

# Voice-based Device Control Using oneM2M IoT Platforms

Isu Jeong\*, Jaeseok Yun\*

## Abstract

In this paper, we present a prototype system for controlling IoT home appliances via voice-based commands. A voice command has been widely deployed as one of unobtrusive user interfaces for applications in a variety of IoT domains. However, interoperability between diverse IoT systems is limited by several dominant companies providing voice assistants like Amazon Alexa or Google Now due to their proprietary systems. A global IoT standard, oneM2M has been proposed to mitigate the lack of interoperability between IoT systems. In this paper, we deployed oneM2M-based platforms for a voice record device like a wrist band and LED control device like a home appliance. We developed all the components for recording voices and controlling IoT devices, and demonstrate the feasibility of our proposed method based on oneM2M platforms and Google STT (Speech-to-Text) API for controlling home appliances by showing a user scenario for turning the LED device on and off via voice commands.

▶ Keyword: Voice command, oneM2M, OCEAN, open source, Google Speech API

## 1. Introduction

우리는 임베디드 시스템과 정보통신기술의 발달로 일상생활에서 사용하는 물건이나 산업환경에서 활용되는 기기들이 서로 연결되어 하나의 거대한 네트워크를 이루는 사물인터넷 (IoT: Internet of Things) 시대를 맞이하고 있다 [1]. 이렇게 주위의 모든 물리 또는 가상 사물들이 인터넷 연결성을 가지고 서로 정보를 공유하고 협업이 가능함에 따라 다양한 산업 분야에서 새로운 가치와 비즈니스 기회를 생성하고 나아가 사회 전반에 걸쳐서 큰 영향을 미칠것으로 예측하고 있다.

사물인터넷 시스템은 초소형 무선 센서부터 대형 클라우드 서버까지 다양한 종류의 하드웨어, 통신 네트워크, 소프트웨어로 구성되며, 이는 산업 도메인 간 데이터 공유를 어렵게 하고 산업별 수직적 시장만을 형성하게 하는 문제를 가진다. 이렇게 시장 파편화 (market fragmentation) 문제를 해결하기 위한 방법으로 표준화된 사물인터넷 플랫폼의 필요성이 대두되었으며 그 방안으로 세계 정보통신 표준화 기구들이 모여 원엠투엠 (oneM2M) 표준을 개발

하였다 [2]. 원엠투엠은 세계 어디서든지 접근이 가능한 사물인터넷 애플리케이션 개발을 위한 표준과 플랫폼의 기능들을 정의하고 있으며, 이러한 장점으로 스마트 오피스, 스마트 팜, 헬스케어 등 다양한 산업 분야에서 활용되어지고 있다 [3-5].

본 논문에서는 원엠투엠 표준 플랫폼을 활용하여 언제 어디서든지 태내에 있는 기기를 음성으로 제어할 수 있는 프로토타입 시스템을 구현하였다. 원엠투엠 오픈 소스 플랫폼을 활용하여 기기 간 표준 연결과 제어 흐름을 구현하고, 구글 Speech to text API를 활용하여 음성 기반 사용자 인터페이스를 구현하였다. 마이크론을 장착한 녹음 장치와 LED 제어 장치는 오픈 소스 하드웨어인 Adafruit Feather M0를 활용하였고 중간에서 텍스트 변환 결과로 이용해 실제 제어를 담당하는 게이트웨이를 라즈베리파이를 활용해 구현하였다. 본 논문에서 보인 시스템을 통해 표준 오픈 소스 플랫폼과 음성 인식 API를 활용하여 확장성과 연결성을 갖춘 커넥티드 홈을 구현할 수 있다.

• First Author: Isu Jeong, Corresponding Author: Jaeseok Yun

\*Isu Jeong (jeong.isu.j@gmail.com), Dept. of Internet of Things, Soonchunhyang University

\*Jaeseok Yun (yun@sch.ac.kr), Dept. of Internet of Things, Soonchunhyang University

• Received: 2019. 01. 25, Revised: 2019. 02. 18, Accepted: 2019. 02. 19.

• This work was supported by the Soonchunhyang University Research Fund (No. 20180077).

• This work was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (NRF-2017R1D1A3B03032676).

## II. Related Works

음성 명령 (voice command)은 두 손이 다른 작업 또는 장애로 인해 불편하거나, 운전과 같이 주의를 분산시키지 않은 채 기기에게 명령을 전달할 수 있는 최적의 사용자 인터페이스이다. 따라서 가정 환경 뿐만 아니라 많은 산업 분야에서 음성 명령을 활용한 기기 제어 방법과 애플리케이션이 개발되어 왔다.

최근 가장 활발히 연구되고 있는 제품 중 하나가 로봇 분야이다. Wang과 Yu는 음성으로 세 가지 자세 (앉기, 눕기, 서기) 제어가 가능한 휠체어를 제안하였다 [6]. Sensory의 음성 인식 마이크로컨트롤러 RSC4128을 이용해 음성 인식 하드웨어를 구성하고, 마이크로폰으로 녹음한 음성 템플릿을 이용해 사용자의 명령에 따라 휠체어의 자세를 바꾸는 실험을 통해 제안된 음성 인식 장치의 효율성을 보였다. Ruzajic는 음성 인식 모듈과 DTW/HMM 알고리즘을 이용한 음성 제어 휠체어를 제안하였다 [7]. Pleva는 미국의 SWAT (special weapons and tactics) 훈련을 위해 사용하는 Jaguar V4 로봇을 제어할 위한 음성 명령 인터페이스 방법을 구현하였다 [8]. 음성 인식을 위해 오픈 소스 음성 인식 엔진인 Julius를 이용하고, 아두이노가 내장된 로봇에게 SWAT 팀원의 음성 명령을 전달하는 방식으로 시스템을 구성하였으며 실제 SWAT 훈련에 적용될 가능성을 보였다.

휴머노이드 로봇과 산업/의료 로봇도 음성 인터페이스를 적용한 주요 영역이다. Ding은 Kinect 마이크로폰 배열을 이용해 음성 기반으로 휴머노이드 로봇 조종하기 위해 SVM, GMM, DTW를 이용해 음성과 화자를 인식하는 알고리즘 개발하여, 인증된 사용자만이 음성으로 휴머노이드 로봇을 제어할 수 있음을 보였다 [9]. Vaz 또한 휴머노이드 로봇에 음성 인식을 적용한 연구를 선보였다 [10]. 특히 Zinchenko는 수술 로봇을 음성으로 제어하는 연구를 선보였다 [11]. 산업과 의학 분야에서는 동작의 모호성 때문에 음성 명령 적용이 더딘 편이나, 음성 명령과 동작 결과를 내시경 이미지를 통해 반복 조정하는 기법을 제안하며, 음성 인식 오픈 소스인 CMU Sphinx와 HMM을 활용해 음성 인식 유닛을 제작하였고 실험을 통해 3.22 sec 이내로 음성 명령 기반으로 수술 로봇이 동작함을 보였다.

음성 명령은 스마트 홈을 위한 직관적 사용자 인터페이스로도 활용되어왔다 [12-14]. 특히 Raj는 스마트 홈을 음성으로 제어하는 가상 물리 시스템 (CPS: cyber physical system)을 제안하였다 [15]. 안드로이드 앱과 구글 Speech to Text API를 활용하여 사용자의 명령을 해석해 맥내 설치된 다양한 센서와 액추에이터를 제어할 수 있는 스마트 홈을 구현하였다.

국내에서도 음성 인식 결과를 활용하는 연구가 활발히 이루어지고 있는데 대표적으로, Lee는 실제 연기자들이 감정을 이입하여 녹음한 대사 샘플들을 통계 프로그램 SPSS를 사용하여 감정에 따른 음질, 운율의 변화를 분석한 결과, 감정 별로 음의 세기, 대역 너비, 높낮이 정보 등이 영향을 받고 이 결과를 감정 음성 인식에 활용할 수 있음을 보였다 [16]. Kang은 음성 정보

를 처리하여 사용자의 스트레스 정도를 측정하기 위한 웨어러블 기기를 제작하고 스트레스 수준에 따라 피드백 경고 또는 3자에게 메시지 전송 등을 통해 사용자의 스트레스를 관리하기 위한 모바일 애플리케이션을 제안하였다 [17].

본 논문에서는 위에서 기술된 관련 연구 중에서 특히 스마트 홈에서 활용되는 음성 명령 인터페이스에 집중하고자 한다. 스마트폰에서 기본으로 제공하는 ‘Siri’나 ‘Google Now’와 같은 음성 비서 (voice assistant)의 대중화와 ‘Alexa’를 탑재한 아마존 에코와 같은 지능형 스피커 보급으로 인해 현재 음성 명령 기반 사용자 인터페이스는 커넥티드 홈과 헬스케어 분야에서 거의 필수적으로 사용되고 있다.

그러나 위에서 기술한 모든 제품들은 비표준 또는 독자적인 인터페이스를 사용함으로써 사물인터넷 시장 파편화 문제를 해결하지 못하고 있으며 오히려 고유의 독자 생태계를 구성하여 1인 개발자나 중소기업들이 해당 생태계에 참여하는데 어려움이 있거나 플랫폼 소유 회사에 대한 종속성과 의존도가 높다. 따라서 본 논문에서는 표준 플랫폼을 활용하여 기기간 확장성과 연동성이 보장된 음성 기반 인터페이스를 제안하고자 한다.

## III. Education Experience

### 1. Overall system

본 논문에서는 원엠투엠 IoT 플랫폼을 활용하여 IoT 기기들에게 표준 연결성을 제공해 공간의 제약 없이 맥내의 기기들을 음성으로 제어할 수 있는 시스템 구현을 목표로 한다. 이를 위해 Fig. 1에서 보이는 바와 같이 시스템 구조를 설계하였다.

맥내 가전기기들은 두 종류로 나눌 수 있는데 계산 리소스가 충분하고 통신 능력을 갖춰 원엠투엠 서버와 직접 통신이 가능한 제품 (예, TV, 냉장고, 세탁기 등)들과 원엠투엠 게이트웨이를 통해서 서버와 통신이 가능한 제품 (예, 온도 센서, 시계 등)으로 분류가 가능하다. 본 논문에서는 두 가지 경우를 모두 보이기 위해 음성 인식 장치는 게이트웨이를 통한 통신, LED를 제어 장치는 서버로의 직접 통신으로 구현하였다.

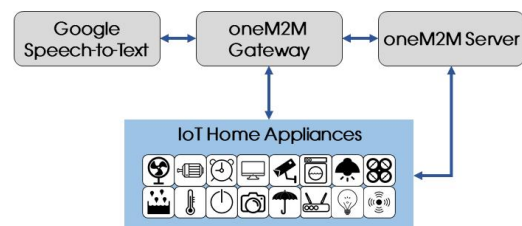


Fig. 1. Overall System Design using oneM2M Platforms

Fig 2에서는 위에서 제안한 설계대로 본 논문에서 구현하고자 하는 프로토타입 시스템의 구조를 보이고 있다. 원엠투엠 서버를 구현하기 위해 잘 알려진 원엠투엠 오픈 소스 플랫폼인

모비우스를 활용하였고, 원애플리케이션 게이트웨이는 오픈 소스 하드웨어인 라즈베리파이, 태내 가진 기기를 모사하기 위한 음성 녹음 장치와 LED 제어 장치는 오픈 소스 하드웨어인 Adafruit Feather M0를 활용하였다. 위에서 설명했듯이 태내 기기가 서버와 통신하는 두 경우를 모두 보이기 위해 음성 녹음 장치는 게이트웨이를 통해서, LED 제어 장치는 직접 서버와 원애플리케이션 API (application programming interface)를 통해 연결하였다.

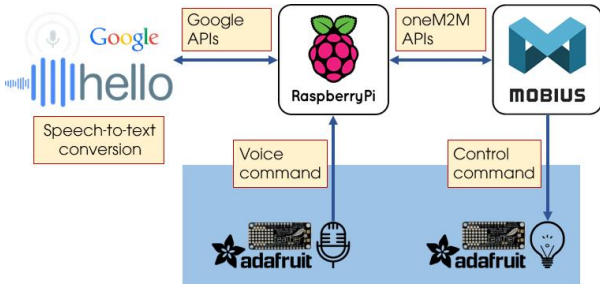


Fig. 2. Prototype System Architecture

구현된 시스템의 전체적인 동작 시나리오와 흐름은 다음과 같다. 사용자가 음성 녹음 장치에서 버튼을 누르며 음성 명령을 녹음하면 생성된 녹음 파일을 라즈베리파이에게 전달된다. 이 녹음 파일은 Google STT (Speech-to-Text) API를 이용하여 Google Cloud로 보내진 후 STT에 의해 텍스트로 변환된 파일을 받는다. 게이트웨이는 수신한 텍스트를 분석하여 해당 디바이스 (즉 여기서는 LED 장치) 제어 명령을 생성하고, 이를 모비우스 서버에 전송한다. 최종적으로 모비우스 서버는 원애플리케이션 API로 라즈베리파이로부터 수신한 데이터를 바탕으로 등록된 LED 제어 장치를 제어한다. 3-1, 2장에서 각각 음성 녹음 장치와 LED 제어 장치 구현 내용을 자세히 다룬다.

2. Voice command to text conversion

음성 녹음 기기는 사용자가 버튼을 눌렀을 때 마이크로폰으로 입력되는 음성을 녹음하고, 녹음된 파일을 Google STT 변환 과정을 통해 텍스트로 변환한다. 라즈베리파이는 변환된 텍스트를 수신하고 이를 기반으로 LED 장치 제어 명령을 생성하여 모비우스 서버로 전송함으로써 등록된 다른 기기들이 명령을 참조할 수 있도록 한다. Fig. 3는 전체 동작 흐름을 표현하고 있으며 각 절차에 대한 자세한 설명은 다음과 같다.

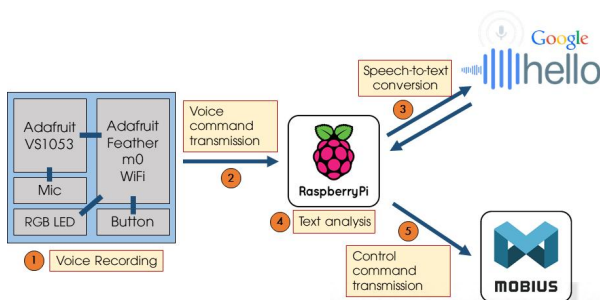


Fig. 3. Voice Command to Text Conversion Procedure

- 절차 1. 음성 녹음
  - 콘텐서 마이크를 이용해 OGG 형식의 음성 파일 생성
- 절차 2. 음성 녹음 장치와 라즈베리파이간 파일 전송
  - 음성 파일은 STT를 위해 Google API를 실행할 수 있는 라즈베리파이로 전송됨
- 절차 3. STT (Speech to Text) 변환
  - 구글 Speech API가 지원하는 형식을 맞추기 위해 FFmpeg를 통해 파일 변환
  - 변경된 음성 파일을 구글 Speech API에 전송 후 STT 진행
- 절차 4: 텍스트 분석
  - 구글 Speech API를 통해 제공된 변환 텍스트를 분석하여 디바이스 제어 명령 생성
- 절차 5: 디바이스 제어 명령 전송
  - 라즈베리파이에서 실행되는 oneM2M 디바이스 플랫폼은 모비우스 서버에 디바이스 제어 명령을 전송
  - 디바이스와 대응하는 컨테이너에 제어 명령과 대응하는 값을 전송

2.1. Voice recording device

음성 명령을 녹음하기 위한 장치는 Adafruit Feather M0 WiFi 보드, Adafruit VS1053 모듈, Li-ion 배터리, 마이크, RGB LED, 푸시 버튼으로 이루어져 있으며, 각 장치들의 기능은 Table. 1에 설명되어 있다.

Table 1. Voice Recording Device Components

Component	Functions
Adafruit Feather M0 WiFi	Microcontroller board
Adafruit VS1053	Audio recording in OGG files
Rechargeable Battery	Power supply
Condenser Microphone	Sound to signal conversion
RGB LED	Recording status indicator
Push button	Voice recording trigger

Adafruit Feather M0는 음성 녹음 장치의 중앙 처리 장치를 담당하며 사용자의 푸시 버튼 입력에 맞추어 마이크로폰에 입력되는 사용자의 음성 명령을 VS1053 코덱 모듈을 이용해 녹음하여 SD 메모리 카드에 저장한다. 사용자는 RGB LED 색상을 통해 녹음 장치의 현재 상태 (대기, 녹음, 완료, 전송 등)를 파악할 수 있다. VS1053 모듈의 SD 메모리 카드에 녹음 파일이 저장되면 Adafruit Feather M0는 녹음 파일을 라즈베리파이에게 전송한다. Fig. 4는 실제 구현한 음성 녹음 장치의 사진과 구성 요소들을 보여준다.

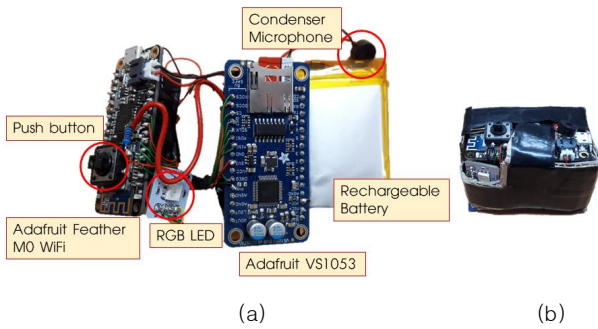


Fig. 4. Implementation of Voice Recording Device (a) All Components (b) Packed in a Container (6\*4\*3.5 cm)

**2.2. Voice recording file transmission**

녹음된 음성 명령을 텍스트로 변환하기 위해 직접 Google Cloud에 전송하여 Speech API를 이용하기 보다는 사물인터넷 플랫폼 간 표준화된 인터페이스를 구현하기 위해서 음성 녹음 장치와 라즈베리파이 간에 파일 전송이 요구된다. 이 파일 전송은 WiFi 무선 환경에서 일반적인 TCP/IP 소켓 통신 방식으로 구현하였다.

**2.3. Speech to text conversion**

본 논문에서 사용한 Adafruit VS1053 모듈은 마이크로프로세서로 녹음된 음성을 OGG 형식 파일로 저장한다. 그러나 Google Cloud가 지원하는 Speech API는 FLAC 형식의 파일을 요구하기 때문에 녹음 파일의 포맷 변환 (즉 OGG -> FLAC)이 필요하다. 이를 위해 라즈베리파이는 이전과 다른 OGG 형식의 파일이 특정 컨테이너에 있을 때 (즉 새로운 녹음 파일이 생성되었을 때), 해당 파일을 FFmpeg를 이용해 포맷을 FLAC으로 변경한 뒤 리샘플링과 리채널링을 진행한다. FLAC 포맷으로 변환된 음성 파일을 Google Speech API에 전송 후 transcript (내용)과 confidence (확신)값을 받아온다.

**2.4. Text analysis**

Google Cloud의 STT를 진행하여 받아온 텍스트는 실질적으로 제어 대상인 LED 장치를 사용자가 원하는 상태로 바꾸기위한 명령을 의미한다. 본 논문에서 사용한 텍스트 명령은 프로토타입 시스템을 위해 단순히 “불 켜”와 “불 꺼”를 사용하였으며 텍스트 명령에서 ‘켜’와 ‘꺼’를 추출하여 LED의 상태를 나타내는 변수로 지정하였다. 따라서 ‘켜’가 추출되었을 때는 ‘불을 켜라’의 의미에서 정수 ‘1’을, ‘꺼’가 추출되었을 때는 ‘불을 꺼라’의 의미에서 정수 ‘0’을 변환된 명령 텍스트 파일에 저장하였다.

**2.5. Device control message transmission**

라즈베리파이는 원앰투엠 서버 플랫폼인 모비우스와 표준화된 인터페이스로 통신을 하기 위해서 게이트웨이 역할을 할 수 있는 디바이스 플랫폼이 반드시 필요하다.

본 논문에서는 사물인터넷 표준 오픈 소스 협의체인 OCEAN (Open alliance for IoT standard)에서 제공하는 원앰투엠 디바이스 플랫폼인 앤큐브 (&Cube)를 활용하였다

[18]. 앤큐브는 원앰투엠 서버 플랫폼인 모비우스에 기기 등록, 데이터 보관/관리, 구독/알림 등 기본적으로 필요한 공통 기능 등을 표준 API로 제공하고 있다. 따라서 관련 API 문서를 참조하여 데이터 구조 (즉 원앰투엠 타입 XML 또는 JSON)를 해석하거나 구성할 수 있다면 개발자가 원앰투엠 표준을 전혀 모르더라도 손쉽게 표준을 준수하는 사물인터넷 시스템을 구성할 수 있도록 해주는 장점이 있다.

라즈베리파이에서 구동되는 앤큐브는 STT를 통해 얻은 텍스트에서 정수로 정리된 텍스트 명령 파일에 접근한다. 텍스트 파일에서 읽어낸 문자열 ('0' 또는 '1')에 따라 모비우스 서버에 전송할 데이터와 데이터가 저장될 컨테이너를 지정한다.

본 논문에서는 읽은 문자열이 '1'일 경우에는 모비우스 서버의 light 컨테이너에 '1'을 '0'일 경우에는 '0'을 업로드 하였다. 이 업로드 동작은 데이터가 저장된 컨테이너의 이름으로 모비우스 서버에 등록된 다른 기기 (즉 여기서는 LED 장치)를 제어할 명령이 된다. 라즈베리파이에 설치된 앤큐브가 전송한 명령 데이터는 Fig. 5.과 같이 전자부품연구원 (KETI)에서 제공하는 원앰투엠 브라우저 (리소스 모니터)를 통해 실시간으로 확인이 가능하다.

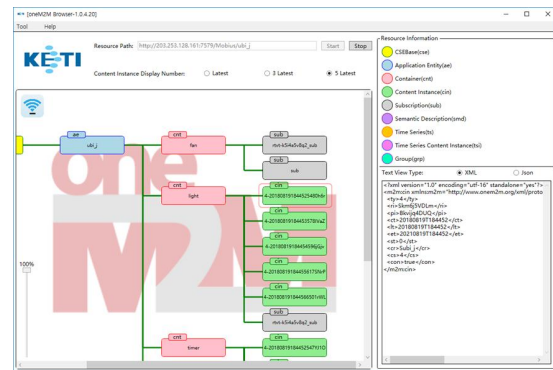


Fig. 5. Containers and Uploaded Data in Resource Monitor

**3. LED Device Control from Mobius**

모비우스는 OCEAN에서 공개하고 있는 원앰투엠 서버를 위한 오픈 소스 플랫폼이다. 공개하고 있는 오픈 소스를 활용하여 실제 원앰투엠 서버를 구축할 수도 있지만 본 논문에서는 표준 인터페이스를 통한 음성 명령 제어 개념만을 보이기 위해 전자부품연구원에서 일반인들에게 실험 및 연구 목적으로 제공하고 있는 공개 서버를 활용하였다.

모비우스는 원앰투엠 게이트웨이를 통한 사물의 간접 연결과 직접 연결 모두를 지원하는데 음성 녹음 장치는 전자를 이용한 거라면, LED 제어 장치는 후자인 모비우스와의 직접 연결을 이용하였다. 따라서 이를 위해 LED 제어 장치는 모비우스에 자신의 이름을 따서 만들어진 light 컨테이너를 생성하며 여기에 업로드된 데이터에 따라 LED 제어 장치의 액추에이터들을 제어할 수 있다.

본 논문에서는 Adafruit Feather M0 WiFi 보드를 사용하여 LED 제어 장치를 구성하였다. 전체적인 장치 시스템 구성도는

Fig. 6에 나와 있으며 각 절차와 이에 따른 동작 내용은 다음과 같다.

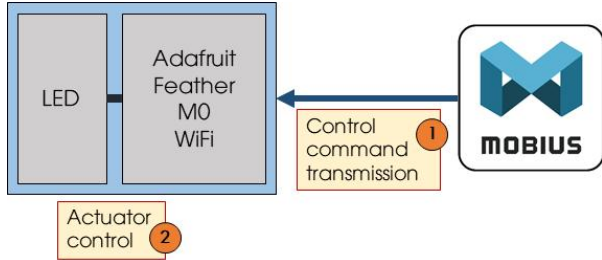


Fig. 6. LED Control Device Architecture

- 절차 1. 디바이스 제어 명령 수신
  - Adafruit Feather MO 보드는 원웹투엠편 디바이스 플랫폼 애플리케이션을 실행 중
  - 모비우스 서버의 특정 컨테이너에 값이 업데이트 되면, 해당 컨테이너를 구독하고 있는 애플리케이션에게 업데이트된 데이터를 전송
- 절차 2. 액추에이터 제어
  - Adafruit Feather MO 보드는 애플리케이션을 통해 전송받은 데이터의 컨테이너 (원웹투엠편 표준, con) 값과, 콘텐츠 인스턴스 (원웹투엠편 표준, cin) 값에 따라 LED를 제어

3.1. LED control device

LED를 제어하는 장치는 Adafruit Feather MO WiFi 보드, LED, Li-ion 배터리로 구성되어 있다. Adafruit Feather MO WiFi 보드는 WiFi 모듈을 통하여 인터넷에 접속한다. 인터넷과 연결된 Adafruit Feather MO WiFi 보드는 애플리케이션을 통해 모비우스 서버로부터 데이터를 수신하고 이를 바탕으로 LED를 제어한다. Fig. 7은 실제 구현한 LED 제어 장치의 사진과 구성요소들을 보여준다.

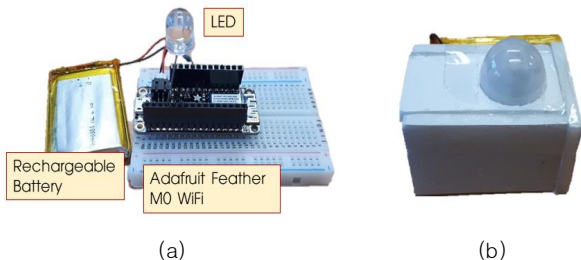


Fig. 7. Implementation of LED Control Device (a) All Components (b) Packed in a Container

3.2. Device control message reception

Adafruit Feather MO 보드 위에서 구동되는 원웹투엠편 디바이스 플랫폼인 애플리케이션은 모비우스 서버의 특정 컨테이너 (여기서는 LED 제어 장치를 위한 컨테이너)를 구독하고 있다. 원웹투엠편 표준에서는 구독/알림 서비스를 다음과 같이 정의하고 있다. 원웹투엠편 리소스는 다른 엔티티 (entity)에서 구독이 가능

하며 구독을 하게 되면 해당 리소스 하위에 'subscription' 리소스가 생성된다. 이후 리소스의 내용이 업데이트되면 구독중인 엔티티들에게 알림 (notification) 서비스를 제공한다.

본 논문에서 이 구독/알림 서비스를 이용해 구현한 내용은 LED 제어 장치의 컨테이너에 신규 데이터가 업데이트될 경우, 모비우스 서버는 해당 컨테이너를 구독 중인 애플리케이션 (즉 LED 제어 장치에 해당하는 Adafruit Feather MO)에게 신규 데이터를 전송한다.

3.3. Actuator control

모비우스에 전송하는 데이터는 Fig. 8에서 보이듯이 XML 형식을 지닌다. 애플리케이션은 전송된 데이터가 저장되어 있던 컨테이너 이름과 콘텐츠 인스턴스 값에 따라 액추에이터를 제어한다. 본 논문에서는 light 컨테이너에 '1' 값이 올라오면 LED를 켜고 '0'이 올라오면 LED를 끄도록 구현하였다.

```
<?xml version="1.0" encoding="utf-16" standalone="yes"?>
<m2m:cin xmlns:m2m="http://www.onem2m.org/xml/protocol"
<ty>4</ty>
<ri>ITaCLn2tKw</ri>
<pi>G4gosu0db3</pi>
<ct>20181222T064303</ct>
<lt>20181222T064303</lt>
<et>20211222T064303</et>
<st>16</st>
<cr>Crtvt-k5i4a5v8q2</cr>
<cs>1</cs>
<con>1</con>
<mni />
</m2m:cin>
```

Fig. 8. Data Format Uploaded to Mobius server

IV. Experimental Results

3장에서 개발한 음성 녹음 장치와 LED 제어 장치를 이용해 음성으로 맥내 가전기기를 제어하는 실험을 진행하였다.

음성 녹음 장치를 손목형 스마트 밴드라 가정하고 Fig. 9-(a)에서 보이듯이 왼쪽 손목에 벨크로 스트랩을 이용하여 고정하였다. 음성 명령을 녹음하기 위해 오른손으로 푸시 버튼을 누르면 LED에 불이 들어오며 사용자에게 녹음 중임을 알려주고 사용자가 녹음을 마치면서 버튼을 다시 누르면 LED 불이 바뀌며 녹음 파일이 라즈베리파이 게이트웨이로 전송되었음을 알린다. 음성 명령이 Google Cloud의 STT를 통해 변환된 결과가 모비우스에 전달되면 Fig.9-(b)에서 보이듯이 무선으로 연결된 LED 제어 장치가 음성 명령에 따라 동작하게 된다.

본 논문에서 개발한 프로토타입 시스템을 위해 사용한 음성 명령은 간단히 “불 켜”와 “불 꺼”이며 ‘켜’와 ‘꺼’를 구분하여 LED를 제어하였다. Fig. 9-(a)에서 보이는 프로토타입 음성 녹음 장치는 비록 손목에 실제로 차고 다니기에는 크기와 부피가 적절하지 않으나 본 논문에서 사용한 원웹투엠편 디바이스 플랫폼 설치만 가능하다면 스마트폰이나 다른 형태의 스마트 밴드에도 적용이 가능하다.

또한 Fig. 9-(b)에서 보이는 프로토타입 LED 제어 장치는 무선으로 댁내 WiFi 공유기에 연결되는 가전기기로 확장될 수 있다. 예를 들어 텔레비전이나 선풍기와 같은 가전기기에 원웹투옴 플랫폼이 탑재되면 LED 제어 장치와 동일한 절차와 방식 대로 제어가 가능하다. 따라서 본 논문에서 제시한 원웹투옴 플랫폼을 활용한 음성 기반 제어 시스템은 컨테이너와 그 하위의 리소스로 기기를 제어할 수 있으므로, 댁내 기기에 원웹투옴 플랫폼을 설치하고 기기의 동작과 기능에 해당하는 음성 명령어만 변경한다면 다양한 기기들을 연결하고 제어할 수 있다.

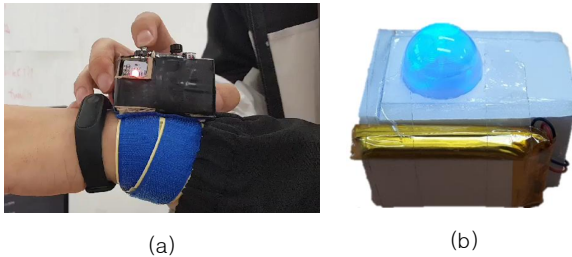


Fig. 9. Demonstration of Voice-based LED Device Control (a) voice recording using Voice Recording Device (b) LED control by received command

## V. Conclusions

본 논문에서는 음성 명령으로 언제 어디서나 댁내 가전기기를 제어할 수 있도록 하기 위한 표준 사물인터넷 (IoT) 플랫폼 활용 방법을 제안하였다. 음성 명령은 스마트폰의 보급과 인공지능 기술의 발달로 댁내 가전기기 뿐만 아니라 산업, 의료 분야에서도 최적의 사용자 인터페이스로 주목을 받고 있다.

그러나 사물인터넷 기기와 애플리케이션의 다양성으로 인해 기기 간 상호연동성 (interoperability)이 떨어지게 되어 음성 명령을 적용할 수 있는 대상이 음성 명령 장치 그룹 (예로서 아마존 알렉사 그룹, 애플 시리 그룹, 구글 나우 그룹 등)으로 제한된다. 이러한 시장 과편화는 사물인터넷 생태계 성장을 가로막고 있으며, 이 문제를 해결하기 위한 방법 중 하나는 표준 사물인터넷 플랫폼과 API의 활용이다.

본 논문에서는 국제 사물인터넷 표준인 원웹투옴을 기반으로 개발된 오픈 소스인 모비우스와 앤큐브를 활용하여 음성 명령 장치와 댁내 가전기기를 모사한 LED 제어 장치를 구현하였다. 구현된 프로토타입 장치들을 이용한 사용 시나리오 실험을 통해 표준 플랫폼 탑재와 API 연동만으로도 손쉽게 음성 명령 인터페이스들을 댁내 가전기기에 적용할 수 있음을 보였다. 무엇보다 표준 플랫폼 사용으로 인해 기기간 연결성과 확장성이 기존 상업 플랫폼들에 비해 매우 높음을 확인할 수 있고, 나아가 제품 제조사나 소프트웨어 개발자의 개발 방법과 관계없이 모든 사물인터넷 제품들에게 음성 명령 인터페이스를 적용할 수 있음을 기대할 수 있다.

비록 사물인터넷 시장 과편화 문제를 해결하기 위해 원웹투옴 등 산업표준들을 개발하고 사용을 장려하고 있으나 여전히 애플, 구글, 삼성전자 등으로 대표되는 초대형 기업들은 앱 스토어 등 고유 플랫폼을 바탕으로 사물인터넷 시장에서 강력한 힘을 발휘하고 있다. 더구나 이들에게 사물인터넷 시스템 간 상호연동성을 높이기 위해 산업표준 사용을 강제하기는 실질적으로 어렵다. 따라서 대학교 창업 동아리, 개발자 커뮤니티, 오픈 소스 기구 등이 주도적으로 협력하여 중소기업 중심으로 원웹투옴 표준을 활용한 사물인터넷 제품개발을 독려할 필요가 있다. 또한 정부 부처와 국가 연구소를 중심으로 단계별 국가 계획을 수립하여 원웹투옴 표준으로 인해 모든 시스템들이 상호연동하는 스마트 시티와 같은 공공 서비스 인프라를 구축해 나가는 계획 등이 필요하다.

## REFERENCES

- [1] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2787-2805, October 2010.
- [2] J. Swetina, G. Lu, P. Jacobs, F. Ennesser, and J. Song, "Toward a standardized common M2M service layer platform: Introduction to oneM2M," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 20--26, June 2014.
- [3] M. Ryu, J. Kim, and J. Yun, "Integrated semantics service platform for the Internet of Things: A case study of a smart office," *Sensors Vol. 15, No. 1*, pp. 2137-2160, 2015.
- [4] M. Ryu, J. Yun, T. Miao, I.-Y. Ahn, S.-C. Choi, and J. Kim, "Design and Implementation of a Connected Farm for Smart Farming System," in *Proceedings of the IEEE Sensors*, pp. 1724-1728, November 1-4 2015.
- [5] S. M. M. Fattah, N.-M. Sung, I.-Y. Ahn, M. Ryu, and J. Yun, "Building IoT Services for Aging in Place Using Standard-Based IoT Platforms and Heterogeneous IoT Products," *Sensors*, vol. 17, no. 10, pp. 2311, October 2017.
- [6] D. Wang and H. Yu, "Development of the Control System of a Voice-Operated Wheelchair with Multi-posture Characteristics," *Proceedings of the 2nd Asia-Pacific Conference on Intelligent Robot Systems (ACIRS '17)*, pp. 151--155, Wuhan, China, June 16-18 2017.
- [7] M. F. Ruzaj, S. Neubert, N. Stoll, and K. Thurow, "Design and Implementation of Low-Cost Intelligent Wheelchair Controller for Quadriplegias and Paralysis Patient," *Proceedings of the IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMII '17)*, pp. 399--404, Herl'any, Slovakia, January 26-28 2017.

- [8] M. Pleva, J. Juhar, A. Cizmar, C. Hudson, D. W. Carruth, and C. L. Bethel, "Implementing English Speech Interface to Jaguar Robot for SWAT Training," Proceedings of the IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI '17), pp. 105--110, Herl'any, Slovakia, January 26-28 2017.
- [9] I.-J. Ding and J.-Y. Shi, "Kinect microphone array-based speech and speaker recognition for the exhibition control of humanoid robots," Computers & Electrical Engineering, vol. 62, no. , pp. 719--729, August 2017.
- [10] J. C. Vaz and P. Oh, "Transformer-DARwIn: A Miniature Humanoid for Potential Robot Companion," Proceedings of the IEEE International Conference on Consumer Electronics (ICCE '18), pp. 1--5, Las Vegas, NV, USA, January 12-14 2018.
- [11] K. Zinchenko, C.-Y. Wu, and K.-T. Song, "A Study on Speech Recognition Control for a Surgical Robot," IEEE Transactions on Industrial Informatics, vol. 13, no. 2, pp. 607--615, April 2017.
- [12] H. Basanta, Y.-P. Huang, and T.-T. Lee, "Using Voice and Gesture to Control Living Space for the Elderly People," Proceedings of the International Conference on System Science and Engineering (ICSSE '17), pp. 20--23, Ho Chi Minh City, Vietnam, September 11 2017..
- [13] R. K. Nath, R. Bajpai, and H. Thapliyal, "IoT Based Indoor Location Detection System for Smart Home Environment," Proceedings of the IEEE, pp. 1--3, Las Vegas, NV, USA, January 12-14 2018.
- [14] D. Sunitha, V. C. Patil, M. Hn, and S. Jebakani, "Digital notice board using Smart Phones- Speech Recognition Voice command," Proceedings of the International Conference on Current Trends towards Converging Technologies, pp. 1--4, Coimbatore, India, March 1-3 2018.
- [15] R. Raj and N. Rai, "Voice Controlled Cyber-Physical System for Smart Home," Proceedings of the Workshop Program of the 19th International Conference on Distributed Computing and Networking (ICDCN '18), pp. Article No. 28 , Varanasi, India, January 4-7 2018.
- [16] Sang-Min Lee and Ho-Joon Lee, "How to Express Emotion: Role of Prosody and Voice Quality Parameters," Journal of The Korea Society of Computer and Information, vol. 19, no. 11, pp. 159--166, November 2014.
- [17] Byeongsoo Kang and Ky Vannroath and Hyun-syug Kang, "A Mobile Stress Management System utilizing Variable Voice Information According to the Wearing Area," Journal of The Korea Society of Computer and Information, vol. 22, no. 6, pp. 95--100, June 2017
- [18] OCEAN, <http://developers.iotocean.org/>

## Authors



Isu Jeong is an undergraduate student with the Department of Internet of Things at Soonchunhyang University. Her research interests include Internet of Things, AI algorithms, and computer networks.



Jaeseok Yun is an assistant professor with the Department of Internet of Things at Soonchunhyang University. Prior to his current position, he worked as a senior researcher with the IoT Platform Research Center at Korea Electronics Technology

Institute (KETI) from 2009 to 2016. He also worked as a postdoctoral research scientist with the Ubiquitous Computing Research Group in the School of Interactive Computing at Georgia Institute of Technology, GA, USA from 2006 to 2009. He received his B.S. degree in Electronics Engineering from Chonnam National University in 1997. He also earned his M.S. and Ph.D. degrees in Mechatronics from Gwangju Institute of Science and Technology (GIST) in 1999 and 2006, respectively. His research interests include ubiquitous computing, Internet of Things, machine learning, and intelligent systems.