Defining Digital Forensic Examination and Analysis Tools

By

Brian Carrier

Presented At

The Digital Forensic Research Conference

DFRWS 2002 USA Syracuse, NY (Aug 6th - 9th)

DFRWS is dedicated to the sharing of knowledge and ideas about digital forensics research. Ever since it organized the first open workshop devoted to digital forensics in 2001, DFRWS continues to bring academics and practitioners together in an informal environment. As a non-profit, volunteer organization, DFRWS sponsors technical working groups, annual conferences and challenges to help drive the direction of research and development.

http:/dfrws.org

Defining Digital Forensic Examination & Analysis Tools

Brian Carrier

<carrier@atstake.com>

Where Security & Business Intersect℠

Definition of Digital Forensic Science

- **As defined at DFRWS 2001:**

  The use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations.

Identification and Analysis

- **We are restricting ourselves to the digital forensic phases of identification and analysis**
- **Using the previous definition, the goal of these phases can be expressed as:**

  To identify digital evidence using scientifically derived and proven methods that can be used to facilitate or further the reconstruction of events in an investigation.
- **All evidence is needed:**
  - Inculpatory Evidence
  - Exculpatory Evidence
  - Traces of tampering

Digital Forensics Complexity Problem

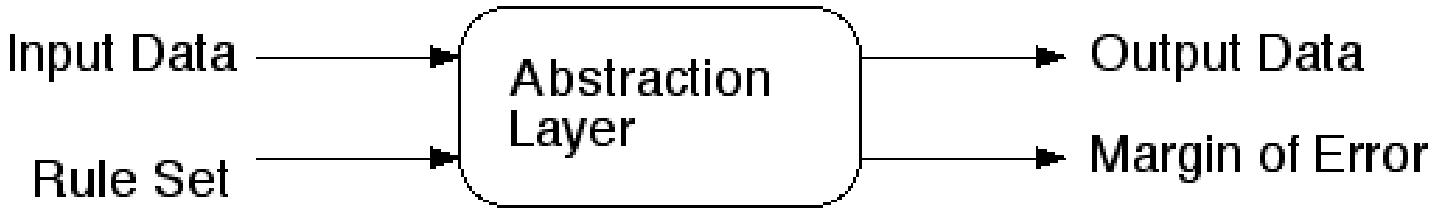
- **Data is typically acquired in its most raw format**
- **This is generally difficult for investigators to understand**
- **This problem has been solved by using tools to translate data through one or more layers of abstraction until it can be understood.**
- **Abstraction Layer Examples:**
  - File System Directories
  - ASCII
  - HTML
  - Network Packets
  - Intrusion Detection Systems (IDS)

Digital Forensic Analysis Tools

- It is proposed that the purpose of digital forensic analysis tools is to accurately present all data at a layer of abstraction and format that can be effectively used by an investigator to identify evidence.
- The needed layer of abstraction is dependent on the case and investigator

Abstraction Layers

- Used by all digital systems to customize generic interfaces
- Function with two inputs and two outputs
- The input rule set is typically the design specification



Tool Implementation Error

- **Errors introduced by bugs in the tools**
- **Examples:**
  - General programming bugs
  - Tool used an incorrect specification
  - Tool used the correct specification, but the original source did not
- **One can assume that the bugs are fixed when identified**
- **To factor in the potential for unknown bugs, a value could be calculated based on the history of a tool**
  - Likely be difficult to maintain for closed source tools that hide bugs that are not made public

Abstraction Error

- **Errors introduced by the abstraction theory**
- **Exists in layers that were not part of the original design**
- **Examples:**
  - Log processing
  - IDS alerts
- **This error can improve with research and better abstraction theories**

Analysis Tool Error Problem

- Data from digital forensic analysis tools will have some margin of error associated with them. This does not include the errors associated with previous tampering, acquisition, or interpretation. It only includes Tool Implementation Error and Abstraction Error.
- Evidence must have a margin of error associated with it and the output must be verified.

Layer Characteristics

- **Abstraction Error: Lossy Layers have an Abstraction Error and Lossless Layers have none**
- **Mapping: A One-to-One Layer can identify the input data given the output data and a Multiple-to-One Layer cannot**
- **Levels: Multiple levels of abstraction can occur, each having several layers of abstraction. A Boundary Layer is the last layer in a level (i.e. file contents).**
- **Tool Types: Translation Tools translate data from one layer to another. Presentation Tools present the layer data in a format that is useful for an investigator:**
  - Directory Entries sorted by directory
  - Directory Entries sorted by MAC times

Tool Requirements

- **Usability: Present data a layer of abstraction that is useful to an investigator (Complexity Problem)**
- **Comprehensive: Present all data to investigator so that both Inculpatory and Exculpatory Evidence can be identified**
- **Accuracy: Tool output must be able to be verified and a margin of error must be given (Error Problem)**
- **Deterministic: A tool must produce the same output when given the same rule set and input data.**
- **Verifiable: To ensure accuracy, one must be able to verify the output by having access to the layer inputs and outputs. Verification can be done by hand or a second tool set.**

Tool Recommendations

- **Read-Only: Because digital data can be easily duplicated, this is not a requirement. Although, to verify the results a copy of the input will be required at a later date.**

Conclusion

- Layers of abstraction are everywhere and have always been used
- Formal discussion of them has not occurred with Digital Forensics
- Lossy layers will be more common as new approaches are developed to decrease analysis time and log processing times
- A Tool Implementation Error value could help quantify the accuracy of a tool