A Case Study in Teaching Forensic Computing

Felix C. Freiling
Universität Mannheim
Lehrstuhl für Praktische Informatik 1

Joint work with Thorsten Holz and Martin Mink
Motivation

• Digital investigations are becoming more and more common
• High demand for trained investigators
• No dedicated degree programme in Germany exists (apart from “standard” computer science)
• Apart from offering good practical training, we need to set academic standards (and then raise them)
• Research and education in forensic computing in Germany has a lot of potential
Online Master in Digital Forensics

- Joint project between Albstadt-Sigmaringen University, University of Tübingen, University of Mannheim, Pädagogische Hochschule Thurgau
- 2 years plus Master’s Thesis
- Blended learning: 75% of course taught offline (good also for part-time students)
- Planned to start in 2009/2010
- For more information ask Steve Kovacs or me
Focus of This Talk

- Connect to other (German) researchers, professors and instructors
- Exchange experiences on teaching forensic computing, in particular
  - Experiences in writing investigation reports
  - Experiences in use of tools
Outline

• Overview of courses
• Definition of forensic computing
• First (2007) course: dead analysis
• Second (2008) course: mobile phone analysis
• Lessons learnt
Two Courses

- For students in computer science (“Informatik”) or business informatics (“Wirtschaftsinformatik”)
- **Forensic Computing** (“Forensische Informatik”), Summer Term 2007
  - Lecture with practical exercises
  - 30 students (4th year diploma)
  - Exclusively focused on forensic computing
  - Exercises: Dead (hard disk) analysis and live (honeypot) analysis
- **Hacking Lab** (“Hacker Praktikum”), Summer Term 2008
  - Lab course
  - 13 students (3rd year bachelor)
  - 30% of course on forensic computing
  - Exercises: hard disk analysis and mobile phone analysis
Other Courses in Germany

• Courses specialized on forensic computing:
  – RWTH Aachen (Dr. Dornseif), 2004
  – TU Chemnitz (Prof. Baumgartl), since 2007
  – FH Offenburg (Prof. Hammer), since 2007
  – FH Ingolstadt (Prof. Hahndel), since 2007

• Many other courses on security offer small parts on forensics
Definition of Forensic Computing

• ... discipline to reconstruct the events which lead to a security policy violation in an information system.

• Particularly interesting: Reconstruction based on **technically unavoidable evidence**
  – in contrast to evidence explicitly generated for reconstruction purposes

• Example: Traces of files in slack space of the file system in contrast to log file entries
Forensic Computing and Computer Security

• Goal: give students a research-oriented introduction into forensic computing
  – Not only a tool for the legal system
  – Also a tool for understanding computer security in general

• Understanding security failures is the basis for improved security in the future
2007 Course Overview

• Two lecture hours per week
• 12 weeks of course
• Three extra meetings to hand out and explain practical exercises
• Four invited talks by practitioners
  – Steven Wood (Alste), Andreas Körner (PwC), Andreas Schuster (Telekom), Knut Eckstein (ESA)
• Course material (including videos of lectures) available online
2007 Course Topics

1. **Course overview**: forensic science and digital evidence
2. **Attack patterns and common computer crime; forensic mindset**
3. **Process models for forensic computing**
4. **Hard disk technology, imaging, integrity preservation**
5. **Disk volumes and disk partitions (DOS partition system)**
6. **File system analysis**: Carrier’s reference model
7. **File system analysis**: FAT
8. **File system analysis**: NTFS
9. **File system analysis**: ext2/3
10. **Network, Internet, Application Forensics**
11. **Commercial tools and legal aspects**
12. **Theoretical basis**: Carrier’s hypothesis-based approach
Exercise 1: Live Analysis

- Paused VMware image of a Linux machine compromised in August 2003
- Source: Forensic challenge of the Honeynet project
- Required skill level: “intermediate to advanced”
Exercise 2: Dead Analysis

- Plan: Have students analyse “real” hard disks
- Role playing exercise: students are investigators and should prepare a report for a court case
- Bought about 50 hard disks on e-bay (1€ each)
- Question: Find out as much as possible about the prior owner!
- Students were free to choose tools
“Court Evidence”
Recommended Report Structure

- Following best practices:
  - Formalities: name of investigator, reference, etc.
  - List of evidence (e.g. serial number), documentation of chain of custody
  - Task description
  - Summary of evidence found
  - Details of acquisition process of evidence
  - Summary of used tools
  - Summary of implications of evidence found
  - Appendix: log files, screen shots, etc.
<table>
<thead>
<tr>
<th>#</th>
<th>Manufacturer</th>
<th>Size (MB)</th>
<th>Reports and their size (in pages)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Western Digital</td>
<td>170</td>
<td>A1 (9), A2 (10), A3 (16), A4 (56), A5 (7)</td>
</tr>
<tr>
<td>B</td>
<td>Seagate</td>
<td>545</td>
<td>B1 (52)</td>
</tr>
<tr>
<td>C</td>
<td>Conner</td>
<td>412</td>
<td>C1 (13)</td>
</tr>
<tr>
<td>D</td>
<td>IBM</td>
<td>4330</td>
<td>D1 (19)</td>
</tr>
<tr>
<td>E</td>
<td>IBM</td>
<td>30700</td>
<td>E1 (14), E2 (13)</td>
</tr>
<tr>
<td>F</td>
<td>Conner</td>
<td>210</td>
<td>F1 (39), F2 (18)</td>
</tr>
<tr>
<td>G</td>
<td>Conner</td>
<td>420</td>
<td>G1 (65), G2 (48)</td>
</tr>
<tr>
<td>H</td>
<td>Seagate</td>
<td>545</td>
<td>H1 (14)</td>
</tr>
<tr>
<td>I</td>
<td>Western Digital</td>
<td>325</td>
<td>I1 (186)</td>
</tr>
<tr>
<td>J</td>
<td>Seagate</td>
<td>546</td>
<td>J1 (29)</td>
</tr>
<tr>
<td>K</td>
<td>Seagate</td>
<td>8400</td>
<td>K1 (15)</td>
</tr>
<tr>
<td>L</td>
<td>Fujitsu</td>
<td>1700</td>
<td>L1 (17)</td>
</tr>
<tr>
<td>M</td>
<td>Quantum</td>
<td>170</td>
<td>M1 (211)</td>
</tr>
<tr>
<td>N</td>
<td>Conner</td>
<td>406</td>
<td>N1 (13)</td>
</tr>
<tr>
<td>#</td>
<td>Manufacturer</td>
<td>Size (MB)</td>
<td>Reports and their size (in pages)</td>
</tr>
<tr>
<td>----</td>
<td>------------------</td>
<td>-----------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>A</td>
<td>Western Digital</td>
<td>170</td>
<td>A1 (9), A2 (10), A3 (16), A4 (56), A5 (7)</td>
</tr>
<tr>
<td>B</td>
<td>Seagate</td>
<td>545</td>
<td>B1 (52)</td>
</tr>
<tr>
<td>C</td>
<td>Conner</td>
<td>412</td>
<td>C1 (13)</td>
</tr>
<tr>
<td>D</td>
<td>IBM</td>
<td>4330</td>
<td>D1 (19)</td>
</tr>
<tr>
<td>E</td>
<td>IBM</td>
<td>30700</td>
<td>E1 (14), E2 (13)</td>
</tr>
<tr>
<td>F</td>
<td>Conner</td>
<td>210</td>
<td>F1 (39), F2 (18)</td>
</tr>
<tr>
<td>G</td>
<td>Conner</td>
<td>420</td>
<td>G1 (65), G2 (48)</td>
</tr>
<tr>
<td>H</td>
<td>Seagate</td>
<td>545</td>
<td>H1 (14)</td>
</tr>
<tr>
<td>I</td>
<td>Western Digital</td>
<td>325</td>
<td>I1 (186)</td>
</tr>
<tr>
<td>J</td>
<td>Seagate</td>
<td>546</td>
<td>J1 (29)</td>
</tr>
<tr>
<td>K</td>
<td>Seagate</td>
<td>8400</td>
<td>K1 (15)</td>
</tr>
<tr>
<td>L</td>
<td>Fujitsu</td>
<td>1700</td>
<td>L1 (17)</td>
</tr>
<tr>
<td>M</td>
<td>Quantum</td>
<td>170</td>
<td>M1 (211)</td>
</tr>
<tr>
<td>N</td>
<td>Conner</td>
<td>406</td>
<td>N1 (13)</td>
</tr>
</tbody>
</table>
Highlight From Report M1
Interesting Points

- Students reverted mostly to open source tools like dd, Sleuthkit, foremost
  - Some used evaluation copy of FTK
- Students often used two independent tools to cross-check evidence found
  - Example: partition table extraction via mmls and foremost
2008 Course Overview

- Laboratory course ("Hacker Praktikum")
- Simulation of a CERT ("PCERT")
- Thirteen students formed four CERT teams
- All had to investigate the same incidents
- Incident types (examples):
  - Malicious website analysis
  - Malware binary analysis
  - Dead analysis of floppy and hard disks
  - Mobile phone analysis
- 30% of course devoted to forensic analysis
Mobile Phone Analysis

- Phones are prime sources of digital evidence
- Large portions of flash memory
- Need special hardware (twister box) to access memory
- Bought 10 mobile phones (mostly Nokia) for around 130 €
- 7 phones were analyzed
Interesting Points

• Phones contain standard file systems, but proprietary file formats
• All teams reverted to evaluation version of the commercial analysis tool Cell Phone Analyzer
  – Use a script to defeat random character obfuscation
Nokia 3510i

- Students were able to recover contact lists, dialed and received call numbers, received and sent SMS
- No pictures (no phone had a camera)
- Still a lot of interesting evidence ...
kommst in fünf stunden?
Ich glaube drei stunden. Bin zweimal ... Ich stelle es mir schön vor so lange mit dir schönen zu haben ... Gehen wir mal wieder spazieren? :-)
Das werden wir schatz ... Wie oft hast du ...?
Einmal ... Wann war das und wo?
Wo zu hause? Im bett oder auf dem sofa? ... An was hast du gedacht?
Die nachricht kam nur halb an ...
Auch an das gleiche ...
Wenn du mir sagst ..., das ... Laß dir was einfallen ...
Magst du das wort ...?
Ist okay ... Keine ahnung, du wirst die richtigen finden ... Mag es wenn du mir sagst wie ...
Wenn du magst ... Du mußt anfangen ...
Wenn das so ist ... Äh ...
... Und ihn
Ja sehr sogar ... Magst du es wenn wird?
Laß uns morgen all das machen was du ... Willst du morgen ...
Überall in der wohnung ... oh ja ...
Lessons Learnt: Tools

- **Bias towards open-source tools in lecture**
  - Most students started using Sleuthkit and foremost
  - 6 students then chose to use evaluation versions of FTK, because evidence could be extracted and analyzed “faster”
  - No real evidence to measure this aspect

- **Open-source tools fail to help in specialized settings (like mobile phone analysis)**
  - After first scans using strings and Hex editors, students quickly reverted to (evaluation versions of) commercial tools

- **Programming experience helped** students to circumvent restrictions of these tools
Lessons Learnt: Documentation

- Report structure lead to mostly good results
  - Chain of custody missing in most reports
  - Only half of the students documented their investigation environment
- Participants of second course had mostly followed first course
  - Documentation was much better
- Identified requirement of quality control
  - Documents need to be versioned
  - Authors responsible for parts should be clearly indicated
- Short “executive summary” for non-technical staff at beginning necessary
- Report should follow standard academic practices (like writing a term paper)
Conclusions and Open Questions

- Good evaluation (1.27 out of 6, standard deviation 0.44)
- We will teach course regularly in summer term (aimed at Master’s degree students)

- How “legal” is the acquisition of dead data?
  - Who owns it? What can we do with it?
- Can we create disk images for future exercises that just “look real” but are artificial?
Contact

Prof. Dr. Felix Freiling
Universität Mannheim
Lehrstuhl für Praktische Informatik 1
68131 Mannheim
Germany

https://pi1.informatik.uni-mannheim.de