Insider Threat Detection

Using Lightweight Media Forensics

Homeland Security
Science and Technology

Cyber Security Division
2012 Principal Investigators' Meeting

October 10, 2012

Nicole L. Beebe, Ph.D.
Assistant Professor
The University of Texas at San Antonio
Nicole.Beebe@utsa.edu
210-458-8040 (210-269-5647 cell)

Outline

- Intro
- Technical Approach
- Milestones, Deliverables, Schedule
- Technology Transition Plan
- Quad Chart

Intro

- TTA #4 Insider Threat
  - Focus on detection (vs. prevention or deterrence)
  - Detect anomalous accumulation of data
    - Indication and warning of exfiltration threat
    - Alert to criminal and/or exploitable behavior
- Research Team
  - PI: Simson Garfinkel, Ph.D.
    - Naval Postgraduate School
  - Co-PI: Nicole Beebe, Ph.D.
    - The University of Texas at San Antonio
    - UTSA Co-Investigator, Dr. Daijin Ko, Ph.D.

Technical Approach
The Problem

Homeland Security
Science and Technology

- Indication and warning of exfiltration threat when
  - Information collected IAW access permissions
  - Data collected, transmitted, and deleted quickly
  - No “signature” exists for behavior
- Past detection efforts focused on
  - Monitoring access patterns
  - Defining signatures
- Detection solutions must
  - Be scalable to large, diverse, dislocated organizations
  - Be lightweight and not impact computational operations
  - Accommodate fluidity and variety in workforce

Technical Approach
The Solution

Homeland Security
Science and Technology

- Detect threatening insiders by detecting deviations in their individual storage profile on their workstation
  - Deviations in storage profile
    - Data focused (not quantity/quota)
    - Types (file/data type and content based)
    - Deviations measured from variety of baselines
      - Individual, role-based, department, organization
  - Tunable anomaly alert sensitivity
    - To balance false positive – false negative trade-off
    - To increase monitoring on individuals or units

Technical Approach
The Solution

Homeland Security
Science and Technology

- Benefits
  - System agnostic
    - Leverages but does not rely on file system
  - Scalable
    - Uses reliable sampling of stored data
  - Does not rely on policy statement or discovery
  - Does not require Internet connectivity
  - Does not rely on insider threat modeling or “signatures”
  - No intellectual property costs (uses GOTS)

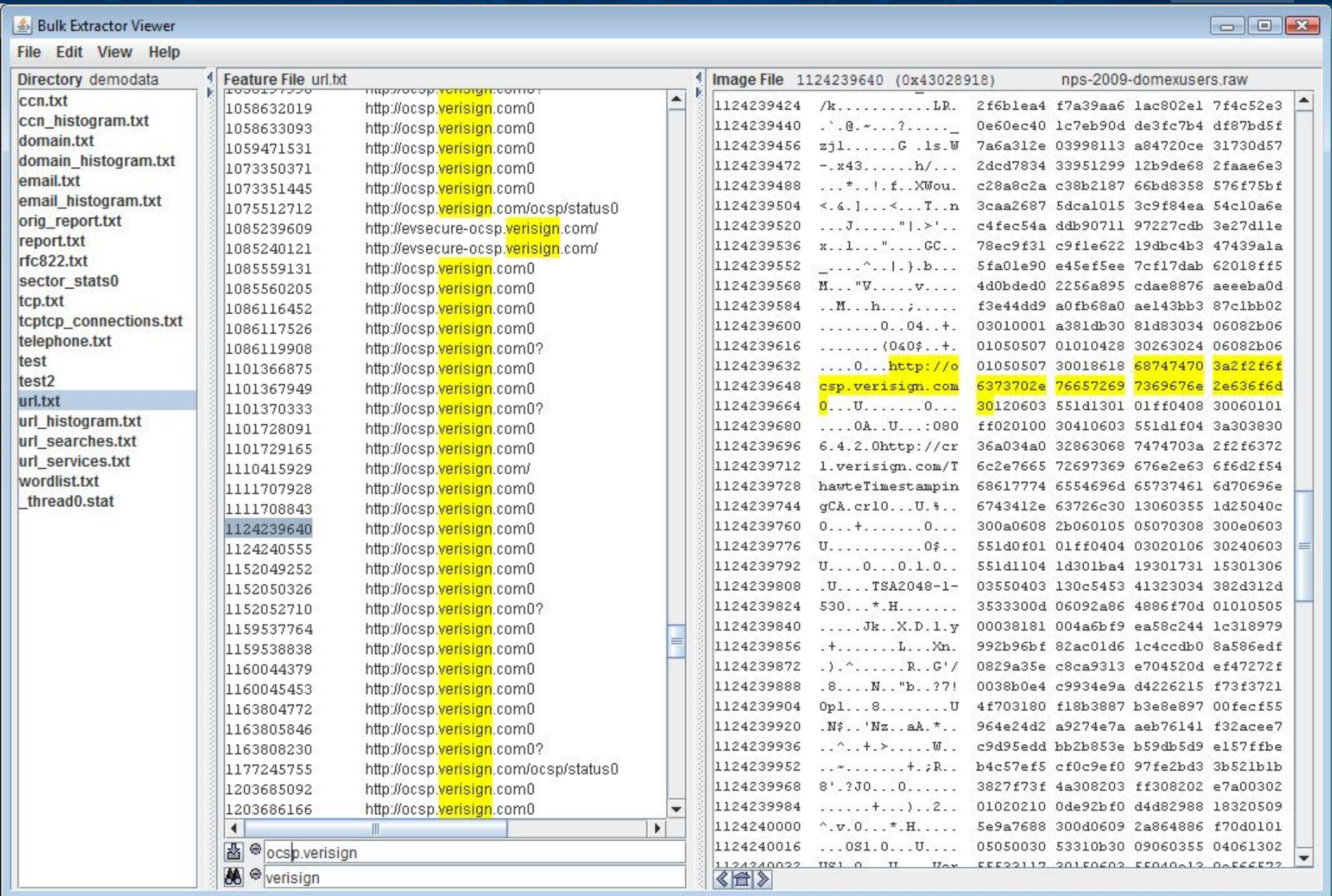
Technical Approach
The Solution

Homeland Security
Science and Technology

- Approach
  - Local, lightweight, covert, secured surveillance agents installed on workstations calculates storage profiles
    - `bulk_extractor`
      - Open-source, demonstrated capability, v1.3 just released
    - Random sampling
      - Of disk data to minimize computational overhead
      - Of interval to minimize anti-forensics efforts
  - Central management console
    - Surveillance agents send line-based profile data for analysis and anomaly detection
    - Statistical cluster analysis detects outliers

Bulk Extractor Viewer Screen Shot

Homeland Security



Technical Approach
Technical Challenges

- Science: Clustering algorithm selection/development
  - Accuracy and speed trade-off of extant algorithms
  - Develop combinatorial algorithm to improve accuracy
  - Need for automated parameter selection amidst noisy data
  - Feature selection

- Engineering: Visualization component
  - Visual representation of anomalies
  - Facilitate analyst drill-down

Milestones & Schedule

- Year 1
  - Milestone: Develop and test lightweight media forensics agent for Windows workstations
    - Done: Bulk Extractor v1.3 released
    - To do: Enhancements and sampling engine
  - Milestone: Develop and test optimal clustering algorithm
    - Leverage past research at UTSA
      - Self-organizing neural nets in digital forensics
      - Automated parameter estimation heuristics
      - Precision INdex (PIN) measure for combining classifiers

Milestones & Schedule

- Year 2
  - Milestone: Develop and test management console tools, data aggregation agent, visualization components on limited scope production network at NPS
    - Currently investigating adaptation of open source network collection and remote management systems
- Year 3
  - Milestone: Final development of outlier detection and visualization technology based on test results and large-scale testing on partner network(s)

Deliverables & Schedule

	Knowledge Products (Tech Report / Academic Pub)	**Software**
Year 1	- Feature Extraction Technique - Anomaly Detection Algorithm	- Workstation Agent (v1)
Year 2	- Visual Component - Small-Scale Network Test	- Workstation Agent (v2) - Server-Side Management Tool (v1)
Year 3	- Large-Scale Network Test	- Workstation Agent (v3 as needed) - Server-Side Management Tool (v2)

- Progress reporting all years
  - Monthly Program Reports
  - Quarterly Status/Financial Reports
  - Quarterly Detailed Technical Reports
  - Annual Program Report
  - Final Technical Report

Technology Transition Plan

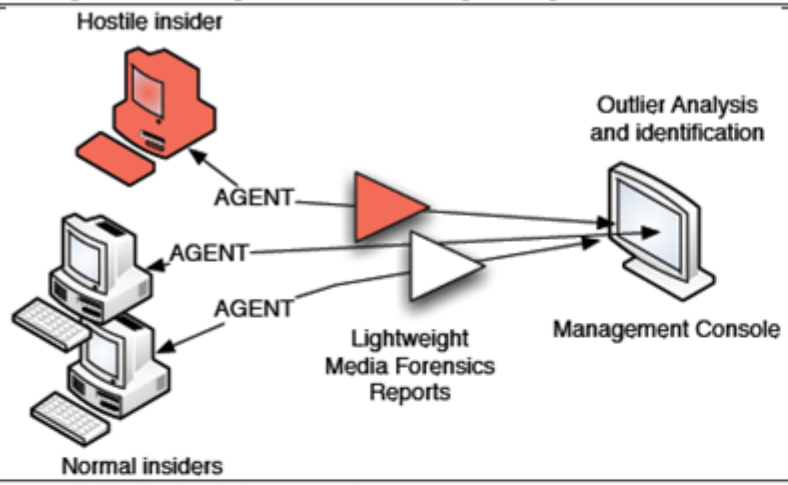
Homeland Security
Science and Technology

- Open source releases
  - Bulk Extractor (BE)
  - BE sampling engine
  - BE plug-in for clustering engine
  - Clustering engine
  - Visualization front-end for clustering engine
- Create GOTS product that integrates all above, plus
  - Secure transmission of data
  - Data aggregation agent
- To get to pilot: production network tests with injects
  - DETER & PREDICT are not applicable
- PI Garfinkel has extensive experience transitioning technology into open-source, commercial, and GOTS products

BAA Number: Cyber Security BAA 11-02

Offeror Name: Simson Garfinkel / NPS

Detecting Threatening Insiders with Lightweight Media Forensics



Operational Capability:

- Performance target: Able to scan every workstation in an organization on a regular basis and find systems that have data divergent from historical trends or organizational norms.
- Key Parameters: Types of data tabulated; Kinds of features extracted; Sensitivity of outlier detection.
- Cost of ownership: No IP cost (GOTS-owned technology); deployed via existing management systems; unknown monitoring costs.
- Proposal addresses TTA #4: Insider Threat.

Proposed Technical Approach

- Hostile insiders are detected by comparing their storage profile with the storage profile of others in their organization and looking for outliers. This is an advance over the signature-based state-of-the-art and accomplishes the BAA's goal of addressing "scale and diversity of the computing infrastructure," since the approach automatically scales and accommodates different storage profiles at different organizations. It also eliminates the need for policy discovery, since instead of looking for policy violations, it merely looks for people who are storing different kinds of data.
- In the base period a prototype agent, aggregation system, and outlier identification algorithms will be developed.
- We have a command-line forensic tool (*bulk_extractor*) that extracts all of the features we plan to use for outlier analysis. The tool is currently used by law enforcement.
- Work to date has demonstrated that outlier analysis is a powerful tool for finding hard drives containing large quantities of financially sensitive information.
- Related effort: We continue to develop *bulk_extractor* for use by law enforcement as a stand-alone forensics tool.

Schedule, Cost, Deliverables, and Contact Info:

- **Year 1:** Demonstration of agent and feasibility of outlier analysis.
- **Year 2:** Demonstration of tool on small network with real data and synthetic hostile insiders.
- **Year 3:** Demonstration of tool on one or more enterprise networks.
- **Deliverables:** Software; Research Papers;

Offeror: Simson L. Garfinkel, Naval Postgraduate School
Monterey, CA 93950
phone: 831.920.8131
slgarfin@nps.edu

Administrative POC: Danielle Kuska
Monterey, CA 93943-5138
831.656.2099 (voice)
dkuska@nps.edu