AMES LABORATORY

NEWS RELEASE
Office of Public Affairs
111 TASF
Ames, IA 50011-3020
http://www.ameslab.gov

Contacts:
Clifford Bergman, 515-294-8137
Jennifer Davidson, 515-294-2941
Kerry Gibson, Public Affairs, 515-294-1405

For release: May 24, 2006

FINDING COMPUTER FILES HIDDEN IN PLAIN SIGHT

Ames Laboratory researchers detect secret files lurking within digital images

AMES, IA – Keeping computer files private requires only the use of a simple encryption program. For criminals or terrorists wanting to conceal their activities, however, attaching an encrypted file to an e-mail message is sure to raise suspicion with law enforcement or government agents monitoring e-mail traffic.

But what if files could be hidden within the complex digital code of a photographic image? A family snapshot, for example, could contain secret information and even a trained eye wouldn't know the difference.

That ability to hide files within another file, called steganography, is here thanks to a number of software programs now on the market. The emerging science of detecting such files – steganalysis – is getting a boost from the Midwest Forensics Resource Center at the U.S. Department of Energy's Ames Laboratory and a pair of Iowa State University researchers.

Electronic images, such as jpeg files, provide the perfect "cover" because they're very common – a single computer can contain thousands of jpeg images and they can be posted on Web sites or e-mailed anywhere. Steganographic, or stego, techniques allow users to embed a secret file, or payload, by shifting the color values just slightly to account for the "bits" of data being hidden. The payload files can be almost anything from illegal financial transactions and the proverbial off-shore account information to sleeper cell communications or child pornography.

"We're taking very simple stego techniques and trying to find statistical measures that we can use to distinguish an innocent image from one that has hidden data," said Clifford Bergman, ISU math professor and researcher on the project. "One of the reasons we're focusing on images is there's lots of 'room' within a digital image to hide data. You can fiddle with them quite a bit and visually a person can't see the difference."

"At the simplest level, consider a black and white photo – each pixel has a grayscale value between zero (black) and 255 (white)," said Jennifer Davidson, ISU math professor and the other investigator on the project. "So the data file for that photo is one long string of those grayscale numbers that represent each pixel."

Encrypted payload files can be represented by a string of zeros and ones. To embed the payload file, the stego program compares the payload file's string of zeros and ones to the string of pixel values in the image file. The stego program then changes the image's pixel values so that an even pixel value represents a zero in the payload string and an odd pixel value represents a one. The person receiving the stego image then looks at the even-odd string of pixel values to reconstruct the payload's data string of zeros and ones, which can then be decrypted to retrieve the secret file.

"Visually, you won't see any difference between the before and after photo," Davidson said, "because the

- more -

shift in pixel value is so minor. However, it will change the statistical properties of the pixel values of the image and that's what we're studying."

Given the vast number of potential images to review and the variety and complexity of the embedding algorithms used, developing a quick and easy technique to review and detect images that contain hidden files is vital. Bergman and Davidson are utilizing a pattern recognition system called an artificial neural net, or ANN, to distinguish between innocent images and stego images.

Training the ANN involved obtaining a database of 1,300 "clean" original images from a colleague, Ed Delp, at Purdue University. These images were then altered in eight different ways using different stego embedding techniques – involving sophisticated transfer techniques between the spatial and wavelet domains – to create a database of over 10,000 images.

Once trained, the ANN can then apply its rules to new candidate images and classify them as either innocent or stego images.

"The ANN establishes kind of a threshold value," Bergman said. "If it falls above the threshold, it's suspicious.

"If you can detect there's something there, and better yet, what method was used to embed it, you could extract the encrypted data," Bergman continued. "But then you're faced with a whole new problem of decrypting the data … and there are ciphers out there that are essentially impossible to solve using current methods."

In preliminary tests, the ANN was able to identify 92 percent of the stego images and flagged only 10 percent of the innocent images, and the researchers hope those results will get even better. An investigator with the Iowa Department of Criminal Investigation is currently field-testing the program to help evaluate its usefulness and a graphical user interface is being developed to make the program more user friendly.

"Hopefully we can come up with algorithms that are strong enough and the statistics are convincing enough for forensic scientists to use in a court of law," Bergman said, "so they can say, 'There's clearly something suspicious here,' similar to the way they use DNA evidence to establish a link between the defendant and the crime."

The project is funded by the Midwest Forensics Resource Center. The MFRC, operated by Ames Laboratory, provides research and support services to crime laboratories and forensic scientists throughout the Midwest.

Ames Laboratory is operated for the Department of Energy by Iowa State University. The Lab conducts research into various areas of national concern, including energy resources, high-speed computer design, environmental cleanup and restoration, and the synthesis and study of new materials.

###