Mailing System using Verbal Communication

Riya Shetty
Computer Engineering
Shah and Anchor Kutchhi Engineering College
Mumbai, India

Neha Gawde
Computer Engineering
Shah and Anchor Kutchhi Engineering College
Mumbai, India

Archita Nawale
Computer Engineering
Shah and Anchor Kutchhi Engineering College
Mumbai, India

Shweta Patil
Assistant Professor
Computer Engineering
Shah and Anchor Kutchhi Engineering College
Mumbai, India

Abstract— In today's world communication has become so easy due to the integration of communication technologies with the Internet. However, the visually challenged people find it very difficult to utilize this technology because using them requires visual perception. Even though advancements have been implemented to help them use the computers efficiently, no naïve user who is visually challenged can use this technology as efficiently as a normal naïve user can do. This project aims at developing an email system that will help even a naïve visually impaired person to use services for communication without previous training. The system will not let the user make use of keyboard. Also, this system can be used by any normal person like the one who is not able to read. The system is completely based on interactive voice response which will make it user-friendly and efficient to use.

Keywords—STT, TTS, NLP, API.

I. INTRODUCTION

The discovery of Internet has made things easier and convenient for all. Internet is one of the mostly used means of communication. There are various tools available like e- mail, Facebook, Skype, WhatsApp etc. which make communication easier over the Internet. Among all the available methods, e-mail is the commonly used, especially in the business world. E-mails being cheaper than the traditional telephone communication, is used by all groups of people to share information and ideas over long distances. They provide functionalities like create, store, organize documents as well as privacy is guaranteed since access to an account is restricted. Above all this, one important requirement is to see what is written on the screen. The currently available systems like screen readers ASR and TTS are an advantage to the visually impaired people but they do not prove to be completely efficient.

1) The development of ASR systems is still in its initial stages. Its performance may lower down due to changing environments.

2) The available screen readers are language and platform dependent. One particular system developed is not applicable in other environments.

3) The available systems are not always free to access and require   specific configuration requirement.

4) Moreover, the available systems are not applicable for mobile environments, which becomes a major drawback since in the current scenario, mobile devices are more popular than desktop PCs.

II. LITERATURE SURVEY

In today's world mailing system is very convenient way of communication as well as sharing information/documents etc. Voice based emailing system uses TTS, STT modules to covert text to speech, speech to text respectively by using this even visually impaired people can easily use emailing system. These are some of the technical literature in engineering and technology where people have tried to implement similar kind of systems with their shortcomings with respect to our application.

[1] Authors Tirthankar Dasgupta, Aakash Anuj, Manjira Sinha, Ritwika Ghose, Anupam Basu created an emailing system that is designed to be used by a visually challenged person as well as by sighted user. Using this system, we can read mail or compose mail. The compose module system will provide user with two options that are 1) Type mail 2) Record a voice message. For recording message user need to press mouse left button anywhere on screen and to stop recording he needs to release the button. Once the recording is over the system will ask for recipient mail id. At the end user either press the "send mail" or middle click on the mouse to send the mail. To access the GUI different mouse-click operations have to perform for example Left, double to compose mail, left, triple to cancel the mail etc. The system allows the person to record the voice and instead of converting speech to text the system directly sense the recorded voice message to the recipient's mail address as an attachment. Drawback of this system is user needs to remember which mouse click does what action.

[2] Authors *Yogita H. Ghadage, Sushama D. Shelke* created a speech to text for multilingual languages representing the following: This project presents a multilingual speech to text conversion system. The system operation is divided into 2 phases- training and testing. In the training phase, the speech utterances of each sentence are recorded. The speech signal is preprocessed and segmented into words. For each word, acoustic features are extracted using MFCC method. Such features for each word forming feature vector is stored for reference. In the testing phase, the speech utterance to be tested is preprocessed, segmented into words and features are extracted for each word. These features are compared with the reference feature vector stored during the training phase. This is done by using a combination of SVM and Minimum Distance Classifier. The word having a minimum difference is given as a recognized word.

[3] Authors *Min-Yuh Day, Chao-Yu Chen* created an automatic text summarization using artificial intelligence the main objectives of this study include:

1. Use artificial intelligence technologies, which are including statistical method, Machine learning and Deep learning to generate candidate titles, and compare the accuracy.
2. Compare the accuracy of different deep learning models. In this, they use a method of evolution approach which was proposed by Lin which states Rogue methods as ROUGE-L uses the Longest Common Subsequence (LCS) to calculate the similarity score.

ROUGE-W is used to improve the disadvantage of ROUGE-L. ROUGE-S uses skip-bigram to generate pairs of words in their sentence order, allowing for arbitrary gaps and calculate the percentage of matched pairs between candidate summary and reference summaries as a similarity. ROUGE-SU combines skip-bigram and unigram to solve the potential problem for ROUGE-S Proposed system uses "Systems Development Research Methodology" from information system research field as their research methodology. Raw data were obtained from the Web of Science Core Collection database. Filtering special characters and convert encode. Essay titles and essay abstracts were extracted, filtered some special characters, convert encoded and converted into the format of "title-abstract" pair.

[4] In this, Khosrow Kaikhah created an automatic text conversion model using neural networks. System is using artificial neural networks to produce summary of a news articles. Corpus of articles act as training data. The neural network is then modified, through feature fusion, and it produces a summary of highly ranked sentences from the article. By using feature fusion technique, the network determines the importance (and unimportance) of various features that calculate the summary-worthiness of each sentence. In this paper, they have used summarization based on extraction approach which involves selecting a number of important sentences from the source text. The system of mainly three phases: training the neural network, feature fusion, and sentence selection.

III. OVERVIEW OF PROPOSED SYSTEM

A. Problem Statement

To create mobile application which is used to perform all the Gmail activities using voice-based commands. The application would provide all the instructions about how to use the application through voice. The user can speak the message which he/she wants to send and it will be converted into text. For the incoming mails, application will read out the mail content for user. This application will also provide the facility for file attachment. User will be provided with the summary of the files which are present in his system. Hence by using this summary user can take decision about file attachment. The user need not remember the file name all the time.

B. Proposed System Framework
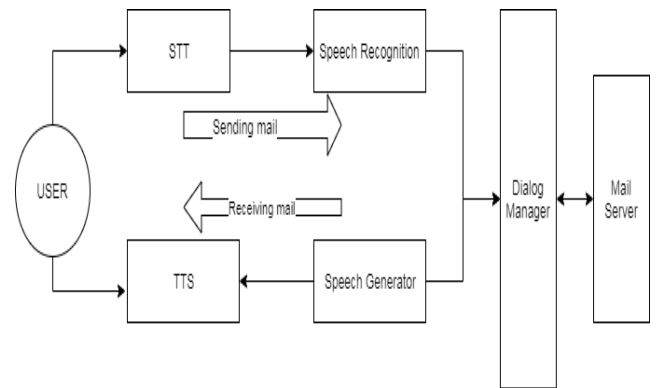
C. System Flow Diagram
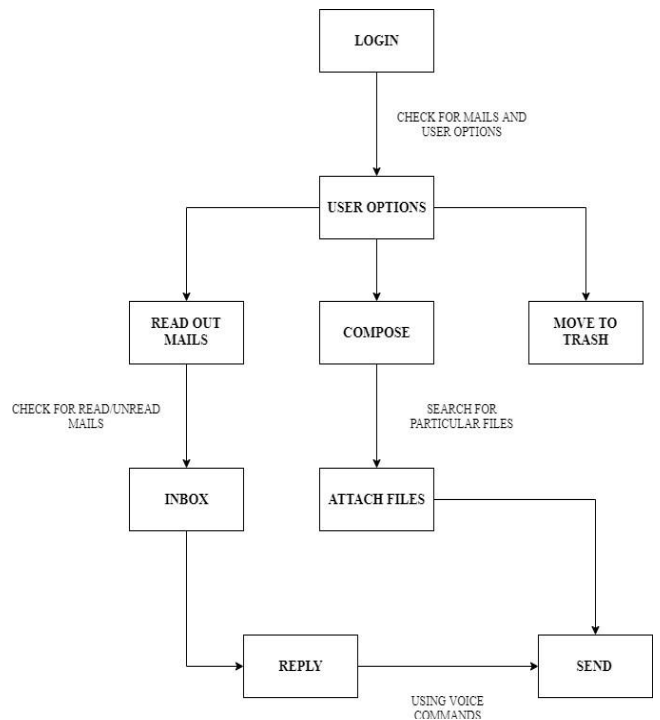
D. Methodology

Proposed system focuses on improving the emailing techniques for the blind by providing them with resources that allow them to speak.

1. Read out the mails.
2. Compose/ Attach files

This model comprises of various techniques like TTS which is used to read the mails and summarization techniques which is used to give a description about the files the user wants to attach as in various formats and then STT which is used by the user to compose the mails.

1. The basic design we plan on is to give a login page to the user where the user can login using his voice which can then be used to linking of the mail account with the app.
2. Once the user has linked the account the user can read

his emails by a converter that is using TTS or can compose emails and attach files using the summarization methods and STT methods.

3. No use of keyboards/mouse enables our system to be

user-friendly where the user only uses his voice as a tool to access his emails.

Text Summarization: TF-IDF method

This method helps in identifying how important a word is in a document. The TF_-IDF value increases proportionally to the number of times a word appears in a document and is balanced by the recurrence of word in the corpus, which modifies for the fact that a few words appear more frequently. Term Frequency (TF) measures how frequently a word occurs in the text. This returns a dictionary of words with their respective TF scores. The term TF is calculated by:

$$TF = \frac{(Frequency\ of\ a\ given\ word\ in\ the\ document)}{(Total\ number\ of\ words\ in\ the\ document)}$$

Inverse Document Frequency (IDF) is an important factor which is used to assign importance to the word. While calculating TF all words are assigned equal importance but there are certain stop words such as 'as', 'that', 'is', etc. which occur frequently but do not give much valuable information. Hence, it is necessary to weigh down certain terms but scale up some rare meaningful words. IDF for all words in an article is calculated using:

IDF = log ((Total number of documents) /

(Number of documents with the given word in it))

These values are aggregated to give a final TF-IDF score for every word. Higher the TF-IDF value, higher is the significance of the word in the document. Both the values are combined by the following formula:

TF-IDF (given word) = TF (given word) * IDF (given word)

E. Algorithm

STEP 1: Open the application.
STEP 2: Authentication is performed using Gmail API.
STEP 3: Select the option to be executed – Read, Compose.

If option is Read, go to STEP 9
Else go to STEP 4.
STEP 4: Compose email
Enter receiver's mail address, subject, body using voice commands.
STEP 5: If user wants to attach file go to STEP 6
Else go to STEP 8.
STEP 6: Control will go to file manager.
Search for the file name user wishes to attach.
STEP 7: using TF-IDF method, generate summary of different files.
TF-IDF ()
STEP 8: Send mail to the appropriate mail address
STEP 9: Control will go to the inbox of user.
The received mails will be read out.
STEP 10: Exit the application.

TF-IDF Algorithm:
STEP 1: Count the term frequency (TF) using the formula
TF = (Frequency of a given word in the document) /
(Total number of words in the document)
STEP 2: Count the term Inverse Document Frequency (IDF)
IDF = log ((Total number of documents) /

(Number of documents with the given word in it))

STEP 3: These values are aggregated to give a final TF-IDF score for every word.

TF-IDF (given word) = TF (given word) * IDF (given word)

IV. SUMMARY

The system allows a visually impaired person to send voice- based e-Mails messages. This will reduce the extensive effort taken by the person to remember the manner in which the characters need to be typed. Further developments in proposed system will also lead to developers building more innovative applications which will be beneficial for the visually challenged people, who also deserve an equal standard in society. We will be using GOOGLE API for speech to text conversion and vice-versa. There will be different options available in the developed application which allows the user to compose, read and delete mails or attach files from their phone's storage. For the summarization purpose system will use TF-IDF method. The user can choose the file using the summery of available files provided by the system. Proposed system is suggested for betterment of the society. It will help the blind people to be a part of growing digital India by using the Internet. Proposed system overcomes some drawbacks that were earlier faced by the blind people in accessing emails. We have eliminated the concept of using keyboard shortcuts along with screen readers. Any user who does not know the location of keys on the keyboard need not worry as keyboard usage is eliminated. The user only needs to follow the instructions given by the IVR accordingly to get the respective services offered. Other than this the user might need to feed in information through voice inputs when specified

V. SCOPE

The scope of proposed system is to allow the users for real time composition of mails using voice commands which lets them to avoid the use of keyboard and mouse which can also act as an advantage for the specially aided people (Blind people).In this we intend to reduce the use of keyboards while allowing the voice commands to take over the system, where the user gets the privilege to hear the mails rather than reading, due to this the user can hear mails even while driving. Attachment of files while sending a mail can be hectic sometimes when using a general mail system, but the privilege that users get while using proposed mail system is that the attachment of files is done using the method of text summarization for identifying similarly named files and voice commands to search for a particular file.

become masters of it. So, we would like to take this opportunity to thank all those who have helped us in the whole journey.

REFERENCES

[1] Tirthankar Dasgupta, A. Anuj, M. Sinha, R. Ghose, and A. Basu, "VoiceMail architecture in desktop and mobile devices for the Blind people," *2012 4th International Conference on Intelligent Human Computer Interaction (IHCI)*, 2012.

[2] Y. H. Ghadage and S. D. Shelke, "Speech to text conversion for multi-lingual languages," *2016 International Conference on Communication and Signal Processing (ICCSP)*, 2016.

[3] M.-Y. Day and C.-Y. Chen, "Artificial Intelligence for Automatic Text Summarization," *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, 2018.

[4] K. Kaikhah, "Automatic text summarization with neural networks," *2004 2nd International IEEE Conference on Intelligent Systems. Proceedings (IEEE Cat. No.04EX791)*.