Nathan Dykstra
SE 4930
Article Review 2
January 11, 2012

**Article Choice**

I read the article *The Failure of Noise-Based Non-Continuous Audio Captchas* by Elie Bursztein, Romain Beauxis, Hristo Paskov, Daniele Perito, Celine Fabry, and John Mitchell. It is from the 2011 IEEE Symposium on Security and Privacy.

**What Was Discussed**

The authors begin by discussing what non-continuous audio captchas are: a sequence of spoken letters and/or numbers distorted by various kinds of noise. Captchas are used on websites to distinguish human users from automated bots attempting to use the system. The goal is to make these tests easy for humans but difficult for computers. Some users require audio captchas in place of visual ones for various reasons (blindness, color blindness, etc).

The team goes on to explain how most of the current audio captchas are based on work over a decade old by a researcher named Kochanksi and that there already exists a two-phase solution for solving these older tests. This new team has developed a new two-phase system that can accurately solve modern audio captchas, often more efficiently than humans. The two-phase systems first extracts portions of the audio containing a digit and then use machine learning algorithms to identify the digit. These algorithms are able to be trained on specific captcha schemes and the team found that these new systems were significantly better at solving audio captchas than speech recognition software.

The authors' new tool, Decaptcha, was tested on a large scale on real-world captchas. They state that their system "is able to solve Microsoft's audio captchas with 49% success and Yahoo's with 45% success, often achieving better accuracy than humans". In addition, it was interesting to note that this tool requires very little training to solve the most difficult captcha schemes. Their system merely requires 20 minutes and 300 labelled captchas and is then able

to defeat the most difficult audio captchas at a rate of tens per minutes on a single desktop system.

The team then describes how their particular algorithm works by separating the audio based on noise intensity and other factors. The system then classifies the segments and works to identify the digits being spoken. They use an algorithm called the Regularized Least Squares Classification (RLSC). This is a binary classifier that uses a complex summation equation on its trained data points, assigning them to positive and negative classes.

The Decaptcha system was tested on tests from Authorize, Digg, eBay, Microsoft, Recaptcha, and Yahoo. Interesting, the team noted that they were unable to test Google's audio captchas because they were too difficult for humans to properly label them for training. Ultimately, the Decaptcha system solved a high percentage of audio captchas ( 89% for Authorize, 41% for Digg, 82% for eBay, 48.9% for Microsoft, 45.45% for Yahoo).

**What I Learned**

Firstly, I learned that captcha is actually an acronym (C.A.P.T.C.H.A) for "Completely Automated Public Turing tests to tell Computers and Humans Apart". I also learned, generally, what a binary classifier does and read about several different types of this system. I also was introduced to Fourier Transforms. I know E.E.s use these significantly but as an S.E. I have not seen nor used this family of equations.

Finally, and probably most importantly, I learned that modern audio captchas are not secure. Visual captchas, when appropriate, are much more secure. The authors mention an interesting conjecture on this as well:

> " One reason that audio captchas might be weaker than visual captchas stems from human physiology: the human visual system consumes a far larger portion of our brains than the human audio processing system. In addition, modern signal processing and

> machine learning methods are fairly advanced. As a result, the difference between human and computer audio capabilities is likely significantly less than the difference between human and computer visual processing."

**How It Will Help**

Currently, I don't work with any web technologies where captchas are used. However, given the findings of this team's research I am now aware of the shortcomings of audio captchas. That being said, it may be difficult to use alternatives like visual captchas due to previously mentioned issues (impaired vision, etc).

**Questions**

None at this time.