Forensic Tools in an Academic Support Role

Jock Forrester, Senior ICT Specialist
Departments of Computer Science & Information Systems
Rhodes University
Email: j.forrester@ru.ac.za

Barry Irwin, Senior Lecturer, CISSP
Department of Computer Science
Rhodes University
Email: b.irwin@ru.ac.za

Forensic Tools in an Academic Support Role

Abstract

The use of digital technology within the Academic environment has brought about significant changes in the manner in which courses are taught and examined. Complications are bound to arise that may benefit from the use of Digital Forensic methodologies and tools. Digital Forensics is a growing discipline and is regarded as a young science especially when compared with other established forensic sciences.

This paper discusses the use of Digital Forensic methodologies and tools in a support role to Academia and examines two assessment scenarios where forensic tools have proven useful. In addition the paper will also evaluate the effectiveness of the approaches taken and provide recommendations for improvements on the process followed and the underlying environment.

Keywords: Plagiarism, Forensic Tools

1 Introduction

Digital Forensics is the process of recovering evidence from any device that stores data in an electronic format as summarized by Palmer (2002). This includes computers, cell phones, PDA's and even digital cameras.

Computers are used extensively within the Information Sciences not only to conduct teaching, but student assessment as well. Assessment exercises will usually take place under one of two scenarios with regards to the computing facilities setup, either the environment is tightly controlled and in a clean state such as those used for exams or the environment is in a less controlled state such as those used for practical sessions. The importance of being able to accurately assess what has happened in an examination scenario is clearly evident, however with the movement towards continuous assessment as well as the increase in the level of plagiarism it is becoming crucial to be able to automatically detect plagiarism in all assessment scenarios.

The two scenarios examined by this paper are from incidents that the Computer Science and Information Systems Departments at Rhodes University encountered in the latter half of 2004. In each scenario the incident, response to the incident and the outcome will be discussed. There will also be recommendations based on the discussion of the incident.

The first scenario is that of an Information System's III practical exam where the student claimed that the computer they were working on had 'lost' their work when saving all the files to the examination server. The second is that of a series of Computer Science II practical assessments where the students were plagiarising work and submitting it as their own.

The use of forensic tools in the two scenarios recovered the student's exam and identified the students that were plagiarizing. There are however improvements that could be made on the processes used before and after the incident, the types of the tools used and how the tools were used which would increase the effectiveness of the response to the scenarios.

A brief overview of literature related to Digital Forensics and its methodologies and tools will be conducted. The Rhodes University AUP (Acceptable Use Policy) and departmental plagiarism policy will also be briefly discussed.

2 Related Literature

2.1 Digital Forensics

Digital Forensics involves the examination of any type of electronic evidence, this may include data on a CDRom, Stiffy Disk, USB flash drive or a Desktop Computer. Digital Forensics as defined by the Digital Forensics Research Workshop:

> 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. (Digital Forensic Research Workgroup, 2001).

There are many differing views on how to conduct a digital investigation. This can be attributed to the way that digital forensics has evolved as a science as well as to the fact that the goals of digital investigations differ according to the body conducting the investigation: Carrier and Spafford (2003) explain that the goal of a digital investigation for law enforcement agencies is to be able to present sufficient evidence on which a successful prosecution can be based. In contrast Giordano and Maciag (2002) report that a military digital investigation needs to occur in real time in order to detect, defend and counter attack the enemy.

The goal of a digital investigation within a business organisation however may be one of several possibilities. The first is to close the gap in the organisation's security then worry about the repercussions of the incident such as what happened and who did it and who and what did it affect. The second is to gain sufficient evidence for an internal hearing within the organisation, such as a disciplinary hearing. The third is to gain sufficient evidence with the intent to prosecute in a court of law.

Davis, Philipp and Cowen (2005) use the following Digital Investigative model in their book name of book???. Their model has been tested legally and technically in American courts of law. Davis *et al* (2005) list and explain the six phases are as follows:

- Assessment:
  The investigator needs to determine the best course of action by determining the scope of the incident, sources of evidence, implement steps to secure the evidence and establish a chain of custody for the evidence.
- Acquisition:
  In this phase the sources of evidence identified above will be acquired in a legally sound manner.
- Authentication:
  The purpose of the authentication phase is to determine whether the evidence that you have copied is an exact duplicate of the original data.
- Analysis:
  The Analysis phase of the investigation is the time consuming phase of the investigation. In this phase the sources of evidence are reduced to facts.
- Articulation:
  In this phase a report is drawn up on the findings of the investigation.
- Archival:
  After the investigation has been completed, the evidence and the results of the investigation need to be stored in the event that the investigation needs to be readdressed.

2.2 Rhodes University AUP

Every student and staff member of Rhodes University has signed the AUP upon registration or commencement of employment. The AUP "is intended to provide the basis of a social contract, protecting both individual users and the University community, in which the expectation of what constitutes reasonable use of Rhodes University's computing/communication resources can be specified" (Rhodes University, 2001:A).

The AUP forms the general framework for the use of the computing facilities at Rhodes University and the conduct that will be tolerated on those computing facilities. Point 6.3 of the AUP, which is relevant for the first case study, states that "No person shall willfully jeopardise the integrity, performance or reliability of computer equipment, software, data or other stored information. Reasonable care should also be taken to ensure that resource use does not result in a denial of service to others" (Rhodes University, 2001:A).

2.3 Rhodes Policy on Privacy and Network monitoring

The Policy on Privacy and Network Monitoring in effect at Rhodes University applies to all users of any computing facilities at Rhodes University (Rhodes University, 2001:B). It is a subset of the AUP and as such all staff and students have agreed to it.

The Policy (Rhodes University, 2001:B) states that whilst monitoring is not actively conducted on the Rhodes network provision is made for access to log files, to monitor network traffic or any other monitoring as deemed necessary by the IT Director in order to maintain the operational status of the Rhodes network. The AUP makes provision for the same powers afforded to the IT Division to be

given to a localised Technical Section as some departments in the University have their own Technical Staff (Rhodes University, 2001:A).

2.4 Plagiarism Policy

Rhodes University simply defines plagiarism as "Taking and using the ideas, writings, works or inventions of another as if they were one's own" (Rhodes University, 2003). This simple definition applies to the lack of referencing for a written assignment,copying a practical or tutorial and copying in an examination.

The Department of Computer Science has the following penalty structure:

- First minor offence: a mark of zero will be awarded for the entire assignment for all students involved.
- Second offence or major offence: loss of the DP (Duly Performed) certificate for the course (Computer Science, 2006).

Any plagiarism offence is dealt with by the Departmental Plagiarism Committee, there is however recourse to the University Plagiarism Committee for both staff and student.

3 Case study 1: A case of the lost Practical exam Submission

During the Information System's III practical exam at the end of the 2004 academic year, a student claimed that the computer on which they were working had lost all of their work and that she had deleted the original files, contrary to the exam instructions. The same student had been implicated in other incidents during the semester in both the departments of Computer Science and Information Systems and was considered to be academically weak.

When the System Administrator in charge of the exam heard the student say that she had "shift deleted" her work, he became suspicious, as it implied that the student specifically did not want their work to be recovered. Pressing shift and delete does not move files to the recycle bin and makes the files harder to recover.

3.1 Examination Setup

The examination venue was configured such that the client machines could only talk to the examination server and to the System Administrator's computer. The client machines could not communicate with each other.

The client machines are configured in the following manner: There is a single physical hard drive in the machine, but it is split into two partitions. The Windows XP install is done on the C: (the 1st partition) and then the user's data (documents and settings) is configured to reside on the D: (the 2nd partition). In preparation for the exam, the C: was ghosted (Symantec Ghost™ Solution Suite) (imaged) with a new install of Windows XP whilst the D: was only deleted.

Each student had a share on the examination server that only they, along with the System Administrator, could access. The examination server is equipped with two physical drives which were mirrored for redundancy.

To facilitate the ease of data transfers from the client machines to the server, network drives were mapped to the source files on the exam server and to the hand-in directory on the examination server. The System Administrator had also prepared batch files to copy source data to the student's machine and to copy the student's submissions to the server once the exam was complete.

Instructions on how to use these batch files were included in the examination instructions.

3.2 The Incident

At the end of the exam the students were asked to stop working, save all their work and lastly to copy their work as per the exam instructions back to the server using the appropriate script.

The student in question then raised their hand to inform the lecturer in charge that they had made a mistake by moving their files to the server and then deleting the files that they thought were on the client machine. The student had actually deleted their work off the server.

At this point, the System Administrator, Balarin (2005) was called and the student was isolated from the computer that they were working on. The System Administrator then asked the student what had happened, the response was as above and that she had shift deleted the files and not moved them to the recycle bin.

3.3 The Response

Once everyone had left the venue, the System Administrator added their laptop to the trusted network initially setup for the exam and attempted a network undelete which was unsuccessful.

Network enabled disk recovery software was installed to the *C:* drive of the student's machine. No data would have been written to the *D:* of the machine, so there was a minimal chance of any of the deleted files being overwritten. A full install of the disk recovery software was done on the server.

At this point the System Administrator started recording his actions and keeping copies of all logs generated by the disk recovery software. These logs included a list of files recovered, time taken,percentage recovered etc.

A complete scan on the server and client was then performed. The client was scanned from the System Administrator's laptop and a local scan was initiated on the server. Once the scans were complete, the System Administrator decided to focus on the student's machine as approximately 5000 files and 500 folders were recovered, as compared to the server where over 17000 files and 2400 folders were recovered. All the recovered files from the server and the student's computer were then copied to the System Administrator's laptop for further analysis.

At this point the System Administrator was confident that he could recover the deleted files. This was an important conclusion to make, as a decision had to be made whether the student would have to start to rewrite the paper, to delay the rewrite until the next day, or set a new paper.

As the *D:* of the machine was only deleted in preparation for the exam, most of the recovered files were from the practical sessions held earlier in the semester.

The disk recovery software recovered approximately 90% of the files, the other 10% were recovered from the backups that Windows XP and Visual Studio .NET created. The Systems Administrator was able to vary the correctness of the files recovered by the structure of the C# Files and Sample files provided by the examiner and recovered folder locations.

The files were first filtered on MAC (modified, accessed, created) dates. This narrowed down the list of possible files, then the System Administrator took a copy of another student's submission for comparison in terms of folder structure and file size so that he would know what files and folder structure he was looking for. The Practical exam was testing the student's knowledge of C#, and the System Administrator was looking for the source code files which were what were going to be marked.

No Binary files (compiled from the source) were recovered, either due to none being created, or due to the recoverable binary file being destroyed, however a compiled binary file (application) was not necessary for submission.

3.4 The Outcome

A complete project attributable to the student in question was recovered. It was clear from the files recovered and in discussion with the examining lecturer that the student had attempted to complete the exam however it was poor and incomplete work. The Faculty then chose to have the recovered files submitted as the student's work for the examination rather than to allow the student to rewrite the exam or charge the student with cheating or with contravening the Rhodes AUP.

The student could have been charged under point 6.3 of the Rhodes University AUP as the exam files could have been considered to belong to the Information Systems Department, in other words the student could have been charged with destroying data belonging to the Information Systems Department.

3.5 Analysis

The System Administrator worked on copies of the recovered files, so if there were any queries on the results, the original recovered files could have been copied again and the documented process repeated.

Disk images should have been made of the student's full system, her *C:* and *D:* and of the server's *C:*, however due to time constraints, the System Administrator was unable to create disk images. All analysis and extraction of recoverable data should have then been done from these images. Saudi (2001) explains that disk images are made in order to keep the original disk in the same state it was in at the time of the incident.

The full install of the disk recovery software to the server could have over written any recoverable files on the server, however as the System Administrator would have been unable to stop to the rest of the students from copying their submissions to the server and as a result any recoverable files would have more than likely already been destroyed.

MD5 hashes of the recovered files should have been made and the recovered files and their hashes written to a write once media, such as CDR OR DVDR.

The student had to leave the exam venue shortly into the investigation, as supper was being served in the Residences, so there was no proper chain of evidence as the System Administrator was left alone with the evidence. The staff member in charge of the examination should have been present with the System Administrator for the duration of the investigation.

The original environment was altered by introducing the System Administrator's laptop and by installing the disk recovery software on the server and the student's computer. Introducing the System Administrator's laptop would not have affected data on the server, or the student's computer, however the installation of the disk recovery software could have potentially overwritten valuable data on the server.

The Systems Administrator did feel that there was a need to recover live evidence from the system as it appeared to be either a case of "finger trouble" or well guided "finger trouble". In retrospect, live evidence should have been recovered in order to make the investigation more complete. However considering the time constraints imposed on the Systems Administrator and the lack of proper tools and training to perform capture it was probably safer not to attempt the capture as there are risks involved (Carrier, 2006).

The Rhodes University AUP and the Policy on Privacy and Network monitoring granted the Systems Administrator the rights to conduct the investigation and to examine the student's workstation. The workstation is the property of Rhodes University and the stance was taken that the output of the examination was also the property of Rhodes University. The investigation was authorised by the Lecturer in charge of the examination who duly had authority from the Head of Department.

3.6 Recommended Actions

The following could have been done in advance to make the examination environment forensically friendly, however these recommendations do come at the expense of time required to configure the examination server and the client workstations as well as licensing costs.

On the client machines, the *D:* should have been wiped using a disc scrubber so that the partition was in a forensically clean state, in other words no ambient data existed except that created on the day of the exam. This would have aided in the recovery of deleted files.

To take the above one step further, the examination server could be setup to have a system partition and a data partition. The data partition then needs to be wiped with a disc scrubber so that it is in a forensically clean state before the examination begins.

The time of the workstations were not synchronized with each other, or with the exam server. In the future the exam server and workstations time needs to synchronized to allow for easier log correlation and file analysis.

As a proactive measure the Network enabled disk recovery software network agent or the network enabled disk recovery software full version should be installed on the client workstations and the server before the examination begins, this will prevent recoverable files being destroyed during the install process.

In place of, or in addition to, the System Administrator's desktop machine being included in the trusted network, a dedicated incident machine should be included as well. This machine should have all the tools necessary to handle any incidents occurring and have enough disk space for multiple disk images.

The level of logging that the server performs should have been increased to include all file writes, changes and deletions for duration of the examination as suggested by Rowlingson (2004) in his paper on Forensic Readiness.

The Technical Division and the Academic Department currently do not have a policy, set of procedures or even a list of best practices for this type of incident. Rowlingson (2004) explains that a documented set of procedures ratified by legal counsel, or the faculty on how to proceed with digital incidents in an examination will provide for a more solid handling of the situation with less errors and uncertainty.

Since the incident the examination server is now the working directory for the exam files, in other words the files are not copied locally, edited and then copied back to the server. The examination server also takes snapshots every thirty minutes of the exam directory. If a student does "loose" their work, there will a half hour old copy.

4 Case study 2: Duplicate Practical Submissions

The lecturer, Siebörger (2005), of the CS202 Object Orientated Course started noticing duplicate practical submissions being submitted for the course. Students were allowed to work in groups, and the students within a group were allowed to submit the group's solution as their own, however there were duplicate submissions from across these groups.

4.1 The Assessment Environment

The students were being assessed on object orientated modeling techniques. They were using Rational Rose to document their models. The completed practicals were submitted electronically.

4.2 The Incident

There were duplicate hand-ins being submitted from different student groups.

4.3 The Response

The lecturer (Siebörger 2005) concerned made a hash of all the submitted files using the MD5 hash.

The MD5 algorithm is designed to give a unique fingerprint to digital data so that it is nearly impossible that two differing input files have the same fingerprint as explained by the RFC 1312 (1992).

Some of the files all had matching hashes, these were system generated files from the application that the students had used to complete the assignment. There were also some matching hashes within groups, however it was the files that had matching hashes across groups that were then examined further.

Of the files that had matching hashes across groups, copies were made and then examined.

The comments entered by the students were copied and a utility that parses files for like strings was then used to search through the files with matching hashes for identical comments. The same comments were found in files with matching hashes.

4.4 The Outcome

Students, from different groups, had submitted duplicate work and were therefore guilty of plagiarism. The plagiarism was proved without any doubt on two levels. The first being the MD5 hash. Secondly those files with matching hashes were examined for comments made by the students and those comments were then searched for, and found, in the other matching files.

4.5 Analysis

The approach taken by the lecturer was an extremely simple one, without the need for commercial tools, nor a large amount of resources and it proved to be effective.

However if the student had opened the file they had received from their accomplice and close it again, a different MD5 hash would have been produced as Rational Rose records when a file was last opened therefore changes the data within the file. Likewise if a student had copied the practical or edited the comments slightly the plagiarism would not have been detected as the hashing process used by the lecturer would only have identified matching files.

4.6 Recommended Actions

The lecturer should have enabled logging on the server that the practical submissions where being uploaded to. Multiple submissions around the same time from the same machine would warrant suspicion.

Also, the lecturer could have made use of tools that compare files on a line by line or binary basis. These tools would be able to generate what percentage of the file is identical to another. This type of tool would catch students who did open and close the document prior to submission.

5 Conclusion

The use of forensic tools in the above Academic scenarios proved to be effective, however with more detailed planning and preparation the tools would have yielded even better results.

The data recovered for the examination scenario is unlikely to have stood up to scrutiny if the incident was to be brought before legal counsel, primarily due to the lack of a chain of evidence and the fact that data was written to the student's computer and the server post incident potentially overwriting sources of evidence. A documented set of procedures that the Support staff should follow when investigating an examination incident is of paramount importance.

The hashing of practical submissions and then further examining the files with matching hashes would have only caught the laziest students, as if they had opened the file and edited one or two lines, a different hash would have been created for the file. The use of software that compares the files on a binary or line by line basis would have picked up more occurrences of plagiarism.

Supporting the Digital Academic process is a challenging task and requires that support staff draw on fields outside standard Technical Support.

6 References

Balarin, J.I. (2005). ICT Manager, Computer Science and Information Systems, Rhodes University, Private Communication.

Carrier, B. (2006). "Risks of Live Digital Forensic Analysis", *Communications of the ACM*, 49(2), 56-61.

Carrier, B., and Spafford, E.H. (2003). "Getting Physical with the Digital Investigation Process," *International Journal of Digital Evidence*, 2(2).

Computer Science. (2006). Department of Computer Science Handbook. Retrieved April 2006 from http://www.cs.ru.ac.za/courses/Handbook/2006/introduction.pdf.

Davis, C., Philipp, A., and Cowen, D. (2005). Hacking Exposed: Computer Forensics Secrets and Solutions. New York: McGraw Hill / Osborne

Digital Forensic Research Workgroup (DFRWS). (2001). A Road Map for Digital Forensic Research, August 2001, http://www.dfrws.org.

Giordano, J., and Maciag, C. (2002). "Cyber Forensics: A Military Operations Perspective," *International Journal of Digital Evidence*, 1(2).

RFC 1312. (1992). Internet Request for Comment IETF, "RFC 1312: The MD5 Message-Digest Algorithm," retrieved April 2006 from ftp://ftp.rfc-editor.org/in-notes/rfc1321.txt.

Kruse, W.G., and Heiser, J.G. (2003). Computer Forensics: Incident Response Essentials. Boston: Addison-Wesley.

Palmer, G.L. (2001). "Forensic Analysis in the Digital World," *International Journal of Digital Evidence*, 1(1).

Rhodes University. (2001:A). "Acceptable Use Policy For Rhodes University", retrieved April 2006 from http://www.ru.ac.za/intranet/policies/rhodes-aup.html.

Rhodes University. (2001:B). "Rhodes Policy on Privacy and Network monitoring", retrieved April 2006 from http://www.ru.ac.za/intranet/policies/monitor.html.

Rhodes University (2003). "Common Faculty Policy On Plagiarism", retrieved April 2006 from http://www.ru.ac.za/intranet/policies/plagiarism_policy.pdf.

Rowlingson, R. (2004). "A Ten Step Process for Forensic Readiness," *International Journal of Digital Evidence*, 2(3).

Saudi, M.M. (2001) "An Overview of Disk Imaging Tool in Computer Forensics," *SANS Reading Room*, retrieved April 2006 from http://www.sans.org/rr/whitepapers/incident/643.php.

D. R. (Private Communication), 2005.

Siebörger, D.R. (2005). Systems Administrator, IT Division, Rhodes University, Private Communication.

Symantec Ghost™ Solution Suite. Symantec Corporation, retrieved April 2006 from http://www.symantec.com/Products/enterprise?c=prodinfo&refId=865.