

Developing Usable CAPTCHAs for Blind Users

Jonathan Holman¹, Jonathan Lazar¹, Jinjuan Heidi Feng¹, and John D'Arcy²

¹Department of Computer and Information Sciences and Universal Usability Laboratory

Towson University

8000 York Road

Towson, MD 21252

Phone: 410-704-2255

e-mail: jon@jonholman.com, jlazar@towson.edu; jfeng@towson.edu, jdarcy1@nd.edu

²University of Notre Dame

Mendoza College of Business

Notre Dame, IN 46556

Phone: 574-631-6183

ABSTRACT

CAPTCHAs are widely used by websites for security and privacy purposes. However, traditional text-based CAPTCHAs are not suitable for individuals with visual impairments. We proposed and developed a new form of CAPTCHA that combines both visual and audio information to allow easy access by users with visual impairments. A preliminary evaluation suggests strong potential for the new form of CAPTCHA for both blind and visual users.

Categories and Subject Descriptors

H.5.2 User Interfaces

General Terms: Documentation, Design, Experimentation, Security, Human Factors, Verification.

Keywords: CAPTCHA, visual impairment, blind user, accessibility, security, usability, Turing test, universal usability

INTRODUCTION

The trade-off between usability and security is well documented [3,4]. However, the issue of usable security for individuals with disabilities has received little attention. It is our position that as more complex security features are added, these features may pose challenges for users with disabilities. In this research, we explore this topic as it pertains to a specific computing population – blind users.

A group of blind users were interviewed to determine the biggest challenges they currently face in regards to security and privacy. CAPTCHAs (Completely Automated Public Turing tests to tell Computers and Humans Apart) [1], were identified as a significant problem for blind users relating to security.

CAPTCHAs are an image with some twisted distorted text, that the user has to identify and type into a text box to be able to proceed (usually submit a form) on a website. The main purpose of a CAPTCHA is to keep web bots from filling out and submitting forms. Web bots usually perform such tasks to obtain free email addresses for spammers, which can then be used to send mass amount of spam or to submit spam to online blogs and forums. Some sort of mechanism needs to be on these sites to prevent these illegitimate users from submitting forms.

Unfortunately the current text-based CAPTCHA method blocks out blind users who cannot see the twisted text because their screen reader accessibility software has no ability to discern these CAPTCHAs and if it did, they would not serve their purpose because the web bots would figure it out as well. There are audio CAPTCHAs as alternatives for blind users, which give a sound clip of text being read, and then the user types in that text. The challenge with audio clips of text is that they can easily be broken by speech recognition software, so they must contain some level of background noise, which the human could filter, but the computer could not. This is in a manner similar to how visual text in a CAPTCHA must be twisted so that visual recognition software could not identify it, but a human (with vision) could.

In this paper, a new form of CAPTCHA is proposed, developed, and evaluated. The results are encouraging in that the blind users were generally satisfied with the new CAPTCHA form and were able to utilize it efficiently. Similar results were obtained from a sample of visual users without disabilities. The results of this preliminary investigation suggest the potential of combined audio/visual-based CAPTCHAs for improving security design for all users.

NEW FORM OF CAPTCHA

The challenge of CAPTCHAs is that they must be easy for users to use, while being hard enough that current image recognition and voice recognition software cannot identify them. We believe that the key is to use a combination of pictures and sounds which are non-textual, that is, pictures of objects, and the sounds that these objects make. Users would then need to identify the objects that they see/hear. This form of CAPTCHA would be much more universally usable as it would provide both visual and audio feedback to the user, so both users with visual impairment and hearing impairment would be able to use the CAPTCHA. In addition, a user without a disability could use either the picture or sound, or potentially both. It would also be more secure as most computers could not easily identify these pictures and sounds. From a security perspective, a potential limitation is the finite number of image and sound combinations that are easy to identify and could be used as authentication mechanisms. It is plausible that an automated script could be written that would test every image/sound combination and thus enable unauthorized access to the particular website. However, traditional text-based CAPTCHAs suffer from a similar limitation, and it is addressed by automatic lockouts after a number of unsuccessful attempts (e.g. lockout after 3 incorrect attempts). Another benefit of this system is that it could theoretically be multi-lingual without much work. While the pictures and sounds could remain the same, only

Copyright is held by the author/owner(s).

ASSETS '07, October 15–17, 2007, Tempe, Arizona, USA.

ACM 978-1-59593-573-1/07/0010.

the correct labels (switching from say, English, to French in the drop-down box) would need to be changed.

One of the systems designed in CMU (Carnegie Mellon University)'s CAPTCHA project uses the concept of the pictures [2]. They display four pictures and have the user select from a drop-down list of options the central concept of the four pictures. From a universal usability perspective, this approach is limited in that it is visually-based and therefore inaccessible to blind users. Our prototype CAPTCHA expands upon this approach by including a combination of pictures and sounds.

IMPLEMENTATION

The prototype tool developed is a web-based application. Upon loading, the tool uses a random number generator to randomly pick an item from the array of possible options. The image and audio file that go along with this answer are then displayed to the user. In the current version, there are 15 possible picture/sound combinations. There were 4 categories of pictures/sounds that were chosen for use in the application: transportation, animals, weather, and musical instruments. These 4 categories were chosen for their universality, and all 4 researchers spent time trying to come up with both pictures and sounds that were easily identifiable. For example, one picture/sound combination was a picture of a train and the sound of a train chugging along. Other examples included bird, cat, drum, and piano. One concern was that for all of these picture/sound combinations, there needed to be only one easily identified name or label. The new form of CAPTCHA was tested with both Window-Eyes and JAWS to ensure that it was fully accessible to work with screen readers.

Our new CAPTCHA tool sorts a drop-down list of possible answers while throwing in some fake selections that would not be a correct answer because there is no image and audio file to go along with them. When the CAPTCHA appears on the screen, the user can look at the image or play the sound effect and then select from the drop-down list what they identify the picture/sound object to be. Upon answering, the application returns to the user if their answer was correct or incorrect and asks them if they would like to attempt another CAPTCHA. Figure 1 displays a visual example of the CAPTCHA tool.

EVALUATION

The new form of CAPTCHA was evaluated with 5 visual users and 5 blind users. All five blind users interact with computers or computer related devices via screen reader and/or Braille keyboards in their daily life. Visual users took an average of approximately 3.5 minutes to answer all fifteen CAPTCHAs (stdev: 35 seconds). No errors were made during the test. Interestingly, although the visual users can answer the CAPTCHAs solely based on the images, most users did listen to the sound clips. Blind users took an average of 8.8 minutes to answer all fifteen CAPTCHAs (stdev: 88 seconds). The longer task time was partly due to the frequent comments and discussions that blind users raised during the task (the visual users made almost no comments).

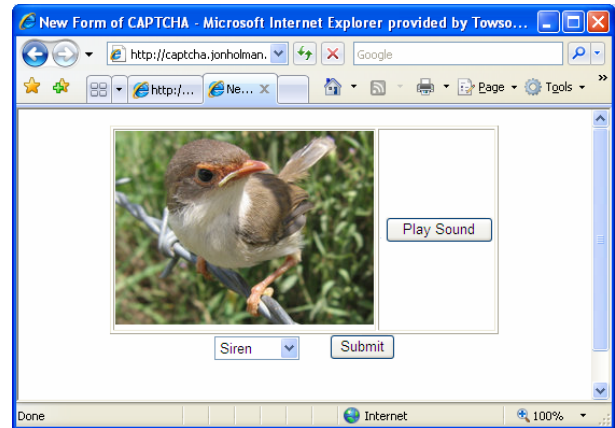


Figure 1. Example of picture/sound form of CAPTCHA

Three of the blind participants made one error during the test. Two blind participants made two errors. All errors were successfully fixed during the second try. Both groups of the users, especially the blind group, provided highly satisfactory feedback regarding the new form of CAPTCHA. All blind users strongly believed that the new form of CAPTCHA was very easy to use and preferred to use the new CAPTCHA over the traditional text-based CAPTCHAs.

CONCLUSIONS

We proposed a new form of CAPTCHA that can be used by both users with and without visual impairments. A preliminary evaluation suggests that blind users can solve this new form of CAPTCHA both efficiently and accurately. This is a major breakthrough compared to traditional text-based CAPTCHAs, which are not suitable for visually impaired users. For future work, the CAPTCHA tool will be improved to include more visual/audio options. The design of the tool may also be altered to increase the cognitive capability required to solve the CAPTCHA.

ACKNOWLEDGMENTS

We would like to thank National Federation of the Blind for their support and assistance in recruiting participants.

REFERENCES

- [1] Ahn, L., Blum, M., and Langford, J. (2004) Telling humans and computers apart automatically. *Communications of the ACM*, February 2004. Vol. 47 (2) 57-60.
- [2] CAPTCHA.net (2007) CAPCHA: Telling humans and computers apart automatically. Last retrieved on June 7, 2007 from www.captcha.net
- [3] Johnston, J., Eloff, J., and Labuschagne, L. (2003). "Security and Human Computer Interfaces." *Computers & Security*, 22(8), pp. 675-684.
- [4] Sasse, M.A., Brostoff, S., and Weirich, D. (2001). "Transforming the Weakest Link - A Human/Computer Interaction Approach to Usable and Effective Security." *BT Technology Journal*, 19(3), pp. 122-130.