center for information technology integration

projects techreports press lab location staff

Steganography Press Information

This page is meant to assist the press in finding information about the ongoing search for steganographic content.

The following paragraphs answer frequently asked questions.

What is this all about?

- Steganography is the art and science of hidden communication.
- In February 20001, the *USA Today* reported that terrorist have been using steganography to hide communication in images on the Internet.
- Motivated by the article, Niels Provos developed a steganography detection framework, which he used to analyze two million images from the Internet auction site eBay. It consist of three tools:
  - crawl - a web crawler that downloads images from the web.
  - Stegdetect/Stebreak - tools that identify images that might contain hidden messages, and then guess the secret key required to retrieve a hidden message if it exists.
  - disconcert - a distributed computing framework that assists stegbreak by running it on a cluster of workstations.
- Not a single hidden message was found.
- Niels Provos is a doctoral candidate at the University of Michigan, working with his advisor Peter Honeyman at the Center for Information Technology Integration.
- The details of the research are outlined in "Detecting Steganographic Content on the Internet" by Niels Provos and Peter Honeyman, NDSS '02.

Why eBay?

- In February 2001, the article Secret Messages Come in .Wavs in *Wired News* mentioned eBay and Amazon as places that carry steganographic content.
- eBay has a very organized web structure that facilitates downloading images pointed to by auctions.

What are the results?

- Not a single hidden message was found in images that were obtained from eBay auctions.
- The recent ABC news coverage about steganography provided the first real steganographic image; see ABC Steganography Trophy.

What about images from USENET?

- To increase the scope of the study, Niels Provos and Peter Honeyman analyzed one million images from USENET archives for hidden messages.
  - The processing rate of the USENET archive was about 370,000 images per day. We analyzed about one million images.
  - The peak performance of the disconcert cluster is 870,000 keys per second. The cluster consists of about two-hundred workstations running OpenBSD, Solaris, Linux and FreeBSD.
- A dictionary attack against the suspicious images revealed no hidden mesages. Our dictionary contains about 1.8 million words and phrases.
- Detailed results from the USENET search are available.

How does dictionary attack work on steganographic systems?

- Steganographic systems embed header information in front of the hidden message. The header contains information about the length of the message, compression methods, etc...
- Dictionary attack with stegbreak chooses a key from a dictionary and uses it to retrieve header information. If the header makes sense, the guessed key is a candidate.
- Our dictionary contains about 1,800,000 words and phrases.
  - The words are from English, German, French, Science Fiction novels, the Koran, famous movies, songs, etc...
- Dictionary attack on JPHide and JSteg-Shell is completely independent of the hidden data. For OutGuess, file magic is used to cut down on false positives.

For further questions, please contact Niels Provos <provos@citi.umich.edu>.

Niels Provos
Last modified: Fri Jan 4 07:12:09 EST 2002

Email information@citi.umich.edu
or call +1 734 763 2929

projects | techreports | press | lab | location | staff