Software Requirements Specification Document

<21, Vivek Ghaisas, Shubham Rathi, Lasya Venneti, Vamshi Palabatla>

Brief problem statement

Our project aims to develop call center enterprise software application that helps analyse the caller's spontaneous reaction (angry, upset, dull) based on speech emotion recognition. Our product is a speech analytics engine that allows call centers to analyze real time audio from calls and educate the customer care representative about the instinctive reaction of the callee. We are constructing an interactive graphical representation of the audio data with the emotional analysis superposed on it, periodically communicating with a server to collect the audio data which generates some basic statistics and sends them to the server. Our product is easy to use, with a friendly user interface and simple integration to existing call center platforms. The UI includes emospeech representation of the audio, ability to span and play the audio and some other features which the client is yet to finalise.

System requirements

Our work involves creating a functional and interactive UI that gets it's data from two input sources: Namely the **Call centre server** and **AksharSpeech back-end.** A desktop application that provides the same has to be created.

WxPython and matplotlib are to be used to plot the initial WAVE file. Libraries used: numpy and scipy.

We have chosen Python to be our primary programming language over other visual plotting languages like processing as the requirement here is functional and not aesthetic. Python is mature, feature rich scientific and has 3rd party libraries for plotting. Common operations such as pulling files from a server and generating statistics are very easy.
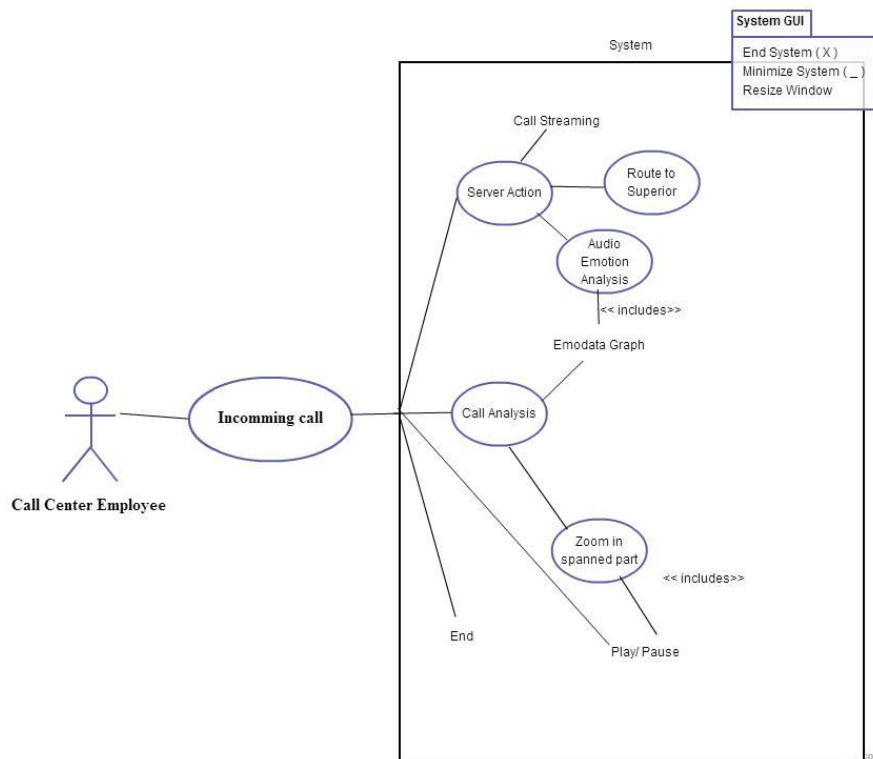
Users profile

The end users for this tool would be call center employees who would be using this tool to differentiate the speech between two basic emotions (for now), Angry and Cool, as the employee might not be capable of handling hot headed customers. In that case, the call can be elevated to a higher official. The employee will be able to easily understand how the customer is reacting to the call and do the necessary. Features such as listening to specific portions of the call again, playing and pausing will be included for the convenience of the call center's employees.

Feature requirements (described using use cases)

No.	Use Case Name	Description	Release
1.	server communication	The server will, at regular intervals, provide short audio files to the application along with the 'emodata'. It will also accept generated statistical data.	R2
2.	Plot audio and associated analysis	The audio file and associated data is rendered as an interactive plot.	R1
3.	zoom plot	User can zoom in/out of the graph, or fill the screen with the selected part.	R1
4.	span plot	The user can span the zoomed graph	R1
5.	route to superior	when the user's angry levels are increasing continuously the call is automatically directed to the manager of the call centre	R2

Use case diagram

Use case description

Use Case Number:	UC_1
Use Case Name:	EmoData
Overview:	End to End System Interaction as mentioned in the brief.
Actors:	Call Center Employee, Server
Pre condition:	Incoming Call and Emodata
Flow:	Main (success) Flow: 1. Call received by Server 2. After a buffer of 5 secs, server branches the call to the Emodata Application 3. Emodata Application displays Emotion Analysis and the audio graph. 4. Unless the Anger frequency reaches beyond the threshold level, the call continues 5. The User can choose to span and zoom into the audio and listen again. If he does not do so, the normal flow continues 6. At the end of the call, the statistics generated is shown to the Call Center employee and also sent to the server 7. The system is prepared for the next call.
	Alternate Flow 1: 1. Call received by Server 2. After a buffer of 5 secs, server branches the call to the Emodata Application 3. Emodata Application displays Emotion Analysis and the audio graph. 4. The Anger frequency reaches beyond the threshold level. 5. The call is routed to the superior 6. The system is prepared for the next call. Alternate Flow 2:

		1. Call received by Server
		2. After a buffer of 5 secs, server branches the call to the Emodata Application
		3. Unless the Anger frequency reaches beyond the threshold level, the call continues.
		4. The User can choose to span and zoom into the audio and listen again.
		5. The call is put on hold and the audio is played after which the call resumes
		6. At the end of the call, the statistics generated is shown to the Call Center employee and also sent to the server
		7. The system is prepared for the next call.
Post Condition:		Last Call Terminated.