www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 5 Issue 8 Aug, 2016 Page No. 17359-17363 DOI: **10.18535/ijecs/v5i8.01**

Understanding Captcha: with its types,for security

Sonali S. pawar[1], Prof. Pravin P. Kalyankar[2]

[1]T.P.C.T.'s College of Engineering ,Dr.B.A.M.University, Aurangabad.
Solapur-Osmanabad Road, Osmanabad. 413501
Sonali.pawar052@gmail.com

[2]T.P.C.T.'s College of Engineering ,Dr.B.A.M.University, Aurangabad.
Solapur-Osmanabad Road, Osmanabad. 413501
kalyankarpravin@rediffmail.com

Abstract: ***CAPTCHAs are short for Completely Automated Public Turing test to tell Computer and Humans Apart.This test is taken to insure that the user is a human being and not the machine or any program.It contains different types of tests which can only be solved by the humans.this is because to avoid bad interface of machines in an application. Captcha as graphical passwords (CaRP) is one of the new security built on Captcha technology. As its name implies that CaRP is the combination of both CAPTCHA and Graphical password scheme. CaRP addresses a number of security problems altogether, it offers reasonable security and usability with some practical applications for improving online security***.

Keywords: CaRP, gimpy,CAPTCHA,Pix,bongo, graphical password**.**

Abstract

To create cryptographic primitives based on hard mathematical problems is a fundamental task in security which are computationally intractable.Captcha is one of the security primitive based on hard AI (Artificial Intelligence) problems. Captcha, which distinguishes human users from computers by presenting a challenge, i.e., a puzzle, which can be only solved by the humans and not by computers ,beyond the capability of computers but easy for humans.

1. INTRODUCTION

A novel family of graphical password systems integrating Captcha technology, which we call *CaRP (Captcha as gRaphical Passwords)*. CaRP is click-based graphical passwords, where a password is derived from the sequence of clicks on an image. Unlike other click-based graphical passwords, images used in CaRP are Captcha challenges. For every login attempt a new CaRP image is generated. Various online services which have threat of major security uses CaRP which offers protection against online dictionary attacks on passwords.This threat is widespread and considered as a top cyber security risk. Defense against online dictionary attacks is a more subtle problem than it might appear. It is now a standard Internet security technique to protect online email and other services from being abused by bots or by automated softwares.

Computers are not as intelligent as humans. Machines have lack the ability to process on visual data. This is because Computers lack the ―Real intelligence‖.Captcha makes the use of this thing and provides a visual test to the user or human. In this password scheme at the time of implementation the user is presented with an image which contains some pattern. This pattern contains distorted or randomly stretched characters which only humans should be able to identify.

1.1 Background

The CAPTCHA is needed to keep out the website/search engine abuse by bots. In 1997, AltaVista sought ways to block and discourage the automatic submissions of URLs into their search engines. Andrei Broder, Chief Scientist of AltaVista, and his colleagues developed a filter. Their method was to generate a printed text randomly that only humans could read and not machine readers. Their approach was so effective that were reduced by 95% and a patent was issued in 2001. In 2000, Yahoo_s popular Messenger chat service was hit by bots which pointed advertising links to annoying human users of chat rooms. Yahoo, along with Carnegie Mellon University, developed a CAPTCHA called EZ-GIMPY, which chose a dictionary word randomly and distorted it with a wide variety of image occlusions and asked the user to input the distorted word. This incident created the urge to use CAPTCHAs for such online polls to ensure that only human users are able to take part in the polls.

2. Types of CAPTCHA

As in CAPTCHAs challenges are presented to the user, Captchas are classified based on what is distorted and presented as a challenge to the user. They are:

- Text based Captcha
- Graphic Captcha

2.1 Text CAPTCHAs

This type of Captcha is generally used ,they are simple to implement. In this approach the user is presented with some questions which only a human user can solve. Examples of such questions are:

1. What are ten minus three?
2. What is the third letter in WATERMELON?
3. Which of Yellow, Thursday and Richard is a colour?
4. If yesterday was a Thursday, what is today?

Such questions are very easy for a human user to solve,as logical thinking is required to solve these which can only be done by humans but it_s very difficult to program a computer to solve them.

These are also friendly to people with visual disability – such as those with colour blindness. Other text CAPTCHAs involves text distortions and the user is asked to identify the text hidden. The various implementations are:

2.1.1 Gimpy

It is a very reliable text CAPTCHA. Gimpy is based on the human has the ability to read extremely distorted text otherwhile the computer programs don‘t have such type of ability to do the same. Gimpy works by choosing ten words randomly from a dictionary, and displaying them in a distorted and overlapped manner. Gimpy then asks the users to enter a subset of the words in the image. The human user is capable of identifying the words correctly, whereas a computer program cannot.

Figure 1: Gimpy Captcha.

2.1.2 Ez Gimpy

This is a extension and simplified version of the Gimpy CAPTCHA, adopted by Yahoo in their signup page. Ez – Gimpy is same as Gimpy in the sense that randomly picks a single word from a dictionary and applies distortion to the text. The user is then asked to identify the text correctly.

Figure 2: Ez-Gimpy Captcha

2.1.3 Baffle Text

This was developed by Henry Baird at University of California at Berkeley. This is a variation of the Gimpy. This doesn_t contain dictionary words, but it picks up random alphabets to create a nonsense but pronounceable text. Distortions are then added to this text and the user is challenged to guess the right word. This technique overcomes the drawback of Gimpy CAPTCHA because, Gimpy uses dictionary words and hence, clever bots could be designed to check the dictionary for the matching word by brute-force.

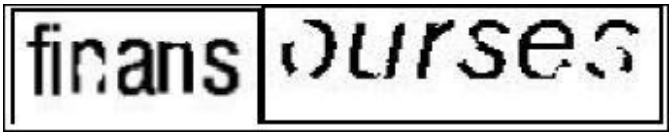

Figure 3 : Baffle Text The Audio Captcha is based on sound. The program picks a word or a sequence of numbers at random, renders the word or the numbers into a sound clip and distorts the sound clip; it then presents the distorted sound clip to the user and

2.2 Graphic CAPTCHAs

Graphic CAPTCHAs are different from text based captchas.Graphic Captcha provides challenges that involve pictures or objects that have some sort of similarity that the users have to guess. They are visual puzzles. Computer generates the puzzles and grades the answers, but is itself unable to solve it. CAPTCHA that requires two steps to be passed. First step visitor clicks elsewhere on the picture that composed of a few images and selects in this way a single image.Second step the selected image is loaded. It is enlarged but much distorted. Also variants of the answer are loaded on the client side. The visitor should select a correct answer from the set of the proposed words.

Figure 4: Graphic captcha

Follwing are some of the types of Graphic Captcha. These Captchas includes some sort of pictures or objects with some properties or characteristics that the user have to guess.

2.2.1 PIX

PIX is a program that has a large database of labeled images. All of these images are pictures of concrete objects (a horse, a table, a house, a flower). The program picks an object at random, finds six images of that object from its database, presents them to the user and then asks the question ―what are these pictures of?‖ Current computer programs should not be able to answer this question, so PIX should be a CAPTCHA.

2.2.2 BONGO

BONGO asks the user to solve a visual pattern recognition problem. It displays two series of blocks, the left and the right. The blocks in the left series differ from those in the right, and the user must find the characteristic that sets them apart.

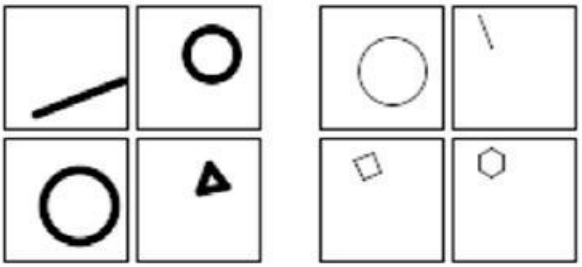

Figure 5 :Bongo Captcha

2.3 AUDIO CAPTCHA

asks users to enter its contents. This CAPTCHA is based on the difference in ability between humans and computers in recognizing spoken language. The idea is that a human is able to efficiently disregard the distortion and interpret the characters being read out while software would struggle with the distortion being applied, and need to be effective at speech to text translation in order to be successful.

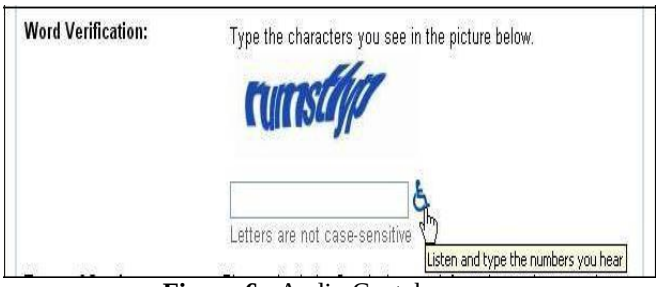

Figure 6 : Audio Captcha

2.4 Challenge in Breaking Captcha

The challenge in breaking the Captcha is really a hard task. It is hard because it is impossible to teach a computer to be think like humans, how to process information in a way similar to how humans think. To make the Computers think like humans ,Algorithms with Artificial intelligence will have to be designed when it comes to recognize the patterns in Captcha images. However, there is no universal algorithm that could pass through and break any Captcha system. Thus, each Captcha algorithm must have to be Tackled individually.

3. Methods in Captcha

3.1 Working of Captcha

Basically Captcha work in following manner

1.Create Random value which are often hard to guess and predict.

2. Generate an image on the basis of the random value created. As text in the images are harder to identify by computers, the image is generated.

3. Different image formation techniques are employed by developer so that it becomes difficult to identify. Some create zig zag lines,insert arcsfor background or twist-and-turn individual character in the pattern.

4.The random string which is generated in first step is stored for matching the user input. Generally session variables are used for storing the patterns because we needed to preserve stored values from one page to another.

5.After storing the pattern next step is to present the user with an image captcha.It may appear on the registration or sign-up form which we want to protect from being abused.

6. The user fills in the form along with the Captcha text and submits it. Now we have the following data,All submitted data ,Captcha string input by user,Captcha string (real one) from session variable.

7. If both Captchas match then it is confirm that the user is human and not the computer, allow to next process. If match is wrong then access is denied.

3.2 Data Flow Diagram

The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of the input data to the system, various processing carried out on these data, and the output data is generated by the system.

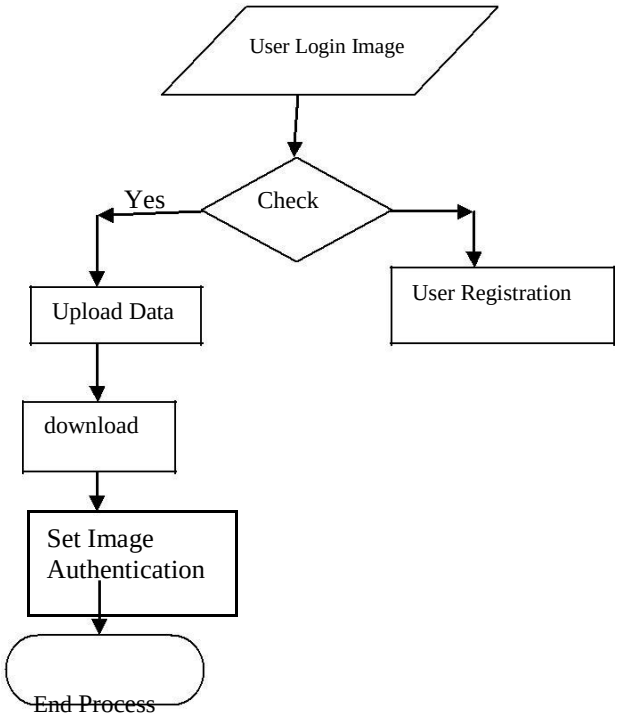

Figure 7. DFD for User.

4. Implementation

The most critical stage in achieving a successful new system and telling the user that the new system will work and be effective. Implementation is the stage of the project when the theoretical design is turned out into a working system.

The implementation stage involves investigation of the existing system and it‘s constraints on implementation, careful planning, designing of methods to achieve changeover and evaluation of changeover methods.

Main Modules for Captcha are:

1. **Graphical Password :**

In this module, Users are having authentication and security to access the detail which is presented in the Image system. Before accessing or searching the details user should have the account in that otherwise they should register first.

2. **Captica in Authentication:**

It was introduced in [14] to use both Captcha and password in a user authentication protocol, which we call *Captcha-based Password Authentication (CbPA) protocol*, to counter online dictionary attacks. The CbPA-protocol in requires solving a Captcha challenge after inputting a valid pair of user ID and password unless a valid browser cookie is received. For an invalid pair of user ID and password, the user has a certain probability to solve a Captcha challenge before being denied access.

3. Thwart Guessing Attacks :

In a guessing attack, a password guess tested in an unsuccessful trial is determined wrong and excluded from subsequent trials. The number of undetermined password guesses decreases with more trials, leading to a better chance of finding the password. To counter guessing attacks,

traditional approaches in designing graphical passwords aim at increasing the effective password space to make passwords harder to guess and thus require more trials. No matter how secure a graphical password scheme is, the password can always be found by a brute force attack. In this paper, we distinguish two types of guessing

attacks: *automatic guessing attacks* apply an automatic trial and error process but *S* can be manually constructed whereas *human guessing attacks* apply a manual trial and error process.

5. Applications

5.1 Secure Website Registration.

As described in above section,Several companies like Yahoo!, Microsoft, etc. offer free email services. Previously, most of these services suffered from attacks called as bots that would sign up for thousands of email accounts every minute. Captchas provides the solution to this problem to ensure that only humans can create their accounts and obtain free accounts. In general, free services should be protected with a CAPTCHA in order to prevent abuse by automated softwares.

5.2 Protecting Email Addresses From Scrapers.

Captchas provides the facility in which you can hide your email address from web scrapers. The idea is to require users to solve a CAPTCHA before showing your email address. It is an effective mechanism to protect your email address and itd abuse.

5.3 Online Polls.

Now a days ,many reality programs are taking their decisions depending on the audience choice. For this reason their votes are collected online. As is the case with most online polls, IP addresses of voters were recorded in order to prevent single users from voting more than once. However, some of people found a way to stuff the ballots using programs that voted for one thousands of times. One of them score started growing rapidly. The next day, another person wrote their own program and the poll became a contest between voting one person and another one. Can the result of any online poll be trusted? Not unless the poll ensures that only humans can vote.

5.4 Dictionary Attacks.

In general password system like text based passwords Dictionary attacks are made to guess the password. CAPTCHAs are used to prevent dictionary attacks in password systems. The idea is simple: prevent a computer from being able to iterate through the entire space of passwords by requiring it to solve a CAPTCHA after a certain number of unsuccessful logins. This is better than the classic approach of locking an account after a sequence of unsuccessful logins, since doing so allows an attacker to lock accounts at will.

5.5 Preventing Search Engine from Bots

Indexed webpages can found easily, however It is sometimes desirable to keep webpages unindexed to prevent others from finding them easily. Search engines bots can be prevented from reading web pages by an html tag. It may work sometimes but not sure that bots won't read a web page. Search engine bots, usually belong to large companies, respect web pages that don't want to allow them in. In this case CAPTCHAs are needed to guarantee that bots won't enter a web site.

5.6 Worms and Spam.

CAPTCHAs also offer a plausible solution against email worms and spam. It means that they will accept the emails only if the email is sent bu human and not by any automated software. This idea is now used by many companies.

6 . Conclusion

CaRP also offers protection against relay attacks, an increasing threat to bypass Captchas protection. Captcha can be circumvented through relay attacks whereby Captcha challenges are relayed to human solvers, whose answers are fed back to the targeted application. Our future work concentrates on improving the login time and memorability.

It offers reasonable security and usability and appears to fit well with some practical applications for improving online security.

- This threat is widespread and considered as a top cyber security risk. Defense against online dictionary attacks is a more subtle problem thanit might appear.
- This paradigm has achieved just a limited success as compared with the cryptographic primitives based on hard math problems and their wide applications.
- Using hard AI (Artificial Intelligence) problems for security, initially proposed in [17], is an exciting new paradigm. Under this paradigm, the most notable primitive invented is Captcha, which distinguishes human users from computers by presenting a challenge.

References

[1] R. Biddle, S. Chiasson, and P. C. van Oorschot, ―Graphical passwords: Learning from the first twelve years,‖ *ACM Comput. Surveys*, vol. 44, no. 4, 2012.

[2] Bin B. Zhu, Jeff Yan, GuanboBao, Maowei Yang, and NingXu_, ―Captcha as graphical Passwords—A New Security Primitive Based on Hard AI Problems‖IEEE RANSACTION SON INFORMATION ORENSICS AND SECURITY, VOL. 9, NO. 6, JUNE 2014.

[3] R. Dhamija and A. Perrig, Deja Vu: A User Study Using Images for Authentication. In the 9th USENIX Security Symposium, 2000.

[4] J. Yan and A. S. El Ahmad. Usability of CAPTCHAs or usability issues in CAPTCHA design. In SOUPS _08, pages 44–52, New York, NY, USA, 2008..ACM.

[5] Greg Mori and Jitendra Malik.Breaking a Visual CAPTCHA. Unpublished Manuscript, 2002.

[6] P. C. van Oorschot and J. Thorpe, ―On predictive models and userdrawn graphical passwords,‖ *ACM Trans. Inf. Syst. Security*, vol. 10, no. 4, pp. 1–33, 2008.

[7] K. Golofit, ―Click passwords under investigation,‖ in *Proc. ESORICS*, 2007, pp. 343–358.

[8] A. E. Dirik, N. Memon, and J.-C. Birget, ―Modeling user choice in the passpoints graphical password scheme,‖ in *Proc. Symp. Usable Privacy Security*, 2007, pp. 20–28.

[9] J. Thorpe and P. C. van Oorschot, ―Human-seeded attacks and exploiting hot spots in graphical passwords,‖ in *Proc. USENIX Security*, 2007, pp. 103–118.

Author Profile

Sonali S. Pawar received the B.E. degree in Computer Science and Engineering from Terana Public Charitable Trust's College Of Engineering, Osmanabad 2011 and pursuing M.E. in Computer Science and Engineering from the same Institution.