Implementation of Colored CAPTCHA

Mandeep Kumar
Assistant Professor,
Dept of Computer Science
S.R.P.A Adarsh Bhartiya College, Pathankot

Dr. Renu Dhir
Associate Professor,
Dept of Computer Science
NIT- Jalandhar

ABSTRACT

CAPTCHAs is an interesting of research are short for **C**ompletely **A**utomated **P**ublic **T**uring test to tell **C**omputers and **H**umans **A**part. The purpose of a CAPTCHA is to block form submissions from spam bots –automated scripts that harvest email addresses from publicly available web forms. We introduce a new CAPTCHA which is based on identifying the color of image, object or background. This actually requires analysis of complex contents of image, which humans usually performs well but machines or robots usually do not. From a large database of images, we select many images which having single color. The main advantages of our CAPTCHA technique over the traditional text-based techniques are that it is language-independent, does not require text-entry.

General Terms

Authentication, Security, Algorithms, Human Ease, Human Interaction Proof

Keywords

CAPTCHA, Image Processing, Spam, Automated Attacks, Character Recognition, Visual Processing, Automated program.

1. INTRODUCTION

CAPTCHAs is an interesting of research are short for **C**ompletely **A**utomated **P**ublic **T**uring test to tell **C**omputers and **H**umans **A**part. The term "CAPTCHA" was coined in 2000 by Luis Von Ahn, Manuel Blum, Nicholas J. Hopper (all of Carnegie Mellon University, and John Langford (then of IBM). They are challenge-response tests to ensure that the users are indeed human. The purpose of a CAPTCHA is to block form submissions from spam bots –automated scripts that harvest email addresses from publicly available web forms. A common kind of CAPTCHA used on most websites requires the users to enter the string of characters that appear in a distorted form on the screen.

Use of INTERNET has remarkably increased **Globally** in the past 10-12 years and so is the need of the **Security** over it. Marketing and Advertisement over INTERNET has seen companies like GOOGLE being made, which at the moment is traded at **181 billion USD** ie. Almost twice of General Motors, McDonalds combined. With an increasing number of free services on the internet, we find a pronounced need to protect these services from abuse. Automated programs (often referred to as bots) have been designed to attack a variety of services. For example, attacks are common on free email providers to acquire accounts. Nefarious bots use these accounts to send spam emails, to post spam and advertisements on discussion boards, and to skew results of online polls.

You're trying to sign up for a free email service offered by Gmail or Yahoo. Before you can submit your application, you first have to pass a test. It's not a hard test -- in fact, that's the point. For you, the test should be simple and straightforward. But for a computer, the test should be almost impossible to solve. This sort of test is a **CAPTCHA**. They're also known as a type of **Human Interaction Proof**(**HIP**). You've probably seen CAPTCHA tests on lots of Web sites. The most common form of CAPTCHA is an image of several distorted letters. It's your job to type the correct series of letters into a form. If your letters match the ones in the distorted image, you pass the test.

1.1 The use of CAPTCHA

The proliferation of the publicly available services on the Web is a boon for the community at large. But unfortunately it has invited new and novel abuses. Programs (bot sand spiders) are being created to steal services and to conduct fraudulent transactions. Some examples:

• Free online accounts are being registered automatically many times and are being used to distribute stolen or copyrighted material.

•Recommendation systems are vulnerable to artificial inflation or deflation of rankings. For example, EBay, a famous auction website allows users to rate a product. Abusers can easily create bots that could increase or decrease the rating of a specific product, possibly changing people's perception towards the product.

•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.

•Online polls are attacked by bots and are susceptible to ballot stuffing. This gives Unfair mileage to those that benefit from it.

In light of the above listed abuses and much more, a need was felt for a facility that checks users and allows access to services to only human users. It was in this direction that such a tool like CAPTCHA was created.

2. LITERATURE SURVEY

1. Today, approximately all the internet users have login accounts for internet sites and these sites require only the registration by human users but unfortunately some automated computer programs to enter these sites and use their resources through false registration. This paper introduces a new TIME-VARIENT CAPTCHA. In this paper we are not focusing on the effective development of Captcha but targeting a display of Captcha over the

webpage for a fixed time, Captcha replaces itself until the final Captcha is filled by user. Refresh process just work with Captcha and don't affect the web page. So, now, automated program has to cover one more area to breach the Captcha: to determine the final entered Captcha.

2. In this second paper to avoid tremendous attack from malicious computer programs, CAPTCHA (Completely Automated Public Turing test to tell Computers and Human Apart) mechanism has been introduced to distinguish humans and computers. They are used to protect various kinds of online services from advertising spam, brute force attacks and denial of service by automatic computer programs. In general the present CAPTCHAS are 2D. Due to the fast development of pattern recognition and artificial intelligence technology, there are increasing safety loopholes concerning traditional 2D static CAPTCHAs, resulting in that certain malicious computer programs could launch serious program attack through breaking such CAPTCHA. So in our project we propose a practical and safe 3-layer dynamic CAPTCHA which is very hard to break and which prevent the attack from malicious computer program. The 3-layered dynamic CAPTCHA can be implemented by using the "layered" concept. Three layers are: Character Layer, Background Interference Layer and Foreground Interference Layer.
3. Atomizing various Web activities by replacing human to human interactions on the Internet has been made indispensable due to its enormous growth. However, bots also known as Web-bots which have a malicious intend and pretending to be humans pose a severe threat to various services on the Internet that implicitly assume a human interaction. Accordingly, Web service providers before allowing access to such services use various Human Interaction Proof's (HIPs) to authenticate that the user is a human and not a bot. Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) is a class of HIPs tests and are based on Artificial Intelligence. These tests are easier for humans to qualify and tough for bots to simulate. Several Web services use CAPTCHAs as a defensive mechanism against automated Web-bots. In this paper, we review the existing CAPTCHA schemes that have been proposed or are being used to protect various Web services. We classify them in groups and compare them with each other in terms of security and usability. We present general method used to generate and break text-based and image-based CAPTCHAs. Further, we discuss various security and usability issues in CAPTCHA design and provide guidelines for improving their robustness and usability.
4. The massive and automated access to Web resources through robots has made it essential for Web service providers to make some conclusion about whether the "user" is a human or a robot. A Human Interaction Proof (HIP) like Completely Automated Public Turing test to tell computers and Humans Apart (CAPTCHA) offers a way to make such a distinction. CAPTCHA is a reverse Turing test used by Web service providers to secure human interaction assumed services from Web bots. Several Web services that include and are not limited to free e-mail accounts, online polls, chat rooms, search engines, blogs, password systems, etc. use CAPTCHA as a defensive mechanism against automated Web bots. In this paper, we present a new clickable image-based CAPTCHA technique. The technique presents user with a CAPTCHA image composed of several sub-images. Properties of the proposed technique offer all of the benefits of image-based CAPTCHAs; grant improved security than that of usual OCR-based techniques, consume less Web page area than most of image-based techniques and at the same time improve the user-friendliness of the Web page.
5. CAPTCHAs are employed on web systems to differentiate between human users and automated programs which indulge in spamming and other fraudulent activities. CAPTCHAs currently in use have been broken and rendered ineffective as a result of continuous evolution in CAPTCHA breaking. Thus, there is a need to employ stronger CAPTCHAs to keep these breaking attacks at bay while retaining ease of implementation on websites and ease of use for humans. In this paper, we introduce Sequenced Picture Captcha (SPC). Each CAPTCHA round comprises of object pictures, each of which may be accompanied by a Tag. The user is required to determine the logical sequence of the displayed object pictures based on the Tags. We identify two generation schemes - one in which object pictures indicate an inherent sequencing and one in which explicit Tags are displayed for determining the sequencing. We also analyze all these schemes. The advantages of high user convenience and simplicity of operation are retained in both generation types.
6. A CAPTCHA is a program that can tell whether its user is a human or a computer. CAPTCHAs are used by many websites to differentiate between the bots (automated program) and the human. The main motive behind using the CAPTCHA is to prevent abuse from "bots," or automated programs usually written to generate spam. The CAPTCHA basically uses images in distorted form that is difficult to read by the bots but they can be read by the human being easily. The websites which provide free services to the users some anonymous user can make false enrollment on the websites with the help of automated computer program. CAPTCAHS are used to prevent the false enrollment. One of the CAPTCHA methods is Collage CAPTCHA. In this method some shapes are shown with distortion and the user is asked to choose a specific object In this paper we increase the resistance of this method to attacks. For this purpose, we show some objects on the left of the screen and some of the objects on the right side of the screen which contain the name corresponding to the images shown on the left side of the screen. Now we ask the user to choose object on the left side and a specific image (containing word) on the right side of the screen. The image on the right side of the screen contains the name of the image. After this user is asked to enter the name of the image in a text box shown below the two images. The user will be passed the test if he chooses the two similar objects correctly and then he/she has entered the name of the image correctly in the text box. In this method because the computer program should also recognize the similar object on the right side and need to enter the name of the image in text box ,the

possibility of passing the test by computer is more reduced.

7. CAPTCHA is widely used security technique employed to avoid automated form submissions or verify user as human. An alternative CAPTCHA method based on pattern matching and number recognition ability is proposed in this paper, which verifies user as human and prevents bots to intervene and spam applications or services. This method is based on user gestures which make it unique and secure. The biggest advantage of this new CAPTCHA technique is that it is simple and easy task conducted by user as it is language independent. It generates a random 4 character string number and shown to user. User should show gesture of particular characters in an order using computer webcam or using a mobile phone. A pattern matching algorithm is applied on those user images to identify gestures, and find matching. This method is very difficult to hack because designing a bot to identify gesture in image is not possible for now. Many experiments we conducted to prove accuracy of our technique
8. Human Interaction Proofs (HIPs) or Completed Automated Public Turing tests to tell Computers and Humans Apart (CAPTCHAs) are systems that allow a computer to distinguish between another computer and a human. Reading based HIP challenges typically comprise a segmentation challenge followed by recognition challenges. The currently available CAPTCHAs have been broken with varying success, using the weakness in the methods used to generate the images. The state of the art of CAPTCHA design suggests that such text based schemes should rely on segmentation resistance to provide security guarantee, as individual character recognition after segmentation can be solved with a high success rate by standard methods such as neural networks. A good CAPTCHA must be not only human friendly, but also robust enough to resist computer programs that attackers write to automatically pass CAPTCHA tests (or challenges). In this paper we propose a multi level approach using hard image segmentation and embedded text to enhance the security of HIPs. The shadow of distorted text is embedded into sparse backgrounds. The effect of different backgrounds, into which the characters are distorted and embedded, is studied

3. PROBLEM FORMULATION

CAPTCHA technology has its foundation in an experiment called the Turing Test. Alan Turing, sometimes called the father of modern computing, proposed the test as a way to examine whether or not machines can think -- or appear to think -- like humans. The classic test is a game of imitation. In this game, an interrogator asks two participants a series of questions. One of the participants is a machine and the other is a human. The interrogator can't see or hear the participants and has no way of knowing which is which. If the interrogator is unable to figure out which participant is a machine based on the responses, the machine passes the Turing Test.

Of course, with a CAPTCHA, the goal is to create a test that humans can pass easily but machines can't. It's also important that the CAPTCHA application is able to present different CAPTCHAs to different users. If a visual CAPTCHA presented a static image that was the same for every user, it wouldn't take long before a spammer spotted the form, deciphered the letters, and programmed an application to type in the correct answer automatically.

Most, but not all, CAPTCHAs rely on a visual test. Computers lack the sophistication that human beings have when it comes to processing visual data. We can look at an image and pick out patterns more easily than a computer. The human mind sometimes perceives patterns even when none exist, a quirk we call pareidolia. Ever see a shape in the clouds or a face on the moon? That's your brain trying to associate random information into patterns and shapes.

But not all CAPTCHAs rely on visual patterns. In fact, it's important to have an alternative to a visual CAPTCHA. Otherwise, the Web site administrator runs the risk of franchising any Web user who has a visual impairment. One alternative to a visual test is an audible one which is already purposed by some researches. An audio CAPTCHA usually presents the user with a series of spoken letters or numbers. It's not unusual for the program to distort the speaker's voice, and it's also common for the program to include background noise in the recording. This helps the wart voice recognition. This software will increase in security level but it is very difficult to implement because in speech recognition process the system is working on basis of two methods one is speaker dependent and another is speaker independent. With this type of software we cannot work in the presence of external noise because if human want to enter voice captcha then he/she need noise less environment to match the voice samples recorded in computer which is not possible all the time.

So there is a need to develop a different and simple captcha which will be very much user friendly to human and difficult for computer machine.

point text, as you see here. Please use sans-serif or non-proportional fonts only for special purposes, such as distinguishing source code text. If Times Roman is not available, try the font named Computer Modern Roman. On a Macintosh, use the font named Times. Right margins should be justified, not ragged.

4. OBJECTIVE & METHODOLOGY

There will a large database of different labeled colored images. All of these images are pictures of concrete objects (a car, a table, a fan, a flower, a computer etc) but within only one color or it be a colored image only. The program picks any image at random basis presents them to the user and with it we will provide a list of all colors to user for their convenience and then asks the question “what is the color in this picture?” Current computer programs should not be able to answer this question, because computer machine is unable to understand he question and judge the color. In this thesis I have proposed a new idea for the CAPTCHA by presenting a method for increasing the resistance of it. Although we can increase the rate of its difficulty in order to improve its resistance against the attacks by applying other effects such as increasing the images present in the screen, this way the test will become more difficult to computer programs. Moreover we can also show different

images or overlapped in some fashion just within one color only that will be difficult task for a computer program to identify the color correctly.

So in this thesis my objective will be to create a database of different colored images, then during turing test the program picks an colored image at random, presents to the use ask the same question "what is color of the image" If user will give correct answer then he/she will pass the test.

5. VARITIONS AND FUTURE SCOPE

CAPTCHA (standing for Completely Automated Public Turing test to tell Computers and Humans Apart) must have seemed like a good idea when it was first invented in 2000. Spam was beginning to become a major problem on the web and a method was needed to fight back. CAPTCHA at first glance seems ideal: a distorted image that would be instantly recognizable by humans yet incomprehensible to machines. Place some letters in the distorted image and get the user to type them back and bingo: you've stopped your spam problem.

Real life though is rarely so easy. The problem is that spam is profitable, and because of that it's worthwhile to write programs that try to crack CAPTCHAs. The original CAPTCHA examples are now trivial for current algorithms to recognise, and the only option that developers had was to increase the complexity of the distortion. Successive CAPTCHA systems have added more distortion, extraneous lines and shapes, fuzz on the letters, multiple colours and different sizes all in an attempt to stay ahead of the spammers. This had lead to the current situation where the CAPTCHAs are so complex that it's difficult if not impossible for a large proportion of humans to recognize any particular one, yet a sizeable proportion of CAPTCHA breaking bots can solve that same one. But the CAPTCHA proposed by me is very to easy to understand by user and he/ she can easily pass the test.

- Future scope of this research work that more efficiently colored CAPTCHA can be developed.
- Multi colored CAPTCHA can be developed.
- Multi colored layered CAPTCHA can be developed.

6. CONCLUSIONS

We have introduced a new CAPTCHA technique that requires users to identify the color of the image or object. This is a technique that will be familiar to many people. And we can even say that people prefer this technique over traditional text-based technique which mostly annoys users. Our technique further improves the traditional text-based CAPTCHAs in that it is language and even written-script independent., and supports keyboard-difficult environments.

We ensure that our CAPTCHA cannot be defeated by state of the art image detection systems. In contrast to traditional image-based CAPTCHAs which usually introduce more noise and distortion as automated character recognition improves, we in our technique do not need to alter or distort the content of the images.

Many of the major pitfalls with other proposed image-based CAPTCHA techniques do not apply to our CAPTCHA technique. A previous knowledge of the image label is not needed. Further it is very difficult for bots to solve than the image-based CAPTCHAs that require a user to identify a common theme across a set of images.

Finally, our technique provides a number of interesting future works. First, the set of images chosen can be more interesting or valuable to the end-user by displaying those that are related to the overall theme of the website of the company. Second, We can even have multiple images and can ask the user to select the color of the particular image among them or we can select the maximum/minimum percentage of the particular color in a image. Third, we can even select 3D models for many applications. .

7. ACKNOWLEDGMENTS

Many thanks to Dr. Ajay Sharma and Dr. Harsh Verma, NIT Jalandhar for providing various research papers and giving their valuable suggestions.. Thanks are also given to Dr. Dinesh, Principal, A.B. College, for his valuable comments.

8. REFERENCES

[1] Yadava P, Sahu C and Shukla S. (2011), "Time-Variant Captcha: Generating Strong Captcha Security by Reducing Time to Automated Computer Programs",JETCIS VOL. 2, NO. 12, pp 701 704.

[2] Babu R, Kumar P and Rao S. (2011), "Implementation of Secure Multilayered CAPTCHA", IJAEST Vol No. 6, Issue No. 2, pp 200-219.

[3] Shah N and Banday T. (2009), "A Study of CAPTCHAs for Securing Web Services " IJSDIA Vol 1, No. 2, pp 66-74.

[4] Shah N and Banday T. (2009), "Image Flip CAPTCHA" ISeCure, Vol 1, No. 2, pp 105-123.

[5] Raj A, Jain A, Pahwa T and Jain Ab. (2010), "Picture CAPTCHAsWith Sequencing: Their Types and Analysis", IJDS, Vol 1, Issue 3, pp 208-220.

[6] Soni R and Tiwari D. (2010), "Improved CAPTCHA Method", IJCA, Vol 1, No. 25, pp 92 -94.

[7] Srinivas B, Kalyan G and Rao K. (2011). "Advanced CAPTCHA technique using Hand Gesture based on SIFT", IJCA Vol 31, No. 11, pp 16-22.

[8] Helena E and Gayathri M. (2009), "IJCSCT", Vol. 2, No. 1, pp 310-315.