

DERBI: Diagnosis, Explanation and Recovery from Break-Ins



Mabry Tyson
Douglas B. Moran
Pauline Berry
David Blei
Jim Carpenter
Ruth Lang

**Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park CA 94025
<http://www.ai.sri.com/~derbi>**



Contrast: Traditional Intrusion Detection Systems



- _ Traditional IDSs monitor events as they occur and produce assessment in real-time
- _ DERBI is invoked by “significant” events
 - May be report from another site days after intrusion occurred
 - Must look backwards for evidence of intrusion that occurred before event was noticed



Adaptive Reaction to Threat



- _ Controlled reaction to perceived threat
 - Level of vulnerability
 - Level of resources
 - Level of attack
- _ Multiple triggering events
- _ Follow threads from each piece of evidence
- _ Detects pieces of attack, even if novelty precludes coordinating them

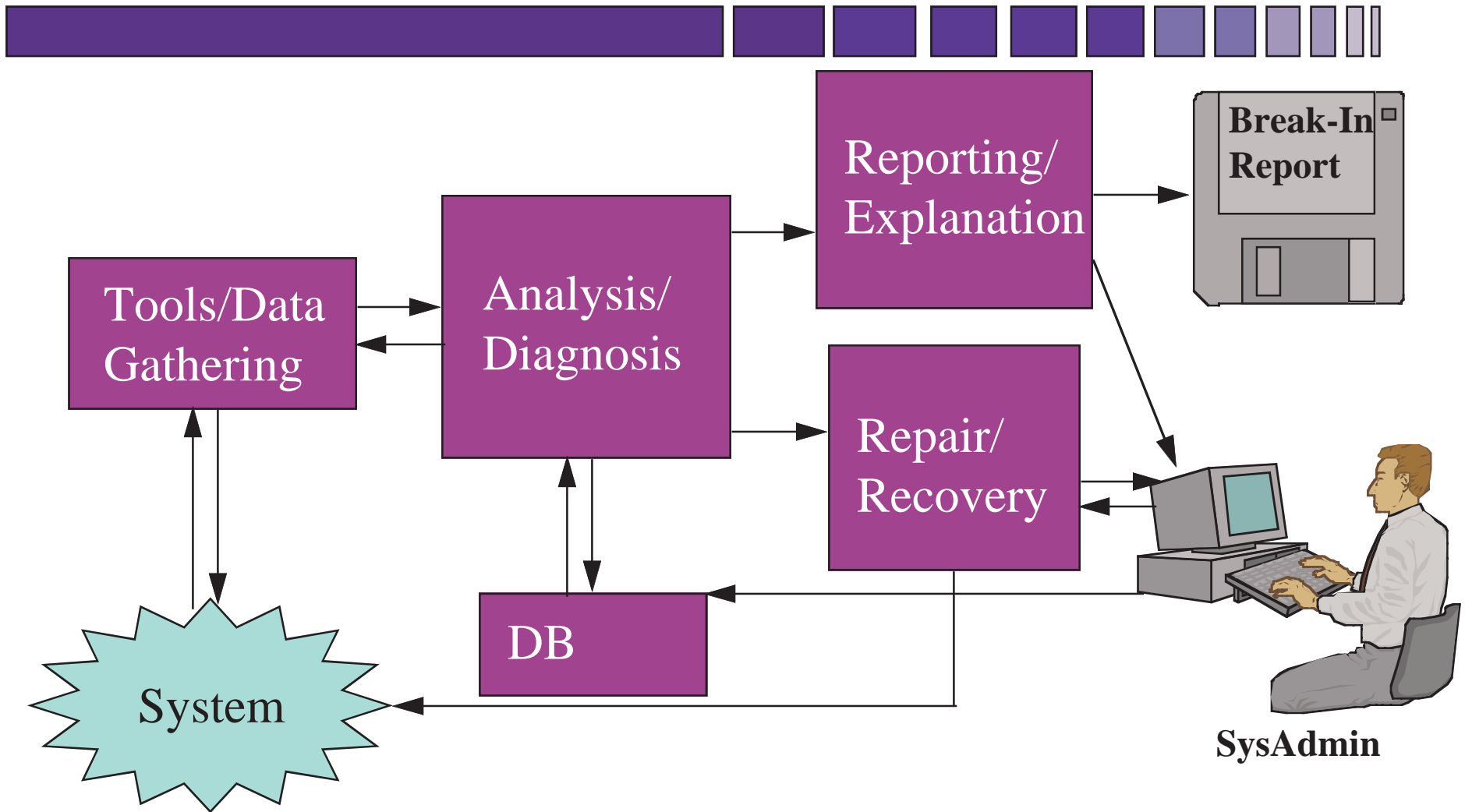


Project Rationale

- _ Issue: limited expertise at most sites
- _ Large commonalties in many intrusions
 - Reusing tools, techniques, tactics
 - Possible variations on shared recipes
- _ Goals:
 - Allow non-expert SysAdmin to understand nature, extent, and recovery of break-in.
 - Faster recovery for typical system
 - Non-intrusive in normal state



DERBI Architecture





Diagnosis: Reasoning from a Model of Intrusion



- _ Model the structure of intrusion event
 - Follow intrusion at abstract level
- _ Relate concrete actions to the abstract stages
- _ Model the relationship of evidence to actions
 - Indirect evidence provides clues to prior and subsequent steps
- _ Explanation to sysadmin of intrusion is based on these models



Benefits of Model



- _ Multiple levels of models provides extensibility
 - _ Models reusable for other platforms
 - _ Novel attack scenarios will give some evidence at different levels
 - _ Models can evolve with exploits
- _ Multiple chains of reasoning promotes robustness



General Model of Intrusion: Components



- _ Point of Entry
- _ Acquire additional privileges (*optional*)
- _ Main Purpose: *Theft, sabotage, publicity, ...*
- _ Camouflage/Concealment
- _ Subsequent activity: *Reentry, data collection*



Camouflage as Indirect Evidence



- _ Hide login by cleaning up wtmp log
 - lastlog inconsistency \Rightarrow *root was compromised*
 - For a user: wtmp/lastlog inconsistency \Rightarrow *which user compromised*
 - For that user: last-access dates on files \Rightarrow *when compromise may have occurred*



Camouflage: Examples of Evidence



- _ Replaced system commands
 - Detection by direct comparison
 - » Checksum: cheaper, spoof-able
 - » Binary comparison: more expensive, original may be unavailable
 - Detection by features
 - » Built from different code base
 - » Hypothesize what it hides and test
 - » Suspicious date (file or directory)



Camouflage: Examples of Evidence



- _ Argument to turn off camouflage?
 - So intruder can see his files and processes
 - Some “conventions”: test if present
- _ Suspicious files and processes
 - Found by other means
 - Confirm modified system command hides suspicious file or process



Reasoning from the Available Evidence



- Replaced system commands
 - Primary function suggests most likely purpose:
ps \Rightarrow hidden process
 - Alternative function (eg, Trojan horse)
 - Replaced system library is an indirect form of replaced system command (harder to identify)



Chain of Reasoning



- _ Modified PS command \Rightarrow hidden process
- _ Hidden process \Rightarrow network sniffer or ...
- _ Network sniffer \Rightarrow info storage or ...
- _ Info storage \Rightarrow retrieval method
- _ Retrieval method \Rightarrow re-entry or transmit
- _ Re-entry \Rightarrow backdoor or ...



DERBI



- _ Reactive tool for analyzing intrusions
- _ Models promote extensibility and robustness
- _ Combine models of intrusion and evidence to guide search for additional information
- _ Model-driven explanation