Volume 5, Issue 6, June 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Security Online Authentication using Captcha

[1]Preetika
M.Tech Student
Jan Nayak Ch. Devilal Vidyapeeth
GJU, Hisar, Haryana, India

[2]Vikas Kamra
Assistant Professor
Jan Nayak Ch. Devilal Vidyapeeth
GJU, Hisar, Haryana, India

Abstract: ***CAPTCHA stands for Completely Automated Public Turing Tests to tell Computers and Humans Apart. Handwritten text offers challenges that are rarely encountered in machine printed text. In addition, most problem faced in reading machine printed textare more severe, in hand written text. A CAPTCHA is a program that protects websites against bots by generating and grading tests that human can pass but current computer programs cannot. A CAPTCHA is a program that generates and grades tests that are human solvable, but intend to be beyond the capabilities of current computer programs. The threats from bots are growing rapidly and also increasing in technical sophistication causing various types of attacks. These attacks are affecting social aspects of human beings. A comparative analysis over different set or different styles of captcha will be tested and suggested for secure online authentication.***

Keywords: Captcha, Threats, Authentication, Security,Web Application.

I. INTRODUCTION

In this modern age, Internet has a high influence in human living in every day's life. Most data are stored and retrieved from time to time over the network across organizations across the world. However, those Internet users are not always authorized persons. Thus, securing system was implemented using various methods. The most elementary one is password authentication system. Text passwords have been extensively used for user authentication. Captchaauthentication system is the oldest life-long serving for humans. Nevertheless, it is not sufficient to fulfilling the security level of organizations. Authentication is certainly at the heart of any secure system. Before a user can be involved in online transactions, enter a secured vault, open a safe or reaches his/her email account, he/she has to be authenticated

The entire security of one system will collapse, if sensitive information or unauthorized access is given to a wrong identity. The fundamental problem of using password is that the login password was hacked by malicious software and the password was exposed. As a result, this intrude program can access the data without permission by emulating human login process. Thus, CAPTCHA, Completely Automated Public Turing test to tell Computer and Humans Apart, was implemented in the year 2004 to protect this unwanted situation. Captcha are intended to permit a computer to determine if a remote client is human or not. CAPTCHA is a program implemented for differentiating between human and automate computer programs by producing questions that can be answered by humans only. Character recognition is a central problem in machine learning. In the perspective of captchasmaybe the most relevant work produced by the machine learning community is on the MNIST database of handwritten digits challenge which targets to recognize (distorted)handwritten digits. From this body of work, the most valuable articlefor captcha security research is which provide a deep analysison how to efficiently recognize digits.

There are malicious bot programs on the Internet (particularly those bots that can sign up for hundreds of accounts a minute with free email service providers, send out thousands of spam messages in an instant, or post plentiful comments in blogs pointing both readers and search engines to unrelated sites). Programs (bots and spiders) are being created to steal services and to conduct fraudulent transactions. E.g.

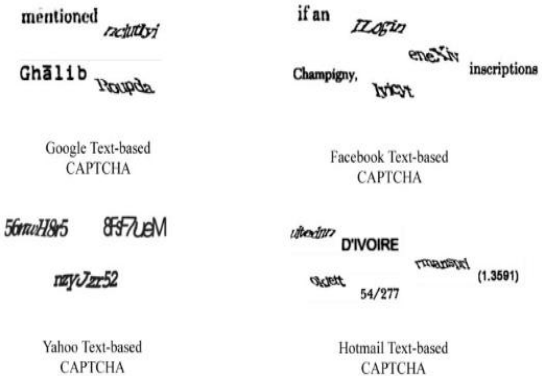

Fig1 Graphical Representation of Text Based Captcha.

1. Free online accounts are being registered automatically a lot of times and are being used to distribute stolen or copyrighted material.
2.Recommendation systems are susceptible to artificial inflation or deflation of rankings. For example, EBay, a famous auction website permits users to rate a product. Abusers can easily create bots that could increase or decrease the rating of a particular product, possibly changing people's opinion towards the product.
3.Spammers register themselves with free email accounts such as those provided by Gmail or Hotmail and use their bots to send unsolicited mails to other users of that email service.
There are some Characteristics of Captcha explained:
1. A good CAPTCHA must not only be human friendly but also robust enough so that it can resist computer programs that attackers write to automatically pass CAPTCHA tests.
2. Easy for human being to pass.
3.Easy for tester machine to generate and grade.
4.Difficult for a software robot to pass.

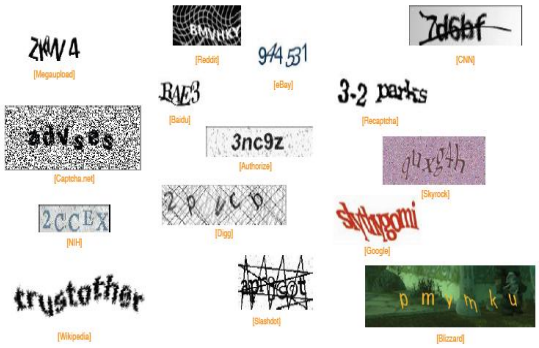

Fig2 Graphical Representation of Text Based Captcha

There are some Attacksclassify in four types includes denial-of-service, worms, DNS, and router attacks.
a. Denial-of-service
DOS attacks use multiple systems to attack one or more target systems with the intent of denying service to legitimate users of the target systems. The degree of automation in attack tools permits a single attacker to install heir tools and control tens of thousands of compromised systems for use in attacks. Intruders often search address blocks known to contain high concentrations of vulnerable systems. Cable modem, DSL, and university address blocks are gradually targeted by intruders planning to install their attack tools. Denial-of service attacks are effective because the Internet consists of limited and usable resources, and Internet security is highly interdependent.
b. Worms
A worm is self-propagating malicious code. Not like a virus, which needs a user to do something to continue the propagation, a worm can propagate by itself. The exceedingly automated nature of the worms coupled with the relatively widespread nature of the vulnerabilities they exploit, allows a large number of systems to be compromised within a matter of hours.
The Code Red worm infected more than 250,000 systems in just 9 hours on 19 July 2001.Certain worms, such as Code Red, include built-in denial-But the biggest effect of these worms is that their propagation effectively creates a denial of service in many parts of the Internet because of the vast amounts of scan traffic generated. Examples include DSL routers that crash and ISPs whose networks are completely burdened, not by the scanning itself but by the burst of underlying network management traffic that the scanning triggers.
c.DNS Attacks
The Domain Name System is the distributed, hierarchical global directory that translates names to numeric IP addresses. The upper two layers of the hierarchy are critical to the operation of the Internet. In the top layer are 13 root name servers. Next are the top-level domain (TLD) servers, which are authoritative for .com and .net, as well as for the country code top level domains (ccTLDs) such as .us, .uk, and so forth. Several threats to DNS include cache poisoning, compromised data, denial of service, and domain hijacking.
d. Cache poisoning
If DNS is made to cache bogus information, the attacker can send traffic intended for a legitimate site to a site in the attacker's control. A latest survey by the CERT Coordination Centerdemonstrates that over 80precents of the TLDs are running on servers that are potentially vulnerable to this form of attack.
e. Compromised data
Attackers compromise vulnerable DNS servers, giving them the chance to modify the data served to users. Several of the TLD servers run a software program called BIND, in which vulnerabilities are discovered regularly. A CERT Coordination Center survey indicates that at least 20 percent of TLDs are running on vulnerable servers.
f. Domain hijacking
By leveraging insecure mechanisms used by users to update their domain registration information, attackers can capture the domain registration processes to hijack legitimate domains. These attacks are affecting social facets of human beings. Human being or legitimate users are cheated by false requests troubling their social life. This paper aims to provide secured authentication by verifying strength of CAPTCHA.

reCAPTCHA is one of the most extensively used CAPTCHAs on websites. Up to this date, it is based on the word recognition problem. Words from scanned books and newspapers are used, most of them are older so that they are subject to an Aging process that has degraded smudged and distorted the words. They can also be misaligned by the scanning process and could be printed in a variety of typefaces of which many could be infrequently used today. The words used for the test have characters that are mostly not separated at all or leave very thin spaces between the characters. In addition to this, they are also distorted artificially to make the AI-problem of identifying these words even tougher. Theuser proves that he is human,By typing in two of this distorted words correctly Figure 3 shows a screenshot of an example CAPTCHA challenge from reCAPTCHA. The CAPTCHAs used by reCAPTCHA change from time to time, thescreenshot was taken in early June 2010.

Fig3Example for a CAPTCHA challenge from reCAPTCHA

II. CLASSIFICATIONS OF CAPTCHA

The CAPTCHA techniques can be classified in three main categories

1. Text-Based

They usually rely on sophisticated distortion of text images rendering them unrecognizable to the state of the art of pattern recognition programs but identifiable to human eyes.

2. Image-Based

Image-based CAPTCHA can be used that stops spam and malicious bots, while remaining easy for people to answer. All your website visitors have to do is click a few images to prove they are human and not a bot. It takes just some seconds to solve.

3. Sound-Based

For the sound-based, this technique is developed for a definite group who has eye vision problem. The image based captcha has been demonstrated as:

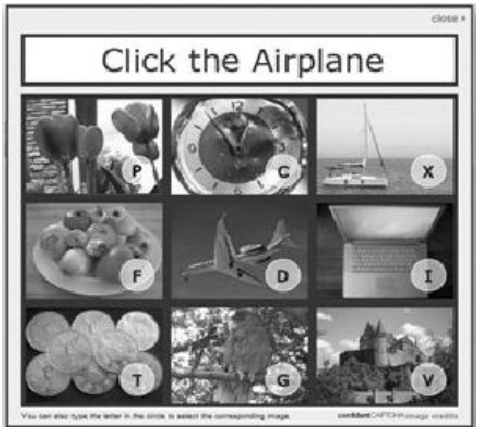

Fig3 Graphical Representation of Text Based Captcha

Each type is appropriate to serve different groups of users. Usually, the text-based mechanism is the easiest and simplest to be implemented but it cannot support blind users or people who have difficulty in reading, mean while, the image-based is not suitable for persons who have skill of thinking or interpreting the image to answer questions.

III. LITERATURE REVIEW

CAPTCHA stands for Completely Automated Public Turing Tests to tell computers and Humans Apart .A CAPTCHA is a program that defends websites against bots –automated scripts by producing and grading tests that humans can pass but current computer programs cannot. For example, current computer programs can't read distorted text but humans can. This paper aims to provide secured authentication by verifying strength of CAPTCHA [1].

CAPTCHAs are also called Human Interaction Proofs (HIPs) in literature. Initially, a CAPTCHA implied a system for which the generator is public (for example it is open source), as the "P" in CAPTCHA stands for public, whereas aHIP does not have this notion. However, now days the terms CAPTCHA and HIP are used as synonyms. The term CAPTCHA is the desired one in new publications and is well established. It is now also used for systems without a public generator.

In this paper we present a novel CAPTCHA that is founded on the current hard AI problem of mixed-text (handwriting and printed-text) segmentation. The proposed CAPTCHA overlays generated handwritten word images on a generated printed-text background. We first propose a modification that allows for character level perturbations on apresent synthetic handwriting generation technique. These perturbations are parameterized permitting for varying levels of

handwritten word complexity. We then use the output from the changed synthetic handwriting generator as the foreground for the mixed-text CAPTCHA. Experiments prove that the proposed approach is effective at effectively distinguishing between humans and machines [2].It has been explained by the author that the CAPTCHA provide a method for automatically distinguishing a human from a computer program, and therefore can protect Web services from abuse by so-called bots. Most CAPTCHA consist of distorted images, usually text, for which a user must provide some description. Unfortunately, visual CAPTCHA limit access to the millions of visually challenged people using the Web. The Audio/Voice based CAPTCHA was created to solve this accessibility issue; however, the security of Audio based CAPTCHA was never formally tested [3].

In this paper we perform a systematic study of existing visual CAPTCHAs based on distorted characters and contribute towards improving the systematic evaluation and design of visual CAPTCHA. We identify a series of recommendations for CAPTCHA designers and attackers, and possible future guidelines for producing more reliable human/computer distinguishers [4].

Author explained that the CAPTCHA implementation is tricky and risky without careful design. In this paper, they gave a study case of the vulnerabilities in current login website using text-based CAPTCHA. Their targeteda website of mainstream bank of china and shown that with some specialized methods, the CAPTCHA scheme in its website can be easily cracked. Finally, they gave some recommendations for CAPTCHA designers to revise our CAPTCHA implementation security in the future.

reCAPTCHAs website reports that over 30 million of such challenges are served every day (as of July 2010). Theintegration of reCAPTCHA into a website is cost free and packages for various web programming languages exist, making itstress-free for webmasters to adopt the technology. It is widely deployed and the reCAPTCHAs website currently states thatover 100.000 websites are using reCAPTCHA. Very popular websites like Facebook, Twitter and StumbleUpon are using

reCAPTCHA since at least 2007

Because reCAPTCHA is a very popular CAPTCHA, as outlined above, it is also anmotivating target for an academicsecurity analysis. The central question in analysing the security of a CAPTCHA is whether it is possible to build anautomated software solver that can solve a non-trivial fraction of the challenges, thereby undermining the premise thatthe CAPTCHA is hard to solve by computers.

IV. PROBLEM STATEMENT

The web application provides the services to the user such emailaccount Creation, Video Based Servicesetc but security of the web service and database need to be more secure.

The automatic scripts generate the overflows at application level as well as database level. There is needed to be identification of the human and automatic scripts. The OCR tool recognizes the CAPTCHA which is used to authentication over the web. The major issue is of distortion of characters.

The CAPTCHA should easily recognize by the human and should be secure which cannot be understandable by the automatic scripts. So the Strengthening the CAPTCHA need to design. The threats from bots are growing gradually and also increasing in technical sophistication causing various types of attacks. These attacks are affecting social aspects of human beings.

V. OBJECTIVES

In research of Captcha, the security dproblems need to be measure and different existing methods can provide the better techniques. The main purpose is of designing approach of Captcha which can be easily understandable and secure from automatic scripts. The different parameters need to be considered for security purpose.

a.Analyze the different parameters for design the Captcha.

b.Design text based Captcha.

c.Implement on Characters with rotation on Characters.

d.Analyze Spacing decision between characters.

e.Dynamic length of Characters for Captcha Generation.

VI. PROPOSED METHODOLOGY

The different steps need to consider for design the security of web application. This section provides the steps to implement the proposed work. The mixture of random character and numbers will be used to generate new captcha.

a.Study current techniques of Web Security including Captcha.

b. Research on these Techniques to identify the issues.

c. Flow Development of new proposed technique.

d.Implementation in any Language.

e.Generate Results.

VII. CONCLUSION AND FUTURE WORK

In this paper, we have been proposed the Captcha based security algorithm and methods which can be used for authentication and provide the protection from automatic scripts and provide the secure communication of the information over the inter-network and intra-network. The proposed is not implemented yet.

The implementation part will be covered in the next paper, which will demonstrate the real working of proposed algorithm.

REFERNCES

[1] Prof. A.A Chandavle, Dr. A.M. Sapkal (2010), "Algorithm for secured online authentication usingCaptcha".

[2] Achint O Thomas, SulabhChoudhary, VenuGovindaraju (2010), "Leveraging the Mixed –Text Segmentation problem to design secure Handwritten Captcha".

[3] Chandavale, A. ;Sapkal, A. (2011), "An Improved Adaptive Noise Reduction for Secured CAPTCHA", Emerging Trends in Engineering and Technology (ICETET), 2011, Page(s): 12 – 17.

[4] ElieBursztein, Matthieu Martin, John C. Mitchell (2011), "Text-based CAPTCHA Strengths and Weaknesses".

[5] Xiao Ling-Zi ; Zhang Yi-Chun (2012), "A Case Study of Text-Based CAPTCHA Attacks", Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), Page(s): 121 – 124

[6] Tamang, T., Bhattarakosol, P.(2012), "Uncover impact factors of text-based CAPTCHA identification", Page(s): 556 – 560.

[7] Xiao Ling-Zi, Zhang Yi-Chun (2012), "A case study of Text-Based Captcha attacks".

[8] Wei-Bin Lee (2012), "A CAPTCHA with Tips Related to Alphabets Upper or Lower Case Broadband", IEEE, 12-14 Nov, Page(s):458 – 461

