Human or bot? Let an audio CAPTCHA decide

Dimitris Gritzalis, Yannis Soupionis
Information Security & Critical Infrastructure Protection Research Laboratory
Dept. of Informatics
Athens University of Economics & Business (AUEB)
{dgrit, jsoup}@aueb.gr
Contents

- Session Initiation Protocol (SIP)
- Spam over Internet Telephony (SPIT)
- CAPTCHA: Development and attributes
- Software bots
- Commercial audio CAPTCHA
- CAPTCHA: Design and evaluation
- Early conclusions and plans
Session Initiation Protocol (SIP)

Spam over Internet Telephony (SPIT)

Bulk unsolicited set of session

...Call initiations

...Instant messages

...Presence requests
CAPTCHA

CAPTCHA: Completely Automated Public Tests to tell Computers and Humans Apart

1. Visual CAPTCHAs: Text or Image based
   The proposed CAPTCHA must be:
   - easy for humans to solve
2. Audio CAPTCHAs: Spoken character based
   - easy for a tester machine to generate and grade
   - hard for a software bot to solve
3. Logical CAPTCHAs: Simple question based

We have a mother and her daughter: Who is the younger one?
CAPTCHA development process

**Decision**: Until \((UR > \bar{a})\) AND \((BR < \bar{e})\)

- **UR**: User Success Rate
- **BR**: Bot Success Rate
- **\(\bar{a}\)**: User Rate Threshold
- **\(\bar{e}\)**: Bot Rate Threshold

---

Select CAPTCHA attributes → Create CAPTCHA → Estimate User Success → Estimate Bot Success → Decision

- **Decision**: No → Final CAPTCHA
- **Decision**: Yes → Final CAPTCHA
Generic CAPTCHA attributes

- **Character**
- **Total CAPTCHA**
- **Between characters**
- **Duration**
- **Vocabulary**
- **Audio CAPTCHA**
- **Noise**
- **Production Procedure**
- **Automated**
- **Reappearance**
- **Language requirements**
- **No. of characters**
- **Data field**
- **Background**
- **Intermediate**

**Generic CAPTCHA attributes**

- **Audio CAPTCHA**
- **Vocabulary**
- **Production Procedure**
Selected CAPTCHA attributes

- Different announcers
- Random positioning of each digit
- Digits only \{0,\ldots,9\}
- Background noise
- Different duration
- Noise between digits

Proposed Audio CAPTCHA
The frequency and energy pick detection bots:

+ have been demonstrated effective
+ are considerably easy to implement
+ require limited time to solve a CAPTCHA
+ waste a limited amount of system recourses

- require a training session, where a human identifies a number of selected CAPTCHA
Experimentation with a bot

Stage 1

UA1

DOMAIN

DOMAIN2

UA2

Captcha Playback

Wait for Answer

answer

timeout

Stage 0

Answer

0

10

20

YES
## Commercial audio CAPTCHA

<table>
<thead>
<tr>
<th>Service</th>
<th>Audio.wav</th>
<th>Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ReCaptcha</strong>&lt;sup&gt;1&lt;/sup&gt;</td>
<td><img src="image1.png" alt="Waveform" /></td>
<td>7 0 8 6 6 1 1 0</td>
</tr>
<tr>
<td><strong>Google</strong>&lt;sup&gt;2&lt;/sup&gt;</td>
<td><img src="image2.png" alt="Waveform" /></td>
<td>3 1 2 2 0 4 5 7 0</td>
</tr>
<tr>
<td><strong>MSN</strong>&lt;sup&gt;3&lt;/sup&gt;</td>
<td><img src="image3.png" alt="Waveform" /></td>
<td>6 7 9 9 8 1 3 2</td>
</tr>
</tbody>
</table>

2. [http://gmail.com](http://gmail.com) (Google, 2008) (Vorm bot access rate: 33%)
3. [https://accounts.services.passport.net/reg.srf](https://accounts.services.passport.net/reg.srf) (Microsoft, 2008) (Vorm bot access rate: 75%)
### Suggested CAPTCHA

<table>
<thead>
<tr>
<th>Stage</th>
<th>Different announcers</th>
<th>Background noise</th>
<th>Intermediate noise</th>
<th>Digit distribution</th>
<th>Training CAPTCHA</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAGE 1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>STAGE 2</td>
<td>7</td>
<td></td>
<td></td>
<td>✓</td>
<td>50</td>
</tr>
<tr>
<td>STAGE 3</td>
<td>5</td>
<td>![Speakers]</td>
<td></td>
<td>✓</td>
<td>100</td>
</tr>
<tr>
<td>STAGE 4</td>
<td>7</td>
<td>![Speakers]</td>
<td>![Speakers]</td>
<td>✓</td>
<td>100</td>
</tr>
<tr>
<td>STAGE 5</td>
<td>7</td>
<td>![Speakers]</td>
<td>![Speakers]</td>
<td>✓</td>
<td>100</td>
</tr>
</tbody>
</table>

Number of users: 22  
Age: 20-30 years old  
Mother tongue: Greek  
Speak English: All
Audio CAPTCHA vs. Bot

Medium Threshold: 13 digits

Low Threshold: 15 digits

High Threshold: 2 digits
Early conclusions and plans

On-going efforts and research…

Low bot success rate was achieved because of the combination of all, referred to, attributes.

Evaluation of the audio CAPTCHA effectiveness will be further tested by audio/speech recognition tools.

Different populations of users will be considered, and further evaluation and testing will be carried out.

Additional bots are required for testing purposes.
References


