IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 2, Ver. I (Mar-Apr. 2014), PP 79-84
www.iosrjournals.org

Security Enhancement in Interactive CAPTCHA

Kumary R Soumya[1], Athira Sasikumar[2], Sancy P Prakash[2]
[1]*Asst. Professor, Department of Computer Science and Engineering Jyothi Engineering College, Cheruthuruthy, Thrissur, India*
[2]*Department of Computer Science and Engineering Jyothi Engineering College, Cheruthuruthy, Thrissur, India*

***Abstract**: CAPTCHA stands for Computer Automated Public Turing Test to Tell Computers and Humans Apart.It is a defense system against attacks on websites.The paper deals with a variant of conventional CAPTCHA named Interactive CAPTCHA which uses a series of user interactions to solve the CAPTCHA.The proposed system also encrypts the CAPTCHA response send to the server so as to avoid attacks like MITM(Man In The Middle attacks).The system uses a Secure Socket Layer(SSL) connection to transmit the encrypted response to the server providing additional security.Encryption and decryption of CAPTCHA response is done by the Rivest Shamir Adleman(RSA) algorithm using a encryption/decryption key pair.*
***Keywords:** CAPTCHA, Instant Messenger CAPTCHA Attack, RSA,SSL.*

I. Introduction

Now-a-days, websites became the identity for many businesses. Many companies have their sites for online business. These companies offer a lot of free services to their users. The only thing is to register on the site. So, many people exploit the offered free services by duplicate registration. They write a computer program that can automatically register on the site and then can use the offered services. CAPTCHA is developed for avoiding such spam registrations on the site and exploitation of the offered free services. CAPTCHA is a distorted image containing short text. It is displayed in such a format so that only human eyes can recognize the alphabets clearly. At the time of registration, such image is displayed on the form and the user is asked to write the same text in given text field. The robots fail to recognize the short text. Thus, website owners can prevent robots from registration and can ensure that all the members using free services are humans [3,4]. Recently the attacks on websites that employ CAPTCHAs by third party human solvers are gaining popularity in the malware industry. The third party human solvers employ technologies like Instant Messenger CAPTCHA Attack Systems (IMCA) to solve CAPTCHAs [2, 7].IMCA utilizes an attack script and IM Connector to solve the CAPTCHA. The CAPTCHA image is scraped from the victim website and its link is forwarded to IM provider via an IM Connector.The attacker clicks on the link to view the CAPTCHA image and responds through the IM Connector thus making false registrations into websites [5].

II. Interactive Captcha

In Interactive CAPTCHA technology [2], user is provided with a set of buttons to solve the displayed CAPTCHA image.Buttons are displayed with obfuscated characters. As the user clicks on the CAPTCHA image a set of buttons appear. One of the buttons will have the character corresponding to the first letter in the CAPTCHA image. User has to click on the correct button. On clicking the response next set of buttons are rendered with one button showing the next character in the CAPTCHA image. This input sequence is continued for each and every letter in the CAPTCHA image. In Interactive CAPTCHA timeout values are set for each button and not the entire CAPTCHA as with conventional CAPTCHA technologies. The time set is so small that a third party human solver who uses IMCA architecture cannot make false registrations into the website.

On the server side, session information is stored about the indices of the correct responses and the indices of the user clicks. When the input sequence is complete, the correct index sequence is compared with the user clicked index sequence. If there is a match, the CAPTCHA has been correctly decoded by the user.

The problem with Interactive CAPTCHA technology is that it does not include any mechanism to encrypt the response text send to the server for verification. This can lead to attacks like Man In The Middle attack (MITM).The attacker is able to intercept all messages going between the user and server and inject new ones .The attacker can eavesdrop the response text send from the user, modify it and submit the same to server impersonating the actual user. This leads to a situation where a user who has correctly decoded the CAPTCHA image is denied accessing the website. This problem can be solved by encryption using RSA and a Secure Socket Layer Connection.

III. Encryption Using RSA

Secure Socket Layer (SSL) is the most popular protocol used in the Internet for facilitating secure communications. The SSL protocol allows encrypted data exchange between the two parties using RSA [1].
Let (e, d) denote the encryption/decryption key-pair.

Let N = pq where p and q are large primes. Then, encrypted message C and plain-text message M are related as follows:

$$\mathbf{C = M^e \bmod N}$$

$$\mathbf{M = C^d \bmod N.}$$

Interactive CAPTCHA sets timeout for each button. This time depends on the kind of connection between sender and receiver. For high latency connections time out is to be set large.This is calculated using dynamic detection threshold algorithm.This algorithm calculates a dynamic threshold D by adding up Round Trip Time (RTT) and average time needed by the user to decode the CAPTCHA image (U_{avg}).

$$\mathbf{D=RTT+U_{avg}} \text{---------- (1)}$$

Round Trip Time (RTT) is calculated as follows: After the user clicks on a button to reply to the first character, the client side Ajax script [11] sends a request to the server for a new set of buttons for the second character. At this moment, the server responds with a ping command and the client side code responds with a pong. This interaction allows the server to calculate RTT.The calculated threshold D is set as the timeout value for each button in the Interactive CAPTCHA. With the introduction of encryption to existing Interactive CAPTCHA time to encrypt the response must also be incorporated into the dynamic threshold set for each button.So the equation for calculating threshold D is

$$\mathbf{D=RTT+U_{avg}+E_c} \text{---------- (2)}$$

E_c is the encryption time for each character in the CAPTCHA image. The key for encryption and decryption are generated during the load time of the website .A key of size 1024 bits is generated and considered enough to provide security in website logins.RSA decryption takes longer time than encryption. But this fact is not to be considered as the time out values for button does in no way depend on the response decryption time. The encryption time RSA is often calculated t to be in milliseconds.

IV. Proposed System

The proposed scheme enhances the security in Interactive CAPTCHA technology by encryption of CAPTCHA response. CAPTCHA is generated by splitting and rotation algorithm. CAPTCHA consist of both English alphabets and numbers withinrange.“ABDEFHKLMNPRSTUVWXZabdefgikmnopqrstuvwxyz0123456789". Initially any five characters are selected at random from this range. Each character has its own bending and size value. Characters are divided into several parts and given random rotation value in a certain angle domain interval. The generated CAPTCHA image is split into k*l matrix with k rows and l columns and the splits are inserted into an array. An interger value is randomized between –d and +d to give random rotation to splits. The proposed scheme selected the value of d to be 3.Each and every split of image is subject to a rotation. After rotation these images are combined to generate the CAPTCHA image.The Interactive CAPTCHA is loaded into the webpage on request to server. As the user clicks on the CAPTCHA image a set of buttons are rendere d. The buttons are dynamically generated. One of the buttons displays the first character of CAPTCHA image. User clicks on the button corresponding to the first character. As the user completes the first click another set of buttons are rendered of which one button has the character corresponding to the second character in the CAPTCHA image. The process is repeated for all characters in the CAPTCHA image as shown in Figure 1.

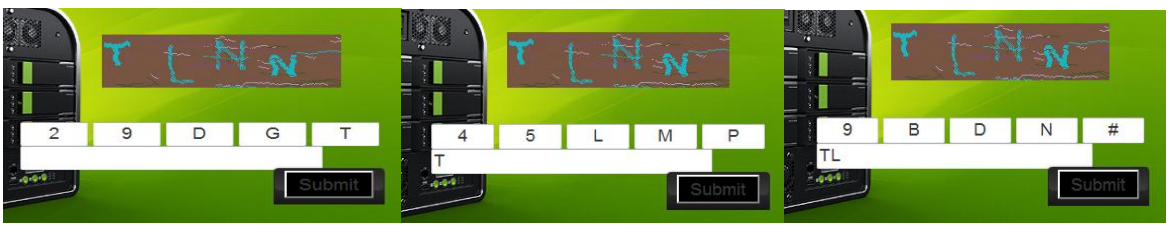

Figure 1.Operation of Interactive CAPTCHA

The CAPTCHA response is send to the server. The server checks if the user selected response matches the correct response sequence stored at the server. If there is a match the CAPTCHA has been correctly decoded and the user is allowed access to the website.

The response is send to the server in an unencrypted form .An attacker can eavesdrop the response and modify it before the response reaches the server. This leads to a situation where the user will be denied access to the website even if he/she has solved the CAPTCHA correctly and in the specified time out. To solve this problem we introduce the concept of a Secure Socket Layer connection which is to be established between the user and the server.SSL encrypts the information send by the user using RSA encryption.RSA encryption make use of a public-private key pair for encryption and decryption at sender and receiver sides respectively.

In RSA e, d denote the encryption/decryption key-pair. Two prime numbers p and q are selected. The value of N=p q. Encrypted message C and plain-text message M are related as follows: $C = M^e \bmod N$ and $M = C^d \bmod N$. The encryption / decryption key pair are generated from values of p and q. User sends the values of N and e to the server along with the HTTP request to the server for the webpage. Server stores the value of N and e. When the encrypted response reaches the server, it is decrypted using the key d. The value of d is calculated from the values of e and N send by the user.

In conventional CAPTCHA technology time out values of CAPTCHA is set for the entire CAPTCHA. This can be exploited by third party human solvers to make false registrations to websites using IMCA[6,8].Interactive CAPTCHA sets timeout for each button.This time out is calculated as the sum of RTT,U_{avg} and E_c as shown in equation 2.

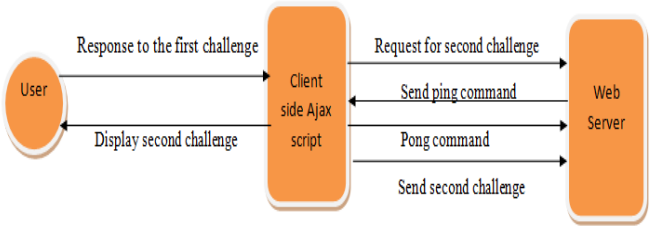

Figure 2.Calculation of Round Trip Time

RTT can be calculated by inserting an automatic round trip in the Ajax[9] code as shown in Figure 2.After the user clicks on a button to reply to the first character, the client side Ajax script sends a request to the server for a new set of buttons for the second character. At this moment, the server responds with a ping command and the client side code responds with a pong. This interaction allows the server to calculate *RTT*

At the server side the response is decrypted using RSA decryption key and it is checked over the correct index sequence stored at the server. If the decrypted response is correct user is allowed to access the website .Otherwise the user is denied access. Server checks the correctness of both the response and timing constraints of the user before denying or allowing access to website. The CAPTCHA image is destroyed after each session. The CAPTCHA image is destroyed when the session is completed. The security enhancement of Interactive CAPTCHA can be summarized by the Figure 3.

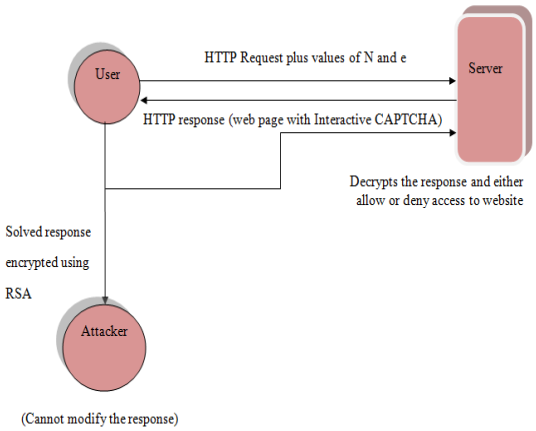

Figure 3.Security Enhancement in Interactive CAPTCHA

VI. PROTOTYPE IMPLEMENTATION

We implemented a prototype of Interactive CAPTCHA. The prototype consists of two components: A website with Interactive CAPTCHA and Apache Tomcat Server. The website was developed using JSP in NetBeans IDE.

Figure 4.Interactive CAPTCHA Website

The developed website displayed a four to six letter CAPTCHA image and a set of five buttons with letters to solve the CAPTCHA test as in Figure 4. An ordinary website as shown in Figure 5 was also developed in order to study the vulnerability of ordinary CAPTCHA to third party human attacks. An Instant Messenger CAPTCHA Attack (IMCA) prototype named CAPTCHA CRACKER as in Figure 6 was developed to study about efficiency of Interactive CAPTCHA website to defend third party human attackers. IMCA prototype was developed using JSP.

Figure 5.Ordinary CAPTCHA Website

Figure 6.CAPTCHA CRACKER (IMCA)

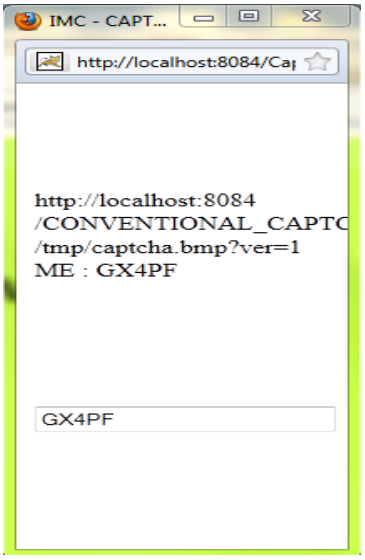

Figure 7.Human Solver's IM Client View

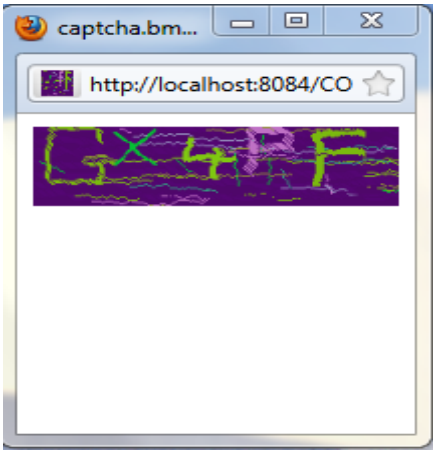

Figure 8. CAPTCHA provided to human solver by IMCA

Figures 7 and 8 show that the 3rd party human solver received a link to the delivered image, viewed the image, and responded with the CAPTCHA text.By exploiting the Instant Messenger infrastructure, IMCA has shown that CAPTCHAs can be delivered to 3rd party human solvers quickly and easily. IMCA delivers CAPTCHA images to 3rd party human solvers so efficiently that it renders the timeout values of ordinary CAPTCHA implementations useless. The response for the CAPTCHA image cannot be delivered to the Interactive CAPTCHA website using IMCA. This is because Interactive CAPTCHA sets the time out for buttons to be so small that IMCA cannot deliver the response within that time. Encryption of CAPTCHA response makes sure that MITM attacks cannot happen with Interactive CAPTCHA.

VII. Performance Evaluation

The performance of Interactive CAPTCHA depends on average time needed to solve the CAPTCHA. In addition to security, ease-of-use is critical for Interactive CAPTCHA to be a successful. It is anticipated that the total response time for an Interactive CAPTCHA would be slower than traditional CAPTCHA. The collected timing data shows that on average a user takes 8.08 seconds to solve a five character Interactive CAPTCHA

versus 6.21 seconds for a traditional CAPTCHA using the same image obfuscation style. When security concepts are added to Interactive CAPTCHA the average time is increased. This effect the timeout values that are intended to be set on buttons of Interactive CAPTCHA.The per character response time of an legitimate user on an average is considered to be 1.62 seconds. An attacker who uses a third party human solver like Instant Messenger CAPTCHA Attack takes on average 5.05 seconds. The encryption time is added to these values when time out for buttons is set. It is assumed that time for encryption is 0.09 seconds. Even though decryption time of RSA is several times more than encryption it has no role to play with button time out. So the button timing for each character is set approximately to 1.71 seconds as shown in Table I.

Table I:Per character response time

Interactive CAPTCHA	Per Character Response Time
With response encryption	1.71
Without response encryption	1.62

VIII. LIMITATIONS AND FUTURE WORKS

The Interactive CAPTCHA has some limitations

1) Users with improper vision: As Interactive CAPTCHA require user to solve the CAPTCHA within a specified time responses of users with improper vision are rejected. Further evaluation is needed to accommodate impaired users.
2) Renegotiation attacks: The SSL connection is vulnerable to renegotiation attacks where an attacker can hijack an https connection to splice its own requests into the beginning of the conversation between the client and web server. The attacker can't actually decrypt the client-server communication, so it is different from a typical man-in-the-middle attack

IX. Conclusion

CAPTCHA plays an important role in protecting internet websites.However these CAPTCHAs is vulnerable to third party human attacks. Most modern techniques employ the use of technologies like Instant Messenger CAPTCHA Attack system to defeat CAPTCHA system and make false registration into websites thus leading to resource wastage. The paper proposes a new kind of CAPTCHA technology that requires a set of user interactions. The multi-step back-and-forth traffic between client and server amplifies the statistical timing difference between a legitimate user and a human solver attack, and hence, provides a better attack detection performance. The scheme also introduces the concept of encrypting CAPTCHA response so as to avoid MITM attacks. Encryption is done using RSA and the technique uses a secure socket layer connection between user and the server.The encryption ensures that an attacker cannot alter the CAPTCHA response sent from the user before it reaches the server for verification. We hope that the proposed scheme equipped with encryption techniques will provide a more reliable CAPTCHA implementation.

References

[1] Krishna Kant, Ravishankar Iyer, Prasant Mohapatra, "Architectural Impact of Secure Socket Layer on Internet Servers: A Retrospect", published by IEEE, 2012, pp 25-26
[2] Huy D. Truong, Christopher F. Turner, Cliff C. Zou,"iCAPTCHA: The Next Generation of CAPTCHA Designed to Defend Against 3rd Party Human", published by IEEE, 2012.
[3] Jeff Yan, Ahmad Salah El Ahmad, "A Low-Cost Attack on a Microsoft CAPTCHA", Proceedings of the 15th ACM conference on Computer and communications security, 2008.
[4] Elias Athanasopoulos, Spyros Antonatos,"Enhanced CAPTCHAs: Using Animation to Tell Humans and Computers Apart". Communications and Multimedia Security 2006: 97-108
[5] Sam Hocevar, "PWNtcha – a Captcha Decoder Website",http://caca.zoy.org/wiki/PWNtcha
[6] Dancho Danchev, "Inside India's CAPTCHA Solving Economy",http://blogs.zdnet.com/security/?p=1835, 2008
[7] Albert E. Whale, "ABS computer technology, inc, Why the CAPTCHA Approach is Doomed", http://www.abscomptech.com/home/headlines/news/why-the-CAPTCHA-approach-isdoomed, 2009.
[8] Byron Acohido, "Cybergangs use cheap labor to break codes on social sites",http://www.usatoday.com /tech/news/computersecurity/ 2009-04-22-captcha-code-breakers_N.htm, 20092003
[9] J.J. Garrett, "Ajax: A New Approach to Web Applications", AdaptivePath Assays, 2005. http://www.adaptivepath.com/ideas/essays/archives/000385.php