International Journal of Computer Science and Engineering (IJCSE)
ISSN 2278-9960
Vol. 2, Issue 1, Feb 2013, 75-80

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

USING HUMAN COGNITIVE ABILITIES TO DISTINGUISH COMPUTERS AND HUMANS FOR PREVENTING BOT ATTACKS

M. KAMESWARA RAO[1] & K. KIRAN KUMAR[2]

[1]Associate Professor, Department of Electronics & Computer Engineering, K L University, Vaddeswaram, Guntur, Andhra Pradesh, India

[2]Professor, Department of Electronics & Computer Engineering, K L University, Vaddeswaram, Guntur, Andhra Pradesh, India

ABSTRACT

Denial of service (DoS) attacks by malicious automated programs (e.g., bots) is becoming a serious problem as masses of Web service accounts are being illicitly obtained. CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) is a test that can tell humans and computer programs apart automatically. The aim is to allow the server to identify the visitor is a human or a computer, and only provide services to human. The goal of CAPTCHA's is to ask questions which human users can easily answer, while current computers programs cannot. In this paper, a new CAPTCHA based on using the human ability to understand comics is proposed .In this Comic CAPTCHA, a seven-panel combination of relevant and irrelevant comics is presented and users that are able to identify the relevant comic panels are identified as human.

KEYWORDS: Bot, CAPTCHA, Denial of Service, Seven -Panel Comic, Turing Test

INTRODUCTION

Security is an essential part of any website, especially large sites holding valuable data and it is not something that can be compromised. The threats from bots is growing rapidly and also increasing in technical sophistication, thereby requiring a depth of defense to safeguard businesses against the risks they present with the attacks they deliver. To prevent these automated tasks, a program called CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) is used to provide an additional layer of security. CAPTCHA is a test to differentiate human users from malicious computer programs over the Internet [1].

They are generally used as security measures to prevent Denial of Service and brute force attacks. In the general Human Interactive Proof (HIP) problem, the goal is to distinguish between a human and machine using a simple test [2]. By focusing on exceptional cognitive abilities such as visual perception, a HIP tries to place a high barrier to machine duplication of human ability. A CAPTCHA is an application of this notion to security problems: it is a security primitive whose hardness assumption is based on a problem in Artificial Intelligence [3]. The remaining text of this paper is organized as follows: Section 2 presents a short review of existing CAPTCHA techniques. Section 3 illustrates the proposed Comic CAPTCHA technique. In section 4 security analysis of the proposed technique is discussed. Finally, we conclude and present future research directions.

BACKGROUND AND RELATED WORKS

Even though CAPTCHA is a very new area of research, it has already attracted researchers from AI, cryptography, signal processing, and computer vision. CAPTCHAs are classified based on their content, such as:

- Text-based.
- Image-based.
- Audio-based.
- Video-based.

The term "CAPTCHA" was originally introduced in 2000 by Luis von Ahn, Manuel Blum, Nicholas J. Hopper and John Langford [4].

Text-based CAPTCHAs: The most commonly used CAPTCHAs are text-based where distorted text is displayed. To solve the CAPTCHA, users must recognize the distorted characters and correctly enter them in a designated space. Textbased CAPTCHAs are easy to generate but are vulnerable to optical character recognition (OCR) attacks [5]. Many text-based Captchas have been developed including CMU's EZ-Gimpy [6], PARC's PessimalPrint [7], Baffle Text [8], Pay pal's Captcha [9], reCAPTCHA [10], Handwritten Word-based CAPTCHA [11], Persian/Arabic CAPTCHA [12] and Microsoft's Captcha [13]. Samples of some text-based CAPTCHA techniques are shown in Figure 1.

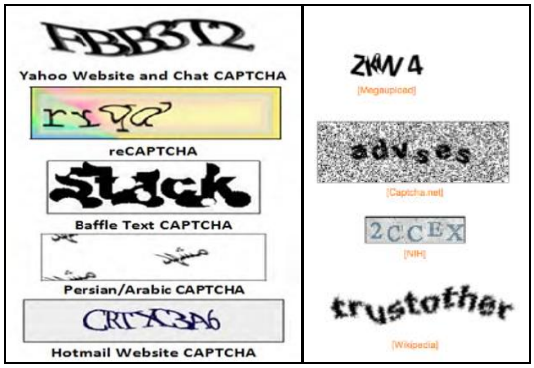

Figure 1: Samples of Some Text-Based CAPTCHA Techniques

Image-based CAPTCHAs: Image-based CAPTCHAs like Image Block Exchange [14] and Face Recognition [15] Captcha generally rely on image classification where users are presented with a series of images and asked to identify the relationship between them. One such CAPTCHA is ESP-PIX, which displays four images and asks users to select a common description from a drop-down list [16]. Since the image categories are selected from a fixed list, there is a high likelihood of random guessing yielding a correct answer. The Asirra image-based CAPTCHA uses a closed database of animals from Petfinder.com [17].Users are asked to select all images of cats from a mixed set of 12 cats and dogs drawn from a large source database of over 3 million images. Samples of some image-based CAPTCHA techniques are shown in Figure 2.

Figure 2: Image-Based CAPTCHA

Audio-based CAPTCHA: For visually-impaired users, some websites have developed audio-based CAPTCHAs. These generally work by playing a recording of a set of words or characters with users being asked to type-in what they hear. Unfortunately, these CAPTCHAs are subject to attacks using speech recognition software [18]. Eco [19] is one such system. Nancy Chan of the City University in Hong Kong has also implemented a sound-based system of this variety [20]. Haichang Gao [21] proposed a new audio CAPTCHA which exploits the gaps between human voice and synthetic voice. Sample of Audio-based CAPTCHA technique is shown in Figure 3.

Figure 3: Audio Based CAPTCHA

Video-based CAPTCHAs: These CAPTCHAs use videos rather than static images or text. In one such CAPTCHA, users are shown You Tube videos and asked to tag them with descriptive keywords. In tests, humans achieved 90% accuracy while computer attack rates were approximately 13% .Sample of video-based CAPTCHA technique is shown in Figure 4.

Figure 4: Video Based CAPTCHA

PROPOSED CAPTCHA

This paper proposes a new type of CAPTCHA based on human ability to understand comics, which is considered one of the most advanced human cognitive processing abilities. As a specific example, this paper will study a CAPTCHA using a seven -panel comic. In this seven panel comic CAPTCHA, a seven-panel comic is presented with the seven panels containing relevant and irrelevant comics arranged randomly, and users that are able to click the relevant comics are identified as human. For a computer, however, it would be difficult to recognize the relevant and irrelevant pictures and utterances in each panel.

The authentication process used in the proposed method is as follows:

STEP 1: A Seven-panel comic is selected at random from a database of comics.

STEP 2: The order of the panels in the selected seven-panel comic is rearranged randomly (shuffled).

STEP 3: The shuffled seven-panel comic is presented to the user.

STEP 4: The user clicks on the relevant comic panels

STEP 5: If the selected (clicked) panels are correct, the user is identified as a human, and if incorrect, the user is identified as bot.

Figure 5 shows an example of the authentication screen used in the proposed method.

Figure 5: Comic CAPTCHA

SECURITY AND USABILITY

In the case of seven-panel comic CAPTCHAs, a seven-panel comic is presented with the seven panels rearranged randomly, and users that are able to respond with the relevant comic panels are identified as human. In this case, there are only 5040 possible combinations. A bot could respond a correct answer at a rate of one out of every 5040 tries. Answering CAPTCHAs is an added annoyance for users, who feel troublesome to prove that they are human at every Web access. If the authentication process itself is a pleasant one for users, however, it is expected that the overall usability level will be acceptable. In the proposed method which is asking the users to read seven-panel cartoons, we believe that the authentication process will be entertaining and enjoyable.

CONCLUSIONS

In this paper, we focused on the user's ability to understand comics, which represents the ultimate in human cognitive processing abilities, and proposed a CAPTCHA that uses seven panel comics. The proposed method is expected to offer a new form of CAPTCHA that feature both security and usability, being difficult for advanced malware to decipher, and at the same time offering entertainment value for users, who will enjoy reading the comics.

REFERENCES

1. M. Blum, L. A. von Ahn, and J. Langford, The CAPTCHA Project, \Completely Automatic Public Turing Test to tell Computers and Humans Apart," www.captcha.net, Dept. of Computer Science, Carnegie-Mellon Univ., and personal communications, November, 2000

2. L. von Ahn et al. CAPTCHA: Using Hard AI Problems for Security. Euro crypt, 2003.
3. L. von Ahn, M. Blum, and J. Langford. “Telling Humans and Computers Apart Automatically,” Communications of the ACM, vol. 47, no. 2, pp. 57-60, Feb. 2004.
4. Luis von Ahn, Manuel Blum, Nicholas Hopper, and John Langford. CAPTCHA: Using Hard AI Problems for Security. In *Eurocrypt* 2003
5. Chellapilla, K., Larson, K., Simard, P.Y., and Czerwinski, M. (2005) Building Segmentation Based Human-Friendly Human Interaction Proofs (HIPs), In: Human Interactive Proofs, (pp. 1-26), Berlin: Springer
6. M. Blum, L. A. von Ahn, and J. Langford, The CAPTCHA Project, \Completely Automatic Public Turing Test to tell Computers and Humans Apart," www.captcha.net, Dept. of Computer Science, Carnegie-Mellon Univ., and personal communications, November, 2000
7. A.L. Coates, R.J. Fateman, and H.S. Baird , Pessimal Print: A Reverse Turing Test, In Proceeding of the 6th Inernational Conference on Document Analysis and Recognition, Seattle, WA,USA, 2001, pp.1154-1158.
8. M. Chew and H.S. Baird, BaffleText: a Human Interactive Proof, Proc., 10th SPIE/IS&T Document Recognition and Retrieval Conf.(DRR2003), Santa Clara, CA, January 23-24, 2003.
9. Paypals-URL" site: www.paypals.com
10. L. V. Ahn, B. Maurer, C. McMillen, D. Abraham, and M. Blum. reCAPTCHA: Human Based Character Recognition via Web Security Measures. Science Express, 321(5895):1465 -1468, 2008.
11. A. Rusu and V. Govindaraju. Handwritten CAPTCHA: Using the Difference in the Abilities of Humans and Machines in Reading Handwritten Words. In Proceedings of the 9th International Workshop on Frontiers in Handwriting Recognition (IWFHR- 9 2004), pages 226{231,Kokubunji, Tokyo, Japan, 2004.
12. M. H. Shirali-Shahreza and M. Shirali-Shahreza. Persian/Arabic Baffletext CAPTCHA. Journal of Universal Computer Science, 12(12):1783{ 1796, 2006.
13. Microsoft 2006. *Microsoft Hotmail*. http://www.hotmail.com/ last visited 5 September 2006.
14. W. H. Liao. A CAPTCHA Mechanism by Exchanging Image Blocks. In Proceedings of the 18th International Conference on Pattern Recognition (ICPR06), volume 1, pages 1179{1183,Hong Kong, 2006.
15. D. Misra and K. Gaj. Face Recognition CAPTCHAs. In Proceedings of the Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services(AICT/ICIW'06), pages 122{127, Guadeloupe,French Caribbean, 2006.
16. C. Pope and K. Kaur. Is It Human or Computer? Defending E-Commerce with Captchas. IEEE IT Professional, 7(2):43{49, 2005.
17. J. Elson, J.R. Douceur, J. Howell, and J. Saul, Asirra: a CAPTCHA that exploits interest-aligned manual image categorization. In Proceedings of the 14th ACM Conference on Computer and Communications Security, Alexandria, Virginia, USA, 2007, 366-374.

18. R. Datta, J. Li, and J. Z. Wang. Imagination:A Robust Image-Based CAPTCHA Generation System. In Proceedings of the 13th Annual ACM International Conference on Multimedia (MULTIMEDIA05), pages 331{334, New York, NY, USA, 2005. ACM Press.
19. Luis von Ahn_ Manuel Blum_ John Langford_ "Telling Humans and Computers Apart (Automatically) or How Lazy Cryptographers do AI"
20. Nancy Chan. Program Byan: http://drive.to/research
21. Haichang Gao, Honggang Liu, Dan Yao, Xiyang Liu "An audio CAPTCHA to distinguish humans from computers" 2010 Third International Symposium on Electronic Commerce and Security July 29-July 31 ISBN: 978-0-7695-4219-5