Stanford computer scientists find Internet security flaw

May 24 2011, By Melissae Fellet

Postdoctoral researcher Elie Bursztein, left, and John Mitchell, a professor of computer science, with colleagues built a computer program that revealed a security flaw in commercial audio captchas used by major Internet companies.

(PhysOrg.com) -- Researchers at the Stanford Security Laboratory create a computer program to defeat audio captchas on website account registration forms, revealing a design flaw that leaves them vulnerable to automated attacks.

Stanford researchers have found an audible security weakness on the Internet.

If you've ever registered for online access to a website, it's likely you were required as part of the process to correctly read a group of distorted letters and numbers on the screen.

That's a simple test to prove you're a human, not a computer program with malicious intent.

Though computers are good at filling out forms, they struggle to decipher these wavy images crisscrossed with lines, known as captchas (short for Completely Automated Public Turing test to tell Computers and Humans Apart).

But there's a second type of captcha, and it may pose more of a security weakness. These audio captchas, designed to help the visually impaired, require users to accurately listen to a string of spoken letters and/or numbers disguised with background noise.

John Mitchell, a professor of computer science, postdoctoral researcher Elie Bursztein and colleagues built a computer program that could listen to and correctly decipher commercial audio captchas used by Digg, eBay , Microsoft, Yahoo and reCAPTCHA, a company that creates captchas.

The researchers presented their results during a symposium on security and privacy in Oakland, Calif.

The Stanford program, called Decaptcha, successfully decoded Microsoft's audio captcha about 50 percent of the time. It correctly broke only about 1 percent of reCAPTCHA's codes, the most difficult ones of those tested, but even this small success rate is considered trouble for websites such as YouTube and Facebook that get hundreds of millions of visitors each day.

Imagine a large network of malicious computers creating many fake accounts on YouTube. This robot network of accounts could highly rate the same video, falsely increasing its popularity and thereby its advertising revenue. "Bot" networks could also swamp email accounts with spam messages.

Decoding sounds

Computers have a tough time attempting to read image captchas, but Mitchell and Bursztein wondered if audio captchas were safe from automated attacks, too.

The researchers taught their program to recognize the unique sound patterns for every letter of the alphabet, as well as numeral digits. Then they challenged their software to decode audio captchas it had never heard before.

The program worked by identifying the sound shapes in the target captcha file, comparing them to those stored in its memory. It worked – the software could to some extent imitate human hearing.

"In the battle of humans versus computers, we lost round one for audio captchas," Bursztein said. "But we have a good idea of what round two should be."

Designing captchas is challenging. The tests must be simple enough for users to answer quickly, yet complicated enough so computers struggle to decipher the patterns. Background noise in an audio captcha can confuse computers, but little is known about the types of noises that trip them up the most.

The researchers generated 4 million audio captchas mixed with white noise, echoes or music, and challenged the program to decode them. After training Decaptcha with some samples, they took it for a test drive.

The program easily defeated captchas mixed with static or repetition, with a 60 to 80 percent success rate, but background music made the task more difficult.

Decaptcha removes the background noise from each audio file, leaving distinctively shaped spikes of energy for each digit or letter in the captcha. The program clearly isolates these spikes from white noise or echoes. But when the captcha contains noises that mimic these energy spikes, Decaptcha is often confused.

Building a program to solve captchas is "an interesting test case for machine learning technology," said Mitchell. "For audio, it's in a realm where machines should do better than humans."

Add meaning

And they do, until they have to think like us. Music lyrics or garbled voices are forms of semantic noise – sounds that carry meaning. Humans can recognize a message mixed with semantic noise, but computers can't distinguish the two clearly. Decaptcha correctly solved only about 1 percent of these captchas.

Of the commercial captchas the team tested, reCAPTCHA was the strongest because it contains background conversation and other semantic noise. Microsoft and Digg have recently changed their audio captchas to use this technology, Bursztein said. But the creation of this latest captcha cracker shows that even the best approach isn't secure enough. "The replacement technology isn't there yet, but we've pinpointed the problem," he said.

Citing data obtained from eBay, the researchers say about 1 percent of people who register at the site use audio captchas. That's enough users to warrant an effort to strengthen this security device.

The researchers suggest programmers tap into our human ability to understand meaning in sounds to improve future captchas. More secure puzzles could include background music or entire words instead of a

string of letters. But the team cautions that programmers need to keep the human user in mind. If the captcha is too complicated, legitimate users won't be able to decode it.

Despite efforts to strengthen audio captchas against computer attacks, they will, like visual captchas, still be vulnerable to crowdsourced attacks by a group of people manually solving captchas for low wages.

Captchas are vital to freedom on the Internet, the researchers say, as the value of many social media sites depends on the assumption that fellow users are humans.

"Captchas are a big inconvenience to people," Mitchell said. "The fact that they're so widely used is evidence of their necessity."

Provided by Stanford University