INF 3510 INFORMATION SECURITY

Guest Lecture @UiO on Digital Forensics
April 14 2011

André Årnes, PhD
andre.arnes@hig.no

Who am I?

- Enterprise Security @Telenor and Assoc Professor @HiG
- PhD and MSc from NTNU / UCSB
- Økokrim / Kripos from 2003 to 2008
- All opinions in this presentation are my own and all facts are based on open sources and state-of-the art research.

Objectives

- What is digital forensics and investigations
- What are the central principles and processes
- Real world examples
- Not a ”computer forensics” course
- Partially based on the book ”Forensic Discovery”

3

Forgery?

http://www.dagbladet.no/kultur/2007/10/30/516705.html

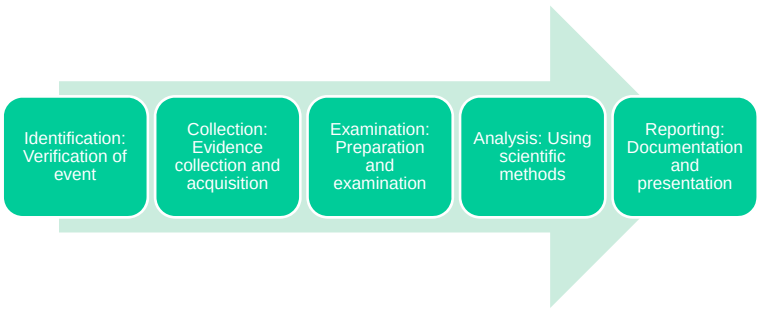
4

Digital Investigations

Central Principles and Definitions

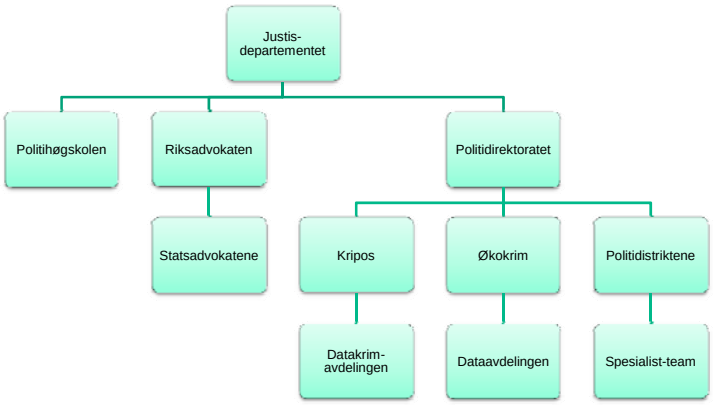
5

Investigation Process



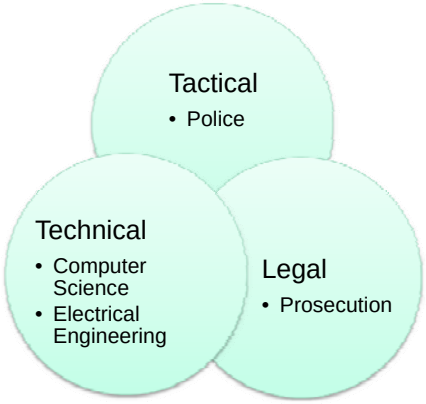
6

Digital Forensics in Law Enforcement



7

Core Competencies of Digital Investigations



8

Digital Forensics

Central Principles and Definitions

9

Forensic Science

- The application of science and technology to investigate and establish facts of interest to criminal or civil courts of law. For example:
  - DNA analysis
  - Trace evidence analysis
  - Firearms ballistics
- Implies the use of scientific methodology to collect and analyse evidence. For example:
  - Statistics
  - Logical reasoning
  - Experiments

10

Digital Evidence

- We define digital evidence as any digital data that contains reliable information that supports or refutes a hypothesis about an incident.
- Evidence dynamics is described to be any influence that changes, relocates, obscures, or obliterates evidence, regardless of intent.

11

Some Terminology

- Digital Forensics
- Computer Forensics
- Network Forensics
- Digital Investigations
- Internet Investigations
- Computational Forensics

12

Evidence Integrity

- Evidence integrity refers to the preservation of the evidence in its original form. This is a requirement that is valid both for the original evidence and the image.
- Write-blockers ensure that the evidence is not accidentally or intentionally changed
  - Hardware
  - Software
- In some cases, evidence has to be changed during acquisition, see discussion of OOV below.

13

Digital Fingerprints

- Purpose is to prove that evidence and image are identical – using cryptographic hash algorithms
- Input is a bit stream (e.g., file/partition/disk) and output is a unique hash (file signature)
- We use cryptographic hash algorithms (e.g., MD5, SHA1, SHA256). These are *non-reversible* and it is *mathematically infeasible* to find two different files that create the same hash.

14

Chain of Custody

- Chain of custody refers to the documentation of evidence acquisition, control, analysis and disposition of physical and electronic evidence.
- The documentation can include laboratory information management systems (LIMS), paper trails, notebooks, photographies, etc.
- Mechanisms:
  - Timestamps and hash values
  - Checklists and notes
  - Reports

15

Order of Volatility (OOV)

- Collect the most volatile data first – this increases the possibility to capture data about the incident in question.
- BUT: As you capture data in one part of the computer, you're changing data in another
- The Heisenberg Principle of data gathering and system analysis: It's not simply difficult to gather all the information on a computer, it is essentially impossible.

16

Dual-tool Verification

- Verification of analysis results by independently performing analysis on two or more distinct forensic tools.
- The purpose of this principle is to identify human and software errors in order to assure repeatability of results.
- The tools should ideally be produced by different organizations/ programmers.

17

Forensic Soundness

- The term *forensically sound* methods and tools usually refers to the fact that the methods and tools adhere to best practice and legal requirements.
- A typical interpretation:
  - Source data is not altered in any way
  - Every bit is copied, incl. empty and unavailable space
  - No data is added to the image.

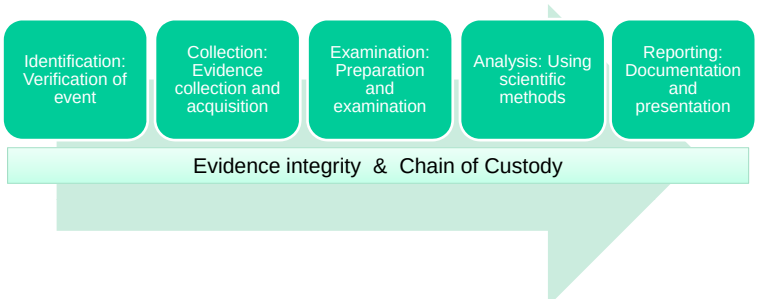
18

ACPO Principles (ACPO p. 6)

1. No action taken by law enforcement agencies or their agents should change data held on a computer or storage media which may subsequently be relied upon in court.
2. In exceptional circumstances, where a person finds it necessary to access original data held on a computer or on storage media, that person must be competent to do so and to be able to give evidence explaining the relevance and the implications of their actions.
3. An audit trail or other record of all processes applied to computer based electronic evidence should be created and preserved. An independent third party should be able to examine those processes and achieve the same results.
4. The person in charge of the investigation has overall responsibility for ensuring that the law and these principles are adhered to.

19

Investigation Process



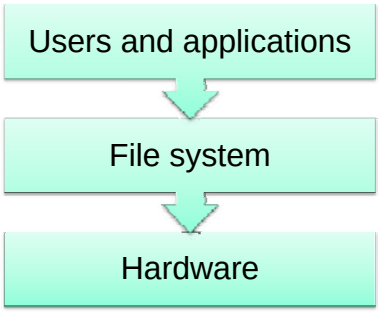
20

Analysing Digital Evidence

21

Analysis and Abstraction Layers

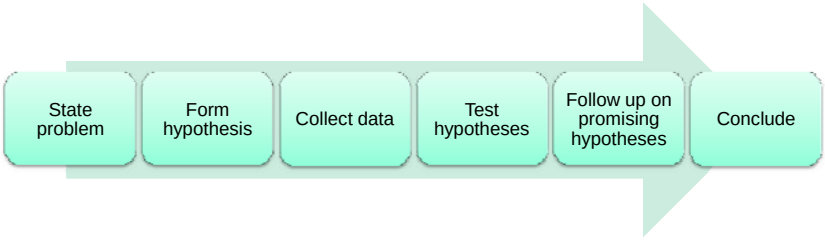
- Unusual activity stands out, e.g.:
  - Location in file system
  - Timestamps (most files are rarely used)
- Fossilization of deleted data
- Turing test of computer forensic analysis
- Digital archaeology vs. geology



22

Crime Scene Reconstructions

Crime Scene Reconstructions is a method to determine the most probable hypothesis or sequence of events by applying the scientific method to interpret the events that surround the commission of a crime. The hypothesis can be tested using statistical or logical reasoning or through experiments.



23

Automated Analysis

- Automated analysis may be implemented using scripting in popular tools, but this is still to some degree an open research problem.
- Automated analysis and reporting can provide increased efficiency and reduces risk of mistakes.
- However, automated analysis can not substitute a human analyst -- an experienced analyst can find important evidence in ways that cannot be formalized as an algorithm.

24

Case Analysis

- Case analysis incorporates both digital, physical and tactical evidence.
- Findings from multiple sources of evidence and information can be managed in a spreadsheet or database.
- Purpose of analysis is to find new links and connections in evidence.
- Data can be visualized to present case for third parties and in court.

25

Reporting

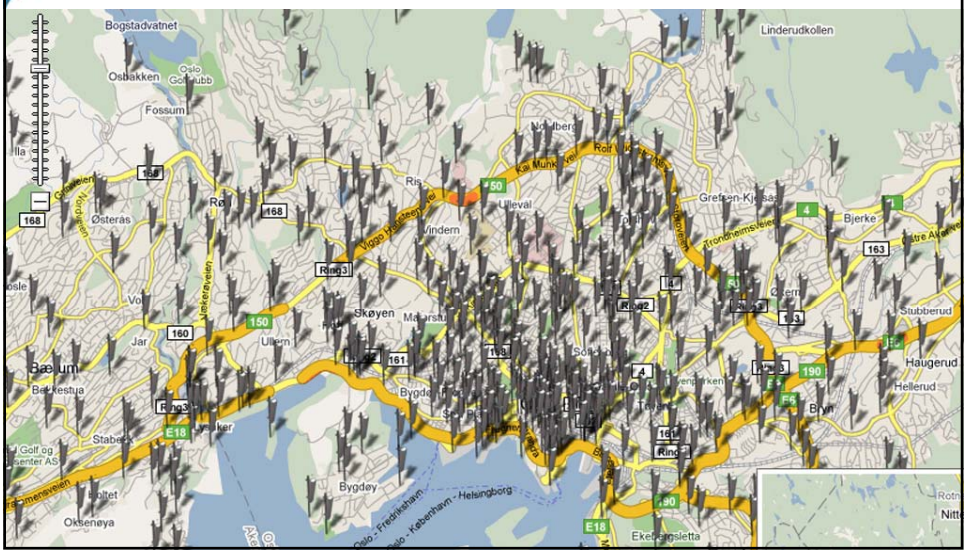
- Chain of custody and evidence integrity
- Document the task given by superviser
- Give a summary for easy access to information
- Document all steps and results for repeatability
- Third parties should be able to repeat all steps in the report and achieve the same results

26

Testimony

- A testimony in court is based on your own observations regarding evidence
- An expert witness can be challenged on the integrity of the evidence and the soundness of the conclusions

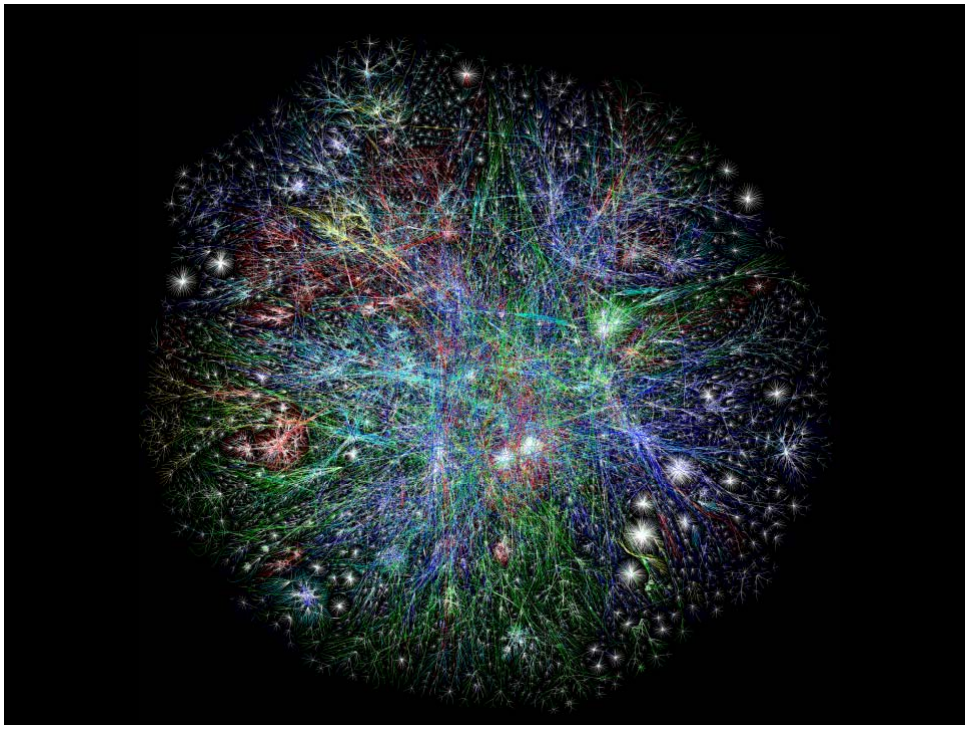
27



Internet Investigations

Tracing and Evidence Acquisition

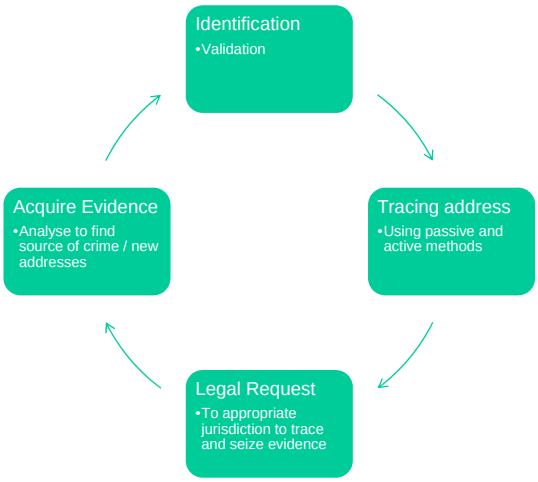
29



The Internet

- A network of networks
- Built to provide reliable and robust connectivity
- No inherent security and traceability
- No inherent user identification
- No inherent accountability or logging

Internet Investigations



Police Cooperation

- Typical Requests (described in Cyber Crime Convention)
  - Identify subscriber information for an IP address
  - Search and seize digital evidence
  - Real-time collection of traffic data
  - Real-time collection of content data
- Prerequisites for effective enforcement
  - Harmonized legal framework
  - Resources to enforce
  - Fast and effective cooperation

Police Cooperation -- Example

- A long lasting investigation of a botnet involved in online bank fraud has finally reached a new step – recent investigation shows that the botnet has been controlled by the IP-address `234.23.34.4` on `June 1st 2010 21:00 CET`.
- Please assist with the following:
  1. Identify subscriber information
  2. Search and seize computer equipment
  3. Perform real-time collection of traffic data prior to search

Police Cooperation - Framework

- Arenas of cooperation
  - Interpol
  - Europol/ Eurojust
  - G8 Subgroup on High Tech Crime
  - Bilateral

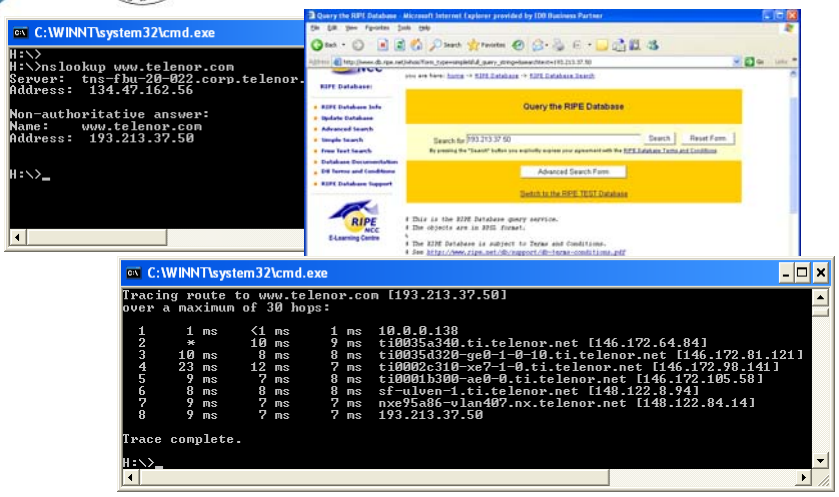
Technical Tracing

Passive Methods

- The use of third party sources to get information about address
- Examples:
  - IP whois (IP and BGP information)
  - DNS whois
  - DNS lookup
  - Reverse DNS lookup

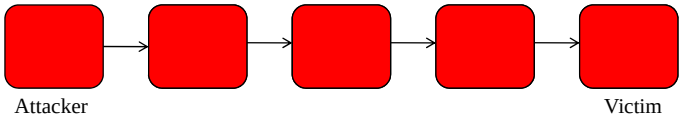
Active Methods

- Connecting to the target host or network to gain further information
- May impact or compromize investigation
- Examples:
  - Ping, traceroute and portscan
  - Connecting to a website
  - Participating in P2P network

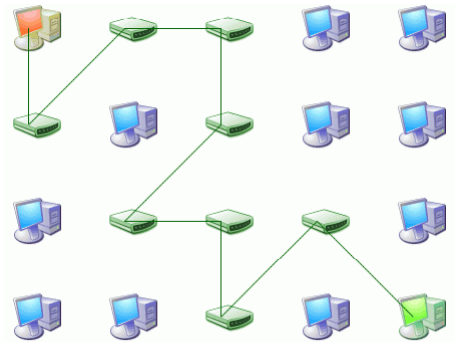


Uncertainties in Tracing (1/2)

- Stepping Stones
  - The perpetrator can use computers in multiple networks to hide his own location. A successful trace will involve multiple jurisdictions.
- Anonymization networks and network tunneling
  - The perpetrator can access the Internet through an anonymization network (e.g., Thor) or encrypted tunnel to hide his own location.
- Network architecture
  - Network architecture elements such as private networks (NAT) and virtual private networks (VPN) can hide the actual address of the perpetrator behind a public address.
- Registration issues
  - The public registers on the Internet (WHOIS registers) may contain incorrect information and it is possible to manipuate WHOIS entries, creating difficulties for tracing attempts.



Uncertainties in Tracing (2/2)



Thor

- Anynymization network
- Encrypts messages
- Randomized hop sequence

Online Evidence Acquisition

- Online evidence should be handled as any other evidence, i.e., by ensuring evidence integrity and chain of custody.
- There are few tools available for this purpose – the investigator must be sufficiently competent to maintain a chain of custody and be able to prove that evidence integrity is preserved

Types of Evidence

- Client data (email, Internet history, malware)
- Domain name and IP addresses
- Network monitoring, intrusion detection, and log data
- Internet content (web, social networks, etc)
- Multimedia streaming data on the Internet
- Online email and calendar accounts
- Online cache (e.g., Google)
- Online archives (e.g., www.archive.org)

Uncertainties and Evidentiary Value

- Who is at the keyboard
  - It can be very hard to prove who was physically using a computer at a particular time
- Uncertainties of origin
  - There are many ways to hide your identity on the Internet, and addresses change over time
- Timestamp inaccuracies
  - There are no standard means of synchronizing and storing timestamps
- Transient nature
  - Evidence changes over time

”The Trojan Did It!”

- Could the perpetrator be a third party with access to the suspect computer using a Trojan?
- UK 2002 arrest in child pornography case
  - Analysis identified 11 trojan horse programs on computer
  - Case acquitted
- UK 2001 DDoS attack on US site
  - No traces of malware detected during analysis
  - Case acquitted due to possibility of trojan

43

Investigating Complex Cases

Intelligent processing and analysis

44

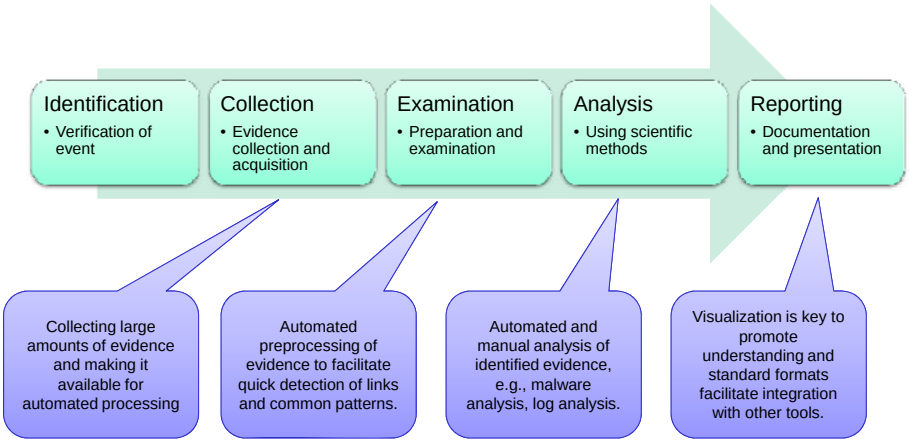
The Challenge

- Scattered evidence across jurisdictions
  - Need to coordinate and synchronize law enforcement across multiple jurisdictions
  - Cooperative efforts from several nations is necessary.
- Large and complex networks of evidence
  - Massive amounts of data
  - Heterogeneous evidence types and format
- No *a priori* knowledge about evidence
  - Relationship between devices not known
  - Access to only subset of potential evidence

One Case – Multiple Sources of Evidence

- Internet
  - E.g., web, social networks, email
- Computers
  - E.g., malware, peer to peer, logs
- Mobile phones
  - E.g., malware, logs, sms, email
- Physical evidence
  - E.g., fingerprints, trace evidence
- Telecommunications and bank transactions

Intelligent Processing and Analysis



47

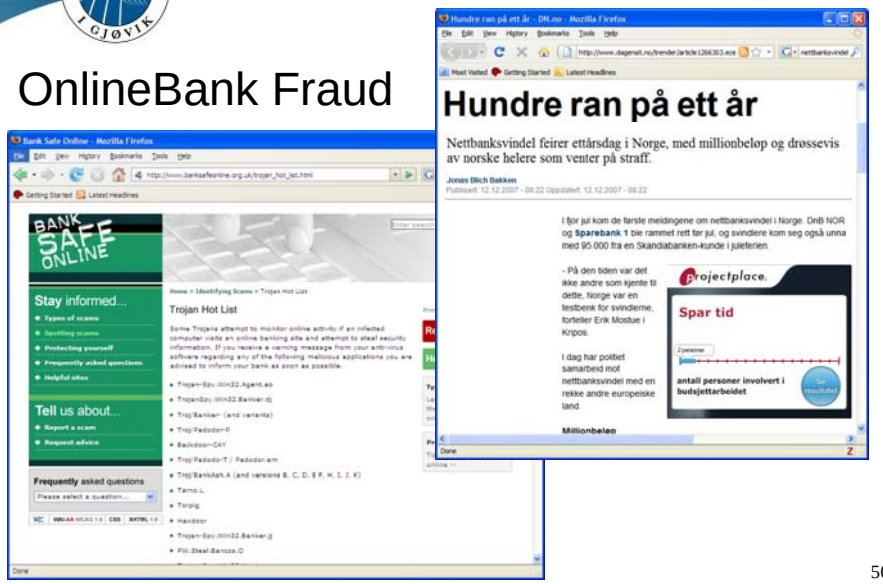
Analysis – Tools and Methods

- Link analysis and data mining
  - Establishing relationships between devices and events
- Timelining physical and logical events
  - Understanding the order of events
- Event based reconstruction
  - Understanding causal relationships based on a hypothesis
- Automated search and file matching
  - Search for known text strings or files

Case: Online Bank Fraud

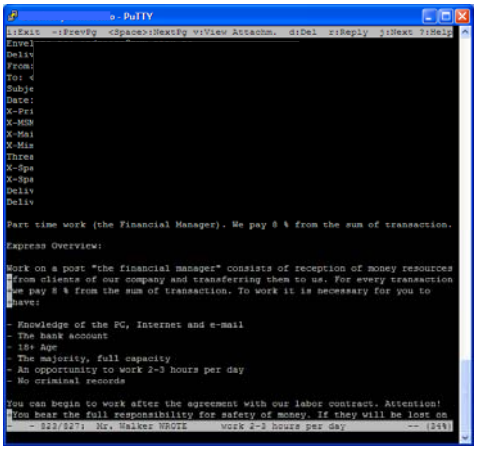
- Online bank fraud involves multiple parties and leaves evidence in many places:
  - Bank transactions from victim
  - Malware on victim host and botnet evidence
  - Server side logs at bank
  - Communication with mule (email and phone)
  - Transactions from mule
  - Network monitoring logs

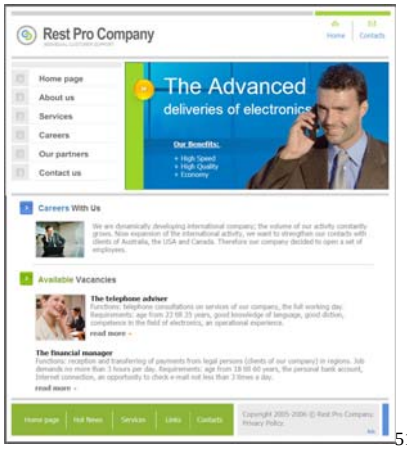
OnlineBank Fraud



50

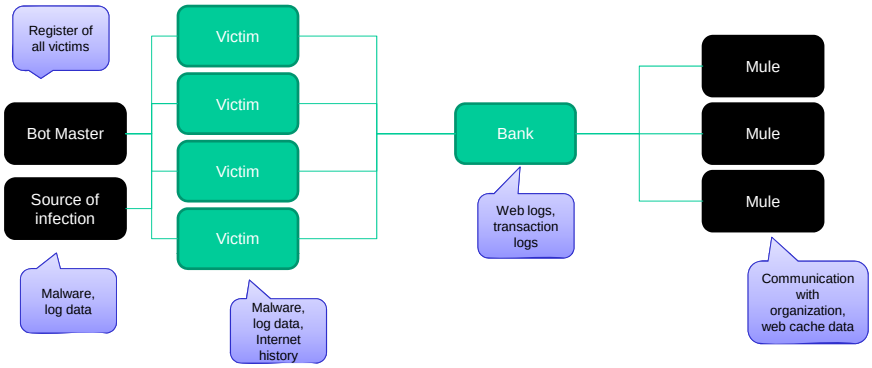
Transaction Agents





51

Evidence Overview



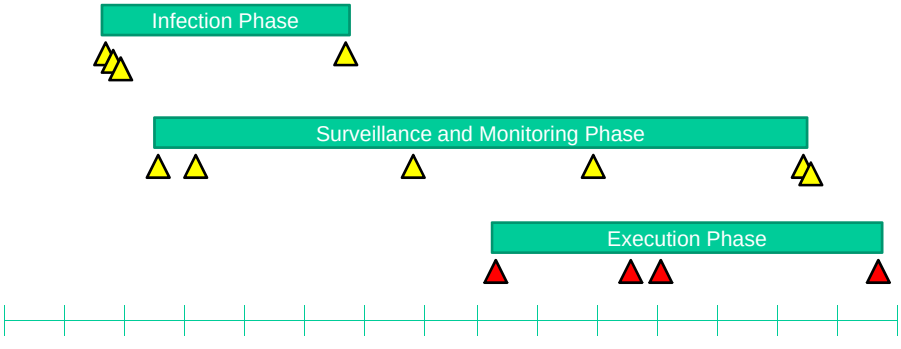
Link Analysis

Found in Victim A
- IP address 11.11.11.11
- DNS my.owned.com
- Signature string: PwNd
- Malware detected: Trojan/BadNews.B

Found in Victim B
- IP address 11.11.11.12
- DNS my.bruteforce.com
- Signature string: PwNd
- Malware detected: Trojan/BadNews.B

Found in Victim C
- IP address 11.11.11.11
- DNS my.bruteforce.com
- Signature string: 1337 H4X0R
- Malware detected: Trojan/BadNews.A

Timeline Analysis



Event Types

Based on digital evidence from seized computers

Based on information provided by bank

Hacking-as-a-service

- ***"For the price of 3,000 dollars, our reporter was offered his personal bank Trojan. In an interview with Computer Sweden, the hacker behind the recent Internet frauds against Sweden's Nordea bank claims responsibility for more intrusions."***
[http://computersweden.idg.se/2.2683/1.93344]

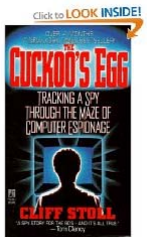
55

Collateral

Some Useful References

1. Brian Carrier, ”File System Forensic Analysis”, Addison Wesley, 2005
2. Keith J. Jones, Richard Bejtlich, Curtis W. Rose, ”Real Digital Forensics – Computer Security and Incident Response”, Addison Wesley, 2006
3. Inger Marie Sunde, ”Lov og rett i Cyberspace”, Fagbokforlaget, 2006
4. US DOJ, ”NIJ Special Report on Forensic Examination of Digital Evidence: A Guide for Law Enforcement”
5. ACPO, ”Good Practice Guide for Computer Based Electronic Evidence”
6. The Honeynet Project; in particular Scan of the month and forensic challenges
7. DOJ, ”NIJ Special Report on Investigations Involving the Internet and Computer Networks” (pages 1-27, excluding ”legal considerations”)

”Cuckoos Egg”



- Sysadmin Cliff Stoll at Lawrence Berkeley National Labs: ”The Cuckoo’s Egg: Tracking a Spy Through the Maze of Computer Espionage”, 1990.
- In 1986 US$ .75 led to detection of computer intrusions and ten months trying to track down the attacker, using session printouts and honeypots.
- Attacker targeted military systems and was looking for password files and documents including terms ”nuclear” and ”SDI”.