Dual Mode Data Acquisition and Analysis Based on Deep Learning for Smart Home Networks

Hui Zhang, Meikun Li, Maoheng Yang, and Chao Zhou

***Abstract*—As the sharp increase in data traffic, except for security, home network and data usage will become two important issues in smart home in the future. On the one hand, aiming to solve the problem of wall jack and room coverage in smart home networks, we develop a set of dual-mode data acquisition device, which supports wired power line communication (PLC) and wireless dual-mode communication. Our PLC module is with the advantage of being pluggable on the debugging board, which can be replaced with a module that supports different PLC protocols as required, such as broadband PLC or narrowband PLC. On the other hand, aiming to solve massive data usage problem acquired by smart home facilities, we propose a data analysis method that integrates the automatic encoder scheme and the *K*-means algorithm. On this basis, we propose a Conv-LSTM-Attn algorithm that is a prediction model of convolutional long- and short-term memory network based on attention mechanism. Through the field test, it illustrates that our system can realize accurate long-term data acquisition and recording, dual-mode control from indoor to long-distance away, the autonomous backhaul of the collected data with low-latency and long-distance transmission. Besides, it can highly flexibly configure various sensor modules to achieve smart home management and control. Moreover, through the analysis of actual sampling data, the proposed Conv-LSTM-Attn algorithm outperforms the existing machine learning schemes in smart home in terms of achieving the prediction values.**

***Index Terms*—Deep learning, power line communication (PLC), smart home, wireless communication.**

I. INTRODUCTION

WITH the rapid growth of big data traffic, smart home is facing challenges from security, home network and data usage. Generally, the smart home refers to the home in which different components have been automated, which can automatically optimize comfort of the residents by using context awareness as well as predefined constraints based on the conditions of the inside home and outdoor environment [1]. Nowadays, the Internet of Things (IoT) plays an important role in smart home area, whose architecture generally includes sensing function, network function and application function [2]. Among them, its sensing function module consists of various sensing devices, and the network function module consists of various network management systems and cloud platforms. The application function module includes smart home and other application scenarios.

Manuscript received 28 December 2022; revised 27 August 2023; accepted 15 September 2023. Date of publication 20 September 2023; date of current version 21 February 2024. This work was supported by the National Natural Science Foundation of China under Grant 61871239 and Grant 61671254. *(Corresponding authors: Hui Zhang; Meikun Li.)*

The authors are with the Tianjin Key Laboratory of Optoelectronic Sensor and Sensing Network Technology and the Department of Communication Engineering, Nankai University, Tianjin 300071, China (e-mail: zhangh@nankai.edu.cn; 2120220369@mail.nankai.edu.cn; 2120200308@mail.nankai.edu.cn; 2120190299@mail.nankai.edu.cn).

Digital Object Identifier 10.1109/JIOT.2023.3317416

In the field of smart home, the IoT usually integrates sensing information and executing commands within the home through the local network [3]. The smart home enables users to remotely control household appliances and equipment through home automation and auxiliary functions, providing a good living environment for residents. The remote monitoring system using network technology is an important component of smart home system [4].

With the convenience of smart home devices, the vulnerabilities and security problems had attracted attention [5], such as IoT security solutions based on machine learning [6]. To solve the secure problems of smart home networks, recently a series of related methods were proposed for smart homes, such as the key agreement scheme by using the secret mismatch problem [7], the privacy-preserving security architecture and anonymous lightweight authentication mechanism [8], the hybrid key agreement scheme by using the Merkle puzzle [9], the credential-less authentication framework “HomeShield” by effectively defending against the attacks without the need for sensitive credentials [10], key-free authentication protocol against subverted indoor smart devices [11], and the security protection in the edge computing environment from the perspective of identity authentication and data security [12].

However, different from the past, the data flow of home networks is changing from bit to gigabit [13]. The massive data application scenarios will inevitably lead to dramatic changes in communication modes and data usage issues. So, it is not enough to focus on the security of smart home. Relatively, the communication mode of smart home networks has received less attention in recent years, but it is urgent to jointly analyze it in the background of big data usage scenario. Although the smart home networks established through power line can overcome the wall jack problem, sharing the electrical wires with appliances also means sharing the noise and interference they create. On the other side for wireless, the WiFi is convenient, but there is a drawback that it could not always reliably blanket every room and reach every device. Especially for each of the above two technologies, there exist unpredictably dropping or retransmitting large amounts of data packets, making it difficult to deliver packets on a predictable schedule. Therefore, today’s and tomorrow’s smart home networks should use multiple connectivity technologies for increased aggregated data rates and better coverage [13].

A. Related Works

In order to enforce optimal energy use, the concept of building energy management systems (BEMSs) was proposed for smart home network applications, making buildings smarter and energy-efficient by using the IoT platform [14]. In such a context, several energy-efficient wireless sensor network (WSN) platforms were designed for home/office automation [15]. As a consequence of data-driven energy transformation, many researchers have dealt with the estimation and forecasting of household appliances usage patterns from an artificial intelligence (AI) perspective, such as machine learning for identifying demand patterns of home energy management systems with dynamic electricity [16]. Moreover, as the sharp increase in data traffic of smart home network IoT devices, a smart gateway was proposed for data collection and awareness to connect the home network and the Internet [17].

In the area of learning from smart home data, the architecture of data processing component of the smart home system was explored in recent years, and the key data mining and machine learning approaches applicable in smart home scenarios were identified from four aspect, respectively, frequent pattern mining, causal relations mining, periodic-frequent pattern mining, pattern classification and clustering [18]. Besides, a survey on five different prediction algorithms that have been or are being used at smart homes were given, such as active Lezi, Flocking, sequence prediction via enhanced episode discovery (SPEED), and Nash H-Learning and Apriori algorithm [19]. As an important method in machine learning, the long short term memory (LSTM) concept was introduced by Hochreiter and Schmidhuber [20], and the autoencoder method was proposed by Kamyshanska and Memisevic [21]. Especially, to analyze the real-time power consumption data collected from a Singapore smart home, a variation of recurrent neural network (RNN) called LSTM network was created to solve long-term dependency forecasting problems. Combine with K-means clustering method and LSTM-RNN model, the electricity prices were predicted in advance with past historical data [22]. Generally, a large number of terminal devices may generate massive data in the IoT system, and machine learning can be used to analyze the data in depth to dig out useful information. However, many different types of terminal devices were often used in the IoT system, and the machine learning used in the existing work often ignored the differences in the data types, data sizes and other parameters collected by different types of terminal devices.

In addition, there were some other smart home network solutions that connected to the Internet by embedding gateways in home appliances [23]. Based on the home appliance control framework of intelligent TV set-top box [24], by installing the home gateway module into the intelligent TV set-top box and then connecting to the Internet, users can control the home appliances through the Web server. Through smart home management system [25], the community managers can unify families in the community for management. Due to the large number of different and highly customized design schemes in the market, the scale of smart home users was limited [18]. However, the solution of smart home cannot be limited to a single type of equipment or manufacturer, in which the compatibility between devices was required. In the scheme of home automation network [26], the ZigBee radio was integrated into the power socket module, which can flexibly control various household appliances. However, to solve the problem that users need to switch applications (APPs) when using smart phones to control different devices, the scheme of smart phone control software was proposed that automatically switches interfaces according to different devices [27]. Generally, the smart home devices mainly rely on wireless communication to complete information transmission [28], including Zigbee, Bluetooth, global system for mobile communications (GSMs), etc. The greenhouse intelligent information monitoring system was established by using Zigbee WSN [29], which can realize the centralized monitoring, display and storage of environmental information. According to the problems of WSN technology in smart home [30], the transmission performance of WSN was closely related to the network coverage. When the line of sight (LOS) transmission was not available, the obvious transmission errors may occur. The Bluetooth technology was used to monitor the time, location and other information of road vehicles, and estimate the road conditions [31]. However, these technologies belonged to short-range wireless transmission area, which cannot meet the requirements of long-distance transmission. In order to overcome the above problems, a joint home management system of GSM and WSN was proposed, which used GSM to communicate with user terminals [32]. Although there were many applications of using wireless and optical fiber technology to build smart home networks in the existing work, the wall perforation was often required when using these technologies, and there still existed wireless coverage holes. In particular, there was still a lack of friendly human-computer interaction interfaces.

Recently, the WiFi has become a hot choice of wireless connection technology for IoT devices, because of its popularity in families and low cost [33]. Smart home devices can share one WiFi access point for Internet and local interconnection [34]. By using WiFi, a household sensor system named CARDEA WiFi extension was established for active assisted living (AAL) [35], in which the power modes of sleep and hibernation were proposed for saving energy of bed occupancy sensors and wearable WiFi devices. As a special case, the power consumption of WiFi technology in low-power sleep and hibernation modes was tested, and the flexibility of using WiFi was analyzed according to its specific situation. Besides, in the smart home control combining WSN and power line communication (PLC), the power line was usually used as the backbone network to connect the WSN in the room, and its control function was realized through PLC transceiver, which focused on the energy-saving effect of the intelligent lighting control algorithm limited to the local rooms, and still lacked long-distance latency test [36].

B. Motivation and Contributions

Based on the above analysis, we propose a smart home network scheme with PLC and wireless dual-mode transmission. We have integrated WiFi technology on the basis of PLC, which can effectively avoid wall perforation and enhance signal coverage. The proposed Conv-LSTM-Attn algorithm can perform cluster analysis on different types of terminal devices and predict different types of sensing information. The designed user interface is easy to operate, and adapts to smartphones.

In our contributions, we combine the PLC module with the WiFi module together to connect the cloud platform, and realize the function of environmental information collection, storage and remote control. Our design can adapt to different housing structures, and this dual mode real-time system has a wide application area, with a good compatibility of intelligent equipment.

The main contributions of this article are summarized as follows.

1) To solve the problem of wall jack and room coverage in smart home, we design a set of dual-mode data acquisition device that can be used in the smart home networks, which adopts the PLC and WiFi dual-mode communication mode, and supports communication standards, such as G3-PLC, 802.11b, etc. Especially, our PLC module has the advantage of being pluggable on the debugging board and can be replaced with a module that supports different PLC protocols as required. Our scheme can realize the autonomous backhaul of the collected data with low-latency and long-distance transmission, moreover it can highly flexibly configure various sensor modules to achieve smart home management and control.
2) To enhance the user experience, we develop a graphical user interface (GUI) that can be used for mobile terminal APP, by means of using the IoT studio tools. the user can conveniently access the cloud through mobile APP to remotely manage smart home facilities.
3) To solve massive data usage problem acquired by smart home facilities, we propose a data analysis method that takes feature fusion method to extract information from collected data features and reduce the dimension of data by using the automatic encoder scheme. Furthermore, after cluster the data by means of the K-means algorithm, we propose a Conv-LSTM-Attn algorithm based on deep learning, which is a prediction model of convolutional long- and short-term memory network based on attention mechanism.
4) Numerical results demonstrate that:
    a) the designed dual mode data acquisition device plays an active role in smart home networks;
    b) the proposed Conv-LSTM-Attn algorithm outperforms the existing machine learning schemes in smart home in terms of achieving the prediction values.

C. Organization

The remainder of this article is organized as follows. The wired and wireless problems in smart home networks are presented in Section II. The system design and implementation are introduced in Section III, including the hardware and software design of the system and the configuration of the cloud platform. The data analysis and algorithm design are presented in Section IV, including data processing, feature fusion, and the proposed Conv-LSTM-Attn algorithm. Section V provides the field test of the system and shows the analysis of the results. The conclusions are finally summarized in Section VI.

II. Wired and Wireless Problems in Smart Home Networks

A. G3-PLC Technology

Fast, safe, reliable and cost-effective communication are critical to the smart grid. According to the size of communication broadband, PLC can be divided into narrowband and broadband PLC [37]. G3-PLC is a narrowband orthogonal frequency division multiplexing (OFDM)-based PLC technology, and it uses lower carrier frequency and slower data rate to provide longer distance and more reliable services, supporting highly reliable power line data links across noisy environments and high-data rate in short distance in buildings. G3/PLC standard has been recognized by many countries and organizations and selected as the basic technology for international standard development, such as IEEE P1901.2, etc. Its standard sampling rate is 400 kHz, and the data transmission rate can reach 33.4 kb/s. It supports BPSK, DBPSK, and DQPSK modulation. It uses an improved carrier sense multiple access with collision avoid (CSMA/CA) media access control protocol to monitor the length of carrier signals and data packets being transmitted before allocation, which is itself a random backoff period. G3-PLC can effectively improve the frequency band utilization and mitigate the multipath power line channel fading, supporting high-speed, highly reliable IP-based communication over power lines. Due to its status as a nonproprietary specification, its high-performance and advanced functions, G3-PLC technology is currently supported by several manufacturers in the smart grid.

B. WiFi Technology

The communication between device to device occupies an important position in smart home. The commonly used short-range wireless communication technologies include WiFi and Bluetooth. WiFi is widely used in daily life, and its transmission rate can reach to 54 Mb/s with low cost. Moreover, WiFi is based on the IEEE.802.11 series of standards, such as 802.11b, 802.11g/a, 802.11n, 802.11ac, etc. Among them, the 2.4-GHz ultra high-frequency (UHF) band is used by the second-generation WiFi based on 802.11b. The 2.4-GHz UHF and 5-GHz super-high-frequency (SHF) industrial scientific medical (ISM) bands are, respectively, used by the third-generation WiFi based on 802.11g/a, and are also used by the fourth-generation WiFi based on 802.11n. The 5-GHz SHF ISM band is only used by the fifth-generation WiFi based on 802.11ac [38]. With the wide application of household accessible WiFi networks, it becomes a trend for building smart home systems through WiFi networks, such as the Wattcher system for distributed energy monitoring in the home [39]. By

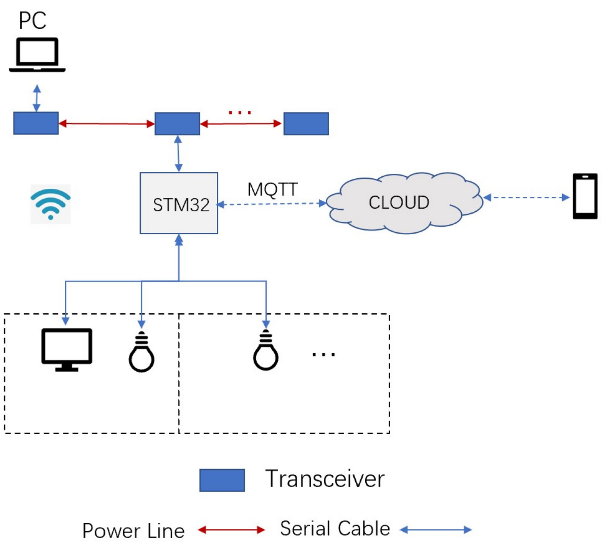

Fig. 1. System block diagram.

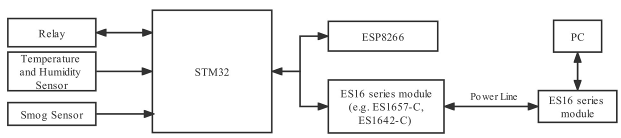

Fig. 2. Hardware system structure.

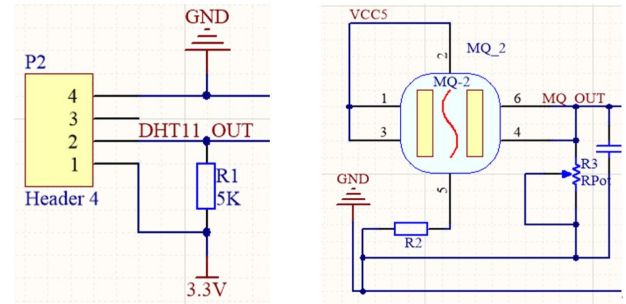

Fig. 3. Sensor circuit module.

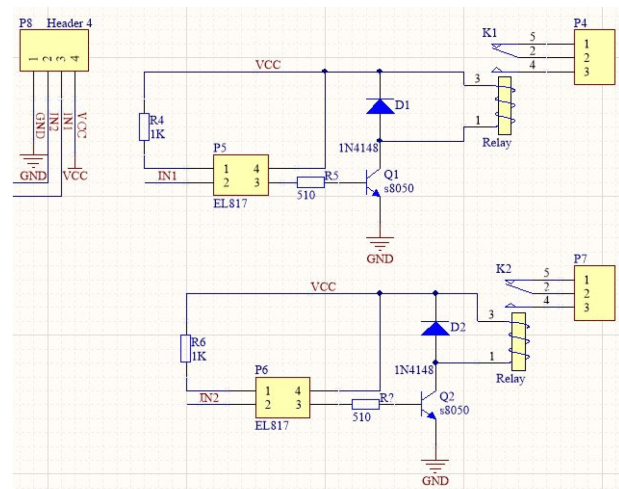

Fig. 4. Circuit diagram of electric relay module.

embedding sensors and collection hubs, the home appliances can regularly transmit data, such as energy to a centralized data repository via a local WiFi network. Moreover, it is convenient in the installation for such a microdeployment, which can adapt well to nonstandard household devices, broadening the concept of smart homes.

III. System Design and Implementation

As shown in Fig. 1, our system design includes four components, respectively, the main control module, communication module, sensor module and cloud platform. We take use of STM32 as the main control module to acquire sensor information, receive and issue instructions, and complete data upload. Through PLC between the main control module and the host computer, the acquisition of information is realized by the power line carrier module. Moreover, we choose the IoT Studio platform under Alibaba cloud as our cloud platform, which provides designers with a variety of development methods. Our design could provide users with an interactive interface by means of mobile application development.

A. Hardware Design

In hardware design, we focus on implementing the functions of information acquisition, transmission, and control. Our hardware system includes the main control single-chip microcomputer, sensor circuit and communication circuit. The hardware system structure diagram is shown in Fig. 2.

1) Main Control Module: We adopt STM32 as the micro control unit (MCU), in which its resources can meet the requirements of external equipment. We use the PC group pins and PA group pins as input and output, respectively, where the PA group serial port communicates with the WiFi module to realize the connection between the device and the cloud platform, and the PB group serial port communicates with the power line carrier chip ES1657-C through the RS232 to TTL module.

2) Sensor Module: We choose DHT11 and MQ-2 as two sensor modules for monitoring the internal environment of the home. The connect circuit is built as shown in Fig. 3. The DHT11 adopts single bus communication mode, with a transmission distance of 20 m. Its types of transmitted information include temperature, humidity and parity bit. The MQ-2 combustible gas sensor can detect various dangerous gases, such as natural gas and methane, with a wide measurement range and high sensitivity.

The electric relay is normally open to connect and limit the load to AC voltage 250 V and DC voltage 30 V. When the output pin is high, the electric relay is closed, the load starts to work, and the weak current signal is used to control the closing of the internal switch state of the home. The electric relay circuit diagram built by the EL817 optocoupler chip is as shown in Fig. 4.

3) Communication Module: We choose Essence ESP-01S as the low-power wireless communication module, which can meet the needs of IoT applications, such as smart homes. The ESP-01S supports a variety of development methods, and we use AT commands for development. Especially, its built-in

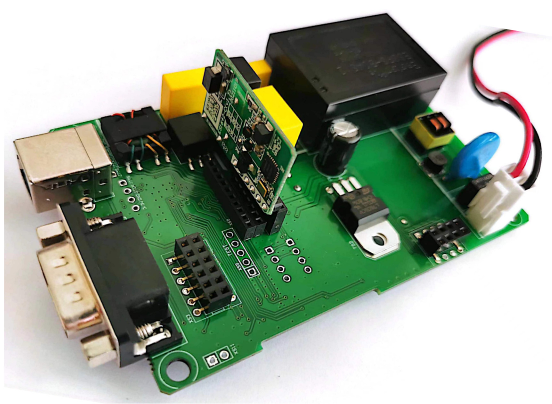

Fig. 5. Power line carrier module.

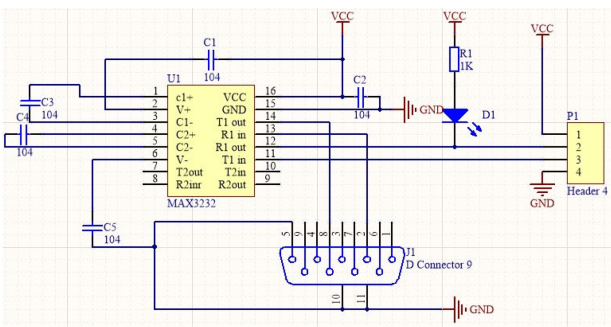

Fig. 6. Conversion circuit.

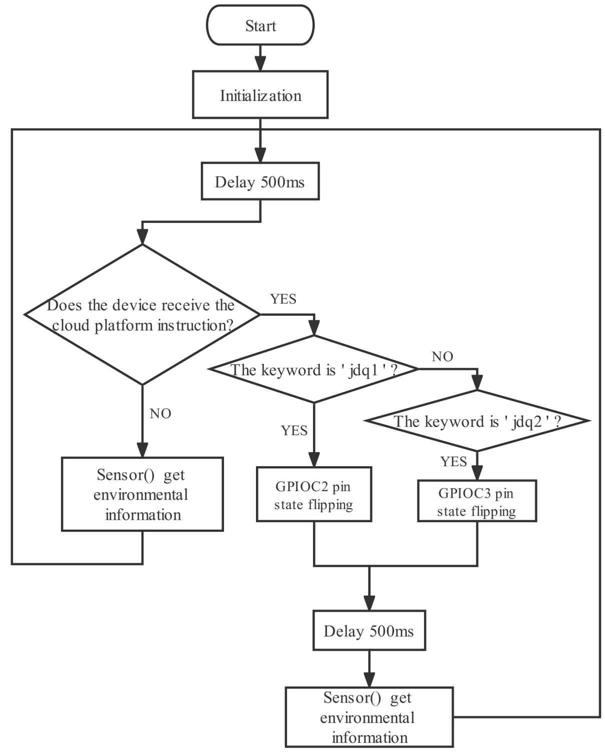

Fig. 7. Flow of the MCU main program.

MQTT firmware can realize the connection between the device and the Alibaba cloud platform. As shown in Fig. 5, we design a PLC debugging board, which can integrate PLC modules. Especially, our PLC module has the advantage of being pluggable on the debugging board and can be replaced with a module that supports different PLC protocols as required. In test, we choose Eastsoft ES1657-C as power line carrier module, which modulates information through DBPSK, and can automatically realize relay. In our design, the PLC debugging baseboard with the built-in ES1657-C module is used to complete the communication with the single-chip microcomputer. As a typical circuit in Fig. 6, the serial port output of the PLC debugging baseboard adopts RS232, and an RS232 to TTL circuit is added to complete the communication, because the TTL level of the single-chip microcomputer requires a protocol conversion circuit.

B. Software System Design

1) Cloud Platform: We use Alibaba cloud as the cloud platform of the IoT to realize the matching of virtual device IDs and actual devices. Moreover, this platform provides data analysis and application development service for various IoT application scenarios and can make the association of the device with the user interface through its inside IoT studio.

2) Microcontroller Unit Programming: After completing the initialization of the STM32 pins of each peripheral device, we send AT commands through the serial port to communicate with the ESP8266 module, connect to the cloud and transmit information from sensors. Considering the communication time of the sensing device, data upload efficiency and update rate, and the size of the cache, we set the MCU processing delay to 500 ms. The main program execution flow of STM32 MCU is shown in Fig. 7. After the initialization, the STM32 enters the receiving instruction function, and the flow of the get function is given in Fig. 8. When receiving the instruction issued by the cloud platform, it uses the pointer to retrieve the keyword of the received information, find the instruction of device attribute setting, and then control the MCU pin state to realize the transition of the electric relay state. When there is no instruction issued at the time, the STM32 reads the status information of each sensor and sends it to the cloud platform through the serial port in sequence.

To realize the control task of this design, we use the serial port interrupt when the STM32 communicates with the ES1657-C PLC module, and the serial port interrupt flow is given in Fig. 9. The corresponding interrupt is triggered according to different instructions issued by the host computer, thereby controlling the working state of the electric relay.

IV. Data Analysis and Algorithm Design

Considering the heterogeneity of IoT systems, such as the differences in terminal types, data types, data sizes, etc., we choose the autoencoder combined with the K-means clustering method to make the unsupervised clustering process for the data. The reason is that the autoencoder can map high-dimensional original data to a lower dimensional representation space, and build encoding and decoding networks by learning the distribution of data. Moreover, the autoencoder can perform nonlinear transformations, and automatically filter out some noise and redundant information during the learning

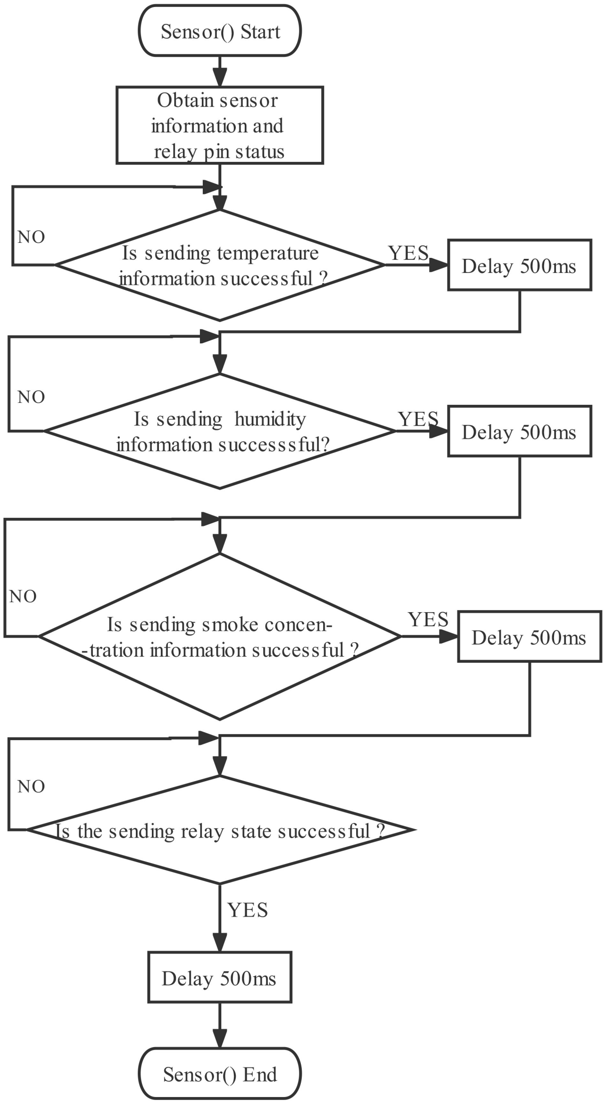

Fig. 8. Flow of the get function.

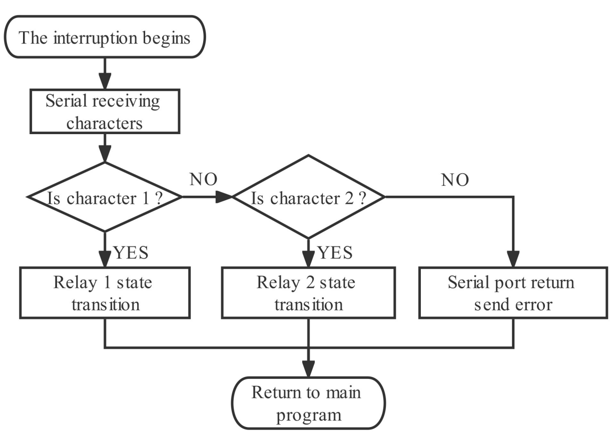

Fig. 9. Flowchart of serial port interrupt.

process, so as to extract more meaningful features. For the features encoded in the low-dimensional space, the K-means algorithm can easily find the cluster structure in data. On this basis, we make a predictive analysis of the data.

TABLE I
DATA CHARACTERISTICS

Name	**Dimension**
Temperature	°C
Humidity	%RH
Smoke	ppm
Time from last information acquisition	s

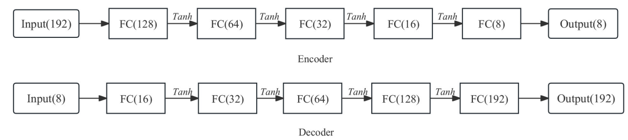

Fig. 10. Structure of automatic encoder.

A. *Unsupervised Clustering of Terminals*

Considering various sensors are distributed in smart homes, the data collected by different sensors are obviously different in terms of acquisition frequency, data size, etc. In order to accurately analyze data, it needs to be classified before analyzing data, and assign terminals with similar behaviors to the same cluster. Moreover, it is necessary to extract key data category features before classification. The features used in this article are shown in Table I.

The features in the table are continuous time series features, with strong correlation in time. In order to extract information from such features, it is necessary to reduce the dimension of features. There are two methods for feature dimension reduction, respectively, feature selection and feature fusion [21]. The feature selection is to obtain a subset of high-dimensional data according to the correlation analysis of feature data, which retains the physical meaning of the original features, but some features are generally omitted in the dimension reduction process, resulting in information loss. Generally, the dimension reduction process is usually complex. The feature fusion is to find the optimal subspace of data distribution, transform and combine data to form new features. All features in feature fusion contribute to the reduced dimension data, which retain most of the information in the features. According to the above analysis, due to the automatic encoder is one of effective feature fusion methods, we choose the automatic encoder method to reduce the dimension of data, and its structure is shown in Fig. 10.

After reducing the dimensions of the data, we use the K-means algorithm to cluster the data. The unsupervised classification process is shown in Fig. 11. First, the original data should be cleaned, such as extracting time stamps, sorting by time stamp order, filling in missing values, and data normalization. Then the data is divided into training set and test set to verify the error of the automatic encoder. The data of temperature, humidity, smoke and other parameters are extracted by sliding window, and the length of each data is 192. After compression, the features are compressed into 8-D features. By using the compressed features and the statistical features of the original data as the input features, the K-means algorithm is finally taken for clustering.

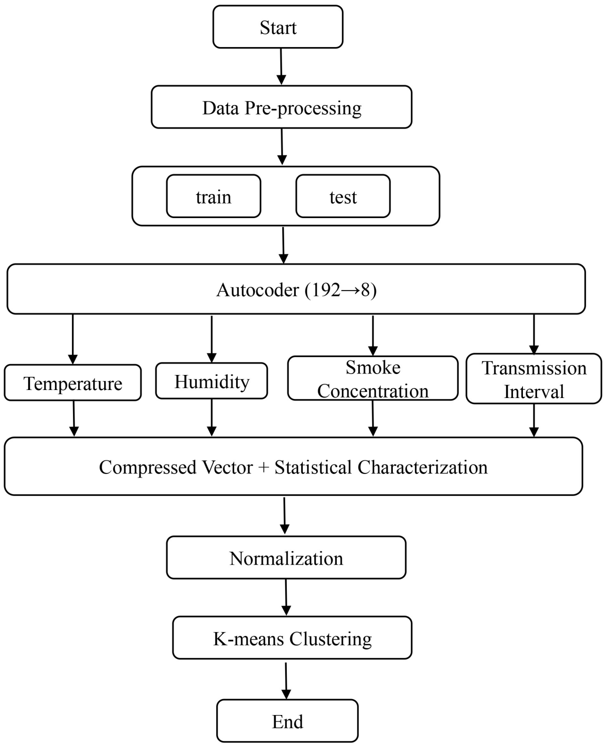

Fig. 11. Unsupervised clustering process.

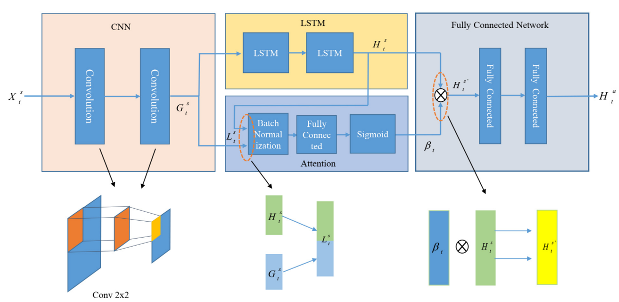

Fig. 12. Conv-LSTM-Attn network structure.

B. Prediction of Data

We propose a prediction model of convolutional long- and short-term memory network based on attention mechanism, written as Conv-LSTM-Attn. The proposed model is mainly composed of convolutional neural network, long- and short-term memory network, and attention mechanism. Its network structure is shown in Fig. 12. The convolutional neural network is composed of two convolutional layers, and the LSTM network contains two LSTM layers. The convolution neural network is used to extract the spatial features of time series data, while the LSTM is used to extract the temporal correlation features of time series data.

The drawback of traditional LSTM network is that the LSTM model cannot detect which time points are important parts. For the information provided by general time series data at different times may not be equally important to the prediction performance of the model, the values of some terminals at different times may have different effects on the prediction results. To solve this problem, we propose an attention mechanism to help the model automatically recognize the importance level of time series at different times, in which the attention mechanism module consists of a batch normalization layer, a full connection layer and a sigmoid function, and its output length is consistent with that of LSTM layer.

In order to extract the temporal and spatial characteristics of time series data, it is necessary to construct the time series data matrix. Let f_p^t represent the measured data of the node p at time t. The measured data of node p at time t to $t-n$ can be expressed as $X_p^t = [f_p^{t-n}, f_p^{t-(n-1)}, \ldots, f_p^t]$. Then, the space-time matrix formed by combining the historical measurement data of the nearest terminal of the same category (m nodes in total) is as follows:

$$X = \begin{bmatrix} X_{t-n}^s \\ X_{t-(n-1)}^s \\ \vdots \\ X_t^s \end{bmatrix}^T = \begin{bmatrix} f_{t-n}^1 & f_{t-(n-1)}^1 & \cdots & f_t^1 \\ f_{t-n}^2 & f_{t-(n-1)}^2 & \cdots & f_t^2 \\ \vdots & \vdots & \ddots & \vdots \\ f_{t-n}^m & f_{t-(n-1)}^m & \cdots & f_t^m \end{bmatrix} \tag{1}$$

where $X_t^s = [f_t^1, f_t^2, \ldots, f_t^m]$ represents the measurement data of the prediction area at time t, and the observation time n will affect the prediction accuracy from two aspects of data length and time step.

The output of Conv-LSTM-Attn at each time step can be expressed as the weighted sum of LSTM network output and attention mechanism network

$$H_t^a = \sum_{k=1}^{n+1} \beta_k H_{t-(k-1)}^s \tag{2}$$

where $n+1$ is the length of the time series, β_k is the attention value at time $t-(k-1)$, and β_k can be expressed as

$$\beta_k = \frac{e^{s_k}}{\sum_{k=1}^{n+1} e^{s_k}} \tag{3}$$

where $s = [s_1, s_2, s_3 \cdots, s_{n+1}]^T$, it represents the importance of each part in the time series, which can be calculated by the following:

$$s_t = V_s^T \tanh\left(W_{hs} G_t^s + W_{ls} H_t^s\right) \tag{4}$$

where V_s^T is the network weight reflecting the overall attention mechanism, G_t^s is the output of the data sample through two 2×2 convolutional layers, and H_t^s is the hidden layer output of LSTM network. W_{hs} and W_{ls} are both the network weights reflecting part of the input attention mechanism [40]. Among them, V_s^T, W_{hs}, and W_{ls} are determined by the corresponding neural network function under Pytorch during the training of the neural network, while G_t^s is determined by the data sample and the convolutional layer. W_{hs} and W_{ls} are used to ensure the unity of G_t^s and H_t^s dimensions, and to achieve data augmentation. The attention value β_k that can be seen as a "selection gate" depends on G_t^s and H_t^s. When β_k is larger, the amount of information flowing through it is larger, indicating that the current vector position is more important.

The error function $F(M_j, P_j)$ is used to compare the differences between predicted value M_j and the observed value P_j. Because the dimensions of different features are various, the

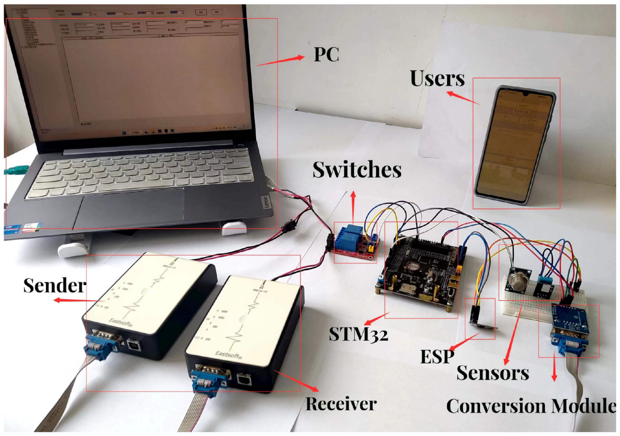

Fig. 13. Test devices.

mean absolute percentage error (MAPE) is taken to measure the prediction error, which can be expressed as

$$F(M_j, P_j) = \text{MAPE} = \frac{100\%}{n}\sum_{j=1}^{n}\left|\frac{M_j - P_j}{P_j}\right|. \quad (5)$$

V. Field Test and Results Analysis

The application scenarios of our design are indoor environments, such as homes and offices. We choose the general household with a house mask of about 90 square meters as an example for testing. Our testbed consists of four parts, namely, the microcontroller module STM32, the communication modules ESP8266 and ES1657-C, the sensor modules DHT11 and MQ-2, and the electric relay module. The actual test devices are shown in Fig. 13.

A. Mobile Client Design

Through use the Alibaba cloud platform, we make personalize the design of the mobile client, and the GUI is as shown in Fig. 14. The user can view the indoor environment and equipment working conditions through the APP on the mobile phone. The acquired monitoring data also can be viewed in the background. As shown in Fig. 14, the user can control by clicking in the control page. When the power line carrier module used in the home, the control information responds quickly and accurately.

B. Indoor Monitoring Test

The sensors used for monitoring are deployed in the living room to monitor the environment. Figs. 15–17 capture the changes of indoor temperature, humidity and smog concentration, respectively. It can be seen that the indoor temperature gradually increases after noon in summer and decreases in the evening. The indoor humidity is between 50%RH and 65%RH, which is a relatively comfortable humidity environment. The smoke concentration increases around noon, but it is still at safer level.

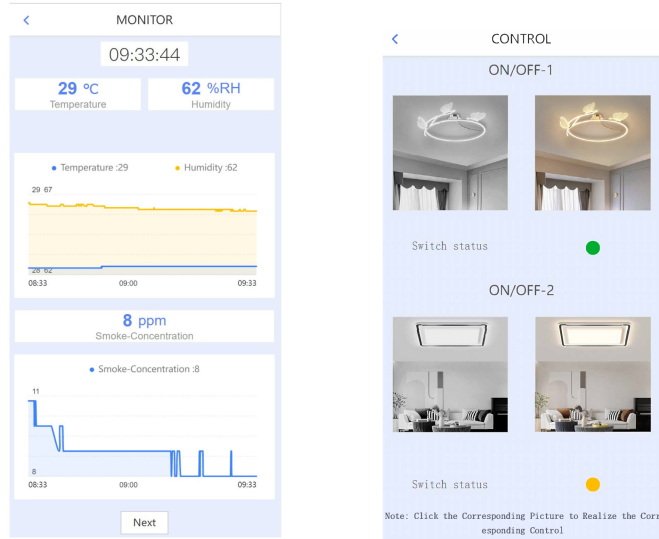

Fig. 14. Monitor and control interface.

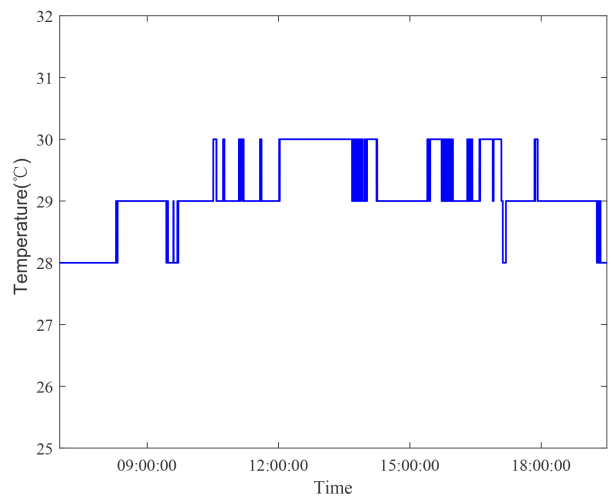

Fig. 15. Indoor temperature (12 h).

TABLE II
Cloud Platform Command Response

Distance	Response Times	Delay
≤ 0.1km	36 responses, 14 no response	5s
≤10km	40 responses, 10 no responses	5s
≤200km	32 responses, 18 no response	5s

C. Wireless Communication Test

The microcontroller module STM32 uploads data that includes electric relay status and environmental information. According to the program settings, the get function sequentially gets the master pin status and completes the upload every 2.5 s. After testing, the STM32 uploading data process is stable, and the data can be uploaded and updated in real time. See Fig. 14 for details. The user can issue control instructions through the mobile APP. Table II records the response of the device after 50 tests were conducted using the APP to issue commands.

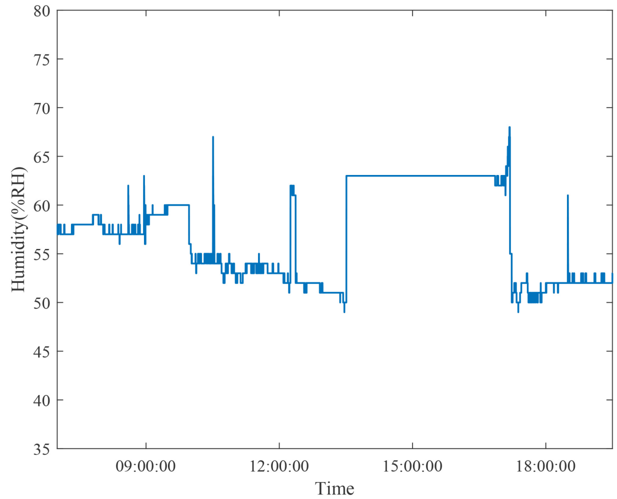

Fig. 16. Indoor humidity (12 h).

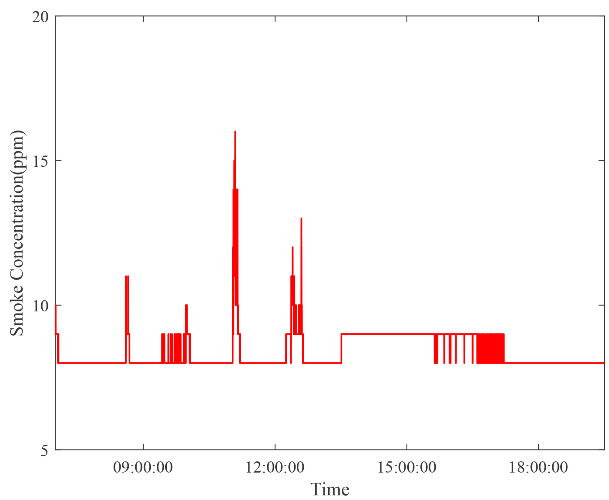

Fig. 17. Indoor smoke (12 h).

It can be observed from Table II that the number of responses within 10 km is more than those within 0.1 and 200 km. The reason is that in the network coverage of different test distances, the number of terminals served by the cloud server is different, the network services are bursty, and the status of the cloud server selected by the terminal is inconsistent. Therefore, the response times depend on the performance of the cloud server selected by the terminals, the number of terminals connected to the cloud server and the size of the network services. Because of these comprehensive factors, it leads to the results shown in Table II.

In summary, using the cloud platform to send and upload data can realize the function of remote monitoring and control of internal devices in the home. The accurate data transmission can be achieved between the device and the cloud platform.

D. Power Line Carrier Communication Test

When testing the power line carrier communication, we select an indoor fixed location as the indoor console, move the location of the receiving node to test the communication performance. Fig. 18 shows the structure of the house used.

Fig. 18. Test house structure.

TABLE III
RESPONSES TO PLC COMMANDS

Location	Response	Delay
A	Response	<1s
B	Response	<1s
C	Response	<1s
D	Response	<1s
E	Unresponsive	–

The position marked in Fig. 18 is selected to verify the accuracy and delay of command issuing using the PLC. The test results are recorded as shown in the Table III.

Based on the above analysis, using PLC, the information transmission response is accurate and fast. The combination of PLC and wireless communication complements each other, making the system more stable.

E. Data Prediction Analysis

We use a data set of 2.3 million samples to train the proposed algorithm, including temperature, humidity, smoke, etc.

1) Simulation Environment: We use Python 3.6 platform to analyze the data from dual-mode smart home sensors. The proposed automatic encoder network and Conv-LSTM-Attn network are built based on the Pytorch platform. Among them, the automatic encoder uses the Adam optimizer to optimize the model, the loss function uses the normalized mean square error (NMSE), the learning rate is 0.01, the iteration is 100 epochs. Except that one part of data set is used to train the model, such as temperature, humidity, and smoke, the other part of data set is used to verify the model accuracy. The data set is extracted by sliding window, 200 data are taken for each ID, the data length is 128, and the feature length is compressed to

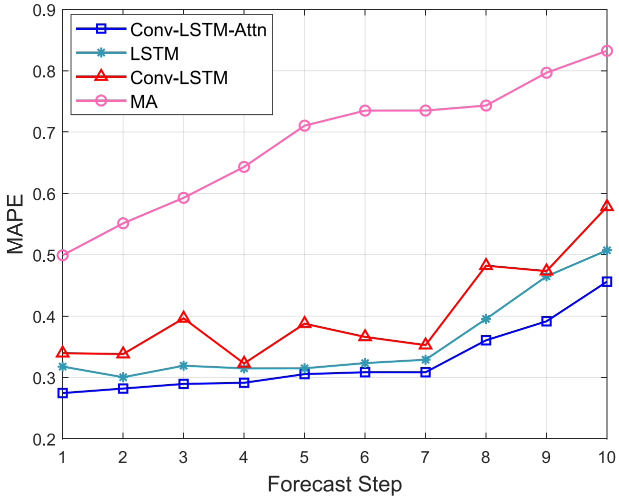

Fig. 19. Comparison of algorithm prediction accuracy.

8 after encoder. The K-means algorithm is implemented based on the scikit-learn package, using the statistical characteristics of temperature, humidity, smoke (maximum, minimum, mean, variance, and standard deviation), and the characteristics compressed by the automatic encoder. The Conv-LSTM-Attn network uses the Adam optimization model, and the loss function uses MAPE. The initial learning rate is 0.05, and the learning rate is reduced to the original 0.25 after every 250 epochs. A total of 1000 epochs are iterated. We use the data of the first 50 steps to predict the data of the next 1–10 steps, and construct the data set by sliding windows. At the same time, in order to improve the prediction accuracy of the model, we train the prediction model for each category of equipment, each equipment has 500 pieces of data for training, and retains the model with the smallest loss function on the verification set.

2) Result Analysis: In order to reflect the performance of different data analysis algorithms, the proposed Conv-LSTM-Attn algorithm is compared with LSTM algorithm using clustering operation, Conv-LSTM algorithm without using clustering operation, and moving average (MA) algorithm. Fig. 19 shows the trend of algorithm prediction error changing with the increase of prediction step size. It can be found that the prediction error of the proposed Conv-LSTM-Attn algorithm is obviously smaller than that of the other three algorithms, and the prediction error of the model increases gradually with the increase of the prediction step. Compared with LSTM, the prediction error of the proposed Conv-LSTM-Attn algorithm is reduced by 8.8% on average. The reason is that our Conv-LSTM-Attn algorithm uses the convolutional neural network to extract the spatial features, which not only uses the historical data of the device itself but also uses the historical data of devices of the same category. Moreover, in Conv-LSTM-Attn algorithm, the weighting of the attention mechanism and the output of the long-term and short-term memory network can help the model focus on the important area of the output vector.

It can be seen from Fig. 19 that the prediction accuracy of Conv-LSTM algorithm fluctuates greatly, and its prediction error is always greater than our Conv-LSTM-Attn algorithm. This is because our Conv-LSTM-Attn algorithm uses unsupervised clustering method to classify terminals collecting similar data into the same type of data and trains the model separately for terminals of the same type of data, effectively improving the accuracy of model prediction. Considering the clustering operation is not adopted in Conv-LSTM algorithm, due to differences in training set data, there are outliers in the data, and the prediction accuracy of Conv-LSTM model is even lower than LSTM, which verifies the effectiveness of the device clustering method proposed in this article. However, the prediction error in MA model is much higher than the other three algorithms due to the accumulation of prediction error when use the MA method.

VI. Conclusion

As the sharp increase in data traffic, except for security, home network and data usage will become two important issues in smart home in the future. On the one hand, aiming to solve the problem of wall jack and room coverage in smart home networks, we develop a set of dual-mode data acquisition device, which supports wired PLC and wireless dual-mode communication. Our PLC module is with the advantage of being pluggable on the debugging board, which can be replaced with a module that supports different PLC protocols as required, such as broadband PLC or narrowband PLC. The main control module of the device can obtain the internal data of the home and upload it to the cloud platform, and it can receive the instructions from cloud platform and power line. Moreover, by means of using the IoT studio tools, we develop a GUI to enhance the user experience, and its advantage is that the user can conveniently access the cloud through mobile APP to remotely manage smart home facilities.

On the other hand, aiming to solve massive data usage problem acquired by smart home facilities, we propose a data analysis method that integrates the automatic encoder scheme and the K-means algorithm. Through this method, we can effectively reduce the data dimension and classify the data obtained by sensors, so as to obtain a unified data analysis format, which is the basis for data prediction. On this basis, we propose a Conv-LSTM-Attn algorithm based on deep learning, which is a prediction model of convolutional long- and short-term memory network based on attention mechanism.

Through the field test, it illustrates that our system can realize accurate long-term data acquisition and recording, dual-mode control from indoor to long-distance away, the autonomous backhaul of the collected data with low-latency and long-distance transmission. Besides, it can highly flexibly configure various sensor modules to achieve smart home management and control. Moreover, through the analysis of actual sampling data, the proposed Conv-LSTM-Attn algorithm outperforms the existing machine learning schemes in smart home in terms of achieving the prediction values.

References

[1] S. P. Mohanty, "Consumer technologies build smart homes," *IEEE Consum. Electron. Mag.*, vol. 9, no. 1, pp. 4–5, Jan./Feb. 2020.

[2] D. C. Nguyen et al., "6G Internet of Things: A comprehensive survey," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 359–383, Jan. 2022.

[3] A. Kakar, S. Agrawal, and G. Sumathi, "Smart home: A new way to life," in *Proc. Innov. Power Adv. Comput. Technol. (i-PACT)*, 2021, pp. 1–6.

[4] M. R. Alam, M. B. I. Reaz, and M. A. M. Ali, "A review of smart homes—Past, present, and future," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 1190–1203, Nov. 2012.

[5] B. D. Davis, J. C. Mason, and M. Anwar, "Vulnerability studies and security postures of IoT devices: A smart home case study," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10102–10110, Oct. 2020.

[6] L. Xiao, X. Wan, X. Lu, Y. Zhang, and D. Wu, "IoT security techniques based on machine learning: How do IoT devices use AI to enhance security?" *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 41–49, Sep. 2018.

[7] Y. Zhang, H. Zhao, Y. Xiang, X. Huang, and X. Chen, "A key agreement scheme for smart homes using the secret mismatch problem," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10251–10260, Dec. 2019.

[8] W. Iqbal et al., "ALAM: Anonymous lightweight authentication mechanism for SDN-enabled smart homes," *IEEE Internet Things J.*, vol. 8, no. 12, pp. 9622–9633, Jun. 2021.

[9] Y. Zhang, X. Huang, X. Chen, L. Y. Zhang, J. Zhang, and Y. Xiang, "A hybrid key agreement scheme for smart homes using the Merkle puzzle," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1061–1071, Feb. 2020.

[10] Y. Xiao, Y. Jia, C. Liu, A. Alrawais, M. Rekik, and Z. Shan, "HomeShield: A credential-less authentication framework for smart home systems," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 7903–7918, Sep. 2020.

[11] Z. Huang, L. Zhang, X. Meng, and K.-K. R. Choo, "Key-free authentication protocol against subverted indoor smart devices for smart home," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1039–1047, Feb. 2020.

[12] X. Li, T. Chen, Q. Cheng, S. Ma, and J. Ma, "Smart applications in edge computing: Overview on authentication and data security," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4063–4080, Mar. 2021.

[13] S. Palm, "Home networks: From bits to gigabits: Lessons learned from the evolution of home networking," *IEEE Consum. Electron. Mag.*, vol. 1, no. 3, pp. 29–35, Jul. 2012.

[14] S. M. Krishna and T. Perumal, "Making buildings smarter and energy-efficient—Using the Internet of Things platform," *IEEE Consum. Electron. Mag.*, vol. 10, no. 3, pp. 34–41, May 2021.

[15] C. S. Abella et al., "Autonomous energy-efficient wireless sensor network platform for home/office automation," *IEEE Sensors J.*, vol. 19, no. 9, pp. 3501–3512, May 2019.

[16] L. Ciabattoni, G. Comodi, F. Ferracuti, and G. Foresi, "AI-powered home electrical appliances as enabler of demand-side flexibility," *IEEE Consum. Electron. Mag.*, vol. 9, no. 3, pp. 72–78, May 2020.

[17] P. Wang, F. Ye, and X. Chen, "A smart home gateway platform for data collection and awareness," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 87–93, Sep. 2018.

[18] M. Antić, I. Papp, S. Ivanović, and M. Matić, "Learning from smart home data: Methods and challenges of data acquisition and analysis in smart home solutions," *IEEE Consum. Electron. Mag.*, vol. 9, no. 3, pp. 64–71, May 2020.

[19] S. Wu et al., "Survey on prediction algorithms in smart homes," *IEEE Internet Things J.*, vol. 4, no. 3, pp. 636–644, Jun. 2017.

[20] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[21] H. Kamyshanska and R. Memisevic, "The potential energy of an autoencoder," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 6, pp. 1261–1273, Jun. 2015.

[22] W. Li, T. Logenthiran, V.-T. Phan, and W. L. Woo, "Implemented IoT-based self-learning home management system (SHMS) for Singapore," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 2212–2219, Jun. 2018.

[23] W. Yan, Z. Wang, H. Wang, W. Wang, J. Li, and X. Gui, "Survey on recent smart gateways for smart home: Systems, technologies, and challenges," *Trans. Emerg. Telecommun. Technol.*, vol. 33, no. 6, 2022, Art. no. e4067.

[24] J. Kim, E.-S. Jung, Y.-T. Lee, and W. Ryu, "Home appliance control framework based on smart TV set-top box," *IEEE Trans. Consum. Electron.*, vol. 61, no. 3, pp. 279–285, Aug. 2015.

[25] Y.-T. Lee, W.-H. Hsiao, C.-M. Huang, and T. C. Seng-Cho, "An integrated cloud-based smart home management system with community hierarchy," *IEEE Trans. Consum. Electron.*, vol. 62, no. 1, pp. 1–9, Feb. 2016.

[26] G. Song, F. Ding, W. Zhang, and A. Song, "A wireless power outlet system for smart homes," *IEEE Trans. Consum. Electron.*, vol. 54, no. 4, pp. 1688–1691, Nov. 2008.

[27] M.-S. Pan and C.-J. Chen, "Intuitive control on electric devices by smartphones for smart home environments," *IEEE Sensors J.*, vol. 16, no. 11, pp. 4281–4294, Jun. 2016.

[28] A. Tzounis, N. Katsoulas, T. Bartzanas, and C. Kittas, "Internet of Things in agriculture, recent advances and future challenges," *Biosyst. Eng.*, vol. 164, pp. 31–48, Dec. 2017.

[29] X. Xing, J. Song, L. Lin, M. Tian, and Z. Lei, "Development of intelligent information monitoring system in greenhouse based on wireless sensor network," in *Proc. Int. Conf. Inf. Sci. Control Eng.*, 2017, pp. 970–974.

[30] D. Dietrich, D. Bruckner, G. Zucker, and P. Palensky, "Communication and computation in buildings: A short introduction and overview," *IEEE Trans. Ind. Electron.*, vol. 57, no. 11, pp. 3577–3584, Nov. 2010.

[31] X. Shen, X. Wang, and M. Jia, "Design and implementation of traffic information detection equipment based on Bluetooth communication," in *Proc. IEEE Inf. Technol., Netw., Electron. Autom. Control Conf.*, 2017, pp. 1595–1601.

[32] J. M. Wang and H. B. Wei, "Design of smart home management system based on GSM and Zigbee," *Adv. Mater. Res.*, vol. 842, pp. 703–707, Nov. 2013.

[33] J. Ding, Y. Wang, H. Si, S. Gao, and J. Xing, "Multimodal fusion-adaboost based activity recognition for smart home on WiFi platform," *IEEE Sensors J.*, vol. 22, no. 5, pp. 4661–4674, Mar. 2022.

[34] B. Tushir, Y. Dalal, B. Dezfouli, and Y. Liu, "A quantitative study of DDoS and E-DDoS attacks on WiFi smart home devices," *IEEE Internet Things J.*, vol. 8, no. 8, pp. 6282–6292, Apr. 2021.

[35] M. Bassoli, V. Bianchi, I. De Munari, and P. Ciampolini, "An IoT approach for an AAL Wi-Fi-based monitoring system," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 12, pp. 3200–3209, Dec. 2017.

[36] M. Li and H.-J. Lin, "Design and implementation of smart home control systems based on wireless sensor networks and power line communications," *IEEE Trans. Ind. Electron.*, vol. 62, no. 7, pp. 4430–4442, Jul. 2015.

[37] L. Farias, M. Barreto, M. Leme, and S. Stevan, "Technical feasibility and performance analysis of G3-PLC standard for monitoring in industrial environment," *IEEE Latin America Trans.*, vol. 14, no. 10, pp. 4241–4248, Oct. 2016.

[38] Y. Yu, F. Han, Y. Bao, and J. Ou, "A study on data loss compensation of WiFi-based wireless sensor networks for structural health monitoring," *IEEE Sensors J.*, vol. 16, no. 10, pp. 3811–3818, May 2016.

[39] S. Barker and D. Parsons, "Smart homes or real homes: Building a smarter grid with 'Dumb' houses," *IEEE Pervasive Comput.*, vol. 21, no. 2, pp. 100–104, Apr.–Jun. 2022.

[40] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017.

Hui Zhang received the B.S. and Ph.D. degrees from Beijing University of Posts and Telecommunications, Beijing, China, in 2005 and 2010, respectively.

He has worked with Nankai University, Tianjin, China, since July 2010, and currently he is an Associate Professor. From April 2013 to April 2014, He worked as a Postdoctoral Researcher with the University of Southampton, Southampton, U.K. From July 2014 to December 2014, He was a Postdoctoral Researcher with Korea Advanced Institute of Science and Technology, Daejeon, South Korea. From July to October 2018, He was a Visiting Associate Professor with the University of Sophia, Tokyo, Japan. His research interests include the Internet of Things, deep learning, big data analysis, and 6G wireless systems.

Dr. Zhang also served as a reviewer for publications, such as IEEE Transactions on Wireless Communications, and served as a TPC Member for 2019 IEEE ICC and 2020 IEEE ICC.

Meikun Li received the bachelor's degree in electronic information engineering from Northeast Forestry University, Harbin, China, in 2022. She is currently pursuing the master's degree in information and communication engineering with Nankai University, Tianjin, China.

Her research interests include Internet of Things and machine learning.

Maoheng Yang received the bachelor's degree in communication engineering from China University of Geosciences, Wuhan, China, in 2020, and the master's degree in information and communication engineering from Nankai University, Tianjin, China, in 2023.

His research interests include Internet of Things and machine learning.

Chao Zhou received the bachelor's degree in communication engineering from the North China University of Water Resources and Electric Power, Zhengzhou, China, in 2018, and the master's degree in information and communication engineering from Nankai University, Tianjin, China, in 2022.

His research interests include machine learning and time-series prediction.