Models of Models: Digital Forensics and Domain-Specific Languages

Daniel A. Ray and Phillip G. Bradford
The University of Alabama
Tuscaloosa, AL
DanielRay@cs.ua.edu, pgb@cs.ua.edu

Outline

- Summary
- Motivation
  - Proactive Forensics
  - Sequential Statistics
- Models for Digital Forensics
  - Different from Classical Forensics
  - Some Digital Forensics Models
  - Leverage Computer Science
- Domain Specific Languages
- Conclusions

Summary

- Modeling the investigative process
  - Different investigation processes for different incidents
    - Classical forensics: different tools and procedures for different incidents
    - Digital forensics: different tools and procedures for different incidents
- Final objective: make the criminal case obvious to a lay-person
  - Depends on the method and procedure of the model
    - A failure on evidence gathering may damage or destroy the case

Motivation: Classical & Digital Forensics

- Computer Security is often *preventative*
  - Focus on preventative measures
    - IDS--anomaly detection may be proactive
- Classical Forensics is *reactive*
  - Post-mortem
- Digital forensics is *reactive*
  - A lot of focus on file recovery from disks
  - Generally reactive
  - Digital Forensics has opportunity to be proactive
- Proactive Forensics!
  - Online Monitoring stakeholders…

Motivation: Proactive Computer-System Forensics

- System structuring and augmentation for
  - Automated data discovery
  - Lead formation
  - Efficient data preservation
- Make these issues proactive
  - How?
- Challenges
  - System resources
  - Exposure
    - Double edged sword…

Proactive Computer-System Forensics

- What data should we capture?
  - Different crimes may require different investigative procedures
    - Static: when and where illicit data was placed on a disk
    - Dynamic: what system states do we document when there is an intrusion?
      - What is being written to logs or disks? Which programs are being run? Where is the smoking-gun?
  - Depending on the nature of our online investigation, we may need to secure evidence in several different models

Crime Types

- Computer *Assisted* Crimes
  - Computers provide basic help in criminal activity
- Computer *Enabled* crimes
  - Computers are a Primary focus on criminal activity
- Focus:
  - Dynamic: computer enabled crimes
    - Range from viruses to spam to sophisticated attacks
  - Static: Computer Assisted Crimes
    - Stolen data, spreadsheets to compute illicit gains, etc.

Variations on Digital Equipment and Software

- Mobility & wireless
  - Cell phones, PDAs, Laptops, etc.
- Enterprise Level Systems
  - Database systems, dynamic Internet sites, large proprietary systems,
- Distributed systems
  - Virtual private networks, network file systems, user mobility, distributed computation, etc.

Gathering Statistics for Proactive Forensics

- Running sequential statistical procedures
  - What data to save?
    - The data we need may change as things progress
      - Proactive not reactive
  - How much data do we save?
  - How costly?

The DFRWS Model

http://www.dfrws.org/2001/dfrws-rm-final.pdf

IDENTIFICATION	PRESERVATION	COLLECTION	EXAMINATION	ANALYSIS	PRESENTATION
Event/Crime Detection	**Case Management**	**Preservation**	**Preservation**	**Preservation**	Documentation
Resolve Signature	Imaging Technologies	**Approved Methods**	**Traceability**	**Traceability**	Expert Testimony
Profile Detection	**Chain of Custody**	Approved Software	Validation Techniques	Statistical	Clarification
Anomalous Detection	**Time Synchronization**	Approved Hardware	Filtering Techniques	Protocols	Mission Impact Statement
Complaints		**Legal Authority**	Pattern Matching	Data Mining	Recommended Countermeasure
System Monitoring		Lossless Compression	Hidden Data Discovery	Timeline	Statistical Interpretation
Audit Analysis		Sampling	Hidden Data Extraction	**Link**	
		Data Reduction		Spatial	
		Recover Techniques			

Figure 1: DFRWS Digital Investigation Framework [29]

Ciardhuain Model
by S. O. Ciardhuain

- Extends DRFWS Model by working on information flows
- Class-based model
  - Authorization activity
  - Planning activity
  - Notification activity
  - Hypothesis activity
  - etc.
- An augmented “waterfall model”
  - supports iterative backtracking between consecutive activities
  - models information flows
  - Feedback critique

Mobile Forensics Platform (MFP)
by F. Adelstein

- To remotely perform early investigations into mobile incidents
- Analyze a live running (mobile) machine
- Maintains original evidence which is verifiable by a cryptographic hash
- Connect to same LAN as the suspect machine

DSLs

- DSLs are, ". . . languages tailored to a specific application domain" Mernik, Heering, and Sloane
- Most Digital Forensics Models
  - Have a good deal in common
    - Evidence verification and storage
    - Flow of investigation
- Pulling together data storage, data modeling and authentication-verification
  - Combining other DSLs: XML, UML, DB Blobs, etc.

DSLs

- May be fairly complex to build a single DSL
  - However, worth investigating
- Must be a very trusted language
  - Numerous cases may depend on the trust-level of the language
- Move from "best practices" to more formal "programming patterns for digital forensics"

Conclusions

- Digital forensics is complex
  - Digital Forensics Models are complex
    - Static and Dynamic
- There may be a need to automatically choose from a diversity of digital forensics models
  - A programming language