PERSONAL DIGITAL ASSISTNT

[Shivangi Mishra[1], Kanchan Maurya[2], Dr. Atul Kumar[3], Dr. Vibha Srivastav[4], Er.Prabhat Kr. Yadav[5]]
DEPARTMENT OF INFORMATION TECHNOLOGY SHRI RAMSWAROOP MEMORIAL COLLEGE OF ENGINEERING AND MANAGEMENT LUCKNOW, UTTAR PRADESH
[mshivangi2001@gmail.com[1] , kanchanmaurya7736@gmail.com[2] , atulverma16@gmail.com[3],
vibhasri2@gmail.com[4], yadav.prabhat@gmail.com[5]]

Abstract-Personal Digital Assistant now- a-days use many multidisciplinary technologies ranging from face Recognition, Signal Processing, speech recognition Natural Language Processing implementing to unite statistical framework. Such systems find a wide area of applications in areas like signal processing problems and many more. The objective of this paper is to present the concepts about Speech Recognition Systems starting from the evolution to the advancements that have now been adapted to the Speech Recognition Systems to make them more powerfully built and precise. This paper has the detailed study of the mechanism, the challenges and the tools to overcome those challenges with an ending note that would ensure that with the advancements of the technologies, this world is surely going to experience revolutionary changes in the future.

KEYWORDS: Speech recognition, personal digital assistant natural language processing, speech understanding research, Hidden Markov Model (HMM).

1. INTRODUCTION

Speech is always a useful and easiest way to communicate. Today, human beings are not communicating within themselves but they started to communicate with machines also. And that's why automatic speech recognition came into light. But this can be only possible if, machine can understand our language. Speech Recognition is a technology built to help machine. Using Speech Recognition machine can acknowledge the spoken words and phrases, and further used to generate text as well as perform the task spoken by the speaker. Speech Recognition System works using techniques popularly termed as acoustic modeling and language modeling. Performance of the machines capable of implementing Speech Recognition technology is evaluated using two parameters i.e. accuracy and speed. Speech Recognition system is something that has been dreamt about and worked for decades. A variety of software products are available for this purpose. Human beings are now able to access function commands, such as opening files, accessing menus, etc. with their voice instructions.
Speech Recognition Systems have now helped many disabled people, who are not able to write. This makes speech recognition system more popular. Now people only speak certain words to get what they need and the system that makes this possible through speech recognition.

2. EVOLUTION OF SPEECH RECOGNITION

The first speech recognition system came in 1952 and focused on numbers, not words. In 1952, Bell Laboratories designed the "Audrey" system which could recognize a single voice speaking digits aloud. Later in 1962, IBM introduced "Shoebox" which understood and responded to 16 words in English.
In 1970s, the US Department of Defense and DARPA ran Speech Understanding Research (SUR) program, one of the largest of its kind in the history of speech recognition.
In 80s the capacity of speech recognition goes from a few hundred words to several thousand words.

One of the breakthroughs came from a statistical method known as the "Hidden Markov Model (HMM)".

BellSouth introduced the voice portal (VAL) which was a dial-in interactive voice recognition system in 1990s.

Approximately 80% accuracy was achieved by the year 2001. For most of the decade there weren't a lot of advancements until Google arrived with the launch of Google VoiceSearch.

In 2011 Apple launched Siri which was similar to Google's Voice Search. Other voice recognition apps were also came into existence. And with Amazon's Alexa, Google Home we've seen consumers becoming more and more comfortable talking to machines.

With the advancements in artificial intelligence and the increasing amounts of speech data that can be easily mined, the speech recognition become feasible and powerful and it is very possible that voice becomes the next dominant interface.

3.MECHANISM OF PERSONAL DIGITALASSISTANT IN DISTRIBUTED REALTIME

The Personal Digital Assistant (PDA) is built on natural language processing (NLP).NLP is the procedure of converting speech into words, sound and text. The input through is breakdown into segments and then identifies the words to make sense. And after analysing, the voice input converted into text and perform the sensible tasks. For example, if we say "take selfie", the system will open camera and capture the picture.

If the PDA needs to say anything back, it will go through same process but in reverse manner. The mechanism of PDA is shown below:
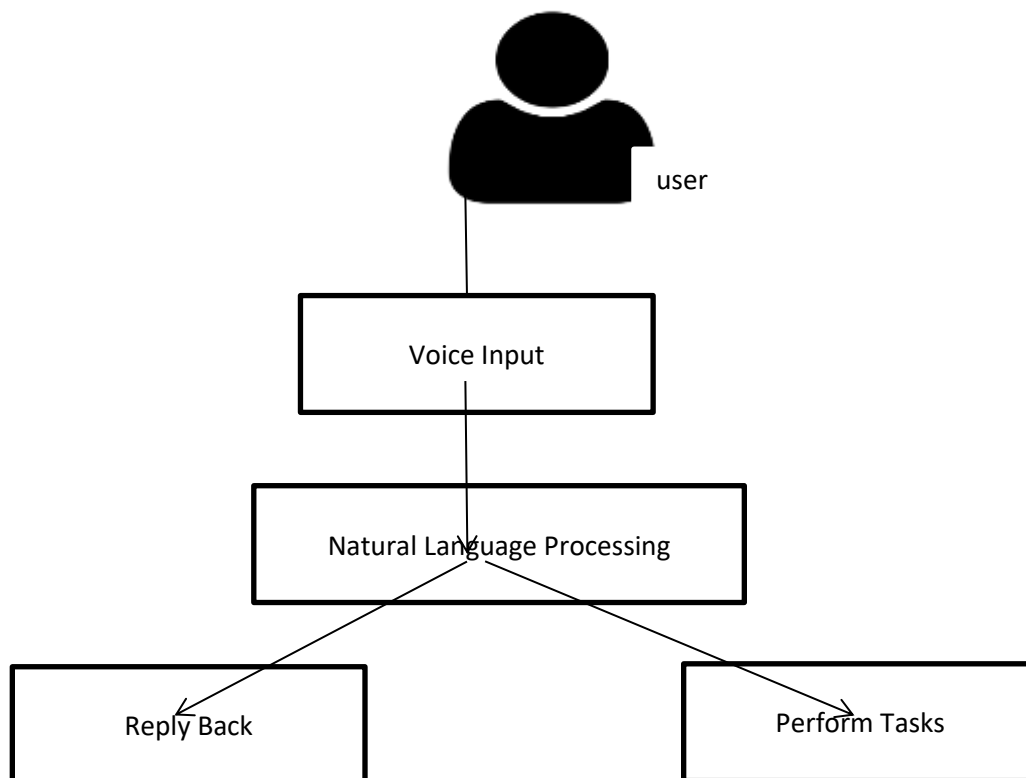
Fig: Mechanism Of PDA

4. APPLICATION OF PERSONAL DIGITALASSISTANT

Various applications of speech recognition domain have been discussed below:

1. The PDA can be used in Telephone directory enquiry without operator assistance.
2. PDA can be used in education sector. IT can useful for those Enabling students who are physically handicapped and unable to use a keyboard to enter text verbally.
3. It can also used in Computer and video games, Gambling, Precision surgery.
4. If this technology implemented in hardware system then it can used in Oven, refrigerators, dishwashers and washing machine, High performance fighter aircraft, Helicopters, Battle management, Training air traffic controllers and can also implemented in robotic technology.

These are the few areas where this technology can be used, but this technology can be used far behind this and prove to be very useful at every aspect.

6. CHALLENGES WITH PERSONAL DIGITALASSISTANT

Speech is an essential mode of communication with computers as well as human beings. Speech Recognition has a wide range of applicability in the domain of computer science, medical science, etc. Some of the key challenges faced with PDA Systems are discussed below:

The drawback affecting most of the Speech Recognition Systems is the environmental noise and its adverse effect on systems performance. It is a challenge for these systems to extract feature during conversion of speech to the on-screen text.

Speech recognition requires the computer's processor to perform a lot of heavy work.

It can be difficult for system to separate continuous segments of speedy speech signals. The speed of speech while speaking may vary from person to person or depend upon situations.

The PDA is not able to recognise punctuation marks while conversation or at the time of conversion of speech to text and vice-versa. Many different strategies are made to overcome this challenge involving the speed dictating these punctuation marks, and so on. But the better solution is yet to come.

The PDA cannot able to understand homophones. Homophones are the words that have different meanings but sounds same when pronounced (e.g., "by" and "buy", "Be" and "Bee"). In Speech Recognition Systems, it is very difficult at the word level to recognize which one is the correct intended word. The current observation say that due to this challenge, PDA cannot able to attain 100% accuracy.

7. TOOLS TO OVERCOME CHALLENGES FACED WITH PERSONAL DIGITAL ASSISTANT

A number of noise reduction techniques have been engineered to extenuate the effect of noise on systems performance.

One of the useful technique to improve performance of speech recognition system is Voice Activity Detector. It can also be used in noisy background.

8. CONCLUSION

Speech is the primary means of the communication among the people and to make the communication with machines easier, the research in the field of PDA and speech recognition is done. This project attracted many researchers with a great enthusiasm over many decades. But we can also see that it has various challenges such as problem with homophones, speed, accuracy etc. The increasing interest of

people lead their potential to find a way to overcome with these challenges.

In future we will be able to add better GUI for user convenience and try to make PDA to work offline or work on LAN.

All the research and development in the field of personal digital assistant and speech recognition will definitely bring the robust and more accurate technology in future.

REFERENCES

[1] Rouse, M. Speech Recognition, Available: https://searchcrm.techtarget.com/definition/ speech-recognition.

[2] Boyd, C. The Past, Present and future of Speech Recognition technology, Available: https://medium.com/swlh/the-past-present- andfuture-of-speech-recognition-technology- cf13c179aaf.

[3]   A short history of Speech Recognition, Available: https://sonix.ai/history-of-speech- recognition.

[4]   van der Valde, N. Speech Recognitiontechnology   overview, Available:https://www.globalme.net/blog/the-present- future-of-speechrecognition.

[5]Deng, L., Hinton, G., Kingsbury, B. 2013. New type of Deep Neural Network learning for speech recognition and related

applications: an overview, IEEE International Conference on Acoustics, Speech and Signal Processing (13886524), ISBN-(978-1-4799-

0356-6), 26-31 May 2013.

[6]Chadha, N., Gangwar, R.C., Bedi, R. 2015. Current Challenges and Applications of Speech Recognition Process using Natural Language Processing: A Survey, International Journal of Computer Applications (0975- 8887), 131(11), 28-31.

[7] Petkar, H. 2016. A review of Challenges in Automatic Speech Recognition", International Journal of Computer Applications (0975- 8887), 151(3), 23-29.

[8]   Ramírez, J., Górriz, J.M., Segura, J.C. 2007. Voice Activity Detection, Fundamentals and Speech Recognition Systems Robustness, University of Granada, Spain.

[9]   Hirsch, H.G., Pearce, D. 2000. The Aurora Experimental framework for the performance evaluation of Speech Recognition System under noisy conditions, ASR-2000 Automatic Speech Recognition: Challenges for the new Millennium Paris, France, 181-188.

[10] https://towardsdatascience.com/how- amazon-alexa-works-your-guide-to-natural- language-processing-ai-7506004709d3

[11]https://www.techtarget.com/searchcustomerexperience/definition/speech-recognition

[12]https://www.researchgate.net/publication/337155654_A_Study_on_Automatic_Speech_Recognition

[13] https://www.ijert.org/research/speech-recognition-IJERTCONV8IS04018.pdf