International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-15

https://doi.org/10.70917/ijcisim-2026-0365
Article

Privacy Protection Mechanism and Residential Security Enhancement Countermeasures for Smart Homes Combined with Internet of Things Technology

He Jiang [1] and Xiaoru Li [1,*]

[1] Department of Computer Science and Technology, Taiyuan University, Taiyuan, Shanxi, 030012, China
* Correspondence author: lixiaorucomputer@163.com

Abstract: As an important application scenario of the Internet of Things (IoT) technology, smart homes realize the interconnection and intelligent management of home devices through sensors, RFID chips and other devices. However, smart home devices face serious privacy leakage risks during data collection, transmission and processing, and sensitive data such as user behavioral data and home environment information are easily acquired and exploited by malicious attackers. In this study, a smart home privacy protection mechanism based on federated learning and differential privacy is proposed for the privacy protection of IoT devices in smart homes. The methodology adopts an adaptive hierarchical differential privacy adding noise algorithm to quantify the layer contribution by calculating the non-zero percentage of activation values and the amount of gradient change in each layer to realize the dynamic privacy budget allocation. Meanwhile, the wireless federated learning system model is established to characterize the channel properties between the base station and the user equipment using the block fading model, and combines the Gaussian and Laplace mechanisms to provide differential privacy protection. The experimental results show that on the MNIST dataset, the performance of this paper's algorithm reaches 95.16%, which is 3.56% higher than the competitive algorithm AUTO-S when the privacy budget takes the value of 10. In the smart home device recognition task, the method achieves an average adversarial rate of 98.625% in the white-box scenario and 89.39% in the black-box scenario. The conclusion shows that the privacy-preserving mechanism can effectively protect user privacy while ensuring the usability of the model, which provides reliable security for the smart home system and has good practicality and popularization value.

Keywords: Internet of Things (IoT) technology, smart homes, federated learning, differential privacy, privacy protection, residential security

1. Introduction

Today's social development in all walks of life has been introduced from different degrees of intelligent design, intelligent design can provide convenient and fast services for social construction, can lay a strong foundation for people's daily work and life comfort [1-3]. With the rapid development of intelligent buildings, intelligent design gradually plays an increasingly important role [4]. The integration and continuous development of the Internet of Things and the building network, so that the modern residential relative to the traditional residential more to meet people's daily needs, through the provision of personalized services for the user to meet the user's comfort requirements for the building

[5-6].

However, with the networking of smart homes with various devices in the community, a large amount of sensitive data is collected and transmitted to the cloud or servers, such as user identity information, living habits, and health conditions [7-9]. If these data protection measures are not in place, then they are easy to be hacked or leaked out from the inside, bringing serious privacy leakage risks to users [10-11].

In addition, smart devices may also unintentionally violate users' privacy when providing convenient services [12]. Smart cameras and microphones, for example, may capture users' private activities or conversations at inappropriate times, and this all raises users' concerns about privacy protection [13-15]. Therefore, it is important to actively research advanced encryption technologies and protection measures to build the information security management system of smart homes and communities to guarantee the security and privacy of user data, so as to enhance users' experience and satisfaction with smart homes [16-18].

The rapid development of information technology in modern society has gradually popularized the intelligent lifestyle, and smart housing, as a product of the deep integration of Internet of Things (IoT) technology and the living environment, is changing people's lifestyles. The current intelligent residential system integrates a large number of sensing devices, control units and communication modules, which can realize a variety of functions such as environmental monitoring, equipment control, security protection, etc., and provide users with a convenient and comfortable living experience. However, smart housing systems generate a large amount of data containing users' private information during operation, including sensitive information such as living habits, behavioral patterns, and equipment usage. Once these data are maliciously acquired or abused, they will pose a serious threat to user privacy and security. The traditional data processing method usually adopts a centralized architecture, where all data are gathered to a central server for unified processing and analysis. Although this method can provide better computing performance, there are problems such as insufficient data privacy protection, high risk of single-point failure, and heavy network transmission burden. Especially in the smart residential environment, users' requirements for privacy protection are increasing, and how to protect users' privacy while fully utilizing the value of data has become a key issue that needs to be solved urgently. Machine learning technology plays an important role in smart residential systems, and through the learning and analysis of user behavior data, it can achieve personalized service recommendation, intelligent control strategy optimization and other functions. However, traditional machine learning methods require access to raw data for model training, which is in fundamental conflict with the requirements of privacy protection.

Based on the above background, this study proposes a privacy protection mechanism for smart homes in combination with IoT technology, which solves the data privacy protection problem by introducing federated learning and differential privacy techniques. First, analyze the current status of the application of IoT technology in smart houses, sort out the technical architecture of data sensing layer, network transmission layer and application layer, and identify potential privacy leakage risk points. Second, a distributed machine learning framework based on federated learning is designed to enable each smart device to collaboratively train models without sharing raw data. Then, an adaptive hierarchical differential privacy noise addition algorithm is proposed to balance privacy protection and model performance by dynamically adjusting the privacy budget allocation of each layer. Finally, the wireless federated learning system model is constructed and the effectiveness of the proposed method in terms of privacy protection and system performance is verified through experiments.

2. Application of Internet of Things technology in smart homes

2.1. Internet of Things (IoT) technology

The Internet of Things (IoT) is a combination of network technology and logistics technology, in which items scattered throughout the world are linked to people's places of residence through the aggregation and transmission of network information. In effect, it is to form a precise point-to-point match between people's demand and the supply of goods. The current stage of the development of the Internet of Things brings together advanced applications such as intelligent sensing systems, identification systems, etc., which is an advanced product of the development of computer science and technology, Internet technology, and information technology to a specific stage [19].

The main physical structure of the Internet of Things contains three layers: the data sensing layer, the network transmission layer, and the actual application residence, which is a closed-loop system in which data generated by different nodes can be transmitted in real time and can be controlled remotely. The data sensing layer is a general term for the first data collection device of IoT to cope with actual

objects. By affixing QR codes, adding RFID chips, etc., to various objects, a unique code number is programmed for the object, and the data is scanned by sensors or determined by proximity, and transmitted to the Internet to achieve the identification and perception of the object.

The main technologies involved are: radio frequency identification (RFID), sensing and control, short-range wireless communication (WIFI, Bluetooth, Zigbee, etc.), the highest frequency of use of RFID and wireless sensor networks (WSD).RFID function is through the chip on the need to control the characterization of the object and localization, the wireless reading device close to the chip (Tag), you can get the basic information of the object to identify. The wireless reading device is close to the chip (Tag), then it can get the basic information of the item and recognize it. Sensor network is through a large number of sensors, as the perception of the node, as long as the target object relative to the sensor has changed, it will be immediately recorded to report to the program, to achieve real-time monitoring, so that the data collected is closest to the real situation, the background through the analysis of the software to be sorted out, you can get to the object's specific location and operability, the establishment of the object in the real world, the only state.

The network transmission layer is the medium for the timely transmission of information, similar to the neural network of an organism, and is one of the basic physical dimensions of the Internet of Things. Today's mobile network with the Internet, the integration of a variety of access devices, communication modules, composed of a high-speed flow of information network system, so that “things - things”, “people - things”, "people - people The three combinations of “thing-thing”, “person-thing”, and “person-human” can freely communicate with each other. The technologies used are: mobile communication networks such as 4G/5G, the Internet, and wide power grids.

The application layer is mainly about the data coming from above, analyzing and organizing them to form standardized data that people can use and identify. This is usually called the application data layer protocol, which can normalize the data obtained from different types of sensors and belongs to the software level closest to the user. Through different programs, the data is processed to provide services to different customers, mainly in the parts of cloud computing, data storage, problem discovery, and intelligent solutions.

2.2. Network transmission technology

Network transmission technology is to connect terminals such as computers in series to realize the sharing and mutual transmission of information resources in terminal devices. Network transmission technology has 3 characteristics.

First, the transmission speed is fast. Network transmission technology provides corresponding technical support for the information stored in different computer devices and improves the efficiency of information transmission.

Secondly, it realizes the rapid transmission and sharing of information. With the help of network transmission technology can be distributed in different locations of the information terminal series connection, the formation of a network, to realize the rapid transmission and sharing of information.

Third, support distributed information processing. Network transmission technology supports distributed information processing, and different terminals share the load of the server.

2.3. Integration of Internet of Things Technology and Building Intelligence Development

(1) Application in intelligent residential buildings

The technical application of the Internet of Things (IoT) provides the technical support of early warning protection for intelligent residential buildings, and the protective ability of residential buildings is more sturdy and the safety performance is more reliable.

First of all, it is the intelligent chip recognition technology, which is an important function of the IoT early warning protection technology. Through the real-time network, the information of residential users is timely transmitted to the system control background for identity verification and matching, so as to facilitate the intelligent identification judgment of unknown persons, avoiding the random entry and exit of unfamiliar persons in residential buildings, and enhancing the security effect of residential entry and exit.

Secondly, it is the intelligent early warning system, which is the combined application effect of the Internet of Things infrared sensor and radio frequency identification technology, which can issue timely early warning prompts within. If applied to the residence, it is equivalent to setting up two layers of early warning alarms for residential users, which can greatly enhance the quality of residential security protection for users.

(2) The application of smart home

The use of the network for remote control of the smart home is currently the most popular application. After the Internet of Things technology equipped with global positioning system, remote sensing technology, can realize the purpose of remote control and remote monitoring of various furniture at home. Even when people are out of work or traveling, they can still manipulate their home appliances through smart devices.

Internet technology also preserves the manual switching of lighting, utilities, and other systems within the home when controlling furniture. The use of IoT technology eliminates the need to remodel traditional switches, which can meet the needs of family members of different ages and different habits. The combination of manual switches and Internet technology not only avoids the embarrassment of local smart devices due to temporary failures or guests not being able to find the controllers, but also realizes the optimal management of the living space according to different living situations.

(3) Intelligent Monitoring System

Based on the Internet of Things technology, the intelligent building can transmit the captured video to the backstage through the communication channel of the network, and utilize the function of intelligent analysis of the video to check the content therein. Finally, it is then reacted to the user through wireless technology, which helps the user to discover the abnormalities in life in time, thus improving the convenience of problem solving. In the current application of intelligent buildings, residents install intelligent cameras at home, and they can use mobile terminals such as cell phones to remotely watch the situation at home, and if an unexpected situation occurs, it will use early warning prompts, and residents can deal with it in a timely manner.

2.4. Smart Home Terminal Technology

Smart home refers to the use of Internet of Things (IoT), sensors, communication technology, intelligent control and other technological means to connect traditional home equipment to the network, such as how to make the old-fashioned home, such as fans, lights, etc., by adding some sensors to be able to connect to the Internet to become smarter and more humane, thus improving the comfort, convenience and safety of home life in an intelligent way of life [20].

The Internet of Things (IoT) has become a solution to this challenge by connecting physical objects using electronics, sensors, software and communication networks. This has transformed the infrastructure of smart cities by introducing various technologies that improve the sustainability, productivity and comfort of city dwellers. The smart home senses environmental information through various sensors, transmits the data to a control center, and then intelligently controls it according to the user's needs and preset conditions to achieve automated management and remote control of home equipment.

The smart home functional structure is shown in Figure 1.

Smart home equipment refers to the ordinary home equipment in the past within the family by connecting to the network, the terminal device data information, user behavior information, equipment and equipment communication information together, so as to realize the function of remote control, automation control and intelligent interaction with the user. Home equipment smart home equipment usually consists of the following components: home equipment terminal, home routing gateway, cloud server, user terminal.

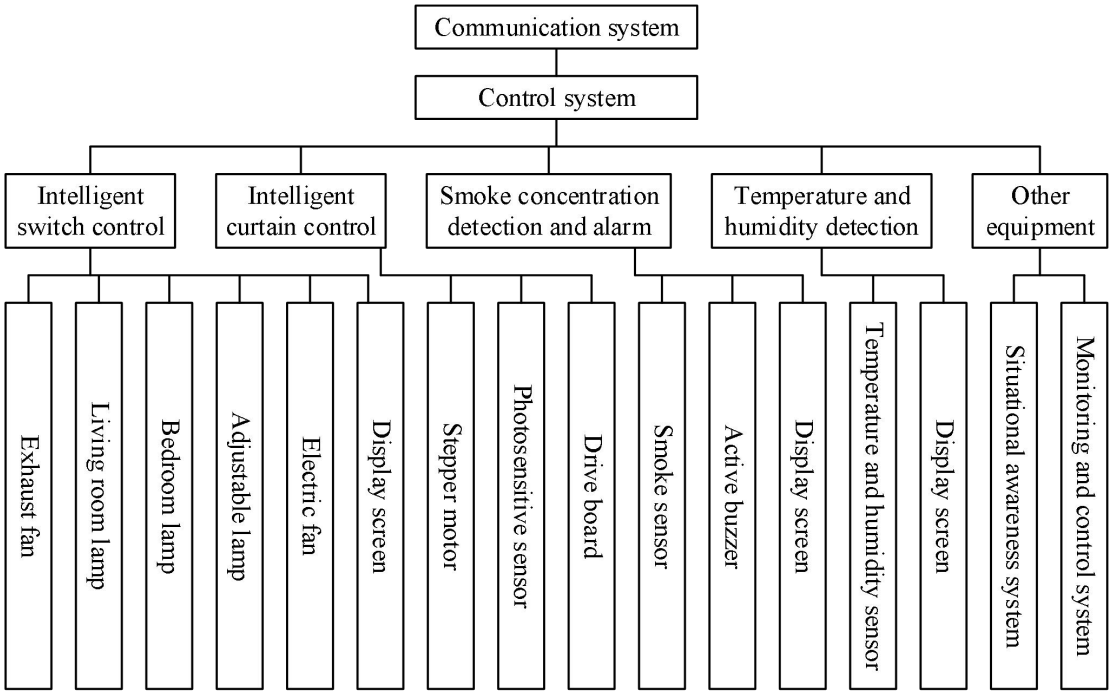

Figure 1. Intelligent home function structure

3. Proposed federated learning algorithms for smart homes

3.1. Federal Learning

Federated learning is a distributed machine learning technique, the goal of federated learning is to enable various participants to train models together to obtain higher quality AI models while ensuring client data privacy and security and legal compliance.

Assuming a total of C participant, a complete federated learning training process can be outlined as follows:

STEP1-Initialization: before starting the federated training, each participant c downloads the initial model w from the central server S to prepare for the model training.

STEP2-Training: In round t, each participant c samples a portion of data D_t from the local dataset D^c for model training to obtain model parameters w_t^c.

STEP3-Model upload: after the training is completed, each participant c uploads the trained model parameters w_t^c to the central server S.

STEP4-Model Aggregation: After the central server S collects the model parameter data w_t^c from each participant, it aggregates the models and obtains the aggregated and updated model w_{t+1}.

STEP5-Model update: Each participant c downloads the updated model w_{t+1} from the central server S to prepare for the next round of training.

STEP2-5 will be repeated until the model converges.

3.1.1. Federal Learning Classification

According to the difference in data distribution between federated learning participants, federated learning can be categorized into horizontal federated learning, vertical federated learning and federated migration learning.

Horizontal federation learning means that there is more overlap of data features between the data of each participant, i.e., the data features are aligned between participants, but there is less overlap of data samples, i.e., fewer identical users.

Vertical federation learning means that more data samples overlap between the data of each participant, i.e., more identical users, and the data samples are aligned between the participants. However, there is less overlap of data features between the participants, i.e., there are almost no identical features.

Federated migration learning means that there are fewer overlapping data samples and data features

among the participants, i.e., there are few identical features and identical users.

3.1.2. Privacy Preserving Techniques in Federated Learning

As federated learning exchanges model parameters and gradient parameters in plaintext between each participant and the central server during the training process, there is no encryption protection mechanism for model parameters and gradient parameters, which enables malicious participants to destroy the model training process or reverse speculate the data distribution information of other participants, thus indirectly leaking the privacy of client data. To address the indirect privacy leakage problem in federated learning, two main approaches are currently adopted to protect the parameter exchange process: information noise addition and information encryption.

The information noise technique aims to protect the data itself, usually through techniques such as differential privacy, adding designed noise during the model training process, so that the trained model maintains usability while avoiding the leakage of parameter information as well as the leakage of sensitive information. In the field of federated learning privacy protection, the differential privacy techniques used include centralized differential privacy techniques and localized differential privacy techniques.

Message encryption techniques mainly protect the process of parameter exchange, mainly through the use of cryptographic tools to provide privacy protection for the transmission process of parameters without destroying the original values and distribution of data. Two methods, homomorphic encryption and secure aggregation, are usually used to provide protection for the parameter transmission process in federated learning application scenarios.

3.1.3. Federal Learning in the Smart Home

Federated learning has been widely used in many fields, including smart home, healthcare, smart transportation, financial services, and so on. And, in the field of smart home, smart devices usually collect a large amount of user data, including home environment data and user behavior data. And traditional machine learning methods need to centralize these data to a central node for training, which has problems such as data privacy leakage and high network bandwidth consumption. The federated learning technology, on the other hand, can realize efficient training and updating of models under the premise of protecting data privacy, and thus has a broad application prospect in smart home.

Federated training process in smart home scenario:

First, federated learning can help smart home devices achieve more personalized intelligent services. Smart home devices can learn and adjust according to the user's personalized needs and habits to provide more intimate and intelligent services.

Second, federated learning can also help smart home devices improve security and privacy protection. Smart home devices usually involve users' private data and home security information, and how to protect these data from leakage and abuse is an important issue. Federated learning technology can realize model training and updating without sharing the original user data, effectively protecting user data privacy.

In addition, federated learning can help smart home devices achieve more intelligent capabilities. Smart home devices usually face problems such as data dispersion and large amount of data, and federated learning can integrate and learn the data in each device, thus enhancing the intelligence of the device.

In summary, federated learning, as an emerging machine learning framework, has a broad application prospect in the field of smart home. Through federated learning technology, smart home devices can achieve personalized services, improve security and privacy protection, and achieve more intelligent capabilities, thus bringing users a more convenient and intelligent life experience.

3.2. Differential Privacy

Differential privacy (DP) is a data privacy protection technique that allows useful information to be extracted from a database while minimizing the invasion of individual privacy.

The core idea of differential privacy lies in the fact that when an information thief attempts to retrieve a piece of data from a database, the original data is obfuscated to ensure that the thief cannot recognize any personal information from the retrieved information, ensuring that the data privacy is secure while maintaining the usefulness of the data. In federated learning, differential privacy resists inference attacks, ensuring that even if an attacker has additional external information, he or she cannot accurately infer the original private data information.

Centralized Differential Privacy (CDP) is defined as follows:

Definition 1: For a privacy mechanism M, the domain of definition is X and the domain of values is

R. If the privacy mechanism M is in any neighboring databases D, $D^{'} \in X$ and the output result $S \subseteq R$, there is:

$$Pr[M(D) \in S] \leq e^{\varepsilon} \cdot Pr[M(D^{'}) \in S] + \delta \tag{1}$$

Then the mechanism M is said to satisfy (ε, δ)-differential privacy. where ε is the privacy budget, which controls the trade-off between privacy and availability, and ε is closer to 0, the better the privacy, but the greater the perturbation, the worse the data availability. δ denotes the slack term, i.e., the probability of privacy protection failure.

Localized differential privacy (LDP) is defined as follows:

Definition 2: For a privacy mechanism M with definition domain X and value domain R, if the privacy mechanism M is in any neighboring databases D, $D^{'} \in X$ and the output result $S \subseteq R$, there are:

$$Pr[M(D) = S] \leq e^{\varepsilon} \cdot Pr[M(D^{'}) = S] + \delta \tag{2}$$

Then the mechanism M is said to satisfy (ε, δ)-local differential privacy.

The Laplace mechanism achieves differential privacy by adding random noise obeying the Laplace distribution to the query result.

Definition 3: For a given query function f, ($\varepsilon, 0$)-differential privacy is achieved by adding noisy data that obeys the Laplace distribution to the output of the function. The formula for adding Laplace noise is as follows:

$$f_p(D) = f(D) + Lap(\frac{\Delta f}{\varepsilon}) \tag{3}$$

where $Lap(\frac{\Delta f}{\varepsilon})$ denotes the Laplace distribution with center at 0 and scale parameter $\lambda = \frac{\Delta f}{\varepsilon}$, ε is the privacy budget, and Δf is the sensitivity of function f.

Definition 4: Sensitivity. Sensitivity in differential privacy is a central concept that quantifies the maximum possible amount of variation in results when the same query is executed on neighboring datasets. For query function f and neighboring dataset $D, D^{'}$, the sensitivity formula is:

$$\Delta s = \max_{D,D^{'}} \| f(D) - f(D^{'}) \| \tag{4}$$

where $\| * \|$ denotes the number of paradigms, with sensitivities using the L_1 -paradigm in the Laplace and exponential mechanisms and the L_2 -paradigm in the Gaussian mechanism. The definition of sensitivity is directly related to how much random noise needs to be added to satisfy a particular differential privacy guarantee, and the magnitude of the sensitivity indicates the maximum effect that a change in a single record in the dataset may have on the query result.

The Gaussian mechanism experiments with differential privacy by adding random noise obeying a Gaussian distribution to the query results. Compared to the Laplace mechanism, the Gaussian mechanism is suitable for providing (ε, δ)-differential privacy, where δ is a small positive number, called the slack term, which indicates the probability of failure of differential privacy. The Gaussian mechanism is defined as follows:

Definition 5: For a given query function f, (ε, δ)-differential privacy can be achieved by adding noisy data conforming to a Gaussian distribution to the output of the function. The formula for adding Gaussian noise is as follows:

$$f_p(D) = f(D) + N(0, \sigma^2) \tag{5}$$

where $N(0, \sigma^2)$ denotes a Gaussian distribution with mean 0 and variance σ^2, and σ is jointly determined by ε, δ and sensitivity Δf, calculated as:

$$\sigma = \sqrt{2 \ln \frac{1.25}{\delta}} \times \frac{\Delta f}{\varepsilon} \tag{6}$$

Two important theorems in the application scenario of differential privacy are are serial combination theorem and parallel combination theorem. The serial combination theorem describes the

accumulation of privacy budget that occurs when a dataset is processed sequentially by multiple differential privacy algorithms, the serial combination theorem is defined as shown below.

Definition 6: Suppose that there are multiple algorithms $f_i(D)(i \in \{1,2,...,k\})$ that satisfy (ε_i, δ_i)-differential privacy on the dataset D. Then their serial combination results in an algorithm $F = \{f_1(D), f_2(D), ..., f_k(D)\}$ that satisfies ($\sum_{i=1}^{k} \varepsilon_i, \sum_{i=1}^{k} \delta_i$)-differential privacy.

The parallel combination theorem is applied to the case when different differential privacy algorithms act on disjoint subsets of the dataset, and the parallel combination theorem is defined as shown below.

Definition 7: Suppose there are multiple algorithms $f_i(D)(i \in \{1,2,...,k\})$ that satisfy (ε_i, δ_i)-differential privacy on disjoint datasets $D_i(i \in \{1,2,...,k\})$ respectively.

Then the algorithm $F = \{f_1(D), f_2(D), ..., f_k(D)\}$ after their parallel combination satisfies ($\max \varepsilon_i, \max \delta_i$)-differential privacy.

3.3. User Privacy Protection Methods in Wireless Federated Learning

In order to be able to effectively protect user privacy without sacrificing model performance, this paper proposes a federated learning privacy preserving method with dynamic noise allocation.

3.3.1. Wireless Federated Learning System Modeling

(1) In this paper, we consider a cellular network model, which consists of a base station (BS) and a user equipment (UE). Each BS is associated with K user equipment through a wireless cellular link, and the location of the BS obeys a uniform Poisson point process with density λ. Φ_s It is assumed that the BS has sufficient computational resources to provide communication and computation services to users, and the UE has sufficient computational resources to accomplish the local federated learning training task.

(2) In wireless communications, signals are usually affected by multipath fading, multi-user interference and noise. In this paper, a block fading model is used to characterize the channel properties between the BS and the UE. In this model, small-scale Rayleigh fading and path loss are considered comprehensively. In the small-scale Rayleigh fading model, the channel gain between BS and UE can be expressed as:

$$h_{bs,ue} = \sqrt{z_{bs,ue}^{-\alpha}} \tilde{h}_{bs,ue} \tag{7}$$

where $\tilde{h}$ is the random scattering component, z is the distance between two cells, and α denotes the path loss index.

3.3.2. Adaptive Hierarchical Differential Privacy Noise Addition Algorithm

Gradient trimming is a commonly used technical tool in differential privacy federated learning systems to prevent gradient explosion on one hand and reduce the risk of privacy leakage on the other hand by restricting the range of gradient.

In this paper, we propose an adaptive layered gradient cropping (ALGC) algorithm. Remembering that the number of data samples is D and the size of each training batch is B, the L_2 -parameter of the gradient of each layer of the local model is computed separately during the local iterative training process, and the cropping threshold is the ς th percentile of the L_2 -parameter of the historical gradient:

$$Clip_{b_i}^{(l)} = [\Psi_0, \Psi_1 \cdots, \Psi_i]_{\varsigma}, i \geq 0 \tag{8}$$

where $Clip_{b_i}^{(l)}$ denotes the i nd batch of data b_i after training is completed. The cropping threshold of the layer l model, Ψ_i is the L_2 paradigm of the gradient. Therefore, the network model can adaptively set the cropping threshold according to the gradient changes in each layer, and the gradient cropping is calculated as:

$$\bar{g}^{(l)} = \frac{g^{(l)}}{\max(1, \frac{\| g^{(l)} \|_2}{Clip_{b_i}^{(l)}})} \tag{9}$$

The L_2-parameter of the gradient reflects the magnitude of the current parameter update.

Considering the impact of the non-zero occupancy of the activation values of each layer and the amount of gradient change on the model output, we propose a new method of calculating the layer contribution and allocating the privacy budget based on the layer contribution. ε By regulating the allocation of the privacy budget, we achieve the improvement of the model usability while guaranteeing the level of privacy protection.

(1) Calculate the percentage of non-zero neurons

The neural network model in this paper chooses $ReLU(x) = max(0, x)$ as the activation function, and defines the variable nzr to represent the proportion of the number of activation values greater than 0 in the total neurons (non-zero ratio), which can be expressed as:

$$nzr^{(l)} = \frac{1}{I \times N} \sum_{i=1}^{I} \sum_{n=1}^{N} 1(o_{b_i,n}^{(l)} > 0) \tag{10}$$

where $nzr^{(l)}$ denotes the non-zero ratio of the layer l model, $I = \frac{D}{R}$ is the total number of iterations in a round of training, N denotes the total number of activation units of the layer l model, and $o_{b_i,n}^{(l)}$ denotes the output value of the i th activation unit of the layer l model after it has been trained by the data b_i. Function $1(\cdot)$ is an indicator function, $1(\cdot)$ is 1 when the input condition is true, and $1(\cdot)$ is 0 when the input condition is false.

(2) Calculate the gradient change

Record the gradient g and g' before and after the model training respectively, to get the gradient change quantity $\Delta g = g' - g$, in order to facilitate the analysis and calculation, this paper uses the mean value of the gradient change quantity $\bar{\Delta} g$ for quantitative analysis, then the gradient change quantity of the layer l model can be expressed as:

$$\bar{\Delta} g^{(l)} = \frac{1}{N} \sum_{n=1}^{N} \Delta g_n^{(l)} \tag{11}$$

(3) Computational Layer Contribution

This section also introduces the gradient change quantity as a measure of the learning speed of each layer, and proposes to use the non-zero ratio and the gradient change quantity together to quantify the computational layer contribution C, which is calculated as follows:

$$C^{(l)} = \xi nzr^{(l)} + (1 - \xi) \overline{\Delta g}^{(l)} \tag{12}$$

where ξ is a moderating factor between 0 and 1 that adjusts for the effect of the two factors on the layer contribution.

(4) Allocating privacy budget

Different privacy budgets are allocated according to the contribution degree of each layer, and layers with higher contribution degrees are allocated more privacy budgets, which in turn reduces the amount of noise introduced in these key layers, and vice versa. In the model with L layers, the formula for allocating the privacy budget is as follows:

$$\varepsilon^{(l)} = \varepsilon \times \frac{C^{(l)}}{\sum_{l=1}^{L} C^{(l)}} \tag{13}$$

where ε is the total privacy budget and $\varepsilon^{(l)}$ denotes the privacy budget of layer l.

In this paper, we use Gaussian noise based differential privacy strategy, then the sensitivity of the gradient of the layer l model on neighboring datasets can be expressed as:

$$\begin{aligned}\Delta s^{(l)} &= \max_{\mathrm{D},\mathrm{D}'} \| \omega(\mathrm{D}) - \omega(\mathrm{D}') \| \\ &= \frac{\eta}{B} \sum_{m=1}^{D/B} \max_{\mathrm{D},\mathrm{D}'} \| g^{(l)}(b_m) - g^{(l)}(b'_m) \| \\ &\leq \frac{2E\eta Clip^{(l)}}{B}\end{aligned} \tag{14}$$

Then the expression for adding noise to the layer l model is:

$$\tilde{\omega}^{(l)} = \omega^{(l)} + N(0, (\Delta s^{(l)} \sigma_n^{(l)})^2) \tag{15}$$

3.3.3. Federated Learning Privacy Preserving Approach with Dynamic Noise Allocation

In order to protect the user privacy security in the aggregation phase, the adaptive differential privacy adding noise algorithm is applied to the wireless federated learning scenario, and the federated learning privacy preservation method with dynamic noise allocation is proposed.

The method is implemented based on differential privacy, and the method satisfies the privacy requirements of differential privacy, i.e., if the standard deviation of Gaussian noise added by UE k to model parameter ω_k satisfies:

$$\sigma_k^{(l)} \geq \frac{2\eta Clip_k^{(l)} \sqrt{2\ln(\frac{1.25}{\delta})}}{B\varepsilon_k^{(l)}} \tag{16}$$

Then the process of data training by UE k using local dataset D_k satisfies ($\varepsilon^{(l)}, \delta$)-differential privacy.

4. Improvement and application of federated learning algorithms in smart home

4.1. Simulation setup and result analysis

In this paper, three datasets, MNIST, Fashion-MNIST, and CIFAR-10, are used for experiments and they are Non-IID preprocessed. All three datasets are 10-classified image datasets, where MNIST and Fashion-MNIST are 28*28 pixels grayscale images and CIFAR-10 is 32*32 pixels color images.

Client data Non-IID processing: in this paper, we consider the scenario of client data Non-IID, for the three datasets of MNIST, Fashion-MNIST, and CIFAR-10, the training data are first categorized and sorted according to numerical labels. Then the data is divided according to the number of clients in the federated system to ensure that each client data contains at most two numeric labels.

4.1.1. Impact of the way the quartiles are taken on the model performance

In order to verify the effectiveness of the algorithm, the following six sets of comparison experiments were conducted under the three datasets. For the fixed quantile values, the values of 0.3, 0.4, 0.5 and 0.6 are taken respectively. For the upper and lower bounds of the dynamic quantile values, the values of [0.3,0.6] and [0.6,0.3] are taken respectively, to compare the accuracy of model classification with different quantile taking methods.

The effect of different quantile values on the model performance is shown in Fig. 2, when the algorithm of this paper takes the value of [0.3,0.6], the classification accuracies in the three datasets are 89.59, 75.26, and 32.18, respectively. the quantile values taken in this paper that gradually increase with iterative training achieve the highest classification accuracies in all the three datasets, with the value of [0.6,0.3], which is the lowest. i.e., the lowest model performance was obtained with the approach of gradually decreasing with iterative training.

As the training proceeds, the model gradually stabilizes and the gradient distribution tends to level off. Therefore, when performing quantile cropping, the quantile values should be adjusted dynamically, and they should be adjusted in a gradually increasing manner so that more gradient information is retained when the model is gradually stabilized.

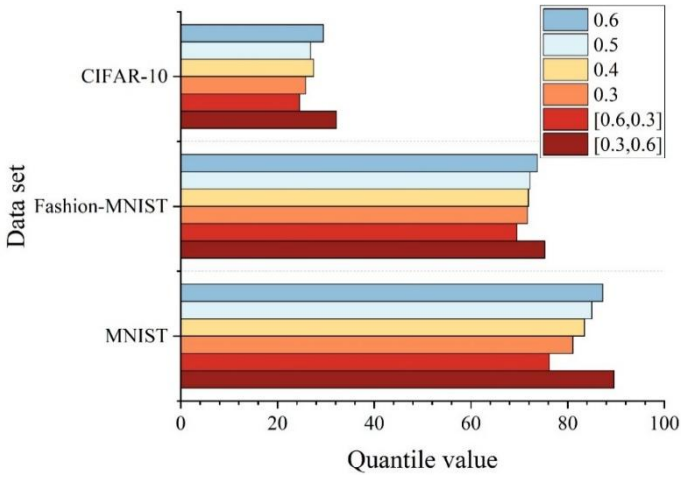

Figure 2. The impact of different quantization values on model performance

4.1.2. Performance comparison under different privacy budgets

In order to further validate the performance of the dynamic noise allocation based privacy preserving method for federated learning, it is compared with DP-SGD, DPAGD-CNN, and AUTO-S, respectively. Among them, DP-SGD is an optimization algorithm that combines privacy preservation and machine learning, and this work aims to solve the problem of privacy preservation in machine learning. In particular, differential privacy is used in deep learning.The basic idea of the DP-SGD algorithm is to introduce a differential privacy protection mechanism into the traditional stochastic gradient descent (SGD) algorithm.DPAGD-CNN is a differential privacy-based adaptive gradient descent algorithm for convolutional neural networks, which optimizes the noise allocation as well as the selection of the step size by dynamically adjusting the privacy budget.AUTO-S is an automatic trimming algorithm that makes model training under differential privacy independent of a specific trimming threshold.

In the experiments ε takes the values of 3, 5, 8, and 10 respectively.The performance comparison under different privacy budgets in the MNIST dataset is shown in Fig. 3. The performance of this paper's algorithm is 86.51%, 90.28%, 92.47%, and 95.16% when ε takes the values of 3, 5, 8, and 10, respectively. It is 3.56% higher than the competitive algorithm AUTO-S when ε takes the value of 3.

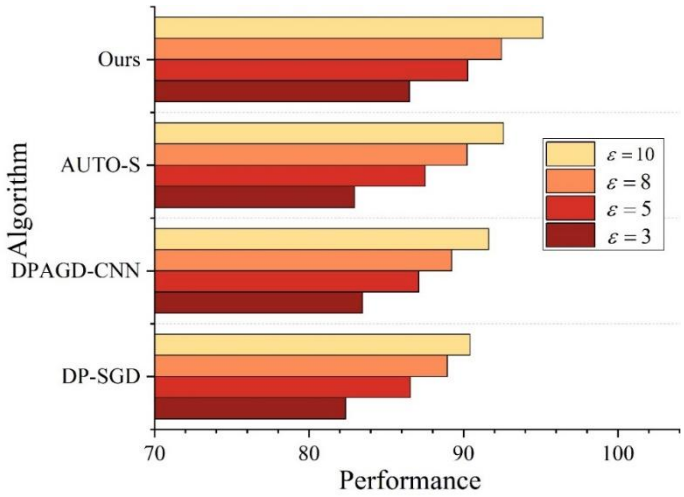

Figure 3. Mnist data focuses on performance comparisons of different privacy budgets

The performance comparison under different privacy budgets in Fashion-MNIST dataset is shown in Fig. 4. The performance values of the algorithms are highest when ε takes the value of 10. The performance of DP-SGD, DPAGD-CNN, and AUTO-S algorithms are 73.2%, 74.79%, and 74.21%, respectively. This paper's algorithm outperforms the comparison algorithms by 2.91%, 1.32%, and 1.9%, respectively. In comparison, the algorithm of this paper has high performance.

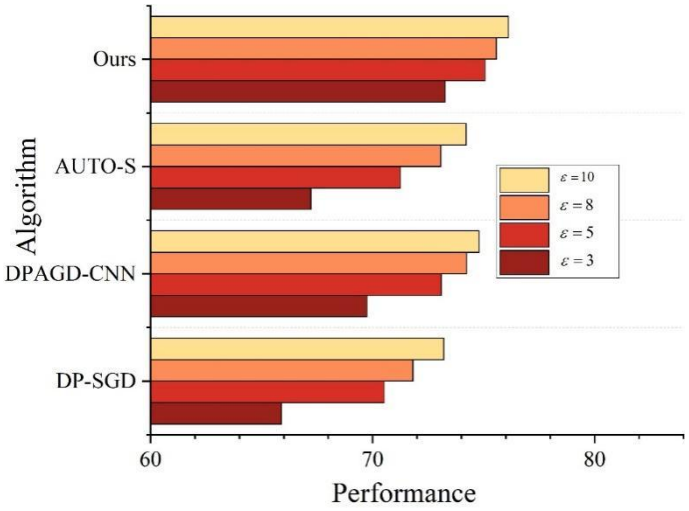

Figure 4. The performance of the fashion-mnist data concentration algorithm

The performance comparison under different privacy budgets in the CIFAR-10 dataset is shown in Fig. 5. The values of ε in the experiment are taken as 3, 5, 8, and 10, respectively. the smaller the privacy budget the higher the privacy protection level and the higher the noise.

With the increase of privacy budget, the noise gradually decreases and the model classification accuracy gradually increases. And under different privacy budgets, this paper's algorithm achieves the highest performance in all three datasets, indicating that this paper's algorithm still has stable and high performance under different privacy protection levels.

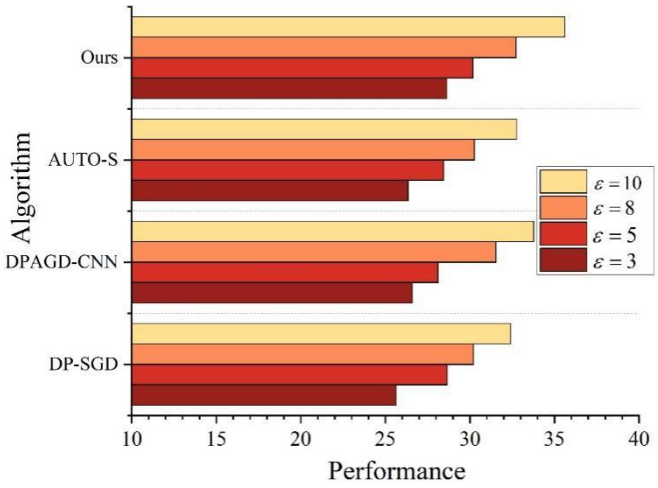

Figure 5. The performance comparison of the cizz-10 data concentration algorithm

4.2. Smart Home Device Identification Privacy Protection

In order to validate the effectiveness of a federated learning privacy preserving approach based on dynamic noise allocation, a large amount of smart device data traffic data needs to be collected. This chapter uses a publicly available smart device dataset that contains traffic packets of 40 smart devices and more than 15,000 device state tags labeled using automatic tagging. It contains smart home devices such as smart cameras and smart door locks.

Since device state identification relies on specific device type information, and defending against device type identification attacks can effectively prevent the leakage of device state information, this chapter uses only the traffic packets of smart devices in the dataset for subsequent experimental validation and testing.

The experiments in this chapter were all done under the Ubuntu server platform and run in a Python and pytorch based environment with Python version 3.8.5, pytorch version 1.7 and CUDA version 11.0.2.

The overall adversarial rate is used to evaluate the defense performance of the adversarial samples against the device identification models to demonstrate the device identification adversarial capability

of the device traffic privacy protection model in both white-box and black-box scenarios. The testing of resnet50 recognition model and vgg16 recognition model belongs to white-box testing, and the testing of the other three recognition models belongs to black-box testing. Meanwhile, in order to demonstrate the superiority of the designed federated learning privacy preserving method based on dynamic noise allocation as a comparison experiment. The resnet50 device recognition model is used as the target model for SparseFool to generate adversarial samples, and its model migration ability in black-box scenarios is tested in the other 3 device recognition models.

The performance of the home device traffic privacy protection model is shown in Table 1. The average confrontation rate reaches 98.625% in the white-box scenario and 89.39% in the black-box scenario, which effectively protects the user's privacy and security.

Table 1. The performance of the privacy protection model of household equipment

Test scenario		**Mean time/s**	**Device identification model**				**Average antagonism/%**
			resnet50 /%	**Resnet18 /%**	**Vgg16 /%**	**Densenet121 /%**	
White box scene	Resnet50	0.00012	98.98	98.42	98.06	99.04	98.625
	Vgg16	0.00026	94.24	95.11	92.17	93.52	93.76
	SpareFool	0.00091	93.25	90.76	92.25	94.25	92.628
Black box scene	Resnet50	0.00012	90.24	89.36	88.52	89.44	89.39
	Vgg16	0.00026	68.15	75.43	60.38	52.12	64.02
	SpareFool	0.00091	63.20	44.23	52.89	30.04	47.59

5. Conclusion

The federated learning privacy protection method based on dynamic noise allocation demonstrates excellent performance in smart residential environments. With the adaptive hierarchical differential privacy noise addition algorithm, the method is able to dynamically allocate the privacy budget according to the contribution of each layer, and the algorithm performance reaches 76.11% when the privacy budget is 10 on the Fashion-MNIST dataset, which is 2.91%, 1.32%, and 1.9% better than the DP-SGD, DPAGD-CNN, and AUTO-S algorithms, respectively. In the smart home device traffic privacy protection experiments, the method achieves an average confrontation rate of 98.625% in white-box attack scenarios and 89.39% in black-box attack scenarios, which effectively defends against device identification attacks and protects the user's device usage privacy. The experimental results further confirm that the algorithm maintains a stable high-performance performance under different privacy budgets on the CIFAR-10 dataset, indicating that the method has good robustness and adaptability. The privacy-preserving mechanism avoids the direct transmission of raw data through the federated learning framework, and combines with the differential privacy technique to provide mathematical privacy guarantee for the transmission of model parameters. The experiments verified that the method can effectively protect user privacy under the premise of guaranteeing model usability, which provides technical support for the safe operation of smart residential systems. This research is of great significance in promoting the healthy development of smart home technology and the protection of user privacy rights and interests, and provides new ideas and directions for the future development of smart home privacy protection technology.

Acknowledgements

1. Project of the 14th Five Year Plan for Education Science in Shanxi Province: Research on the Application of Sensor Technology System in the Smart Transformation of University Libraries (Fund No. GH-220170);
2. 2023 Shanxi Province Higher Education Teaching Reform and Innovation Project: "Research and Practice on the Achievement oriented Education Model of 'Specialized Innovation Integration' - Taking the Course of 'Sensor Technology' as an Example (Fund No. J20231425)";
3. “Supported by Open Project Foundation of Key Laboratory of Computation Intelligence and Chinese Information Processing of Ministry of Education and Key Laboratory of Data Intelligence and Cognitive Computing of Shanxi Province”;
4. National Computer Education Research Project for Higher Education Institutions: OBE Driven Blended Learning Model for IoT Technology Courses: Innovative Practice of Applied Undergraduate Education (Fund No. SXCF202411);
5. Special Project for Implementing the Spirit of the Third Plenary Session of the 20th Central Committee of the Communist Party of China in the Philosophy and Social Sciences Planning of Shanxi Province in 2024: "Promoting Innovative Research on Emotional Computing and Social Governance - Group Behavior and Public Opinion Analysis Based on Multimodal Data Analysis (Fund No.: 2024QH061 (YB));

6. Taiyuan University Project: Application and Research of Internet of Things Technology in Smart Campus (Fund No.: 23TYYB03).

References

1. Eini, R., Linkous, L., Zohrabi, N., & Abdelwahed, S. (2021). Smart building management system: Performance specifications and design requirements. Journal of Building Engineering, 39, 102222.
2. Jia, M., Komeily, A., Wang, Y., & Srinivasan, R. S. (2019). Adopting Internet of Things for the development of smart buildings: A review of enabling technologies and applications. Automation in construction, 101, 111-126.
3. Al Dakheel, J., Del Pero, C., Aste, N., & Leonforte, F. (2020). Smart buildings features and key performance indicators: A review. Sustainable Cities and Society, 61, 102328.
4. Aliero, M. S., Asif, M., Ghani, I., Pasha, M. F., & Jeong, S. R. (2022). Systematic review analysis on smart building: Challenges and opportunities. Sustainability, 14(5), 3009.
5. Stojkoska, B. L. R., & Trivodaliev, K. V. (2017). A review of Internet of Things for smart home: Challenges and solutions. Journal of cleaner production, 140, 1454-1464.
6. Padmanaban, S., Nasab, M. A., Shiri, M. E., Javadi, H. H. S., Nasab, M. A., Zand, M., & Samavat, T. (2023). The role of internet of things in smart homes. Artificial intelligence‐based smart power systems, 259-271.
7. Song, T., Li, R., Mei, B., Yu, J., Xing, X., & Cheng, X. (2017). A privacy preserving communication protocol for IoT applications in smart homes. IEEE Internet of Things Journal, 4(6), 1844-1852.
8. Magara, T., & Zhou, Y. (2024). Internet of things (IoT) of smart homes: privacy and security. Journal of Electrical and Computer Engineering, 2024(1), 7716956.
9. Yao, Y., Basdeo, J. R., Kaushik, S., & Wang, Y. (2019, May). Defending my castle: A co-design study of privacy mechanisms for smart homes. In Proceedings of the 2019 chi conference on human factors in computing systems (pp. 1-12).
10. Mocrii, D., Chen, Y., & Musilek, P. (2018). IoT-based smart homes: A review of system architecture, software, communications, privacy and security. Internet of Things, 1, 81-98.
11. Geneiatakis, D., Kounelis, I., Neisse, R., Nai-Fovino, I., Steri, G., & Baldini, G. (2017, May). Security and privacy issues for an IoT based smart home. In 2017 40th international convention on information and communication technology, electronics and microelectronics (MIPRO) (pp. 1292-1297). IEEE.
12. Hui, T. K., Sherratt, R. S., & Sánchez, D. D. (2017). Major requirements for building Smart Homes in Smart Cities based on Internet of Things technologies. Future Generation Computer Systems, 76, 358-369.
13. He, J., Xiao, Q., He, P., & Pathan, M. S. (2017). An adaptive privacy protection method for smart home environments using supervised learning. Future Internet, 9(1), 7.
14. Zeng, E., Mare, S., & Roesner, F. (2017). End user security and privacy concerns with smart homes. In thirteenth symposium on usable privacy and security (SOUPS 2017) (pp. 65-80).
15. Lin, C., He, D., Kumar, N., Huang, X., Vijayakumar, P., & Choo, K. K. R. (2019). HomeChain: A blockchain-based secure mutual authentication system for smart homes. IEEE Internet of Things Journal, 7(2), 818-829.
16. Tao, M., Zuo, J., Liu, Z., Castiglione, A., & Palmieri, F. (2018). Multi-layer cloud architectural model and ontology-based security service framework for IoT-based smart homes. Future Generation Computer Systems, 78, 1040-1051.
17. Ali, B., & Awad, A. I. (2018). Cyber and physical security vulnerability assessment for IoT-based smart homes. sensors, 18(3), 817.

18. Zaidan, A. A., Zaidan, B. B., Qahtan, M. Y., Albahri, O. S., Albahri, A. S., Alaa, M., ... & Lim, C. K. (2018). A survey on communication components for IoT-based technologies in smart homes. Telecommunication Systems, 69, 1-25.
19. Seyed Payam Fatemi,Nahideh Derakhshanfard,Fahimeh Rashidjafari & Ali Ghaffari. (2025). Distributed data storage using decision tree models and support vector machines in the Internet of Things. Sustainable Computing: Informatics and Systems,46,101134-101134.
20. Hala Elhadidy,Hassan Badran,Eslam Atef,Ahmed Essam,John Ramez,Mohammed Ayman... & Amira Elsonbaty. (2025). A Systematic Literature Review of Smart Power Monitoring and Controlling Systems for Smart Homes. Journal of Engineering Research and Reports,27(4),438-445.