The IoT Service Agent Model based on Federated Learning to Improve Service Quality

Tse-Chuan Hsu [1]*, William Cheng-Chung Chu[2] , and Shyh-wei Chen[2]
[1]Dept. of Computer Science & Information Management, Soochow University, Taipei, Taiwan
[2]Dept. of Computer Science, Tunghai University, Tiachung, Taiwan
tchsu@scu.edu.tw, cchu@thu.edu.tw, chensw@thu.edu.tw
*corresponding author

***Abstract*—The combination of learning analysis technology and terminal equipment can provide a rapidly imitating learning method between devices, and cognitive intelligence affects the learning effect. To enhance automated learning services, we can use federated learning methods to enable the training of different devices and enhance each other through devices. Using the training database to improve the quality of automated learning services. In this study, a novel agent-assisted active detection and data collection framework is designed. Monitoring agents can learn from each other to establish intelligent models, and through mutual communication between devices. Can check if established data can be applied to machine data model to get data. It can be used for intelligent manufacturing in the future. The agent may learn methods of learning and managing between devices having different properties. Obtaining experimental simulation and control data, and using machine learning to analyze growth progress and results allow for a deeper analysis of associated adjustments and anticipated changes.**

Keywords-component; Federated Learning, Dectection Design, Agent, Automated Learining, IoT

I. INTRODUCTION

In the past, when discussing the application of big data analysis, acquiring pertinent data from scratch was an intimidating task. If the content of the resulting data is entered manually, its accuracy and stability are essential for data quality. There are many controversies, including human error, missing values, etc. Problems derived from missing data values are included in big data analysis, and judgment rules are affected by missing data values. Increase the influence on the possible result of the analysis.

Therefore, explore the device's machine learning, automatically collect relevant data, and further through imitation learning, provide additional terminal nodes for sampling verification. In the present study, we will discuss how to build an agent machine learning model. Simulate and verify using joint agents and complete a learning correction and verification on the status of changes to sampled information variables. In the present document, Chapter 2 summarizes our assessment and analysis of related studies. Chapter 3 provides a description of the structure of the study. The fourth chapter is verified by simulation experiments, and the fifth chapter presents an overview of our study and conclusions.

II. REDLATE WORK

In the past, when we wanted to do information related to machine learning, the challenge we faced was that it was not easy to get data from the source. The characteristics of the Internet of Things are its ease of access to data. Intelligent learning models can be quickly built by comparing current data sources with known digitized and passable features and submitting that data to AI for learning. By comparing and multiplying the results of different terms, potential solutions approach the truth. The accompanying study on technological development, such as cloud-fog fusion, the Internet of Things system, machine learning and alliance learning, is as follows:

A. Cloud-Fog-Edge Computing:

Cisco proposed fog computing. Its application model is mass storage with the assistance of one or more end-user clients or near-user edge devices.[1] Where the data is associated with the appropriate reporting mechanism, the result of the data analysis may be returned. By the automatic construction of the network environment of communication between the devices, without having to a basic Internet platform, to realize the mutual communication between the mutual devices.[2][3]

It can be used to control, configure, measure and manage service monitoring and data collection. Edge computing further incorporates the concept of distributed computing, decentralizing intelligence to edge devices, allowing real-time processing and analysis of data near the source of data collection. In advanced computing, data does not need to be uploaded directly to the cloud or to a centralized data processor.

B. IoT data collection and learning:

The Internet of Things focuses on the last mile of device control at the end node. Due to the limitation of the operation and control of the Internet of Things, only one individual control of related material may be carried out.

When IoT is discussed in greater detail, domestic intelligence, industrial intelligence and agricultural intelligence. Including communication constraints, appliance constraints, energy consumption constraints, etc., the dilemma of applying high construction costs is quite common. From the research results presented by Shajulin Benedict 2020 [4][5], in a large number of IoT devices, a more complex communication system is required. It was discussed that in a wide variety of device communication and communication,

2693-9371/22/$31.00 ©2022 IEEE
DOI 10.1109/QRS-C57518.2022.00113

there will be related problems in communication, power supply, transmission safety, etc.

Therefore, In Shajulin study[4], proposed a blockchain network which combines the three levels of on-board computing, fog and cloud, and established a set of methods using a server-less architecture to facilitate communication between developers.

C. Machine Learning:

Machine learning is part of artificial intelligence and is also widely used in visual speech recognition and robotics. Given the problem scope and training data, feature selection can be performed from the data, and the constructed model selection can be used as the predicted learning [6]. Therefore, in the research and development of machine learning, it is necessary to connect the various steps in series to speed up each cycle and further shorten the development time.

For machine learning, many statistical theorems are used as the main core basis for reasoning. In the field of computer science, an efficient algorithm will be required to solve optimization problems and handle large data spaces. When a model is machine learned, the reasoning ability of the algorithm must also be more efficient than before. From disk space and time complexity is also as important as prediction accuracy.

D. Federated Learning:

In response decentralized data sources and data security issues, Google proposed federated learning in 2016. The data don't need to leave the device, each train the model on its own device [7]. Federated learning enables endpoints to collaboratively learn and share predictive models. Keeping all training data on-device at the same time decouples the power of machine learning from the need to store data in the cloud.

Federated learning requires a global training model [8][9], a general learning method that can be applied to relevant data in the field and gradually adjust the learning model according to the learning results. The training process can be performed on the user side using private data, and modified according to the parameters, weights, gradients and other characteristics of the adjusted data model.

E. Machine to machine communication:

Machine-to-Machine is a broad project that can be used to describe any technology that enables networked devices to exchange information and perform actions without human assistance. M2M communication is often used for remote monitoring. For example, a vending machine might send dispenser information when a particular item is running out. M2M communication is an important aspect of warehouse management, remote control, robotics, traffic management, logistics services, supply chain management and fleet management. The need for devices to communicate with each other forms the basis of a concept known as the Internet of Things.

Based on the OM2M technical standard [10], OM2M is used to provide a RESTful API to enhance interoperability with device endpoints[11]. At the same time, the design of the agent combined with the M2M management communication mechanism provides a more flexible application. In addition to the communication and communication between the node status and the cloud, it can further realize the mutual communication control between machines and machines. Control the nodes through the agent recording operation. Service process and state control management.

III. DESCRIBES THE DESIGN OF THE STUDY STRUCTURE E

In order to build an automated learning experimental environment, we design a set of IoT experimental environment. Collects relevant data and compares equipment control and operation conditions, establishes intelligent agents for component control and learning analysis, and establishes machine learning comparison data with data analysis. Using the agent framework design an application data learning environments in different domains whose role is to control the collection of data.

Combined with the federated machine learning capabilities, device endpoints have learning data that can be turned into valuable knowledge information. This study takes plant growth monitoring learning as an example. Through monitoring learning, knowledge of plant growth parameters can be obtained, and mutual learning between edge computing devices can be derived. Based on computer measurements, predictions can be made, and relevant growth management parameters adjusted. Control all factors that affect plant growth. Combine machine learning models to expand the Internet of Things from limited local computing to the cloud for core data collection and machine learning training calculations

This study adopts the quasi-experimental research method, selects samples to participate in the selection of planting and crop growth IoT sensor equipment information, and constructs data collection, learning and data analysis. The samples were divided into experimental group and control group. The first part compares traditional planting and parametric automatic planting validation learning. The second part adopts automatic planting and automatic planting after data factor analysis for verification and learning comparison.

The discussion of the core research structure can be divided into three main items for discussion.

A. IoT data collection and learning:

Design a cloud-edge agent system, create and expand an IoT experimental data collection environment, establish automated sampling data collection and sampling for planting parameters and plant growth information.

B. Verification of edge computing operations:

Build leading-edge computer models for smart IoT environments. The parameters are sampled based on the traditional plantation learning factor, and then the parameters of the automatic control equipment are adjusted by the edge calculation. Complement the experiment of correlation between traditional and automated planting and verification of control group planting results.

C. Analysis of node machine learning parameter model:

In response to environmental changes, in different reference natural environments, how to use the data to predict and adjust the factors necessary for plant growth for data analysis. We support using the Apriori algorithm for conducting a correlation analysis and selecting factors. IoT data for analyzing planting edge devices and comparing and contrasting group and treatment group results.

IV. RESEARCH IMPLEMENTATION

A. Agent Service Sampling Method:

The Arduino development board is used to obtain sensor data. For this study, vegetative growth was used as a validation case study. If vegetation detection data are sampled every 10 seconds according to traditional frequency of detection. Data does not change significantly after detection, and a large quantity of identical data is obtained for the same factor. [12][13]. The sampling method uses an algorithm to compute an automatic sampling time history.

In addition to calculating the correction frequency, the Frontal Smart Agent automatically calculates the digital simulation compensation value when the result is returned to the cloud system. When farmers perform soil watering operations, soil humidity changes and increases..

When farmers perform soil watering operations, soil humidity changes and increases. Generally, it takes between 2 and 3 hours for the humidity to decrease progressively. At present, the digital information returned by the moisture sensor is of no importance.

```
#include <DHT.h>
#define DHTPin 2
#define DHTType DHT11
DHT dht11(DHTPin,DHTType);
int SoilPin = A0;
void setup() {
  Serial.begin(9600);
}

void loop() {
  int result = analogRead(SoilPin);
  delay(10000);
  float h = dht11.readHumidity();
  float t = dht11.readTemperature();
  if (isnan(h) || isnan(t)) {
    Serial.println("Error");
    return;
  }else{
    Serial.print(t);
    Serial.print("C  ");
    Serial.println("");
    Serial.print(h);
    Serial.print("%");
    Serial.println("");
  }
  delay(10000);
}
```

```
//MQTT
const char* mqtt_server = "192.168.30.44"; //ip ex:192.127.51.27
const char* mqtt_user = ""; //null
const char* mqtt_password = "";//null
const char* clientID = "trasmit_client"; //unique_ID ex:trasmit_client

//VARS
const char* topic = "temp";

WiFiClient espClient;
PubSubClient client(espClient);


void callback(char* topic, byte* payload, unsigned int length) {
  Serial.print("Message arrived by ");
  Serial.print(topic);

  String s = "";
  for (int i=0;i<length;i++) {
    Serial.print((char)payload[i]);
    s += (char)payload[i];
  }
```

Figure 1. The system automatically samples the data information returned by the agent

B. Cloud-edge integrated data exchange and data analysis architecture

As a result of the agent's construction, the established device is able to communicate with the service platform independently. At the same time, under the surveillance status of the platform, all devices can be managed on time. Including monitoring the state of the device, obtaining data exchange records as part of the surveillance state of the device, and assisting the terminal device in processing delivery calculation requirements. Enable the platform to communicate directly with the device and deliver software information exchange services in both directions.

In the research, Raspberry Pi is used to build the agent base service. When the device itself has basic computing power, it can send analysis or data collection tasks through the platform. A server that manages computing services.

In our environment, we use a service agent, which can be used for quickly configuring roles and assigning tasks using the platform. The server side sends the task events to the device side, including software system update messages, boot software service messages on the device side, etc. The device-side system will facilitate computation and analysis and return the results to the device-side platform to enhance the performance of the object. The calculation of the processing of the node information is decided. The agent analyzes and judges, and starts the communication control event, so that the node can communicate directly with the node, and the effectiveness of the service is improved.

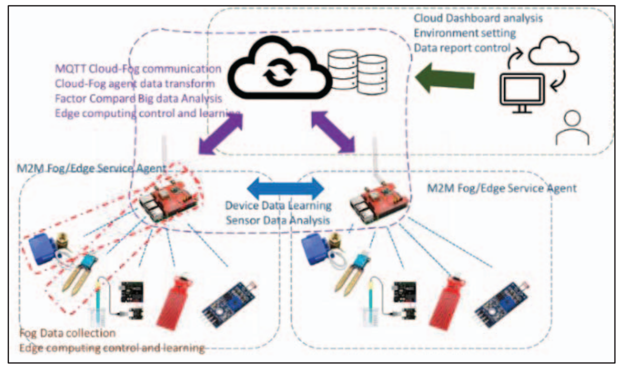

Figure 2. Experimental Simulation Environment Construction Architecture

V. EXPERIMENTAL SIMULATION RESULTS

Use the device's medium-range federated learning model to conduct diagnostic analysis of learning outcomes on the various sides. The architecture of the study design system utilizes officers who monitor information about learning data over time. Data on the trajectory of last-mile correlation data to support analysis of mother sample data.

The detection and calculation of the frequency shall be carried out by the detection apparatus. The device is a triggered detection equipment to collect frequency data based on the different conditions of the experiment, in order to perform the extraction and exploration of the environmental data association. Various parameters can have an impact on growth results, service analysis, etc. The same data is organized, consolidated and stored according to data engineering models, and the long-term and short-term data are then analysed using basic collective intelligence techniques. Further plant growth-related data for long-term and short-term trajectory analysis.

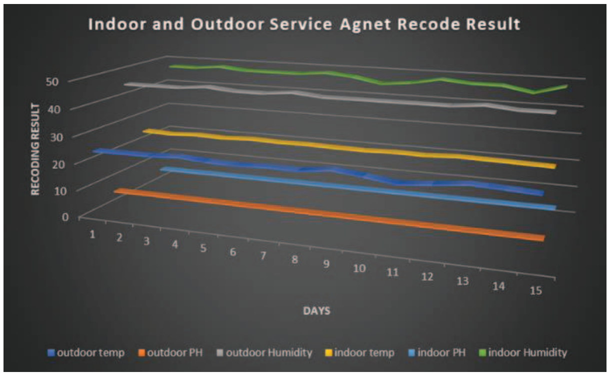

Figure 3. Automatic monitoring of data learning situations in two different environments

Figure 4. Data Collection Agent Sensing Device (include LoRa Mini +YL69+ DS18B20)

In the research, through automatic data collection, the data collected every 10 seconds every day is collected and returned by the agent, and the daily average value is automatically calculated.

This study compared the exterior and interior control records and analysed the corresponding temperature, humidity and pH values. After 15 days of sampling, there was no significant change in soil pH, and all of them were in the new soil added state. Utilizing automated Edge planning and verification. We try to observe that there were still some differences between indoor and outdoor plants in the same place, particularly in terms of temperature and moisture.

At present, the sampling days are small, and it is impossible to compare the significant differences in the impact of the environment on plant growth. If observed continuously, the influence coefficient of numerical change on plant growth will be different.

Combined with the federated learning model under study, device endpoints successfully communicate parameter values back to the system. In our research, we combine machine learning and federated learning to define N data owners $\{F_1,...F_N\}$ who all want to train machines by combining their respective models of data $\{D_1,...D_N\}$ study.

In this study, we put all the data together and then use $D = D_1 \sim D_N$ to train the model M_{sum}. A federated learning system is a learning process in which data owners collaboratively train a model M_{fed}. During this process, any data owner F_i will not submit its data Di to other devices. After repeated learning, the accuracy results of M_{fed} marked as V_{fed} should be very close to the values of M_{sum} and V_{sum}. The final delta is a non-negative real number, as in (1), so the device can provide single-point optimization parameters through the analysis results of different endpoints, and further improve the environmental parameter information of planting conditions.

$$|Vfed - Vsum| < \delta \quad (1)$$

Comparative analysis of experimental results related to edge federated:

Through the linear aggregation analysis of support and reliability, the correlation linear correlation graph is sampled, and the cross-correlation comparison of improvement degree is carried out interactively. Scatter plots are used to filter out attributes that exhibit linear or suspicious clustering relationships, strong or weak correlations. It is necessary to look at the relationship between the X-axis and the Y-axis (attribute value), establish the data factor distribution results for feedback learning, and improve the correlation between the endpoints in the automated plant growth curve for comparison, discussion and verification.

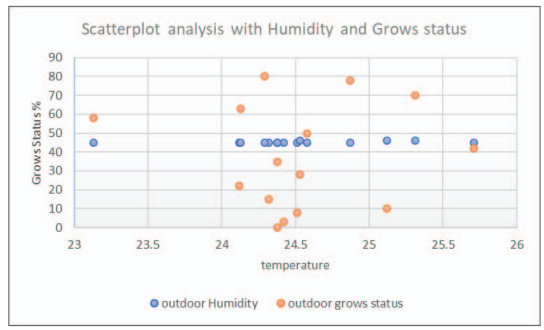

Figure 5. Outdoor soil humidity scatterplot with grows status result

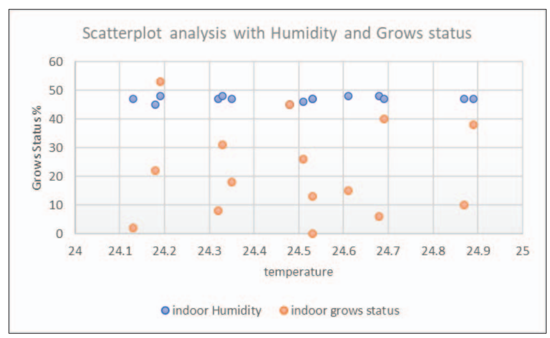

Figure 6. Outdoor soil humidity scatterplot with grows status result

Based on the correlation analysis, the data show that soil moisture in the growth cycle of external plants is directly linked to plant growth in Figure 5. Observing temperature and humidity, by outside air and exposure to sunlight. When the temperature is between 24°C and 25°C, the growth cycle of plants is more suitable, and soil moisture and average temperature are also more conducive to plant growth plant.

In relative terms, we can see that the correlation between plant growth is more scattered in the same indoor environment. The main reason is that we suppose it can be because the time of the sun. It is only used when the lights are on indoors, which affects plant growth and interferences with moisture influences.

Furthermore, the data show that although soil moisture and temperature are not very different from the outside soil. Plant growth is more diverging and may not be as stable as outside growth. It is observed that in order to better control the domestic service, it may be necessary to determine the varying factors. Which interfere with the stable growth of plants, including other interference resources such as sunlight and walkers. Adjusting the growing conditions of plants indoors is compatible with outdoor planting, which enhances the quality of the IoT monitoring service.

VI. Conclusion

By analyzing the results of this study, many studies discuss the control and management of IoT-based device parameters. Many applications on the Internet of Things in the past can be analysed by data analysis, but the issue of end-state data has not been discussed in the past.

In this study, the agent uses to learn the data throughout the devices. Learning outcomes are combined with joint training to hide relevant data, provide different nodes to obtain information on knowledge between nodes, and provide a basis for managing other terminals. The analysis of results improves the quality of the IoT application data management service and the follow-on audit cycle. Automated intelligent learning is not confined to one final point, but reinforces the collective formation of federation knowledge. At the same time makes intelligent the discovery of other data that can be explored and analyzed.

In this research experiment, the data collection and analysis of edge computing has been completed. Design automated monitoring is carried out through the agent, and the alliance model has been established, so that the devices can learn the required data from each other. Automatic control, federated information on learning trends, automatic control of relevant factors using Edge, construction of models and applications for managing intelligent agent services, automated data comparison analysis and monitoring.

Acknowledgment

This research was funded by National Science and Technology Council of the Republic of China grant number NSTC 111-2221-E-031 -006 -.

References

[1] Anshul Gandhi, Parijat Dube, Alexei Karve, Andrzej Kochut, Li Zhang "Providing Performance Guarantees for Cloud-Deployed Applications" IEEE Transactions on Cloud Computing, pp. 269 – 281, Volume 8, Number 1, 2020

[2] Michaela Iorga,Larry Feldman,Robert Barton, Michael J. Martin, Nedim Goren, Charif Mahmoudi "The NIST Definition of Fog Computing" NIST special publication, pp. 800-191, 2018,

[3] Enzo Baccarelli, Michele Scarpiniti, Alireza Momenzadeh, Sima Sarv Ahrabi "Learning-in-the-Fog (LiFo): Deep Learning Meets Fog Computing for the Minimum-Energy Distributed Early-Exit of Inference in Delay-Critical IoT Realms" IEEE Access Vol. 9, P.P. 25716 – 25757 08 February 2021

[4] Kunpeng Yang, Baofeng Zhang, Jinwei Zhang, Junchao Zhu " Design of Remote Control Inverter Based on MQTT Communication Protocol " 2021 IEEE International Conference on Mechatronics and Automation (ICMA), Takamatsu, Japan, Aug. 2021.

[5] Pengfei Hu, Wai Chen, Chunming He, Yiping Li, Huansheng Ning "Software-Defined Edge Computing (SDEC): Principle, Open IoT System Architecture, Applications, and Challenges" IEEE Internet of Things Journal, pp. 5934 – 5945, Volume 7, Number 7, 2020

[6] Salman, Ola, et al. "A machine learning based framework for IoT device identification and abnormal traffic detection." Transactions on Emerging Telecommunications Technologies 33.3 (2022): e3743.

[7] Alam, Tanweer, and Ruchi Gupta. "Federated Learning and Its Role in the Privacy Preservation of IoT Devices." Future Internet 14.9 (2022): 246.

[8] Rey, Valerian, et al. "Federated learning for malware detection in iot devices." Computer Networks 204 (2022): 108693.

[9] Latif U. Khan, Walid Saad, Zhu Han, Ekram Hossain, Choong Seon Hong "Federated Learning for Internet of Things: Recent Advances, Taxonomy, and Open Challenges" IEEE Communications Surveys & Tutorials, Vol. 23, Issue: 3, thirdquarter 2021

[10] Christian Makaya, Ming-Yee Lai, Fuchun Joseph Lin "Over-the-air remote management and control of IP-based M2M devices" 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT), Milan, Italy, Dec. 2015

[11] Paolo Bellavista, Luca Foschini, Nicola Ghiselli, Andrea Reale "MQTT-based Middleware for Container Support in Fog Computing Environments", 2019 IEEE Symposium on Computers and Communications (ISCC)

[12] Dong-Meau Chang, Yao-Hong Tsai, Tse-Chuan Hsu, and Hsien-Wei Tseng "Novel OneM2M Communication Mechanism Based on Labeling of IoT Devices" Sensors and Materials, Volume 33, Number 2(3) (2021)

[13] Tse-Chuan Hsu, Hongji Yang, Yeh-Ching Chung. Ching-Hsien Hsu "A Creative IoT Agriculture Platform for Cloud Fog Computing" Sustainable Computing: Informatics and Systems, Volume 28, December 2020, 100285