INTRODUCTION

- Cyber Investigation Evolves
- Defining Cyber Investigation
- Distinguishing Between Cyber Forensics & Cyber Investigation
- DFRWS Framework Classes

A Note on Etymology (added by Kabay)

fo-ren-sic [fô rén'sik, fô rén'zık] adjective
1. crime-solving: relating to application of science to decide questions arising from crime or litigation; forensic evidence
2. of debating: relating to debate & formal argumentation; forensic oratory

[Mid-17th century. < Latin forensis “of legal proceedings” < forum “forum” (as a place for discussion)]

Microsoft® Encarta® 2008. © 1993-2007 Microsoft Corporation. All rights reserved.

Defining Cyber Investigation (1)

- Rogers, Brinson & Robinson establish cyber forensics as an ontology [on tōlājē] (plural on-tol-ō-gies). noun
  - 1. study of existence: most general branch of metaphysics, concerned with nature of being
  - 2. theory of existence: a particular theory of being

[Early 18th century. < modern Latin, “study of being” < Greek ont- “being” (see onto-)]

Microsoft® Encarta® 2008. © 1993-2007 Microsoft Corporation. All rights reserved.

- Stephenson’s ontology focuses on defining unique aspects of computer-related crime that can be studied

Topics

- Introduction
- End-to-End Digital Investigation
- Applying Framework & EEDI
- Using EEDI & Framework
- Motive, Means, & Opportunity: Profiling Attackers
- Some Useful Tools
- Concluding Remarks
### Defining Cyber Investigation (2)

- Cyber investigation relies on **taxonomy** (tax·on·o·my) [tak suhn-uh-mee] (plural tax·on·o·mies) noun
  1. grouping of organisms: science of classifying plants, animals, & microorganisms into increasingly broader categories based on shared features. Traditionally, organisms were grouped by physical resemblances, but in recent times other criteria such as genetic matching have also been used.
  2. principles of classification: practice or principles of classification
  3. study of classification: study of rules & practice of classifying living organisms

(From 19th century. < French taxonomie < Greek taxis (see taxis))

Microsoft® Encarta® 2008. © 1993-2007 Microsoft Corporation. All rights reserved.

---

### Rogers’ Taxonomy (2)

- Two major classes
  - Profession – structure of human endeavors
  - Technology – subjects of investigation
- Benefits
  - Supports understanding of concepts
  - Each additional sub-category supports more detail in analysis
  - Framework encourages thorough attention to details
  - Can serve as a checklist to avoid overlooking evidence
  - Supports analysis of cyber crime

---

### DFRWS Framework Classes

- Digital Forensics Research Workshop (2001)
- Framework for digital investigation
- Supports end-to-end digital investigation (EEDI)
- Each class comprises elements

---

### DFRWS Class: Identification

- How investigator is notified of potential incident
  - half of reports of possible security breaches turn out not to be crimes
- Framework classes in Identification
  - Event/crime detection: direct evidence (e.g., discovery of unauthorized access)
  - Resolve signature: intrusion detection/prevention systems, gateway security devices using pattern recognition
  - Profile detection: heuristic pattern recognition; attack scenarios, attack profiles
  - Anomalous detection: deviation from observed norms
  - Complaints: person reports event or results of event
  - System monitoring: situational awareness processes
- Audit analysis: analysis of log files
DFRWS Class: Preservation
- Management of evidence ensuring integrity
- Case mgmt: notes, process controls, quality controls, procedural issues
- Imaging tech: making bit-for-bit image copies of evidence
- Chain of custody: preventing unauthorized access to & modification of evidence – preservers evidentiary value
- Time synchronization (normalization):
  - Ensuring that all time records use a common base time
  - No evidence modified
  - Determine offsets from a baseline (e.g., "- 0:00:07.6 GMT-5" for 7.6 seconds behind GMT-5)

DFRWS Class: Collection (1)
- Approved methods:
  - General acceptance by courts
  - E.g., qualifying under Daubert rule for admission of technical information – see CSH5 Ch 73
  - Or qualified under current case law
- Approved software: source code identical to that of tool that has qualified in courts (see above)
- Approved hardware: same principles as above
- Legal authority: policy (e.g., for owner of equipment), subpoena, warrant
- Lossless compression: provable fidelity

DFRWS Class: Collection (2)
- Sampling: demonstrated validity & safety for data
- Data reduction:
  - Valid, repeatable, provable results
  - Applied only to copies of evidence
- Recovery techniques
  - Extraction of useful data from data repositories
  - Comply with all court-permitted techniques
  - Forensic investigators must keep up to date with current case law

DFRWS Class: Examination (1)
- Traceability or chain of evidence
  - Clear documentation of reasoning linking evidence to other evidence (not conclusions)
  - Traceability & continuity of chain of evidence crucial to credibility of conclusions
  - Distinct from chain of custody!
- Validation techniques
  - Corroboration
  - May involve demonstration of internal consistency
  - Resistance to claims that evidence has been modified or fabricated

DFRWS Class: Examination (2)
- Filtering techniques
  - Sometimes source filtering (e.g., IDS) eliminates some data in stream
    - Must supply courts with evidence of techniques used
    - Demonstrate validity of remaining records
  - Also refers to extraction of relevant data types (e.g., images) from data
    - May include comparison using hashes
    - All such tools & techniques must be understood by investigator / examiner
    - Understanding includes clear grasp of appropriate usage & a reasonable grasp of underlying principles (see Daubert Rule)

DFRWS Class: Examination (3)
- Pattern matching
  - Finding potential events by matching signatures & other patterns
  - E.g., Intrusion-detection & anti-malware systems
- Hidden data discovery
  - Deleted but recoverable
  - Stored outside a file system's control (e.g., slack space)
  - Encryption
  - Steganography
- Hidden data extraction
  - Getting reliable data from sources described above
DFRWS: Analysis

- “Fusion, correlation & assimilation of material for reasoned conclusions.”
- Tying together evidence into coherent & probably correct scenario of events
- Ideally use accepted standards for processes of deduction & induction
  - Deduction: reaching a conclusion by applying rules of logic
  - Induction: forming a generalization based on observed evidence

DFRWS: Presentation

- Reporting facts with organization, clarity, conciseness, & objectivity
  - Organization: using a comprehensible structure
  - Clarity: unambiguous, easily understood communication
  - Conciseness: using fewest words possible to supply necessary information
  - Objectivity: free from bias, not trying to convince anyone of a particular interpretation
- See CSH5 Ch 73 for recommendations on being an expert witness in court

END-TO-END DIGITAL INVESTIGATION
1. Collecting Evidence
2. Analysis of Individual Events
3. Preliminary Correlation
4. Event Normalizing
5. Event Deconfliction
6. Second-Level Correlation
7. Timeline Analysis
8. Chain of Evidence Construction
9. Corroboration

Collecting Evidence

- Approved tools & techniques
- Trained technicians
- Time sensitive
- Events must be considered in context of prior, concurrent & following events
  - Events are most granular element of incident
  - Incidents are collection of events that lead or could lead to a compromise
  - Incident becomes a crime when laws are broken
- Critical data collection includes
  - Images of affected computers
  - Logs of intermediate devices (esp. Internet)
  - Logs of affected computers
  - Logs & data from intrusion detection systems, firewalls etc.

Analysis of Individual Events

- Events may leave records in multiple places
- Analysis assesses value of events to investigation
- Tie events into each other
- Aim to understand incident
  - Put events into coherent narrative

Preliminary Correlation

- Correlation distinguishes among
  - Evidence that stands alone (unique events)
  - Evidence recorded in different ways & located in different places
  - Evidence that supports other information located elsewhere
- Corroboration supports formulation of chain of evidence
  - Consistent description of incident
  - Time sequences are called timelines
  - Causal sequences impute causes & effects
Event Normalizing

- Combine evidentiary data from multiple sources
- Eliminate duplications to ensure each unique event is correctly represented once in timeline & causal sequence

Event Deconfliction

- Some events have multiple repetitions of identical or near-identical steps
  - E.g., denial-of-service attacks may have 1000s of similar or identical packets flooding perimeter
  - These may be defined as subevents
- If reasonable, may define multiple subevents
  - E.g., probes
  - That occur in a defined time period
  - E.g., 48 seconds
  - As a single event
    - E.g., “Denial-of-service”

Second-Level Correlation

- Normalization & deconfliction should support creation of a coherent picture of events
- Second-level correlation of remaining data establishes a basis for building chains of evidence

Timeline Analysis

- Use normalized, deconflicted data to create a sequence (timeline) of events
- Expect to update constantly
- Iterative process
  - Event analysis
  - Correlation
  - Deconfliction
  - Timeline analysis

Chain of Evidence Construction

- Ideally
  - Each link in chain supported by evidence
  - Leads to next link
- In reality
  - Often gaps in chain
  - Must infer links
    - Not evidence: a lead
    - May point to legitimate evidence
  - May also corroborate missing or dubious link
    - If all corroboration points to same link, may be acceptable

Corroboration

- Match every element of chain of evidence
  - With other, independent evidence
  - Using correlated & uncorrelated data
- Best evidence
  - Developed using digital methods
  - Corroborated using traditional investigative methods
- Final evidence chain
  - Digital & traditional evidence
- Similar process in investigation vs postmortem analysis
**APPLYING THE FRAMEWORK & EEDI**

- Overview
- Supporting EEDI Process
- Investigative Narrative
- Intrusion Process
- Describing Attacks
- Strategic Campaigns

**Supporting EEDI Process**

- Traditional investigators often resist process
- Prof Stephenson’s research finds practice conforms to his recommendations
- Thus DFRWS Framework & EEDI can serve traditional investigators entering world of cyber investigation
- Provide guidance on sequence of actions in investigation

**Investigative Narrative**

- Investigator’s detailed notes
- EEDI supports construction of investigation using framework(s)
- DFRWS Framework helps focus attention on all elements of situation
- E.g., DFRWS Collection class refers to authorized/approved methods
  - Therefore must be careful to use accepted, standard software, hardware & methods
  - Basis is case law – acceptance by courts

**Intrusion Process**

- Details of specific attacks vary – increasingly *blended*
- But in general, attacks include
  - *Information gathering*: research, locating IP addresses, superficial scans
  - *Footprinting*: scanning IP addresses for visible devices
  - *Enumerating*: probes/scans to document operating systems & other details of exposed systems
  - *Probing for weaknesses*: vulnerability scans or social-engineering attacks
  - *Penetration*: obtaining unauthorized access
  - *Backdoors, Trojans, rootkits*: payload deposited for immediate or later exploitation
  - *Cleanup*: wiping tools, altering logs, generally covering tracks

**Describing Attacks (1)**

- Various attack taxonomies available
  - But no generally accepted language
- Howard’s Taxonomy (CSH5 Ch 8)
  - Simple, concise
  - Good starting point
Describing Attacks (2)

- **Description of attack**: events, targets, vulnerabilities
- **Type of attack**: exploit, denial-of-service, reconnaissance
- **Attack mechanism**: how accomplished
- **Correlations**: comparison with other attacks, current attacks
- **Evidence of active targeting**: generic or specific
- **Severity** = Target Criticality + Attack Lethality – System countermeasures
  - Rough guesses
  - Usually lowest 1 to 5 highest
  - Heuristic purposes only – not analytical or rigorous

Describing Attacks (3)

**Informal template for early interviews**

1. Nature of incident?
2. How to be sure there really was an incident?
3. What was/were entry point(s) to system?
4. What kind of evidence are we looking for in this context?
5. What monitoring systems may have collected evidence?
6. What legal issues are relevant?
7. Who could have caused or allowed incident?
8. What security was in place at time?
9. What nontechnical (business) issues may have affected attack?
10. Who knew about attack – & when?

Strategic Campaigns (1)

- Attack may be isolated
- But may be a tactic in a larger strategy; e.g.,
  - Spam
  - Identity theft
  - Hacktivism
  - Cyber war
- Differences between tactical attack & strategic campaign
  1. Single objective vs ongoing objectives
  2. Low-hanging fruit vs sustained efforts
  3. Trivial vs complicated targets & objectives

Strategic Campaigns (2)

**Distinct phases**

1. Mapping & battle space preparation
2. Offensive & defensive planning
3. Initial execution
4. Probes & skirmishes
5. Adjustment & sustainment
6. Success & termination

MOTIVE, MEANS, & OPPORTUNITY: PROFILING ATTACKERS (1)

Threat agents deliver a threat
- What are benefits of attack? (motive)
- Does agent have capability for attack? (means)
  - Access
  - Inhibitors vs amplifiers affect planning
- When is best time to attack? (opportunity)
  - Catalysts are variable factors that affect decision
Motive (1)

- Understanding motive may help
  - Understand/analyze attack
  - Narrow down field of possible attackers
  - Identical attacks may have different motives
- Outcomes may differ significantly
  - Seeking revenge: embarrass victim
  - Seeking profit: extort money from victim
- Groups may behave differently from individuals

Motive (2)

Adversarial matrix can help refine picture of motives

<table>
<thead>
<tr>
<th>States</th>
<th>Cyber mercenaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theft of intellectual property</td>
<td>Personal gain</td>
</tr>
<tr>
<td>Disruption</td>
<td>Gangs, new-age gangs</td>
</tr>
<tr>
<td>Organized crime and terrorists</td>
<td>Financial gain</td>
</tr>
<tr>
<td>Money laundering</td>
<td>Lone hackers and hacker &quot;clubs&quot;</td>
</tr>
<tr>
<td>Theft of trade secrets for resale</td>
<td>Peer reputation and power</td>
</tr>
<tr>
<td>Disruption</td>
<td>Occasionally knowledgeable</td>
</tr>
<tr>
<td>Competitors</td>
<td>Disgruntled employee</td>
</tr>
<tr>
<td></td>
<td>Revenge</td>
</tr>
<tr>
<td></td>
<td>Personal gain</td>
</tr>
</tbody>
</table>

Motive (3)

Jones' motivation taxonomy:

- Political
- Secular
- Crime
- Personal gain
- Revenge
- Financial
- Knowledge/information

Means (1)

- Tools & techniques used in attack
- Relate means to skill of attacker
  - Potential divergence between sophistication of attack tools & competence of attacker
  - Script-kiddies classic example
- More productive to exclude suspects who cannot be attacker

Means (2)

<table>
<thead>
<tr>
<th>Category of Offenders</th>
<th>Motivation</th>
<th>Personal Characteristics</th>
<th>Potential Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups</td>
<td>Intellectual challenge</td>
<td>Highly intelligent individuals, lack of expertise in support of a cause</td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>Intellectual challenge</td>
<td>Moderately to highly intelligent individuals, lack of expertise in support of a cause</td>
<td></td>
</tr>
</tbody>
</table>

Episodic:

- Money & a chance to attack the network
- Problem solving, ability to solve problems
- Hacking as a hobby
- Criminals become greedy for more information and then become careless
- Become greedy and lose motivation

Means (3)

<table>
<thead>
<tr>
<th>Category of Offenders</th>
<th>Training Skills</th>
<th>Equipment Needed</th>
<th>Support Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups</td>
<td>High level of technical training</td>
<td>Basic computer equipment</td>
<td>Peer group support</td>
</tr>
<tr>
<td>Individuals</td>
<td>Expertise gained through experience</td>
<td>Advanced computer equipment</td>
<td>Peer group support</td>
</tr>
</tbody>
</table>

Experiential:

- Experience-based expertise
- Computer equipment with nodes, more experience
- Computer equipment with nodes, more experience
- Support may come from sponsoring intelligence agency

Means (4)

<table>
<thead>
<tr>
<th>Category of Offenders</th>
<th>Training Skills</th>
<th>Equipment Needed</th>
<th>Support Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups</td>
<td>Basic computer</td>
<td>Basic computer</td>
<td>Peer group support</td>
</tr>
<tr>
<td>Individuals</td>
<td>Expertise gained through experience</td>
<td>equipment with nodes</td>
<td>Peer group support</td>
</tr>
</tbody>
</table>

Experiential:

- Some programming experience
- Computer equipment with nodes, more experience
- Support may come from sponsoring intelligence agency

Means (5)

<table>
<thead>
<tr>
<th>Category of Offenders</th>
<th>Training Skills</th>
<th>Equipment Needed</th>
<th>Support Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups</td>
<td>Basic computer</td>
<td>Basic computer</td>
<td>Peer group support</td>
</tr>
<tr>
<td>Individuals</td>
<td>Expertise gained through experience</td>
<td>equipment with nodes</td>
<td>Peer group support</td>
</tr>
</tbody>
</table>

Experiential:

- Some programming experience
- Computer equipment with nodes, more experience
- Support may come from sponsoring intelligence agency
Opportunity

➤ Opportunity helps determine if suspect is credible perpetrator
➤ Includes knowledge of victim system
➤ Insiders or confederates of insiders should be examined
➤ External groups may be involved
  ❯ E.g., Anonymous or LulzSec

SOME USEFUL TOOLS

➤ The Usual Toolkit
➤ Link Analysis
➤ Attack-Tree Analysis
➤ Modeling
➤ Statistical Analysis

The Usual Toolkit

▪ Computer forensic imaging and analysis
▪ Network forensic/log aggregation and analysis
▪ Malware discovery
▪ Media imaging (without analysis)
▪ Network discovery
▪ Remote (over-the-network) computer forensic analysis and imaging

➤ Well known & accepted
➤ See product evaluations; e.g., in SC Magazine
  ❯ May 2011 edition in particular
  ❯ http://www.scmagazine.com/lets-go-analyze-something/article/200541/ or
  ❯ http://tinyurl.com/bunu4aB

Link Analysis (1)

➤ Link analysis immensely useful
  ❯ Analyze large data sets
  ❯ Find non-obvious relationships
  ❯ Applied to fraud, drugs, terrorism, organized crime
➤ Core theory
  ❯ Pairs of related items; e.g.,
    ・ People/address
    ・ Source/destination IP addresses
    ・ Alias/realname
  ❯ Pairs can lead to further linkage

Link Analysis (2)

➤ Example: linking data about cyber attacks
  ❯ Hacker alias / realname
  ❯ Alias / group
  ❯ Alias / attack
  ❯ Group / attack
➤ Clusters
  ❯ Group of entities bound more closely to each other by links than to surrounding entities
  ❯ Cluster analysis simplifies link maps

Link Analysis (3)
Link Analysis (4)

Attacker Tree Analysis (1)

- Method for analyzing possible attack scenarios
  - Define goal as root
  - Hypothesize attack method as leaves
  - Look at probabilities of scenarios
  - Eliminate impossible sequences

Attacker Tree Analysis (2)

- Can assign any Boolean (logical) value to nodes; e.g.,
  - Easy/difficult
  - Legal/illegal
  - Special equipment req’d/not req’d
- Even quantitative variables can be assigned; e.g., cost

Attacker Tree Analysis (3)


Attacker Tree Analysis (4)

- "Attack trees provide a formal methodology for analyzing security of systems & subsystems. They provide a way to think about security, to capture & reuse expertise about security, & to respond to changes in security. Security is not a product -- it’s a process. Attack trees form basis of understanding that process."

Modeling: CPN (1)

- Simulating attack behavior
- Coloured Petri Nets (CPN) useful language
  - Invented by K. Jensen at Aarhus University in Denmark
  - Transferred to Eindhoven University of Technology, The Netherlands (2010)
  - Good overview at <http://cs.au.dk/CPnets/>
Modeling: CPN (2)

- Graphical language -- constructing models of concurrent systems & analyzing properties
- Foundation of graphical notation & basic primitives for modeling concurrency, communication, & synchronization
- Standard ML -- definition data types, describing data manipulation, & creating compact models
- Typical application domains: communication protocols, data networks, distributed algorithms, embedded systems, business processes, workflows, manufacturing systems, & multi-agent systems
- Simulation-based performance analysis -- delays, throughput, & queue lengths in system are investigated

http://cs.au.dk/CPnets/

Modeling: CPN (3)

Statistical Analysis

- Statistical methods & probability analysis of great value
- Look for anomalies -- events with low probability if not related to crime & high probability if related
- Calculate probabilities of sequences of events; e.g., if faced with n events, each with probability $p_i$,
  - Probability that all events would occur simultaneously or in sequence by chance alone: $P(\text{all}) = p_1 \times \cdots \times p_n$ for identical $p_i$
  - Probability that at none of events would occur simultaneously or in sequence by chance alone: $P(\text{none}) = (1 - p_1) \times \cdots \times (1 - p_n)$ for identical $p_i$
  - Probability that at least one of events would occur simultaneously or in sequence by chance alone: $P(\geq 1) = 1 - (1 - p_1) \times \cdots \times (1 - p_n)$ for identical $p_i$

DISCUSSION