;login:

THE USENIX MAGAZINE

December 2003 • volume 28 • number 6

Panel: Electronic Voting Security

Dan S. Wallach (Rice University) – Moderator

Jim Adler (VoteHere)
David Dill (Stanford University)
David Elliott (Washington State, Office of Sec. of State)
Douglas W. Jones (University of Iowa)
Sanford Morganstein (Populex)
Aviel D. Rubin (Johns Hopkins University)

inside:

SECURITY

BOOK REVIEWS AND HISTORY

USENIX NEWS

CONFERENCE REPORTS

Focus Issue: Security

Guest Editor: Rik Farrow

USENIX
The Advanced Computing Systems Association

Evidence Enhancing Technology

by Erin Kenneally

Erin Kenneally is a Forensic Analyst with the Pacific Institute for Computer Security (PICS), San Diego Supercomputer Center. She is a licensed attorney who holds Juris Doctorate and Master of Forensic Sciences degrees.

erin@sdsc.edu

Bridging the Techno-Legal Gap with Secure Audit Logging

Got Logs?

Computer logs may be used as evidence. Computer logs are like footprints for traditional crime scene investigators or financial ledgers for auditors. All of these objects are time machines, containing answers to questions related to IT processes, unlawful acts, and economic transactions, respectively. Inherent in each is the ability to reconstruct the who, what, when, where, why, and how of an IT, legal, or financial dispute. As crimes and social wrongs increasingly involve or target the use of computers, and as business relies on information systems to function, logs have become the digital eyewitnesses to transactions between computers and humans.

Realizing that eyewitnesses are only as valuable as their perception, memory, and cognition, so too are logs in their ability to paint a picture of digital events. Similarly, just as persons cannot predict or prepare for eyewitness events, it is difficult to foreknow which digital transactions will necessitate recreation in resolving a dispute. However, we can engineer reliable perception, memory, and cognition into our digital eyewitnesses through the process of secure audit logging (SAL).

SAL provides a general model for collecting and storing digital event data in accordance with legal admissibility standards and in compliance with the specific audit needs of systems administrators, law enforcement, and businesses. During legal disputes, investigators – system administrators, forensic examiners, regulators, private and public law enforcement – will often rely on audit and transaction logs as a source of evidence to prove/disprove their claims. These logs can contain virtually any type of data that a computer system is programmed to capture.

The SAL model facilitates the automated, centralized, and trustworthy collection and storage of any audit data that is dictated by a chosen policy. Information assurance and the ability to maintain the integrity of digital data for the purposes of legal proof are continually challenged by the nature of network computing, system bugs and vulnerabilities, and constantly changing technology. These features have conspired to facilitate confusion surrounding the admissibility of log records. The secure audit logging model is being designed with evidentiary standards and presumptions in mind. As such, SAL raises the bar for successful challenges to integrity of log records by including assurances of credibility – authenticity and reliability.

Secure Audit Logging: A Forensics Enhancing Technology

Many of the emerging applications for auditing and investigation are focused on data collection and monitoring within an organization's intranet. Unfortunately, existing tools for facilitating such capabilities require system administrators within these networks, who are not versed in legal principles, to deploy the technology that should enable business process within the context of its policies and legal directives. Such deployments frequently become mired in the difficulties of supporting user functionality, configuring hardware and software for compatibility, and providing other utility services, all within volatile distributed environments. The design purpose for logs produced by computers originated from utility service requirements. For instance, logs were system administrators' way of debugging, or troubleshooting, computers for various technical performance reasons. Absent this data collection mechanism, the ability to detect suspicious behavior and computer intrusions and distinguish between inadvertent machine error and malicious human tampering was implausible. As logs are increasingly being used for purpose of corporate governance, regulatory and legal forces are driving nontechnical folks to turn to log data to substantiate and defend against dispute claims. Because logs are humans' link to the who, what, when, where, and how of computer functioning and usage, logs are being thrust from the annals of computer "techdom" to the adversarial realm of jurisprudence. Like all other evidence offered for legal proof purposes, they must meet certain evidentiary standards.

In general, the standard for the admissibility of evidence is that it is shown to be relevant, authentic, and reliable.

Anatomy of the Law Applied to Computer Logs

LEGAL SEMANTICS AND PURPOSE OF EVIDENTIARY STANDARDS

In general, the standard for the admissibility of evidence is that it is shown to be relevant, authentic, and reliable. This includes that the evidence must not contain hearsay, unless it falls within an exception to the hearsay prohibition. These preliminary determinations can occur under the auspices of the Federal Rule of Evidence requirement that the matter in question is what it is claimed to be, or via the more demanding showing of reliability for scientific, technical, or specialized evidence. The purpose of this initial screening is to ensure that the evidence is reliable enough to go before a fact-finder, whose job it is to decide what weight that evidence should carry in resolving the issue at hand. In other words, a basic evidentiary tenet governing admissibility determinations is that there are guarantees of trustworthiness attached to the evidence so that a jury is not unduly confused or prejudiced.

AUTHENTICATION AND LOG EVIDENCE

Authentication standards are meant to ensure that the evidence is what it purports to be, and how rigorous a foundation is needed to make this finding depends on the existence of something that can be tested in order to prove a relationship between the document and an individual, and control against the perpetration of fraud.

The degree of scrutiny applied to determine whether or not computer log evidence is admissible is unsettled. This determination may turn on how a court categorizes the log evidence: computer-generated, computer-stored, or some hybrid. To date, there is no overarching prescription for establishing how computer logs should be categorized, thus leaving admissibility open to case-by-case determinations.

Generally, the authenticity control is established by testimony that the computer program which generated the record was functioning properly. It is important to keep in mind that this can rebutted if the source, method, or circumstance of preparation indicates lack of trustworthiness.

While increasing automation will diminish the number of witnesses qualified to authenticate computer-generated evidence like logs, inconsistencies at the human-computer interface when collecting, processing, and storing logs may provide fodder for log opponents to rebut the low threshold of proving authenticity and reliability and force proponents of log evidence to offer more solid foundational proof.

SECURE AUDIT LOGS AS THE DIGITAL CHAIN-OF-CUSTODY

The level of scrutiny and legal categorization of computer logs is ambiguous. To be clear, although the legal standard is unwavering – relevant, authentic, original – the application of the standard to log evidence is unsettled. While not comforting for those seeking black letter law on whether logs can be used as a sword or shield in legal disputes, there are controls that courts use to measure the reliability of evidence which can serve as a blueprint for attempts to ensure log admissibility. Arguably one of the most recognized reliability controls is chain-of-custody, and it is this concept that the SAL model mirrors.

Chain-of-custody (COC) is one of the controls used by courts to implement reliability standards. That is to say, authenticity of physical evidence is tested by accounting for the who, what, when, where, and how of a given piece of evidence from its initial discovery, to its collection, access, handling, storage, and eventual presentation at trial. COC has been institutionalized as a procedure for the seizure of physical evidence by law enforcement, as well as for the handling of digital evidence by computer forensic examiners as a measure of evidence integrity. The SAL ensures a digital chain-of-custody so as to minimize the challenges that digital evidence has been created, lost, damaged, or modified. SAL minimizes the manual human interfaces during collection and storage, as well as providing the metrics upon which legal determinations of reliability can be made.

SAL replicates the general procedures followed by computer forensic examiners to establish authenticity of physical evidence. These include:

- refraining from altering the original evidence
- documenting procedures used in collection, storage, and analysis and explaining any changes that may have been made to the evidence. These procedures should be auditable.
- maintaining the continuity of evidence; making a complete copy of data in question using a reliable copy process (independently verifiable; hashing)
- employing security measures (tamperproof storage, write protection)
- properly labeling time, date, source (tracking # and tagging)
- limiting and documenting the persons with access to data

LOG EVIDENCE RELIABILITY CONTROLS – WHAT IS THE LYNCHPIN OF CREDIBILITY?

Is the lynchpin of credibility for log data derived from the technology (computer and software producing the log), or from the person who reads and interprets the log data? In other words, who the real witness is should dictate what should be examined to measure the trustworthiness of statements in the logs. The nature of log evidence, unlike instances where a human is putting a pen to paper, suggests that the "real witness" is the chain of digital events surrounding the creation, transportation, and storage of logs. As such, courts should insist upon controls that measure the reliability with as little abstraction as possible.

Do the controls applied by courts to adjudge reliability log records ensure that evidence standards prescribed by the F.R.E. regulations and policy are being met? Controls are the guarantees of trustworthiness that enable an audit event to be measured against a standard or principle. The value of SAL lies in its ability to provide more direct guarantees of trustworthiness of log records, thereby reducing the uncertainty of legal risks. Even though the reliability controls for paper records are an abstraction of the controls for witness credibility, the underlying metric is the same: time (chronology), distance (location), and computation (cognition). SAL provides controls that more directly measure the lynchpin of credibility – the technology producing the logs – against the relevant standard.

When a witness takes the stand to testify, the audit event is what he is being asked to testify about – i.e., the accident he saw, what he did with the evidence being offered, or whether the computer was functioning – and is manifest by testing the witness's perception, memory, or narration/bias. Audit tools such as oath, personal presence at trial, and cross-examination are used to measure the credibility/trustworthiness of the witness's account of the audit event. The reliability metric the audit tools are measuring can be distilled into time (chronology), distance (location), and cognition.

As logs are increasingly used to resolve legal disputes and become the lynchpin of proof, focus will shift from presumptively ushering in the digital traces of business activities to disputing the logs used to buttress claims. Attempts to discredit logs will accompany this shift, and the technical folks who understand the mutability associated with current log data will be tapped for their knowledge that alterations (insertion, deletion, modification) are not only possible but probable, and oftentimes impossible to detect. This will be exacerbated by the emergence of software programs that expand data alteration capabilities to anyone with point-and-click capabilities, in contrast to the present state of affairs where log data alteration is limited to a small number of persons with the knowledge and skills to manually weave through log data and manipulate certain bits to reflect factual changes. The evidentiary significance is that continued reliance on controls such as proper functioning of the computer producing the logs do not speak to the threats to log integrity. Indeed, one's IDS, spreadsheet program, or email program may be working in tip-top shape, but that does not address the risk that the data it produces was altered by virtue of the interconnections or vulnerabilities posed by other persons and programs.

Two recent cases have turned this conjecture into reality. Log evidence was the subject of scrutiny in the acquittal of a U.K. teen accused of launching a DDoS attack that knocked out IT systems at the Port of Houston in Texas. The striking aspect of this case is that logs were used as both a sword and shield to support the increasingly popular, "unknown third party" defense. On one hand, the defense leveraged server logs showing regular probing of the defendant's computer to assert that it was possible the system could have been compromised and wielded by a remote hacker to perpetrate the crime. Simultaneously, the accused decried that the log files found on his system that implicated him in the attack were unreliable because his system was unpatched and thus susceptible to manipulation.

A similar tactic was used successfully to persuade a jury in a Montgomery County Circuit Court that an accountant charged with tax evasion was not guilty. Here the defendant blamed tax return inaccuracies on an unnamed computer virus. Despite evidence showing that the alleged virus did not affect the tax returns of clients prepared on the same computer, the defendant averted the maximum 33 years in prison and up to $900,000 in penalties.

Whether or not these outcomes are merely exceptional or the tip of an iceberg, they illustrate increasing reliance on digital evidence to fortify a particular rendition of the "truth." As the possibility of backdoors and vulnerabilities in systems challenge litigants to prove a negative in presenting or defending a claim, it becomes all the more

A battle of the logs will demand that the data contained therein is reliable.

important to establish the reliability of the digital footprints that paint the real picture of the “truth.” Because logs can be authoritatively persuasive for either party in a dispute, a battle of the logs will demand that the data contained therein is reliable. As such, control mechanisms must be employed to avoid finding reasonable doubt or a preponderance based on logs shrouded by conjecture.

By relying exclusively on humans as the only witness in addressing the reliability of log evidence, courts are not addressing the threats and vulnerabilities attendant to electronic evidence. They overlook the reality that computer hardware, software, and their interconnections converge to produce log evidence that is susceptible to events that render logs unreliable. This is simply an inadequate control to measure reliability. It is similar to claiming that all the cells in one’s body can be labeled as trustworthy because of the fact that the body’s organs and systems are healthy and functional.

If the law continues to use controls that provide second-order indicia of reliability, business reliance will be the control used to safeguard trust in logs. However, businesses run on commodity technology. The problem with relying on commodity technology to satisfy legal reliability standards is that it is driven by time-to-market forces and not built with legal standards in mind. This is not satisfactory when the costs of mistakes and errors are economically high and socially detrimental. Further, taking judicial notice of a process’s accuracy (i.e., that computers produce logs that can be relied on) may be confused with taking notice that a particular result is accurate (i.e., that logs submitted as evidence are trustworthy).

SAL addresses the log challenges by engineering the collection and preservation of logs with the principles and procedures of forensic integrity in mind. SAL offers more empirical evidence of the sequence of events surrounding log collection and storage, as well as minimizing the error that accompanies human interaction with log processes. By performing a digital chain-of-custody , the SAL model better fits the evidence whose trust is attempted to be measured.

The Timing Is Right

The development of this secure audit logging technology is motivated by the need to facilitate a just legal framework for establishing the trustworthiness of digital log records and for recognizing the fundamental uncertainties in the processes involved in utilizing these logs as evidence. These uncertainties are not being addressed by current information assurance and product development processes. IT departments lack meaningful guidance on how to utilize technology to comply with legal/regulatory standards, and reliance on vendors to know and foster the enabling technology is misguided at best. Further, these uncertainties risk being perpetuated if the assumptions underlying legal interpretations of the standards are institutionalized without proper measurement.

This article is an abbreviated version of a much more detailed work in progress. An extended version of this paper, complete with references, can be found at *http://security.sdsc.edu/*.