# Modeling Interactive Voice Response Systems – A Review

**Nakul Vashishth, Jaibir Singh**

**Abstract-***In today's world, IVR Systems has become more then a necessity. Whether it's a company, an education organization, a hospital, a Bank or any other organization, without IVR System neither of these can survive. IVR System helps organization to automate various processes with human interaction and gives a feel to caller as if caller is interacting with a human being and not a machine and thus has become need of modern world. An IVR System consist of many technology and components like Speech Synthesizer, Speech Recognition, DTMF encoder and decoder, database and Hardware upon which entire system operates. This paper intends to give overview of how an IVR System is modelled and implemented.*

**Keywords:** IVR System, DTMF, Speech Synthesizer, Private Automatic Branch Exchange, Audio Response Unit, VoiceXML, TTS, ASR.

## I. INTRODUCTION

IVR System is a technology that makes interaction between human being and Computer as if there are two human beings interacting. An IVR System masks a Computer or Computing device and present itself as if that is a human being. IVR allows customers to interact with a company's host system via a telephone keypad or by speech recognition, after which they can service their own inquiries by following the IVR dialogue. IVR systems can respond with pre-recorded or dynamically generated audio to further direct users on how to proceed. IVR applications can be used to control almost any function where the interface can be broken down into a series of simple interactions. IVR systems deployed in the network are sized to handle large call volumes.

This is a common need of IVR System in all organization. Whether it's a company or a university or a college or a hospital or a bank or a call centre. Without IVR System no business can sustain.

For a company IVR System is required for attending its client as IVR serves as Automated Attendant. A university or college typically use IVR System for dealing with students queries, a hospital need IVR System as attendant for patients and care takers. Call centres use IVR System for both answering customers as well as for calling customers for marketing and sells purpose. Banks now a days has a feature of Phone Banking and IVR System is what making that successful.

Some of the benefits of an IVR System are:

- Enables automation of processes.

- Improve customer service.

- Lower implementation cost.

- Service with chosen language

- Call logs available for auditing, performance report and future enhancement.

- Provide 24*7 availability for customers.

- Provides secure interaction between computer and human being and thus secure business like banking or account management becomes possible.

- Out bound dial gives a great customer experience in marketing campaigns.

- Call Centre and BPOs are there just because of IVR Systems

## II.    IVR SYSTEM

INTRODUCTION TO TELEPHONY

To understand architecture and implementation of conventional IVR System a study of telephony system is prerequisite.

The first telephones were connected directly with each other by cable. As this technology continued evolving, the necessity of finding a more scalable solution arose. One centralized system uses $n$ instead of $n*(n-1)$ connections between $n$ participants. This enormous reduction led to the invention of an electronic circuit system called **Telephone Exchange** or **Telephone Switch**. Such a switched system began to play the role of telephone the connecting part between two or more telephony end points. Later it was called **Exchange** the **Central Office**. Every telephone set is connected to the central office by a pair of cables, defining this way the so called **Local Loop**. At the beginning the destination number was sent by dial pulses as a result of measuring how far the rotary dial is, before being released.

At the beginning the destination number was sent by dial pulses as a result of measuring how far the rotary dial is, before being released. Nowadays, this method is replaced by sending a sequence of touch-tones, known as **Dual-Tone Multi Frequency (DTMF).** There are sixteen keys (numbers in the range from 0-9, the letters A, B, C, D and the symbols ∗ and #) on the telephone keypad. Pressed, each of them produces two tones, which are sent to the central office as audio signal. The keys are ordered in a 4x4 matrix. Each row of this matrix represents low and each column high frequency. For example, if you press 5, two frequencies will be sent to the central office to determine which key is associated with these frequencies.

For simplification purposes, the architecture of a phone system adopted later concepts like **trunking** which can be explained as a tree structure with a main transmission channel (**The Trunk**) beginning from the switch centre and branching at the opposite end into smaller lines. The amount of all possible telephone exchanges forms the so called **Public Switched Telephone Network (PSTN)**. This network has a hierarchical structure roughly divided into classes for long distance switching and those for connection to the end subscribers. There are three kinds of networks defined: **Local Network, Exchange Area Network and Long-Haul Network**.

The granularity of the telephone structure can be further refined with the introduction of so called **Private Branch Exchange (PBX)** or **Private Automatic Branch Exchange (PABX).** It plays the connecting part between a private telephone network of an organization and the PSTN. With the extremely fast development of the Internet in the 1990's, this concept evolved to **VoIP PBX (Voice over Internet Protocol)**. VoIP describes a technology that allows delivery of voice signals through a packet-switched network (usually Internet) able to communicate with other networks through the **Internet Protocol (IP)**. VoIP is often referred to as **Internet Telephony or Voice over Broadband (VoBB)**.

Later it was obvious that not every small company has the resources and competence to manage its own PBX. Therefore the telephone service providers offered to host and manage PBXes, delivering all features and calls via the Internet. As a consequence of this evolution, nowadays there are four types of PBX:

|  | Circuit Switched | Packet Switched |
|---|---|---|
| Private | Traditional PBX | IPPBX(VoIP PBX) |
| Hosted | Virtual PBX | Virtual IP |

IVR SYSTEM ARCHITECTURE

A typical IVR application consist of an IVR System exposed over PSTN and connected to Application Server which is connected to database system.



**Figure 1:A Typical IVR System**

A caller want to know information related to his account could call on customer care number from mobile phone or landline phone. Call would be routed to specified number through PSTN. An IVR System is hosted at called number along with PABX which would then play a pre-recordedwelcome message using Pre-recoded Audio Player and would either route the call to available line or would ask caller to wait, on getting an available line it would then transfer call to attendeeor may ask user to input more choices. Choices would be decoded by DTMF Decoder so that IVR System can understand and act upon. Thenthese provided user choice would be processed by IVR System. IVR System then request application server to fetch or calculate details based on input choice. Application server using database would generate response and send back to IVR System. IVR System then synthesis speech using Speech Synthesizer for the response returned by Application Server. The IVR System would then play the same to caller (Explained in Figure 1).

IVR SYSTEM COMPONENTS

Entire IVR System can be understood as multiple layers which perform designated tasks. Figure 1 explains basic functioning and data flow of IVR System, figure 2 presents detailedlayeredarchitectureof components of IVR System:

**Automatic Call Forwarding**

This is typically task of PABX or PBX to land the caller call to available line. It lands the call to lines based on predefined algorithm and if all lines are busy IVR may play a message stating "All our representatives are busy, please hold on." PABX is most important component of the IVR System as this enhance the system to serve multiple users at a time.

**Telephony Interface**

The Telephony Interface are APIs, which provide computer telephony integration and enables PC to use telephone services. Different implementation and versions of these APIs are available that allows applications to control telephony functions between a computer and telephone network for data, fax, and voice calls. It includes basic functions, such as dialling, answering, and hanging up a call. It also supports supplementary functions, such as hold,transfer, conference, and parked call found in PBX, ISDN, and other telephone systems. These APIs are very important as in modern IVR system these have replaced many hardware components.
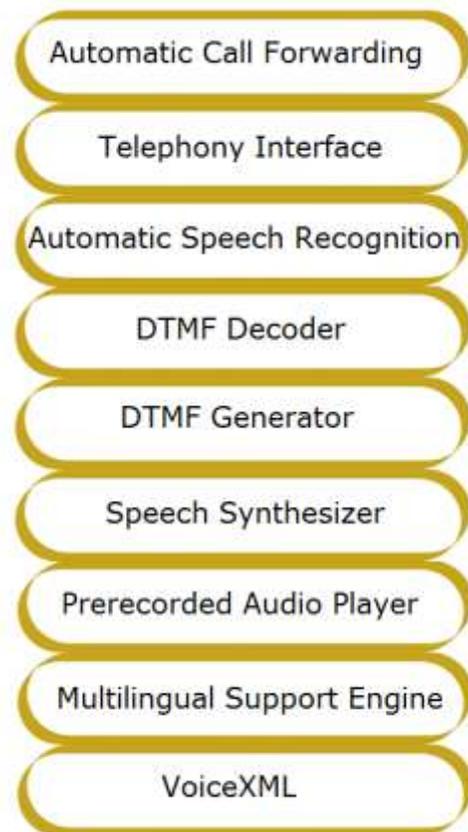


**Figure 2: A Layered Architecture of IVR System**

**Automatic Speech Recognition (ASR)**

ASR is used for building speech-driven user interfaces and provides an efficient and intuitive spoken alternative to touch-tone (DTMF) applications. However ASR is every IVR Systems feature but there are three main reasons that reduce the accuracy of ASR systems:

**Ambient Noise**: In many cases the caller is calling from a noisy environment and the ambient noise is reducing the accuracy of speech recognition engines.

**Accent**: Not all ASR systems are able to efficiently interpret different accents. Vendors of ASR systems keep improving their engines, sometimes by adding

vertical solutions – i.e. solutions that are tuned for specific market segments.

**Network quality**: Low network quality is causing interruptions in the audio signal that impact the ability for proper speech recognition.

### DTMF Decoder

This is either a software component or a hardware component depending on the implementation of IVR System that decoded input of the caller based on touch-tone or button pressed.In modern IVR Systems these are typically software components.DTMF was originally decoded by tuned filter banks but in modern world *Digital Signal Processing* became the predominant technology for decoding. DTMF decoding algorithms often use the *Goertzel Algorithm* to detect tones.

### DTMF keypad frequencies (with sound clips)

|         | 1209Hz | 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       |

### DTMF Generator

Like DTMF Decoders DTMF Generators are also can be Software or Hardware components based on implementation of IVR System. In modern IVR Systems these are typically software components that generates DTMF tones and that can be given back as response to caller. This are a required feature in case caller wanted to integrate response with some other IVR or Voice Command System.

### Speech Synthesizer

In order to implement an IVR system, components capable of interacting with the people are necessary. The basic system modules are an input component (either recognition of a spoken language or DTMF control signal) and an output component (the synthesis of speech). The text-to-speech (TTS) unit is responsible for transforming an input text into TTS an audio file. TTS engines are often referred to as speech synthesizers. There are different ways for synthesizing speech. Some systems use a unit selection principle, while others produce completely synthesized speech by replacing the formants with pure tones. An example of unit selection speech synthesis is when phones, diaphones, syllables etc. are segmented from a large database with recorded utterances and then concatenated with each other in the process of synthesis. TTS engines built by unit selection sound more natural than parametric ones but they rely on gigabytes of recorded speech which could be a problem. A different approach is the parametric speech synthesis which does not use human speech but artificially produced output by configuring parameters such as fundamental frequency, voicing and noise levels of a waveform.

### Pre-recorded Audio Player

Despite of many researches and developments in Speech Synthesizer still human voice is better than synthesised voice. Therefore every IVR System has some pre-recorded messages that it plays while interacting with caller. Only data part is in general synthesised and that is concatenated to pre-recorded message. The components that plays these messages is known as Pre-recorded Audio Player and is commonly a software components.

### Multilingual Support Engine

There are cases when callers calling at customer care, belongs to different region and languages and thus an IVR System must support multiple languages. The software component that enables IVR System to support multiple languages is Multilingual Support Engine.

### VoiceXML

VoiceXML (VXML) is a digital document standard for specifying interactive media and voice dialogs between humans and computers. It is used for developing audio and voice response applications, such as banking systems and automated customer service portals. VoiceXML applications are

developed and deployed in a manner analogous to how a web browser interprets and visually renders the Hypertext Markup Language (HTML) it receives from a web server. VoiceXML documents are interpreted by a voice browser and in common deployment architectures, users interact with voice browsers via the public switched telephone network (PSTN).

The VoiceXML document format is based on Extensible Markup Language (XML). It is a standard developed by the World Wide Web Consortium (W3C).

**IVR Request-Response Tree**

A typical IVR System presents choice to a caller and on the basis of choice provided either it would serve the response or would again present choices. This is in general a configurable part in any IVR System and typically looks like Figure 3.
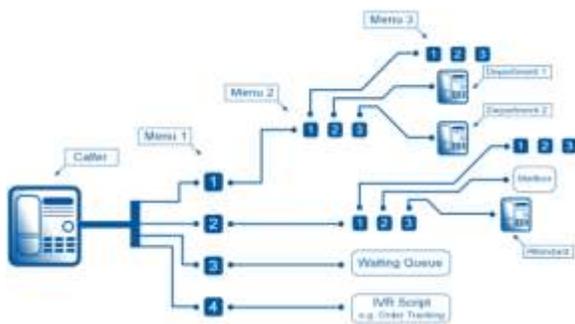


**Figure 3: IVR System Request-Response Tree**

**Application Server**

This is a separate machine or server which is responsible for doing all processing on data. In some implementation this server is connected with HTTP Server and several web applications are hosted on these. This server can be used for catering reporting need or can provide another channel as web application for IVR System.

**Database Server**

This is typically a Relational Database server or LDAP Server which is used for storing all the data of all the customers that could be asked by caller. Information about all the inbound and outbound calls, call duration and other information's of call are also stored in database and may be used for reporting

purpose. In some cases report engines are also implemented.

## III.      CONCLUSION

In this paper we figured out importance of IVR System as this is one of the most important breaks through in the field of telecommunication and software engineering. IVR System provide a voice response to the customers and guide them to the information they require anywhere and anytime. The customers can call up any institute such as banks, universities, tourism industry and obtain any information by simply pressing certain button on his telephone as per the guidance of the voice fed into the computer. IVR System is an electronic devicethrough whichinformation is available related to any topic. IVR System is usually employed to know more about the organizations and can be modified to respond to voice of the customer for better response customer satisfaction. IVR System can be employed in customer services there by improving its flexibility and efficiency.

We also studied several layers and components of a typical IVR System like PAXB, Automated Speech Recognition, DTMF Decoders and Generators, Speech Synthesizer or Text-to-Speech and other components like Application Server which is responsible for doing all processing on data and Database for persisting data.

We also studied that IVR Systems can have different implementation in terms of Software and Hardware components. As PAXB, DTMF Decoder and Generator can be implemented by Software as well as Hardware.

## IV.      REFERNCES

[1] "Dodd, Annabel Z. The essential guide to telecommunications. Prentice Hall PTR, 2002"

[2] "Anil Kumar, S. Niranjan: Design, Development and Implementation of an Automated IVR System with feature based TTS using Open Source Tools,International Journal of Engineering Research & Technology, Vol.1 - Issue 3 ( May- 2012)".

[3] "Ankita Bhondge, Aditi Bhatkar, Sapna Fender, Sonali Thakre, Megha Goel: Interactive Voice Response System by Using Asterisk, International Journal of Innovative Research in Computer and

Communication Engineering, Vol. 3, Issue 3, March 2015".

[4] "Hang-Xing QI, Qing-Dong Du, "An smart IVR system based on application gateways" ninth international conference on hybrid intelligent systems, 2009."

[5] "ShanchuanXu, Wanlin Gao, Zhen Li, Shuliang Zhang, Jianing Zhao, "Design of hierarchical and configurable IVR system" 2nd International conference on computational intelligent and natural computing 2010."

[6] "Ms Seema P Mishra, Ms Apeksha S. Chavan, Swapnil S. Gourkar, "Interactive voice response system for educational institution" international journal of advanced engineering technology, April 2012."

[7]"Santosh A. Kulkarni, Dr. A.R. karwankar, "IVRS for college automation" international journal of advanced research in computer andcommunication engineering vol. 1, issue 6, August 2012."

[8] "http://the-asterisk-book.com/1.6/"

[9] "http://solicall.com/automatic-speech-rcognition-asr-in-ivr-systems/"

[10] "http://www.voicexml.org/"

**Nakul Vashishth,** B.Tech in Computer Science and Engineering from NIT Kurukshetra with a 9 years of industry experience in Software Development.

**Jaibir Singh,** Assistant professor at Shanti Niketan College of Engineering, Hisar.