Tring! Tring! - An Exploration and Analysis of Interactive Voice Response Systems

Siddhartha Asthana, Pushpendra Singh, Ponnurangam Kumaraguru, Amarjeet Singh, Vinayak Naik
Indraprastha Institute of Information Technology
New Delhi, India
(siddharthaa, psingh, pk, amarjeet, naik)@iiitd.ac.in

ABSTRACT

In developing regions like India, voice based telecommunication services are one of the most appropriate medium for information dissemination as they overcome prevalent low literacy rate. However, voice based Interactive Voice Response (IVR) systems are still not exploited to their full potential and are commonly considered as frustrating to use. We did a real world experiment to investigate the usability issues of a voice based system. In this paper, we report analysis of our experimental IVR and interface difficulties as experienced by the user. We also highlight the user behavior towards accessing critical and non-critical information over multiple information media vis-a-vis IVR, web and talking to a human on the phone. The findings suggests that an IVR which can adapt its behavior will prove to be more efficient and provide a better user experience. We believe that our results can be used for efficient development of next-generation adaptable IVR systems.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: Voice I/0 ; H.1.2 User/Machine Systems:Human factors

General Terms

Design, Experimentation, Human Factors

Keywords

Interactive Voice Response (IVR), Real-world experiment, India, Information dissemination, Study in wild

1. INTRODUCTION

Telecommunication services are a vital medium for information exchange and communication. They have a huge impact in developing countries, where reach of the Internet connection is relatively low compared to developed countries. At present, there are more than 5 billion mobile phone users [1] compared to 2 billion Internet [8] users in the entire world. According to 2011 statistics released by Telecom Regulatory Authority of India (TRAI), there are 885 million [17] land-line and mobile connections in India whereas the count of the Internet users is around 100 million [8]. Similar trends can be seen in other BRIC[1] nations [2]. Brazil, Russia and China have less than 42% of the total population connected to the Internet. The number of subscribers of cellular mobile in terms of total population for Brazil, Russia and China are 90.5%, 162.4% and 64.4% respectively. These figures clearly show high penetration of telecommunication services in the developing countries as compared to the Internet. Thus in developing regions, a phone based medium has much more reachability than the Internet. Apart from reachability, any technological intervention for large scale information dissemination also has to deal with the literacy rate of the end users. Researchers have shown that voice based systems can overcome the barrier of low literacy, as voice is natural and an accessible medium for many people who often have limited formal education [10]. Additionally, any voice based information dissemination system, accessible through regular phones, will easily penetrate in existing infrastructure and will seamlessly integrate with what people routinely use.

Use of Interactive Voice Response (IVR) has already been advocated for information dissemination in developing regions [6, 10]. IVR is a voice based system that is accessible through any mobile or land-line phone. IVR has been extensively used in industry for call automation while providing customer service. It is also argued that IVR is easier to use when compared to the Internet, as Internet use requires certain skills and training. In comparison to traditional information dissemination media like Television or Radio, communication on traditional media is passive whereas IVR enables interactivity for active communication.

Present-day IVR systems are developed and deployed with an information database at the back-end. The information is made available to the end-users over the interface of simple key presses or voice interaction using a pre-defined menu structure to answer user queries. This menu structure and the information that it extracts from the database usually remain fixed throughout the life-time of an IVR system and any change requires manual intervention. Every information, provided by the 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. For example, in case of the Indian Railway IVR system, the menu structure to access information remains fixed through 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

[1]BRIC refers to Brazil, Russia, India and China.

pattern.

In spite of significant work to improve the interaction with the IVR systems, this technology has not been utilized to its full potential. Most of the work in this area is confined to studying a group of users and proposing a system solution that is common for everyone [10, 14, 15]. By common system solution, we mean that all callers[2] are handled in the same way regardless of their knowledge, experience, navigation skills and willingness to use the automated system. However, we argue that every individual has a unique set of properties that differentiate him/her from the other individuals and the same holds true for the caller of a voice based application. A human operator is good at handling these intricacies and is therefore typically preferred over the IVR [3].

Similarly, the nature of information sought from a system by its users change with time and also depends on external contextual factors. For example; FAQs database on the sites of colleges and universities websites provide answers to frequent questions that a prospective student may have in mind. The FAQs grow with time and usually end up as a long list of questions and associated answers. However the nature of questions, that are accessed at a given time, also depend on external factors; for example, before the entrance exam for admissions, queries are mostly related to entrance exam dates, syllabus, eligibility criteria etc. while after the exam, questions change for results, fees structure, hostel accommodation 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 systems 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 waste in time adds to frustration of the user. In current IVR systems, such shortcoming can 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. Here we are going farther than single user intricacies and talking about changes in global pattern of usage. This leads us to requirements for global or system-wide adaptation of IVR systems to take into account external factors affecting experience/behaviors of all or a set of users.

Both the issues, single user intricacies and global pattern of usage, motivate us to study, in detail, the design issues of IVR systems that can accommodate individual differences as well as global usage pattern leading towards better user experience. To explore this idea, we conducted a real world experiment in an academic institute in Delhi (hereafter referred to as IIIT-Delhi or IIIT-D). We designed, implemented and deployed an IVR system at IIIT-Delhi campus for admissions in undergraduate courses for the year 2011. The experiment helped us in understanding the behavior of users of IVR system which can be crucial in designing next-generation IVR systems.

In this paper, we present analysis of usage of our IVR system used by 2,211 applicants (referred as users in the rest of the paper) over 13 days of deployment. Since multitudes of information was accessible through multiple mediums - IVR, web interface and talking to a call executive over phone, we also present the comparative analysis of information dissemination through multiple mediums. Based on the results, we propose a hypothesis for an adaptable IVR system that can change itself both based on specific caller requirements for interaction as well as global pattern of usage.

The remainder of the paper is organized as follows: In the next section, we describe related work on IVR. We then describe the design, features, and implementation of IIIT-D IVR system. Deployment section outlines the IIIT-D IVR deployment, followed by the results and observations: we describe statistics for the usage of IVR content and difficulty in accessing the information over the IVR in this section. Based on these findings, we discuss some design implications for an IVR system with improved user experience. We conclude with directions for future work.

2. RELATED WORK

Studies in the area of IVR and automated telephone services have proposed various approaches to improve the usability of the system. Perugini et al. [11] discussed about three design dimensions for automated telephone services that could be used to study design issues in IVR. These three design dimensions in their conceptual design space were, nature of the user in terms of addressable[3] input (in-turn vs. out-of-turn), input modality (touch vs. text vs. voice), and interaction style (menu-based vs. natural language). In their work, they studied *out-of-turn interaction:* a different nature of user addressable input. Their work reflected user's model of the task while navigating through the menus. They showed that this type of approach is only possible for an IVR with speech recognition. However, its use may be limited in developing regions wherein speech technology is not advanced for native languages.

IVR system's input modalities for VUI (Voice User Interfaces) have also been extensively explored. Lee and Lai [5] compared dial interfaces with speech and showed that dial interfaces are preferred for linear tasks and speech is preferred for non-linear tasks. Patel et al. [9] showed that dial interface outperformed speech in terms of task completion rate and learning ability. They also showed that dial interface was relatively easier to use than speech.

Navigation is an essential component of any menu based IVR. Multiple studies have proposed their own methods to improve navigation in menu based IVR. Skip and Scan is one such navigation approach for menu based IVR, wherein caller could easily navigate back and forth through menus without first listening to all of the prompts for a particular menu [12]. Zap and Zoom [4] is another proposed approach for navigation in menu based IVR, that improves over Skip and Scan approach by allowing users to jump directly to a location in IVR using short-cuts. However, Zap and Zoom requires that the caller should be aware of options in IVR menu beforehand otherwise they will not be able to take the decision of the location they want to jump upon in the IVR menu.

Apart from technological aspects, human factors also impact system design significantly. Grover et al. [14] proposed a dialog design model for low literate users. They discussed that socio-cultural and domestic environment of users may affect the usability of IVR

[2]Callers, users and applicants are interchangeably used in this research.

[3]By addressable information authors mean the information which the system can accept from the user or in other words, the information that the user can supply. They do not mean information that the system indexes (addresses).

systems. Some users preferred DTMF for privacy reasons. They also emphasized that techniques like speech recognition and profiling cannot work in cases where mobile sets have multiple ownerships . In a country like India, people generally share mobile phones among family members therefore leading to common existence of multiple ownership. BlindSight [7], a prototype voice application for mobile phones allows users to achieve eyes free (using the phone without looking at it) error rate below 5%. The experiment revealed that overhead for eyes-free use of mobile phone key pad is only 200 ms per keystroke compared to sighted use. Bernhard Suhm [16] studied cognitive limitations relevant to voice user interface design.

IVR technology has also been used in the developing world to study different aspects of system design. Patel et al. [10] studied the use of IVR as a social media with traditional media as radio. Sherwani et al. [15] deployed phone based speech interface to access health information by low-literate community health workers. Sambasivan et al. [13] designed a phone broadcasting system for urban sex workers in India.

So far, most of the studies are confined to proposing solutions by understanding a group of users. To the best of our knowledge, IVR systems are not designed to differentiate and adapt to individual caller requirements. To fill the gap in system design to accommodate individual requirement, in this study, we explore the user behavior for information access over IVR and try to compare it with web and interaction with human (or call executive).

3. SYSTEM BACKGROUND

IIIT-Delhi (or IIIT-D) is a state university in New Delhi, India. For the academic session 2011-12, the annual intake in an undergraduate course (for which the IVR was developed and deployed) was of 120 students. This study was conducted for the session 2011, which had 2,211 applicants for the admission. Applicants appeared for a written exam based on which a merit list was created. All the relevant information regarding admissions was provided on institute's official website. To resolve any queries, applicants were also provided with a telephone number on institute's official website. A call executive was appointed to answer these queries during 9 AM to 5 PM. Some of the calls went unattended as the call executive was busy in some other work during the office hours of 9 AM to 5 PM. There were also significant number of unattended calls from people calling in non-office hours. We deployed IIIT-D IVR with an objective to get real world data of IVR usage.We took explicit permission from our call executive to record her calls.

4. DESIGN PROCESS

The information content for the IIIT-D IVR was prepared with the help of experienced and proficient staff of the admission department of IIIT-D. To ensure the validity of information content and usability of the system, IIIT-D IVR went through a series of demonstration to the individuals involved in the admission and academic processes. Voice quality and information content were the primary issues reported during the demonstrations. The initial system was tested with various Text to Speech (TTS) engines for delivering the information content of the system. We tried our system with the voices available in Cepstral, TextAloud, AT&T Natural Voices and Microsoft Voice. Due to unsatisfactory performance of voices available in TTS (both foreign as well as Indian accent voices), we ultimately opted for human recorded voice. We tested our system with 3 human (2 male and 1 female) voices. After taking inputs from multiple people on different aspects of voice, including clarity, we decided to use the female voice for our deployment. The information content was made short and concise, which also helped in optimizing the time duration for each call. All information presented through the IVR was supported with appropriate URLs (announced at the end of voice prompt), to facilitate accessing the specific content in detail on the IIIT-D website.

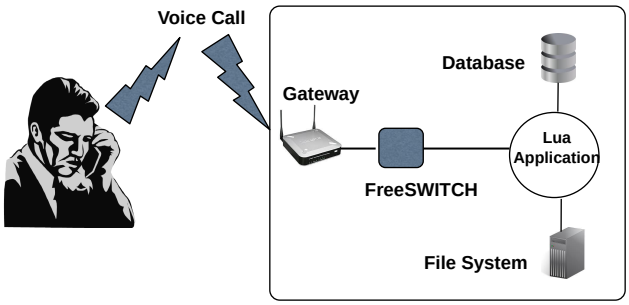

Figure 1: A user accessing information over IVR: user dials the number of IVR on his phone. The call is first received by the gateway and is then forwarded to FreeSWITCH (Telephony platform). An application written in Lua answers the call by playing appropriate voice prompts. Lua application accesses database and file-system for storing and retrieving information.

5. IMPLEMENTATION

IIIT-D IVR was implemented as a Voice application written in Lua[4] and hosted on FreeSWITCH (an open source telephony platform). Applicants calling to IIIT-D were connected through a gateway as shown in Figure 1. Linksys SPA 3102, single line device with capability to handle one call at a time, was used as the gateway. The gateway converted the analog signal of telephone line to the Session Initiation Protocol (SIP) and forwarded the request to the FreeSWITCH. User input in the form of Dual-tone multi-frequency (DTMF) key presses were forwarded by the FreeSWITCH to the application written in Lua.

6. DEPLOYMENT

After deploying the IVR in IIIT-D, we conducted the study in two phases for a total of 13 days (6 days for first phase and 7 days for second phase). The first phase was pre-examination phase. In this phase, IVR presented information related to examination and online application process. The second phase was Post-examination phase and it was used to deliver results of the entrance examination along with other information like counseling schedule and fee structure.

The pre-examination phase was set up to handle calls in the absence of a human attendee. Institute published a telephone number on its official website to resolve the queries of applicants related to admission procedure. For this purpose a call executive was also appointed to respond to those queries. We integrated our IVR system with the telephone line of the call executive. Callers were not aware about existence of IVR a priori. The call was transfered to

[4] http://www.lua.org/

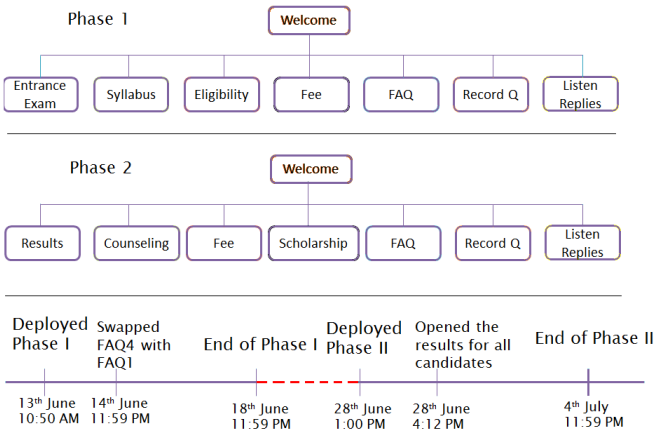

Figure 2: Menu options available in each phase of IVR (1^{st} phase corresponds to pre-examination and 2^{nd} phase corresponds to Post-examination). The time-line shows the modification done in the system during the experiment.

IVR only if it was not attended by the call executive within the first 16 seconds.

Figure 2 shows the menu structure of the IVR system deployed at IIIT-D together with the changes we performed in the IVR with time. Pre-examination phase started on 13^{th} June 10:50 AM and ended on 18^{th} June 11:59 PM.[5] We continuously monitored the call recording as well as interviewed the call executive on a daily basis. Based on the feedback from call executive, who specified that majority of the received queries were related to admit card, we made the first change in the running experiment on 14^{th} night. The 4^{th} item in the FAQ menu that was related to admit card queries was swapped with the 1^{st} item in the FAQ menu. Written exam for the admission process at IIIT-D was conducted on 19^{th} June, that marked the end of first phase.

The second phase started with results declaration on 28^{th} June. Results were declared, simultaneously on web and IVR, only for the selected candidates. Selected candidates comprised of top 247 ranks amongst the total of 2,211 candidates who appeared in the exam. In this phase, we used a separate telephone line for IVR and call executive. There was no option to call the executive and get the results checked by her. Unlike the first phase wherein no specific phone number for IVR was mentioned (and the executive line defaulted to IVR when not attended in time), this time the telephone number to access information through IVR was explicitly published on the official website of IIIT-D. In the first few hours of second phase, call executive received huge number of telephonic queries related to the result of unselected applicants. Due to this, we made a change, both on web and IVR on 28^{th} June at 4:12 PM wherein we declared the results of all the candidates, providing their ranks for all the students who appeared in the exam which earlier were restricted to selected candidates.

7. DATA COLLECTION

We collected data from five sources:

- *Log of system navigation* - We recorded every interaction between the IVR system and the caller. It included DTMF key press along with the time stamp and the voice prompts played by system.
- *Audio recording of calls* - We performed audio recording for all the calls to capture the callers' interaction with the call executive. From these recordings the context of the call was analyzed to understand if any systemic changes are required in the IVR system.
- *Log of Linksys SPA3102* - We collected log of PSTN gateway (Linksys) that includes information on call duration, call forwarding and caller ID.
- *Web log* - We logged every web request for results received on our website. It includes the 9-digit application number and the web server time at which request was made.
- *Survey* - We conducted a post-study survey to collect responses for user experiences with IIIT-D IVR. People who came for the counseling[6] participated in this survey.

We also logged various parameters associated with the call. These include time-stamp of DTMF key-presses, hangup-cause of calls, duration of ring before the call was picked up by IVR or call executive, among others. The telephony platform (FreeSWITCH) used by us had 7 different levels of logging. Log levels are from most critical to least critical. We enabled the most detailed level of logging in FreeSWITCH that helped us to log all the events. We also collected a detailed log of every SIP packet exchange between PSTN gateway and FreeSWITCH. Detailed logs from multiple sources enabled cross validation of logged parameters.

8. STUDY FINDINGS

Pre-examination phase involved comparison of information access on IVR and call executive. Post-examination phase involved comparison of information access over IVR and web interface. In this section we present findings from our experiment that helped us to form new hypotheses.

8.1 Traffic Overview

In the pre-examination phase, 421 calls were made to IIIT-D telephone line in a span of 6 days. Thirty-three calls went unattended because neither the call executive nor IVR picked the call as the call executive was busy and the caller disconnected the call within 16 seconds. For the remaining 388 calls, 180 calls were attended by call executive and 206 calls were handled by the IVR. Two call recordings were corrupt due to malfunctioning of the system.

Post-examination phase was focused towards delivering the results of applicants and provided other useful information like counseling schedule, fee structure etc. Over the seven days of deployment, 405 calls were made to IVR and 42,420 requests were received over the web. Figure 5 shows breakdown of 405 calls based upon selected user input on IVR (wherein "Other" correspond to an option other than checking for results was selected by the caller). As expected, the option for checking the result was most accessed among the IVR menu in this phase. It was interesting that 42,420 requests were received on the web interface for getting the results for 2,211 applicants. This indicates that on an average people have accessed the result for more than 20 times.

[5]All times mentioned in this paper are in IST (Indian Standard Time).

[6]Counseling is a part of admission process where an applicant gets admission in a course based on merit list.

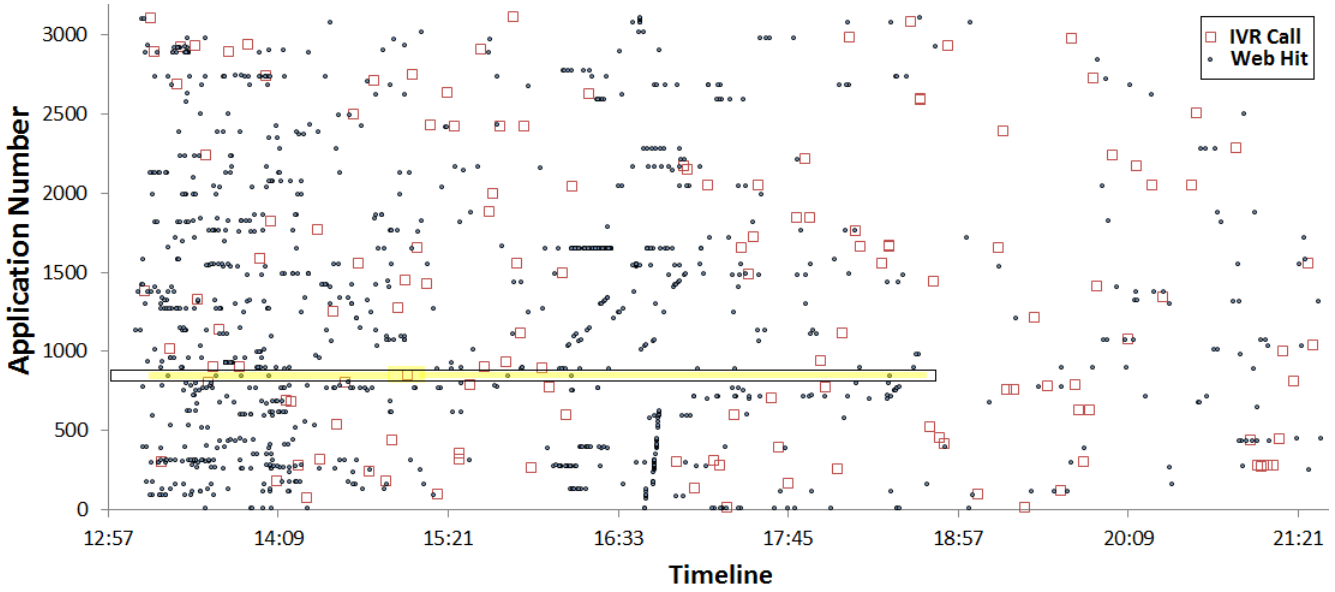

Figure 3: Pattern for access of information on the first day of Post-examination Phase for both web and IVR interface. Y-axis represent anonymized application number. It is evident that there were multiple access by same application number. Multiple access of information for anonymized application number 850 is shown in shaded region.

8.2 IVR Content and Usage

As shown in Figure 2, IVR had seven options in each phase; the first five options contained information specific to the corresponding topic.The sixth option allowed a user to record his / her question if the user could not find the sought after information on the IVR. The seventh option, *Listen Replies*, allowed the IVR caller to listen to any response that the call executive may have recorded for the previously asked query in the system. The order of the first five options presented in the IVR menu was based on the discussion with the administrative staff experienced with handling admission procedure for the past three years. The objective was to present the options in order of their relevance as suggested by the experienced staff. This ordering made us believe that the number of hits (from callers calling into IVR) on each option will be in order of their position in the menu structure.

Option	1	2	3	4	5	6
Phase1	62	17	10	8	32	18
Phase2	284	16	5	4	7	10

Table 1: Number of hits on each option in first and second phase of IVR. Fifth menu item in pre-examination phase of IVR had more hits than the second, the third and the fourth option.

Table 1 shows the number of hits on each menu item in each of the two phases. Contrary to our earlier belief, fifth menu item in pre-examination phase of IVR had more hits than the second, the third and the fourth option. Similarly fifth option in post-examination phase has more hits than the third and the fourth item. This shows that in spite of taking inputs from experienced staff, the order of relevance based upon human belief (from prior experience) was different from what was experienced by accurately logging the system usage. It also suggests that it is not easy to predict the correct order of options and it may make sense for the system to learn it based on past usage in an automated manner. In the discussion section, we highlight the importance of having correct order of options.

In order to get information access pattern, we logged every interaction of the caller with the IVR. We also analyzed the call recordings to know about information received by the caller from the call executive. For repeated callers, we observed that once they got a specific piece of information from call executive they did not try to access the same over the IVR in their successive calls. We also observed that callers did not ask the call executive for the information they had already received over the IVR in their previous calls. This leads to an assumption that once the caller was informed about any piece of information through any medium (call executive or IVR), they will not try to access it again or over another medium. Implications of this assumption may impact the relevance of an information for a repeated caller - it may be assumed that once an option has been accessed, it will not be accessed again. However, the analysis of Post-examination phase provides us contrasting outcomes (discussed next).

Post-examination phase was focused towards delivering the results of applicants that can be considered as the most critical piece of information in the admission process. Applicants had the option to access their result both on web and IVR. We logged every request for results. From the logged data, we observed that applicants checked their results multiple times and through multiple interfaces vis-a-vis on web and on IVR. Figure 3 shows requests for results made by those applicants who accessed this information on both web and IVR interface on the first day of Post-examination Phase with a time line. Y-axis represents anonymized application number. The data points encircled in the Figure 3 shows that an applicant with application number “850" has made 4 web request followed by 1 IVR call and again 3 web requests. In our dataset we found that 189 application numbers were common among queries made by users on IVR and web interface. It was also interesting to observe that there were 8 applicants who checked their result on IVR only.

Figure 4 shows the trend over the seven days among 189 applicant who accessed their result both on web and IVR. It shows high

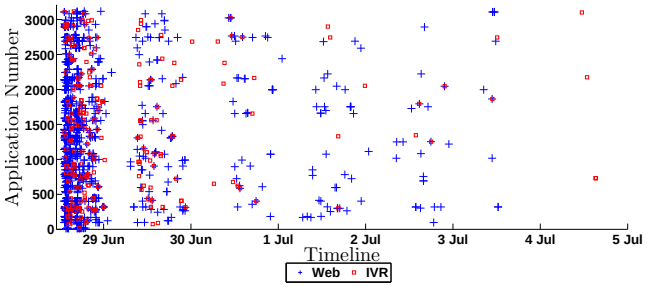

Figure 4: Common result seekers: Applicants who accessed the information both on IVR and web interface.

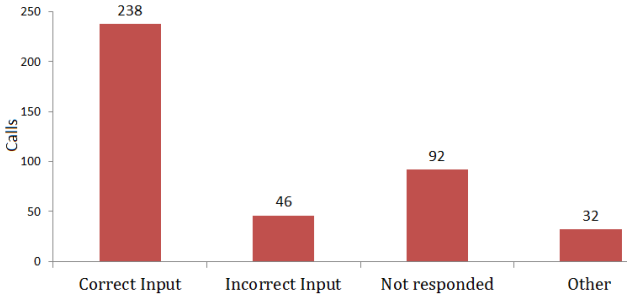

Figure 5: Key press accuracy: 238 Callers correctly able to input the 9 digit application number to access there results where 46 callers were not able to input the same in given time limit.

amount of traffic on the first day. Figure 3 displays the detailed information for these 189 applicants for the first day. This data helped us to capture user behavior towards critical information like results. It shows that people have a tendency to cross check a critical piece of information like results through multiple sources.This behavior is contrary to the behavior observed in the pre-examination phase. Based on these two observations (pre-exam and post result) we may hypothesize that a critical piece of information is accessed multiple times whereas a non-critical piece of information is not accessed multiple times via available multiple mediums (i.e. Web and IVR).

As shown in Figure 2, we made a small change in the information content in post-examination phase on 28^{th} June at 4:12 PM. Initially the results announcement was confined to successful candidates (247 out of the total of 2,211 who appeared). Unsuccessful candidates were told as “ineligible for counseling". We observed a different pattern of access among successful and unsuccessful candidates. The successful applicants accessed their result on an average 6.2 times compared to 1.7 times of unsuccessful applicants. This may be because successful applicants took pride in the result or felt happy in accessing the information or shared it with others also. Repeated access of same information does indicate that once such information is identified, it should be handled understanding that it will be accessed repeatedly.

8.3 Input error

In the post-examination phase, a user had to enter a 9-digit application number in an assigned time limit to access the results. We also analyzed the usage pattern for correct and incorrect inputs as provided by the users. We observed that many people were not able to enter their application number due to frequent time-out. Call distribution in terms of key press accuracy of caller is shown in Figure 5. Out of 284 callers who chose the option to check their results from the IVR, 238 callers were able to check their results. The remaining 46 callers tried multiple times to enter their 9-digit application but failed because of time constraints. We set a reasonable time limit of 3 seconds inter-digit time-out for pressing the DTMF key for inputting the application number. It was thoroughly tested while system testing and was found to be enough (also evident from the fact that 238 callers out of 284 callers were successful in checking their results). An important thing to note is that although the time limit was reasonably good but once the time limit is set it is a hard time constraint and all the callers had to respond within that time limit. Even though callers failed multiple times to enter the input correctly the system did not adapt to relax the time constraint for repeatedly failing callers. If the system had adapted itself to allow longer time for entering the application number, throughput of IVR would have increased.

8.4 Survey

We conducted a survey among 122 people who attended the counseling (next phase of getting admissions) at IIIT-D. These were the applicants along with parents attending the counseling. The survey comprised of 10 Questions, none of which was compulsory to answer. We asked them to rate IIIT-D IVR on various parameters on a 7-point Likert scale. The summary of the responses are shown in Table 2.

Likert Scale	**1**	**2**	**3**	**4**	**5**	**6**	**7**
User friendliness 1: Very easy 7: Very hard	9	18	13	7	4	4	0
Voice clarity 1: Very Good 7: Very Bad	11	16	11	5	8	3	0
Comparison for ease 1: IVR is easy 7: Internet is easy	6	4	9	10	7	13	26

Table 2: Survey data: Number of survey participant rated IVR for parameter mentioned in leftmost column on a 7-point Likert scale. Below each parameter the description of 1^{st} and 7^{th} point of Likert scale is also mentioned.

One of the questions in the survey was, which information source will user trust more in case of contradicting information available on different sources. In response to this question, 37 people voted for the Internet, 34 for Human, 13 for IVR and 8 for the medium where information is most pleasing to them.

Given that our target population was from a metropolitan city and had easy access to Internet. It was good to know that still IVR system was used and appreciated by some of them. We believe that in case of users (rural or illiterate) with poor access to computers, support for IVR will increase.

9. OBSERVATIONS

In this section we report some other observations from the data collected during our experiment.

Pre-examination Phase (IVR vs. Human)

In this phase, callers were not aware about the existence of IVR

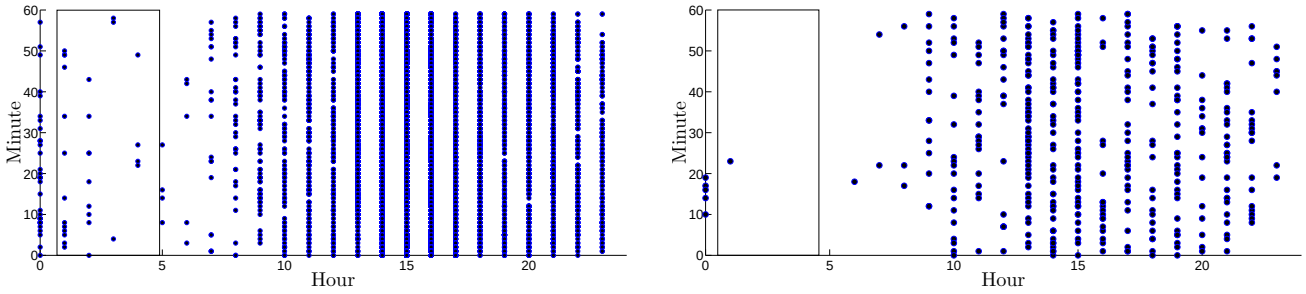

Figure 6: Time of the day for web requests and IVR calls. Left: A data symbol at co-ordinate (x,y) represents a web request received at x Hour and y minute. Queries on the web interface even during the night time. Right: A data symbol at co-ordinate (x,y) represents a IVR call received at x Hour and y minute. Queries on IVR during the night time were close to zero showing that people tend to not use the telecommunication interface during this time period.

prior to calling. We observed that several people disconnected the call when they encountered the IVR. An intuitive hypothesis is that a caller calling during the office hours will most likely expect the call to be picked up by a call executive and the call to be picked up by an IVR during non office hours. Higher expectation of talking to a call executive will most likely result in higher percentage of early disconnects of IVR calls during office hours than in non-office hours. Table 3 shows the comparison of early disconnected calls during office and non-office hours. Forty-four out of 145 calls (30.3%) were disconnected early by callers when they encountered an IVR during the office hours whereas 19 out of 61 calls (31.1%) were disconnected early by callers during the non-office hours. It shows insignificant difference in early disconnected calls between office and non-office hours proving the intuition wrong. However, it is clear that majority of users disconnected without even trying to get the information, this does indicate that the present-day IVR systems (if presented without making users aware) do not inspire confidence in users to use them. It may come from past experience of users of interacting with different IVR systems.

	Office hours	**Non-office hours**
Early disconnected call	44 (30.3%)	19 (31.1%)
Other calls	101	42
Total calls	145	61

Table 3: Comparison of early disconnected calls during office and non-office hours. Table shows insignificant difference in early disconnected calls between office and non-office hours.

We analyzed the call recordings of IVR and call executive to identify the calls that were related to admission in undergraduate course. Out of the total of 421 calls received during this phase, we identified 264 legitimate calls that were related to the admission process. The rest 157 calls were either not related to B.Tech. admissions or disconnected too early. Table 4 shows the distribution of calls attended by call executive and IVR.

	Office hours	**Non-office hours**
IVR	102 (45.9%)	42
Call executive	120 (54.1%)	NA
Total calls	222	42

Table 4: IVR Vs Call executive: Distribution of calls received in office and non-office hours.

Out of the 264 legitimate calls, 144 (54.5%) calls were attended by IVR and 120 calls by call executive. During the office hours, the IVR received 102 (45.9%) calls and call executive received 120 (54.1%) calls out of the total 222 calls made during this time period. Forty-two (29.1%) calls out of the total of 144 received by IVR were during the non-office hours. Therefore, IVR not only helped in reducing the number of unattended calls both during office and non-office hours but also reduced the burden of the call executive during office hours.

Post-examination Phase (IVR vs. Web)

This Phase was primarily focused towards delivering the results of the examination. We received a total of 405 calls on IVR and 42,420 queries on web interface during the 7 days of experiment in this phase. We observed that for 238 out of 405 calls, callers were successfully able to check their results on IVR. On analyzing the remaining 167 calls, we observed that people either did not respond or were not able to correctly input the requisite details (application number) to fetch the desired information. On the web interface, we received 42,420 requests out of which 1,281 were invalid queries due to incorrect input by the user.

Ninety-Five percent of queries on web interface were received on the first day itself whereas on IVR we received 56% of the total calls on the first day. Looking closely, Figure 6 presents the traffic on the web and IVR interface only for the first day of Post-examination phase. In both these figures, X-Axis represents Hour value of the timing instant and Y-Axis represents Minute value of timing instant. A diamond symbol at co-ordinate (x,y) represents a web request and IVR call received at x Hour and y minute. The rectangular regions in Figure 6 represents the time of 1 AM to 4 AM.

A key observation from Figure 6 is how people adopt different ethics for usage of technology. Queries on IVR during the night time were close to zero showing that people tend to not use the telecommunication interface during this time period. On the other hand, several people checked their result on the web interface even during the night time. One may argue that the number of data points for the IVR and the web interface are different and that may have resulted in different usage pattern during the night time. However, keeping into consideration that the traffic on both the interfaces was generated by the same population of users (those checking their examination results) and there is no resource availability problem for IVR at the night time (the line is mostly available), different traf-

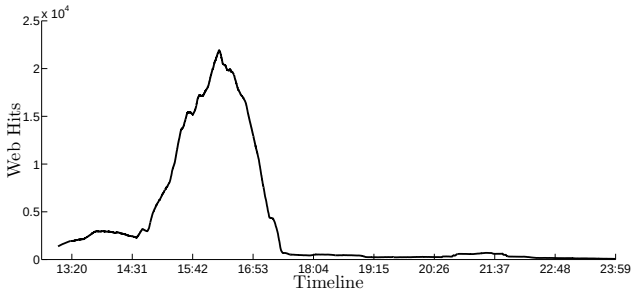

Figure 7: Traffic load on web: A data point at coordinate (x,y) represents y number of hits on web in duration of one hour in x $\pm$ 30 minutes.

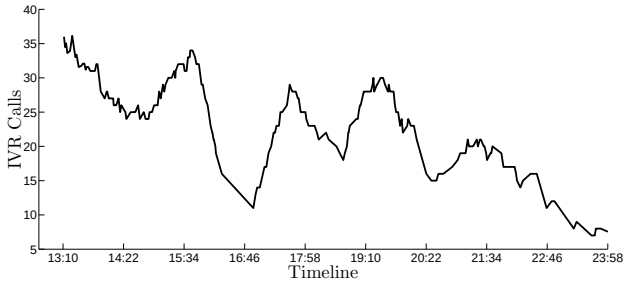

Figure 8: Traffic load on IVR: A data point at coordinate (x,y) represents y number of calls on IVR in duration of one hour in x $\pm$ 30 minutes.

fic pattern on web and IVR interface point towards different ethics for usage of web and telecommunication channels for information access.

Let us now look closely at the traffic pattern for the first day in the Post-examination phase when most of the queries were received for both the web and IVR interface. Figure 7 and 8 capture the load of web and IVR traffic respectively for the first day.

We released our results on web and IVR simultaneously around 1 PM on 28^{th} June 2011. It is interesting to observe that for the web interface there exists a single peak in the traffic load pattern. However, for IVR interface, we observe multiple peaks in the traffic load. It is important to note that for the web interface unlimited number of parallel connections are available whereas for the IVR interface, only a single telephone line is available for all the users. As the news of examination results spread, traffic on the web (with unlimited parallel connections) increases. After a certain time (when most people would have checked their results), the traffic load on the web interface decreases. It is, therefore, intuitive to have a single peak in the web interface traffic load.

With a single connection available in the case of IVR interface, when people started making calls over IVR, an existing call will block other users from being able to access the information on IVR as well. Due to lack of free available channel over IVR people may start backing off from it, with an intention to try calling in again after a certain delay. This backing off leads to a fall in the traffic to the IVR. Further, as a result of this backing off load on the IVR channel reduces and the channel will again become more accessible leading to a rise in traffic. This back-off phenomenon intuitively explains the existence of multiple peaks in the IVR traffic load. As a result, one can argue that with multiple channels of IVR being available to the users, one would have probably seen a similar traffic pattern on the IVR interface as on the web interface. Extending this argument further, such an analysis can be used by businesses relying on the IVR interface for customer service to estimate the number of parallel channels required to optimally serve their clients.

10. DISCUSSION

In this section we discuss some of the possible implications of our findings from the experiment and the analysis performed thereof. We try to formulate hypothesis for the design of next-generation IVR systems, highlighting possibly new research challenges in this domain.

In the previous section, we showed that relying on human feedback for predicting the relevance order of information, to decide on the menu structure of IVR is not straight forward. Our results demonstrate that an automated analysis of traffic pattern and adaptation will probably give a better relevance order than relying on human expertise. IVR systems have been typically found to be frustrating as they may provide information that is not relevant in the current context of the call, thus requiring higher amount of time to deliver the relevant information.

This problem can be addressed by an adaptive IVR that adapts to every caller as well as global pattern of usage, e.g. by rearranging the information based on the order of relevance as learned by the system.

10.1 Order of relevance

We now discuss,, how rearranging the menu items in the order of relevance in an IVR system can help reduce the amount of time spent by the user. Table 5 gives the exact number of hits on each of the menu item in our experimental IVR during the pre-examination phase. If suppose for every menu item, the IVR takes *c* seconds to announce about it (while delivering the complete menu structure) and *d* seconds to play specific information contained inside the item to the caller, then minimum amount of time required by ordering of option as mentioned in Table 5 will be

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

Table 5: Number of hits on each option in first phase.

$$62(c+d) + 17(2c+d) + 10(3c+d) + 8(4c+d) \\ +32(5c+d) + 18(6c+d) = 147d + 426c$$

However, if instead the IVR had realigned the menu items in the decreasing order of their hits, resulting minimum amount of time required will be

$$62(c+d) + 32(2c+d) + 18(3c+d) + 17(4c+d) \\ +10(5c+d) + 8(6c+d) = 147d + 346c$$

Therefore, a simple adaptation, of automated rearrangement of the menu items in the order of their relevance, by IVR could possibly lead to significant time saving. Similarly we have observed that very few people listened to the non-critical information option that

they had already listened to in their previous encounters with the IVR system. An adaptable IVR can put such information at the end of the menu structure. Further, adaptable IVR systems can also account for accuracy of the key-press. We discussed earlier that 46 callers were not able to fetch the results because of time constraint. This happened in spite of rigorous testing priori with the IVR to ensure that enough time is given to input the application number which is also evident by 238 callers being able to successfully input the application number. For the unsuccessful callers, the system could have adapted by increasing the time so as to allow them more time the next time they try to call.

Based on above discussion we hypothesize that an adaptable IVR will prove to be better than present-day IVR systems and adaptation should be an integral part of design in next-generation IVR systems. What the actual adaptation would be will depend on the type of IVR and will vary from one IVR system to another IVR system.

10.2 Take away from the study

Automated rearrangement of the menu items in the order of their relevance can form a feature of adaptable IVR system. Order of relevance for a caller can vary from one another based on whether the caller is a repeated caller or the first time caller.The rearrangement can be based upon number of hits historically received by a given menu item. The option receiving more hits can be presented to the caller before other options in IVR menu.

Further analysis of our data showed that among the repeated callers, very few people listened to the non-critical information option that they had already listened. This can form additional input while rearranging the menu items wherein for a repeated caller the options previously listened to (for such a service primarily meant for delivering static information) by the caller can be put at the end of the menu structure. To this effect, we propose the following guidelines:

- Divide the menu items of IVR in two information set of *accessed* and *never-accessed*.
- Individually calculate the order of relevance for each set based on the global order of relevance calculated as per the historical usage of IVR system (not specific to any user).
- Order of relevance for a repeated and non-repeated caller therefore should be *never-accessed* in its order of relevance followed by *accessed* in its order of relevance and vice versa, respectively. In case of first time caller menu item falling in set *accessed* will be a null set.

It is important to note that the above guideline is proposed for delivery of static information content over IVR.

11. LIMITATIONS AND FUTURE WORK

We would also like to mention some of the limitations of our experimental study:

- We made two changes in the system (one in each phase) to better serve the end users (considering the criticality of the live information dissemination system catering to admission process of the institute) that may have had an effect on usage of the system.
- We only used a single telephone line for our IVR system that may have blocked several callers trying simultaneously. Outcomes of the study could have been more enriching if we had used multiple telephone lines.
- In our analysis we assumed that request corresponding to each application number is from the same individual, which may or may not be the case in reality. However, such assumption does not critically affect any of the analysis or discussion mentioned in the paper.

In the future, we plan to collect more data to overcome several of the limitations mentioned earlier. We aim to build IVR systems which have ease of access like the Internet while adaptable enough to match human-like ability in dealing with caller intricacies. Future experimental studies will be in the direction to develop and validate new hypothesis for improved usability of IVR systems based upon self-adaptability of these systems. We also plan to extend our study to the admission process next year and validate several of our findings in the live system used for the same scenario.

12. ACKNOWLEDGMENT

The research of first author is supported by TCS (Tata Consultancy Service) Research Scholarship. We appreciate the effort of Ramjot Singh, Akashdeep and Pooja in helping us develop the IVR system. We express our thanks to Niharika Sachdeva for providing audio prompts in her voice, played on IVR. We would also like to thank Aditi Gupta and all members of PreCog research group and MUC group at IIIT-Delhi for their valuable inputs and immense support.

13. REFERENCES

[1] Global mobile statistics 2011. Article, July 2011. http://mobithinking.com/mobile-marketing-tools/latest-mobile-stats#subscribers.

[2] National Bureau of Statistics of China, Joint Statistical Publication by BRICS Countries 2011, April 2011. http://www.stats.gov.cn/english/statisticaldata/otherdata/brics2011/P020110412517148811510.pdf.

[3] Dean, D. H. What's wrong with IVR self-service. *Managing Service Quality 18* (2008), 594–609.

[4] Hornstein, T. Telephone Voice Interfaces on the Cheap. In *Union Bank of Switzerland, UVK Informatik* (Zurich, 1994), 134–146.

[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), 343–360.

[6] 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.

[7] Li, K. A., Baudisch, P., and Hinckley, K. Blindsight: eyes-free access to mobile phones. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, ACM (New York, NY, USA, 2008), 1389–1398.

[8] mobiThinking. Internet world stats. Article, July 2011. http://www.internetworldstats.com/top20.htm.

[9] Patel, N., Agarwal, S., Rajput, N., Nanavati, A., Dave, P., and Parikh, T. S. A comparative study of speech and dialed input

voice interfaces in rural India. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, ACM (2009), 51–54.

[10] 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.

[11] Perugini, S., Anderson, T. J., and Moroney, W. F. A study of out-of-turn interaction in menu-based, IVR, voicemail systems. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '07, ACM (New York, NY, USA, 2007), 961–970.

[12] Resnick, P., and Virzi, R. A. Skip and scan: cleaning up telephone interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '92, ACM (New York, NY, USA, 1992), 419–426.

[13] Sambasivan, N., Weber, J., and Cutrell, E. Designing a phone broadcasting system for urban sex workers in india. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI '11, ACM (New York, NY, USA, 2011), 267–276.

[14] Sharma Grover, A, S. O., and Lubensky. Designing interactive voice response (IVR) interfaces: localisation for low literacy users. In *Proceedings of Computers and Advanced Technology in Education*, CATE 2009 (St Thomas, US Virgin Islands, 2009).

[15] Sherwani, J., Ali, N., Mirza, S., Fatma, A., Memon, Y., Karim, M., Tongia, R., and Rosenfeld, R. Healthline: Speech-based access to health information by low-literate users. In *Information and Communication Technologies and Development, 2007. ICTD 2007. International Conference on*, IEEE (dec. 2007), 1 –9.

[16] Suhm, B. *Human Factors and Voice Interactive Systems Second*, 2nd ed. Springer, 2008, ch. IVR Usability Engineering Using Guidelines and Analyses of End-To-End Calls.

[17] 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`.