IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁑ Impact Factor 7.918 ⁑ Vol. 12, Issue 2, February 2023

DOI: 10.17148/IJARCCE.2023.12213

How can we make IoT Applications better with Federated Learning- A Review

Hari Gonaygunta[1], Deepak Kumar[2], Surender Maddini[3], Saeed Fazal Rahman[4]

University of the Cumberlands, KY, USA[1,2,3]

Wilmington University, USA[4]

Abstract: Since its inception by Google, Federated Learning (FL) has been instrumental in improving the performance of a wide range of applications. Android's Gboard for predictive text and Google Assistant are two of the most well-known and widely used FL-powered applications. FL is a configuration that enables on-device, collaborative Machine Learning. A diverse body of literature has investigated FL technical considerations, frameworks, and limitations, with several works presenting a survey of the prominent FL literature. Prior surveys, however, have focused on FL's technical considerations and challenges, and there has been a limitation in more recent work that presents a comprehensive overview of FL's status and future trends in applications and markets. We introduce the fundamentals of FL in this review,describing its underlying implementation of technologies, pros and cons, and recommendations along with privacy- preserving methods. More importantly, this work contributes to the understanding of a wide range of FL current applications and future trends in technology and markets today.

Keywords: Federated Learning, Cybersecurity, IoT, Decentralized networks, NIST.

I. INTRODUCTION

IoT applications are growing rapidly and are becoming an important part of many businesses. However, many IoT applications are still relatively primitive and lack the sophistication needed for more advanced uses. Learning algorithms are a powerful way to improve the effectiveness of IoT applications by improving the performance of machine learning algorithms, without requiring significant changes to the application itself. Using federated learning to train a machine learning model can improve the performance of an IoT application.

Federated learning is a technique that allows a set of compute nodes in a network to work together to train a machine learning model [1]. This technique is different from regular distributed machine learning because multiple nodes in the network can share the training data and provide training samples for one another. With this method, each node can be updated as it trains a new model, reducing the time to retrain the model when a change needs to be made. An integrated system that can handle federated learning allows an IoT application to be much more flexible as business needs change and can be a valuable tool for adapting and improving IoT applications over time [2,7].

II. ENABLED APPLICATIONS

A. Healthcare

As IoT devices become more prevalent in people's daily lives, the privacy of the data collected becomes increasingly important. IoT E-health is an example of a privacy concern [3,4]. Smart wearable devices are now used to monitor patients' health statuses such as heart rate, blood pressure, and glucose level. In comparison to other types of data, personal healthcare data is the most sensitive to privacy and is heavily regulated by the government for any type of datasharing [5]. FADL [6] first proposes a federated training framework that allows each hospital to participate in learning part of the model using its own medical data source in order to mitigate the aforementioned challenges. Yuan et al. [9] present a similar FL-enabled collaborative healthcare framework for medical IoT devices. FL's potential was recently discovered during the fight against the COVID-19 pandemic. Liuet al. [10] present a FL-based system that enables hospitals to collaborate in training models for identifying CT scans of COVID-19 patients without transmitting raw data, thereby overcoming the data isolation problem in IoT E-health [11,12, 18].

B. Cybersecurity

The potential traffic volume of IoT-based DDoS attacks is reaching unprecedented levels, as seen during the Mirai botnet attack leveraging infected webcams and home routers [13]. Attacks like this have raised awareness of the importance of IIoT risk assessment and security, particularly in fields like healthcare. FL can provide an alternative approach to IoT cybersecurity by protecting the network from malicious attacks as close to the edge devices as possible. DIoT [14] is the first system to use a FL approach to anomaly detection-based intrusion detection in IoT gateways. Another similar framework is IoTDefender [15], which obtains personalized models on each local device by fine-tuning the global model trained with FL to filter malicious traffic. The FedML Ecosystem recently proposed the first FL platform designed specifically for IoT devices, as well as the FedDetect algorithm to detect anomalous traffic. Their Raspberry Pi experiments demonstrate the efficiency and feasibility of deploying FL on IoT devices.

C. Industry 4.0

The Industrial Internet of Things (IIoT) is rapidly evolving, bringing several advances in information technology applications for the manufacturing field. The concept of Industry4.0, also known as the fourth industrial revolution, has been proposed in response to the growing importance of IIoT-based interconnectivity and access to real-time data [16]. With unprecedented connectivity, Industry 4.0 will provide greater insight, control, and data visibility for many industries' supply chains. However, there are some real-world issues that are impeding the implementation of Industry 4.0. First, the amount of data generated by a single factory may not be sufficient for comprehensively training a reliable model. Second, the data collected by IIoT devices is highly related to commercial value, making privacy protection critical. Eavesdroppers, for example, may infer manufacturing capacity from industrial IoT users' electricity usage. Lu et al. [17] proposed a blockchain and FL-based data-sharing framework in IIoT to address the aforementioned challenges. Its empirical findings demonstrate the effectiveness of FL usage in IIoT.

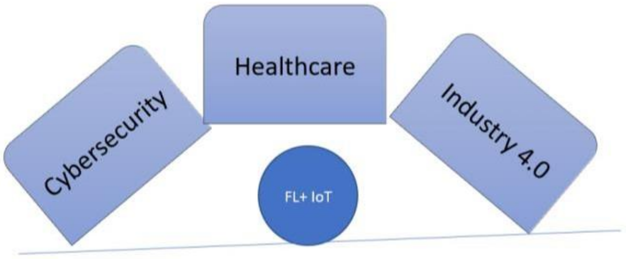

Fig .1 *Enabled Applications*

III. IMPLEMENTATION OF FEDERATED LEARNING WITHIN IOT APPLICATIONS

The beauty of a federated learning model is that it is not dependent on a centralized cloud server for training. Instead, each device in the networked IoT devices is trained independently using data from the other devices in the network rather than directly from the central server. This allows for faster training and updates if a new algorithm is desired or if a problem with the existing algorithm is identified.

As the cost of computing power continues to decline, it becomes more practical to equip more devices with powerful computers that can be integrated into the IoT network. This enables larger networks with more connected devices that can all share in the work of training and updating models to improve the overall performance of the system. The ability to quickly retrain and update a model is critical to the success of an IoT application along with cybersecurity training inbuild, that and can cause significant problems if the update process is not completed in a timely manner [18]. In real-world environments, it can be difficult to ensure that every device is up-to-date and the system is operating properly. Having fewer centralized servers in the network increases the number of potential points of failure and makes it more difficult to maintain a consistent level of performance. Federated learning can reduce these problems by making it possible to update the system in a more decentralized manner. In addition, some devices that may not be able to perform complex computing tasks can also be included within the network without creating a significant impact on the overall system performance. Having the ability to share some of the workload makes it easier to create a versatile system that can meet a wide range of needs.

One of the challenges of implementing an IoT system is making sure that all of the devices are functioning properly at all times. Poorly designed systems can lead to unpredictable behavior that is difficult to troubleshoot. This can lead to major problems that require extensive repairs or result in costly downtime. Decentralized networks like the ones created with federated learning eliminate this problem since it allows each individual device to communicate with the others in the network and train them using the data that is available. This ensures that the system will be fully functional and will not exhibit unexpected behavior that could potentially cause damage or other problems for the users of the system [19].

While many of the benefits of federated learning are similar to those offered by other artificial intelligence techniques, there are some unique benefits provided by this implementation that make it an excellent choice for a variety of

DOI: 10.17148/IJARCCE.2023.12213

applications. One of the most important benefits of federated learning is that it provides a more flexible learning environment that allows it to be used for a wide variety of applications. This means that it has the potential to improve the performance and functionality of virtually any device or device that is connected to the network. This has the potential to create a number of innovative new applications as well as to improve the performance of existing devices and systems that are already being used on the network using digitalization [20].

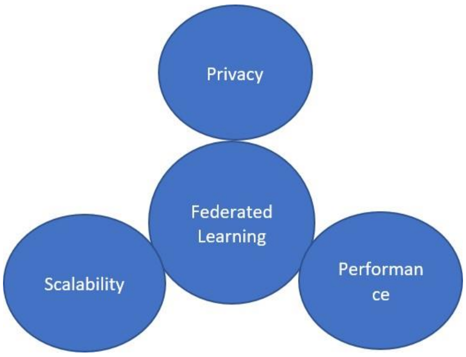

Fig .2 *Federated Learning Benefits*

Another important aspect of federated learning is that it helps to create a more diverse and adaptive network that is better equipped to adapt to changing environmental conditions and other challenges that are encountered by the devices on the network. Since most of the devices on the network are self-sufficient and can learn from each other, they have the ability to better respond to the challenges that they face without having to rely on a central server or other centralized source of information that can limit their capabilities or prevent them from performing as well as they could otherwise. This results in a more adaptable and responsive network that can handle a wider range of problems and challenges that may emerge during prevention from cyber attacks without requiring additional assistance from a centralized server [21]. This type of adaptability and flexibility is one of the key features of most computer networks and can make a big difference in the way the network performs and how effectively it is able to handle the challenges it is faced with on a daily basis. This type of network performance can go a long way towards improving the overall performance of the network and increasing its overall efficiency and functionality. Another benefit of federated learning is that it provides users with a more individualized experience that allows them to customize the system to their needs and preferences. This is especially important when it comes to mobile devices that are constantly moving between different physical locations and may be subject to different types of connectivity and performance standards that may be in place in those locations. With federated learning, devices can communicate with each other and coordinate their activities in a way that allows them to maintain a consistent level of performance regardless of where in the network they are located or what specific types of access conditions they may be operating under. This not only makes using the devices on the network more convenient and intuitive, but also makes it possible for them to maintain more consistent levels of performance and performance quality regardless of how they are being accessed or where they are being used at any given time [22].

IV. PROS AND CONS FOR IOT APPLICATIONS WHILE USING FEDERATED LEARNING

There are many potential advantages and disadvantages to using IoT applications that rely on federated learning. Here are a few pros and cons to consider:

A. Benefits of using federated learning for IoT applications

It is more efficient and enables faster processing of the data collected by IoT devices than a centralized approach, enabling much faster processing of data gathering from a large number of devices. - Federated learning allows training data to be retained on the devices that collect it, rather than sending the data to the cloud for processing. This reduces the amount ofdata that needs to be sent over the network, reducing bandwidth usage and lowering the cost of data transmission. - In centralized systems, all of the data is collected and stored in one location, so if there is a problem with the network or with the data storage hardware, all the collected data is at risk. But in a federated system, the data is stored only on the devices where it is generated, so it is less vulnerable to interruptions in the network or to problems with data loss and storage hardware. - Using federated learning means that enterprises can reduce the cost of data storage because the data is not centrally stored but rather stored on each individual device [23]. This allows organizations to save a substantial amount of money by not having to pay for cloud-based storage capabilities. - Because federated learning reduces the amount of data that needs to be transferred over the network, it can improve network security by reducing the amount oftraffic that passes through the network. This will improve the overall security of the network by ensuring that there is less risk of unauthorized access to the network and that the risk of network disruption is minimized. - Data collected by IoT devices needs to be processed quickly in order to derive meaningful insights from the data [24]. By using the decentralized approach of federated learning, data can be analyzed and processed much more quickly than is possible with a centralized system.

B. Disadvantages of using federated learning for IoT applications

The distributed architecture of federated learning increases the complexity of the system and requires additional technical support and monitoring to ensure that the system is functioning correctly. A centralized system is easier to manage because all of the data is gathered in a single location, whereas a federated system requires additional resources to ensure that each node in the system is functioning properly. - Although the decentralized architecture of federated learning offers the advantage of improved data security and reduced network congestion, there are disadvantages to this approach as well [25]. Because the system is distributed across multiple nodes, it becomes more difficult to access the data if the nodes become disconnected or otherwise inoperable. There is also a greater chance of human error because of the increased complexity associated with the distributed architecture. This means that there is a greater risk of errors and inconsistencies in the data that is being collected and analyzed by the system. - Federated learning systems are vulnerable to security breaches when individuals attempt to gain access to the system through unauthorized means [26]. For example, an individual may be able to bypass the security of the system and gain access to the data stored on it if it is stored in an insecure location or is transmitted over an unsecured network. - In federated learning systems, the accuracy of results depends on the inputs that are provided to the system. If the input data is inaccurate or incomplete, the results will be unreliable and will not be accurate representations of the underlying data. It is also possible that the system will produce incorrect results if the inputs change over time and do not match the actual behavior of the device being measured. - Federation learning systems work best when the participating nodes have direct access to the devices that are being measured and the data that is generated as a result of the measurements. It is much more difficult to collect data from multiple remote devices than to collect data from a single device located in the laboratory where the researchers can access it directly.

V. RECOMMENDATIONS FOR IOT APPLICATIONS WHILE USING FEDERATED LEARNING

IoT has the potential to change our lives in innumerable ways. From increased efficiency in manufacturing to improved safety and security, IoT is poised to have a profound impact on the way we live and work. Unfortunately, many of these benefits will not be realized until the security and privacy concerns around IoT devices are addressed. Unlike traditional computing devices like laptops and smartphones, IoT devices often lack the necessary security controls to protect user data and ensure data privacy. And many IoT devices are designed to operate without an internet connection, which can make device management and updates difficult [27].

Recognizing the urgent need to address these security concerns, federal agencies including the National Institute of Standards and Technology (NIST) and the Software Engineering Institute (SEI) at Carnegie Mellon University have established a consortium focused on improving the security of IoT devices through the creation of standards and guidelines. In 2018, these agencies launched the Federated Learning framework, an open-source project that builds on the work of other consortiums and community organizations to create a standardized architecture for securely deploying and managing connected devices. The goal of the project is to develop a series of best practices that will improve the security and privacy of connected devices while reducing the cost and complexity associated with managing IoT deployments [28].

With a standardized framework for securely deploying and managing IoT devices, the federal government will be better positioned to harness the benefits that IoT has to offer and deliver those benefits to the American people. Here are a few recommendations for implementing this framework in your own organization:

- Develop and implement standards and best practices that ensure the safety and security of all devices in your IoT deployment.
- Implement strong cybersecurity practices to protect your devices and data from unauthorized access or malicious activity.
- Develop a process for regularly updating your devices to ensure that they remain secure and up-to-date at all times.
- Consider incorporating artificial intelligence and machine learning capabilities into your IoT devices to help improve their performance and manage them more efficiently [29].
- Conduct periodic reviews and audits to ensure that your devices are secure and perform as expected.

DOI: 10.17148/IJARCCE.2023.12213

- Leverage third-party security vendors to test and validate the security of your devices and systems on a regular basis.

The NIST-IoT program is a collaborative effort that brings together federal agencies, industry partners, and other stakeholders to advance the development of secure IoT technologies and promote the use of these technologies in government. This program is part of the larger Cybersecurity Framework initiative under Executive Order 13636, which was developed to enhance the federal government's ability to protect federal information and information systems.

VI. CONCLUSION

Federated learning has emerged as a critical privacy-preserving machine learning paradigm. Because of its distributed nature, it is well suited for deployment in a wide range of Internet of Things applications. In this article, we summarized the benefits of federated learning for IoT and elaborated on some key applications. We also identified and detailed some key research challenges for FLon IoT's future development. We hope that this article will draw more attention to the field of federated learning and inspire the creativity and innovation required to build the FL-IoT ecosystem.

REFERENCES

[1] Dash, B., Sharma, P., & Ali, A. (2022). Federated Learning for Privacy-Preserving: A Review of PII Data Analysis in Fintech. International Journal of Software Engineering & Applications (IJSEA), 13(4).

[2] T. Zhang, C. He, T. Ma, L. Gao, M. Ma, and S. Avestimehr, "Federatedlearning for internet of things: A federated learning framework for on-device anomaly data detection," ArXiv, vol. abs/2106.07976, 2021

[3] Mohammed, M. A., Mohammed, M. A., & Mohammed, V. A. (2022). Impact of Artificial Intelligence on the Automation of Digital Health System. International Journal of Software Engineering & Applications (IJSEA), 13(current). https://doi.org/10.5121/ijsea.2022.13602

[4] Thatikonda, R., Padthe, A., Vaddadi, S. A., & Arnepalli, P. R. (2023). Effective secure data agreement approach-based cloud storage for a healthcare organization. International Journal of Smart Sensor and Adhoc Network., 60–70. https://doi.org/10.47893/ijssan.2023.1232

[5] Dash, B., & Ansari, M. F. (2022). Self-service analytics for data-driven decision making during COVID-19 pandemic: An organization's best defense. Academia Letters, 2.

[6] Vaddadi, S. A., Arnepalli, P. R., Thatikonda, R., & Padthe, A. (2022). Effective malware detection approach based on deep learning in Cyber-Physical Systems. International Journal of Computer Science and Information Technology, 14(6), 01–12. https://doi.org/10.5121/ijcsit.2022.14601

[7] Gonaygunta, Hari (2023) "MACHINE LEARNING ALGORITHMS FOR DETECTION OF CYBER THREATSUSING LOGISTIC REGRESSION," International Journal of Smart Sensor and Adhoc Network: Vol. 3: Iss. 4,Article 6.Available at: https://www.interscience.in/ijssan/vol3/iss4/6

[8] D. Liu, T. Miller, R. Sayeed, and K. D. Mandl, "Fadl: Federated-autonomous deep learning for distributed electronic health record,"ArXiv, vol. abs/1811.11400, 2018.

[9] B. Yuan, S. Ge, and W. Xing, "A federated learning framework forhealthcare iot devices," ArXiv, vol. abs/2005.05083, 2020

[10] B. Liu, B. Yan, Y. Zhou, Y. Yang, and Y. Zhang, "Experimentsof federated learning for covid-19 chest x-ray images," ArXiv, vol.abs/2007.05592, 2020 [11]. Dash, B., Sharma, P., & Ali, A. (2022). Federated Learning for Privacy-Preserving: A Review of PII Data Analysis in Fintech. International Journal of Software Engineering & Applications (IJSEA), 13(4).

[11] Mohammed, M. A., Mohammed, M. A., &amp; Mohammed, V. A. (2022). Application of Data Analytics to Improve Patient Care: A Systematic Review. International Research Journal of Engineering and Technology, 9(11), 197–203.

[12] Vaddadi, S. A., Arnepalli, P. R., Thatikonda, R., & Padthe, A. (2022). Effective malware detection approach based on deep learning in Cyber-Physical Systems. International Journal of Computer Science and Information Technology, 14(6), 01–12. https://doi.org/10.5121/ijcsit.2022.14601 [13]. Dash, B., Sharma, P., & Ali, A. (2022). Federated Learning for Privacy-Preserving: A Review of PII Data Analysis in Fintech. International Journal of Software Engineering & Applications (IJSEA), 13(4).

[13] E. Bertino and N. Islam, "Botnets and internet of things security,"Computer, vol. 50, no. 2, pp. 76–79, 2017.

[14] T. Nguyen, S. Marchal, M. Miettinen, H. Fereidooni, N. Asokan, and A.-R. Sadeghi, "D˙Iot: A federated self-learning anomaly detection systemfor iot," 07 2019, pp. 756–767

DOI: 10.17148/IJARCCE.2023.12213

[15] Y. Fan, Y. Li, M. Zhan, H. Cui, and Y. Zhang, “Iotdefender: A federatedtransfer learning intrusion detection framework for 5g iot,” in 2020 IEEE14th International Conference on Big Data Science and Engineering(BigDataSE), 2020, pp. 88–95

[16] L. D. Xu, E. L. Xu, and L. Li, “Industry 4.0: state of the art and futuretrends,” International Journal of Production Research, vol. 56, no. 8,pp. 2941–2962, 2018

[17] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, “Blockchain andfederated learning for privacy-preserved data sharing in industrial iot,”IEEE Transactions on Industrial Informatics, vol. 16, no. 6, pp. 4177–4186, 2020.

[18] Dash, B., & Sharma, P. (2022, December). Impact of Digitalization on Shaping Consumer-Centered Smart Healthcare System-A Comprehensive Study. In 2nd International Conference on Advances in Computing & Information Technologies (CACIT 2022).

[18] Ansari, M. F. (2022). A quantitative study of risk scores and the effectiveness of AI-based Cybersecurity Awareness Training Programs. International Journal of Smart Sensor and Adhoc Network, 3(3), 1..

[19] T. Li, A. K. Sahu, A. Talwalkar and V. Smith, "Federated learning: Challenges methods and future directions", IEEE Signal Process. Mag., vol. 37, no. 3, pp. 50-60, May 2020. [20]. Sharma, P. Sustainable Smart Cities and Associated Risks–A Review.

[20] Ansari, M. F., Dash, B., Sharma, P., & Yathiraju, N. (2022). The Impact and Limitations of Artificial Intelligence in Cybersecurity: A Literature Review. International Journal of Advanced Research in Computer and Communication Engineering.

[21] Ansari, M. F., Panigrahi, A., Jakka, G., Pati, A., & Bhattacharya, K. (2022, November). Prevention of Phishing attacks using AI Algorithm. In 2022 2nd Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON) (pp. 1-5). IEEE.

[22] Sharma, P. (2023). Cloud Computing for Supply Chain Management and Warehouse Automation: A Case Study of Azure Cloud.

[23] L. Ashiku and C. Dagli, "Cybersecurity as a centralized directed system of systems using SoS explorer as a tool", Proc. 14th Annu. Conf. Syst. Syst. Eng. (SoSE), pp. 140-145, 2019.

[24] Sharma, P., Chetti, P., Dash, B., & Ansari, M. (2023). Data Modeling Best Practices Key to Data Mining and Data Standardization. Available at SSRN 4337595.

[25] L. Ashiku and C. Dagli, "Cybersecurity as a centralized directed system of systems using SoS explorer as a tool", Proc. 14th Annu. Conf. Syst. Syst. Eng. (SoSE), pp. 140-145, 2019.

[26] J. Ren, H. Wang, T. Hou, S. Zheng and C. Tang, "Federated learning-based computation offloading optimization in edge computing-supported Internet of Things", IEEE Access, vol. 7, pp. 69194-69201, 2019.

[27] Mohammed, M. A., & Sharif, M. H. U. (2022). A literature review of financial losses statistics for cyber security and future trend. A Literature Review of Financial Losses Statistics for Cyber Security and Future trend16, 16(2).

[28] Arnepalli, P. R. R., Vaddadi, S. A., Padthe, A., & Thatikonda, R. (2022). Impact of emerging technology to improve the network aggregation for Business Organizations. IJARCCE, 11(12). https://doi.org/10.17148/ijarcce.2022.111204.

[29] Mohammed, M. A., & Sharif, M. H. U. (2022). A literature review of financial losses statistics for cyber security and future trend. A Literature Review of Financial Losses Statistics for Cyber Security and Future trend16, 16(2). https://doi.org/10.30574/wjarr.2022.15.1.