUDY AND CHALLENGES

Federated learning approaches have needed to adapt to use within IoT architectures. This is necessary given the constraints inherent in IoT environments, including transfer cost, latency, privacy and incompatibility [18]. Although Federated Learning is an efficient distributed learning process for privacy preservation, it does have limitations, specifically as it creates performance and security points of failure as it relies on a single centralised server. Furthermore, it uses intelligent IoT devices with high-functioning designs to decrease the number of communication rounds during a model's training where limited IoT devices are not feasible [19]. Typical distributed learning systems offer several advantages. A summary view of the Federated learning frameworks with advantages from an IoT perspective is presented in Table 2.

Table 2. Federated Learning frameworks in IoT perspective.

Frame-works	Langu-age	Open-source	Implements	Advantage
PySyft	Python	✓	•Private deep learning •Private data decoupling •Homomorphic encryption •Differential Privacy	•Used for data storage location and owner •Can be fetched from virtual worker
FedML		✓	•Separate distribute communication module •Separate training module by PyTorch	•Topology Manager supports to execute of several federated learning algorithms
TensorFlow Federated		✓	•experiments on distributed dataset	•Supports high-level interface set for existing TensorFlow •Establish lower-level interfaces for new federated learning algorithms
LEAF	Federated settings		•Mocha •FedAvg •minibatch SGD	•Supports meta-learning, multitask and federated •Easier implementation of diverse experimental scenarios
Paddle FL		✓	•Natural language •Processing computer vision	•Easily replicate various federated learning algorithms for large scale distributed clusters •Easily deployed on full-stack open-source software

They have non-independent, identically distributed (Non-IID) training data requiring less IoT device communication to preserve privacy. Table 2 presents a summary of Federated Learning Personalised frameworks with their advantages and disadvantages in terms of the IoT. Also, while most current research focuses on Federated Learning's convergence time rather than its trustworthy global aggregation, an edge aggregator is expected to improve the trustworthiness of its framework without compromising the model's accuracy.

5. PRIVACY PRESERVATION

Privacy preservation is essential for users, clients and service providers. Data may be sourced from the IoT, social networks, intelligent business applications, databases, documents, the Internet, etc. Such information can include personal or private information derived from it. To protect digital asset privacy, cyber security mechanisms are defined as Privacy preservation. There are several techniques used to ensure aspects of privacy preservation, and these broadly are partitioned into several categories[20]: encryption; perturbation; authentication; differential privacy (DP); data distribution; anonymisation (K-anonymity, l-diversity, t-closeness etc.); randomisation; Multidimensional Sensitivity Based Anonymisation (MDSBA); condensation; cryptographic techniques; and combined lightweight artificial intelligence.

A wide range of research is being undertaken to provide new frameworks and schemes and enhance privacy preservation.

Edge intelligence has brought several advantages to IoT infrastructure, including mechanisms to preserve privacy using federated learning. Federated learning was developed to enhance the intelligent privacy preservation of end-users sensitive information in edge nodes. According to their objectives, existing federated learning-based privacy-preserving methods are classified into two main categories:

1) data privacy and 2) content privacy[21]. Homomorphic encryption (HE) and secure multiparty computing (SMC) are essential for preventing direct access to users' data and establishing data privacy. On the other hand, to ensure content privacy, researchers modify the original data using several methods such as DP[22], generalisation and perturbation.

Differential privacy There are approaches in which an eavesdropper or data analyst does not determine whether a specific individual's sensitive data is employed in a computation[23]. In contrast, privacy-preservation techniques are strongly motivated by Differential privacy. There are several reasons DP is increasing in popularity, including:

1. the ability for DP to protect any individual's sensitive information without determining an attacker's intention; and
2. DP is not concerned with what an attacker knows about datasets. Also, data analysts may use those datasets because individual information remains protected.

Differential privacy can be described as protecting an individual's sensitive private information through a process that takes that information as input and returns the processed output. DP can be achieved through statistical computation, anonymisation techniques, ML, or other methods. It might add randomness and/or noises or remove information. Based on our observations, DP can be classified as local and global differential privacy, as shown in Figure 2. This figure also describes the key features of both differential privacy types.

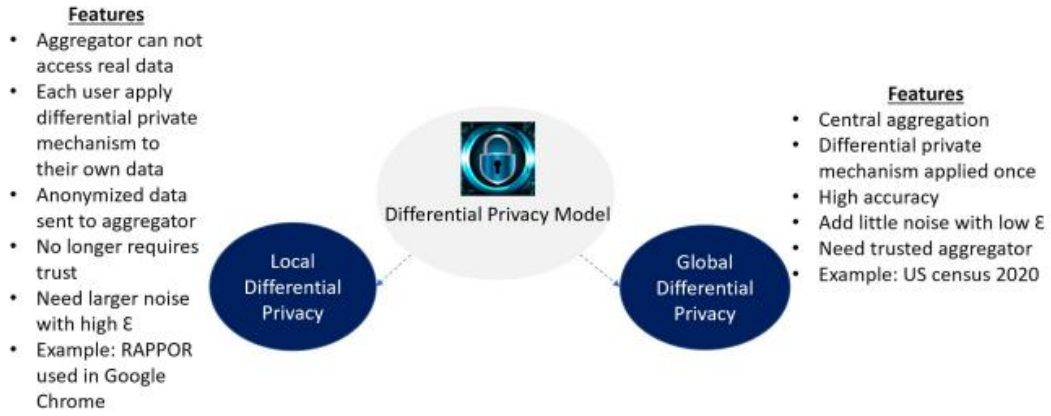

Fig. 2. Differential privacy model with characteristics

Local Differential Privacy A local differential privacy model can be defined as an aggregator that cannot retrieve actual data, with every client individually using a differentially private mechanism for its data. In this model, every user sends only anonymised data to the aggregator. After collecting the anonymised data, the aggregator may compute and disseminate statistics. RAPPOR, a system created to capture differential privacy data in Google Chrome, is perhaps the most recent and well-known example of this approach. Apple uses it to collect data on the iOS keyboard. The benefit of this technique is that it eliminates the requirement for trust because users secure their data. Therefore, even if the aggregator is malicious, data remains safe. This model is suitable if trust is challenging to obtain, but it has a significant limitation in that anonymising user data produces a final dataset that contains considerable noise levels. However, practical applications frequently use high values of ϵ to overcome this difficulty.

Global Differential Privacy A global DP model can be defined as one in which only the central aggregator can read the true data. The aggregator might be a facility or a research association collecting information about individuals. In this model, the aggregator collects real data without noise and transforms it using a differential privacy mechanism. It has the one significant advantage of accuracy because, as a DP mechanism is applied only once at the end, combining more noise is unnecessary to obtain a better result with minimal ϵ. However, a primary concern regarding this model is that the central aggregator needs to be trusted; otherwise, all the data can be hacked and leaked. The US Census is arguably the best-known real-world illustration of global differential privacy.

6. PRIVACY PRESERVATION IN FEDERATED LEARNING IOT PERSPECTIVE

FL was developed to improve the privacy preservation of end-users sensitive information. According to the objects to be protected, existing FL-based privacy preserving methods are classified into two categories: 1) data privacy and 2) content privacy [24]. Homomorphic encryption and secure multi-party computing are essential features for preventing direct access to users' data to establish data privacy. On the other hand, to ensure content privacy, researchers modify the original data using several methods, such as DP, generalisation and perturbation.

Table 3. Federated Learning Personalized Schemes in Terms of IoT.

Personalised Learning Schemes	Data Distribution	Proposed Model	Pros	Cons
Federated transfer learning	Non-iid	FedPer	•Less computation •Less communication communication overhead	•Requires model pruning and compression techniques
Federated meta-learning	Non-iid	Personalised FedAvg	•Flexible to combine with model representation •Can learn and adapt quickly from only a few data samples	•It has higher implementation complexity not suitable for massive data
Federated multitask learning	iid	MOCHA	•Robust to fault tolerance •Great significance for intelligent IoT applications	•Produces one model per task
Federated distillation	iid	FedMD	•Significantly reduce the communication cost •Exchanges not the not the model parameters but the model outputs	•A public dataset is required
Data augmentation	iid	FAug	•Can train a more personalised and accurate model for classification	•FAug should be trustworthy

Unauthenticated IoT nodes with anaemic behaviour might cause harm to the global network. A lightweight fingerprint Federated Learning approach isolates related devices and eliminates hidden or illegal ones from the web. It is also capable of identifying spoofed devices. Federated Learning protects privacy in IoT networks by avoiding raw data-sharing in a model's training. However, sensitive information still faces data leakage from the model's updates. Traditional cryptographic methods are inappropriate for solving this problem in distributed IoT settings. Homomorphic encryption enables the discovery of possible cryptographic solutions for a distributed model's transmission.

The complicated heterogeneity of IoT systems poses significant challenges for conventional Federated Learning. Researchers are trying to find an intelligent way of solving this heterogeneity issue in cloud-edge architectures by providing personalised Federated Learning frameworks [22] with new trends of different types for privacy preservation. This literature review introduces a new privacy paradigm

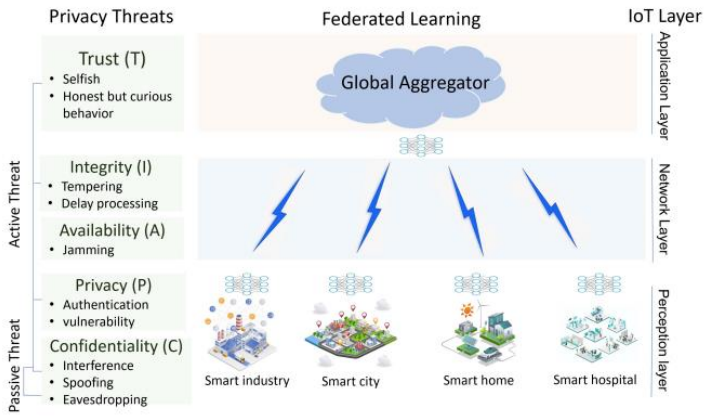

Fig. 3. Summary of IoT-related security and privacy attack issues PTCIA

called PTCIA (privacy, trust, confidentiality, integrity and availability) for attacks in IoT networks. Its concept combines Data confidentiality, Privacy, and Trust - collectively known as DPT [25] and Confidentiality, Integrity, and Availability - the ”CIA triad” [26]. The proposed privacy threat paradigm PTCIA (Privacy, Trust, Confidentiality, Integrity, Availability) and their threat actions in federated learning from an IoT perspective are shown in Figure 3.

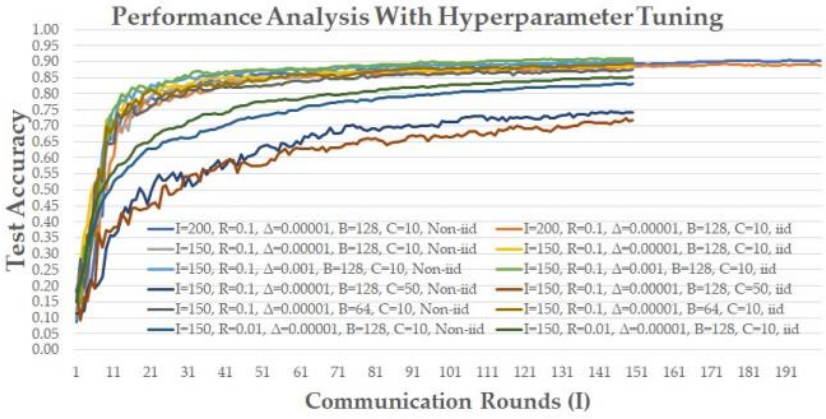

Fig. 4. Test set accuracy vs. communication rounds with MNIST dataset by tuning Hyperparameters of Federated Learning and Privacy Preservation

Federated learning models with differential privacy are implemented and evaluated on the MNIST dataset (Modified National Institute of Standards and Technology database) with different IoT client numbers C=(10, 50) in both iid and Non-iid with different δ=(0.0001, 0.001) noise variation in various iteration I=(150, 200), Learning rate η=(0.1, 0.01), and batch size B=(64, 128) in Figure 4. Here we can see that δ=0.001 noise addition in the same IoT client number got almost 90%

7. RESEARCH CHALLENGES AND FUTURE DIRECTIONS

Privacy preservation is linked to categorising various privacy threat issues, such as PTCIA, and their operations in recent research. Heterogeneity data is generated from questions about an IoT system’s node authentication and confidentiality problems related to integrity attacks that illegally sniff, steal, and/or modify the original data. The following paragraphs discuss the open challenges related to the security and privacy of federated learning in IoT networks.

– Distributed Federated Learning in Edge Computing requires parameter exchanges among the edge nodes that consume bandwidth which the Distributed Hierarchical Tensor Deep Computation Model tried to prove using the STL-10 dataset. However, it still faces many challenges in real-world applications where Federated Learning relies on a server for the aggregation of a local model. This model has a scalability issue, and Federated Learning has a few constraints, such as slow and unstable communication and varying heterogeneous resources. Consensus-based algorithms can be applied to overcome these issues. A deep Federated Learning framework ensures the privacy and ownership of users’ sensitive healthcare data and performs better using the Atlas Dermatology dataset. However, it is unsuitable for data-sharing as it increases a model’s conversion time.

– Handling communication delays between federated learning clients and the central server aggregating their models is still under-explored. To overcome this problem, an edge computing-

based joint client selection and networking scheme for vehicular IoT can be applied. However, it needs to hold raw data locally in federated settings.

– Responding to the challenges of long training times and the consumption of a considerable amount of communication resources, the researcher proposed resource-efficient federated learning with hierarchical aggregation (RFL-HA) methods [27]. Although they are usually suitable for only static networks, they could be extended to dynamic ones.

Module-based neural-structure-aware resource management has been proposed to conduct resource optimisation. However, this framework supports only model partitioning in terms of width, depth, and kernel size. Therefore, sub-model structures that are flexible in these aspects can be implemented.

8. CONCLUSION

This paper has discussed Comparative referenced IoT architectures with the categorised PTCIA privacy attack paradigm. It presented an overall summary of Federated learning (FL) and its categories and investigated the potential privacy threat of federated learning from an IoT perspective. Finally, several recent research challenges and their causes were described, and recommendations for solving them were provided. Dispersed federated learning will be a crucial strategy for future IoT applications involving many end devices.

REFERENCES

[1] Wei Yang Bryan Lim, Nguyen Cong Luong, Dinh Thai Hoang, Yutao Jiao, Ying-Chang Liang, Qiang Yang, Dusit Niyato, and Chunyan Miao. Federated learning in mobile edge networks: A comprehensive survey. IEEE Communications Surveys & Tutorials, 22(3):2031–2063, 2020.

[2] Xhafer Krasniqi and Edmond Hajrizi. Use of iot technology to drive the automotive industry from connected to full autonomous vehicles. IFAC-PapersOnLine, 49(29):269–274, 2016.

[3] Ibrar Yaqoob, Ejaz Ahmed, Ibrahim Abaker Targio Hashem, Abdelmuttlib Ibrahim Abdalla Ahmed, Abdullah Gani, Muhammad Imran, and Mohsen Guizani. Internet of things architecture: Recent advances, taxonomy, requirements, and open challenges. IEEE wireless communications, 24(3):10 16, 2017.

[4] Jackie Fenn and Mark Raskino. Gartner's hype cycle special report for 2011. Stamford, CT: Gartner, 2011.

[5] Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic, and Marimuthu Palaniswami. Internet of things (iot): A vision, architectural elements, and future directions. Future generation computer systems, 29(7):1645–1660, 2013.

[6] Rafael Angarita. Responsible objects: Towards self-healing internet of things applications. In 2015 IEEE International Conference on Autonomic Computing, pages 307–312. IEEE, 2015.

[7] Martin Feldhofer, Sandra Dominikus, and Johannes Wolkerstorfer. Strong authentication for rfid systems using the aes algorithm. In Cryptographic Hardware and Embedded Systems-CHES 2004: 6th International Workshop Cambridge, MA, USA, August 11-13, 2004. Proceedings 6, pages 357–370. Springer, 2004.

[8] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 10(2):1–19, 2019.

[9] Md Mamunur Rashid, Shahriar Usman Khan, Fariha Eusufzai, Md Azharuddin Redwan, Saifur Rahman Sabuj, and Mahmoud Elsharief. A federated learning-based approach for improving intrusion detection in industrial internet of things networks. Network, 3(1):158–179, 2023.

[10] Muneerah Al Asqah and Tarek Moulahi. Federated learning and blockchain integration for privacy protection in the internet of things: Challenges and solutions. Future Internet, 15(6):203, 2023.

[11] Latif U Khan, Walid Saad, Zhu Han, Ekram Hossain, and Choong Seon Hong. Federated learning for internet of things: Recent advances, taxonomy, and open challenges. IEEE Communications

Surveys & Tutorials, 23(3):1759–1799, 2021.
[12] Maria Rigaki and Sebastian Garcia. A survey of privacy attacks in machine learning. ACM Computing Surveys, 56(4):1–34, 2023.
[13] Xuefei Yin, Yanming Zhu, and Jiankun Hu. A comprehensive survey of privacy-preserving federated learning: A taxonomy, review, and future directions. ACM Computing Surveys (CSUR), 54(6):1–36, 2021.
[14] Jahoon Koo, Giluk Kang, and Young-Gab Kim. Security and privacy in big data life cycle: a survey and open challenges. Sustainability, 12(24):10571, 2020.
[15] Henry Muccini and Mahyar Tourchi Moghaddam. Iot architectural styles: A systematic mapping study. In Software Architecture: 12th European Conference on Software Architecture, ECSA 2018, Madrid, Spain, September 24–28, 2018, Proceedings 12, pages 68–85. Springer, 2018.
[16] Somayya Madakam, Ramya Ramaswamy, and Siddharth Tripathi. Internet of things (iot): A literature review. Journal of Computer and Communications, 3(5):164–173, 2015.
[17] Rong Yu and Peichun Li. Toward resource-efficient federated learning in mobile edge computing. IEEE Network, 35(1):148–155, 2021.
[18] Ulrich Matchi A¨ıvodji, S´ebastien Gambs, and Alexandre Martin. Iotfla: A secured and privacypreserving smart home architecture implementing federated learning. In 2019 IEEE security and privacy workshops (SPW), pages 175–180. IEEE, 2019.
[19] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In Artificial intelligence and statistics, pages 1273–1282. PMLR, 2017.
[20] Marwa Keshk, Benjamin Turnbull, Nour Moustafa, Dinusha Vatsalan, and Kim-Kwang Raymond Choo. A privacy-preserving-framework-based blockchain and deep learning for protecting smart power networks. IEEE Transactions on Industrial Informatics, 16(8):5110–5118, 2019.
[21] Yoshinori Aono, Takuya Hayashi, Lihua Wang, Shiho Moriai, et al. Privacy-preserving deep learning via additively homomorphic encryption. IEEE transactions on information forensics and security, 13(5):1333–1345, 2017.
[22] Qiong Wu, Kaiwen He, and Xu Chen. Personalized federated learning for intelligent iot applications: A cloud-edge based framework. IEEE Open Journal of the Computer Society, 1:35–44, 2020.
[23] Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science, 9(3–4):211–407, 2014.
[24] Shiho Moriai. Privacy preserving deep learning via additively homomorphic encryption. In 2019 IEEE 26th Symposium on Computer Arithmetic (ARITH), pages 198–198. IEEE, 2019.
[25] Yassine Maleh, Abdellah Ezzati, and Mustapha Belaissaoui. Security and privacy in smart sensor networks. IGI Global, 2018.
[26] Javed Asharf, Nour Moustafa, Hasnat Khurshid, Essam Debie, Waqas Haider, and Abdul Wahab. A review of intrusion detection systems using machine and deep learning in internet of things: Challenges, solutions and future directions. Electronics, 9(7):1177, 2020.
[27] Zhiyuan Wang, Hongli Xu, Jianchun Liu, He Huang, Chunming Qiao, and Yangming Zhao. Resource-efficient federated learning with hierarchical aggregation in edge computing. In IEEE INFOCOM 2021-IEEE conference on computer communications, pages 1–10. IEEE, 2021.