

# Database Assurance

Dr. Peter Mork

(703) 983-1465 • [pmork@mitre.org](mailto:pmork@mitre.org)

MSR

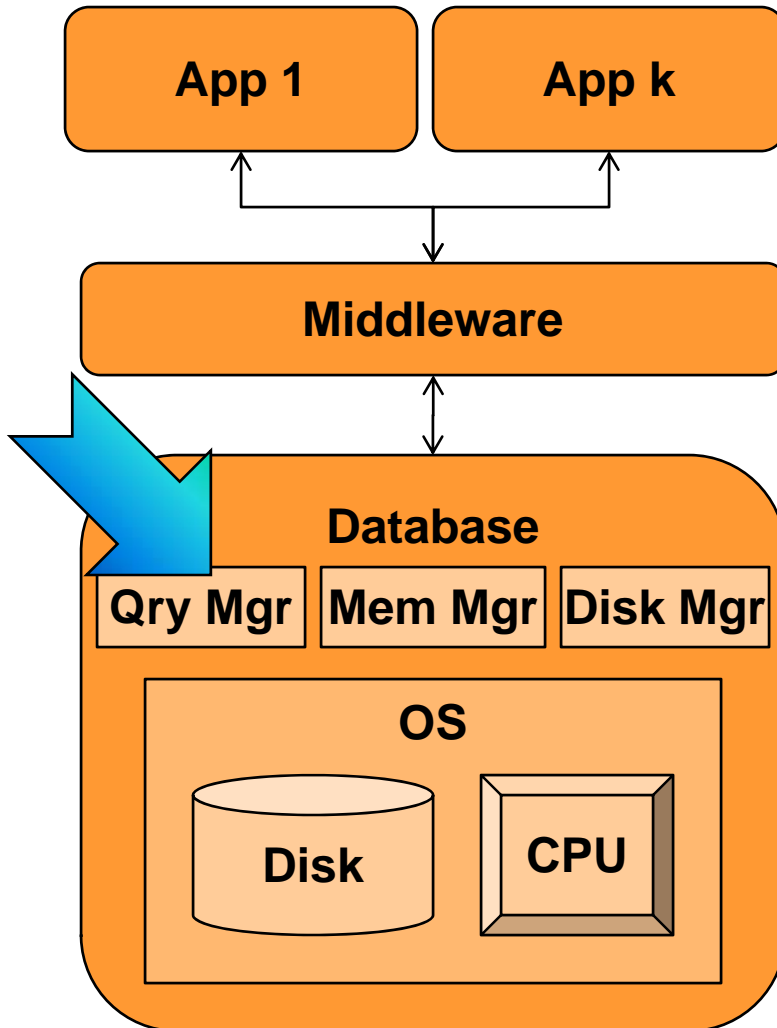


# Problem



- **Critical applications often store data in relational databases.**
- **Database administrators need to know when data assets are being compromised**
  - **Exfiltration: database contents leaked by adversary**
  - **Manipulation: database contents modified by adversary.**
- **Hardening database system vs. external threats insufficient:**
  - **Hard to identify correct security configuration,**
  - **Adversaries often have insider access, and**
  - **Adversaries constantly invent new attack vectors.**

# Background



## Top 10 Database Threats

1. Excessive Privilege Abuse
2. Legitimate Privilege Abuse
3. Privilege Elevation
4. ~~OS Vulnerabilities~~
5. **SQL Injection**
6. ~~Weak Audit Trail~~
7. ~~Denial of Service~~
8. Protocol Vulnerability
9. Weak Authentication
10. ~~Backup Exposure~~

# Objective



**Establish techniques for monitoring database query activity to flag suspicious actions by:**

- **Obtaining history of legitimate queries**
- **Automatically identifying profiles of legitimate query access plans**
- **Comparing incoming queries against profiles**
- **Flagging** anomalous queries as suspicious.

**Success Criterion (Year 1): Identify at least 80% of illegitimate activity with minimal false positives and performance degradation.**

# Activities

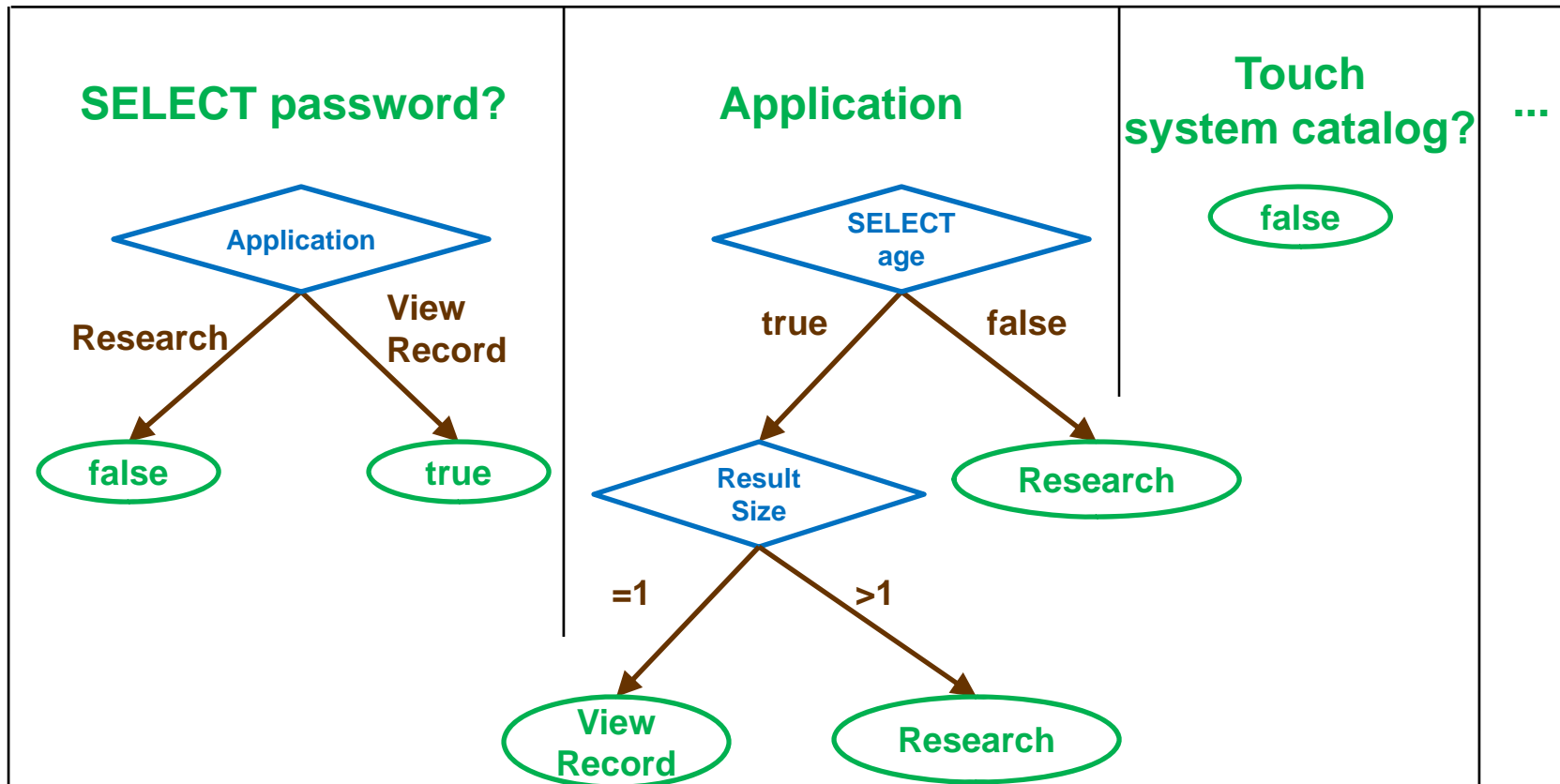


- Identified suite of relevant features, such as:
  - How many tuples from each relation,
  - Which attributes are selected/projected/joined,
  - Expected result size, etc.
- Built **profiling software** based on Cross-Feature Analysis
  - Complication: All samples assumed to be legitimate
  - Predict each feature based on other features
  - Previously used to identify mobile network intrusions.
- Built **monitoring software**
  - Monitor inserted between optimizer and execution engine
  - Features extracted and compared to profiles
  - Monitor can suppress query if desired.

# Highlight

- **Cross-Feature Analysis**

- Machine learning with only positive examples
- For each **feature**, learn a model in which it is the **class**.



# Demonstration



MITRE Health Record System  
MHR S

HOME :: DX FREQUENCY :: February 22, 2009

My Patient Record

Patient ID:  Password:

**Adversary crafts malicious input: 99' ; --**

**Without security monitor: Adversary bypasses password and retrieves record 99!**

**With security monitor: Query omits password field—identified as anomaly and denied.**

# Impacts



- **Our work applies whenever critical data are stored in relational database systems, for example:**
  - **MITRE's Time Reporting System**
  - **TRANSCOM's Global Transportation Network and Global Decision Support System**
  - **Electronic Medical Record systems.**
  
- **Research of interest to academic communities:**
  - **Intrusion detection (new system to protect)**
  - **Database community (new topic for additional research)**
  - **Machine learning (new unsupervised learning algorithm).**

# Future Plans

- Identify **exfiltrated** records
- Use schema to find **related** records
- Combined report hints at adversary's intent.

