Provenience-based cross-verification of digital forensic artifacts applied to NTFS

Learning objective

This presentation will demonstrate a methodology for comparing the subject data exploratory coverage of two digital forensic processes that share some in-common goals in their reporting. After attending this presentation, attendees will better understand the distinction between identifiers, and identifying characteristics, when describing artifacts recovered and interpreted in the course of digital forensic analysis. Attendees will see how two metadata summary languages, Digital Forensics XML (DFXML) and Cyber-investigation Analysis Standard Expression (CASE), can implement a strategy to compare the results of two file system analysis tools, and how a previous implementation of the strategy had relied on an artifact of unstable referential integrity, showing a need for using identifying characteristics with less chance of ambiguity.

Impact Statement

This presentation will impact the forensic science community by demonstrating cross-verification of digital forensic tool results, and how choice of artifact identifiers can impact the ability to compare results algorithmically.

Abstract

In digital forensics, file system analysis is a precursor task to event reconstruction. Often, unallocated content within a file system is content of interest to an investigation, and thus recognition, extraction, and ascription of unallocated files are typical intermediary steps en route to interpreting file system contents. The results of this general workflow form comprise a set of intermediary results worth cross-verification, due to potential impact on later interpretations of initial evidence. However, unallocated files often lack stable identifiers, presenting subtle challenges that can foil algorithmic comparison.

Unallocated content recovery requires careful understanding of the storage medium, or storage format, from which the content is recovered (Casey et al., 2019). This work focuses on a model of a file system where an allocated file's definition comprises at least three (Carrier, 2005) key dimensions: The data structure housing its metadata such as timestamp and owner, often referred to as an inode; the data structure compactly housing its location within the file system's namespace, often implemented as a directory entry; and the range within the file system that houses its contents, which might be discontiguous.

Prior work has used this model to implement differential analysis, both for comparing changes in a file system's state across time (Garfinkel et al., 2012), and across parse results when using multiple tools against the same subject image (Nelson et al., 2014). While the three-dimensional file model enables comparison of allocated content with seemingly little difficulty, some attempts to verify some POSIX-required characteristics of the allocated content show a weakness in the three-dimensional model that impacts interpretation of allocated and unallocated files.

We present a strengthening of the three-dimensional model, emphasizing a geometric representation of the three file dimensions as a first-order concern. This pattern extends in applicability beyond file system analysis, but is presented initially in the context of New Technology File System (NTFS) file system analysis. We demonstrate corrections over a model improvement previously proposed by Casey et al. (2019), and show results from extending two independently-developed open source tools to enable geometric comparability between their NTFS results. Using the tool-agnostic languages Digital Forensics XML (DFXML) (Garfinkel, 2012) and Cyber-investigation Standardized Analysis and Expression (CASE) (Casey et al., 2017), the geometry-based identifier strategy corrects a previous measurement of unallocated content.

References

* Carrier, B. File system forensic analysis. Addison-Wesley Professional, 2005.
* Casey, E., Barnum, S., Griffith, R., Snyder, J., van Beek, H., and Nelson, A. Advancing coordinated cyber-investigations and tool interoperability using a community developed specification language, Digital Investigation, Volume 22, 2017, Pages 14-45.
* Casey, E., Nelson, A. and Hyde, J. Standardization of file recovery classification and authentication. Digital Investigation, 31:100873, 2019.
* Garfinkel, S. Digital forensics XML and the DFXML toolset. Digital Investigation 8.3-4 (2012): 161-174.
* Garfinkel, S., Nelson, A.J., and Young, J. A general strategy for differential forensic analysis. Digital Investigation, 9:S50–S59, 2012. The Proceedings of the Twelfth Annual DFRWS Conference.
* Nelson, A.J., Steggall, E.Q., and Long, D.D.E. Cooperative mode: Comparative storage metadata verification applied to the XBox 360. Digital Investigation, 11:S46–S56, 2014. Fourteenth Annual DFRWS Conference.