Monitoring DBMS Activity for Detecting Data Exfiltration by Insiders

Purdue:
Elisa Bertino
Lorenzo Bossi
Syed Rafiul Hussain
Asmaa Sallam

Northrop Grumman:
David Landers
R. Michael Lefler
Donald Steiner

Customer Need: Detect and Respond to Insider Threats

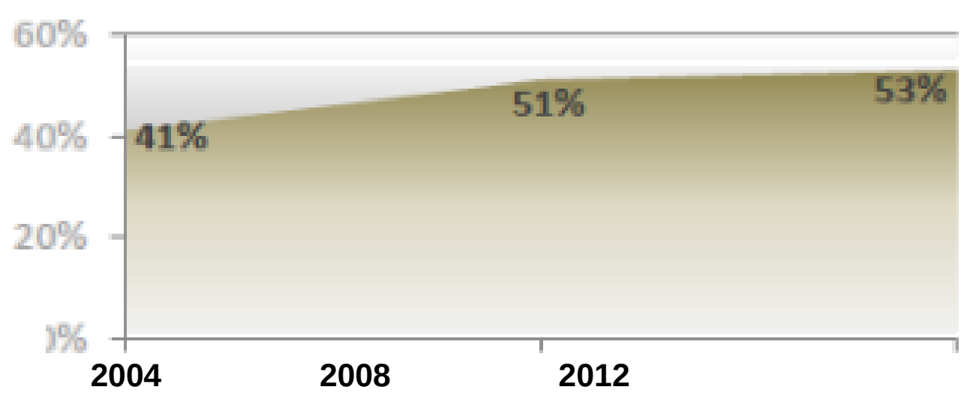
Corporate Awareness

53% Known Incident

33% No response plan

54% Threat has become harder

Insider Threat Occurrences



Average Time to Detect: 32 Months

Damage to Enterprise

Cost

Reputation

Operations

Lives

Types of Insiders

Malicious

Unwitting

Rule Bender

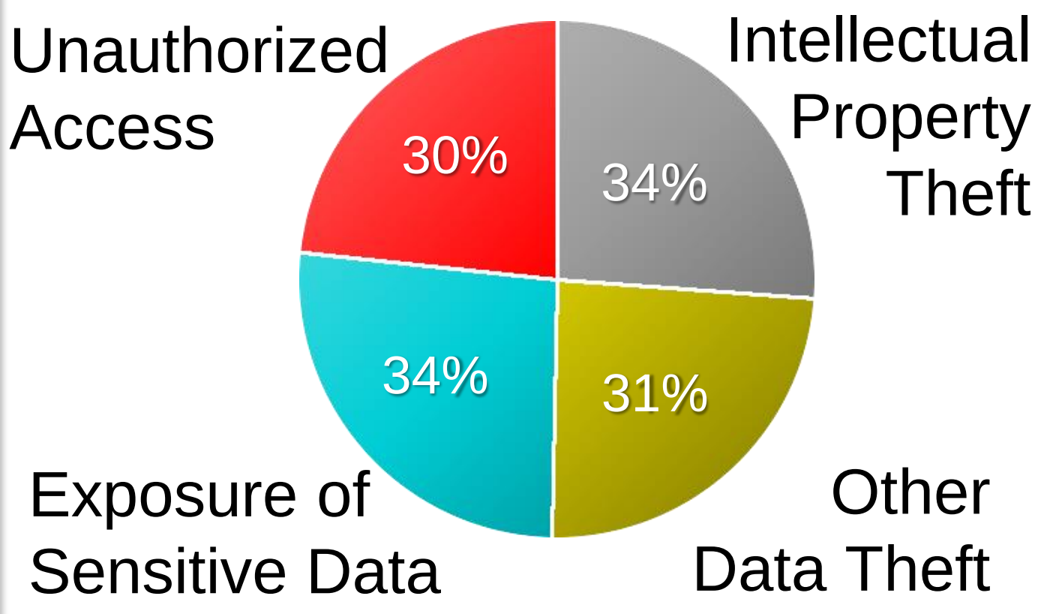
Employee Behavior

51% OK to take data if policies not enforced

37% Use online share-sites without permission

20% Have stored corp. IP on personal devices

Types of Breaches



Background

Hypothesis

Exfiltration causes an anomalous state that can be distinguished from the legitimate actions executed in a DBMS system.

Challenge

Identify the events that represent signs of cyber-insider actions:

- “How do we define and identify user queries that are anomalous?”
- “Which data sources does an insider target?”
- “What information should be collected to detect such actions?”

Approach (Technical)

- Build accurate DBMS access profiles (patterns of normalcy) using Role Based Access Control (RBAC) model
- Detect and respond to anomalous user behavior and events
  - Observe deviations from profiles in real-time
  - Alert security operators
  - Respond according to set policies and forensics

Approach (Programmatic)

Research
PURDUE UNIVERSITY
CERIAS
Develop
NORTHROP GRUMMAN
Evaluate
NORTHROP GRUMMAN

- Provide core algorithms
- Develop Proof of Concepts
- Test in lab environment

- Transform PoCs to “shrink-wrapped” prototypes
- Integrate COTS products
- Maintain during evaluation

- Run prototype in realistic controlled environments
- Collect performance and accuracy measurements

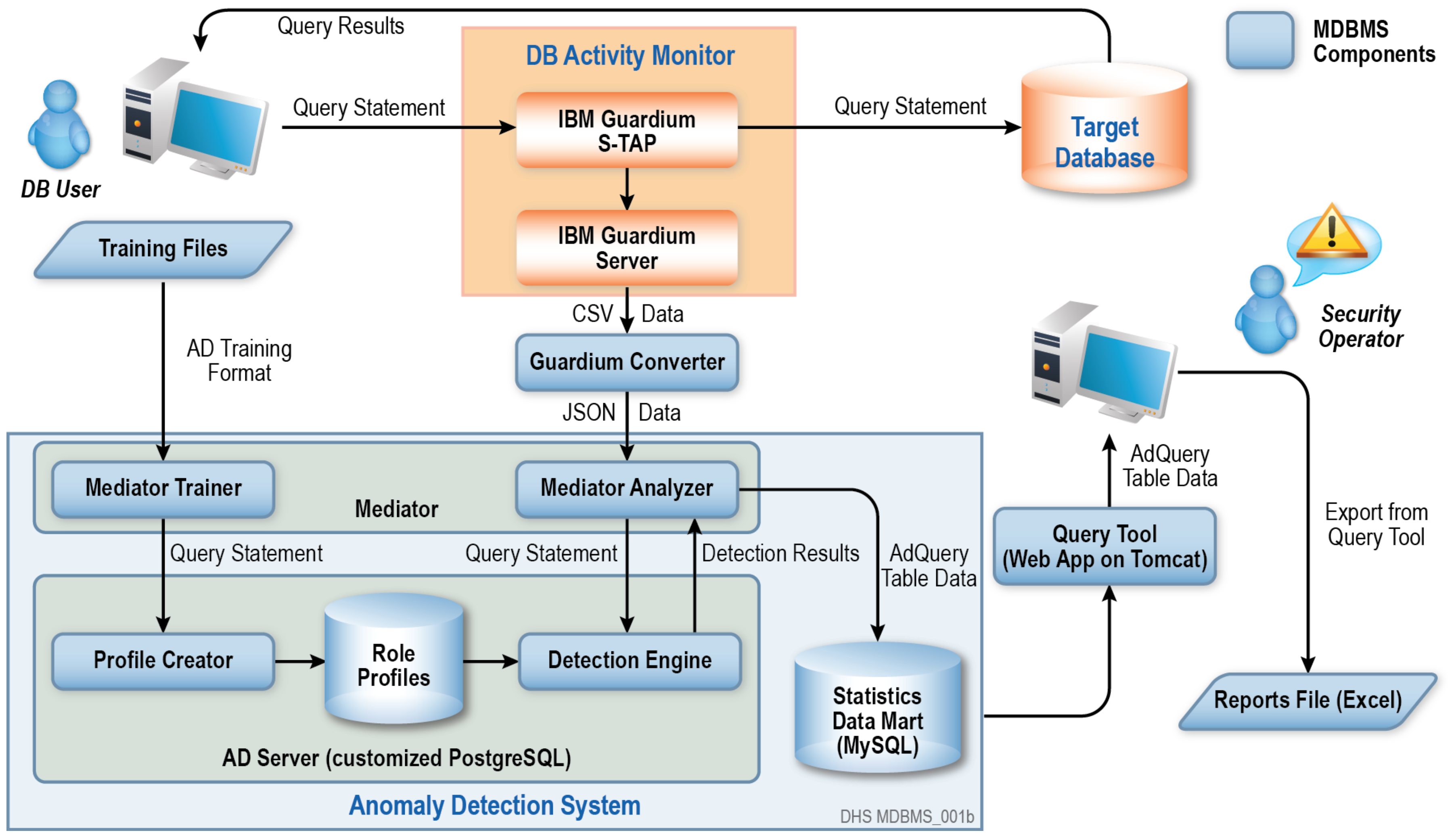
Three Phases over three years

- **Prototype 1**: Initial key features in controlled lab environment
- **Prototype 2**: Expanded features in controlled lab environment
- **Pilot**: Operational environment at select government agency

Benefits

- Dynamic and automated generation of behavioral profiles
- Near-real time alerts of anomalous database activity
- Policy-defined (automated) response
- History and explanation for forensics

Current Status (Prototype 1)



Evaluation Results

Summary Using All Available Data

Detector Type	Evaluation Method	True Positives	False Positives
		Average Values	Average Values
Baseline	Human Evaluation	0.00%	0.00%
Bayesian Detector	AD Score - Alerts only	41.73%	14.54%
	AD Score - Alerts and Warnings	60.93%	25.15%
	Human Evaluation	39.31%	8.50%
Binary Detector	AD Score	66.37%	55.72%
	Human Evaluation	48.79%	12.75%
For Reference		100.00%	100.00%

Summary Using Only Parsed Data

Detector Type	Evaluation Method	True Positives	False Positives
		Average Values	Average Values
Baseline	Human Evaluation	0.00%	0.00%
Bayesian Detector	AD Score - Alerts only	61.04%	19.78%
	AD Score - Alerts and Warnings	88.79%	34.84%
	Human Evaluation	58.14%	11.81%
Binary Detector	AD Score	89.20%	75.68%
	Human Evaluation	65.21%	18.08%
For Reference		100.00%	100.00%

Next steps (Prototype 2)

- Role profiling
  - Enhanced machine learning algorithms
  - Analysis of query optimizers for use in profiling the selectivity of role queries (e.g. for data-based anomaly detection)
- Application program profiling
  - Profile and monitor application programs with respect to their database accesses
  - Use concolic testing to capture the application behavior.
- Response mechanisms