

Voice based Email System for Visually Impaired People

Sonali Malap¹ Vaishali Shirke² Mrunali Chalke³ Rohan Jain⁴ Sonali Pakhmode⁵

^{1,2,3,4}Student ⁵Professor
^{1,2,3,4,5}Department of Information Technology
^{1,2,3,4,5}PVPPCOE, India

Abstract— A decisive role can be played by communication in promoting human development in today's new climate of social change and of all the different communication technologies available, email is considered as one of the most pervasive forms of communication. However all these technologies can be of no use to the people who are visually impaired as all activities that can be performed on the computer are based on visual perception. 285 million people are estimated to be visually impaired worldwide: 39 million are blind and 246 have low vision[1]. Therefore it becomes necessary to make available internet facilities for communication for them also. This paper aims to develop voice based email system for visually challenged people to use email system as efficiently as normal user. The system will not let the user to make the use of keyboard instead will work on speech to text conversion and vice-versa. The complete system is based on Text-to-speech and Speech-to-text API's.

Key words: Visually Challenged People, Speech to Text Converter, Text to Speech API's

I. INTRODUCTION

In today's information age, computer has become an integral part of our daily life. Much of the communication takes place through the internet. Internet is considered as a major storage of information. No single work can be completed without the help of Internet. And out of all the available communication technologies, email is one of the most common and popular medium of communication especially in the business world. The number of worldwide email accounts continues to grow from over 4.1 billion accounts in 2014 to over 5.2 billion accounts by the end of 2018. The total number of worldwide email users, including both business and consumer users, is also increasing from over 2.5 billion in 2014 to over 2.8 billion in 2018[2].

In order to access internet and computer handling, one should know what is written on the screen. For this he/she should have visual capabilities. This means that a visually challenged person cannot take the benefits of the facilities provided by the internet. This makes internet a completely useless technology for the visually impaired and illiterate people. Even the systems that are available currently like the screen readers and ASR do not provide full efficiency to the blind people so as to use the internet. Thus, it turned out to be the moral responsibility to give something back to society and use our technical knowledge to build an application for their benefit.

As we are developing voice based email system, user need not to have basic information regarding keyboard shortcuts or location of keys. So, it is very easy for any type of user to use the system, as the system itself will be prompting them. As a result, system will reduce the cognitive workload of user and help visually impaired people to access email in a hassle free manner.

II. EXISTING SYSTEM

The most common mail services that are available today are of no use to visually impaired people. This is because these systems do not provide any audio feedback. As they cannot visualize what is present on the screen, it becomes very difficult for them to perform required operations. The systems available now a day's uses screen readers which read information displayed on desktop in sequential manner or it prints information on Braille printer. ASR (Automatic speech recognizer) is not very much useful for small scale application for E-mail. Although these technologies are being improved continuously, some major problems still persist which make them unusable for visually impaired people. These systems have following drawbacks:

- 1) With the help of screen readers it is difficult for visually challenged person to access E-mail system and computer operating easily. Because if there is noise or some other sound in the room (e.g. the television or a kettle boiling), the number of errors will increase[3].
- 2) Speech Recognition works best if the microphone is close to the user (e.g. in a phone, or if the user is wearing a microphone). More distant microphones (e.g. on a table or wall) will tend to increase the number of errors[3].
- 3) Visual layout: Also, screen readers cannot survey the entirety of a web page as a visual user might do. A visual user can look at a web page and quickly realize how the page is organized, then zero in on the most important content. A screen reader is not able to do this. It reads in a linear fashion, one word at a time. It cannot always intelligently skip over extraneous content, such as advertisements or navigation bars[4].
- 4) These available systems require use of keyboard which is very difficult for visually challenged people to recognize and remember characters of keyboard. This problem can be solved by using Braille keyboards, but these Braille keyboards are very costly.

To overcome all these drawbacks of existing systems, we come with this project which is based on voice inputs.

III. PROPOSED SYSTEM

The key opinion kept into consideration while developing the proposed system was accessibility. If such applications will be developed then that will be used efficiently by anyone whether he is able or disable. The most common mail services that we are using in our day to day life cannot be used by visually impaired people. This is because the common mail services do not provide any facility so that the person in front can hear out the content of the screen and they can perform their operations easily. For a visually impaired person using a computer for the first time is not that convenient as it is for a normal user though it is user

friendly. The current systems focus more on user friendliness than accessibility. Thus the system which we are developing will be helpful to all types of people including normal people, visually impaired people as well as illiterate people. Component diagram for proposed system is as shown in Fig.1.

- 1) STT - Accepts speech from the user and produces text.
- 2) Language Understanding Component - Extracts semantics from a text string by using a pre-specified grammar.
- 3) Context Interpreter - Enhances the semantics from the Language Understanding Module by obtaining context information from a dialog history For example; the Context Interpreter may replace a pronoun by a noun to which the pronoun referred.
- 4) Dialog Manager - Prompts the user for input, makes sense of the input, and determines what to do next according to instructions in a dialog script specified.
- 5) Language Generator - Accepts text from the Dialog manager and prepares it to the user as spoken voice via a text-to-speech synthesizer (TTS).
- 6) Text-to-Speech Synthesizer (TTS) - Accepts text from the Language Generator and produces acoustic signals which the user hears as a human-like voice .

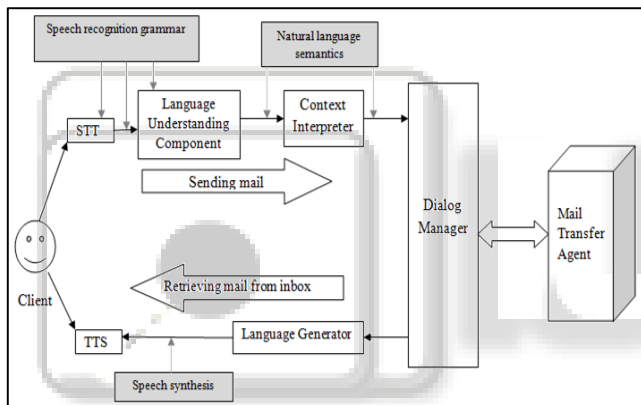


Fig. 1: Component Diagram

IV. METHODOLOGY

The Software Development Life Cycle includes models such as Waterfall Model, Prototype Model, and Object-oriented Model, etc. for developing the correct software. The Waterfall Model is the earliest method of structured system development.

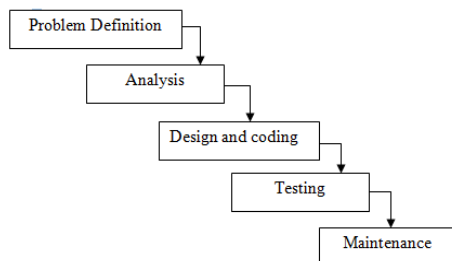


Fig. 2: Waterfall Model

The software model used by our system is the waterfall model. Waterfall model is a systematic and sequential approach to the software development. This includes system engineering and modeling which establishes requirements for all the system elements and allocating some subset of these requirements to software. System engineering and analysis encompass requirement gathering

at the system level with small amount of top-level design. Analysis information engineering encompass requirement gathering at the strategic business level and at the business area level.

A. Problem definition:

This is the very first stage to develop any project. It actually defines the aim and the concept of the project. The aim of “Voice based email system” is to envisage providing effective accessing capabilities to people having reading disabilities and visual impairments.

B. Analysis:

Existing systems such as screen readers and ASR are analyzed. Care is taken to ensure that all the drawbacks of the existing browsers were overcome.

C. Design and Coding:

It is necessary to get the logical flow of the software.

D. Testing:

Testing will involve working of individual module and after integration.

E. Maintenance:

Maintenance of the system to check if it is accepting all speech commands as expected and dynamic additions to the grammar are not causing any problems.

V. DESIGN AND IMPLEMENTATION

We are developing Voice based email system currently. ER (Entity-Relationship) diagram and activity diagram are shown in Fig.3 and Fig.4 respectively.

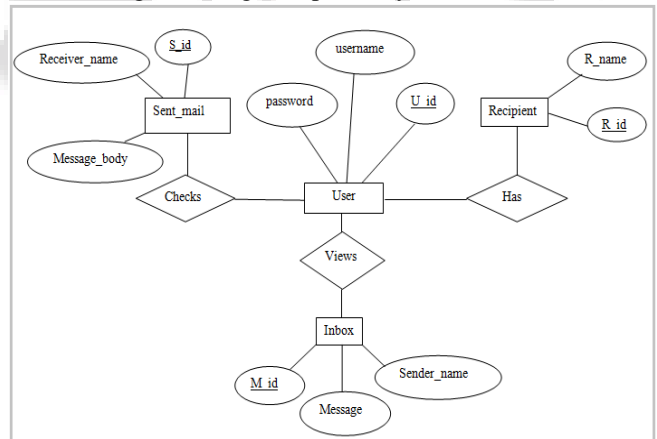


Fig. 3: ER diagram

And modules of this system are as follows:

A. Registration:

This is the first module of the system. The user who wants to take benefits of our system, he/she needs to register themselves. In this module, system will be prompting what details needs to be entered. Accordingly, user will provide details through voice inputs. Before submitting details, system will confirm the information by prompting. After registration, all the information along with username and password will be stored in database.

B. Login:

Once the registration is done, user can login to system by providing username and password for accessing his/her

account. This module is used for authentication. It will accept username and password in speech format and convert it to text. Then this text will be used by system to decide whether that user is valid or not. If he/she is authorized user then system will redirect that user to next GUI.

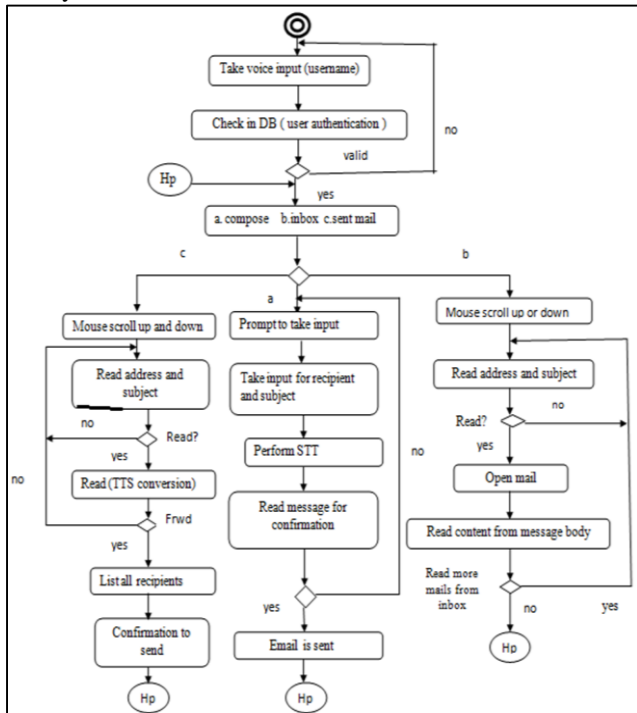


Fig. 4: Activity diagram

After successful login, system prompts for action to be performed such as:

- 1) Compose mail
- 2) Inbox
- 3) Sent mail

C. Compose Mail:

In this module, user can create mail he/she wants to send. It works as follows:

- 1) Then system will ask for recipient's name/s, subject and message body.
- 2) User will provide this information through speech and this speech is converted into text by the system.
- 3) After that system will read recipient name, subject, and message body and ask for confirmation.
- 4) Once the confirmation is done mail will be sent to respective recipient/s.

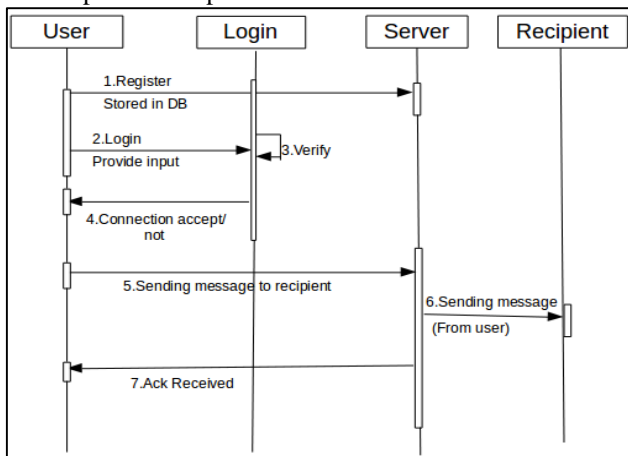


Fig. 5: Sequence Diagram for Sending Mail

D. Inbox:

This option helps user to view all mails received to his/her account. After selecting this option, most recent mails get loaded. Then system will prompt recipient's name and subject of each mail, if user wants to listen to mail then user has to perform operation specified by prompt. In order to navigate to next mails he has to give appropriate voice command. The user also has option to delete received mails.

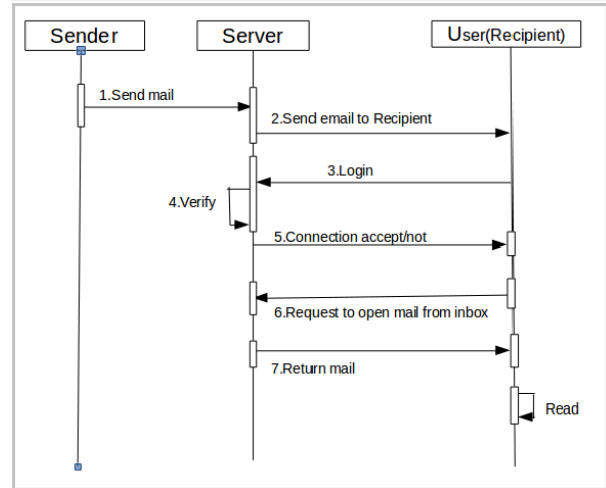


Fig. 6: Sequence diagram for reading mail from inbox

E. Sent Mail:

This option keeps track of mails sent by the user. In order to access the sent mails user needs to perform actions provided by the prompt to navigate between mails. This will help the user in efficiently extracting and forwarding the required mail.

VI. FUTURE SCOPE

The system we are developing will be working only on desktops. As use of mobile phones is increasing day-by-day, there is a need to include this facility as an application in mobile phones also.

Also security features can be implemented during login phase to make the system more secure.

VII. CONCLUSION

Voice based email system helps visually challenged people to access email services efficiently. It has been observed that nearly about 60% total blind population across the world is present in India[1]. This system overcomes difficulties faced by visually impaired people as well as illiterate people. This will reduce the drawbacks of existing system such as software load of using screen readers and Automatic Speech Recognizer(ASR). The system will be guiding the user what needs to be performed for obtaining desired results by prompting. Hence this reduces the user's load of remembering keyboard shortcuts and location of keys. The user needs to follow the instructions given by the system.

REFERENCES

- [1] The WHO website. [Online]. Available: <http://www.who.int/mediacentre/factsheets/fs282/en/>.
- [2] The Radicati website. [Online]. Available:

- <http://www.radicati.com/wp/wp-content/uploads/2014/01/Email-Statistics-Report-2014-2018-Executive-Summary.pdf>.
- [3] http://www.match-project.org.uk/resources/tutorial/Speech_Language/Speech_Recognition/Rec_6.html
- [4] <http://webaim.org/articles/visual/blind>
- [5] <https://developers.google.com/gmail/api/?hl=en>.
- [6] T. Dasgupta and A. Basu. A speech enabled indian language text to braille transliteration system. In Information and Communication Technologies and Development (ICTD), 2009 International Conference on, pages 201–211. IEEE, 2009
- [7] Jagtap Nilesh, Pawan Alai, Chavhan Swapnil and Bendre M.R. “Voice Based System in Desktop and Mobile Devices for Blind People”. In International Journal of Emerging Technology and Advanced Engineering (IJETA), 2014 on Pages 404-407(Volume 4, issue 2).
- [8] G. Shoba, G. Anusha, V. Jeevitha, R. Shanmathi. “AN Interactive Email for Visually Impaired”. In International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), 2014 on Pages 5089-5092.(Volume 3, Issue 1).

