IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⧖ Impact Factor 7.39 ⧖ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.115148

Data Collection and Analysis in a Smart Home Automation System

Mr. Krishna. M. Patel[1], Mr.L.N. Yadav[2], Mr.V.M. Rakhade[3]

Student, Department of Computer science and engineering, Shri Sai College of Engineering and Technology, Bhadravati, India[1]

Asst. Professor, Department of Computer science and engineering, Shri Sai College of Engineering and Technology, Bhadravati, India[2]

Asst. Professor, Department of Computer science and engineering, Shri Sai College of Engineering and Technology, Bhadravati, India[3]

Abstract: We implement a basic solution for accumulating and obtaining user data within the existing home automation system in this paper. Data collection modules that operate on the home automation gateway and within the home automation cloud are implemented, allowing us to connect to the already existing big data middleware platform. It's the first step toward constructing the extensive data storage and analysis component for the existing housing automation solution. The collected data can be used to enable various scenarios of interest to end-users, such as the detection of emergencies and system irregularities.

Keywords: Data Gathering, Big Data, Machine Control, Cloud Computer.

I. INTRODUCTION

Nowadays, industries that have access to information and data have a significant advantage over their competitors. As a result, various mechanisms for collecting and storing system usage data are being developed and deployed. However, in a world where information is not difficult to obtain, it is frequently difficult to distinguish between relevant and irrelevant data. The next step in the process would be data analysis after deciding which data should be collected and implementing data collection mechanisms. This is where the real difficulty lies. Data processing and analysis tools must be extremely dependable, efficient, and secure while remaining relatively fast. For security and efficiency, data is typically stored across multiple clusters and replicated.

In this paper, we concentrate on data collection and analysis for the existing home automation (HA) system. The devices in the HA system are classified into two types: actuators and sensors. Actuators are devices that system users can control, such as lights, dimmers, plugs, and so on. Temperature, humidity, motion, smoke, and other sensor device properties, on the other hand, can only be monitored by the user. Sensor devices in the HA system generate massive amounts of data because they are constantly collecting new information. Actuators provide more information about the user's daily habits and routines because they not only monitor environmental parameters but are also controlled directly by the user. As a result, distinguishing between relevant and irrelevant information is critical.

Our goal is to collect data about system usage to analyze it and enable some interesting scenarios for end users. For example, the system learns about the user's daily habits by storing and analyzing device consumption and sensor data in the smart home environment. The system can recommend actions or HA system settings to the user based on the detected patterns. Irregular behavior can also be detected and interpreted as an emergency. Furthermore, recorded HA system history can be simply replayed within the specified time frame (days/weeks). To prevent intrusions, this enables the creation of scenarios that mimic regular user behavior when the household is empty.

We want to add the cloud data collection service to the existing HA system as the first step toward building the extensive data analysis framework and enabling the various scenarios mentioned above. The implemented data collection service must continuously monitor all events in the environment and record them in the database. Large amounts of data will be collected and analyzed over time to enable features such as automatic adaptation of the system to user habits and detection of emergencies. For the time being, we will limit data analysis to simple monitoring of history within the specified timeframe.

DOI: 10.17148/IJARCCE.2022.115148

In this paper, we first present the existing HA system's architecture. Then, we build a basic data collection and analysis solution using existing big data technologies and an IoT middleware platform. Finally, we test the solution's functionality.

II. THE ARCHITECTURE OF A HOME AUTOMATION SYSTEM

In this section, we will first describe the existing HA system's architecture. Then, we identify the system points where the interface to the data collection service can be implemented.

Figure 1 depicts the architecture of the existing system. As previously stated, devices within the HA system can be either sensors or actuators. All of the devices in the HA system are referred to as nodes. The gateway is connected to all of the system's nodes . IP, Z-Wave, and ZigBee protocols are used to communicate between devices and the gateway. The HA gateway manages all local network devices and communicates with the HA cloud . The HA cloud allows for remote control of the HA system via the web interface and mobile applications.

Our goal is to create a cloud-based data collection service that all system users can use. Because the nodes in the HA system are made by different manufacturers, it is not possible to extend their functionality and enable them to communicate directly with the data collection service. Furthermore, this would be extremely inefficient. As a result, either the HA gateway or the HA cloud should collect all system usage information, format it, and send it to the data collection service. Because the gateway device has limited internal storage and memory, it is preferable to delegate this task to the HA cloud.The log upload strategy is configured so that the HA cloud communicates with the Kaa server.

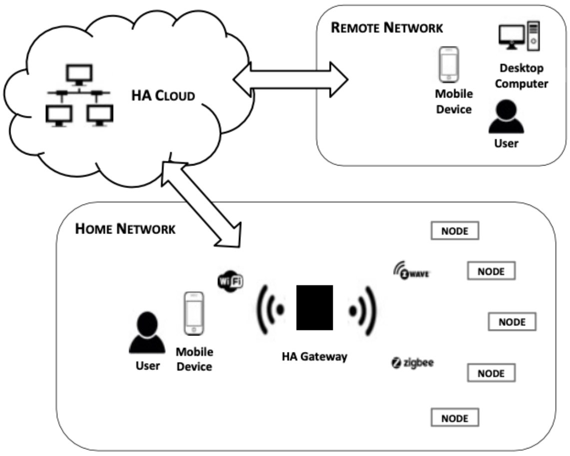

Fig. 1 – Architecture of the existing HA system

III. IMPLEMENTATION

This section describes how to integrate the data collection service into the existing HA system.
The implementation diagram is shown in Fig. 2. The data collection service was implemented within the data collection cloud, which operates independently of the existing HA cloud. The Kaa IoT platform is used as middleware in the implemented data collection cloud . This platform provides us with a highly customizable and dependable server that can run in a cluster. The server handles database communication and provides interfaces to Cassandra, MongoDB, MariaDB, and PostgreSQL. We chose a NoSQL database because HA data is of variable length and structure. We chose Cassandra over MongoDB in particular because of its higher throughput. We used Zeppelin to perform simple data retrieval and log monitoring from the desired timeframe.

To enable endpoint communication with the Kaa server, data collection modules had to be installed on each endpoint. In our case, we did not want individual home network nodes to communicate with the Kaa server. Instead, as shown in Fig. 2, we implemented data collection modules only within the HA cloud that communicate with the HA gateway.
The Kaa platform includes SDKs for C, C++, Java, and Objective C endpoint applications by default. Node.js is used to write the existing HA cloud. In order to build the data collection module for HA cloud, we created a Node.js add-on that exposed the C++ SDK functionality.

DOI: 10.17148/IJARCCE.2022.115148

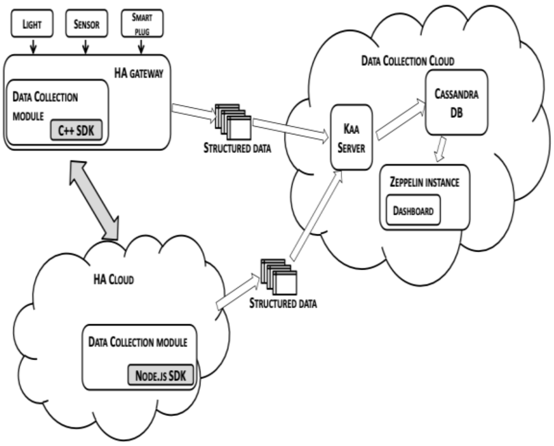

Fig. 2 – HA system with data collection and analysis cloud

IV. TESTING

We performed a simple test of logging the temperature from the sensor to validate the implementation. The HA cloud-based data collection module checked the temperature from the sensor on a regular basis and sent it to the data collection cloud. Aside from the actual sensor reading, all logs included the timestamp, user information, and the unique ID of the sensor device within the HA system. The logs were successfully delivered, and the log history was visible in the Zeppelin dashboard. The average log delivery time was 5.95 ms, with a range of 6 to 10 ms.

V. CONCLUSION

We were able to quickly enable and validate the data collection mechanism within the existing HA system by utilizing the Kaa middleware platform. Further work will be focused on expanding the existing HA system's extensive data collection and analysis component.

REFERENCES

[1]. Asst.Professor, Mr L.N.Yadav , “Predictive Acknowledgement using TRE system reducing costs and bandwidth” IJRECE VOL. 7 ISSUE 1 (JANUARY- MARCH 2019) ISSN: 2348-2281
[2]. Asst.Professor, Mr V.M. Rakhade, “ Reducing Routing Distraction in IP Networks using Cross-Layer Methodology” (ICRTEST 2017) Volume: 5 Issue: 1(Special Issue 21-22 January 2017) ISSN:2321-8169
[3]. M. Tucic, R. Pavlovic, I. Papp and Dj. Saric, “Networking layer for unifying distributed smart home entities", Proc of IEEE Telfor, 2014 .
[4]. V. Moravcevic, M. Tucic, R. Pavlovic and A. Majdak, “An approach for uniform representation and control of ZigBee devices in home automation software”, Proc. of IEEE 2015 ICCE– Berlin, 2015
[5]. M.Sekulic, I. Lazarevic, M.Bjelica and V. Pekovic, “Asynchronous application programming interface library for distributed home automation software”, Proc. of IEEE 2015 ICCE– Berlin, 2015
[6]. H. Dai, S. Zhang, L. Wang and Y. Ding, “Research and implementation of big data preprocessing system based on Hadoop”, Proc. of IEEE 2016 ICBDA, 2016
[7]. Apache Zeppelin [Online] Available: https://zeppelin.apache.org/
[8]. Kaa Project, [Online] https://www.kaaproject.org/overview
[9]. C++ Addons, [Online] Available: https://nodejs.org/api/addons.html