Something about audio CAPTCHAs

Elie Bursztein, Romain Bauxis, Daniele Perito, Hristo Paskov, Celine Fabry, John Mitchell
Elie Bursztein
@elie
Stanford
Researcher at Stanford University. Tweeting about web security, mobile security, and offensive technologies. No I don't wear a black hat :)
http://elie.im
Elie Bursztein (@elie) - Stanford

Researcher at Stanford University. Tweeting about web security, mobile security, and offensive technologies. No I don’t wear a black hat :)
http://elie.im

About @elie

Following 21

Followers 151

Similar to You - view all

SoftwareMaven
Travis Jensen

asteingruebl
Andy Steingruebl

frsqzdlemons
Nicole

Marsmensch
Florian Maier
Twitter Follower Packages

Please Select One Of Our Targeted Follower Pages

Silver Package

- 1000 Targeted Followers
- Guaranteed REAL, Targeted People Interested In Your Business
- Added To Your Page Within 25 days
- Targeted To Your Business/Niche
- Select The Country/s Where You Want Your Followers From
- No Automatic Bots/Programs To Get Followers, We Proudly Target 100% Of Your Followers Manually

$49.99

Order Now »

Gold Package

- 5000 Targeted Followers
- Guaranteed REAL, Targeted People Interested In Your Business
- Added To Your Page Within 40 Days
- Targeted To Your Business/Niche
- Select The Country/s Where You Want Your Followers From
- No Automatic Bots/Programs To Get Followers, We Proudly Target 100% Of Your Followers Manually

$139.99
Join the Conversation

Already use Twitter on your phone? Finish signup now.

Full name: workshop
Your full name will appear on your public profile

Username: eliedemo
Your public profile: http://twitter.com/eliedemo

Password: ************

Are you human?

Before we create your account, we need to make sure you're not a computer.

Type the words above

Create my account

I want the inside scoop—please send me email updates!
CAPTCHAS

Elie Bursztein (@elie)

The Failure of Noise-Based Non-Continuous Audio Captchas

http://ly.tl/p18
The Failure of Noise-Based Non-Continuous Audio Captchas
The Failure of Noise-Based Non-Continuous Audio Captchas
The Failure of Noise-Based Non-Continuous Audio Captchas
Are you human?

Before we create your account, we need to make sure you’re not a computer.

Type the words above

Finish

Can’t read this?

- Get two new words
- Hear a set of words

Powered by reCAPTCHA.
Help
The Failure of Noise-Based Non-Continuous Audio Captchas
Audio capchas

The not-so-fine print

For added security, please enter the verification code hidden in the image.

Refresh the image | Listen to the verification code

stop spam. read books.
The Failure of Noise-Based Non-Continuous Audio Captchas

For added security, please enter the verification code hidden in the image.

The not-so-fine print

stop spam. read books.
Outline

• Audio captchas background
• Breaking audio captchas
• Evaluation results
• Demo
Creating audio captcha

Captcha Maker
Creating audio captcha

Voices

Captcha Maker
Creating audio captcha

Captcha Maker

Noises
Creating audio captcha

Super secure captcha
Type of noise

- Additive noise i.e. white noise
- Convolutive noise i.e. echo
- Semantic noise i.e. music
Noise intensity (RMS/SNR)

Authorize

Digg

Microsoft
Sound representation

WAV → DFT → Cep → TFR → TDC

TCR
Breaking audio captchas
Solving an audio captcha
Solving an audio captcha
Solving an audio captcha
Solving an audio captcha
Solving an audio captcha

Elie Bursztein (@elie)
Solving an audio captcha
Solving an audio captcha
Dealing with random noise

- Statistical learning
- Supervised learning
- RLS (Regularized least square) classifier

5:
- Authorize
- eBay
- Recaptcha

J:
- Authorize
- Digg
Solver efficiency

Solver accuracy = Coverage \times Precision^{\text{length}}

Coverage: Segmentation

Precision: Recognition rate
Decaptcha
Decaptcha overview

Web Site → Captcha scraping → Sound processing → Discretized and segmented captcha → Classifier → Answers

Mechanical Turk users → Captcha labels

Elie Bursztein (@elie)
The Failure of Noise-Based Non-Continuous Audio Captchas

http://ly.tl/p18
This section reports Decaptcha’s performance on a synthetic corpus that was generated following methodology introduced by [?]. The synthetic corpus uses nine types of noise described in Table III.

We measured Decaptcha’s performance on 2fififi captchas for each noise type at SNRs ranging from $-5$ to $30$ dB. Each captcha is composed of six spoken digits spaced randomly between $0.5$ and $4$ seconds.

<table>
<thead>
<tr>
<th>Family</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant Noise</td>
<td>White</td>
<td>White Gaussian noise.</td>
</tr>
<tr>
<td></td>
<td>buzz</td>
<td>Sine waves at 700 Hz, 2100 Hz and 3500 Hz.</td>
</tr>
<tr>
<td>Regular noise</td>
<td>pow</td>
<td>10 ms bursts of white Gaussian noise repeated every 100 ms.</td>
</tr>
<tr>
<td></td>
<td>rnoise</td>
<td>Every 100 ms, a section of the signal is replaced by white noise of the same RMS amplitude.</td>
</tr>
<tr>
<td></td>
<td>lofi</td>
<td>Add distortion, cracks, bandwidth limiting and compression. Simulates old audio equipment.</td>
</tr>
<tr>
<td></td>
<td>echo</td>
<td>The signal starts to echo at 0.6, 1.32, and 1.92 seconds.</td>
</tr>
<tr>
<td></td>
<td>disintegrator</td>
<td>Amplifies random half-cycles of the signal by a multiplier. Simulates a bad audio channel.</td>
</tr>
<tr>
<td>Semantic noise</td>
<td>chopin</td>
<td>Chopin Polonaise for Piano No. 6, Op. 53.</td>
</tr>
<tr>
<td></td>
<td>gregorian</td>
<td>Gregorian chant.</td>
</tr>
<tr>
<td></td>
<td>nina</td>
<td>“Just in time” by Nina Simone.</td>
</tr>
</tbody>
</table>
Synthetic evaluation

![](/content/fig.png)

- SNR (dB)
- Per-Captcha Precision (%)
- white
- buzz
- gregorian
- nina
- chopin
- pow
- echo, lofi, noise, disintegrator
The Failure of Noise-Based Non-Continuous Audio Captchas
## Captcha features

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Authorize</th>
<th>Digg</th>
<th>eBay</th>
<th>Microsoft</th>
<th>Recaptcha</th>
<th>Yahoo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Length</strong></td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td><strong>Type of voice</strong></td>
<td>Female</td>
<td>Female</td>
<td>Various</td>
<td>Various</td>
<td>Various</td>
<td>Child</td>
</tr>
<tr>
<td><strong>Background Noise</strong></td>
<td>None</td>
<td>Constant (random)</td>
<td>Constant (random)</td>
<td>Constant (random)</td>
<td>Constant (random)</td>
<td>None</td>
</tr>
<tr>
<td><strong>Intermediate noise</strong></td>
<td>None</td>
<td>None</td>
<td>Regular (speech)</td>
<td>Regular (speech)</td>
<td>Regular (speech)</td>
<td>Regular (speech)</td>
</tr>
<tr>
<td><strong>Charset</strong></td>
<td>0-9a-z</td>
<td>a-z</td>
<td>0-9</td>
<td>0-9</td>
<td>0-9</td>
<td>0-9</td>
</tr>
<tr>
<td><strong>Avg. duration</strong></td>
<td>5.0</td>
<td>6.8</td>
<td>4.4</td>
<td>7.1</td>
<td>25.3</td>
<td>18.0</td>
</tr>
<tr>
<td><strong>Sample rate</strong></td>
<td>8000</td>
<td>8000 8000</td>
<td>8000</td>
<td>8000</td>
<td>8000</td>
<td>22050</td>
</tr>
<tr>
<td><strong>Beep</strong></td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Coverage</th>
<th>Digit</th>
<th>Captcha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authorize</td>
<td>5</td>
<td>100</td>
<td>97</td>
<td>89.2%</td>
</tr>
<tr>
<td>Digg</td>
<td>5</td>
<td>100</td>
<td>76</td>
<td>41.4%</td>
</tr>
<tr>
<td>eBay</td>
<td>6</td>
<td>85.6</td>
<td>92.5</td>
<td>82.9%</td>
</tr>
<tr>
<td>Microsoft</td>
<td>10</td>
<td>80.6</td>
<td>89.6</td>
<td>48.9%</td>
</tr>
<tr>
<td>Recapcha</td>
<td>8</td>
<td>99.9</td>
<td>40.5</td>
<td>1.5%</td>
</tr>
<tr>
<td>Yahoo</td>
<td>7</td>
<td>99.1</td>
<td>74.7</td>
<td>45.4%</td>
</tr>
</tbody>
</table>
Recaptcha semantic noise
Recaptcha semantic noise

![Graph showing the temporal distribution of audio noise in decibels (DB) over time in seconds. The graph has a horizontal line representing the noise level, with peaks and troughs indicating variations in noise intensity.](http://ly.tl/p18)
Confusion matrices

Microsoft

Recaptcha

Yahoo

Predicted Digit

Actual Digit

Percent

http://ly.tl/p18
How many captchas do you need?

![Graph showing the relationship between corpus size and captcha precision for different websites.](http://ly.tl/p18)
Conclusion

• Non-continuous based captchas are broken
• Urgent need to come-up with the next generation of audio captchas
Questions?

Thanks

http://ly.tl/p18

Twitter: @elie