P3FL: A Privacy-Preserving Personalized Federated Learning Framework for Collaborative Smart Home Predictions and Decision-Making

Ye Wang, *Member, IEEE*, Hansong Xu, *Senior Member, IEEE*, Kun Hua, *Senior Member, IEEE*, Yang Bai, *Member, IEEE*, Jianqi Yu, *Member, IEEE*, Wenyin Zhu, *Member, IEEE*, Lixing Chen, *Member, IEEE*, Bo Yang, *Senior Member, IEEE*, and Xinping Guan, *Fellow, IEEE*

***Abstract*—Smart homes depend on collaborative sequential prediction tasks to optimize energy consumption and appliance scheduling. Federated learning (FL) offers a promising approach by enabling decentralized model training to balance privacy and usability. Yet, standard FL techniques fail to effectively address data diversity and individual user preferences in smart home contexts. To address these issues, we propose P3FL: a Privacy-Preserving Personalized Federated Learning framework that integrates tailored model training and privacy enhancements for federated collaborative predictions and decision-making. Our framework introduces the Personalized Collaborative Decision-Making (PCDM) algorithm, which dynamically adapts to different household environments while ensuring privacy and personalization. P3FL combines a global model for knowledge aggregation with a personalized adaptation module to provide fine-tuned predictions based on user preferences, environmental factors, and device configurations. Theoretical convergence bounds analysis confirms the robustness and efficiency of PCDM under conditions of strong convexity, smoothness, and bounded variance. Extensive experiments on real-world smart home datasets demonstrate that P3FL outperforms state-of-the-art methods, with PCDM achieving a training accuracy of 92.14%. Our approach enhances operational efficiency and ensures personalized user satisfaction, privacy enhancement in smart homes.**

***Index Terms*—Federated Learning, Personalization, Privacy Preservation, Collaborative Decision-Making, Smart Homes**

I. Introduction

The integration of advanced information and communication technologies (ICTs) has transformed modern smart homes into intelligent cyber-physical systems (CPS), enabling seamless automation and optimization of daily tasks [1, 2]. Central to this evolution is the fully automated management of interconnected devices, such as lighting, heating, and appliances, which hinges on collaborative sequential predictive task management to orchestrate operations across heterogeneous environments [3, 4]. For example, in energy-efficient smart homes, such systems leverage real-time multi-source data, including appliance energy consumption, solar generation, and ambient temperature, to dynamically optimize appliance scheduling. A dishwasher might be activated during peak solar production hours, while a washing machine's operation aligns with user-defined preferences and real-time energy pricing, thereby minimizing costs and maximizing renewable energy utilization. This synergy of predictive analytics and adaptive decision-making not only enhances operational efficiency but also lays the foundation for sustainable smart home ecosystems [5, 6, 7].

Traditional collaborative decision-making methods in smart homes are increasingly augmented by decentralized machine learning (DML), which excels in handling large-scale, high-dimensional data environments inherent to modern smart homes. DML improves model training efficiency and system-wide performance by distributing computational workloads across interconnected devices. However, this paradigm introduces significant security and privacy risks [8, 9]. The necessity for data sharing in distributed architectures, such as transmitting behavioral logs, occupancy patterns, and appliance-specific energy consumption data, may expose sensitive user information, including personal identities and daily routines. These vulnerabilities threaten not only individual privacy but also the integrity of smart home ecosystems, which often function as critical infrastructure. Ensuring secure information interoperability is therefore paramount to maintaining system reliability and user trust of modern smart home CPS [10, 11, 12].

Federated Learning (FL) has emerged as a privacy-aware distributed learning paradigm, particularly effective in handling non-independent and identically distributed (non-IID) data from heterogeneous smart home devices [13, 14, 15]. Unlike conventional centralized approaches, FL enables decentralized model training by eliminating the need to centralize raw data, thus mitigating risks associated with data breaches. In this framework, local devices (e.g., smart meters, thermostats) train personalized models using their private datasets, generating private model parameters instead of transmitting sensitive raw data. These parameters are securely aggregated

Ye Wang, Yang Bai, Bo Yang and Xinping Guan are with the Department of Automation, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: ywang@sjtu.edu.cn, ybai@sjtu.edu.cn, bo.yang@sjtu.edu.cn, xpguan@sjtu.edu.cn)

Hansong Xu and Lixing Chen are with the Shanghai Key Laboratory of Integrated Administration Technologies for Information Security, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: hansongxu@sjtu.edu.cn, lxchen@sjtu.edu.cn)

Kun Hua is with the Department of Electrical Engineering, California Polytechnic State University, San Luis Obispo, CA 93407, USA (e-mail: kuhua@calpoly.edu)

Jianqi Yu is with the National Innovation Institute of High-end Smart Appliances, Qingdao 266071, China (e-mail: yuja@gczx.cn)

Wenyin Zhu is with the Qingdao Haier Smart Technology R&D Co., Ltd., Qingdao 266101, China (e-mail: zhuwy@haier.com)

Corresponding authors: Yang Bai and Xinping Guan.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3572090

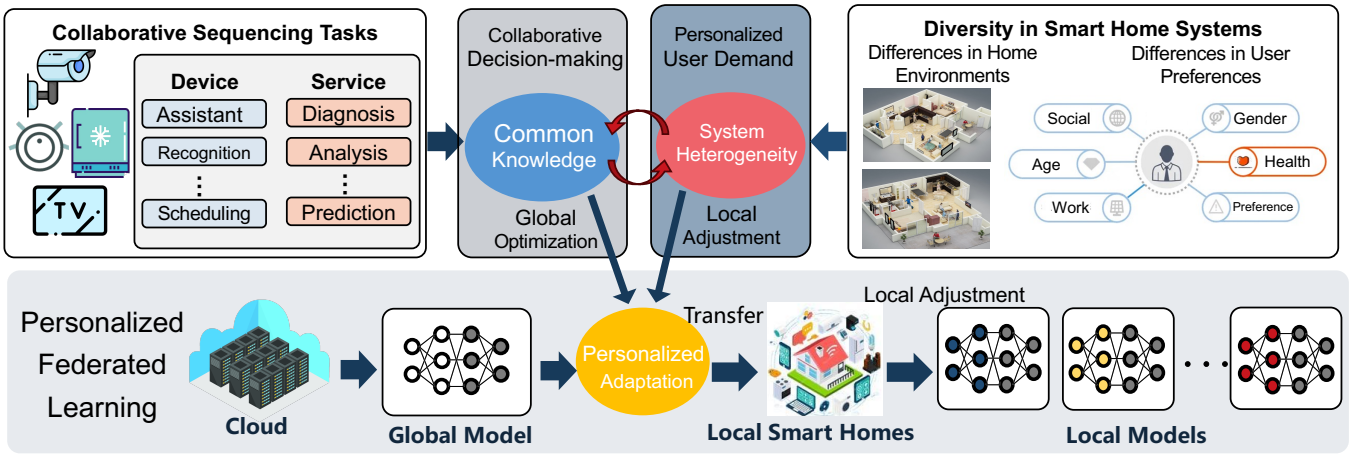


Fig. 1: The system architecture of a privacy-preserving personalized federated collaborative decision-making model for smart homes

by a central server to iteratively refine a global model, which captures system-wide patterns while preserving individual user privacy. By decoupling model training from data sharing, FL not only addresses the statistical challenges of non-IID data but also establishes a foundational privacy-preserving mechanism for collaborative intelligence in smart homes, which is critical for applications ranging from energy consumption forecasting to anomaly detection [16, 17].

Effectively incorporating user preferences requires resolving complex multi-agent coordination challenges in smart homes. Personalized federated decision-making addresses this need by unifying federated learning with adaptive decision-making, enabling decentralized systems to dynamically balance individual user requirements with global operational constraints. Unlike conventional FL frameworks, personalized federated decision-making introduces a two-tiered architecture: (1) a privacy-preserving federated layer that aggregates local model updates to capture system-wide patterns, and (2) a personalized adaptation layer that fine-tunes decisions based on real-time user preferences, device states, and environmental feedback. This dual-layered approach not only synchronizes operations across devices but also ensures that privacy-sensitive data (e.g., occupancy patterns, appliance usage habits) remain localized, thereby achieving a principled equilibrium between personalized intelligence and data confidentiality.

As illustrated in Figure 1, our proposed Privacy-Preserving Personalized Federated Learning (P3FL) framework bridges global optimization and local adaptation to address the dual challenges of data heterogeneity and user preference diversity in smart homes. The framework operates through two synergistic layers: (1) Global Federated Layer: Smart home devices train local models on their private datasets (e.g., appliance usage logs, occupancy patterns) and upload encrypted model parameters, instead of raw data, to a central aggregator. The global model is periodically updated via federated averaging and serves as a foundational reference for all devices, ensuring consistency in collaborative tasks (e.g., peak load shaving). (2) Local Adaptation Layer: Each device dynamically fine-tunes the global model using real-time contextual data, such as user-defined preferences (e.g., "laundry after 8 PM"), energy pricing signals, and device-specific constraints (e.g., thermostat operational limits). Local adaptations enable personalized decision-making, such as deferring a dishwasher cycle to prioritize a user's evening laundry schedule, without compromising global energy optimization goals.

This hierarchical design not only mitigates data heterogeneity but also harmonizes user-centric flexibility with system-wide efficiency. The personalized adaptation module tailors services by jointly optimizing static user profiles (e.g., health conditions, device capabilities) and dynamic contextual factors (e.g., time-varying preferences, energy pricing), integrating a federated learning framework to ensure privacy-aware customization. It employs similarity metrics to dynamically align local user models with the global consensus, assigning adaptive weights to balance individualized needs (e.g., a household's unique temperature thresholds) with collective efficiency goals (e.g., grid-wide energy reduction).

The convergence analysis establishes that the framework's robustness and efficiency are anchored in four key theoretical properties: strong convexity, smoothness, bounded gradient variance, and bounded client diversity. First, strong convexity guarantees the existence of a unique optimal solution and ensures exponential convergence rates, which are vital for time-critical smart home applications. Second, smoothness enforces Lipschitz continuity of gradients, stabilizing optimization by preventing erratic updates across heterogeneous clients. Bounded gradient variance further safeguards against destabilization caused by non-IID data noise, ensuring consistent training progress even under stochastic sampling. Finally, bounded client diversity quantifies the maximum divergence between local and global objectives, enabling personalized adaptations without sacrificing convergence speed. These theoretical conditions collectively explain the framework's empirical efficiency in real-world deployments, where approximations of these properties demonstrate practical resilience despite relaxed theoretical assumptions.

In our real data experiments, the proposed PCDM algorithm significantly outperforms baseline algorithms, including the

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3572090

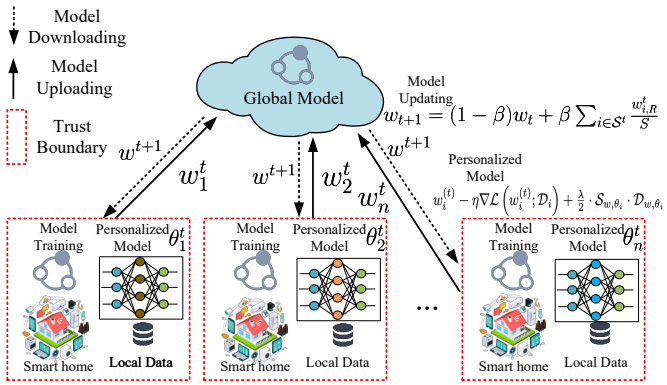


Fig. 2: A personalized privacy-preserving federated learning collaborative decision-making framework

local client approach using LSTM, which shows the lowest accuracy. By leveraging federated learning to combine data from multiple clients, the PCDM algorithm achieves a training accuracy of 92.14%, surpassing FedAvg (89.99%), PerAvg (90.13%), and pFedMe (91.06%). Finally, the framework enhances adaptive task sequencing in energy-efficient smart homes, ultimately improving operational efficiency, user satisfaction, and privacy protection. The contributions of this paper are as follows:

- A privacy-preserving personalized federated learning framework is developed for enhanced energy forecasting capabilities in smart homes, ensuring diverse privacy-preserving collaborations.
- Theoretical convergence bounds are established, demonstrating the robustness and reliability of the proposed algorithm in real-world applications.
- The framework is evaluated on real-world energy consumption datasets. Comprehensive experimental results validate the effectiveness of our approach, showing significant improvements in prediction accuracy and personalized adaptability across varying scenarios.

The remainder of this paper is organized as follows: In Section II, we present our approach to personalized federated collaborative decision-making in smart homes. In Section III, we present the algorithm design and theoretical convergence bounds analysis. We present the experimental evaluation in Section IV. We review the related works in Section V. Finally, the article is concluded in Section VI.

II. Our Approach

In this section, we elaborate on the personalized federated learning stability prediction framework, as shown in Figure 2, that includes smart home device users and system coordinators. Device Users: Let $i = 1, ..., N$ represent the collection of different scenarios (locations) in smart homes, such as schools, residential areas, industrial zones, etc. Each device user $i \in N$ possesses various heterogeneous energy data created by their smart meters, appliances, solar panels, electric vehicles, and so on. Let $\mathcal{D}_i$ denote the private energy dataset owned by each device user $i \in N$. System Coordinators: They are primarily responsible for publishing federated learning tasks and aggregating local knowledge from different locations to achieve federated collaborative decision-making. The federated collaborative decision-making process leverages machine learning to train on smart home datasets, including load demands and voltage variations, to predict system stability. It enables automatic feature learning, handles nonlinear problems, and supports continuous improvement through online learning. The process includes local model training, parameter uploading, model aggregation, global model distribution, and personalized adjustments.

- *Local Model Training:* Each device user $i \in \{1, ...N\}$ trains a model locally using their private dataset D_i. The model learns the mapping relationship between input features x (such as load demand, motor power, and voltage variations) and the system stability label y (stable/unstable).
- *Parameter Upload:* Device users upload their trained model parameters to the system coordinator, ensuring data privacy.
- *Model Aggregation:* The system coordinator aggregates the uploaded model parameters from all device users using methods like the Federated Averaging algorithm (FedAvg) to generate a global model.
- *Global Model Distribution:* The system coordinator distributes the updated global model to all device users, who can then use this global model for predictions and decision-making.
- *Personalized Adjustment:* Each device user can make personalized adjustments based on the global model to better fit their local environment. For example, they can fine-tune the global model to adapt to specific energy usage patterns in their location.

A. Federated Learning Model and Training Process

Federated learning is a secure distributed machine learning practice that allows for local model training while only transmitting model updates, effectively reducing the risk of data leakage and protecting privacy. In a federated learning system, there is an aggregation server and N clients, each possessing their sensitive dataset D_n. All selected clients send their gradients to the server for model aggregation. The federated training process includes steps such as model initialization, local model training and updates, and global model aggregation and updates.

The federated training process includes the following steps:

- *Model Initialization:* The aggregation server initializes the global model parameters w and distributes them to all selected clients.
- *Model Initialization:* Each client trains the model locally using its private dataset D_n and computes the gradients ∇w_n.The client updates its local model parameters using these gradients.
- *Gradient Transmission:* Each client sends its computed gradients ∇w_n to the aggregation server. To protect privacy, secure communication protocols or encryption techniques can be employed during this transmission.

- *Global Model Aggregation and Updates:* The aggregation server receives the gradients ∇w_n from all clients and aggregates them using methods like the Federated Averaging algorithm (FedAvg) to update the global model parameters w.
- *Global Model Distribution:* The aggregation server distributes the updated global model parameters w to all clients, initiating a new round of local training and updates.

B. Personalized Model Aggregation in Smart Homes

In federated learning frameworks for smart home systems, the personalized model aggregation mechanism addresses the inherent tension between collaborative learning and individualized optimization. Traditional federated averaging (FedAvg) methods treat all local models uniformly during aggregation, which risks suppressing unique device-specific patterns and degrading performance in heterogeneous environments. To overcome this limitation, the proposed mechanism introduces a dual-phase process: compatibility evaluation and adaptive contribution weighting, designed to harmonize global knowledge integration with local personalization.

1) Compatibility Evaluation: The mechanism first quantifies the alignment between each local model and the global model using two complementary metrics:

- *Parameter Distance:* Computed via Euclidean or cosine similarity between model weight vectors, this metric identifies devices whose optimization trajectories closely align with the global model.
- *Task-Specific Divergence:* For applications like energy optimization, domain-specific metrics (e.g., difference in predicted energy savings) supplement geometric distances to evaluate functional compatibility.

2) Adaptive Contribution Weighting: Based on compatibility scores, a dynamic weighting function assigns aggregation weights to each client i at iteration t. Clients exhibiting high compatibility receive elevated weights, ensuring their strategies disproportionately influence the global model. Conversely, devices with low compatibility—often reflecting legitimate personalization needs (e.g., medical equipment requiring strict temperature ranges)—retain full local autonomy while contributing minimally to global updates. The weighting function is periodically recalibrated to adapt to temporal shifts, such as seasonal energy demand variations or changes in user behavior.

In practical applications, this mechanism excels in collaborative tasks among smart home devices, such as intelligent appliance control, environmental monitoring, and energy optimization. For instance, in air conditioning systems, it enables households to generate personalized regulation strategies based on unique preferences while maintaining efficient collaboration for enhanced overall performance. Key advantages of this approach include: 1) dynamic balancing of global and local models to improve adaptability across diverse scenarios; 2) efficient device collaboration without raw data sharing, enhancing privacy trust; and 3) higher flexibility and scalability for supporting a wider range of smart home applications in the future.

TABLE I: The main symbols and descriptions

Symbols	Description
$\mathcal{N}$	The smart home device users
x	The input features (load demand, motor power, and voltage variations)
y	The system stability label (stable/unstable)
θ_i	The personalized model of client i
w	The global model parameters
∇w_n	The gradients of local models
$\mathcal{L}_i(w)$	The training cost on the data distribution of client i
$\mathcal{D}_i$	A randomly selected data sample of client i
λ	The degree of model personalization
η	The learning rate of clients.
β	The contribution ratio of clients to the aggregated model on the server
d	The dimensionality of the search space
ν	The accuracy level
$\mathcal{S}_{w,\theta_i}$	The gradient similarity
$\mathcal{D}_{w,\theta_i}$	The Euclidean distance to calculate the diversity
$\mathcal{O}(\cdot)$	The hidden constants
S^t	The server randomly samples the client clusters
δ	The sampling noise adjustment

III. Algorithm Design and Analysis

We assume that energy consumers and producers are honest but curious, adhering to the federated learning (FL) protocol while attempting to extract information from other entities. Threats considered include information leakage and model inversion attacks. We model the personalized federated collaborative intelligent decision-making for smart homes as a server aggregation optimization problem. In traditional federated learning, $\mathcal{N}$ clients upload model parameters to the server, raising concerns about parameter updates that should not reveal sensitive information, performance optimization to enhance the global model's accuracy, and privacy preservation through mechanisms that secure client data during aggregation. This reformulation allows us to develop strategies to improve both the efficiency and security of the federated learning process in smart home environments. The main symbols and descriptions in this paper are shown in Table I.

$$\min_{w\in\mathbb{R}^d}\left\{\mathcal{L}(w)=\frac{1}{\mathcal{N}}\sum_{i=1}^{\mathcal{N}}\mathcal{L}_i(w)\right\} \tag{1}$$

The equation represents the expected training cost $\mathcal{L}_i(w)$ on the data distribution of client i.

$$\mathcal{L}_i(w)=\mathbb{E}_{\mathcal{D}_i}\left[\tilde{\mathcal{L}}_i\left(w;\mathcal{D}_i\right)\right] \tag{2}$$

where $\mathcal{D}$ is a randomly selected data sample from the data of client i.

Based on traditional federated learning training problems, this paper constructs the training objectives for personalized federated learning in smart homes:

$$\begin{aligned}&\min_{w}\mathcal{L}(w)=\frac{1}{\mathcal{N}}\sum_{i=1}^{\mathcal{N}}\mathcal{L}\left(w;\mathcal{D}_i\right)\\&w_i^{(t+1)}=w_i^{(t)}-\eta\nabla\mathcal{L}\left(w_i^{(t)};\mathcal{D}_i\right)+\frac{\lambda}{2}\cdot\mathcal{S}_{w,\theta_i}\cdot\mathcal{D}_{w,\theta_i},\end{aligned} \tag{3}$$

Where $\mathcal{S}$ is the similarity of gradient between local clients and server.

As presented in equation (3), gradient similarity $\mathcal{S}_{w,\theta_i}$ is utilized as a key metric to assess individual model personalization. θ_i is the personalized model of client i. The parameter

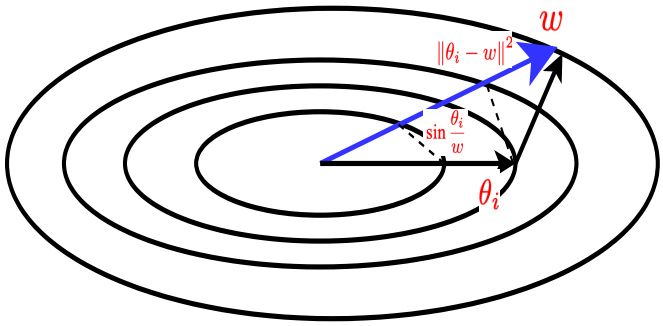


Fig. 3: The personalized gradient iteration direction in the model aggregation

ratio, $\frac{\theta_i}{w}$, serves as an indicator of the client's model personalization. In the paper, sine similarity $\mathcal{S}_{w,\theta_i}$ is considered to measure the directional consistency between global gradients and local gradients. As observed from Figure3, if the sine similarity is high, it indicates that the two directions are more consistent; if the sine similarity is low, it suggests a greater discrepancy in their directions. The sine similarity can be expressed as:

$$\mathcal{S}_{w,\theta_i} = \sqrt{1 - \left(\frac{w \cdot \theta_i}{\|w\| \, \|\theta_i\|}\right)^2} \tag{4}$$

Then, we consider the gradient distance between local client and server. In this paper, we use the *Euclidean* distance to calculate the diversity.

$$\mathcal{D}_{w,\theta_i} = \|\theta_i - w\|^2 \tag{5}$$

By combining the information from global gradients and local gradients, we can better adapt to the local data distributions of each client and enhance the performance of personalized models. The adjustment method based on sine similarity and distance demonstrates high flexibility, enabling dynamic adjustments to the gradient iteration direction according to the heterogeneity of client data. However, this method has limitations, such as increased computational overhead caused by calculating sine similarity and distances, particularly when there are many clients. In Federated Learning, frequent communication between clients and servers is required to exchange gradient information, which may lead to additional communication overhead. Therefore, in future research, we will consider improving training efficiency.

A. Algorithm Design

In line 6 of the algorithm, adopt δ-approximation of $\hat{\theta}_i\left(w_{i,r}^t\right)$, represented as $\tilde{\theta}_i\left(w_{i,r}^t\right)$, satisfying Assumption 3. Then, using $\lambda\left(w_{i,r}^t - \tilde{\theta}_i\left(w_{i,r}^t\right)\right)$to approximate $\nabla F_i\left(w_{i,r}^t\right)$. In actual training, by sampling the dataset $\mathcal{D}_i$ using mini-batch samples, the unbiased estimation of $\nabla\mathcal{F}_i\left(\theta_i\right)$ is as follows:

$$\nabla\tilde{\mathcal{F}}_i\left(\theta_i, \mathcal{D}_i\right) := \frac{1}{|\mathcal{D}_i|} \sum_{\xi_i \in \mathcal{D}_i} \nabla\tilde{\mathcal{F}}_i\left(\theta_i; \xi_i\right) \tag{7}$$

Algorithm 1: Personalized Collaborative Decision-Making Algorithm (PCDM)

1 **Input:** Training time T, Training Round R, client set S, hyperparameters $\lambda, \eta_l, \eta_g, \beta$ and the initial gradient w^0
2 The server sends the gradient parameters w_t to all clients.
3 **for** $i = 1 \to N$ **do**
4 $w_{i,0}^t = w^t$
5 **for** $r = 0 \to R-1$ **do**
6 Sample a small dataset $\mathcal{D}_i$, to minimize equation (3) as follows:
$$\min_w \mathcal{L}(w) = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} \mathcal{L}\left(w; \mathcal{D}_i\right) \tag{6}$$
7 Each client's gradient update: $w_i^{(t+1)} = w_i^{(t)} - \eta\nabla\mathcal{L}\left(w_i^{(t)}; \mathcal{D}_i\right) + \frac{\lambda}{2} \cdot \mathcal{S}_{w,\theta_i} \cdot \mathcal{D}_{w,\theta_i}$
8 **end**
9 **end**
10 The server randomly samples the client cluster S^t, and collects each client's gradient parameters.
11 Server-side personalized collaborative decision model aggregation: $w_{t+1} = (1-\beta)w_t + \beta\eta_g \sum_{i \in \mathcal{S}^t} \frac{w_{i,R}^t}{S}$

where $\mathbb{E}\left[\nabla\tilde{\mathcal{F}}_i\left(\theta_i, \mathcal{D}_i\right)\right] = \nabla\mathcal{F}_i\left(\theta_i\right)$. Then, using a first-order iterative method, we obtain $\tilde{\theta}_i\left(w_{i,r}^t\right)$ approximation, as follows:

$$\tilde{\mathcal{H}}_i\left(\theta_i; w_{i,r}^t, \mathcal{D}_i\right) := \tilde{\mathcal{F}}_i\left(\theta_i; \mathcal{D}_i\right) + \frac{\lambda}{2} \cdot \sin\frac{\theta_i}{w_{i,r}^t} \left\|\theta_i - w_{i,r}^t\right\|^2 \tag{8}$$

To better optimize the above problem, assume that $\tilde{\mathcal{H}}_i\left(\theta_i; w_{i,r}^t, \mathcal{D}_i\right)$is strongly convex, which can be approximated to the optimal value using gradient descent, satisfying:

$$\left\|\nabla\tilde{\mathcal{H}}_i\left(\tilde{\theta}_i; w_{i,r}^t, \mathcal{D}_i\right)\right\|^2 \leq \nu \tag{9}$$

It satisfies: $K := \mathcal{O}\left(\kappa\log\left(\frac{d}{\nu}\right)\right)$ (resp. $\mathcal{O}\left(\sqrt{\kappa}\log\left(\frac{d}{\nu}\right)\right)$) [18], where d represents the dimensionality of the search space, ν denotes the accuracy level, and $\mathcal{O}(\cdot)$ indicates hidden constants. In the algorithm, the computational complexity for each client is K times that of FedAvg. By adjusting the data size $\mathcal{D}$ through the following lemma helps control the sampling noise adjustment δ and the accuracy level.

B. Convergence Analysis

In this section, we give the theoretical convergence bounds analysis of PCDM under conditions of strong convexity, smoothness, and bounded variance. In order to derive the convergence bound, we give the following assumptions:

Assumption 1 (Convex and smoothness [19]) *For each $i = 1, \ldots, n$, the function $\mathcal{F}_i$ is continuously differentiable. There exist constants $L_w, L_{w'}$ such that for each $i = 1, \ldots, n$:*

$\nabla_w \mathcal{F}_i(w)$ is L-Lipschitz with respect to w, and $\nabla_w \mathcal{F}_i(w)$ is μ-strongly convex with respect to w.

Assumption 2 (Bounded variance [20]) *The stochastic gradients in Algorithm 1 is unbiased and have the bounded variance, that is, $\forall w$:*

$$\mathbf{E}\left[\widetilde{\nabla}_w \mathcal{F}_i(w)\right] = \nabla_w \mathcal{F}_i(w), \tag{10}$$

Furthermore, there is a random variable ξ, for the stochastic gradient $\nabla \mathcal{F}(w;\xi)$ with a bounded variance σ^2, the variance of stochastic gradients in each client is bounded:

$$\mathbb{E}_{\xi_i}\left[\|\nabla \mathcal{F}_i(w;\xi_i) - \nabla \mathcal{F}_i(w)\|^2\right] \leq \sigma^2, \quad \forall i, w$$

where σ^2 is an upper bound on the variance, $||\cdot||$ is the L_2 norm of the vector.

The Assumption 2 is a standard bounded variance assumption for stochastic gradients of each client.

Assumption 3 (Bounded diversity [21]) *The variance of local gradients to global gradients is bounded*

$$\mathbb{E}\left[\|\nabla \mathcal{F}_i(w) - \nabla \mathcal{F}(w)\|\right] \leq \delta^2, \quad \forall i, w$$

where δ^2 is the upper bound of diversity. The smaller δ^2 is, the smaller the differences between clients or tasks. $||\cdot||$ is the L_2 norm of the vector.

Theorem 1 *Let $\tilde{\theta}_i\left(w_{i,r}^t\right)$ be a solution to equation (3), we have the following two cases:*

(a) if Assumption 1 (a) holds:

$$\mathbb{E}\left[\left\|\tilde{\theta}_i\left(w_{i,r}^t\right) - \hat{\theta}_i\left(w_{i,r}^t\right)\right\|^2\right] = \frac{2}{(\lambda+\mu)^2}\left(\frac{\sigma^2}{|\mathcal{D}|} + \nu\right)$$

(b)if Assumption 1 (b) holds, and $\lambda > L$:

$$\mathbb{E}\left[\left\|\tilde{\theta}_i\left(w_{i,r}^t\right) - \hat{\theta}_i\left(w_{i,r}^t\right)\right\|^2\right] = \frac{2}{(\lambda-L)^2}\left(\frac{\sigma^2}{|\mathcal{D}|} + \nu\right)$$

Proof *For the case (a) in Assumption 1, $\mathcal{H}_i\left(\theta_i; w_{i,r}^t, \mathcal{D}_i\right) := \mathcal{F}_i\left(\theta_i;\mathcal{D}_i\right) + \frac{\lambda}{2}\cdot\sin\frac{\theta_i}{w_{i,r}^t}\left\|\theta_i - w_{i,r}^t\right\|^2$, then $(\lambda+\mu)$-strongly convex with the solution $\hat{\theta}_i(w_{i,r}^t)$ can be obtained as*

$$\begin{aligned}
&\left\|\tilde{\theta}_i\left(w_{i,r}^t\right) - \hat{\theta}_i\left(w_{i,r}^t\right)\right\|^2 \\
&\leq \frac{1}{(\lambda+\mu)^2}\left\|\nabla \mathcal{H}_i\left(\tilde{\theta}_i; w_{i,r}^t\right)\right\|^2 \\
&\leq \frac{2}{(\lambda+\mu)^2}\left(\left\|\nabla \mathcal{H}_i\left(\tilde{\theta}_i; w_{i,r}^t\right) - \nabla \tilde{\mathcal{H}}_i\left(\tilde{\theta}_i; w_{i,r}^t, \mathcal{D}_i\right)\right\|^2 + \right. \\
&\left.\left\|\nabla \tilde{\mathcal{H}}_i\left(\tilde{\theta}_i; w_{i,r}^t, \mathcal{D}_i\right)\right\|^2\right)
\end{aligned} \tag{11}$$

where

$$\nabla \mathcal{H}_i\left(\tilde{\theta}_i; w_{i,r}^t\right) = \nabla \tilde{\mathcal{F}}_i\left(\tilde{\theta}_i;\mathcal{D}_i\right) + \frac{\lambda}{2}\cos\frac{\hat{\theta}}{w_{i,r}^t}\|\theta_i - w_{i,r}^t\|^2 \tag{12}$$

From Assumption 3, we can get the gradient boundary:

$$\|\nabla \mathcal{F}_i(w) - \nabla \mathcal{F}(w)\| = \|\theta_i - w_{i,r}^t\| \leq \delta^2, \quad \forall i, w. \tag{13}$$

Thus, the gradient of function $\mathcal{H}_i\left(\theta_i; w_{i,r}^t, \mathcal{D}_i\right)$ can be transferred as

$$\nabla \mathcal{H}_i\left(\tilde{\theta}_i; w_{i,r}^t\right) = \nabla \tilde{\mathcal{F}}_i\left(\tilde{\theta}_i;\mathcal{D}_i\right) + \frac{\lambda}{2}\cos\frac{\hat{\theta}}{w_{i,r}^t}\delta^2 \tag{14}$$

The Equation 11 can be transformed into the following

$$\begin{aligned}
&\left\|\tilde{\theta}_i\left(w_{i,r}^t\right) - \hat{\theta}_i\left(w_{i,r}^t\right)\right\|^2 \leq \frac{2}{(\lambda+\mu)^2}(\left\|\nabla \tilde{\mathcal{F}}_i\left(\tilde{\theta}_i;\mathcal{D}_i\right) + \right. \\
&\left.\frac{\lambda}{2}\cos\frac{\hat{\theta}}{w_{i,r}}\delta^2 - \nabla \mathcal{F}_i\left(\tilde{\theta}_i\right) - \frac{\lambda}{2}\cos\frac{\hat{\theta}}{w_{i,r}}\delta^2\right\|^2 + \nu) \\
&= \frac{2}{(\lambda+\mu)^2}\left(\left\|\nabla \tilde{\mathcal{F}}_i\left(\tilde{\theta}_i;\mathcal{D}_i\right) - \nabla \mathcal{F}_i\left(\tilde{\theta}_i\right)\right\|^2 + \nu\right) \\
&= \frac{2}{(\lambda+\mu)^2}\left(\frac{1}{|\mathcal{D}|^2}\left\|\sum_{\xi_i\in\mathcal{D}_i}\nabla \tilde{\mathcal{F}}_i\left(\tilde{\theta}_i;\xi_i\right) - \nabla \mathcal{F}_i\left(\tilde{\theta}_i\right)\right\|^2 + \nu\right)
\end{aligned} \tag{15}$$

The following equation is obtained by solving for expectations on both sides of the equation (15):

$$\begin{aligned}
&\mathbb{E}\left[\left\|\tilde{\theta}_i\left(w_{i,r}^t\right) - \hat{\theta}_i\left(w_{i,r}^t\right)\right\|^2\right] \overset{(a)}{\leq} \frac{2}{(\lambda+\mu)^2}(\frac{1}{|\mathcal{D}|^2} \\
&\sum_{\xi_i\in\mathcal{D}_i}\mathbb{E}_{\xi_i}\left[\left\|\nabla \tilde{\mathcal{F}}_i\left(\tilde{\theta}_i;\xi_i\right) - \nabla \mathcal{F}_i\left(\tilde{\theta}_i\right)\right\|^2\right] + \nu) \\
&\overset{(b)}{\leq} \frac{2}{(\lambda+\mu)^2}\left(\frac{\sigma^2}{|\mathcal{D}|} + \nu\right) \leq \delta^2,
\end{aligned} \tag{16}$$

where the inequality (a) is obtained, since

$$\mathbb{E}\left[\left\|\sum_{i=1}^{M} X_i - \mathbb{E}\left[X_i\right]\right\|^2\right] = \sum_{i=1}^{M}\mathbb{E}\left[\|X_i - \mathbb{E}\left[X_i\right]\|\right]^2 \tag{17}$$

And for the unbiased estimate:

$$\mathbb{E}\left[\nabla \tilde{\mathcal{F}}_i\left(\tilde{\theta}_i;\xi_i\right)\right] = \nabla \mathcal{F}_i\left(\tilde{\theta}_i\right) \tag{18}$$

For the case (b), we can get the similar result based on $(\lambda - L)$-strongly convex for the equation $\mathcal{H}_i\left(\theta_i; w_{i,r}^t\right)$, if Assumption 1 (b) holds, and $\lambda > L$:

$$\mathbb{E}\left[\left\|\tilde{\theta}_i\left(w_{i,r}^t\right) - \hat{\theta}_i\left(w_{i,r}^t\right)\right\|^2\right] = \frac{2}{(\lambda-L)^2}\left(\frac{\sigma^2}{|\mathcal{D}|} + \nu\right) \leq \delta^2 \tag{19}$$

IV. Experimental Evaluation

In this section, we conduct extensive experiments to validate our approach.

A. Data Preparation

We use real-world smart home datasets from [22, 23] in our experiments. The smart home datasets contain readings from smart meters for household appliances in kilowatts over 350 days at one-minute intervals, along with weather conditions for the specific region.

The smart home dataset consists of a total of 503,910 data entries, each containing 32 features, including time,

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3572090

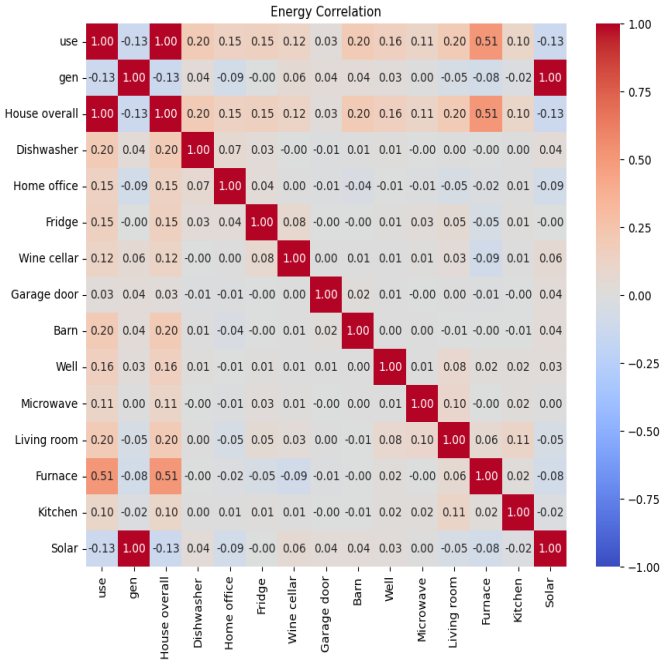


Fig. 4: Correlation Indicators of Features in the Smart Home Dataset

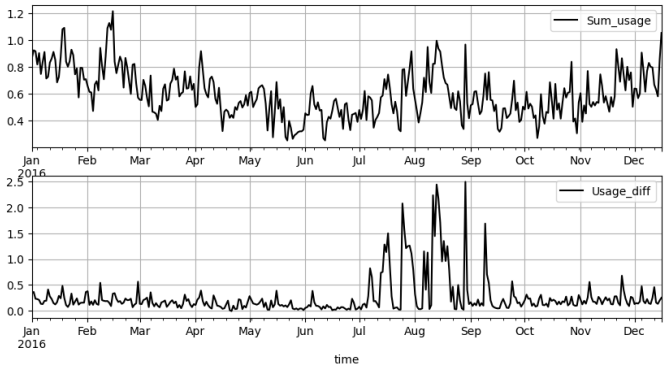


Fig. 5: Variation curve of total household electricity consumption

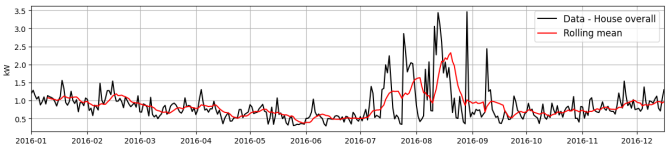


Fig. 6: Fitted variation curve of total household electricity consumption

use, gen, House overall, Dishwasher, Furnace 1, Furnace 2, Home office, Fridge, Wine cellar, Garage door, Kitchen, Barn, Well, Microwave, Living room, Solar, temperature, icon, humidity, visibility, summary, apparent temperature, pressure, wind speed, cloud cover, wind bearing, Precipitation Intensity, dew Point, Precipitation Probability.

To assess the quality of the training dataset, correlation metrics for the 32 feature values were evaluated. As illustrated in Figure 4, correlation graph analysis helps uncover relationships between features, providing insights for optimizing feature selection. Highly correlated feature pairs can introduce redundancy, requiring appropriate handling. In the experiments, a de-redundancy process was applied to manage highly correlated features within the dataset.

For example, the correlation between the "gen" and "solar" data is 1, indicating a strong relationship. These two data groups will be removed to ensure the effectiveness of model predictions in the experiment, as shown in Figure 4.

Figure 5 depicts the yearly trends in total electricity consumption and peak electricity consumption. As shown, this experiment utilized the federated learning model for training, based on the dataset features, to predict electricity consumption. By analyzing and comparing the trends in Figure 5, it was determined that predicting peak power consumption is more practical than total power consumption. Consequently, the experiment focused on forecasting instantaneous peak electricity consumption using the smart home dataset to maintain the stability of the overall power consumption system. Figure 6 presents the peak electricity consumption time series data after smoothing. This was achieved by applying a sliding window to calculate the mean, effectively eliminating noise or short-term fluctuations while preserving the long-term trend.

B. Metrics and Baseline Algorithms

In the experiment, the metrics include mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and R^2 Score. The details are as follows:

- MSE: This metric defines the average of the squares of the errors, serving as an evaluation criterion for prediction models or predictors. A lower MSE value indicates a better predictor. The MSE parameter considers the relationship between variance (differences among predictions) and bias (the distance between predicted values and actual values), calculated as follows:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2. \tag{20}$$

- RMSE: This metric defines the square root of the mean squared error and serves as an extension of the MSE indicator. Like MSE, it penalizes larger errors more heavily. RMSE is a positive value, with smaller values indicating better model prediction performance. An advantage of RMSE is that its value is in the same units as the predicted values. It is calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}. \tag{21}$$

TABLE II: Comparison of the experimental results of the LSTM model and ARIMA model

	MSE	RMSE	MAE	MAPE	MASE	R^2 Score
LSTM model	0.06930	0.263	0.174	0.226	0.710	0.095
ARIMA model	0.06967	0.26394	0.176	0.311	0.710	0.104

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3572090

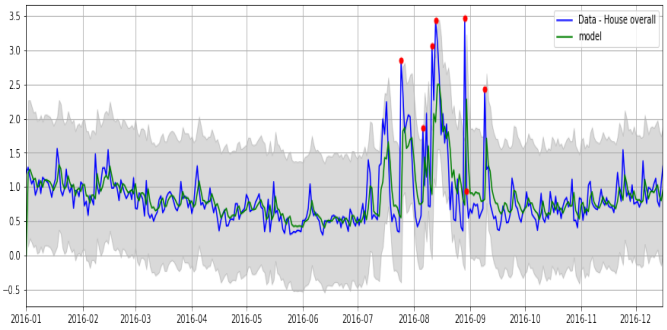


Fig. 7: Experimental results of the LSTM local model

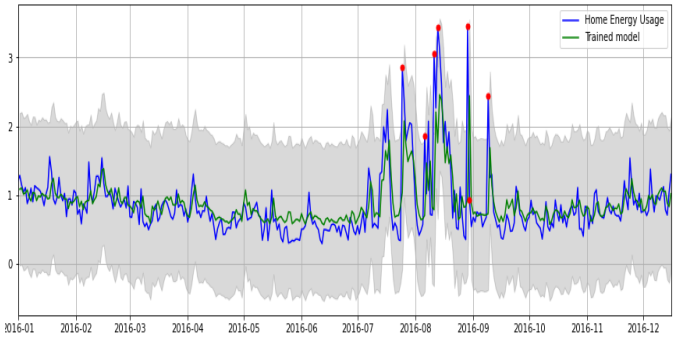


Fig. 8: Experimental results of the ARIMA model

- MAE: This metric defines the average absolute error, representing the arithmetic mean of the absolute differences between the predicted values and the actual values. The calculation formula is as follows:

$$\text{MAE} = \frac{\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|}{n} = \frac{\sum_{i=1}^{n}|e_i|}{n}. \quad (22)$$

- MAPE: This metric defines a relative measure of error between data points, using absolute values to avoid cancellation between positive and negative errors. The calculation is as follows:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right|. \quad (23)$$

- R^2: This metric measures the proportion of variance in the dependent variable that the independent variables can explain. The calculation method is as follows:

$$R^2\text{score} = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2} \quad (24)$$

To evaluate the accuracy of the time-series prediction model, we tested the LSTM (Long Short-Term Memory) and ARIMA (Autoregressive Integrated Moving Average) models, as detailed in Table II. The results indicate that the LSTM model outperforms the ARIMA model, particularly in the MAPE metric, where the improvement is more pronounced. Figures 7 and 8 illustrate the prediction performance of the LSTM and ARIMA models on the stability of the overall power consumption, respectively.

The LSTM model proposed in this study demonstrates excellent predictive performance in the smart home scenario. In the subsequent federated learning framework, the local model leverages the LSTM model for training. Once training is complete, the local models are uploaded to the server for aggregation, resulting in a personalized model tailored to individual clients.

The LSTM model addresses the short memory issue of traditional recurrent neural networks (RNN). LSTM utilizes a series of ”gates” (each with its own RNN) that can probabilistically retain, forget, or ignore data points. By basing the LSTM model on this framework, sequences can be divided into multiple input/output patterns, transforming the sequences into a supervised learning problem. From the data partitioning of input/output patterns, the model learns the input patterns and outputs, producing single-step predictions.

In our design, the LSTM model consists of 50 units in its layers, uses the ReLU activation function, includes a Dense(1) layer, employs the Adam optimizer, and utilizes mean squared error (MSE) as the loss metric, resulting in a total of 10,451 parameters. Figure 7 illustrates the results of predicting electricity consumption data using LSTM locally, demonstrating that LSTM can achieve excellent prediction accuracy. The experiments are executed on an Ubuntu system with an Intel i9 12900K CPU and GeForce RTX 3090 Ti GPU.

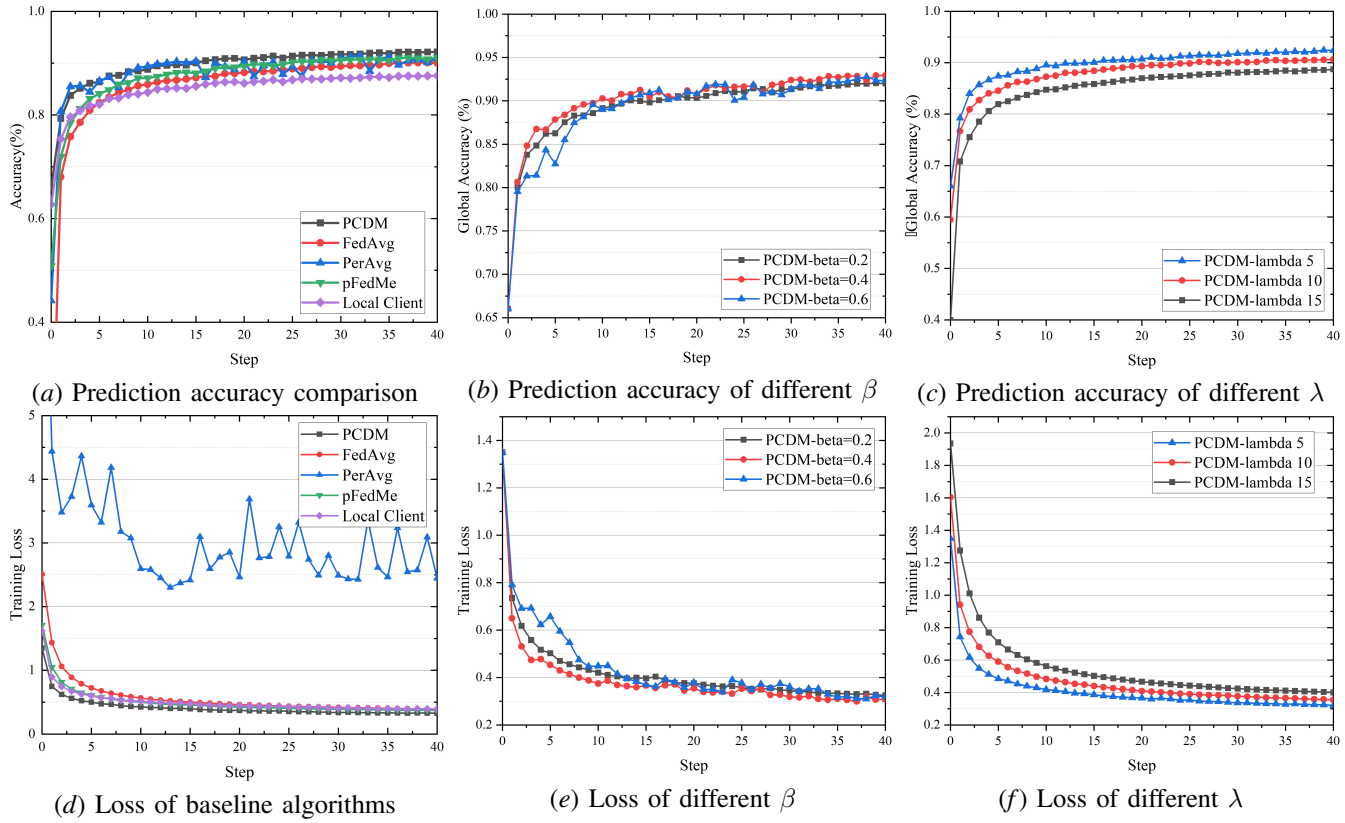
Using the proposed PCDM algorithm, the LSTM models are trained locally on multiple clients. After training, the local models are uploaded to the server, where they are aggregated into an enhanced smart home prediction model. Here, we compare the prediction accuracy and training loss of PCDM to other state-of-the-art federated learning algorithms. The details of those algorithms, such as FedAvg, PerAvg, pFedMe are defined as follows:

- FedAvg: The baseline algorithm for federated learning.
- PerAvg: A personalized federated learning approach that combines meta-learning model MAML to find a global model that performs well after updates on each node's specific loss function [24].
- pFedMe: A personalized federated learning scheme that incorporates Moreau envelopes optimization based on client loss functions. This allows clients to build a global model while also optimizing their personalized models [25].
- Local client: This algorithm is without the FL architecture and is trained locally based on the LSTM model to verify the superiority of the FL algorithm.

C. Evaluation Results

1) Baseline algorithms comparison: In Figure 9 (*a*) and (*d*), the comparison of model prediction accuracy across different algorithms is displayed, showing that local training yields the lowest accuracy. The black line represents our proposed PCDM algorithm, with the local client algorithm serving as the baseline in the experiments. Specifically, the baseline algorithm employs LSTM during local training and can be viewed as a centralized training approach. The local client algorithm was chosen to highlight the performance improvement

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3572090



(*a*) Prediction accuracy comparison (*b*) Prediction accuracy of different β (*c*) Prediction accuracy of different λ

(*d*) Loss of baseline algorithms (*e*) Loss of different β (*f*) Loss of different λ

Fig. 9: Performance comparison of baseline algorithms, β and λ.

achieved by using the PCDM-based FL algorithm. Unlike the local client approach, the FL algorithm addresses the limitations of relying solely on local datasets, resulting in enhanced model performance. By integrating data from different clients through federated learning, the model's prediction accuracy is significantly improved. The personalized federated learning algorithm PCDM designed in this study achieves superior training accuracy and lower training loss at 92.14%, compared to FedAvg at 89.99%, PerAvg at 90.13%, and pFedMe at 91.06%.

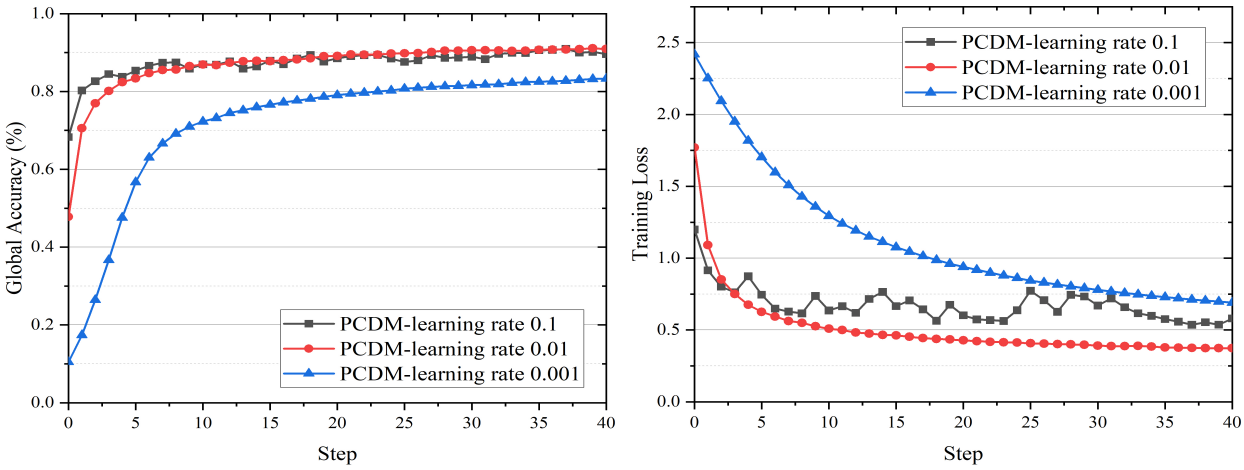
2) Hyperparameter β comparison: We evaluated the prediction accuracy and model training loss of the PCDM algorithm under different hyperparameters. Figure 9 (*b*) and (*e*) illustrate the impact of varying β on the performance of the PCDM algorithm. The parameter β measures the contribution ratio of different clients to the aggregated model on the server. The experimental results indicate that a higher contribution ratio from clients does not necessarily lead to better performance. Due to the significant variability in data quality among clients, the performance of the models trained by clients may differ. If a client's data quality is poor, increasing that client's contribution ratio could actually degrade the performance of the final aggregated model. Therefore, it is essential to strike a balance between the contributions of clients and the performance of both local and server models.

As shown in the figure, increasing the parameter β from 0.2 to 0.4, which enhances the contribution of the client model, leads to a significant improvement in model accuracy during training rounds 0–10. This is because, at the initial stage of FL training, the knowledge reserve of the local client models is greater than that of the server model. Thus, appropriately increasing β, i.e., the weight of the local model contributions—during this phase enhances both the performance and the training speed of the FL-trained model.

3) Hyperparameter λ comparison: Figure 9 (*c*) and (*f*) illustrate the impact of varying the parameter λ on the performance of the PCDM algorithm, where λ indicates the degree of model personalization. As shown in the figure, as the value of λ increases, the predictive performance of the PCDM algorithm decreases continuously. There is an inverse relationship between client personalization and model aggregation accuracy. In personalized training in FL, the level of personalization on the client side affects the overall model performance; a higher degree of personalization inevitably comes at the cost of reduced model accuracy. As shown in the figure, the model accuracy drops from 92.43% to 88.71% as the parameter λ increases from 5 to 15. This indicates that higher levels of personalization can negatively impact the model's overall accuracy.

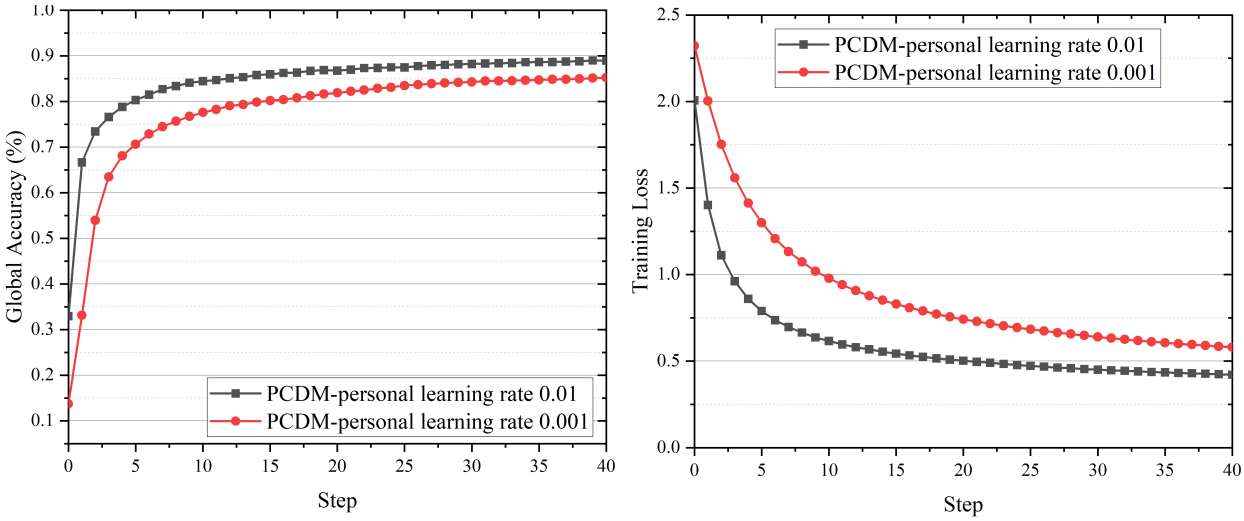
4) Hyperparameter η comparison: Figure 10 and 11 demonstrate the impact of varying learning rates on the performance of the PCDM algorithm in the server and clients, respectively. In Figure 10, as the learning rate increases, the predictive accuracy of the PCDM model improves effectively; however, after reaching a certain threshold, further increases in the learning rate yield diminishing returns on model performance. While an appropriate adjustment in learning rate can significantly enhance the federated learning model's ef-

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3572090



(*a*) Prediction accuracy of different learning rates η_g (*b*) Loss of different learning rates η_g

Fig. 10: Performance comparison of different learning rates for PCDM



(*a*) Prediction accuracy of different personal learning rates η_l (*b*) Loss of different personal learning rates η_l

Fig. 11: Performance comparison of different personal learning rates for PCDM

fectiveness, higher learning rates also result in longer training times.

When the learning rate η_g is set to 0.1, the model exhibits poor training robustness, with significant fluctuations in the training accuracy and loss curves. This instability arises because a high learning rate causes excessively large weight updates during optimization, making it difficult for the model to converge to or stabilize near the optimal point. While a larger learning rate can speed up training, it often leads to overly large updates, inefficient loss reduction, and even risks of gradient explosion. These issues can be addressed by lowering the learning rate, employing a learning rate scheduler, using adaptive optimizers, or applying gradient clipping techniques.

V. Related Work

Research in the smart home sector is experiencing robust growth globally, particularly with an increasing focus on collaborative task sequencing for integrating automated devices for lighting, heating, and appliance management [26, 27]. Prior work in this area has highlighted the benefits of task coordination for energy optimization, where machine learning models can schedule high-energy-consuming appliances based on real-time data such as energy prices, weather forecasts, and user habits. Such approaches allow for the effective utilization of renewable energy sources and reduced electricity costs while accommodating user preferences [6, 28]. However, such models encountered limitations with scalability and flexibility, especially when managing the high-dimensional and real-time data typical in smart home environments [29, 30].

The use of federated learning improves performance and model aggregation security in collaborative decision-making for smart homes. Wen *et al.* in [31] developed a federated learning-based security monitoring system for smart homes, using a federated learning model to jointly analyze data from home security devices and make intelligent decisions to address potential safety hazards. Liu *et al.* in [32] introduced

a federated machine learning technique that allows for model training on user devices, thus avoiding the transmission of sensitive user data to servers and enhancing privacy protection. Shahsavari *et al.* in [33] leveraged federated learning technology to enable the sharing of medical data across multiple healthcare institutions while safeguarding patient privacy.

Moreover, much current research focuses on federated learning across homogeneous devices, which may not fully address the diversity of devices within smart homes. There has been a growing interest in personalized privacy-preserving federated learning frameworks, which allow smart home systems to adapt to individual user preferences and optimize device interactions based on localized data [34, 35]. Unlike traditional federated learning, which assumes a uniform global model, personalized federated learning addresses the unique heterogeneity of smart home environments by accommodating varying user needs, device types, and data distributions [36, 37].

For example, Lu *et al.* in [38] proposed a personalized energy management system based on federated learning, which utilizes data from smart meters and user behavior patterns to intelligently predict and regulate household energy consumption while employing differential privacy techniques to protect user privacy. Yang *et al.* in [39] enhanced federated learning for smart homes by improving robustness to device heterogeneity and non-IID data, enabling personalized, privacy-preserving, and accurate decision-making within smart home systems. Unlike existing approaches, this paper proposes a privacy-preserving personalized federated learning (P3FL) framework for stability predictions in smart homes, enabling collaborative, privacy-preserving decision-making across diverse environments while adapting to individual user preferences in real-time.

VI. Final Remarks

This paper presents a novel Privacy-Preserving Personalized Federated Learning framework (P3FL) designed to address the challenges of collaborative sequential prediction in smart home environments. By leveraging federated learning, P3FL overcomes the limitations of localized data and significantly enhances energy consumption forecasting and appliance scheduling. Central to this framework is the Personalized Collaborative Decision-Making (PCDM) algorithm, which dynamically adapts to diverse household environments, ensuring both personalization and privacy preservation. The theoretical convergence bounds analysis of PCDM provides robust convergence guarantees under conditions of strong convexity, smoothness, and bounded variance. Extensive experiments on real-world smart home datasets demonstrate the superior performance of P3FL. These results underscore the framework's ability to improve operational efficiency, enhance user satisfaction, and deliver personalized privacy-preserving solutions in smart home applications.

Future works will focus on enhancing scalability and cross-domain adaptability of P3FL to support large-scale, heterogeneous smart home ecosystems while enabling privacy-preserving knowledge transfer across domains like smart grids. We aim to integrate advanced privacy-utility trade-off mechanisms (e.g., adaptive differential privacy paired with explainable AI) to address evolving adversarial threats without sacrificing prediction reliability or user transparency. Additionally, we will develop energy-aware dynamic optimization strategies to reduce the computational costs of federated training, adapting to real-time behavioral shifts and energy market fluctuations through lightweight algorithm refinements.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grant 62302301. This work was Sponsored by Shanghai Rising-Star Program under Grant 24QB2700400.

References

[1] A. Alsufyani, O. Rana, and C. Perera, "Knowledge-based cyber physical security at smart home: a review," *ACM Computing Surveys*, vol. 57, no. 3, pp. 1–36, 2024.

[2] H. Xu, J. Wu, Q. Pan, X. Guan, and M. Guizani, "A survey on digital twin for industrial internet of things: Applications, technologies and tools," *IEEE Communications Surveys & Tutorials*, 2023.

[3] K. Zhang, Y. Shi, S. Karnouskos, T. Sauter, H. Fang, and A. W. Colombo, "Advancements in industrial cyber-physical systems: An overview and perspectives," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 716–729, 2022.

[4] R. Chataut, A. Phoummalayvane, and R. Akl, "Unleashing the power of iot: A comprehensive review of iot applications and future prospects in healthcare, agriculture, smart homes, smart cities, and industry 4.0," *Sensors*, vol. 23, no. 16, p. 7194, 2023.

[5] X. Guo, Y. Zhang, F. Luo, and Z. Y. Dong, "User-centric recommendations on energy-efficient appliances in smart grids: A multi-task learning approach," *Knowledge-Based Systems*, vol. 284, p. 111219, 2024.

[6] M. Z. Fakhar, E. Yalcin, and A. Bilge, "A survey of smart home energy conservation techniques," *Expert Systems with Applications*, vol. 213, p. 118974, 2023.

[7] S. Constantinou, C. Costa, A. Konstantinidis, P. K. Chrysanthis, and D. Zeinalipour-Yazti, "A sustainable energy management framework for smart homes," *IEEE Transactions on Sustainable Computing*, 2024.

[8] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Generation Computer Systems*, vol. 115, pp. 619–640, 2021.

[9] S. Hu, X. Chen, W. Ni, E. Hossain, and X. Wang, "Distributed machine learning for wireless communication networks: Techniques, architectures, and applications," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1458–1493, 2021.

[10] M. C. Marin, M. Cerutti, S. Batista, and M. Brambilla, "A multi-protocol iot platform for enhanced interoperability and standardization in smart home," in *2024 IEEE 21st Consumer Communications & Networking Conference (CCNC)*. IEEE, 2024, pp. 1–6.

[11] A. Pundir, S. Singh, M. Kumar, A. Bafila, and G. J. Saxena, "Cyber-physical systems enabled transport networks in smart cities: Challenges and enabling technologies of the new mobility era," *IEEE Access*, vol. 10, pp. 16 350–16 364, 2022.

[12] H. Xu, J. Wu, Q. Pan, X. Liu, and C. Verikoukis, "Digital twin and meta rl empowered fast-adaptation of joint user scheduling and task offloading for mobile industrial iot," *IEEE Journal on Selected Areas in Communications*, 2023.

[13] D. C. Nguyen, Q.-V. Pham, P. N. Pathirana, M. Ding, A. Seneviratne, Z. Lin, O. Dobre, and W.-J. Hwang, "Federated learning for smart healthcare: A survey," *ACM Computing Surveys (Csur)*, vol. 55, no. 3, pp. 1–37, 2022.

[14] K. Sun, H. Xu, K. Hua, X. Lin, G. Li, T. Jiang, and J. Li, "Joint top-k sparsification and shuffle model for communication-privacy-accuracy tradeoffs in federated learning-based iov," *IEEE Internet of Things Journal*, 2024.

[15] K. Sun, H. Xu, X. Zhang, K. Hua, and J. Li, "Time-sensitive local differential privacy-based federated learning for vehicular digital twin networks," in *International Symposium on Intelligent Computing and Networking*. Springer, 2024, pp. 105–118.

[16] M. Le, T. Huynh-The, T. Do-Duy, T.-H. Vu, W.-J. Hwang, and Q.-V. Pham, "Applications of distributed machine learning for the internet-of-things: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, 2024.

[17] Z. Lu, H. Pan, Y. Dai, X. Si, and Y. Zhang, "Federated learning with non-iid data: A survey," *IEEE Internet of Things Journal*, 2024.

[18] S. Bubeck *et al.*, "Convex optimization: Algorithms and complexity," *Foundations and Trends® in Machine Learning*, vol. 8, no. 3-4, pp. 231–357, 2015.

[19] G. Wang, S. Lu, Y. Hu, and L. Zhang, "Adapting to smoothness: A more universal algorithm for online convex optimization," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 04, 2020, pp. 6162–6169.

[20] L. Bottou, F. E. Curtis, and J. Nocedal, "Optimization methods for large-scale machine learning," *SIAM review*, vol. 60, no. 2, pp. 223–311, 2018.

[21] Z. Yang, L. Li, X. Xu, S. Zuo, Q. Chen, P. Zhou, B. Rubinstein, C. Zhang, and B. Li, "Trs: Transferability reduced ensemble via promoting gradient diversity and model smoothness," *Advances in Neural Information Processing Systems*, vol. 34, pp. 17 642–17 655, 2021.

[22] "Temperature Forecasting Analysis — kaggle.com," https://www.kaggle.com/code/pavankumar20/temperature-forecasting-analysis, [Accessed 22-11-2024].

[23] "Smart Home IoT-EDA,ARIMAs,LSTM and more — kaggle.com," https://www.kaggle.com/code/piergiacomofonseca/smart-home-iot-eda-arimas-lstm-and-more/notebook, [Accessed 22-11-2024].

[24] A. Fallah, A. Mokhtari, and A. Ozdaglar, "Personalized federated learning: A meta-learning approach," *arXiv preprint arXiv:2002.07948*, 2020.

[25] C. T Dinh, N. Tran, and J. Nguyen, "Personalized federated learning with moreau envelopes," *Advances in neural information processing systems*, vol. 33, pp. 21 394–21 405, 2020.

[26] J. Xi, G. Xu, S. Zou, Y. Lu, G. Li, J. Xu, and R. Wang, "A blockchain dynamic sharding scheme based on hidden markov model in collaborative iot," *IEEE Internet of Things Journal*, 2023.

[27] H. Chen, C. Gouin-Vallerand, K. Bouchard, S. Gaboury, M. Couture, N. Bier, and S. Giroux, "Enhancing human activity recognition in smart homes with self-supervised learning and self-attention," *Sensors*, vol. 24, no. 3, p. 884, 2024.

[28] E. Al-Masri, A. Souri, H. Mohamed, W. Yang, J. Olmsted, and O. Kotevska, "Energy-efficient cooperative resource allocation and task scheduling for internet of things environments," *Internet of Things*, vol. 23, p. 100832, 2023.

[29] S. S. Gill, M. Golec, J. Hu, M. Xu, J. Du, H. Wu, G. K. Walia, S. S. Murugesan, B. Ali, M. Kumar *et al.*, "Edge ai: A taxonomy, systematic review and future directions," *Cluster Computing*, vol. 28, no. 1, pp. 1–53, 2025.

[30] K. Ren, J. Liu, Z. Wu, X. Liu, Y. Nie, and H. Xu, "A data-driven drl-based home energy management system optimization framework considering uncertain household parameters," *Applied Energy*, vol. 355, p. 122258, 2024.

[31] J. Wen, Z. Zhang, Y. Lan, Z. Cui, J. Cai, and W. Zhang, "A survey on federated learning: challenges and applications," *International Journal of Machine Learning and Cybernetics*, vol. 14, no. 2, pp. 513–535, 2023.

[32] J. C. Liu, J. Goetz, S. Sen, and A. Tewari, "Learning from others without sacrificing privacy: Simulation comparing centralized and federated machine learning on mobile health data," *JMIR mHealth and uHealth*, vol. 9, no. 3, p. e23728, 2021.

[33] Y. Shahsavari, O. A. Dambri, Y. Baseri, A. S. Hafid, and D. Makrakis, "Integration of federated learning and blockchain in healthcare: A tutorial," *arXiv preprint arXiv:2404.10092*, 2024.

[34] M. M. Fouda, Z. M. Fadlullah, M. I. Ibrahem, and N. Kato, "Privacy-preserving data-driven learning models for emerging communication networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, 2024.

[35] X. Wang, J. Lyu, J. D. Peter, and B.-G. Kim, "Privacy-preserving ai framework for 6g-enabled consumer electronics," *IEEE Transactions on Consumer Electronics*, 2024.

[36] T. M. Mengistu, T. Kim, and J.-W. Lin, "A survey on heterogeneity taxonomy, security and privacy preservation in the integration of iot, wireless sensor networks and federated learning," *Sensors*, vol. 24, no. 3, 2024. [Online]. Available: https://www.mdpi.com/1424-8220/24/3/968

[37] S. D. N, A. B, S. Hegde, C. S. Abhijit, and S. Ambesange, "Fedcure: A heterogeneity-aware personalized

federated learning framework for intelligent healthcare applications in iomt environments," *IEEE Access*, vol. 12, pp. 15 867–15 883, 2024.

[38] C. Lu, J. Cui, H. Wang, H. Yi, and C. Wu, "Privacy preserving user energy consumption profiling: From theory to application," *IEEE Transactions on Smart Grid*, 2023.

[39] Z. Yang, Y. Liu, S. Zhang, and K. Zhou, "Personalized federated learning with model interpolation among client clusters and its application in smart home," *World Wide Web*, vol. 26, no. 4, pp. 2175–2200, 2023.

Ye Wang (Member, IEEE) received the Bachelor's degree in Heating, Ventilation, and Air Conditioning Engineering from Qingdao Architectural Engineering Institute in China in 1997. He is currently pursuing a Ph.D. degree at the Department of Automation, Shanghai Jiao Tong University. He also serves as the CTO of Haier Group Corporation, General Manager of the National Innovation Institute of High-end Smart Appliances, and Deputy Director of the State Key Laboratory of Massive Personalized Customization System and Technology. He is in charge of the planning and development of smart home appliances, smart homes, and core components, leading the development of original IoT and digital key technologies for home appliances and their applications in Haier's products. Furthermore, he spearheads independent research efforts in core appliance components, resulting in the industrial application of self-controlled solutions such as inverter chips, IoT modules, and CA certification platforms. These efforts significantly contribute to the development of the upstream and downstream industry chain of home appliances.

Hansong Xu (Senior Member, IEEE) is currently a research Associate Professor of School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, China, where he was a postdoctor, and research Assistant Professor. He obtained his Ph.D. degree from the Department of Computer and Information Sciences at Towson University, MD, USA, in 2020 and received the Doctorial Research Fellowship. He was a recipient of the Shanghai Pujiang Talent Program award, Shanghai Rising-Star Program award, and Special Support from China Postdoctoral Science Foundation. His research interests include internet of things and machine learning.

Kun Hua (Senior Member, IEEE) is currently an Assistant Professor of the Electrical Engineering Department, at California Polytechnic State University. Before that, he has been an Associate Professor of the Electrical and Computer Engineering Department, at Lawrence Technological University, Southfield, MI, USA since 2010. He received the B.Sc. (with First Class Honors) and M.Sc. degrees from Electrical and Computer Engineering, Xi'an Jiao Tong University, China, in 1999 and 2004 respectively. He earned his Ph.D. degree in computer and electronic engineering from the University of Nebraska Lincoln, Nebraska, USA in 2008. He continued his research at the University of Nebraska Lincoln as a Post-doctoral researcher in 2009. His current research interests are in the areas of wireless communication, digital signal processing, and Internet of Things. Prof. Hua is a senior member of IEEE. He has served as guest editor of international journals and chair of several conferences.

Yang Bai (Member, IEEE) received the B.S. degree from Northeastern University, Shenyang, China, the M.S. degree and the Ph.D. degree from the College of Engineering, University of Miami, USA, in 2015 and 2021, respectively. She is an assistant professor in the Department of Automation, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, China. Her primary research interests include intelligent edge systems and industrial IoT.

Jianqi Yu (Member, IEEE) received the Bachelor's degree in Computer Science from Grenoble Alpes University in 2004, and her Ph.D. in Computer Software Engineering from the same university in 2010. She currently serves as the Chief Architect at the National Innovation Institute of High-end Smart Appliances, she is responsible for leading research and development efforts in foundational software, lightweight AI engines, and novel intelligent interaction frameworks for smart home domain. Jianqi Yu is deeply involved in numerous provincial and ministerial-level research projects. Her current research interests encompass intelligent fusion technologies for edge computing, multimodal perception, and decision-making, as well as integrated intelligent computing systems for edge-cloud synergy in AIoT applications within smart home environments.

Wenyin Zhu (Member, IEEE) received the Bachelor's degree in Electronic Information Engineering from Changchun University of Science and Technology in 2004. She completed her Master's degree in Business Administration from Southwest University in 2018, and since 2023, she has been pursuing a Ph.D. in Electronic Information at Ocean University of China. Currently, she holds positions as the General Manager of Technology Resources Platform at Haier Group's Technology Committee, General Manager of Haier Smart Home Policy Platform, Secretary-General of the Academic Committee of the National Key Laboratory of Large-Scale Personalized Customization Systems and Technologies, and Supervisor of the China Household Electrical Appliances Association. Her current research interests include innovative applications of artificial intelligence and service-oriented manufacturing.

Lixing Chen (Member, IEEE) received the BS and ME degrees from the College of Information and Control Engineering, China University of Petroleum, Qingdao, China, in 2013 and 2016, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Miami, in 2020. He is a tenure-track Associate Professor with the Institute of Cyber Science and Technology, Shanghai Jiao Tong University, China. His primary research interests include mobile edge computing and machine learning for networks.

Bo Yang (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the City University of Hong Kong, Hong Kong, in 2009. He is currently a Full Professor with Shanghai Jiao Tong University, Shanghai, China. From 2009 to 2010, prior to joining Shanghai Jiao Tong University, Shanghai, China, in 2010, he was a Postdoctoral Researcher with the KTH Royal Institute of Technology, Stockholm, Sweden, and a Visiting Scholar with the Polytechnic Institute of New York University, in 2007. His research interests include optimization and control for energy networks and Internet of Things. Dr. Yang is an Associate Editor for IEEE Transactions on Network Science and Engineering. He has been the Principle Investigator in several research projects, including the NSFC Key Project. He was a recipient of the Ministry of Education Natural Science Award 2016, the Shanghai Technological Invention Award 2017, the Shanghai Rising-Star Program 2015, and the SMC-Excellent Young Faculty Award by Shanghai Jiao Tong University.

Xinping Guan (Fellow, IEEE) is currently a Chair Professor with Shanghai Jiao Tong University, Shanghai, China, where he is the Dean of the School of Electronic, Information and Electrical Engineering, and the Director of the Key Laboratory of Systems Control and Information Processing, Ministry of Education of China. He has authored or co-authored four research monographs, over 270 papers in IEEE TRANSACTIONS and other peer-reviewed journals, and numerous conference papers. His current research interests include industrial cyber-physical systems, wireless networking and applications in smart city and smart factory, and underwater sensor networks. Dr. Guan was a recipient of the Second Prize of the National Natural Science Award of China in 2008 and 2018, the First Prize of the Natural Science Award from the Ministry of Education of China in both 2006 and 2016, the First Prize of the Technological Invention Award of Shanghai Municipal, China, in 2017, the IEEE Transactions on Fuzzy Systems Outstanding Paper Award in 2008, the National Outstanding Youth Honored by the NSF of China, the Changjiang Scholar by the Ministry of Education of China, and the State-Level Scholar of New Century Bai Qianwan Talent Program of China. As a Principal Investigator, he has finished/been working on many national key projects. He is the Leader of the prestigious Innovative Research Team of the National Natural Science Foundation of China. He is an Executive Committee member of the Chinese Automation Association Council and the Chinese Artificial Intelligence Association Council.