ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 2, Issue 4, pp: (83-86), Month: October - December 2014, Available at: www.researchpublish.com

Technique to CAPTCHA Recognition

Prof. G. K. Gaikwad, Prof. B. S. Chaudhary

Department of Computer Engineering, Lecturer in Sandip Polytechnic, Nashik, India

Abstract: **A CAPTCHA is a program that can generate and grade tests that humans can pass but current computer programs cannot. For example, humans can read distorted text as the one shown below, but current computer programs can't. The term CAPTCHA (for Completely Automated Turing Test to Tell Computers and Humans Apart) was coined in 2000 by Luis von Ahn, Manuel Blum, Nicholas Hopper and John Langford of Carnegie Mellon University. At the time, they developed the first CAPTCHA to be used by Yahoo. ReCAPTCHA improves the process of digitizing books by sending words that cannot be read by computers to the Web in the form of CAPTCHAs for humans to decipher. More specifically, each word that cannot be read correctly by OCR is placed on an image and used as a CAPTCHA. This is possible because most OCR programs alert you when a word cannot be read correctly.**

Keywords: **CAPTCHA, Automated, Digitizing, Decipher, OCR.**

I. INTRODUCTION

The potential difficulty of differentiating humans from computers pretending to be humans was addressed at least as early as 1950, when Alan Turing described his now-famous Turing test. (His test was not automated.) The first discussion of automated tests which distinguish humans from computers for the purpose of controlling access to web services appears in a 1996 manuscript of Moni Naor from the Weizmann Institute of Science, entitled "Verification of a human in the loop, or Identification via the Turing Test".

A simple CAPTCHA had been developed in 1995 by Anton Lam of The Chinese University of Hong Kong, in a voting application written for Radio Television Hong Kong. The public were able to vote for their favorite singers and songs online for the first time in the annual "Top Ten Chinese Songs Award". To prevent robotic submissions, users were required to input a 6-digit number, which was displayed in an image, correctly.

A CAPTCHA system is a means of generating new challenges which:

- Current computers are unable to accurately solve.
- Most humans can solve.
- Does not rely on the attacker never having seen the given type of CAPTCHA before. For example, although a checkbox "check here if you are not a bot" might serve to distinguish between humans and computers, it is not a CAPTCHA because it relies on the fact that an attacker has not spent effort to break that specific form.
- Is able to automatically generate new challenges that require artificial intelligence techniques to solve.

In practice, the algorithm used to create the CAPTCHA does not need to be made public, though it may be covered by a patent. Although publication can help demonstrate that breaking it requires the solution to a difficult problem in the field of artificial intelligence, deliberate withholding of the algorithm can increase the integrity of a limited set of systems (see security through obscurity).

Hackers

If a captcha requires private data in order to be secure, then this captcha is vulnerable to hacker attacks: once a hacker breaks into the system and finds this secret data, the captcha ceases to be secure forever. We want captchas to generate tests that are hard to pass for all computers; not that are hard to pass for computers that don't know a particular secret.

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 2, Issue 4, pp: (83-86), Month: October - December 2014, Available at: www.researchpublish.com

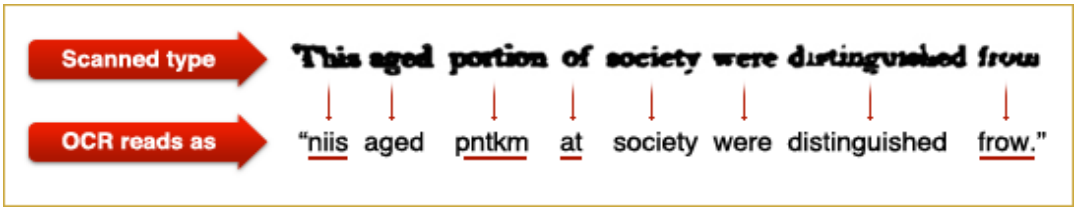


Fig.1. Text Scanned by OCR

About 60 million CAPTCHAs are solved by humans around the world every day. In each case, roughly ten seconds of human time are being spent. Individually, that's not a lot of time, but in aggregate these little puzzles consume more than 150,000 hours of work each day. What if we could make positive use of this human effort? ReCAPTCHA does exactly that by channeling the effort spent solving CAPTCHAs online into "reading" books.

To archive human knowledge and to make information more accessible to the world, multiple projects are currently digitizing physical books that were written before the computer age. The book pages are being photographically scanned, and then, to make them searchable, transformed into text using "Optical Character Recognition" (OCR). The transformation into text is useful because scanning a book produces images, which are difficult to store on small devices, expensive to download, and cannot be searched. The problem is that OCR is not perfect.

Fig.2. The word is selected from book & distorted

ReCAPTCHA technique improves the process of digitizing books by forwarding words that cannot be read by computers to the Web in the form of CAPTCHAs for humans to decipher it. More specifically, each word that cannot be read correctly by OCR is placed on an image and used as a CAPTCHA. This is possible because most OCR programs alert you when a word cannot be read correctly.

II. LITERATURE SERVEY

Actually we were thinking for developing the project based on web security or brute force attacks that are currently dangerous problem in web world.While surfing on net we had seen the simple image in small sized square in many famous websites like Google mail which was distorted & containing some image or text, one question always came into our mind what is the motto behind that distorted image. Then we searched many websites & we came to know the name of that image i.e. nothing but known as CAPTCHA. Then we started to collect the information about the CAPTCHA. We also visited captcha.net website and seen the free demo of CAPTCHA system. And slowly we came across many security related terms like Brute force attack, bots, spam we found the drawbacks, vulnerability of currently available system with solutions that avoid the hacking, viruses, blogs, spam which is nothing but CAPTCHA. With the help downloaded PDF files and research papers we decided to develop CAPTCHA system. We decided to design new type of CAPTCHA i.e.

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 2, Issue 4, pp: (83-86), Month: October - December 2014, Available at: www.researchpublish.com

Lyla CAPTCHA and after deciding to develop different algorithms we thought that there should be something for mobile devices also, so we started our search for mobile application development & for that J2ME was the only option at that time so we choose JAVA as front end of our Project & well-known Tomcat-Apache as Back-end then from one website we realized that it is also possible to run this daemon on processors like 8086 & typically used processors for door locking security system But the problem was that the screen size was very small & we have to communicate through Serial port only. The goal of steganography is to hide information from humans that is accessible to machines. The so-called CAPTCHAs have the opposite goal: hide information from machines that is accessible to humans. This viewpoint suggests non-traditional applications, including instructions, for example online or in a battlefield that a software agent will not be able to decode and may not even notice, but that a human or agent directly under human control will discover. Current proposals for CAPTCHAs suffer from the increasing ability of machines to discover the hidden information. The response of increasing the complexity of the CAPTCHAs makes them harder for humans to decode, negating their usefulness. This article presents three new approaches to the design of CAPTCHAs that may enlarge the gap between machine and human abilities.

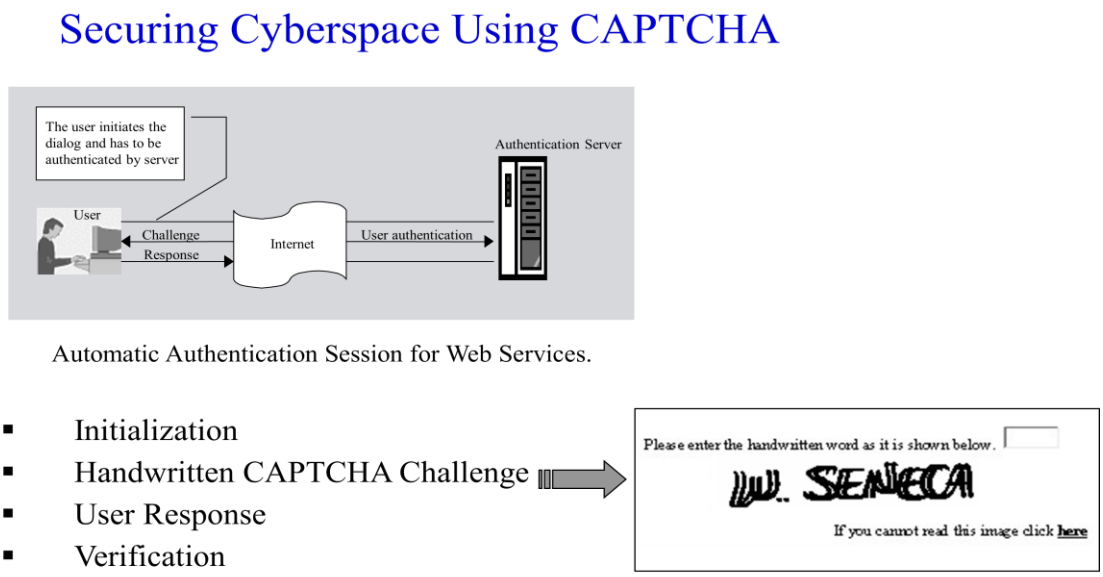
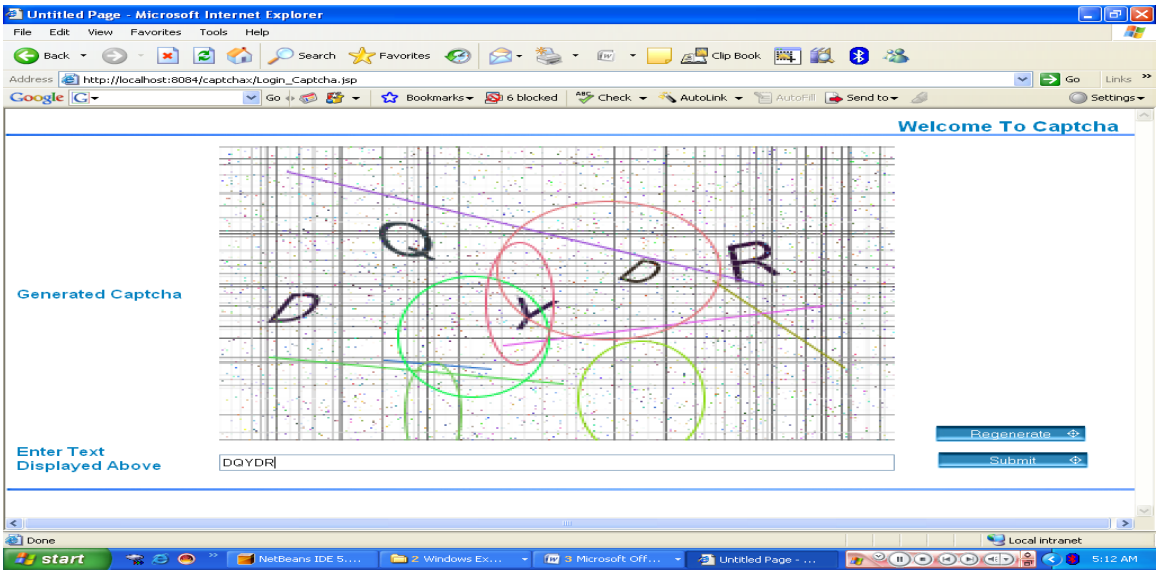


Fig.3. Block diagram of CAPTCHA System

Lyla CAPTCHA web page



III. CONCLUSION

In any case, clearly, the security of a CAPTCHA password authenticated system increases as the gap between human and machine performance on the test widens. The TGC CAPTCHA can thus serve as a modest cross-disciplinary challenge in the fields of pattern recognition, system security, computer graphics, and even psychology. Through friendly competition, we hope to encourage not only new OCR algorithms, but also a better understanding of the strengths and weaknesses of the human visual system relative to the best present-day machine vision systems. We believe that the fields of cryptography and artificial intelligence have much to contribute to one another. Captchas represent a small example of this possible symbiosis. Reductions, as they are used in cryptography, can be extremely useful for the progress of algorithmic development. We encourage security researchers to create captchas based on different AI problems.

REFERENCES

[1] http://www.captcha.net\ reCAPTCHA Security.htm

[2] www.lyalacaptcha.com\Lyla Captcha.htm

[3] www.IEEE.com\The Captcha Project@Everything2_com.htm

[4] http://en.wikipedia.org/wiki/Captcha.html

[5] http://www.kessels.com/captcha/security.html