

A Novel Gesture-based CAPTCHA Design for Smart Devices

Nan Jiang
Software Systems Research Centre
Bournemouth University
Fern Barrow, Talbot Campus
Poole, UK
BH12 5BB
njiang@bournemouth.ac.uk

Feng Tian
Creative Technology Research Centre
Bournemouth University
Fern Barrow, Talbot Campus
Poole, UK
BH12 5BB
ftian@bournemouth.ac.uk

CAPTCHAs have been widely used in Web applications to prevent service abuse. With the evolution of computing environment from desktop computing to ubiquitous computing, more and more users are accessing Web applications on smart devices where touch based interactions are dominant. However, the majority of CAPTCHAs are designed for use on computers and laptops which do not reflect the shift of interaction style very well. In this paper, we propose a novel CAPTCHA design to utilise the convenience of touch interface while retaining the needed security. This is achieved through using a hybrid challenge to take advantages of human's cognitive abilities. A prototype is also developed and found to be more user friendly than conventional CAPTCHAs in the preliminary user acceptance test.

CAPTCHA. Human Interactive Proof (HIP). Gestural interaction. Mobile devices. Ubiquitous computing.

1. INTRODUCTION

CAPTCHAs (Completely Automated Public Turing test to tell Computers and Human Apart) have become a very popular security measure adopted by web applications for preventing service abuse since the term was firstly coined in 2000 (Von Ahn et al., 2004). As the name implies, it is a type of Turing test utilising challenge-response authentication process to determine whether the user of a web application is a real person or a program. CAPTCHAs are used as Human Interactive Proof (HIP) hence the design goal of a CAPTCHA challenge is that it should be difficult for a computer but easy for a human to solve. Von Ahn et al. suggest that an effective CAPTCHA should reflect an AI-complete or AI-hard problem (2003) which cannot be solved by computers alone. For this reason, since humans are very good at visual perceptions, visual recognition challenges are widely used in forming CAPTCHA puzzles. An example can be found in Fig. 1 where users need to recognise and type the characters (words) shown on a randomly generated image or video.

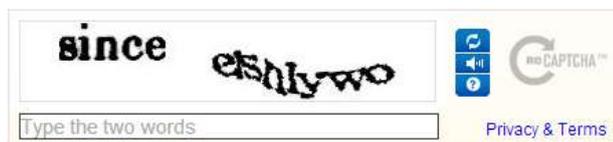


Figure 1: an example of ReCAPTCHA

Like ReCAPTCHA, most commercial CAPTCHAs are text recognition-based. These CAPTCHAs require additional text entry as the means of verification process when integrated into web applications. For example, in an account registration form where a CAPTCHA is used, a user must move their mouse cursor into the CAPTCHA's input field and then type the characters with keyboard after providing their registration details. This process is generally fine with computers and laptops where two-handed interaction style is popular (Buxon and Myers, 1986). However, it is not convenient when the user is using smart devices where single-handed, touch-based interaction style is dominant especially if they are moving (Poslad, 2011). With the evolution of computing environment from desktop computing to ubiquitous computing, more and more users are accessing web applications on these devices, which have made the need for designing user-friendly CAPTCHAs in such environments become more demanding.

This paper proposes a novel CAPTCHA design for use in ubiquitous computing environments, which offers intuitive touch based interaction support for smart devices. This is achieved through using a

hybrid challenge combining traditional text recognition puzzles with shape and motion puzzles to take advantage of human's cognitive abilities while keeping the response process simple.

The rest of the paper is organised as follows. Section 2 discusses related work and Section 3 describes the novel design. Section 4 provides a prototype and Section 5 shows preliminary test results. Section 6 is the conclusion and future work.

2. RELATED WORK

Most CAPTCHA designs are based on recognition challenges as humans are very good at visual perceptions. Popular visual recognition challenges used in CAPTCHAs include text-recognition (Chew and Baird, 2003; Imsamai and Phimoltares, 2010), object/image recognition (Elson et al., 2007; Gao et al., 2010; Matthews, et al., 2010) and audio recognition (Sauer et al., 2008) where text-recognition is the most widely deployed. Fig. 2 shows an example of Microsoft CAPTCHA where users need to recognise a combination of randomly generated and visually distorted characters (i.e., HMG6M3C8) on the image.

Help us make sure you're not a robot

Enter the characters you see

[New](#) | [Audio](#)



Figure 2: an example of Microsoft CAPTCHA

This reveals a major design issue with recognition-based CAPTCHAs as the majority of them require additional text entry as the means of the verification process. This process requires too much attention on smart devices especially if the user is moving. For example, Lin et al. (2011) found existing text recognition-based CAPTCHA systems presented in a web browser on mobile devices were awkward for users to select, zoom and respond to due to the varied screen size, resolution and unavailability of physical keyboard. In some cases when the characters required for human recognition are case sensitive and/or contain a combination of letters, numbers and symbols, the process will become even more complex.

An approach to simplify user text entry is integrating a reduced alphabet on-screen keyboard into a CAPTCHA system to replace its original text input field. This will accelerate the text entry process on the mobile devices as users only need to tap or click a limited range of characters on the

screen (Lin et al., 2011). However, recent studies reveal that humans are actually not good at solving this kind of challenges even when some hints are already provided (Bursztein et al., 2010; Fidas et al., 2011). A generally agreed explanation is that the usability of character recognition based CAPTCHAs has been significantly affected by the improved security measures (Yan and El Ahmand, 2008; Kluever and Zanibbi, 2009; Bursztein et al., 2011; Lee and Hsu, 2011; Yan, 2012; El Ahmand et al., 2012). To address this issue, new types of image recognition CAPTCHAs with enhanced interactive features have been proposed. Examples include Drawing CAPTCHA (Shirali-Shahreza and Shirali-Shahreza, 2006), puzzle-based CAPTCHA (Gao et al., 2010), semantic-based CAPTCHA (Vikram et al., 2011) and commercially adopted CAPTCHA systems like KeyCAPTCHA and AreYouHuman. On one hand, these CAPTCHAs offer better interaction support on smart devices as the response processes are purely based on direct input. In other words, they map to mouse interactions and touch based interactions perfectly. On the other hand, they bring some old and new issues. For example, the security of these CAPTCHAs are generally not as good as text recognition-based CAPTCHAs (SpamTech, 2012). In addition, the oversized visual presentation of these new CAPTCHAs will still require extra user attentions on smart devices due to device display limitations.

The novel design proposed in this paper considers a hybrid challenge by taking advantage of text-recognition CAPTCHAs for additional security measures and interactive CAPTCHAs for the convenience of interaction styles. This reflects the common pursuit of balancing the needs of security and usability in an effective CAPTCHA design.

3. DESIGN

The novel CAPTCHA design featuring a hybrid challenge is described in this section. The challenge, which contains two processing stages, combines text, object recognition and motion puzzles to improve the security of interactive CAPTCHAs while keeping the response process simple, fast and intuitive on smart devices. The first stage contains a text-recognition based reading comprehension task where a user needs to recognise the text presented on the screen and understand its meaning as it provides instruction for the next challenge. The second stage combines some shape and motion puzzles based on the context presented in the first stage. The whole process design is shown in Fig. 3. It should note that the motion puzzle is the core CAPTCHA challenge as it determines how users will respond and interact with the CAPTCHA on smart devices

while other puzzles are mainly introduced for security considerations.

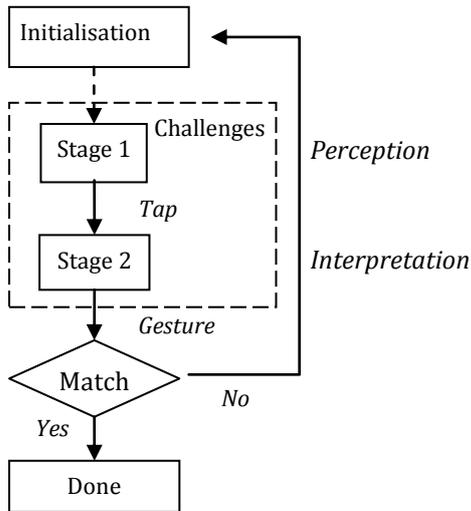


Figure 3: the process design

In Stage 1, the instruction is processed and presented in similar way as characters in a text recognition-based CAPTCHA as shown in Fig. 4. This is used to reflect a typical AI-complete problem domain: natural language understanding. The idea is: humans may not be good at recognising a set of randomly transformed characters without a context but they can easily perceive words from a context and understand the meaning of it (Rusu and Docimo, 2010). Certainly, machines may be able to identify transformed characters and even words by using advanced AI and image processing techniques. However, they cannot easily understand the meaning of a statement formed by words especially if (1) some characters in a word can be wrongly recognised and (2) it contains logical puzzles. Compared to pure clear text instructions, this text recognition based reading comprehension task will provide added security.

Drag the **largest** shape on the left
 into the **lightest** shape on the **right**
 without **touching** any other shapes

Figure 4: the processed instruction

In Stage 2, a shape and motion challenge is provided based on the instruction above where various transformed shapes are drawn on the screen and transformed using Geon principles (Biederman, 1987). This approach is similar to the shape CAPTCHA design proposed by Rusu and Docimo (2009). The benefits are obvious. First, the motion response can be easily achieved by using gestural interactions on touchscreens and mouse based interactions on conventional screens.

Second, this is more language-independent than (English) recognition challenges so it can be used worldwide with localised instructions. Last, this is not prone to the issues of many existing OCR-based CAPTCHAs (Yan and El Ahmad, 2008) (Xu et al., 2012). Fig. 5 shows an example of the shape and motion challenge. According to the instruction, the square numbered with 1 is the largest shape on the left and the square numbered with 2 is the lightest shape on the right. So the goal is to drag square 1 and drop it into square 2 without touching any other shapes. This is an easy task for users once they understand the instruction as the ability to understand shapes and basic graph structures is something that most humans can do easily.

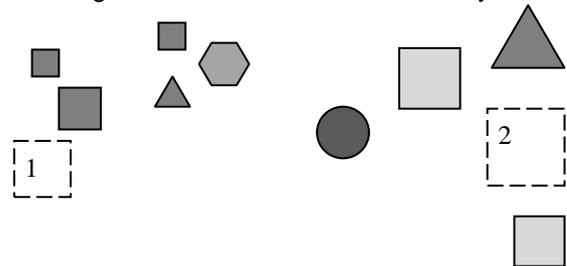
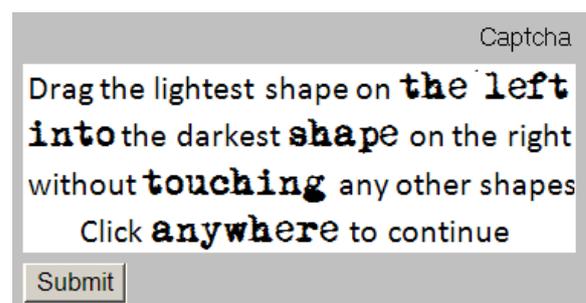


Figure 5: the shape and motion challenge

3. PROTOTYPING

A proof-of-concept called TapCHA is developed for demonstrating this novel design approach. The front end of TapCHA was implemented with HTML 5 Canvas and AJAX frameworks so it is compatible with all modern web browsers, which can be found on all smartphones and tablets. The CAPTCHA-like instruction was generated by using a customised CAPTCHA generator. The shapes and colours of available objects were dynamically generated and their positions on each side were also randomly decided to minimise potential security issues.

As shown in Fig. 6, a user needs to read the transformed text, understand the instruction and click or tap to load the real challenge. Once they performed the required motion, the CAPTCHA test will be passed. Otherwise, a new test will be loaded. In this example, the user actually needs to move the largest square on the left over the triangle without touching any other shapes in between them.



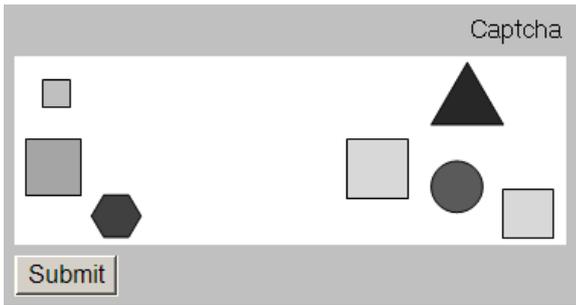


Figure 6: TapCha demonstration

4. USER ACCEPTANCE TEST

In order to understand whether TapCHA is easy to use, a preliminary user acceptance test was conducted. 17 final year students were volunteered for this study. They were asked to use TapCHA on tablets, smartphones and desktop computers for 15 minutes and complete an acceptance survey afterwards. The survey focused on three aspects of this novel design: (1) Ease of understanding, (2) Ease of completion/operation and (3) Speed of completion. The results are shown in Fig. 7. In general, despite two negative responses for (2) and (3), the general responses were positive. In detail, all 17 subjects think TapCHA is easy to understand, 15 subjects (7 highly positive and 8 positive) think it is easy to use and 7 (positive) think it is quick to complete.

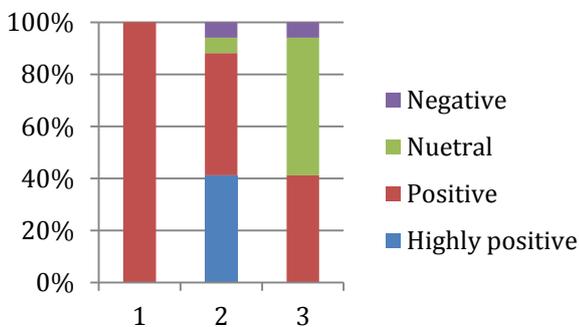


Figure 7: user acceptance survey of TapCha

5. DISCUSSION

TapCHA presents a novel CAPTCHA design which features a hybrid challenge combining text-recognition based reading comprehension task and shape and motion puzzles. By doing this, it transfers the recognition problem presented in many CAPTCHAs to an even larger problem domain where machine recognition is still low while human recognition remains strong. Compared to other existing CAPTCHAs, this design offers the following benefits. First, this CAPTCHA design does not require large databases as required by some image recognition CAPTCHAs such as Asirra

and IMAGINATION (Datta et al., 2005) as the shape and motion challenge can be randomly generated. Second, the display effects of instruction can be customised and localised with an existing CAPTCHA generator solution. This means it can be used with any language in the world. Third, it supports both smart devices and conventional devices such as computers and laptops as the response process is done through mouse or touchscreen interactions. An issue with TapCHA is that it currently only supports HTML 5 compliant browsers. A fallback solution for legacy browsers may need to be considered in the future. A common approach is to use Flash, which can be seen in KeyCAPTCHA. Another area of improvement is the accessibility of CAPTCHAs as the current design process presupposes intended human users do not have serious legibility and mobility issues. Additionally, a full usability test is also needed to compare the proposed solution to existing systems for a better understanding.

6. CONCLUSION

In this paper, we present a novel CAPTCHA design for overcoming the usability and security issues seen in the most widely used commercial CAPTCHAs when they are used in ubiquitous computing environments. This is achieved through using a hybrid challenge that leverages some of humans' advantageous areas of cognitive abilities such as understanding the context from transformed text, recognising shapes and moving them based on certain logical criteria. Some cognitive principles have also been exploited in the design to ensure that the CAPTCHA is readable, understandable and operable by humans, but not by machines. The preliminary test results for the proof-to-concept TapCHA show that real users feel positive to the design as they can easily understand and complete the challenge on smart devices in a relative short time frame. Certainly, full usability testing will be needed in order to improve the design for real-life deployment. It should also note that the main contribution of this paper is the design process rather than the CAPTCHA test itself. That is, the shape and motion puzzle is only used to demonstrate how different problem areas can be combined and reflected in a CAPTCHA challenge for balancing the needs of usability (for real user) and security (for computers). This means the approach can also be taken further by investigating the technical aspects of the CAPTCHA challenge for better optimisation. For example, how long the description should be or how many shapes should be used and so on. Moreover, the authentication process in the background which reflects the security of a CAPTCHA is very important to make sure such a

system can work flawlessly. Further study on this area should also be considered.

7. ACKNOWLEDGEMENTS

The authors are grateful to James Bliss for inspiring this research and to Dr. Huseyin Dogan for helpful discussions. Further thanks to the final year students at Bournemouth University who participated in the experiment. Thanks should also be given to BHCI 2013 reviewers for their constructive feedbacks and comments.

8. REFERENCES

- Biederman, I. (1987) Recognition-by-components: A Theory of human image understanding. *Psychological Review*, 94(2), pp.115-147.
- Bursztein, E., Bethard, S., Fabry, C., Mitchell, J.C. & Jurafsky, D. (2010) How good are humans at solving CAPTCHAs? a large scale evaluation. *2010 IEEE Symposium on Security and Privacy*, Oakland, CA, USA, 399-413.
- Buxton, W. & Myers, B.A. (1986) A Study in Two-Handed Input. *In Proceedings of the ACM CHI 86 Human Factors in Computing Systems Conference*, Boston, MA, USA, 321–326.
- Chew, M. & Baird, H.S. (2003) Baffletext: A human interactive proof. *In Electronic Imaging 2003*, 305-316.
- Datta, R., Li, J. & Wang, J.Z. (2005) IMAGINATION: a robust image-based CAPTCHA generation system. *In Proceedings of the 13th annual ACM international conference on Multimedia*, 331-334.
- El Ahmad, A.S., Yan, J. & Ng, W.Y. (2012) CAPTCHA Design: Color, Usability, and Security. *Internet Computing*, 16(2), 44-51.
- Elson, J., Douceur, J.R., Howell, J. & Saul, J. (2007) Asirra: a CAPTCHA that exploits interest-aligned manual image categorization. *CCS 7*, 366-374.
- Fidas, C.A., Voyiatzis, A.G. & Avouris, N.M. (2011) On the Necessity of User-friendly CAPTCHA. *In Proceedings of the 2011 annual conference on Human factors in computing systems (CHI)*, Vancouver, BC, Canada, 2623-2626.
- Gao, H., Yao, D., Liu, H., Liu, X. & Wang, L. (2010) A novel image based CAPTCHA using jigsaw puzzle. *In 2010 IEEE 13th International Conference on Computational Science and Engineering (CSE)*, Hong Kong, China, 351-356.
- Imsamai, M. & Phimoltares, S. (2010) 3D CAPTCHA: A next generation of the CAPTCHA. *In 2010 International Conference on Information Science and Applications (ICISA)*, Fukuoka, Japan, 1-8.
- Kluever, K. A. & Zanibbi, R. (2009) Balancing usability and security in a video CAPTCHA. *In Proceedings of the 5th Symposium on Usable Privacy and Security*, Mountain View, CA, USA, 14.
- Lee, Y.L. & Hsu, C.H. (2011) Usability study of text-based CAPTCHAs. *Displays*, 32(2), 81-86.
- Lin, R., Huang, S., Bell, G.B. & Lee, Y. (2011) A new captcha interface design for mobile devices. *In ACSW 2011: Australasian User Interface Conference*, Curtin, Australia.
- Matthews, P., Mantel, A. & Zou, C. (2010) Scene tagging: image-based CAPTCHA using image composition and object relationships. *In Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security*, Beijing, China, 345-350.
- Poslad, S. (2011) *Ubiquitous computing: smart devices, environments and interactions*. Wiley.
- Rusu, A. & Docimo, R. (2009) Securing the Web using human perception and visual object interpretation. *In 13th International Conference on Information Visualisation*, 613-618.
- Rusu, A., Docimo, R. & Rusu, A. (2010) Leveraging cognitive factors in securing WWW with CAPTCHA. *In Proceedings of the 2010 USENIX conference on Web application development*, Boston, MA, USA, 5.
- Sauer, G., Hochheiser, H., Feng, J. & Lazar, J. (2008) Towards a universally usable CAPTCHA. *In Proc. of the 4th Symp. On Usable Privacy and Security (SOUPS'08)*, Pittsburgh, PA, USA. 2008.
- Shirali-Shahreza, M. & Shirali-Shahreza, S. (2006) Drawing captcha. *In 28th International Conference on Information Technology Interfaces*, 475-480.
- SpamTech (2012) Cracking the AreYouAHuman Captcha. <http://spamtech.co.uk/software/bots/cracking-the-areyouhuman-captcha/> (August 9, 2013)
- Vikram, S., Fan, Y. & Gu, G. (2011) SEMAGE: a new image-based two-factor CAPTCHA. *In Proceedings of the 27th Annual Computer Security Applications Conference*, Orlando, FL, USA, 237-246.
- von Ahn, L., Blum, M., Hopper, N. & Langford, J. (2003) CAPTCHA: Using Hard AI Problems for Security. *In Proceedings of Eurocrypt*, Warsaw, Poland, 294-311.
- von Ahn, L., Blum M. & Langford, J. (2004) Telling Humans and Computers Apart Automatically. *Communications of the ACM*, 47(2), 56-60.
- Xu, Y., Reynaga, G., Chiasson, S., Frahm, J.F., Monroe, F. & Van Oorschot, P.C. (2012) Security and usability challenges of moving-object CAPTCHAs: decoding codewords in motion. *In 21st USENIX Security Symposium*, Bellevue, WA, USA.
- Yan, J. & El Ahmad, A.S. (2008) A Low-cost Attack on a Microsoft CAPTCHA. *In Proceedings of the 15th ACM conference on Computer and communications security*, Alexandria, VA, USA, 543-554.