Babylonian Journal of Internet of Things
Vol.2023, **pp**. 1–7
DOI: https://doi.org/10.58496/BJIoT/2023/001; ISSN: 3006-1083
https://mesopotamian.press/journals/index.php/BJIoT

Research Article

Federated Learning in IoT: A Survey on Distributed Decision Making

Rasha Talal Hameed[1], Omar Abdulwahabe Mohamad[1,*],

[1] *Computer Science, College of Education, Al-Iraqia University, Iraq.*

ARTICLE INFO

Article History

Received 17 Nov 2022
Accepted 04 Jan 2023
Published 10 Jan 2023

Keywords

Federated Learning

Internet of Things (IoT)

Distributed Decision Making

Decentralized Machine Learning

IoT Applications

ABSTRACT

The proliferation of Internet of Things (IoT) devices has ushered in an era of unprecedented data generation, necessitating innovative approaches to extract valuable insights while respecting privacy and resource constraints. Federated Learning (FL) operates as a transformative decentralized model training solution for networks of multiple edge devices. This paper provides an extensive review of Federated Learning usage within IoT frameworks while mainly discussing distributed decision making obstacles and resolution approaches. This study introduces basic principles about Federated Learning and IoT along with their integration and outlining the advantages of implementing FL in decentralized IoT networks. The paper details various implementations of Federated Learning across Internet of Things fields which cover smart cities and healthcare systems and industrial IoT and smart grid systems. The study demonstrates how FL solves IoT data distribution issues consisting of heterogeneous information and privacy vulnerabilities and network limitations. The survey devotes a major part to analyze federated learning algorithms with their methodologies for distributed decision making in IoT applications. The research covers federated optimization approaches and provides details about algorithms that reduce communication while ensuring privacy protection. Edge computing serves as a fundamental component for running efficient FL in IoT networks due to the resource limitations of edge devices according to the survey evaluation. The paper performs a review of contemporary federated learning framework and platform solutions developed for IoT environments. This section assesses real-life problem-solving abilities and deployment scalability of distributed decisions from available frameworks. Future research in federated learning for IoT requires examination of ongoing challenges along with directions for enhancement to include new algorithms together with strong security measures and standardized platforms according to the survey outcome. This paper presents a thorough insight into IoT Federated Learning by assessing its capabilities and addressing its difficulties as well as solutions especially regarding distributed decision processes. The information presented enables researchers along with practitioners and policymakers to understand FL in IoT better to support progress in this dynamic field.

1. INTRODUCTION

Smart devices integrated with Internet-of-Things (IoT) [1] networks propagate distributed computing environments that spread data generation and storage across numerous fundamental nodes.[2]The distributed framework of IoT systems creates difficulties for implementing conventional data-driven machine learning approaches that need substantial unified datasets. Federated learning is a distributed interaction approach that avoids the need to move particular data among places by allowing several entities to train machine learning algorithms across different edge devices or servers. The technique offers a tempting way for IoT networks to manage the massive amount of private user data that is gathered from sensors and connected devices.

In federated learning for IoT platforms, localized data from devices can be used to cooperatively train a common global model without the need for direct data sharing or pooling with consumers. [3] The technique protects user privacy while allowing network-wide knowledge of statistics collection to continue to benefit users. Devices are then given access to the trained global model so they can use it to make informed decisions based on fresh local data.

*Corresponding author. Email: engit2020@gmail.com

While showing promise, federated learning opens stimulating research questions around efficiency, robustness, and security when put into practice - especially for real-world IoT deployments at population-scale. This survey reviews federated learning developments for IoT contexts with a lens towards distributed decision making in particular. We summarize the current landscape of algorithms, applications, and open challenges for effectively coordinating intelligent decisions across decentralized IoT devices and users via federated learning's distributed model training approach.

The organization of this paper is as follows. First, we outline the fundamentals of federated learning and its distinguishing components. We then detail popular applications of federated learning specifically using IoT network data across domains like smart homes, smart cities, and healthcare. Focusing on decision making, we next categorize different consensus algorithms and other distributed techniques investigated in the literature to coordinate decisions across IoT devices informed by shared models from federated learning. Finally, we suggest promising directions and open problems to advance the convergence of federated learning and decentralized decision making tailored to practical IoT networks with potential real-world societal impact.

2. BACKGROUND

2.1 Federated Learning

Federated Learning (FL)[4] represents a paradigm shift in the field of machine learning, particularly in the context of distributed and decentralized data environments. It addresses the challenges posed by the increasing ubiquity of connected devices, such as smartphones, Internet of Things (IoT)[5] devices, and edge computing nodes. The traditional model of centralized machine learning, where data is collected and processed in a central server, faces limitations in scenarios where data privacy, communication bandwidth, and real-time processing are critical concerns. The fundamental premise of Federated Learning is to bring the machine learning model to the data, rather than centralizing the data for model training. The decentralized model optimization method becomes essential for situations with sensitive large-scale or dispersed data especially when operating through restricted device networks. Model training in FL occurs throughout numerous devices without transmitting raw information to central servers which simultaneously resolves privacy matters and streamlines data transfer requirements.

The programming model of FL operates through decentralized training of models that happens directly on individual devices but only distributes updates to gradients to central servers or neighboring devices. This technique's distributed infrastructure minimizes the need to transfer large amounts of data, which lessens the susceptibility of publicly available raw private information. Programs in the healthcare, financial, and personal device sectors need the highest level of privacy protection possible. Because it becomes necessary for operational procedures, FL is highly dependent on the protection of personal privacy. Because FL uses model updates and retains raw data on users' computers throughout training, data privacy is protected.

FL improves communication efficiency by enabling its devices to transmit unique model information. The model update distribution reduces bandwidth requirements and speeds up replies, FL enables devices with few resources to function efficiently, enabling real-time applications. The FL's principles of operation align well with the distributed computing paradigm, which processes data at short distances from its origin, It fits in perfectly with edge computing. Productive model training using edge devices, such as sensors and Internet of Things devices, takes place on-site without relying on a large cloud infrastructure, resulting in quick answers and requiring fewer external servers.

The FL works in a number of areas, including the banking and healthcare sectors, as well as smart city operations and industrial Internet of things applications. By facilitating application-based learning and proactive maintenance of machines through peer cooperation, as well as by creating patient-specific healthcare assistance systems while protecting user privacy, FL exemplifies its usefulness. Ongoing research focuses on solving multiple challenges that affect model aggregation as well as security and robustness and scalability in Federated Learning. FL demonstrates potential as a promising solution because it fits the requirements of modern decentralized systems that need privacy protection.

2.2 Internet of Things (IoT)

Different sources define the Internet of Things (IoT)[6] as the connected infrastructure of physical devices equipped with sensors and actuators which enable them to gather and transfer information.[7]Connected devices under IoT classification perform various monitoring and task execution operations through inter-networked communication with central systems and other IoT devices.

The IoT concept developed throughout recent years because technological innovations combined with the need for efficient connected systems.

Key Components of IoT:

IoT devices contain sensors which collect environmental data while built-in actuators execute directed actions following sensor-obtained information. The combination of temperature sensors along with motion sensors and cameras function as common sensors yet actuators consist of motors or valves or alternative control elements. The sharing of data requires IoT devices to connect to networks according to reference 8. Different communication protocols including Wi-Fi, Bluetooth, Zigbee, cellular networks and Low Power Wide Area Networks (LPWAN) enable the achievement of this technology. The selected connectivity system should match the requirements of data transfer speed alongside its operational range along with power usage needs.

IoT devices usually conduct data processing either through onboard local capabilities or by transmitting information to centralized cloud databases for evaluation. Edge computing has become important for IoT because it allows data processing near the data source which improves real-time responses and decreases latency. The massive data output of IoT devices depends on cloud platforms for both data storage and analysis functions. Cloud computing systems deliver crucial computing tools for managing the extensive and frequently huge datasets that originate from IoT ecosystems.

When it comes to IoT security stands as the most vital issue because of its interconnected nature. IoT devices present exposure to cyber-attacks which leads to serious outcomes when devices become compromised. Physical protection for IoT systems consists of encrypted information and authentication protocols and frequent software updates that handle security weaknesses. Seamless communication and collaboration require proper management of diverse IoT devices and systems interoperability. Standards development for common communication protocols together with data format standards continue to advance in order to achieve interoperability across various IoT devices.

3. SIGNIFICANCE OF FL IN THE CONTEXT OF INTERNET OF THINGS (IOT)

The integration of Federated Learning (FL) with the IoT environment FL plays a key role due to its ability to revolutionize the way decentralized training of machine learning models functions in interconnected systems. [9,10] A combination of factors makes FL essential for IoT applications. The following elements highlight how crucial FL is in the context of IoT:

1. Decentralized Data Governance:

The Internet of Things network generates large data collections that disperse among numerous devices without ensuring privacy requirements. By avoiding data centralization, FL's distributed algorithm training abilities preserve end device privacy and satisfy data compliance requirements.

2. Preserving Data Locality:

FL makes it possible to install models near data sources, allowing devices to do analytical tasks on-site without worrying about network overload or security threats .Because it presents security issues and necessitates substantial computer resources, sending large, unprocessed IoT datasets to central servers is difficult.

3. Mitigating Communication Overhead:

When centralized servers are utilized, standard machine learning algorithms perform inefficiently, especially in situations with bad connectivity or limited capacity. The FL system is a great technique for IoT installations since it distributes model changes rather than full data, optimizing data transmission efficiency.

4. Privacy-Preserving Machine Learning:

The IoT networks handle a wide variety of sensitive and personal user data kinds, user privacy security becomes a critical concern. Federated smoothing and encryption approaches, which allow devices to develop systems with private data.

5. Edge Intelligence and Real-time Decision-Making:

Real-time applications operating on IoT platforms suffer when traditional cloud-based options are used because they cause latency in decision-making. By enabling algorithms to be updated locally through FL's compliance with edge computing, contextual decision-making carried out directly on IoT devices is accelerated.

4. CHALLENGES IN CENTRALIZED MACHINE LEARNING FOR IOT

Classical server-based machine learning algorithms encounter a number of difficulties when interacting with IoT environments due to their standardized methods of gathering data. Appropriate solutions for these implementation challenges are necessary for the successful deployment of machine learning in IoT systems.

The primary challenges include:

1. Data Privacy and Security:

Various sensitive elements of information, such as geographic locations and private health information, as well as specifics about industrial operations, are included in the data produced by IoT devices. All devices must send data to main servers, which puts users' privacy at risk when data is sent among the gadgets and a server.

The attractiveness of centralized structures makes them excellent targets for cyber-attacks. Each piece of collected data is subject to significant privacy and security threats in the event of a single server's security breach or failure.

2. Communication Overhead:

Networks utilized for transferring data from large IoT devices to central server operations frequently result in restrictions that reduce capacity for bandwidth and cause data transfer issues as well as longer transmission times.

Data transport times to central processing nodes pose time-related issues with performance for real-time processing operations.

3. Scalability Issues:

The volume of data processed by the network increases exponentially as the number of IoT gadgets rises. Scalability issues brought on by moving centralized systems might result in decreased performance and resource constraints.

Performance issues arise from centralized processing since it requires a lot of processing power, which runs out as device connectivity rises.

4. Data Heterogeneity:

The IoT networks involve devices that use multiple formats and concepts and preserve diverse data structures, they contain a variety of data providers. Underlying hierarchical models, a diversity of data sources creates an adaptation difficulty that lowers throughput outputs.

5. Reliability and Availability:

The failures of servers create problems with service, network delays, and repair delays, centralized platforms that depend on a central server have significant risks. Relationships between various system components make service outages and decreased reliability conceivable.

6. Regulatory Compliance:

Data globalization rules, which specify precise regional boundaries for data processing and storage, must be followed by businesses. Centralized systems for data face difficulties as a result of mandated territorial data processing and storage requirements.

7. Energy Consumption:

Data transmission from IoT gadgets to a main server consumes energy in a way never seen before, especially when dealing with low-power IoT gadgets. Devices that run on batteries may use so much more energy that their overall operating time is shortened.

Effective fixes for these problems are necessary for the effective incorporation of machine learning in the Internet of Things. Federated Learning, a developing technique that offers autonomous learning abilities, security for privacy, and a decreased communication burden, improves flexibility and compatibility with different IoT scenarios.

5. APPLICATIONS OF FEDERATED LEARNING IN IOT

The FL is a novel technological method for the IOT that provides independent learning of system operation and algorithms that guarantee secrecy. Federated Learning's adaptability is demonstrated by the variety of IoT domains in which it operates programs in various settings.

Among the noteworthy applications are:

1. Smart Cities:

With FL traffic management benefits from predictive traffic operations which processes data through traffic cameras and sensors and mobile devices. Real-time traffic movement adjustments are made possible by FL's decentralized management. The application of FL allows smart grids to optimize their energy consumption which leads to both efficient energy distribution and minimization of wastage.

2. Healthcare:

FL enables the development of custom healthcare systems which maintain patient privacy during remote patient monitoring. Health information from wearable technology Stronger prediction models for early disease detection can be produced by combining sensors with electronic health data. FL enables drug discovery acceleration through collective knowledge aggregation of research centers and pharmaceutical organizations while preserving information security.

3. Industrial IoT (IIoT):

FL provides predictive maintenance capabilities through its ability to collect industrial machinery sensor information. The identification of potential equipment faults becomes timely through FL leading to reduced maintenance expenses and fewer machine downtime periods.

As a result of distributed decision-making in FL organizations can run quality control procedures through real-time data analysis from production line sensors.

4. Smart Agriculture:

Precision Farming gets enhanced through FL because it unifies different agricultural IoT devices including soil sensors alongside drones and weather stations. Farmers can make superior choices about irrigation practices plus fertilization procedures and crop management operations by using FL.

5. Edge Devices and IoT Sensors:

Local training of edge devices through FL enables the optimization of IoT device energy consumption. The solution extends battery life for devices because it reduces the requirement for server communication.

The detection of anomalies becomes possible through FL which allows sensors to combine collective insights across networks. Application of this technique in environmental monitoring and security surveillance functions particularly useful.

6. Smart Homes:

Through FL users can customize automation models because the system learns their device preferences by monitoring their habits. The system adapts lighting and temperature controls along with additional features according to personal routines of users.

7. Finance:

The financial sector benefits from FL by using collaborative learning of multiple banks to uncover fraudulent transactions while safeguarding customer information confidentiality. FL decision systems help risk assessment models achieve better accuracy through decentralized consensus-building activities from varied financial institutions.

The several examples demonstrate how Federated Learning improves communication effectiveness and decentralized decision tasks in Internet of Things systems while resolving data privacy concerns. FL technology development will expand the range of novel IoT implementations.

6. OPEN RESEARCH CHALLENGES AND FUTURE DIRECTIONS

Researchers and authorities from the Internet of Things (IoT) domain space are now required to focus on a number of open research issues related to Federated Learning (FL). In order for FL to reach its full potential in the Internet of Things, these current problems must be resolved.

Prospective study strategy guidelines provide FL improvement with a route upward, leading to general interaction and sustained innovation. The main research challenges in FL for IoT are listed in the part that follows, along with potential evaluation paths:

When working with distributed IoT setups that employ FL examples, protecting against hostile attacks is an essential challenge. To achieve widespread adoption, strong defenses against model infection attacks and privacy violations are essential.

The constant issue is to optimize methods of communication to convey less data while maintaining model accuracy. Effective strategies must be created to handle situations in which bandwidth is less than anticipated and connectivity fails.

Deploying FL solutions on a broad scale necessitates methods for problem solving that can manage numerous IoT gadgets at once. Scalability must become a crucial component of linked systems due to the growing number of IoT gadgets.

There are now no established standards, common protocols and structures for FL in IoT systems are unable to create interoperability among devices. Reliable protocols need to be created in order to promote cooperation between different electronic platforms and gadgets.

FL adoption on energy-constrained edge gadgets requires power-saving strategies. Energy savings and model accuracy must be balanced, particularly when working with battery-powered Internet of Things devices.

Hybrid algorithms that incorporate FL with other methods of machine learning, such as reinforcement learning or transfer learning, are being researched in order to enhance their accuracy and flexibility in complex IoT situations.

Understandable FL algorithms are looking at approaches to make FL models more transparent and accessible in order to get beyond the "black-box" aspect of certain machine learning algorithms in distributed systems.

7. CONCLUSION

In conclusion, the current study has examined the dynamic relationship among Federated Learning (FL) and the Internet of Things (IoT), highlighting the significance of autonomous decision-making in IoT contexts. The study estimates IoT programs and methodologies alongside handling upcoming advances, since it demonstrates the manner in which FL allows answers to IoT's numerous difficulties across its variety of gadgets. Federated learning appears to be a powerful solution to the issues brought on by centralization algorithms in the Internet of Things, according to an architecture that promotes data privacy, communication performance, and computing edge flexibility. Applications of FL across a range of industries, such as healthcare, smart cities, and industrial IoT, demonstrate its adaptability and promise to transform the way insights are extracted from dispersed IoT data sources.

Research and innovation should remain a priority because the identified issues involving data privacy, security, communication overhead and scalability need resolution. The solution of these challenges remains essential for achieving smooth integration of FL within IoT domains because it creates trust among users and supports decentralized decision-making models. Looking ahead, the outlined future directions provide a roadmap for researchers and practitioners to explore hybrid models, enhance explainability, promote cross-domain collaboration, and navigate ethical considerations. These avenues pave the way for the evolution of FL applications in IoT beyond theoretical frameworks, fostering real-world implementations that contribute to the advancement of both fields. As Federated Learning in IoT matures, collaboration among academia, industry, and policymakers becomes increasingly vital. Standardization efforts, interdisciplinary research, and the establishment of regulatory frameworks will play crucial roles in shaping the ethical, legal, and practical aspects of FL deployment in IoT ecosystems.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

No funding is provided and no financial support is received to carry out the research presented in this paper.

Acknowledgment

The author would like to express gratitude to the institution for their invaluable support throughout this research project.

References

[1] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, "Federated learning for internet of things: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1622-1658, 2021.

[2] A. Imteaj, U. Thakker, S. Wang, J. Li, and M. H. Amini, "A survey on federated learning for resource-constrained IoT devices," *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 1-24, 2021.

[3] A. Imteaj and M. H. Amini, "Distributed sensing using smart end-user devices: Pathway to federated learning for autonomous IoT," in *2019 International conference on computational science and computational intelligence (CSCI)*, 2019, pp. 1156-1161: IEEE.

[4] L. Li, Y. Fan, M. Tse, and K.-Y. Lin, "A review of applications in federated learning," *Computers & Industrial Engineering*, vol. 149, p. 106854, 2020.

[5] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, "A survey on federated learning," *Knowledge-Based Systems*, vol. 216, p. 106775, 2021.

[6] S. Madakam, V. Lake, V. Lake, and V. Lake, "Internet of Things (IoT): A literature review," *Journal of Computer and Communications,* vol. 3, no. 05, p. 164, 2015.
[7] C. Srinivasan, B. Rajesh, P. Saikalyan, K. Premsagar, and E. S. Yadav, "A review on the different types of Internet of Things (IoT)," *Journal of Advanced Research in Dynamical and Control Systems,* vol. 11, no. 1, pp. 154-158, 2019.
[8] K. Ashton, "That 'internet of things' thing," *RFID journal,* vol. 22, no. 7, pp. 97-114, 2009.
[9] L. Farhan, S. T. Shukur, A. E. Alissa, M. Alrweg, U. Raza, and R. Kharel, "A survey on the challenges and opportunities of the Internet of Things (IoT)," in *2017 Eleventh International Conference on Sensing Technology (ICST)*, 2017, pp. 1-5: IEEE.
[10] N. Hossein Motlagh, M. Mohammadrezaei, J. Hunt, and B. Zakeri, "Internet of Things (IoT) and the energy sector," *Energies,* vol. 13, no. 2, p. 494, 2020.