CMAD IV

COMPUTER MISUSE & ANOMALY DETECTION

PRESENTATION SLIDES AND PAPERS

MONTEREY, CALIFORNIA
NOVEMBER 12-14, 1996

Presented by Jim Anderson, Marvin Schaefer, Salvatore Stolfo, Dai Vu, Raymond Yip, E. Eugene Schultz, Steve Smaha, Kathleen Jackson, Steve Hofmeyr, Richard Lippman, Christoph Schuba, Simson Garfinkle, Hai-Ping Ko, Douglas Moran, Y. Frank Jou, Mark Crosbie, JF Mergen, Katherine Price, Tom Haigh, Ab Kader, Kevin Ziese, Mike Neuman, Mark Schneider, Gene Spafford and Mary Ellen Zurko. Publication assistance by Mary Brown.

Table of Contents

SESSION 1: POLICY-DRIVEN INTRUSION DETECTION AND THE INSIDE THREAT

Misuse
JIM ANDERSON, James Anderson Co.

Auditing for Database Systems and Applications
MARVIN SCHAEFER, Arca Systems, Inc.

Concept Learning and Searching Over Networks Using Java Agents for Meta-learning
SALVATORE J. STOLFO, Columbia University

Distributed Security Policy Database
DAI VU, Lockheed Martin

Misuse Detection in Database Systems
RAYMOND YIP, University of California, Davis

Detecting Insider Attacks
E. EUGENE SCHULTZ, SRI Consulting

SESSION 2: INTRUSION DETECTION TECHNOLOGY FOR SMALL SCALE SYSTEMS

Haystack Labs, Inc. Product Lines
STEVE SMAHA, Haystack Labs, Inc.

A NADIR Progress Report
KATHLEEN A. JACKSON, Computing, Information, and Communications (CIC) Division

Immunology and Computer Security
STEVEN HOFMEYR, University of New Mexico

Lincoln Laboratory Intrusion Detection Research
RICHARD P. LIPPMAN, MIT Lincoln Laboratory

SESSION 3: NEW ATTACKS AND NEW TWISTS ON EXISTING ATTACKS

ATM Firewall Technology: Lessons for Intrusion Detection
CHRISTOPH L. SCHUBA, Purdue University

Denial-of-Service Attacks
SIMSON GARFINKLE

Attacks on Cellular Systems
HAI-PING KO, GTE Laboratories Incorporated

SESSION 4: INTRUSION DETECTION IN THE LARGE

Miscellaneous Papers from Participants

Internetwork Security Monitor: An Intrusion-Detection System for Large-Scale Networks
L. T. HEBERLEIN, B. MUKHERJEE, K. N. LEVITT, UC Davis

Analysis and Response for Intrusion Detection in Large Networks
PETER G. NEUMANN, PHILLIP A. PORRAS, ALFONSO VALDES, SRI International

Distributed Detection of Distributed Attacks
DOUGLAS B. MORAN, SRI International

Scalable Intrusion Detection for the Emerging Network Infrastructure
Y. FRANK JOU, MCNC

Autonomous Agents
MARK CROSBIE, Hewlett-Packard/COAST

Network Management and Operations
JF MERGEN, BBN

SESSION 5: NEW ENVIRONMENTS FOR INTRUSION DETECTION

Thoughts About Susceptibility to Data Driven Attacks
MARVIN SCHAEFER, Arca Systems, Inc.

The Need for a Standard for the Format and Content of Audit Trails
KATHERINE PRICE, Purdue University

Auditing on Sidewinder
TOM HAIGH, Secure Computing Corp.

Information Security and the Electric Power Industry
AB KADER, EPRI

SESSION 6: TOOLS FOR INVESTIGATIVE SUPPORT

Computer Based Forensics – A Case Study – U. S. Support to the U. N.
CAPT. KEVIN J. ZIESE, AF Information Warfare Center

Interactive Intrusion Detection
MIKE NEUMAN, En Garde Systems, Inc.

SESSION 7: NEW IDEAS

CMAD IV Summary
MARK SCHNEIDER, Office of INFOSEC Research

Some Thoughts
GENE SPAFFORD, Purdue University

New Ideas: Borrowing from other Areas
MARY ELLEN ZURKO, Open Group Research Institute

Misuse

JIM ANDERSON

James Anderson Co.

Auditing for Database Systems and Applications

MARVIN SCHAEFER

Arca Systems, Inc.

Auditing for Database Systems and Applications

Marvin Schaefer

Arca Systems, Inc.

CMAD IV Workshop

November 12-14, 1996

Audit Goals

- **Accurately Record Events**
  - **Protection Critical**
    - **Policy-Based**
    - **Who, What, When**
  - **Integrity Critical**
    - **Transaction-Based**
    - **Who, What, When, Why**
  - **Private**
    - **from normal users, applications**
    - **from "outsiders"**
  - **Protected**
    - **from unauthorized users**

Levels of Abstraction

- **User**
  - Intention?
- **Application**
  - HOL Expression
  - Compiled Expression
    - Optimisation
- **Server**
  - Input
  - Mediation
  - Result
- **O/S - IDS**
  - Gazintas
  - Gazoutes

Points of Confusion

- OS/TCB excel in syntactic stuff
  - Does not address change in contents
- Applications excel in semantic richness
  - Does address internal content and structure
- OS/TCB relatively old, archaic, stable, planned
- Applications often flexible, up-to-date, once planned, intuitive or hypeful
  - Maintenance & planning vs $49 update

Concept Learning and Searching Over Networks Using Java Agents for Meta-learning

SALVATORE J. STOLFO

Department of Computer Science

Columbia University

Concept Learning and Searching Over Networks Using Java Agents for Meta-learning

———-

THE JAM PROJECT

———-

Application: FRAUD AND INTRUSION DETECTION IN FINANCIAL INFORMATION SYSTEMS ———-

———-

CMAD IV

Salvatore J. Stolfo
Department of Computer Science
Columbia University

——

Electronic Commerce on the WEB provides New Challenges

- More data and services are available everyday on the WEB
- We seek a new way to search and LEARN FROM very large and remote databases
- Electronic Commerce provides new opportunities for Electronic FRAUD
- We seek a new way to LEARN about FRAUD on the WEB
- Proposal: Build an IMMUNOLOGICAL Capability for the WEB to DETECT FRAUD
- Learn SELF (Good Transactions) from NON-SELF (Bad Transactions)

A New Information Extraction Paradigm

- Empower the User with *Data Mining* Tools to Learn Knowledge from Data
- Agent Proxies that Learn Knowledge over Remote Data
- Agent Proxies that Learn Collective Knowledge over Remote Agents
- Agent Proxies Use Learned Knowledge to Search Other Data

Terminology

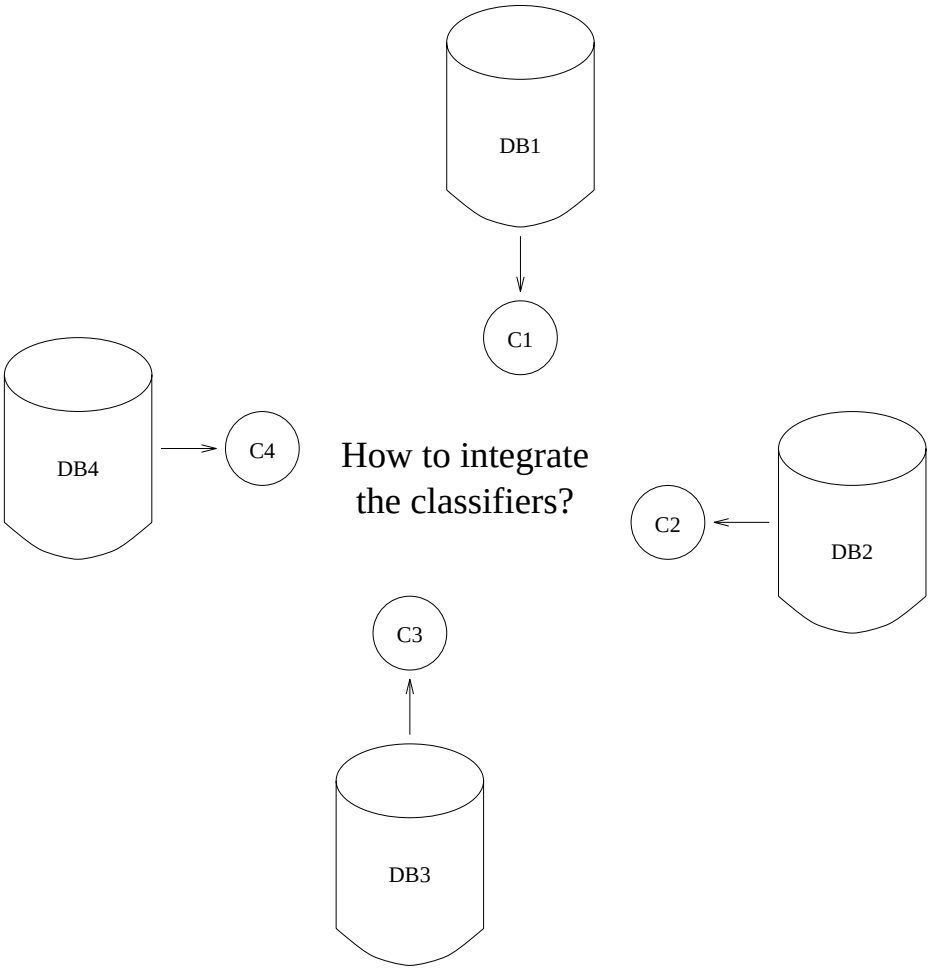
- Data Mining: Scalable Machine Learning Applied to Verly Large Databases
- Learning Agent: A Machine Learning program launched to and applied at a remote source of data
- Classifier Agent: A derived program learned over some remote site of data, labels or tags data with class labels
- Meta-Learning Agent: A Machine Learning program that Learns how to combine a number of remote classifier agents, the result is a single classifier agent

Meta-learning: An Algorithm-independent Technique for Scalable and Accurate Inductive Learning

Salvatore J. Stolfo
Department of Computer Science
Columbia University and
Philip Chan
Florida Institute of Technology

Learn and Integrate Classifiers

- Large datasets are partitioned into subsets
- Distributed databases are inherently partitioned
- Collective knowledge is harvested from individual knowledge sources

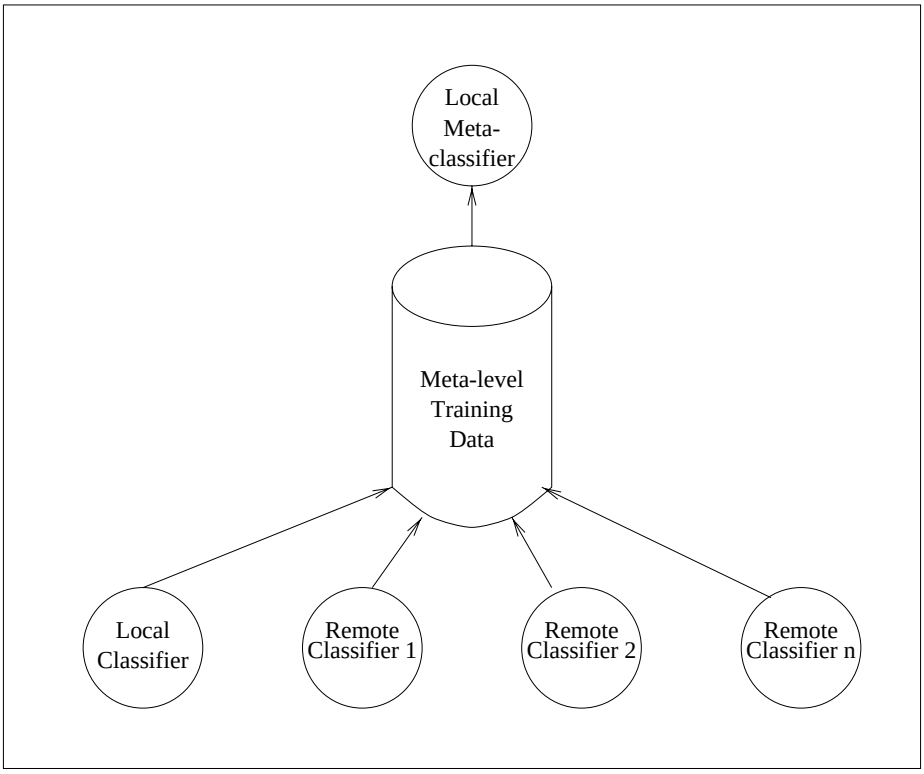


Integrating Classifiers

- Integrating the *concept descriptions languages* (a logical cross-bar switch)?
  - different representations: probabilities, hyperplanes, logical expressions
  - difficult if not impossible to accurate map all representations into one standard

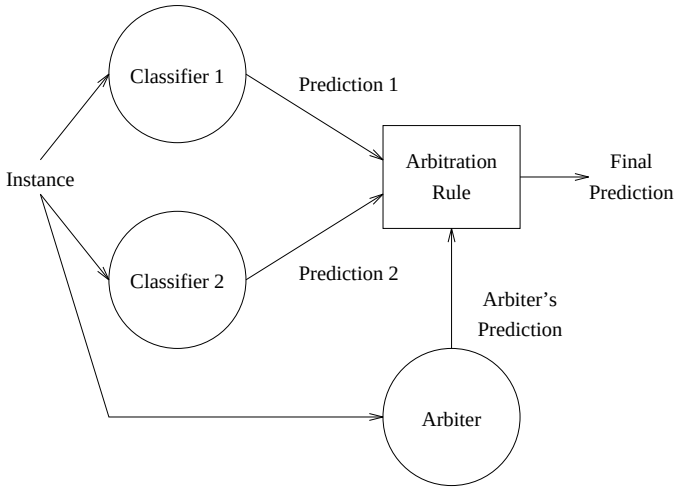
- Integrating the behavior of classifiers (their predictions)?
  - algorithm/representation-independent
  - existing and new algorithms can be plugged in with ease
  - voting and statistical techniques abound
  - meta-learning:
    - arbitration: conflicting predictions are resolved by a learned arbiter
    - combining/coalescing: learn a function over classifiers' predictions

SHARING REMOTE CLASSIFIERS

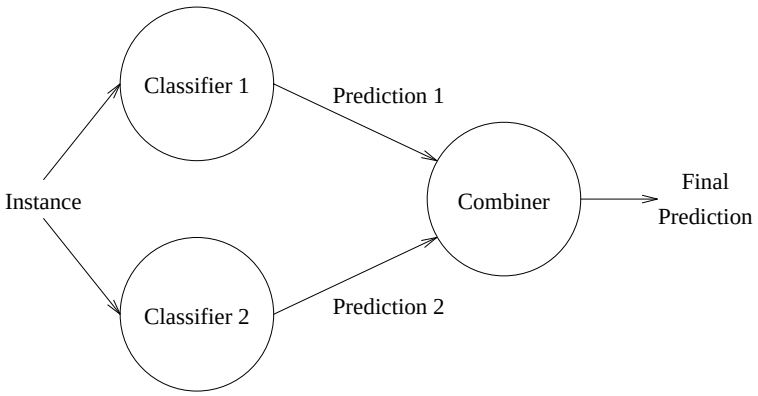


SHARING KNOWLEDGE WITHOUT SHARING DATA

Meta-learning: Arbiters and Combiners



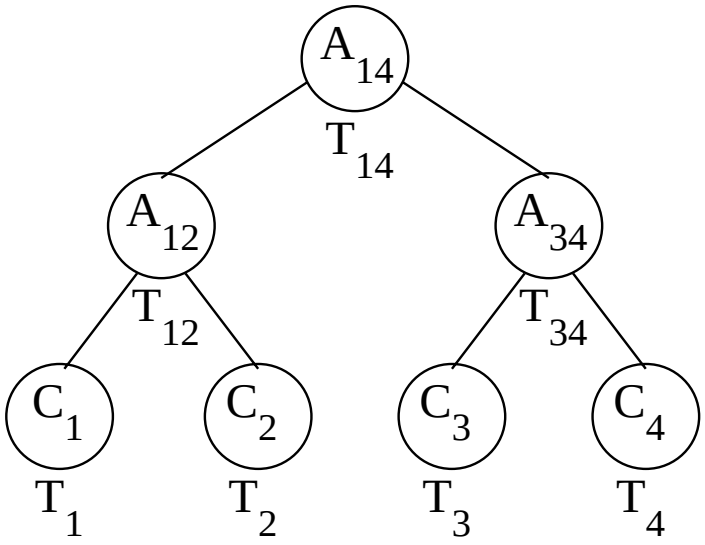
- The *arbiter* Resolves conflicting predictions (disagreements)



- The *combiner* makes a final prediction based on the base predictions

Hierarchical Meta-learning in Agent Infrastructures

- *Arbiter tree*

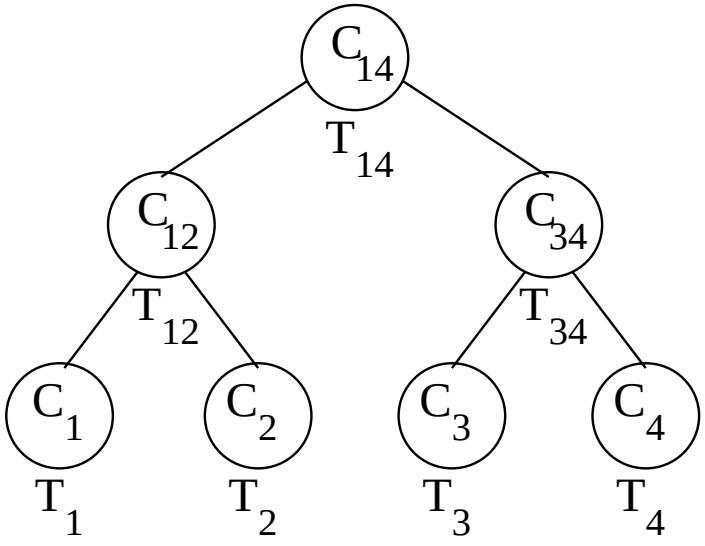


Arbiters

Classifiers

Training data subsets

- *Combiner tree*



Combiners

Classifiers

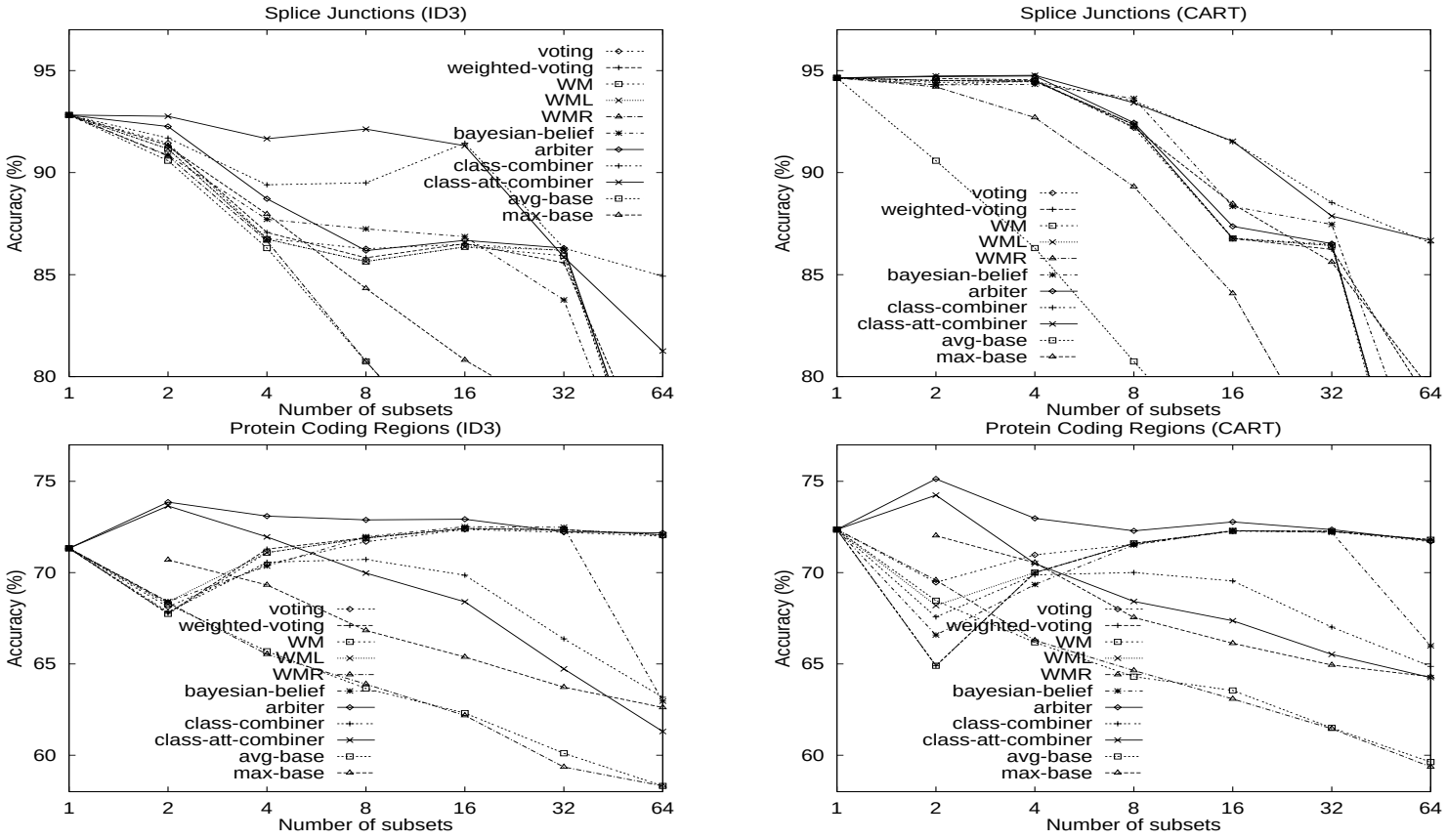
Training data subsets

Evaluation Studies

- Many issues exist and are addressed by various experiments
- Main focus is on prediction accuracy
  - disjoint training and test sets
  - 10-fold cross validation
  - 2 to 64 data subsets
  - global classifier (whole dataset or 1 data subset)
- "Off-the-shelf" learning algorithms
  - ID3 (Quinlan 86)
  - CART (Breiman et al. 84)
  - BAYES (Clark & Niblett 87)
  - WPEBLS (Cost & Salzberg 93)
- "Off-the-shelf" learning tasks
  - DNA splice junctions (3,190) (Towell et al. 90)
  - Protein coding regions (21,625) (Craven & Shavlik 93)
  - Protein secondary structures (20,000) (Qian & Sejnowski 88)

Subsets and Sampling

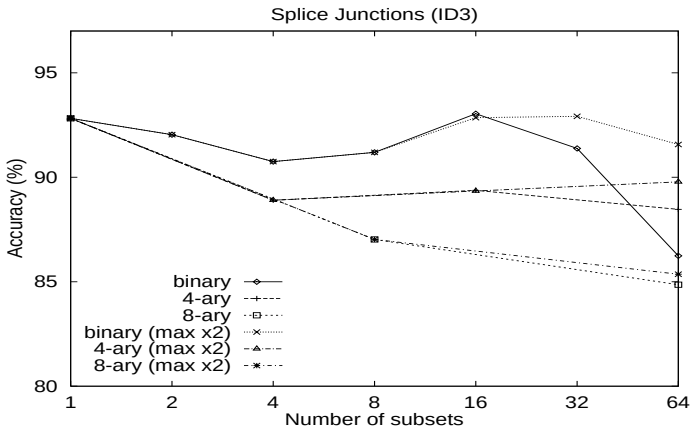
- How do the # of subsets and subset size affect accuracy?
- Is random sampling of a subset sufficient?

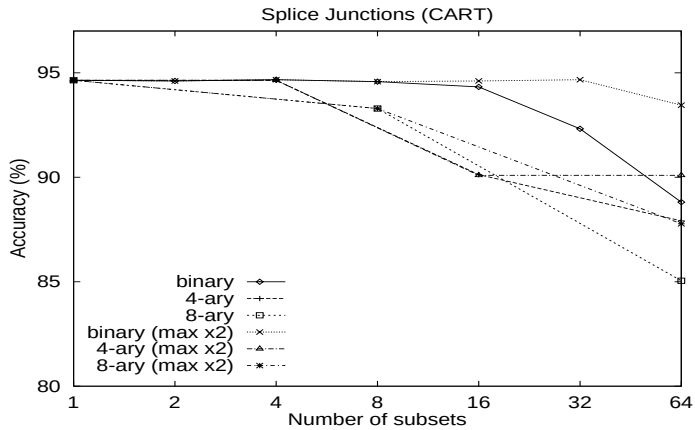


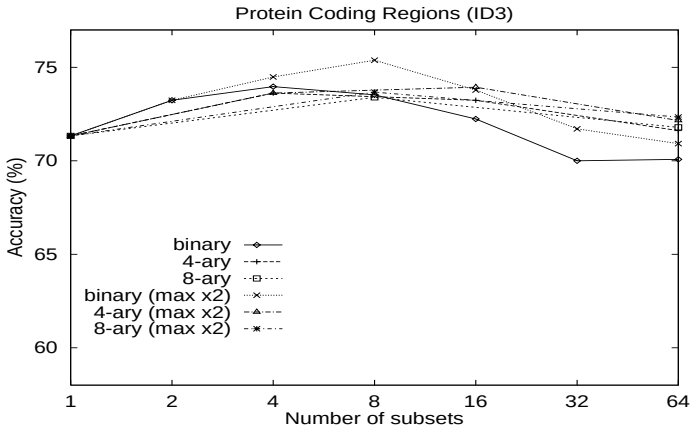
- Subsets can't be too small to generate reasonable classifiers
- Random sampling is not sufficient; combining is necessary

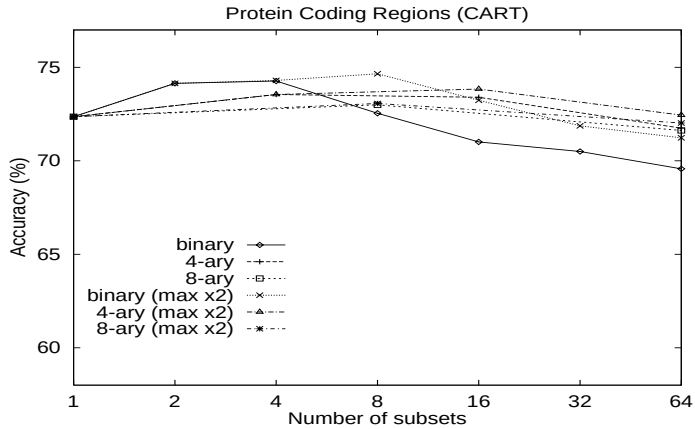
Arbiter Trees

- Is hierarchical meta-learning necessary?
- How do the order of the arbiter trees and training set size limit affect the accuracy?





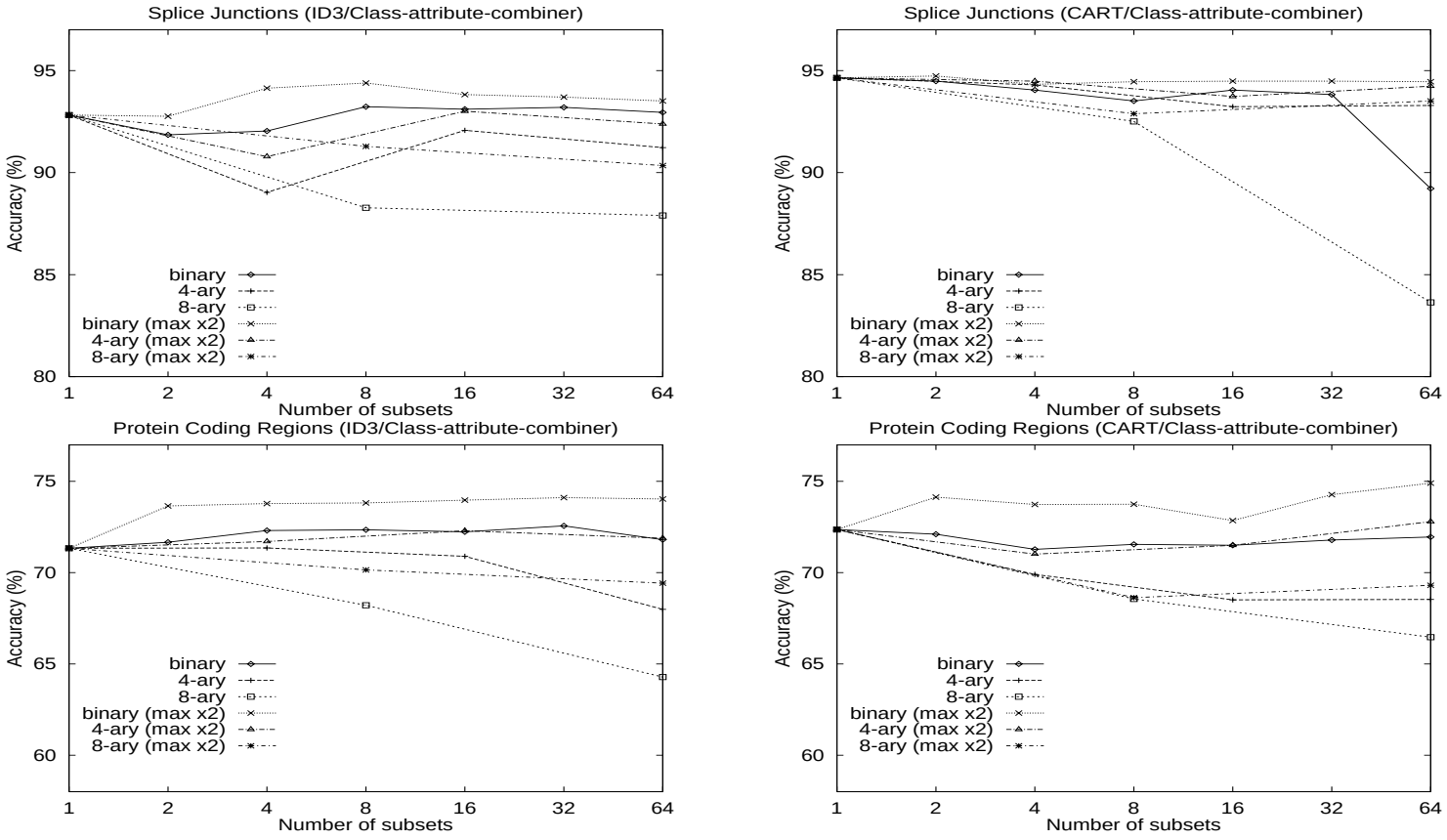




- Lower order trees are more accurate
- Doubling the arbiter training set size maintains accuracy

Combiner Trees

- How does the combiner trees fare?
- Class-attribute-combiner strategy



- Statistically significant and consistent improvement in the PCR dataset beyond the original accuracy

Summary of Meta-learning Results

- Random sampling is not sufficient
- Existing voting and statistical combining techniques are not sufficient
- "One-level" meta-learning outperforms the voting and statistical techniques
- Hierarchical meta-learning can sustain high accuracy
- Meta-level training set size needs only to be twice the subset size
- Proportional distribution of classes in the data subsets is beneficial
- Lower-order trees are more accurate than higher-order trees
- Combiner trees can boost accuracy beyond the global classifier's
- Data replication does not improve accuracy

An Illustration: Distributed DNA Sequence Databases

SITES 1 and 2:

DNA sequence #	Nucleotide sequence
1	...CCAGCTGCATCACAGGAGGCCAGCGAGCAGGTCTGTTCCAAGGGCCTTCGAGCCAGTCTG...
2	...GAGAGAGAGACCAGAAATAATCTTGCTTATGCTTTCCCTCAGCCAGTGTTTACCATTGCA...

DNA sequence #	Nucleotide sequence
1	...ACAGGCTTTTCACAGCCTCCAGCGAGGCATGTACTGATTCCAGGCCTCGGAGCCAGTCTG...
2	...TAGCCGAGACAAAGGATAAGTCTTGATGTATGCTTACCACAGTCTAATGCTTCCCATACT...

Sample SPLICE JUNCTION sequences at SITE 3

Junction	p_{-30}	p_{-29}	$p_{-28} \cdots p_{-3}$	p_{-2}	p_{-1}	p_1	p_2	$p_3 \cdots p_{28}$	p_{29}	p_{30}
intron-exon (IE)	C	T	..TAATAACATTCTTAT	A	G	G	G	..ATCCATTCATGTGAAT	A	T
exon-intron (EI)	G	A	..GCCCGTCATAAAATC	T	G	G	T	..GAGACTCATGCCCAGC	T	C
neither (N)	T	A	..CTATCCACAGACAGT	A	G	G	A	..TGCCCGCCTCTGGGCA	A	A

An ID3 Decision Tree Learned Over SJ Data at SITE 3

```
p-1 = A:
|   p2 = A:  N
|   p2 = C:  N
|   p2 = G:  N
|   p2 = T:
|   |   p5 = A:  N
|   |   p5 = C:  N
|   |   p5 = G:
|   |   |   p1 = A:  N
|   |   |   p1 = C:  N
|   |   |   p1 = G:  EI
|   |   |   p1 = T:  N
|   |   p5 = T:  N
p-1 = C:  N
p-1 = G:
|   p2 = A:
|   |   p-2 = A:
|   |   |   p-3 = A:  N
|   |   |   p-3 = C:  IE
|   |   |   p-3 = G:  N
|   |   |   p-3 = T:  IE
|   |   p-2 = C:  N
|   |   p-2 = G:  N
|   |   p-2 = T:  N
|   p2 = C:
|   |   p-2 = A:  IE
|   |   p-2 = C:  N
|   |   p-2 = G:  N
|   |   p-2 = T:  N
```

A (logic-based) rule equivalent of the first branch at the top of the ID3 Decision tree is:
"If ($X.p_{-1} = A$) and ($X.p_2 = A$) then the center doesn't have a junction, i.e. X.Junction = N."
A rule equivalent to the second branch is:
"If ($X.p_{-1} = A$) and ($X.p_2 = C$) then the center doesn't have a junction, i.e. X.Junction = N."

Sample Sequences To Be Extracted

Classifier Agent Sent to SITE 1:

Select X. From DNA-Sequence Where* $C_{ID3-1}(X.p_{-30}..X.p_{30}) = EI.$

C_{ID3-1}	Meta-classifier	p_{-30}	$p_{-29}\cdots p_{-3}$	p_{-2}	p_{-1}	p_1	p_2	$p_3\cdots p_{29}$
EI	EI	A	CCAAGAAGGGATCTATCACCTCTGTAC	A	G	G	T	AAGAAAAATTACATAGATGAAGATCTG
EI	EI	T	GGCGACTACGGCGCGGAGGCCCTGGAG	A	G	G	T	GAGGACCCTGGTATCCCTGCTGCCAGT
N	EI	G	GAGCTGCCAGACACGGAGGAGAGCCAT	G	A	G	T	AAGTGGGCCCAGCTGAGGGTGGGCTGG
N	N	A	TTCTACTTAGTAAACATAATTTCTTGT	G	C	T	A	GATAACCAAATTAAGAAAACCAAAACA
N	N	A	GGCTGCCTATCAGAAGGTGGTGGCTGG	T	G	T	G	GCTGCTGCTCTGGCTCACAAGTACCAT

A Sample Meta-Classifier Learned From 4 Base Classifiers

```
c-id3-1 = EI:  EI
c-id3-1 = IE:
|   p-3 = A:  N
|   p-3 = C:  IE
|   p-3 = G:  N
|   p-3 = T:  IE
c-id3-1 = N:
|   p1 = A:  N
|   p1 = C:  N
|   p1 = G:
|   |   p5 = A:  N
|   |   p5 = C:  N
|   |   p5 = G:
|   |   |   p2 = A:  N
|   |   |   p2 = C:  N
|   |   |   p2 = G:  N
|   |   |   p2 = T:  EI
|   |   p5 = T:  N
|   p1 = T:  N
```

A Host Meta-Learning Environment

- Partitioning and Distributing data,
- Invoking Different Meta-Learning Strategies In Parallel,
- Pairing Classifiers to Reduce Intermediate Training Sets for Meta-Learning,
- Filtering and Communication of Training and Testing Data Between Processors, and,
- Instrumentation to Gather Statistics Used in Formulating or Designing Specific Meta-Learning Architectures.
- LAUNCHING OF ENCAPSULATED LEARNING AND META-LEARNING AGENTS OVER NETWORKS

Future Research: The JAM PROJECT

- Specialized representations (new attributes/predicates) and algorithms for meta-learning
- New meta-learning strategies and training-set composition rules
- Agent computing: collaboration with FSTC in field-testing learning agents on the Internet:
- – Acquisition of TRANSACTION DATABASES with FRAUD LABELS
  - Demonstration of Remote Learning and Meta-Learning Agents
  - Exchange of Learned Classifiers
  - Installation of Learned Classifiers as SENTRIES to warn of FRAUD

JAM Prototype: One coordinator, multiple data sites

- Coordinator
  - Dispatches agents to different data sites
  - Multithreaded for concurrent service
  - Simple error recovery from data sites crashes
- Data Site
  - Accepts and executes agents
  - Agent Independent
- Agent: the ID3 machine learning algorithm
- Platform Independent (Java)
- Simple Graphical User Interface

Data Schema and Stats for (Fraud) Transaction Data Sets

- Number of Attributes: $30 + -\Delta$ (all numeric)
  - Many fields are categorical (i.e.numbers represent a few discrete categories)
  - Developed over years to capture important information
- Size: Fixed 137 bytes per transaction
- Type of Information:
  - A (jumbled) account number (no real identifiers)
  - Scores produced by a COTS authorization/detection system
  - Date/Time of transaction
  - Past payment information of the transactor
  - Amount of transaction
  - Geographic information: where the transaction was initiated, the location of the merchant and transactor
  - Codes for validity and manner of entry of the transaction
  - An industry standard code for the type of merchant
  - A code for other recent "non-monetary" transaction types by transactor
  - The age of the account and the card
  - Other card/account information
  - Confidential/Proprietary Fields (other potential indicators)
  - Fraud Label (0/1)
- .5MM records by each Bank:
  - sampling 50,000 per month
  - Span 11/95 - 10/96

VISIT with your favorite Browser:

- http://www.fstc.org - and click on Fraud Page
- http://www.cs.columbia.edu/šal
- http://www.cs.columbia.edu/šal/JAM/PROJECT

SUPPORTED BY:

- NYSSTF Polytechnic Univeristy CATT
- NSF CISE KMCS and DBES Programs
- DARPA ITO Intrusion Dection Program

Distributed Security Policy Database

DAI VU

Lockheed Martin

LOCKHEED MARTIN
Communication Systems

Distributed Security Policy Database

Dai Vu

CMAD IV
November 11, 1996

Topics of Discussion

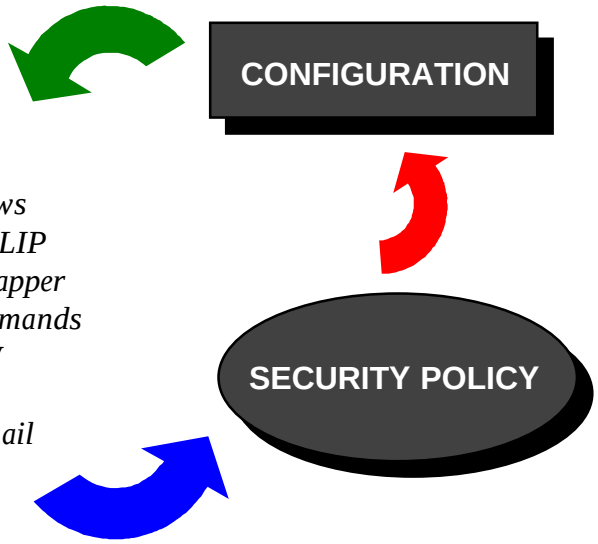
LOCKHEED MARTIN
Communication Systems

- Need for Policy Database
- Object Classes
- Regions and Spheres of Influence
- Database in Operation

Need for Policy Database

LOCKHEED MARTIN Communication Systems

- *FTP*
- *Cron*
- *Finger*
- *NIS*
- *Netnews*
- *PPP/SLIP*
- *Portmapper*
- *R-commands*
- *WWW*
- *NFS*
- *Sendmail*



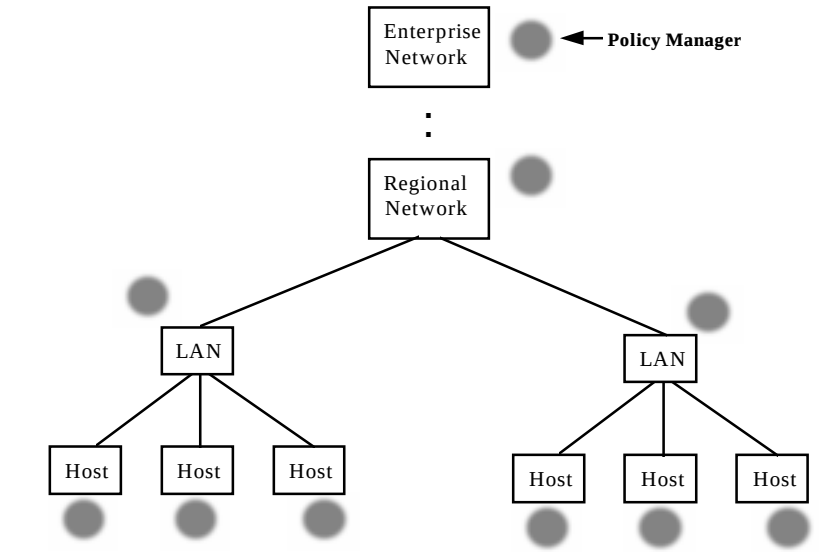
Object Classes

LOCKHEED MARTIN Communication Systems

- Threat - What possible Threats
- Policy - What is allowed / disallowed
- Resources - Users, Hosts, PID, etc.
- Functionality - What is required for policy
- Configuration - Enable / disable
- Regions - Organization of processes

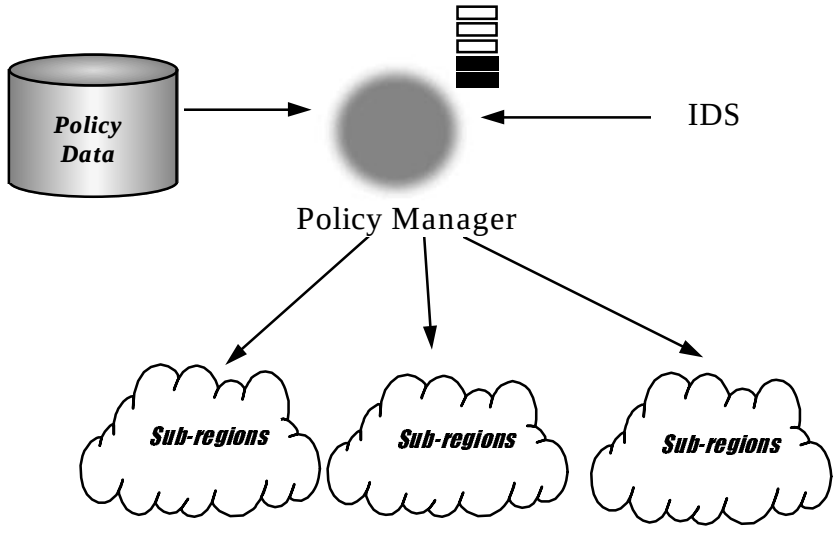
Regions and Spheres of Influence

LOCKHEED MARTIN
Communication Systems



Database in Operation

LOCKHEED MARTIN
Communication Systems



Misuse Detection in Database Systems

RAYMOND YIP (SPEAKER)

KARL LEVITT

University of California, Davis

Detecting Insider Attacks

E. EUGENE SCHULTZ

SRI Consulting

Haystack Labs, Inc. Product Lines

STEVE SMAHA

Haystack Labs, Inc.

Computer Misuse and Anomaly Detection - IV (11/96)

Steve Smaha

President
Haystack Labs, Inc.
10713 RR 620 North, Suite 512
Austin, Texas 78726
(512) 918-3555 (voice)
(512) 918-1265 (fax)
smaha@haystack.com
http://www.haystack.com

How To Form Your Very Own Silicon Valley Startup

by Laura Lemay

1. Go to Menlo Park. Find a tree.
2. Shake the tree. A venture capitalist will fall out.
3. Before the venture capitalist regains its wits, recite the following incantation: “Internet! Electronic Commerce! Distributed Enterprise-Enabled Applications! Java”
4. The venture capitalist will give you four million dollars.
5. In 18 (12? 6? 3?) months, go public.
6. After you receive your check, go back to Menlo Park. Find a tree.
7. Climb it. Wait.

Haystack Labs, Inc.

- Founded in 1989 & based in Austin, Texas
- 25 employees, 3 offices
- Current product development began in 1991
- R&D work for intelligence agencies
- University of Texas Technology Incubator Graduate
- Venture funded - Venrock, Trellis

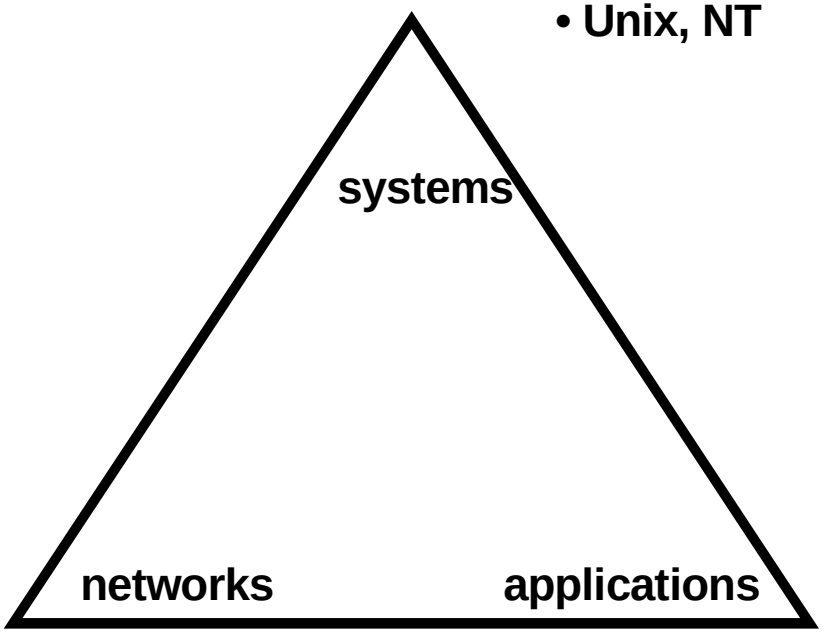
Partners

- **Sun, IBM, Storage Technologies (Network Systems Corp.), AT&T, European VARs, Ascend**
- **Coming soon: firewall vendors, PC/NT vendors**

Product Lines

Stalker

• Unix, NT



NetStalker

• NSC routers/switches
• Embedded in firewalls

Application Stalker

• WebStalker
• CommerceStalker
• More...

Underlying Technology

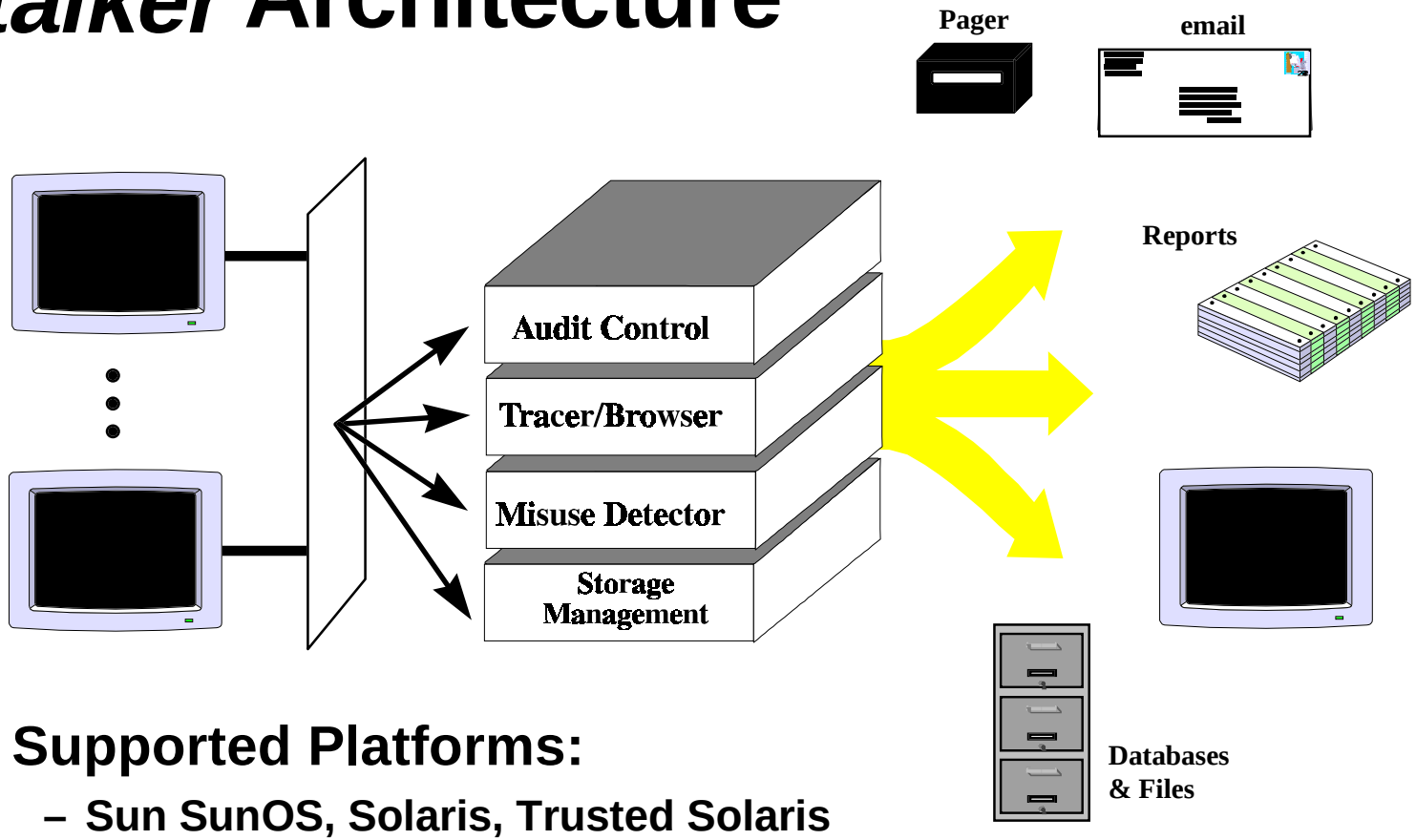
- **Generic signature recognition approach**
  - **Developed in 1992-93 after delivering and installing statistical and AI-based systems**
  - **Applying compiler/parser techniques to look for security-relevant patterns in audit trails, network event logs, and other security logs**
  - **U.S. patent #5,557,742 issued 17 Sep 96, other countries pending**
- **Engine + database model**
- **Significant use of outcomes analysis as “safety net”**

What's In The Patent?

ABSTRACT

A processing system intrusion and misuse detection system and method utilizes instructions for and steps of processing system inputs into events and processing the events with reference to a set of selectable misuses in a misuse engine to produce one or more misuse outputs. The system and method convert processing system generated inputs to events by establishing an event data structure that stores the event. The event data structure includes authentication information, subject information, and object information. Processing system audit trail records, system log file data, and system security state data are extracted from the processing system to form the event data structure. A signature data structure stores signatures that the misuse engine compares and matches to selectable misuses. The signature data structure includes an initial state for each selectable misuse, an end state for each selectable misuse, one of more sets of transition functions for each selectable misuse, and one or more states for each selectable misuse, which can include the end state or the initial state. Furthermore, a misuse output and an index are utilized so that for each selectable misuse element there is a mechanism for loading the signature data structure.

Stalker Architecture



- **Supported Platforms:**
  - **Sun SunOS, Solaris, Trusted Solaris**
  - **IBM AIX**
  - **HP- UX**
  - **NT 4.X Soon**

Misuse Detector: What *Stalker* Detects

Insider and outsider activities:

Known attacks

- **"doorknob rattling"**
- **rdist**
- **rlogin bin**
- **ICMP**
- **login trojan horses**
- **NFS mounts**
- **YP/NIS maps**
- **RPC portmapper**
- **Password "sniffer"**
- **SATAN**

Attempts to exploit known vulnerabilities

- **bugs in the code**
- **design flaws**
- **unexpected interactions with other system components**
- **affects operating systems, network protocols, applications**
- **example: "Internet worm" of 1988**
- **SATAN**

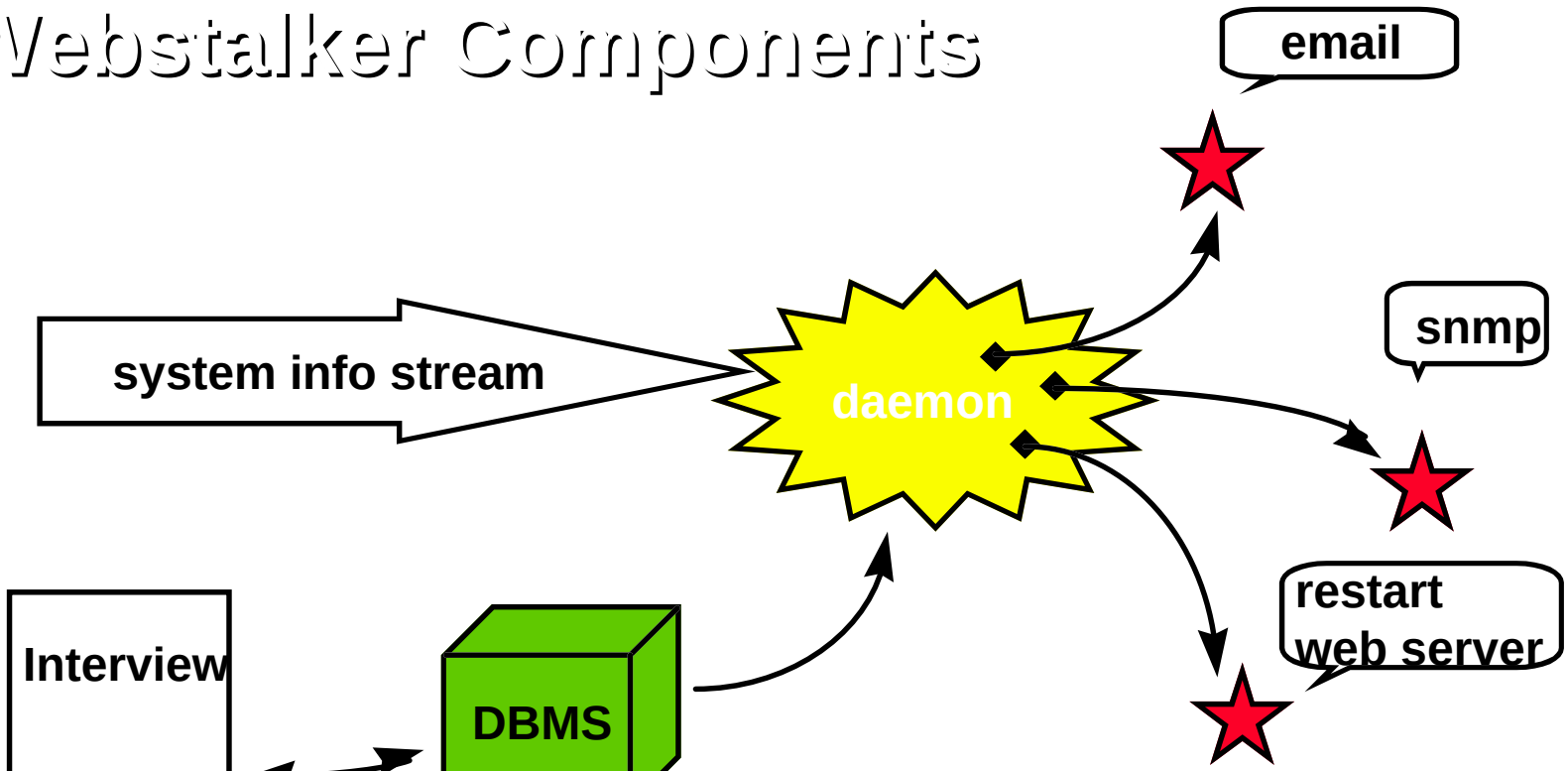
Known attack outcomes

- **Detecting these outcomes provides a "safety net" for trapping new hacker techniques.**
  - **Privilege escalation**
  - **Monitors disabled**
  - **Special files modified**

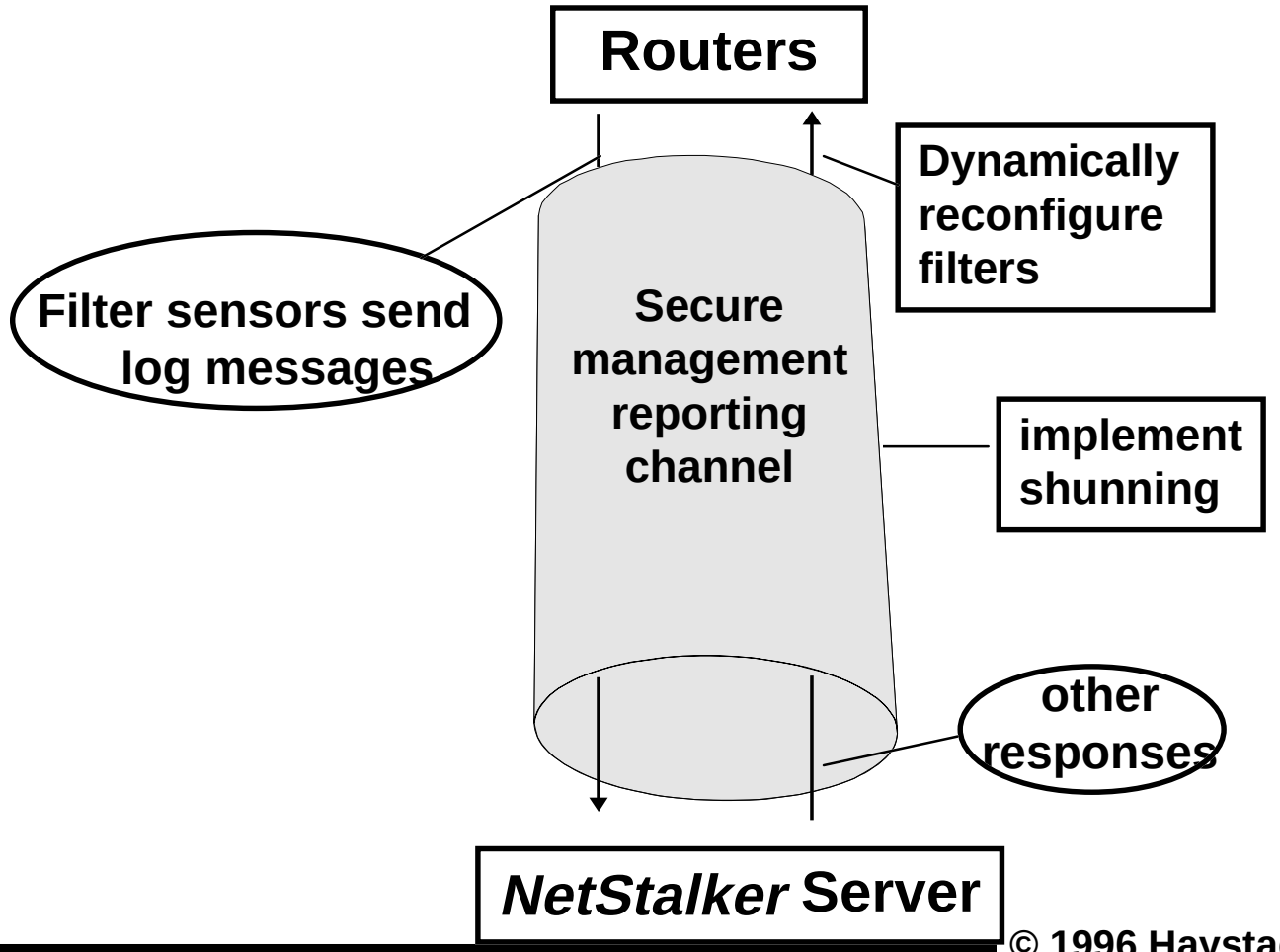
Introducing...

Real-time security monitoring for Web servers, ensuring 7 x 24 availability and integrity

Webstalker Components



NetStalker for NSC: Architecture



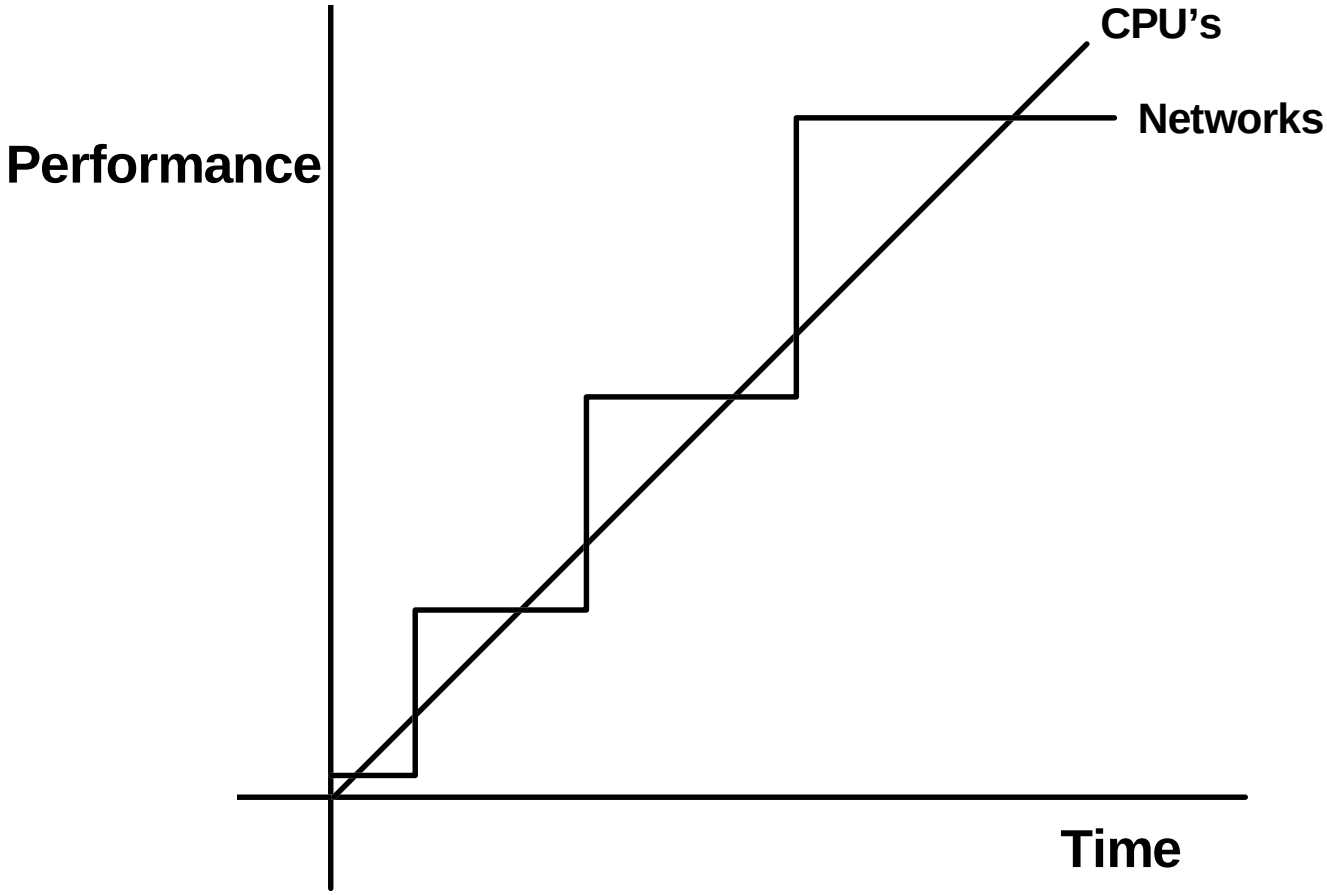
Overview

- **The new threat**
- **Challenges of increased bandwidth**
- **Signs of hope: toasters**
- **Suggested roles**

The Biggest Problem: Vendors

- **Outsiders -> Insiders -> Vendors**
- **Mass-market software is designed to satisfy 80% of the market's needs, and to do so NOW!**
- **3-4 major releases a year:**
  - **How much testing before your users download it?**
  - **Security flaws published in minutes on the Internet!**
- **Security products are mostly Band-Aids™.**
- **Large PC vendors don't give any special priority to security problems reported by governments.**

Fundamental Problem of Network Security Monitoring



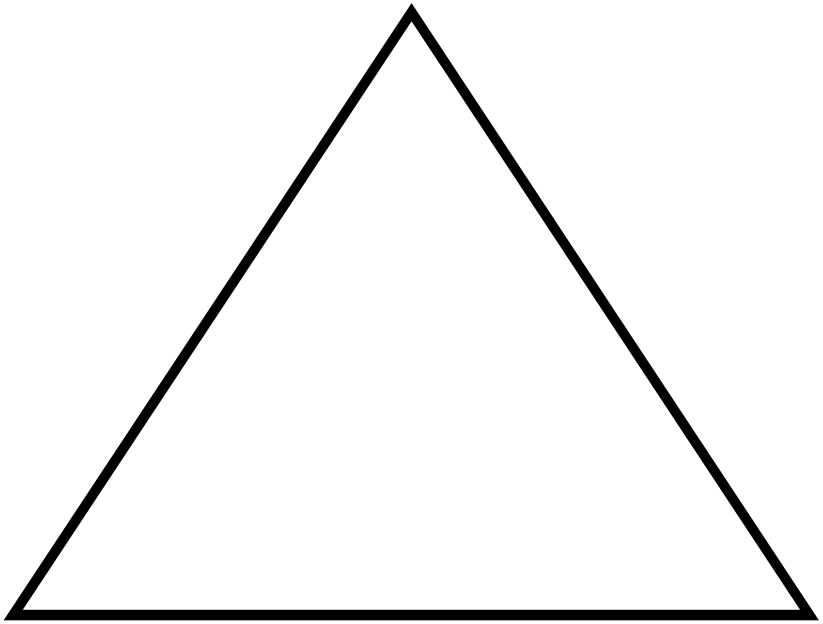
Toasters Are A Good Thing.

- **Distributed computing with cheap boxes allow specialization of functions.**
  - **Divide and conquer … sounds object-oriented!**
  - **Fewer general purpose computers**
  - **Do one thing and do it well: e.g. serve Web pages.**
- ***SOME* may be built on a recycled MLS/CMW base, <u>but not many</u>!**
- **Major research issues:**
  - **How to state security attributes of components?**
  - **How to compose pieces into bigger systems?**

Active Security Needs

Availability

- 7 x 24
- Restart



Paranoia

- Confidentiality
- Integrity

Accountability

- Audit Requirements
- Reconstruction
- Traceability

A NADIR Progress Report

KATHLEEN A. JACKSON

Division Security Office

Computing, Information, and Communications (CIC) Division

A NADIR Progress Report

Kathleen A. Jackson
Team Leader, Division Security Office
Computing, Information and Communications Division

What is NADIR?

- Network Anomaly Detection and Intrusion Reporter
- Los Alamos-developed system, operational since 1990
- Accredited by the DOE
- Looks for attempted ICN intrusion and misuse
- Monitors several critical systems on LANL's network
- Uses three approaches
  - ~ automated audit record analysis
  - ~ vulnerability testing
  - ~ active probing for signs of misuse
- Processes data in near realtime
- Uses an expert system approach

Target Network

- The Integrated Computing Network (ICN), the main computing network at Los Alamos
- Consists of two separate networks; Open (Unclassified) and Secure (Classified)
  - ~ Approximately 9000 users
  - ~ 5 Cray supercomputers (4 Y-MPs, T3D)
  - ~ Over 10,000 smaller computers and workstations
  - ~ Connects to 5 external networks (e.g,. the Internet)
- Used by both Laboratory employees and others

Goals

- Deterrence
  - ~ increase difficulty in undertaking misuse
  - ~ increase perceived odds of being caught
- Detection
  - ~ discover act of misuse
  - ~ manage investigation
- Accountability
  - ~ trace activities to responsible individuals
  - ~ hold them responsible for their actions
  - ~ collect evidence suitable for prosecution

Functions

- A near realtime method by which to detect a range of security relevant events
  - ~ attempted break-ins to the ICN by outsiders
  - ~ invalid activity or abuses by insiders
- The capability for ad-hoc analysis of past ICN user activity
  - ~ useful for on-going investigations, background examinations, and audits
- Long term maintenance of a record of audit analysis
  - ~ for documenting compliance with DOE security directives

Strategy

- Monitor selected set of critical network systems
- Do not monitor network traffic
- Currently monitors
  - ~ UNICOS Cray supercomputers
  - ~ IBM-based data archiving system (the Common File System)
  - ~ UNIX-based Kerberos (network authentication system)

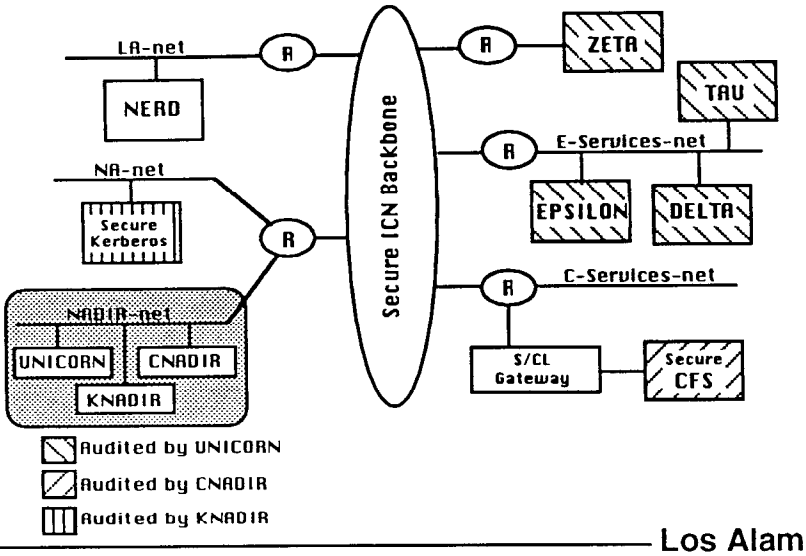
Distributed Design

- Online - for each target system
  - ~ target system-based client
    - pre-process audit data
    - search for signs of misuse and vulnerabilities
    - transmit data to server (push only)
  - ~ workstation-based server
    - summarize target system data into profiles
    - analyze overall system and individual user activity
    - produce reports and alarms
- Offline - investigate anomalous users

Why the Distributed Design?

- Functional protection
- Isolate data analysis and alarm functions from the target systems
- Results in greater level of trust in the detection system
  - ~ less opportunity for tampering by users
- Activity correlation
  - ~ capability to correlate activity from several target systems
  - ~ increased sensitivity to distributed misuse
- Increased security and flexibility is well worth the cost in terms of hardware and software interface development

NADIR in the Secure ICN



Profiles

- Profiles provide a statistical summary of activity on each target system
- Individual user profiles
  - ~ one for each system user
  - ~ activity that can be attributed to that user
- Composite (system) profile
  - ~ one for each system
  - ~ combination of all user activity on the system
  - ~ misuse not attributable to a single user
  - ~ vulnerable configuration information

Event Detection

- Expert rules
  - ~ are applied to profiled data
  - ~ describe interesting behavior
- If behavior is found
  - ~ one or more rules are "triggered"
  - ~ an anomaly score for user or system is set
- Stored for each user and for the whole system
  - ~ anomaly score
  - ~ list of rules triggered

Funding

- The production NADIR has been funded entirely by LANL
  - ~ FSS Division (S&S funding)
  - ~ CIC Division (operational funding)
- Staffing
  - ~ has ranged from 3 to 5 FTEs over the last six years
  - ~ currently 4 FTEs
- Classified extension funded outside LANL

General benefit

- The electronic equivalent to a police officer patrolling a neighborhood, which provides an opportunity to
  - ~ get an overall impression of current conditions
  - ~ spot and evaluate specific problems
  - ~ get to know the neighborhood residents
  - ~ become known in the neighborhood
- Similarly, NADIR
  - ~ provides a summary of network operation
  - ~ points out suspicious users and events
  - ~ creates an opportunity for security officers to meet and talk with users

Specific benefits

- Detects *many* more events than did manual auditing
- These events are detected more quickly
- Follow-up investigations are more timely, systematic, complete, and fully documented
- Event detection and investigation takes fewer personnel
- System has enhanced security awareness in the user community
- Improved understanding of how the network really works
- More effective, and less expensive, response to external audits and requests for special reports

Attack handling

- NADIR compares individual and composite activity to typical or valid activity
- Attacks that require frequent repetition are detected easily
  - ~ by comparing current usage to normal past usage
- It also recognizes violations of computer policies
  - ~ like improper accesses
  - ~ illegal combinations of events
- Second order anomalies, like being repeatedly being "almost interesting", are missed

False positives and negatives

- How many false positives?
  - ~ few enough that they can easily be investigated and eliminated by a half-time investigator
  - ~ getting fewer
  - ~ invested a considerable effort to improve detection accuracy, using automated statistical tuning over a significant period of past usage
- False negatives are hard to prove
  - ~ we do not know of any significant event missed by NADIR (but found by other means) since the current system was implemented

Tuning

- NADIR was designed and tuned for the LANL user population
- We chose NOT to implement self-learning to avoid the potential weaknesses of that method
- We pre-characterize the user population, followed by periodic re-characterizations

Fielding the system

- Normal business constraints limit our ability to do everything we'd like to do
  - ~ i.e., we've never had the funding to all we'd like
- Development/maintenance costs are on-going and seemingly never ending
  - ~ monitored systems constantly change
  - ~ five workstations must be maintained/upgraded etc.
  - ~ resource intensive (3 to 5 developers/administrators)
- Running costs are low
  - ~ 9000 users on eight network systems are monitored
  - ~ with one half-time investigator
- We have a proven, well-functioning system

Further work/research

- Hope to advance the technology through collaboration and research funding
- Interested in expanding to
  - ~ look more at Internet activity
  - ~ develop a characterization of Internet usage
- Investigating other promising detection methodologies that LANL has used for IRS, Social Security, and credit card fraud applications
- Have obtained additional funding
  - ~ one FTE and one post-doc
  - ~ currently hiring

Immunology and Computer Security

STEVEN HOFMEYR (SPEAKER)

ANIL SOMAYAJI

STEPHANIE FORREST

University of New Mexico

Lincoln Laboratory Intrusion Detection Research

RICHARD P. LIPPMAN (SPEAKER)

HAROLD M. HEGGESTAD

MIT Lincoln Laboratory

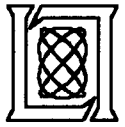
LINCOLN LABORATORY INTRUSION DETECTION RESEARCH

RICHARD P. LIPPMANN
617-981-2711 rpl@sst.ll.mit.edu

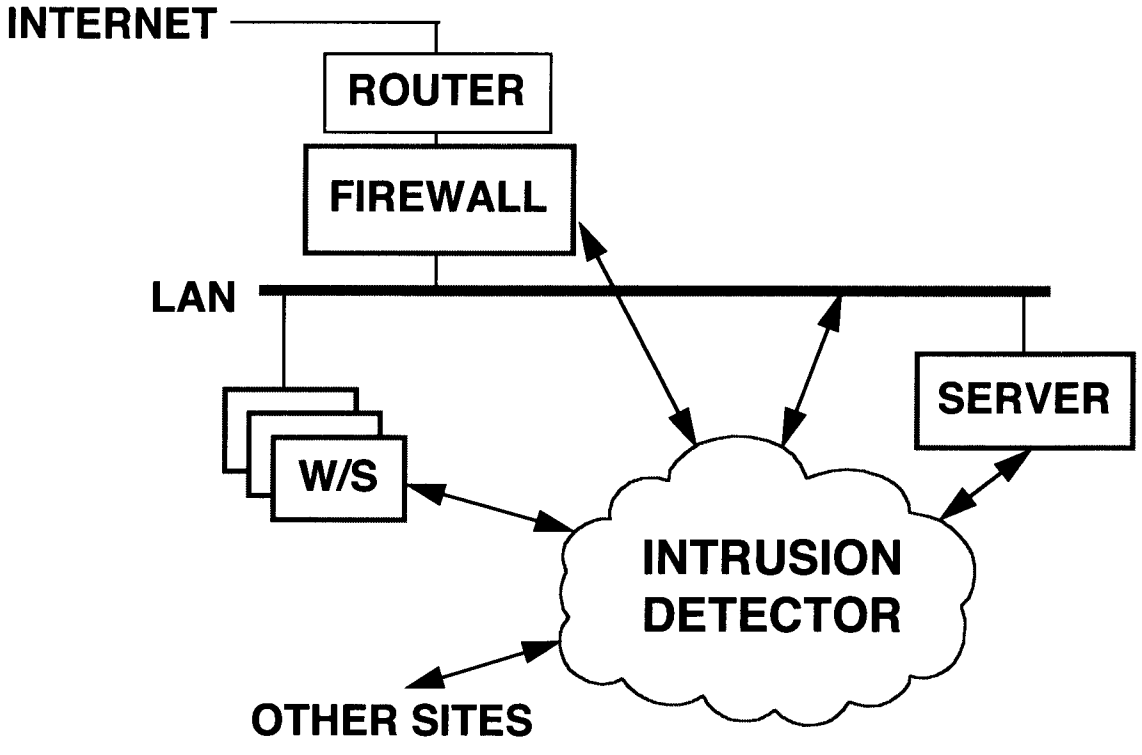
HAROLD M. HEGGESTAD
617-981-4014 hal@xn.ll.mit.edu

MIT LINCOLN LABORATORY
LEXINGTON, MA 02173

Presented at AFIWC
23 October 1996, San Antonio TX



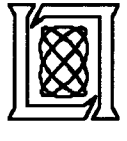
A TEST AND EVALUATION ENVIRONMENT IS REQUIRED TO VERIFY THE PERFORMANCE OF INTRUSION DETECTION SYSTEMS



- FEW OBJECTIVE COMPARISONS BETWEEN SYSTEMS
- FEW OPERATIONAL PERFORMANCE ANALYSES
- NO STANDARD COMPARISON METRICS
- FEW STANDARD INTERFACES
- FEW MODERN SYSTEMS IN OPERATIONAL USE

GOALS OF TEST AND EVALUATION WORK

- DEVISE OBJECTIVE APPROACH TO EVALUATE NEW INTRUSION DETECTION SYSTEMS
  - R&D RESULTS ARE DIVERSE AND INCOMMENSURABLE
  - HARD TO ASSESS SUITABILITY FOR DEPLOYMENT
- FOSTER INTEGRATION OF COMPLEMENTARY ID TECHNOLOGIES
  - IDENTIFY MUTUALLY SUPPORTIVE IDEAS
  - PERFORM EVALUATIONS AND ANALYSES
- EXPEDITE MIGRATION OF NEW TECHNOLOGIES INTO OPERATIONAL ID TOOLKITS
  - PROVIDE BRIDGE BETWEEN RESEARCH AND OPERATIONS
  - PERFORM TECHNOLOGY INSERTION AND DEMONSTRATION



APPROACH

- PERFORM UNBIASED COMPARISONS OF RESEARCH SYSTEMS
- DEVELOP AND APPLY STANDARD METRICS AND INTERFACES
- TEST IN REALISTIC GOVERNMENT APPLICATIONS WITH VARIED TYPES OF ATTACKS AND MISUSE MODELS
- CONTINUALLY INTERACT WITH THE RESEARCH COMMUNITY
- TRANSITION TO REALISTIC OPERATIONAL ENVIRONMENTS

TECHNICAL APPROACH

STEP 1: IMPLEMENT A TEST ENVIRONMENT

- ARCHITECTURE
- PERFORMANCE METRICS
- TEST DATA SET COLLECTION EXAMPLE
- OBTAIN BASELINE PERFORMANCE OF OPERATIONAL SYSTEM

STEP 2: TEST A SINGLE-SITE R&D SYSTEM

STEP 3: TEST ADDITIONAL SINGLE-SITE R&D SYSTEMS

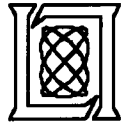
STEP 4: TEST MULTI-SITE R&D SYSTEMS

•ONGOING:

- FORM AND CHAIR A WORKING GROUP
- TEST ADDITIONAL SYSTEMS

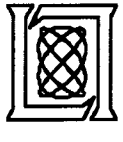
•LONGER-TERM GOALS:

- INSTALL R&D PRODUCTS IN GOVERNMENT APPLICATIONS
- TRANSITION THE TEST AND EVALUATION ENVIRONMENT TO AN OPERATIONAL NATIONAL ASSET

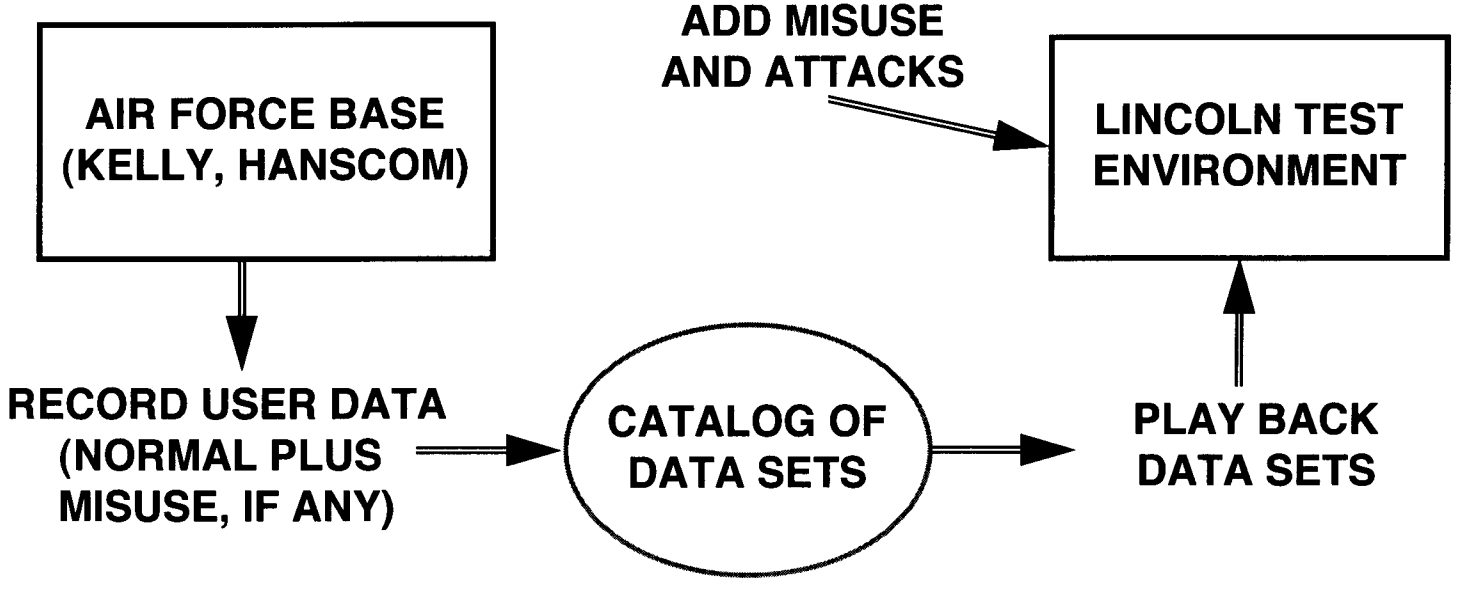


TECHNICAL APPROACH, STEP 1: IMPLEMENT A BASELINE TEST ENVIRONMENT USING CURRENT TECHNOLOGY

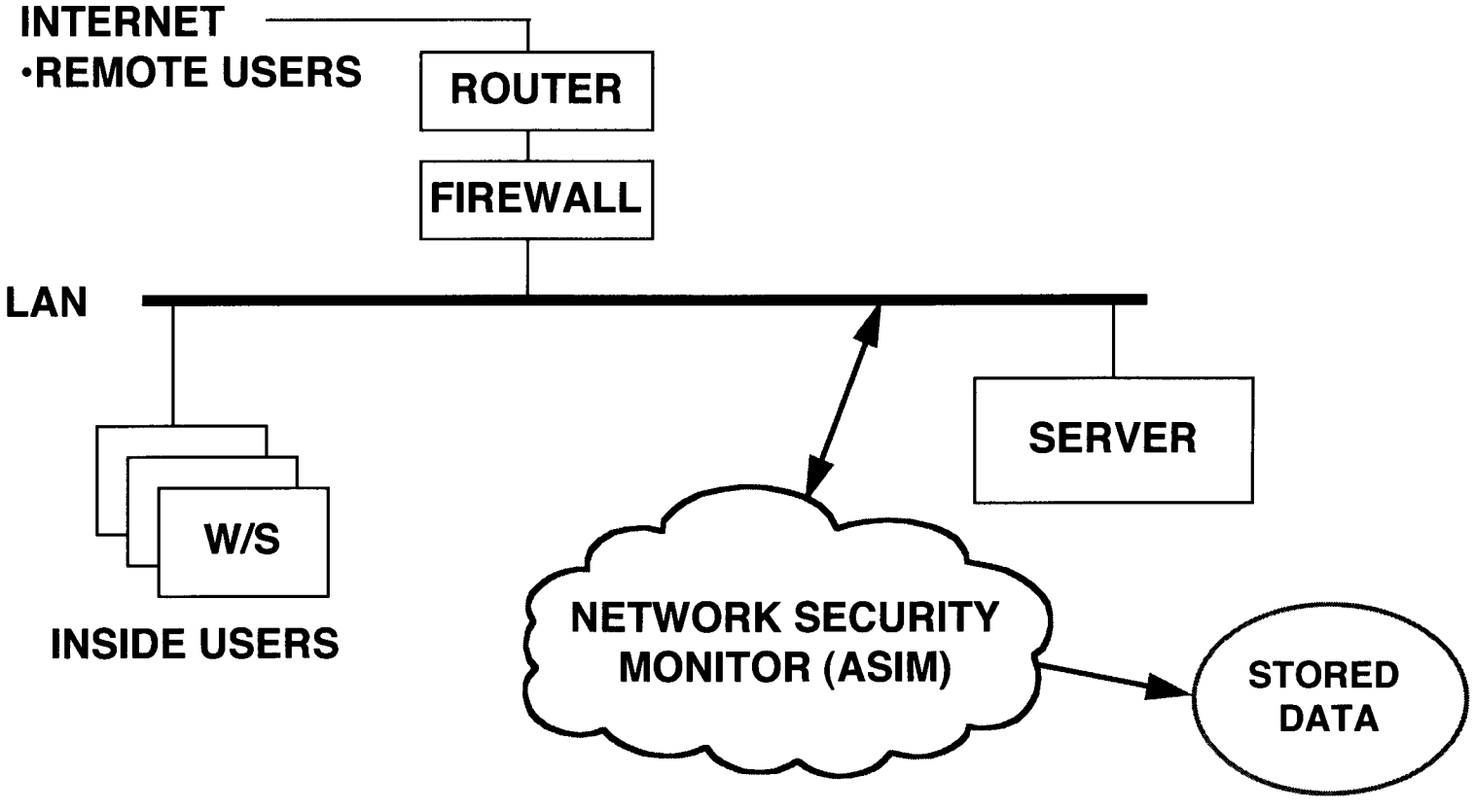
- CREATE A TEST ENVIRONMENT AT LINCOLN LABORATORY
- USE AN EXISTING INTRUSION DETECTION TOOL (ASIM)
- BRING UP ASIM IN THE TEST ENVIRONMENT
  - EXPERIMENT WITH ITS FUNCTIONS AND CONTROLS
  - FIX ANY INTERFACING PROBLEMS
- APPLY RECORDED DATA FROM OPERATIONAL SITE
- DEVELOP AND APPLY ATTACK AND MISUSE MODELS
- EVALUATE BASELINE PERFORMANCE



DATA COLLECTION FROM OPERATIONAL AIR FORCE BASES

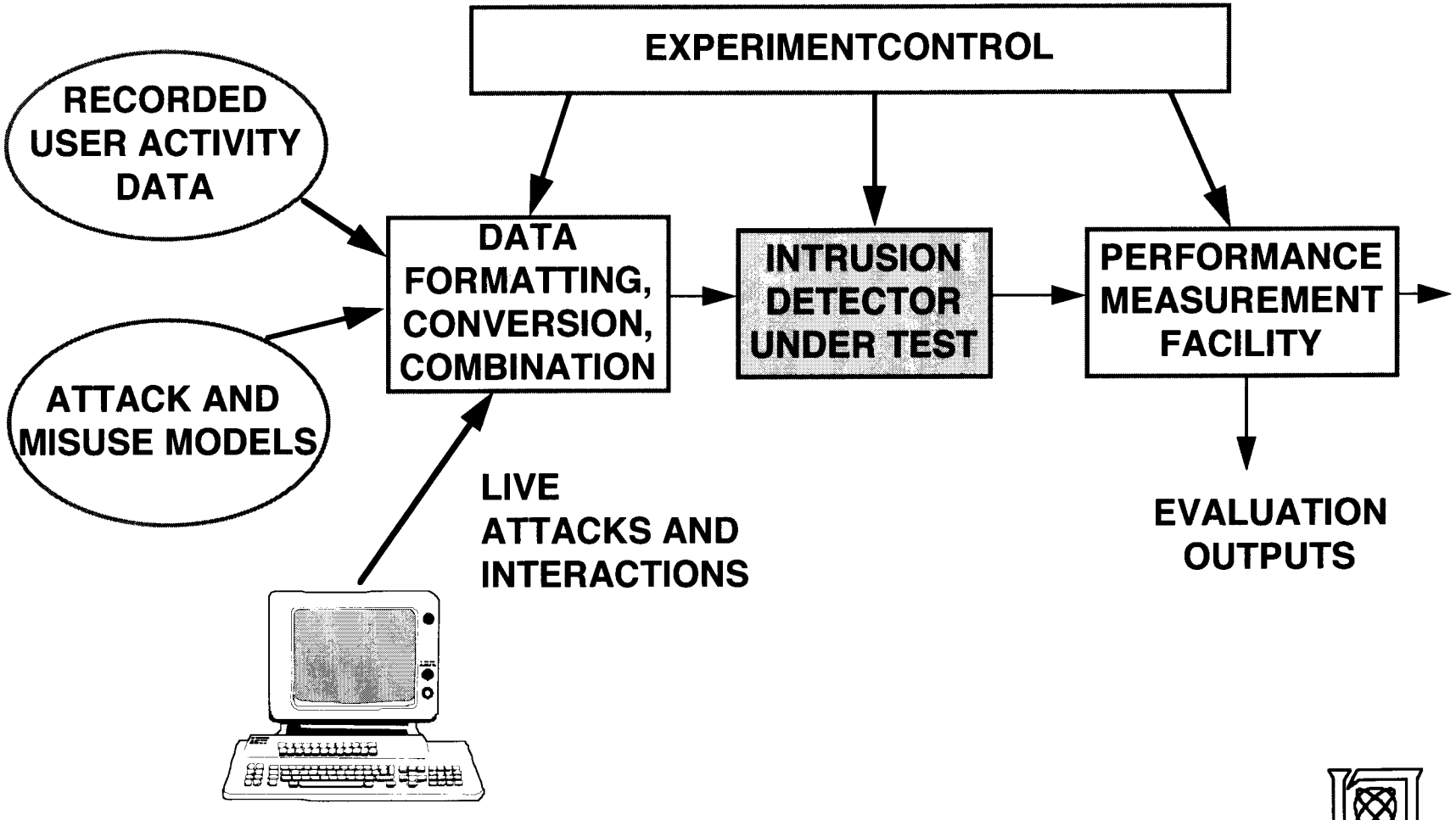


ASIM INTRUSION DETECTION ENVIRONMENT ON AIR FORCE BASES



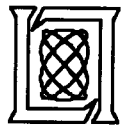
•ASIM EXAMINES ALL TCP/IP PACKETS FROM LOCAL TO REMOTE SITES
•STORES PACKET INFORMATION AND CONTENTS

LOCAL LINCOLN/ROME TEST ENVIRONMENT

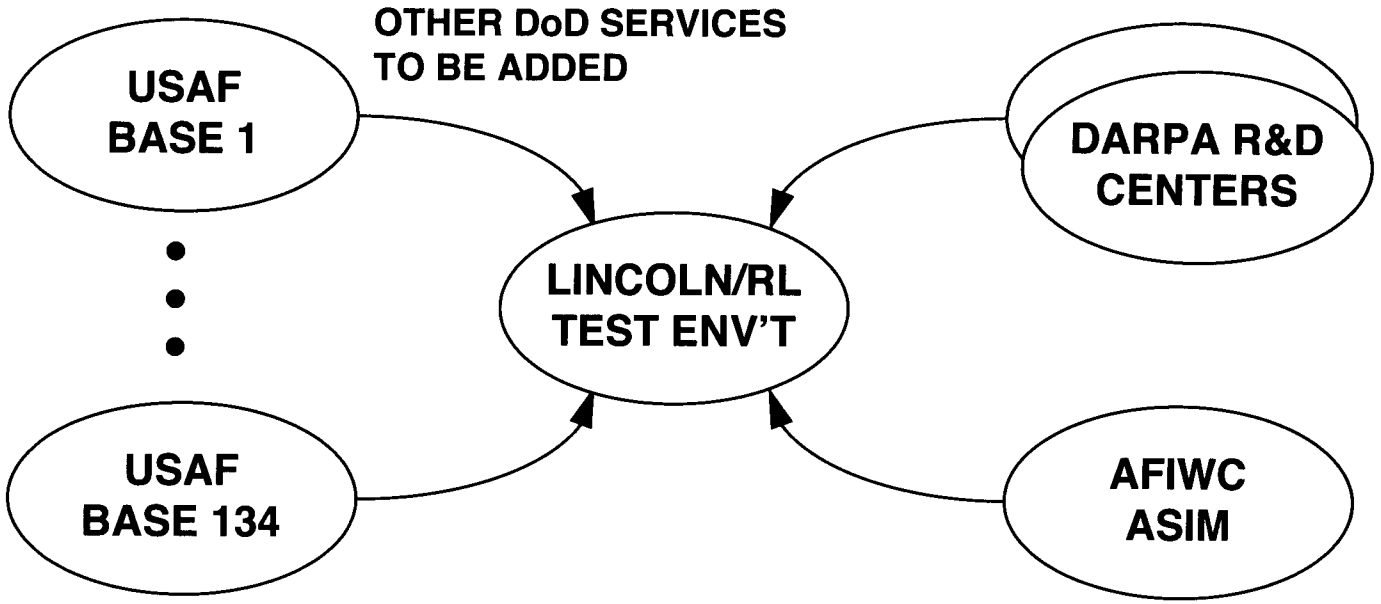


ATTACK AND MISUSE MODELS

- SOURCES OF ATTACKS
  - INCIDENTAL EVENTS IN NORMAL DATA
  - COMPUTER SECURITY ASSESSMENT TEAMS
  - DARPA R&D CONTRACTORS
  - RESEARCH AND COMMERCIAL SCANNERS (COPS, SATAN, Internet Security Systems Internet Scanner)
- GENERATING NEW ATTACKS
  - NEW REAL ATTACKS CAN BE ADDED DURING PROGRAM
  - PRESENT HISTORICAL SEQUENCE (CERT Advisories) OF ATTACKS, DISABLE ATTACK-SPECIFIC RULES
- SOURCES OF MISUSE
  - AIR FORCE MONITORS AND SYSTEM ADMINISTRATORS
  - SIMPLE BASELINE (Swap Users, Move Users Between Groups)



TEST ENVIRONMENT AND DATA SOURCE RELATIONSHIPS



- NORMAL OPERATION
- INCIDENTAL ATTACKS AND FAULTS
- EXERCISES
- RED TEAM ATTACKS

DATA BASE ISSUES

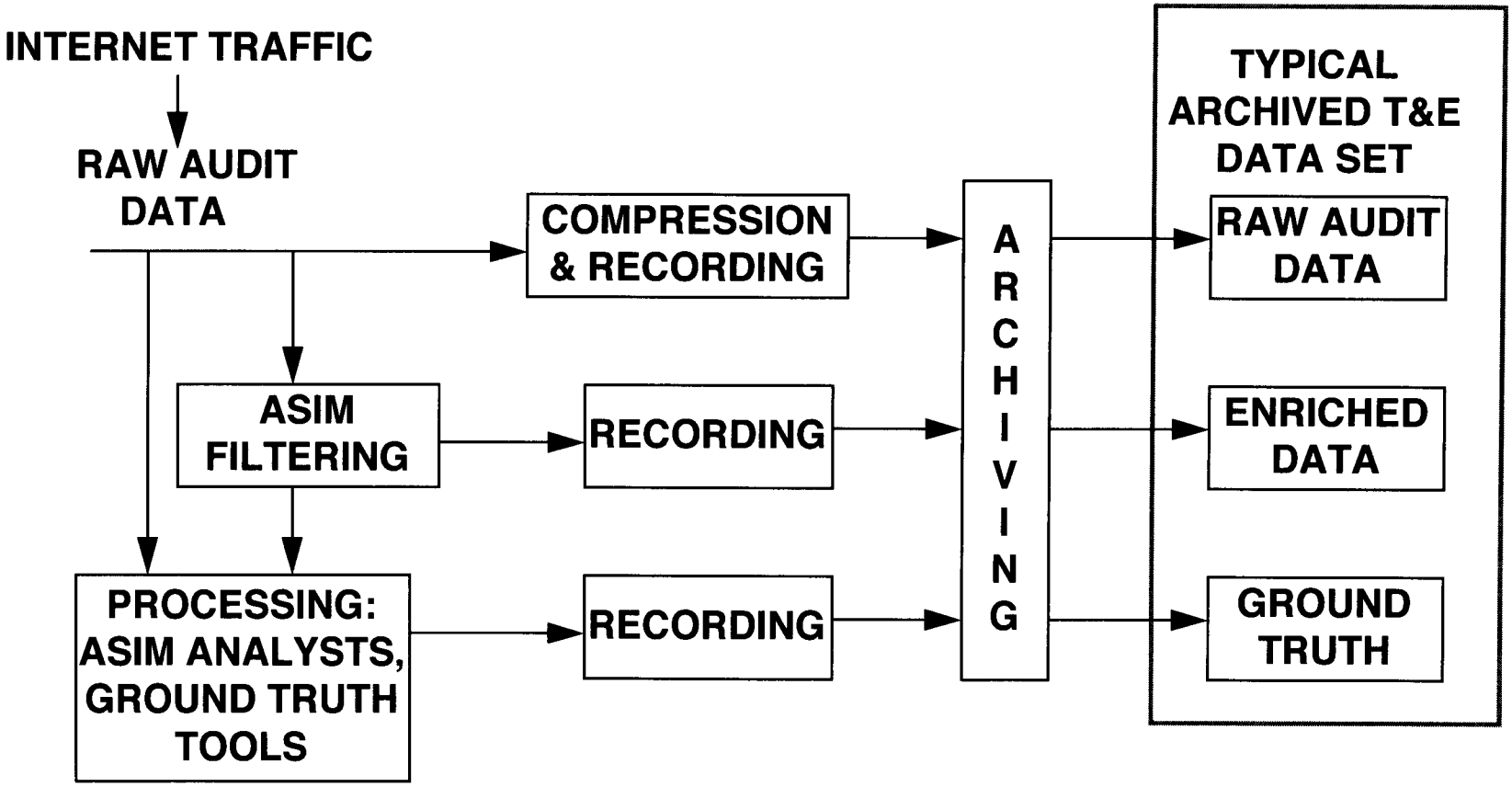
- VALIDITY OF SAMPLING (Location, Date/Time, Activities, System, System Load, System Configuration)
- OBTAINING GROUND TRUTH (Are Attacks or Misuse Hidden in the Data?)
- SELECTING TRAINING AND TEST DATA
- STATISTICAL SIGNIFICANCE OF RESULTS (Attacks and Misuse are Infrequent)
- TYPES AND FREQUENCY OF OCCURRENCE OF ATTACKS

POTENTIAL PERFORMANCE METRICS

- DETECTION PROBABILITY AND FALSE ALARM RATE (KNOWN AND NEW ATTACKS)
- RESOURCE UTILIZATION BY DETECTOR
  - CPU, MEMORY, FILE SIZE, NETWORK LOAD
- LATENCY OF DETECTION
- VALIDITY OF DIAGNOSES AND RECOMMENDED ACTIONS

- EASE OF EXTENSION TO DETECT NEW ATTACKS
- PORTABILITY, EASE AND COST OF INSTALLATION
- QUALITY OF TOOLS FOR INFORMATION REPRESENTATION AND EVALUATION
- WORKLOAD AND EFFICIENCY LEVERAGE

TEST DATA SET COLLECTION EXAMPLE: AIR FORCE SITE MONITORED BY ASIM



TECHNICAL APPROACH

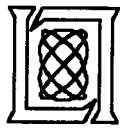
STEP 1: IMPLEMENT A TEST ENVIRONMENT

- ARCHITECTURE
- PERFORMANCE METRICS
- TEST DATA SET COLLECTION EXAMPLE
- OBTAIN BASELINE PERFORMANCE OF OPERATIONAL SYSTEM

STEP 2: TEST A SINGLE-SITE R&D SYSTEM

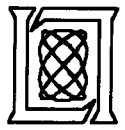
STEP 3: TEST ADDITIONAL SINGLE-SITE R&D SYSTEMS

STEP 4: TEST MULTI-SITE R&D SYSTEMS

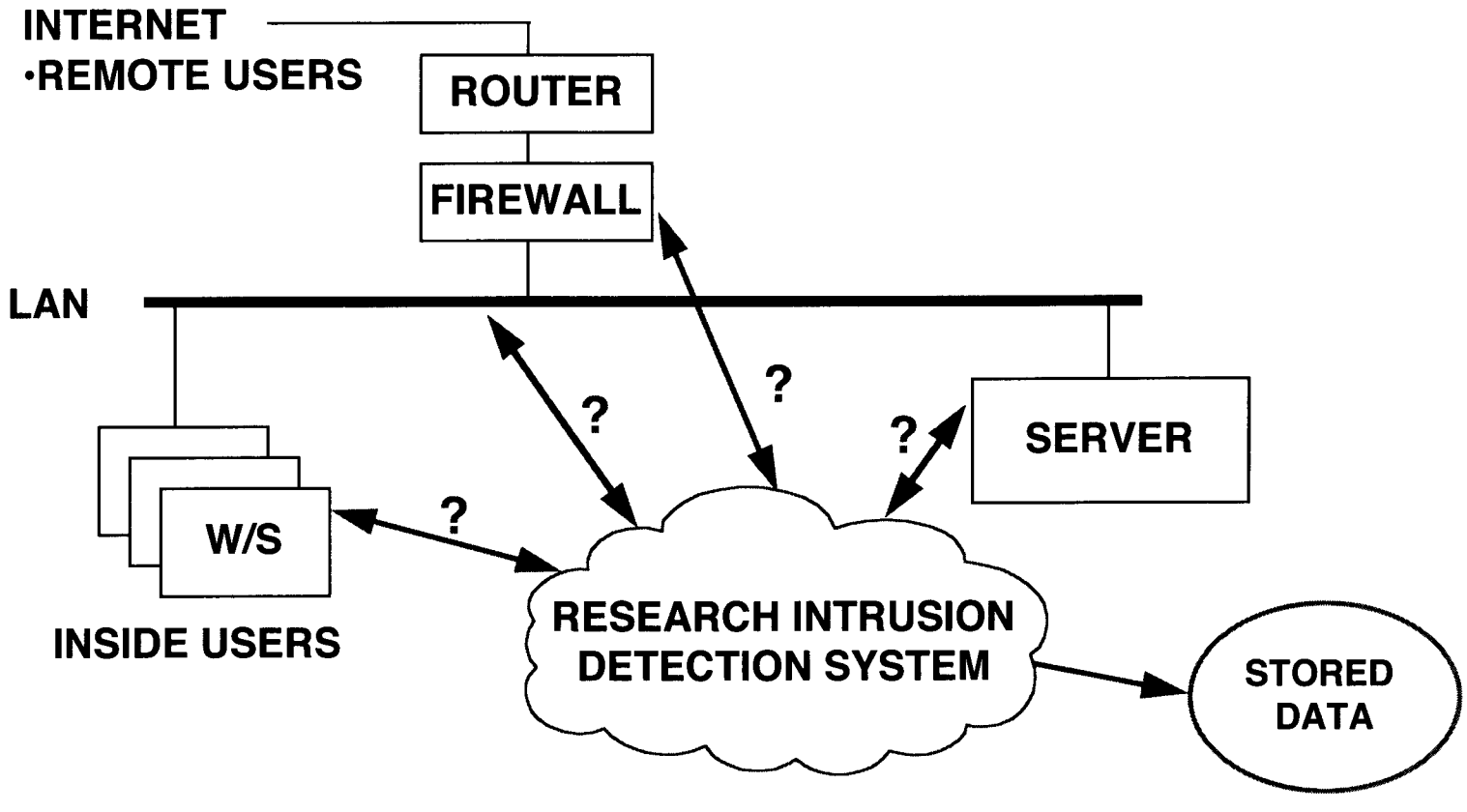


TECHNICAL APPROACH, STEP 2: TEST AN INTRUSION DETECTION R&D PRODUCT

- SELECT A SUITABLE SYSTEM FROM THE R&D COMMUNITY
- CUSTOMIZE FACILITIES IN THE LOCAL TEST ENVIRONMENT
  - DATA FORMATTING
  - PERFORMANCE MEASUREMENT
- MODIFY AF BASE DATA COLLECTION AS NECESSARY
- BRING UP THE SYSTEM TO BE TESTED AT LINCOLN
  - EXPERIMENT WITH ITS FUNCTIONS AND CONTROLS
  - FIX ANY INTERFACING PROBLEMS
- APPLY RECORDED DATA FROM OPERATIONAL SITE
- APPLY ATTACKS AND VARIOUS MISUSE MODELS
- EVALUATE PERFORMANCE AND COMPARE TO BASELINE

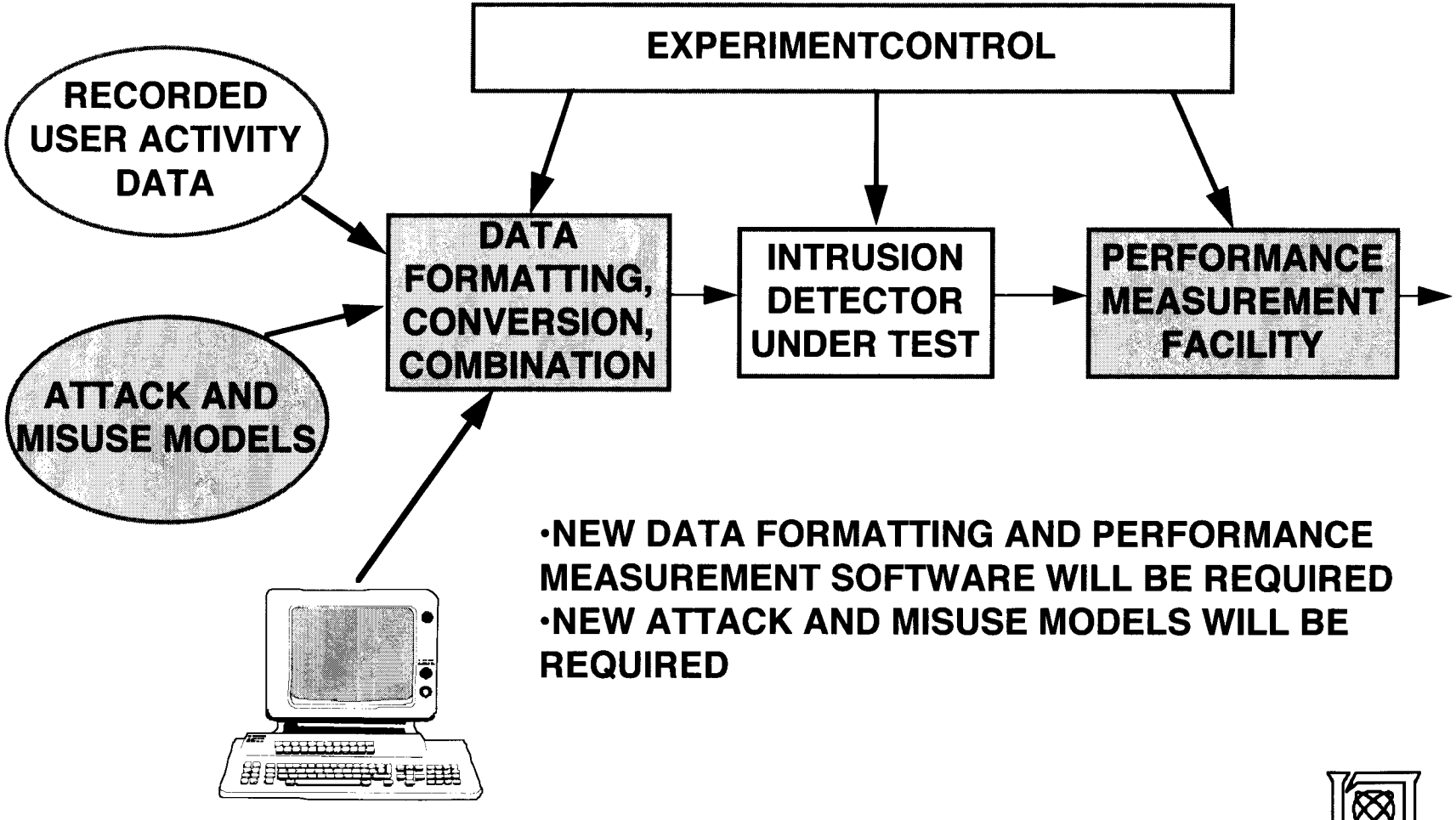


INSTALLING A RESEARCH INTRUSION DETECTION SYSTEM ON AIR FORCE BASES



•NEW SOFTWARE WILL HAVE TO BE INSTALLED IN WORKSTATIONS, FIREWALL, AND/OR SERVERS TO OBTAIN DATA

STEP 2 EXTENSIONS REQUIRED FOR LOCAL LINCOLN/ROME TEST ENVIRONMENT



- NEW DATA FORMATTING AND PERFORMANCE MEASUREMENT SOFTWARE WILL BE REQUIRED
- NEW ATTACK AND MISUSE MODELS WILL BE REQUIRED

TECHNICAL APPROACH

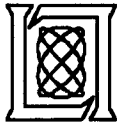
STEP 1: IMPLEMENT A TEST ENVIRONMENT

- **ARCHITECTURE**
- **PERFORMANCE METRICS**
- **TEST DATA SET COLLECTION EXAMPLE**
- **OBTAIN BASELINE PERFORMANCE OF OPERATIONAL SYSTEM**

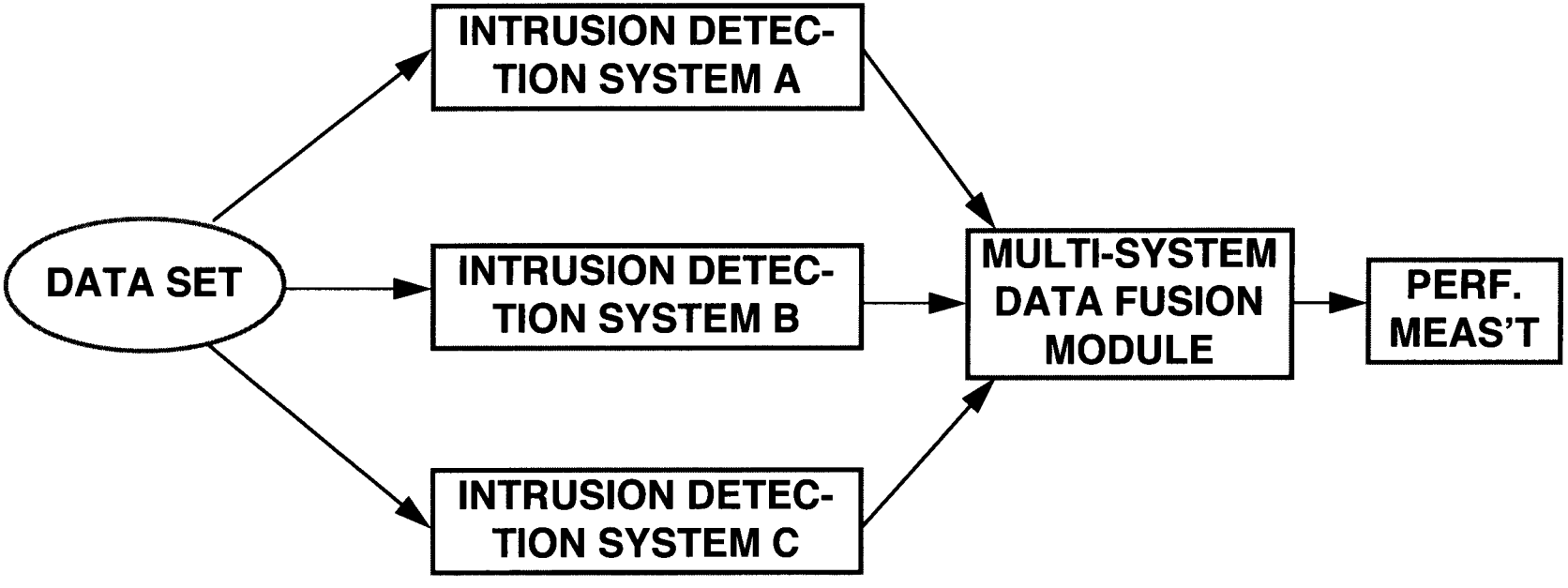
STEP 2: TEST A SINGLE-SITE R&D SYSTEM

STEP 3: TEST ADDITIONAL SINGLE-SITE R&D SYSTEMS

STEP 4: TEST MULTI-SITE R&D SYSTEMS



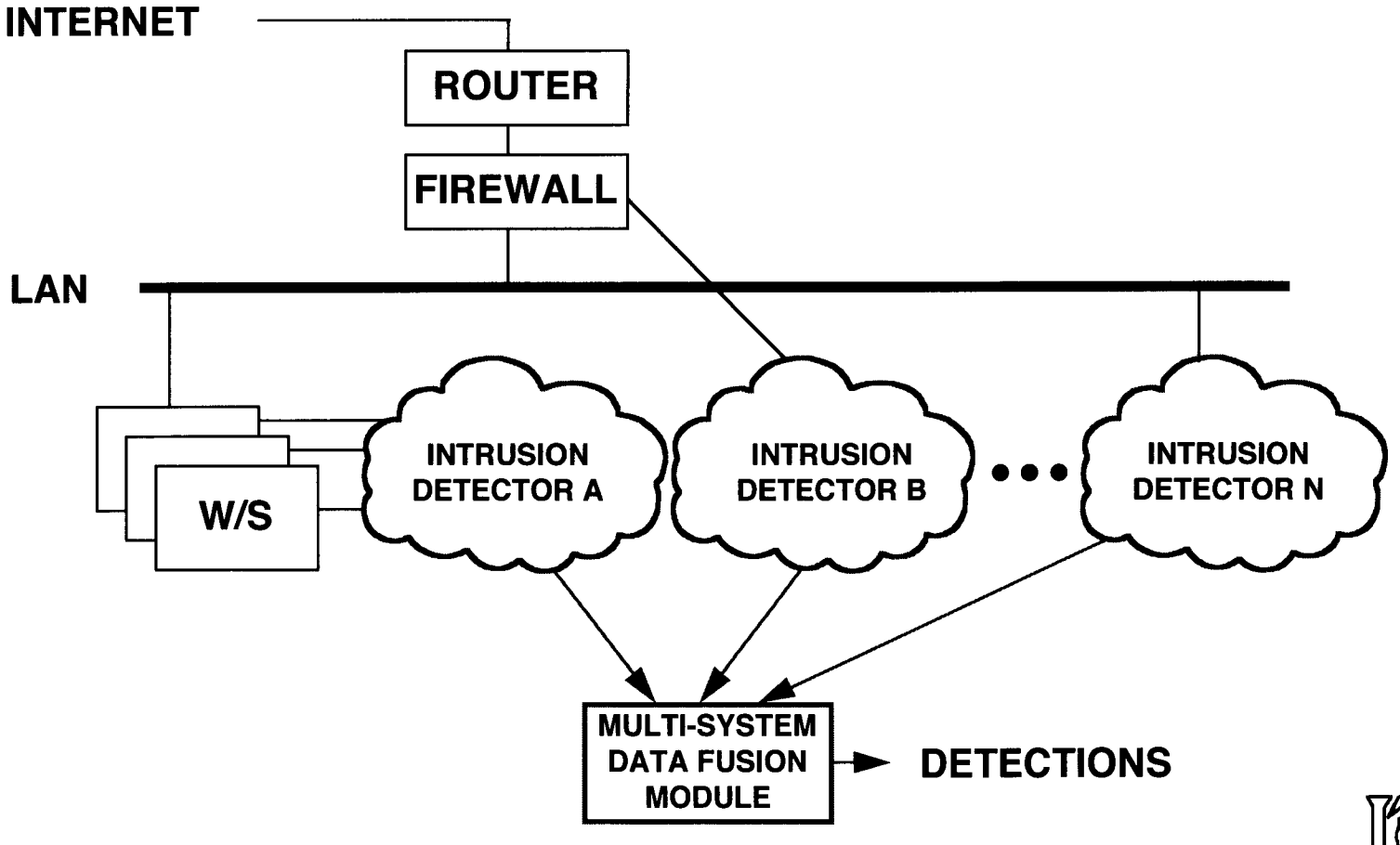
TECHNICAL APPROACH, STEP 3: TEST ENVIRONMENT FOR COMBINATIONS OF INTRUSION DETECTION SYSTEMS



GOALS:

- COMPARE APPROACHES ON IDENTICAL DATA SETS
- FIND MOST EFFECTIVE DETECTION INPUT MEASURES AND ALGORITHMS
- COMBINE TO PROVIDE IMPROVED PERFORMANCE AT LOWER OPERATIONS COST

INTEGRATED INTRUSION DETECTION SYSTEM ENVIRONMENT (FOR STEP 3)



TECHNICAL APPROACH

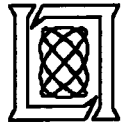
STEP 1: IMPLEMENT A TEST ENVIRONMENT

- **ARCHITECTURE**
- **PERFORMANCE METRICS**
- **TEST DATA SET COLLECTION EXAMPLE**
- **OBTAIN BASELINE PERFORMANCE OF OPERATIONAL SYSTEM**

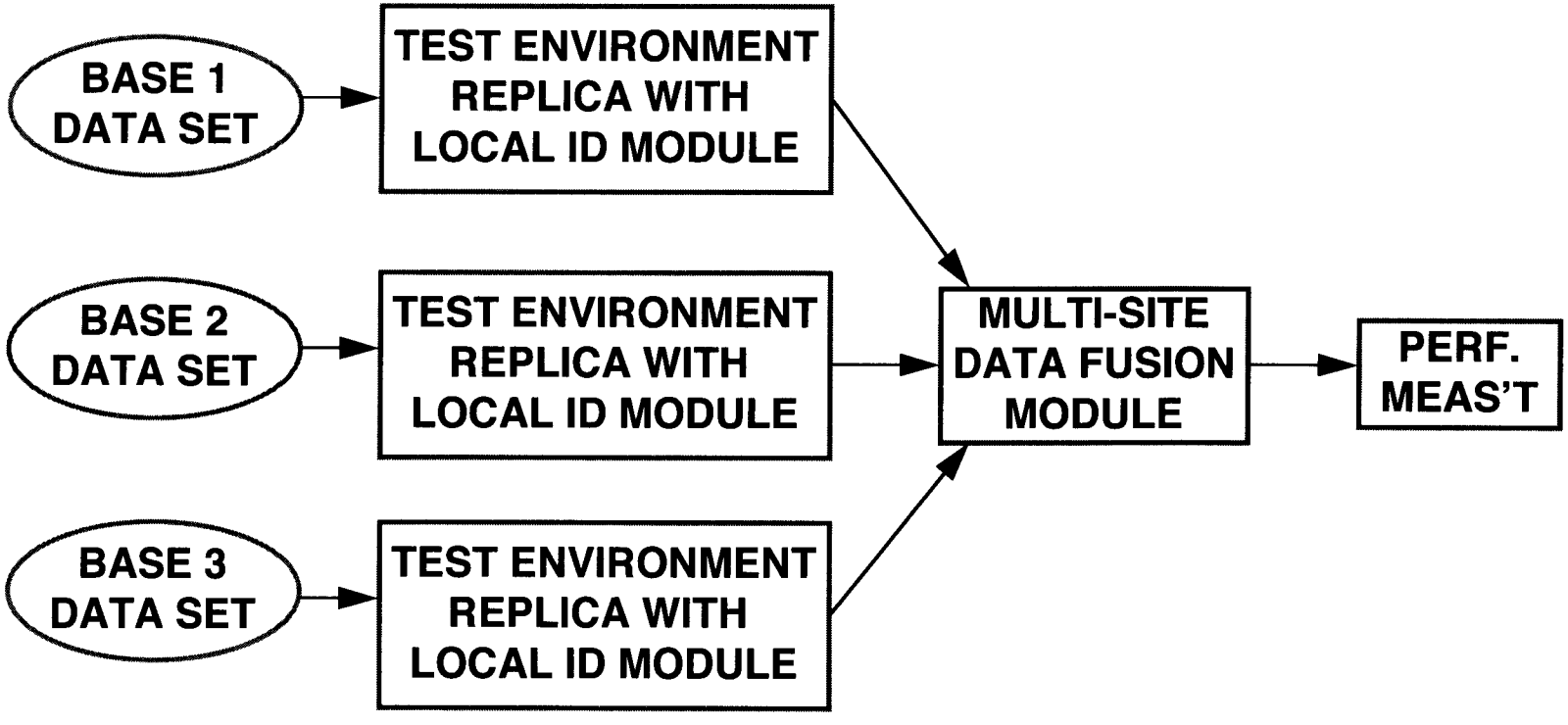
STEP 2: TEST A SINGLE-SITE R&D SYSTEM

STEP 3: TEST ADDITIONAL SINGLE-SITE R&D SYSTEMS

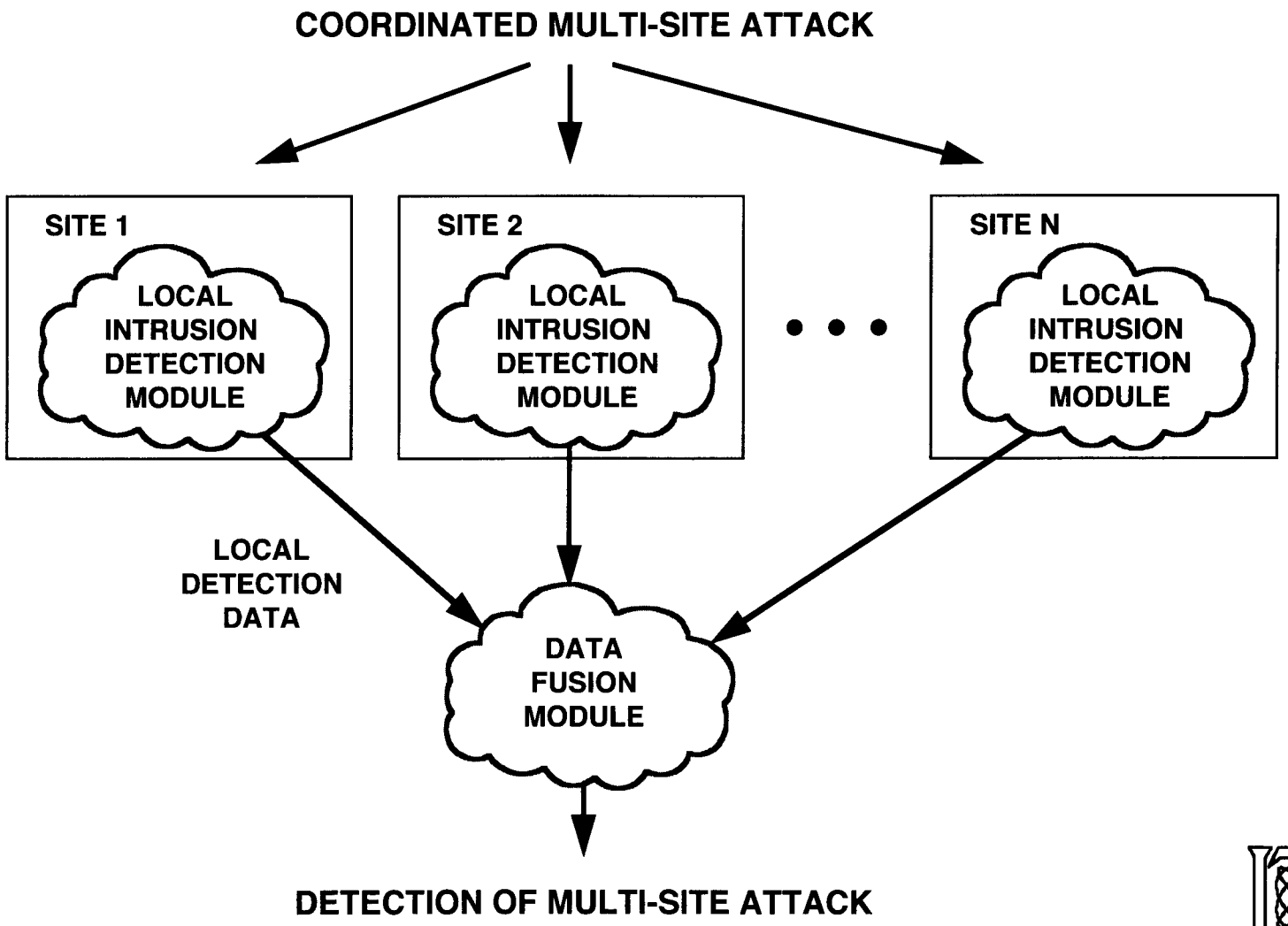
STEP 4: TEST MULTI-SITE R&D SYSTEMS



TECHNICAL APPROACH, STEP 4: TEST ENVIRONMENT FOR MULTI-SITE INTRUSION DETECTION SYSTEMS

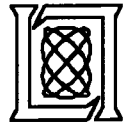


MULTI-SITE ATTACK ENVIRONMENT (FOR STEP 4)



A LARGE REAL CONNECTION DATA BASE IS REQUIRED TO EVALUATE ASIM (NSM)

- SELECT A FEW REPRESENTATIVE BASES (e.g. Wright Patterson, Hanscom, ...)
- OBTAIN SIX MONTHS OF DATA
  - RAW SNIFFED PACKET LOGS STORED ON BASE
  - CONNECTION SCORES STORED AT AFIWC
  - HIGH-SCORING CONNECTION TRANSCRIPTS STORED AT AFIWC
  - INCIDENTREPORTS ISSUED FROM AFIWC
  - INFORMATION ABOUT RED-TEAM AND BASE EVALUATION ACTIVITIES
- STORE DATA AT LINCOLN TO PLAY BACK AND EVALUATE INTRUSION DETECTION SYSTEMS
  - USE ON LOCAL NET WITH NO EXTERNAL CONNECTIONS
  - INSIDE BUILDING THAT REQUIRES CLEARANCE TO ENTER



OR TEST AND MENT

C (San Antonio)

OF ALL CTIONS

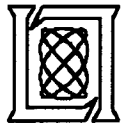
3. TRANSCRIPTS FOR TOP SCORING CONNECTIONS

DENCE ORTS

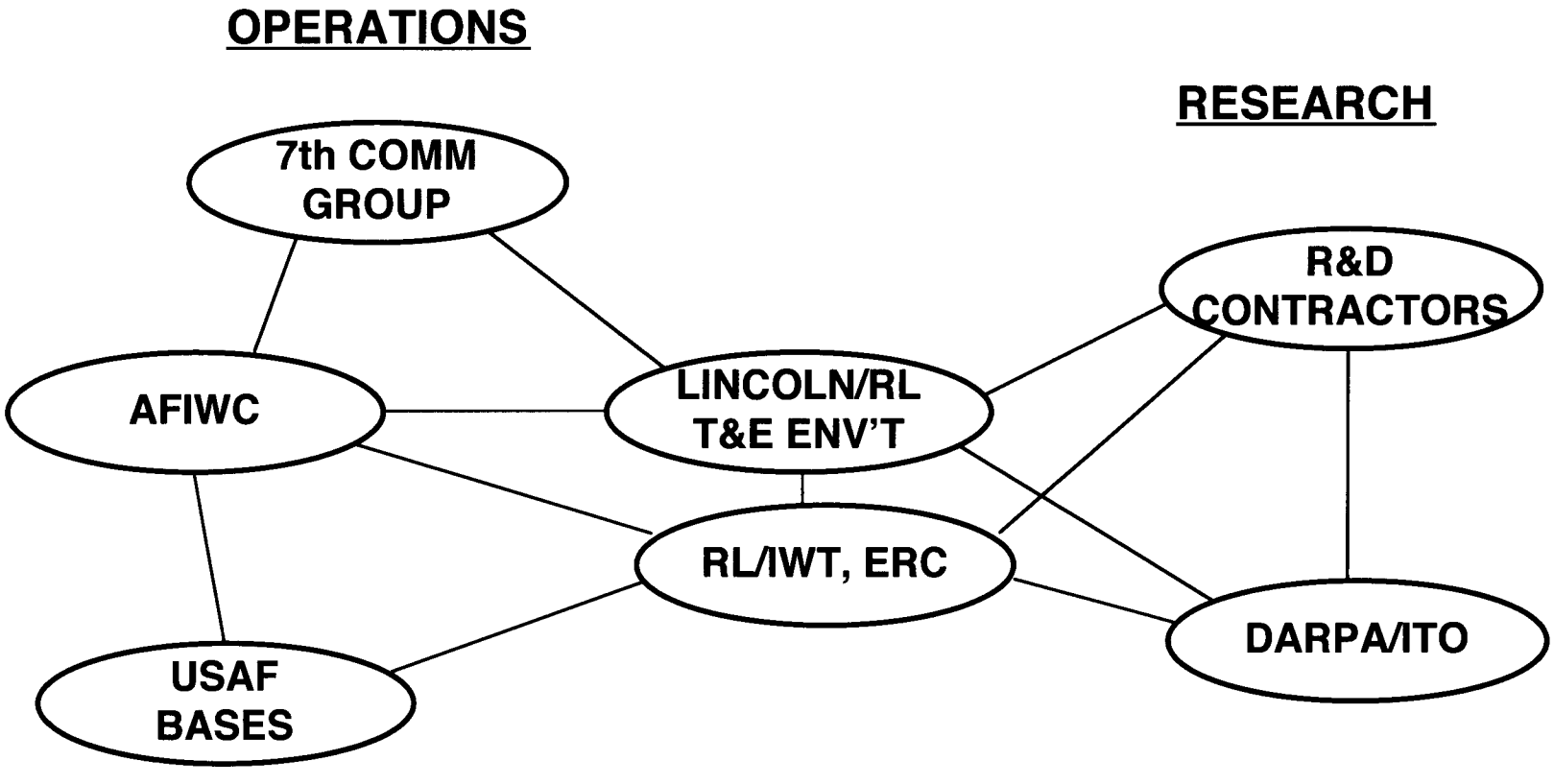
5. ATTACK SCRIPTS

BASE LUATION ATES

D USED TO ND ALSO OTHER

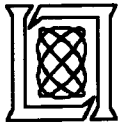


KEY PARTICIPANTS



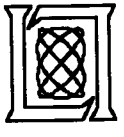
NEAR-TERM ACTIVITIES

- PROCEED WITH TECHNICAL APPROACH, STEP 1
  - IMPLEMENT THE TEST ENVIRONMENT FOR ASIM
  - COLLECT DATA SETS AND GROUND TRUTH
  - GENERATE MISUSE AND ATTACK MODELS
  - PERFORM EVALUATIONS
- PROVIDE UPDATES TO THE R&D COMMUNITY
  - TWO-WAY FLOW OF ADVICE AND PROGRESS REPORTS
  - ANALYSIS AND EVALUATION REPORTS
  - PLANNING OF STEP 2 AND BEYOND
- FORM AND CHAIR A WORKING GROUP
  - DEFINE TEST AND EVALUATION METHODOLOGY
  - DEFINE THE TEST ENVIRONMENT AND PERFORMANCE METRICS



SUMMARY OF TEST EVALUATION WORK

- LINCOLN AND ROME LABORATORIES ARE DEVELOPING AN ENVIRONMENT TO EVALUATE INTRUSION DETECTION SYSTEMS
  - UNBIASED EVALUATION
  - MODEL ACTUAL GOVERNMENT OPERATIONS
  - ACTUAL ATTACK AND MISUSE MODELS
  - OBJECTIVE EVALUATIONS
- INITIAL BASELINE WORK WILL USE ASIM SOFTWARE
- RESEARCH SYSTEMS WILL THEN BE EVALUATED



ATM Firewall Technology: Lessons for Intrusion Detection

CHRISTOPH L. SCHUBA

COAST Laboratory

Purdue University

ATM Firewall Technology: Lessons for Intrusion Detection

Workshop on Computer Misuse and Anomaly Detection (CMAD) IV
Monterey, CA

November 12-14, 1996

Christoph L. Schuba

Purdue University
COAST Laboratory
1398 Department of Computer Sciences
West Lafayette, IN 47907-1398

schuba@cs.purdue.edu

Overview

Problems

ATM Firewall Technology

Lessons

Problems

Quality of Audit Data in Large Systems

- Level of detail vs. amount of data:

  >compression, reduction/aggregation, deduction

- Context of data:

  >users, connections, actions,. ..

- Value of data:

  > authenticity, integrity

E.g., IP, ATM addresses (low level access, e.g., /dev/ip)

Integration of Intrusion Detection and System Design

- Design of large scale distributed systems is *hard*
- Getting designers to include security is *harder*
- Adding intrusion detection support mechanisms is _____

ATM Firewall Technology

Goal

Develop Model for ATM Firewall Technology

Instantiation of Model (Implementation):

- Proof of concept
- Gaining practical experiences

Background and Definitions

Definition Firewall Technology:

Mechanism to help enforce access policies about communication traffic entering or leaving networks.

ATM Technology

- Developed for use in B-ISDN
- Switching of small fixed-length packets (cells)
- Pt-to-pt, pt-to-mpt communication
- Connection-oriented

- permanent connections: administrative mechanisms
- switched connections: connection establishment protocol

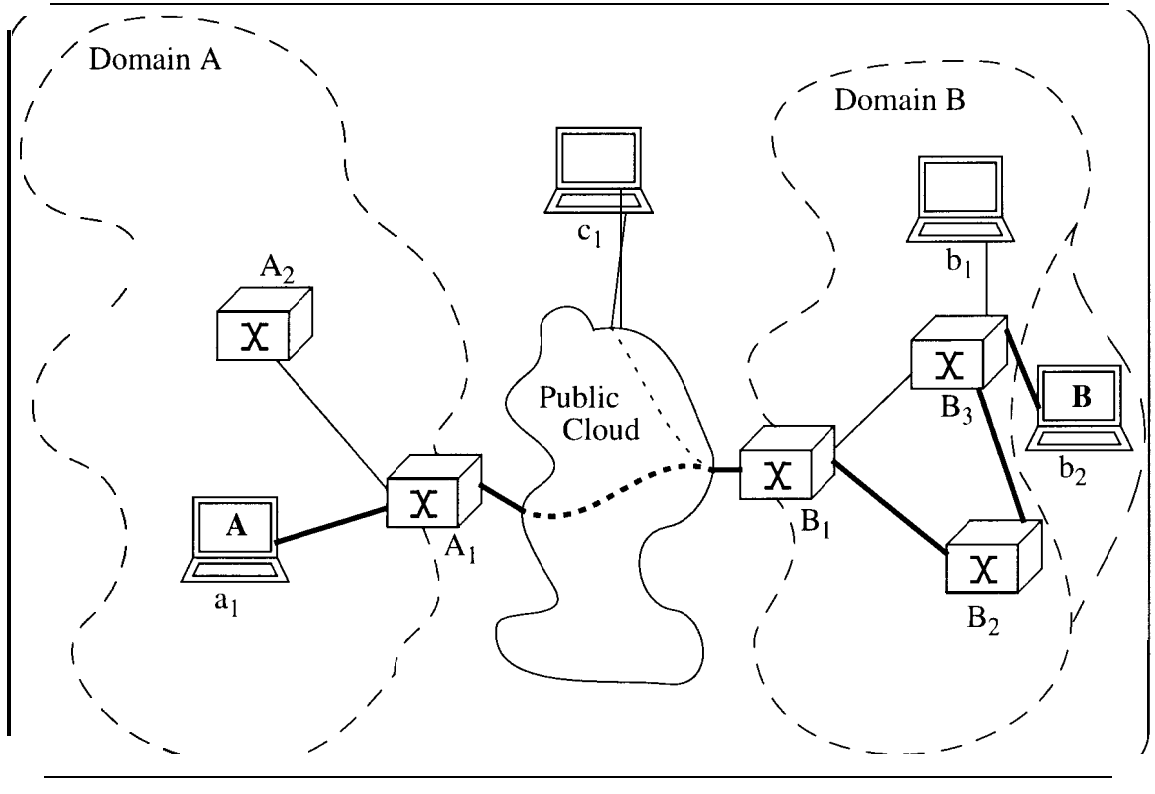
- Quality of service guarantees

IP over ATM

Interesting case for the purpose of this workshop session:

- ATM: spans local-wide area networks systems
- ATM: still room for standard improvement
- IP: legacy system baggage

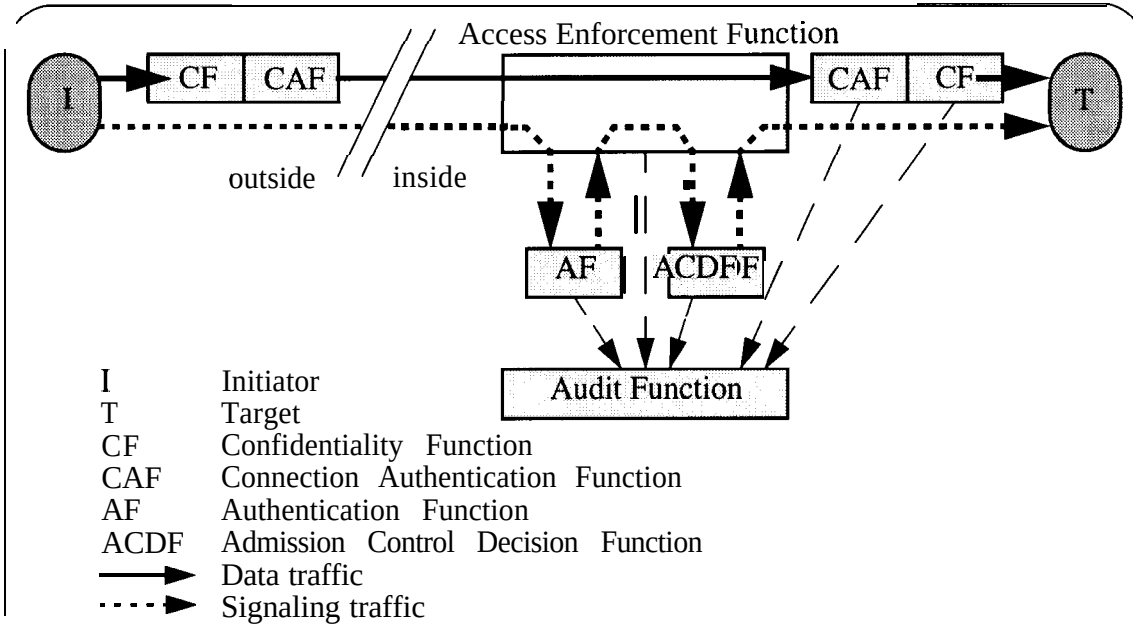
Example



Assumptions

- Connection oriented character of communication
- Secure public key infrastructure, name service
- Secure binding between principals and keys
- Integrity of trusted computing base
- Strength of cryptographic algorithms

Reference Model



Essential Elements

- Endpoint authentication
- Domain based call admission control
- Connection authentication (authenticity and integrity)
- Audit
- Centralized policy with distributed service and enforcement

Contributions

- Concept of firewall technology is viable in connection-oriented highspeed networks
- Five elements are essential for a reference model of firewall technology
- Few additions to signaling protocol and system are necessary and sufficient for implementation

Lessons

(Quality of Audit Data)

1.) Authenticity

- Lack of authenticity - see ATM firewall architecture
- Context establishment problem - security context
- Level of detail - e.g., information elements

(Integration of ID and System Design)

2.) Functional Dependencies

Between *authentication* and *access control*

Between *audit* and *all other security services!*

Now, who *acts* accordingly?

3.) Prevention vs. Detection/Recovery

There should be no tension between *prevention* and *detection*

There should be an *integrated approach*, where

- Preventive mechanisms operate under the assumption that they will fail in certain circumstances
- Preventive mechanisms should provide as much help for detection mechanisms as possible

4.) **Intrusion Detection List of Mechanisms**

What basic *mechanisms* are necessary (e.g., audit; secure, reliable communication)?

Make certain this list becomes second nature for system designers.

5.) Motivation for Businesses

Leverage off advantages for other industries

- Telecommunication carriers want nonpudiable billing information
- Identical mechanisms required for billing and ID

Pay close attention to justifying our case not for the sake of ID alone, but also different business needs that can be fulfilled.

Denial-of-Service Attacks

SIMSON GARFINKLE

Practical UNIX & Internet Security

Denial-of-Service Attacks

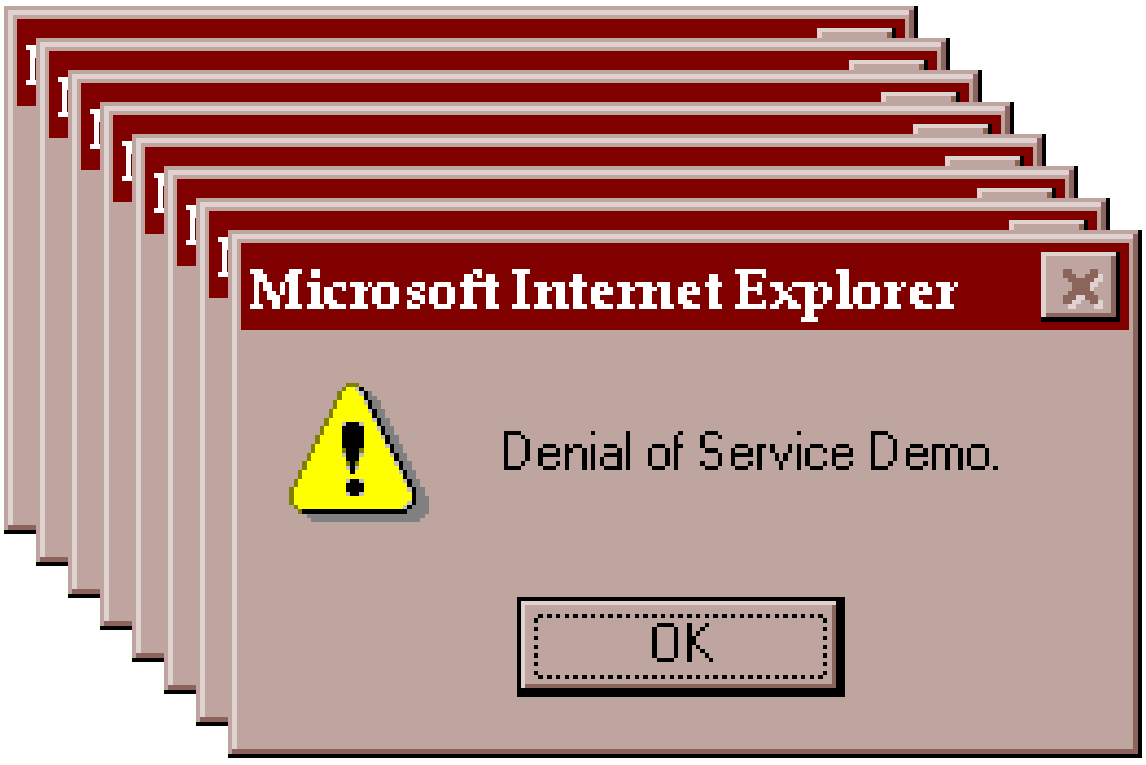
- HTML
- JavaScript
- ActiveX
- Programs & Attachments

HTML

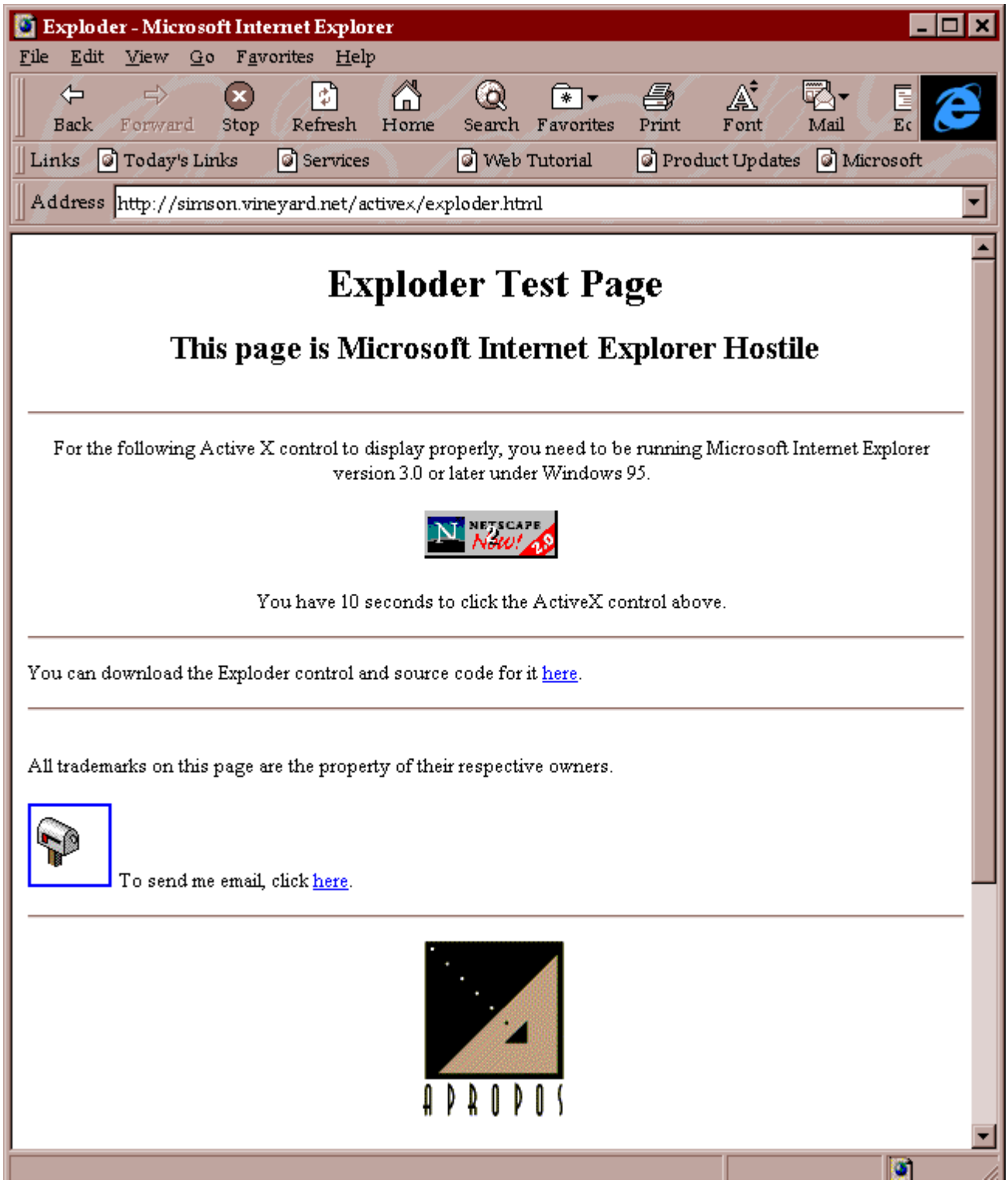
```
<html><head>
<title>Table Attack</title>
</head>
<body><h1>Table Demo</h1>
<hr>
This tests your browser's ability to deal with HTML
   tables...<p>

<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
```

JavaScript

```
<script lang="JavaScript">
while(1){
    alert("Denial of Service Demo.");
}
</script>
```



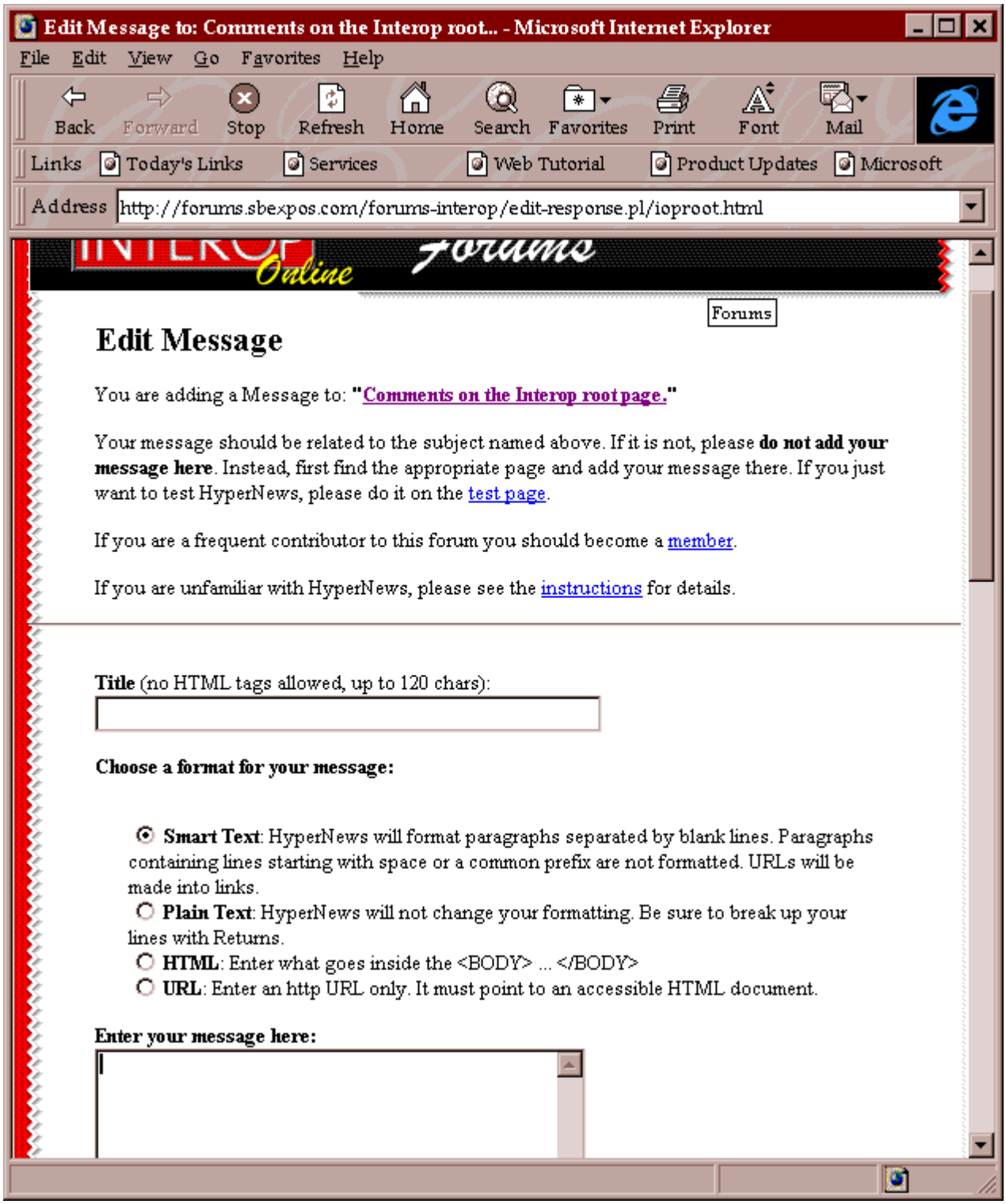
ActiveX Exploder



Delivery by Email

```
Date: Wed, 6 Nov 1996 23:27:03 -0500 (EST)
From: "Simson L. Garfinkel" <simsong>
Message-Id: <199611070427.XAA20319@vineyard.net>
To: anybody@anywhere.net
Mime-Version: 1.0
Content-Type: text/html; charset="us-ascii"

<html><head>
<title>Table Attack</title>
</head>
<body><h1>Table Demo</h1>
<hr>
This tests your browser's ability to deal with HTML
   tables...<p>

<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
<table><tr><th>This is your header...
```

Delivery by Forum



It’s actually remarkably easy to get people to run programs on their computers...

More Malicious Email

ate: Fri, 08 Nov 1996 11:33:02 +0000
From: Warrick Taylor <warrick@wuthmann.com>
Reply-To: warrick@wuthmann.com
Organization: Wuthmann Associates
MIME-Version: 1.0
To: cbermant@mci.com, murometz@aol.com, mike_drips@msn.com, kmfields@cris.com,
bronwynf@aol.com, simsong@vineyard.net, gametheory@aol.com,
fluxman@flux.com, owl@bigfoot.com, morg@li.net, marshalr@pipeline.com,
"pscisco@nr.infi.net4968469"@mcimail.com, chris_shipley@infoworld.com,
urbfutur@interramp.com, pp002580@interramp.com, Ksiegmann@aol.com,
71333.2623@compuserve.com, newsproj@aol.com, askdrj@aol.com,
skatz@ap.org, atworks@instorm.net, meast@axcess.com, mgb@tiac.net
Subject: Postcards from the Net

This is a multi-part message in MIME format.

----------256E5F274A12
Content-Type: text/plain; charset=iso-8859-1
Content-Transfer-Encoding: quoted-printable

NETWORK SOUND & LIGHT, INC. ANNOUNCES
POSTCARDS FROM THE NET
-E-MAIL JUST GOT COOL-

Network Sound & Light, Inc. (http://www.coolcards.com) is excited to introd=
uce Postcards from =

the Net, a whole new way to communicate by e-mail. PLEASE OPEN THE ATTACHME=
NT OR SCROLL TOP THE =

BOTTOM TO SEE A SAMPLE.

Attacks on Cellular Systems

ROBERT A. MCKOSKY

CHRIS CARROLL

HAI-PING KO (SPEAKER)

GTE Laboratories Incorporated

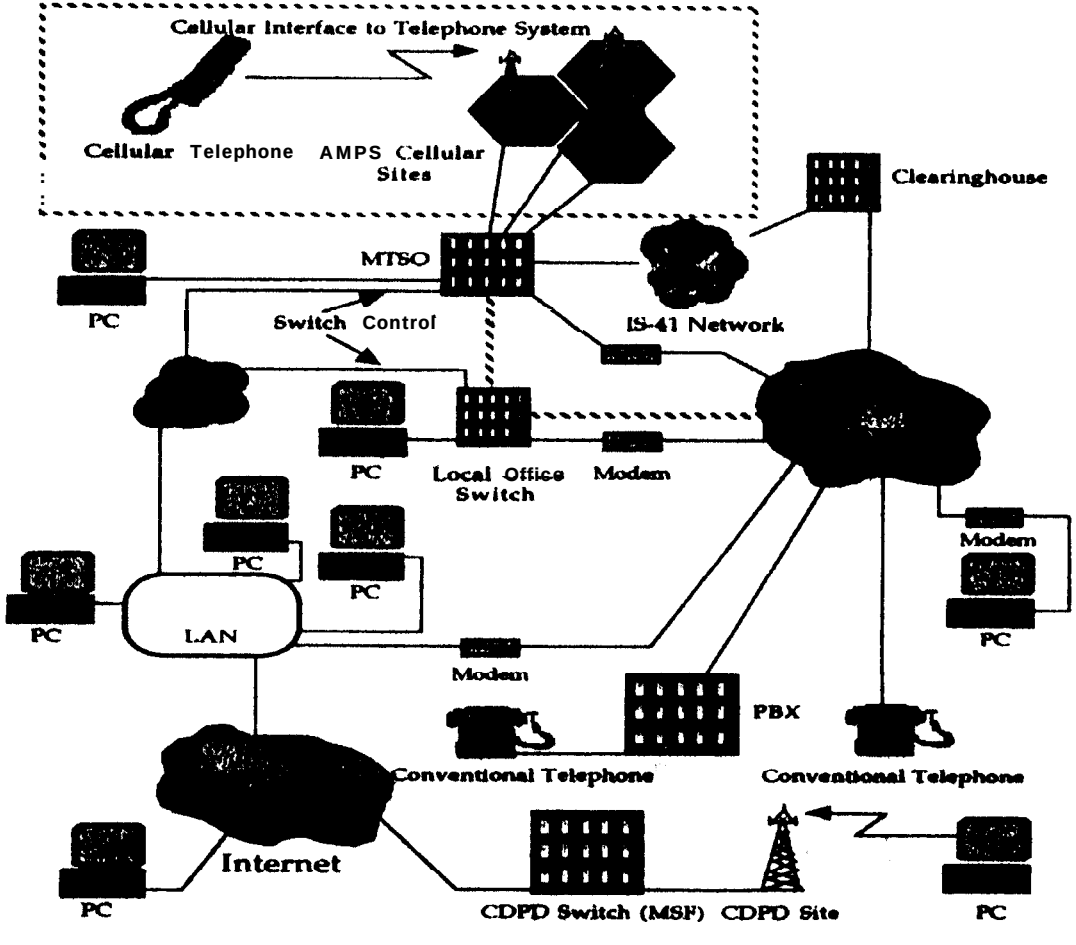
Attacks on Cellular Systems

Robert A. McKosky, Ph.D., CISSP
Chris Carroll, Co-Principal Investigator
Hai-Ping Ko, Ph.D. (Speaker)

November 13, 1996

GTE Laboratories Incorporated

Overview


Cellular Interface to Telephone System
Cellular Telephone
AMPS Cellular Sites
Clearinghouse
MTSO
PC
Switch Control
IS-41 Network
PC
Local Office Switch
Modem
PC
PC
PC
LAN
Modem
Modem
PC
PBX
Conventional Telephone
Conventional Telephone
Internet
PC
CDPD Switch (MSF)
CDPD Site
PC

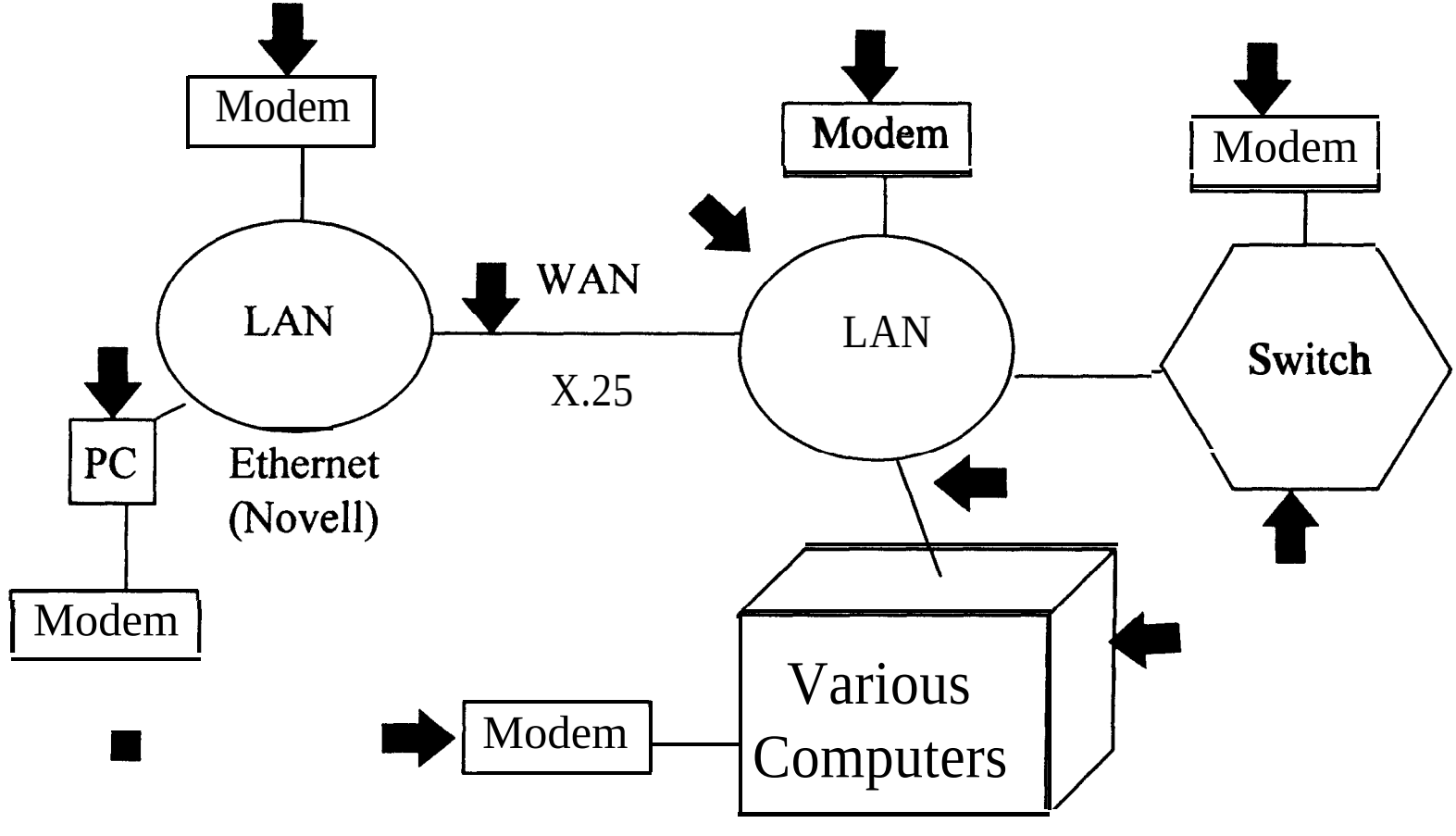
Types of Attack

- $6○○M L○ss Per Year -

- Air
  - Human Fraud or Hijack
  - Clone (Tumbling, Simple, Tumbler)
    - ESN (Electronic Serial Number)
    - MIN (Mobile Identification Number)
- Land
  - Network Attack on Multiple Points
    - Switch
    - Modem, LAN, WAN, PC, various computers

Attack Points

Indicated possible point of attack



Land Tiger Team Results

- remote access to the switch computer
  - obtained /etc/passwd, MINs/EINs, billing, ...
- physical access to offices, computer room,...
- beat SecureID
- clone phone
- Trojan Horse on a PC
- NOT DETECTED

Comments

- "Our biggest problem is the access from the business systems."
- "Our biggest problem is access from personal modems."
- "We try to do a good job here of controlling access, but other areas of the company are not as conscientious."

Comments (Cont.)

- "I don't know who I would call if we had a security problem."
- "Nobody looks at the log on a regular basis."
- "We only look at the logs when we think there has been a problem."
- "I didn't know you could do that!"

Comments ...

- "The modem thief can steal more with a computer than with a gun. Tomorrow's terrorist may be able to do more damage with a keyboard than a bomb . . . To date, we have been remarkably lucky . . . (A)s far as we can tell, there has been no systematic attempt to subvert any of our critical computing systems. Unfortunately, there is reason to believe that our luck will soon run out."
  - National Research Council, 1991

- "Neither AT&T, nor the local exchange telephone companies, nor anyone else can tell you what is connected to the public network fabric today."
  - John C. Wohlstetter, 1993

- "If [senior management] really understood the potential liability and the potential risks to corporate assets and to their reputations, they might shut down all networks and computer centers."
  - Kenneth Weiss, chairman of the computer security division of the American Defense Preparedness Association

Solutions (Partial)

- Air
  - Authentication, Encryption, Clone Detection
- Land
  - Security Owner
  - Security Policy
  - Training
  - Enhanced Audit
  - Encryption

Attacks on Cellular Systems

Hai-Ping Ko
GTE Laboratories Incorporated
Waltham, MA 02254

The cellular industry is growing quickly but so is the fraud. For instance, based on the surveys conducted by the Cellular Telecommunications Industry Association (CTIA), the number of U.S. cellular subscribers, cell sites, and total revenue have grown 112, 27, and 36 times, respectively, over the past ten years. At least 38.2 million, 14.5% of the entire U.S. population, have subscribed to the wireless service and another subscriber is added approximately every 2.8 seconds. There are more than 300 cellular carriers in the States now, but only a small group of entrepreneurs ten years ago. Cellular fraud cost the cellular industry $365 millions in 1994 and at least $500 millions in 1995, consistently 2.56% and 2.62% over the total revenues, in respective years. [1,2,3]

GTE has a special responsibility to understand and to make recommendations on the security problems and solutions of the cellular systems. In 1993, GTE Laboratories was selected by CTIA as the industry's technical analysis laboratory for fraud detection, control, and prevention. Most cellular attack methods were understood. In 1994, the GTE Laboratories succeeded a tiger team attack to a cellular switch station. The switch station was severely compromised without detecting the attacks. The GTE Laboratories consequently was invited by CTIA to conduct a vulnerability study of the cellular industry in general and proposed security policy recommendations and standards to the cellular industry. [1,4,5,6]

I will briefly describe some known attacks on the cellular phone systems, based on years of work of C. Carroll and R.A. McKosky at the GTE Laboratories. I will also briefly describe my sense of computer security and intrusion detection at one of the largest telecommunication companies, GTE.

The attacks on the cellular systems can take place through air (wireless) or through wirelines. To understand this, it is important to know that every connection from a cellular phone to a regular telephone involves the following types of communication: (1) air communication between the cellular phone to a nearest cell base station, (2) wirelined communication between the cell base station and a cellular switch station, and (3) wirelined communication between the cellular switch and the destination through the conventional Public Switched Telephone Network (PSTN). The cellular switch stations are the brains of the cellular systems. With networked computers, they control and direct all the requested connections. These switch stations are connected with the cell base stations and the Public Switched Telephone Network under various protocols and agreements and make sure together that the requested cellular connection can be serviced without interruption when cellular phones move from one location to another. [1,7,8]

The most severe attack to the cellular systems through the air is phone cloning. Unlike a regular telephone which can be recognized by a uniquely distinguishable wire, a cellular

phone is only recognized by a pair of uniquely assigned numbers: ESN (Electronic Serial Number) and MIN (Mobile Identification Number). Such pairs of numbers are transmitted to a cell base station through the open air whenever the cellular phone is powered on. These numbers can be easily read by equipments at a price from $700 to $2000. With an equipment of $7000, one can even possibly find the physical location of any powered-on cellular phone. It is illegal to clone cellular phones with such ESN/MINs, but the cloning methods are freely available from the Internet and phone cloning has become a cottage industry. Some cellular phones are equipped with PINs (Personal Identification Numbers). In such cases, when placing a call, the PIN will need to be sent through the assigned voice channel after ESN and MIN are sent through a control channel. Such cellular phones are less likely to be cloned. However PINs are vulnerable to eavesdropping as well. In fact, there are equipments which can be used to trace the transmitted ESN/MIN/PINs in real-time. [1,3]

Another possible attack through the air is hijacking. Once a voice channel is established between a cellular phone and a cellular base station, a counterfeit cellular phone may seize the voice channel by increasing its power level above that of the legitimate cellular phone. A criminal could then make an illegal cellular call. [1]

The cellular switch stations need the tightest security against any electronic or physical attacks on the cellular systems. These switch stations not only control the cellular connections but also maintain all the registered ESN/MINs and the billing information. The cellular switch computers are vulnerable to all types of network attacks. They are accessible from the Public Switched Telephone Network, which was in turn accessible via the Internet. They are physically connected to modems, various computers, LANs, and WANs, directly or indirectly. Any loose security on the modems, computers, or links will make one or more cellular switch stations vulnerable.

In 1994, one of the cellular switch stations accepted the challenge of a tiger team attack. Only ordinary hacking techniques were used, such as looking for an open port access and cracking weak passwords. The tiger team easily gained the root privilege remotely, altered the password file, obtained the highly confidential information about ESN/MINs and customer billing. The tiger team intentionally left obvious footprints in the hope of being caught, but was not detected. The tiger team also used social engineering gaining physical access to the offices and the computer room, beating the SecureID mechanism, and placing a Trojan horse program on an office PC.

Switch stations of other cellular carriers are not too much different from the switch station under the tiger team attack. It was confirmed in 1995 that several other cellular switch stations of different cellular carriers were equally vulnerable. Even though the switch station under the controlled attack has tightened its computer and physical security since 1994, the overall cellular connections remain vulnerable.

The wirelined attacks are as real as phone cloning. As published in the New York Times of 9/12/95, among the arrested attackers, two actually broke into the computer systems of cellular phone companies.

There have been actions taken to combat the attacks on the cellular systems. For instance, for the phone cloning and hijacking problems, the following methods are being used or considered: voice verification, radio frequency fingerprint verification, dynamic PINs, call pattern analysis, authentication and voice encryption. Securing the cellular networks involves considerations of security ownership, security policy, personnel training, enhanced auditing, and again authentication and encryption on remote connections. More than 30 security issues were identified for the wireless systems and networks by the GTE Laboratories in 1996. Security guidelines were developed by the GTE Laboratories for the cellular industry shortly afterwards. I will not get into any further details here.

Since my employment with GTE beginning early 1995, I observed that GTE is sensitive to computer security problems. Secure architectures were carefully designed and reviewed for every development of company product. Audit records at application level were generated. Some part of GTE are sensitive to external penetrations only and other parts of GTE carefully keep a record of all attacks and observed 80% of them originated internally. In any case, the actual adoption of automated audit analysis and intrusion detection is relatively new and experimental. Well-tested intrusion detection tools have captured unexpected attacks after they are properly installed. GTE understood the existence of potentially fierce attacks and appreciated the value of automated intrusion detection. The cellular fraud problem is well understood and put in good hands at GTE.

References

[1] Cellular Fraud Training Manual for the United States Secret Service, prepared by Cellular Telecommunications Industry Association Technical Analysis Laboratory at GTE Laboratories Incorporated, April 11, 1995

[2] CTIA's Semi-Annual Data Survey, INFOFAX from CTIA (Cellular Telecommunications Industry Association), June, 1996

[3] Fraud and Countermeasures, Part II: Clones, Private Line #10, copyright by 1995 Cellular Networking perspectives, 1996

[4] McKosky, R.A., Security Guidelines, Prepared for CTIA, 1996

[5] McKosky, R.A., Vulnerability Assessment of the Wireless Industry, 1996 Security Seminar, CTIA Network Vulnerability Solutions Committee, June, 1995

[6] Summary of Security Recommendations for Wireless Systems and Networks, Prepared for CTIA by GTE Laboratories Incorporated, 1996

[7] Tanenbaum, A.S., Computer Networks, 3rd Ed, Prentice-Hall, Inc., 1996

[8] Telecommunications Engineer's Reference Book, edited by F. Mazda, Butterworth-Heinemann Ltd., 1993

Miscellaneous Papers from Participants

Internetwork Security Monitor: An Intrusion-Detection System for Large-Scale Networks

L. T. HEBERLEIN, B. MUKHERJEE, K. N. LEVITT, UC Davis

Analysis and Response for Intrusion Detection in Large Networks

PETER G. NEUMANN, PHILLIP A. PORRAS, ALFONSO VALDES, SRI International

Distributed Detection of Distributed Attacks

DOUGLAS B. MORAN

Artificial Intelligence Center

SRI International

Intrusion Detection in the Large:
Distributed Detection of Distributed Attacks

Douglas B. Moran
Artificial In telligencc Center
SRI International
moran@ai.sri.com
http://www.ai.sri.com/~moran/

Distributed Attacks

- **Distributed Target**
  - Distributed System
    - Distributed File System
    - Database
    - Agent Systems
  - Shared privilege
- **Distributed Source**
- **Distributed over time**
- **Data Fusion Problem**
  - Loose clusters
  - Massive overlap
  - No hierarchy: flexible & dynamic organizations
    - task force
    - business process re-engineering
    - out-sourcing
- **Task Model**
- **Human Factors**

Distributed Detection

- **Partial Evidence per Intrusion**
- **Merge Evidence from Multiple Sites**
  - Matching incidents
  - Reliability/Competence of reporter
  - Terminological and procedural uncertainty and inconsistency
- **Sites Under Attack Directly Communicate**
- **Reporting Problems**
  - Confidentiality/Sanitize
  - Security
  - Feedback to cracker
  - Under-reporting

Improved Reporting

- Create Automated Security Manual (shortage of human expertise)
- Catalogue of Known Intrusion Scenarios and Techniques
  - Confidentiality issue
- Customizable to Site
  - Better diagnosis
  - Reduced consistency

Goals of Project:

- Short-term Goal
  - Improved diagnosis
  - Assisted recovery
- Long-term Goal
  - Automated report generation
  - Multilevel reports
    - trustworthiness of recipient
    - current situation

AI Technology

- Reactive (PRS)
  - Event driven
  - Automated manual
  - Short horizon
- Look-ahead Planner (SIPE)
  - resource usage
  - info retrieval conflicts
- Common Representation Formalism
- Each Domain Requires its own Extensions and Customizations
- Intelligent, Adaptive Scheduler of Tasks (threads)

SRI International
AI Center

PRS-CL
A Procedural Reasoning Reactive Execution System

TECHNOLOGY

- Reasoning based upon predefined procedural knowledge
- Reactive and goal driven
- Real-time response
- Meta-level reasoning
- Multiple cooperating agents
- Interactive, menu-driven, graphical interface

APPLICATIONS

- Space shuttle fault diagnosis
- Aircraft maintenance
- Air battle management
- Mobile robot control
- Communications network management
- Joint military operations
- Sonobuoy deployment

Design Issues

- Phased Response
  - Are there dependable cues
  - Limit: avoid becoming denial-of-service (computer or human)
- Building up Catalogue of Attack Scenarios
  - Reuse of attack components
  - Ease of specifying
- Ability to Identify
  - Variants
  - New attacks using some known components
- Distributed Attack in small Cluster of Computers
- Single Platform Type

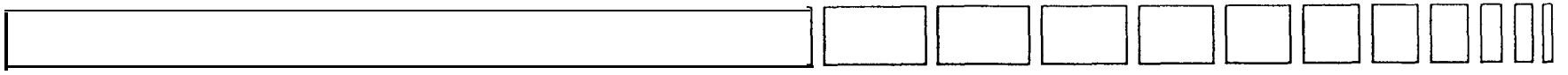
Scaling-Up

- Filtering and Routing Info
  - Little relevant structure in network
  - Trust vs. need-to-know
- Incomplete Info
  - Too little for meaningful report
    - request info from “authorities”
    - reanalyze
  - Enough to report
    - clearing house
    - involved hosts
    - siblings

Thresholds for above??

- Automatic Processing of Reports
- Determine what can reasonably be shared with whom

User in Loop vs. Uses at end of a pipe



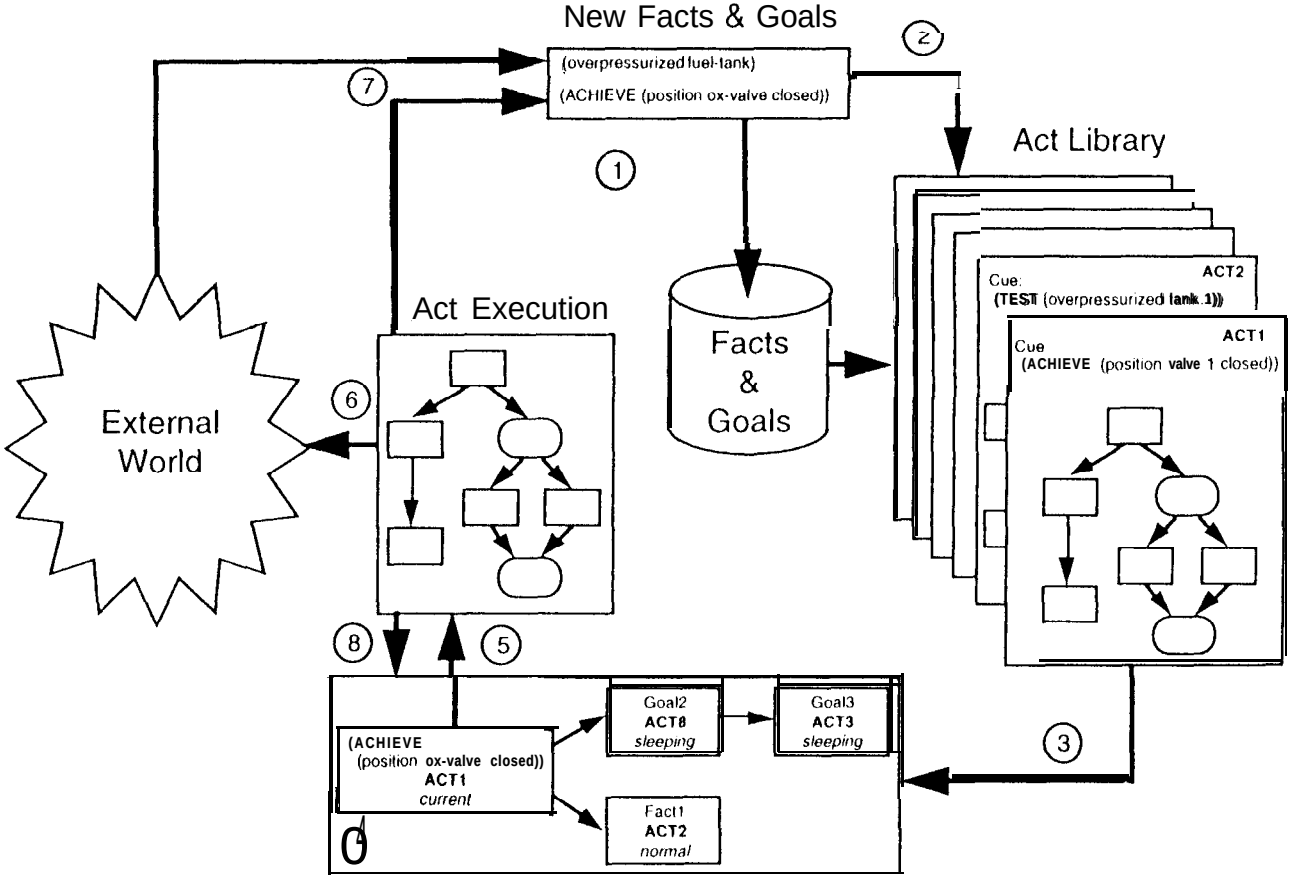
- User of security system is major knowledge source
  - Often unavailable
  - Mobile
  - Different user interfaces
- Backup with automated reasoning system
- Collaboration of Humans and Automated Systems
- Agent-based Architecture

AI Center

PRS-CL Architecture

Execution Cycle

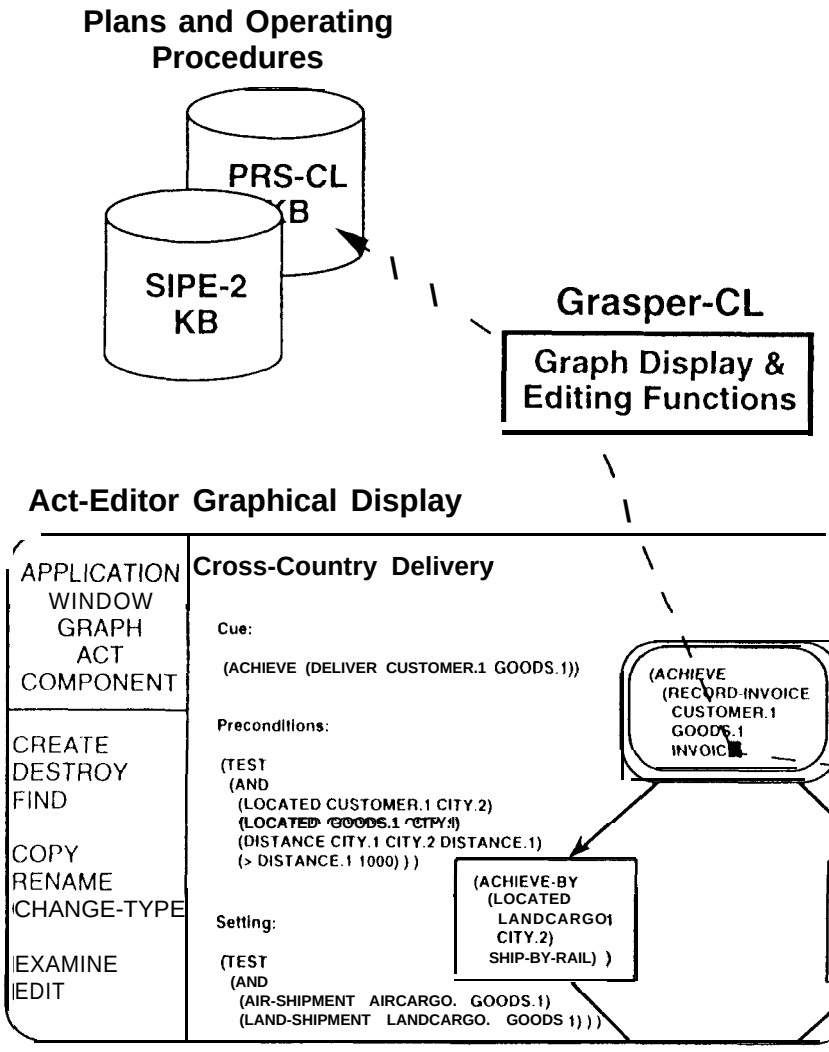
1. New information arrives that updates facts and goals
2. Acts are triggered by new facts or goals
3. A triggered Act is intended
4. An intended Act is selected
5. That intention is activated
6. An action is performed
7. New facts or goals are posted
8. Intentions are updated



AI Center

Act-Editor

Procedural Knowledge Browser/Editor



- Graphically browse procedures
- Edit procedures through direct pictorial manipulation
- Uniformly manage plans and operating procedures
- Verify against dictionaries of predicates and objects

Scalable Intrusion Detection for the Emerging Network Infrastructure

Y. FRANK JOU

HPCC Research

Information Technologies

MCNC

Autonomous Agents

MARK CROSBIE

Hewlett-Packard/COAST

Autonomous Agents

A solution for Large Scale Intrusion Detection ?

Mark Crosbie

Hewlett-Packard/COAST

Critical Problems

- Distribution of configuration information.
- Allowing local configuration changes.
- Putting “local wisdom” in reports.
- Data acquisition for trend analysis and risk management.
- Tool evaluation in an enterprise-wide setting.

Distributing Configurations

- How do we distribute configurations across administrative domains?
- Push or Pull model?
- Automated or human driven?
- Diverse user groups - not everyone is an expert!
- Need a background propagation mechanism.

Autonomous Agents

- Lightweight, mobile code modules.
- Migrate and replicate across network - implicit “push” model.
- Background task - no need for human intervention.
- Can interact with local “wisdom stores” when generating reports.

Reporting Problems

- Reporting - how do we get the right information to the right people?
- Will they know what to do with the report?
- Each group has a local “wisdom store”.
- Agents interact with wisdom store to provide reports tailored for the group.
- Relieves burden on central security “expert”

Evaluating a large IDS

- A System that attempts to break into itself.
- Automate attack capture.
- Replay attacks across the enterprise.
- Evaluate detection relative to enterprise-wide security policy.
- Feedback of test results into configuration.

Problems that remain

- Do we want automated intrusion responses? *Active Intrusion Detection.*
- How does the IDS integrate with enterprise reporting and issue tracking tools?
- Allowing local configuration changes, but remaining within enterprise policy.

Conclusions

- Problems are often to do with humans, not technology.
- Can’t change the world - must integrate with existing technologies.
- Automate tasks - humans are not always “experts”.
- Use “push” models for distributing configurations.

Network Management and Operations

JF MERGEN

BBN

Thoughts About Susceptibility to Data Driven Attacks

MARVIN SCHAEFER (SPEAKER)

GARY R. GROSSMAN

Arca Systems, Inc.

The Need for a Standard for the Format and Content of Audit Trails

KATHERINE PRICE

COAST Computer Security Laboratory

Purdue University

The Need for a Standard for the Format and Content of Audit Trails

Katherine Price

Purdue University

COAST Computer Security Lab

kep@cs.purdue.edu

Topics of Discussion

- Problem of No Widely Accepted Standard
- Difficulties with Lack of Content
- Difficulties with Tool Migration
- Difficulties with Data Reconciliation
- Some Proposed Standards
- Current Work to Develop a Standard

No Widely Accepted Standard

- Each audit source creates its own ad-hoc standard for format and content
  - the format for the audit trails varies greatly from system to system
  - each system gathers different data based on what the developer believed was important
- Disparity in format and content of audit data impedes progress in intrusion detection

Impediments to Progress in Intrusion Detection Methods

- Three major difficulties face intrusion detection techniques
  - difficulties with lack of content
  - difficulties with tools migration
  - difficulties with data reconciliation
- A standard for format and content would help overcome these impediments

Difficulties with Lack of Content

- Many current auditing systems do not supply enough data
  - lack of record activities
  - lack of detail
- Intrusions are not being detected because of insufficient evidence in audit trail

Bindings and Lack of Detail

- Audit data often does not contain enough information to resolve bindings
  - files names are transient bindings that may change over the life of the file
  - file descriptors, such as inode numbers in UNIX, are fixed throughout the life of the file
- Race condition attacks often take advantage of binding resolution problems

Difficulties with Tool Migration

- Many detection tools are designed for a particular audit source
- Difficulties in changing audit source
  - disparity in types of data available
    - » algorithms tailored for particular data may become ineffective
  - converting between formats is difficult
- Disparity in audit data makes it difficult to migrate tools to new audit sources

Difficulties with Data Reconciliation

- Detection systems must analyze data from multiple sources to uncover new, sophisticated attacks
- Many possible sources of information
  - applications and operating systems
  - firewalls and routers
- Disparity in audit data makes it difficult to reconcile multiple audit sources

Some Proposed Standards

- Standards for Content
  - C2 Level Audit
    - » standard is too broad
    - » different interpretations of “security relevant event”
- Standards for Format
  - ASAX’s NADF
  - Bishop’s Format
    - » Both handle UNIX OK, but difficulties may arise with other sources, especially with hierarchical data

Auditing on Sidewinder

TOM HAIGH

Secure Computing Corporation

Information Security and the Electric Power Industry

AB KADER (SPEAKER)

RON SKELTON

EPRI

Information Security and the Electric Power Industry

A Presentation to the Fourth Workshop on Computer Misuse and Anomaly Detection (CMAD-IV)

Ab Kader

Ron Skelton

Electric Power Research Institute (EPRI)

Presentation Overview

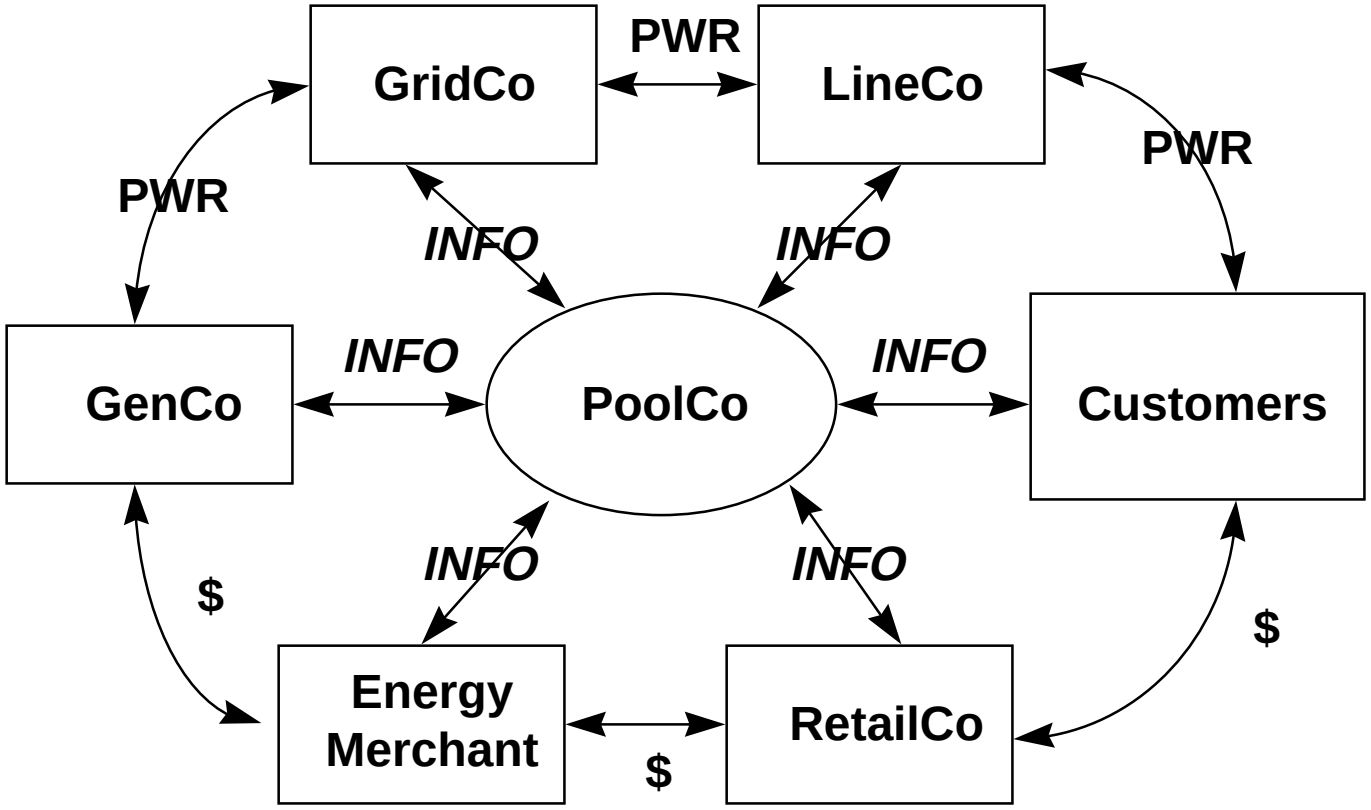
- The Challenge
  - Why do Electric Utilities have a security problem?
- The Response
  - What is EPRI doing about it?
- Future Work
  - Where do we go from here?

Utility Information Networks

Utility Information Networks

- Corporate: generic (& utility specific) back office processing.
- Power Plant: generation control & communication systems.
- Control Center: interface between generation & transmission.
- Transmission: SCADA and EMS.
- Distribution Automation: remote monitoring and control of distribution substations.
- Customer Interface: remote communication with devices at customer sites.
- External: other utilities, power pools, vendors etc..

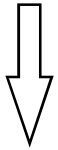
Utility Industry "Future"



Richard Abdoo, CEO WEPCo
Public Utilities Fortnightly, 2/15/95

“Future” Is At Hand

- Federal Energy Regulatory Commission (FERC) 889
  - information on transmission availability and prices.
  - equal access for wholesale sellers and purchasers.

EPRI 

- Open Access Same Time Information Systems (OASIS)
  - internet based information system.
  - encryption and digital certificate based security.

OASIS Nodes

EPRI Security Initiatives

- Information Security Workshop
  - Utility Security Survey (NSTAC)
  - Utility Security Assessment (Battelle)
  - Utility Security Policies (EPRI)
  - Security Tutorial (MIS Training)

- Information Security Applications
  - Power System Security (LANL)
  - Residential Customer Security (LANL)

Security Survey Highlights

- Willing to share security incident information.
- Believe “private nets” are secure.
- Trend towards less secure “public nets”.
- Concerned more about internal threats.
- Widespread lengthy electric grid disruptions unlikely.
- Security protection and audit practices inadequate.
- Internal priorities limiting attention to security concerns.
- 90% expressed a desire of ongoing EPRI involvement.

Security Assessment Conclusions

- Growth and reliance on information technology increases security threats.
- Business climate does not foster adequate security protection measures.
- Electric utility industry trends introduce new ill understood security vulnerabilities.

Security Policies Universe

Inter Control Center Communications Protocol (ICCP)

Internet Based Home Energy Management Pilot

Next Steps

- Real time intrusion detection
  - research techniques for protecting power dispatching and trading, utility customer communications....
- Incident response handling
  - security incident reporting, resolution, and information dissemination (anonymously, if so desired).
- Security testing center
  - penetration testing and security auditing services customized for electric utilities.

Computer Based Forensics
A Case Study – U. S. Support to the U. N.

CAPT. KEVIN J. ZIESE

AF Information Warfare Center

Computer Based Forensics
- A Case Study -
U.S Support To The U.N.

Capt Kevin J. Ziese

AF Information Warfare Center

ziese@mailcenter.cmet.af.mil

Overview

- Presentation Strategy
- Working Definitions
- Problem Description
- Field Prototypes
- Major Shortfalls
- Top Five Christmas Gifts

Presentation Strategy

- Describe A, Serious, Real-World Problem
- Present The Low-Level Technical Issues
- Identify The Relevant Solution Criteria
- Generate, Focused, Expert Discussion
- Synthesize Potential R&D Directions
- Generate Potential COTS Opportunities
- Improve Overall Forensics Process

Computer Forensics

VALID TOOLS AND TECHNIQUES APPLIED AGAINST COMPUTER NETWORKS, SYSTEMS, PERIPHERALS, SOFTWARE, DATA, AND/OR USERS -- TO IDENTIFY ACTORS, ACTIONS, AND/OR STATES OF INTEREST

RELATED TO TRADITIONAL BIOLOGICAL, CHEMICAL, AND PHYSICAL SCIENCES

Valid Tools & Techniques

TOOLS AND TECHNIQUES THAT CAN BE APPLIED AS REQUIRED AND DO NOT REQUIRE RECASTING THE PROBLEM TO BE USED EFFECTIVELY

ARE CONFIGURATION DRIVEN WHICH MEANS THEY ARE, WHERE POSSIBLE, NOT COUNTRY OR OPERATING SYSTEM CENTRIC

Problem Description

- International Problems
  - Defines, And Complicates, The Solution Space
- Operational Problems
  - There Are Deltas Between “The Lab” & “Reality”
- Technical Problems
  - Effectiveness Always Comes Before Efficiency
- Legal Issues
  - Just Because It’s Legal Doesn’t Mean It’s Right

International Problems

- Is Iraq Violating U.N. Sanctions?
- Are Computers Supporting That Activity?
- Is Iraqi Compliance Real Or Feigned?
- How Reliable Are The Team's Findings?
- Did We Protect Iraq's Right?
- Did We Act As Good International Citizens?
- Where Are The 16 (?) Missing SCUDs?

Operational Problems

- How Did Computers Support NBC Activity?
- How Do You Protect Search Methods?
- How Do You Search Ancient Hardware?
- How Do You Search Hostile Systems Safely?
- How Do You Protect Tools & Data?
- When Should You Confiscate Hardware?
- How Long Can You Search ‘In Situ?’

Technical Problems

- Non-English Search Terms
- Non-Symmetric Language(s)
- Binary Application Interfaces
- Proprietary Storage Techniques
- Semantic Representation Of Data
- Information Hiding Techniques
- Search Tools Can Aggravate A Tense Situation

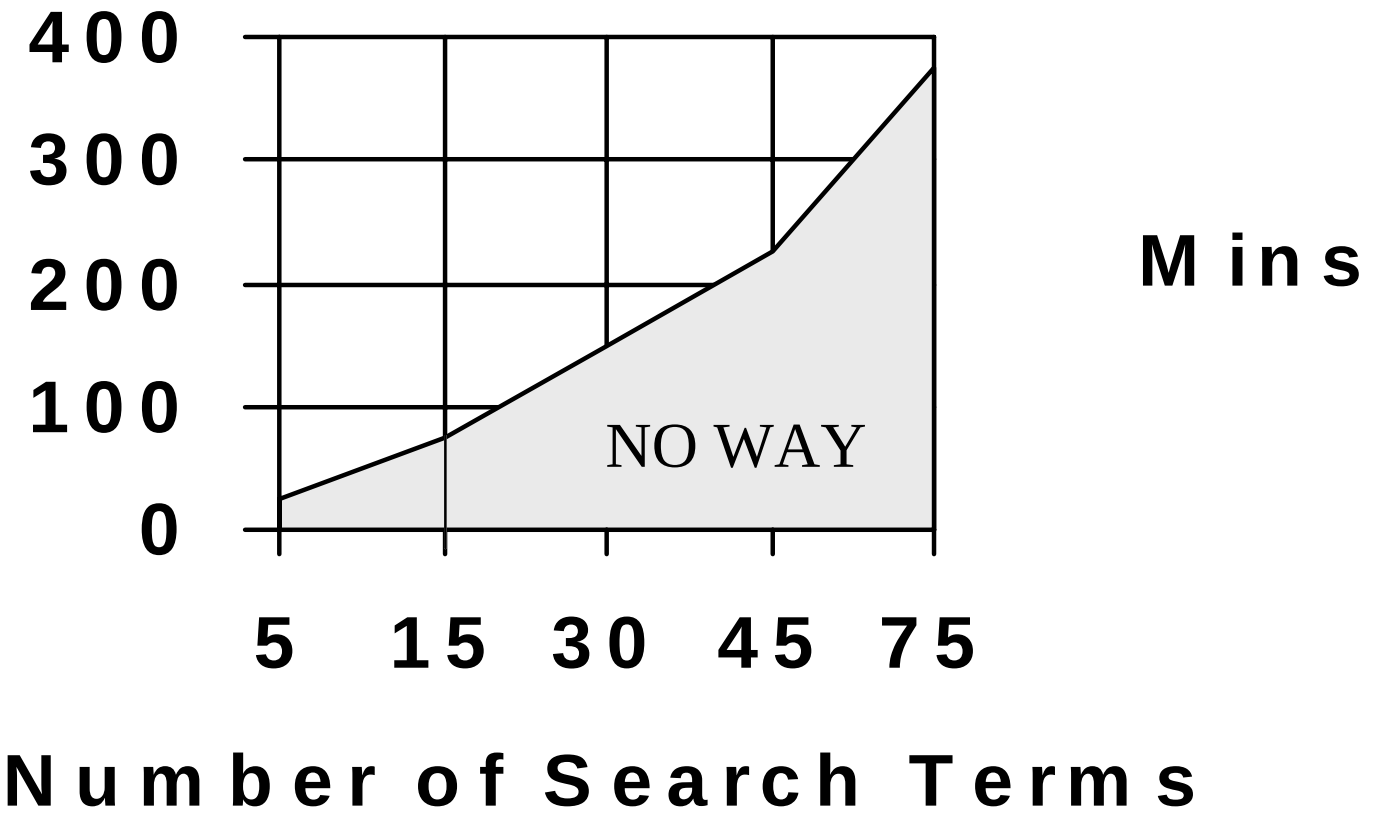
Non-English Search Terms

- Strings Are Not Easily Visualized
  - CONTRACT = UR]
  - CREDIT = HUJLA]
- Strings Change On Context
  - “HUJLA]” OR “MDXM” OR “HGU;”
- Often Mimic Binary Code Stubs
  - High False Positive Rate
  - Defies Many US-Centric Tools (STRINGS)

Non-Symmetric Languages

- Language Order Is Right -> Left
  - “ESUOH” vs “HOUSE”
- There Can Be Holes In The Language
  - “ESUxyOH” + “ESUxxyyzOH”
- Expressed Words Vary
  - “HOUSE” or “ABODE” ???
- Non-REGEX Searches Increased In Step-Linear Time
  - Time =(terms*3 mins) + (int(terms/5))* 10 mins
  - Best “Device” Tool Didn’t Support REGEX Searches

Search Times vs Search Terms



Binary Application Interfaces

- One Application Processes Data
- One Application Displays Data
- Common In Non-English Computers
- Data Was Stored As Huffman Encoded Trees

Semantic Data Representation

- Search Term Representation
  - MISSILE
- Context Representation

  “...MISSILE IN YOUR PATH...”
- Semantic Meaning
  - AUTOEXEC.BAT
  - REM Put Missile In Your Path
  - REM To Play Missile Commander Vers 2.1.3

Legal Issues

- People, And Countries, Have Rights!
- How Long, And Hard, Can You Search?
- What If Your Results Are Indeterminate?
- How Reproduceable Are Your Findings?
- Is Privacy Violated When Data Is Held?
- Is Freedom Violated When Data Is Held?

Operational Issues Today

- Do Manual Searches Endanger Privacy?
- Searches Are Long; Often Involve Confiscation
- Tools Are Not Standardized Or Validated
- Examiners Are Not Standardized Or Validated
- Good Forensics Can Enhance Personal Freedom
- Poor Forensics Can Erode Personal Freedom
- Today’s Forensics Need *LOTS* Of Work

Generic Search Shortfalls

- Tools/Data Must Be Secure In Transit
  - No Tool To Install Encrypted Payloads
- Findings Must Be Secure In Transit
  - No “Encrypt While Copying” Function
- Tools Require Positive Control
  - No “Permissive Action Link” Function(s)
- Tools Are OS Dependent

Prototyped Solutions

- Secure Delivery Tools
- Permissive Action Links
- Device Driven Tools

Secure Delivery

- Tools, And Terms, Encrypted On Floppy
- Floppy Is Mastered With Serial Number
- Decryption Requires
  - Decryption Key
  - Valid Serial Number
  - Operator Authorization
- Only Then Can Search Begin

Permissive Action Links

- Two-Passwords To Execute
- Aperiodic “Attributes” Check
- Destroy On Failed Test Return

Device Driven Tools

- Search Files, Slack Space, Erased/Swap
- Search Logical, Network, Devices
- Search Logical Filesystem Partitions
- Search Raw Device Filesystems

Major Shortfalls

- Technical Shortfalls
- Privacy Shortfalls
- Tomorrow’s Shortfalls

Technical Shortfalls

- Tools Tend To Be Time Inefficient
- Tools Tend To Be US-Centric
- Tools Tend To Be OS-Centric
- What About Information Hiding Techniques?
- We Need ‘dd’ For Every OS

Tools vs Time Efficiency

- Overfocus On Graphic Interfaces
- No Focus On Efficiency/Performance Impact
- Too Little Focus On Semantic Representation
- Don’t Scale Well To Disjoint Text Patterns
- Very High False Positive Rates
- Still Very Much “Caveat Emptor”

US-Centric Tools

- Strings & Egrep Are Efficient -- Not Effective
- Filtering Templates Would Be Better
  - Allow Users To Define “Strings”
  - Allow Users To Define “Operators”
- Other “Languages” Problem Mimics Encryption
  - What About Encryption…
  - Was Encryption Used? What Types?

OS-Centric Tools

- We Need Device Oriented Searches
- We Need User Definable Data Views
  - User Specifies Disk Geometry
  - User Specifies /etc/magic Relations
  - User Specifies User Views
- We Don’t Need “UNIX” Solutions…
- We Do Need “Cross-Platform” Solutions...

Information Hiding Techniques

- Painfully Slow
  - Good Graphics, Limited Functionality
- Few Choices And Limited Envrionments
  - Sound Files And Graphics In DOS primarily
  - What About .AU files What About JPEG?
- More Anecdotal & Notional
  - A Smart “Attacker” Will Use Them…
  - …We Don’t Usually Catch The Smart Ones

One ‘DD’ For Unix, DOS, Mac

- No Less Than 5 Backup Methods
- No Less Than 15 Reload Procedures
- One Source Tree With One Makefile
- Low-Level, Configurable, Disk Backup
- Ability To “Model” One System On Another
  - Simulation Environment To Analyze X on Y
  - Ability To Model One Executable X on OS Y
  - Backups/Reloads, Static, and Dynamic Analysis

Top Five Christmas Gifts

- Encrypted File Systems For Unix, DOS, Win95
- /etc/magic For All Unix, DOS, Mac, Etc
- Fast, CLI, Search Tool For Unix, DOS
- To Be Home For The Holidays
- The 16 Remaining SCUD Missiles

Interactive Intrusion Detection

MIKE NEUMAN

En Garde Systems Inc.

CMAD IV Summary

MARK SCHNEIDER

Office of INFOSEC Research

Introduction

Mark Schneider

Office of INFOSEC Research

Computer Science Research Division

mss@tycho.ncsc.mil

CMAD IV Summary

- **:-)**
- Intrusion Detection research is maturing
- Intrusion Detection research is advancing
- Perception Management needed

Research Challenges

- Anomaly Detection
- Enterprise IDS
- Prevention/Response
- Infrastructure Support

Some Thoughts

GENE SPAFFORD

Purdue University

Academic Differences

- Limits to experience
- Limits to scope
- Different environments
- Different equipment
- Different policies
- Limited continuity & event horizon

- Sometimes larger view
- Fewer constraints
- Requirement to be “clever”
- Access to different sources of data and information

Next Steps

- **Understand policy in clear manner**
- **Understand expected environment**
- **Vision of goals**

- **More clearly identify opportunities & strengths.**

Some Thoughts

What Do We *Really* Want?

- Intrusion Detection
- Misuse Detection
- Anomaly Detection
- Performance Analysis
- Forensic Examination
- Easy to Use
- Infinitely Scalable
- Finds Unknown Conditions
- Easy to Maintain
- Updates Itself
- Standardized Testing
- Completely Portable
- Free of Charge

Redefine the Problem

We want understanding

- Of program interaction
- Of system interaction
- Of user behavior
- Of fault and exception consequences

Note: is there any way to distinguish a bug from a fault from a mistake from a violation of security, in general?

Maybe we've been asking too narrow a set of questions?

New Ideas: Borrowing from Other Areas

MARY ELLEN ZURKO

Open Group Research Institute

New Ideas

Borrowing From Other Areas

Mary Ellen (Mez) Zurko

Open Group Research Institute

http://www.osf.org/~zurko/

zurko@osf.org

User Centered Intrusion Detection?

- Add usability testing
  - **To ID test suites**
  - **To determine what to audit**
- Enable users to use their knowledge to detect intrusion and misuse
  - **The watchers watch the watchers**
  - **Minimize adversarial relationship**
  - **Distribute trust between technology and people**
    Example: Misuse of CPU

User Centered Intrusion Detection?

- ID checks that policy is being properly enforced
  - Share authorization policy database
- Policy language work in both areas
  - Share syntax or semantics
- Use audit data for history-based authorization policies

User Centered Security

- Deep synthesis of security and usability
- Computer Human Interface (CHI)
  - Emphasis on understanding the end user
  - Users do things for a reason
  - Reliance on training is suspect
  - Iterative usability testing is the back bone
    - Early and often

Other Aspects of Security Management

- Trust management
  - Merges Authentication and Authorization
  - Policy Maker, SDSI
- Merge aspects of intrusion
  - Detection and authorization