

ABSTRACT

WANG, TING. A VoIP anti-Spam System based on Reverse Turing Test. (Under the direction of Dr. Peng Ning.)

A reverse Turing test based anti-voice-Spam scheme in conjunction with black/white listing is proposed, demonstrated and verified to mitigate the Spam threats to the security of a VoIP network. The first part of this work demonstrates the feasibility of VoIP Spam generation and routing without going through a proxy server. The second part studies the implementation and evaluation of using a reverse Turing test to detect the spams, as well as the Turing-like challenge design considerations.

It is demonstrated in this thesis that VoIP spam can be automatically generated and routed without the involvement and control of a proxy server or registrar. Without specific SIP configuration knowledge of the target phone, a fake INVITE message was sent to the target and a SIP session was successfully established to broadcast spam messages. This was also verified by the test result from monitor and simulation software SIPp. In the process, an automatic UDP port scanner was implemented to find the listening port of the victim machine.

The VoIP spam detection system consists of two components, a regular SIP based VoIP softphone and a challenge selection & grading program. The challenge is a randomly picked voice question from a pool of pre-recorded questions designed by the user. A set of challenge design guidelines was discussed particularly for the application to a VoIP environment. The reverse Turing test was implemented and evaluated for

usability, correctness and performance merits. The time requirement for installation and configuration is short. About 75% of evaluators used less than 10 minutes to install and configure our tool. The experimental results show the program works well with a high passing rate for human users and a low passing rate for the machine users. Over 86% evaluators who speak English as the 2nd language passed the English audio challenge based VoIP anti-Spam filter, whereas around 96% of native speakers pass the tests. On the other hand, the A.I. chatter robot in our experiments failed 98% of the challenges. Also the resource consumption of this system is very little. According to the reports we received from users, the average disk space consumes about 4.63 MB.

A VOIP ANTI-SPAM SYSTEM BASED ON REVERSE TURING TEST

By

TING WANG

A thesis submitted to the Graduate Faculty of
North Carolina State University
In partial fulfillment of the
Requirements for the degree of
Master of Science

COMPUTER SCIENCE

Raleigh, NC

2007

Approved By:

Professor Peng Ning
Chair of Advisory Committee

Professor Douglas S. Reeves

Professor Ting Yu

Dedication

To my parents and my fiancé ...

Biography

Ting Wang was born in Chong Qing, China. She received her Bachelor's degree in Computer Science from University of Nevada Las Vegas, USA in 2003. In March of 2004, she worked as Software Developer in Triangle Research Collaborative, Inc. Since August of 2004, she has been a Master student in the Department of Computer Science at North Carolina State University. In Fall of 2005 and Spring of 2006 she worked at Cisco Systems, RTP, NC, as a Network Support Engineer for her co-op.

Acknowledgements

I acknowledge the efforts of Dr. Peng Ning, my thesis advisor, in providing me all necessary direction, guidance and support. He helped me remain focused and persistent. Working with him has been a pleasure and honor. The training I received from Dr. Ning is priceless.

I am grateful to my committee members Dr. Douglas S. Reeves and Dr. Ting Yu for devoting time and providing valuable inputs.

I want to thank Yi Tang, Zhe Zhang, Qinghuan Zhang, Panos Kampanakis, Kyuyong Shin, Ting Yi, Liping Li for inputs and spending time on my defense rehearsal.

My colleagues in the Cyber Defense Laboratory supported me in my research Work. I want to thank Dingbang Xu, An Liu, Chongkyung Kil, Emre Sezer, Zhi Chen, John Stille, Sangwon Hyun for their help during my study in the lab.

This work would not be possible without the 45 evaluators to by my testing object. Thank them for devoting time on my experimental study.

Finally, I would also like to thank my parents and my lovely fiancé. This thesis would not be possible without their support.

Table of Contents

List of Figures	vii
List of Tables	viii
Chapter 1 Introduction.....	1
§1.1 Voice over IP (VoIP)	1
§1.2 Spam	3
§1.3 What is VoIP Spam?.....	4
§1.4 Proposed Research and its Contribution	6
§1.5 Organization.....	8
Chapter 2 Related Work	9
§2.1 Text based Spam detection	9
2.1.1 Content Filtering	10
2.1.2 Black and White Lists.....	11
2.1.3 Turing Test & Reverse Turing Test.....	12
§2.2 Text Spam vs. VoIP Spam.....	13
§2.3 Review of Current VoIP Anti-spam Research Works	16
2.3.1 Black and White List Family	16
2.3.2 Turing Tests Family.....	18
Chapter 3 Feasibility of Automatic VoIP Spamming.....	20
§3.1 Background on Session Initiation Protocol.....	20
3.1.1 Introduction of Session Initiation Protocol (SIP).....	20
3.1.2 SIP Authentication and Potential Security Concerns.....	25
§3.2 Automatic Spam Generation.....	28
Chapter 4 Mitigating Automatic VoIP Spams using Reverse Turing Test	38
§4.1 System Overview	38
§4.2 Challenge-based VoIP Spam Detection.....	40
§4.3 Security Threats and Analysis	42
§4.4 Considerations for Diversified Challenge Design	44
Chapter 5 Experimental Evaluation.....	50
§5.1 Usability Study.....	50
5.1.1 Usability Study for Callees	51
5.1.2 Callee Evaluation Data Analyses and Discussions.....	53
5.1.3 Usability Study for Callers.....	58
5.1.4 Caller Evaluation Data Analyses and Discussions	59
§5.2 Correctness Study	63
5.2.1 Correctness Study for English-as-Second-Langrage (ESL) Users	64
5.2.2 Correctness of humans on Mother language audio question testing.....	66
5.2.3 Analysis of machine passing rate on English challenge questions	66
§5.3 Performance Study.....	69
Chapter 6 Conclusions and Future Work	70
§6.1 Conclusions.....	70
§6.2 Limitations and Future Work.....	71

Appendix.....	76
Appendix A User Instruction Guide.....	77
Appendix B Questionnaires.....	94
Questionnaire A -- Receiver Version.....	94
Questionnaire B -- Caller Version	95
Appendix C Asterisk PBX Configuration.....	96
SIP Configuration	96
Dial Plan.....	106
Appendix D Evaluator's recorded challenge questions.....	112

List of Figures

Figure 1-1 VoIP Revenue Forecast.....	3
Figure 3-1 an example of an INVITE message	22
Figure 3-2 User-to-User Session Initiation.	23
Figure 3-3 User-Proxy Server-User Session Initiation.	24
Figure 3-4 Asterisk SIP configuration	31
Figure 3-5 Asterisk extensions configuration.....	32
Figure 3-6 The signaling process of a regular SIP Call in our system. All messages except ACK are routed through Asterisk.....	32
Figure 3-7 An example of INVITE message designed to go directly to the target user without interaction with the PBX, for spamming purposes.....	33
Figure 3-8 The signaling process of an automatically generated VoIP spam message in our system. None of the messages go through the PBX as expected.	34
Figure 3-9 Screenshot of the simulation results from SIPp of three spam calls generated and all of them went through the complete INVITE to ACK to BYE circle successfully.....	35
Figure 3-10 Screenshot of the statistics result of SIPp of three spam calls	36
Figure 3-11 Results of VoIP spamming by MJSIP.....	37
Figure 4-1 System Overview	40
Figure 4-2 Flow of Anti-VoIP Spam Tool.....	41
Figure 5-1 Time for Installation & Configuration.....	54
Figure 5-2 Comfortable Rate of User's Feeling on Tool Installation & Configuration ...	55
Figure 5-3 Time to Design Challenge Questions.....	57
Figure 5-4 Time to Record Audio Questions.....	57
Figure 5-5 Screenshot of Phone Pad of the java softphone.	59
Figure 5-6 How Hard Users Feel about the Audio Questions	61
Figure 5-7 How difficult does user feel to answer the questions.....	61
Figure 5-8 User Responding Time for Each Audio Challenge Question	62
Figure 5-9 How annoyed does user feel about the challenge questions	63

List of Tables

Table 2-1 Email Spam vs. VoIP Spam.....	14
Table 5-1 Background of our test participants.....	51
Table 5-2 Audio Question Recorded Longest/Shortest/Average Time for Each Phone Number	60
Table 5-3 disk space consumed by the pre-recorded audio questions	69

Chapter 1 Introduction

§ 1.1 Voice over IP (VoIP)

VoIP refers to voice over Internet protocol, which is also known as IP telephony. Since the first Voice over Internet Protocol (VoIP) application was introduced by an Israeli company named VocalTec in 1995 [6], VoIP usage has been expanded dramatically over the last twelve years. By 1998, VoIP traffic represented 1% of all voice traffic in the United States. [7] In the year of 2000, networking manufacturers such as Cisco Systems and Nortel started to produce VoIP equipment capable of switching and routing VoIP traffic, which greatly facilitated the development of VoIP as a commercial communication system. By the end of 2000, 3% of all voice traffic had become Internet based. IDC, the premier global provider of market intelligence for information technology, telecommunications, and consumer technology markets [8], reported that about 10% of all voice calls were made with VoIP technologies in 2005. [9]

The required equipments to set up VoIP call are simple, a PC, or a special VoIP phone or traditional telephone with a special adaptor (Analog Telephone Adapter-ATA), high-speed Internet access are enough. The cost is low, compared with the traditional phone, especially for long distance calls. Most VoIP services charge as low as \$25 per month for all features (such as caller id, call waiting, call transfer, long distance, etc). Some VoIP companies even provides unlimited international calling for \$25 a month, such as iTalkBB. However, traditional phone companies are more costly with separate tariffs for local and long distance calls.

VoIP is not only inexpensive, but also flexible. VoIP calling can be dynamical binding, which means a user can make IP phone call from anywhere in the world as long as s/he has an internet connect. An Analog Telephone Adapter (ATA) is often used to convert a regular telephone into a VoIP phone. On the other hand, connecting to a traditional telephone service usually needs the layout of telephone line into a residence, which is both cost-inefficient and inflexible.

In addition, a VoIP service user can own virtual telephone numbers in VoIP services covered places. That is, if your friend or family in another state or country wants to call you without any additional cost or just pay the local call cost, what you have to do is to get a virtual telephone number in that state or country and have the respective call routed to your ATA irrespective of your current physical location. This allows the user's family to call him/her without paying for the long distance even international call cost. Besides, taxing is less on VoIP services. VoIP only pays Federal Excise Tax while traditional phone carries numerous taxes and other surcharges [10].

With all the advantages mentioned above, demand of VoIP is growing fast. According to the survey of *Business Week onLine*, by the year of 2009, the total VoIP subscribers are going to reach 133,633,938. [11] Researchers at iSuppli Corp. (El Segundo, CA) have projected that the number of residential VoIP subscribers worldwide will rise to 197 million by 2010 [12], a dramatic increase over the estimated about 7 million consumers utilizing VoIP tools in 2006. Figure 1-1 (reprinted from telephonyonline.com) shows the dramatic growth of VoIP revenue from 2002 to the estimated revenue of 2008. [13] In terms of annual revenue, the North American VoIP service revenue is expected to soar to \$19.9 billion in 2009. [14] Compared with

traditional phone call, traditional phone service revenue plunges 88% while IP phone service revenue forecast to grow 82%. [15] We can see that the future of VoIP is promising.

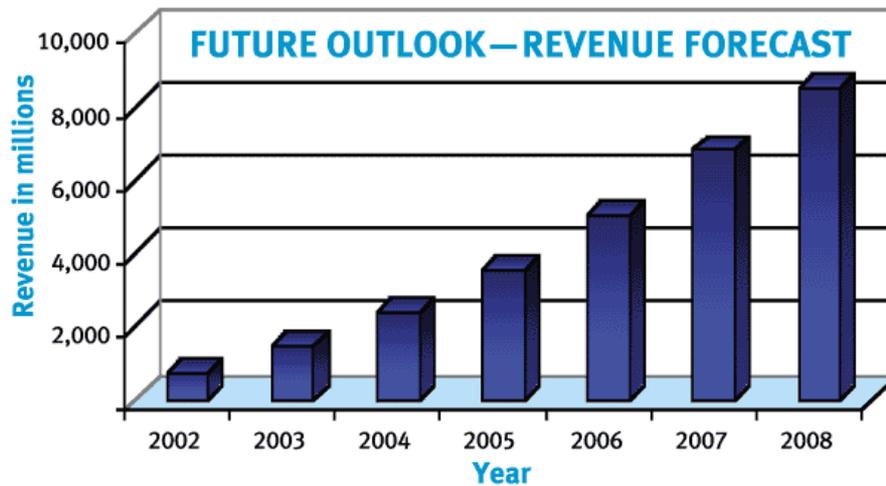


Figure 1-1 VoIP Revenue Forecast

§ 1.2 Spam

Spam is a message that has a large amount of copies flooding into the Internet and forcing people to receive it; whether they want to receive it or not. If a message qualifies the following conditions, it is considered as a Spam [2]:

- The content is irrelevant with recipient's personal identity since the same message is sending to many different unrelated recipients.
- Recipient does not grant the right for the sender to send messages to him/her and receiving the message is unwanted for the receiver.

Spam is an issue with consent, not content. The content of Spam can be varied from advertisement, free food to any commercial means. If it is unsolicited and sent indiscriminately to many individuals, mailing lists or newsgroups, it can be called Spam.

Forms of Spam are varied. The most widely recognized one is e-mail Spam, while there are many others, such as instant message Spam, mobile phone messaging Spam, Usenet newsgroup Spam, etc. No matter what kind of Spam, it is not only annoying but also can be dangerous. It may contain computer virus, spywares, or offer illegal goods. This monster also costs humans a lot of money and time. Each e-mail Spam consumes human's time to read, which is time wasting. Spam cost United States organizations alone more than \$10 billion in 2004, including lost productivity and the additional equipment, software, and manpower needed to combat the problem [3]. Some countries, such as Australia and USA, have declared that Spam is illegal [4].

Since e-mail is very cheap to send, it motivates spammers to saturate the Internet with junk e-mails. According to Ref. [5], it is virtually no cost to spammers. The cost of cable Internet connection is about \$45 per month in USA. Other Internet Service Provider can cost less, such as DLS service or dial up ISP. Anyone can send advertisements to others by e-mail through the Internet with a PC and the pre-paid Internet access account. As a result, e-mail users all over the world are being bothered by e-mail Spam.

§ 1.3 What is VoIP Spam?

With the rapid worldwide growth of VoIP services, the Spam issue in VoIP systems becomes a bigger and bigger concern [16]. A VoIP Spam or Spam over Internet Telephony (SPIT) is defined as a bulk unsolicited set of session initiation attempts (e.g. INVITE requests in SIP) trying to establish a voice communication session. If the user answers the call, the spammer proceeds to relay their message over the real time media. This is the classic telemarketer Spam, applied to a VoIP environment.

Call Spam itself is not a new problem – telemarketer calls have been around for decades. Although these calls are annoying, they do not arrive in the same kind of volume as email Spam. The difference is cost; it costs more for the spammer to make a phone call than it does to send email. This cost manifests itself in terms of the cost for systems, which can perform telemarketer call, and in cost per call.

Both of these costs are substantially reduced by the new VoIP technology. As a result, the cost of VoIP spam is much lower compared to the cost for traditional telemarketing. In particular, the cost per call for VoIP is extremely low since while the VoIP Spam can send many copies at one time whereas a traditional residential phone line allows only one call to be placed at a time. A SIP call Spam application is easy to write. It is just a user agent client that initiates, in parallel, a large number of calls.

For example, some industry observers found that the open nature of a VoIP phone call makes it easy for spammers to send VoIP Spam in much the same way they carpet bomb e-mail inboxes today. Any open, IP-based phone system could be a target of a spammer [17]. The cost to make SPIT is low. A computer and high-speed Internet access are enough to make SPIT. The low cost of making Spam has been discussed in section 1.1. It is easy to make SPIT. Engineers in Qovia lab only used two hours to create a VoIP Spam generator program, which knocked out a call manager in Qovia labs capable of handling 100,000 phones [18].

There are two ways of sending out VoIP spam calls: by a human caller or by a computer program. SPIT by a human is similar to the traditional telemarketing spam in that one caller can only make one phone call at a time. This is not an effective way for VoIP spamming because it can not take advantage of the capability of VoIP system to

make a large volume simultaneously. Automatic SPIT by a computer program is more dangerous because of its potential huge volume. Therefore, we focus on the feasibility of SPIT automation and the possible schemes to mitigate its threats in this study.

§ 1.4 Proposed Research and its Contribution

Motivated by the above reasons, we have developed a VoIP anti-Spam system. Our purpose is to mitigate automated VoIP Spam.

Black & white list is a good method to filter VoIP Spam. However, the system has two problems:

- How to treat a first time caller?
- Since the lists need constant human maintenance, how to keep them up to date?

Above problems can be solved by reverse Turing test method together with the black & white list. If the call's ID is neither in black list nor white list, it needs to be filtered by reverse Turing test. Only the one, who passed reverse Turing test challenge, can establish the call. Caller, who failed the reverse Turing test, is treated as spammer and the call is torn down. Consecutively challenge failed user IDs will be automatically added on black list while frequently test passing user IDs are automatically put on white list. If the reverse Turing test challenges are recorded by attacker, attacker can launch Spam by the learned challenge pattern. Therefore, diversity question is a good solution to solve this problem. It is hard for attackers to learn all challenge patterns only based on a few recorded questions if the diversity challenges are used.

Besides above problems, it is possible to send VoIP Spam without routing proxy server or registrar. Server based Spam detection system has little effect on this issue,

because the abnormal behavior pattern of a spammer is harder to be detected on the server side. Therefore, a client mode anti-spam program based on reverse Turing test is necessary. Our anti-Spam program is embeded in softphone, which is client side and can be used to solve this problem.

The contributions of this thesis work is shown below:

- 1) It is feasible to send automatic VoIP Spam. In this thesis work, it is demonstrated that VoIP spam can be automatically generated and routed without the involvement and control of a proxy server or registrar by using SIPp simulation. We implemented a VoIP spam generator, which repeated our observations from SIPp simulation and shows it is feasible to send automatic VoIP Spam.
- 2) Our system can mitigate VoIP Spam no matter it routs through proxy server or not since it is built on client side.
- 3) Reverse Turing test is a good supplementary to the existing black & white list on VoIP anti-Spam system. As we mentioned earlier in this section, black & white list has problems on first caller's identification and maintenance. Reverse Turing test is a good method to solve those problems.
- 4) Diversity questions is extention to reverse Turing test on VoIP anti-Spam system. Diversity challenges make it harder for attack to track all reverse Turing test challenge patterns.

- 5) It is an implementation of reverse Turing test in a voice context, including challenge design. Our experiments show that the reverse Turing test is good to be used on VoIP anti-Spam system.

§ 1.5 Organization

The remainder of this thesis is organized as follows. The next chapter discusses the related work on VoIP Spam. Chapter 3 describes the feasibility of automatic VoIP Spamming, which presents VoIP Spam generation. Chapter 4 describes the method to mitigate automatic VoIP Spams using reverse Turing test. Chapter 5 presents the experimental results of our system. Chapter 6 concludes this thesis and talks about some future work. Appendix A is the user instruction guide, which shows how to install and configure our tool. Appendix B is the questionnaire, which is used for our testing objectives to report usage information. Appendix C is configuration of our Asterisk PBX. Appendix D is the evaluator's recorded challenge questions.

Chapter 2 Related Work

The term “Spam” refers to the unsolicited and undesired bulk messages, which have been automatically generated to send to the personal accounts of the users of electronic messaging systems for marketing or other purposes. Spamming is economically viable because the barrier to entry is so low and the advertiser has virtually no operating costs other than the management cost of their mailing lists. Due to the large volume of Spam messages and their social impact on the Internet users, there has been a great deal of attention recently on the research of anti-Spam security mechanics.

While the most widely recognized form of Spam is email Spam, the term is applied also to similar abuses in other media, such as instant messaging Spam, chat room Spam, mobile phone messaging Spam and discussion forum Spam. As one of the fast growing on-line messaging systems, Voice over IP (VoIP) system is also vulnerable to the similar type of Spam threats, with the cost efficiency being the major drive for VoIP spamming. For example, Spam messaging for telemarketing purpose through a VoIP system is going to be much cheaper as compared to the traditional telemarketing.

In this chapter, the anti-Spam approaches for text based Spam, including content filter, black and white list and Turing tests are first discussed in section §2.1. In §2.2, the difference of text-based and VoIP spam is addressed. The last section reviews the current status of research work on VoIP spam detection.

§ 2.1 Text based Spam detection

Most anti-Spam studies so far target the text-based Spam messaging, especially email Spam. In this section, the investigations on three most popular anti-Spam solutions,

Black and White Lists, Content Filtering and *Turing Tests*, are reviewed in separate parts of this section.

2.1.1 Content Filtering

The nature of Spam determines that the fundamental difference between a Spam message and a regular message is their content. For example, it is easy for a human recipient to recognize a Spam email for advertising a product because it most likely will contain some characteristic words associate with this product. The challenge is how to train a computer program to make the same judgment. [19]

One popular email Spam detection technique is called Bayesian Spam filtering [20], which is the process of using a Naïve Bayes classifier to identify Spam email. A Bayesian Spam filter take advantage of Bayes’s theorem, which says, in the context of Spam, can be expressed mathematically as [21]

$$P(spam | words) = \frac{P(words | spam) \times P(spam)}{P(words)} \quad (2.1)$$

where $P(spam | words)$ is the probability of an email being a spam given that it has certain words in it, $P(words | spam)$ is the probability of finding those certain words in spam email, $P(spam)$ is the probability that any email is spam and $P(words)$ is the probability of finding those words in any email.

The probabilities of some particular words existing in a spam email and in a regular email are very different. For instance, most email users will frequently encounter the phrase “FREE!!!” in spam email, but will seldom see it in other emails. The filter does not know these probabilities in advance, and must first be trained so it can build

them up. To train the filter, the user must manually indicate whether a new email is spam or not. For all words in each training email, the filter will adjust the probabilities that each word will appear in spam or legitimate email in its database.

After training, a Bayesian spam filter will search an incoming email for the “sensitive” words and use the corresponding word probability (also known as “likelihood function”) are used to compute the probability of the email being a spam. Each word in the email contributes to its spam probability based on Bayes's theorem or equation (2.1). If the total spam probability of an incoming email exceeds a certain threshold (say 95%), the filter will mark the email as a spam and automatically move it to a "Junk" email folder.

2.1.2 Black and White Lists

Blacklist is an access control anti-spam technique based on a list of identified spammers to block unsolicited Spam messages. The opposite is white list, which contains the addresses or domains of trusted users to allow the legitimate messages through. A well-maintained black or white list is very effective to block the known sources of Spam [23]. The list can be updated either manually or automatically. Manual update of the black list or white list requires the human user to specify whether a message sender is a spammer or not. On the other hand, the automatic list maintenance works based on the assumption that certain action towards an email indicates it is a Spam email. For example, if the emails from a particular sender are constantly deleted after only reading the title, the sender is most likely a spammer. Such updating (or learning) process is “reactive” and has limited efficiency -- it may take up to several days to identify a spam

source [24]. Also, it is possible for the black lists to block a friendly incoming message (a false positive).

A question related to the white list is how to treat the first greeting user, or so called introduction problem [25]. Since the white list needs to “learn” if the first meeting user is a friendly one, he/she has to be included in the white list at the beginning, which creates potential security concerns. Another common issue is the cheating behavior of a new user, which means that a malicious spammer can behave friendly at the very beginning to gain the reputation to remain in the white list and then starts to send Spam messages.

2.1.3 Turing Test & Reverse Turing Test

As initially described by Professor Alan Turing [31] in 1950, Turing Test is a test of a machine’s capability to perform human-like conversations. Assuming a human judge engages in a natural language conversation of two parties, one a human and the other a machine that tries to appear human, if the judge can not reliably tell which is which, then the machine is considered passing the test.

Recently there has been a great deal of interest in applying the variation form of Turing Test or so called Reverse Turing Test to separate machine spammers from regular human users of an anti-Spam system. The term “Reverse Turing Test” is a somewhat ambiguous term, used to describe various situations on the Turing test in which the objective and/or one or more of the roles have been reversed between computers and humans. In the spam detection case, it refers to a Turing test that is administered by a computer rather than a human judge, to achieve automatic generation and grading of Turing-like challenges.

This kind of computer administrated reverse Turing test in the spam protection world is also called a CAPTCHA (Completely Automated Public Turing Test to Tell Computer and Humans Apart). A CAPTCHA, initially introduced by Luis von Ahn *et. al.* in 2000, is a computer program that can generate and grade tests that: (a) most human can pass, but (B) current computer programs cannot pass [32][33]. In the acronym, the P for Public means that the code and the data used by a CAPTCHA should be publicly available.

As a general concept, CAPTCHA can be implemented in many different ways. Examples of CAPTCHAs are Gimpy [32], Bongo [34], PIX, ECO [35], Byan [36], and ARTiFACIAL[37]. Among them, the most popular program is Gimpy, which is widely used by many Web services including Yahoo! [38]. GIMPY works by picking up some words out of a dictionary, rendering a distortion containing the word, and then presenting it to a human user for recognition test. While most users have little difficulty recognizing these words, the tests appear very hard to current computer programs. Breaking a CAPTCHA generally requires some effort specific to that specific CAPTCHA implementation (e.g., OCR or Optical Character Recognition for GIMPY), and an abuser may decide that the benefit granted by automated bypass is negated by the effort required to engage in abuse of that system in the first place.

§ 2.2 Text Spam vs. VoIP Spam

Voice over IP systems, like email and other text-based Internet applications, are susceptible to abuse by malicious parties who initiate unsolicited and unwanted communication. VoIP Spam, sometimes also referring as Spam over Internet Telephony (SPIT), certainly shares a lot of commonality with email spam. For example, both of

them contain bulk of undesired messages, which are both annoying and can flood a user's mail (or voice mail) box.

Table 2-1 Email Spam vs. VoIP Spam

	Email Spam	VoIP Spam
Protocol	SMTP/POP3	SIP/RTP
Session	Asynchronous	Synchronous
Source of Spam	SMTP clients	Softphones
Result	Annoying/email box full	Annoying/call server outage /voice mailbox full
Time to delete	Deciding to delete after seeing subject line	Deciding to delete after hearing message

However, voice Spam is quite different in several aspects from email Spam, as summarized in Table 2-1 [30]. A practical protection scheme against voice Spam needs to have some unique features compared to the general anti-Spam solutions in the previous section. All three existing email spam detection methods fail in the new VoIP application, as discussed below:

First of all, Bayesian content filtering in an email system is not applicable to the case of VoIP spasm because email Spam is asynchronous while VoIP Spam is synchronous in terms of session control. As opposed to email protocols such as SMTP (Simple Mail Transfer Protocol), VoIP does not tolerate any negotiation or screening of content in the signaling step. The call recipient or an anti-spam software can only know the content of a phone call when it is picked up, which is too late for a filter to take any action.

Second, a black/white list for VoIP network needs even more human maintenance and intervention than its email counterpart because it is more difficult for a computer program to learn from a phone call receiver's action towards a caller to know if that caller is a malicious spammer or not. In the case of email, a human user can tell easily from the title and sender's name if a message is a Spam. So if he/she keeps deleting messages from a certain source without reading the content of these messages, the source is most likely a spammer. However, in the current VoIP systems, the user will not be able to tell a spammer from regular callers until he/she picks up the call (simple Caller ID function is not sufficient for many people to reject a phone call). As a result, a traditional black list or white list for VoIP Spam protection requires initial human input and constant maintenance. Also, given that a regular telephone is used together with a VoIP adapter as the client device in many situations, a special function needs to be implemented on the regular phone to add a caller to the white/black list, which can be cost inefficient.

Finally, the existing reverse Turing test implementations in an email Spam protection system, such as Gimpy and PIX, may not work since they usually involves the transfer of an image or other content. In the world of VoIP systems, since most of the calls are voice, the technique needs to be adapted to voice [25]. This process actually may bring additional advantages in reverse Turing test anti-Spam solution. The voice challenges are closer to natural language environment, which make the challenges harder for computers and easier for humans. The other difference is that a VoIP user can only enter the answer to a reverse Turing test on a phone pad, while in the case of email a keyboard is often available. This causes some limitations to the choice of challenge types for voice Spam detection.

In a word, due to the unique nature of a VoIP system, it is necessary to put more research effort in the field of VoIP anti-spam, which is the main emphasis of this thesis. The related work in this field that has already been carried out by other research groups is introduced in the next section.

§ 2.3 Review of Current VoIP Anti-spam Research Works

As mentioned in §2.2, although Bayesian Spam filtering technique has been proven to be able to mitigate the email Spam problem successfully, the algorithm highly relies on the fact that the email systems are text based and the filter program can know the content of an email before the email account users. However, in a VoIP system, it is almost impossible to analyze the content of an audio communication before the phone call is established [22] and it is already too late for any filtering when the human recipient answers the call. Therefore, there is few research work done on content filtering solution for anti-VoIP-spamming. Most of the proposed anti-spam schemes for VoIP applications focus on either black and white list approach, or Turing tests, both of which are discussed separately in the rest of the section.

2.3.1 Black and White List Family

Ram Dantu and Prakash Kolan [27, 28] presented a voice Spam detection system based on the feedback from users utilizing both blacklisting and trust & reputation functions. The software tracks each caller's past history and the user's reaction to each and every caller. Reputation system is a special implementation of the black or white list method to improve the performance and efficiency of Spam detection. Reputation system provides a mechanism that collects, distributes, and aggregates feedback about

participants' past behavior [26]. The algorithm then calculates the corresponding reputation value through Bayesian learning. If the caller's reputation is below a certain threshold, the call is classified as a Spam and blocked. It has been shown in their studies that this idea is able to detect the Spam messages with much higher accuracy and efficiency. Another reputation-based mechanism that builds trust between users using a social networks approach enhanced with some reputation ratings is also proposed to cope with SIP Spam problems [29].

In a recent paper published by Shin *et. al.* [30], a new Spam protection algorithm called Progressive Multi Gray-Leveling (PMG) method is discussed. PMG monitors only call patterns from each caller and determines the Spam calls based on these patterns, instead of requiring user's feedback all the time. The algorithm calculates the "gray level" of a caller (the level that determines if the caller is a likely Spam source or not). Unlike the traditional black lists, the status of a caller is changing progressively based on its previous call patterns. When a call is received, PMG calculates two gray levels, one for long-term gray level and the other for short term. If the summation of the two levels is less than a given threshold, the caller is considered a regular user and the call is connected. Otherwise, it is regarded as a Spam and blocked. PMG, however, gives callers multiple chances to atone for their undesired behaviors. If a blocked user stops sending out spams, his gray level is going to decrease and eventually the block will be removed. But if the user tries to attack again, PMG can block him at a quicker rate because of the Spam history of this user.

However, in the context of VoIP or voice Spam, unlike the content filtering techniques, using only blacklisting (including reputation system and PMG) by itself is

hard to effectively stop Spam phone calls or Spam over Internet Phony (SPIT). There are some disadvantages of using these methods alone in a VoIP system. First of all, they are somewhat “reactive”, meaning that the black list, or reputation or gray level builds over time and some of them need constant maintenance. Especially, most of blacklisting and PMG algorithms block the SIP spammers through their URI (Uniform Resource Identifier, e.g. username@proxy.com), instead of using their actual IP addresses [30]. So if the malicious user can get a new user name or change it easily, he/she can appear to the Spam protection system as a new user. This requires another learn cycle to identify the same spammer and thus make the system much less effective. Secondly, black or white lists often generate larger false positive percentage compared to filtering.

2.3.2 Turing Tests Family

Since using only black and white list approach is not sufficient for VoIP spam detection, reverse Turing test based solutions become necessary. With most of the CAPTCHA program targeting mainly toward on-line text messaging services, ECO was proposed as a sound version of Gimpy [35]. The program picks a word or a sequence of numbers at random, renders the word or the numbers into a sound clip and distorts the clip. It then presents the distorted sound clip to its user and asks the user to type in the contents of the sound clip. Although the concept of ECO is proposed in Ref [32], it has not been implemented and evaluated in a VoIP system.

In January of this year, NEC announced that they would debut its new Turing test based SPIT blocker, called VoIP SEAL. This company did not give the details of how VoIP SEAL works but gave three features of this product: [39]

- Calls arising from spam-generating-software and calls from real individuals are separated by a Turing test.
- By adopting a module structure, VoIP SEAL enables rapid response to new kinds of SPIT attacks, without adjusting the system, by adding and updating modules to respond to new and different kinds of SPIT.
- The adoption of a module structure also realizes response to a broad range of applications by enabling flexible and easy customization of systems to meet the needs of a variety of hardware, such as SIP servers, SBC home network equipment and terminal equipment.

NEC claimed that 99% of SPIT was detected and blocked, but there is no indication of when this commercial product will be released.

By definition, CAPTCHAs are completely automated, requiring little human administration and maintenance, and therefore produce benefit in cost and reliability. Also, properly designed CAPTCHA challenges can screen computer programs from human user effectively with very high accuracy.

In the context of VoIP spamming, a sound based CAPTCHA test is applicable, although ECO appears to be not directly suitable for VoIP spam detection. How to implement a voice spam protection scheme and the evaluation of this anti-spam method is the main topic of the current research work.

Chapter 3 Feasibility of Automatic VoIP Spamming

In order to study the anti-Spam schemes for a VoIP system, the feasibility of automatic VoIP Spam needs to be discussed first. That is, whether it is possible to automatically make the Spam phone calls to a group of VoIP users without the knowledge of IP configuration or the phone number of each specific user. Through the tests carried on our automatic VoIP spamming system, the answer has been proven to be positive. The same system also forms the basic platform for the anti-Spam studies.

This chapter starts with the introduction on Session Initiation Protocol (SIP) and its key elements in section 3.1.1. Section 3.1.2 discusses common attacks in a SIP network, SIP authentication and related potential security concerns. The last section of this chapter shifts the focus on VoIP spamming as one of the security concerns of a SIP based system. The implementation and demonstration of an automatic VoIP generator in our lab is discussed based on two different scenarios, depending on whether the Spam needs to go through the proxy servers in the VoIP system or not.

§ 3.1 Background on Session Initiation Protocol

3.1.1 Introduction of Session Initiation Protocol (SIP)

The Session Initiation Protocol (SIP) is an application-layer control (signaling) protocol for creating, modifying, and terminating sessions with one or more participants. These sessions include Internet telephone calls, multimedia distribution, and multimedia conferences." [40] As a signaling protocol widely used for Voice over IP, SIP has the following features:

- Light weight, in that SIP has only six methods, reducing complexity;
- Transport-independent, because SIP can be used with UDP, TCP, ATM etc.
- Text based, allowing for humans to read SIP message.

There are three major kinds of basic components (roles) in SIP:

- UA: works on behalf of users to set up calls;
- Proxy Server: keeps track of location of end-points;
- Registrar: each UA registers to inform current location and preferred reachability information.

As a text-based application layer protocol, SIP works on the basis of human readable SIP messages. The message syntax and header files are similar to those of HTTP/1.1, although SIP is not an extension of HTTP. SIP also uses a client server model, which means that the client initiates SIP messages called Request Messages to the server, while the server “answers” these requests with SIP messages called Response Messages.

A request message begins with a SIP method name, which is used to define the request type. Some important SIP methods include INVITE (initiate a session), ACK (acknowledge session initiation), OPTIONS (query server capacities), BYE (terminate a session), CANCEL (cancel a pending session) and REGISTER (register user’s location). SIP messages require six basic elements: SIP method, Via, Max-Forwards, To, From, Call-ID. As the first request INVITE method (also called INVITE message) is used to initiate a session. A simple example of an INVITE message is shown in Figure 3-1, which contains the six required elements for a SIP message. In this figure, *INVITE* is the message method. *Via* contains the address at which the caller is expecting to receive responses. *Max-Forwards* limits the number of hops on the way to the destination. *To*

field contains the destination address to which the caller sends the message. *From* field has the caller's address. *Call_ID* is the globally unique identifier, referring to this particular call session.

```
INVITE sip: 100@Asterisk.csc.ncsu.edu  
Via: SIP/2.0/UDP Asterisk.csc.ncsu.edu  
Max-Forwards: 70  
To: phone0_2006 <sip: 100@Asterisk.csc.ncsu.edu>  
From: phone1_2006 <sip: phone1_2006@Asterisk.csc.ncsu.edu>  
Call_ID: 123456789@Asterisk.csc.ncsu.edu
```

Figure 3-1 an example of an INVITE message

Response message is a SIP message sent from a server to a client, for indicating the status of the server's attempt to process a request. Each and every response message contains a 3-digit integer status code representing the outcome of an attempt to understand and satisfy a request. The first digit of the status code defines the class of response. SIP/2.0 allows six values for the first digit:

- 1xx: Provisional -- request received, continuing to process the request;
- 2xx: Success -- the action was successfully received, understood, and accepted;
- 3xx: Redirection -- further action needs to be taken in order to complete the request;
- 4xx: Client Error -- the request contains bad syntax or cannot be fulfilled at this server;
- 5xx: Server Error -- the server failed to fulfill an apparently valid request;
- 6xx: Global Failure -- the request cannot be fulfilled at any server.

Now let us look at an example. **Figure 3-2** illustrates how a VoIP session is initiated successfully through a series of SIP signaling in the simple user-to-user

communication case (no proxy involved). At the beginning, UAC (User Agent Client) on the caller side start the negotiation by sending out an INVITE message, ask the UAS (User Agent Server) on the callee side to establish a call. When UAS receive the request, it responds with a 180 Ringing message and prompt the callee to ask if he/she would accept or decline the call, normally by ringing the IP phone. If the callee decides to take the call, UAS will then send a 200 OK response back to the caller. The caller confirms that with an ACK message and the session control negotiation is done. If one of the users wants to hang up the phone, it sends a BYE message.

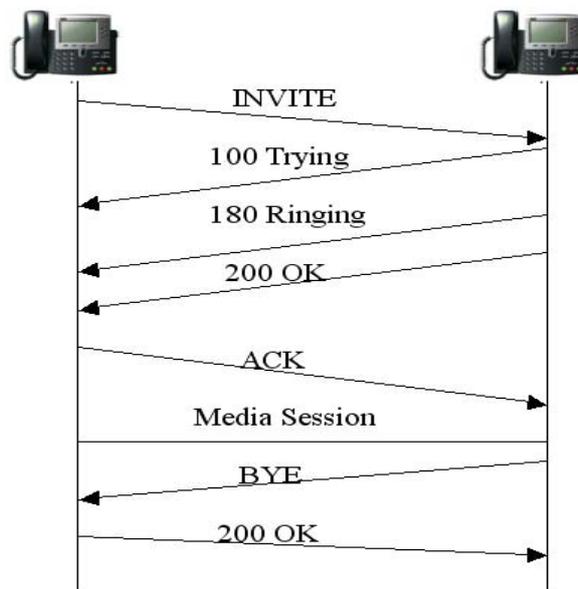


Figure 3-2 User-to-User Session Initiation.

The example of **Figure 3-2** is a simplified case, which is based on the assumption that the caller knows the current IP address of the callee. In reality, the same user may move between different locations and therefore can be available at various physical IP addresses. A user is usually recognizable and reachable by using his SIP identity, a type of Uniform Resource Identifier (URI) called a SIP URI. It has a similar format to an

email address, which typically contains a user name and a host name, for example, sip:phone100@Asterisk.csc.ncsu.edu. The host name can be the domain name of a user's SIP service provider.

The participants of a SIP call may come from different SIP service providers and have no knowledge of each other's IP address (or where another user is) prior to the SIP session initiation. In this case, proxy servers and registers are necessary. Since the sender (caller) does not know the location of the target receiver (callee) or the SIP server in the target domain, the soft phone sends the INVITE message to the proxy server that serves his own domain (eg. caller_domain.com) first, as shown in Figure 3-3. The address of this server could have been configured in the caller's soft phone, or it could have been discovered by DHCP, for example.

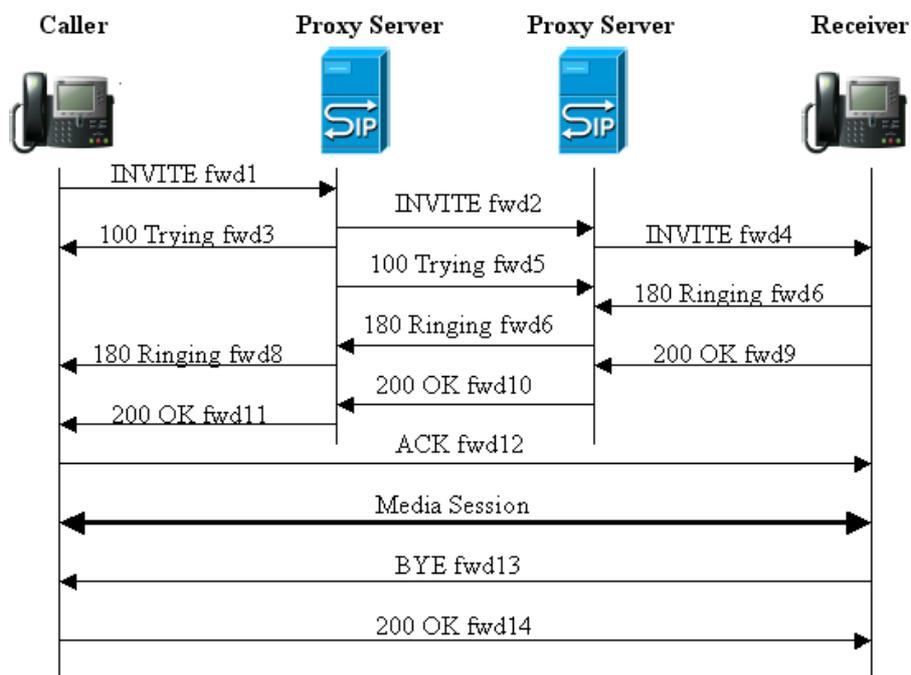


Figure 3-3 User-Proxy Server-User Session Initiation.

The proxy server receives the INVITE request from the caller and responds with a 100 Trying message, which indicates that the request is received and the proxy server is working on the requestor's behalf to route the INVITE message to the destination. The next thing that this proxy server (for caller_domain.com) will do is to perform some type of DNS (Domain Name Service) lookup and locate the proxy server that serves the target domain (callee_domain.com). As a result, it forwards the INVITE message to the callee_domain proxy server, which responds with another 100 (Trying) response back after receiving the INVITE message. The target proxy server consults a database containing the current IP address of the callee, generically called a location service, and forwards the INVITE message to its final receiver. The corresponding 180 Ringing and 200 OK responses are routed in the similar fashion, only this time the address lookups are mostly not performed since they have already been stored in the Via field during the initial routing.

Registration is another common operation in SIP. A user's soft phone usually sends a REGISTER message to a registrar server at periodic time intervals. The REGISTER message associates a user's SIP URI dynamically to the computer or IP phone to which he is currently logged on. The location service, which typically contains this association, or dynamic binding, is used by the proxy server to route the requests. Often, a registrar server for a domain is co-located with the proxy for that domain. Obviously, the distinction between types of SIP servers is logical, not physical.

3.1.2 SIP Authentication and Potential Security Concerns

SIP is subjectable to several kinds of security attack, which is why SIP authentication is needed [41]. The common security threats associate with SIP include

replay attack, registration hijacking, request spoofing, server impersonating etc.. Replay attack is a retransmission of a genuine message by a malicious user to set up an authorized session with the victim entity. Registration hijacking means that an attacker impersonates a party and changes the registration information (or location service) to redirect all the requests for the affected user to his device. Request spoofing is used by the malicious person to fool the legitimate recipient through a forged identity and message because he/she does not want to use his/her real address for various reasons (eg. cheating or billing etc.). If an attacker can successfully impersonate a proxy or registration server, he/she is able to gather the user information as well as redirect the legitimate requests to an insecure entity.

SIP provides authentication methods to verify that a request sender and/or receiver are legitimate. In a SIP based network, authentication can take place when a user agent client (UAC) initiates a request to a server (proxy, registrar or UAS), where the server requires credentials from the client before processing the request. A user agent client, on the other hand, can also require authentication of a server. Authentication is needed for the registration (REGISTER), session setup (INVITE), session modification (RE-INVITE) and termination (BYE) processes.

SIP authentication adopts the challenge-based Digest Authentication from HTTP, which uses an MD5 hashing function on the username/password combination. In the digest authentication scheme, when a UAC tries to establish a conversation with a server, the server responds with a 401 Unauthorized message (for a UAS, registrar or redirect server) or a 407 Proxy Authentication Required (for a proxy server) to challenge the

UAC. Until UAC resends the original message with the correct Credentials, its request will not be processed by the server.

When a server sends the the 401 Unauthorized or 407 Proxy Authentication Required response, UAS defines an implementation dependent “nonce” value to authenticate the client. A recommended digest scheme should include the client’s IP address and a time-stamp in the “nonce” value. This is designed to prevent replay attack since in this case the attacker has to get the sender’s IP address correct before the time-stamp of the “nonce” expires. The UAC, on the client side, will generate the credentials based on this “nonce” value and the MD5 hashing of the legitimate user’s username and password, along with some other information (eg. digest URI). Usually the client should also let the server know the username. Due to the nature of SIP, these credentials are included in the client authentication header and transmitted over the internet with plain text format (without encryption).

Upon receiving the client authentication header, the server may check its validity by looking up the corresponding password to the submitted username (ref: RFC 2617). The server must then perform the same digest algorithm again as the client did, and compare the results. Note that the password of the user in plain text format is not always necessary. As long as the hashing function of the username/password is available to the server, it can perform the verification.

SIP authentication provides a certain degree of security against malicious attacks. However, there are also limitations. First of all, in the latest version of SIP documents (RFC 3261), SIP authentication is only an option (keyword MAY), which is not mandatory. In many commercial softphones (e.g. Windows Messenger), digest

authentication is not implemented at the client side for direct user-to-user communication. As a result, the attacker can easily reach the victim entity without being challenged by the 401 Unauthorized response, which is demonstrated in the following section. Secondly, digest authentication is not enough and some authors recommended Shiboleth [42]. Finally, in the particular case Internet spamming, a legally registered SIP user can be a spammer too, which means that SIP authentication can hardly stop these kinds of attacks. Therefore, a supplementary spam detection scheme is necessary, which is the topic of this thesis.

§ 3.2 Automatic Spam Generation

In the previous sections, we discussed SIP attacks and the potential security concern of SIP authentication in general . This section focuses on a particular type of security attack – VoIP spamming. As mentioned before, spamming is the abuse of an electronic message system (email, forum, instant messaging or VoIP etc.) to send unsolicited bulk messages, which are universally undesired. In the current study, a VoIP spam detection and filtering scheme based on the so called Reverse Turing Test is proposed as a supplementary security measure for SIP authentication. As the foundation of this work, the implementation and demonstration of an automatic VoIP spam generation system based on a SIP network is addressed in this section to further justify the necessity of a new method for VoIP spam detection.

There are two basic scenarios of automatic VoIP Spam generation, depending on whether the Spam needs to go through the proxy servers of the VoIP network or not. In the VoIP system established in our lab, Open Asterisk is used as a PBX (Private Branch

Exchange, traditionally a telephone switch to connect telephone calls), which acts like both the proxy server and registrar in the SIP network.

Sending out the Spam messages through a proxy server is not an optimized idea for VoIP spamming. If the spammer is a registered user of the PBX, a simple program can be implemented to just auto-dial a range of numbers and flood the VoIP Spam out through PBX. However, in this case, the big volume of voice mail coming from one registered user will pull up the attention of the proxy server. PBX can easily notice and detect the unusual communication behaviors with other Spam detection methods discussed in Chapter 2. However, in the case that the attacker chooses to send out the Spam messages without going through the proxy server (or PBX), these Spam detection schemes based on the server side will not be able to protect the SIP network any more. In the real world, the smart spammers tend not to route their voice spams through a specific proxy server.

It is most likely and challenging that if the spammer, on the other hand, is not a registered user of the PBX and the phone number of the target remains unknown, the attacker can still automatically send out the VoIP Spam by searching IP addresses with the knowledge of binding open SIP port. In other words, if an attacker wants to send VoIP Spam to a victim VoIP phone, getting the IP address is not enough, he also needs to find out the binding open SIP port to make the Spam sending successful. An open SIP port is a basic requirement to successfully establish an SIP channel. For example, a range of the IP addresses, such as 152.1.2.3 to 152.1.2.100 is first checked for an open SIP port, such as the standard SIP port of 5060. If an open port is detected, a VoIP Spam automatically floods into that IP address bound VoIP phone. Since the spammer is not an

unregistered user and the messages are not routed through the PBX, it has no control over the spamming situation. Compared to the previous case, it is much harder to detect the attacks based on the second strategy, which is going to be the main approach we adopted in our system and also the focus of the rest of this section.

A VoIP Spam automatic generation and testing system has been set up in our lab. The basic VoIP network consists of a PBX, a registered client machine and a spammer machine. The PBX is the open source software of Open Asterisk, version 1.4.0-beta3. The software provides voice mail over IP services in many protocols and is configured in SIP mode in our system. Registered phones can make calls to each other through Asterisk. Asterisk runs on various operating systems such as Linux, BSD and MacOSX. In our case, it is Linux Fedora 5.

For the user agent (UA) roles, the commercial soft phone in our VoIP system used as the target entity for spamming is Windows Messenger 5.0, a real time communication client. The attacker program or the VoIP auto-spamming program is implemented on top of MjSip Version 1.6, which is a java-based open source soft phone.

Open Asterisk, Windows Messenger and the MjSip-based auto-spamming program form the three essential elements needed in our testing system. When the three key devices are chosen successfully for the VoIP automatic Spam generation system, the next step is to connect and configure each individual element together to be a whole system. In order to make Windows Messenger be an alive phone client, it needs to be registered with Asterisk. Thus Asterisk configuration for Windows Messenger is necessary.

There are two types of basic configurations for Asterisk: SIP configuration and dial-plan configuration. The SIP configuration file of Asterisk is named as *sip.conf*, part of which looks like Figure 3-4. Note that in this figure, comments are added behind each program line to explain its function. The file starts with a *[general]* section, which includes the general default settings for all user channels, including the context, authentication realm, SIP port, DNS and domain. DNS SRV refers to the Domain Name System Service record, which binds the actual IPs for outbound calls to the logic addresses. *[phone0_2006]* section contains the information of a specific user registered to the Asterisk PBX, such as the user name, password, user category etc.. In this example, the user phone0_2006 belongs to the “friend” category with the capability of sending and receiving calls.

```
[general]
context = default           ; Default context for incoming calls
realm = Asterisk.csc.ncsu.edu ; Realm for digest authentication
bindport = 5060             ; UDP port to bind to (SIP standard port is 5060)
bindaddr = 0.0.0.0         ; IP address to bind to (0.0.0.0 binds to all)
srvlookup = yes            ; Enable DNS SRV lookups on outbound calls

[phone0_2006]
type = friend               ; Enable sending and receiving calls
secret = testvoip000        ; Password of the user
context = from-sip          ; The from-sip context controls what we can do
host = dynamic              ; This device registers with us
username = phone0_2006     ; User name of this softphone
```

Figure 3-4 Asterisk SIP configuration

The second part of setting up an Asterisk system is the configuration of dial plan, which routes every incoming and outgoing call in the system from its source through various applications to its final destination [ref: The Asterisk handbook]. In Asterisk, a dial plan is coded in the file of *extensions.conf*, which contains a collection of extensions. The basic syntax of an extension looks like:

exten => name, priority, application()

where name is the user name or phone number of a user, application() is applications or arguments to run, and priority refers to the order in which the series of application is executed. Figure 3-5 shows an example of a part of Asterisk extensions configuration. Here the phone number “100” is bound to two applications that need to be executed by Asterisk towards the SIP user “phone0_2006”: Dial() and Hang-up(). In the Dial() application, it contains two arguments, with the first one specifying a reachable destination named *phone0_2006* in a SIP system and the second argument setting the time-out limit of this application to be 60 seconds. This small paragraph of configurations tells Asterisk to keep on dialing a SIP user called phone0_2006 for 60 seconds before canceling and running the next application Hangup() application to hang up the phone.

```
[From-sip]
exten => 100,1,Dial(SIP/phone0_2006,60)
exten => 100,2,Hangup()
```

Figure 3-5 Asterisk extensions configuration.

As we mentioned previously, it is more effective to send the Spam message directly to the target user without routing through the PBX. In order to achieve this, the technical feasibility needs to be examined first. For example, Figure 3-1 shows what a normal INVITE message should look like in the simplest form, which only includes the six required fields -- SIP method, Via, Max-Forwards, To, From, Call-ID.

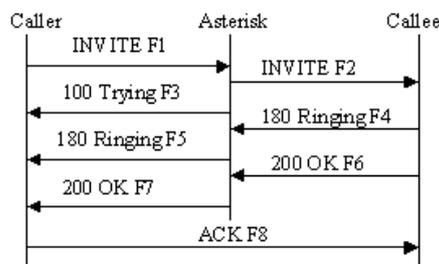


Figure 3-6 The signaling process of a regular SIP Call in our system. All messages except ACK are routed through Asterisk.

Figure 3-6 illustrates the signaling process of setting up a regular IP phone call in our system. The initial INVITE message is sent to the PBX asking for the callee by his URI, such as phone0_2006@Asterisk.ncsu.edu. PBX looks up this URI in its location service and forwards the INVITE message to the callee. The callee responds to this message with a 180 Ringing message to the server name in the *Via* field, which is forwarded back to the caller by Asterisk. This is in fact a simplified version of Figure 3-3, since the caller and callee are both in the same domain and thereby share the same proxy server (PBX).

```
INVITE sip:152.1.2.3 <target IP address >  
Via: SIP/2.0/UDP spam.com < spammer's IP address >  
Max-Forwards: 70  
To: 152.1.2.3 <target IP address >  
From: spam.com < spammer's IP address >  
Call_ID: 123456789@spam.com
```

Figure 3-7 An example of INVITE message designed to go directly to the target user without interaction with the PBX, for spamming purposes.

Note that in the negotiation of a regular SIP phone, Windows Messenger as the UA responds to the basic INVITE message without any authentication challenge during the session initialization process. This indicates that it is possible to send the Spam message to the target client without going through PBX. To do this, an INVITE message similar to Figure 3-7 is designed. The major difference of the two INVITE messages in Figure 3-1 and Figure 3-7 is that without PBX no location service is available any more so the actual IP addresses rather than the SIP URI are needed in the *To* and *From* fields etc. The corresponding signaling process is illustrated in Figure 3-8. No PBX participation is expected.

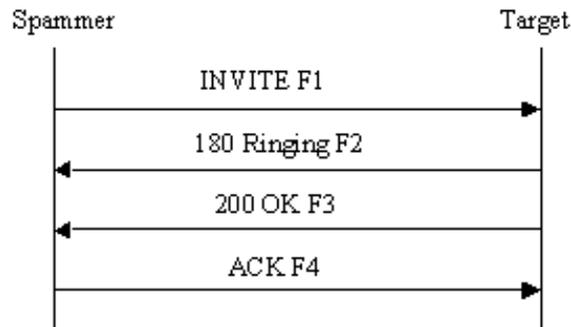


Figure 3-8 The signaling process of an automatically generated VoIP spam message in our system. None of the messages go through the PBX as expected.

Based on the above analyses, it appears to be possible to send a VoIP Spam to the target's SIP-based softphone without going through PBX. The next thing to do is to prove it in our experimental VoIP system. Here simulation tool SIPp is used to help the demonstration as well as the data analysis in this experiment. SIPp is an open source performance-testing tool, which can generate traffic for the SIP protocol. We generated three Spam phone calls from SIPp UAC without registration with Asterisk and sent the traffic to a registered Windows Messenger. It is clearly seen in Figure 3-9 that all three calls have successfully gone through the INVITE to ACK to BYE circle and accepted by the UAS (Windows Messenger). Figure 3-10 is the statistical analysis result of this simulation, which shows that the number of outgoing created calls is the same amount as the successful call. These results prove that the automatic VoIP Spam generation is feasible without the participation of the PBX.

With the feasibility of automatic VoIP Spam generation without PBX proven, the remaining question is how to find out which IP address and which port number of this IP is bonded to a VoIP softphone. Among the Internet protocols, SIP is somewhat similar to HTTP in that it is based on the request/response transaction model. When a request is

transmitted to the right port of an SIP server, at least one response can be expected. Apparently this also indicates that finding an open SIP port on this particular server is required. Based on this nature of SIP, a probing INVITE message is designed and sent to every port of a target IP address. If a SIP response is coming back, we can tell that a SIP-based softphone is bonded with that IP address and also which port is the dedicated SIP port. VoIP Spam can be sent to that SIP phone through the responding port number. This is also the basic idea of the port scanner embedded in our auto Spam tool.

```

C:\WINDOWS\system32\cmd.exe
C:\Program Files\SIPp>sipp 152.14.92.175:16271 -i 152.14.92.232 -p 5064 -t un -t
race_err -m 3 -trace_msg -r 3 -rp 6000
Warning: open file limit > FD_SETSIZE; limiting max. # of open files to FD_SETSI
ZE = 64
----- Scenario Screen ----- [1-4]: Change Screen --
Call-rate<length>      Port      Total-time  Total-calls  Remote-host
3.0<0 ms>/6.000s      5064      12.91 s     3            152.14.92.175:16271(UDP)

Call limit reached <-m 3>, 0.910 s period 1 ms scheduler resolution
0 concurrent calls <limit 1>          Peak was 1 calls, after 2 s
0 out-of-call msg <discarded>
1 open sockets

Messages  Retrans  Timeout  Unexpected-Msg
INVITE ----->      3         0         0
100 <-----      3         0         0
180 <-----      3         0         0
183 <-----      3         0         0
200 <----- E-RTD  3         0         0
ACK ----->      3         0
Uar Pause [      0ms]  3
BYE ----->      3         0         0
200 <-----      3         0         0
----- Test Terminated -----

```

Figure 3-9 Screenshot of the simulation results from SIPp of three spam calls generated and all of them went through the complete INVITE to ACK to BYE circle successfully.

Our auto VoIP Spam tool is built on the top of MjSip. MjSip is a SIP-based open source softphone, written in java. This Spam generator has two functions: one is to look for VoIP phone, which is implemented by our java-based port scanner program (UDP_SIP_PortScanner); the other feature is to send the audio Spam to the victim, which is implemented by the MjSip.

```

C:\WINDOWS\system32\cmd.exe
----- Statistics Screen ----- [1-4]: Change Screen -----
Start Time      : 2007-03-04 22:16:34
Last Reset Time : 2007-03-04 22:16:46
Current Time    : 2007-03-04 22:16:47
-----
Counter Name    : Periodic value      : Cumulative value
-----
Elapsed Time    : 00:00:00:912         : 00:00:12:996
Call Rate       : 0.000 cps           : 0.231 cps
-----
Incoming call created : 0                : 0
OutGoing call created : 0                : 3
Total Call created  : 0                : 3
Current Call       : 0                :
-----
Successful call    : 1                : 3
Failed call       : 0                : 0
-----
Response Time     : 00:00:01:042         : 00:00:02:168
Call Length       : 00:00:01:170         : 00:00:02:301
-----
Test Terminated

```

Figure 3-10 Screenshot of the statistics result of SIPp of three spam calls

The VoIP Spam generation experiment sets up as the following. As we mentioned earlier, we need to find out the VoIP phone first. The factor to find out if the IP address is attached with a VoIP phone is if there is an open SIP port for VoIP phone. UDP-SIP_PortScanner works on scanning the listening SIP port of the victim. Since our PBX Asterisk only supports UDP, the listening SIP port of the attached VoIP client must be UDP port and we just need to scan UDP port. The UDP_SIP_PortScanner borrowed some codes from Bill Clagett's TCP simple port scanner [43]. UDP_SIP_PortScanner sends SIP INVITE message to every UDP port of the specific IP address ranged from port number 1 to 65535. If SIP responses (i.e. 180 Ringing) are detected from any port,

that port number will be returned to the auto Spam tool. Thus we know there is VoIP phone attached with the probing IP address.

After finding the VoIP phone, the remaining item is to send voice mail to this victim. MjSip can handle it as regular phone call. It sends the INVITE message to the target. Once the target picked up the phone or mailbox was activated, VoIP spam can flood in.

Our experiment has observed that the VoIP Spam has been successfully sent from spammer to our target Windows Messenger. Figure 3-11 is the print screen of our Spam software console, which shows the successful performance. *Test call to 152.x.x.x* means the victim's IP address is 152.x.x.x. *Local host* shows the IP address of the spammer. *Port is 7151* carries the information of target's SIP listening port number. *UAC: CALLING sip: 152.x.x.x: 7151* shows that spammer sends INVITE message to victim, whose IP address is 152.x.x.x through SIP port 7151. *UA: RINGING* means the receiver's phone is ringing. The last line *UA: ACCEPTED/CALL* indicates the information of that the other party has accepted this call, which implies that Spam is flooding in.

```
Graphical MJSIP UA 1.0
test call to 152.x.x.x < IP address of Windows Messenger >
Local host is : 152.x.x.x < IP address of spammer >
Quote of the Moment: port is 7151 received
UA: UAC: CALLING sip:152.x.x.x:7151 < IP address of Windows
                               Messenger >
UA: RINGING
UA: ACCEPTED/CALL
```

Figure 3-11 Results of VoIP spamming by MJSIP.

Chapter 4 Mitigating Automatic VoIP Spams using

Reverse Turing Test

With the feasibility of automatic VoIP Spam generation addressed previously, we proceed in this chapter to the discussion of a proposed anti-Spam solution using reverse Turing test. Since the general concepts about Turing test and reverse Turing test were included in 2.1.3 for mainly text-based messaging systems, the first section of this chapter starts with the introduction of the overview of our proposed anti-VoIP-spam system to mitigate the automatic VoIP spams, as well as how it is used in conjunction with the traditional black and white list method. Section §4.2 focuses on the challenge-based VoIP Spam detection solution using reverse Turing test, which fits seamlessly into SIP infrastructure. As a challenge-based HIP (Human Interactive Proof) system, the challenge design is the most important consideration to guarantee the success of differentiating humans from machines. The general guidelines about question design and their application in a VoIP system are discussed in section §4.4. Section §4.3 evaluates the security advantages of using customized challenge question in the reverse Turing test VoIP Spam detection system.

§ 4.1 System Overview

In Chapter 1, we proposed that a Turing-like challenge based VoIP Spam detection scheme can be applied as the supplementation to the conventional black and white list (including reputation) system to mitigate the Spam problems in a VoIP network. This

section focuses on the system overview of how reverse Turing test and black/white lists could be working for Spam detection.

The proposed system used in reality is overviewed in Figure 4-1, where white list is used to keep the trusted IDs and black list to keep the blocked IDs. As mentioned previously, the disadvantage of black/white listing includes:

1. The system would not be able to know how to treat a first time caller;
2. The lists need constant human maintenance to keep them up to date.

These two situations can both be improved by using our reverse Turing test method together with the black and white lists.

First of all, if a caller is requesting phone connection for the first time and the ID is not in either black or white list, it should be challenged by the reserve Turing tests before being allowed through. A reverse Turing test is triggered to check the “unknown” IDs, either a first timer or a caller who has not gained enough trust to be put into the white list. That is, a voice challenge is automatically sent to the caller. If the caller responds to the challenge with the correct answer, the phone call is put through; otherwise it is rejected.

Second, this Spam detection scheme allows the black and white lists to be managed both manually by a human user and automatically according to the result coming from the reverse Turing tests. If a caller has been rejected (connected) consecutively within a certain period, it should be put into the black (white) list. For this kind of automatic list management, which would be the primary option, the threshold for putting a user into either black or white list should be set fairly high. Only those well-known callers (a legal regular caller or a notorious spammer) is immediately connected or turned down.

Therefore, combining reverse Turing test and black/white list in a VoIP anti-Spam system allows us to take advantage of the benefit of both approaches. Since black/white listing has been widely considered for VoIP [23-25], this study focuses only on how to implement reverse Turing test in a voice IP network and the related evaluation. The details of how the reverse Turing test program works are described in next section.

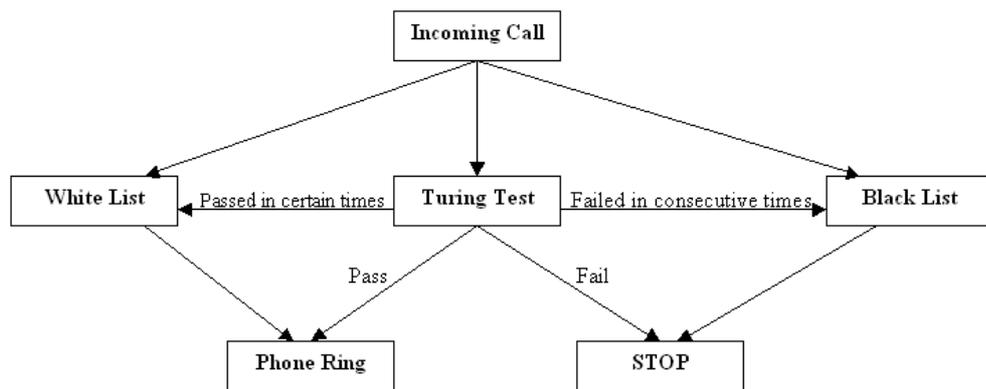


Figure 4-1 System Overview

§ 4.2 Challenge-based VoIP Spam Detection

As mentioned in the last part of §2.2, challenge-based VoIP Spam detection requires the implementation of reverse Turing test using voice calls. The procedure is as follows: when a VoIP call is received by the client softphone on the receiver side, an anti-Spam program is going to take over and send a pre-recorded voice challenge back to the caller; the caller is asked to answer the question by entering it on his phone pad within a certain time out limit; If the answer is verified successfully, the anti-Spam software is going to quit the process and normal calling process takes over. Otherwise, if the answer

is verified as wrong or time is out, the session is torn down. The flow of the above description is shown as Figure 4-2.

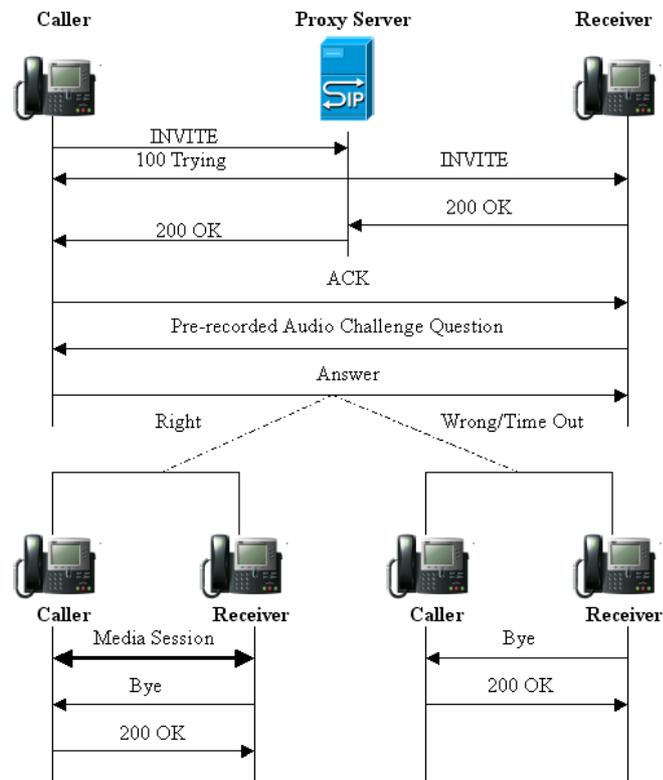


Figure 4-2 Flow of Anti-VoIP Spam Tool

The anti-spam solution described above should be categorized as one of the reverse Turing test algorithms in that the objective of the challenges is to distinguish human from machine rather than to improve the intelligence of a computer program. Instead of pulling sound clips of words out of a dictionary and distorting them with noises as proposed by ECO [32], we focus on the study of how to use voice challenges pre-recorded by a human user to conduct reverse Turing test. The difficulty for a computer program to answer these questions is not to recognize the words in a sentence but to understand the *meaning* this sentence trying to convey. To break this kind of anti-Spam system requires much more sophisticated AI, indicating potentially very high

accuracy in telling spammers apart from regular users using our system. The key to the success of our proposed Spam protection scheme is to design the appropriate challenges, which is further discussed in the next two sections.

§ 4.3 Security Threats and Analysis

In this section, the possible security threats related to VoIP threats are discussed and the methods to mitigate these threats are proposed and analyzed. Four major categories of threats are considered:

First of all, it is proved in Chapter 3 that making spam calls without routing through a proxy server is feasible. As a result, server based filtering technique can be less effective if no server is involved in the spamming process. In order to mitigate VoIP spam no matter it routes through a proxy or not, a client mode control scheme is needed, which in our study is the reserve Turing test based anti-spam program.

The second possible threat is that the attacker records can the audio challenges used for spam detection and then launch attack after learned our challenge patterns. The challenge design is the essential element of our system. If the challenge is easy to be interpreted by the attacking machine, our system has a big security hazard. Making challenge questions hard for machine is our security concern.

The solution we proposed in this study to protect this system is to introduce diversified questions in reserve Turing tests, i.e. to let users to design the challenge questions. Each user needs to design ten challenge questions, which means each softphone has ten different patterns. Millions of users have millions of different patterns. No two users will have exactly the same ten questions since human's thoughts are varied. It is hard for the attacker to catch all. Attacker can record all the challenge questions from

one user and find out the answer to attack the specific one, but attacker cannot use it to attack the others. Otherwise, attacker needs to record all the user's questions in order to attack all, which is meaningless. If the challenges are pre-recorded by us and the softphones are delivered to all the users with the same challenge pool, once the attacker records all the patterns from one user, the attacker can use it to attack all the users. Thus, it is good to have users to design and record the audio questions. Besides, another advantage to design and record the audio questions by users are that accent is varied. Everyone has his own natural accent. English has London, American, Australia, Indian, etc. accent. Even for Americans, the people from South or from North speak different accents.

The third possible security concern is no-effect attack, which refers to the attacks that can achieve a high hacking accuracy rate without solving hard AI questions. A yes-or-no question allows a spam program to "guess" the answer without solving an A.I. question and still get 50% chance to get the correct response. To solve this problem, all answers to the questions should be combination from 0 to 9. This may effectively forbid no-effort attacks. From the report we got from our testers, above 50% of the answers are between 0-9. To further reduce the susceptibility of our system to no-effect attacks, an answer required at least two number digits should be recommended. Because all the answers are entered by the caller using a phone pad, which implies that if the correct answer is only one digit, a spammer has a 10% of change of passing the test by just guessing one single-digit number every time.

Finally, the possible consequence of brute force attack should be considered because a spam software could be programmed to try all possibilities to get the answer.

Although the spammer may not eventually get through, a lot of precious network resources are consumed during this kind of brute force attacks. If a softphone is based a portable device that runs on battery power, the power consumption due to brute force attacks may also not be tolerable.

The damage caused by brute force attacks can be mitigated by introducing black & white listing techniques. Any ID or IP address that is failed in a certain consecutive times should be blocked in the black list. By doing this, one could limit the connection attempts from the same IP address or URI. For example, a user could setup the anti-Spam system so that five consecutive failed connection attempts will cause a caller's IP address or URI to be blocked. This would not only reduce the passing rate of brute force attacks, but more importantly makes it risky for a spammer to choose brute force attack and therefore save network bandwidth and power etc.

§ 4.4 Considerations for Diversified Challenge Design

A research area called Human Interactive Proofs (HIP), whose goal is to defend Internet services from malicious user by differentiating computer bots from human users, is attracting more attention recently. In this section, the general guidelines that ensure an HIP system to be useful and secure are discussed, along with how these guidelines affect the challenge designs for our VoIP Spam detection system.

Generally speaking, a successful HIP should have the six following desired properties [32, 34, 37]:

- The test should be automatically generated and graded by computer;
- The test should be user-friendly;
- The test should be easy for human users;

- The test should be hard for machine to pass;
- The test should be resistant to no-effort attacks;
- The test should be robust when database is publicized;
- The test should be universal for global users.

In this section, we explain these principles and apply them to an anti-voice-Spam system to consider the better practice for voice challenge design. As mentioned before, it is preferable that a softphone user can enter the answer to a voice challenge on a phone pad, which limits the question to be a mathematical problem, or a multiple choice question etc. This is the hardware constraint. On the other hand, a well-designed audio question should also satisfy some other criteria.

1) The test should be automatically generated and graded by computer:

The voice challenges in the proposed VoIP Spam detection system is pre-designed, recorded and stored along with the answers in a database, similar to the image challenges in CAPTCHAs such as PIX and Animal-PIX [35]. When an IP phone call is received, one of the challenges will be automatically pulled out and played. Depending on the implementation, it is optional to distort the corresponding sound clip by adding noise or other post processing steps. Once the caller answers a particular challenge, the response is compared by the auto-grading program with the correct answer, which determines whether the incoming call gets through or not.

2) The test should be user friendly:

An optimized challenge for a VoIP anti-Spam software should be user friendly because it is partially designed for human users. Whether a human user feels that he/she is treated fairly or not during a test will directly affect the usability of a proposed anti-

Spam system. Inevitably a challenge-based Spam detection scheme will cause delay for regular users. Therefore, the question should be short and clear. For one question, it should take less than 30 seconds, including the time needed for the callee side to ask the question and for the caller side to respond. Also, too much information in a question tends to confuse the regular caller – if the human user has to ask for the replay of a voice clip, it most likely will double the delay time.

A user-friendly question also implies that if noise or other distortion is to be applied to the original sound clip, it needs to stay in a moderate level to avoid causing unpleasant feeling to a human user. A voice message could still be recognizable (easy for humans) but becomes so annoying that it affects the customer experience of a VoIP system user. There is a fine line that needs to be defined in this case.

3) The test should be easy to human users:

A well-design anti-spam challenge for VoIP should be easy enough for a human user to answer correctly in the matter of seconds. If the question is a mathematical problem, it should not require too complicated calculation. If the question is a multiple choice one, it should not require knowledge other than common sense. This guideline is brought up to make sure that virtually all human users pass the test. As long as the user understands what the question is, there should be no additional effort or knowledge necessary to answer it. The test is made difficult for a machine in terms of understanding the question, which is discussed in the next part.

4) The test should be hard for machines to pass:

The test should fail virtually all current computer programs (ideally also for predictable future programs) with setup of a time out limitation. The nature of a VoIP

system allows us to ask vocal questions in a natural language environment, which is of course easy for humans. A lot of unique techniques beyond physical distortion of sound can be applied to make the questions almost impossible for a machine to understand. For example, one may use interchangeable concepts and expressions, or imperfect grammar to confuse the computer program. Putting things into context is another option. Speaking in an accent, which is more natural and less annoying than simple distortion of a sound clip, would also make automatic speech recognition much more difficult.

Interchangeable concepts and expression: In a natural language system, multiple words could refer to the same or similar things. For instance, it is natural for a human to realize that kids, boys and girls and children are the same concept. On the other hand, a look-up table is needed for a machine to learn the connection between interchangeable concepts and expressions, which requires a lot more effort.

Imperfect Grammar: When people speak in a certain language, the imperfection in grammar generally can be tolerated and will not affect the effectiveness of communication. For a machine, since it relies on the recognition of grammar or other patterns to catch the meaning of a sentence, an intentional mistake confuses the A.I. program.

Speaking in an accent: Instead of adding noise or an unnatural distortion to the pronunciation of some words, simply speaking a question in an accent would make it harder for a machine to recognize but still reasonably easy for humans. For example, in a diversified society with many ethnic groups such as the United States, most people need to communicate with others who speak English in a different accent, like British,

Chinese, Indian, or Spanish accents etc. For an automatic speech recognition system, however, a heavy accent is still one of the major factors causing mistakes.

Putting things in context: One unique feature of the way humans communicate is that we normally put things in context. First of all, the same word in a different context could refer to a different thing. Most natural language terms-such as verbs, nouns, conjunctions, prepositions-have a context-dependent meaning. For example, Mercury in English could mean a chemical element, a planet, a car brand, a record label and many other things. Secondly, in a real conversation, people often use a pronoun to substitute something they mentioned in the previous statement. If there are multiple things in that sentence, it is going to be very hard for even state-of-the-art A.I. programs to figure out which concept a pronoun refers to.

Apparently, there are many other things in a natural language other than what have been discussed here to make a voice challenge harder for a machine to comprehend and answer. Besides, if several techniques are applied simultaneously in a challenge, it would possibly become more effective in terms of differentiating computer bots from human users in a VoIP system.

5) The test should be resistant to no-effort attacks:

No-effort attacks are attacks that can achieve a high hacking accuracy rate without solving hard AI questions. For example, CAPTCHA program Bongo can easily be guessed with 50% accuracy without working on any hard AI problem. Bongo is classified into two categories – left and right. If an attacker tries the same side every time, he/she has a 50 percent chance of getting through. Even the spammer does not break the system;

a lot of precious Internet resources have already been consumed during the brute force attack processes.

It is important to make sure that our proposed anti-Spam system to be less susceptible to this kind of brute force attack. In the VoIP case, all the answers are entered by the caller using a phone pad. This implies that if the correct answer is only one digit, a spammer has a 10% of change of passing the test by just guessing one single-digit number every time.

To overcome this problem, one could limit the connection attempts from the same IP address or URI. For example, a user could setup the anti-Spam system so that five consecutive failed connection attempts will cause a caller's IP address or URI to be blocked. This would not only reduce the passing rate of brute force attacks, but more importantly makes it risky for a spammer to choose brute force attack and therefore save network bandwidth etc.

6) The test should be robust when the database is publicized:

The test should be difficult to attack even if the database from which the test is generated is publicized. [37] Even our questions are publicized; the factors as accent, noisy background also makes it to be hardly understood by machine.

7) The test should be universal for global users:

The purpose of this system is to detect VoIP spammer, but language cannot be global. Therefore, this condition cannot be applied.

Chapter 5 Experimental Evaluation

This chapter focuses on the experimental evaluation of the proposed reverse Turing test based VoIP Spam detection solution. Section §5.1 discussed the usability of our anti-Spam system, or how easy it is for a human user to install, configure and use the software. In §5.2, we evaluate the percentage rate for human users and computer programs to pass the tests. The performance issues are addressed in §5.3, with the emphasis on how much resources are required for the implementation and operation of an anti-VoIP-spam program.

§ 5.1 Usability Study

A good anti-spam system needs to be user-friendly, which means that it should require minimum installation and confirmation efforts. Also, the delay caused by the software to accept a regular phone call and reject a spam call should not exceed a certain limit. To evaluate the reverse Turing challenge based VoIP anti-spam system, we invited 45 people from different backgrounds, such as major, degree, occupation and gender, to form the test volunteer group. Table 5-1 summarizes the various backgrounds of the participants for evaluation tests. They are from eleven different majors, with one of them as a co-majored student. The current occupation includes student (most of them are from North Carolina State University), visiting scholar, post-doc researcher, industrial professional and homemaker.

Each of these volunteers participated in the evaluation as both the roles of receiver (callee) and sender (caller). The experimental details of call receiver evaluation

are presented in §5.1.1. Based on the data collected from each and every user, we analyze the usability study for call receivers in section 5.1.2. As a caller, one evaluator is asked to call three to five different VoIP telephone numbers, and report the percentage rate of passing the reverse Turing test as well as the delay of the each call due to the test. The detailed caller evaluation procedure is described in §5.1.3 and the corresponding results are analyzed and discussed in §5.1.4.

Table 5-1 Background of our test participants

DEGREE	MAJOR	TESTER #
PHD	CIVIL ENGINEERING	3
PHD	COMPUTER SCIENCE	9
PHD	ELECTRICAL ENGINEERING	7
PHD	INDUSTRY ENGINEERING	3
PHD	MECHANICAL ENGINEERING	1
PHD	OPERATIONS RESEARCH	1
PHD	PHYSICS	5
PHD	STATISTICS	2
PHD	TEXTILE ENGINEERING	2
PHD	ZOOLOGY	1
MS	COMPUTER SCIENCE	5
MS	COMMUNICATION	1
MS	ELECTRICAL ENGINEERING	2
MS	INDUSTRY ENGINEERING	1
MS	STATISTICS	1
BS	COMPUTER SCIENCE	1
BS	COMMUNICATION	1

5.1.1 Usability Study for Calleees

The evaluation process started with the participants acting as the call receiver (callee), whose task is to install and configure the anti-Spam tool in a lab VoIP system, as well as design and record ten voice questions as anti-Spam Turing challenges. A user instruction and a challenge design guideline were written by the authors and provided to the test participants to walk them through the installation/configuration and question recording steps. The users were then asked to report how long it takes for them to install

& configure the VoIP Spam detection program and how difficult it is to complete the task using a receiver version questionnaire (Questionnaire A, Appendix B).

The callee version of the usability evaluation experiment proceeds as follows:

1. A user first needs to check if the pre-required software programs has been installed properly on the callee side computer. These softwares include JRE (Java Runtime Environment version 3.5.0 or higher) and JMF (Java Media Framework) used for audio data transfer. Both JRE and JMF are easily downloadable from the website of Sun company (JRE 6 Update 1 <https://sdlc6d.sun.com/ECom/EComActionServlet;jsessionid=337DF60C14A4278705ECACD4BDC325DE>; JMF <http://java.sun.com/products/java-media/jmf/2.1.1/download.html>). Note that the version of JRE is accessible by type “java-version” in the command prompt. Also, if the network firewall is armed on a computer, one needs to make sure that it is not blocking JRE.
2. The user instruction guide for our anti-Spam softphone client is attached in Appendix A, in which the detailed steps for the installation and configuration of the software is described. After all pre-requirements are satisfied, the user needs to install and configure our software. Installation is quite simple; it only requires unzipping our software in a stand-alone folder. There are two requirements for configuration – fill up the IP address and answers of the ten pre-recorded vocal questions on configuration sheet. How to find the IP address is described in the User Instruction. The vocal questions will be discussed in step 3. User instruction has all the detailed explanations to walk user through this step.

3. User needs to record ten challenge audio questions. The main idea of our method is to provide the pre-recorded challenge questions to caller to distinguish human from machine. The pre-recorded challenge questions are recorded by our testing subjects. Challenge design guideline is included in the user instruction to help user to design and record the ten vocal questions.

5.1.2 Callee Evaluation Data Analyses and Discussions

During the experimental procedures and details described in the previous part, each and every test participant was asked to fill out a questionnaire designed to collect information about the call receiver's background and user experience using the anti-spam tool. The questionnaire (Questionnaire A) consists of eight questions, most of which are multiple choice, and can be found in Appendix B.

For the callee usability study, we are interested in how much time it takes for a user to setup our anti-Spam environment. It is clearly seen in Figure 5-1 that most of the evaluators spend only less than 10 minutes to install and configure the software with the help of the user instruction. To be specific, 87% of the users take less than ten minutes for the install step and 49% take even less than five minutes, which indicates that the time required for general users to setup the anti-Spam software falls into a reasonable range. According to our record, the degree or major background does not affect the installation and configuration time. Among the users who took more than 10 minutes, most of them have a PhD degree and 33% of them are computer science majors.

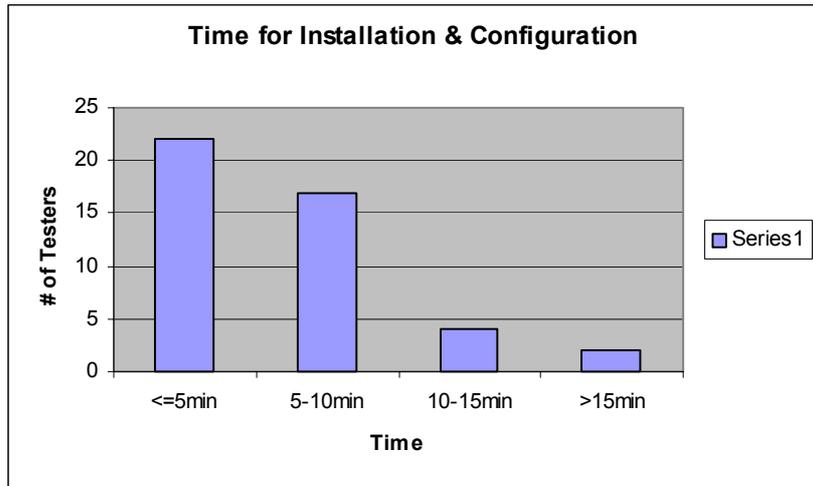


Figure 5-1 Time for Installation & Configuration

The questionnaire also includes a survey on the difficulty level for a callee to install and configure our VoIP Spam detection system. Only one user out of the 45 evaluators finds the software to be “a little bit hard” to setup and no one considers the installation step to be very difficult. 56% of the users feel it is easy to do the whole system setup and 42% of them feel it is acceptable. Based on the reports we received, difficulty level for installation & configuration plays fair game for our all kinds of background testers. The only one who feels it is a little bit hard is a computer science PhD student. Half of the computer science majored testing objects feel it is easy while the other half feel it is not bad.

For a challenge-based VoIP spam detection system, the voice challenge design and record step is an import part of user experience. If the user feels this step is difficult and time-consuming, the anti-Spam solution then becomes unrealistic. Similar to the user instruction, which walks a call receiver through the environment setup, a challenge design guide was put together to help the users to design the effective and proper challenge questions. The general ideas of designing a good reverse Turing test for VoIP

Spam detection have already been discussed in Chapter 4. The challenge design guide (Appendix A) concerns mainly the application of these ideas into a VoIP system and gives some specific easy-to-follow advice, including:

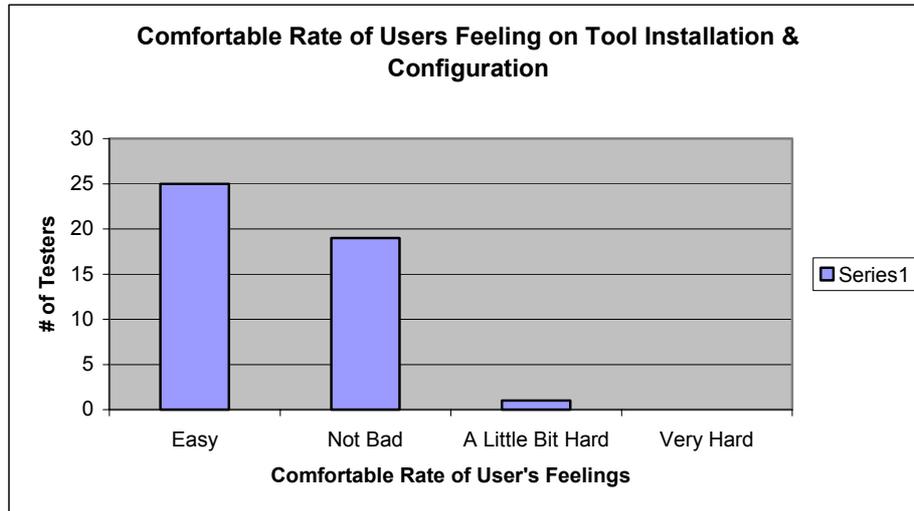


Figure 5-2 Comfortable Rate of User's Feeling on Tool Installation & Configuration

- Recorded question should be less than 30 seconds. If it is too long, the caller will feel that it is wastage of time. Usage degree of comfort, in using the software, will be degraded.
- Recording question should be computationally easy. The advantage of machines is the strong power of computation. Humans have a limited computational ability. Any calculation question needs to be simple and easy to be answered by human in a very short time, approximately less than 30 seconds.
- Recorded question should have a fuzzy logic. It is easy for the machine to follow the pattern of *plus* or *minus*. If a non-numerical scenario is added, machine can be easily fooled. An example question can be: there are five clouds in the sky. Wind

blows away two. How many clouds are left. Another example, I have 12 cats and 999 roses. How many flowers do I have?

- Based on the characteristics that our phone pad has, short letter answers are also accepted, besides numbers. However, the letters have to be short, no more than 10 letters. For example, the question can be: *what's my last name?* Try to avoid asking your first name when you have one official first name and one so-called name.
- Multiple-choice questions should be avoided. These kinds of questions can easily be attacked by brute force attack, since multiple-choice question lists cannot be long; if the lists are long, they are hard to be remembered by caller.
- Questions that require knowledge need to be avoided. Normally machine is more knowledgeable than people. Some examples of questions that should be avoided are: how many days are there in January. How many stars are there in the American flag?
- Try to speak loudly to record the challenge questions. Soft voice may cause difficulty for the caller to catch your question.

The challenge design often takes longer compared to the system setup. Figure 5-3 shows the time needed to design the challenge questions is less than 30 minutes for 89% of users. Only 5 users out of the total 45 (11%) spend 30 to 40 minutes on the voice question design. Since the user designed questions can be varied a lot more than phone company pre-designed questions. Millions of users can have more than millions different

challenge patterns, which makes it hard for the attacker to catch the design pattern.

Although it is a little bit time consuming, it is worth it with the security concern.

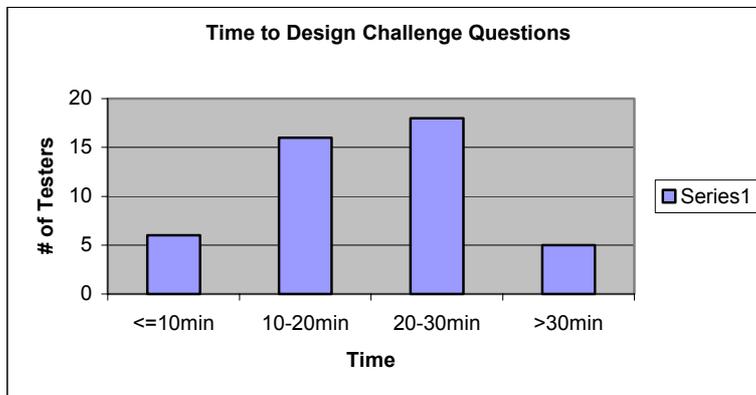


Figure 5-3 Time to Design Challenge Questions

Figure 5-4 is the time needed for users to record the ten audio questions. Since we ask in the challenge design guide for these audio questions to be kept as short and clear as possible, there are 76% of users who finished the recording within 15 minutes. And only two persons took more than 20 minutes, indicating this task can be done in a practical time frame.

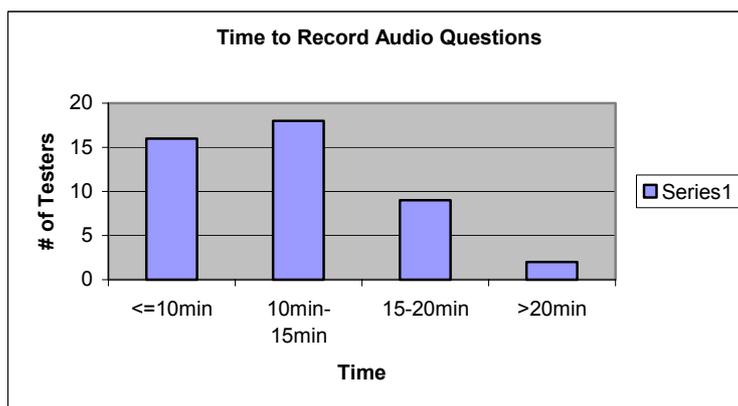


Figure 5-4 Time to Record Audio Questions

5.1.3 Usability Study for Callers

When each test participant finished his/her evaluation work as a call receiver, a VoIP softphone with the embedded anti-Spam capability has been installed and configured, and ten challenge questions have been designed and recorded into the system. In order to conduct the usability evaluation for callers, a unique user name and phone number were then assigned to every softphone using the regular Asterisk extensions configuration described in §3.2. Each caller needs to make three to five phone calls to different users, whose softphone contains different pre-recorded challenge questions. The caller study proceeds like this:

1. The VoIP softphone on the receiver side was loaded and logged in with the proper user name corresponding to one particular set of challenges. We control the “who is calling whom” scenario so that all of the questions were equally accessed. The appearance of our VoIP softphone looks like Figure 5-5.
2. The caller dialed a given phone number, which is assigned to another user with his/her designed challenges.
3. After the phone call was connected, caller would hear an instruction voice message: “Please answer my question. When you finish, press *send* to submit the answer”.
4. The softphone then started to play a random audio question from ten challenges pre-recorded by the call receiver. Caller responds this answer by punching the keys on phone pad.

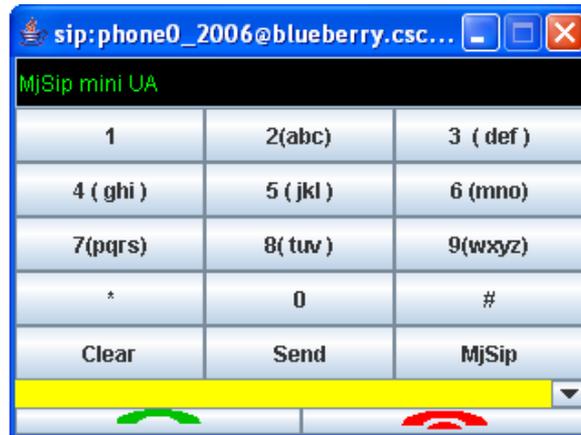


Figure 5-5 Screenshot of Phone Pad of the java softphone.

5. Repeat step 4 for another two to four different phone numbers.
6. The callers were then asked to fill out another questionnaire (Questionnaire B, Appendix B), which is used to collect the data for caller usability analyses.

5.1.4 Caller Evaluation Data Analyses and Discussions

The impact of VoIP Spam detection system on the caller's user experience is larger than that on the callee's in that every time when a regular user calls a VoIP softphone equipped with anti-Spam tool, he/she needs to pass the reverse Turing test. Therefore, as mentioned previously, the reverse Turing tests have to be user-friendly. The challenge questions should be short and clear. Among the 450 audio challenge questions (10 questions per user and 45 users) studied in this experiment, the longest pre-recorded audio question lasts 39 seconds and the shortest one ends within 3 second. Only two of the 450 questions take more than 30 seconds, which satisfies the delay time criteria we proposed in §4.4. Table 5-2 summarizes that for the 10 audio questions designed and recorded by each user (phone number), how much time the longest and shortest ones take and what the average question time is in the unit of second. One can also notice that the

average question times for most users range from 10 to 20 seconds, which is short enough. The overall average for all the challenge questions is about 14 seconds. This 14 sec is considered as the time spent on listening to the pre-recorded audio questions in our anti-Spam solution.

Table 5-2 Ten Audio Question Recorded Longest/Shortest/Average Time for Each Phone Number

Tel #	Longest Time	Shortest Time	Average Time	Tel #	Longest Time	Shortest Time	Average Time	Tel #	Longest Time	Shortest Time	Average Time
103	17	13	15.5	118	22	16	18.3	133	17	28	20.4
104	19	12	15.1	119	21	14	17	134	24	13	18.3
105	9	3	5.4	120	12	10	11.5	135	12	9	10.5
106	15	10	12.4	121	17	9	11.1	136	15	10	11.9
107	14	9	10.5	122	11	8	9.5	137	16	9	12.4
108	19	11	15.8	123	16	12	14	138	14	11	12.3
109	15	9	11.9	124	19	11	15.5	139	20	13	15.5
110	17	12	14.4	125	24	14	19.1	140	14	12	13.7
111	19	13	15.9	126	17	13	14.8	141	8	14	10.9
112	12	9	9.5	127	16	13	14.3	142	16	13	14.2
113	18	12	15.3	128	39	21	26.1	143	16	13	14.2
114	20	15	16.7	129	18	13	15.2	144	19	11	14
115	18	12	15.9	130	15	10	12.4	145	18	10	12.8
116	26	16	20	131	19	13	14.9	146	14	8	10.6
117	18	12	13.7	132	30	14	19.7	147	15	10	13.5

How much time is needed for the users to listen and understand the audio question is only one aspect for caller evaluation of our software. How long does it take for them to answer these audio questions? How difficult is it to understand and answer the questions? How long is required for answering them? These are also studied in the current work. Figure 5-6 presents the difficulty level of understanding the challenge questions. 54% of the callers feel the questions are easy to understand and 34 percent think they are understandable, which means that most of the challenges are properly designed to avoid confusion on the caller side.

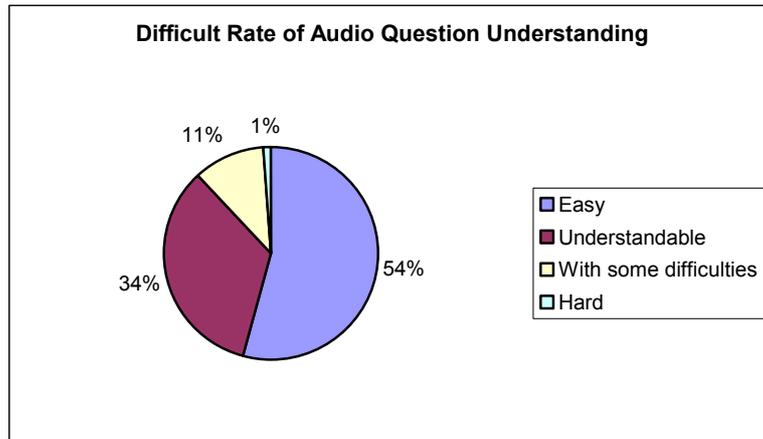


Figure 5-6 How Hard Users Feel about the Audio Questions

The difficulty to answer the challenge questions once they are understood is also surveyed and shown in Figure 5-7. According to the survey, 68% of the test participants found the questions with which they were challenged are easy to answer; and another 24 percent think the difficulty is acceptable. Only 2% of the users feel it is hard to answer the challenge questions. The issues associated with understanding and answering the pre-design challenges are further discussed in the next section.

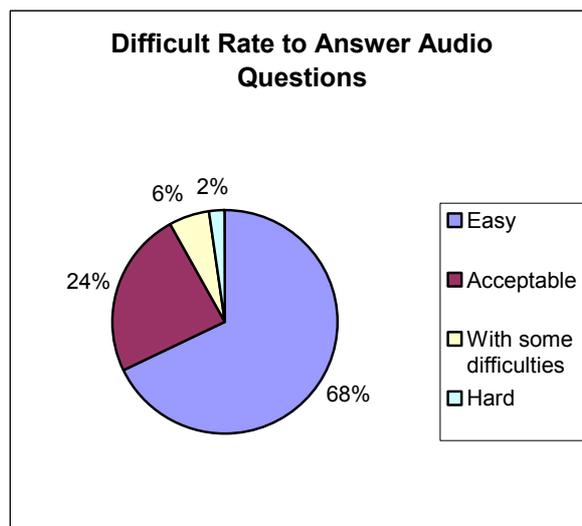


Figure 5-7 How difficult does user feel to answer the questions

Generally speaking, the total delay time caused by our Spam detection software consists of two parts, the time needed to play an audio challenge question (as summarized in Table 5-2) and the time needed for the caller to understand and answer it (as shown in Figure 5-8). From this figure, one notices that 70% of the callers can understand and answer their questions within a second and 93% will not take more than 3 seconds, which indicates that as long as a caller heard his/her given challenge, as a human user, he/she is able to answer it right away.

Another important requirement of a challenge is that it should not be annoying. A question can become annoying because it is too long or not clear enough, or for some other reasons. This is overall how comfortable the caller feel about calling someone who installed our software. Figure 5-9 shows that 52% of the evaluators think it is comfortable to use the software and 46% feels the experience is acceptable. Nobody consider the user experience to be unacceptable.

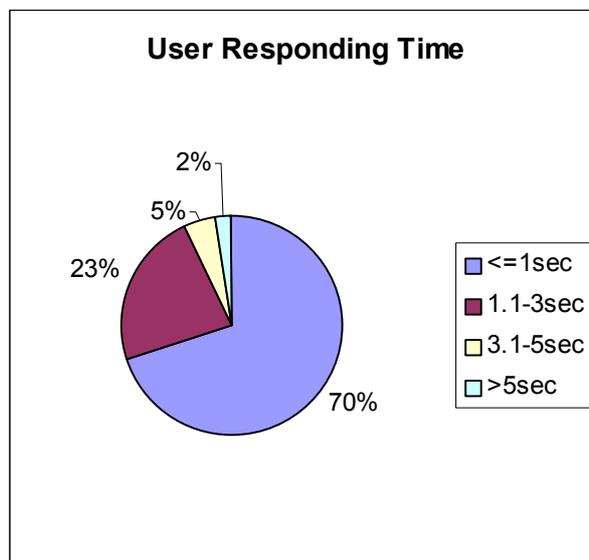


Figure 5-8 User Responding Time for Each Audio Challenge Question

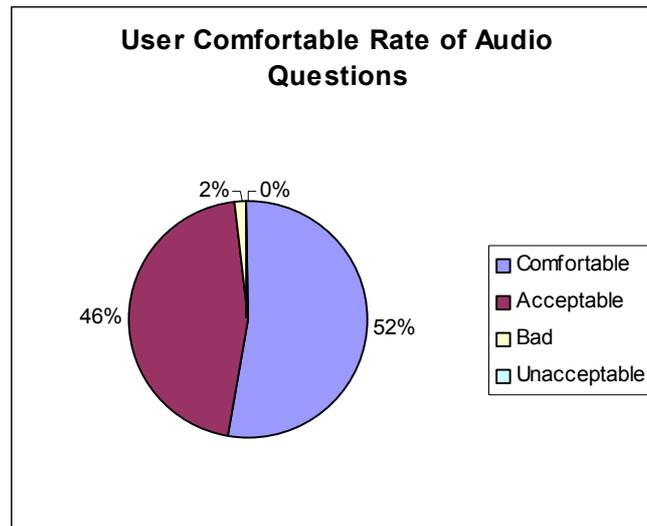


Figure 5-9 How annoyed does user feel about the challenge questions

§ 5.2 Correctness Study

The VoIP Spam detection method in this work is based on the theory of reverse Turing test to differentiate human users from machine bots. The success of this proposed method is measured by the percentage of humans passing the test and that of computer programs failing the challenge – the higher the better in both cases.

The nature of an audio question based reverse Turing test determines that it is always implemented on one particular language. Since the VoIP anti-Spam system in our experiment uses English environment, the human users who speak English as their first language inevitably have certain advantage in taking the test. Beside native speakers of English, the group of volunteer evaluators also contains people from a diverse ethnic background. There are volunteers originally from China, Korea, Greece, India and Turkey, who speak English as their second (or third etc.) language. The reverse Turing test passes percentage from both English as second and as first language groups are

collected and analyzed, with the results and discussion in section 5.2.1 and 5.2.2 respectively. Besides, the fail percentage for machine user taking the challenges in our system is also investigated using an artificial intelligence (A.I) program called SmarterChid.

5.2.1 Correctness Study for English-as-Second-Langrage (ESL) Users

Among our 45 evaluators, over 90% speak English as their second language. Since each test participant was ask to make three to five phone calls, we got 192 sample calls, out of which 118 calls should be categorized as test calls from users who speak English as second language and the rest 74 calls are from their native language.

From the test results we collected, 86 out of the 102 (~86%) English as second language users passed the test challenges. As the caller evaluation, the callers were also asked what they think is the reason for failing each challenge question. There are multiple causes to make those failed test calls happen. Four of the failed calls are due to the call receivers accent in the audio questions they recorded. Although accent might be a powerful leverage to increase the difficulty of Turing test for computer program, a thick accent could also affect the caller's understanding to a question, which caused around 4 percent failure in our experiment. For example, some callers reported in a question the pronunciation of "Z" sounds like "C". The 4 percent failure is an acceptable rate and observed only among non-native speakers. None of the native English speakers found that accent is the cause for being rejected by the anti-Spam tool.

The second reason to fail a test is that the caller had a hard time hearing a question due to the soft voice. Two failed calls were reported for this reason. One solution to this problem is to state in the user instruction that the callee should speak

loudly when recording their challenge questions. Also, a callee could raise the voice column of recorder when recording.

The third reason is the English problem. Five failed calls are reported with this reason for failing. Since some of our testers are new comers to USA and not familiar with English, they still have some problems on simple English understanding. If they are tested in their own language, this should not be a problem.

The fourth reason is mind wandering and missed part of the question. Two failed calls belong to this reason. It is a common human mistake.

The fifth reason is the wrong usage of a word, framed incorrectly. It is another English related problem. The recorder is not from English speaking country and uses the wrong phrase. The question is *how many cycles in one bicycle?* The right one should be *how many wheels are there in one bicycle?* Since human has the fault tolerance, most of the wrong-framed English questions are understandable for users. However, it is hard for machine to know the real meaning with the wrong phrase. This is the only one that has been reported as this reason to fail calls.

The sixth reason for call fail is that there is too much noise in the background. One fail call reports this reason. In our user instruction, there is one recommendation to make some background noise. It may not be a good idea. Some voice recognition machines have this feature to isolate noise and the useful voice; it may just fool some humans. This instruction should be cancelled from our user instruction.

The last reported failing reason is that the answer is not a number. If the user pays a little bit attention, s/he should notice that our phone pad provides the character settings on outlook as it shows on Figure 5-5. If the answer is characters, user only needs to input

the character corresponding number. For example, if the answer is apple. User needs input 27753 and send it. It does not matter the answer is number or characters with the phone pad we provide.

5.2.2 Correctness of humans on Mother language audio question testing

The rate for passing English as the 2nd language audio challenge calls is 86% as we mentioned in section 5.2.1. There is some English understanding problem involved for call failings. When language is not the failing issue, the passing rate is dramatically increased. There are only three calls are failed among 71 Mother language audio question testing. 96% of the calls are set up successfully. The only reason reported on failure call is that user understands the question, but do not know the answer. Three failed calls reported the same reason. We checked the three failed calls questions. Two of the three failure calls are failed on the same question -- *What does number stand for with X in Greek letters?* Actually, the question itself is wrong. The tester who recorded this question wants to ask *what does number stand for with X in Roman letters?* The other failed call is because that the pre-recorded question is in Korean, but answer requires English. It confuses caller. All mother language recorded questions should assume all answers are in the same language.

5.2.3 Analysis of machine passing rate on English challenge questions

Knowing that human users pass the challenges with high percentage rate completes only part of the correctness evaluation for our anti-Spam system. Another important evaluation is to study the machine passing rate for the designed Turing test, which requires computer programs with artificial intelligence. Since A.I. is out of the

scope of the current work, a commercial MSN chatterbot called SmarterChild™ is used as the machine evaluator of our collected challenges.

A chatterbot is a computer artificial intelligent program, which is designed to participate in an intelligent conversation with human by his/her recognizing audio or text input. SmarterChild is available on AOL instant messenger, MSN instant messenger and ICQ networks. Unlike the previous chatter bots that just mimicked human conversation, this chatterbot uses a combination of complex pattern recognition and simple natural language recognition to take user questions and comments and translate them into database queries. Thus, it can converse in a variety of topics ranging from philosophy to the perverse and become one of the ideal evaluators for our VoIP spam detection system.

As discussed in §4.4, we made the Turing challenges harder for machines by putting the question into context. To break our system, an A.I. program should not only be able to use speech recognition to convert a sound clip to the correct words, but also can understand what these words mean. As a result, the challenge questions were brought to machine evaluator in our experiment through text chatting to see how well SmarterChild can understand a question. It is reasonable to assume that if the possibility for a machine to pass a reverse Turing test in this controlled experiment is low, the passing rate in a real world situation should be even lower due to the errors caused by speech recognition.

Three out of the 10 questions designed by each human caller was randomly picked to form a test question pool with 135 questions for the chatterbot. It was observed that SmarterChild failed 133 out of the 135 questions (~98%). The only two questions it can answer correctly are:

1. How many hours in a day?
2. How many hours are there in a day?

which are about common sense and can easily be caught by a computer program.

In addition to evaluate the machine's failing rate for our Turing-like challenges, the chatterbot was also used to find out what is a more difficult question for an A.I program, which continues the discussion in §4.4 about challenge design guidelines. To test the limitation of SmarterChild, we started with the simple math problems. For example, it can answer questions like "what is 234 plus 158?", or harder ones like "what is 33 power 5?" without any problem, as long as it consist a key word, such as add, plus, minus etc. and a bunch numbers.

SmarterChild works based on the recognition of only a key pattern of a question and the rest of it can be varied. For instance, as long as a question contains "2 plus 2", the chatterbot will not be distracted by either it is "What is 2 plus 2?" or "2 plus 2 is what?" or "2 plus 2 equals to?". It can even correct some typos (e.g. it treats "plus" and "puls" the same way).

However, the chatterbot might not be able to recognize a new variation. A straight forward example is that it does not understand a question like "2 plus what is 4?". Also, SmarterChild will not be able to realize a question is about math unless it is asked explicitly with number alone. If a question as "what is 2 apples plus 2 apples?" is asked, the computer program by default is going to think "apples" is important in this sentence, rather than the two numbers. The answer will then be centered on the concept "apple". Moreover, if the question is put in a context or a story, it makes the chatterbot even more

confusing. These experiments indicate that the general guidelines about challenge design in Chapter 4 are effective against current artificial intelligence.

§ 5.3 Performance Study

Our software, except audio file, occupies 3.43MB. Since each user has different length of pre-recorded questions, the disk space for audio file is varied. Table 5-3 shows us the disk space consumed by each tester. The unit of the disk space is megabytes. From Table 5-3, we can get the average disk space as 1.2 MB. Thus the total space required is about 4.63 MB. This is small space consumed software, even requires less disk space than some single MP3 format song.

Table 5-3 disk space consumed by the 10 pre-recorded audio questions

Tel. #	Disk Space								
103	1.21	112	0.8	121	0.905	130	1.99	139	1.22
104	1.18	113	1.2	122	0.769	131	1.17	140	1.06
105	0.468	114	2	123	1.09	132	1.53	141	0.887
106	0.98	115	1.24	124	1.21	133	1.58	142	1.11
107	0.86	116	1.57	125	1.48	134	1.5	143	1.11
108	1.2	117	1.08	126	1.19	135	0.94	144	1.1
109	0.992	118	1.5	127	1.14	136	0.969	145	1.6
110	1.12	119	1.33	128	2.01	137	1	146	0.865
111	1.26	120	0.905	129	1.19	138	1	147	1.07

Chapter 6 Conclusions and Future Work

§ 6.1 Conclusions

In this study, we proved in a VoIP network that it is feasible to send automatic VoIP spam with and without going through a proxy server. In order to mitigate the threats due to VoIP direct user-to-user spamming action, diversified question Turing test based on client end is proposed as an extension to reverse Turing test and also as a good supplementary to the exist Black/White listing scheme for anti-VoIP-spam systems. Such a diversified question Turing test system was set up in our lab and evaluated by 45 evaluators to study its correctness, usability and resource consumption.

The experiments results have proven that we have achieved around 96% passing rate among human users who took the tests in our anti-spam tool using their mother language. Also, about 86% of testers, who speak English as the 2nd language, passed the English audio challenge based VoIP anti-spam filter. These results indicate that the proposed anti-spam solution is effective enough in terms of passing human users through. As for machine programs, out of 135 diversified challenge questions designed by the evaluators, only two (<2%) can be answered correctly by the A.I. chatterbot SmaterChild, indicating that the tests are difficult enough to block automatic machine spammers.

Concerning usability study on user's experience with our anti-spam tool, the time requirement for installation and configuration is easy short. About 75% evaluators used less than 10 minutes to install and configure our tool. According to the reports we got from users, the average disk space consumes about 4.63 MB, which is even less than some single MP3 format song.

§ 6.2 **Limitations and Future Work**

Since the challenges are designed by customers, good side is that it is hard for the spammer to find the challenge pattern since millions of people have millions of different challenges. Bad effect is that it is hard to control the quality of every challenge. We make the suggestions on user's guide, but we cannot guarantee that users follow exactly what we say.

Due to the characteristics of language that it is not global, the test is not universal for global users. When caller takes the challenge of the reverse Turing test, delay is inevitable. Those are the limitations for our work.

For one of our further studies, an A.I. speech recognition program could be used to investigate the effects of noise and accent on the correctness of our system in differentiating human and machine users. Another good future study would be the implementation and evaluation of an anti-spam system including both the diversified question reverse Turing tests and black & white lists.

References:

- [1] http://en.wikipedia.org/wiki/History_of_spamming#_note-0
- [2] <http://www.spamhaus.org/definition.html>
- [3] “SPAM: Get the Facts”, Jan. 2007
- [4] <http://www.tradersnetwork.org/intelligence/content/view/1313/45/>
- [5] Andrew Leung, “Spam, The Current State”, Aug.8, 2003
- [6] <http://www2.rad.com/networks/2001/voip/history.htm>
- [7] http://www.whichvoip.com/voip/articles/voip_history.htm
- [8] <http://www.idc.com/about/about.jsp>
- [9] “Voice Over IP Gains Increasing Market Share”, Apr.26, 2005,
www.infocenter.cramsession.com
- [10] <http://ezinearticles.com/?VoIP-Services-vs.-Traditional-Phone-Services&id=478473>
- [11] http://www.businessweek.com/technology/tech_stats/VoIP060419.htm
- [12] Matt Hines, “Experts Say Applications will Drive VoIP Growth”, Sep.20, 2005,
www.eWeek.com
- [13] http://telephonyonline.com/voip/metrics/voip_revenue_forecast_021705/
- [14] Johanne Torres, “Special Report: VoIP Revenue to Skyrocket to \$20B in 2009”,
May 5, 2005, TMCNET
- [15] Sarrel, Matthew D. “VoIP Sends the Message”, PC Magazine; 10/3/2006, Vol. 25
Issue 17, p100-100, 1p, 1 diagram
- [16] J. Rosenberg, C. Jennings, J. Peterson, “The Session Initiation Protocol (SIP) and
Spam”, draft-ietf-sipping-spam-02, March 6, 2006.
- [17] http://www.voipaction.com/voice_spam.php

- [18] Jim Bartbold, "SPITting Mad: Qovia Stopping Voice Mail Spam in its Tracks", Jun, 2004, pp.10
- [19] Paul Graham, "A plan for spam", Aug. 2002
- [20] A. Dornan, "Bayesian Spam Filtering", IEEE Netowrk, no. 3, vol. 19, 2004, pp. 64-65.
- [21] http://en.wikipedia.org/wiki/Bayesian_spam_filtering
- [22] Markus Hansen, Marit Hansen, Jan Moller, Thomas Rohwer, Carsten Tolkmit, Henning Waack, "Developing a Legally Compliant Reachability Management System as a Countermeasure against SPIT", 3rd Annual VoIP Security Workshop
- [23] Andrew Conry-Murray, "Fighting the spam monster - And winning", Network Magazine, v 18, n 4, April, 2003, p 24-29
- [24] Bertrand Mathieu, Quentin Loudier, Yvon Gourhant, François Bougant, Mathieu Osty, "SPIT Mitigation by a Network-Level Anti-Spit Entity", Third Annual VoIP Security Workshop, Jun. 01, 2006.
- [25] J. Rosenberg, C. Jennings, J. Peterson, "The Session Initiation Protocol (SIP) and Spam", draft-ietf-sipping-spam-02, March 6, 2006.
- [26] Paul Resnick, Richard Zeckhauser, Eric Friedman, and Ko Kuwabara, "Reputation Systems", Communications of the ACM, 43(12), December 2000, pages 45-48.
- [27] Ram Dantu and Prakash Kolan, "Preventing Voice Spamming", VoIP Security Challenges and Solution, IEEE BLOBECOM 2004, Dallas, TX, Dec. 2004
- [28] Ram Dantu and Prakash Kolan, "Detecting Spam in VoIP Networks", USENIX SRUTI 2005 Wksp., Cambridge MA, July 2005, pp. 31-37

- [29] Dorgham Sisalem and Yacine Rebahi, “SIP Service Provides and the Spam Problem”, 2nd Wksp., Securing Voice over IP, Washington DC, June 1-2, 2005.
- [30] Dongwook Shin, Jinyong Ahn, Choon Shim, “Progressive multi gray leveling: a voice spam protection algorithm”, Network IEEE, Sept.-Oct. 2006, Volume: 20, Issue: 5, page 18-24.
- [31] Turing AM (1950), “Computing Machinery and Intelligence,” Mind, 59:236, pp.433-460
- [32] Luis von Ahn, Manuel Blum, and John Langford, “Telling Humans and Computer Apart (Automatically) or How Lazy Cryptographers do AI”, to appear in *Communications of the ACM*
- [33] Luis von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford, “CAPTCHA: Using Hard AI Problems For Security”, in Proceedings of Eurocrypt’03 International Conference on the Theory and Applications of Cryptographic Techniques, LNCS 2656, pp. 294-311, Springer-Verlag, Berlin Heidelberg, 2003.
- [34] Mikhail M. Bongard. (1970). Pattern Recognition. Rochelle Park, N.J.: Hayden Book Co., Spartan Books.
- [35] <http://chaptcha.net>
- [36] Nancy Chan. Program Byan: <http://drive.to/research>
- [37] Y. Rui, Z. Liu, “ARTiFACIAL: Automated Reverse Turing test using FACIAL features”, International Multimedia Conference Proceedings of the eleventh ACM international conference on Multimedia 2003, Berkeley, CA, USA, 2003, pp 295 – 298.
- [38] Tim Converse, “Captcha Generation as a Web Service”
- [39] <http://www.nec.co.jp/press/en/0701/2602.html>

- [40] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler, "SIP: Session Initiation Protocol", June 2002
- [41] Qi Qiu, "Study of Digest Authentication for Session Initiation Protocol", Dec. 2003
- [42] Samir Chatterjee, "SIP, Security & Threat Models", SURA/ViDe 4th Annual Workshop, Apr. 2002
- [43] Bill Clagett, <http://www.32768.com/bill/java/portscanner/index.html>
- [44] Baird HS, Popat K (2002), "Human interactive proofs and document image analysis", Proceedings of Document Analysis Systems 2002, Princeton, NJ, August 2002, pp 507-518
- [45] Luca Veltri, MjSip-Mini-Tutorial, Apr.2005

Appendix

Appendix A User Instruction Guide

Our tool is an anti-VoIP(Voice Over Internet Protocol) Spam softphone, which means users can make phone call by Internet. VoIP Spam is similar as e-mail Spam in that senders use an IP network to send unsolicited abundant of calls for marketing purposes. The purpose of our tool is to distinguish humans from machines and block the machine calls. The main idea to make it success is to pre-record some audio challenge questions that human can answer but machine cannot pass. This user instruction will guide you to install and use our tool step by step. There are three parts to compose this user guide --pre-requirements, tool installation and user configuration.

Software Pre-requirements:

Windows 2000/XP

JRE (Java runtime environment)

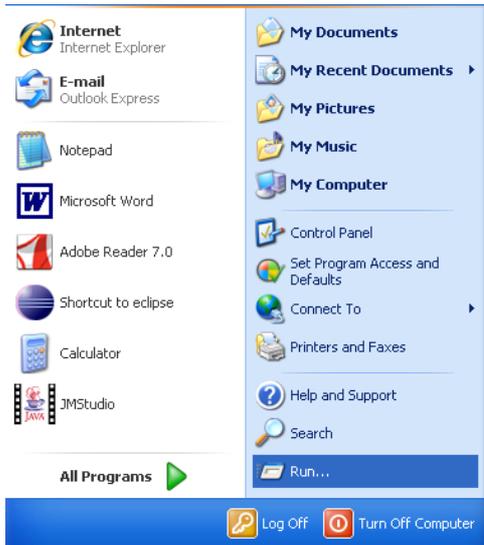
JMF (Java media framework API)

I. JRE

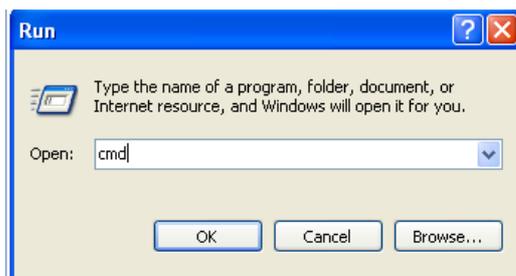
Part A: Check JRE installation

JRE is Java runtime environment. You may check as the following:

- 1) Open command line from start window. Click on run.

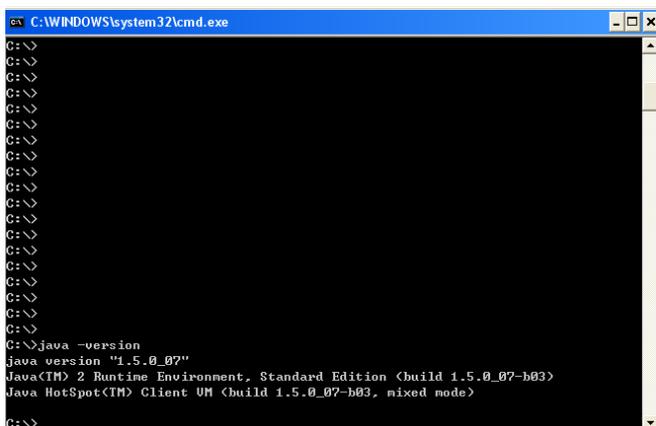


2) Type **cmd** in *run* window and click **OK**



3) Type **java -version** in a Command prompt (Windows) or Terminal console (Unix). If you receive an “Unknown Command” error message then you need to install the latest edition of JRE, which is free and can be downloaded from

<http://java.sun.com/javase/downloads/index.jsp> .



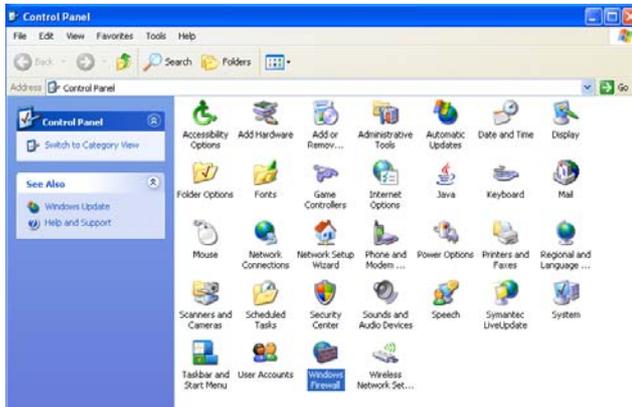
After installation of JRE, if your firewall is on, make sure your firewall doesn't block JRE. You need check it.

Part B: Instruction of checking if JRE has the permission to pass firewall.

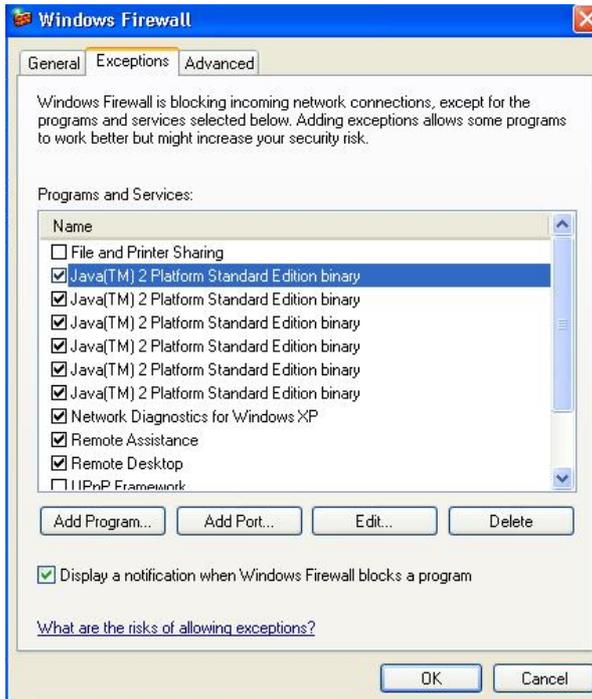
1) From start window, click on *Control Panel*.



2) Double click on *Firewall*.



3) Choose *Exceptions*. If JRE is in the exceptions category, it will show as Java TM 2 Platform Standard Edition binary.



4) Double click on Java TM 2 Platform Standard Edition binary; you should see

<dir> \javaw.exe as the bellowing



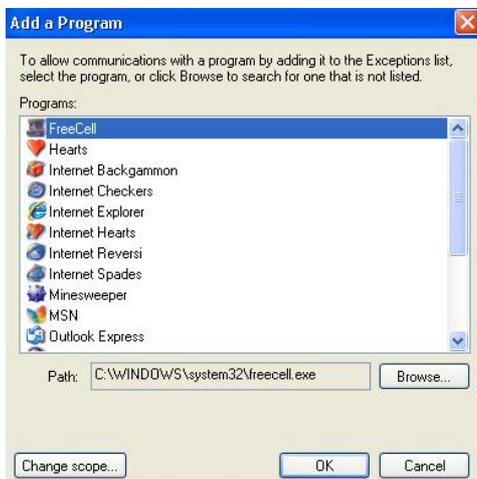
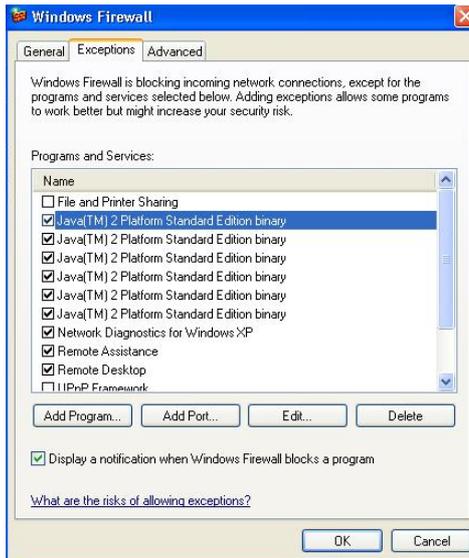
Or <dir> \java.exe as the bellowing



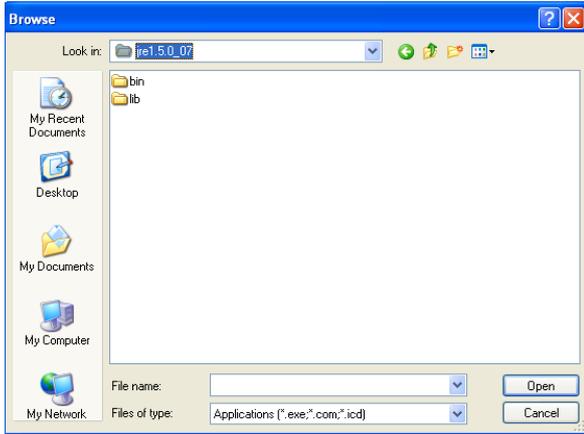
Part C: Instructions of unblocking java or javaw on firewall

If you don't have <dir>\javaw.exe and <dir>\java.exe, you need add both. The instruction of how to do it is showing bellow

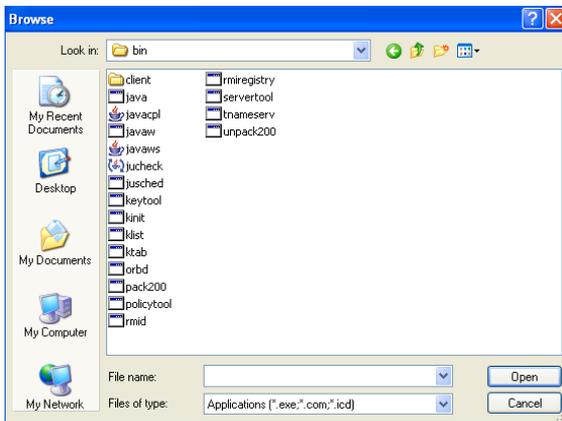
- 1) After you click on Exception of firewall, double click on **Add Program**.



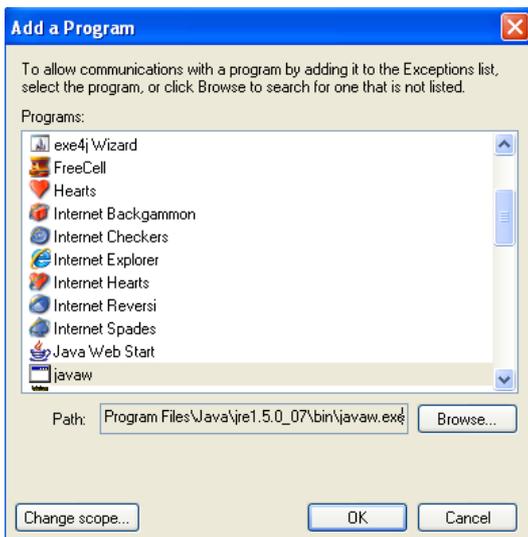
- 2) Double click on **Browse**; go to the folder where you downloaded JRE and double click on your downloaded folder. There are two subfolders – bin and lib.



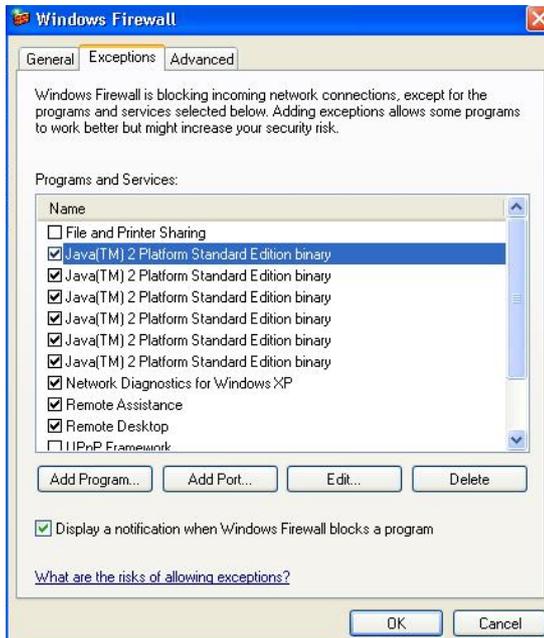
3) Double click on *bin*.



4) Click on *java* (or *javaw*) and click *Open*.



5) Click on **OK**.



6) Click on **OK** again. JRE is unblocked in firewall.

II JMF

JMF is java media frame API, which can be used to enable audio, video and other time-based media to be added to Java applications and applets. It can be downloaded for free at <http://java.sun.com/products/java-media/jmf/2.1.1/download.html> .

Tool Installation

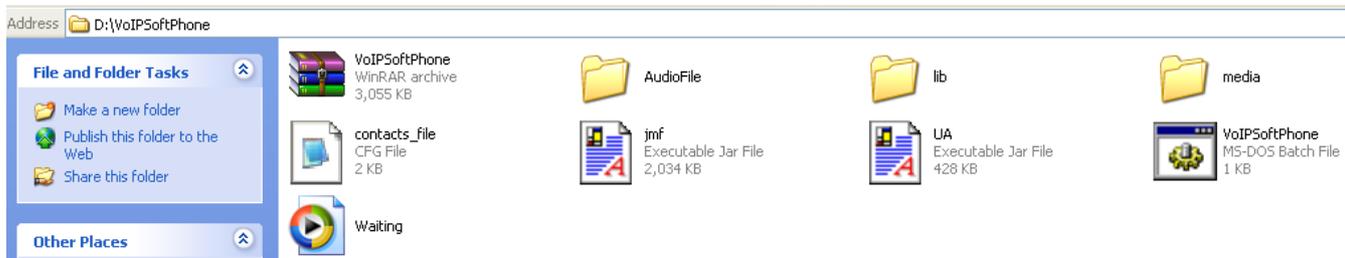
After all the pre-requirements are done, it's ready to install our anti-VoIP spam tool.

The instruction of downloading is described below:

1) Download VoIPSoftPhone in a folder



- 2) Unzip it and you'll get eight files. They are Audio; lib; media; contacts_file.cfg; jmf.jar; UA.jar; VoIPSoftPhone.bat; Waiting.wav.



User Configuration

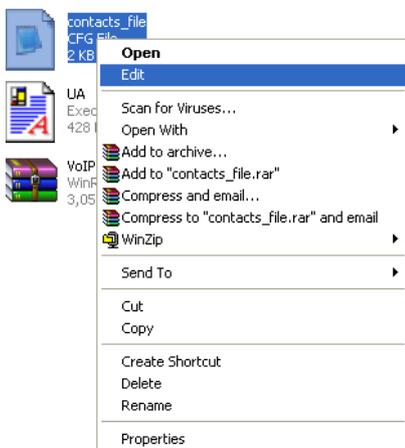
Before starting to run our tool, we need do two configurations. The first is to configure your machine's IP address. The second is to record ten audio questions. The instructions are described below.

I Configure IP Address

- 1) Open the configuration file – contacts_file.cfg with Notepad.



Right click on contacts_file.cfg; click **Edit**.



2) Fill up `contact_url` with your own IP address.

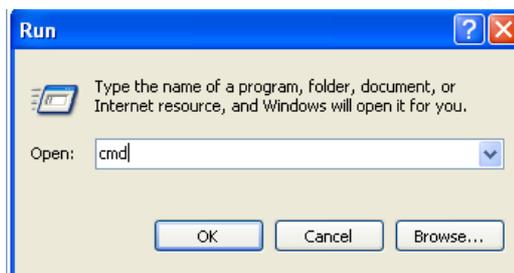
`contact_url =`

Here is the instruction of how to check your own IP address.

Step 1: Open command line pomp from  *run* as



Setp2: Type `cmd` in *run* window and click *OK*

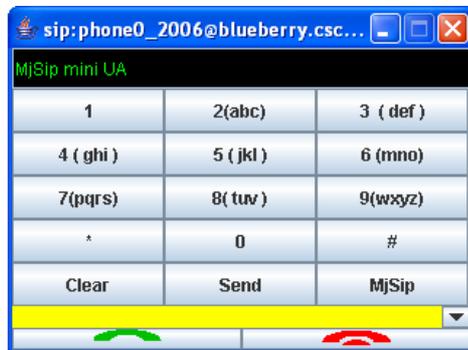


Format of beginning: When user begins to record, user needs to mention sender to press send to transfer the answer back to receiver. For example, please answer my question. When you finish, please press send to send the answer. This is the format of the beginning of the answer.

Format of body: Body of each vocal question is the main part of our challenges. The purpose of question body design is to make questions humanities and hard for machine to understand. The following are some advices to make challenge human understandable but machine cannot pass.

- Recording question should be less than 30 seconds. If it's too long, caller will feel it's time wasting. Comfortable using degree will be degraded.
- Recording question should be computationally easy. The advantage of machines is the strong power of computation. Humans have the limitation on mathematics. Any calculation question needs to be simple and easy to be answered by human in a very short time as less than 30 seconds.
- Recording question should be fuzzy logic question. It's easy for the machine to follow the pattern of *plus* or *minus*. If a scene is added, machine can be easily fooled. For example, there are five clouds in the sky. Wind blows away two. How many clouds left. Another example, I have 12 cats and 999 roses. How many flowers do I have?
- Based on the characteristics that our phone pad has, short letter answers are also accepted besides numbers. However, only short letter answers are accepted, not more than 10 letters. By example, the question can be *what's my last name?* Try

to avoid asking your first name when you have one official first name and one so-called name.



- Multiple-choice questions should be avoided. This kind of question can easily be attacked by brute force attack, since multiple-choice question lists can't be long; otherwise, it's hard to be remembered by caller.

Part B: Audio Formats & Recording Guideline with JMStudio

After user got their challenge question design, the next step is to record the ten audio questions. The audio question should be saved as .wav format with sample rate as 8000Hz; encoding as ULAW; bits per sample as 8 bit; channels as mono. Since JMStudio is a stand-alone Java application, which can play, capture, transcode, and write media data by means of JMF (Java Media Framework API), user can use JMStudio to record the challenges. All the recorded questions must be saved in AudioFile folder

 , and name each audio question as Q1.wav; Q2.wav;...; Q10.wav. Be careful, the name Q must be capitalized and the wav must be small letters. We can use JMStudio to record the audio questions. After the installation of JMF, your machine has automatically installed JMStudio.

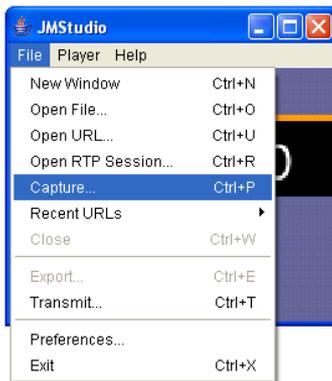
After downloading JMF, if it's stored in program files, you can open JMStudio as the following.



1) Double click on *JMStudio*



2) Select *File* and *Capture*.

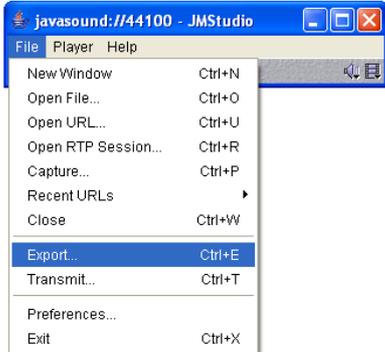


3) Select Capture Device as the picture below. Check *Use Audio Device* and click *OK*.

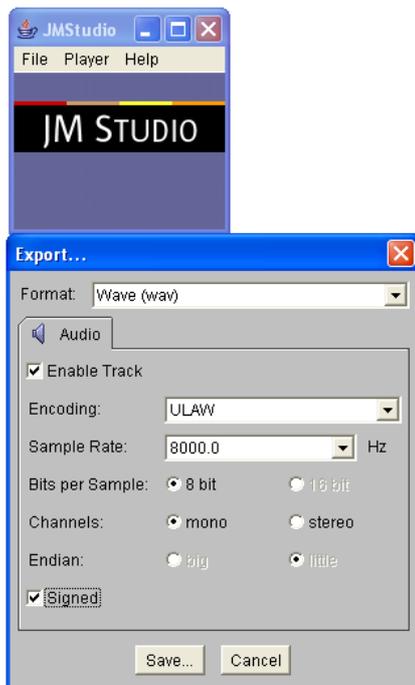


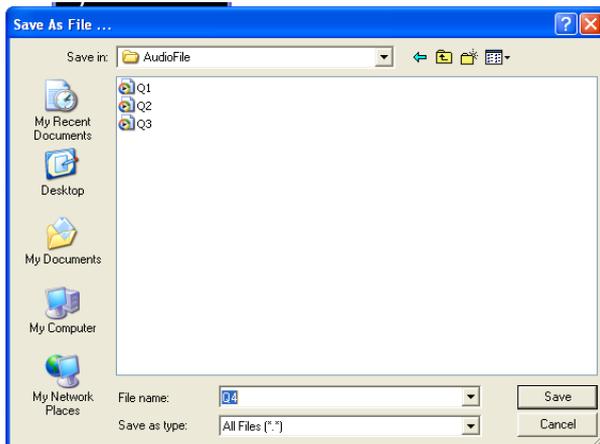


4) Click **File** and choose **Export**.

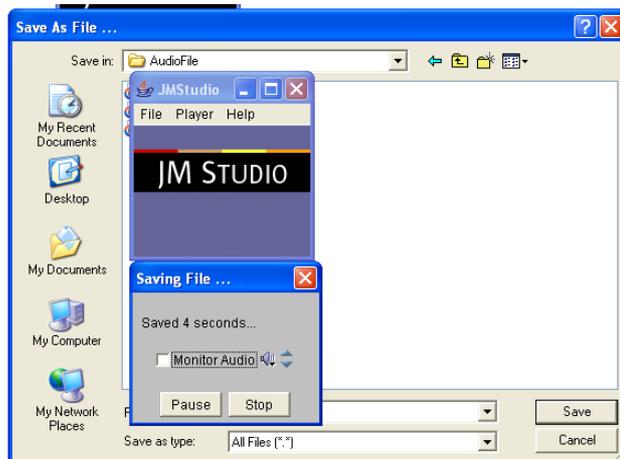


5) Choose the elements as the following; choose **format** as *wave*; check **enable track**; **encoding** as *ULAW*; **sample rate** as *8000Hz*; **bits per sample** as *8 bit*; **channels** as *mono*; **Save** the audio file in MyIPPhone/AudioFile.  Be careful, the name of each recorded challenge must be Q1.wav; Q2.wav;...;Q10.wav. Q must be capitalized and wav must be small letters.





6) You may record your questions now. When you click **stop**, the audio you recorded is saved in the MyIPPhone/AudioFile, which was the file you selected to store the audio file in step 5.



After the ten audio questions have been stored in AudioFile folder , the last configuration is to input the answers of the ten questions on contacts_file.cfg



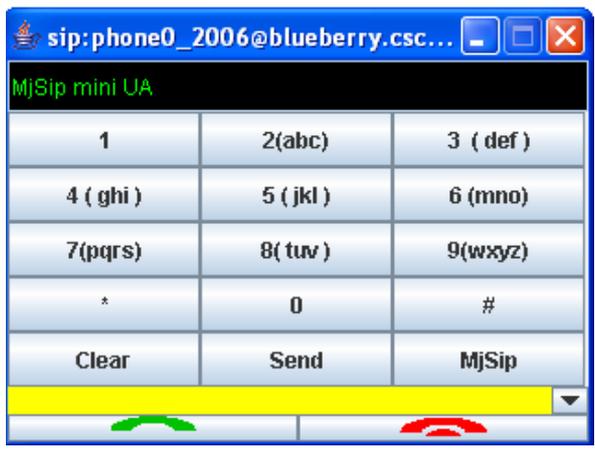
. Open contacts_file.cfg with notepad. Fill up each challenge's corresponding answer.

```
# Answer of requested questions. (Only numbers here)
Q1=
Q2=
Q3=
Q4=
Q5=
Q6=
Q7=
Q8=
Q9=
Q10=
```

Only numbers are allowed. If your answer is letters, put the corresponding numbers on it.

When ten answers have been filled up, save this new contacts_file.cfg. It's ready

to run our tool. Double click on VoIPSoftPhone.bat  . Softphone pad is shown.



Descriptions of some phone pad keys

When user wants to make a phone call, he/she needs to dial the number first. The number will be shown on  and then punch . If user wants to hang up, press . Here are more descriptions of the above phone pad keys.

: Display screen of dialing number or the answers you are sending to the receiver.

: Restart number punching. When user dials the wrong number, s/he can punch this key to start over and re-enter the number.

: Send the answer to the corresponding audio question.

: Show *MjSip mini UA* on display screen. Our anti-VoIP spam softphone is built up on MjSip, which is a VoIP free softphone.

: Dialing pat key.

: Hang up pat key.

Appendix B Questionnaires

Questionnaire A -- Receiver Version

Receiver is the user who needs install, configure and set up the SIP soft-phone and record the questions in his/her SIP soft-phone. This is the questionnaire that a receiver needs to fill up.

Name: _____ Major: _____
Degree: BS / BA / MS / MA / Ph.D
Background: Student / Employed / Non-Employed
First Language: _____ Second Language: _____
Age: _____ Gender: Male/Female

Questionnaires:

1. How long does it take you to install and configure the tool?
A. 0-5 min B. 5-10 min C. 10-15 min D. More than 15 min
2. How long does it take you to design the ten challenge questions?
A. <10 min B. 10-20 min C. 20-30 min D. > 30 min and used about __ min
3. How long does it take you to record the ten challenge questions?
A. <10 min B. 10.1-15 min C. 15.1-20 min D. > 20 min and used about __ min
4. What platform do you use?
A. Windows 98 B. Windows 2000 C. Windows XP D. Unix/Linux
5. Is this user instruction readable?
A. Easy to follow B. Understandable C. With some difficulty D. Hard
6. Is the question design guideline readable?
A. Easy to follow B. Understandable C. With some difficulty D. Hard
7. How hard is it to install and configure this tool?
A. Easy B. Not bad C. A little bit hard D. Very hard
8. How much is the disk space needed for the ten-recorded questions?

Questionnaire B -- Caller Version

Caller is the one who makes the call. He/she needs to answer the recorded questions to verify him/her is not VoIP spammer. Receiver is the one who is called.

Name:

Caller's phone #:

Receiver's phone #:

1. How much is the delay from when you heard the challenge until when you hear the other party's phone rings or rejected tone?
A. <1 sec B. 1.1-3 sec C. 3.1-5 sec D. > 5 sec

2. Do you pass the audio challenge question?

A. Pass B. Fail

If you failed on this number, what's the reason?

A. Don't understand the question due to the accent.

B. Have hard time to hear the question due to the soft voice.

C. Understand the question but don't know the answer

D. If the above reason doesn't apply to you, what's the reason

-
3. How annoyed do you feel about the audio challenge question?

A. Comfortable B. Acceptable C. Bad D. Unacceptable

4. How difficult do you feel to understand the questions?

A. Easy B. Understandable C. With some difficulties D. Hard

5. How difficult do you feel to answer the questions?

A. Easy B. Acceptable C. With some difficulties D. Hard

6. What language do you hear?

A. English B. Chinese C. Korean D. Hindi

Appendix CAsterisk PBX Configuration

SIP Configuration

```
[general]
context=default ; Default context for incoming calls
realm=BLUEBERRY.csc.ncsu.edu ; Realm for digest authentication
; defaults to "asterisk"
; Set this to your host name or domain name
bindport=5060 ; UDP Port to bind to (SIP standard port is 5060)
bindaddr=0.0.0.0 ; IP address to bind to (0.0.0.0 binds to all)
srvlookup=yes ; Enable DNS SRV lookups on outbound calls
; Note: Asterisk only uses the first host
; in SRV records
; Disabling DNS SRV lookups disables the
; ability to place SIP calls based on domain
; names to some other SIP users on the Internet

domain=BLUEBERRY ; Set default domain for this host
; If configured, Asterisk will only allow
; INVITE and REFER to non-local domains
; Use "sip show domains" to list local domains
;domain=152.0.0.0/8 ; Add IP address as local domain
; You can have several "domain" settings

disallow=all ; First disallow all codecs
allow=ulaw ; Allow codecs in order of preference
allow=ilbc
allow=gsm
allow=pcm
;allow-wav

musicclass=random ; Sets the default music on hold class for all SIP
calls

language=en ; This may also be set for individual users/peers
; Default language setting for all users/peers
; This may also be set for individual users/peers

registerattempts = 0
registertimeout = 20

[blueberry]
type=peer
host=152.14.92.31
canreinvite=yes
qualify=no
```

port=6001
nat=no

[phone0_2006]
type=friend
secret=testvoip000
context=from-sip
host=dynamic
canreinvite=yes
username=phone0_2006
nat = no

[phone1_2006]
type=friend
secret=testvoip001
context=from-sip
host=dynamic
canreinvite=yes
username=phone1_2006
nat=no

[phone2_2006]
type=friend
secret=testvoip002
context=from-sip
host=dynamic
canreinvite=yes
username=phone2_2006

[phone3_2006]
type=friend
secret=testvoip003
context=from-sip
host=dynamic
canreinvite=yes
username=phone3_2006

[phone4_2006]
type=friend
secret=testvoip004
context=from-sip
host=dynamic
canreinvite=yes
username=phone4_2006

[phone5_2006]
type=friend
secret=testvoip005
context=from-sip
host=dynamic
canreinvite=yes
username=phone5_2006

[phone6_2006]
type=friend
secret=testvoip006
context=from-sip
host=dynamic
canreinvite=yes
username=phone6_2006

[phone7_2006]
type=friend
secret=testvoip007
context=from-sip
host=dynamic
canreinvite=yes
username=phone7_2006

[phone8_2006]
type=friend
secret=testvoip008
context=from-sip
host=dynamic
canreinvite=yes
username=phone8_2006

[phone9_2006]
type=friend
secret=testvoip009
context=from-sip
host=dynamic
canreinvite=yes
username=phone9_2006

[phone10_2006]
type=friend
secret=testvoip010
context=from-sip
host=dynamic
canreinvite=yes

username=phone10_2006

[phone11_2006]
type=friend
secret=testvoip011
context=from-sip
host=dynamic
canreinvite=yes
username=phone11_2006

[phone12_2006]
type=friend
secret=testvoip012
context=from-sip
host=dynamic
canreinvite=yes
username=phone12_2006

[phone13_2006]
type=friend
secret=testvoip013
context=from-sip
host=dynamic
canreinvite=yes
username=phone13_2006

[phone14_2006]
type=friend
secret=testvoip014
context=from-sip
host=dynamic
canreinvite=yes
username=phone14_2006

[phone15_2006]
type=friend
secret=testvoip015
context=from-sip
host=dynamic
canreinvite=yes
username=phone15_2006

[phone16_2006]
type=friend
secret=testvoip016
context=from-sip

host=dynamic
canreinvite=yes
username=phone16_2006

[phone17_2006]
type=friend
secret=testvoip017
context=from-sip
host=dynamic
canreinvite=yes
username=phone17_2006

[phone18_2006]
type=friend
secret=testvoip018
context=from-sip
host=dynamic
canreinvite=yes
username=phone18_2006

[phone19_2006]
type=friend
secret=testvoip019
context=from-sip
host=dynamic
canreinvite=yes
username=phone19_2006

[phone20_2006]
type=friend
secret=testvoip020
context=from-sip
host=dynamic
canreinvite=yes
username=phone20_2006

[phone21_2006]
type=friend
secret=testvoip0021
context=from-sip
host=dynamic
canreinvite=yes
username=phone21_2006

[phone22_2006]
type=friend

secret=testvoip022
context=from-sip
host=dynamic
canreinvite=yes
username=phone22_2006

[phone23_2006]
type=friend
secret=testvoip023
context=from-sip
host=dynamic
canreinvite=yes
username=phone23_2006

[phone24_2006]
type=friend
secret=testvoip024
context=from-sip
host=dynamic
canreinvite=yes
username=phone24_2006

[phone25_2006]
type=friend
secret=testvoip025
context=from-sip
host=dynamic
canreinvite=yes
username=phone25_2006

[phone26_2006]
type=friend
secret=testvoip026
context=from-sip
host=dynamic
canreinvite=yes
username=phone26_2006

[phone27_2006]
type=friend
secret=testvoip027
context=from-sip
host=dynamic
canreinvite=yes
username=phone27_2006

[phone28_2006]
type=friend
secret=testvoip028
context=from-sip
host=dynamic
canreinvite=yes
username=phone28_2006

[phone29_2006]
type=friend
secret=testvoip029
context=from-sip
host=dynamic
canreinvite=yes
username=phone29_2006

[phone30_2006]
type=friend
secret=testvoip030
context=from-sip
host=dynamic
canreinvite=yes
username=phone30_2006

[phone31_2006]
type=friend
secret=testvoip031
context=from-sip
host=dynamic
canreinvite=yes
username=phone31_2006

[phone32_2006]
type=friend
secret=testvoip032
context=from-sip
host=dynamic
canreinvite=yes
username=phone32_2006

[phon33_2006]
type=friend
secret=testvoip033
context=from-sip
host=dynamic
canreinvite=yes

username=phone33_2006

[phone34_2006]
type=friend
secret=testvoip034
context=from-sip
host=dynamic
canreinvite=yes
username=phone34_2006

[phone35_2006]
type=friend
secret=testvoip035
context=from-sip
host=dynamic
canreinvite=yes
username=phone35_2006

[phone36_2006]
type=friend
secret=testvoip036
context=from-sip
host=dynamic
canreinvite=yes
username=phone36_2006

[phone37_2006]
type=friend
secret=testvoip037
context=from-sip
host=dynamic
canreinvite=yes
username=phone37_2006

[phone38_2006]
type=friend
secret=testvoip038
context=from-sip
host=dynamic
canreinvite=yes
username=phone38_2006

[phone39_2006]
type=friend
secret=testvoip039
context=from-sip

host=dynamic
canreinvite=yes
username=phone39_2006

[phone40_2006]
type=friend
secret=testvoip040
context=from-sip
host=dynamic
canreinvite=yes
username=phone40_2006

[phone41_2006]
type=friend
secret=testvoip041
context=from-sip
host=dynamic
canreinvite=yes
username=phone41_2006

[phone42_2006]
type=friend
secret=testvoip042
context=from-sip
host=dynamic
canreinvite=yes
username=phone42_2006

[phone43_2006]
type=friend
secret=testvoip043
context=from-sip
host=dynamic
canreinvite=yes
username=phone43_2006

[phone44_2006]
type=friend
secret=testvoip044
context=from-sip
host=dynamic
canreinvite=yes
username=phone44_2006

[phone45_2006]
type=friend
secret=testvoip045
context=from-sip
host=dynamic
canreinvite=yes
username=phone45_2006

[phone46_2006]
type=friend
secret=testvoip046
context=from-sip
host=dynamic
canreinvite=yes
username=phone46_2006

[phone47_2006]
type=friend
secret=testvoip047
context=from-sip
host=dynamic
canreinvite=yes
username=phone47_2006

[phone48_2006]
type=friend
secret=testvoip048
context=from-sip
host=dynamic
canreinvite=yes
username=phone48_2006

[phon49_2006]
type=friend
secret=testvoip049
context=from-sip
host=dynamic
canreinvite=yes
username=phone49_2006

[phone50_2006]
type=friend
secret=testvoip050
context=from-sip
host=dynamic

```
canreinvite=yes
username=phone50_2006
```

```
[phone51_2006]
type=friend
secret=testvoip051
context=from-sip
host=dynamic
canreinvite=yes
username=phone51_2006
```

```
[phone52_2006]
type=friend
secret=testvoip052
context=from-sip
host=dynamic
canreinvite=yes
username=phone52_2006
```

Dial Plan

```
[general]
static=yes
writeprotect=no
autofallthrough=yes
clearglobalvars=no
priorityjumping=no
```

```
[globals]
CONSOLE=Console/dsp           ; Console interface for demo
IAXINFO=guest                 ; IAXtel username/password
TRUNK=Zap/g2                  ; Trunk interface
TRUNKMSD=1                    ; MSD digits to strip (usually 1 or 0)
```

```
[from-sip]
exten => s,1,Answer()
exten => s,2,Background(enter-ext-of-person)
exten => 100,1,Dial(SIP/phone0_2006,60)
exten => 100,2,Playback(vm-nobodyavail)
exten => 100,3,Hangup()
exten => 101,1,Dial(SIP/phone1_2006,60)
exten => 101,2,Playback(vm-nobodyavail)
exten => 101,3,Hangup()
exten => 102,1,Dial(SIP/phone2_2006,60)
exten => 102,2,Playback(vm-nobodyavail)
```

exten => 102,3,Hangup()
exten => 103,1,Dial(SIP/phone3_2006,60)
exten => 103,2,Playback(vm-nobodyavail)
exten => 103,3,Hangup()
exten => 104,1,Dial(SIP/phone4_2006,60)
exten => 104,2,Playback(vm-nobodyavail)
exten => 104,3,Hangup()
exten => 105,1,Dial(SIP/phone5_2006,60)
exten => 105,2,Playback(vm-nobodyavail)
exten => 105,3,Hangup()
exten => 106,1,Dial(SIP/phone6_2006,60)
exten => 106,2,Playback(vm-nobodyavail)
exten => 106,3,Hangup()
exten => 107,1,Dial(SIP/phone7_2006,60)
exten => 107,2,Playback(vm-nobodyavail)
exten => 107,3,Hangup()
exten => 108,1,Dial(SIP/phone8_2006,60)
exten => 108,2,Playback(vm-nobodyavail)
exten => 108,3,Hangup()
exten => 109,1,Dial(SIP/phone9_2006,60)
exten => 109,2,Playback(vm-nobodyavail)
exten => 109,3,Hangup()
exten => 110,1,Dial(SIP/phone10_2006,60)
exten => 110,2,Playback(vm-nobodyavail)
exten => 110,3,Hangup()
exten => 111,1,Dial(SIP/phone11_2006,60)
exten => 111,2,Playback(vm-nobodyavail)
exten => 111,3,Hangup()
exten => 112,1,Dial(SIP/phone12_2006,60)
exten => 112,2,Playback(vm-nobodyavail)
exten => 112,3,Hangup()
exten => 113,1,Dial(SIP/phone13_2006,60)
exten => 113,2,Playback(vm-nobodyavail)
exten => 113,3,Hangup()
exten => 114,1,Dial(SIP/phone14_2006,60)
exten => 114,2,Playback(vm-nobodyavail)
exten => 114,3,Hangup()
exten => 115,1,Dial(SIP/phone15_2006,60)
exten => 115,2,Playback(vm-nobodyavail)
exten => 115,3,Hangup()
exten => 116,1,Dial(SIP/phone16_2006,60)
exten => 116,2,Playback(vm-nobodyavail)
exten => 116,3,Hangup()
exten => 117,1,Dial(SIP/phone17_2006,60)
exten => 117,2,Playback(vm-nobodyavail)
exten => 117,3,Hangup()

exten => 118,1,Dial(SIP/phone18_2006,60)
exten => 118,2,Playback(vm-nobodyavail)
exten => 118,3,Hangup()
exten => 119,1,Dial(SIP/phone19_2006,60)
exten => 119,2,Playback(vm-nobodyavail)
exten => 119,3,Hangup()
exten => 120,1,Dial(SIP/phone20_2006,60)
exten => 120,2,Playback(vm-nobodyavail)
exten => 120,3,Hangup()
exten => 121,1,Dial(SIP/phone21_2006,60)
exten => 121,2,Playback(vm-nobodyavail)
exten => 121,3,Hangup()
exten => 122,1,Dial(SIP/phone22_2006,60)
exten => 122,2,Playback(vm-nobodyavail)
exten => 122,3,Hangup()
exten => 123,1,Dial(SIP/phone23_2006,60)
exten => 123,2,Playback(vm-nobodyavail)
exten => 123,3,Hangup()
exten => 124,1,Dial(SIP/phone24_2006,60)
exten => 124,2,Playback(vm-nobodyavail)
exten => 124,3,Hangup()
exten => 125,1,Dial(SIP/phone25_2006,60)
exten => 125,2,Playback(vm-nobodyavail)
exten => 125,3,Hangup()
exten => 126,1,Dial(SIP/phone26_2006,60)
exten => 126,2,Playback(vm-nobodyavail)
exten => 126,3,Hangup()
exten => 127,1,Dial(SIP/phone27_2006,60)
exten => 127,2,Playback(vm-nobodyavail)
exten => 127,3,Hangup()
exten => 128,1,Dial(SIP/phone28_2006,60)
exten => 128,2,Playback(vm-nobodyavail)
exten => 128,3,Hangup()
exten => 129,1,Dial(SIP/phone29_2006,60)
exten => 129,2,Playback(vm-nobodyavail)
exten => 129,3,Hangup()
exten => 130,1,Dial(SIP/phone30_2006,60)
exten => 130,2,Playback(vm-nobodyavail)
exten => 130,3,Hangup()
exten => 131,1,Dial(SIP/phone31_2006,60)
exten => 131,2,Playback(vm-nobodyavail)
exten => 131,3,Hangup()
exten => 132,1,Dial(SIP/phone32_2006,60)
exten => 132,2,Playback(vm-nobodyavail)
exten => 132,3,Hangup()
exten => 133,1,Dial(SIP/phone33_2006,60)

exten => 133,2,Playback(vm-nobodyavail)
exten => 133,3,Hangup()
exten => 134,1,Dial(SIP/phone34_2006,60)
exten => 134,2,Playback(vm-nobodyavail)
exten => 134,3,Hangup()
exten => 135,1,Dial(SIP/phone35_2006,60)
exten => 135,2,Playback(vm-nobodyavail)
exten => 135,3,Hangup()
exten => 136,1,Dial(SIP/phone36_2006,60)
exten => 136,2,Playback(vm-nobodyavail)
exten => 136,3,Hangup()
exten => 137,1,Dial(SIP/phone37_2006,60)
exten => 137,2,Playback(vm-nobodyavail)
exten => 137,3,Hangup()
exten => 138,1,Dial(SIP/phone38_2006,60)
exten => 138,2,Playback(vm-nobodyavail)
exten => 138,3,Hangup()
exten => 139,1,Dial(SIP/phone39_2006,60)
exten => 139,2,Playback(vm-nobodyavail)
exten => 139,3,Hangup()
exten => 140,1,Dial(SIP/phone40_2006,60)
exten => 140,2,Playback(vm-nobodyavail)
exten => 140,3,Hangup()
exten => 141,1,Dial(SIP/phone41_2006,60)
exten => 141,2,Playback(vm-nobodyavail)
exten => 141,3,Hangup()
exten => 142,1,Dial(SIP/phone42_2006,60)
exten => 142,2,Playback(vm-nobodyavail)
exten => 142,3,Hangup()
exten => 143,1,Dial(SIP/phone43_2006,60)
exten => 143,2,Playback(vm-nobodyavail)
exten => 143,3,Hangup()
exten => 144,1,Dial(SIP/phone44_2006,60)
exten => 144,2,Playback(vm-nobodyavail)
exten => 144,3,Hangup()
exten => 145,1,Dial(SIP/phone45_2006,60)
exten => 145,2,Playback(vm-nobodyavail)
exten => 145,3,Hangup()
exten => 146,1,Dial(SIP/phone46_2006,60)
exten => 146,2,Playback(vm-nobodyavail)
exten => 146,3,Hangup()
exten => 147,1,Dial(SIP/phone47_2006,60)
exten => 147,2,Playback(vm-nobodyavail)
exten => 147,3,Hangup()
exten => 148,1,Dial(SIP/phone48_2006,60)
exten => 148,2,Playback(vm-nobodyavail)

exten => 148,3,Hangup()
exten => 149,1,Dial(SIP/phone49_2006,60)
exten => 149,2,Playback(vm-nobodyavail)
exten => 149,3,Hangup()
exten => 150,1,Dial(SIP/phone50_2006,60)
exten => 150,2,Playback(vm-nobodyavail)
exten => 150,3,Hangup()
exten => 151,1,Dial(SIP/phone51_2006,60)
exten => 151,2,Playback(vm-nobodyavail)
exten => 151,3,Hangup()
exten => 152,1,Dial(SIP/phone52_2006,60)
exten => 152,2,Playback(vm-nobodyavail)
exten => 152,3,Hangup()

exten => t,1,Playback(vm-goodbye)
exten => t,2,Hangup()

[local]

exten => 100,1,Dial(SIP/phone0_2006,,r)
exten => [phone0_2006@blueberry.csc.ncsu.edu](tel:phone0_2006@blueberry.csc.ncsu.edu),1,Dial(SIP/phone0_2006,,r)
exten => 101,1,Dial(SIP/phone1_2006,,r)
exten => [phone1_2006@blueberry.csc.ncsu.edu](tel:phone1_2006@blueberry.csc.ncsu.edu),1,Dial(SIP/phone1_2006,,r)
exten => 102,1,Dial(SIP/phone2_2006,,r)
exten => [phone2_2006@blueberry.csc.ncsu.edu](tel:phone2_2006@blueberry.csc.ncsu.edu),1,Dial(SIP/phone2_2006,,r)

exten => 103,1,Dial(SIP/phone3_2006,,r)
exten => [phone3_2006@blueberry.csc.ncsu.edu](tel:phone3_2006@blueberry.csc.ncsu.edu),1,Dial(SIP/phone3_2006,,r)

exten => 104,1,Dial(SIP/phone4_2006,,r)
exten => [phone4_2006@blueberry.csc.ncsu.edu](tel:phone4_2006@blueberry.csc.ncsu.edu),1,Dial(SIP/phone4_2006,,r)
exten => 105,1,Dial(SIP/phone5_2006,,r)
exten => [phone3_2006@blueberry.csc.ncsu.edu](tel:phone3_2006@blueberry.csc.ncsu.edu),1,Dial(SIP/phone3_2006,,r)
exten => 106,1,Dial(SIP/phone6_2006,,r)
exten => [phone6_2006@blueberry.csc.ncsu.edu](tel:phone6_2006@blueberry.csc.ncsu.edu),1,Dial(SIP/phone6_2006,,r)

exten => 107,1,Dial(SIP/phone7_2006,,r)
exten => [phone7_2006@blueberry.csc.ncsu.edu](tel:phone7_2006@blueberry.csc.ncsu.edu),1,Dial(SIP/phone7_2006,,r)

exten => 108,1,Dial(SIP/phone8_2006,,r)
exten => [phone8_2006@blueberry.csc.ncsu.edu](tel:phone8_2006@blueberry.csc.ncsu.edu),1,Dial(SIP/phone8_2006,,r)

exten => 109,1,Dial(SIP/phone9_2006,,r)
exten => [phone9_2006@blueberry.csc.ncsu.edu](tel:phone9_2006@blueberry.csc.ncsu.edu),1,Dial(SIP/phone9_2006,,r)

exten => 110,1,Dial(SIP/phone10_2006,,r)
exten => [phone10_2006@blueberry.csc.ncsu.edu](tel:phone10_2006@blueberry.csc.ncsu.edu),1,Dial(SIP/phone10_2006,,r)

exten => 111,1,Dial(SIP/phone11_2006,,r)
exten => 112,1,Dial(SIP/phone12_2006,,r)
exten => 113,1,Dial(SIP/phone13_2006,,r)
exten => 114,1,Dial(SIP/phone14_2006,,r)
exten => 115,1,Dial(SIP/phone15_2006,,r)
exten => 116,1,Dial(SIP/phone16_2006,,r)
exten => 117,1,Dial(SIP/phone17_2006,,r)
exten => 118,1,Dial(SIP/phone18_2006,,r)
exten => 119,1,Dial(SIP/phone19_2006,,r)
exten => 120,1,Dial(SIP/phone20_2006,,r)
exten => 121,1,Dial(SIP/phone21_2006,,r)
exten => 122,1,Dial(SIP/phone22_2006,,r)
exten => 123,1,Dial(SIP/phone23_2006,,r)
exten => 124,1,Dial(SIP/phone24_2006,,r)
exten => 125,1,Dial(SIP/phone25_2006,,r)
exten => 126,1,Dial(SIP/phone26_2006,,r)
exten => 127,1,Dial(SIP/phone27_2006,,r)
exten => 128,1,Dial(SIP/phone28_2006,,r)
exten => 129,1,Dial(SIP/phone29_2006,,r)
exten => 130,1,Dial(SIP/phone30_2006,,r)
exten => 131,1,Dial(SIP/phone31_2006,,r)
exten => 132,1,Dial(SIP/phone32_2006,,r)
exten => 133,1,Dial(SIP/phone33_2006,,r)
exten => 134,1,Dial(SIP/phone34_2006,,r)
exten => 135,1,Dial(SIP/phone35_2006,,r)
exten => 136,1,Dial(SIP/phone36_2006,,r)
exten => 137,1,Dial(SIP/phone37_2006,,r)
exten => 138,1,Dial(SIP/phone38_2006,,r)
exten => 139,1,Dial(SIP/phone39_2006,,r)
exten => 140,1,Dial(SIP/phone40_2006,,r)
exten => 141,1,Dial(SIP/phone41_2006,,r)
exten => 142,1,Dial(SIP/phone42_2006,,r)
exten => 143,1,Dial(SIP/phone43_2006,,r)
exten => 144,1,Dial(SIP/phone44_2006,,r)
exten => 145,1,Dial(SIP/phone45_2006,,r)
exten => 146,1,Dial(SIP/phone46_2006,,r)
exten => 147,1,Dial(SIP/phone47_2006,,r)
exten => 148,1,Dial(SIP/phone48_2006,,r)
exten => 149,1,Dial(SIP/phone49_2006,,r)
exten => 150,1,Dial(SIP/phone50_2006,,r)
exten => 151,1,Dial(SIP/phone51_2006,,r)
exten => 152,1,Dial(SIP/phone52_2006,,r)

Appendix D Evaluator's recorded challenge questions

Name: Wu, Dong

Tel. 103

- 1) I have five books. I lent 2 to my friend. How many books do I have? 3
- 2) I'm 9 years old. How old was I last year? 8
- 3) How many bottles of beer in one box? 12
- 4) If you have two brothers and one sister, how many children do your parents have?
4
- 5) There are one dozen of birds on the tree. Only five of them are female. How many birds are male? 7
- 6) If you are 20 years old, I am 3 years younger than you, how old am I? 17
- 7) If two bottles of milk are \$6 totally. How much to buy 1 bottle of milk? 3
- 8) If I drink twice one day and each time I drink two bottles of juice, how many bottles of juice do I need to buy for one week? 28
- 9) I am a teacher; my two daughters are teacher too; my son is a doctor. How many teachers are in my family? 3
- 10) I have 50 dollar and I pick up \$10, but I lent 65 to my friend. Do you think it is possible? No

Name: Liu, Shufang

Tel. 104

- 1) I'm 16. How old will I be next year? 17
- 2) I had four pens. I lost one, how many pens do I have? 3
- 3) My parents have 4 children. How many siblings do I have? 3
- 4) Today is Sunday. What day is tomorrow? Monday
- 5) My father and I have the same last name. Mine is Bush. What's my father's last name? Bush
- 6) I'm 16. My brother is two years younger than me. How old is my brother? 14
- 7) I have one blue pen, one yellow pen and one red pen. How many pens do I have? 3

- | | |
|---|-------|
| 8) My breakfast was bread. What did I eat this morning? | Bread |
| 9) My only cat gives birth to 2 kittens. How many cats do I have? | 3 |
| 10) I have two cats and a dog. How many pets do I have? | 3 |

Name: Li, Jiangtian

Tel. 105

- | | |
|--|----|
| 1) There are three cats and two dogs. How many animals are there? | 5 |
| 2) There are seven chairs in the room. One chair is removed from the room. How many chairs are left? | 6 |
| 3) How many vertexes are there in the triangle? | 3 |
| 4) There are five chickens. How many legs are there? | 10 |
| 5) There are four dogs. How many legs are there? | 16 |
| 6) There are five chickens and three dogs. How many legs are there? | 8 |
| 7) How many letter are there in the letter APPLE? | 5 |
| 8) How many seconds are there in a minute? | 60 |
| 9) How many minutes are there in an hour? | 60 |
| 10) How many hours are there in one day? | 24 |

Name: Lin, Shiyong

Tel. 106

- | | |
|--|-----|
| 1) Each apple costs 1 dollar. How much are 9 apples? | 9 |
| 2) I cooked 3 eggs for myself. If I eat one egg, how many eggs are left? | 2 |
| 3) One person has 150 lbs. If he can lose 20 lbs by exercise, what will be his final weight? | 130 |
| 4) A student checked out 10 books from library. After he returns 3 overdue books, how many books will he have? | 7 |
| 5) A father has 10 cookies for his two kids. How many cookies can each kid get equally? | 5 |
| 6) David bought 4 books from bookstore yesterday. Today he bought 3 more books. Totally how many books did he buy within two days? | 7 |

- 7) A company has totally 1000 employees. 600 are males. How many are females? 400
- 8) A person is 25 years old now. How old is he after 20 years? 45
- 9) Tommy has one dog and 2 cats. How many pets does he have? 3
- 10) The price for a T-shirt is \$10. If a custom can get a 50% off discount on the original price, how much should she pay for this T-shirt? 5

Name: Fu, Sai

Tel. 107

- 1) I have one daughter called Mary and one son called Jack. How many children do I have? 2
- 2) Reverse the string "1 2 3". 321
- 3) How many days in January? 31
- 4) How many "o"s in word "google"? 2
- 5) In your cell phone keyboard, what's the upper side of number 5? 2
- 6) What's the emergency number? 911
- 7) I have one penny and one quarter. How much do I have? 26
- 8) How many stars in the Chinese flag? 5
- 9) Which one is the biggest number among 3,4,5? 5
- 10) What's the integer between 1 and 3? 2

Name: Xue, Xiaoqiang

Tel. 108

- 1) Yesterday was Thursday. What day is it today? Friday
- 2) I have 5 pens, while Mr. Russ has 4 pens. So, totally how many pens do we have? 9
- 3) The price for an apple is 1.5 dollar. How much for 4 apples? 6
- 4) It was sunny yesterday, and it is the same today as yesterday. What's the weather is it today? Sunny
- 5) There were 7 people in the bar, 4 of them are male. How many females in the bar? 3

- | | |
|---|----|
| 6) How many cycles for a bicycle? | 2 |
| 7) There are 3 red apples and 5 green ones on the desk. How many apples totally? | 8 |
| 8) Gym closes at 6:00pm, and now it is 5:00pm. How long will the gym close? | 1 |
| 9) It is 10 mile driving to school from home. How many miles from home to school? | 10 |
| 10) I have 5 bananas and 4 pears in the bad. How many pears in my bag? | 4 |

Name: Li, Jie

Tel. 109

- | | |
|--|----------|
| 1) What is the abbreviation of the United States of America? | USA |
| 2) How many players of the NBA games on the court for each team? | 5 |
| 3) Today is Friday, what day is the day after tomorrow? | Saturday |
| 4) Tony's last name is James. What is the last name of his father? | James |
| 5) John has five apples. He gives his brother two apples. How many apples he left? | 2 |
| 6) Who is the current president of the United States? | Bush |
| 7) Kate was 5 years old on 2000. How old is she in 2003? | 8 |
| 8) How many inches make one foot? | 12 |
| 9) How many legs does dog have? | 4 |
| 10) What is the traditional food in Thanksgiving? | Turkey |

Name: Zhou, Shanshan

Tel. 110

- | | |
|---|----|
| 1) There are two beers. How many left after I drink one. | 1 |
| 2) Twenty students in my class. Ten are boys. How many girls? | 10 |
| 3) I have apples, bananas and oranges. How many kinds of fruits do I have? | 3 |
| 4) I have apples, bananas and breads. How many kinds of fruits do I have? | 2 |
| 5) Ten gallon tank. How much gas should I refill when I have 2 gallon left? | 8 |
| 6) How many eyes do I have? | 2 |

- | | |
|---|-----|
| 7) What is the other color of the traffic light besides yellow and green? | Red |
| 8) What is the letter between a & c? | b |
| 9) Who is older? Mom or child? | Mom |
| 10) Which is colder. 73 degree or 42 degree? | 42 |

Name: Xia, Heng

Tel. 111

- | | |
|---|-----|
| 1) It's sunny today. The probability of raining tomorrow given sunny today is 10%.
What's the probability of not raining tomorrow? | 90 |
| 2) There're two cats and three ducks. How many legs do they have? | 10 |
| 3) It takes me 1 hour to go to RBC center from home. How many minutes do I need
to come back home from RBC center? | 30 |
| 4) I like all the fruits. Do you think I like apples? | Yes |
| 5) I have a corolla. My parents have a BMW. How many cars do we have? | 2 |
| 6) Macy's is having 15% off discount on all the merchants. I want to buy a \$100
coat. How many dollars can I save? | 15 |
| 7) The first letter is R. The second letter is E. The third letter is D. What's the word? | Red |
| 8) I bought 10 oranges yesterday. I ate 2. My roommate ate 1. How many oranges
left? | 7 |
| 9) Google Check gives \$10 off your purchase if you're a new user. I'm not a new
user. Can I get this offer? | No |
| 10) I like all the songs by Maria Carrey. She has a song called "Hero". Do you think
if I like this song? | Yes |

Name: Chao, Jingyi

Tel. 112

- | | |
|--|------|
| 1) I walked to school. Did I take the bus this morning? | No |
| 2) I ordered the Pizza for dinner yesterday. Did I eat last night? | Yes |
| 3) I have three sisters. How many children that my parents have? | 4 |
| 4) I am a single-mother. Do I have the children? | Yes |
| 5) I am a boy or girl? | Girl |

- | | |
|---|---------|
| 6) It is raining. The cloud is white or black? | Black |
| 7) It is sunny. The cloud is white or black? | White |
| 8) Today is Saturday. Do I need to work now? | No |
| 9) I have ten dollars. It takes two dollars to buy the coffee? How much I left? | 8 |
| 10) Which operator system is more expensive? Linux or Windows? | Windows |

Name: Li, Wenjun

Tel. 113

- | | |
|---|----|
| 1) Mr. Brown has a twin of girls and one boy. How many children does Mr. Brown have? | 3 |
| 2) It is May. How many days does the next month have? | 30 |
| 3) I bought a dozen of oranges and ate 5. How many oranges are left? | 7 |
| 4) Mary and John held a party and invited 5 people to come. How many people were at the party? | 7 |
| 5) How many possible outcomes if a dice is thrown? | 6 |
| 6) All the tickets are sold out for the show. How many tickets will Mary get if she wants to buy 2? | 0 |
| 7) A car can hold four people at maximum. How many cars are needed for a group of 10 travelers? | 3 |
| 8) I've been to New York, Boston, Washington DC and Philadelphia during my trip. How many cities did I visit? | 4 |
| 9) There are 5 people in our lab. All but Mary are married. How many are married among all people in our lab? | 4 |
| 10) I should have been there at 10am, but I was an hour late. When did I get there? | 11 |

Name: Ni, Kai

Tel. 114

- | | |
|---|---|
| 1) I bought ten eggs and ate 3. How many eggs left? | 7 |
| 2) Cilice had her first child in 2005 and she wants her second child in 2008. How many children does Cilice have? | 1 |

- 3) Five little monkeys swing in the tree. One of them left and two another join. How many monkeys in the tree now? 6
- 4) There are four balls in the bowl. One is blue, two are red and the others are green. How many kinds of colors are there? 3
- 5) I am born in 1980 and I got my MS degree two years ago. How old am I when I get my MS? 25
- 6) A car with three people ran into another car with four people. Everyone are hurt except one of the drivers. How many people are injured in this accident? 3
- 7) My father is a doctor. My mother is a teacher. What's the job of my dad's wife? teacher
- 8) Mary likes swimming and tennis but not chess. Which sports Mary dislike? Chess
- 9) Curing couple has two sons and one daughter. How many people in Curing's family? 5
- 10) Gail like red jacket and hates green one, but she wears a blue cloth today. What's Gail's the favorite color? Red

Name: Tang, Yi

Tel. 115

- 1) My son has 2 car toys. He buys a new today. How many cars toy does he have? 3
- 2) The difference between 8 and 6 is? 2
- 3) My son is 8 years this year. How old was he last year? 7
- 4) I see two dogs and three cats on the way home. How many pets do I see? 5
- 5) Jim scored 5 points in the 1st half and added 9 points in the second. How many points he scored in the game? 14
- 6) There are 10 apples in the bag. I take out 4. How many apples left in the bag? 6
- 7) The first number of my phone is? 1
- 8) The last number of my phone is? 5
- 9) The minimized number between 5 and 7 is? 5

10) The maximized number between 4 and 6 is? 6

Name: Hyun, Sangwon

Tel. 116

- 1) 10 cars are in a parking lot, which is full. Two cars of them just leave. How many empty slots are there now? 2
- 2) I have 6 apples, 2 pineapples, and 4 cookies. How many fruits do I have? 8
- 3) What is the reverse of the letters in my last name? nuyh
- 4) It's 11:45 am. What is the remaining time to noon? 15
- 5) I have one ballpoint pen, one desktop and one laptop. How many computers do I have? 2
- 6) How many days are in Jan and March? 62
- 7) How many e's are in the word, blueberry? 3
- 8) There are two birds and 3 dogs. How many of them can fly? 5
- 9) How many e's are in the word, employee? 3
- 10) What is the name of the nation, which is commonly contained in the words, plus and minus? us

Name: Hu, Jingzhen

Tel. 117

- 1) My parents have two kids: John and me. Who is my brother? John
- 2) I raised two cats. All died. How many cats left? 0
- 3) There are 2 birds on the tree. One left. How many birds are on the tree? 1
- 4) A lion and a sheep meet together one day. Who will be killed? Sheep
- 5) I have 4 dogs and 3 cats. Dogs killed all cats. How many cats left? 0
- 6) I slept at 11am and got up at 2:00pm. How long did I sleep? 3
- 7) I had 3 apples and 2 piece of bread for my breakfast. How many fruits did I eat? 3
- 8) Four stars around one star. How many stars in the center? 1
- 9) The price of apple is 2 dollars. How many apples for 6 dollars? 3

10) I have a dozen of eggs. Two of them are thrown. How many left? 10

Name: Harsh Upadhyay

Tel. 118

- 1) If alphabet A is equal to 1; alphabet B is equal to 2; then alphabet Z is equal to what number? 26
- 2) How many gallons of oil in a tank with a capacity of 3 gallons? 3
- 3) If I pay 100 cents for a candy that is worth \$1. How much more do I need to pay? 0
- 4) What is the last letter of the word “everything”? g
- 5) If I have a laptop and a desktop. How many computers do I own? 2
- 6) Three out of the five lambs were eaten by the wolf. How many lambs were left alive? 2
- 7) If Jack is one kilometer away from Jill. How far is Jill from Jack? 1
- 8) I’m my father’s only son. How many siblings do I have? 0
- 9) I have 2 pigeons and 2 chairs. How many birds do I have? 2
- 10) Captain Haddock lost an eye in the 2nd world war. How many eyes does he have now? 1

Name: Devasthali Vinayak

Tel. 119

- 1) I have 5 flowers in my hand. I removed two of them. How many flowers are now remaining in my hand? 3
- 2) I had 10 dollars. My sister gave me 2 dollars. Now how many dollars do I have? 12
- 3) My brother eats 2 apples everyday. How many apples does he eat in 3 days? 6
- 4) I have 2 roses and 3 oranges. How many fruits do I have? 2
- 5) My sister has 2 white pens and 1 black pen. I have 2 black pens. How many black pens do we have? 3

- 6) There are 5 boys and 2 girls in my class. How many students are there in my class? 7
- 7) I started walking from my home at 11:00in the morning. I reached office after 2 hours. What is the time when I reach office? 1
- 8) I have 3 books and 4 bags. I gave 2 books to my sister. How many bags do I have? 4
- 9) My uncle has 2 cats and 3 dogs. How many animals does my uncle have? 5
- 10) I have to read 4 books. I read 2 books everyday. How many days do I need to read all books? 2

Name: Li, Dongxia

Tel. 120

- 1) How many hours in a day? 24
- 2) How many cycles in one bicycle? 2
- 3) You have two apples in hand. How many apples are left after you eating a banana? 2
- 4) How to call your father's sister? Aunty
- 5) Which flower is for Valentine's Day? Rose
- 6) How many eyes in your face? 2
- 7) What is the traditional food for Thanksgiving? Turkey
- 8) How many days in a week? 7
- 9) How many states in United States? 50
- 10) What is the abbreviation of Los Angeles? LA

Name: Li, Nan

Tel. 121

- 1) I have breakfast, lunch and supper everyday. How many meals do I have each day? 3
- 2) Which month is the shortest in a year? February
- 3) How many months are there in a year? 12

- | | |
|---|-----|
| 4) How many days are there in a week? | 7 |
| 5) How many days are there in a year? | 365 |
| 6) How many days are there between two new years? | 365 |
| 7) How many hours are there in a day? | 24 |
| 8) How many minutes are there in an hour? | 60 |
| 9) How many seconds are there in a minute? | 60 |
| 10) I have an apple and my sister has two apples. How many apples do we have? | 3 |

Name: Liu, Cui

Tel. 122

- | | |
|--|----------|
| 1) How many hours are there in a day? | 24 |
| 2) Which month is the shortest month in a year? | February |
| 3) How many ears do you have? | 2 |
| 4) I have three apples and one orange. How many apples do I have? | 3 |
| 5) How many legs does our cat have? | 4 |
| 6) There are four oranges in the box. How many oranges are there in the box? | 4 |
| 7) How many inches make one foot? | 12 |
| 8) How many months in a year? | 12 |
| 9) Today is Monday. What day is tomorrow? | Tuesday |
| 10) What's emergency number in USA? | 911 |

Name: Liu, Xiang

Tel. 123

- | | |
|--|----|
| 1) The gas price today is 3 dollars per gallon. How many dollar do you pay for 2 gallon gas? | 6 |
| 2) You need 10-gallon gas to drive 200 miles. How many gallon gases do you need to travel 400 miles? | 20 |
| 3) You pay \$5 to wash car once. How many dollars do you pay to wash twice? | 10 |
| 4) I had one apple and one orange today. How many fruits do I have today? | 2 |
| 5) I left home at 7 in the morning and I arrived school at 8. How many hours I spent on the road? | 1 |

- 6) I receive one check every month. How many checks do I get every year? 12
- 7) I sleep every day. How many times do I sleep every week? 7
- 8) I read one page every minute. How many pages do I read every hour? 60
- 9) I drink one beer every day. How many drink do I have every week? 7
- 10) I watch one movie every day. How many movie do I watch every week? 7

Name: Shin, Kyuyong

Tel. 124

- 1) I graduated from KAIST in 2000. How many years have so far? 7
- 2) I have 2 sons with my wife. How many people are there in my family? 4
- 3) I wake up 7 in the morning and it normally takes 2 hours before going to school.
What time do I left my home? 9
- 4) I was born in 2000. How old am I? 7
- 5) I have 5 apples. If I eat three of them, how many apples are there? 2
- 6) My favorite color is green but I like red apple. What's my favorite fruit? Apple
- 7) I go to bed at 12:00 midnight and get up 7:00am in the morning. How long do I sleep? 7
- 8) There are 7 children in the playground. Now 2 children join the group but at the same time 3 of them left. How many children are there? 6
- 9) There are 3 butterflies and 2 rabbits in my backyard. How many insects are there? 3
- 10) There were 3 bees on the flower but a bird ate 2 of them. How many bees are there? 1

Name: Park, Younghee

Tel. 125

- 1) I have 6 friends. Three friends left me. How many friends do I have? 3
- 2) It was 10 o'clock. The time passed one hour. What time is it now? 11
- 3) I have 5 hairpins and 2 headbands. How many accessories do I have? 7
- 4) I awake at 7 o'clock today. It takes me for 2 hours to get school everyday. What time can I arrive at school? 9
- 5) I have 10 pens, but 7 pens are out of work? How many pens left? 3

- 6) I have to review 5 papers. Each of them takes 1 hour. How many hours does it take to finish? 5
- 7) A desk needs a chair. I have 2 desks. How many chairs do I need? 2
- 8) I have 3 cell phones, but one is out of service. How many cell phones work? 2
- 9) Today is March 8th. What is the date of tomorrow on March? 9
- 10) There are ten tablets. I take 3 tablets. How many left? 7

Name: Miao, Huiping

Tel. 126

- 1) Peter is 5 years old. Mary is 7. Who's older? Mary
- 2) I had milk for breakfast, and rice for dinner. What did I have in the morning? Milk
- 3) I have 2 sisters and a brother. How many children are there in my family? 4
- 4) Mike has brown hair. Peter has the same eye color. What color are Peter's eyes? Brown
- 5) I had 11 dumplings. There're 4 dumplings left. How many did I eat? 7
- 6) Peter stayed at home today. Mary went to school. Who missed class today? Peter
- 7) I have 2 dogs and 7 roses. How many pets do I have? 2
- 8) Mike's Chinese. Chris is British. Who's Asian? Mike
- 9) I sent a letter via Fedex and another letter via UPS. How many letter did I send? 2
- 10) I bought 5 hats. I returned 3. How many hats do I have now? 2

Name: Kampanakis, Panos

Tel. 127

- 1) If now I am 25 years old. How many is it going to take me to become 26? 2
- 2) What if I had 4 friends and 2 of them left to go abroad. With how many friends do I hang out with now? 2

- 3) What is the number that is bigger than 1, smaller than 5 and is in the middle of them? 3
- 4) The names of my cats are Mary, Jenny, Cary and Alice. How many cats do I have? 4
- 5) If I work for 10 hours, sleep for 8 and eat for 1 a day. How many do I have to shop? 5
- 6) There is a train every 10 minutes. I lost the last one for 4 minutes? In how many minutes is the next train coming? 14
- 7) If I work for 70 years and change jobs every 10, how many jobs am I going to change? 7
- 8) I have 15 apples. I give 3 to my daughter and 4 to my son. How many are there left to make an apple-pie? 8
- 9) My aunt has been married 3 times and gave birth to 3 children in each marriage. How many children does she have? 9
- 10) I run 10miles a day? Charlotte is 100 miles away? How many days do I need to go there? 10

Name: Liu, An

Tel. 128

- 1) Alice is the daughter of Bob. Bob is the son of John. What is the relationship between Alice and John?
 - 1-Alice is John's granddaughter.
 - 2-Alice is John's mother
 - 3-John is the son of Alice.
 - 4-John is the teacher of Alice.
 - 5-John is a professor.
- 2) In Lake Johnson Park, I saw 3 ducks, my wife saw 2 dogs. How many animals we saw in Lake Johnson Park? 5
- 3) There are 3 birds in my garden. I walk into my garden. One bird fly away. How many birds are still in my garden? 2

- 4) I have 10 goldenfishs in my living room. My cat eat all of them. How many goldenfish do I have now? 0
- 5) Three boys play on the parking lot. One car comes. The driver park the car and leave. How many people are there on the parking lot? 3
- 6) Today is Thursday. The paper is due this Sunday. How many days do I have for revising paper including today and Sunday? 4
- 7) I have 10 dollars. Then I buy three apples in Foodlion. Each of them cost 1 dollar. How much money do I have now? 7
- 8) Tiger eats sheep or sheep eats tiger?
6 for tiger eats sheep
9 for sheep eats tiger 6
- 9) I prepare 10 once milk for my baby. My baby drink 7 once. How many once milk is left in the bottle? 3
- 10) It's 1:00 pm. I have a meeting after 3 hours. When is the meeting? 4

Name: Zhang, Qinghua

Tel. 129

- 1) How many meals does a human eat everyday? 3
- 2) There are 5 people on a bus, two people get off, and three people get on. How many people are there on a bus? 6
- 3) How many toes do three normal people have? 6
- 4) A one-year-old rat has three one-month baby rat. How many parents does this father rat have? 2
- 5) It's three o'clock in the morning. When will it be 23 hours later? 2
- 6) Mr. Smith has one daughter, one son and no wife living with him. How many people are there in his family? 3
- 7) Two cats eat three rats. One rat run away. How many rats died? 3
- 8) How many meals does a normal people eat in two days? 6
- 9) A thief steals 2 dollars from me and three dollars from Tom. How many dollars do I lose? 2

10) I got two monitors and one computer. How many monitors can be used for the computer? 2

Name: Stille, John Cosgrove

Tel. 130

- 1) I have three pairs of dirty socks and one pair of clean socks, how many individual socks do I have? 8
- 2) There are 4 quarters in a dollar so how many quarters does it take to make 2 dollars? 8
- 3) I am 21 years old and my sister is 23 years old, how many years older is my sibling? 2
- 4) My car is red and my bike is blue, what color is my automobile? Red
- 5) There are 6 dogs, how many pairs of dogs exist? 3
- 6) My dog eats 2 bowls of food a day, how many bowls of food will the dog eat in 1 week? 14
- 7) My phone has 10 keys on it, what is the biggest number on the phone? 9
- 8) A beer has 100 calories and a cookie has 40 calories, how many calories are in the alcoholic beverage? 100
- 9) I draw a square on the board, how many sides does it have? 5
- 10) My mother had 5 children and 2 dogs, how many siblings do I have? 4

Name: Li, Xin

Tel. 131

- 1) I have two cat and three dogs, how many pets do I have? 5
- 2) I have two sisters and one brother, how many kids my parents have? 4
- 3) A mule has four legs, how many legs do they have if I own 10 mules? 40
- 4) It is ten forty in the morning, what would the time be after twenty minutes? 11
- 5) I spent 35 dollars on baseball, each ball worth 3.5 dollars, how many ball did I get? 10
- 6) My family has 3 twins, how many children do we have? 6

- 7) I borrow my friends 10 dollars, and he return me 8 later, how much did he still owe me? 2
- 8) A rabbit has two ears, how many rabbits I got if I have total 36 ears? 12
- 9) I break a rod into half, then I got 2 rods. How many rods will I get if I bread the rod twice? 4
- 10) I have 25 students, each of them has 2 parents, how many parents totally do they have? 50

Name: Dai, Guanglin

Tel. 132

- 1) There are 20 students sitting in 5 rows. How many students are there in 1 row? 4
- 2) My car can fill 10 gallon gas. The gas price now is \$2.5/gallon. How much do I have to spend if I fill up my oil tank? 25
- 3) There are 10 people in my office. All but me are Americans. How many Americans are there in my office? 9
- 4) It takes 5 hours to fly from Raleigh to Los Angelis, and 12 hours to fly from Los Angelis to Shanghai. How many hours will it take to fly from Raleigh to Shanghai if only one stop is made in Los Angelis? 17
- 5) James was born on Mar. 1, 2000. Today is Mar.1, 2007. How old is James? 21
- 6) Lee's family has a car and a truck. They bought a new van last week. How many vehicles does Lee's family have? 3
- 7) 30 students took a speech test. Only 2 failed. How many students passed the test? 28
- 8) Sam makes \$400 / month. How much does he make in 10 months? 4000
- 9) Kay is 120 pounds now. She lost 20 pounds since she started working out. How heavy was she before? 140
- 10) The meeting started at 10:00 am and ended at 2:00 pm. How long did it last? 4

Name: Han, Yang

Tel. 133

- 1) I cost 4 bucks out of my 10 bucks to buy a pen. How much money left? 6
- 2) The gas is \$2.5/gallon now. How much is for 10 gallons? 25
- 3) Today is March 22. How many days left in this month? 9
- 4) I will graduate on May 12 and start my professional career beginning from May 31. How many days in maximum can I have for my vacation after my graduation? 19
- 5) Crabtree mall has big deal! You can get 10% off on a skirt whose original price was \$100. How much money can you save? 10
- 6) I booked two flight tickets; each of them costs \$180. How much is the total price of the tickets? 360
- 7) I have two cats. My friend gives me a dog today as a gift. How many pets do I have now? 3
- 8) It's 30 miles from NC State to UNC. How many miles do I need to drive for a round trip between the two schools? 60
- 9) I register 9 credits each semester. What's my total credits in an academic year? 18
- 10) There were 6 apples on the table. I eat 2. How many left? 4

Name: Pan, Liming

Tel. 134

- 1) A student is allowed to choose 5 courses each semester, Mike has chosed a math course and a physics cours-es. How many more courses could he choose? 3
- 2) Each day it takes tom 3 hours in tutorial center as a teach assistant, how many hours does he spend for a whole week? 21
- 3) Lee has 6 pencils first and gives 3 to Rose, how many left? 3
- 4) The price of beef is \$10.00 a pound, how much I should pay for 3 pounds? 30
- 5) I hope to have a trip to Europe this July, now it is March, so how many months should I wait? 4
- 6) My friend's baby is 10 months old, so how long will it have the first birthday present? 2

- 7) I borrowed \$100 to invest in stock market, and the price for one share of IBM stock is \$10, how many shares could I buy? 10
- 8) Mike's birthday is Feb. 29th, how many birthday parties did he have for each 4 years? 1
- 9) Jim's salary is \$20 per hour, and he works 8 hours a day. How about the total he can earn each day? 160
- 10) Matt has 8 apples and get 3 more from Lisa, how many apples does he have now? 11

Name: Zhang, Zhe

Tel. 135

- 1) When the clock is 3pm, how much degree is the angle between the hour arrow and the minute arrow? 90
- 2) Which month in Chinese calendar is the Lantern Festival? 1
- 3) Which day in a month is the mid-autumn festival in Chinese calendar? 15
- 4) In apple, pear and beef, how many kinds of fruit do we have? 3
- 5) How many legs do 2 crabs have in total? 16
- 6) Which rank does the year of dog have in the 12 animal years counting backwards? 2
- 7) In 2 pencils, 5 notebooks and 10 apples, how many pieces of stationery do we have? 7
- 8) How many legs do 3 crabs and 5 fishes have? 24
- 9) A person is 8 years old this year, how old will he/she be in 2009? 10
- 10) In 30 boats and 50 cars, how many pieces of water transportation vehicle do we have? 30

Name: Li, Liping

Tel. 136

- 1) I'm 26 years old this year. How old was I three years ago? 23
- 2) I have a laptop and one desktop. How many computers do I have? 2
- 3) I'm the only child of my family. How many boys do my parents have? 0
- 4) I bought a ring for myself and another for my husband? How many rings do I buy today? 2

- 5) I went to classroom A five minutes ago. I'm in classroom B now. How many classrooms have I ever been? 2
- 6) I flew from Raleigh to Dallas and then from Dallas to Denver. How many stops have I changed from Raleigh to Denver? 1
- 7) I have 1000 suits. I dressed one for one party. How many suits are there in my closet now. 999
- 8) I have two novels, three journals and a newspaper. How many kinds of readings do I have? 3
- 9) The diameter of earth is 6000 kilometers. I traveled half. How much left? 3000
- 10) I'm drunk and see 10 moons. How many actually are there? 1

Name: Leng, Wei

Tel. 137

- 1) There are 12 groups in a class, and there are 4 students in each group. How many students are there in 5 groups? 20
- 2) Suppose there are four doors open in one room. How many doors are open after I closed one door? 3
- 3) I put three oranges and four eggs in an empty basket. How many kinds of fruit are in the basket? 1
- 4) I constructed 12 buildings and destroyed one third of them. How many buildings are left? 8
- 5) The length and width of a box is 4 centimeters and 5 centimeters, what is the area of the box? 20
- 6) I have lunch and dinner everyday, how many lunches do I have in one week? 7
- 7) I write seven numbers in a blackboard and erased them, how many numbers are left on the blackboard? 0
- 8) I bought five apples and four oranges. How many apples do I still have after I eat three oranges? 5
- 9) Desks have four legs and chairs have three legs, how many legs do two chairs have? 6

- 10) One employer has three employees, how many employees does the employer have after he dismissed one employee and recruits four new ones? 6

Name: Anush

Tel. 138

- 1) I bought five mushroom pizzas, two spinach pizzas and three beef pizzas. How many vegetarian pizzas did I buy? 2
- 2) I sleep before nine pm every day except for the weekends, when I sleep at midnight? What time do I sleep on Saturday? 12
- 3) I bought 15 computer monitors and 5 VCRs. How many items did I buy that did not have monitors? 5
- 4) I play basketball, tennis and racquetball. How many sports do I play? 3
- 5) I wear glasses on Mondays and Saturdays and wear lenses on the rest of the days. How many days do I wear lenses? 5
- 6) I play drums and guitar. How many instruments do I play? 2
- 7) I married five times and divorced twice. How many times did I not divorce? 3
- 8) I have sixty houses. I sold a third. How many do I have left? 20
- 9) I ate 70 burgers and 16 apples. How many fruits did I eat? 16
- 10) I have 25 cars and 100 buses. How many automobiles do I have? 125

Name: Arhun, Vinod

Tel. 139

- 1) I sleep at 11:00 pm and I wake up at 7 am everyday. How many hours a day do I sleep? 8
- 2) I have 15 apples and 20 bananas. How many fruits do I have in total? 35
- 3) I have 20 chocolates and 5 pairs of shoes. How many eatable items do I have? 20
- 4) This month is March. My birthday is in January. How many months should I wait for my birthday? 10

- 5) I have visited India, America and Africa. How many continents have I been to? 3
- 6) I take 60 minutes to go to school from my home. Now the time is 9 am. I need to be at school at 11 am. When should I leave home to reach school at the correct time? 10
- 7) I have 2 slices of bread for breakfast, 4 slices for lunch and 4 slices for dinner. How many slices of bread do I have in a day? 10
- 8) Today I bought 5 gallons of juice. I finished 2 gallons. How many gallons of juice am I left with? 3
- 9) I drink 1 gallon of milk, 1 gallon of water and 1 gallon of juice everyday. How many gallons of liquid do I drink in a day? 3
- 10) I bought 1 gallon of milk on March 5. Milk would stay good only for 10 days, after which I would not be able to drink it. What day of March should I next buy milk? 16

Name: Liang, Jing

Tel. 140

- 1) I need 10 colors for dyeing the fabric; I have 3, how many more colors I need? 7
- 2) I have three bags of polymer, each weights 30g, how much weight of these three bags of polymer is? 90
- 3) Amma is 3 years old, how old she will be after 70 years. 73
- 4) I have 400 Falaria, and 200 Camry, and 300 BMW, how many cars I have? 900
- 5) I have 3 dogs, 2 cats and 3000 roses, how many flowers I have? 3000
- 6) My grandfather is 120 years old, and my brother is half of his age, how old is my brother? 60
- 7) I have 12 uncles, and each uncles has 3 children, how many children my uncles have? 36
- 8) I bought 3 apples, 4 oranges, and 5 kiwis, how many fruits I have? 12
- 9) I waked up at 10:00 AM, and I went to school at 2:00 PM, how many hours between? 4

10) Now it is March, and my child will be born in September, how many months I need wait? 6

Name: Sezer, Emre Can

Tel. 141

- 1) I have one cat two dogs and one monkey. How many pets do I have?
3
- 2) Susan is 7 years old and Mary is 2 years old. Which one is older?
Susan
- 3) If Mary can solve 2 math questions a day, how many can she solve in a week?
14
- 4) If I'm 28 years old and my sister is 2 years younger, how old is she?
26
- 5) If I buy 2 oranges 3 apples and then two more oranges, how many oranges will I have in the end?
4
- 6) This morning I had 20 dollars in my pocket. I spent half of it at lunch and another dollar on snacks. How much money do I have now?
9
- 7) If I complete a lap around the track in 1 min and 12 seconds, how many seconds does it take for me to complete a lap around the track?
72
- 8) If I were to assign digits for the letters in the alphabet starting with 1 for A, 2 for B and so on, what would be the code for the word "cab"?
312
- 9) If John is 5 years old and his sister is a year younger than him, what is their total age?
9
- 10) What is the only prime number between 6 and 9?
7

Name: Kil, Chongkyung

Tel. 142

- 1) I bought three mackerels and two-hair tail today. How many fish do I have now?
5
- 2) I read three cartoons and nine fairy tales. How many books did I read today?
12

- 3) Today, I had a slice of pizza for lunch. For dinner, I had three slices of them. How many slices did I have today? 4
- 4) Chulsu has five oranges originally and give three oranges to Younghee. How many oranges does Chulsu have now? 2
- 5) Chulsu eats three buns everyday. How many buns does Culsu eat for five days? 15
- 6) Dongsu and Chulsu play dice. Donsu rolls the dice and he got 5. Chulsu also rolls the dice and he got 6. What is the sum of Chulsu's and Dongsu's? 11
- 7) Chulsu brings three science finctions to the school today. Younghee brings two fairy tale books. How many books we have now? 5
- 8) Younghee has three red skirts, three white pants, and three pink shirts. How many pieces of cloths does she have? 8
- 9) Younghee gave three eggs to Chulsu and five eggs to Dongsu. Now she got two eggs in her hand. How many eggs Younghee had originally? 10
- 10) Dongsu solves two math problems and three English problems daily. How many questions Dongsu will try for a week? 35

Name: Kariath, Riya Raju

Tel. 143

- 1) I was 10 years old in 1996. How old am I at the end of 2007? 11
- 2) I have 10 roses and 15 petunias in my hand. How many lilies do I have in my hand? 15
- 3) I woke up today at 8 am. The time now is 2 pm. How long have I been awake? 6
- 4) I own 4 cars. 2 of them are Ferraris. The rest are Cadillac. How many Cadillac do I have? 2
- 5) I have 15 bananas in my storage. I ate 3 of them. How many do I have left? 12
- 6) I have 6 chairs and 5 tables in my office. Each table has a laptop. How many laptops do I have in my office? 5
- 7) A human being makes approximately 1000 decisions a day. How many decisions does a human make in a day? 1000

- 8) A chessboard has 64 boxes. Half of them are white and the rest are black. How many black boxes are there in a chessboard? 32
- 9) I have 2 apples in my storage. I bought 4 more apples. How many apples do I have now? 6
- 10) N C state lost to UNC by 9 points in a basketball game. UNC scored 89 points. What was NC State's score? 80

Name: Du, Yan

Tel. 144

- 1) What are my phone number's last 2 digits? 44
- 2) If daytime saving time starts today, how many hours do we gain or lose? 1
- 3) If my dad has only one brother and no siblings, how many aunts from my father side do I have? 0
- 4) If my work place is number 1, my home is number 2, where do I sleep each night? 2
- 5) I have 3 apples, 4 peaches, 5 pair of shoes, how many fruits do I have? 7
- 6) I have following msn accounts, yu@msn.com, hi@msn.com, how many msn accounts do I have? 2
- 7) I went to Denver for a conference; i stayed in conference center all week. How many days did I ski in Colorado? 0
- 8) If today is one, tomorrow is 2, what's number for "the day after tomorrow"? 3
- 9) If BP is one, the oil price at BP is 2.39 for 87, Shell is two; the oil price at Shell is 2.47 for 87, which one do I choose? 1
- 10) What does number stand for with X in Greek letters? 10

Name: Yi, Ting

Tel. 145

- 1) Jason has 3 apples and 40 oranges. How many fruit do Jason have? 43
- 2) Rong has 50 peaches and 6 grapes. How many fruit do Rong have? 56

- 3) I have 20 banana and 7 cakes. How many fruit so I have? 20
- 4) We have eaten 20 apples and 50 peaches. How many fruit have we eaten?
70
- 5) I have 10 girlfriends and 15 boyfriends. How many friends do I have?
25
- 6) Ting has 26 sucks and 15 shoes. How many pairs of sucks does ting have?
13
- 7) I will go to china this Christmas. Today is January. How long will I go to china
from now on? 11
- 8) There is 30miles from my home to office. Now I have droved 10 miles. How far
will I go to office? 20
- 9) There are 64 boys and 11 girls in my class. How many students are there in my
class? 75
- 10) Ting has gone to American for 9 years and she will stay here for another 5 years.
So how many years have ting been in American? 14

Name: Jin, Qing

Tel. 146

- 1) There are four birds in the tree. How many heads do they have? 4
- 2) If there are 17 balls, how many balls are needed at least to be more than 50% of
the total? 9
- 3) There are three cats and two chickens. How many legs do they have?
16
- 4) Two liters of water is in a cup. How many liters of water will be left if I drink a
half of it? 1
- 5) If I have one red ball and three blue balls, what is the total number of these balls?
4
- 6) The distance from the start to destination is 10 km. How many hours it will cost
me if walking at a speed of 2 km per hour? 5
- 7) I have quarters, how many coins are needed to constitute 1dollar exactly?
4

- 8) If I want to exchange 1-dollar currency for quarters, how many coins can I get? 4
- 9) There are one bicycle, one car. How many wheels are there totally? 6
- 10) I have three apples and two peaches. How many apples are there if I eat one peach. 3

Name: Huang, Yushen

Tel. 147

- 1) I have 3 sons and 1 daughter, how many children do I have? 4
- 2) I have 2 dogs and 4 cats, how many kinds of pets do I have? 2
- 3) Mary has 3 apple trees and 200 peach trees, how many trees does Mary have? 203
- 4) Mary's father gave 24 roses to Mary, how many dozes of roses do marry have? 12
- 5) Mary has 3 dresses and 12 skirts, how many kinds of clothes does Mary have? 2
- 6) Mary just had her 3rd wedding, how many ex-husbands did her have? 2
- 7) Mary has 100 dogs and 200 blankets, how many pets does Mary have? 100
- 8) I like swimming, tennis, bowling and golf, how many kinds of sports do I like? 4
- 9) I have 1 brother and 2 sisters, how many children do my parents raise? 4
- 10) I broke 3 cups and 6 bowls, how many pieces did I break? 9