

# Intelligent Virtual Assistant System (IVAS) for Air Traffic Controllers

Anusha DILHANI, Ruminda J WIMALASIRI

*(Department of Mechanical Engineering, The Open University Sri Lanka, Nawala, Nugegoda, Sri Lanka)*

**Abstract:** Air traffic control is an essential obligation in the aviation industry to have safe and efficient air transportation. Year by year, the workload and on-job-stress of the air traffic controllers are rapidly increasing due to the rapid growth of air traveling. Controllers are usually dealing with multiple aircrafts at a time and must make quick and accurate decisions to ensure the safety of aircrafts. Heavy workload and high responsibilities create air traffic control a stressful job that sometimes could be error-prone and time-consuming, since controlling and decision-making are solely dependent on human intelligence. To provide effective solutions for the mentioned on the job challenges of the controllers, this study proposed an intelligent virtual assistant system (IVAS) to assist the controllers thereby to reduce the controllers' workload. Consisting of four main parts, which are voice recognition, display conversation on screen, task execution, and text to speech, the proposed system is developed with the aid of artificial intelligence (AI) techniques to make speedy decisions and be free of human interventions. IVAS is a computer-based system that can be activated by the voice of the air traffic controller and then appropriately assist to control the flight. IVAS identifies the words spoken by the controller and then a virtual assistant navigates to collect the data requested from the controllers, which allows additional or free time to the controllers to contemplate more on the work or could assist to another aircraft. The Google speech application programming interface (API) converts audio to text to recognize keywords. AI agent is trained using the Hidden marko model (HMM) algorithm such that it could learn the characteristics of the distinct voices of the controllers. At this stage, the proposed IVAS can be used to provide training for novice air traffic controllers effectively. The system is to be developed as a real-time system which could be used at the air traffic controlling base for actual traffic controlling purposes and the system is to be further upgraded to perform the task by recognizing keywords directly from the pilot voice command.

**Keywords:** Artificial Intelligence, Air Traffic Controller, Intelligent Virtual Assistant

## 1 Introduction

Air traffic controlling is an advisory service provides to the pilot by the Air Traffic Controllers (ATC), which operates from the ground. Air traffic control is a well-trained, highly professional job in which the safety of the passenger is the primary responsibility. ATC direct or guide the aircraft by giving commands

and necessary information to the pilot. The international civil aviation organization (ICAO) defines two major duties of ATC, which are, preventing collisions between aircrafts and between aircraft with other obstructions, and expediting and maintaining an orderly flow of air traffic<sup>[1]</sup>.

Every day, over two million people travel across the skies while ATC are working vigilantly to keep all

air travelers safe<sup>[2]</sup>. Twenty-four hours a day, seven days a week, ATC are on duty across the countries throughout the world, keeping a watchful “eye to the sky”. Generally, ATC are guiding airplanes which travel overhead at 500 miles an hour.

The heavy workload along with the high responsibilities makes air traffic controlling a strenuous job<sup>[3]</sup>. Stress is a part of everyday life and is a physiological stimulus associated with the interaction of human and environment. Stress-at-work can be generated by job demands, environmental conditions, organizational aspects and human relations, and will affect to job satisfaction, performance efficiency and health. These affects could be differ depending on the psycho-physical characteristics and coping resources of individuals, as well as on the social support received. However, it can become a harmful risk factor for health when it is perceived as an imbalance between an excess of demands and the individual ability to meet them.

ATC are generally considered as one of the working groups having to deal with a highly demanding job responsibility. In fact, this entails a complex set of tasks requiring a very high level of knowledge and experience, as well as the practical application of certain skills related to different cognitive domains.

According to a job analysis of ATC six main activities can be identified (which are, situation monitoring, resolving aircraft conflicts, managing air traffic sequences, routing or planning flights, assessing weather impact, managing sector/position resources), which include 46 sub-activities and 348 distinct tasks. For example, the relevant cognitive/sensory attributes required for high performance levels at radar workstations are spatial scanning, movement detection, image and pattern recognition, prioritizing, visual and verbal filtering, coding and decoding, inductive and deductive reasoning, short- and long-term memory, and mathematic and probabilistic reasoning<sup>[3]</sup>.

The controlling unit must constantly reorganize its flight information processing system, changing its operating methods (in particular, cognitive processes, conversation, coordination with helpers, foresight and

problem solving) as they arise and interact with each other. This is done through the precise and efficient application of rules and procedures. However, flexible settings are required depending on different circumstances, often in conditions of time pressure. ATC job can sometimes be error-prone and time-consuming as control and decision-making are entirely dependent on experience, human intelligence, and capabilities.

Nowadays, more and more people are provided with on-job assistance by appropriate technical systems such as assistants/decision support systems which can be found in most workplaces and on-vacation environments. Latest apps such as Apple siri<sup>[4]</sup>, and Google voice search<sup>[5]</sup>, use speech recognition as input interface for direct communication between a person and a machine to run a specific action.

The specific virtual assisting systems are developed to provide on-job assistance to perform the assigned duty as and when requested by the worker. The technology has evolved to improve the efficiency and extends beyond the tasks traditionally required by humans. Speech recognition is been used in industries such as automotive<sup>[6]</sup>, banking and healthcare. In automotive industry, speech recognition technology has been used for developing and improve safety of their products. To provide effective assistance to ATC, an Intelligent virtual assistant system (IVAS) is developed which can be used to support reducing the burden/on job stress. IVAS is a computer system that can be activated by a voice signal of ATC. IVAS is cable of providing rapid assistance which allows free time to ATC to think on the job or to assist to yet another aircraft or to perform additional tasks such as, script maintaining, situation monitoring, routing planning.

## 2 Approach and Method

### 2.1 Current System

Currently, most airports use an Automated terminal radar system (ARTS) to track aircraft when any aircraft enters their aerospace industry<sup>[7]</sup>. Fig.1 shows the ARTS system.

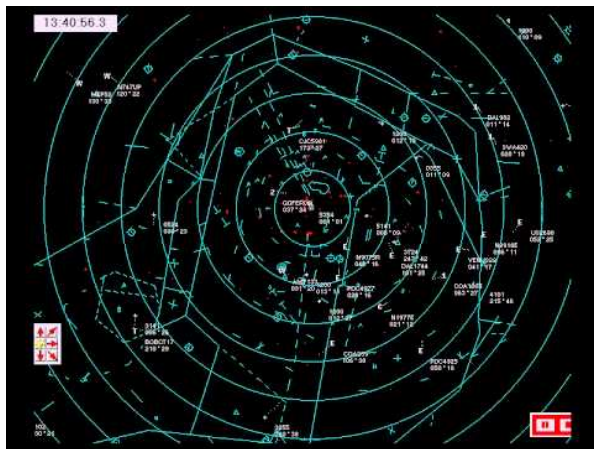


Fig.1 ARTS System

Standard terminal automation replacement system (STARS) was developed to replace ARTS in 2010 (Fig. 2). It is a color display with associated other inputs, computers and networks. STARS provide the controller with aircraft positional information with the weather, flight data and other ATC information well. The ARTS and STARS systems are designed to identify each aircraft by matching its transponder code with flight plan data, provided either by the flight data processing computer located in the Terminal radar approach control (TRACON). Once the aircraft identified, the ARTS computer system maintained constant identification and predicted the aircraft's future location.

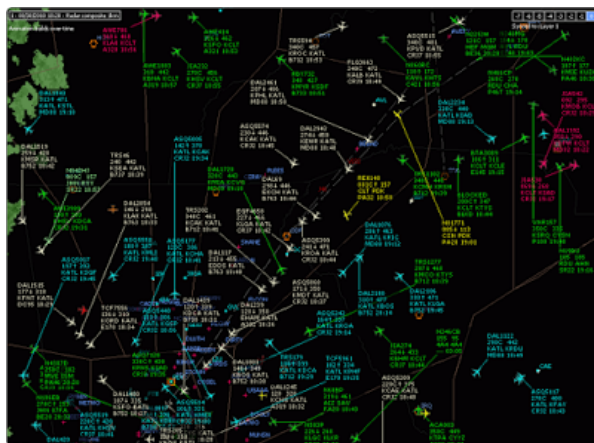


Fig.2 STARS System

Both systems are look the same and locating/finding information of an aircraft manually is difficult.

ATC is trying to locate aircraft position and vital details on this radar screen when the pilot request for information over voice. Most pilots use voice communication to receive instructions or flight information from ATC. Rarely do they use a text message system to ask for details. This text message system is linked to the radar screen and automatically detects the message details and displays the relevant information on the screen. But the voice communication system is not directly connected to the ARTS or STARS systems. Therefore, ATC must first locate the aircraft on the radar screen. It takes a considerable amount of training to find the aircraft and its information on the screen. It is obvious that this job is an urgent matter and finding an aircraft on the screen will take a considerable time, as there are thousands of aircrafts display on the screen at a given instant.

## 2.2 Design

IVAS is a computer-based application which can be installed on ARTS or STARS systems. It has main two components; voice assistant and texting system where both can be used to fulfill requested tasks in separate ways. IVAS can be recognized the ATC voice command and after recognition it assists to gather related information on the radar screen. Then the information send to ATC in the form of voice (by words). In addition, IVAS can highlight relevant information found on the radar screen.

The voice assistant system is activated by the ATC' voice command and the entire conversation between IVAS and ATC is displayed on the screen. ATC also can ask for any help from the IVAS texting system by sending a text message and then again IVAS will assist to do the requested work. Instead of a radarscope, the 'live screen Radarbox.com was used. It is a desktop application with four main parts which are the voice recognition system, intelligent navigation system, graphical interface system and voice assistant system.

Fig.3 illustrates the block diagram of the developed desktop application for the air traffic control service and it represent the tasks that the system performs.

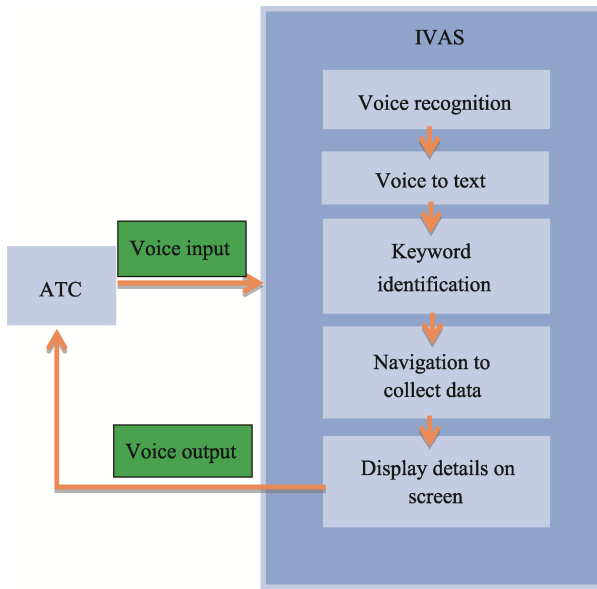


Fig.3 Block Diagram of the System

### 2.3 Technology and Libraries

IVAS provides contextual information and performs actions such as finding requested information or sending a message. ATC could also type requests to IVAS if they do not wish to get the assistance of voice inputs. To perform its functions, IVAS relies on artificial intelligence (AI) technologies and concepts such as Natural language processing (NLP) and Machine learning (ML) to understand what the user is saying and to make suggestions or act on that language input. IVAS was programmed to understand what a person is saying through tools like natural language processing and machine learning. IVAS also benefits from machine learning technology, which focuses on systems

that teach machines to learn from experience and make decisions accordingly, rather than being programmed for every single task. In other words, machine learning is about giving the machines the rules and leaving them to learn from trial and error without explicit instructions in programming codes. IVAS instantly processes the word inputs by ATC and understands exactly what ATC is saying. It then responds by constructing clear sentences as a human being would base on what it has learned. To implement AI aspirations such as IVAS, requires a stable, flexible; programming language with necessary tools. Python with its rich technology stack consists of extensive set of libraries for ML. Therefore, ‘python’ is used to develop the IVAS because it offers all required features.

Speech recognition is new to the air traffic controlling field. Speech recognition is a technology that enables the recognition and translation of spoken language into text by computers. Speech recognition works using algorithms through acoustic and language modeling. Acoustic modeling represents the relationship between linguistic units of speech and audio signals; language modeling matches sounds with word sequences to help distinguish between words that sound similar. Fig.4 illustrates the working nature of speech recognition in IVAS as below.

Speech recognition consists of two main modules which are feature extraction, and feature matching. The purpose of the feature extraction module is to convert speech waveform to some type of representation for further analysis and processing, this extracted information is known as the feature vector. In feature

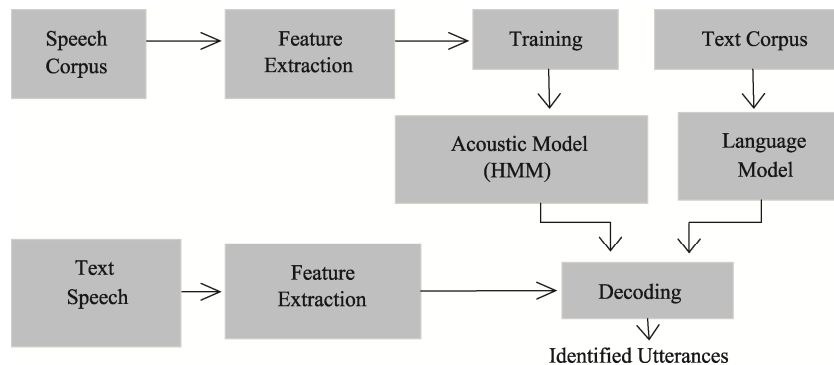


Fig.4 Speech Recognition Model of the System

matching, the extracted feature vector from an unknown voice sample is scored against the acoustic model, the model with max score wins, and its output is considered as a recognized word. The mel-frequency ceptrum coefficient (MFCC) and linear predictive coding (LPC) are the methods for extracting feature factors that were used in developing the IVAS system. Once the feature vector is obtained IVAS builds the acoustic model. The acoustic model is used to score the unknown voice sample. In speech recognition, the basic unit of sound is a phoneme. A phoneme is a minimal unit that serves to distinguish between meanings of words. To recognize a given word, IVAS should extract a phoneme from a voice sample. Hidden markvo model (HMM) neural network algorithm was used to build an acoustic model. IVAS deals with the ATC voice, and they have their own phraseologies and unique phonemes word set. IVAS is trained to recognize a set of 167 words and these words act as the keywords of a system. Each word, or each phoneme, have a different output distribution; a hmm for a sequence of words or phonemes is made by concatenating the individual trained Hmm for the separate words and phonemes.

## 2.4 Model Building

In this research 167 words were used as a keyword. Some of general and others are standard words. All these words are taken for standard phrases which in use ATC services as well. Those 167 words are divided into three main categories as below. Word list one has included 102 words which are called numerical word (eg: zero, one, hundred) set. Second world list has included 35 words which are called general list (eg: I, you, flight, altitude). Third word list has included 30 words which are called standard word (eg: bravo, alpha, roger).

As an example, when pilot request current altitude details from ATC, in the communication process is listed below.

**Pilot:** Flight EY five seven need altitude

ATC can get support from the IVAS for finding requested information. Below phraseology can be used to

request assistant from IVAS. IVAS can identify what ATC request.

**ATC:** Flight EY five seven altitude.

ATC and IVAS conversation flows as below.

**ATC:** Flight EY five seven altitude please.

**IVAS:** Callsign please.

**ATC:** Flight EY five seven.

Then the IVAS attend to the keywords what it recognize and then IVAS navigate to gather requested details.

**IVAS:** EY five seven altitude

In this case, phoneme-based system was used. The main advantage in the phoneme-based recognition is that the number of phonemes is limited in any language, wherever they have been itemized, as opposed to the unlimited number of words that can be built from phonemes.

There are five separate phonemes which are E, Y, 5,7 and Altitude. According to the HMM probability equation, probability matrix for the given example shows in Table 1.

Markovian property,

$$P(q_1, \dots, q_n) = \prod_{i=1}^n P(q_i | q_{i-1})$$

**Table 1 The Probability Matrix**

	E	Y	Five	Seven	Altitude
E	0.7	0.2	0.3	0.1	0.2
Y	0.1	0.8	0.1	0.3	0.3
Five	0.1	0.5	0.7	0.4	0.1
Seven	0.3	0.1	0.1	0.6	0.4
Altitude	0.4	0.2	0.2	0.1	0.6

HMM model for the IVAS is illustrated in Fig.5 and signal graph for the “EY Five Seven Altitude” is shown in Fig.6.

The system is trained with taking 8 speakers phonemes. Three from them are male and others are women. Total sentences that were used to train the model were 37.

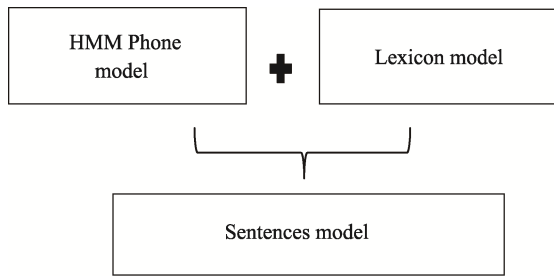


Fig.5 HMM Model for the IVAS

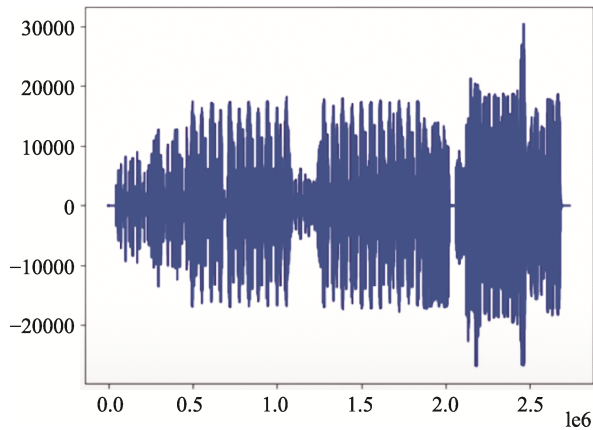


Fig.6 Signal Graph for the Selected Phrase

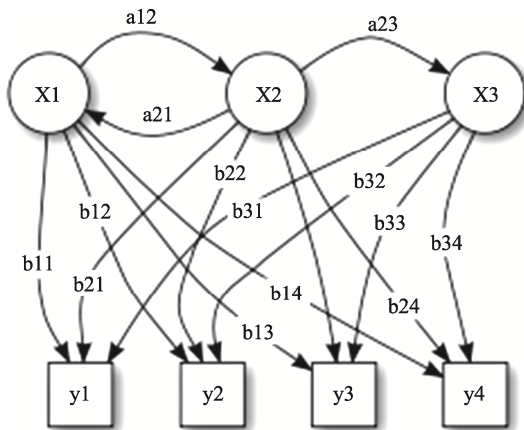


Fig. 7 Probabilistic Parameters for Model

Fig. 7 illustrates the probabilistic parameters behavior of the model for the total system. Indicates of X, y, a and b are below.

X: State

Y: Possible observations

a: State transition probability

b: Output probability

In air navigation services, there are several nationalities available. Their pronunciations are differed one to other, and speech style is varying with age, ascent and situation. To avoid that variation of speech recognition that phoneme based HMM model was built for the IVAS.

### 3 Results and Discussions

The utmost performance indicator of the IVAS is the similarity between the messages of IVAS to the ATC.

The Word error rate (WER) is the concept which used to measure the quality of the speech recognition system. The WER is defined as follow:

$$WER = \frac{S+D+I}{N}$$

Where,

S: number of substitutions

D: number of deletions

I: number of insertions

N: number of words in the references

The system recorded WER is shown in table 2. IVAS has trained with separate three word category. They are,

Word set 1: Numerical words

Word set 2: General words

Word set 3: Standard words

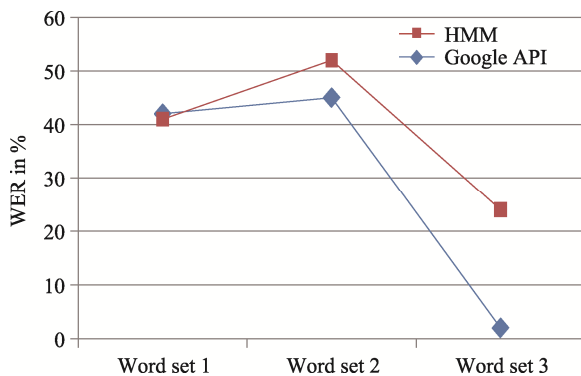
Result of each word test as below.

System	WER %		
	Word set 1	Word set 2	Word set 3
Based on Google API	42	45	1.2
Trained HMM	41	52	24

Initially, IVAS showed a quick response when recognizing general words (e.g.:You, I). But it took a long processing time to identify the standard words use by ATC (e.g.: roger). When occurring phoneme variation, comparing with trained audio clips IVAS could

not recognize the voice command in the first attempt. The Command needs to repeat again and again and took considerable time to navigate for collecting data.

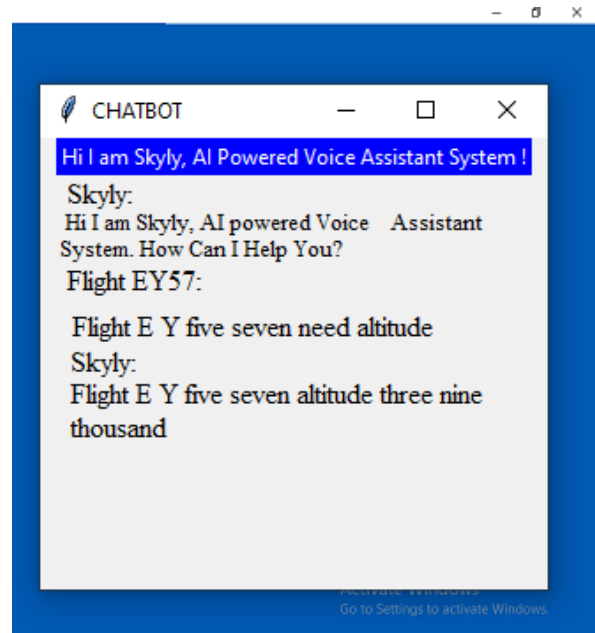
Two separate models were developed for recognizing voice command and to increase accuracy level of the system. To identify words that use in day today life, IVAS obtained support from Google API platform, otherwise also from HMM model. The system takes more time than anticipated to give the output (to assist). Both Google API and HMM models were used to maintain a level of system accuracy and provide redundancy which IVAS synchronized with both models. Behavior of the both models are shown in Fig.8.



**Fig.8 Behavior of the Different Two Model (Google API & HMM) Used in IVAS**

In HMM response more accurately to recognize word which is spoken by ATC. Different phrases with different phoneme more correctly identified by HMM. Therefore, it is shown good behavior than using Google API model. But considering time consuming for recognizing words in phrases, HMM is taken little bit more than Google API.

Each ATC has their own login password and username to IVAS. Fig.9 illustrates a voice assistant window. Each window has a voice activated emergency function. There are two functions in the selection window, voice assistant or messaging assistant, which ATC could select one of them as desire. In the messaging assistant, the ATC can put the desired information in text. Voice assistant consists of two functions which it active with voice command and all the conversation between IVAS and ATC displayed on the screen.



**Fig.9 Voice Assistant Systems**

## 4 Conclusion

This research was performed to build AI assistant for ATC to get help for reducing their workload when traffic gets high. As a first step, a prototype was developed which to be introduced for ATC. The proposed IVAS prototype could be used for training purpose and it is capable of collecting data which requested from ATC. IVAS acts as human being with activation ATC voice command and identified correct keywords from the voice command. It also provides answer to ATC in the form of a voice words accurately.

If IVAS to be effectively used in real time ATC assistance, it needs to be improved in several aspects. Current system needs to train more and more by using further ATC voice command clips. It also requires additional training by using audio clips that are in various ranges of phonemes. These developments are currently performing to make IVAS a system which is completely fit to actual real time conditions. The ultimate aim would be developed IVAS and implement it for air traffic controlling effectively, minimizing the stress and burden of ATC and it will assist to the pilot directly under the minimal supervision of controllers.

## References

- [1] *Procedures for air navigation services Air traffic management*, 6<sup>th</sup> ed, International civil aviation organization, QC, Canada, 2016.
- [2] *The world of air transportation 2019*, International Civil Aviation Organization 2019, [online], Available: <https://www.icao.int/annual-report-2019/Pages/the-world-of-air-transport-in-2019.aspx>
- [3] Helmke H, Ohneiser O, Mühlhausen T, Wies M, 2016 Reducing Controller Workload with Automatic Speech Recognition, *35th Digital Avionics System Conference, Sacramento, CA, USA*.
- [4] SRI International “Siri-based virtual personal assistant-technology”, Available: <http://www.sri.com/engage/ventures/siri>
- [5] J. Schalkwyk, D. Beeferman, F. Beaufays, B. Byrne, C. Chelba, M. Cohen, M. Kamvar, and B. Strope, “Google search by voice: A case study,” in *Advances in Speech Recognition: Mobile Environments, Call Centers and Clinics*, pp. 61-90. Springer, 2010.
- [6] *Automotive voice recognition system market size*, Acume research and consulting, 2020 [online, Available: <https://www.globenewswire.com/news-release/2020/04/14/2015959/0/en/Automotive-Voice-Recognition-System-Market-Size-Worth-Around-US-39-Bn-by-2027.html>
- [7] M.S.Nolan *Fundamentals of Air Traffic Control*. 5<sup>th</sup> ed. Canada: Cengage learning, 2011.

## Author Biographies



**Anusha DILHANI**, received her B.Eng. (Hons) degree from the OUSL. Her main research interest includes the design of an intelligent air traffic control system.

Email: [anushadilhara@gmail.com](mailto:anushadilhara@gmail.com)



**Ruminda J WIMALASIRI**, received his B.Eng. (Hons) degree from the OUSL and Ph.D. degree from Sheffield Hallam University, United Kingdom. He is currently a Senior Lecturer in Mechanical Engineering at OUSL. His main research

interest includes structural integrity of machine elements.

Email: [rjwim@ou.ac.lk](mailto:rjwim@ou.ac.lk)



**Copyright:** © 2021 by the authors. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).