

*Making ~~Forensic Attack~~
Event/forensic Analysis as Simple
as Possible and No Simpler*

Sean Peisert, UC Davis

Given at Schloss Dagstuhl — July 22, 2008

Electronic Voting Machines

- Need to be able to count ballots
- Need to be able to determine if and how a machine failed.
- Cannot allow a voter to indicate to an auditor who they are (vote selling)
- Cannot allow an auditor to determine who a voter is (voter coercion)
- This leads to a direct conflict. So how do we balance this?
 - Add noise
 - Enforce regularity

Existing Technical Solutions and the Insider Problem

- Access Control
- Intrusion Detection
 - Anomaly-Based Detection
 - Misuse-Based Detection
 - Signature-Based Detection

Optimistic Access Control

- Security and **usability** are in conflict.
- Ideally, a system should block all forbidden actions and permit all allowed actions. (**This is not feasible.**)
- Policies can be **binary** (block access) or **flexible** (perform this countermeasure).
- Policies can be **static** (always do this) or **dynamic** (uh oh—an intruder)
- Many possible countermeasures exist
 - log
 - checkpoint/replay
 - make a particular partition read-only
- Many possible dynamic approaches exist
 - Use an a standard IPS
 - Incorporate external factors

So we need to focus on
non-binary (e.g., *post
mortem* analysis).

What is forensic Analysis?

- forensic analysis is the process of answering the questions:
 - **How** did an event take place?
 - **What** was the nature of the event?
 - What were the **effects** of the event?
- forensic analysis applies to arbitrary events. This can include attacks, but is not limited to attacks (e.g., **mistakes**).
- forensic analysis is **not** intrusion detection.
 - The goal of intrusion detection is to determine whether an attack occurred.

Transparent Society

(abbreviated from David Brin's ideas)

- Anyone can know anything.
- There is no privacy.
- It's better if everyone knows everything than if a few people know everything.
- “Watching the watchers”
- R. Heinlein: “privacy laws’ only make the bugs smaller.”

Audit trails are...

- Is it is not well understood what forensic data is necessary, and there is no general solution to find that data.
- Data is often **redundant, missing, vague, or misleading**.
- **Forensic analysis is worthless with bad data.**
- We're wasting time, drawing bad conclusions, and making bad decisions.
- **We need better data.**
- A **systematic approach** to forensic **logging** gives better data and **better analysis**.

Current State

- Decent tools, but **what problem do they solve?**
 - file & filesystem analysis (Coroner's Toolkit, Sleuth Kit, EnCase, FTK)
 - syslog, tcpwrappers, Windows event logs
 - BSM
 - process accounting logs
 - IDS logs
 - packet sniffing

Forensics

- What do we need?
- What are we missing?

What are the assumptions for using current forensic tools?

- Often that there's only one person who had access to the machine.
- Often that the owner of the machine was in complete control (as opposed to malware).
- Probably a lot of other assumptions that we have no clue about.

For forensics, we need to...

- go back to the beginning.
- understand what the purpose of the analysis is
- understand what data can answer that purpose, with $X\%$ accuracy, and under a set of Y assumptions
- log the data
- give tools and techniques to an analyst to analyze that data

Art & Science

- But computer science can only answer part of it.
- Forensic analysis is an art, but there *are* scientific components. What are they?
 - Determining what to log
 - Determining relevance of logged data
 - what is relevant?
 - what is not relevant?
 - under what circumstances something might be relevant?
 - Using the results to constrain and correlate data.
 - *This can be measured, systematized and automated.*

Logging

- Two options:
 - Log everything (e.g., all non-deterministic events), and capture upon replay
 - Log selectively
 - *Ad hoc*
 - Systematic (e.g., based on security policies)

A Systematic Approach is Better

- Given system S , that records data D , what intrusions I_D can we understand with the data we have?
- Given intrusions I' , what additional data $D_{I'}$ do we need to record to analyze those intrusions?
- Given an arbitrary system defined by certain specifications, what information must be logged to detect violations of those specifications?

Laocoön

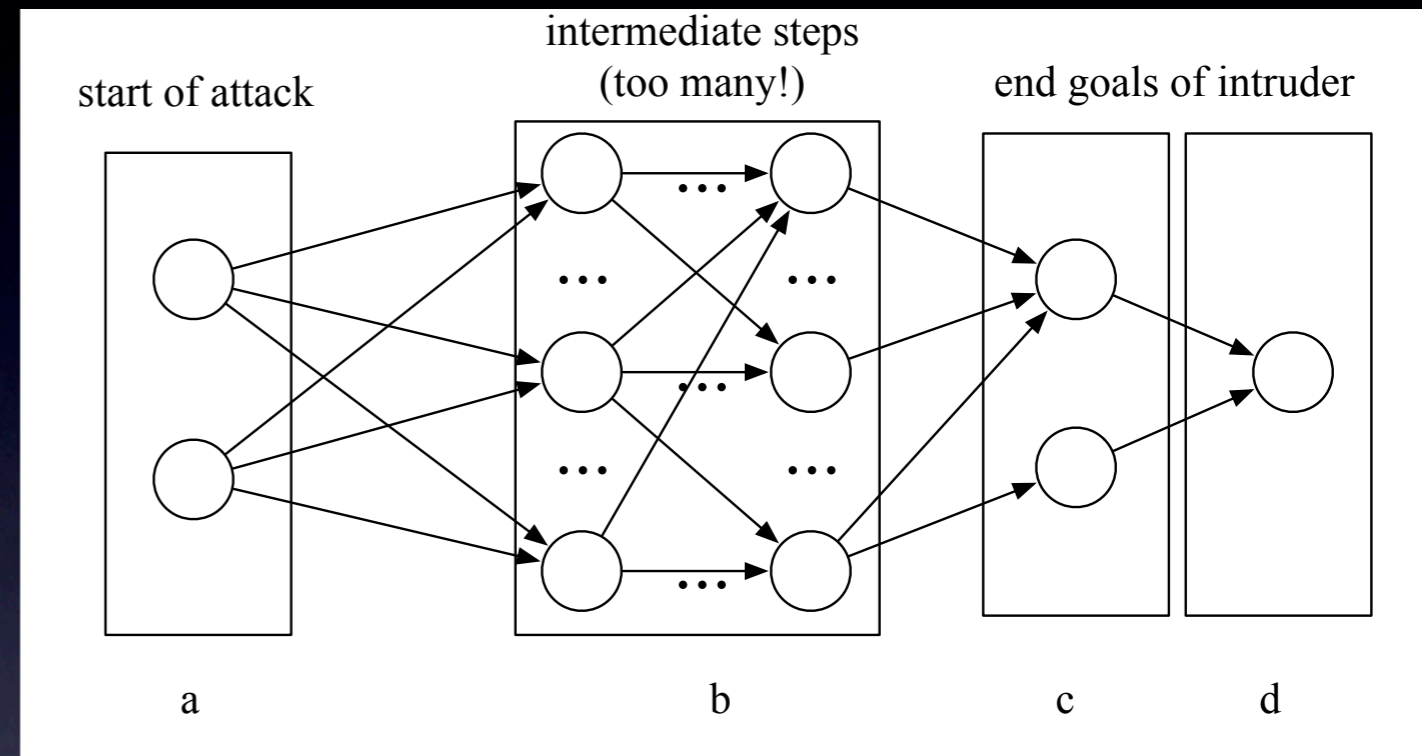
- *Laocoön: A Model of Forensic Logging*
- Attack graphs of goals.
- Goals can be attacker goals or defender goals (i.e., “security policies”)
- Pre-conditions & post-conditions of events to accomplish goals.
- Method of translating those conditions into logging requirements.
- Logs are in a standardized and parseable format.
- Logged data can be at arbitrary levels of granularity.

Goals

- Premise: compute resources are cheap, **human time is expensive**.
- Understand the **scope** of the possible data, analyses, and conclusions.
- Be able to define (or place bounds on) what necessary information is present and **what is missing**.
- Assuming all potentially relevant information is recorded (e.g., by extrospection of a virtual machine), be able to **correlate and prune** the information necessary for a human to analyze.

Attack Graphs

- Intruder goals can be enumerated.
- Vulnerabilities, attacks, and exploits cannot (or in many cases, we would patch them, or they would inhibit usability).
- Defender goals can also be enumerated. They are called security policies.



Security Policies

- Legal policies (HIPAA, Sarbanes-Oxley)
- Formal policies (Bell-LaPadula, Chinese Wall)
- Actual metrics
 - Severity (path length, time, difficulty)
 - Attack Surface Metric
 - Historically known vulnerabilities

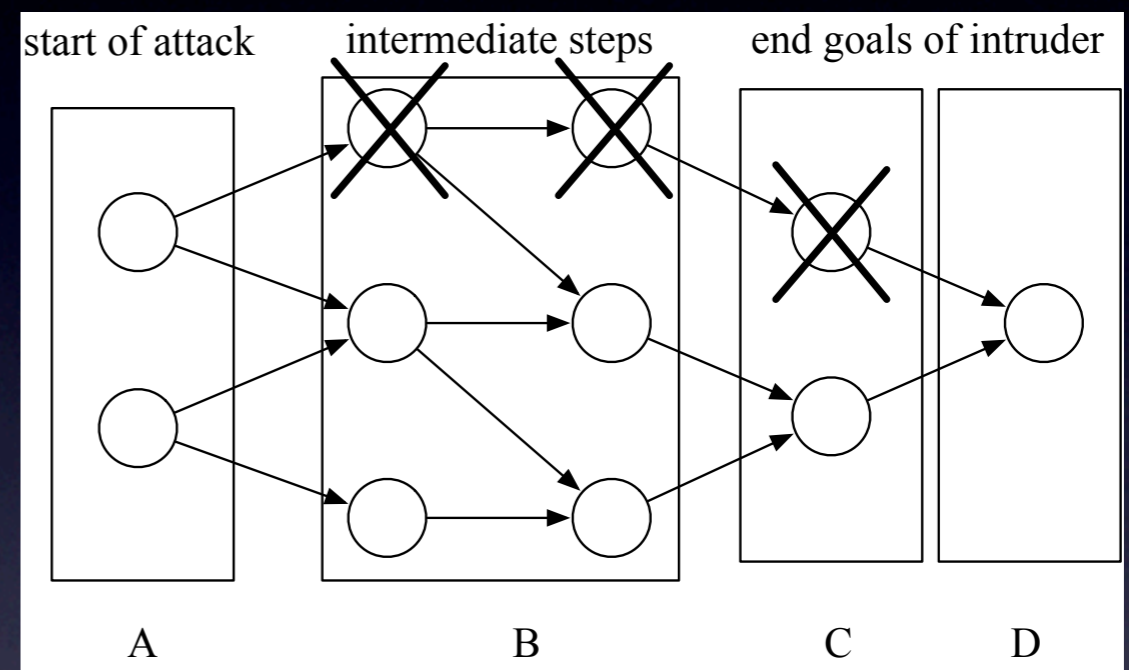
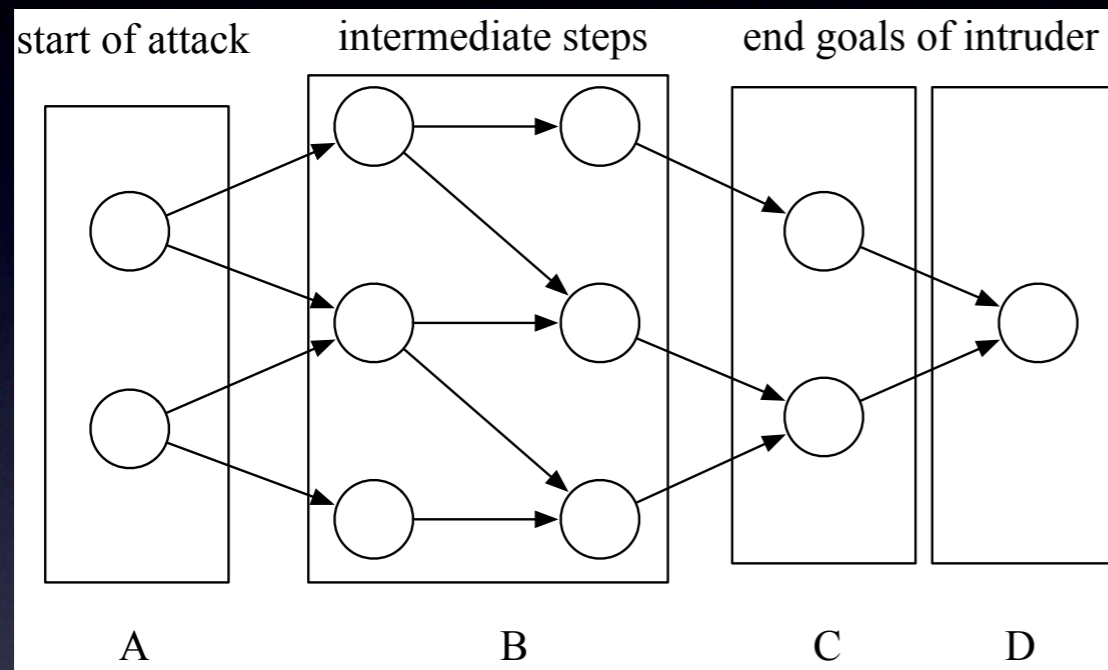
Security Policies

- Security policies can be reverse-engineered or enforced, automatically.
- i.e., determine the current policy, and modify.
 - Policies can be **binary** (block access) or **flexible** (log something).
 - Policies can be **static** (always do this) or **dynamic** (uh oh—an intruder)
- Assumptions get in the way of security. What are they?

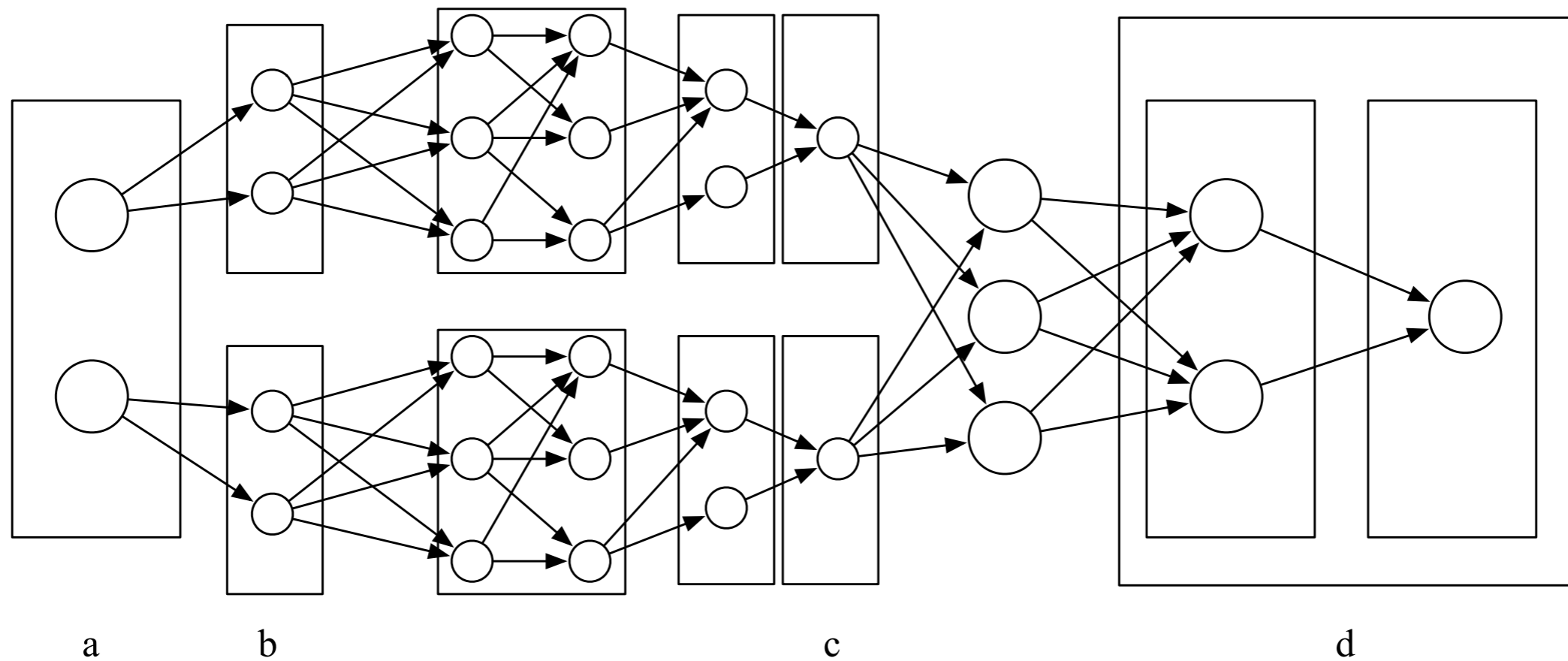
Applying Security Policies

- Applying Laocoön to security policies guides where to place instrumentation and what to log.
- The logged data needs to be correlated with a unique path identifier.
- Branches of a graph unrelated to the attack can be automatically pruned.
- Avoid recording data where events can be recreated because they are deterministic.

Pruning Paths



Complex Attack Graph



Summary

- **Forensics**, attack analysis, logging, and auditing are **broken**.
- We have developed methods to **correlate and constrain data that needs to be analyzed**.
- We have developed methods for **logging based on known vulnerabilities**.
- We have developed methods for integrating **societal needs** (e.g., law) with forensic logging and auditing capabilities.

Thank you

- Questions?
- Sean Peisert
 - peisert@cs.ucdavis.edu
 - <http://www.sdsc.edu/~peisert/>