National Institute of
Standards and Technology
Technology Administration
U.S. Department of Commerce

NISTIR 7490

Digital Forensics at the National Institute of Standards and Technology

James R. Lyle
Douglas R. White
Richard P. Ayers

NISTIR 7490

Digital Forensics at the National Institute of Standards and Technology

James R. Lyle
Douglas R. White
Richard P. Ayers

Software Diagnostics and Conformance Testing

Software Diagnostics and Conformance Testing Division
Information Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899-8970

April 2008

U.S. Department of Commerce
Carlos M. Gutierrez, Secretary

National Institute of Standards and Technology
Dr. James M. Turner, Acting Director

Reports on Computer Systems Technology

The Information Technology Laboratory (ITL) at the National Institute of Standards and Technology (NIST) promotes the U.S. economy and public welfare by providing technical leadership for the Nation's measurement and standards infrastructure. ITL develops tests, test methods, reference data, proof of concept implementations, and technical analysis to advance the development and productive use of information technology. ITL's responsibilities include the development of technical, physical, administrative, and management standards and guidelines for the cost-effective security and privacy of sensitive unclassified information in Federal computer systems. This Interagency Report discusses ITL's research, guidance, and outreach efforts in computer security, and its collaborative activities with industry, government, and academic organizations.

National Institute of Standards and Technology Interagency Report
7 pages (2008)

Abstract

There are three digital forensic science projects: National Software Reference Library (NSRL), Computer Forensic Tool Testing (CFTT), Computer Forensic Reference Data Sets (CFReDS) currently providing resources for the digital investigator underway at the National Institute of Standards and Technology (NIST) Information Technology Laboratory – Software Diagnostics and Conformance Testing Division. These projects are supported by the U.S. Department of Justice's National Institute of Justice (NIJ), federal, state, and local law enforcement, and the National Institute of Standards and Technology Office of Law Enforcement Standards (OLES) to promote efficient and effective use of computer technology in the investigation of crimes involving computers. Numerous other sponsoring organizations from law enforcement, government, and industry are also providing resources to accomplish these goals. The "Digital Forensic at the National Institute of Standards and Technology" paper provides an overview of the before mentioned projects and methodologies.

Overview

There are three digital forensics projects currently providing resources for the digital investigator underway at the National Institute of Standards and Technology (NIST). These projects are supported by the U.S. Department of Justice's National Institute of Justice (NIJ), federal, state, and local law enforcement, and the National Institute of Standards and Technology Office of Law Enforcement Standards (OLES) to promote efficient and effective use of computer technology in the investigation of crimes involving computers. Numerous other sponsoring organizations from law enforcement, government, and industry are also providing resources to accomplish these goals. The three projects are the following:

- National Software Reference Library (NSRL)
- Computer Forensic Tool Testing (CFTT)
- Computer Forensic Reference Data Sets (CFReDS)

NSRL

The NSRL is designed to collect software from various sources and incorporate file profiles computed from this software into a Reference Data Set (RDS) including hashes of known files created when software is installed on a computer. The law enforcement community approached NIST requesting a software library and signature database that meets four criteria:

- The organizations involved in the implementation of the file profiles must be unbiased and neutral.
- Control over the quality of data provided by the database must be maintained.
- A repository of original software must be made available from which data can be reproduced.
- The database must provide a wide range of capabilities with respect to the information that can be obtained from file systems under investigation.

The primary focus of the NSRL is to aid computer forensics examiners in their investigations of computer systems. The majority of stakeholders are in federal, state and local law enforcement in the United States and internationally. These organizations typically use the NSRL data to aid in criminal investigations. Other stakeholders include businesses and other government agencies which may use the NSRL as part of their routine IT operations.

The NSRL has three components:

- A collection of over 9,000 original software packages.
- A database containing detailed information about the files in those software packages.
- A public NSRL Reference Data Set (RDS) which contains a subset of the metadata held in the database. The RDS is published and updated quarterly, as NIST Special Database 28.

The collection of original software allows NIST to investigate file metadata that may be called into question. The collection allows new algorithms to be applied against the files in the future, to address cryptographic breakthroughs or other investigative needs. This software includes virtually any type available, such as operating systems, database management systems, utilities, graphics images, component libraries, etc., in many different versions. The collection contains software dating back to the early 1980's.

The NSRL database contains metadata on computer files which can be used to uniquely identify the files and their provenance. For each file in the NSRL collection, the following data are published:

- Cryptographic hash values (MD5 and SHA-1) of the file's content. These uniquely identify the file even if, for example, it has been renamed.
- Data about the file's origin, including the software package(s) containing the file and the manufacturer of the package.
- Other data about the file, including its original name and size.

The RDS is used by law enforcement, government, and industry organizations to review files on a computer by matching file profiles in the RDS using an automated system. The reference data is used to rapidly identify files on computer systems, based solely on the content of the files.

In most cases, NSRL file data is used to eliminate known files, such as operating system and application files, during criminal forensic investigations. This reduces the number of files that must be manually examined and thus increases the efficiency of the investigation. The current distribution method of known file metadata is becoming unwieldy, and the NSRL is researching more effective methods.

The RDS is a collection of digital signatures of known, traceable software applications. Currently, metadata and hash values for over 40 million files are available in the RDS. There are applications in the NSRL which may be considered malicious, i.e. steganography tools and hacking scripts. The RDS is intended to be used as a filter of *known* file signatures, not *known good*. When used in this manner, the process is fail-safe; unknown files will remain for review by an investigator. There are no instances of illicit data, i.e. child abuse images. Further details are available at http://www.nsrl.nist.gov.

CFTT

The goal of the CFTT project at NIST is to establish a methodology for testing computer forensic software tools through the development of general tool specifications, test procedures, test criteria, test sets, and test hardware. The results provide the information necessary for toolmakers to improve tools, for users to make informed choices about acquiring and using computer forensics tools, and for interested parties to understand the tools capabilities.

The testing methodology developed by NIST is functionality driven. The activities of forensic investigations are separated into discrete functions, such as hard disk write protection, disk imaging, string searching, etc. A test methodology is then developed for each category. After a test methodology is developed it is posted to the web and can be used by anyone to test the specified function implemented in a computer forensic tool.

After a tool category is selected the development process is as follows:

1. NIST staff and law enforcement representatives develop a specification document that sets forth requirements that the forensic tool should meet.
2. The specification is posted to the web for peer review by members of the computer forensics community and for public comment by other interested parties.
3. Relevant comments and feedback are incorporated into the specification.
4. A test methodology is developed and an assertions and test plan document that specifies how to implement the test methodology is produced.
5. The test plan document is posted to the web for peer review by members of the computer forensics community and for public comment by other interested parties.
6. Relevant comments and feedback are incorporated into the specification.
7. A test environment with support software is designed and implemented for the test plan.
8. NIST posts support software to the web.

Once a tool is selected for testing, the test process is as follows:

1. NIST acquires the tool to be tested.
2. NIST reviews the tool documentation.
3. NIST selects relevant test cases depending on features supported by the tool.
4. NIST develops test strategy.
5. NIST executes test cases.
6. NIST produces test report.
7. Steering Committee reviews test report.
8. Tool vendor reviews test report.
9. NIJ posts test report to web. (http://www.ojp.usdoj.gov/nij/topics/ecrime/cftt.htm)

NIJ has published test reports on several forensic imaging tools, several software write block tools, and a variety of hardware write block devices. Currently specifications and

test methodologies for deleted file recovery and string searching tools are in development. In addition to forensic tools for acquisition and analysis of digital data on desktop and laptop computers, CFTT is also developing test methodologies for mobile devices.

Data acquisition performed on cellular devices operating over Global System for Mobile Communications (GSM) and non-GSM networks has proven not only frustrating but extremely tedious due to the rapid rate at which new cellular devices are introduced into the market. Software vendors specializing in cellular forensics are forced to continuously provide updates to software and associated hardware in order to maintain support and provide examiners with solutions for the latest technologies. Mobile device forensic research performed at the National Institute of Standards and Technology (NIST) Information Technology Laboratory (ITL) has produced numerous reports on tools capable of acquiring data from Personal Digital Assistants (PDAs), smart phones, and cellular devices operating over GSM and non-GSM networks.

NIST has presented at numerous conferences world-wide providing software vendors, forensic specialists, incident response team members, and law enforcement an overview of the current capabilities and limitations of forensic applications capable of acquiring data from cellular devices as well as suggestions on preservation and handling of digital data. Research conducted over the past two years has produced the following publications: NISTIR 7250 *Cell Phone Forensic Tools: An Overview and Analysis*, SP800-101 *Guidelines on Cell Phone Forensics*, NISTIR 7387 *Cell Phone Forensic Tools: An Overview and Analysis Update*, *Forensic Software Tools for Cell Phone Subscriber Identity Modules*.

In addition to the NIST reports and conference articles produced our research has provided extensive involvement with software engineers from various manufacturers troubleshooting potential issues, providing suggestions on product improvement and overall dependability, which have played a key role in the evolution of cellular forensics software. Research conducted and shared materials have provided academia with a starting point for education materials and continue to inform law enforcement and forensic examiners of expectations of the interaction between numerous devices and tools.

CFReDS

The **Computer Forensic Reference Data Sets (CFReDS)** provide to an investigator documented sets of simulated digital evidence for examination. Since CFReDS has documented contents, such as target search strings seeded in known locations, investigators can compare the results of searches for the target strings with the known placement of the strings. Investigators can use CFReDS in several ways including validating the software tools used in their investigations, equipment check out, training investigators, and proficiency testing of investigators as part of laboratory accreditation. The CFReDS site is a repository of images. Some images are produced by NIST, often from the CFTT (tool testing) project, and some are contributed by other organizations. In addition to test images, the CFReDS site contains resources to aid in creating test images. These creation aids are in the form of interesting data files, useful software tools and procedures for specific tasks. The CFReDS web site is http://www.cfreds.nist.gov.