

Implicit CAPTCHAs

Henry S. Baird^a and Jon L. Bentley^b

^a Lehigh University, CSE Dept
19 Memorial Dr West
Bethlehem, PA 18017 USA

E-mail: baird@cse.lehigh.edu
URL: www.cse.lehigh.edu/~baird

^b Avaya Labs Research
233 Mt. Airy Road
Basking Ridge, NJ 07920
E-mail: jbentley@avaya.com

ABSTRACT

We propose a design methodology for “implicit” CAPTCHAs to relieve drawbacks of present technology. CAPTCHAs are tests administered automatically over networks that can distinguish between people and machines and thus protect web services from abuse by programs masquerading as human users. All existing CAPTCHAs’ challenges require a significant conscious effort by the person answering them — *e.g.* reading and typing a nonsense word — whereas implicit CAPTCHAs may require as little as a single click. Many CAPTCHAs distract and interrupt users, since the challenge is perceived as an irrelevant intrusion; implicit CAPTCHAs can be woven into the expected sequence of browsing using cues tailored to the site. Most existing CAPTCHAs are vulnerable to “farming-out” attacks in which challenges are passed to a networked community of human readers; by contrast, implicit CAPTCHAs are not “fungible” (in the sense of easily answerable in isolation) since they are meaningful only in the specific context of the website that is protected. Many existing CAPTCHAs irritate or threaten users since they are obviously tests of skill: implicit CAPTCHAs appear to be elementary and inevitable acts of browsing. It can often be difficult to detect when CAPTCHAs are under attack: implicit CAPTCHAs can be designed so that certain failure modes are correlated with failed bot attacks. We illustrate these design principles with examples.

Keywords: *human interactive proofs, CAPTCHAs, abuse of web sites and services, implicit CAPTCHAs, usability, legibility*

1. INTRODUCTION

In 1997 Andrei Broder and his colleagues at the DEC Systems Research Center developed a scheme to block the abusive automatic submission of URLs to the AltaVista web-site [Bro01,LBBB01]. Their approach was to ask each user to read an image of printed text formed specially so that machine vision (Optical Character Recognition, or OCR) systems could not read it but humans still could. Since then many such CAPTCHAs — Completely Automated Public Turing tests to tell Computers and Humans Apart — have been developed, including CMU’s EZ-Gimpy [BAL00, HB01], PARC’s PessimialPrint [CBF01] and BaffleText [CB03], Paypal’s CAPTCHA (www.paypal.com), Microsoft’s CAPTCHA [SSB03], and Lehigh University’s ScatterType [BR05]. In addition to these, many have been developed and fielded but have not been described in the literature.

CAPTCHAs have encountered resistance from some human users who find them unpleasant, distracting, and a waste of time [BAL01]. Some users feel threatened, believing that they are being set up to fail. All CAPTCHAs known to us take the form of tests of skill that can be passed without any knowledge of the context in which the CAPTCHA is presented, and so it is technically feasible to “automate” their solution by distributing them to a community of human users enlisted for this purpose (perhaps rewarded by small advantages such as access to porn sites). Some of the first generation of CAPTCHAs have already been broken (see, *e.g.*, [CB03,MM03]) and so the next generation must be made stronger and thus, unless we are careful, are likely to be even more difficult and awkward for human users.

We propose a family of *implicit CAPTCHAs* to relieve these and related problems.



Figure 1. An unconscious CAPTCHA in the form of a link on the text “MORE PHOTOS,” rendered as an image. Such link text can be rendered deliberately to be difficult for machine readers, to form one in a series of effective implicit CAPTCHAs.

2. DESIGN PRINCIPLES

We propose these design principles for implicit CAPTCHAs:

1. challenges are disguised as necessary browsing links;
2. challenges can be answered with a single click while still providing several bits of confidence;
3. challenges can be answered only through experience of the context of the particular website; and
4. challenges are so easy that failure indicates a failed robot attack.

Each of these is considered in a separate subsection.

2.1. Disguise as Necessary Browsing Link

Figure 1 shows a fragment of a web page with a link on the text “MORE PHOTOS”. This text is rendered as an image and is, effect, a primitive (and, here, an unconscious) CAPTCHA. An interested human finds it easy to click on the relevant text, while a bot would have to surmount several obstacles, including the OCR and understanding enough semantics not to click on the phrase “Head to ...”. The relevant text occupies approximately 1/4 of the width (from the left) and 1/16 of the height (from the bottom) of the image. An attacker that makes a random mouse click on the image therefore has a probability of success of 1/64, so we say that this implicit test gives us 6 bits of confidence that such a click is not from a lucky random attacker. Because there are two potential jump phrases, we retain one bit of confidence against an attacker with both OCR and substantial semantic analysis.

2.2. Single-Click CAPTCHAs

When a human user finishes filling out a a web form, the typical final step is to click on a “Submit” button. An Implicit CAPTCHA replaces the the standard button with an image like that in Figure 2. It is a schematic example of a single-click implicit CAPTCHA which is nearly trivial for people, but the wide variety of slang synonyms for “wrong answer,” together with the variations in randomized placement and rendering of each alternative, provides several bits of confidence if it is answered correctly. Since the “Submit” image is roughly 1/32 of the area of the image, we have 5 bits of confidence against a random clicker. And because the image contains 8 different clusters of letters, we retain 3 bits of confidence against an attacker that can group words.



Figure 2. A variety of options can be presented as text rendered as a complex image so that, although it's answerable with a single click, several bits of confidence are gained. The varying styles of the rendering can make this a challenging test for OCR programs, while it remains simple enough for humans that failure is a strong indicator of an attempted robot attack.



Figure 3. A friendly greeting page is an implicit one-click CAPTCHA.

Figure 3 shows a screen that would appear to many mountain climbers to be a friendly greeting to a climbing web site. It is also a one-click implicit CAPTCHA. If we define the "mountain top" to be a square of about 20 pixels per side, or about 1/16th of each dimension of the image, a click in that area yields 8 bits of confidence that the visitor is not a randomly clicking robot.

Figure 4 illustrates the principle of a single-click implicit CAPTCHA which is keyed to a number of questions referring to parts of a natural image. Thus one image can supports several CAPTCHA challenges, such as, in

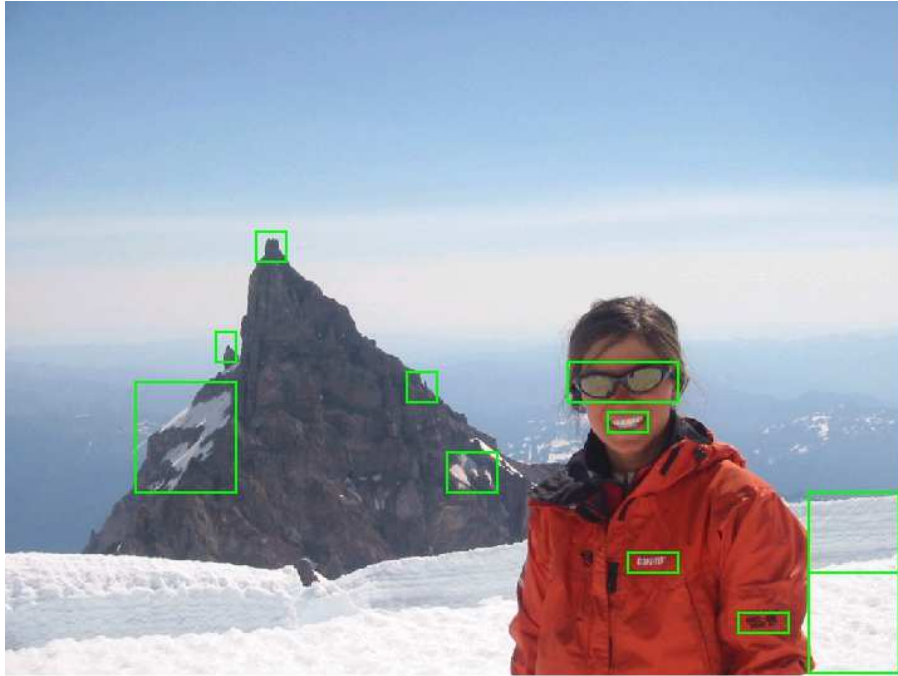


Figure 4. This natural scene offers a large number of reference objects for CAPTCHA challenges such as “Click on the climber’s glasses.”

this case, “Click on the climber’s glasses,” or “Click on the logo on the climber’s arm.” The parts of the image would be labeled manually, each part linked to a set of queries. The image itself can be randomized in many ways through cropping and degradations.

2.3. Contextual CAPTCHAs

In November, 1999, slashdot.com conducted a poll to determine the “best” graduate program in Computer Science [SD99]. The only defense against robotic voting was a check on IP addresses. Groups at Carnegie Mellon University (CMU) and at the Massachusetts Institute of Technology (MIT) rose to the implicit challenge and automated the process of casting votes for their schools. By the close of the poll, MIT had 21,156 votes, CMU had 21,032 votes, while all other schools had fewer than a thousand votes. This event is often used to motivate traditional entry CAPTCHAs: a voter must pass an explicit test before casting a vote.

Figure 5 illustrates an alternative approach to ensuring human voters. The (say) thirty schools are presented as thirty buttons in a random order, changed for each voter (here, only two are shown). Each button contains the textually degraded name of a school, also changed for each voter. A human voter finds it easy to choose a favorite school, while an attacking robot would have to solve a problem almost as difficult as a traditional CAPTCHA.

This approach applies to a broad class of online polls. Consider, for instance, a nationwide poll for a contest such as television’s American Idol (www.idolonfox.com). To cast a vote for or against a contestant, a human could select a button that contains an image of the person. Rather than using one static image per person, the web site could choose one of several scenes from a motion picture, and then automatically transform the chosen scene by cropping, smoothing, etc.

2.4. Trivial CAPTCHAs

Permuting poll buttons and transforming their images are strategies for making it difficult to tip a poll in a chosen direction. Unfortunately, a robotic attack can still invalidate a poll by drowning it in noise. For instance, an attacking robot could cast so many random votes as to swamp all the real information.

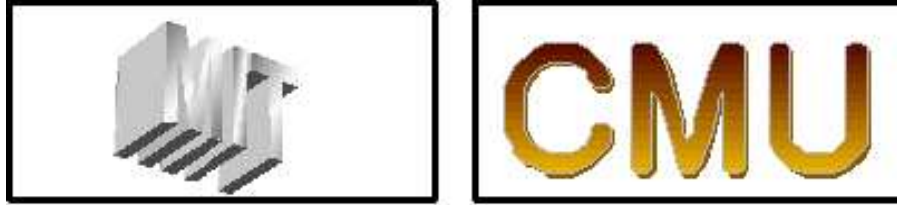


Figure 5. Two of the alternatives in a poll for the best CS Graduate School, CMU and MIT, could be represented as image CAPTCHAs.

We can provide a partial defense against such an attack by replacing the traditional “Submit” button in a poll with an image such as Figure 2. A human finds such an image trivial to navigate. A guessing robot, on the other hand, has only a chance of a few percent of clicking in the right place (because, as we noted earlier, the “Submit” area is only about 1/32 of the image). While a few percent of the random votes do sneak through, the lion’s share of invalid attempts alerts the web site to a probable attack. At that point, more traditional CAPTCHAs might be deployed to defend the site.

3. DISCUSSION

Traditional CAPTCHAs present test images ranging from a common word (about 10 bits for a list of 1000 words) to a string of eight random alphanumeric characters (about 40 bits). The implicit CAPTCHAs that we have sketched so far are easier and less off-putting to users, but provide only between 3 and 8 bits of confidence. When is this tradeoff useful?

Boundaries in the physical world range from a string between wooden stakes to a chain link fence to a brick wall topped by barbed wire. Each successive barrier provides more protection than its predecessors, but each is also more expensive to build and is less friendly to passersby. Wooden stakes with string between them are appropriate devices to keep pedestrians off of recently seeded grass, and brick walls topped by barbed wire are appropriate to defend high security installations. Our implicit CAPTCHAs will not replace the traditional CAPTCHAs, but will instead easily extend their use to important new domains.

This paper has summarized the principles underlying Implicit CAPTCHAs and has illustrated those principles with several images. We have already extended the principles to a much larger set of images, which we cannot illustrate here because of the page limitations. An entrance to a web site for a technical community, for instance, might contain a group photo of a recent conference and directions such as “Click on Fred Jones, who is the right half of the second row, wearing the red sweater.” While this task is very difficult to automate, it is easy for a human, and is possibly interesting for many users. (“Oh, that’s what he looks like!”) We have explored other images suitable for “one-click CAPTCHAs”, such as instrument panels (“To enter HorrendouslyFastCars.com, click on 112 mph on that speedometer”) and cartoon scenes and cartoon maps.

We have also constructed Implicit CAPTCHAs that can extract many more bits of confidence by constructing a “story” contained in a sequence of related images. The user is given an instruction in each image to click on a given subregion for about 8 bits of confidence; a sequence of five such images can then give a total of 40 bits, which compares favorably with current explicit CAPTCHAs.

4. ACKNOWLEDGMENTS

We are grateful for stimulating conversations with Gary Elko, Joe Hall, Dan Lopresti, Larry O’Gorman and Terry Riopka.

References

- [BAL00] M. Blum, L. A. von Ahn, and J. Langford, *The CAPTCHA Project*, “Completely Automatic Public Turing Test to tell Computers and Humans Apart,” www.captcha.net, Dept. of Computer Science, Carnegie-Mellon Univ., and personal communications, November, 2000.
- [BK02] H. S. Baird and K. Popat, “Human Interactive Proofs and Document Image Analysis,” *Proc., 5th IAPR Int’l Workshop on Document Analysis Systems*, Princeton, NJ, Springer-Verlag (Berlin) LNCS 2423, pp. 507–518, August 2002.
- [Bro01] AltaVista’s “Add-URL” site: altavista.com/sites/addurl/newurl, protected by the earliest known CAPTCHA.
- [BR05] H. S. Baird and T. Riopka, “ScatterType: a Reading CAPTCHA Resistant to Segmentation Attack,” *Proc., IS&T/SPIE Document Recognition and Retrieval Conf*, San Jose, CA, January 16–20, 2005 [in the present proceedings].
- [CB03] M. Chew and H. S. Baird, “BaffleText: a Human Interactive Proof,” *Proc., 10th SPIE/IS&T Document Recognition and Retrieval Conf. (DRR2003)*, Santa Clara, CA, January 23–24, 2003.
- [CBF01] A. L. Coates, H. S. Baird, and R. Fateman, “Pessimistic Print: a Reverse Turing Test,” *Proc., IAPR 6th Intl. Conf. on Document Analysis and Recognition*, Seattle, WA, September 10-13, 2001, pp. 1154-1158.
- [HB01] N. J. Hopper and M. Blum, “Secure Human Identification Protocols,” In: C. Boyd (Ed.) *Advances in Cryptology, Proceedings of Asiacypt 2001*, LNCS 2248, pp.52 -66, Springer-Verlag Berlin, 2001
- [LABB01] M. D. Lillibridge, M. Abadi, K. Bharat, and A. Z. Broder, “Method for Selectively Restricting Access to Computer Systems,” U.S. Patent No. 6,195,698, Issued February 27, 2001.
- [MM03] G. Mori and J. Malik, “Recognizing Objects in Adversarial Clutter: Breaking a Visual CAPTCHA,” *Proc., IEEE CS Society Conf. on Computer Vision and Pattern Recognition (CVPR’03)*, Madison, WI, June 16-22, 2003.
- [NS96] G. Nagy and S. Seth, “Modern optical character recognition.” in *The Froehlich / Kent Encyclopaedia of Telecommunications*, Vol. 11, pp. 473-531, Marcel Dekker, NY, 1996.
- [Pav00] T. Pavlidis, “Thirty Years at the Pattern Recognition Front,” King-Sun Fu Prize Lecture, 11th ICPR, Barcelona, September, 2000.
- [RNN99] S. V. Rice, G. Nagy, and T. A. Nartker, *OCR: An Illustrated Guide to the Frontier*, Kluwer Academic Publishers, 1999.
- [RJN96] S. V. Rice, F. R. Jenkins, and T. A. Nartker, “The Fifth Annual Test of OCR Accuracy,” ISRI TR-96-01, Univ. of Nevada, Las Vegas, 1996.
- [SCA00] A. P. Saygin, I. Cicekli, and V. Akman, “Turing Test: 50 Years Later,” *Minds and Machines*, 10(4), Kluwer, 2000.
- [SD99] Slashdot site www.slashdot.com; for details of the poll, cf. www.captcha.net.
- [SSB03] P. Y. Simard, R. Szeliski, J. Benaloh, J. Couvreur, I. Calinov, “Using Character Recognition and Segmentation to Tell Computer from Humans,” *Proc., IAPR Int’l Conf. on Document Analysis and Recognition*, Edinburgh, Scotland, August 4–6, 2003.
- [Tur50] A. Turing, “Computing Machinery and Intelligence,” *Mind*, Vol. 59(236), pp. 433–460, 1950.