Dynamically Automated Interactive Voice Response System for Smart city Surveillance

Saqquaf S S M, Sarvesh Araballi, Bhagyalakshmi P, Shruti S Mahadeek, Varsha B M

***Abstract*— Interactive Voice Response System (IVRS) is an automated system which is multi-lingual and makes use of text-to-speech (TTS) technology. This paper describes the implementation of an Interactive Voice Response System personalized especially for students to obtain any information desired by them in a reliable and timely manner by just placing a normal call, thus being a building block of a smart city. It works on a 24/7 basis and is very user friendly. The IVRS system comprises of ring detector hardware, dual tone multi frequency decoder (DTMFD), microcontroller, serial interface unit, computer and cell phone. It includes a ring detector hardware which identifies the caller and connects the identified caller to a computer. When a caller presses a key from their cell phone the input is produced in the form of dual tone multi frequency signal. The computer generates voice response dynamically based on the key pressed by the caller. A "Text to Speech Converter" converts the database into a voice format as per the response given by the caller.**

Keywords—interactive voice response, text to speech, smart city, ring detector, dual tone multi frequency decoder, microcontroller, SIP account.

I. INTRODUCTION

Interactive Voice Response System (IVRS) in vogue today relies on cell phone communication backbone that allows computers identify voice and touch tones while establishing an ordinary phone call to a registered number [1]. The IVR system responds with a context specific pre-recorded message to direct the caller on how to proceed further. The interface is broken down into a series of simple menu driven options of which the caller selects any one as desired. This system stimulates a response to the caller asking him to respond either by voice or by pressing a touchtone key existing on the phone and based on the response; it supplies the caller with requisite information. This information is either played as a pre-recorded script or as a pre-recorded message, which can be sent to the caller's phone. This IVR system can be designed to handle large number of callers, once it is properly fabricated. This can be used as a small part in the making of a smart city. A smart city is a vision to develop urban areas which integrates in handling several information and communication technology solutions to manage a city's assets such as, schools libraries, water supply networks, hospitals, transportation systems etc. The object of constructing a smart city is to improve the quality of life by means of advanced technology and to improvise the productivity of services to cater the needs of various applications. It helps to enhance the essence of life.

The proposed system can be used in schools and other educational institutions to acquire information in a timely manner thus being one of the activities for the salubrious connectivity in a smart city.

This system is basically used in mobile operator's support number, customer support of our bank, e-booking of gas etc. An interactive Voice Response (IVR) system makes it easier for different callers to interact with this communication system over the telephone [2]. As in the case of a student enquiry system in an educational institution, if the announcements regarding any event, their agenda or any such information are provided manually, then there is a chance of misinterpretation or unavailability of the information to every individual. In such cases, interactive voice response system comes handy. IVR system also enables the caller to retrieve required information from database, and to upgrade information into a database, or both.

Dual Tone Multi-Frequency (DTMF) signals are created when a caller presses a key of his/her telephone set, or a press of Touch-tone keys form the system inputs. The sequence of messages to be communicated is context-specific to the user input and is determined dynamically according to an internal menu driven structure archived within the IVR database. The pre-designed IVR hub provides information for establishing start-up organizations and to the already existing small firms. This efficient way not only replaces the existing dialogic card system, but also proven to be more economical. However it requires maintenance and regular upkeep from time to time.

So designed IVR system consisting of simple components like microcontroller and normal application chips interfaced to a PC using simple software running in the background with other jobs running on the front end.

II. METHODOLOGY

Student Enquiry System, for example, can provide automated voice services over the phone to the students such as:

- Information about attendance, grades
- Schedules of classes and examinations
- Notifications on the fee paid or due
- Request for certificates
- Library services
- Any other general information

A. Hardware Requirements

- Personal Computer

Saqquaf S S M, Supercomputer Education and Research Center, Indian Institute of Science, Bangalore, India (saqquaf@serc.iisc.ernet.in)

Sarvesh Araballi, Department of Information Science and Engineering, Sri Venkateshwara College of Engineering, Bangalore, India.

Bhagyalakshmi P, Department of Information Science and Engineering, Sri Venkateshwara College of Engineering, Bangalore, India.

Shruti S. Mahadeek, Department of Information Science and Engineering, Sri Venkateshwara College of Engineering, Bangalore, India.

Varsha B M Department of Information Science and Engineering, Sri Venkateshwara College of Engineering, Bangalore, India.

978-1-5090-0774-5/16/$31.00 © 2016 IEEE

- Microcontroller
- DTMF Decoder
- Ring Detector
- Primary Rate Interface

B. Software Requirements

- Visual Studio
- MySQL
- Device Drivers

The various components used and their uses are as follows [3].

1. *Relay*: These are used as swapping tool between the ring detector and the DTMF decoder.
2. *Ring Detector*: It is used to detect the occurrence of incoming calls.
3. *DTMF Decoder*: It is used to convert the DTMF tones to 4 bit BCD numeric codes.
4. *Microcontroller*: This converts BCD calls, and process them to transmit them serially to the PC [4].
5. *Level Translator:* This acts as an interface between PC and microcontroller.
6. *Personal Computer:* PC to store the database and is used to carry out the text to speech conversion.
7. *Audio Amplifier*: This provides audio amplification to standard output and to act as a buffer between the telephone line and sound card.

III. Implementation of Student Enquiry System

A. Sequence of events followed in IVRS [5]:

1. The caller dials the IVRS service number.
2. The computer recognizes a specified number for ringing tones and the connection is established.
3. A pre-recorded voice message is communicated to the caller.
4. A menu is provided to the caller in the voice mode, to help caller select one of the many options.
5. If the information to be given is confidential, then the system may also ask the dialler, to demand a password authentication which can be a number, USN or a student registration number.
6. Required information is provided after fetching necessary information from the database.
7. The same information is relayed to the user in a voice/audio form.
8. The caller is normally given an option to :
   a) To request a repeat of information that was voiced to him.
   b) To repeat various menu choices.
   c) Or to terminate the call.

B. Touch-Tone Keypad

By touching a button it generates a 'tone'. This tone generates two frequencies, one from lower band and other from upper band. For e.g. pressing push button '6' transmits 770 Hz in lower band and 1477 Hz in upper band, as illustrated in Table 1 [3].

Digits from 0 to 9 are used for calling a number on phone. This generates frequencies in Dual Tone Multi Frequency (DTMF) signals. The signal so generated is of sinusoidal in nature.

Rows and columns in the table give the frequencies. The keypad can be made to work with different frequencies but only two frequencies are transmitted at a time. The signals coming from such type of telephone are called Dual Tone Multi Frequency (DTMF) signals.

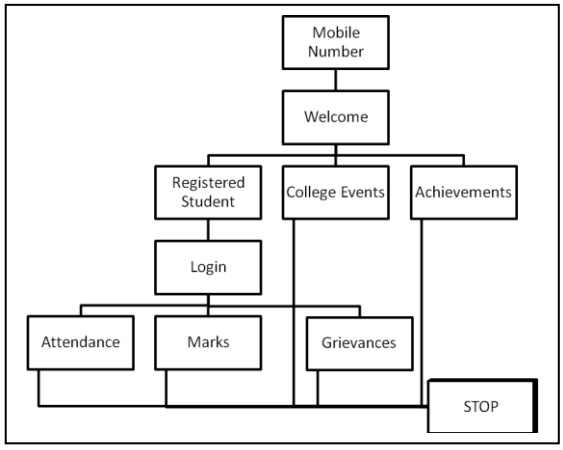

Fig. 1: Flowchart of IVRS

TABLE 1: DTMF KEYPAD FREQUENCIES

	DTMF Keypad Frequencies		
	1209 Hz	**1336Hz**	**1477 Hz**
679 Hz	1	2	3
770 Hz	4	5	6
852 Hz	7	8	9
941 Hz	*	0	#

C. Telephone Interface

The following block diagram depicts the schematic representation for the sequential flow of events in IVRS and various hardware components.

D. Voice over Internet Protocol (VoIP)

VoIP technology is used for transferring phone call signals by IP telephony [6]. It basically refers to the communication protocols, methodologies, and transmission techniques involved in the delivery of voice communications and multimedia sessions over Internet Protocol (IP) networks. The steps involved are media channel setup and signalling. This digitization of the analog voice signal, encoding, packetization, and transmission as IP packets are used in a packet-switched network.

The earlier service providers of Voice over IP services offered only for business models and technical solutions similar to the architecture of the existing telephone network. Subsequently, the second generation service providers, like Skype have built private networks for users, offering the

benefit of free calling facility. The major disadvantage being that calling to other service providers is denied.

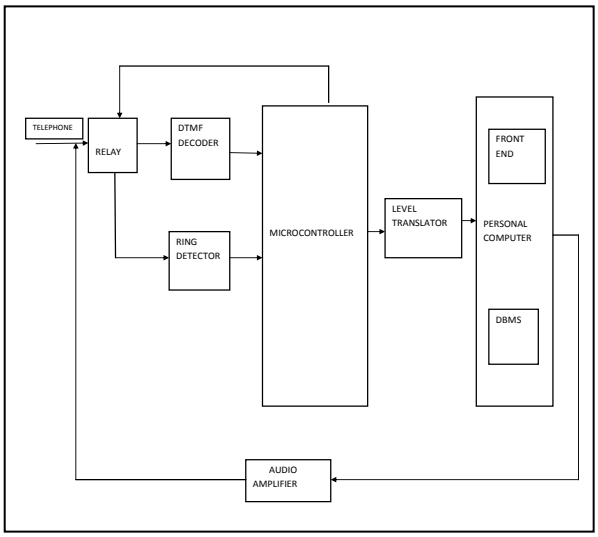

Fig. 2: Block diagram of Telephone Interface.

VoIP systems controls the set-up, tear-down of calls and audio codecs responsible to encode speech allowing transmission over an IP network as digital audio via an audio stream, by making use of session control protocols. Based on application requirements and network bandwidth, the choice of codec would vary. Some applications rely on narrowband and compressed speech, while others support high fidelity stereo codecs.

Low cost is the biggest single advantage of VoIP, which has over average telephone systems. In addition, international calls over VoIP are usually very inexpensive and are usually free between various VoIP users. But, VoIP telephone systems are susceptible to attacks like active or passive attacks like any other internet connected devices.

E. Text to speech system

A text-to-speech (TTS) system is used to convert any normal language text into speech. Concatenation of different pieces of recorded speech stored in a database, gives rise to synthesized speech [7]. Systems differ in the size of the stored speech units. A system that stores phones or di-phones 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.

In other words, to generate a completely synthetic voice output, a synthesizer can incorporate a model of the vocal tract and human voice characteristics. The speech synthesizer's similarity to the human voice and its ability to be clearly understood are few of the things on which the quality of the speech synthesizer can be judged.

The main advantage is that an intelligible text-to-speech service allows people with disabilities such as visual impairments or eye sight disabled persons to listen to written words easily on a personal computer.

A text-to-speech system is usually composed of two parts: a front-end system and a back-end system. The front-end has two major tasks namely to convert raw text containing symbols like numbers and abbreviations into the equivalent of written-out words and then assign transcriptions to each word, and divides and marks the text into phrases, clauses, and sentences.

This process is often refers as text normalization, pre-processing, or tokenization process. The advanced process uses the technique of assigning phonetic transcriptions to words is called text-to-phoneme or grapheme-to-phoneme conversion. The output by the front-end is the combination of phonetic transcriptions and prosody information which organizes to make up the symbolic linguistic representations. The back-end often referred to as the synthesizer is responsible to convert the symbolic representation to sound.

IV. Results and Discussions

The call flows in the manner as shown below in Fig 3. On dialling a number, the welcome message is selected and played in one default language from a multi-lingual system. After this, the caller can choose any of these languages as per his/her choice for the information required. Subsequently, if we consider a student enquiry system in an engineering university, the caller can select the study discipline (Civil, Electronics, Electrical, Mechanical, Computer Science, Information Science etc.) and then picks year (I, II, III, IV) to which the student belongs. After this the caller can indicate either attendance or result or any other option provided in the menu which is played out to the caller. The caller needs to enter the USN number of the student for verification purpose. The system then plays the information retrieved from the database for attendance, result etc. of the student. The information can be sent via texts if needed too. After getting the required information the call is disconnected [8].

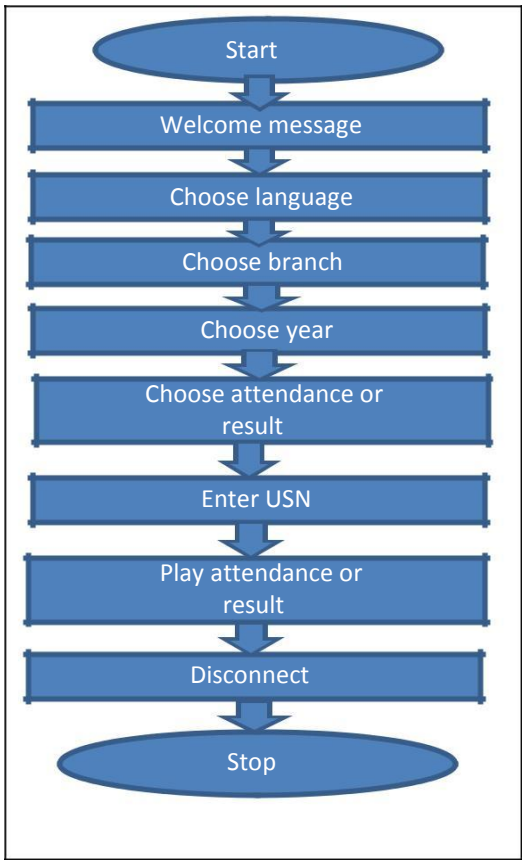

Fig. 3: Execution Flowchart

The following figures show the implementation protocol of IVRS in any institution. Fig 4 depicts the setting up of a

Session Initiation Protocol (SIP) account and registering our personalized softphone on a private network. The registration confirmation is shown in Fig 5

Snapshots

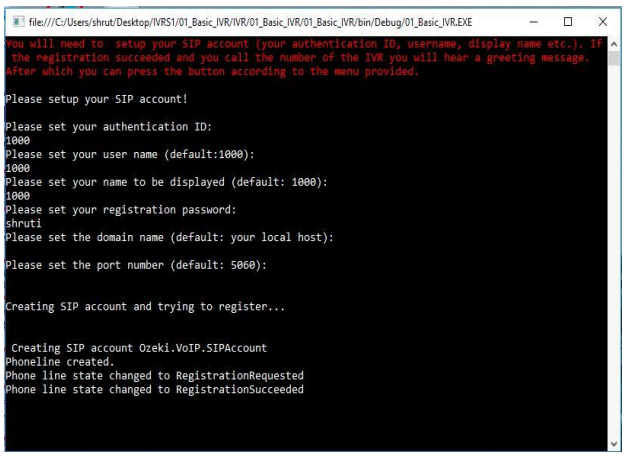

Fig. 4: Registering our softphone with VoIP PBX

The various advantages of this system includes-

- Better user contact and satisfaction.
- Cost effective and reduces human inefficiencies.
- There is no entry for hackers hence it provides security for the data.
- Saves time and money for small and large organizations.
- An IVR system can be made available 24/7 basis to field questions and help callers with simple tasks.
- Increased automation freezes the customer service agents from any routine administrative tasks thus reduces cost associated to customer service staffing.

Fig. 5:Registration confirmation

Fig. 6: Incoming call to softphone

Fig. 7: Call status

Fig. 8: Call completion

The call status i.e. the incoming call, in call and the call completion is shown in Fig 6, Fig 7 and Fig 8 respectively.

V. FUTURE SCOPE

In the future, the concept of Interactive Voice Response System can be made available for various smart city developments, for example it can be used in various transport departments, Railways and Airports administrations. This gives a fast and reliable response catering to the needs. The end users' needs be to provide customers with an easily accessible information system providing necessary type of quarries like:

- •Information Enquiry
- •Schedule Enquiry
- •Tele ticketing etc.

So, in near future, all the information regarding available routes, schedules, fares etc. will be known through the Interactive Voice Response System. This concept may be implemented in Cinema halls and Multiplexes where the caller will get to know the timings of ones favourite movies as well as one can book tickets through this IVR system.

In this student enquiry system, various schedules and tasks taking place in the institution along with grievances are highlighted.

VI. CONCLUSION

The aim of this project is to provide the entire necessary information to the end user in an easy way. The application has an access that is compatible to computers, laptops, mobile phones, smart TVs and/or any other suitable device.

Owing to this project the traditional manual way of handling the customer queries will be handled in a technological, timely and automated fool-proof manner. The operations performed by this system are similar to that of a human telephone operator.

The advantage of this study is its relevance to the field of telephony and the inherent low cost to build using simpler and easily available components. This will act as a basic building block of the *e*-infrastructure of a smart city in the future development of the society.

ACKNOWLEDGEMENT

The authors would like to place on record their gratitude to the authorities of Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, India for the encouragement and support during the entire course of this work. The authors would also thank their college SVCE and the faculties of department of ISE for their immense support.

REFERENCES

[1] http://en.wikipedia.org/wiki/Interactive_voice_response

[2] Ankita Bhondge, Aditi Bhatkar, Sapna Fender, and Sonali Thakre, Megha Goel, "Interactive Voice Response System by Using Asterisk," Vol. 3, Issue 3, March 2015. (*references*)

[3] Andreas Velonis, Theofilos Milonas, George Papazidis, George Dimitriou, "VoiceCampus: An Automated Interactive Voice Response System for Students,"

[4] Douglas V.Hall / Microcontroller and Interfacing /New York/TMH Publishing

[5] Ms Seema P Mishra, Ms Apeksha S.Chavan, Swapnil S. Gourkar, "INTERACTIVE VOICE RESPONSE SYSTEM FOR EDUCATIONAL INSTITUTION", International Journal of Advanced Engineering Technology, E-ISSN 0976-3945

[6] Santosh A. Kulkarni, Dr. A.R.Karwankar, "IVRS FOR COLLEGE AUTOMATION", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 1, Issue 6, August 2012.

[7] D. Sasirekha, E. Chandra, "TEXT TO SPEECH: A SIMPLE TUTORIAL", International Journal of Soft Computing and Engineering, Vol. 2, Issue 1, 2013.

http://www.ivrsdevelopment.com/ivrseducation.htm