Poster: Adaptcha: An Adaptive CAPTCHA for Improved User Experience

Brian M. Powell[1], Richa Singh[1,2], Mayank Vatsa[1,2] and Afzel Noore[1]
[1]West Virginia University, [2]IIIT-Delhi
{brian.powell, afzel.noore}@mail.wvu.edu, {rsingh, mayank}@iiitd.ac.in

Abstract—**CAPTCHAs are an important security tool for preventing automated attacks against online systems. However, they can be an undesirable source of friction in the user experience. In this poster, we propose a novel approach to dynamically improve the CAPTCHA experience for users by individually learning which types of CAPTCHA tests each user is best at solving and adapting future tests presented based on this knowledge.**

1. Introduction

"Completely Automated Public Turing Tests to Tell Computers and Humans Apart" (CAPTCHAs) are a type of security test that can be used to distinguish between human users and automated attackers. They use a challenge-response mechanism where would-be users are presented with tests that are intended to be easy for humans to solve but difficult for software bots to complete [1]. Unfortunately, the line between easy and difficult has become blurred as automated attacks have become more adept at performing the tasks necessary to solve CAPTCHAs. One approach to ensure resiliency against automated attack is to increase the difficulty of CAPTCHA tests; however, this results in a reduced user experience for legitimate human users. The resulting time and inconvenience for users, known as *user experience friction*, can discourage individuals from accessing systems protected by CAPTCHAs [2], [3].

An alternative approach to reducing friction is to simplify or eliminate CAPTCHA tests. Google's recent No CAPTCHA reCAPTCHA and Invisible reCAPTCHA are examples of this approach, presenting users with a simplified test or skipping the test altogether if the system believes the user is human [4], [5]. This approach results in a low friction experience but it also eliminates most of the security benefits of CAPTCHAs. Software bots have been shown to have high accuracy in attacking reCAPTCHA, with even electromechanical robots having been recorded defeating the system [4], [6].

This poster proposes Adaptcha, a novel approach designed to provide a low friction user experience while not compromising the security benefits of CAPTCHAs. Adaptcha works by learning about how each user completes CAPTCHA tests and using that knowledge to identify types of CAPTCHAs and specific CAPTCHA tests the user can likely solve successfully and quickly on the first attempt. This reduces friction and improves the user experience. In preliminary evaluation, Adaptcha improves success rates by 8 percentage points while reducing the time required to solve CAPTCHA tests by more than one-third.

2. Proposed Approach

Adaptcha's CAPTCHA selection process involves two phases: user performance initialization and adaptive test selection. In the initialization (first) phase, selected CAPTCHAs are used to learn the individual user response and customize future tests for each user. As the user completes more CAPTCHAs, in the second stage, the algorithm learns more about individual user performance and shows tests for which the user is more likely to easily solve. As compared to other solutions for adaptive CAPTCHAs, the proposed Adaptcha does not require any data from outside the CAPTCHA for its work in contrast to proposals by Belk, which use cognitive style tests [7], and reCAPTCHA, which relies on external data collected from Google services [8].

2.1. User performance initialization

When a user first starts using Adaptcha, the algorithm lacks information needed to make intelligent decisions about which CAPTCHA tests to select for the user. Thus, as the user access resources protected by Adaptcha, the algorithm begins by presenting randomly selected tests from each of the n CAPTCHA types used by the Adaptcha instance, where n is generally between 3 and 5 types. This continues until, in the normal course of use, the user has successfully completed 3 tests from each type. This provides an initial base of knowledge about user performance on which Adaptcha can make intelligent decisions for its adaptive selection process.

2.2. Adaptive test selection

Once $n \times 3$ tests have been successfully completed, Adaptcha is ready to begin adaptively selecting the tests presented to each user. Adaptive selection is phased in so users begin to see the benefit of adaptive selection while the algorithm continues to learn about the user's performance characteristics. Whether a user sees an adaptively selected or a randomly selected test is determined by

$$Selection = \begin{cases} Adaptive, & \text{if } rand[0,1) \leq min(\frac{c}{50}, 0.95) \\ Random, & \text{otherwise} \end{cases} \tag{1}$$

where, c is the number of tests the user has completed. At least 5% of tests are selected at random to ensure some variability in tests as a guard against automated attacks. If adaptive selection is used, a fitness value is calculated for each CAPTCHA type such that

$$f = (0.8s_{average}) + (0.2t_{average}) \tag{2}$$

Here $s_{average}$ represents the average success rate for completed tests of that type, $t_{average}$ is the average completion time globally normalized to $(0,1]$, and f is the resulting fitness value. The CAPTCHA type is selected by a roulette wheel-based process using these calculated fitness values.

Once a CAPTCHA type has been selected, an individual test from that type must be chosen for the user. If the CAPTCHA tests are keyword-tagged, a fitness value is calculated for each keyword as in Eq. 2 and a roulette wheel-based selection process is used to choose a keyword. A test associated with the selected keyword is chosen at random. If the CAPTCHA tests are not keyword-tagged, the test is chosen at random from all of the tests of the selected type. Fitness values are calculated anew for each test. Adaptcha continually learns and improves its selections to provide a better user experience as users complete more tests.

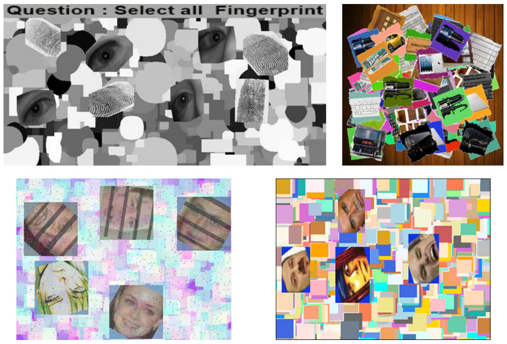

Figure 1. CAPTCHA tests used in Adaptcha experiments.

3. Experimental Results and Analysis

A group of 51 volunteers participated in evaluating performance of Adaptcha using a variety of computing devices. A total of 3,319 tests were recorded using CAPTCHAs from the fgCAPTCHA [9], FaceDCAPTCHA [10], MB-CAPTCHA [11], and aiCAPTCHA [12] types shown in Fig. 1. However, the proposed Adaptcha can be generalized to work with any type of CAPTCHA test. Table 1 shows the results in three phases: initialization with 12 attempts from each user, testing with first 35 attempts, and beyond 35 attempts. There is improvement in both the success rates and average completion time as more tests are completed. This is a result of Adaptcha's adaptive selection process being able to select better CAPTCHA tests as it has more performance data available to analyze for each user. Compared to the initialization results, success rates improved by 8 percentage points and average completion time reduced by 35% as the proposed Adaptcha became better adapted for each user. This improvement in accuracy and speed marks a reduction in the amount of friction caused by the CAPTCHA.

TABLE 1. PERFORMANCE OF ADAPTCHA (SUCCESS RATES AND TIME TO COMPLETE)

Number of Tests Completed by User	Recorded Tests	Success Rate	Avg Completion Time (seconds)
Initialization (12 attempts)	612	79.4%	14.5
Adaptive Adaptcha (1-35 attempts)	1,785	86.7%	10.6
Adaptive Adaptcha (beyond 35 attempts)	923	87.4%	9.5

Adaptcha's resilience against automated attack is determined by the underlying CAPTCHA types used in its tests. Adaptcha is at least as resilient as its weakest underlying CAPTCHA type. The expected automated attack success rate against Adaptcha can be represented as

$$AttackSuccessRate \leq 0.05 * average(A) + 0.95 * max(A) \tag{3}$$

where, A represents the set of attack success rates for the CAPTCHA types used in a specific Adaptcha instance. In the case of four CAPTCHA types used in the human performance evaluation above [9], [10], [11], [12], the expected attack success rate is not more than 0.2304%.

4. Conclusion

This poster presents a novel approach to reduce the friction caused by CAPTCHAs without compromising security. Adaptcha's adaptive selection process reduces the time for users to complete CAPTCHA tests while boosting success rates, a combination which makes for an improved user experience.

References

[1] S. Shirali-Shahreza and M. H. Shirali-Shahreza, "Bibliography of works done on CAPTCHA," in *Proc. 3rd ISKE*, Nov. 2008, pp. 205–210.

[2] E. Bursztein, S. Bethard, C. Fabry, J. C. Mitchell, and D. Jurafsky, "How Good Are Humans at Solving CAPTCHAs? A Large Scale Evaluation," in *Proc. 2010 S&P*, 2010, pp. 399–413.

[3] "Distil Networks Study Reveals CAPTCHAs Have Negative Impact on Web Traffic and Leads," Mar. 2015. [Online]. Available: http://bit.ly/2nNSltr

[4] S. Sivakorn, I. Polakis, and A. D. Keromytis, "I am Robot: (Deep) Learning to Break Semantic Image CAPTCHAs," in *Proc. 2016 EuroS&P*, Mar. 2016, pp. 388–403.

[5] R. Verger, "Google just made the internet a tiny bit less annoying," *Popular Science*, Mar. 2017. [Online]. Available: http://www.popsci.com/google-invisible-recaptcha

[6] S. Holmes, “Googly-eyed robot beats ’I am not a robot’ security system,” *Daily Mail*, Jan. 2017. [Online]. Available: http://dailym.ai/2kfvrIk

[7] M. Belk, C. Fidas, P. Germanakos, and G. Samaras, “Do Cognitive Styles of Users Affect Preference and Performance Related to CAPTCHA Challenges?” in *Proc. CHI ’12 Extended Abstracts*, May 2012, pp. 1487–1492.

[8] L. O’Reilly, “Google’s new CAPTCHA security login raises ’legitimate privacy concerns’,” *Business Insider*, Feb. 2015. [Online]. Available: http://read.bi/17YxEyH

[9] B. M. Powell, G. Goswami, M. Vatsa, R. Singh, and A. Noore, “fg-CAPTCHA: Genetically Optimized Face Image CAPTCHA,” *IEEE Access*, vol. 2, pp. 473–484, Apr. 2014.

[10] G. Goswami, B. M. Powell, M. Vatsa, R. Singh, and A. Noore, “FaceDCAPTCHA: Face Detection based Color Image CAPTCHA,” *Future Generation Comput. Syst.*, vol. 31, pp. 59–68, Feb. 2014.

[11] B. M. Powell, A. Kumar, J. Thapar, G. Goswami, M. Vatsa, R. Singh, and A. Noore, “A Multibiometrics-based CAPTCHA for Improved Online Security,” in *Proc. 8th BTAS*, Sep. 2016.

[12] B. M. Powell, E. Kalsy, G. Goswami, M. Vatsa, R. Singh, and A. Noore, “Attack-Resistant aiCAPTCHA using a Negative Selection Articial Immune System,” in *Proc. 38th S&P, 2nd Workshop on BioSTAR*, May 2017.