



## Secured Speech Industrial Automation Based on Raspberry Pi and IoT

**Bhagirathi Bai Vittal Rao** <sup>1\*</sup>

**Hanume Gowda Chandrakanth** <sup>2</sup>

<sup>1</sup>*Department of Mechatronics Engineering, Acharya Institute of Technology, India*

<sup>2</sup>*Department of Electronics and Communication, Sambhram Institute of Technology, India*

\* Corresponding author's Email: bhagirathistelkar@gmail.com

---

**Abstract:** Today, Internet is the best place to communicate and share information among the people throughout the world and gives an endless support of knowledge and entertainment. The main objective of Internet technology is to increase efficiency and decrease human effort. With the introduction of Internet of Things (IoT) in the last decade, we have been pushing for ubiquitous computing in all spheres of life. Physically challenged people are also using the Internet with the help of Speech commands (SC). The main objective of this paper is to minimize the effort and increase efficiency of the Voice recognition and IoT based secured automation, which is named as Raspberry pi and IoT based Speech automation (RASP-IoT-SA) technique. This RASP-IoT-SA system has only worked for authenticated users which improves the security in home and industrial automations. This RASP-IoT-SA system consists of four different Process such as secured speaker prediction, voice data transmission, web update and voice data Receiver processing. In this research work, Mel frequency Cepstral Coefficient (MFCC) technique is used for Feature Extraction (FE). Artificial neural network (ANN) technique is used for two different process such as, speaker prediction and speech data recognition (SR). The experimental result shows that the MFCC and ANN based technique provides better results in terms of accuracy, precision, false measure and recall. The RASP-IoT-SA technique delivers 99% of speech recognition accuracy, which was higher than the existing Speech automation technique.

**Keywords:** Artificial neural network, Internet of things, Mel frequency cepstral coefficient, Speaker prediction, Raspberry pi and secured speech automation.

---

### 1. Introduction

Speech recognition is the inter-disciplinary sub-field of computational linguistics that develops technologies and methodologies for the recognition and translation of spoken words into text by computers. The Renowned market research and market intelligence reported that the 4 million home automation products sold in 2013. The same organizations also reported that 90 million Home automation products sold till the end of 2017 [1]. Automation is defined as an automatic working of system without the help of human interaction. The electrical switch Circuit is positioned in various rooms of home, which makes it very difficult for the members in home, especially to physically challenged people, so that some automated system introduced in the market. [2-3]. The existing

automation technologies consists of Blue-tooth and Zig-bee [4-5]. The major limitation of the Blue-tooth and Zig-bee based methodology is, only the user can access the data up to a particular distance. To overcome this problem, it requires three major things such as internal network, intelligent control and automation with wired or Wireless Communications (WC) through the internet [6].

IoT applications generally connects to the sensors like Humidity, fire temperature for monitoring the system. So that the automation technique is suitable to certain situations and in-turn it makes comfortable for the user. Internet based automation technology's also utilized the machine-learning techniques to discover the daily activities of the user, and generate an automation rules and actions that mimic these actions [7]. The SR automation basically depends upon the Voice recognition device and the controller. There are different varieties of automation controller

are available in the market for example ATmega8, ATmega16, Micro controller 89c51 and etc. [8-9]. Likewise, there are different variety of Voice recognition device used in the existing methods such as HM2007, SPCE061A and etc. The separate automation controller and Voice recognition device make the system too larger [10-15].

In this paper, to overcome the existing drawback and to improve the speaker recognition and Speech Recognition Accuracy, RASP-IoT-SA Technique is introduced. The proposed methodology: RASP-IoT-SA consist of two different recognition systems such as speaker recognition and speech recognition. Speaker recognition system is used for identifying the person and speech recognition is used for identifying speech data. The RASP-IoT-SA normalization based MFCC is used for Feature Extraction (FE) and artificial neural network (ANN) technique is used for the classification purpose. For testing the speaker recognition and speech recognition of RASP-IoT-SA a separate database is stimulated. The RASP-IoT-SA system is also compared with the existing system.

The paper is organized as follows. Section 2 provides a brief description of the related works. Section 3 focus on RASP-IoT-SA method. Section 4 discussed about Hardware and Software. In section 5, comparative study of the results and discussion is presented. Section 6 gives a summary of this paper.

## 2. Literature review

H. Bharathi, U.Srivani, M.D.Azharudhin, M. Srikanth and M. Sukumarline [16] designed a remote household appliance control methodology using RASP and android mobile phone. By using this methodology, it is possible to control any electrical and electronic device in home and office whether the person is available or not. There was no monitoring methodology other than controlling options which was the major disadvantage of this system.

T. Obaid, H. Rashed, A.A.E. Nour, M. Rehan, M.M. Saleh, and M. Tarique [17] has introduced the Zig-bee and voice controlled based wireless smart home methodology. In this methodology voice commands were used to convert the voice data into data format and it was sent to the receiver (via the WC). The wireless receiver performed desired switching operations based on the received data. The proposed method speech recognition accuracy is average.

G. Muthuselvi, and B. Saravanan [18] has introduced SR Based Real Time Automation methodology. This methodology is basically in standby mode, once an input is detected, it is analyzed by the SR module. If a known automation

command is sensed, at the time SR methodology sends detected data to the micro-controller (MC). The MC process the data signals (compares them with a database) and identifies the command state. According to the load state determined, control signals are sent to the respective relay circuits for actuating the appropriate loads. The proposed system doesn't satisfy for long distance communication and also the proposed automation is a time consuming process.

M.R. Kamarudin, and M.A.F.M. Yusof [19] designed and developed the less cost SR based smart automation. The Visual Basic software based Graphical User Interface (GUI) used the Microsoft SR engine as an input source for controlling and monitoring home appliances and controlled through wireless. These methodologies have an application in the field of RF communication, MC and computer programming but, not suitable for long distance communication and also the speech recognition accuracy is average.

F. Baig, S. Beg, and M.F. Khan [20] has introduced Zig-bee and speech command based home appliances controlling using hand held devices. The mobile application converts the voice command into text data and initialize the command on GSM network via SMS. The GSM based technology is too costly and it has some limitations.

S.X. Ding, P. Zhang, S. Yin, and E.L. Ding [21] introduced the integrated voice based control of smart systems in the industry using micro controller with the visual basic interface. RFID based multilevel user access to provide high security authentication. Staff attendance and data logger were the main features of the project (VC-SS-IA). The common functionalities in the industry are autonomously controlled based on voice commands providing a luxury of e-lifestyle for the operator, production, maintenance engineer and workers of several other departments using low power consumption. This method used Zig-bee for communication which is not sufficient for long distance communication and the accuracy of the system is need to be improve.

H.J.R.S.R Sushma [22] has designed and developed a computer based interactive system with the help of SR in order to monitor and control the power stations. SR is the translation of spoken words into text data. Speaker authentication is the process, which automatically identify the speaking person based on the individual information included in speech waves. MFCC is used for speech FE. The authors don't mention about the accuracy of the system and the classification problem is also not yet resolved.

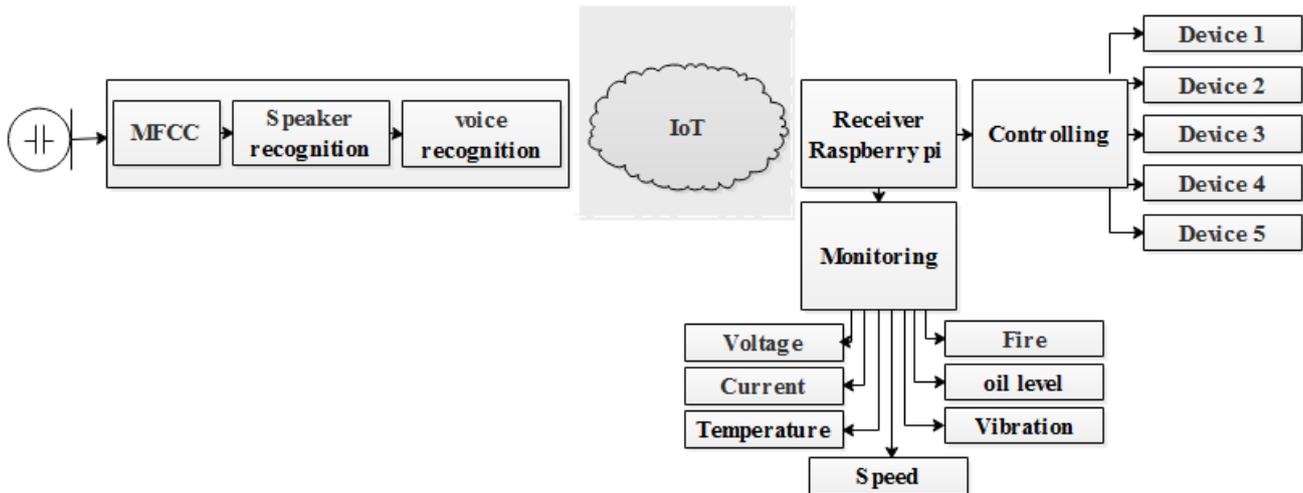


Figure.1 RASP-IoT-SA system block diagram

S. Goyal, and N. Batra [23] has done a comparative analysis of voice recognition feature extraction and classification. The author has considered three different kind of feature extraction (such us, Perceptual Linear Prediction (PLP), Linear Predictive Code (LPC) and MFCC) and four different kind of classifiers (such us HMM (Hidden Markov Model), DTW (Dynamic Time Wrapping), VQ (Vector Quantization), SVM (Support Vector Machine)). Finally, the author has concluded that the combination of MFCC and PLP feature extraction and VQ based classification provided 98% accuracy. But the major limitation of the MFCC-PLP-VQ is database which occupies more database storage.

### 3. RASP-IoT-SA system

The speech based automation system are very helpful for physically challenged people, but most of the home and industrial automations systems [1-14] doesn't provide better accuracy and which doesn't have a security system for automation. Hence to moreover the existing limitation and improve the automation security a RASP-IoT-SA system has introduced. The block diagram of the RASP-IoT-SA system is shown in the Fig. 1. RASP-IoT-SA system has three major sections such as 1). Voice control based RASP (Transmitting), 2). RASP Monitoring and controlling (Receiving) and 3). Internet based web-page updating. In the Transmitting section voice signal is given as an input signal which is recognized by the Microphone. The features are extracted from the voice data with the Help of MFCC, speaker of the voice is recognized. If the speaker authorizes the voice signal, then it converts into data signal. That

data is updated in the web-page over the Internet. After the data updation in the web page, the Monitoring and controlling section is under work where the monitoring section measures the Voltage, current, Temperature, speed, Vibration, and oil level. Controlling section controls the device which is connected to the different electrical devices like Transformer, light, Motor, Fluid Pump and etc.

#### 3.1 Voice control based RASP

In this block microphone detects voice signal, the detected voice signal is given to the RASP. The Python programming extracts the MFCC Speech features which is recognized by using an ANN. The recognized Data are updated on the web-page through the Internet.

##### 3.1.1. MFCC

MFCC is the best signal FE technique. Frequency bands are placed logarithmically to monitor the human system response more closely than any other system. From the spoken word, only sixteen coefficients of MFCC are corresponding to the Mel scale frequencies of speech Cepstrum.

The Fig. 2 shows the first step of pre-processing in which the signals are pre-processed before MFCC FE. In time domain, signals are divided into amount of frames by applying hamming window technique. That frame is converted into frequency domain with the help of the Discrete Fourier Transform (DFT). DCT is used to convert mel spectrum coefficient into time domain [15].

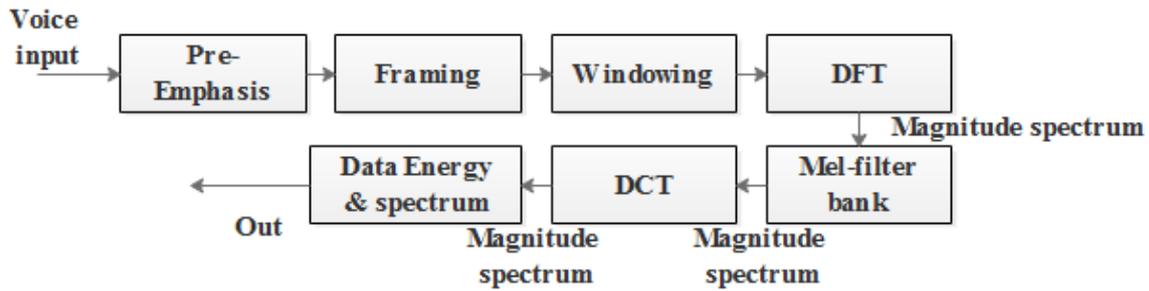


Figure.2 Block diagram of MFCC FE

**3.1.2. Artificial neural network (ANN)**

Artificial neural networks (ANN) are generally known as a technique offering another way to resolve the difficulties and ill-defined issues. In case if the fault is lenient then it can able to handle noisy data, imperfect data and data with nonlinear issues; once trained, it can perform prediction and simplification at high speed. It can be extensively used in diverse applications such as forecasting, control, medicine, robotics, pattern recognition, power systems, optimization, and signal processing, manufacturing, and social/psychological sciences. Specifically, it is very valuable in system modeling such as utilizing complex mappings and scheme identification. The NN consist of two major steps such as testing and training which has been defined below.

**3.1.3. Neural network training**

NN basically contains a nonlinear functional block named as neurons, which is connected to each other based on parallel synaptic weights as well as each neuron has some corresponding threshold value. To create the activation of a neuron, the weighted average of the inputs is formed and the threshold is subtracted. There are also hidden neurons present inside the network, these neurons play an internal role in the network. In training, the neuron is represented by MFCC feature values. Manually the NN is trained based on MFCC voice information. The weight of the neurons is adjusted while performing the training operation. The NN training Structure is shown in the Fig. 3.

**3.1.4. NN testing**

MFCC feature extracts voice data from the NN training. In the NN testing, the extracted voice data of a particular persons are cross checked with the trained data. Finally, the NN gives the exact Optimized voice data signal which is updated in Web-page using the Internet. The mathematical representation of the NN is derived below.

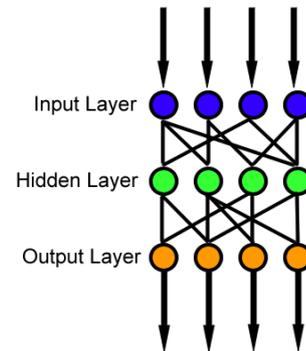


Figure.3 NN training structure

The input  $j$  is given to the NN in the layer of  $l + 1$ , which is expressed as follows Eq. (1).

$$m^{l+1}(j) = \sum_{k=1}^{S_i} w^{l+1}(j, k) b^l(k) + c^{l+1}(j) \tag{1}$$

The output of the NN is expressed in Eq. (2).

$$b^{l+1}(j) = f^{l+1}(m^{l+1}(j)) \tag{2}$$

The matrix form of system equations for a  $N$  layer network is denoted as  $b^0$  and it is expressed as follows,

$$b^0 = q \tag{3}$$

$$b^{l+1} = f^{l+1}(W^{l+1} b^l + c^{l+1}), l = 0, 1, \dots, N - 1 \tag{4}$$

The Eq. (3) and (4) are used to forward the input to the NN. It absorbs associations among the defined set of input and output pairs such as  $\{(q_1, u_1), (q_2, u_2), \dots, (q_R, u_R)\}$ .

An approximate decent rule is employed in back propagation algorithm. The approximate form of the performance index is  $V$  and it is expressed as follows Eq. (5):

$$\hat{V} = \frac{1}{2} e_r^U e_r \tag{5}$$

Where  $e_r$  is the error of the  $r$ th input and is specified as  $e_r = u_r - b_r^N$  and the squared errors for a single input/output pair is exchanges the total sum of errors.

The approximate steepest (gradient) algorithm is specified in Eq. (6) and (7),

$$\Delta w^l(j, k) = -\beta \frac{\partial \bar{V}}{\partial w^l(j, k)} \tag{6}$$

$$\Delta c^l(j) = -\beta \frac{\partial \bar{V}}{\partial c^l(j)} \tag{7}$$

Where the learning rate is denoted as  $\beta$ .

The performance index sensitivity is to changes the input  $j$  at the layer of  $l$  and it is expressed as following Eq. (8):

$$\gamma^l(j) = \frac{\partial \bar{V}}{\partial m^l(j)} \tag{8}$$

By considering Eq. (1), (5) and (8)

$$\frac{\partial \bar{V}}{\partial w^l(j, k)} = \frac{\partial \bar{V}}{\partial w^l(j)} \frac{\partial w^l(j)}{\partial w^l(j, k)} = \gamma^l(j) b^{l-1}(k) \tag{9}$$

$$\frac{\partial \bar{V}}{\partial c^l(j)} = \frac{\partial \bar{V}}{\partial w^l(j)} \frac{\partial w^l(j)}{\partial c^l(j)} = \gamma^l(j) \tag{10}$$

The weight and offset updation process is accomplished by using Eq. (7), (8), (9) and (10). The recurrence relation is satisfied by the sensitivities and it is specified as  $\gamma^l$ .

$$\gamma^l = F^l(m^l) W^{l+1} \gamma^{l+1} \tag{11}$$

The initialized final layer recurrence relation as below:

$$\gamma^N = -F^N(m^N)(u_r - b_r) \tag{12}$$

Where the Eq. (11) and (12) is propagate the sensitivities back.

The back propagation is a steepest algorithm is executed and the LM algorithm is used to approximate the newton's method. The function  $V(y)$  is minimized with respect to the parametry. The Newton's method written as Eq. (13).

$$\Delta y = [\nabla^2 V(y)]^{-1} \nabla V(y) \tag{13}$$

Where the hessian matrix is represented as  $\nabla^2 V(y)$ , the gradient is represented as  $\nabla V(y)$ ,  $y$  is expressed as Eq. (14).

$$y = [w^1(1,1) w^1(1,2) .. w^1(S1, Q) c^1(1). c^1(S1) w^2(1,1) .. c^N(SN)]^T \tag{14}$$

If the  $V(y)$  is the sum of squares function in Eq. (15)

$$V(y) = \sum_{j=1}^M e_j^2(y) \tag{15}$$

Where  $M$  is expressed as follows Eq. (16),

$$M = R \times SN \tag{16}$$

The hessian matrix and gradient is expressed as follows in Eq. (17) and (18),

$$\nabla V(y) = J^T(y) e(y) \tag{17}$$

$$\nabla^2 V(y) = J^T(y) J(y) + S(y) \tag{18}$$

Where  $J(y)$  is defined as the Jacobian matrix. The Jacobian matrix is computed by the simple modification in the back propagation algorithm as well as it is used for avoiding the mapping problem in  $y$  of Eq. (13) in the neural network. The  $J(y)$  expressed in the form matrix as given below

$$j(y) = \begin{bmatrix} \frac{\partial e_1(y)}{\partial y_1} & \frac{\partial e_1(y)}{\partial y_2} & \dots & \frac{\partial e_1(y)}{\partial y_m} \\ \frac{\partial e_2(y)}{\partial y_1} & \frac{\partial e_2(y)}{\partial y_2} & \dots & \frac{\partial e_2(y)}{\partial y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial e_m(y)}{\partial y_1} & \frac{\partial e_m(y)}{\partial y_2} & \dots & \frac{\partial e_m(y)}{\partial y_m} \end{bmatrix}$$

$$\text{and } S(y) = \sum_{j=1}^M e_j(y) \nabla^2 e_j(y) \tag{19}$$

Eq. (19) are leads to the Levenberg-Marquardt method,

$$\nabla y = [J^T(y) J(y) + \phi I]^{-1} J^T(y) e(y) \tag{20}$$

If the  $V(y)$  is increased, the parameter  $\phi$  is multiplied some factor that is  $\alpha$ . While the parameter  $\phi$  is small, the process becomes the LM algorithm or else it is executed as a steepest descent algorithm which is denoted in the Eq. (20).

The terms calculated by the standard propagation is represented in Eq. (21)

$$\frac{\partial \bar{V}}{\partial w^l(j, k)} = \frac{\partial \sum_{r=1}^{SN} e_r^2}{\partial w^l(j, k)} \tag{21}$$

The standard propagation algorithm with one modification at the final layer is,

$$\Delta^N = -F^N(m^N) \quad (22)$$

In Eq. (22), the back propagation is done at each column of matrix that is sensitivity vector for produce the one row of the jacobian.

### 3.2 RASP based Monitoring and controlling (Receiving)

RASP-IoT-SA system consist of two major blocks: controlling and monitoring.

#### 3.2.1. Monitoring

The Monitoring RASP-IoT-SA system consists of five Sensors such as Temperature, Voltage, Current, Oil Level, Vibration and speed. The temperature sensor helps to monitor the temperature of electrical device. Voltmeter and current meters find the Current and voltage level of the electrical devices. The sensor values updated in the web-page, are obtained from the transmitting section based on voice commands.

#### 3.2.2. Controlling

In controlling block, the electrical devices are connected with the RASP. Depend upon the Transmitting section command, the switching operation of the electrical devices are controlled.

## 4. Hardware and software description

The Hardware design of RASP-IoT-SA system is sub-divided into two parts due to various functions in the RASP-IoT-SA system. The main use of this technique is to receive the commands and apply action according to it. The first part is a transmitting unit consist of RASP-PI-3 with Raspbian Jessie OS with Python programing including Microphone (3.5mm HYPO clip-MIC).

In the second part of RASP-IoT-SA is monitoring and controlling unit; which consist of same RASP-PI-3 with Raspbian Jessie (Which of consist inbuilt Bluetooth) Operating system including Python programing.

## 5. Result and description

The RASP-IoT-SA methodology has ability to convert voice commands into text which is updated on the web page via the Internet. The controlling and monitoring functions are processed by using the web-

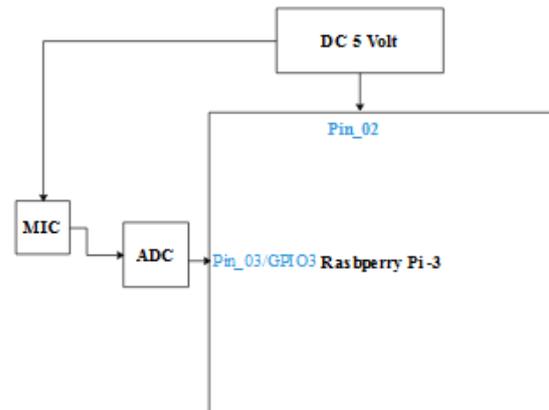


Figure.4 Circuit diagram for RASP-IoT-SA system transmitting section

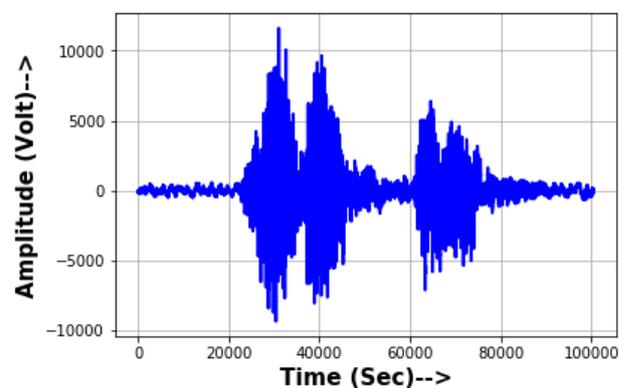


Figure.5 Input audio signal

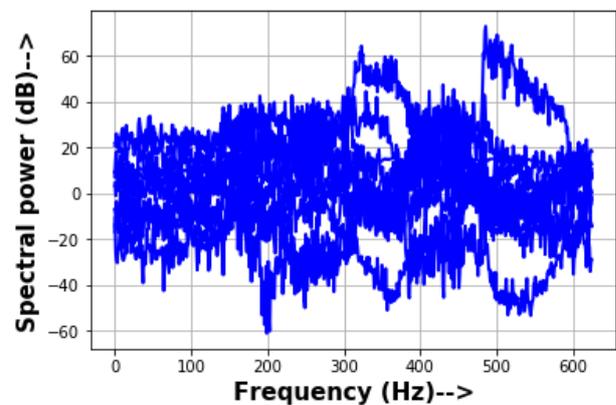


Figure.6 MFCC audio Feature extracted signal

page update. The route diagram of the RASP-IoT-SA transmitting methodology is shown in below Fig. 4.

In this section, voice command is received from the 3.5mm HYPO clip-MIC Microphone in the form of an analog signal. It's converted into the digital signal with the help of AD converter. The converted digital signal is given to the RASP-PI-3 RASP which is shown in the Fig. 5.

By using the python language, voice commanded MFCC feature values are extracted in the RASP-PI-3 which is shown in the Fig. 6.

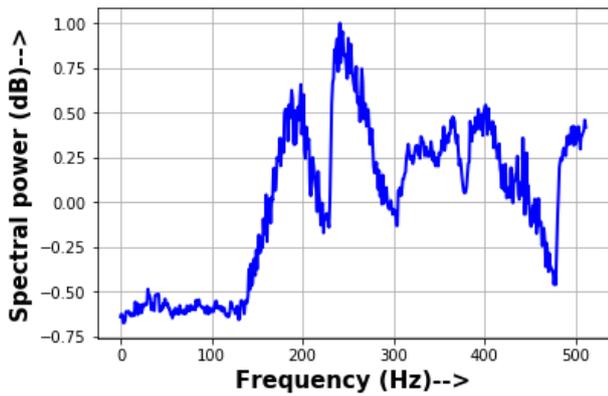


Figure.7 MFCC audio feature normalized signal

MFCC provides the Delta cepstrum feature values which consist of  $1116 \times 26$ . This 1116 represents the velocity and acceleration of energy + MFCC and 26 means that MFCC filter bank. This kind of features does not provide better accuracy, so that the feature values are normalized which is shown in the Fig. 7.

The feature values are given to the ANN, in this RASP-IoT-SA system two different levels of ANN are employed. The First level ANN helps to identify the speaking person and if the person is authenticated, the second level of ANN extracted the voice data.

In the RASP-IoT-SA system, the ANN consist of two sections such as testing and training. In the training portion, voice commands for different persons are recorded and edited with the help Audacity tool and the audio edited files are saved as wave data format. The wave data files are updated in the RASP, then MFCC feature values are extracted for all signals and it trained to the ANN. In the RASP-IoT-SA system training purpose 30 different people voices with 15 different classes are recorded. The 15 different classes commands are mentioned in the below Table 1.

Table.1 ANN training table

Class	Device condition
Class-1	Motor Temperature sense
Class-2	Motor Current sense
Class-3	Motor Speed sense
Class-4	Motor Voltage sense
Class-5	Motor Vibration sense
Class-6	Transformer Voltage sense
Class-7	Transformer Temperature sense
Class-8	Transformer Oil Level sense
Class-9	Motor Off
Class-10	Motor On
Class-11	Speed level 1
Class-12	Speed level 2
Class-13	Speed level 3
Class-14	Transformer Tap 1
Class-15	Transformer Tap 2

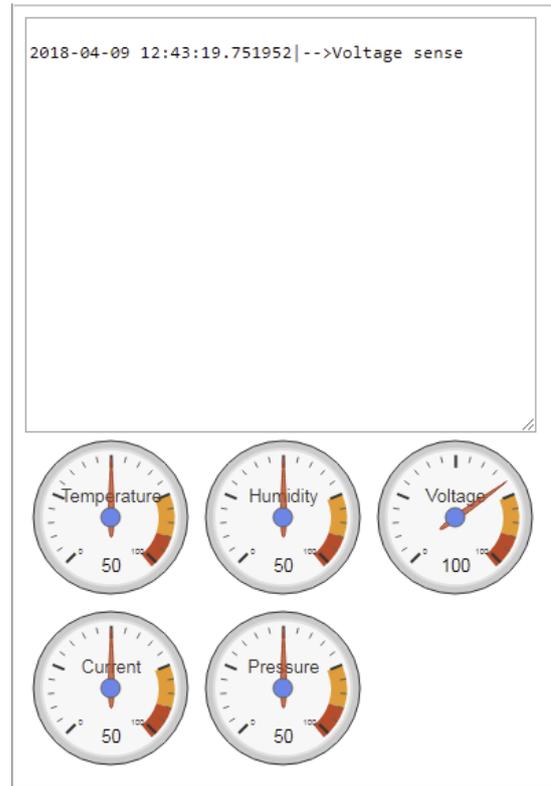


Figure.8 RASP-IoT-SA system drop box Text file

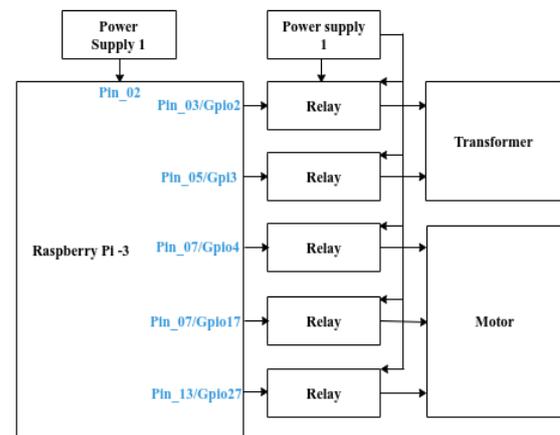


Figure.9 Circuit diagram for RASP-IoT-SA system receiving section

In the testing section input command is passed through the ANN's, which analyzed the audio signal and provide the predicted output class. The Python program declares the output (Text) that is written in the text file. The output text is saved in the text file connected with an Internet. The web page screen shot is shown in the Fig. 8.

From the web page, the receiving section is processed. For example, if the voice command is TEMP sense, then the transmitter update the command in the drop box file, at the same time Receiver Section read the text file and execute the

	precision	recall	f1-score	support
1	1.00	1.00	1.00	11
2	1.00	1.00	1.00	11
3	1.00	1.00	1.00	11
4	1.00	1.00	1.00	11
5	0.92	1.00	0.96	11
6	1.00	1.00	1.00	11
7	1.00	1.00	1.00	11
8	1.00	1.00	1.00	11
9	1.00	1.00	1.00	11
10	1.00	1.00	1.00	11
11	1.00	1.00	1.00	11
12	1.00	1.00	1.00	11
13	1.00	1.00	1.00	11
14	1.00	1.00	1.00	11
15	1.00	1.00	1.00	11
16	1.00	0.91	0.95	11
17	1.00	1.00	1.00	11
avg / total	1.00	0.99	0.99	187
Accuracy	0.9946524064171123			

Figure.10 STC Performance

particular Monitoring and controlling process. By using this RASP-IoT-SA system only eight devices are controlled and monitored but we can able to control and monitor more than 30 devices. The route diagram of the Receiver section is shown in the Fig. 9.

The RASP-IoT-SA system has tested for the 17 different persons the results of that has represented in the Fig. 10. The system performance in terms of Precision, recall, false score and accuracy.

The voice recognition, voice to text conversion performance are calculated by using the Eq. (23-26).

$$Accuracy = \frac{tp+tn}{tp+fp+tn+fn} \tag{23}$$

$$Recall = \frac{tp}{tp+fn} \tag{24}$$

$$Precision = \frac{tp+tn}{tp+tn+fp+fn} \tag{25}$$

$$FalseMeasure = \frac{2.R.P}{R+P} \tag{26}$$

Where, *tp* is named as True positive, *tn* is named as True negative, *fp* is named as False Positive, *fn* is named as False negative, *R* is named as a Recall and *P* is named as Precession.

The comparison between the Existing systems voice recognition and RASP-IoT-SA system is mentioned in the Fig. 11 and Table 1. From the analysis, it is concluded that the RASP-IoT-SA system provide better accuracy compared to the VC-SS-IA [15], BV-RSS [17], MFCC-VQ [23], MFCC-DTW [23], MFCC-PLP-HMM [23] and MFCC-PLP-DTW [23].

Table 1. Comparison between the existing and RASP-IoT-SA system

Technique	Accuracy
BV-RSS [17]	75%
MFCC-VQ [23]	96%
MFCC-PLP-DTW [23]	92%
MFCC-DTW [23]	91%
MFCC-PLP-HMM [23]	98%
VC-SS-IA [15]	95%
RASP-IoT-SA	99.46%

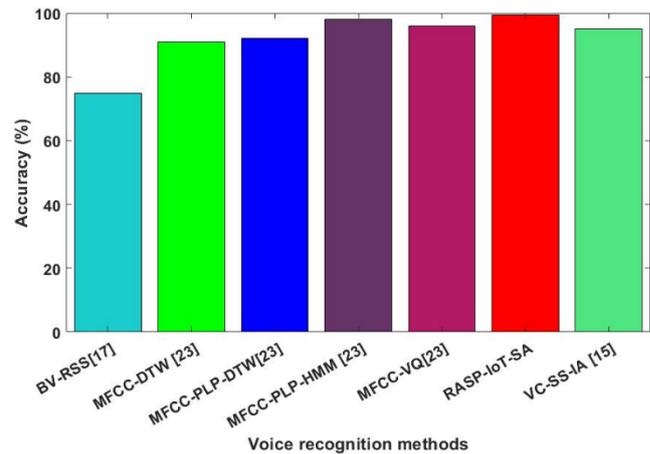


Figure.11 Accuracy Comparison between the Existing and RASP-IoT-SA system

Smartphone or laptop via Raspberry Pi to control light switches, fans and A.C. system is introduced in [15]. The author has mainly concentrated on the home automation. The major limitation of the system is security (any one can access the system while knowing the IP address) and physically challenged people easily couldn't be able to access system.

Compared to the other existing MFCC base technique such as MFCC-VQ [23], MFCC-DTW [23], MFCC-PLP-HMM [23] and MFCC-PLP-DTW [23]. They have only considered with any one of log filter bank signal. In this RASP-IoT-SA system all 26 log filter bank signals are average and which is used as a feature values. This process makes the RASP-IoT-SA system better compare to others. Compared to the other existing systems [1-23] another one greater advantage of the RASP-IoT-SA system is speaker person identification system which provided 96.32% of accuracy.

## 6. Conclusion

An IoT based automation system is introduced for an automatic working of electrical devices in home and industries. The voice based automation technique helps to access the information through Internet in remote areas. The designed technique not only monitors the sensor data, like temperature, Humidity, light, condition motion sensors, but also activates the

process according to the requirements. The MFCC and ANN based voice data processing system helped to improve the efficiency of the RASP-IoT-SA system which provided 99.46% of accuracy. From the Comparative analysis RASP-IoT-SA system provided much better results compared to other existing systems. The RASP-IoT-SA system contains speaker identification which provided 96.32% of accuracy. Furthermore, this system is used as a framework for security based automation to expand various applications such as capturing the photo of a person moving around the house and storing it on the cloud. This reduced the data storage instead of using CCTV camera which records at all time and stores it.

## References

- [1] B. Yuksekkaya, A.A. Kayalar, M.B. Tosun, M.K. Ozcan, and A.Z. Alkar, "A GSM, internet and speech controlled wireless interactive home automation system", *IEEE Transactions on Consumer Electronics*, Vol.52, No. 3, pp.837-843, 2006.
- [2] S.I. Azid and S. Kumar, "Analysis and performance of a low cost SMS based home security system", *International Journal of Smart Home*, Vol.5, No.3, pp.15-24, 2011.
- [3] P.J. Berkelman, L.L. Whitcomb, R.H. Taylor, and P. Jensen, "A miniature microsurgical instrument tip force sensor for enhanced force feedback during robot-assisted manipulation", *IEEE Transactions on Robotics and Automation*, Vol.19, No.5, pp.917-921, 2003.
- [4] A.Z. Alkar and U. Buhur, "An Internet based wireless home automation system for multifunctional devices", *IEEE Transactions on Consumer Electronics*, Vol.51, No.4, pp.1169-1174, 2005.
- [5] T. Obaid, H. Rashed, A.A.E. Nour, M. Rehan, M.M. Saleh, and M. Tarique, "ZigBee based voice controlled wireless smart home system", *International Journal of Wireless & Mobile Networks*, Vol.6, No.1, pp.47, 2014.
- [6] M. Chan, D. Estève, C. Escriba, and E. Campo, "A review of smart homes—Present state and future challenges", *Computer methods and programs in biomedicine*, Vol.91, No.1, pp.55-81, 2008.
- [7] P. Rashidi and D.J. Cook, "Keeping the resident in the loop: Adapting the smart home to the user", *IEEE Transactions on systems, man, and cybernetics-part A: systems and humans*, Vol.39, No.5, pp.949-959, 2009.
- [8] I.A. Anderson, T.A. Gisby, T.G. McKay, B.M. O'Brien, and E.P. Calius, "Multi-functional dielectric elastomer artificial muscles for soft and smart machines", *Journal of Applied Physics*, Vol.112, No.4, p.041101, 2012.
- [9] C.L. Hsu, S.Y. Yang, and W.B. Wu, "3C intelligent home appliance control system—Example with refrigerator", *Expert Systems with Applications*, Vol.37, No.6, pp.4337-4349, 2010.
- [10] H.N.M. Shah, M.Z. Ab Rashid, M.F. Abdollah, M.N. Kamarudin, K.L. Chow, and Z. Kamis, "Biometric voice recognition in security system", *Indian Journal of Science and Technology*, Vol.7, No.2, pp.104-112, 2014.
- [11] H. Rajput, K. Sawant, D. Shetty, P. Shukla, and A. Chougule, "Implementation of Voice Based Home Automation System Using Raspberry Pi", *International Research Journal of Engineering and Technology*, Vol.5, No.5, pp.2771-2776, 2018.
- [12] L. Basyal, S. Kaushal, and G. Singh, "Voice Recognition Robot with Real Time Surveillance and Automation", *International Journal of Creative Research Thoughts*, Vol.6, No.1, pp.11-16, 2018.
- [13] I.U. Khan, M. Arif, M.F. Khalid, R. Ali, Q. Khan, and S. Ahmad, "Voice controlled home automation system", *IJRCCT*, Vol.6, No.5, pp.148-154, 2017.
- [14] K. Pampattiwar, M. Lakhani, R. Marar, and R. Menon, "Home Automation using Raspberry Pi controlled via an Android Application", *International Journal of Current Engineering and Technology*, Vol.7, No.3, pp.962-967, 2017.
- [15] A. Mane, N. Thopate, P. Pawar, M. Kanse, B.E. Student, S.N. SEC, and P. Baramati, "A Survey on: Home Automation through Voice Recognition Application and Messaging on Social Sites", *International Journal of Engineering Science*, Vol.7, No.12, 15772-15775, 2017.
- [16] H. Bharathi, U. Srivani, M.D. Azharudhin, M. Srikanth, and M. Sukumarline, "Home Automation by Using Raspberry Pi and Android Application", In: *Proc. of International Conf. On Electronics, Communication and Aerospace Technology*, Vol.2, pp.687-689, 2017.
- [17] T. Obaid, H. Rashed, A.A.E. Nour, M. Rehan, M.M. Saleh, and M. Tarique, "ZigBee based voice controlled wireless smart home system", *International Journal of Wireless & Mobile Networks*, Vol.6, No.1, pp.47, 2014.
- [18] G. Muthuselvi and B. Saravanan, "Real Time Speech Recognition Based Building Automation System", *ARNP Journal of Engineering and Applied Sciences*, Vol.9, pp.2831-2839, 2014.

- [19] M.R. Kamarudin and M.A.F.M. Yusof, “Low cost smart home automation via microsoft speech recognition”, *International Journal of Engineering & Computer Science*, Vol.13, No.3, pp.6-11, 2013.
- [20] F. Baig, S. Beg, and M.F. Khan, “Zigbee based home appliances controlling through spoken commands using handheld devices”, *International Journal of Smart Home*, Vol.7, No.1, pp.9-26, 2013.
- [21] S.X. Ding, P. Zhang, S. Yin, and E.L. Ding, “An integrated design framework of fault-tolerant wireless networked control systems for industrial automatic control applications”, *IEEE Transactions on Industrial Informatics*, Vol.9, No.1, pp.462-471, 2013.
- [22] H.J.R.S.R Sushma, “Wireless Monitoring for Industrial Automation Using Speech Recognition”, *International Journal of Engineering and Innovative Technology*, Vol.3, No.11, pp.133-139, 2014.
- [23] S. Goyal, and N. Batra, “Issues and Challenges of Voice Recognition in Pervasive Environment”, *Indian Journal of Science and Technology*, Vol.10, No.30, pp.1-4, 2017.