ASSESSING THE USABILITY ISSUES OF CAPTCHA

Ragavi.V[1] and Dr.G. G.Geetha[2]

[1]Department of Computer Applications, [2]Department of Computer Science Engineering
[1]Sathyabama University,Chennai, India. [2]Lovely Professional University, Phagwara, India.
[1]ragaviprabhu@gmail.com, [2]gitaskumar@yahoo.com

ABSTRACT

Denial of Service (DoS) attacks has become a large problem for users of computer systems connected to the Internet. CAPTCHA is now almost a standard security technology, and has found widespread application in commercial websites. These CAPTCHAs day by day become very harder for humans to pass. We agree that the CAPTCHAs must be robust enough for the machines to break, but making it tougher for humans cannot be justified. Usability is critical for any application. If the users can't use it, it does not work. Here is our attempt in assessing the usability issues of CAPTCHA. This paper is written with the intention to help CAPTCHA designers to design CAPTCHA technology that would pass the basic test of usability.

Keywords : CAPTCHA, DoS, Security, Usability

1. INTRODUCTION

With an increasing number of free services on the internet, we have found a pronounced need to protect these services from abuse. Attackers hijack secondary victim systems using them to wage a coordinated large-scale attack against primary victim systems. To thwart automated attacks, services often ask users to solve a puzzle before being given access to a service. These puzzles, first introduced by von Ahn et al. in 2003 [1], were CAPTCHAs:

Completely Automated Public Turing Tests to Tell Computers and Humans Apart (CAPTCHAs) are one of the important branches of HIP systems. This technology is now almost a standard security mechanism for addressing undesirable or malicious Internet bot programs (such as those spreading junk emails and grabbing thousands of free email accounts instantly) and has found widespread application on numerous commercial web sites including Google, Yahoo, and Microsoft's MSN.

Usability and robustness are two fundamental issues with CAPTCHA, and they often interconnect with each other. It is widely accepted that a good CAPTCHA must be both robust and usable. The robustness of a CAPTCHA is its strength in resisting adversarial attacks, and this has attracted considerable attention in the research community (e.g. [2,3,4,5]). However, it is strikingly surprising that there has been little study of the usability aspects of CAPTCHA, although by definition, a CAPTCHA that is unusable for human should have no reason to exist. Attackers are constantly developing new methods to circumvent these countermeasures, so the puzzles get still harder and more burden is placed on humans.

CAPTCHA was introduced to address the availability problem but it in turn pose accessibility problem to users who are blind, who have low vision, or have a learning disability

such as dyslexia. In [6], [7] K.Chellapilla et al have discussed about the design of a user friendly CAPTCHA. In addition, some usability issues of CAPTCHAs were touched in [8,9,10,11]. However, solving a CAPTCHA requires a substantial human cognitive effort. Based on the type of cognitive effort required to solve CAPTCHA, CAPTCHAs can be classified into three categories.

1. *Text based CAPTCHAs:* They require users to read and type distorted text rendered in an image.
2. *Audio based CAPTCHAs:* They rely on sound or speech recognition by the users.
3. *Image based CAPTCHAs:* They ask users to perform an image recognition task.

In this paper, our discussion will largely focus on text and image-based CAPTCHAs, for the following reasons.

First, text-based CAPTCHA and image based CAPTCHAs are the most widely deployed.

Second, it can have a large and positive impact for the society to improve the usability of such popular and well-claimed CAPTCHAs by identifying issues that should be addressed in these schemes.

Lastly, although our discussions are focused on text and image-based schemes, they can also be relevant to other types of CAPTCHAs.

This paper is organized as follows: In section 2 we bring out a simple framework for usability related issues. In Section 3 we discuss in detail about the usability issues of both text and image based CAPTCHA, Section 4 gives the evaluation of the two Schemes and in Section 5 concluding remarks are presented.

2. A SIMPLE FRAMEWORK

As said by Jakob Nielsen [12], usability is defined by the following five quality components:

- *Learnability*: How easy is it for users to accomplish basic tasks the first time they encounter the design?
- *Efficiency*: Once users have learned the design, how quickly can they perform tasks?
- *Memorability*: When users return to the design after a period of not using it, how easily can they re-establish proficiency?
- *Errors*: How many errors do users make, how severe are these errors, and how easily can they recover from the errors?
- *Satisfaction*: How pleasant is it to use the design?

There are many other important quality attributes. A key one is utility, which refers to the design's functionality: Does it do what users need? Usability and utility are equally important: It is not good if the system can hypothetically do what is needed, but one can't make it happen because the user interface is too difficult. To study the design utility, in this paper, we will only consider the quality components. The nature of CAPTCHAs determines that the following usability criteria are applicable to address efficiency, errors and satisfaction:

- *Accuracy*: how accurately can a user pass a CAPTCHA challenge? For example, how many times she has to try in order to pass a test?
- *Response time*: how long does it take for a user to pass the test?
- *Perceived difficulty/satisfaction of using a scheme*. How difficult to use do people perceive a CAPTCHA is? Are users subjectively satisfied and would they be willing to use such a scheme?

This set of criteria can be key for (quantitatively) evaluating the usability of CAPTCHAs. However, this set offers little specific guidance on *how* to improve accuracy, response time or perceived difficulty / satisfaction.

Instead, the following four-dimensional framework for examining the usability of CAPTCHAs.

- *Distortion*. This dimension examines the form of distortions employed by a CAPTCHA and their impact on usability.
- *Content*. This dimension examines contents embedded in CAPTCHA challenges (or tests) and their impact on usability. For example, how should the content be organised, and is the content appropriate?
- *Presentation*. This dimension examines the way that CAPTCHA challenges are presented and its impact on usability.
- *Understandability*. This dimension examines how fast the user is able to understand and do the necessary task.

With this framework, specific elements of a CAPTCHA can be pinpointed and improved so as to enhance the usability of the scheme as a whole. This framework is applicable to image based CAPTCHA.

3. USABILITY ISSUES OF TEXT AND IMAGE-BASED CAPTCHAS

In this section, we discuss usability issues in text and image-based CAPTCHAs under the framework proposed in Section 2. Table 1 summarizes all the usability issues in text based CAPTCHA and Table2 summarizes the usability issues of the image based CAPTCHA. These will be discussed in the following sections.

Table 1. Usability issues with text-based CAPTCHAs

Category	**Usability issue**	
Distortion	Distortion method and level	
	Confusing characters	
Content	Character Set	
	String Length	How Long?
		Predictable or not?
	Offensive word	
Presentation	Use of colour	
	Font type and size	
	Integration with web pages	

Table 2. Usability issues with image-based CAPTCHAs

Category	**Usability issue**
Distortion	Distortion method and level
	Resolution
Content	Image Set

	Time taken.
	Offensive image
Presentation	Image size
	Integration with web pages
Understandability	What to do?

3.1 Distortion related issues

Distortion has a clear impact on the usability of CAPTCHAs, since human users would find it difficult or impossible to recognise over-distorted characters and images. To cope with usability problems caused by distortion, a system will have to allow multiple attempts for each user. Typically a new challenge is used for each attempt. This will not only annoy users, but also lowers the security of the system by a factor of the number of allowed attempts.

3.1.1. Distortion method and level. The most intuitive usability concern of a text-based and an image-based CAPTCHA is its recognizability, which can be largely determined by what distortion methods are used and how much distortion is applied to texts and images. A Microsoft team [6] examined the following common distortion methods, among others, and empirically determined the level of distortion for each method that will not make it difficult for human users to recognize distorted images.

- *Translation*: moving characters / images either up or down and left or right by an amount
- *Rotation*: turning characters / images either in a clockwise or counter clockwise direction
- *Scaling*: stretching or compressing characters / images in the x direction and y-direction
- *Warp*: elastic deformation of CAPTCHA characters / images at different scales

These methods of distortion are used and the characters and images are distorted very badly so that it is very difficult for the users to identify the text and images. Users with low vision suffer a lot. However, distortion is absent in some image based schemes (e.g. Assira [13] and Bongo [1]) - for these schemes, only the dimensions of content and presentation matter.

3.1.1.1 Text-Based CAPTCHA- Distortion

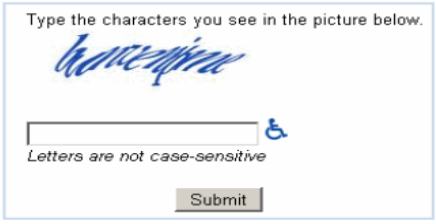

Fig 1. Distorted text based CAPTCHA

3.1.1.2 Image-Based CAPTCHA- Distortion

Fig.2 and Fig 3. Distorted Image CAPTCHA

3.1.2. Confusing Characters in Text-Based CAPTCHA.

Distortion often creates ambiguous characters, where users cannot be sure what they are. Although some characters have very different shapes, after distortion, they become hard to tell apart from each other. Certain letters are confused with digits such as o with 0, b with 6, s with 5, z with 2, l with 1 and so on. Certain letters when appearing together are confusing, letters such as vv is mistaken as w, cl is taken as d, nn as m, rn as m.

Fig. 4 Confusing Characters in text-based CAPTCHA

3.1.3. Resolution in Image-based CAPTCHA. If the resolution of the system used by the user is not high, then the image displayed will not be clear. This again pose a problem in addition to the disorted image. If the image has a high resolution and if the users system has

a colour quality of 16 bits only then the distorted image will be very badly damaged such that the user will not be able to identify and perform the necessary task.

Fig. 5 Image CAPTCHA with bad Resolution

3.2. CONTENT RELATED ISSUES

The choice of content materials used in each CAPTCHA challenge can also have significant impact on usability.

3.2.1. Character set for Text Based CAPTCHA. The size of the character set used in a CAPTCHA matters for security. Typically, the larger the character set, the higher resistance to random guessing attacks each challenge can have. However, a larger character set can also imply a higher number of characters that look similar after distortion, causing confusion.

3.2.2. Image set for Image Based CAPTCHA. The size of the image set used in a CAPTCHA matters for security. Typically, the larger the image set, the higher resistance to random guessing attacks each challenge can have. It is difficult for a small site to acquire a large dictionary of images which an attacker does not have access to and without a means of automatically acquiring new labelled images, an image based challenge does not usually meet the definition of a CAPTCHA. KittenAuth, by default, only had 42 images in its database.[14]. This problem is addressed by employing huge public image databases to randomly select a couple of images of a random object and asks user to identify it. Unfortunately, huge public image databases, such as images.google.com, often contain mislabeled images yielding false images for any specific object.

3.2.3 String Length for text CAPTCHA. The length of the text string used in each challenge also matters for security. If both the character set size and the string length are small, random guessing would have a high chance of passing the CAPTCHA. Typically, the longer the string is used in a challenge, the more secure is the result. If the length is predictable then security is lost, if the length is random then security is more but difficult for the user if the string is too long.

3.2.4. Time Taken for Image Based CAPTCHA. The content of the image CAPTCHA matters a lot. If any form containing this CAPTCHA is to be displayed then the time taken to download the image CAPTCHA is really notable. At times the time expires also. Users are frustrated in filling up such forms and quit half way.

3.2.5. Offensive Word in Text Based Scheme. Whether the content of the string used in each challenge is appropriate can affect user satisfaction, and thus is another usability issue. For example, it would be offensive to present a challenge showing words such as “negro”. Offensive content can occur in either random or dictionary words based schemes.

3.2.6. Offensive Images. Whether the content of the image used in each challenge is appropriate can affect user satisfaction, and thus is another usability issue. There are images which are offensive and images which are not user friendly. The idea of CAPTCHA has also been taken to more frivolous places, such as a system based on the somewhat dubious "hot or not" tests, shown here. Some may find that version funny, others may find it offensive. Either way, it's no use as a genuine authentication system. The answers are arbitrary, and in any case, they can be mined programmatically.

Fig. 6. Offensive Image CAPTCHA

3.3. PRESENTATION RELATED ISSUES

The way that a CAPTCHA presents its challenges (or tests) has usability concerns.

3.3.1 Use of Colours in Text Based CAPTCHA. Colour is extensively used in user interfaces. When used properly, colour can much enhance user interface design [8]. Using colour has also been common in text-based CAPTCHAs Colour can facilitate recognition, comprehension and positive affect and can make CAPTCHA images compatible with the colour of web

pages and make them look less intrusive[17]. But now a days colours are disturbing. Fig 7 shows some of the CAPTCHAS were colour are used extensively.

Fig 7. Use of colours in Text Based CAPTCHA

3.3.2 Image Size. Basically any image based CAPTCHA occupy nearly one fourth of the page as we can see the IMAGINATION CAPTCHA, Assira etc., The space occupied by the images is an usability issue which is to be addressed. If one tries to reduced the space the image becomes still more complex to be identified.

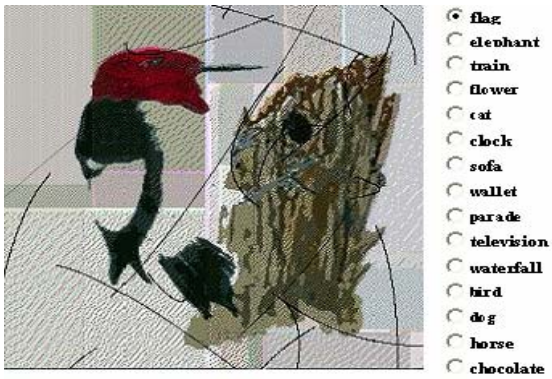

Fig 8. Image CAPTCHA occupying more space

3.3.3. Integration with a web page/form for both Text and Image CAPTCHA. The integration of CAPTCHA challenges with web pages can also have usability concerns. For example, until very recently, the "type the 3 words" box in the popular CAPTCHA scheme was not automatically enabled (see Figure). So users had often input their answers to nowhere, unless they manually activated the box in advance. But it certainly increased the users' burden by forcing them to enable the text box before they could enter an answer. To avoid annoying end users, a CAPTCHA should be integrated into a web page with care to minimize their burden.

Fig 9. Image CAPTCHA User Interface

3.4 UNDERSTANDABILITY RELATED ISSUES.

Understanding the CAPTCHA challenge is a major issue for Image Based CAPTCHA.

3.4.1. What to do? Users panic when they see certain image CAPTCHA. This proves to be a heavy burden for the end users. It is very difficult to grasp the task and perform it correctly the very first time. User needs many tries which annoys him. This issue has to be taken in hand for certainty.

Fig 10. Image CAPTCHA with understandability problem

In the above image CAPTCHA one has to perform the task of clicking the nearest geometric center of any one of the eight composite tiled image. But people need explanation to find the geometric center of any one image and they find it very difficult to get it right in the first attempt. They need atleast two or three try to understand the fact that the geometric center is for one of the image in the composite image. Next, it is hard for the users to recognize the edges of the images especially when the background colours of two adjacent images are same. An analysis of this CAPTCHA is done in [15].

In Fig. 11. the question given is very confusing. The user searches for the third object, which is in the grid. He needs explanation, thus needs more tries.

Fig 11. Image CAPTCHA 2 with understandability problem

Fig. 12 displays two series of blocks, the left and the right series. The blocks in the left series differ from those in the right, and the user must find the characteristic that sets the two series apart. A possible left and right series are shown below:

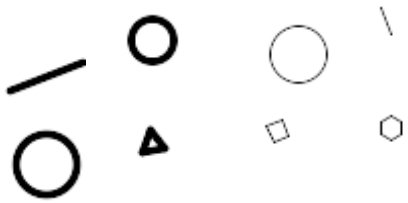

Fig 12. Image CAPTCHA 3 with understandability problem

After seeing the two series of blocks, the user is presented with four single blocks and is asked to determine whether each block belongs to the right series or to the left. The user passes the test if he or she correctly determines the side to which all the four blocks belongs, which sounds really baffling.

4 . ASSESSING CAPTCHA

To assess CAPTCHA, we conducted user survey for text-based and image-based CAPTCHA respectively. The analysis was conducted with rigorous investigation and the results are drawn out. The various dimensions in which the evaluation was considered are

- Finish time: the total time taken to solve a CAPTCHA.
- Rate of misidentification: the number of CAPTCHAs identified wrongly with divided by the number of all CAPTCHAs.
- Rate of timeout: the number of timeout CAPTCHAs divided by the number of all CAPTCHAs. A CAPTCHA is regarded as "timeout" if it is targeted but the player did not have any further action in 3 seconds.
- Rate of giving up: the number of CAPTCHAs given up divided by the number of all CAPTCHAs. User may wish to skip to the next CAPTCHA.

4.1 Text-Based CAPTCHA. The survey was conducted with more than 100 people. Various age groups were included. Many text-based CAPTCHAs like Google, Yahoo, Microsoft, CoolCAPTCHA, TgCAPTCHA were taken into consideration. These CAPTCHA schemes use various strategies, such as background patterns, distortions, and noises, to generate images. The graph given below shows the results of the ordeal.

Fig.13 Survey Results of Text-Based CAPTCHA

Lessons Learned:

- Many of them had typing error as they were unable to identify the distorted letters.
- Many found, letters were confusing which led to time out.
- Others tried to give another attempt.
- Works definitely have to concentrate for low vision people, as they find it very difficult.

4.2 Image-Based CAPTCHA. The survey was conducted with more than 100 people. Various age groups were included. Many image-based CAPTCHAs like Assira, Bongo, IMAGINATION, Collage CAPTCHA, 3D CAPTCHA etc were taken in the test . These CAPTCHA schemes use various strategies, such as rotation, distortions, translation , wraping, scaling

and noises, to generate images. The graph given below shows the results.

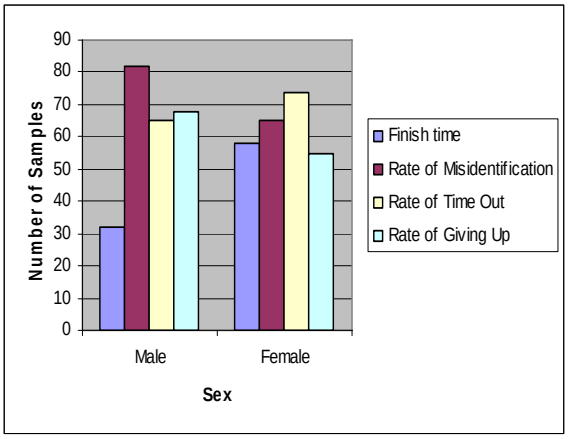

Fig.14 Survey Results of Image-Based CAPTCHA

Lessons Learned:

- Understanding what has to be done was the major problem identified.
- Many images are very badly distorted and it was very difficult to identify.
- As it was very tricky, time out and many attempts have to be made.
- Low vision people suffer a lot.

5. CONCLUSION

We have discussed usability aspects and evaluated the CAPTCHA design. We do not claim the list of usability issues we have discussed is complete, and hence encourage researchers to identify more of them using our framework, and we are sure, a lot more can be explored. As spammers advance, the puzzles get progressively harder and more of a burden is placed on humans. What is required is a new test that does measures user heuristics without requiring additional user input.

6. REFERENCES

[1] L von Ahn, M Blum and J Langford. "Telling Humans and Computer Apart Automatically", CACM, V47, No2, 2004.

[2] Greg Mori and Jitendra Malik. "Recognising Objects in Adversarial Clutter: Breaking a Visual CAPTCHA", IEEE Conference on Computer Vision and Pattern Recognition.

[3] J Yan and A S El Ahmad. "Breaking Visual CAPTCHAs with Naïve Pattern Recognition Algorithms", in *Proc. Of the 23rd Annual Computer Security Applications Conference (ACSAC'07).* FL, USA, Dec 2007. IEEE computer society. pp 279-291.

[4] J Yan and A S El Ahmad. "A Low-cost Attack on a Microsoft CAPTCHA", School of Computing Science Technical Report, Newcastle University, England. Feb, 2008.

[5] J Yan and A S El Ahmad. "Is cheap labour behind the scene? - Low-cost automated attacks on Yahoo CAPTCHAs", School of Computing Science Technical Report, Newcastle University, England. Apr, 2008

[6] K Chellapilla, K Larson, P Simard and M Czerwinski, "Designing Human friendly human interaction proofs", ACM CHI'05, 2005.

[7] K Chellapilla, K Larson, P Simard and M Czerwinski, "Building Segmentation Based Human-friendly Human Interaction Proofs", 2nd Int'l Workshop on Human Interaction Proofs, Springer-Verlag, LNCS 3517, 2005.

[8] HS Baird, MA Moll and SY Wang. "A highly legible captcha that Resists segmentation attacks". Proc. of Second Int'l Workshop on Human Interactive Proofs (HIP'05), ed. by HS Baird and DP Lopresti, Springer-Verlag. LNCS 3517, Bethlehem, PA, USA, 2005.

[9] M Chew and HS Baird. "BaffleText: a human interactive proof". Proc. of 10th IS&T/SPIE Document Recognition & Retrieval Conference, 2003.

[10] AL Coates, H S Baird and RJ Fateman. "PessimalPrint: A Reverse Turing Test", Int'l. J. on Document Analysis & Recognition, Vol. 5, pp. 158-163, 2003.

[11] T Converse, "CAPTCHA generation as a web service", Proc. of Second Int'l Workshop on Human Interactive Proofs (HIP'05), ed. by HS Baird and DP Lopresti, Springer-Verlag. LNCS 3517, Bethlehem, PA, USA, 2005. pp. 82-96

[12] Jakob Nielsen. Usability 101: Introduction to Usability, 2003. Available at http://www.useit.com/alertbox/20030825.html.

[13] J Elson, JR Douceur, J Howell and J Saul. "Asirra: a CAPTCHA that exploits interest-aligned manual image categorization". Proceedings of the 14th ACM conference on Computer and communications security (CCS), 2007.

[14] The Cutest Human-Test: KittenAuth from The PCSpy.com

[15] Ragavi.V, Geetha.G, An Analysis of IMAGINATION CAPTCHA, 2nd International Conference on Signal and Image Processing , ICSIP 2009, pp 649-651, August 2009.

[16] J. Yan and A. S. El Ahmad. Usability of CAPTCHAs or usability issues in CAPTCHA design. In *SOUPS '08: Proceedings of the 4th symposium on Usableprivacy and security*, pages 44–52, New York, NY, USA, 2008. ACM.

[17] K Chellapilla, K Larson, P Simard and M Czerwinski, "Building Segmentation Based Human-friendly Human Interaction Proofs", 2nd Int'l Workshop on Human Interaction Proofs, Springer-Verlag, LNCS 3517, 2005.