ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

IVRS FOR COLLEGE AUTOMATION

Santosh A. Kulkarni[1], Dr. A.R.Karwankar[2]

Dept. of Electronics and Telecommunication Engineering, Govt. College of Engineering, Aurangabad, (M.S.), India[1,2]

ABSTRACT—***The interactive voice response (ivr) system serves as a bridge between people & computer by connecting the telephone network with instructions. The telephone user can access the information from anywhere at anytime simply by dialing a specified number and following an automated instruction when a connection has been established. The IVR system uses pre-recorded or computer generated voice responses to provide information in response to an input from a telephone caller. The input may be given by means of touch-tone or Dual Tone Multi-Frequency (DTMF) signal, which is generated when a caller presses a key of his/her telephone set, and the sequence of messages to be played is determined dynamically according to an internal menu structure (maintained within the IVR application program) and the user input. The IVR System which will be designed to provide an ideal platform for the operation of start-ups and existing small concern . It will be a highly economical & efficient way to replace the Dialogic card which is very costly and requires a high maintenance and regular up gradation. The IVRS system which will be designed will consist of simple components like microcontroller and some basic application chips interfaced to a PC which will have small software running in the backend while the other jobs are performed on the front end.***

Keywords— Dual Tone Multi-Frequency (DTMF), Goertzel algorithm, Speech synthesis, Voice Over Internet Protocol.

I. INTRODUCTION

In telephony, Interactive Voice Response, or IVR, is a phone technology that allows a computer to detect voice and touch tones using a normal phone call. The IVR system can respond with pre-recorded or dynamically generated audio to further direct callers on how to proceed. IVR systems can be used to control almost any function where the interface can be broken down into a series of simple menu choices.Once constructed IVR systems generally scale well to handle large call volume.

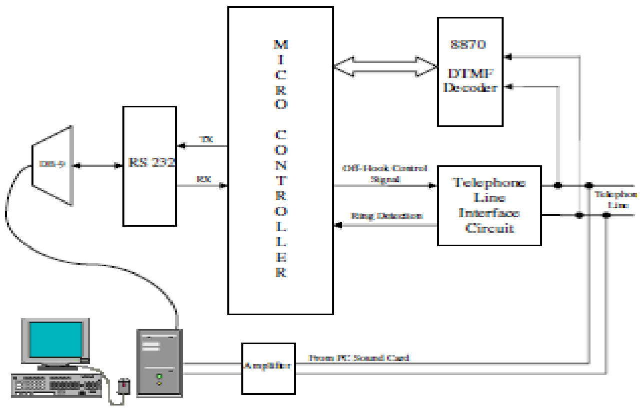
Taking advantages of IVRS we are developing the system for college automation using voice over internet protocol (VOIP).Which is described in next section.

A caller dials a telephone number that is answered by an IVR system. The IVR system executes an application which is tied to the number dialed DNIS (Dialed number information service). As part of the application, prerecorded audio files or dynamically generated Text to Speech (TTS) audio explain the options available to the caller. The caller is given the choice to select options using DTMF tones or spoken word. Speech recognition is normally used to carry out more complex transactions and simplifies the application menu structure.

Sequence followed in the IVRS service

• Caller dials the IVRS service number.
• The computer waits for ringing tones at the end of which, the connection is established.
• The connection is established by lifting the handset of telephone base from ON-HOOK condition
• Now, a pre-recorded audio greets the caller conforming that the number dialed corresponding to the particular service.
• Next, the menu is presented to the caller again in the voice form, giving him the various options to choose from.
If the information to be relayed back is confidential, then the system may even ask the dialer, to feed in a password number.
• The database is accordingly referenced and the necessary information is obtained.
• Next, the same information is put across to the user in voice.
• The caller generally given the option to :
a. Repeat whatever information was voiced to him.
b. Repeat the choices
c. Break the call by restarting ON-HOOK condition

The following figure shows the overall system block diagram of IVRS system. in which includes microcontroller,database, amplifier, DTMF decoder etc.



ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

Fig.1. System Block Diagram

II.METHODOLOGY

We are developing the college automation system using voice over internet protocol (VOIP). Which the major part of the system software design. The system software development includes the technologies Goertzel algorithm, dual-tone multi-frequency signaling (DTMF), speech synthesizer etc.

When caller dial the number then the technique used for identifying frequency components of a signal is Goertzel algorithm. That is for Dual Tone Multi-Frequency (DTMF) detection or decoding. A text-to-speech (TTS) system converts normal language text into speech. For that Speech synthesis is used.

A. Goertzel algorithm:

The Goertzel algorithm is a digital signal processing (DSP) technique for identifying frequency components of a signal.

The Goertzel algorithm implementation examines the energy of one of the two tones from an incoming signal at eight different DTMF frequencies to determine which DTMF frequency is present. To do this evaluation, the input signal is transformed to the DTMF frequencies, which are computed by the modified Goertzel algorithm. The matched filter concept is used for each DTMF frequency to determine the frequency at which the incoming signal has maximum energy. Since maximum energy corresponds to DTMF frequency, this procedure enables us to detect the DTMF frequency. It is important to choose the right algorithm for detection to save memory and computation time.

The Goertzel algorithm is the optimal choice for this application because it does not use many constants, which saves a great deal of memory space. Also, only eight DTMF frequencies need to be calculated for this application, and the Goertzel algorithm can calculate selected frequencies. This saves computation time. The DTMF frequency is transformed to a Discrete Fourier Transform (DFT) coefficient

IMPLEMENTATION:

Since the telephone industry has preset the sampling frequency to 8 kHz and the DTMF frequencies to 697, 770, 852, 941, 1209, 1336, 1477 and 1633 Hz the filter length must be large enough to find the desired value that corresponds to the DTMF frequencies. Therefore, there is a trade off to be considered between the computation burden and better resolution. For this application report, the Filter length N was chosen as 105 which is the smallest value that can fulfill DTMF detection.

The following flowchart shows the implementation of goertzel algorithm.

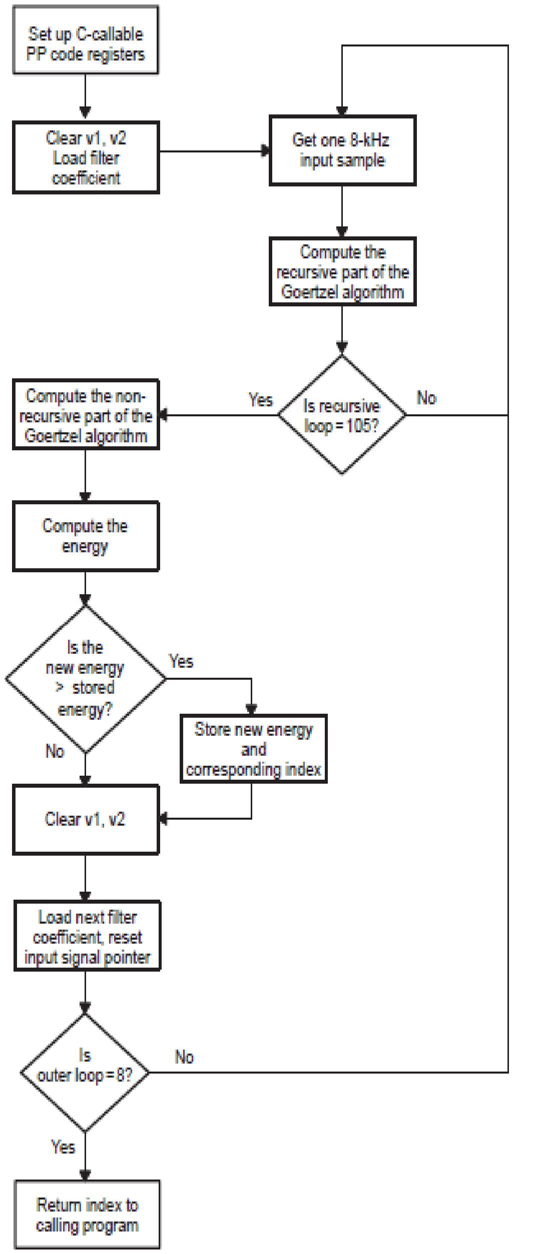


Fig.2. Flow Chart of Goertzel algorithm

B Dual Tone Multi-Frequency (DTMF):

Dual-tone multi-frequency signaling (DTMF) is used for telecommunication signaling over analog telephone lines in the voice-frequency band between telephone handsets and other communications devices and the switching center. The version of DTMF that is used in push-button telephones for tone dialing is known as Touch-Tone.

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

Prior to the development of DTMF, Automated telephone systems employed pulse dialing or loop disconnect (LD) signaling to dial numbers. It functions by rapidly disconnecting and re-connecting the calling party's telephone line, similar to flicking a light switch on and off. The repeated interruptions of the line, as the dial spins, sounds like a series of clicks. The exchange equipment interprets these dial pulses to determine the dialed number. Loop disconnect range was restricted by telegraphic distortion and other technical problems, and placing calls over longer distances required either operator assistance (operators used an earlier kind of multi-frequency dial) or the provision of subscriber trunk dialing equipment.

Fig.3. DTMF Keypad Layout.

The above figure shows DTMF keypad which is explained below.

The DTMF keypad is laid out in a 4×4 matrix, with each row representing a low frequency, and each column representing a high frequency. Pressing a single key (such as '1') will send a sinusoidal tone for each of the two frequencies (697 and 1209 hertz (Hz)). The original keypads had levers inside, so each button activated two contacts. The multiple tones are the reason for calling the system multi frequency. These tones are then decoded by the switching center to determine which key was pressed.

DTMF Keypad Frequencies (with sound clips)				
	1209 Hz	**1336 Hz**	**1477 Hz**	**1633 Hz**
697 Hz	1	2	3	A
770 Hz	4	5	6	B
852 Hz	7	8	9	C
941 Hz	*	0	#	D

The tone frequencies, as defined by the Precise Tone Plan, are selected such that harmonics and inter modulation products will not cause an unreliable signal. No frequency is a multiple of another, the difference between any two frequencies does not equal any of the frequencies, and the sum of any two frequencies does not equal any of the frequencies. The frequencies were initially designed with a ratio of 21/19, which is slightly less than a whole tone. The frequencies may not vary more than ±1.8% from their nominal frequency, or the switching center will ignore the signal. The high frequencies may be the same volume as – or louder than – the low frequencies when sent across the line. The loudness difference between the high and low frequencies can be as large as 3 decibels (dB) and is referred to as "Twist". The duration of the tone should be at least 70 ms, although in some countries and applications DTMF receivers must be able to reliably detect DTMF tones as short as 45ms

D .Speech synthesizer:

Speech synthesis is the artificial production of human speech. A computer system used for this purpose is called a speech synthesizer, and can be implemented in software or hardware. A text-to-speech (TTS) system converts normal language text into speech; other systems render symbolic linguistic representations like phonetic transcriptions into speech. Synthesized speech can be created by concatenating pieces of recorded speech that are stored in a database. Systems differ in the size of the stored speech units; a system that stores phones or diphones provides the largest output range, but may lack clarity. For specific usage domains, the storage of entire words or sentences allows for high-quality output. Alternatively, a synthesizer can incorporate a model of the vocal tract and other human voice characteristics to create a completely "Synthetic" voice output.

The quality of a speech synthesizer is judged by its similarity to the human voice and by its ability to be understood. An intelligible text-to-speech program allows people with visual impairments or reading disabilities to listen to written works on a home computer.

A text-to-speech system (or "engine") is composed of two parts:] a front-end and a back-end. The front-end has two major tasks. First, it converts raw text containing symbols like numbers and abbreviations into the equivalent of written-out words. This process is often called text normalization, pre-processing, or tokenization. The front-end then assigns phonetic transcriptions to each word, and

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

divides and marks the text into prosodic units, like phrases, clauses, and sentences. The process of assigning phonetic transcriptions to words is called text-to-phoneme or grapheme-to-phoneme conversion. Phonetic transcriptions and prosody information together make up the symbolic linguistic representation that is output by the front-end. The back-end—often referred to as the synthesizer—then converts the symbolic linguistic representation into sound. In certain systems, this part includes the computation of the target prosody (pitch contour, phoneme durations), which is then imposed on the output speech. The following diagram shows overview of typical text-to-speech (TTS). Which consist of text analysis, linguistic analysis and wave form generation etc.

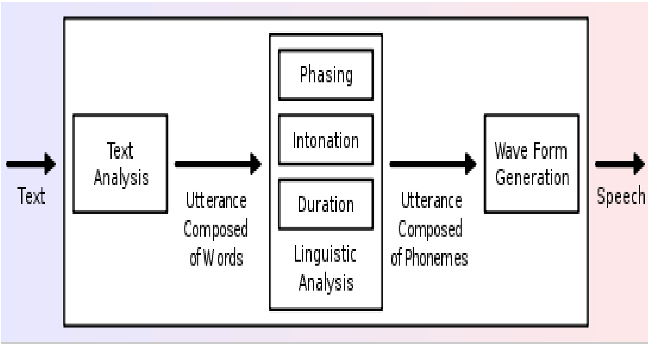


Fig.4. Overview of a typical TTS system

D. Voice over internet protocol (VOIP):

VoIP is a technology used by IP telephony as a means of transporting phone calls.commonly refers to the communication protocols, technologies, methodologies, and transmission techniques involved in the delivery of voice communications and multimedia sessions over Internet Protocol (IP) networks, such as the Internet. Other terms commonly associated with VoIP are IP telephony, Internet telephony, voice over broadband (VoBB), broadband telephony, and broadband phone. The steps involved in originating a VoIP telephone call are signaling and media channel setup, digitization of the analog voice signal, encoding, packetization, and transmission as Internet Protocol (IP) packets over a packet-switched network. Early providers of Voice over IP services offered business models (and technical solutions) that mirrored the architecture of the legacy telephone network. Second generation providers, such as Skype have built closed networks for private user bases, offering the benefit of free calls and convenience, while denying their users the ability to call out to other networks. VoIP systems employ session control protocols to control the set-up and tear-down of calls as well as audio codecs which encode speech allowing transmission over an IP network as digital audio via an audio stream. The choice of codec varies between different implementations of VoIP depending on application requirements and network bandwidth; some implementations rely on narrowband and compressed speech, while others support high fidelity stereo codecs.

The biggest single advantage VoIP has over standard telephone systems is cost. In addition, international calls using VoIP are usually very inexpensive. One other advantage, which will become much more pronounced as VoIP use climbs, calls between VoIP users are usually free. VoIP telephone systems are susceptible to attacks as are any Internet-connected devices.

IV.RESULT AND DISCUSSION

The call flows in the following manner and accordingly caller will get the information.

When caller dial number, caller listen the welcome message that is in three languages (Marathi, Hindi, English). After that callercan choose any of thes language for the information. After this caller can choose the field and then branch (Civil, Electronics, Electrical, Mechanical, Computer, IT). Then choose year (i, ii, iii, iv). After this the caller can choose either attendance or result. Then caller have to enter Roll no. And PIN no. Then system plays for attendance or result of the student. After getting information call is disconnected.

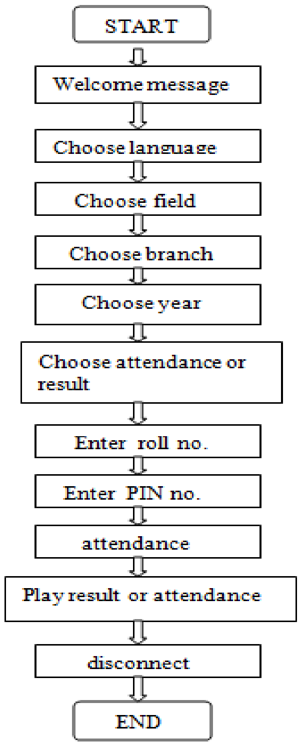


Fig.5 Flow Chart of Call

ISSN : 2278 – 1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 1, Issue 6, August 2012

V.CONCLUTION

In today's world everything needs to be done from the comfort of one's home or office. For this application is prepared in such a way that they can be easily accessed through computers. In the same way our project's aim is to provide the entire information to the user at the tip of his fingers. Due to this project the traditional manual way of handling the customer queries.

Will be handled in a more technological and automated way. This type of system performs operations similar to that of a human telephone operator. The USP of the project is its relevance to the field of telephony and its cost that will be bearable even by a small concern due to its simpler and easily available components.

REFERENCES

[1] C. Marven, "General-Purpose Tone Decoding and DTMF Detection," in Theory, Algorithms,

[2] and Implementations, Digital Signal Processing Applications with the TMS320 Family, Vol. 2,

[3] Literature number SPRA016, Texas Instruments (1990).

[4] J. G. Proakis and D. G. Manolakis, Digital Signal Processing: Principles, Algorithms, and

[5] Applications, 2nd ed., Macmillan, New York, NY (1992). .http://www.dattalo.com/technical/theory/sinewave.html.

[6] Schenker, L (1960), "Pushbutton Calling with a Two-Group Voice-Frequency Code"

[7] .ITU's recommendations for implementing DTMF services (PDF)

[8] .Allen, Jonathan; Hunnicutt, M. Sharon; Klatt, Dennis (1987). From Text to Speech: The MITalk system. Cambridge University Press. ISBN 0-521-30641-8.

[9] 8. Rubin, P.; Baer, T.; Mermelstein, P. (1981). "An articulatory synthesizer for perceptual research". Journal of the Acoustical Society of America70 (2): 321–328. doi:10.1121/1.386780.

[10] Van Santen, J. (April 1994). "Assignment of segmental duration in text-to-speech synthesis". Computer Speech & Language **8** (2): 95–128.doi:10.1006/csla.1994.1005.

Biography

Santosh A. Kulkarni Woking in the Department of Electronics and Telecommunication Engineering Govt College of Engineering, Aurangabad, (M.S.), India

Dr. A.R.Karwankar is Working in the Department of Electronics and Telecommunication Engineering Govt College of Engineering, Aurangabad, (M.S.), India