The Usability of Captchas on Smartphones

Gerardo Reynaga and Sonia Chiasson
School of Computer Science, Carleton University, Ottawa, Canada
{gerardor, chiasson}@scs.carleton.ca

Keywords: CAPTCHAS: SECURITY: USABILITY: MOBILE: SMARTPHONES.

Abstract: Completely Automated Public Turing tests to tell Computers and Humans Apart (CAPTCHA) are challenge-response tests used on the web to distinguish human users from automated bots (von Ahn et al., 2004). In this paper, we present an exploratory analysis of the results obtained from a user study and a heuristic evaluation of captchas on smartphones; we aimed to identify opportunities and guide improvements for captchas on smartphones. Results showed that existing captcha schemes face effectiveness and user satisfaction problems. Among the more severe problems found were the need to often zoom and pan, and too small control buttons. Based on our results, we present deployment and design guidelines for captchas on smartphones.

1 INTRODUCTION

CAPTCHAs (denoted captcha) typically display distorted characters which users must correctly identify and type in order to proceed with a web-based task such as creating an account, making an internet purchase, or posting to a forum (von Ahn et al., 2004).

We use web services for a wide range of activities from banking to sharing data and socializing. The importance of web services is by now well established. More over, mobile devices such as smartphones and tablets have become a primary means of accessing these online resources for many users, but existing captchas do not properly fit mobile devices and lead users to abandon tasks (Asokan and Kuo, 2012). Finding an alternative captcha that addresses the usability issues while maintaining security has potential uses for any mobile website concerned with spam and bots.

Captchas have become sufficiently hard for users to solve that some web sites are actively looking at alternatives. A recent, February 2013, example is TicketMaster's decision to stop using traditional character-recognition captchas and move to a cognitive-based captcha (BBC, 2013). For users of smartphones, the problem is compounded by various factors: a reduced screen size can lead to typing mistakes (Kjeldskov, 2002), and loss of position (Bergman and Vainio, 2010). Environmental conditions and device handling positions also have an impact on the user experience (MacKenzie and Soukoreff, 2002).

In order to propose alternatives, it is important to discover where most of the problems lie. In this paper, we present an exploratory analysis of the results obtained from a user study and a heuristic evaluation of captchas on smartphones.

The main contributions of this paper are empirical results exploring the usability of four existing captcha schemes on mobile devices followed by a discussion of design recommendations applicable to future proposals. We collect our quantitative and qualitative results using two complementary evaluation methods to ensure a broader coverage. We find that existing schemes have significant usability problems that frustrate users and lead to errors. In their present state, captchas are unsuitable for mobile devices. Devising a suitable alternative remains an open problem but we hope that our findings help to guide such designs.

2 BACKGROUND

Captchas can be categorized according to the type of cognitive challenge presented. Character-recognition (CR) captchas involve still images of distorted characters; Audio captchas (AUD) use words or spoken characters as the challenge; Image-recognition captchas (IR) involve classification or recognition of images or objects other than characters; Cognitive-based captchas (COG) include puzzles, questions, and other challenges related to the semantics of images or language constructs. For both CR and IR, we further subdivide then into dynamic subclasses. That is, the CR-dynamic class encompasses dynamic movement

(a) reCaptcha (Google, Inc., 2013) (b) Asirra (Microsoft Inc., 2012) (c) NuCaptcha (NuCaptcha, Inc., 2012) (d) Animated (Vappic, 2012)

Figure 1: Target Schemes. reCaptcha (CR), Asirra (IR), NuCaptcha (MIOR), and Animated (MIOR).

of text as the challenge and the IR-dynamic class uses moving objects as the challenge. These two can be grouped as a cross-class category: moving-image object recognition captchas (MIOR) (Xu et al., 2012). While captchas have existed for some time and usability analysis has been done (*e.g.,* (Yan and El Ahmad, 2008; Bursztein et al., 2010; Wismer et al., 2012)), only limited work has been carried out to evaluate captchas for mobile device usage. To our knowledge, Wismer *et al.*(Wismer et al., 2012) provide the only evaluation of existing captchas on mobile devices and they found significant problems. Their evaluation focuses on voice and touch input using Apple's iPad.

Captcha proposals for mobile devices. Chow *et al.*(Chow et al., 2008) introduce the idea of presenting several CR captchas in a grid of clickable captchas. The answer is input by using the phone's (NOKIA 5200) keyboard and selecting the grid elements which satisfy the challenge. For example, the user may have to identify in the grid a subset of captchas with embedded words, as opposed to random strings. Since the answer consists of selection by clicking, this scheme could be used on mobile devices. Despite showing benefits, this captcha scheme has not been made public or implemented.

Gossweiler *et al.* (Gossweiler et al., 2009) present a IR captcha scheme that, although not designed for mobile devices, could be adapted for mobile usage. Their scheme consists of rotating an image to its upright / natural position with a slider. They suggest that the mobile version would allow direct image rotation with finger gestures.

Lin *et al.* (Lin et al., 2011) introduce two captcha schemes for mobile devices. The first is an IR scheme called "captcha zoo". It asks users to discriminate certain target animals from a set of containing two types of animals. For example, displaying dogs and horses, and the user clicks on the horses. The images are 3D models. The second proposal, a CR scheme, presents a four-character challenge with distorted letters and provides a small set of buttons with characters that include the answer.

3 OUR EVALUATION

Our motivation to conduct an evaluation of captchas on smartphones was to identify usability problems in representatives of the main categories of existing schemes. Rather than a summative evaluation, our evaluation is a formative evaluation to explore the gaps, identify opportunities, and guide improvements for captcha schemes for mobile devices.

We conducted two types of evaluations on four different captcha schemes. The first evaluation consisted of a user study. The second evaluation was a heuristic evaluation. The goals of the studies were to assess the following aspects: 1) the effectiveness of captcha schemes on smartphones, and 2) the user's experience of captchas on smartphones. The four captcha schemes are described below, Figure 1.

The schemes selected for evaluation were chosen because they are a good representation of each of the main captcha categories: CR, IR, and MIOR.

reCaptcha (Google, Inc., 2013) is a free service that is widely deployed on the Internet. The CR challenge consists of recognizing and typing two words.

Asirra (Microsoft Inc., 2012) is a research IR captcha from Microsoft and it is provided as a free captcha service. The challenge consists of asking users to identify images of cats and dogs.

NuCaptcha (NuCaptcha, Inc., 2012) is a commercial MIOR scheme. The challenge consists of either reading alphanumeric characters that overlap as they swing independently left to right (statically pinned at the centre of each letter), or reading a code word in a phrase that loops endlessly in the captcha window.

Animated captcha, (Vappic 4D) (Vappic, 2012), is an experimental captcha. The MIOR challenge typically consists of six alphanumeric characters arranged in a patterned cylinder that rotates in the centre of the captcha screen. The similarly patterned background portrays what could be the floor (or base) where the cylinder sits; this floor swivels up and down.

4 USER STUDY

Study Design. The user study was done in a controlled environment. Each participant completed a one-on-one session with the experimenter and the session was video and audio taped. The participants responded to a demographics questionnaire and a satisfaction survey. Their performance measurements were limited to noting the number of successes, skips/refresh, and errors while answering the challenges. A within-subjects experimental design was used, where each participant attempted ten challenges for each scheme. Participants received random challenges from the respective demo sites. Participants were paid $15 honorarium for their cooperation. The solving order for the schemes was determined by a 4×4 Latin Square.

Participants. Ten participants were asked to complete challenges on either a provided smartphone or their own smartphone. The participants (5 females, 5 males) were graduate and undergraduate students with diverse background, university staff, a private company IT employee, faculty members and a freelance employee. They ranged in age from 18 to 44, mean age of 32 years old. None had participated in any prior captcha studies. The average self-reporting expertise using smartphones was 6.33 out of 10. The average phone ownership was 3.3 years. All except two had encountered captchas before the study.

Procedure. The study protocol consisted of the following steps: 1) Briefing session. We explained the goals of the study, detailing the study steps, and asking them to read and sign the consent form. 2) Demographics questionnaire. Before solving the challenges, participants answered a demographic questionnaire. 3) Captcha testing. Participants visited a host page with links to the four schemes located on third party demo sites from the smartphone. 4) Satisfaction questionnaire. After each scheme, participants completed an online satisfaction questionnaire collecting their satisfaction and opinion of that scheme.

Equipment and software. Seven participants used an Android OS (ver. 2.3.6) smartphone and three used iOS (iOS 4.0). The demographics and satisfaction questionnaires were implemented using Limesurvey[1]. We chose not to implement our own version of the schemes due to two main reasons: first, visiting the original demo sites allowed testing of the latest version of the schemes; second, we did not have access to implementation and deployment details which could impact the behaviour of the schemes.

Ethics approval. This research has been approved by our institution's Research Ethics Board.

Audio Captcha Pilot Test. We pilot tested audio schemes from several major websites. We realized that audio schemes are currently unusable on smartphones due to their high operational complexity and strong need for recall, so discontinued them from our tests. Specifically, we found that the audio would open on different window or tab, the audio would open on a different application, or the audio decoder was not supported.

[1] LimeSurvey http://www.limesurvey.org/

5 USER STUDY RESULTS

We now present the results from our usability study. From this study we collected performance data, usability problems and perceived qualitative indicators. We do not report statistical analysis because our goal was to formatively identify strengths and weaknesses, not to compare the schemes against each other.

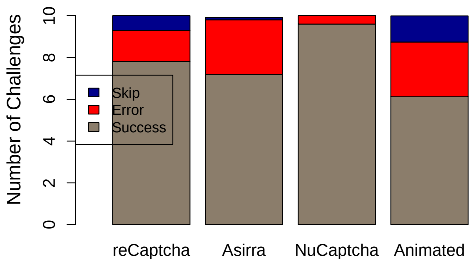


Figure 2: User Study. Mean number of success, error and skipped outcomes.

5.1 Performance

We counted the number of successes, skips/refresh, and errors while answering the challenges (Figure 2). We counted a *success* when a user's answer to the challenge is deemed correct by the demo site. An *error* was counted when the user's response did not match the challenge's solution and was indicated as incorrect by the demo site. A *skipped* outcome was counted when the participant pressed the "Request new images", "Get A New Challenge" or "Skip" button and was presented with a different challenge.

NuCaptcha shows the most successful outcomes compared to the other schemes, followed by reCaptcha. A possible explanation is that challenges for NuCaptcha consisted of only three characters with no distortion, while reCaptcha uses distortion on only one of the two words. However, we noted some participants were flipping the phone from landscape to portrait mode a few times, attempting to find the best fit to see and answer challenges without panning. Asirra requires selecting images which expand and obscure other images, forcing users to pan across the challenge; we observed that this was the cause of a large number of errors. Animated demands considerable attention from users. We noticed participants often shifting their sitting position and handling of the phone while solving this scheme and verbally indicating their discomfort. We observed that Animated's movement exacerbates the known issue of confusable characters and thus participants were prone to typing errors and requests for new challenges.

5.2 Usability Problems

Two researchers watched and coded the videos of the testing sessions. Usability problems were identified and summarized through an iterative process where the researchers reached mutual consensus of the main categories of problems and identified the most serious issues. We group the usability problems uncovered by the user study in six groups:

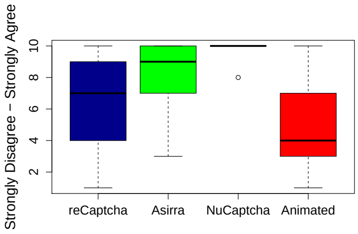
1. Small buttons: Participants found control buttons (skip, audio, help) too small and sometimes they pressed these by mistake.
2. Interface interaction: Input interaction can interfere with answering challenges. Some IR schemes require tapping on images. While solving challenges in Asirra, participants found the scheme's zooming mechanism obscured other thumbnails. We believe that most of the usability problems with this scheme are due to the scheme's auto-zoom feature that blocks other images and forces unnecessary panning and zooming.
3. Confusing characters/images: Captchas are by nature somewhat confusing to solve, but the problem is compounded on small screens. We observed participants confusing characters (*e.g.,* `1/i/l`) and confusing images of dogs with those of cats primarily because the small image made it difficult to identify details.
4. Inefficient schemes: Several participants pointed out that the challenges were so small that they needed to zoom and pan across the screen to locate and reply to them. Some tasks were tedious, time consuming, and frustrating to solve on a smartphone. CR schemes sometimes mixed alphanumeric characters, forcing users to swap between input keyboards.
5. Data plans: Several participants were concerned about data transfer due to costly data plans. Schemes that are image or video intensive are probably not good options for mobile devices.
6. Lack of instructions: We observed, and heard from, participants not knowing if CR challenges were case sensitive or not, or being unsure if spaces were required for challenges with two words, being unsure how to clear previous image selections, and being confused about where the challenge started. Deselecting images in Asirra required double tap on the image under iOS, where as Android required a single touch for selecting or deselecting. Instructions were not immediately apparent to users as they struggled with the interface problems.

In summary, the most severe problems were found due to the small buttons, the interface interaction (input mechanisms) and confusing characters/images.

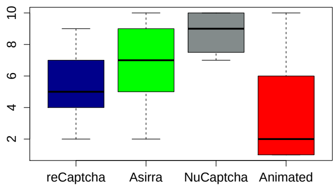
5.3 User Satisfaction

Users answered a number of Likert scale questions about each scheme. The collected satisfaction results are graphed in Figure 3. The results show that users clearly favoured the NuCaptcha scheme and rated the others lower on all subjective measures. We speculate that NuCaptcha is favoured over the rest due to its lack of distortion and short challenges which were considerably easier than the other schemes (although also least secure). Animated was clearly the most disliked scheme.

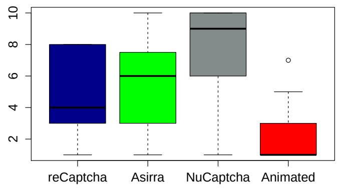
Participants had the opportunity to provide free-form comments about each scheme and offer verbal comments to the experimenter. We highlight a few comments about each scheme.
reCaptcha: "Text entry on smart phones needs to be mastered better", "Challenges are long to type for a mobile device keyboard"
Asirra: "The number of images presented became crowded on my phone", "it was to big! I want to see things on one screen, don't like to move so much".
NuCaptcha: "...the letters didn't move at all so it was very ease for me and the attackers!", "When you enter the text I can press the keypad enter or the captcha button, didn't know which one to press at first."
Animated: "It hurts my head - it requires too much thought...", "The captcha controls and smartphone input mechanism were overlapping."



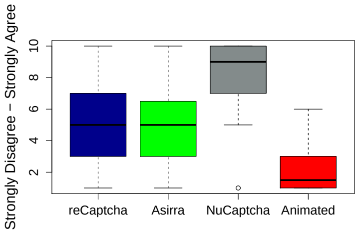
(a) Easy to understand



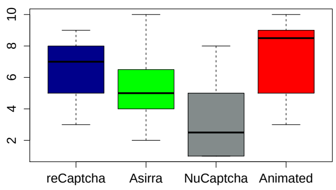
(b) Pleasant to use



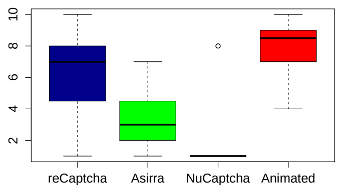
(c) Preference over other schemes



(d) Good candidate for smartphones



(e) More error-prone



(f) Harder to solve challenges

Figure 3: User Study. Likert-scale responses: 1 is Strongly Disagree, 10 is Strongly Agree.

Overall participants preferred schemes that involve quick, simple challenges and little or no distortion. Participants disliked ambiguity on the challenge itself or while replying to challenges.

6 HEURISTIC EVALUATION

Background. A heuristic is an abstraction of a guideline or principle that can provide guidance at early stages of design; or be used to evaluate existing elements of a user interface (Sharp et al., 2007). According to Nielsen, "heuristic evaluation is the most popular of the usability inspection methods" (Nielsen, 2013). A heuristic evaluation (HE) does not require the researcher to be present while the evaluation is ongoing. The heuristic evaluation includes the following steps: preparing the target software (*i.e.,* captcha schemes), briefing session (the experts are told what to do, using a prepared script), evaluation period (experts go over the system a few times using the heuristics as guide to evaluate, note the usability problems found and rate their severity), debriefing session (when possible experts get together and discuss their findings, reassign priorities if needed and suggest solutions). In our case, the experts were geographically dispersed and we did not want to use more of their time, so two experimenters completed the last step. This approach is often taken when conducting HE in a research environment. Typically, 5 - 10 experts participate in a HE (Nielsen, 2013).

HE Design. To conduct the heuristic evaluation, we first developed a set of seven domain specific heuristics. The heuristics cover the usability and deployability of captchas. Usability heuristics evaluate issues such as challenge obstruction, typing, and restricted screen space. Deployability deals with language, culture and universality. The evaluation was done by requesting expert evaluators to use our heuristics to evaluate the four captcha schemes described above (§3). To recruit experts, we sent e-mails to a list of known people with Human Computer-Interaction (HCI) and security background. Once experts agreed, they were sent an e-mail introducing the heuristics and providing instructions on how to conduct the evaluation. Experts' self-assessed mean for security was 4.1 and for HCI was 4.2 (out of 5, $N = 9$). Experts were not given an honorarium, they volunteered their time to conduct the evaluation.

Procedure. Our HE design allowed experts to solve challenges and explore the overall interface. Nine experts completed the evaluation of at least one scheme. Experts used the same host web page and live versions of the captcha schemes as the user study. Each expert conducted his/her assessment independently. For each scheme visited on the expert's smartphone, they assessed its merits based on the heuristics, noted in Limesurvey any problems uncovered, rated each problem's severity, and provided an overall rating of the problems found based on the heuristic.

Equipment and software: All of the experts completed the evaluations on their own smartphones, and the environment of their choosing. There was one Nexus S, one Galaxy Nexus, 2 iPhones 4, 4 iPhones

Table 1: Unique problems for HE and user study.

Scheme	Only HE	Matching	Only User Study
reCaptcha	32	11	11
Asirra	30	7	15
NuCaptcha	18	3	9
Animated	30	9	14

4s, and one Samsung Focus (SGH-i917). Limesurvey was used as the tool to collect experts' feedback and severity ratings.

7 HEURISTIC EVALUATION RESULTS

To tabulate the set of usability problems obtained from the HE, we summarized the usability problems identified by each expert and then generated an aggregate list of problems per scheme. We used a variant of Grounded Theory (Charmaz, 2006) to synthesize, consolidate, and categorized the reported problems.

7.1 Unique Problems and Severity Ratings

Table 1 depicts the number of unique problems for the heuristic evaluation and the user study. Matching problems are those that both the HE and the user study found. We note that HE proved to be more effective at finding issues than the user study. We believe that this is due to the heuristics motivating experts to inspect in more detail than simply solving challenges since their task was specifically to find problems. In contrast, the user study participants' task was to solve challenges. To see which method was most effective, we intentionally kept the number of participants similar in the two evaluations.

The mean severity ratings assigned by experts to unique problems is as follows. reCaptcha: 2.5, Asirra: 2.78, NuCaptcha: 3.1, Animated: 2.17. Where 1 represents critical usability issues and 5 represents minor issues. Expert evaluators rated NuCaptcha as having less severe problems and uncovered fewer unique problems for this scheme.

7.2 Highlighted Problems

Below list samples of the unique problems uncovered by experts. The problems are mainly grouped as in §5.2 to help with comparisons.

1. Small buttons: Experts found that typing in input fields zooms on the text box and this obscures the challenge. Experts had difficulty zooming to the right level to see entire challenge. Others pointed out that there was no deselect-all option and that there was insufficient control over speeds, orientation, and position.
2. Interface interactions: Experts remarked that auto-correct sometimes mistakenly "fixes" non-english words. They noted that it is hard to click on small images, and that the input box is small and users may hit other buttons by mistake. Once experts started typing they could not see the captcha challenge and type at the same time.
3. Inefficient schemes: The problems that experts found include needing excessive zooming and panning, selecting the input box is time consuming when is out of the screen, and some challenges are long to type for a mobile device keyboard.
4. Confusing characters/images: Experts observed that there is difficulty recognizing images or challenges without zooming due to the small screen.
5. Localization and context of use: Experts remarked that some challenges may be difficult to solve in direct sunlight. Experts also found that some challenges had non Roman characters, and low colour contrast.
6. Lack of instructions: Experts uncovered problems such as schemes having no instructions about case sensitivity or no indicator that the audio prompt words differ from the image. In some schemes instructions displayed on a new window which is challenging to navigate on a mobile browser.

8 DISCUSSION

While the user study provided insight into user's satisfaction of the schemes, the HE gave us more detailed feedback on the problems found when using the schemes on smartphones. We found that the issues raised by the two studies were similar and confirmed each other even though they may have been expressed differently.

Regarding the user study, we observed differences in the participants' outcomes, with NuCaptcha scheme being most successful, and Animated resulting in the least successful outcomes. The most skipped outcomes were observed for the Animated scheme. NuCaptcha was found the most pleasurable, while Animated was rated the least. We note that the satisfaction results for the user study are only a reflection of users' comparison among these four schemes;

positive scores are not necessarily an indication that schemes do not have usability problems on smartphones.

Experts found the most severe problems in areas relating to efficiency of use, and supporting interface interactions (input mechanisms) for easier response. Most severe problems relate to zooming and panning to be able to fully see and answer the challenge thus affecting the efficiency of the captcha. Experts also indicated that restrictions on input mechanisms considerably hinder the usability of the evaluated captchas.

Although NuCaptcha's outcomes for the user study and the HE showed favourable results, we remark the following standing issues. Regarding its security, NuCaptcha has recently been broken, along with several potential improvements to the scheme. It is not advisable to use it as a security mechanism at this time (Xu et al., 2012). NuCaptcha provides a clear example of a security mechanism that meets usability criteria but does not provide adequate security, therefore failing to meet its intended purpose. When designing security mechanisms that involve users, both usability and security must be given equal attention. In some cases usability problems lead to decreased security as users find ways to circumvent the security system. In other instances, such as with captchas, usability problems lead users to abandon the related primary task which is equally problematic for websites who lose business as a result.

We have developed recommendations for captcha deployment and design. Besides usability, security guidelines always have to be followed and evaluated before deploying or adopting any scheme. We separate them for discussion but some of the recommendations are applicable to both deployment and design.

Deployment recommendations. For administrators of any mobile website concerned with bots, it is more efficient to deploy an existing captcha scheme than develop a new one. Thus we list deployment recommendations to consider before adopting a scheme:

- Avoid keyboard switching and confusable characters (*e.g.,* `1/l, 6/G/b, 5/S/s, nn/m, rn/m`) since these are specially problematic on smartphones.
- Take into account browser capabilities and limitations (*e.g.,* past Flash support on Apple's devices).
- Avoid current audio captcha schemes. As discussed in §4 these are unusable on smartphones.
- Render challenges appropriate for mobile devices. Large challenges will cause the user to lose overview, while small challenges force zooming.
- Test on a wide variety of configurations since the differences in hardware and OS impact usability.

Design recommendations. In addition to reviewing past security design recommendations for captchas (Bursztein et al., 2011; Zhu et al., 2010; Yan and El Ahmad, 2008), we recommend the following considerations specifically for mobile captcha design:

- Follow HCI standards to give the user adequately-sized targets for touch interactions. Captcha controls should follow established mobile standards.
- Consider the ever-changing usage context of mobile devices such as using while standing, sitting, or walking. The device may also be operated with one or two hands. Lighting conditions have particular impact on low contrast challenges. These factors impact the input process, therefore they lead to input mistakes.
- Instructions need to be minimal due the real-estate constraints of smartphones.
- Follow known interaction standards, when possible maintain consistency between platforms so that users may transfer experience with desktop captchas to the mobile environment.
- Take into account network and bandwidth usage for challenge and reply transmissions.
- Avoid designing schemes that require the user to zoom and pan.

Based on our experience and study results we believe that these are valid recommendations. However, as future work includes confirming their applicability.

Implementing and designing a captcha from scratch is not a trivial task. Moreover, the design and implementation of schemes by non-experts is typically weak. This occurs because of the lack of knowledge on current threats and flaws in the scheme's design. Furthermore, subscription to captcha services or installing libraries that provide captcha schemes may not be flexible or configurable enough to adapt to multiple environments (*i.e.,* mobile devices). These services and implementations are commonly one-size-fits-all solutions. Finding a suitable alternative for mobile devices remains an open problem. We hope that this work helps to guide possible solutions.

9 CONCLUSIONS

This paper presents the results of two usability studies, a user study and a heuristic evaluation, of captchas on smartphones. This work is an important step aimed at understanding user frustration common to existing and deployed captchas on smartphones. Our results suggest that participants preferred schemes that involve quick, simple challenges with little or no distortion. Unfortunately existing captcha schemes that were preferred by users fail to provide adequate security.

Participants had some success with completing the challenges on all four schemes, but struggled with more complex challenges. User feedback, the HE and our analysis of the session videos indicate frustration with inappropriately sized interface elements: controls that are too small and challenges that are larger than the available screen size.

This paper represents the first empirical work identifying the main usability issues with existing captchas on smartphones. Considering the prevalence of these devices for web access, it is important to address this compelling usable security issue. We identify what works, what does not, and provide recommendations for the next generation of mobile captchas.

ACKNOWLEDGEMENTS

Sonia Chiasson holds a Canada Research Chair in Human Oriented Computer Security acknowledges the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding the Chair and a Discovery Grant. The authors also acknowledge funding from NSERC ISSNet and thank P.C. van Oorschot for his valuable feedback on this project.

REFERENCES

Asokan, N. and Kuo, C. (2012). Usable mobile security. In *Distributed Computing and Internet Technology*, volume 7154 of *Lecture Notes in Computer Science*, pages 1–6. Springer Berlin Heidelberg.

BBC (Accessed: Feb 2013). Ticketmaster dumps 'hated' captcha verification system. Available from http://www.bbc.co.uk/news/technology-21260007.

Bergman, J. and Vainio, J. (2010). Interacting with the flow. In *International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '10, pages 249–252, NY, USA. ACM.

Bursztein, E., Bethard, S., Fabry, C., Mitchell, J. C., and Jurafsky, D. (2010). How good are humans at solving CAPTCHAs? A large scale evaluation. In *IEEE Symposium on Security and Privacy*, pages 399–413. IEEE Computer Society.

Bursztein, E., Martin, M., and Mitchell, J. C. (2011). Text-based captcha strengths and weaknesses. In *ACM Conference on Computer and Communications Security*, pages 125–138. ACM.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage Publications Limited.

Chow, R., Golle, P., Jakobsson, M., Wang, L., and Wang, X. (2008). Making captchas clickable. In *Workshop on Mobile computing systems and applications*, HotMobile '08, pages 91–94, NY, USA. ACM.

Google, Inc. (2013). reCaptcha: Stop Spam, Read Books. http://www.google.com/recaptcha.

Gossweiler, R., Kamvar, M., and Baluja, S. (2009). What's up CAPTCHA?: a CAPTCHA based on image orientation. In *International conference on World wide web*, WWW '09, pages 841–850, NY, USA. ACM.

Kjeldskov, J. (2002). "Just-in-Place" information for mobile device interfaces. *Lecture Notes in Computer Science*, 2411:271–275.

Lin, R., Huang, S.-Y., Bell, G. B., and Lee, Y.-K. (2011). A new captcha interface design for mobile devices. In *ACSW 2011: Australasian User Interface Conference*.

MacKenzie, I. and Soukoreff, R. (2002). Text entry for mobile computing: Models and methods, theory and practice. *Human–Computer Interaction*, 17(2-3):147–198.

Microsoft Inc. (2012). Asirra (Animal Species Image Recognition for Restricting Access). http://research.microsoft.com/en-us/um/redmond/projects/asirra/.

Nielsen, J. (2013). Heuristic evaluation. Available from http://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/.

NuCaptcha, Inc. (2012). Available from http://www.nucaptcha.com/resources/whitepapers. White paper: NuCaptcha and Traditional Captcha.

Sharp, H., Rogers, Y., and Preece, J. (2007). *Interaction Design: Beyond Human-Computer Interaction*. John Wiley & Sons, Indianapolis, IN, 2 edition.

Vappic (2012). 4D CAPTCHA. http://www.vappic.com/moreplease.

von Ahn, L., Blum, M., and Langford, J. (2004). Telling humans and computers apart automatically. *Commun. ACM*, 47:56–60.

Wismer, A. J., Madathil, K. C., Koikkara, R., Juang, K. A., and Greenstein, J. S. (2012). Evaluating the usability of captchas on a mobile device with voice and touch input. In *Human Factors and Ergonomics Society Annual Meeting*, volume 56, pages 1228–1232. SAGE Publications.

Xu, Y., Reynaga, G., Chiasson, S., Frahm, J.-M., Monrose, F., and Van Oorschot, P. C. (2012). Security and usability challenges of moving-object CAPTCHAs: Decoding codewords in motion. In *USENIX Security Symposium*, Berkeley, USA. USENIX Association.

Yan, J. and El Ahmad, A. S. (2008). Usability of CAPTCHAs or usability issues in CAPTCHA design. In *Symposium on Usable Privacy and Security*, SOUPS '08, pages 44–52, New York, NY, USA. ACM.

Zhu, B. B., Yan, J., Li, Q., Yang, C., Liu, J., Xu, N., Yi, M., and Cai, K. (2010). Attacks and design of image recognition captchas. In *Computer and Communications Security*, CCS '10, pages 187–200, New York, NY, USA. ACM.