Computer Assisted Telephony System

Anurag Rana

Computer Science and Engineering

Indraprastha Institute of Information Technology, Delhi (IIIT-D), India

A Thesis Report submitted in partial fulfilment for the degree of

M.Tech Computer Science and Engineering

August 2015

1. Dr. Pushpendra Singh (Thesis Adviser)

2. Konstantinos Kazakos (External Examiner)

3. Dr. Vikram Goyal (Internal Examiner)

Day of the defense: 29 July 2015

Signature from Post-Graduate Committee (PGC) Chair:

Abstract

Telephones have high reach both in developed and developing countries that have made it a preferred tool for collecting and disseminating information. Voice based information collection systems offers data collection at low cost and save time that is needed to reach a human. Due to high teledensity, voice based systems have been very effective and have emerged as promising tools for facilitating data collection and community interaction in developing nations where the Internet penetration and literacy level is low.

Such voice based telephony systems can be improved by assistance from the computer system integrated with the data collector's telephone. We are explaining the development and deployment of computer-assisted telephony system (CATS) to conduct interviews with patients at *Grant Medical College and Sir Jamshedjee Jeejeebhoy Group of Hospitals* (a.k.a. JJ Hospital) Mumbai, India. This work presents the benefit of CATS over standalone telephonic surveys. In order to further take the leverage of CATS, we propose to develop automated telephonic data collection system. A prototype of same is evaluated in a lab study with low-literates. We discuss the results of the study conducted, requirements and benefits of fully automated system.

Acknowledgements

Foremost, I would like to thank my mentor Siddhartha Asthana for his constant support, suggestions, encouragement and help during the course of my thesis. His guidance, invaluably constructive criticism and friendly advice helped me in all the time of research and writing of this thesis. This thesis has been possible because of you.

Next, I would like to thank my advisor Dr. Pushpendra Singh for the continuous support of my thesis, for his patience, motivation, enthusiasm, and immense knowledge. I could not have imagined having a better advisor for my MTech thesis.

Declaration

This is to certify that the M.Tech Thesis Report titled **Computer Assisted Telephony System** submitted by **Anurag Rana** for the partial fulfillment of the requirements for the degree of *M.Tech* in *Computer Science and Engineering* is a record of the bonafide work carried out by him under my guidance and supervision at Indraprastha Institute of Information Technology, Delhi. This work has not been submitted anywhere else for the reward of any other degree.

Dr. Pushpendra Singh
Indraprastha Institute of Information Technology, New Delhi

Contents

List of Figures

1

Introduction

Due to deep reach, low cost and time requirement, data collection processes like survey and interviews are shifting from face to face or web to telephonic interviews and survey wherever possible[1]. In developing countries telephony network have deep reach compared to the Internet. In India overall Internet penetration is still less than 19.19% [2]. This figure is even lesser (6.7%) if we talk about rural areas only [3][4]. On the other hand if we talk about telephony network, It have a deep reach in rural areas. As of May, 2015 overall and rural teledensity in India was 79.67% and 48.60% respectively [5]. Automated telephone-based services in South Africa has helped filled the information gap as mobile phone penetration and usage have experienced massive growth. Africa have more phone connections than USA and Canada combined [6]. If we compare teledensity with Internet penetration, we can clearly say that phones are the best medium to reach maximum number of people in rural areas. Some research studies have investigated the metrics other than time and cost to judge the effectiveness of phones as medium for collecting information as compared to mail or web surveys. Fan W. and Yan Z. [7] claim that response rate of telephonic surveys is 10% more than web surveys.

American Speech-Language-Hearing Association (ASHA) conducted survey of 4000 speech-language pathologists (SLPs) and educational audiologists in school settings. The survey was designed to provide information about school-based service delivery[8]. This survey was conducted via mail. As per a report released in 2013 by SAII (Supreme audit institute of India), 38376 beneficiaries (26,115 males and 12,261 females) were interviewed from 3,837 gram panchayats in 27 states and three Union Territories under MNREGA (Mahatma Gandhi National Rural Employment Guarantee Act) scheme[9].

During a survey in 1999-2000 and 2004-05 by CSEI (center for science and Environment, India) to find the increase in employement rate 124680 households were covered[10]. These surveys were conducted by interviewing beneficiary face to face. There are many other such surveys which are conducted by government and other institutions, costing them lot of time, effort and money. Voice based technologies has helped in efficient delivery of healthcare services. Several mobile based solutions are being used in healthcare domain such as education and awareness, health data collection and patient monitoring and disease tracking. Voice based technologies has been used to give timely intervention and reminders to diabetic and CVD patients. Voice based systems have been developed to help children learn with the help of game [11]. A voice based entertainment system went viral in Pakistan. The system Polly was designed for entertainment purpose but was also used for social contact [12].

Web based social networking platforms being text rich are hard to use by illiterate people. Voice based social media systems have been developed to ease the communication between low-literate or illiterate users in rural areas [13][14]. Voice based discussion forums have been developed where once user can post their question and others can answer that. Such systems have helped farmers in regions where the Internet penetration is low [15]. Since low-literate people face great difficulty accessing the information which needs text input, Voice based systems have promising future and such systems have been developed which outperform the text-based input systems [16]. Reasons like low literacy level, poor financial conditions, low internet penetration and fear of technology also prevent rural population in developing nations from accessing information.

Many voice based systems have been developed to help conduct surveys and interviews. Such systems have been proven to be cost effective as compared to face to face interviews [17]. Mobile phone based system are found to be better alternative to conduct surveys on sensitive topics such AIDS. It was found that users are more willing to disclose personal information to automated agents.

Voice based information systems can be classified broadly in two domains. Voice based information collection systems and voice based information dissemination systems. Telephonic surveys conducted by government, telephonic interviews conducted by a firm and collecting health related data from patient by hospital are some examples of voice based information collection systems. Voice based information dissemination

systems are used to spread awareness among mass population, to advertise about some product or service, to send timely reminders and to help people learn about technology. Voice based information collection and dissemination systems can be used in combined fashion where people produce information which is consumed by other people. Discussion forums are good example of such systems. Combined system or technology mediated communication systems are used to engage people with each other to share local knowledge.

After extensive literature review we found that a lot of work has been done in the field of voice based information dissemination systems. Relatively lesser work is done in the field of automated voice based information collection systems. In this work we focused on voice based information collection system. In the next chapter we did literature review. In chapter 3 we explained the design and development of a voice based information collection system, also called as CATS which is deployed in *Grant Medical College and Sir Jamshedjee Jeejeebhoy Group of Hospitals, Mumbai*(Hereafter referred as JJ hospital) and is used by hospital staff to conduct interviews for monitoring patients which underwent heart surgery. Chapter 4 conclude our work, its limitations and discuss future work.

2

Background

There has been many studies related to usability, effectiveness and usefulness of automated systems. Automated system have been very competent specially in field of data collection by conducting interviews, surveys and disseminating information. We have categories the related work in three sections namely voice based systems for data collection, voice based systems for data dissemination and a combination of both the systems where people consume the information produced by other people.

2.1 Voice Based Systems For Information Collection:

Conducting automated interview and surveys to collect the data are widespread methods using telephony based automated systems. There have been studies to understand the different aspects of telephonic survey. Thomas R. et al. suggest that telephonic interviews have become commonplace to collect data but has not replaced face to face interviews yet[18]. Grover et al. worked on developing a voice based system for getting feedback regarding meal from school children[19]. Chakraborty et al. conducted telephonic interview using interactive voice response system and suggested that automated systems are the viable option to conduct interviews[17]. Stritzke et al. conducted the study to find the feasibility of interactive voice response technology to know the attitude of children towards alcohol and tobacco use [20]. Study concluded that retention rate among children was fairly high(91%) and slightly lower than adults(97%). Pollina et al. studied the feasibility of using automated system to conduct applicant's interviews for federal security clearance[21]. They used computer-generated agent to administer

the interview. Results suggested that national security interviews can be conducted using computer-generated agent.

2.2 Voice Based Systems For Information Dissemination:

There have been several studies which suggest that voice based systems can be used efficiently in data dissemination domain to reduce operational cost and to save time. Joshi et al. deployed a system TAMA, i.e., Treatment Advice by Mobile Alerts, using IVR and conducted studies on people living with HIV[22]. Studies suggested that it is possible to improve the adherence using interactive voice response systems. Raza et al. used voice based system Polly with entertainment as vehicle to disseminate information to low-literate population in Pakistan[12]. Raza et al. developed another system which was build over Polly to spread job ads via voice based service[23]. Gupta et al. developed an audio visual based system known as KrishiPustak for poora farmers[14]. Farmers could register themselves on this system with the help of a mediator and then could access the information shared by others. Asthana et al. developed an adaptive interactive voice response system [24] which was deployed in IIIT Delhi to disseminate the college admission related information. System was heavily used and received total 1120 calls over a period of 40 days. Larson et al. developed a voice based system for rural children which uses Cricket game to help them learn English words[11]. Grover et al. developed a multilingual voice based information dissemination system that enables managers at local community center to broadcast information to community workers[25].

2.3 Technology-Mediated Communication Systems:

Many systems have been developed by research community which tries to overcome the drawbacks such as illiteracy and fear of technology. Patel et al. developed an interactive voice application named *avaaj otalo* to be used by farmers as a forum to ask questions [15]. This system was deployed in Gujarat state of India. In this system farmers could ask question or listen to answer to other question by navigating using mobile keys. Gupta et. al. design and deployed an audio-visual social networking application, KrishiPustak, in rural area near Mysore, India [14]. This system overcame

the limitation of non-usability of systems due to illiteracy of user. It was not text-rich application. This system uses audio and images to convey information. Farmers registered using mediator's mobile device. This system was used to mirror village life where everyone knows everyone hence every registered user was connected to everyone else. Raza et. al. developed a system telephone-based voice-based system Polly [12]. Purpose of this system was to let the inexperienced user be familiar with technology using entertainment as motivator but this was put to an unintended use to share voicemails and group messages. Lobo et. al. developed GappaGoshti, a mobile based multimedia social networking platform for farmers in rural India [13]. Accessing any information from the Internet is almost impossible for a farmer due to reasons like low literacy, poverty and fear of technology. Rural populace is comfortable with mobile devices. In this system they developed a visual user interface which offers to record or playback an audio message.

Every system designed so far supports some or other use cases. Few systems were used to conduct telephony interviews while other systems were used to facilitate the data dissemination. Few systems were used as voice based social networking platforms where one user post content and other user consume that content. Most of the systems which support conducting telephone-based interview and surveys requires a human operator to conduct such interviews. We did exhaustive literature survey and found that no fully automated voice based data collection system has been designed. Here in this work we focused on facilitating data collection through voice based systems. We also proposed the design of fully automated telephony interview system based on results of lab studies. In the next chapter we explained in detail the design, development and deployment of Computer Assisted Telephony System.

3

Voice Based Information Collection System

Any data collection process like interview or survey is conducted using a set of questionnaire presented to study participants. In a manual process where telephone system is used to conduct a survey, surveyor marks the participant's responses on a paper sheet. A computer-assisted telephone system may offer a wide range of feature that can ease-of the burden of surveyor. Responses can be marked with a simple click on a GUI application on the computer system that will save both time and effort of surveyor.

Moreover, in several scenario the process of conducting a survey is quite repetitive in nature. The order and number of question are often fixed, and sometimes may differ based on participant response to a particular question in the set. However, in both the scenario order of survey question is deterministic in nature and hence can be automated. In an automated telephonic system, a computer generated or recorded voice may presents the question on participant phone. The responses may be recorded through touch-tone, voice recognition or simply audio recording of the calls. This can effectively minimize the need of human intervention in conducting a survey study, and large number of participants can be surveyed simultaneously to reduce the time to conduct the survey. In this chapter we explained the design and deployment of a voice based information collection system named as Computer Assisted Telephony System (CATS).

3.1 Computer-Assisted Telephony System (CATS)

Founded in 1845 *Grant Medical College and Sir Jamshedjee Jeejeebhoy Group of Hospitals, Mumbai*, also known as JJ hospital is counted among the premier medical college of India. It is always ranked in top 10 colleges in India. It bears annual load of 1200000 out and 80000 indoor patients [26]. Central and state government runs many economical health care programs in many departments of this hospital. Department of cardiovascular disease (CVD) is one of them. Patients undergoes a surgery in this department under a government scheme. These patients need constant health monitoring for some time after operation.

3.1.1 Background

We began this study at JJ hospital Mumbai. Doctors conduct interviews with patients who underwent an operation in that hospital to keep a check on vital health parameters. This process was again repeated after few months. Before we deployed the computer-assisted telephone systems (CATS) at JJ Hospital Mumbai, all the patient follow up, monitoring, interviewing and data entry task was manual. A dedicated staff of nurses and sisters will call the patients one by one, will interview them and will note down the important points on paper. Interview process follow a decision tree structure, Hence number of questions keep varying from patient to patient, reaching up to more than 100 in few cases. After that another staff member will digitize the content into Excel sheets. Then there were other formalities which will wrap up the whole process. More than 1900 patients were interviewed using manual process.

3.1.2 Problems

We conducted interviews with hospital staff involved (3 nurses and 1 doctor) in this activity and figured out few issues with this system.

1. **Time:** One whole interview process takes 20 to 30 minute. So one staff member was able to finish two to three interview in an hour. Digitizing the data collected from interviewing the patient needs another 10 to 15 minutes. Wrapping up the process requires 5 to 10 minutes. So whole process would take from 35 to 55 minutes. Hence we can say that process is very time consuming.

2. **Error in data handling:** Hospital staff use to call patients using landline phones and note down the things on paper. Data was filled from paper to Excel sheets later on. There were many human errors and inconsistencies in the data. Some times if interviewer forget to note down some vital piece of information, he need to call the patient again. Also there were very limited validations in Excel sheet because of that also chances of error increases. And not to mention that a small error in medical field may result in big loss. Although there were no such cases because this process was completed very carefully and validated again after completion but it again was very time consuming.

3. **Cost:** As discussed previously, completion of the interview and digitalization of data requires at least two staff members or man-hours equivalent to two employees hence increasing the cost of interviewing process.

4. **decentralize data:** Fourth problem we figured out was that the medical data collected was not centralized. If staff need to share the data, they will email it, another user will make changes and will email back the data. Multiple copies of same data were created. So there was a need to established a central repository to dump whole data at one place so that any authorized person can access, change and update the data.

To mitigate the issues discussed in section 3.1.2, we developed a computer-assisted telephony system according to hospital staff requirement and deployed it. Structure and implementation is discussed in detail in next sections.

3.1.3 Design

System consist of a central telephony server, soft-phones (Figure 3.2), Database and web-based graphical user interface to enter the data(Figure 3.3). A detailed structure is shown in Figure 3.1. Telephony server consist of many modules named as *user validation module*, *Dial plan module*, *call recording module* and *call hangup module*. Working of these modules is described in next section. Soft-phones are connected to central telephony server. Server save data to database which can be accessed and updated through web-based GUI(Graphical user interface).

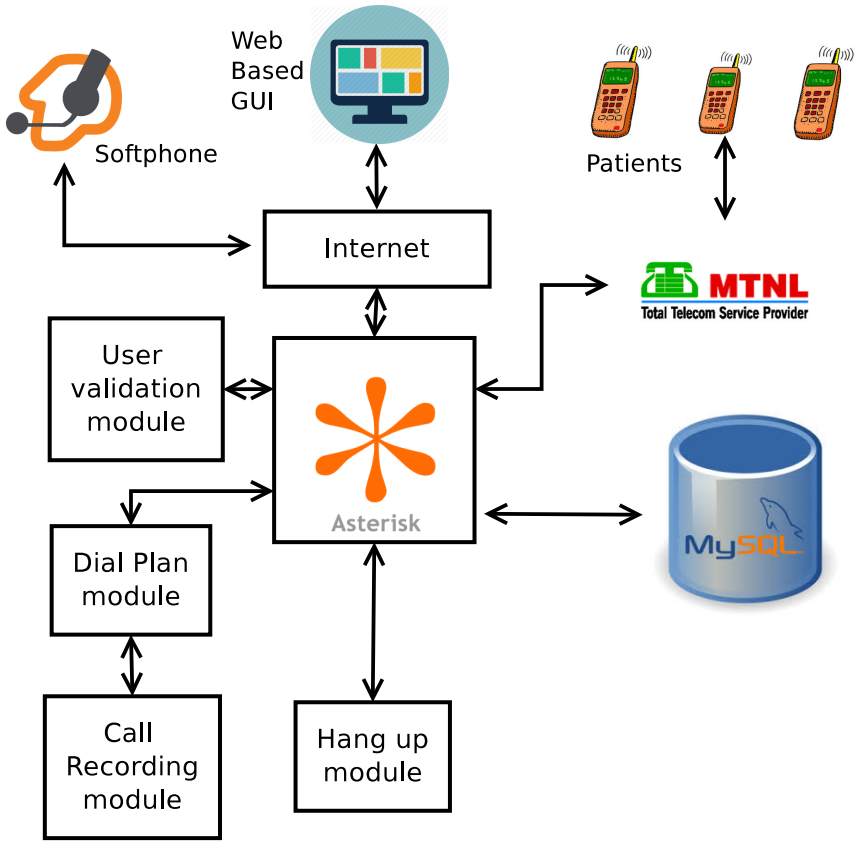

Figure 3.1: Computer assisted telephony system - detailed structure.

3.1.4 Implementation

System uses a tool called soft-phone to make outbound calls. This software is installed on user's PC and is connected to central telephony server. An active Internet connection is required to make a call. Staff member will call on the patient's mobile using this soft-phone. Request to make call is sent to telephony server which initiate the call and user conducts the interview.

Central telephony server consist of different modules handling separate tasks. When a call is made to some patient, first *user validation module* validate the number. If it belongs to a patient then only call is made, otherwise not. This validation is done to prevent making an invalid call, saving time and call cost both. *Dial Plan module* will make the call to patient if number dialled is valid. *Call recording module* start audio recording if call is bridged. *Call hangup module* comes into picture when call

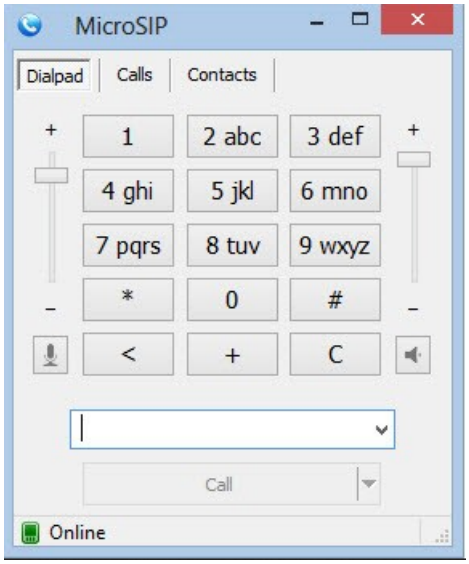

Figure 3.2: MicroSip Soft-phone: A tool used to connect to telephony server and to make calls.

is disconnected and is responsible to store all the responses into database. All these modules interact with databased using the Asterisk Gateway Interface (AGI).

Data can be filled while interview is going on or later by referring the recording saved in database (Figure 3.3). Different validation like valid value, valid length of data value etc, are done on data before it is saved to database. To enter the data, user need to open the web-based GUI (graphical user interface) in the Internet browser after entering his/her credentials. So the only authorized person can access the patient related data, honoring the privacy of patient. Other users like senior doctors can see the data from their end using their credentials.

3.1.5 Evaluation

After a month of work using CATS by hospital staff members, we asked them to provide feedback about the usability and level of comfortable with the system. Below are the major points made by them.

- It is more comfortable to conduct long interview using headphone and mic on PC instead of landline phone.

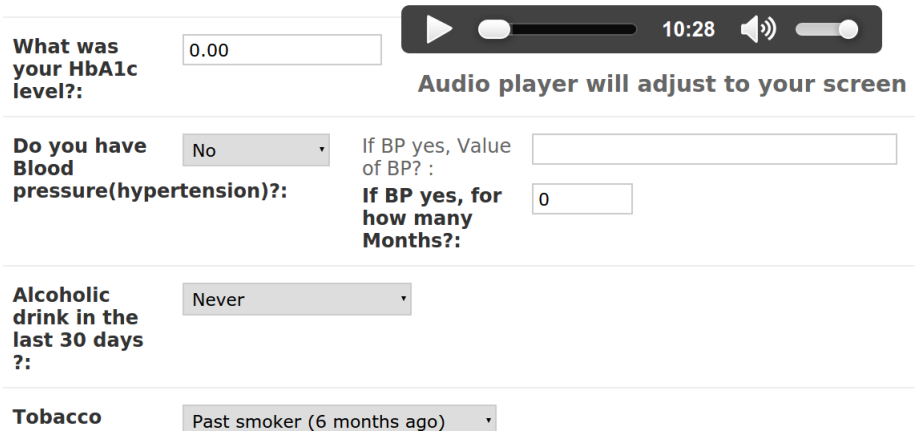

Figure 3.3: A zoomed in screenshot of web-based GUI, which is used to fill and save medical data.

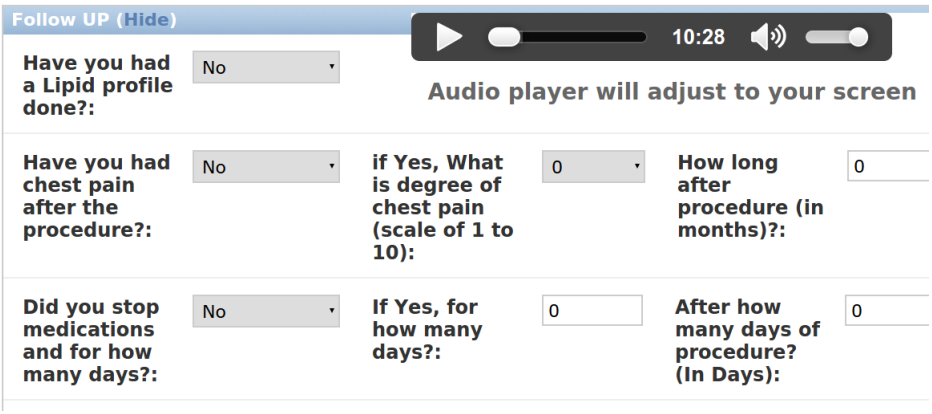

Figure 3.4: Audio player floats on web-based GUI. Notice that background changed (Refer Figure 3.3) but audio player remained at top right corner. It help the Nurses in playing and pausing the audio while transcribing data.

- Data entry is easier as we can enter data while interview is going on.
- Floating audio player makes data entry easier.
- No error in data entry due to validation rules. Consitency in data format mitigated the requirement of data cleaning.
- Excel sheets crashes many time, generally in case of large data. There is no such problem in case of database.

- System ensured that the information collected is authentic as there are audio recordings available for each interview hence mitigating the chaces of data fabrication.

So overall users were very comfortable with the system and acknowledged that no technical skill is required to work with this system.

CATS was able to mitigate three out of four main problems faced in manual process. As per data entry operator, digitizing data using web based GUI was easy and faster. Also interviewer can fill the data during the interview, saving the time per interview process hence reducing operational cost. Because of validation rules in web based GUI, errors and inconsistencies in data were mitigated. Also all the data was kept at central location, making accessibility and updation of data very easy. keeping data in database added the addition feature of security and reliability against data loss. Still the problem of lengthy interview process remained unaddressed. Although a small amount of time was saved in CATS as compared to manual process, but it was not significant enough. To solve this issue we conducted staff member's interview after around one month of deployment of CATS, asked for their feedback about it and conducted a separate study which is discussed in detail in next section.

3.2 Automating Voice Based Information Collection System: A Lab Experiment

To mitigate the fourth problem i.e. large time being consumed in taking interviews and data filling, we are proposing the system which will be fully automated and will be next efficient version of the system deployed previously. To understand the requirement completely we interviewed the staff about the challenges they face while conducting an interview. We also listened the recordings of interviews. Main points we figured out are -

- Call gets disconnected sometimes. Interview breaks up in multiple parts.
- Sometimes user do not pick up the call in first attempt. Staff members need to call multiple times in order to schedule an interview.

- Sometimes call is made at an odd time. User might be busy somewhere else and wants to reschedule the call.
- If a user miss the call then they try to call back.
- Out of around 120 questions being asked, most of them are multiple choice questions. Few questions have very short answer and very few questions have large answers.

3.2.1 User Study

Based on the above points we conducted a study in controlled environment. We designed a prototype automated system which will initiate the call without human intervention and ask the questions once call is bridged and store the answers in database. Although interviews conducted by automated systems are lengthier than interviews conducted by Human operator[17], we planned to conduct multiple interviews simultaneously hence reducing the average time per interview.

Study Design: There has been many studies on the usability of IVR systems. We referred few of them[17][27] [22], specially the ones which focus on low-literate participants because a large fraction of patients in our case is not well educated. We conducted the study on the participants with same socio demographics as of patients. Since this was an experimental study, we do not wanted to disturb the patients for this. So to test the prototype automated system we targeted those participants in our vicinity who have educational qualification, income and age comparable to most of the patients. Since this study was conducted inside IIIT Delhi campus, participants were members of housekeeping staff, gardeners, mess/canteen employees, security guards and supervisors. So the group of participants have a blend of all type of users, i.e. illiterates, low-literate(majority) and moderately educated. A total of 30 users participated in this study. To know the user experience, we conducted short interviews after the call with each participant. Since incentive is also an affecting factor in surveys[7], we promised to offer mess food coupon to all the participants for their honest participation.

Experiment Design and Implementation: System is the prototype version of fully automated system. It consist of all the modules of CATS plus a module to schedule a call and a module to initiate the call. Set of questions consist of 10 questions. All the questions were recorded in Hindi language. Type of questions vary from simple

yes/no answer to providing a large input like mobile number. We provided a list of numbers to be called to the system. System will schedule the call to each number. In case call is not picked up in first attempt, system will retry after 60 seconds. If even second attempt fails, call is scheduled to some other day. System was made to handle simultaneous calls. Once call is bridged, questions are played to participant. There was an option to press 9 to replay the question. If user doesn't enter anything for a specified period of time question will be played again, but this will happen only once. Second time also user doesn't enter anything, question will be skipped (Refer Figure 3.5). Responses were saved in the database once call is finished. So basically whole interview process is automated and data is saved to database, saving a lot of time and human operator cost.

Results: Most of the participants performed well and answered questions correctly. We conducted face to face interview later on with all the participants to know 1) the reasons of mistakes they made, 2) the difficulties they faced and 3)their feedback. Major findings of this study are listed below. A problem faced is labeled as 'P', reason(if any) is labeled as 'R' and proposed solution is labeled as 'S'.

- **P:** Three users disconnected the call after few seconds.
  R: Participants thought it is some marketing call and pressing any key may result in balance deduction.
  S: Before asking to press any key, make clear that this call will not cost you anything.

- **P:** Two users entered wrong age.
  R: For questions which require multi-digit input as answer, we gave an example on how to enter the input. For example we said 'If your age is thirty-four years, press three and four'. Participants got confused and inputted the age provided as an example.
  S: We asked such questions without an example and other users gave the correct input this time. Just need to tell users in the call introduction that 'Provide answers by pressing keys on your mobile keypad.'

- **P:** A user although provided all the answers but later on he revealed that answers were not correct.

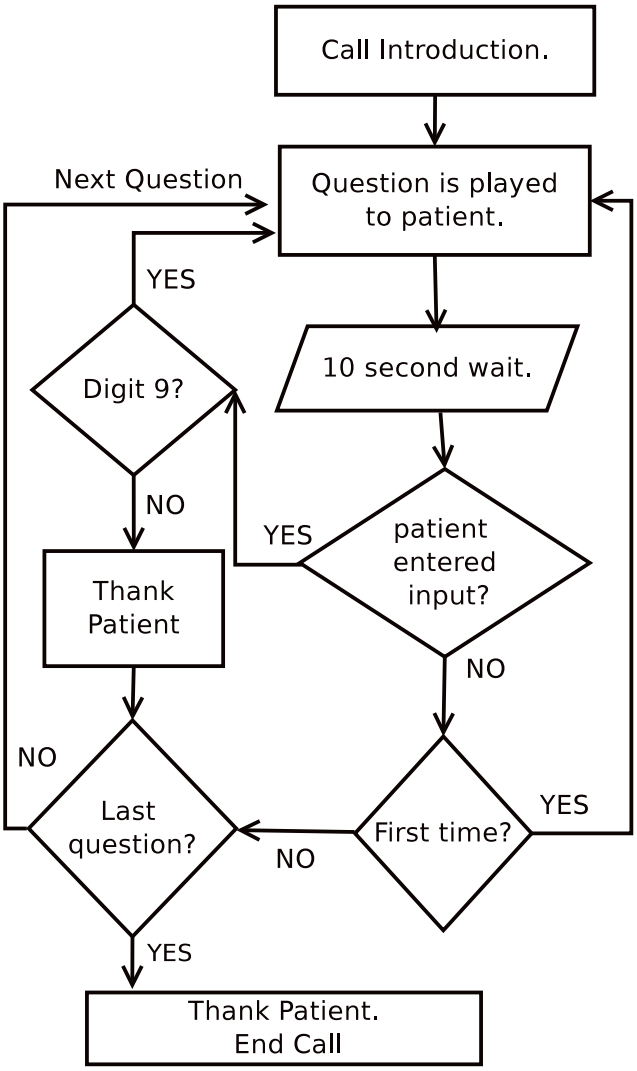

Figure 3.5: Flowchart of interview process

R: He thought this call may fetch him a job.

S: State the motive of call clearly in call introduction.

- **P:** One user didn't entered his marital status.

  R: The Hindi word for 'married' we used was not so common.

  S: Use most commonly used words/phrases. Use dialect specific words.

- **P:** A user didn't entered his mobile number.

  R: He didn't remembered his mobile number. By the time he searched for it,

question timedout.

S: Identify questions which might require some extra time in case user need to search the answer.

- **P:** Two users were not able to input anything.

  R: They were using smart phones which do not have physical keypad. Onscreen keypad gets locked and require more time to input the answer. Low-literate users are not very comfortable with smart phones. This leads to question timed out.

  S: Identify the type of phone being used by user and set the time out limit accordingly.

We also tested the system by placing three simultaneous calls. Interviews were completed successfully without any issue. Number of simultaneous calls can be increased to desire value based on the availability of PRI (Primary rate interface) lines.

Graph in Figure 3.6 shows the time taken by participants in automated interview. This interview consisted of 10 questions and was initiated automatically by prototype model. Three participants disconnected after around 40 seconds in call due to reason explained above. Two participants were not able to enter any input. Average time taken by participant (above five participants excluded) in automated interview was 282 seconds.

To calculate the average time in face-to-face interview, consisting of same questions, we conducted interviews with second set of participants. Different set of participants was chosen to avoid any biased result as first set already knew the questions. Average time in face-to-face interview was 120 seconds. For interviews consisting more question, this time will increase and may vary depending on type of questions.

Although our study reported that automated interviews are slowers than face-to-face interviews, we conducted three interviews simultaneously, bringing average time to 94 seconds, a saving of 21.67% compared to face-to-face in interview. With an increase in degree of parallel interviews, this saving will be even larger, hence saving a lot of time for interviewer and making the automatic interview process faster.

3.2.2 Automated System

Requirements: We identified few requirements for this system to function in an efficient way.

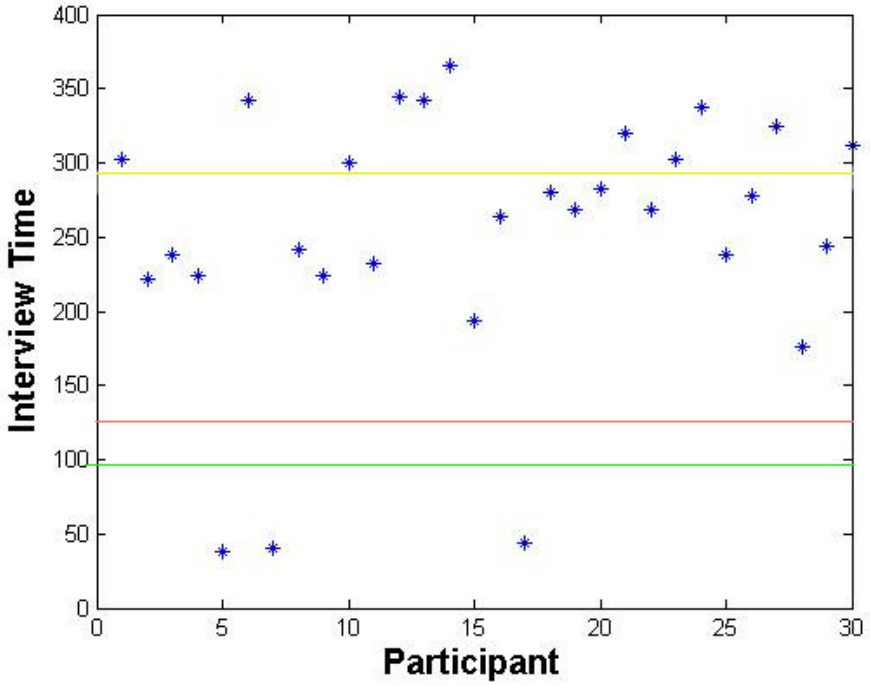

Figure 3.6: Automated interview time in seconds for participants. Yellow line show the average time. Red line depicts the average manual interview time. Yellow line shows the average time for three automated interviews simultaneously

- Recording of the questions in regional language/dialect. Use of dialect specific words and phrases in questions.
- Identifying the questions which can be answered by pressing keys. Multiple choice type questions fall under this category. For example 'Do you have chest pain?' can be answered in Yes or No. So this question can be recorded as 'If you have chest pain, press 1 and if you do not have chest pain, press 2'. Response of such questions can be directly stored in database after confirmation from patient.
- Identifying the questions which are NOT multiple choice type questions. Such questions need long answers. Example 'What medicines are you taking, please name them'. For such question user will answer through speech.
- Identifying the questions for which patient might need more time than usual. To answer such questions patient may need to refer some documentation like

prescription provided by Doctors.

- Development of call scheduling and call initiation module.
- Development of Data transcribe module which will parse the user input enter from keypad and will save to the database.

Features: Here we are presenting a list of features Automated System will have-

- Capability to work without human intervention. Hospital staff just need to give the list of mobile phone numbers to the system.
- System will be able to schedule and initiate the call at appropriate time.
- If call is not received by patient at first attempt, system will try again after 60 second. Even if second attempt fails, call will be scheduled for next day.
- Once a call is bridged, system will present the option to either continue the call or reschedule it for some later time. System will proceed, based upon the user input.
- If call is disconnected in between, system will recall the patient and interview will resume from where it was left.
- If patient miss the call and then tries to call back, system will identify the number and if it belongs to a patient then system will call back the patient and complete the interview process.
- All the response of multiple choice type questions will be saved to database directly. Since most of the questions fall in this category, it will save a lot of effort of data entry.
- System will offer an option to choose the language at the start of interview. Since most of our patients are from Maharashtra State in India, we will offer to choose Hindi or Marathi.
- System will be able to make multiple calls at the same time.

- Staff member can monitor the progress of ongoing interviews from their end. System will have an option for patients to have (in case they need) staff member's attention.

We will keep updating the system from time to time based on the requirements provided by staff members and based on the feedback from patients.

Design and Working: In addition to CATS, this fully automated system will have a call scheduling and call initiation module. Call scheduling module takes one number at a time from list and will schedule the call at some appropriate time. Call initiation module will keep checking if there is some call scheduled at this time. If yes then module will initiate the call. If call is not bridged, initiation module will inform scheduling module and call will be queued for later time. Once call is bridged, process will be handled by main module which will introduce the call, play the question and will send the response to data transcribe module. This module will parse the input and store it in database. For every response, patient will be acknowledged and response will be stored only if patient confirm the answer. In case patient wants to change the response, he can do so by pressing the specified undo key.

Web-based GUI will be more feature rich. In addition to change and update the data, staff member can input the list of contact numbers from GUI. Ongoing interviews will be monitored and if any patient need human operator's assistance any time during the call, patient can press the specified key and staff member will be notified through web-based GUI.

Figure 3.7 shows structure overview of fully automated system. Main module structure is same as shown in Figure 3.1. Flow inside main module is almost same as shown in Figure 3.5. Flowchart in Figure 3.8 shows the flow of call in a complete automated system.

Challenges: We did a feasibility study and found that all the things listed in requirement section can be achieved. However for the question which need open ended answers we need to parse the speech using some speech to text conversion tool. Efficient tools are available for languages like English, but for regional languages like Marathi no such efficient tool is available. For such question hospital staff needs to manually transcribe the data. Since number of such questions is very less, it will not affect the overall efficiency of system.

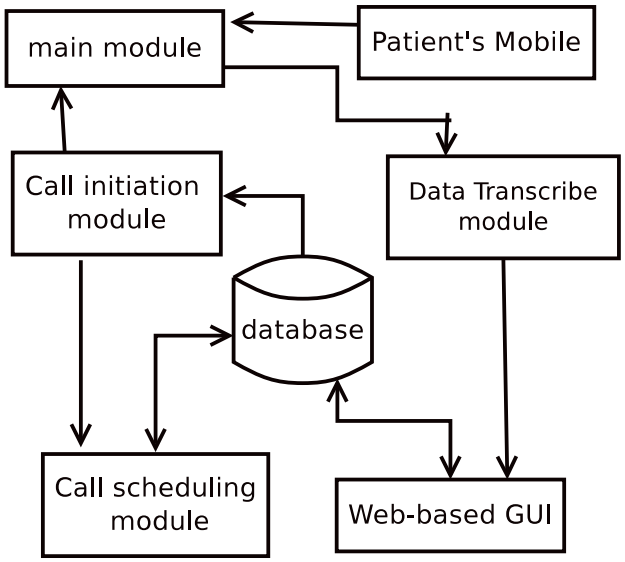

Figure 3.7: Structure of a fully automated system.

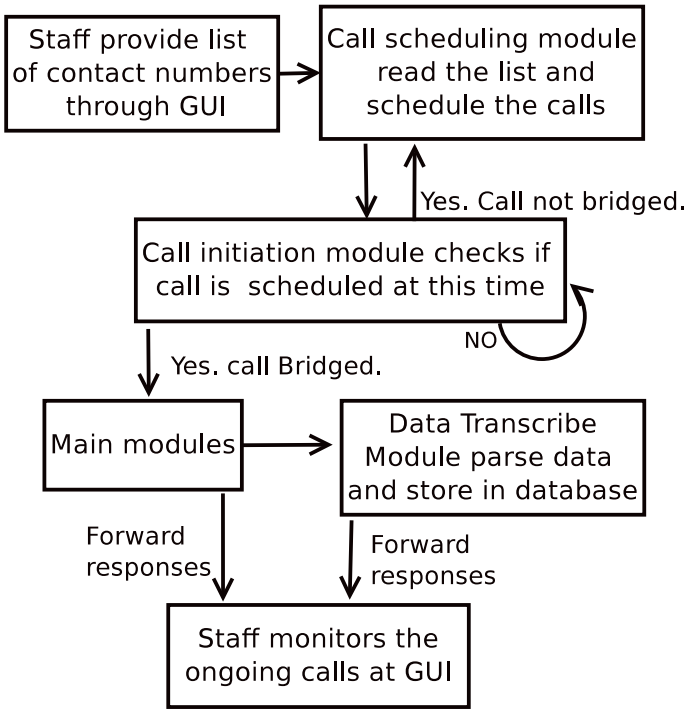

Figure 3.8: Process flow in fully automated system.

Benefits: As discussed in CATS section, this fully automated system will reduce the average time per interview as calls to more than one patient can be made simultaneously. Majority of the questions are multiple choice questions hence data will be transcribed automatically reducing the time consumed in data entry up-to a large extent. Because of automatic data entry, there will be no human error in data handling. Now multiple calls can be made and monitored at the same time, it will add to cost savings by eliminating the need of human operator.

4

Conclusion and Future Work

After working on CATS for over a year and conducting 575 interviews we concluded that computer assisted voice telephony systems are better than standalone telephony systems. CATS made the interview and information collection process easy, fast and efficient. We were able to reduce the operational cost, errors and inconsistencies in data and average interview time. After conducting lab studies and working with prototype model in simulated environment, we can say that fully automated system is totally feasible in terms of development and deployment. Since this system doesn't require any technical skills apart from basic computer operations, any non-technical user can also work on this system. Hospital staff is already familiar with CATS hence they will not require any training for this.

There has been many studies on interactive voice response systems and healthcare [27] [22] [28], but CATS is of its kind. Our main focus is on reducing cost, time and errors in data handling which our system is achieving efficiently. Although we are developing this system for healthcare domain, this can be used in any field which require interviewing large group of people on the set of same questions. For example same system can be used by placement agencies after a slight modification. Interactive voice response systems have reduced operational cost and human labor. Now it is evident that with a proper design of questionnaire and with the help of technique used by our system, IVR systems can save time too. Once fully developed and deployed it will save a lot of time, money and effort. Since there will be automatic data entry, there will be no error in data handling hence chances of potential loss reduces to a large extent.

Future work involves developing the fully automated computer assisted telephony system which can conduct interview without human operator. An intelligent module will be developed in future which based on user interaction behaviour would be able to detect if a user is struggling during the interview and will switch to manual mode in such case. In manual mode a human operator will take over the system and assist the user. Since speech recognition is hard for Indian regional languages, all the question which need speech based input will be rearranged at the end of the interview so that user can answer all the touch-tone based question first and then system will switch to human-operator mode.

References

[1] http://www.cleveland.com/business/index.ssf/2011/03/phone_interviews_are_repl **Phone interviews are replacing in-person ones requiring job seekers to learn new techniques**, 2011. "http://www.cleveland.com/business/index.ssf/2011/03/phone_interviews_are_replacing.html". 1

[2] internetlivestats.com. **Internet penetration India 2015**, 2015. "http://www.internetlivestats.com/". 1

[3] mobientech.com. **Rural India internet penetration still hovers low at 6.7%**, 2014. "http://www.mobientech.com/blog/rural-india-internet-penetration-still-hovers-low-6-7/". 1

[4] Economic Times. **Rural internet users to rise to 85 million by June 2014: IAMAI**, 2014. "http://articles.economictimes.indiatimes.com/2013-10-22/news/43288777_1_internet-and-mobile-association-iamai-rural-india". 1

[5] TRAI. **Telecom regulatory authority of India**, 2015. "http://www.trai.gov.in/WriteReadData/PressRealease/Document/PR-32-TSD-May-15_10072015.pdf". 1

[6] Richard Heeks. **Emerging markets IT and the world's bottom billion**. *Communications of the ACM*, **52**(4):22–24, 2009. 1

[7] Weimiao Fan and Zheng Yan. **Factors affecting response rates of the web survey: A systematic review**. *Computers in Human Behavior*, **26**(2):132–139, 2010. 1, 14

[8] ASHA. **Survey Methodology, Respondent Demographics, and Glossary**, 2014. "http://www.asha.org/uploadedFiles/2014-Schools-Survey-Demographics-Methodology-Glossary.pdf". 1

[9] SAII. **Findings of the Beneficiary Survey**, 2013. "http://saiindia.gov.in/english/home/Our_Products/Audit_Report/Government_Wise/union_audit/recent_reports/union_performance/2013/Civil/Report_6/chap_13.pdf". 1

[10] CSEI. **The National Rural Employment Guarantee Act (NREGA). Opportunities and Challenges**, 2006. "http://cseindia.org/challenge_balance/readings/NREGA.pdf". 2

[11] Martha Larson, Nitendra Rajput, Abhigyan Singh, and Saurabh Srivastava. **I Want to Be Sachin Tendulkar!: A Spoken English Cricket Game for Rural Students**. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, pages 1353–1364, New York, NY, USA, 2013. ACM. 2, 5

[12] Agha Ali Raza, Mansoor Pervaiz, Christina Milo, Samia Razaq, Guy Alster, Jahanzeb Sherwani, Umar Saif, and Roni Rosenfeld. **Viral entertainment as a vehicle for disseminating speech-based services to low-literate users**. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development*, pages 350–359. ACM, 2012. 2, 5, 6

[13] Sylvan Lobo, Pankaj Doke, and Sanjay Kimbahune. **GappaGoshti: a social networking platform for information dissemination in the rural world**. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, pages 727–730. ACM, 2010. 2, 6

[14] Nakull Gupta, Indrani Medhi-Thies, Pedro Ferreira, Jacki O'Neill, and Edward Cutrell. **KrishiPustak: A Social Networking System for Low-Literate Farmers**. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing*, pages 45–48. ACM, 2015. 2, 5

[15] Neil Patel, Deepti Chittamuru, Anupam Jain, Paresh Dave, and Tapan S Parikh. **Avaaj otalo: a field study of an interactive voice forum for small farmers in rural india**. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 733–742. ACM, 2010. 2, 5

[16] Jahanzeb Sherwani, Sooraj Palijo, Sarwat Mirza, Tanveer Ahmed, Nosheen Ali, and Roni Rosenfeld. **Speech vs. touch-tone: Telephony interfaces for information access by low literate users**. In *Information and Communication Technologies and Development (ICTD), 2009 International Conference on*, pages 447–457. IEEE, 2009. 2

[17] Dipanjan Chakraborty, Indrani Medhi, Edward Cutrell, and William Thies. **Man Versus Machine: Evaluating IVR Versus a Live Operator for Phone Surveys in India**. In *Proceedings of the 3rd ACM Symposium on Computing for Development*, ACM DEV '13, pages 7:1–7:9, New York, NY, USA, 2013. ACM. 2, 4, 14

[18] Roger Thomas and Susan Purdon. **Telephone methods for social surveys**. *Social research update*, **8**:1–6, 1994. 4

[19] Aditi Sharma Grover, Karen Calteaux, Etienne Barnard, and Gerhard van Huyssteen. **A voice service for user feedback on school meals**. In *Proceedings of the 2nd ACM Symposium on Computing for Development*, page 13. ACM, 2012. 4

[20] Werner GK Stritzke, Justine Dandy, Kevin Durkin, and Stephen Houghton. **Use of interactive voice response (IVR) technology in health research with children**. *Behavior research methods*, **37**(1):119–126, 2005. 4

[21] Dean A Pollina and Allison Barretta. **The effectiveness of a national security screening interview conducted by a computer-generated agent**. *Computers in Human Behavior*, **39**:39–50, 2014. 4

[22] Anirudha Joshi, Mandar Rane, Debjani Roy, Nagraj Emmadi, Padma Srinivasan, N Kumarasamy, Sanjay Pujari, Davidson Solomon, Rashmi Rodrigues, DG Saple, et al. **Supporting treatment of people living with HIV/AIDS in resource limited settings with IVRs**. In *Proceedings of*

the 32nd annual ACM conference on Human factors in computing systems, pages 1595–1604. ACM, 2014. 5, 14, 23

[23] Agha Ali Raza, Farhan Ul Haq, Zain Tariq, Mansoor Pervaiz, Samia Razaq, Umar Saif, and Roni Rosenfeld. **Job opportunities through entertainment: Virally spread speech-based services for low-literate users**. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2803–2812. ACM, 2013. 5

[24] Siddhartha Asthana, Pushpendra Singh, and Amarjeet Singh. **Design and Evaluation of Adaptive Interfaces for IVR Systems**. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 1713–1718, New York, NY, USA, 2013. ACM. 5

[25] Aditi Sharma Grover and Etienne Barnard. **The Lwazi Community Communication Service: Design and Piloting of a Voice-based Information Service**. In *Proceedings of the 20th International Conference Companion on World Wide Web*, WWW '11, pages 433–442, New York, NY, USA, 2011. ACM. 5

[26] http://www.gmcjjh.org/. **Official website of Grant Medical College and Sir JJ Group of Hospitals**, 2015. `"http://www.gmcjjh.org/"`. 8

[27] Philip C Cooley, Heather G Miller, James N Gribble, and Charles F Turner. **Automating telephone surveys: using T-ACASI to obtain data on sensitive topics**. *Computers in Human Behavior*, **16**(1):1–11, 2000. 14, 23

[28] Brian DeRenzi, Gaetano Borriello, Jonathan Jackson, Vikram S Kumar, Tapan S Parikh, Pushwaz Virk, and Neal Lesh. **Mobile Phone Tools for Field-Based Health care Workers in Low-Income Countries**. *Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine*, **78**(3):406–418, 2011. 23