



Introduction to Computer forensics

Security and Privacy
VSE Prag, 9 - 13.6.2008

Institute for Information Processing and
Microprocessor Technology (FIM)
Johannes Kepler University Linz, Austria

E-Mail: sonntag@fim.uni-linz.ac.at
<http://www.fim.uni-linz.ac.at/staff/sonntag.htm>



- What is computer forensics?
 - When and where is it used?
 - Who may use such techniques?
- Computer forensics vs. encryption
- Computer forensics vs. steganography
- Securing evidence
 - Running systems
 - "Inert" systems
- What information can be obtained in which circumstances?
- Legal aspects:
 - Classifying information to look for according to crimes
 - Admissibility of evidence



What is "Computer Forensics"?

- Computer Forensics (CF) is obtaining digital evidence
 - » Analogue evidence is usually not considered here: Use "ordinary" forensics to gather/evaluate
 - Analogue computers are almost non-existing today!
 - This may come from running systems or parts of them
 - » Hard disks flash drives, PDAs, mobile phones, telephones etc.
 - Can be evidence for computer crimes (computer fraud, hacking, ...) or any other crime (documents with plans for x) or for various other uses
- One indispensable issue is "data integrity"
 - Data is easily changeable:
 - Evidence is **then and only then** usable in proceedings, if it is **ensured**, that it **has not been changed!**



What is "Computer Forensics"?

- Other definitions:

- Analytical techniques to identify, collect, preserve, and examine evidence/information which is magnetically stored or encoded
 - » Problem: "magnetically" → Flash disks, running systems?
 - » Better: "in computerized systems and their parts"
- We define computer forensics as the discipline that combines elements of law and computer science to collect and analyze data from computer systems, networks, wireless communications, and storage devices in a way that is admissible as evidence in a court of law.
 - » Focus on legal proceedings; there are many other uses as well!
 - Note that this is the "highest" form: If evidence is sufficient for criminal proceedings, it can be used for everything else as well!
- A technological, systematic inspection of the computer system and its contents for evidence or supportive evidence of a crime or other computer use that is being inspected.



The basic principles of CF

- No action to secure/collect evidence should affect its integrity
 - It becomes much less worth/completely worthless!
- Examiners should be trained
 - Only investigate as far as your knowledge goes
- All activities should be logged
 - Seizure, examination, storage, and transfer
 - » Complete chain of custody (including its security measures)
 - Documented, preserved, and available for review
 - » Proof for the chain of custody
- Investigations must be accurate and impartial
 - Computer forensic \neq prosecutor/attorney/judge
 - » Describe what was actually found
 - And what should have been found, but was missing!
 - » Describe how reliable these facts are
 - » Describe what conclusions can reasonably be drawn from it



When to use CF?

- To provide digital evidence of specific activity
 - In general, proving non-activity might also be the goal, but this is more difficult and only sometimes possible!
- For legal proceedings
 - Criminal cases: Child pornography, computer fraud, ...
 - Civil cases: Hacking, information theft, industry espionage, ...
- Recovering data
 - (In)advertently deleted information
- Identifying weaknesses
 - After a break in, identify the method employed to prevent it in the future
- Identifying the attack/attacker
 - Verify, whether an incident actually happened and who was responsible for it

When to use CF?

Concrete examples



- Misuse of ICT by employees
 - Unauthorized disclosure of data
 - Internet (WWW, E-Mail, ...) abuse
 - Deleted/damaged information
- Exploiting ICT
 - Industrial espionage
 - Hacking of systems
 - Infiltration (zombie, trojans, viruses, ...)
- Damaging ICT
 - Web page defacements
 - Denial of Service attacks
 - Crashing computers
- Use of ICT
 - Storing data on various (planned) crimes



Who should/may use CF?

- Authorization required for accessing data
 - See privacy laws!
- Live monitoring, hacking, password cracking etc. tools are legally very "dangerous"!
 - Possession alone might be criminal
 - » Good explanation and evidence for its necessity is required!
- Personnel to "do" CF:
 - System administrators in their own area
 - Experts for courts or private investigations
 - Everyone on their own system
 - » Note: A second person (→ e.g. husband/wife) uses the system means, that consent by this person is indispensable!

The sequence of actions in CF (1)



- Secure and isolate
 - Remove all other personnel
 - Keep reliable witness (police, other third persons)
 - » To protect against "The investigator added this data!"
- Record the scene
 - Photograph, write down
 - » Example: Mouse on left or right side? → Left-/Right-handed
 - » How are the systems connected (WLAN!)?
 - » What is the current state (running; screen content; ...)
- Conduct a systematic search for evidence
 - Especially: Notes with passwords, hints for online services used, storage mediums (USB sticks, flash cards etc.)
 - » More "conventional" search, but important
 - » E.g. steganography impossible without programs → Disks, ...
 - Printouts in waste paper basket, ...

The sequence of actions in CF (2)



- Collect and package evidence
 - Keep it safe (no loss/destruction) and secure (no changes)
 - » Secure wrapping; external influences
 - » Especially: Magnetic media and magnet fields
 - Modern harddisks are resilient, but not all media are as safe!
 - » Flash cards, USB sticks, etc.: Static electricity
- Maintain chain of custody
 - Keep log on who has access and restrict this access
- Inspect and evaluate data
 - The main aspect we are going to cover here!
 - Perhaps triage: Immediate brief investigation
 - » What to impound, already some illegal material found → arrest
 - Detailed investigation + report in lab (from copy of media!)
- Present the results
 - In a report and/or before the court



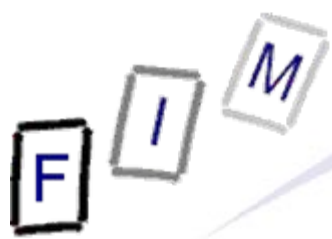
Chain of Custody

- Guaranteeing identity and integrity of the evidence
- Requirements:
 - ❶ Making sure, the piece of evidence introduced is the same as was taken from the suspect/scene of crime/....
 - » Serial numbers → All harddisks/USB/... look exactly the same!
 - ❷ Making sure there was no tampering with it
 - » Witnesses of actions, trust in the person
 - ❸ Making sure of the transition to the next custodian
 - » Who got it next, i.e. when was a chance for tampering
 - Lying around somewhere? Handed to a untrusted person? ...
- Repetition of ❷ and ❸ until the presentation in court
- Note: Digital evidence has a very nice property here: Hash values can nicely prove the "no tampering"!
 - Acquire as early and trustworthy as possible
 - Store it "securely", e.g. on paper with signature of third person



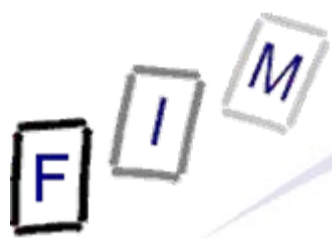
The main problems of CF

- Anything done to a system changes it
 - Especially problematic for running systems
 - Usually not a problem for hard disks
 - » Reading data may change the content microscopically ...
- You can never trust the system under investigation
 - It may be hacked, modified by the owner etc.
- Proving you did **not** change anything is difficult
 - You must be "above suspicion" and take precautions
- The past can never be known
 - We can only find hints what might have possibly been
 - » The content could have been manufactured by someone!
 - » This can be pretty good evidence, but no absolute proof
- Not everyone knows everything
 - Every forensic examination is limited through the examiner!



An increasing problem of CF: Networking & Security

- Today much data is not stored on "the" computer anymore
 - FTP server, bulletin boards, "online harddisks"
 - » Example: RapidShare and similar services
 - Webmail accounts
 - Remote harddisks
 - VPN networks to other systems
- Obtaining a copy of one system is often not enough today!
 - Find traces of the existence of remote information
 - Find traces of the remote information itself
 - » Caches, paging file, file slack, ...
 - Try to access this remote information
 - » By seizure, copying, access over the network, ...
- Encrypted disks are difficult
 - Obtain keys from memory of running system if possible
 - See also TPM (Trusted Platform Module)



The order of volatility

- Registers, caches
- Memory
- Network state (routing configuration, estab. connections)
- Running processes
- Media in use: Disks in use
- Backup media: Disks not in use, tapes
- WOM: CD-ROMs, DVDs
- Analogue material: Paper, fingerprints, DNA, ...

Evidence should be secured/collected in this order !

- Power management (e.g. sleep) can be a great help here
→ Used also normally, so the likelihood of delete-scripts is low)!



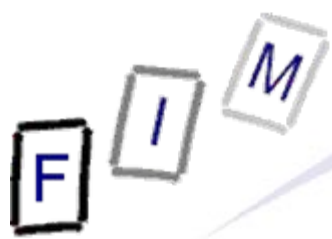
Computer forensics vs. encryption

- CF does work, but doesn't bring **usable** results if the data dis-/recovered is encrypted
 - Depends strongly on the kind of encryption!
- For some programs, decryption software is readily available
 - Especially the integrated encryption of MS Office and Zip!
 - Sometimes based on weaknesses or short keys
 - » But otherwise just brute force attacks: High computing power, special software, and long time may be necessary!
- If **really good** encryption is used, there is almost no chance of decryption without the key (or brute force)
 - One of the reason for the hidden searches: Get at the data before/after it has been en-/decrypted!
 - But: Very often passwords are known words (→ lists!), are written down somewhere, stored somewhere, ...
 - » Important to search the environment for any clues!



Data hiding methods

- Numerous approaches to hide data exist :
 - Through the operating system
 - » Mark as "hidden", "system", ...; use ADS
 - File extension modification: "order.txt" → "cmd.com"
 - RAM slack: End of file → End of sector
 - File slack: End of file → end of cluster
 - Partition slack: End of partition → end of track
 - Disk slack: End of last partition → end of disk
 - Unallocated/bad sectors
 - Delete file / partition; format disk
 - Steganography
- Attention: Several methods are "unstable", i.e. further actions might destroy the data → Using such methods is complex!
- Many approaches require special programs (Hints!)



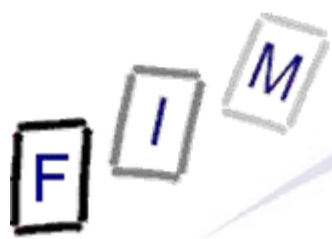
Introduction to Steganography

- Steganography: Hiding messages
 - The intention is, that there is no sign, that data exists at all
- Typical "recipients": graphics, HTML, text, executables
 - Usual problem: Only a small part of the content data can be used for hiding information → Large "cover" for little "content"!
- Usage areas:
 - Where encryption is illegal
 - When the fact of communication itself should be hidden
- Combining encryption and steganography
 - Makes detection through statistics much harder!
- Relation to computer forensics:
 - Hiding data in "inaccessible" places is steganography too
 - Examples: Various slack spaces, alternate data streams
 - » Rather easy to uncover, if presence is known!



Problems of Steganography

- Not very resilient:
 - Data hidden in images is easily destroyed through recoding
 - Text can be reformatted
- Not all base data is suitable:
 - Many files are exactly "known": E.g. OS files cannot be used to hide data within them
 - » See also the problems caused by signed code!
- Complicated to use: Additional tools necessary
 - These can be found on the computer, disks, USB sticks, ...
 - » But need not necessarily be installed!
- Large pieces of seemingly important base material needed
 - This is not always available, or is a hint to hidden data
- Requires a high level of knowledge to be "good"
 - Free tools are available, but these are often easily detected!



CF vs. Steganography

- In practice, Steganography seems to be rather rare
 - There are much easier methods for hidden communication!
 - » E.g. the personal ad columns with certain pre-defined texts
 - » If the text to hide is very long (or multiple pictures, videos), Steganography is still problematic
- Still, looking for hints that it has been applied should be part of every investigation
 - Are there any traces of Steganography programs?
 - Is there suspicious data?
- Brute force attacks, e.g. using steganalysis programs on all images on a computer, are probably less useful
 - Takes very long and it's not very probable to find anything
 - » Mostly, the programs only "support" specific tools!



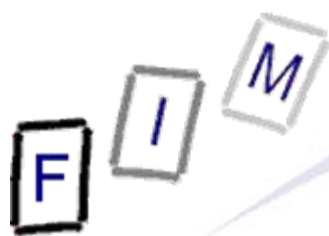
Securing evidence: General considerations

- Evidence must be secured in a "trustworthy" way
 - Nobody should later be able to question the authenticity
- Evidence should be collected as fast as possible, but without destroying anything
 - This might mean, keeping some devices powered, others without power
 - » Keep with power: mobile phones, PDAs, fax machines, ...
 - » Store without power: Flash disks, hard drives, computers
 - Disconnect any communication to/from the device
 - Attention: Not necessarily immediately!
 - » E.g. mobile phones: Shielding (no powering off!)
 - » Computers: Network cables, phone lines, serial lines etc.
 - Check with other forensic experts: Fingerprints
 - » Obtaining traces can damage electronic media!



Securing evidence

- Secure the scene
 - Preserve potential fingerprints, ensure personnel safety
 - Immediately restrict access to computers
 - » Physically; electronically comes next!
 - Document current state (hardware & software)
- Secure the computer as Evidence
 - If the computer is "OFF", do not turn it "ON"
 - » Disconnect all power sources; unplug from wall AND computer
 - » Place evidence tape over each drive slot
 - » Photograph/diagram and label back of components with existing connections
 - » Label all connectors/cable end to allow reassembly as needed
 - » Package components and transport/store components as "fragile"
 - » Keep away from magnets, radio transmitters, heated seats, etc.
- Interview all persons/witnesses



Securing evidence: Online computers (1)

→ If the computer is "ON"

» Stand-alone computer (non-networked)

- Consult computer specialist
- If specialist is not available

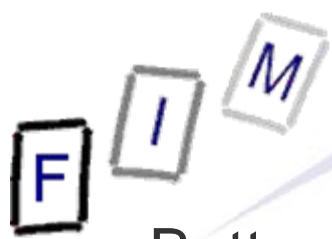
- » Photograph screen
- » Disconnect all power sources; unplug from wall AND computer
- » Continue as with offline computer!

» Networked or business computers / Routers

- Consult a Computer Specialist for further assistance, because pulling the plug could:
 - » Severely damage the system
 - » Disrupt legitimate business
 - » Create officer and department liability

● Please note: Typical procedure for non-experts

→ Experts will (try to) acquire the runtime-state first!



Securing evidence: Online computers (2)

- Better: Obtain as much information from the running system as possible; only then "shutdown" the system
 - General rule: Do not alter the state (On → On, Off → Off)!
- ① Obtain a copy of the complete state
 - Copy of the complete memory
 - » With as little changes as possible!
 - Some additional software MUST be started for transfer!
 - Output of various "state" commands, e.g. running processes, open network connections, open files/shares, ...
- ② Remove power cable from computer
 - » In general, some files might be destroyed, so the computer might not boot anymore. But much less data is lost/changed in this way than when shutting it down!
 - "Delete paging file on shutdown", "Clear privacy data when I close Firefox", ...
 - Not from wall socket: There might be a UPS somewhere!
 - Laptops: Remove accumulator (both if present) as well



Pulling the plug

- Note: Other recommendations are bit more sophisticated
 - Servers: Shutdown
 - » Much data can be destroyed when a file/database/E-Mail server is "killed", which can be a problem for companies
 - Data is lost, computer must be reinstalled/backups restored, ...
 - » Little danger of deletion/modification scripts
 - These might be shut down at any point in time by someone else
 - Workstations: Pull plug
 - » Little damage to be done by killing
 - » Usually full control by a single person → Traps much likelier
 - » Restore much quicker and easier
 - » Affects only a single private person, not a huge company!
 - Appliances: Pull plug
 - » They typically are built to survive this without any damage
 - » The runtime data must be copied before, of course!



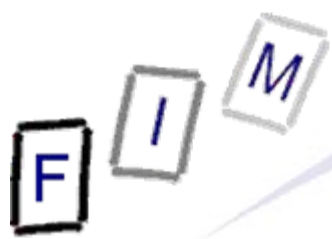
The Heisenberg principle - Analogon

- It is impossible to completely capture an entire running system at any point in time
 - Every kind of "copying the state" will change the state itself!
- The goal to reach:
 - With as little changes as possible
 - Without distortion (like installing additional software)
 - Without bias (like adding hardware/software)
 - » With additional hardware, the data state alone can be captured completely and without its modification!
- Decisions are necessary, what to do (and with that tools!)
 - Generally, try to obtain as much information as possible without changing too much
 - Examples: Display the running processes and photograph the output on the screen
 - » Even better: Use your own (statically linked) program from a CD



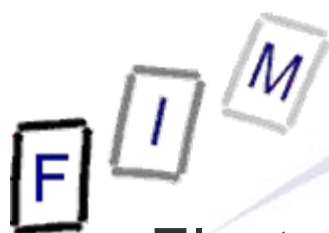
Interviewing personnel/witnesses

- Very important: Encryption, Steganography!
- Information to obtain:
 - Owner
 - User names, passwords
 - » PW: Account, BIOS, E-Mail, configuration, network, ISP, applications, token codes, ...
 - Procedures for access (log in method)
 - E-Mail addresses, online services/applications used, ISP
 - Purpose of the system, persons using it
 - Security schemes (self-destruct systems; e.g. delete scripts)
 - Offsite data: Backups, online replications, ...
 - Documentation of the system: Version numbers
- Note also when information is **not** provided!
 - » Or what turns out to be incorrect
 - Won't help the investigation, but may be important in court



Guiding the search for information

- The aim of the search is most important
 - Is it a search for "something illegal", a specific crime, or whether the image "xyz.jpg" is present on the computer?
 - Uncovering **all** information that is recoverable is possible, but also a lot of work (and therefore very expensive!)
- Assessing the proficiency of the suspect
 - What "hiding" can reasonable be expected?
 - » If unknown, **always** assume the **worst**, i.e. expert techniques!
- When to stop:
 - If something matching has been found or must all/the most of such data be recovered?
 - Monetary considerations (expenses)



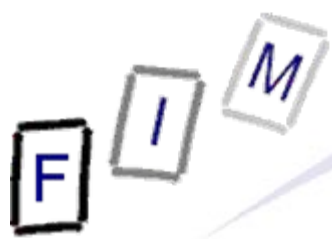
Information according to crimes

- Electronic intrusion
 - Configuration files
 - Executable programs and source code/scripts
 - Open ports, running processes (esp. servers)
 - Logs: Activity, connection, programs, communication, ...
- Fraud
 - Address books, calendars: Physical, E-Mail etc.
 - Images: Cheques, currency, Western Union, signatures, products, ...
 - Credit card data, esp. CVC
 - Office documents: Letters, spreadsheets, databases
 - Banking/accounting software: Dedicated and online
 - Internet activity: Logs, caches, cookies, ...
 - Account information: eBay, banks, ...
 - Communication history: E-Mails, chat logs



Information according to crimes

- Undesirable communication (threats, spam, mobbing)
 - Address information: E-Mail, telephone, ...
 - Documents: Background information, diaries, legal etc.
 - Communication: Letters, E-Mails, SMS, chat logs, ...
 - Internet activity: Cache, logs, cookies
 - Accounts: Online communication facilities
 - Images: Person, products, fakes
 - Software: Mass mailers, text/image/PDF generators
 - Financial information: Accounts, banking
- Violence: Child abuse/pornography, domestic v., death
 - Images, especially hidden ones, and videos
 - Date and time stamps
 - Internet activity: Cache, logs, cookies, access time, searches
 - Software: Communication, photo, P2P
 - Address information and communication: E-Mails, chats, tel.
 - Documents: Legal, medical



Information according to crimes

- Identity theft

- Personal information: Name, address, credit card, ...
- Communication: Especially copies of other person's, obtaining/buying information online
- Software: Generators (names, credit card numbers), imaging (scanner, photo modification)
- Images: Certificates, forms, signatures
- Documents: Forms, letters, orders, ...
- Electronic signatures
- Internet activity: Cache, logs, searches



Information according to crimes

- Copyright

- Software: P2P, CD/DVD-burning, encryption, recoding, key generators, cracks
- Documents: Serial numbers, authorization information
- Internet activity: Cache, logs, searches, cookies
- Images: Covers, license forms
- Communication information: E-Mail, chat
- Accounts: Web-Sites, FTP, shops
- Date and time stamps

Admissibility of evidence (1)



- Digital information is no evidence as such alone
 - Illegal image on disk? How did it come to be there? Unknown!
 - » Was it the accused, someone else with his account, the police, a hacker who broke in over the network, ... ?
 - » Additional information can help if present
 - Physical access to computer, logon-history, encryption etc.
- One very important aspect is the person collecting and interpreting the evidence
 - If this person is trusted, then no later modifications took place
 - When a conclusion is stated as a fact, the person will not be very useful, as judges will not believe them
 - » Fact = Observable
 - Example: Car braking took x meters (measured on asphalt)
 - » Conclusion = Fact + interpretation/general rules
 - Example: Start speed was y km/h because of known friction of tires, weight of car, laws of physics, ...



Admissibility of evidence (2)

- Continental law:
 - Generally all evidence is admissible, regardless how obtained
 - » But what evidence is worth depends on
 - How it was collected and stored
 - By whom it was collected
 - Who analyzed it
 - How it was analyzed
 - Whether the conclusions are supported by facts
 - Whether the conclusions are "state of the art"
 - Typically the judge (or a jury) decides
- Common law:
 - Facts might also be fixed by parties!
 - » If agreed upon, judge/jury cannot discuss it any more
 - Esp. USA: "fruit of the poisonous tree" doctrine
 - » Evidence obtained unlawfully may not be used



Documenting actions

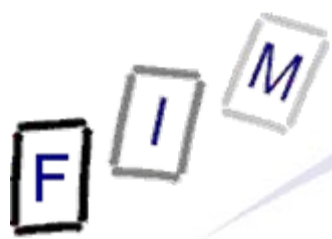
- All actions during an investigation must be documented
 - This starts with acquiring the evidence!
 - » Writing down and photographing when/how the computer was found, which state it was in, etc.
- Running systems: Every single command entered must be documented with the time and the complete results
 - Ideally the log and the result should be stored as a file with a checksum to verify its integrity
- Offline systems:
 - The state must be exactly documented, e.g. checksums over the whole disk
 - Every step of the examination should be documented like in a running system
- Generally: Document also tools (make, version, ...) used!



Final report: General information

- Identity of the examiner
- Identification of the case, e.g. case numbers
- Subject of examination
 - List of and serial numbers of disks/components/...
- Procedural history
 - When was what piece of evidence received, examined, passed on, reported upon, ...
 - Description of the examination: Who did what when
- Results and conclusions
 - Facts (see next slide): What was found
 - Conclusions: What can be derived from that?
 - » This must conform to a very high degree and state assumptions!
 - Example: Time of computer matches "real" time, file access date is 12.12.06 (facts) → File was accessed at that time

» Note: Changing the clock, who used the computer, network connections, ...?



Final report: Content

- Summary of findings (non-technical language!)
- Detailed findings:
 - Specific files matching the search
 - » And other files supporting the findings
 - String searches, keywords searches, and text string searches
 - Internet-evidence: Web traffic analysis, chat logs, cache files, E-Mail, newsgroup activity, ICQ/Skype/... activity
 - Graphic image analysis
 - Ownership status of all files found
 - » Who of the users owned them/when were they created/accessed
 - Techniques used to hide data or limit access to it
 - » Steganography, encryption, hidden attributes/partitions/streams
 - » Incorrect file names (e.g. JPEG files with ".bin" extension)
- Annex: Printouts, digital copies, documentation



Conclusions

- Obtaining some information from hard disks is easy
 - Ensuring it is **complete and usable** in courts is difficult!
 - There is only **a single chance** ...
- A wide variety of hardware exists, which must be treated differently and contains various information
 - Specialization is needed for in-depth investigation
- The huge amount of data on modern computers is a problem
 - Try to reduce the scope of investigation
 - » Lists of "known good" files
 - Automate examination
 - » Keyword searches, deleted file recreation etc.
- Expensive software needed
 - Some investigation also possible with cheaper tools
 - Open source software available partly

F I M

Questions?

Thank you for your attention!



- NIJ Report: Forensic Examination of Digital Evidence: A Guide for Law Enforcement. <http://www.ojp.usdoj/nij>
- NIJ Report: Electronic Crime Scene Investigation: A Guide for First Responders. <http://www.ojp.usdoj/nij>
- dns: An introduction to: Computer Forensics
<http://www.dns.co.uk/NR/rdonlyres/5ED1542B-6AB5-4CCE-838D-D5F3A4494F46/0/ComputerForensics.pdf>