Breaking E-bay audio captcha

Elie Bursztein Steven Bethard
Stanford University
Outline

• Breaking an audio captcha
• E-bay audio captcha
• What is next?
Bad or good user?
Bad or good user?
Bad or good user?
Bad or good user?

bots

users

Breaking E-bay audio captcha
Completely Automated Public Turing test to tell Computers and Humans Apart

Term coined in 2000 [VBHL]
Breaking the registration process

Scraping
Breaking the registration process

Scraping

Solving
Breaking the registration process

Scraping

Solving

Registering
Breaking captcha

Pre-processing
Breaking captcha

Pre-processing

Segmentation
Breaking captcha

Pre-processing

Segmentation

Classification
captcha breaker efficiency metric

- Coverage
- Precision
- Accuracy

Accuracy = Coverage x Precision

Captcha goal = 0.01% accuracy [CLC05]
Breaking E-Bay captcha
The registration page

Choose your user ID and password - All fields are required

Create your eBay user ID
Use letters or numbers, but not symbols. Learn more about creating great user IDs.

Create your password
case sensitive. Learn about secure passwords.

Re-enter your password

Pick a secret question
Select your secret question...

Your secret answer
If you forget your password, we'll verify your identity with your secret question.

Date of birth
Month Day Year
You must be at least 18 years old to use eBay.

The not-so-fine print
For added security, please enter the verification code hidden in the image.

287207

Listen to the verification code

I agree that:

- I accept the User Agreement and Privacy Policy.
- I may receive communications from eBay and can change my notification preferences in My eBay.
- I'm at least 18 years old.

Continue

Elie Bursztein, Steven Bethard

Breaking E-bay audio captcha

http://dontrythisathome.com/decaptcha/
Registration Security

• Allow to download/try multiple time the same captcha

• Limit poorly the number of captcha download
Fetching performance

Elie Bursztein, Steven Bethard
Breaking E-bay audio captcha

http://dontrythisathome.com/decaptcha/
Fetching performance

- Set 4
- Set 5
- Set 6
- Set 7
- Set 8
E-Bay captcha

3 1 1
E-Bay captcha
E-Bay captcha

3 1 1 7 4
E-Bay captcha
• Fixed number of digit
• Multiple downloads with different voices
Building a corpus

- We downloaded **26,000** captchas overall
  - **8x 2000** samples (training set)
  - **50 x 200** samples (evaluation set)
Building a breaker

- Two options
  - Use state of art speech recognizer
  - Build a specialized classifier
Sphinx 4

- State of art speech recognizer
- Available models
  - HUB4: 64,000 words model
  - Isolated digits
  - Connected digits
### Attempt 1 result

<table>
<thead>
<tr>
<th>Model</th>
<th>Acc</th>
<th>Prec</th>
<th>Cov</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIDGITS</td>
<td>0.3%</td>
<td>9.6%</td>
<td>3.6%</td>
</tr>
<tr>
<td>HUB4</td>
<td>1%</td>
<td>28.9%</td>
<td>3.6%</td>
</tr>
</tbody>
</table>
Attempt 2: build our own classifier
Attempt 2: build our own classifier

Buggy / Ugly C code!
How Decaptcha works
How Decaptcha works

[Diagram showing a graph of audio frequencies]

Cursor: 94 Hz (F#2) = -11 dB  Peak: 88 Hz (F2)

Elie Bursztein, Steven Bethard
Breaking E-bay audio captcha
http://dontrythisathome.com/deaptcha/
How Decaptcha works

Elie Bursztein, Steven Bethard

Breaking E-bay audio captcha

http://dontrythisathome.com/decaptcha/
# How Decaptcha works

Decaptcha works by analyzing the audio signal of the captcha challenge. The audio signal is sampled at a high frequency, and the amplitude of each sample is quantified. These amplitudes are then visualized as a spectrogram, which shows the frequency content of the audio signal over time.

The spectrogram in the image above shows a range of frequencies, with the peaks indicating the most prominent frequencies in the audio signal. The spectrogram is color-coded, with different colors representing different frequencies.

The text below the image provides a list of sample frequencies from the audio signal, which can be used to identify the predominant frequencies in the audio.

### Sample Frequencies

- 11Hz
- 22Hz
- 44Hz
- 88Hz
- 175Hz
- 351Hz
- 702Hz
- 1KHz
- 3KHz
- 6KHz
- 11KHz

**Frequency Analysis**

- **Cursor**: 94 Hz (F#2) = -11 dB
- **Peak**: 88 Hz (F2)

### Sample Frequencies List

```
0: 26645 67191 119061 89244 54427 65733 123753 34290 228898 99908 201288 241704 187527
187337 355112 131920 67751 152565 129169 143524 76310 76563 202572 143785 225241
187323 350394 94288 209828 270356 279296 120812 268639 140728 98185 223023
269431 277125 236198 246403 132344 289054 102437 105509 212886 295669 197393 318710
431137 785228 843192 3190283
24248 399464 830038 880644 111876 760679 469023 78097 297356 146268 298207 257063
556285 606488 352779 159524 229851 164489 102580 201637 81613 112075 249878 226535
477520 866948 902598 4626601
148464 208997 592865 1087574 74635 509526 382653 238383 202327 243430 268997 606916
301822 358115 208761 111799 119561 142713 172740 202756 76463 66374 146260 274962
246128 982006 1654120 5740142
45956 258133 1005985 1098416 57227 828851 1339151 314516 513288 179292 550734 519032
799286 231594 498671 342832 202360 197459 185155 315956 209138 234810 160295 300425
329821 1353146 2571746 4564688
81712 161237 417391 506251 185897 119133 338142 140221 337865 401714 465674 350169
270428 365631 253015 139859 124894 180852 114128 147683 100545 138302 202106 381131
388385 74749 1390303 3786484
```

Elie Bursztein, Steven Bethard

Breaking E-bay audio captcha

<table>
<thead>
<tr>
<th>Model</th>
<th>Acc</th>
<th>Prec</th>
<th>Cov</th>
</tr>
</thead>
<tbody>
<tr>
<td>3x100</td>
<td>53.4%</td>
<td>65.7%</td>
<td>81.3%</td>
</tr>
<tr>
<td>3x500</td>
<td>59.6%</td>
<td>73.3%</td>
<td>81.3%</td>
</tr>
<tr>
<td>6x100</td>
<td>65.6%</td>
<td>80.7%</td>
<td>81.3%</td>
</tr>
<tr>
<td>8x1000</td>
<td>75.1%</td>
<td>92.3%</td>
<td>81.3%</td>
</tr>
</tbody>
</table>
### More captcha or more samples?

<table>
<thead>
<tr>
<th>Size</th>
<th>Accuracy</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x50</td>
<td>44%</td>
<td>55%</td>
</tr>
<tr>
<td>25x20</td>
<td>76%</td>
<td>95%</td>
</tr>
<tr>
<td>50x10</td>
<td>72%</td>
<td>90%</td>
</tr>
</tbody>
</table>
What’s next
Building a generic attack framework

- What is **difficult** to computer ?
- What is **easy** for human ?
- Difference between **visual** and **audio** captcha ?
Visual captcha Hardness

- What is hard for a computer?
- Decomposition is hard
- Recognition is easy

<table>
<thead>
<tr>
<th>Characters under typical distortions</th>
<th>Recognition rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Characters" /></td>
<td>~100%</td>
</tr>
<tr>
<td><img src="image" alt="Characters" /></td>
<td>96%</td>
</tr>
<tr>
<td><img src="image" alt="Characters" /></td>
<td>100%</td>
</tr>
<tr>
<td><img src="image" alt="Characters" /></td>
<td>98%</td>
</tr>
<tr>
<td><img src="image" alt="Characters" /></td>
<td>~100%</td>
</tr>
<tr>
<td><img src="image" alt="Characters" /></td>
<td>95%</td>
</tr>
</tbody>
</table>
Decomposition hardness

• Separation should be harder for some type
  • Separated digits
  • Letters
  • Numbers
  • Words
  • Continuous speech
Recognition hardness

- Introduce noise
- Introduce distortion
- Larger corpus
What about user?

Recognition vs. Difficulty

Human

AI?
Visual vs Audio

• It is all about data pool:
  • Infinite for visual captcha
  • Finite for audio captcha
• Things we are currently doing
  • Building a generic attack framework
  • Defining a security metric
  • Studying usability
  • Breaking more captcha scheme :)

• Things we hope to do
  • Find a new way to do audio captcha
  • Going after visual captcha
Thanks / questions

• Thanks to
  • Jerry Louis and the E-bay security team
  • Celine Fabry

• Questions?

If you want to try decaptcha

http://www.dontrythisathome.com/decaptcha/