

Towards Identifying Criteria for the Evidential Weight of System Event Logs

A. Ahmad

Department of Information Systems, University of Melbourne
atif@unimelb.edu.au

A.B. Ruighaver

Department of Information Systems, University of Melbourne
tobias@staff.dis.unimelb.edu.au

Abstract

Despite the widespread use of computing in almost all functions of contemporary society and the consequently large number of forensic investigations where computing has been involved, there has been little progress made in adapting the primary mechanism by which computers record past activity, namely event logs to facilitate computer forensic investigation. From an evidence point of view system event logs do not readily conform to the requirements of a forensic investigation. We identify two criteria – Accuracy, and Completeness, and a third criterion - Utility that can be used to assess the evidential weight of system event information derived from event logs and to identify the desirable qualities of a forensically suitable event log.

Keywords

Digital Forensic Investigation, System Event Logs, Event Logs, Digital Evidence

INTRODUCTION

One of the primary aims of a traditional forensic investigation is to reconstruct past events in an attempt to answer the ‘what happened’ question. To achieve this aim forensic investigators treat the scene as the witness examining the environment as a source of trace evidence (Chisum and Turvey, 2000). In principle a witness is potentially a useful source of evidence. The witness may be able to recount the sequence of events that took place thereby assisting the reconstruction of the scenario for the investigation.

In the digital world where activity is conducted by processes, the ‘scene’ is the entire computing system including the processor, the memory, secondary storage devices, applications and so forth. Historically, investigators have typically studied storage devices like hard disks as they are usually the only source of preserved evidence. Although interactions among computing processes do not drop bodily properties akin to hair or blood, their interaction with and use of resources may leave traces that can be used to help reconstruct past events. In trying to ask ‘what happened’, computer forensic investigations tend to concentrate on the state of the filesystem including slackspace and virtual memory space for traces of deleted data or indications of the nature of programs previously run on the system (Yasinsac 2001).

From a forensic point of view perhaps the most significant advantage of the computing scene over the real world scene is the computing system’s provision of an event log. The log is an ongoing record of events taking place in the operating system. In addition, since event logs are collected as part of the routine course of system operation they are generally considered ‘direct evidence’ and may be admissible in court (Casey 2000, p.46).

At first the provision of a readily available history of computing activity appears to have the capacity to resolve the problem of reconstructing past events. However, the event logging mechanism of computing systems has proven to be largely unsuitable for forensic purposes and rarely used in litigation.

Evidentially, the weight of the information contained in the event log does not readily conform to the requirements of a forensic investigation. In fact although the weighting criteria of the investigator’s evidence extraction process has been somewhat discussed (Sommer 1998), the weighting of the system’s own evidence extraction facility (event logs) has been relatively left unexplored in scientific research.

To investigate the weighting of evidence from event logs we draw on two key references. Sommer discusses the criteria for the weight for non-testimonial evidence but is aiming his discussion at the traditional forensic investigation where the forensic investigator confiscates digital equipment, preserves the environment, then analyses it and draws conclusion based on the evidence found in the digital environment.

The second reference is an analysis of a case study outlining a number of investigations involving audit trails. Allinson (2003) makes a series of conclusions about the desirable qualities of event logs that are pivotal to forensic investigations. From these two references we identify a set of criteria that can be used to assess the evidential weight of information derived from event logs and at the same time identify the desirable qualities of a forensically suitable event log.

ADMISSIBILITY AND WEIGHT OF EVIDENCE

According to Sommer (1998), evidence in general must satisfy two criteria. The first being ‘admissibility’ and the second being ‘weight’. Sommer suggests that “weight” is a largely non-scientific concept. In fact that weight and admissibility may be easily confused especially in the case of digital evidence. Interestingly, Sommer refers to two cases (not related to the digital environment) - R v Doheny (1996) in the UK and Daubert V Merrel Dow in the US (1993).

In the UK case, the expert presenting DNA evidence neglected to mention the basis of the empirical data and did not inform the jury of the frequency with which matching DNA may be found in the population at large. The expectation being that it was up to the jury to decide whether the DNA evidence was conclusive. The case was subsequently overturned because the expert was considered to have overstepped the boundary by preventing the jury from consciously making the decision to accept the scientific basis or to reject it. In this scenario the jury decides whether to accept or reject the scientific basis of the evidence in effect judging both its admissibility and its compelling nature.

In the US however, the judge remains the gatekeeper of evidence admissibility. Once the judge decides the evidence is admissible, the jury must consider its implications. In this scenario the admissibility is separated from the weight and the jury considers the implication of the evidence only once the judge decides to admit the evidence.

In Daubert vs Merrel Dow a series of tests was applied (termed the Daubert test) after which the actual evidence was presented to the jury.

The Daubert tests are:

- whether the theory or technique can be (and has been) tested
- the error rate associated with the method
- publication in a peer-reviewed journal
- whether the technique has gained widespread acceptance

The two cases illustrate the confusion between admissibility and weighting. The decision to accept the evidence can be considered an admissibility problem or a weighting problem depending on the legal process adopted.

We adopt the English Common Law Model where the scientific basis for the acceptance of evidences in general (and its compelling attributes) is a weighting issue.

In this paper weighting is a qualitative measurement of evidence that is proportional to the ability of the evidence to convince the authority one way or the other (this includes the scientific basis of the evidence as well). In a sense, the weighting of evidence is its compelling nature. We are interested in identifying criteria that can be used to assess the weighting of digital evidence. In particular, evidence contained within host-based system event logs i.e. generated by computers and stored in digital environments.

SOMMER'S CRITERIA

With the traditional forensic investigation process in mind we present Sommer's criteria for the weighting of non-testimonial evidence. Sommer identifies three main attributes – authenticity, accuracy and completeness (Sommer, 1998 cites Miller, 1992):

- (1) Accurate: free from any reasonable doubt about the quality of procedures used to collect the material, analyse the material if that is appropriate and necessary and finally to introduce it into court – and produced by someone who can explain what has been done..
- (2) Complete: tells within its own terms a complete story of particular set of circumstances or events
- (3) Authentic: specifically linked to the circumstances and persons alleged

Sommer expands these attributes for more technical types of evidence and presents five main tests designed to assess the reliability of the evidence derived from digital environments (Sommer 1998).

1. Computer's Correct Working Test

Sommer argues that the computer must be shown to be behaving "correctly" or "normally". In cases where the computer is acting simply as an information store then such a requirement may be easy to satisfy. However if the computer is providing a service such as a database query function and given the investigation is related to precisely that function then it must be tested and shown to be "correct" or "normal".

2. Provenance of Computer Source Test

The evidence collected that is deemed relevant to the investigation must be proven to be taken from the specific computer and from nowhere else.

3. Content/Party Authentication Test

The evidence collected must be relevant i.e. linked to the incident or parties accused in the investigation.

4. Evidence Acquisition Test:

The information evidence must have been gathered accurately, must be free from contamination, and must be complete (note this refers back to the three main attributes of non-testimonial evidence).

5. Continuity of Evidence/Chain of Custody Test

A full account of what happened to the retrieved evidence after it was extracted must be provided. Frequently, all of the individuals involved in the collection and transportation of evidence may be requested to testify in court. Thus, to avoid confusion and to retain complete control of the evidence at all times, the chain of custody should be kept to a minimum (Casey 2000 cites Saferstein, 1998, p 58)

Although these rules relate to digital evidence, they are aimed at the process of evidence extraction from a static system and are not suited to the content of logs that are constantly recording information from a live system. In this sense Sommer's rules only passively govern content. This is not surprising given the roots of digital forensics in traditional media analysis and the evidence extraction process as followed by law enforcement authorities applied on confiscated equipment.

The scope of our discussion is limited to the process of evidence acquisition by the event log within the computing system which corresponds to Sommer's test number 4 and does not include extraneous circumstances like those covered by tests 1 to 3 and 5.

We present the three attributes within the scope of evidence acquisition:

Accuracy: When a forensic investigator states the evidence collected is an accurate copy (in digital terms) s/he means it is identical to the original. Another method of expression is to state the copy is "correct". In this scenario the term "accuracy" implies "correctness" and 'non-contamination'. We can interpret

accuracy as being the extent to which the evidence is correct in relation to the original in the digital environment.

Completeness: When a forensic investigator states the evidence collected is a complete account it is implied that all the relevant evidence from the environment has been preserved (relevant to the subject of the investigation). We can interpret completeness as being the extent to which all the relevant evidence from the digital environment has been collected.

Authenticity: When a forensic investigator states the evidence collected is authentic it is implied that the evidence was collected in a way that the link between the evidence and the environment from where it was obtained remains intact.

Reed tries to explain authentication as:

“Authentication means satisfying the court that (a) the contents of the record have remained unchanged, (b) that the information in the record does in fact originate from its purported source, whether human or machine, and (c) that extraneous information such as the apparent date of the record is accurate. As with paper records, the necessary degree of authentication may be proved through oral and circumstantial evidence, if available, or via technological features in the system or the record” (Sommer 1997).

Authenticity is concerned with a specific circumstance rather than a general one. The above definition further clarifies the intention behind the term by describing authentication as reassurance that the contents of the evidence have not changed (since extraction), that the evidential record does pertain to the implied circumstance and that other information that links the record to the circumstance like the date of the record is in fact correct.

Authenticity also includes an element of reliability (in this sense authenticity is similar to reliability and in fact the two terms have been used interchangeably). Whether the document admitted into evidence is the one purported to be, and whether the document itself is indeed reliable or trustworthy (Sommer 1997)

It is important to separate authenticity from accuracy by noting a digital record may have been preserved accurately and completely but may not be reliable or trustworthy.

Weighting of authenticity relies largely on the forensic investigator’s testimony regarding the circumstances surrounding the extraction of evidence and is addressed by Sommer’s tests 1 to 3 and 5. Event logs can play a constructive role in the testimony of the circumstances surrounding evidence extraction. This aspect is covered in Utility later in this paper.

Testimony to the authenticity of digital evidence is specific to the circumstance in which the evidence is extracted and therefore cannot be directly improved or increased through the use of event logging although certain useful information regarding general circumstances may be captured (see utility). Hence this evidential criterion is outside the scope of the aim of this paper.

We present the remaining two terms within the scope of evidence extraction:

1. Accuracy: The extent to which the evidence collected is correct and not contaminated
2. Completeness: The extent to which all the relevant evidence is collected

EVENT LOGGING: AN INFORMATION COLLECTION SERVICE

Forensic investigations ultimately revolve around the occurrence of a single activity or series of activity. In the case of digital forensic investigations a critical activity may be the sending of an email, the access of unauthorized information, and so forth. Traditional forensic investigations attempt to extract incriminating evidence like the contents of a file or an email. As the investigation almost always begins long after the events occurred there is seldom an opportunity to preserve some of the critical evidence that helps investigators reconstruct the activities that lead up to the incident in question (Ahmad 2002).

However, most computing systems retain event logs. The basic aim of an event log is to record information about system activity. A log of events consists of one or more entries. The aim of a single log entry is to record significant facts regarding an event by preserving descriptive attributes relating to that event in as

much useful detail as possible. A log therefore is a series of entries each containing descriptive attributes about a past event.

Event logs have rarely been perceived as “evidence-collection” mechanisms. In other words event logs typically do not make copies of files or dump parts of memory for use in future investigations. Instead, event logs only extract useful information about system events. In that sense, in terms of information - event logs are extracting information from the system thereby creating evidence and storing it in the event log. Although historically this has not been done for forensic purposes event logs can be configured with this purpose in mind.

In principle, the same three attributes of evidence collected by the forensic investigator can be applied to information collected by an event log. In this sense the event log is an evidence collection mechanism recording information about system activity.

From a forensic point of view, event logs are information collection services that provide an account of the system’s activities and may be admissible in court. As a service that is constantly in the “act of collection”, the evidential criteria of accuracy and completeness apply to the content of event logs

ALLINSON’S CRITERIA

Allinson comments on the analysis of audit trails in relation to several legal cases involving computers. She makes two key observations about evidence regarding the circumstances surrounding these cases. Firstly, that the analysis must determine whether all the relevant events were recorded:

“Audit trail content requires both analysis of what is recorded within each event/activity record and whether or not all event/activity records are recorded in the audit trail..”

because:

“Without full recording of all event/activity it is not possible to state or prove with a degree of certainty that particular actions were or were not performed. Therefore, there is a requirement for a high level of assurance that all activity is recorded.”

And secondly, that the “reliability” of each event must be analyzed in particular:

“The validity, relevance and reliability of each piece of data is the second area of scrutiny.”

Combining the Sommer and Allinson criteria, we present the case for the evidential weight of event log content derived from digital environments along with a re-interpretation of existing terminology in the context of event logs.

Given any event or set of events occurring in a digital environment the criteria for the assessment of the evidential weight of the information content related to that event or set of events that is derived from an event log is:

Process-centric description

- Accuracy: The extent to which the event log mechanism collected the relevant information correctly and did not contaminate it
- Completeness: The extent to which the event log mechanism collected all the relevant information needed to describe each event

Information-centric description

- Accuracy: The extent to which the information description of each event is correct and not contaminated
- Completeness: The extent to which all the relevant information needed to describe each event is collected

EVIDENTIAL ATTRIBUTES IN TERMS OF INFORMATION COLLECTION

There is a fundamental difference however, between the use of the terms accuracy and completeness in the traditional evidence extraction process as opposed to the information collection process. We now examine the use of these terms in their new context to identify any misinterpretations.

Accuracy

Accuracy, in terms of extraction of evidence tests whether the evidence is identical to the original exhibit. In this case it can be assumed that an exact replica would be an “accurate” copy. However, in the case of an event log process, accuracy refers to the event log recording a correct description of a transpired event. Individual operating system types have their own implementations that generate event descriptions. The correctness of the event process by which the description is generated depends on the infrastructure’s design and implementation.

In this sense, an attack on the operating system infrastructure can result in contamination of the evidential description thereby impacting the accuracy of information.

Protection from modification and fabrication whether it be a malicious attack launched by a third party or due to a design flaw in the event log or operating system, these two must be prevented to ensure the accuracy of event log information.

Assuming the event log is secure from modification and fabrication of information, the information in the log entry must correctly describe the transpired event. However, to illustrate a further property other than ‘correctness’ we present an event log entry. It describes a logon event:

Date: 24/06/2004

Time: 11:32

User: John

Type: Success

Source: Security

Event ID: 576

Is the record of the login accurate? The original meaning of the term refers to the attributes being copied from the digital environment correctly. In the information context however we can understand this condition to mean whether the information is correct – i.e. is this a correct description of what happened (if the process of copying is correct then this information is correct). Note that although the record says ‘John’ it is unable to confirm that John was actually the source of the event. Similarly, the time is reliant on the system clock which may or may not be a reliable source. In the case of the time field there appears to be an additional issue – that of precision. In the case of a digital environment several events may occur within the space of a very short interval affecting the investigators understanding of the timing and sequence of events. We may add the test ‘is the timestamp precise’?

Accuracy is the extent to which the information description of each event is correct (and timing information is precise) and has not been wholly or partially modified or fabricated.

Completeness

From our definition, completeness has two properties. Firstly, that all of the events under investigation have been described in the event log. Secondly, that all of the relevant detail regarding each such event is also recorded in the event log.

Allinson is referring to evidence regarding activities under investigation. However in a system when such an activity occurs there are numerous low-level system events taking place that may bear some relation to and/or may explain what happened on the system. Completeness may be difficult to achieve given the large number of events. In addition, each entry describing each low-level event may not necessarily record all of the information needed by the investigator about that specific event.

Most computing systems only record the event that signifies the activity took place. For example a single login activity records a single login event. Event entries on such systems typically record the same (common) attributes for all events like the date/time, event-type, event source, filename and so forth. Other forensically useful information (eg. the full path of the object) may not be included as well.

Completeness is the extent to which all the relevant information needed to describe each relevant event is collected.

Utility

Sommer presented five criteria for the assessment of the weight of evidence preserved by forensic investigators. In essence the forensic investigator acts as an expert witness providing testimonial evidence to the court that the manner in which the investigation was conducted in no way invalidates the conclusions reached. One of these criteria, the Evidence Acquisition Test governs the “act of collection”. Although the forensic investigator will be testifying to the accuracy, authenticity, and completeness of his collection processes, the event log in particular may assist in upholding the burden of proof by describing system activity in a forensically acceptable manner.

We present utility as a desirable attribute of evidence that assists the process of forensic investigation. Although Sommer’s criteria is directed towards the forensic investigator we now investigate the possibility of collecting information that is not necessarily related to system activity but can be used to complement the testimony of the forensic investigator.

The first criteria tests whether the digital environment was behaving “correctly” or “normally”. This coincides with the primary reason why event logs were originally designed – for troubleshooting purposes. Event logs can collect information from the environment that can be used to ascertain whether the system was working “correctly” or “normally”.

The second criteria tests whether the evidence used by the forensic investigation is taken from the digital environment where it claims to be from. The burden of proof here is largely on the forensic investigator to show the traceable link from the extraction of the evidence to the presentation of evidence in court. However, there may be technical solutions to this problem. One such solution may involve the event log, acting as information collector recording information identifying the host environment where it has been collected within its evidential record. Event logs have not been designed to record host identification information in each event entry, adding such a feature would reduce the burden of proof on forensic investigators using evidence from event logs although this is not considered an optimal solution from a performance point of view. The problem of traceability can be solved in the larger context of a network environment where management activities like log centralization can be made responsible for tracing the origins of event information.

The relevance of information to the incident and/or parties cannot be addressed by event logs as their identity is not known at the time of information collection. However, if particular events that may become the subject of forensic investigations can be anticipated then maximizing identification detail of parties using the system and resources like files may help to create a more useful context within which the relationship between the incident or parties accused in the investigation is more apparent.

Continuity of evidence examines what happened to the evidence after it was extracted from the host environment. Although the chain of custody is typically understood to begin after the evidence leaves the digital environment there is an obvious extension to this principle that applies to event logs. Event information is frequently collected in centralized logs and later archived – i.e. information is moved from the point of information collection in the environment to where the description may be ultimately stored on a different host (in a networked environment). This information may be later written to a magnetic tape drive or burnt to a compact disc. Regardless, the chain of custody principle can be extrapolated to include these circumstances.

We present ‘utility’ as a desirable quality of event logs that makes the evidence more compelling:

1. the extent to which the information in the event log testifies to the correct working nature of the relevant system or systems

2. The extent to which the information in the event log identifies the host system in as much detail as possible
3. The extent to which the information in the event log identifies parties and resources in as much detail as possible
4. The extent to which the information in the event log identifies all systems that have hosted the event log or any information contained in the event log

CONCLUSION:

Event logs are essentially information collection mechanisms that make invaluable forensic evidence collection devices. This paper investigates the weighting of evidence from event logs by adapting existing evidential criteria governing the extraction of evidence from computing environments by forensic investigators to event logs. Two criteria – Accuracy and Completeness are identified as the primary attributes of evidence in event logs. A further criterion – utility makes the evidence from event logs more compelling by complementing the testimony of forensic investigators about the circumstances in which the system was operating.

REFERENCES:

Ahmad, A. (2002), The Forensic Chain of Evidence Model: Improving the Process of Evidence Collection in Incident Handling Procedures, Proceedings of the 6th Pacific Asia Conference on Information Systems, Tokyo, Japan.

Allinson, C. (2003), Audit Trails In Evidence: Analysis of a Queensland Case Study, Journal of Information, Law and Technology.

Casey, E. (2000), Digital Evidence and Computer Crime, Academic Press.

Chisum, W.J. & Turvey, B.E. (2000), Evidence Dynamics: Locard's Exchange Principle & Crime Reconstruction, Journal of Behavioral Profiling, Vol. 1, No. 1.

Sommer, P. (1998), Intrusion Detection Systems as Evidence. RAID 98, Louvain-la-Neuve, Belgium.

Sommer, P. (1997), Downloads, Logs and Captures: Evidence from Cyberspace, *Journal of Financial Crime*, October, 1997, 5JFC2 138-152.

Yasinsac, A. (2001), Policies to Enhance Computer and Network Forensics, IEEE Workshop on Information Assurance and Security, West Point, NY.

COPYRIGHT

A.Ahmad, A.B. Ruighaver ©2004. The author/s assign the We-B Centre & Edith Cowan University a non-exclusive license to use this document for personal use provided that the article is used in full and this copyright statement is reproduced. The authors also grant a non-exclusive license to the We-B Centre & ECU to publish this document in full in the Conference Proceedings. Such documents may be published on the World Wide Web, CD-ROM, in printed form, and on mirror sites on the World Wide Web. Any other usage is prohibited without the express permission of the authors