M2M Device Cooperation Method
Using iHAC Hub and Smart Speaker

Kosuke Hayashi
Graduate School of Science and Technology
Meijo University
Aichi, Japan 468–8502
kosuke.hayashi@ucl.meijo-u.ac.jp

Hidekazu Suzuki
Graduate School of Science and Technology
Meijo University
Aichi, Japan 468–8502
hsuzuki@meijo-u.ac.jp

Abstract—Smart speakers with AI assistant, such as Amazon Echo and Google Home, and compatible IoT devices are emerging. The authors have developed an iHAC Hub that realizes coordinated operation of IoT devices that can be controlled by communication protocols. This paper proposes a cooperation method for M2M devices by using iHAC Hub and a smart speaker. In the proposed method, we provide a framework for cooperating environment sensing IoT devices and smart speaker compatible IoT devices. The iHAC Hub generates voice commands for controlling IoT devices by using the cloud service and reproduces them toward the smart speaker to realize cooperative control of the smart speaker compatible IoT device based on the environment information.

Index Terms—IoT, M2M, Smart Speaker, Cooperation

I. INTRODUCTION

With the advent of smart speakers equipped with artificial interactive intelligence (AI) assistants such as Amazon Echo and Google Home, users can control smart devices by talking to the smart speakers. Furthermore, with the spread of Internet of Things (IoT) devices, environmental information such as temperature, humidity, illuminance, air quality, etc. that humans have physically felt are acquired as digital data and has come to be used in various IT services. However, current IoT systems consisting of smart speakers and corresponding IoT devices require human beings to utter control instructions, and have not reached the realization of true Machine-to-Machine (M2M) system in which those devices cooperate autonomously.

On the other hand, iHAC Hub, which we have proposed so far [1], can control smart home appliances based on environmental information sensed by IoT devices installed in a home network. However, smart appliances that can be controlled by the iHAC Hub must support communication protocols such as ECHONET Lite and DLNA. Many commercially available smart speaker devices do not have these protocols, and the iHAC Hub could not control these products directly.

Therefore, in this paper, we propose an M2M device cooperation method without human intervention by collaborating iHAC Hub and smart speaker. The iHAC Hub generates and plays voice instructions for controlling smart speakers using a cloud service, and realizes smart control of IoT devices without communication protocol by using smart speakers.
connect between iHAC Hub and smart speakers is required. Therefore, we propose a method to achieve M2M device cooperation by iHAC Hub uttering voice instructions instead of human and asking the smart speaker to control IoT devices.

Fig. 1 shows the extended iHAC framework and an overview of the proposed M2M device cooperation system. The functions to be added in the iHAC framework are the voice command generator and the voice command player. First, IoT devices sense environmental information and upload to the cloud using MQTT. Next, a user creates a recipe as shown in Fig. 2 using an iPad on which the iHAC application is installed and sends it to the iHAC Hub.

The recipe engine in the iHAC Hub analyzes the received recipe and acquires sensing data from the cloud using process such as MQTT Subscriber according to the method specified by “fetch” key. If the received sensing data satisfies the “condition” object described in the recipe, the iHAC Hub executes the processing specified by the “action” object. Here, if “VoiceInstruction” is specified in the “controlMethod” key, the iHAC Hub selects the wakeup word from the type of smart speaker indicated by “controlDeviceId” key, combines the wakeup word with the text data set with the “utterance” key, and converts it into voice command data. Furthermore, the iHAC Hub instructs the smart speaker by playing the generated voice command data. After that, the smart speaker controls the IoT device based on the recognized voice instruction.

By the above processing, it becomes possible to control the smart speaker compatible IoT devices based on the environmental information sensed by other IoT devices.

### IV. IMPLEMENTATION AND EVALUATION

#### A. Implementation

We developed a prototype of iHAC Hub using Raspberry Pi 3. In this prototype, we adopted Google Cloud Text-to-Speech [4] as a cloud service for generating voice commands. The voice command generator and player functions added to the iHAC framework were implemented in Python. Since iHAC framework is developed in C language, it was implemented to call Python function from C language.

We verified the operation of the proposed method using two types of smart speakers (Google Home and Amazon Echo) and a Wi-Fi smart plug supporting both models and an electric fan. As a result, as shown in the recipe in Fig. 2, when the room temperature reaches 30 degrees or more, iHAC Hub plays a voice command saying “OK Google, turn on the electric fan” and we confirmed that the fan turned on and worked.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>input (utterance)</td>
<td>Alexa, Turn on the electric fan. OK Google, Turn on the electric fan.</td>
</tr>
<tr>
<td>languageCodes</td>
<td>en-US</td>
</tr>
<tr>
<td>name</td>
<td>en-US-Wavenet-A</td>
</tr>
<tr>
<td>ssmlGender</td>
<td>MALE</td>
</tr>
<tr>
<td>speakingRate</td>
<td>0.25–4.00 (step: 0.25)</td>
</tr>
<tr>
<td>pitch</td>
<td>-20–20 (step: 5)</td>
</tr>
</tbody>
</table>

#### B. Evaluation

We evaluated whether voice commands generated using the Google Cloud Text-to-Speech API can be correctly recognized by commercially available smart speakers. A voice audio file generated by changing speaking rate and pitch with settings such as TABLE I was played 10 times each to the smart speaker of both Amazon Echo and Google Home.

Fig 3 and Fig 4 show the results of speech recognition rates for Amazon Echo and Google Home. From these results, we confirmed that when speakingRate is in the range of 1.0 to 1.25 and pitch is in the range of -5 to 5, the recognition rate of both Amazon Echo and Google Home is 100% even for the speech generated by Google Cloud Text-to-Speech.

In addition, we played the voice commands 100 times generated from the settings that both the recognition rates of Google Home and Amazon Echo were 100%, and investigated the recognition rate. As a result, the recognition rate of both Amazon Echo and Google Home was almost 100%. Therefore, it was not necessary for the human to speak, and it was proved...
that the M2M device cooperation system is possible by the iHAC Hub talking to the smart speaker.

V. CONCLUSION

In this paper, we proposed an M2M device cooperation method without human intervention by collaborating iHAC Hub and smart speaker. As a result of evaluating the recognition rate of smart speakers for mechanically generated voice, it was shown that IoT devices can be controlled with practically no problem.

REFERENCES