A Seminar Report
On
CAPTCHA
Contents

• Definition
• Generic CAPTCHAs distort letters & numbers
• What is Turing Test
• Types
• Applications
• Constructing CAPTCHA
• CAPTCHA logic
• Conclusion
Introduction

• CAPTCHA – Completely Automated Public Turing test to tell Computers & Humans Apart.

• Invented in 2000 by Luis von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford at Carnegie Mellon University used by Yahoo.

• It is a program that is a challenge response test to separate humans from computer programs.
Generic CAPTCHAs distort letters & numbers

• Distorted characters are presented to the user.

• User has to recognize the distorted letters.

• If the guessed letters are correct, the user is inferred to be a human & allowed access.
Contd...

• Humans can read the distorted & noisy text.

• Current OCRs (Optical Character Recognition) cannot read them.
What is Turing Test?

- Proposed by Alan Turing.
- To test a machine’s level of intelligence.
- Human judge asks questions to two participants, one is a machine & the other human.
- The judge doesn’t know which is which.
- After listening to the answer, if the judge fails to recognize which one is the machine, then the machine passes the test.
Contd...

- **CAPTCHA** employs a Reverse Turing Test.
- Judge = **CAPTCHA** program, participant = user
- If the user passes **CAPTCHA**, he is human otherwise it is a machine.
Types of CAPTCHA

1. Text Based
2. Gimpy
3. EZ-Gimpy
4. MSNs passport service CAPTCHA
5. Graphic based CAPTCHA
6. Audio CAPTCHA
1. Text Based

- Simple, normal questions:
  - What is the sum of three & thirty-five?
  - If today is Saturday, what is day after tomorrow?
  - Which of mango, table & water is a fruit?
- Very effective, needs a large question bank.
- Cognitively challenged users find it hard.
2. Gimpy

- Designed by Yahoo & CMU(Carnegie Mellon University)
- Picks up 10 random words from dictionary & distorts, fills with noise.
- User has to recognize at least 3 words.
- If the user is correct, then he is admitted.
3. EZ-Gimpy

- A modified version of Gimpy.
- Yahoo used this version in Messenger.
- Has only 1 random string of characters.
- Not a dictionary word, so not prone to dictionary attack.
- Not a good implementation, already broken by OCRs (Optical Character Recognition).
4. MSNs passport service CAPATCHAs

- Provided for Microsoft’s MSN services.
- Use of 8 characters.
- Warping is used to distort.
- Very strong implementation, hasn’t been broken.
- It is segmentation-resistant.
5. Graphic based CAPTCHAs

1. BONGO
   - After M.M. Bongard, pattern recognition expert.
   - User has to solve a pattern recognition problem.
   - Has to tell the distinct characteristic between two sets of figures.
   - Then tell to which set a given figure belongs to.
Contd...

2. PIX

Uses a large database of labelled images.

- It shows a set of images, user has to recognize the common feature among those.

- Eg :- pick the common characteristic among the following 4 pictures = “aeroplane”.
6. Audio CAPTCHAs

- Consists of downloadable audio clip.
- User listens & enters the spoken word.
- Helps visually disabled users.
- Below is the Google’s audio enabled CAPTCHA
Applications

• Protect Online polls.
• Prevent web registration abuse, protect passwords from brute-force attack.
• Prevent e-mails spam such as the webmail services of Gmail, Hotmail, and Yahoo! Mail.
• Stop automated posting to blogs, forums and wikis
• E-ticketing, prevent scalping.
• Abusing free Online accounts.
Contd...

- Verify digitized books: “RE-CAPTCHA”
  - Used in Google books project.
  - Two words are shown, the program knows the first word.
  - If the user enter the first word correctly, it assumes that the second unknown word will also be entered correctly.
  - Second word becomes “known”.

![RE-CAPTCHA Example](image-url)
Constructing CAPTCHAs

- Things to keep in mind :-
  - Don’t store CAPTCHA solution in web page’s metadata.
  - A CAPTCHA is no good if it doesn’t distort.
  - Need a large database of different CAPTCHA questions.
  - Avoid repetition of question.
  - Java script may be needed to be activated it in browsers.
  - Some may need Adobe Flash Plugin.
CAPTCHA logic

- Generate the question
- Persist the correct answer
- Present the question to the user
- Evaluate the answer, if incorrect start again-
  Generate a different CAPTCHA
- If correct allow the access to the user
Conclusion

- It uses for human verification
- Free services should be protected with a CAPTCHA in order to prevent abuse by automated programs.
- It is an effective way to counter bots & reduce spam.
Thanks To All