# Context-aware IVR: Challenges and Opportunities

Siddhartha Asthana, Pushpendra Singh, Amarjeet Singh, and Vinayak Naik
Indraprastha Institute of Information Technology, New Delhi, India
Email: {siddhartha,psingh,amarjeet,naik}@iiitd.ac.in

***Abstract*—Voice based telecommunication services are emerging as most appropriate medium for information dissemination in developing regions, as they overcome prevalent low literacy rate and are more reachable than any other medium like internet. However, present-day Interactive Voice Response (IVR) systems are not able to capture user requirements and present a rigid interface to its users, and thus the IVR systems are commonly considered as frustrating to use.**

**In this paper, we are proposing an architecture to build context-aware IVR systems, which can adapt themselves to match user requirements and provide a better experience, thus enabling their widespread use. We did a real world deployment to understand the requirements for the new architecture. We intend to make use of work in context-aware systems and adaptive systems to create next-generation IVR systems that have human-like ability in tackling caller intricacies with ease of access like internet.**

***Index Terms*—adaptive computing; context-aware; information system; IVR;**

## I. Introduction

In recent years, the Internet has emerged as a primary source of diffusing information across the world. Ease of use and scalability have been the prime factors behind its success. In developed countries, where literacy is a norm and most of the people have access to computers, the Internet has reached to every segment of the society and is enormously popular. On the contrary in developing countries like India, the Internet is limited to only urban and rich segment of the society. Latest TRAI (Telecom regulatory authority of India) reports suggest that in a country of over 1 billion, regular Internet use is limited to around 30 million people only, majority of whom live in metro cities, thus making a penetration of less than 5% [1]. Illiteracy (especially with respect to English which is the primary language of Internet) coupled with fear of technology and weak economic status compared to their western counterparts are some of the main reasons behind such a low penetration among the masses in developing countries. With growing need of providing an effective information dissemination platform for the masses, use of Interactive Voice Response (IVR) system is growing.

Unlike Internet, IVR system provides interaction in a natural language interface - often local language - over telephone to make information accessible. This overcomes the language barrier and since access to IVR is through phone, accessibility is not an issue for mass usage even in a developing country like India. Access through phone also results in requiring people to only have numeral literacy limited to the ability to recognize a number such as on the key pad of a phone. A good example of IVR system is Indian Railway Information System which provides information about running trains and is used across India from all sections of the society everyday. IVR systems are also fast emerging as an alternative way of providing other services, e.g. telephone banking, tele-voting, etc. Latest TRAI report indicates that mobile phones subscription has reached 650 million and is on increase thus making mobile phone the most accessible technological platforms. With constantly falling prices for phones as well as phone call charges, accessing IVR system is in reach of common people. These observations has led to advocate use of IVR systems for information dissemination in developing regions [3], [4].

Currently, An IVR System is developed with an information database at the back-end, which is then made available to the users over the interface of simple key presses using a pre-defined menu structure to answer user queries. This menu structure and the information that it extracts from the data-base usually remain fixed throughout the life-time of IVR System and any change requires manual intervention. For example, in case of Indian Railway IVR system, the menu structure to access information remains fixed thought out the year and offers same interface for all the users, thus not taking into account contextual factors like nature of queries, caller abilities to interact with the system, or interaction pattern.

Every information given in IVR System has a unique contextual value. The IVR System tries to capture it at a very coarse-grained level by providing a fixed menu for traversal. The fixed menu structure by its very nature ignores the dynamic nature of context and provides information as envisioned by the developer of the system at the beginning. However in reality, the nature of information sought from a system changes with time and depends heavily on external contextual factors. Let us take an example from real-life: most of the colleges and universities websites provide FAQs on their website which usually ends up as a long list of questions and associated answers. However the questions that get most attention varies with time for example before entrance exams, queries are mostly related to syllabus, eligibility criteria etc. while after the exam, questions change for results, fees etc. Here, it is quite clear that an external event, i.e., exam has a direct impact on the questions that get most attention.

The menu structure of most of the present-day IVR system is also designed to answer commonly occurring questions same as FAQs. Thus IVR Systems with static content end up with either providing non-relevant information or overloading user with the task of navigating to right information in a highly complex structure with lots of menu items requiring key presses and repetition of instructions which are not always very clear. This results in wasting the time of the user in finding the information they need and adding to the frustration of using the IVR system. In current IVR System, such shortcoming can only be overcome by manual intervention, i.e., by either changing the menu structures or editing the information base from time to time by constantly monitoring the system and then changing it accordingly, however this solution is neither scalable nor dynamic enough to be employed in real-life systems which have high availability requirements.

Current context must play an important role in determining the behavior of IVR system to make it easier for the end-user to find the relevant information. As IVR systems become mainstream medium of providing information, a context-aware IVR System which can adapt itself is need of the hour. Such a system should be able to

1) monitor current usage of the system, e.g . the questions that are being asked, most accessed menu items, etc.
2) take cognizance of external factors related to the system, e.g. actual exams in case of a college admission IR system
3) recognize and learn other contextual factors, e.g. access patterns such as time taken in pressing the keys, ethnography of users of the system, different times at which system is most used, history of the usage, etc.

In this paper we present design challenges of such an IVR system which is context-aware and can self-adapt itself without manual intervention.

## II. Experiment

We developed and deployed an IVR system to provide information specific to admission process in the bachelor of technology program at Indraprastha Institute of Information Technology (IIIT-Delhi). Aim for this deployment was to get real-world usage of IVR system to see how context-awareness can help improve the effectiveness of an IVR system.

More than 2,000 application were received for this course. The IVR menu had six options which user can select one at a time by pressing the key on the keypad of mobile phone. Each option was designed to give information about a topic related to the admission process, e.g. issue of admit card, fees of the program, syllabus for entrance exam, etc. The information content for each option were designed with the help of IIIT-Delhi's office staff who had experience in handling admission related queries. These six options were ordered in such a way that the information which is most likely to be queried by user is put as first option followed by the second most likely information as second option. The aim was to allow callers to access the information fast and efficiently over the IVR.

The table below shows the actual usage of each option as accessed by callers.

| Option | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Hits | 62 | 17 | 10 | 8 | 32 | 18 |

TABLE I
Number of hits on each option.

The actual usage data in Table I shows that $5^{th}$ and $6^{th}$ option should have been the $2^{nd}$ $3^{rd}$ option in the menu. This shows that despite consulting experienced staff, it is impossible to guess user behavior in real settings. Placing the options of IVR in the order where more relevant information or more likely information is put before the less relevant or less likely information can have great impact on throughput of the system in handling number of calls per unit time and also on user experience of using the system. Providing a good user experience is essential for the widespread use of such systems.

IVR systems announce a main menu structure and once a menu item is selected, specific information about the option is provided; after finishing the specific information, again the main menu structure is announced. Let us assume that IVR takes $x$ seconds to announce the complete menu structure and $y$ seconds to play specific information contained inside an item to the caller. For the sake of simplicity, we are assuming that all menu items contains information of equal length and we want to access all the information i.e. starting from first option to last option. Then, the minimum amount of time required by ordering of option as mentioned in Table I will be

$$62(x+y)+17(2x+y)+10(3x+y)+8(4x+y)$$
$$+32(5x+y)+18(6x+y)=147y+426x$$

However, if the IVR adapts itself and changes the menu items in the order of user preference, resulting minimum amount of time required will be:

$$62(x+y)+32(2x+y)+18(3x+y)+17(4x+y)$$
$$+10(5x+y)+8(6x+y)=147y+346x$$

From this simple example, it is easy to see that a context-aware self-adaptable IVR could provide better user experience through significant times savings at the user's end besides helping in achieving a better system throughput since the call duration will be reduced leading to more users able to interact with the systems. However, it is important to see that our simple hypothesis only captures a small part of the context and evaluates its role in a simplistic manner. To make an IVR system truly context-aware and self-adaptive, we intent to understand, capture, and evaluate various environmental factors surrounding an IVR system. User personalization has been tried before to give better services to a particular user, however our goals are much broader and beyond user personalization. We intend to propose IVR systems that are as easy to use as talking with a human operator while scalable like internet with minimal human intervention.

## III. Design

Current-day IVR System designs are very crude in nature. Most of the IVR systems have fixed and same functionality for all of its callers. In this section we present a novel design of context aware IVR system which can learn and predict the callers requirement and adapt itself to serve accordingly.

Our design for context aware IVR system is based upon the prior work in complex adaptive systems. A complex adaptive system can be modeled as: Multi-agent system model, Building block Model or Boundary and Signal Model.

1) A Multi-agent system model is simply defined as a system composed of multiple interacting agents.
2) A building block model visualizes a system as organized structure of finite elements known as building blocks. Each of these building blocks can have different state. While adaptation such systems change from one such state to another state for some building block. For illustration, an eye and a mouth can be a building block in recognizing a face. A different types of eyes and mouth can lead to same face look differently.
3) In Boundary and signal model, a signal is a sensory information or data passed by a processing element known as boundary to another boundary. As the signal passes by these boundary, each of these boundaries modifies the signal for a specific purpose.

We propose using multi-agent system model for context-aware self-adaptive IVR systems. We prefer multi-agent system because a multi-agent system can acquire the properties of other two model. A multi-agent system when viewed as strict and defined structure and hierarchy can provide the functionality of building block models where as one group of agents interacting with another group of agents can serve as boundary and signal model

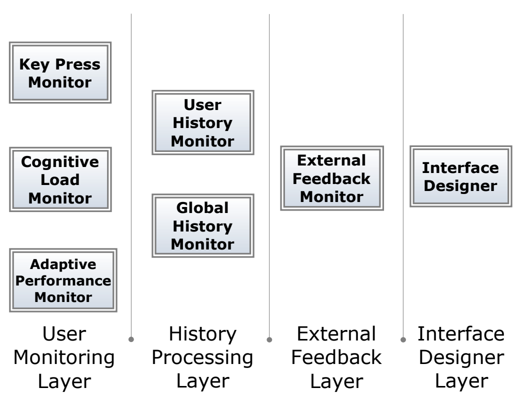


Fig. 1. 4-Layer Architecture Diagram for Adaptive IVR system based on Agents Models

In Figure 1, we present an architecture diagram using 4-layer multi agent model for our context-aware self-adaptive IVR system. Each layer in the architecture comprises of one or more agents. Each agent in a layer works on conceptually the same information entity but looks for different information aspects of the entity. These agents monitor the attributes specific to a caller or a group of caller where a caller is a user of the IVR system. A wide variety of user models and analysis techniques have been developed to support specific applications [2]. Typical attributes (which we also intend to capture as part of context) maintained in a user model are:

- User preferences, interests, attitudes and goals
- Proficiencies (e.g. task domain knowledge, proficiency with system)
- Interaction history (e.g. interface features used, tasks performed/in progress, goals attempted/achieved, number of requests for help)
- User classification (stereotype)

The general architecture can be extended to any number of layers based upon types of adaptation and the extent to which context sensitive information related to attributes of user model needs to be sensed by the IVR system. The four layers in the architecture are:

1) User Monitoring Layer: This layer is responsible for monitoring user attributes. Specific values for the attributes may be explicitly specified by the user, captured directly from user actions, or derived by the analysis engine. Our model of adaptive IVR system has defined 3 agents in User Monitoring layer which are looking for different user attributes. The sensory information from these agents and the agents in other layer will capture attributes mentioned above for the user model.
   - Agent 1 (Key Press Monitor): Interaction with IVR systems is through speech or key presses with research finding the key-press interface much simpler to use [5]. This agent will monitor the pattern of key-press while taking a DTMF (Dual-tone multi-frequency) input from the user. Different patterns of key presses can be mapped to how accustomed is user with the system. This contextual feature can help system in deciding certain adaptation features, e.g. the ideal inter-digit time out need to be set for different users.
     A similar agent can be developed for speech-based IVR systems.
   - Agent 2 (Cognitive Load Monitor): IVR systems exposes some set of option to its user to decide upon. Literature suggests the ideal number of option to expose at any level in IVR menu should be in between 4 to 7. This general guideline may not be true when dealing with users of different age group. An elderly person may not have the same ability to remember the options compared to a younger person. This agent will look for parameters related to cognitive load of the user and will try to identify the optimum value for each user.
   - Agent 3 (Adaptive Performance Monitor): As the system adapts itself, it is important to continuously

measure its performance to ensure that adaptation is not causing a negative effect. This agent will monitor the performance of user after every adaptation of the system.

2) History Processing Layer : Agents for this layer will continuously monitor the historical data to make intelligent predictions for the caller. This layer will primarily have 2 agents:
   - Agent 1 (User History Monitor): This agent will look for historical data pertaining to the current caller. A repeated user must get a experience more suited to his or her needs and though previous work exists on user personalization, external contextual factors have not been taken into account.
   - Agent 2 (Global History Monitor): This agent will look for historical usage data from all the users of the system to adapt system for its better use by the future users, e.g. reordering of menu as we observed in our experiment.

   Many interesting questions will arise, like when to adapt, how much to learn, what to learn, what constitute for a user experience versus a global experience, etc. that we intend to answer in our future work.
3) External Feedback Layer: This layer is provided to accept features from outside world which may impact the validity or expiry of the information content. Users of the system are continuously influenced by outside world and their activity in the system may have relation to events in outside world. For example an IVR system for news content of a country may suddenly start getting high number of hits on Sports section in comparison to politics section because of some important Sporting event happened in that country or exams for a college admission IVR system.
4) Interface Design Layer: This layer is mainly to consume the service of all the agents presents in other layer. It will take intelligent decisions on how to accommodate all the decisional information reported by other agents to adapt the system interface for its users. User interface has its own challenges, e.g. how to present information, what to present, etc. that will be addressed at this layer.

We are proposing an agent-based architecture with 4 layers to capture the requirements of a context-aware self-adaptive IVR system. We believe that the architecture is extensible and scalable to allow us any modifications that we may require in the course of the development of our system.

## IV. Future directions

Developing and information disseminations system for masses has its own challenges and when it comes to developing countries, additional challenges of illiteracy, accessibility, etc. come up. IVR systems are most suitable because of their ease of use and higher accessibility across all sections of the society. So far a major limitation in the widespread use of IVR systems has been their static and crude way of delivering information.

We are proposing to use the knowledge of context-aware systems to make IVR systems self-adaptive. This will allow IVR systems to capture dynamic nature of usage, thus delivering better user experience to the callers. To understand the design requirements, we developed and deployed an IVR system and captured data reflecting its real-world usage. Currently, we have deployed same IVR system in a different setting where we are applying adaptation manually - in the line of Wizard of Oz type experiment - to further validate our approach.

We have already started the development of a new IVR system based on our proposed architecture. Once, we have a prototype of new IVR system ready, we will be deploying it to evaluate its usage and performance in real-world. This will allow us to better understand role of context and adaptiveness in improving IVR systems. We intend to make an IVR system which is easy to use like internet while adaptable enough to understand caller intricacies and match human-like ability in providing answers to callers.

## References


[1] TRAI. Press Release June11. Information Note to the Press, August 2011. http://www.trai.gov.in/WriteReadData/trai/upload/PressReleases/835/Press\%20Release\%20June11.pdf.

[2] http://otal.umd.edu/UUGuide/wmk/

[3] Lerer, A., Ward, M., and Amarasinghe, S. Evaluation of IVR data collection UIs for untrained rural users. In *Proceedings of the First ACM Symposium on Computing for Development*, ACM DEV '10, ACM (2010), 2:1–2:8.

[4] Patel, N., Chittamuru, D., Jain, A., Dave, P., and Parikh, T. S. Avaaj otalo: a field study of an interactive voice forum for small farmers in rural india. In *Proceedings of the 28th international conference on Human factors in computing systems*, CHI '10, ACM (New York, NY, USA, 2010), 733–742.

[5] Lee, K. M., and Lai, J. C. Speech versus touch: A comparative study of the use of speech and dtmf keypad for navigation. International Journal of Human Computer Interaction 19, 3 (2005), 343360.