Secure Captcha Input based Spam Prevention

L. Devi Priya, S. Karthik

Abstract*— Spam is one of the principal stumbling blocks of Internet, which consistently emerge as redundant or unsolicited messages in various communication areas. For example in VoIP, spam appears as unnecessary calls generated by the computers and other auto call generating BOTs. In WEB servers like Gmail, FTP servers like Rapid Share spam appears as fake accounts creation and it sometimes lead to jam of server process due to the BOTs. We have anticipated a resolution to thwart the spam using SECURE CAPTCHA INPUT (SCI) system. CAPTCHA (Completely Automated Public Turing Test to Tell Computers and Human Apart) method aims to determine whether the call is coming from a human or a machine. A CAPTCHA is a type of challenge-response test used in computing to ensure that the response is not generated by a computer. The process usually involves one computer (a server) asking a user to complete a simple test which the computer is able to generate and grade. Because other computers are unable to solve the CAPTCHA, any user entering a correct solution is presumed to be human. Thus, it is sometimes described as a reverse Turing test, because it is administered by a machine and targeted to a human, in contrast to the standard Turing test that is typically administered by a human and targeted to a machine. CAPTCHA that is deployed here is a 3D model developed with action script that can overcome RT-MITM attack which is another milestone in the CAPTCHA security.

***Index Terms*— CAPTCHA, SPAM, RT-MITM attack.**

I. INTRODUCTION

Human race has shrunk our planet within a computer with the advancement of the networks with vast database called 'Internet'. Most traditional communications media including telephone, music, film, and television are being reshaped or redefined by the internet. Newspaper, book and other print publishing have adapted to Web sites and blogging. The Internet has enabled or accelerated new forms of human interactions through instant messaging, Internet forums, and social networking.

CAPTCHAs, in an attempt to block automated interactions with their sites. These efforts may be crucial to the success of these sites in various ways. For example, Gmail improves its service by blocking access to automated spammers, eBay improves its marketplace by blocking bots from flooding the site with scams, and Facebook limits creation of fraudulent profiles used to spam honest users or cheat at games. The most widely used CAPTCHA schemes use combinations of distorted characters and obfuscation techniques that humans can recognize but that may be difficult for automated scripts. There are different types of spam:

Manuscript received May, 2013.

L.Devi Priya, ME-Software Engineering, SNS College of Technology, Coimbatore, Tamil Nadu, India.

Dr.S.Karthik, DEAN/CSE, SNS College of Technology, Coimbatore, Tamil Nadu. India.

Call Spam: Number of calls to attempt user. If user answers the call, the user hears a recorded message and the call ends when the message finishes. This type is used by spammers on PSTN for marketing and is used widely by telemarketers as well. This type of spam is also known as Spam over Internet Telephony (SPIT).

IM Spam: This type, which is similar to email spam, describes the bulk of unsolicited messages where the spammer uses Instant Messaging (IM) to send spam to the user. This type of spam is also known as Spam over Instant Messaging (SPIM)

Presence Spam: This is similar to IM Spam since it consists of a large number of unsolicited set of presence requests. It means that the message is trying to get authenticated or 'white listed' by the user. This is also known as Spam over Presence Protocol (SPPP).

Captcha

Completely Automated Public Turing tests to tell Computers and Humans Apart (CAPTCHAs) are often the first line of defense in many online services. Their purpose is to protect such services and resources from automated misuse by malicious programs.

Fig. 1 Captcha

Captchas are sometimes called "reverse Turing tests" because they are intended to allow a computer to determine if a remote client is human or not. Many websites use CAPTCHAs in an attempt to block automated interactions with their sites. These efforts may be crucial to the success of these sites in various ways. For example, Gmail improves its service by blocking access to automated spammers.

Secure Captcha Input (Sci)

CAPTCHA is the use of hard AI problem to distinguish human and BOT apart which was originally evolved from visual authentication and identification. The primary use of CAPTCHA is to fight against BOT in account registration and click fraud. Also, it application can be used to authenticate a group of peoples sharing common knowledge or abilities. However, CAPTCHA itself is hard to authenticate specific person by asking personal or professional question. In fact, visual human verifiable techniques are vulnerable to MITM attack. Also, careless CAPTCHA implementation can lead the application fail to achieve its mission.

II. LITERATURE SURVEY

There are number of method which has been implemented to solve spam problems using captcha. To block the unwanted calls the author suggested some of the methods that are stated.

CAPTCHA Strengths

A systematic study of existing visual CAPTCHAs based on distorted characters that are augmented with anti-segmentation techniques. Applying a systematic evaluation methodology to 15 current CAPTCHA schemes from popular web sites, we find that 13 are vulnerable to automated attacks. Based on this evaluation, we identify a series of recommendations for CAPTCHA designers and attackers, and possible future directions for producing more reliable human/computer distinguishers. As visible in figure 2.1, real-world Captchas exhibit a lot of variation in their design. By analyzing how each scheme is constructed we grouped the security defenses used in these schemes into the following ten techniques. These techniques were assigned into the anti-recognition or the anti-segmentation category. We assigned to the anti-recognition category every feature that didn't directly prevent segmentation.

The anti-recognition techniques considered are:

1. Multi-fonts - Using multiple fonts or font-faces.
2. Charset - Which charset the scheme uses.
3. Font size - Using variable font size.
4. Distortion - Distorting the captcha globally using attractor fields.
5. Blurring - Blurring letters.
6. Tilting Rotating - Characters with various angles.
7. Waving rotating - The characters in a wave fashion.

The anti-segmentation techniques considered are

1. Complex background - Try to hide the text in a complex background to "confuse" the solver.
2. Lines - Add extra lines to prevent the solver from knowing what the real character segments.
3. Collapsing - Remove the space between characters to prevent segmentation.

To mitigate the Spam Problem

CAPTCHA (Completely Automated Public Turing Test to Tell Computers and Human Apart) uses the Turing Test approach to determine whether the user or caller is human or machine. Email is one application that uses the CAPTCHA method to prevent spam entering the user's mailbox. The strength of this method and its success in preventing spam in email made it appealing as a means of preventing spam in VoIP applications. The basic idea of this method is to implement a challenge response application for call establishment in application.

Most applications are open source at the present time. As mentioned before, the approach for spam solution in is to be used in the client application software. One of the open source has been selected to implement this research to see whether this approach is suitable on each open source VoIP client. For this research, we have selected MjUA from MjSIP application as a client for implementing the approach and integrate it. The main outcome of this project research is the program is able to block spam calls without user interference. The program automatically blocks any attempted call that is suspected to be spam with a challenge program embedded in the VoIP client. Each call that attempts the VoIP client is not permitted to establish or request a call unless the sender provides the correct CAPTCHA answer when challenged.

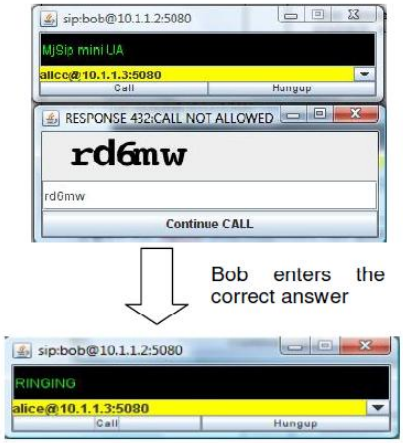

Fig 2 client establish a call

Figure 3 shows the connection to Alice finally established after Bob answer the CAPTCHA correctly. In contrast to Figure 3, Figure 4 shows the client is blocked from establishing a call with Alice because the answer to the CAPTCHA is not correct. A pop-up window has appeared to alert the user that the call is not authorized.

The client is blocked from establishing a call with Alice because the answer to the CAPTCHA is not correct. A pop-up window has appeared to alert the user that the call is not authorized.

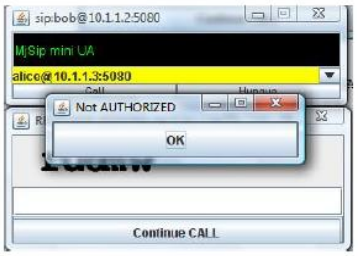

Fig 3 call not established

3D Captcha

CAPTCHA is a way how to verify that web application user is human. Using of CAPTCHA is a need in many web applications. Today it is commonly used by types based on the recognition of alphanumeric characters. The problem is that these tests are becoming more and more complicated for people but on the other hand they are becoming even easier for bots.

Creating 3D captcha system the major emphasis on innovation and user friendliness. The CAPTCHA is based on the human imagination. The basic idea is rotation of a special 3D model and finding the correct position of rotation. 3D can be created from 2D image. The task of user is to rotate the model to find the right observation point and solve captcha.

III. SYSTEM ANALYSIS

Existing System

CAPTCHA which are used from third party vendors are subjected to various attacks such as RT-MITM and OCR attacks which are still in use.

A successful CAPTCHA smuggling attack is performed with a malicious component on the victims' host needs to intercept the user interactions with an online service (e.g., Facebook) and delay their execution until the victim successfully solved a CAPTCHA challenge. .

Proposed Work

CAPTCHA (Completely Automated Public Turing Test to Tell Computers and Human Apart) method aims to determine whether the call is coming from a human or a machine. SCI have been used to prevent automated access to specific authenticated online services. Miscreants have tried to circumvent these protection mechanisms. CAPTCHA that is deployed here is a 3D model that can overcome RT-MITM attack which is another milestone in the CAPTCHA security.

IV. METHODOLOGY

System Design

The client is trying to establish a call with user and user challenge client with CAPTCHA images to prove client is a human rather than a machine. Client needs to response with the correct answer of the CAPTCHA in order to establish calls with Mug. The main purpose of using the challenge-response is to prevent the spammer from establishing call with the VoIP users. In order to implement this approach, the CAPTCHA application needs to be integrates with the existing UA clients.

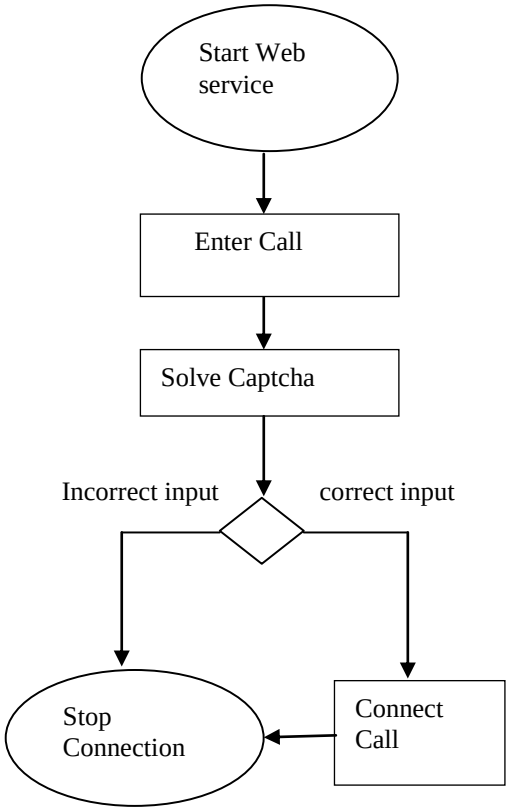

Fig4: Data Flow Diagram

The CAPTCHA application has been implemented as a pop-up application in the existing UA client.

Module Description

Analysing Spam problems and examine the CAPTCHA behaviour

Various dynamic methods have been developed to avoid the attacks. The primary defensive purpose is recognizing human users and computer programs in reading images of text. The hacking computer programs run dynamically, so the security systems should run dynamically against them. It is essential to develop automatic defense systems with continuously updated features for blocking the programmed attacks.
This method should be done automatically because examining of a large amount of registration on the Internet web sites by human power requires a load in terms of time and cost. CAPTCHA is a challenging puzzle used to determine whether a user is human or not and it's a cryptographic protocol whose underlying hardness assumption is based on an AI problem. It is a program that can generate and grade tests that most human can pass but current computer programs cannot pass.
CAPTCHA is basically used to produce the image of words that Optical Character Recognition (OCR) cannot understand. OCR programs can recognize the texts with high quality; hence; it is the fundamental thought to develop text images with low quality that is distortion in images in so many ways. Moreover CAPTCHA reduces the new hardware cost and their maintenance charges for providing security like biometric level securities. Admin can add Products for user using Product ID .When Authenticated user can process and create account. After creating an Admin will activate their Account. After Activation only the user can processed using that User id. After Activation Admin can view the details of purchased items with user ID and Product ID. After Purchasing, Admin will give away the delivery status like when the product will reach the user.

CAPTCHA Designing

The CAPTCHA has been designed in the way to foil the OCR attack and RT-MITM attack. It was developed using the Flash Action Script having a rotator and a cube. The cube has been made to move according to the usage of the rotator by the user.

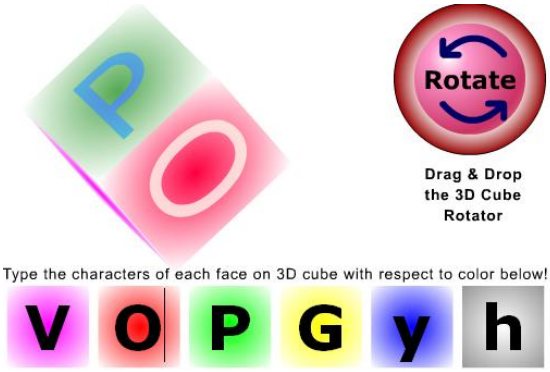

Fig 5 Captcha Design

Implementation of CAPTCHA in Webpage

In this module, the modules II and III are integrated. The CAPTCHA had been embedded with the client software which is used by the users to communicate each other. When a user wants to call another user he will be prompted to solve the CAPTCHA when the user presses the call button. If the user wants to continue the call he/she should solve the CAPTCHA. If the user fails to solve then the call will be aborted.

V. CONCLUSION

We have anticipated a resolution to thwart the spam in using SECURE CAPTCHA INPUT (SCI) system. CAPTCHA (Completely Automated Public Turing Test to Tell Computers and Human Apart) method aims to determine whether the call is coming from a human or a machine. SCI have been used to prevent automated access to specific authenticated online services. Miscreants have tried to circumvent these protection mechanisms. CAPTCHA that is deployed here is a 3D model that can overcome RT-MITM attack which is another milestone in the CAPTCHA security. The strength of this method and its success in preventing spam in email made it appealing as a means of preventing spam in applications. The spam prevention in Webpage that has been implemented on this project provides an advantage especially from the user's point of view. We have demonstrated that the CAPTCHA method can be applied in the applications and the CAPTCHA used here is capable of foiling OCR and RT-MITM attacks.

ACKNOWLEDGMENT

The authors would like to thank the editor-in-chief, the associate editor and anonymous referees for their comments.

REFERENCES

1. Ahmad El Ahmad, Jeff Yan and Wai-Yin Ng **"**CAPTCHA Design-Color, Usability, and Security".
2. CAPTCHA Official Site
3. Elie Bursztein, Matthieu Martin, John C. Mitchell, "Text-based CAPTCHA Strengths and Weaknesses", ACM Computer and Communication security 2011.
4. Elie Bursztein, Steven Bethard, John C. Mitchell, Dan Jurafsky, and Celine Fabry. How good are humans at solving captchas? A large scale evaluation. In Security and Privacy, 2010.
5. Ismail Ahmedy, Marius Portmann, "Using Captchas To Mitigate The VoIP Spam Problem", Second International Conference on Computer Research and Development 2010.
6. Jeff Yan and Ahmad Salah El Ahmad," Captcha Robustness: A Security Engineering Perspective", Research feature, New Castle University.
7. L. von Ahn, M. Blum, N. J. Hopper, and J. Langford, (2003) "Captcha: Using hard AI problems for security," in EUROCRYPT, pp. 294–311.
8. L. von Ahn, M. Blum, and J. Langford, "Telling Humans and Computer Apart automatically", ACM, vol. 47, no. 2, 2004, pp. 57-60
9. MacIntosh, R.; Vinokurov, D., (18-19 April 2005) "Detection and mitigation of spam in IP telephony networks using signaling protocol analysis," Advances in Wired and Wireless Communication", 2005 IEEE/Sarnoff Symposium on , vol., no., pp.49-52.
10. Robert MacIntosh and Dmitri Vinokurov, "Detection and Mitigation of Spam in IP Telephony Networks using Signaling Protocol Analysis".
11. Sajad Shirali-Shahreza, Ali Movaghar," A New Anti-Spam Protocol Using CAPTCHA", Proceedings of the 2007 IEEE International Conference on Networking, Sensing and Control, London, UK, 15-17 April 2007.
12. S.Y. Huang, Y.K. Lee, G. Bell, and Z. Ou. A projection-based segmentation algorithm for breaking MSN and YAHOO CAPTCHAs. In Proceedings of the World Congress on Engineering, volume 1. Citeseer, 2008.
13. T. Converse, "CAPTCHA Generation as a Web Service," Proc. 2nd Int'l Workshop Human Interactive Proofs", Springer, 2005, pp. 82-96.
14. 3D Captcha

AUTHORS PROFILE

L.Devi Priya, received B.TECH degree on Information Technology from Amrita University, Coimbatore, Tamilnadu, INDIA in 2011 and pursuing M.E Software Engineering in SNS College of Technology affiliated to Anna University, Chennai .Her Research includes in Information Security.

Dr.S.Karthik is presently Professor & Dean in the Department of Computer Science & Engineering, SNS College of Technology, affiliated to Anna University- Coimbatore, Tamilnadu, India. He received the M.E degree from the Anna University Chennai and Ph.D degree from Ann University of Technology, Coimbatore. His research interests include network security, web services and wireless systems. In particular, he is currently working in a research group developing new Internet security architectures and active defense systems against DDoS attacks.

Dr.S. Karthik, published more than 35 papers in refereed international journals and 25 papers in conferences and has been involved many international conferences as Technical Chair and tutorial presenter. He is an active member of IEEE, ISTE, IAENG, IACSIT and Indian Computer Society.