B13 A Strategy for Testing Graphic File Carving Tools

James R. Lyle, PhD, NIST, 100 Bureau Drive, Mail Stop 8970, Gaithersburg, MD 20899; and Richard Ayers, MS, 100 Bureau Drive, MS 8970, Gaithersburg, MD 20899-8970*

After attending the presentation, attendees will be made aware of some of the issues revealed by creating and using data sets for testing computer forensic tools for carving deleted files from unallocated file space.

The presentation will impact the forensic science community by increasing awareness of tool test strategies and their ability to reveal anomalies in tool behavior. This presentation will aid the forensic practitioner in recognizing the limitations of file carving tools.

The Computer Forensic Tool Testing (CFTT) project at NIST develops methodologies for testing computer forensic tools. This presentation reports on creating data sets for testing file carving tools and their behaviors.

File carving is widely used in digital investigations to extract deleted files from unallocated storage. Usually file carving is applied to file types with a recognizable structure so that unallocated space can be scanned for file components that are then reassembled into complete files. If the file has easily identified beginning and ending content and is contiguously allocated then carving is simple. However, the reality of file fragmentation complicates the task considerably.

Categories of files that are common targets of file carving include:

Graphics: JPG, GIF, PNG, BMP, TIF & PCX
Videos: MP4, AVI, MOV, MPG, FLV & WMV
Audio: MP3, WAV, AU & WMA
Document: DOC, DOCX, XLS, XLSX, PDF, PPT & PPTX
WEB: HTML, SQLite & chat
Archive: ZIP, RAR, 7Z, GZ & TAR
Misc: exec, logs, etc.

A common tool testing strategy is needed to help investigators characterize and understand tool behaviors, to compare tools, and to create test data with known content for investigator practice and training.

Data sets are available for download with each being available as a single file as would be created using the UNIX **dd** command to image a storage device.

A set of complete and contiguous graphics files with no intervening data between files – the last sector of each file is padded with zeros until the end of the sector. This test set reveals problems finding the beginning and the end of graphics files.

A set of complete and contiguous graphics files – the gap between files is filled with varied data in common cluster sizes, i.e., 1, 2, 4, 8, 16, 32, or 64 sectors. The gap content is varied over zeros, constant value, random data and various text formats. This test set reveals problems triggered by the presence of non-graphics files.

Files are placed such that each file does not begin on a sector boundary, but is offset from the sector beginning – this test differentiates algorithms that require the file to start on a sector boundary from those that can find an embedded graphics file.

Simple Fragmentation – a set of sequential fragmented graphics files. This test identifies how a tool deals with fragmentation.

Complex Fragmentation – the test set has some fragments out of order and some fragments intertwined. This test identifies algorithms that deal with complex fragmentation.

Incomplete Fragmented Files – some fragments are missing. This test set reveals complications when encountering incomplete files.

In this presentation, existing file carving test sets are examined to identify the underlying assumptions that have guided the creation of the test sets. From this, a new set of test images have been developed that can help investigators characterize and understand tool behaviors, compare tools, satisfy laboratory accreditation requirements and create test data with known content for investigator practice and training.

Certain trade names and company products are mentioned in the text or identified. In no case does such identification imply recommendation or endorsement by the author or the author's employer, nor does it imply that the products are necessarily the best available for the purpose.

The test images and image layout documentation are available at the CFReDS project web site:

** Presenting Author*

http://www.cfreds.nist.gov/FileCarving/.

Test reports on specific tools are available from the Department of Homeland Security Cyber FETCH web site: https://www.cyberfetch.org/.

Digital, Software, Testing