The Voice Controlled Internet of Things System

Chien-Hung Lai
Department of Electronic Engineering
National Taipei University of Technology
Taipei, Taiwan
laisan86@gmail.com

Yuh-Shyan Hwang
Department of Electronic Engineering
National Taipei University of Technology
Taipei, Taiwan
yshwang@ntut.edu.tw

***Abstract*—This paper presents an Internet of Things (IoT) system which is controlled by human voice and natural language, including appliance and indoor environment sensor. The appliances will be bundled with a socket controlled by relay and a MediaTek LinkIt 7697, which is marketed Arduino compatible development board. The environment sensor was DHT11, a marketed indoor environment sensor, which also bundled with a MediaTek LinkIt 7697. The protocol to communicate with both MediaTek LinkIt 7697 is Message Queuing Telemetry Transport (MQTT). The gateway for this IoT application is a QBoat Sunny, it is home IoT gateway of Taiwan QNAP Corp. which has built-in Linux and numerous functions. Based on these functions, the IoT system can access voice recognition and natural language parsing functions of Google to accomplish the jobs.**

Keywords—Voice recognition, MQTT, Natural language

I. INTRODUCTION

There are many occasions in life that are perfect for voice notifications or reminders, such as many people who do not carry a smart phone with them at home, and for some bedridden people, it's easier to communicate what they're going to do by voice. For instance, Amazon's "Axela" are very welcome in home automation market now [1].

Our research decided to create an IoT system is similar to "Axela", which is able to interactive with users by natural speech. This IoT system will record user's voice into a media file in memory, after set up duration, it uploads the media file to Google's Speech to Text (STT) service to get plain text [2].

After the IoT system get the returned plain text from Google, will call the on line natural language understanding engine "DialogFlow", which is also owned by Google now [3].

The two-pass procedure produce instructions as defined in "DialogFlow" at last, our IoT system recognize the returned instructions to perform action like turn on and off the light, or answer the temperature and humidity to user by synthesis voice generated by "eSpeak" this open-source package [4-10].

II. DESIGN AND IMPLEMENTATION

The circuit diagram of IoT node is shown in Fig. 1(a) and Fig. 1(b), both MediaTek LinkIt 7697 use Wi-Fi to link with our IoT system. Fig. 1(a) is the node scribe to MQTT service on our IoT system for control the light is on or off, Fig. 1(b) is the node published to MQTT service on our IoT system for report indoor temperature and humidity sensed by DHT11 every 2 seconds. The infrastructure of whole study is shown in Fig. 2.

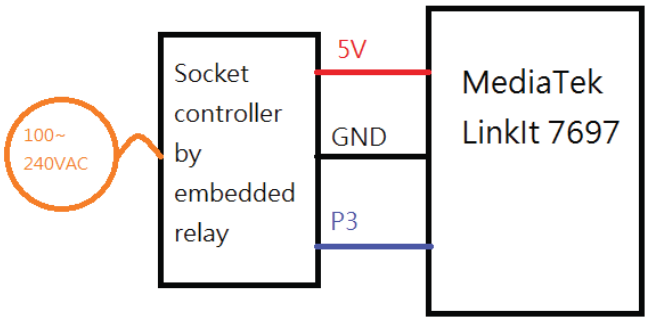

Fig. 1(a). The IoT node to control AC socket.

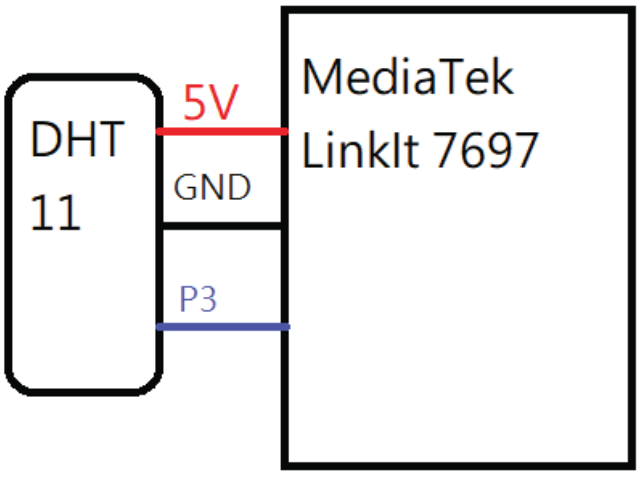

Fig. 1(b). The IoT node to sense the environment.

978-1-5386-1445-7/18/$31.00 ©2018 IEEE

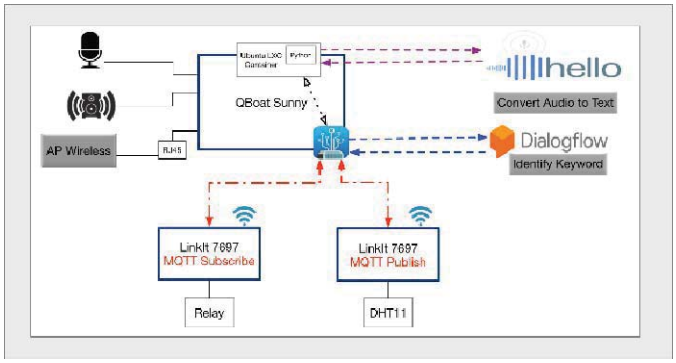

Fig. 2. The proposed infrastructure.

The IoT system contains two programs, one is written in Python to import "Speech_Recognition" package to get plain text by upload the recorded voice to Google STT service. After get the returned plain text from Google, this program written in Python will send the plain text to our rules defined on "DialogFlow" to get clear instructions. After that, the instructions will pass to our second program.

The clear instructions return from "DialogFlow" will be passed to our second program written in Node-red. This second program will determine which IoT node shown in Fig. 1(a) and Fig. 1(b) shall be invoked. For the node shown in Fig. 1(a), the second program publish "0" means turn off the light, "1" means turn on the light. The rules we defined for node shown in Fig. 1(a) has presented in Fig. 3.

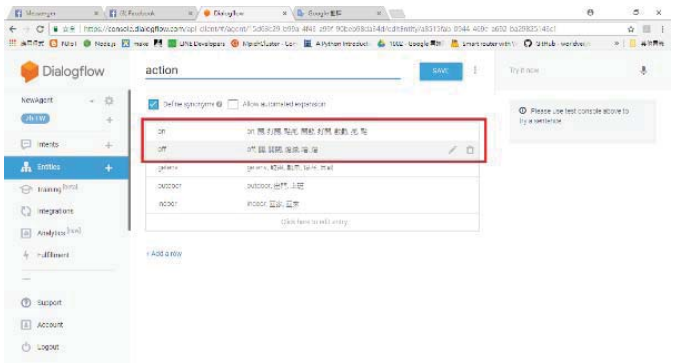

Fig. 3. Rules for node to turn on and off the light.

For the node shown in Fig. 1(b), it publish indoor temperature and humidity, so our second program will read the one of the values as instruction specified, pass the value to the first program by a callback function. The callback function will trigger a string to "eSpeak" package for produce synthesis voice answer. The rules we defined for node shown in Fig. 1(b) has presented in Fig. 4.

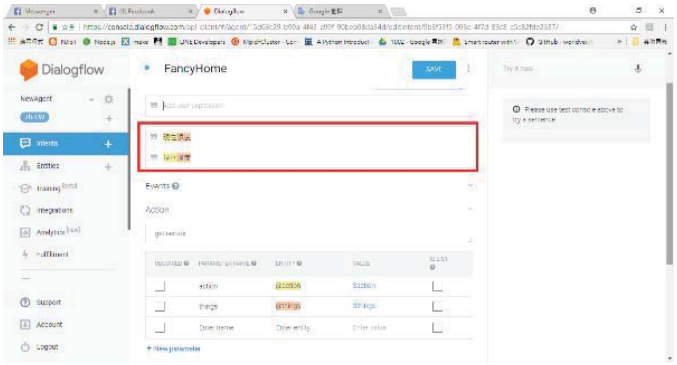

Fig. 4. Rules for indoor environment sensor.

III. EXPERIMENT

The chart we drawn by Node-red shown in Fig. 5, the microphone to record voice will be triggered by the block marked in red circle.

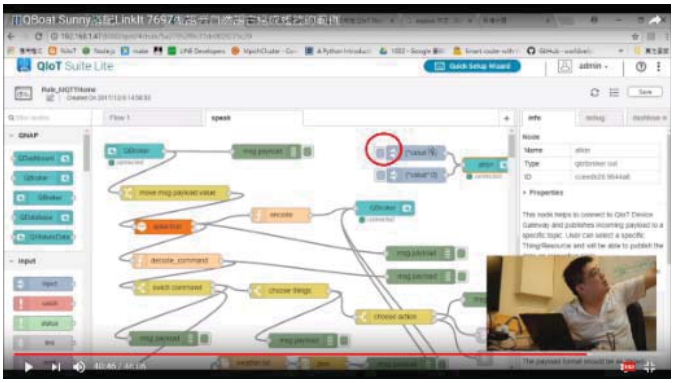

Fig. 5. The block diagram in Node-red.

We can command the system by microphone, and get the response by speaker, like shown in Fig. 6.

Fig. 6. Command the system by microphone.

IV. DISCUSSION

Our ideal is to make everything easy by develop artificial intelligence (AI) applications. This is the first demonstration we provided to Taiwan QNAP Corp. There are still many different methods to interactive with people, such as hand-gestures, wearable devices, image recognition…and so on.

With this experience, Taiwan QNAP Corp. plans to develop their own AI platform system, after that more applications will be released one after another.

ACKNOWLEDGMENT

We are grateful for the support of Taiwan QNAP Corp., valuable suggestions and technical supports during the system development and the provision of test samples.

REFERENCES

[1] Alexa Voice Service Overview, https://developer.amazon.com/docs/alexa-voice-service/api-overview.html [online].

[2] CLOUD SPEECH API, https://cloud.google.com/speech/?hl=zh-tw [online].

[3] Dialog flow -- Build natural and rich conversational experiences, https://dialogflow.com [online].

[4] eSpeak text to speech, http://espeak.sourceforge.net [online].

[5] QBoat Sunny, https://www.qnap.com/zh-tw/news/2017/qnap-%E5%85%A8%E6%96%B0%E6%8E%A8%E5%87%BA-qboat-sunny%E5%B0%88%E7%82%BA%E7%89%A9%E8%81%AF%E7%B6%B2%E9%96%8B%E7%99%BC%E8%80%85%E9%87%8F%E8%BA%AB%E6%89%93%E9%80%A0%E7%9A%84%E5%96%AE%E6%9D%BF-iot-%E5%BE%AE%E5%9E%8B%E4%BC%BA%E6%9C%8D%E5%99%A8 [online].

[6] MediaTek LinkIt 7697, https://labs.mediatek.com/en/chipset/MT7697 [online].

[7] Node-red, https://nodered.org [online].

[8] SpeechRecognition 3.8.1, https://pypi.python.org/pypi/SpeechRecognition [online].

[9] DHT 11 Humidity & Temperature Sensor, http://www.micropik.com/PDF/dht11.pdf [online].

[10] The demonstration this paper presented https://www.youtube.com/watch?v=ZM_-7LldXTs [online].