TECHNISCHE UNIVERSITÄT DARMSTADT

CAPTCHAs: The Good, the Bad, and the Ugly

ISSE-GI SICHERHEIT 2010

Paul Baecher*, Marc Fischlin*, Lior Gordon, Robert Langenberg, Michael Lützow, Dominique Schröder*

* TU Darmstadt, supported by DFG Emmy Noether Program

Emmy Noether-Programm
Deutsche Forschungsgemeinschaft
DFG

Introduction

What Are CAPTCHAs?

TECHNISCHE UNIVERSITÄT DARMSTADT

- **C**ompletely **A**utomated **P**ublic **T**uring test to tell **C**omputers and **H**umans **A**part
  - “reverse” Turing test, term coined by [vABHL03]
- challenge/response protocol where
  - response should be easy to observe for humans
  - response should be hard to compute for machines
- application: protect online services from automated use

image: cryptographp

What Are CAPTCHAs?

TECHNISCHE UNIVERSITÄT DARMSTADT

- **C**ompletely **A**utomated **P**ublic **T**uring test to tell **C**omputers and **H**umans **A**part
  - “reverse” Turing test, term coined by [vABHL03]
- challenge/response
  - response should be easy to observe for humans
  - response should be hard to compute for machines
- application: protect online services from automated use

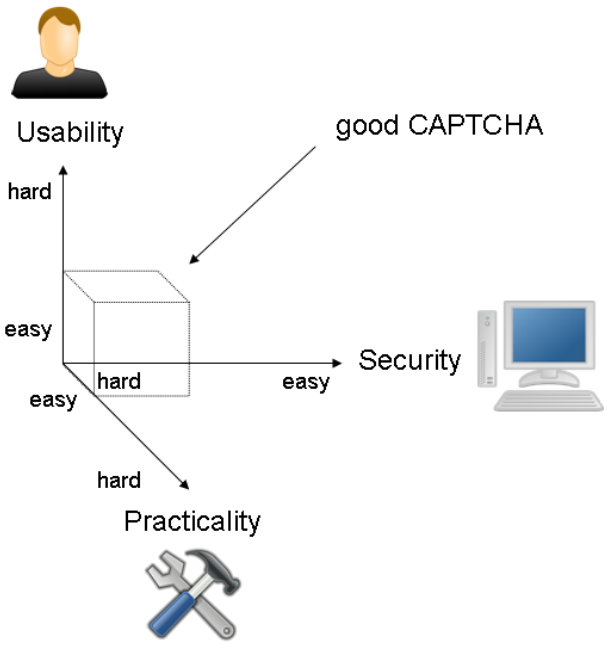
0.01% according to [CLSC05, vAMM+08]

image: cryptographp

A Third Dimension

TECHNISCHE UNIVERSITÄT DARMSTADT

- easy for humans, hard for machines
- what about practicability?
  - small display dimensions
  - varying input devices/methods
  - different media formats and support thereof
  - acceptance by users
  - environmental aspects (audio CAPTCHAs in a shared office. . .)



Bad CAPTCHAs

Breaking Bad CAPTCHAs

Three Bad CAPTCHAs

TECHNISCHE UNIVERSITÄT DARMSTADT

- Bundesamt für Wirtschaft und Ausfuhrkontrolle (BAFA)
  - “Umweltprämie”, economic stimulus program

- Bundesrepublik Deutschland - Finanzagentur GmbH – Bundeswertpapiere
  - online banking interface to governmental bonds

- Sparda-Banken
  - online banking interface

One Approach to Break Them All

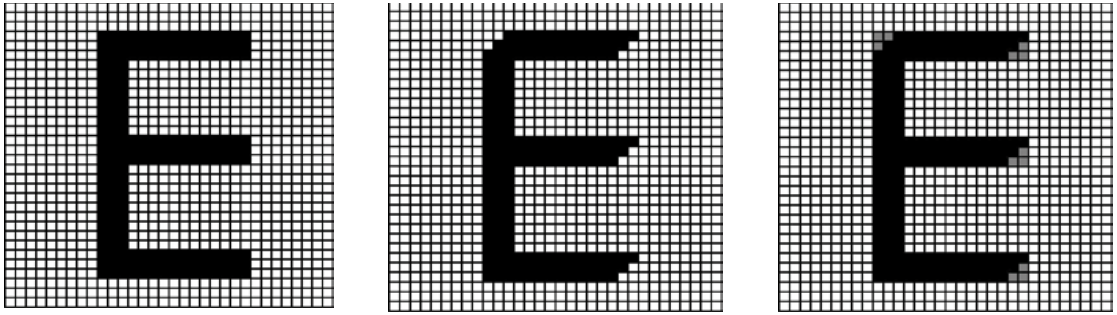
TECHNISCHE UNIVERSITÄT DARMSTADT

1. preprocess the images
   - the grid is static: rather trivial to remove
   - the line always starts in the same location, follow and remove
2. segment characters
   - easy, since they do not touch each other
3. detect individual characters
   - use a k-means clustering algorithm to learn mean characters
   - see next slide. . .

From Characters to Vectors

TECHNISCHE UNIVERSITÄT DARMSTADT

- k-means clustering operates on d-dimensional vectors
- obtain a 1024-dimensional vector for each character
  - scale character to a 32×32 pixels bounding box
  - normalize brightness of each pixel to $[0, 1]$
  - traverse pixels in a unique sequence



Breaking a CAPTCHA

TECHNISCHE UNIVERSITÄT DARMSTADT

- offline (training) phase
  - obtain a set of training data CAPTCHA challenges
  - preprocess and run k-means algorithm (Lloyd's algorithm)
    - use labels to correct a few errors
  - save mean characters
- online (query) phase
  - preprocess and find nearest cluster

Results

TECHNISCHE UNIVERSITÄT DARMSTADT

- experimental results of our implementation:

	“Umweltprämie”	68%
	Bundeswertpapiere	70%
	Sparda-Banken	87%

- 5% is considered broken according to [vAMM+08]

using tesseract OCR

Better CAPTCHAs

Designing Good CAPTCHAs

TECHNISCHE UNIVERSITÄT DARMSTADT

- use random challenge strings
  - dictionary words help the attacker
  - interpolate partially detected word fragments
  - make an offline-decision
- use monochromatic images

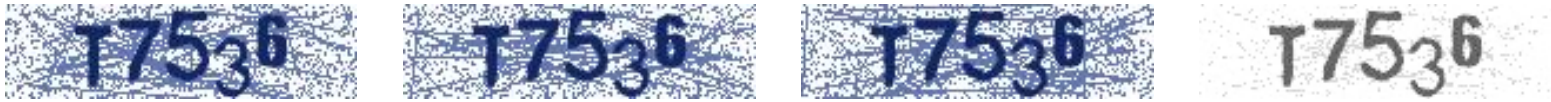
- require segmentation
  - mere recognition is not enough [CLSC05]
- apply distortions with many degrees of freedom

Implementation Pitfall

TECHNISCHE UNIVERSITÄT DARMSTADT

- one version per challenge
  - `digg.com`

  - `quoka.de`



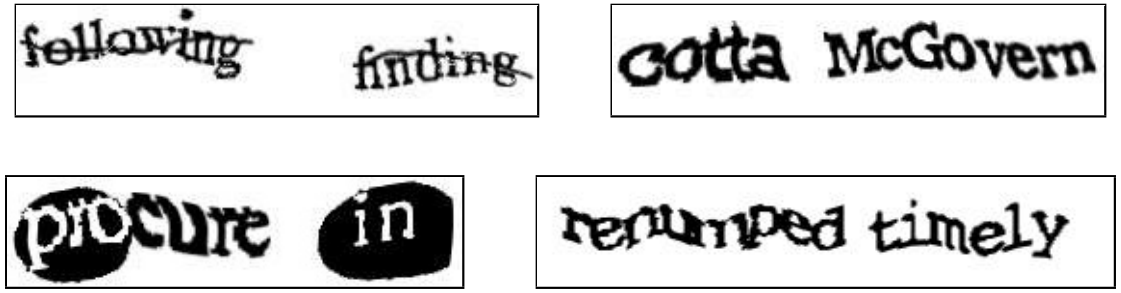
  - consider an attacker that is able to recognize one randomly chosen character

reCAPTCHA

reCAPTCHA

TECHNISCHE UNIVERSITÄT DARMSTADT

- unique concept
  - human OCR system
  - verification words, scan words
- proprietary but free centralized service
- very popular (facebook, . . .)
- secure?



major revisions of reCAPTCHA

reCAPTCHA Considered Broken

TECHNISCHE UNIVERSITÄT DARMSTADT

- first generation, early 2008
  - broken by Wilkins using erode/dilate and OCR [Wil09], 5%*
- second generation, until December 2009
  - broken by Wilkins, 5%*; our results: 6–10%
- third generation, until August 2010
  - broken by Houck [Hou10], 10%; our results: ca. 6%
- fourth (current) generation
  - broken by Houck, 30%

Conclusions

TECHNISCHE UNIVERSITÄT DARMSTADT

- the majority of all CAPTCHAs can be broken easily
- not hard to avoid most common errors
- rely on segmentation task
- reCAPTCHA is (was?) a good choice
- designing a robust CAPTCHA seems extremely difficult

The End

TECHNISCHE UNIVERSITÄT DARMSTADT

Thank you!

?

References

TECHNISCHE UNIVERSITÄT DARMSTADT

Kumar Chellapilla, Kevin Larson, Patrice Y. Simard, and Mary Czerwinski.
Building Segmentation Based Human-Friendly Human Interaction Proofs (HIPs).
In *HIP*, volume 3517 of *Lecture Notes in Computer Science*, pages 1–26. Springer-Verlag, 2005.

Chad W. Houck.
Decoding reCAPTCHA.
`http://www.n3on.org/projects/reCAPTCHA/docs/reCAPTCHA.docx`, 2010.

Luis von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford.
CAPTCHA: Using Hard AI Problems for Security.
In Eli Biham, editor, *Advances in Cryptology – EUROCRYPT 2003*, volume 2656 of *Lecture Notes in Computer Science*, pages 294–311, Warsaw, Poland, May 4–8, 2003. Springer, Berlin, Germany.

Luis von Ahn, Benjamin Maurer, Colin McMillen, David Abraham, and Manuel Blum.
reCAPTCHA: Human-Based Character Recognition via Web Security Measures.
Science, 321(5895):1465–1468, 2008.

Jonathan Wilkins.
Strong CAPTCHA Guidelines v1.2.
2009.