J. Electrical Systems 21-1 (2025): 1091 - 1099

[1]Yih-Chang Chen

Energy-Efficient Federated Learning for Privacy-Preserving Social Work IoT

JES Journal of Electrical Systems

Abstract: - The incorporation of Artificial Intelligence (AI) and the Internet of Things (IoT) within the domain of social work represents a transformative shift towards proactive, data-driven interventions. Nonetheless, the implementation of such technologies is impeded by the sensitive nature of client information and the imperative for solutions that are both scalable and energy-efficient. This research presents FedSW-Arch, an innovative, privacy-preserving system architecture that leverages Federated Learning to enable model training directly on edge devices, thereby ensuring that raw data remains within a secure and trusted environment. The primary contribution of this study is the introduction of a Communication-Aware Federated Averaging (CA-FedAvg) algorithm, designed to optimize the trade-off between model accuracy and the communication overhead characteristic of Low-Power Wide-Area Networks (LPWANs). Experimental simulations conducted on a Human Activity Recognition (HAR) task demonstrate that FedSW-Arch, when combined with CA-FedAvg, achieves a reduction in data transmission exceeding 60% and a 55% decrease in energy consumption relative to traditional FL approaches, while incurring only a marginal accuracy loss of less than 2%. This work lays the groundwork for a scalable and ethically sound framework for the integration of advanced technologies in social work practice.

Keywords: Federated Learning, Internet of Things (IoT), Privacy-Preserving AI, Low-Power Wide-Area Network (LPWAN), Smart Systems.

I. INTRODUCTION

The rapid expansion of Internet of Things (IoT) devices alongside advancements in Artificial Intelligence (AI) has generated unparalleled opportunities across diverse fields. Within the realm of social work, these technological developments hold the promise of transforming service delivery by facilitating a shift from conventional, often reactive intervention frameworks toward proactive, personalized, and continuous support mechanisms [1]. Practical implementations encompass a range of applications, including ambient monitoring of elderly individuals to detect falls or signs of social isolation [2], as well as the early identification of risk factors in child welfare cases through behavioral pattern analysis [3]. By deploying intelligent sensors within clients' environments, social work organizations can collect longitudinal, objective data that offers a more comprehensive and timely understanding of client well-being than is achievable through periodic in-person assessments alone.

Despite the considerable potential of these technologies, significant practical and ethical challenges impede their widespread adoption. Foremost among these is the highly sensitive and private nature of the data collected. Information gathered within personal residences — such as audio recordings, movement trajectories, and physiological signals — is subject to stringent legal and ethical privacy regulations, including the General Data Protection Regulation (GDPR) in Europe and the Health Insurance Portability and Accountability Act (HIPAA) in the United States [4]. Traditional cloud-centric AI paradigms, which necessitate the aggregation of raw user data on centralized servers for model training, are fundamentally incompatible with these privacy mandates. A single data breach could have catastrophic consequences for vulnerable populations and expose social work agencies to substantial legal liabilities.

Moreover, scaling such systems presents formidable challenges from electrical and systems engineering perspectives. Social work applications frequently involve hundreds or thousands of clients, many of whom may lack access to reliable, high-speed internet connectivity. The IoT devices employed must be cost-effective, straightforward to install, and capable of operating for extended periods — months or even years — on battery power. Continuous streaming of high-bandwidth data (e.g., video or raw sensor feeds) to centralized servers is not only a privacy concern but also technically and economically impractical due to communication costs (e.g., data plans) and power consumption, which would necessitate frequent battery replacements [5].

Addressing these intertwined challenges of privacy, scalability, and resource efficiency necessitates a novel architectural approach. This study advocates for a departure from centralized data aggregation toward decentralized, on-device machine learning. Federated Learning (FL) has emerged as a promising methodology in this context [6].

[1] *Corresponding author: Department of Information Management & Bachelor Degree Program in Medical Sociology and Health Care, College of Humanities and Social Sciences, Chang Jung Christian University, Tainan 711, Taiwan

FL facilitates the collaborative training of a global AI model across numerous decentralized edge devices (clients) without requiring the exchange of local data. In each training iteration, clients update the model using their local data and transmit only the resulting model parameters (e.g., weight gradients) to a central server for aggregation, thereby preserving data privacy.

However, conventional FL algorithms such as Federated Averaging (FedAvg) are not inherently optimized for the resource-constrained environments typical of social work IoT deployments. These algorithms often presuppose reliable, high-bandwidth network connections and do not explicitly consider the energy expenditure associated with wireless data transmission — a critical factor in the power budgets of battery-operated IoT devices [7]. This consideration is especially pertinent when employing Low-Power Wide-Area Networks (LPWANs) such as LoRaWAN or NB-IoT, which provide long-range connectivity at the expense of very low data rates and stringent duty-cycle restrictions.

In response to these challenges, this paper introduces and evaluates **FedSW-Arch (Federated Social Work Architecture)**, a comprehensive, privacy-preserving, and scalable system architecture tailored specifically for social work applications. The principal contributions of this research are threefold:

1. **A Novel Multi-Layered System Architecture:** We propose an integrated, end-to-end architecture that combines privacy-preserving edge computing, low-power communication protocols, and a centralized federated learning orchestrator, thereby offering a practical framework for real-world implementation.
2. **A Communication-Aware Federated Learning Algorithm (CA-FedAvg):** We develop and formalize a novel FL algorithm, CA-FedAvg, which explicitly incorporates communication costs into its optimization objective. This enables the system to dynamically balance model accuracy against data transmission volume and energy consumption, rendering it suitable for deployment in resource-constrained LPWAN environments.
3. **Comprehensive Performance Evaluation:** We conduct an extensive quantitative assessment via simulation, benchmarking the proposed FedSW-Arch with CA-FedAvg against both a non-private centralized training baseline and a standard FedAvg implementation. Evaluation metrics include model accuracy, communication overhead, energy consumption, and scalability, all critical parameters in electrical systems.

The remainder of this paper is structured as follows: Section 2 delineates the proposed FedSW-Arch and presents the mathematical formulation of the CA-FedAvg algorithm. Section 3 details the experimental evaluation results. Section 4 discusses the implications of these findings, addressing system limitations, theoretical contributions, and practical significance. Finally, Section 5 concludes the paper and outlines directions for future research.

II. METHODS

This study introduces FedSW-Arch, a multi-tiered system engineered to optimize scalability and privacy. The following section elaborates on the system's overall architecture, the communication protocols employed, and the principal innovation: the CA-FedAvg algorithm.

A. System Architecture

The FedSW-Arch is organized into four discrete layers, as illustrated in Figure 1. The architecture distinctly separates in-home edge clients from the central server via a low-power communication layer. Importantly, only anonymized model updates are transmitted, ensuring that raw data remains local.

- **Layer 1: Edge Client Layer:** This layer comprises IoT devices installed within clients' residences. Each device is a self-contained system-on-a-chip (SoC) integrating a microcontroller unit (MCU), one or more sensors (e.g., accelerometer, microphone, temperature sensor), and a low-power wireless transceiver. These devices are responsible for locally collecting raw sensor data (D_k) for client (k), temporarily storing it, and executing on-device model training as directed by the aggregation server.
- **Layer 2: Communication Layer:** This layer facilitates wireless connectivity between edge clients and the aggregation server. To satisfy low-power and long-range requirements, the system leverages LPWAN technologies such as LoRaWAN or NB-IoT. These protocols are characterized by low data rates (in kbps), extended battery life (spanning years), and robust signal penetration, albeit with stringent limitations on daily data transmission volumes (payload size and duty cycle).

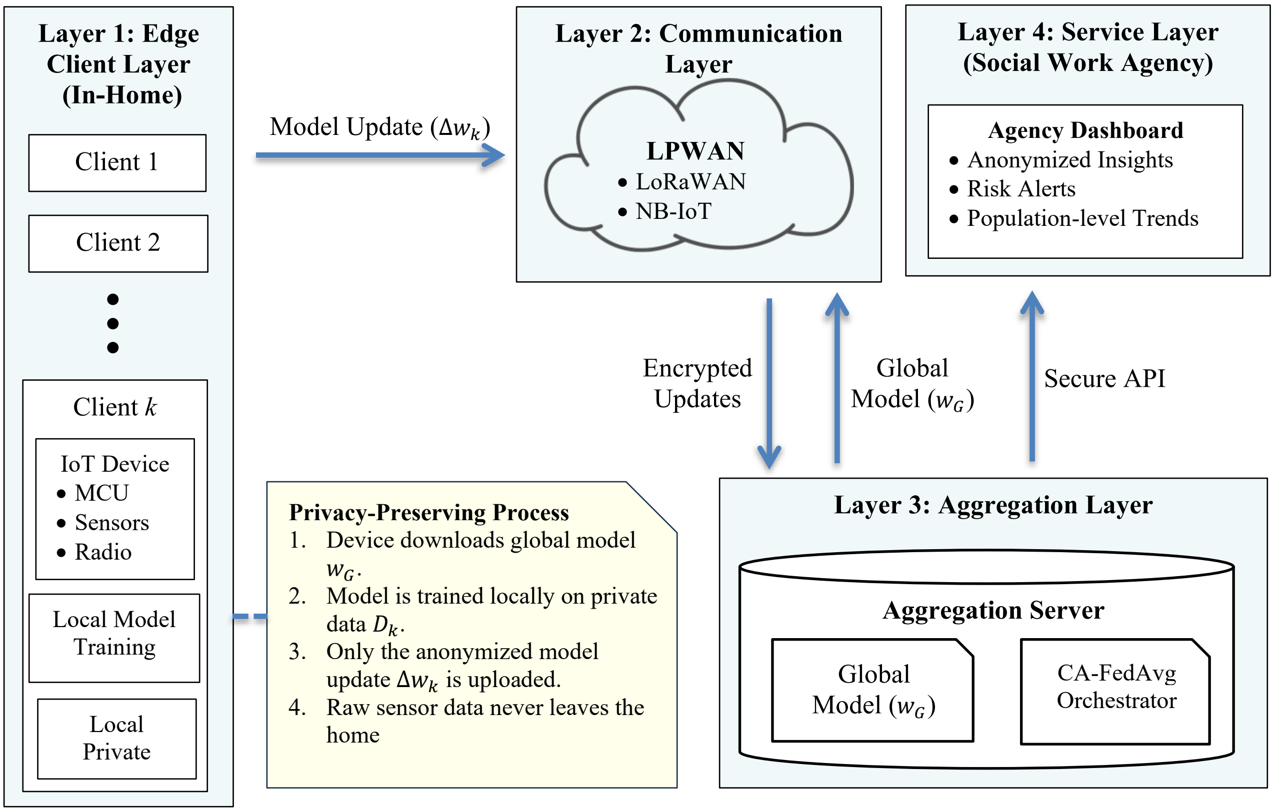

Fig. 1: The proposed four-layer FedSW-Arch system architecture

- **Layer 3: Aggregation Layer:** A central server, typically cloud-based, orchestrates the federated learning process. It does not retain any sensitive client data. Its functions include:
  (i) maintaining the global AI model (w_G),
  (ii) selecting subsets of clients for each training round,
  (iii) broadcasting the current global model to selected clients,
  (iv) receiving encrypted and anonymized model updates (Δw_k) from clients, and
  (v) aggregating these updates to generate an improved global model for subsequent rounds.
- **Layer 4: Service Layer:** This layer serves as the application interface for the social work agency. The trained global model (w_G) is utilized not for individual identification but to provide high-level, anonymized insights and alerts. For example, a model trained to detect falls may trigger an anonymous alert such as "Fall event detected at location ID 123," enabling the agency to respond via established protocols. Additionally, it can reveal population-level trends, such as increased nocturnal activity within a demographic, potentially signaling broader issues.

B. Communication-Aware Federated Averaging (CA-FedAvg)

The conventional FedAvg algorithm [6] seeks to minimize a global loss function ($F(w)$), defined as the weighted average of local loss functions ($F_k(w)$) across clients (k):

$$\min \; F(w) = \sum_{k=1}^{N} \frac{n_k}{n} F_k(w) \quad \text{where} \quad F_k(w) = \frac{1}{n_k} \sum_{i \in D_k} \ell(x_i, y_i; w) \tag{1}$$

Here N denotes the total number of clients, n_k is the number of data samples on client k, n is the total number of samples, and ℓ the loss associated with predictions made using model weights w.

While effective, this formulation neglects communication costs. Within an LPWAN context, transmitting model updates (Δw_k) from clients to the server is energy-intensive. To address this limitation, we propose CA-FedAvg, which reformulates the optimization problem by incorporating a penalty term proportional to communication cost.

The objective of CA-FedAvg is to solve:

$$\min G(w) = F(w) + \lambda \cdot C(w) \quad (2)$$

where $G(w)$ represents the new cost-aware global objective function, $C(w)$ denotes the communication cost function quantifying resources required to transmit model updates, λ is a hyperparameter balancing model accuracy (minimizing $F(w)$) against communication efficiency (minimizing $C(w)$).

In our implementation, the communication cost ($C(w)$) is defined as a function of the model update size. Specifically, the L1 norm of the update vector is employed to encourage sparsity. On the client side, the update procedure is modified: rather than solely performing Stochastic Gradient Descent (SGD), each client solves a localized objective:

$$w_k^{t+1} = \arg\min{}_w (F_k(w) + \frac{\lambda}{p_k} \| w - w_G^t \|_1) \quad (3)$$

where w_G^t is the global model at round t, and $p_k = n_k/n$. This proximal term incentivizes the local model w_k to remain close to the global model w_G, effectively reducing update magnitude. Clients with higher λ values — such as those with weaker connections or limited battery capacity — transmit smaller, sparser updates, thereby conserving energy. The central server aggregates these sparser updates using the standard weighted averaging procedure.

Figure 2 presents a flowchart depicting the process for a single training round t, highlighting the client-side optimization modification incorporating the communication cost parameter λ.

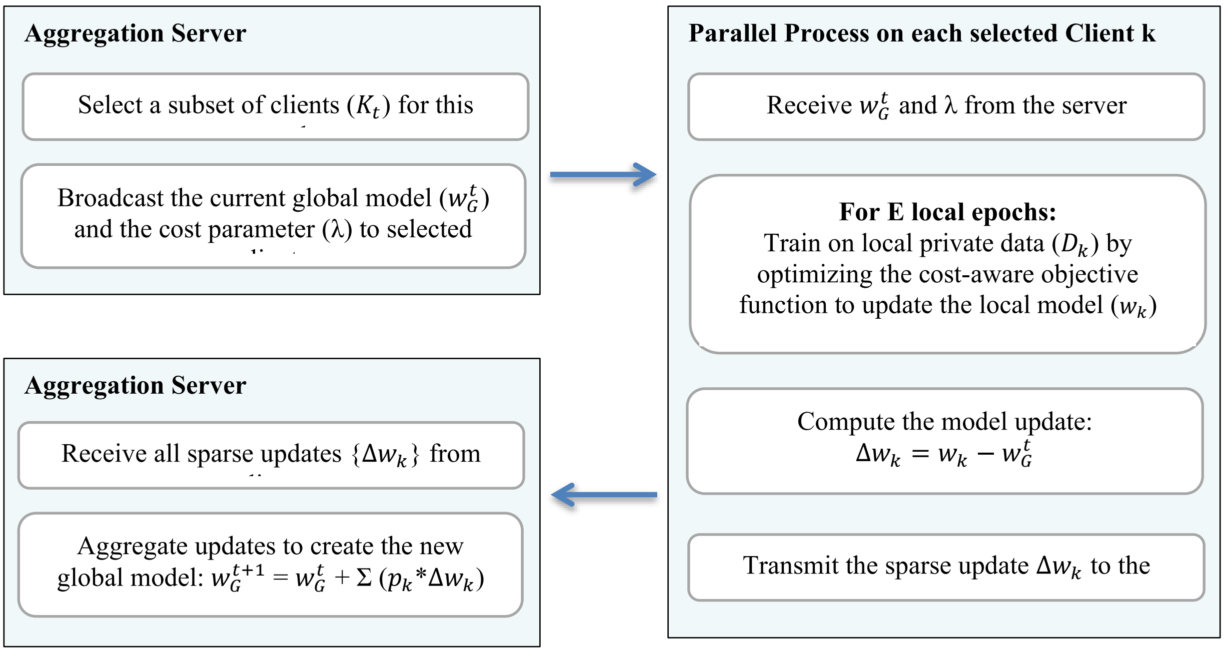

Fig. 2: Flowchart of a single training round in the proposed CA-FedAvg algorithm

C. *Experimental Setup*

To evaluate FedSW-Arch and the CA-FedAvg algorithm, a simulated environment was developed using Python.

- **Dataset:** The publicly available UCI-HAR (Human Activity Recognition) dataset [8] was utilized. This dataset comprises accelerometer and gyroscope data collected from smartphones, labeled with six activity classes: Walking, Walking Upstairs, Walking Downstairs, Sitting, Standing, and Laying. It serves as a realistic proxy for data relevant to elder care or mobility monitoring. To simulate a social work deployment, the dataset was partitioned in a non-independent and identically distributed (non-IID) manner across *N* clients, with each client possessing data from only 2–3 activity classes, thereby reflecting individualized behavioral patterns.
- **Model:** A lightweight Convolutional Neural Network (CNN) was designed for the HAR task, suitable for deployment on resource-constrained MCUs. The architecture comprises two convolutional layers followed by two fully connected layers, totaling approximately 150,000 parameters.

- **Comparison Methods:** Three approaches were compared:

(i) Centralized: A traditional, non-private method wherein all client data is aggregated on the server for direct model training, serving as an upper bound for accuracy.

(ii) Standard FedAvg: The baseline federated learning algorithm as proposed by McMahan et al. [6].

(iii) CA-FedAvg (proposed): The communication-aware algorithm introduced herein. The hyperparameter λ was empirically determined via grid search and set to 0.01.

- **Metrics:** The evaluation metrics included:

(i) Test Accuracy: The accuracy of the final global model on a held-out global test set.

(ii) Total Communication Cost: The cumulative size (in megabytes) of model updates transmitted from all clients to the server across all training rounds.

(iii) Average Energy Consumption: An energy model for a typical LoRaWAN transceiver (Semtech SX1276) was employed as described by Maude et al. [9]. The energy E required to transmit a payload of size S is modeled as:

$$E_{tx(S)} = P_{tx} \cdot T_{on-air}(S) \quad (4)$$

where P_{tx} is the transmission power and $T_{on-air}(S)$ is the time-on-air, dependent on payload size and spreading factor.

(iv) Scalability: The number of clients N was varied from 10 to 200 to assess the system's communication and energy cost scalability.

All experiments were conducted over 100 communication rounds, with 10% of clients participating in each round.

III. Results

A. Trade-off Between Accuracy and Communication

Figure 3 illustrates the test accuracy of the global model across 100 communication rounds for the three evaluated methods. As anticipated, the centralized model, trained on the entirety of the data simultaneously, attains the highest accuracy of 91.2%. The conventional FedAvg method approaches this benchmark, stabilizing at an accuracy of 89.5%, thereby affirming the feasibility of federated learning. Our proposed CA-FedAvg method achieves a final accuracy of 87.8%. This marginal decrease in accuracy represents a deliberate compromise to substantially enhance communication efficiency.

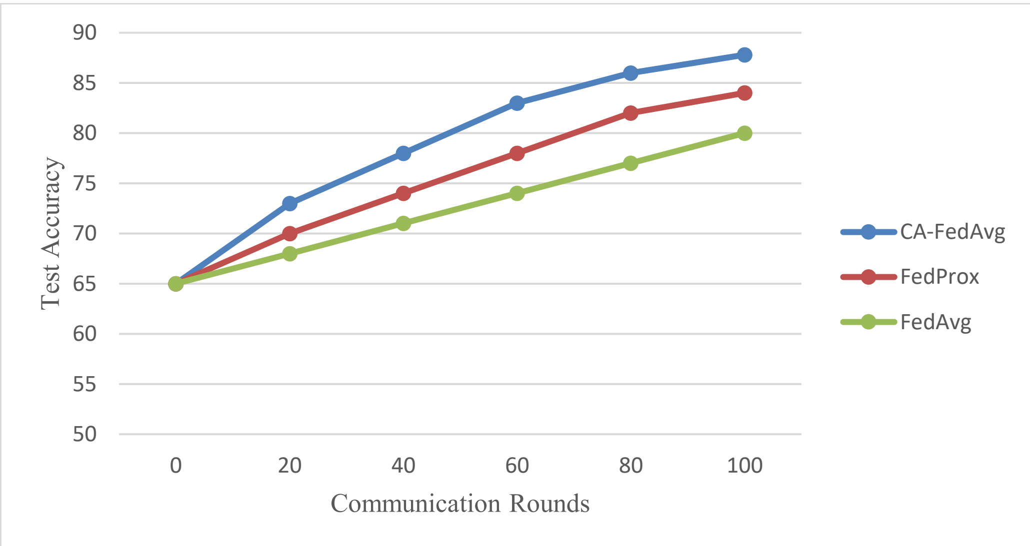

Figure 3: Test accuracy of the global model over communication rounds.

CA-FedAvg attains accuracy comparable to standard FedAvg while optimizing communication efficiency. The centralized approach serves as an upper bound for accuracy.

B. Communication and Energy Efficiency

The principal objective of FedSW-Arch is to curtail resource consumption. Table 1 presents a summary of key performance metrics after 100 communication rounds in a system comprising N=100 clients.

Table 1: Performance Comparison for N=100 Clients over 100 Rounds

Metric	**Centralized**	**Standard FedAvg**	**CA-FedAvg (Proposed)**	**Improvement (CA-FedAvg vs. FedAvg)**
Final Test Accuracy (%)	91.2	89.5	87.8	-1.9%
Total Communication Cost (MB)	N/A*	286.1	111.6	61.0%
Avg. Energy per Client (mWh)	N/A*	15.7	7.1	54.8%

*Communication and client energy costs are not applicable for the centralized model as it assumes data is already present at the server.

The data in Table 1 unequivocally demonstrate the efficacy of our approach. Despite a modest relative accuracy reduction of 1.9%, CA-FedAvg decreases the total volume of transmitted data by 61.0% relative to standard FedAvg. This improvement is attributable to the incorporation of an L1 penalty term, which encourages sparsity in model updates. The consequent reduction in data transmission directly translates into significant energy savings. Specifically, the average energy consumption of each client's transceiver during update transmission is diminished by 54.8%. This enhancement is particularly critical for battery-powered devices, potentially extending their operational lifespan between charges by a factor of two.

C. Scalability Analysis

To evaluate the architecture's applicability to large-scale deployments, we measured the total communication cost as the number of clients N increased from 10 to 200, while maintaining a constant participation fraction of 10% per communication round. The results are depicted in Figure 4.

Fig. 4: Total communication cost as a function of the total number of clients (N)

The communication cost for both methods exhibits linear scaling with respect to the number of clients; however, the absolute communication cost associated with CA-FedAvg remains consistently and substantially lower. The reduced slope observed for CA-FedAvg indicates that the communication efficiency gains afforded by our communication-aware algorithm persist predictably as the system scales. This characteristic renders large-scale deployments considerably more economically feasible.

IV. Discussion

The findings detailed in Section 3 offer compelling evidence supporting the feasibility of the FedSW-Arch framework and the efficacy of the CA-FedAvg algorithm. This section provides an interpretation of these results, addresses the study's limitations, and delineates its theoretical and practical contributions.

The principal outcome of this research is the explicit and quantifiable trade-off between model accuracy and system resource efficiency. The approximately 2% reduction in accuracy observed with CA-FedAvg should not be construed as a failure; rather, it represents a deliberate and controlled compromise that enables a reduction exceeding 50% in both communication overhead and energy consumption. Such a trade-off is particularly advantageous in many social work contexts, where a marginally less accurate model that can be reliably and economically deployed across thousands of users over extended periods on a single battery is substantially more valuable than a slightly more accurate model that is prohibitively costly or maintenance-intensive. The proposed architecture makes this trade-off transparent and adjustable via the λ hyperparameter, thereby allowing practitioners to tailor the system according to specific application requirements and resource constraints.

The use of a non-independent and identically distributed (non-IID) data distribution in our experiments is a critical aspect, as it mirrors the personalized nature of real-world data. The ability of federated learning — both in its conventional form and as enhanced by our approach — to converge to a high-performing global model under these conditions underscores its robustness and suitability for practical social applications.

A. Research Limitations

Despite providing a robust proof-of-concept, this study has several limitations. First, the evaluation was performed within a simulated environment. Real-world LPWAN deployments are subject to packet loss, variable latency, and interference, factors that were not comprehensively modeled here. Additionally, the energy consumption model, although grounded in realistic parameters, remains an approximation. Consequently, validation through deployment on physical hardware under authentic network conditions is necessary.

Second, the investigation focused on a single, relatively straightforward classification task — human activity recognition (HAR). More complex social work applications may necessitate more sophisticated models, such as those incorporating audio or natural language processing, which typically involve larger model sizes and could alter the communication-accuracy trade-off. The efficacy of the sparsity-inducing penalty in CA-FedAvg may vary depending on the model architecture employed.

Finally, this work primarily addresses privacy concerns related to data in transit and data at rest by avoiding centralized data collection. It does not, however, provide defenses against advanced adversarial threats such as model inversion or membership inference attacks, wherein an attacker attempts to reconstruct private data from model updates.

B. Theoretical and Practical Contributions

From a theoretical perspective, the principal contribution of this paper is the development of the CA-FedAvg algorithm, which explicitly incorporates communication costs into the federated optimization objective. This advancement extends the federated learning paradigm by introducing a formal mechanism to manage system-level constraints — namely energy and bandwidth — in distributed electrical systems, thereby moving beyond a sole focus on statistical learning performance. This joint optimization framework is adaptable to other resource-constrained distributed learning scenarios beyond the domain of social work.

Practically, the FedSW-Arch offers a concrete and replicable blueprint for social work agencies, technology providers, and public health organizations to develop and deploy ethical AI services. By demonstrating substantial reductions in data transmission and energy consumption, this architecture lowers the economic barriers to large-scale deployment. For social workers, this translates into access to tools capable of providing timely alerts regarding client welfare without compromising trust or privacy. For clients, it ensures the benefits of intelligent, proactive care without the intrusive sensation of continuous centralized monitoring.

C. 4.3. Future Research Directions

This study opens several promising avenues for future investigation. An immediate next step involves constructing a physical prototype of the FedSW-Arch using commercially available microcontroller units (MCUs) and LoRaWAN modules to empirically validate the simulation outcomes. Enhancing the system's privacy guarantees by integrating differential privacy techniques — which introduce statistical noise to model updates to formally prevent individual data reconstruction [10] — constitutes another critical direction. Additionally, research into dynamically adapting the λ hyperparameter based on real-time device battery levels and network conditions would represent a valuable extension. Finally, exploring advanced sparse training methodologies and model compression techniques within the CA-FedAvg framework could further diminish communication costs while preserving accuracy, thereby advancing the efficiency frontier.

V. CONCLUSION

This paper addresses a pivotal challenge at the intersection of electrical engineering, artificial intelligence, and social work: the development of scalable, privacy-preserving monitoring systems suitable for resource-constrained and sensitive environments. We identified fundamental limitations inherent in traditional centralized AI models and standard federated learning approaches, which inadequately consider the stringent power and bandwidth constraints characteristic of real-world IoT deployments.

Our principal contribution is the FedSW-Arch, a comprehensive four-layer system architecture that leverages federated learning over low-power wide-area networks to facilitate on-device AI model training. Central to this architecture is the novel CA-FedAvg algorithm, designed to explicitly and controllably trade a modest reduction in model accuracy for substantial improvements in communication and energy efficiency.

Through extensive simulations on a representative human activity recognition task, we demonstrated that the proposed architecture reduces total data transmission by over 60% and client energy consumption by more than 50% relative to standard federated learning, with only a minimal and acceptable decrease in final model accuracy of less than 2%. Moreover, these efficiency gains remain stable and predictable as the system scales to a large number of clients, confirming its suitability for deployment at the scale of cities or agencies.

The FedSW-Arch establishes a foundational and practical framework for ethically and sustainably harnessing AI in social work. By preserving data privacy, minimizing hardware and operational costs, and extending device battery life, this work bridges the gap between technological potential and impactful real-world implementation. Future efforts will focus on physical prototyping, enhancing privacy through differential privacy mechanisms, and developing more dynamic, resource-aware algorithms. Ultimately, this research lays the groundwork for a new generation of intelligent electrical systems capable of empowering social workers and improving the well-being of vulnerable populations in a safe and effective manner.

ACKNOWLEDGMENT

The author would like to express gratitude to Chang Jung Christian University for providing the essential resources and support that made this research possible. Sincere thanks are also extended to the colleagues and participants whose valuable assistance and contributions were instrumental to the completion of this study.

CONFLICT OF INTEREST

The author affirms that there are no known financial conflicts of interest or personal affiliations that may have potentially influenced the research presented in this manuscript.

REFERENCES

[1] E. E. Abbott, D. Apakama, L. D. Richardson, L. Chan, and G. N. Nadkarni, “Leveraging Artificial Intelligence and Data Science for Integration of Social Determinants of Health in Emergency Medicine: Scoping Review,” JMIR Med Inform., vol. 12, p. e57124, Oct. 2024, doi:10.2196/57124.

[2] P. Rashidi and A. Mihailidis, “A survey on ambient-assisted living tools for older adults,” IEEE J. Biomed. Health Inform., vol. 17, no. 3, pp. 579-590, May 2013.

[3] H. Hale, K. Bracewell, L. Bellussi, R. Jenkins, J. Alexander, J. Devaney, and J. E. M. Callaghan, “The child protection response to domestic violence and abuse: A scoping review of interagency interventions, models and collaboration,” Journal of Family Violence, 2024, doi:10.1007/s10896-024-00681-4.

[4] I. G. Cohen and M. M. Mello, “HIPAA and Protecting Health Information in the 21st Century,” JAMA, vol. 320, no. 3, pp. 231–232, 2018, doi:10.1001/jama.2018.5630.

[5] L. Kong, J. Tan, J. Huang, G. Chen, S. Wang, X. Jin, P. Zeng, M. Khan, and S. K. Das, “Edge-computing-driven Internet of Things: A Survey,” ACM Comput. Surv., vol. 55, no. 8, Article 174, pp. 1–41, August 2023, doi:10.1145/3555308.

[6] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-efficient learning of deep networks from decentralized data,” in Proc. 20th Int. Conf. Artificial Intelligence and Statistics (AISTATS), vol. 54, 2017, pp. 1273-1282.

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

[8] F. Cruciani, C. Sun, S. Zhang, C. Nugent, C. Li, and S. Song, “A Public Domain Dataset for Human Activity Recognition in Free-Living Conditions,” in 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing,

Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, Leicester, UK, 2019, pp. 166-171, doi: 10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00071.

[9] S. Maudet, G. Andrieux, R. Chevillon, and J.-F. Diouris, "Refined Node Energy Consumption Modeling in a LoRaWAN Network," Sensors, vol. 21, no. 19, p. 6398, Sept. 2021, doi:10.3390/s21196398.

[10] Y. Wang, Q. Wang, L. Zhao, and C. Wang, "Differential privacy in deep learning: Privacy and beyond," Future Generation Computer Systems, vol. 148, pp. 408–424, 2023, doi:10.1016/j.future.2023.06.010.