ABSTRACT

We argue that a set of usability heuristics are needed for easy and quick evaluation of Captchas implementations. With this set of heuristics we contribute to sustain the Captcha Mantra: “Easy for humans, hard for machines”. In particular, the usability of Captcha schemes change radically when utilized on mobile environments. We are developing a set of heuristics for use by practitioners wishing to evaluate which Captcha scheme is most appropriate for their website.

1. POSITION

We argue that a set of usability heuristics are needed for easy and quick evaluation of Captcha implementations. A Captcha is a program that generates and grades challenges that are human solvable, and should be unsolvable by current computer programs [19, 21]. They are typically used on websites to deter automatic programs (e.g., bots) from abusing web applications, to prevent e-mail harvesting, to avoid automated voting in Internet polls, and other applications that may require online automatic human verification [15, 11, 9, 10, 12, 13]. A challenge refers to a single Captcha puzzle to be solved by the user.

Attacks on Captcha schemes and new proposals are frequent and common [20, 13, 6, 8, 11]. Diligent site administrators may want to update their Captcha challenges based on news of such attacks, but it can be difficult to choose an appropriate replacement. Our goal is to provide an evaluation methodology to help administrators make such decisions. However, heuristic evaluation cannot quantify the security of Captcha schemes. Therefore, an acceptable level of security of a Captcha scheme has to be evaluated as part of the overall decision process.

Small changes to Captcha schemes may not cause obvious problems, nevertheless these changes may affect their overall usability in ways that are unexpected. In particular, changes may be acceptable for desktop or laptop usage, but may cause difficulties for other modalities such as smartphones. With the increase use of mobile devices [12], website designers must consider the usability impact of design choices in this growing segment of users.

The strength of heuristic evaluation is that it is cheap, quick and easy to carry out. It does not require a user study with a large number of users. A few people, knowledgeable in both the domain area and interaction design, are recruited to conduct the evaluation and no special facilities are needed for the heuristic evaluation.

While existing heuristics, such as Nielsen’s [16], Jaferian’s [14], and Zhou’s [22], provide a good set of heuristics, these are insufficient to evaluate Captchas. Nielsen’s heuristics are too general, and in a mobile environment they may find more cosmetic problems rather than critical problems. In fact, Nielsen suggests developing domain-specific heuristics that apply to a specific category of products. Jaferian’s and Zhou’s, although developed for the security domain, evaluate security management tools and intrusion detection systems, respectively. As opposed to these software programs, Captchas are used in a variety of environments, including mobile, causing a set of problems that are substantially different. The usability work done for Captchas focuses mainly on challenge design betterment, or design of innovative scheme proposals other than text-based Captchas [7, 8, 10, 21, 2].

We are developing a set of domain-specific heuristics for evaluating Captcha schemes. The main goal of the proposed heuristics is to assess Captcha scheme deployment targeting smartphones. The proposed heuristics cover the usability and deployability of Captcha schemes. For example, usability heuristics may include Input mechanisms and Solvability. Deployability includes Consistency with user’s localization and environment. Usability heuristics evaluate issues such as challenge obstruction, typing, restricted screen space. Deployability deals with language, culture and universality.

We have started an evaluation of Captchas on mobile devices using our proposed set of heuristics with evaluators having expertise from the usability and security. In addition, we are running a small user study, not involving the heuristics, to compare results between the usability study and the expert evaluation.

2. ACKNOWLEDGMENTS

I am working on this research in collaboration with my advisors Sonia Chiasson and Paul van Oorschot. The author acknowledges the NSERC Internetworked Systems Security Network (ISSNet) funding.
3. REFERENCES


