International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering
Vol. 12, No. 1, 2025, pp. 7-11.

ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

Federated Learning for Healthcare Privacy-Preserved Artificial Intelligence in Distributed Systems

Aruna Pal [a], Devina Chhabra [b]

[a] *Computational Epidemiology Research Centre, Asansol Public Health Tech Institute, India.*
[b] *Computational Epidemiology Research Centre, Asansol Public Health Tech Institute, India.*

Abstract

Integrating the Internet of Things (IoT) is prominently embraced in healthcare and is referred to as the Medical IoT (MIoT). MIoT is transforming healthcare by offering numerous advantages for patients and healthcare professionals. The utilization of MloT is rapidly increasing, producing substantial volumes of IoT data that necessitate thorough analysis to derive significant insights. This has resulted in the implementation of Artificial Intelligence (AI) methods, including Machine Learning (ML) and Deep Learning (DL) algorithms, to comprehend the significance of the underlying health information, with the process of learning typically occurring in cloud or healthcare systems. The rapid proliferation of IoT sensors and the extensive distribution of private MloT data sets render centralized learning AI systems increasingly challenging to implement for these jobs. In this context, Federated Learning (FL) is becoming recognized as a viable device learning approach without transferring sensitive and confidential information to a central server. The terminal gear and the central server in FL only exchange learning model upgrades to maintain the confidentiality of sensitive data. Despite its emergence as a viable research domain, no contemporary studies have been undertaken. This paper synthesizes recent work and advancements in FL to enhance FL-driven MIoT medical applications and services. This research enables stakeholders in both academia and business to comprehend the advantages of the most sophisticated privacy-preserving MloT platforms utilizing FL.

Keywords: Federated Learning, Healthcare, Artificial Intelligence, Privacy.

DOI: 10.71086/IAJSE/V12I1/IAJSE1202
Received: 11 December 2024/Revised: 14 January 2025/Accepted: 11 February 2025/Published: 31 March 2025

Introduction

The transformation in the healthcare sector has generated numerous Medical Internet of Things (MIoT) applications (Akhtar et al., 2021), each capable of accessing vast data sets and necessitating thorough evaluation. Over time, MIoT, or Internet of Things (IoT)-driven technologies in medical care (Ajagbe et al., 2022), have demonstrated significant potential across various medical applications, including identifying illnesses, patient condition tracking, remote health surveillance, smartwatches, health clubs, urgent care, elderly care, and epidemic administration. Daily, influenced by circumstances such as the current COVID-19 epidemic, the increase in persistent illnesses, and the aging population, the utilization of these MIoT services has surged, resulting in the development and accumulation of substantial health information. In a conventional MIoT configuration, patient pathologies and ecological data are initially gathered by MIoT gadgets and transmitted to the cloud via base stations to train statistical machine learning algorithms for data interpretation and prediction, commonly called Artificial Intelligence (AI) (Ajagbe et al., 2022). Previously, rudimentary data transaction frameworks were employed.

In applying these AI technologies, one entity gathers data and communicates it to another entity, which then processes and integrates the data. An outside organization will utilize collected information to create models for others. The challenge is that the information is sequestered like islands, and the research is restricted from aggregating, integrating, and utilizing data from disparate sites for AI processing. AI practitioners encounter significant challenges in legally addressing data dispersion and isolation issues.

The exponential proliferation of MloT activities has resulted in the corresponding escalation of technology, culminating in the generation and accumulation of substantial data volumes (Allioui & Mourdi, 2023). The practicality of clinical research is frequently obstructed by data communication and availability challenges, particularly concerning privacy concerns, as health data typically contains Personally Identifiable Information (PII) and sensitive unhealthy details specific to patients, which also contravene rules and regulations. The Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Rules (GDPR) advocate for data confidentiality and PII safeguarding.

Recent statistics indicate that the proliferation of IoT devices will yield an economic boost between 5 to 13 trillion US dollars. At the same time, the IoT healthcare sector is projected to attain approximately 530 billion dollars by 2025, thereby offering a forecast of the information volume generated by the expanding healthcare in the natural world. To address this issue, including the previously described fragmentation of information and isolation challenges, FL facilitates distributed optimization approaches for privacy-preserving data analysis while ensuring data security and independence (Wei et al., 2021). Over time, the enhanced proliferation of IoT networks and escalating privacy concerns in standard IoT settings jeopardize the privacy of collected information when conventional AI technologies depend on collecting and consolidating any information in a singular location for further analysis. This form of federated learning appears as a collaborative and dispersed AI approach that facilitates teaching decentralized IoT devices sans data sharing and safeguarding data privacy.

Background

2.1 IoT Healthcare

IoT medicine involves collecting patient medical data through IoT devices, such as sensors, which transmit the information to the cloud for tracking, analysis, and remote customization. Individuals are more inclined to embrace disease monitoring and treatment at home through IoT devices (Mohammed et al., 2020). For instance, individuals can utilize smartwatches to consistently monitor health metrics (e.g., blood glucose, cholesterol levels, etc.) and their physical activity. The researchers also focused on handling the acquired user health information. In the context of severe illnesses such as Alzheimer's disease, IoT healthcare demonstrates significant benefits over traditional screening evaluations and neuropsychological evaluations, as it does not necessitate complex medical devices and facilitates remote therapy.

2.2 Security Risks Encountered by IoT Healthcare

The IoT healthcare system remains a significant security hazard (Affia et al., 2023). The weaknesses of smart homes are increasingly uncovered, including data leaks and security flaws in interactions. The VMask Challenge can compromise spoken interfaces by generating adversarial instances to manipulate voice commands and control smart IoT devices. Secondly, the health information users provide is susceptible to theft and manipulation during transmission by a hacker. The assailant can execute a man-in-the-middle assault and a side-channel attack to jeopardize the safety of the devices (Thankappan et al., 2022). The Wind Talker can assault and expropriate private material

using channel state details derived from the Wi-Fi signal. The research addresses privacy-related concerns, deferring other safety concerns for further investigation.

Proposed Federated Learning (FL) based Healthcare Application

3.1 Objective of Design

Compared to current Alzheimer's Disease (AD) detection systems utilizing IoT healthcare, AD Identifier (ADI) is engineered to accomplish the following objectives.

1) Simple Deployment: In ADI, the user needs to engage in conversations with commercialized voice assistants (e.g., Amazon Echo) without necessitating additional analysis or installation of complex devices by the user.

2) High Effectiveness: ADI must attain accuracy superior to current acoustic or linguistic-based AD detection methodologies. It must also attain minimal time costs to implement ADI with the growing user base.

3) Private-Preserving: The health-related private data, encompassing the user's raw information and attributes transmitted for categorization, must remain undisclosed during the data flow transfer of identifying anomalies.

3.2 Threat Model

This article examines attackers who can access the link between the user's IoT gadgets and the information evaluation cloud by executing a man-in-the-middle assault. As seen in Fig. 1, whatever data is supplied by each element of the ADI, including the processed characteristic, is susceptible to acquisition by an attacker. ADI should implement a privacy-preserving detection framework (i.e., FL in this study) and data-oriented privacy protection mechanisms (e.g., encryption). This article does not presume the attacker can infiltrate the user's network and retrieve the original data. Directly targeting the user's IoT device is a significant assumption beyond the present research's scope.

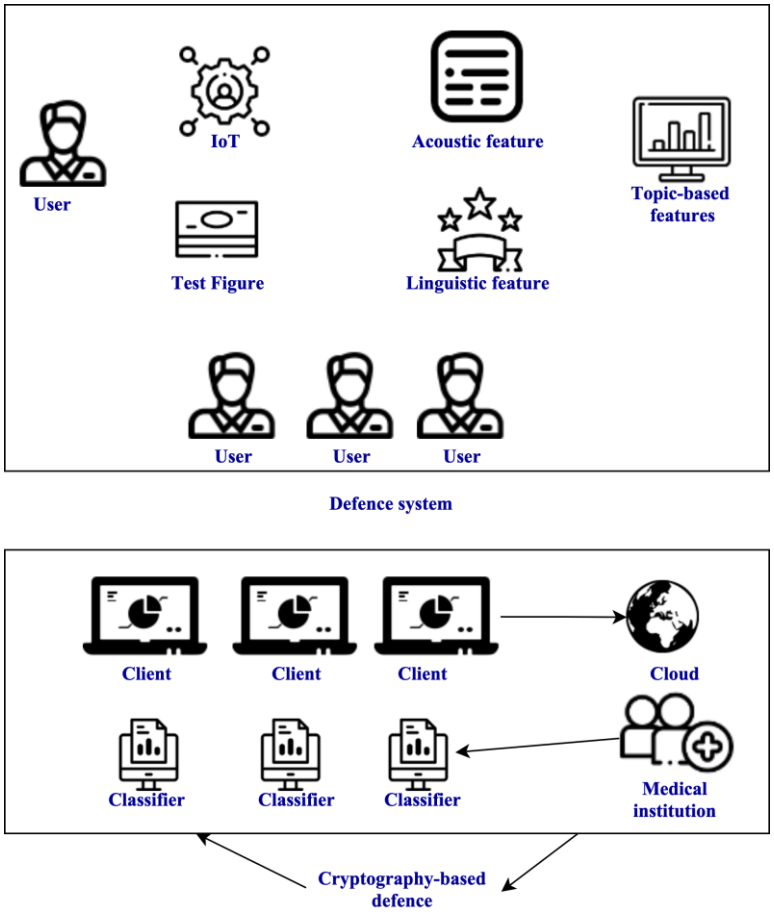


Fig. 1: System architecture

System Overview

To fulfill the specified design objectives and tackle security difficulties, the research has developed ADI with a three-tier architecture comprising the user, detecting client, and cloud levels. ADI shall consist of four components. At the user level, the Data Collecting Module prompts the user to provide speech samples for Alzheimer's disease identification. At the same time, the Feature Extraction Unit derives characteristics from both auditory and language dimensions. The FL-Based Decision Unit utilizes the FL structure, wherein user characteristics are initially allocated to various detection customers, followed by the optimization of AD categorization through the interaction among the detection consumers and the cloud service. The FL architecture facilitates the retention of raw data at the user stage. At the same time, differential privacy safeguards the transmission of information between the user and the detecting client levels. The Concurrent Privacy-Preserving Aggregation Unit is ultimately designed to guarantee the validity and secrecy of interactions across the detecting clients' and the cloud levels.

Discussion

4.1 Feature Selection

The duration allocated to feature generation constitutes a significant portion of the anomaly detection process. Therefore, it is imperative to identify more effective and beneficial features to enhance detection speed while preserving high precision in the future.

4.2 Feature Sizes

The research does supplementary experiments by configuring the number of dimensions of the linguistic characteristic. It is important to note that, throughout the trials, in addition to fi, the settings of the ADI adhere to the specified configuration.

It is noted that a rise in the dimensionality of the training characteristics enhances their identification efficacy progressively. Once the linguistic feature dimension exceeds 50, efficiency experiences a slight improvement; however, the time expenditure per user escalates significantly. To balance time complexity and success, the research established the measure of the language characteristic data at 50 in this investigation.

4.3 Client Identification Number in Florida

The accuracy significantly declines when the customer number exceeds three. Due to the tradeoff between speed and time expenses, the research established the client number at 3, resulting in each client including 33 users, a standard setup for federated learning. The general outcome with client numbers ranging from 2 to 9 is satisfactory, achieving approximately 80% following the implementation of all security measures.

4.4 Selection of Classification

The classifier's effectiveness is contingent upon the attributes of the generated characteristics. The research presents several attributes for the identical Challenge database. It employs distinct classification algorithms, such as Probability Linear Discriminant Assessment (PLDA) and Support Vector Machines (SVM), to attain the highest efficiency in IoT medicine. Due to the feature variations, studies employ various classifiers (i.e., PLDA and SVM) on the identical AD detection database. This paper presents the innovative feature derived from the ADReSS Initiative database. The efficacy of Logistic Regression (LR) surpasses that of other classifiers, such as SVM and Naive Bayes (NB), due to the qualities of the retrieved features. All three classifiers achieve accuracy exceeding 75%, demonstrating the feature's efficacy.

4.5 Data Set

Due to the high-level encryption of health-related user information, the only fully released large-scale database for AD diagnosis is the Challenge database. The research was assessed based on prior studies utilizing the Competition database. The efficacy of more unpublished records within the IoT healthcare context is now under investigation and will be addressed in subsequent studies. The research asserts that ADI is both practical and beneficial for alternative datasets. The finding indicates that the disparities in auditory and linguistic aspects between healthy and people with AD are substantial, representing a widespread problem.

Conclusion

This paper examines federated learning in the MIoT, a unique learning paradigm that offers enhanced privacy for learning data relative to conventional AI methods. The data security and confidentiality legislation prohibits parties

from directly accessing personally identifiable data to facilitate collaborative research and development on shared challenges between medical academics and business. FL enables various stakeholders to gain insight from confidential medical information while fostering a cooperative learning environment that safeguards the privacy of the information that underlies it. This diverges significantly from conventional AI learning, wherein data is centralized. This federated learning paradigm will imminently be implemented across all facets of IoT-based healthcare, including diagnosis of illnesses, clinical research, and medication discovery, to extract insights from data and enhance IoT-based health to a new echelon. The study has examined the distinctive attributes of FL and the associated challenges in the MIoT, as the application of FL to healthcare is expanding swiftly due to increasing demand. The research has analyzed the taxonomies, outlining the pertinent works, their primary contributions, and distinctive characteristics.

The research has identified significant concerns that merit deeper investigation and additional study. Collaboration between the two sectors might be advantageous in addressing these difficulties. Given that this is the inaugural review in this domain, to the greatest extent of the understanding, this study will serve as a valuable reference for FL in the MIoT field and facilitate subsequent investigations.

References

Akhtar, N., Rahman, S., Sadia, H., & Perwej, Y. (2021). A holistic analysis of Medical Internet of Things (MIoT). *Journal of Information and Computational Science*, *11*(4), 209-222.

Ajagbe, S. A., Awotunde, J. B., Adesina, A. O., Achimugu, P., & Kumar, T. A. (2022). Internet of medical things (IoMT): applications, challenges, and prospects in a data-driven technology. *Intelligent Healthcare: Infrastructure, Algorithms and Management*, 299-319. https://doi.org/10.1007/978-981-16-8150-9_14

Allioui, H., & Mourdi, Y. (2023). Exploring the full potentials of IoT for better financial growth and stability: A comprehensive survey. *Sensors*, *23*(19), 8015. https://doi.org/10.3390/s23198015

Wei, K., Li, J., Ding, M., Ma, C., Su, H., Zhang, B., & Poor, H. V. (2021). User-level privacy-preserving federated learning: Analysis and performance optimization. *IEEE Transactions on Mobile Computing*, *21*(9), 3388-3401 https://doi.org/10.1109/TMC.2021.3056991.

Mohammed, M. N., Desyansah, S. F., Al-Zubaidi, S., & Yusuf, E. (2020, February). An internet of things-based smart homes and healthcare monitoring and management system. In *Journal of physics: conference series* (Vol. 1450, No. 1, p. 012079). IOP Publishing. https://doi.org/10.1088/1742-6596/1450/1/012079

Affia, A. A. O., Finch, H., Jung, W., Samori, I. A., Potter, L., & Palmer, X. L. (2023). IoT health devices: exploring security risks in the connected landscape. *IoT*, *4*(2), 150-182. https://doi.org/10.3390/iot4020009

Thankappan, M., Rifà-Pous, H., & Garrigues, C. (2022). Multi-Channel Man-in-the-Middle attacks against protected Wi-Fi networks: A state of the art review. *Expert Systems with Applications*, *210*, 118401. https://doi.org/10.1016/j.eswa.2022.118401