A Multi-Disciplinary Approach for Countering Insider Threats

Robert DelZoppo, Eric Brown, Matt Downey: ***Syracuse Research Corporation***

Michael D'Eredita, Elizabeth D. Liddy, Joon S. Park, Anand Natarajan, Svetlana Symonenko, Shuyuan M. Ho : ***Syracuse University***

Secure Knowledge Management (SKM 2004)

September 23-24, 2004

Marriott Buffalo-Niagara

Amherst, NY USA

ARDA

Insider Threat

- Mission-critical information = High-value target
- Threatens US Intelligence Community (IC), other Government organizations and large corporations
- Probability is low, but impact is severe
- Types of Threat posed by malicious insiders
  - Denial of service
  - Compromise of confidentiality
  - Compromise of integrity
- High complexity of problem
  - Increase in sharing of information, knowledge
  - Increased availability of corporate knowledge online
  - "Low and Slow" nature of malicious insiders

ARDA

Malicious Insider, examples

Robert Hanson: (1985-2001)

- Compromise: Exfiltrated over 6000 pages of classified material
- Impact:
  - Divulged Intel capabilities of FBI and other agencies
  - Identified three Soviet double agents (1 imprisoned, 2 killed)
- Cyber Activities:
  - Frequent need-to-know "violations"
  - Frequent queries looking for signs of an investigation targeting him

Brian Patrick Regan: (1999-2001)

- Compromise: Removed and hid over 800 pages of classified material, email contact to leaders in Iraq, Libya, and China
- Impact:
  - Suspected acquisition of classified imagery and reports to Iraq
- Cyber Activities:
  - Frequent need-to-know "violations"
  - High volume printing; Encrypted emails

Characteristics of Malicious Insider Behavior (current, projected)

- Technically competent to highly-skilled
- Attempts to cover up, destroy evidence
- Sophisticated search / query techniques
- Abuses security clearance to gain access to information (violates "need to know")
- Downloads data to new devices (e.g., USB thumb drive)
- Encrypts data
- Changes system logs to hide activity
- Uses "stealthy" techniques to communicate with handlers (e.g., encrypted email)

Approach

- **Staged**: Detect anomalies in user behavior from cyber observables and, based on these anomalies, assess the risk of malicious insider behavior
- **Multi-Perspective**: Detect anomalies in user behavior considering *user-to-user, user-to-content, user-to-resource* relationships
- **Multi-Disciplinary:**
  - **Social Network Analysis** (SNA) - Apply concepts from SNA to detect anomalies in social behavior [user-to-user]
  - **Semantic Analysis** (SA)- Leverage Natural Language Processing (NLP) and machine learning techniques to analyze the textual data associated with insiders at a semantic (conceptual) level [user-to-content]
  - **Composite, Role-based Monitoring** (CRBM) – Analyze insider activity based on the organizational, application and operating system roles. [user-to-resource]

Research Objectives

- Advance the state-of-art in Insider Threat Countermeasures by developing techniques to:
  - Model behavior of insiders operating in an IC-based context
  - Distinguish between expected and anomalous user behavior
  - Detect indicators of malicious insider behavior (MIB)
  - Assess indicators of MIB for potential threat to the confidentiality and integrity of information.
- To reduce the overall effort in countering threat from malicious insiders:
  - Reduce the size of the problem space to a manageable number of indicators a system security / assurance administrator would need to look at
  - Provide early awareness of risk elevating situations

ARDA

Research Objectives, cont'd

- To provide a robust solution which:

Has Breadth	Incorporates a wide range of observable types and can assess multiple types of risk
Has depth	Can analyze observables at fine-grained levels (e.g., semantics)
Is scalable	Can model behavior at multiple levels (e.g., insider, role) and is minimally impacted as # of insiders increases
Is extensible	Can be extended to incorporate new threat scenarios and other sources of indicators (e.g., anomaly detectors)
Is reusable	Modules could be reused in another system or context

Assumptions

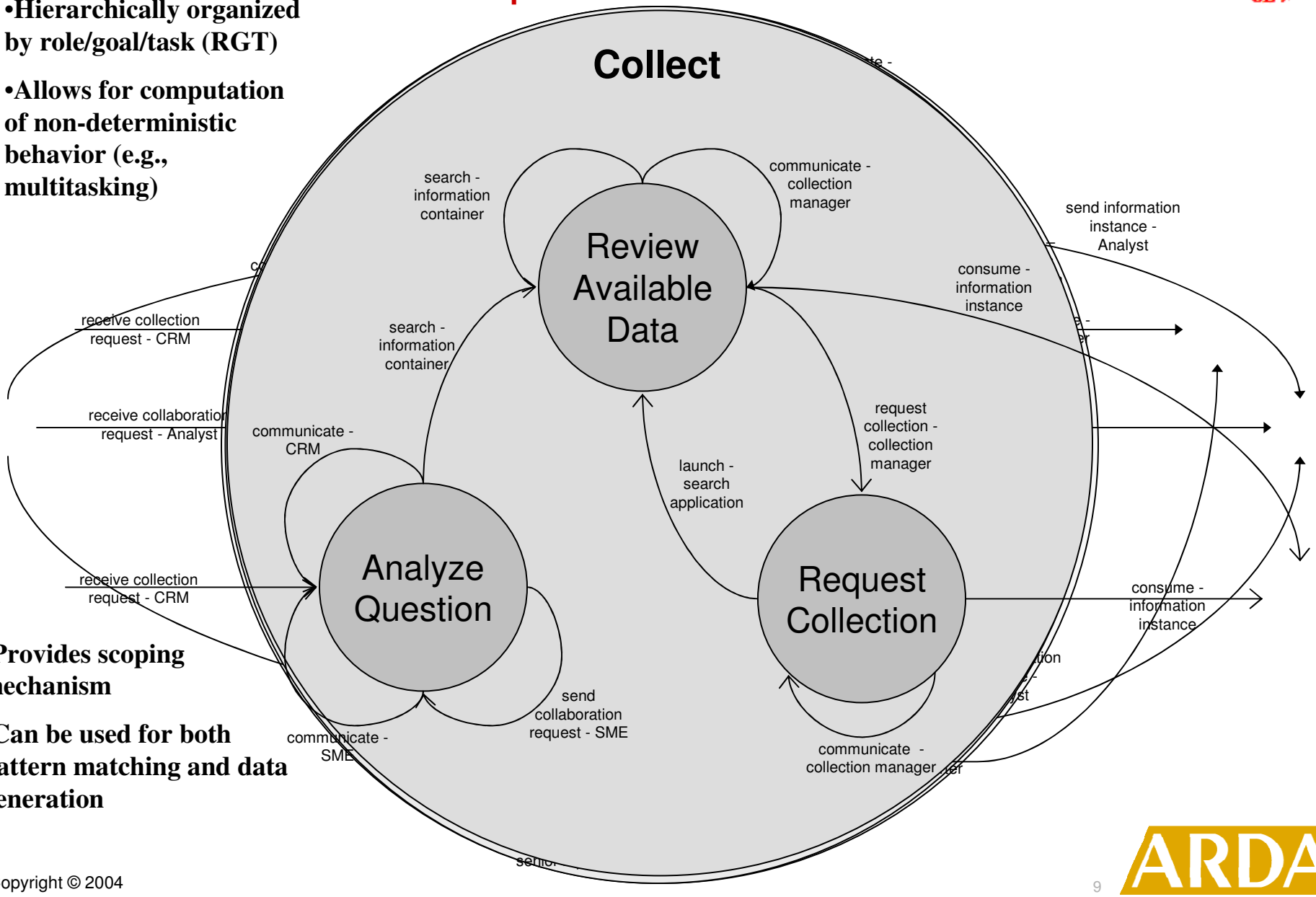
- Insiders with similar roles, goals and tasks will have similar behavior.
- Malicious insider behavior will differ, to a measurable degree, from behavior of typical insiders.
- Insiders' actual behavior will be discernable through cyber-observations from sensors which currently exist or could be constructed.
- Anomaly-based or signature-based methods, by themselves, are insufficient for identification of Insider Threats.

ARDA

Approach/Methodology

Expected Behavior Model

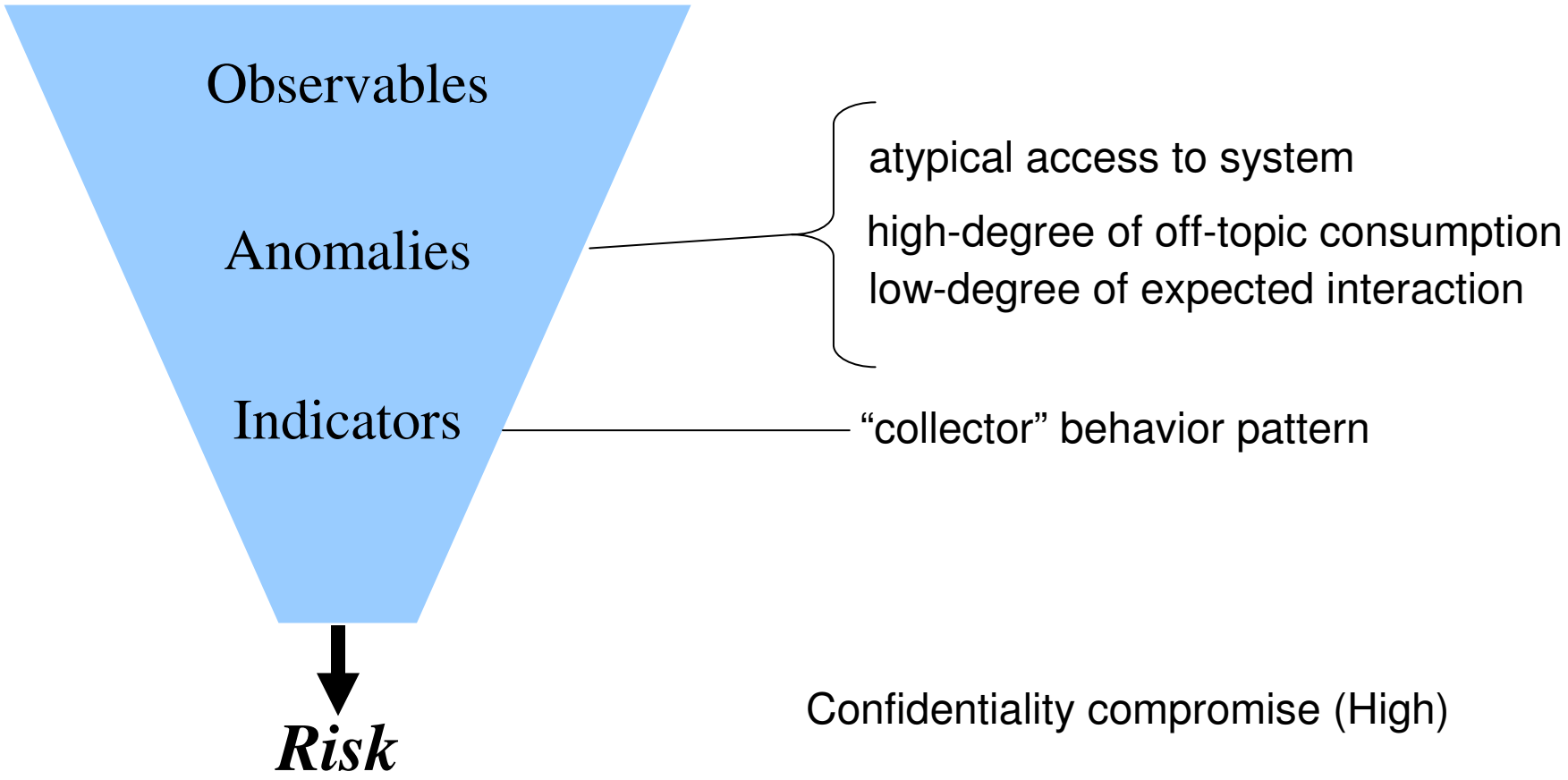
- **Hierarchically organized by role/goal/task (RGT)**
- **Allows for computation of non-deterministic behavior (e.g., multitasking)**



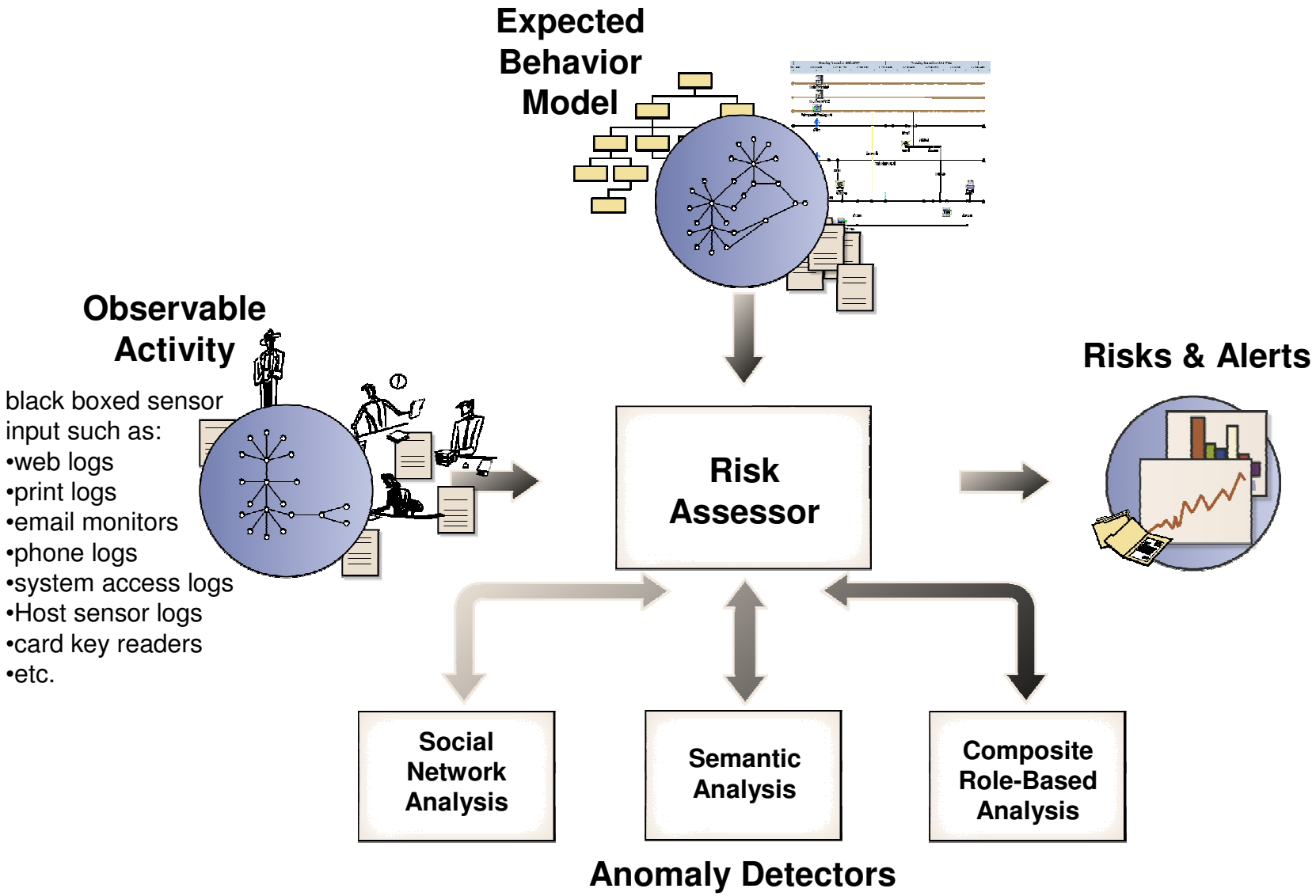
- **Provides scoping mechanism**
- **Can be used for both pattern matching and data generation**

Approach/Methodology:
Risk Assessment

Risk is identified as indicators are asserted; indicators are asserted from the anomalies detected



System Overview



Current Work: Relational Matrix Analysis Tool (user-to-user, user-to-resource)

Observables (from Scenario)

```
<Observable>
  <Name>Terry</Name>
  <Role>analyst</Role>
  <Toi>Biological Weapons</Toi>
  <Aoi>Russia</Aoi>
  <Task>Report</Task>
  <Method>leave VM</Method>
  <ResourceLabel>Smith</ResourceLabel>
  <ResourceType>senior reporter</ResourceType>
  <Time>1071032734</Time>
</Observable>
```

Resource Restrictions: TOI, AOI, task

Insider Restrictions: role, TOI, AOI, task

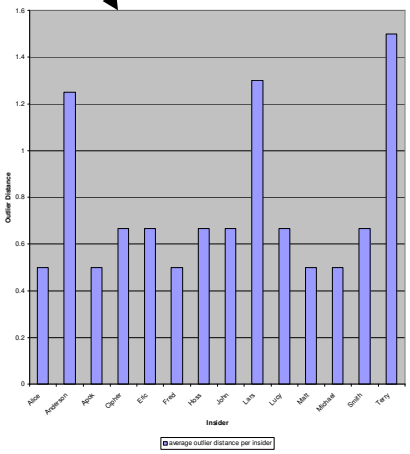
Method Restrictions

Relational Matrix Analysis Tool

Insider vs. Resource Matrix

	datasource2	datasource4	datasource6	datasource7	datasource8	datasource9
Alice	1	0	0	0	0	0
Anderson	0	0	2	4	1	0
Apok	0	0	0	0	0	0
Cipher	0	0	1	0	0	1
Eric	0	0	1	0	0	1
Fred	0	0	0	0	0	0
Graham	0	4	0	0	0	0
Herb	0	0	0	0	0	0
Hoss	1	1	0	1	0	1
Joe	0	0	4	0	0	0
John	1	0	0	0	1	0
Lars	0	0	4	2	0	0
Lucy	0	1	0	0	1	0
Matt	0	1	0	0	0	0

Outlier Indicators and Analysis

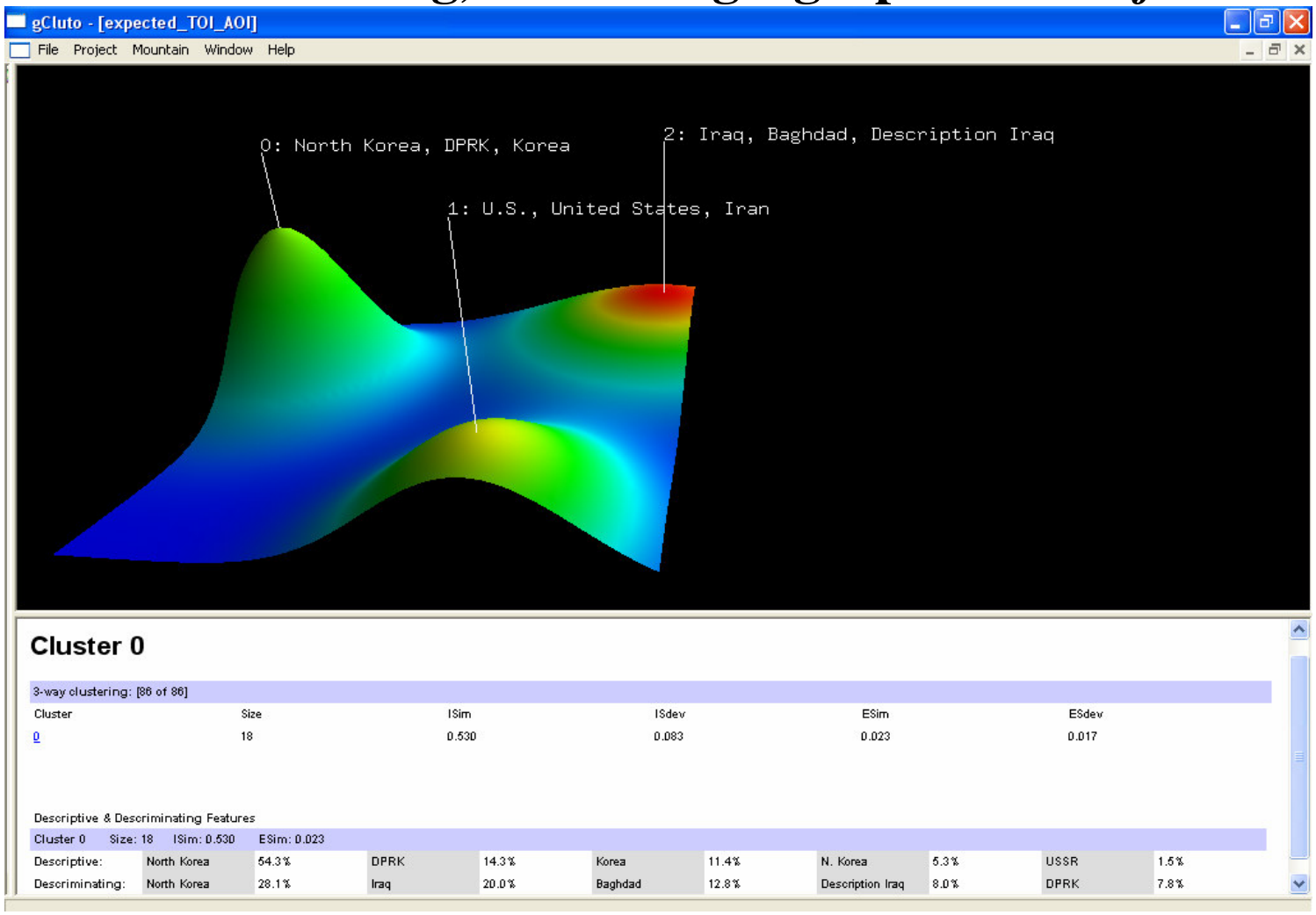


Given:

- ❑ Observables
- ❑ Method Restrictions
- ❑ Insider Restrictions
- ❑ Resource Restrictions
- ❑ **Generate Relational Matrices**
  - Based on insider (constrained by RGT) versus a hierarchy of resources, goals, and interaction methods
  - Comparison level: specific (explicit resource) or generic (resource type)
- ❑ **Perform Outlier Analysis**

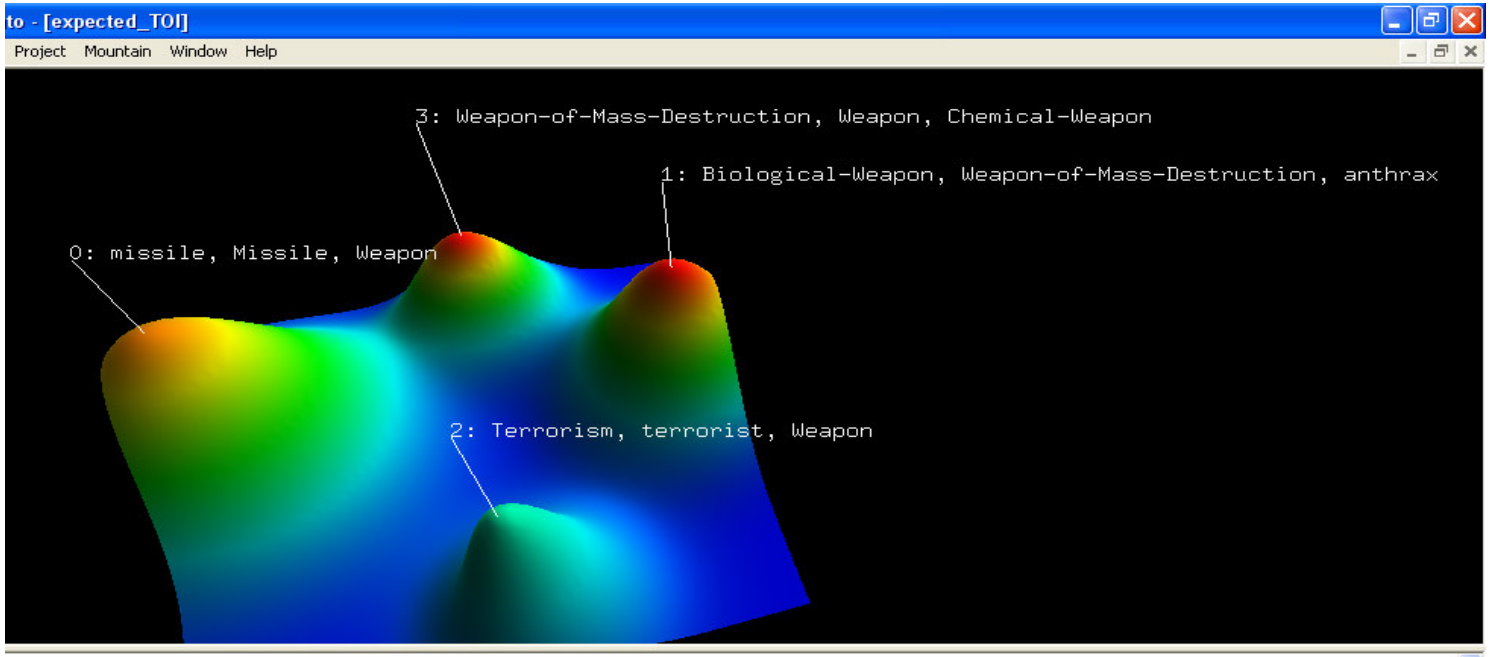
Current Work: Semantic Analyses (user-to-content)

Document clustering, based on geographic *area-of-interest*



Current Work: Semantic Analyses (user-to-content)

Document clustering, based on *topic-of-interest*



Cluster 0

4-way clustering: [31 of 31]

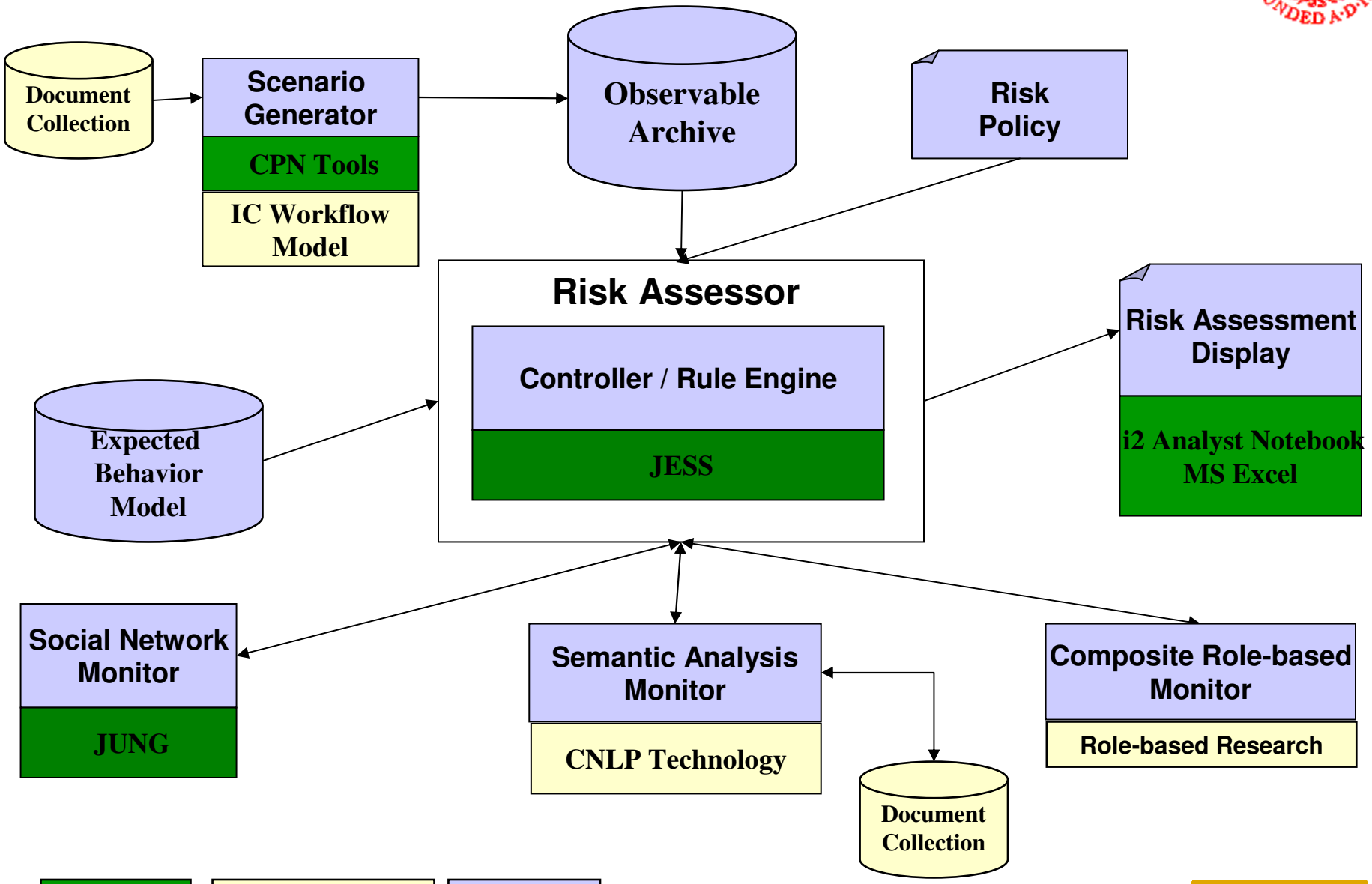
Cluster	Size	ISim	ISdev	ESim	ESdev
0	15	0.936	0.057	0.254	0.108

Descriptive & Descriminating Features

Cluster 0 Size: 15 ISim: 0.936 ESim: 0.254

Descriptive:	missile	37.7%	Missile	37.7%	Weapon	23.5%	Weapon-of-Mass-D	0.4%	nuclear_weapon	0.3%
Descriminating:	missile	23.9%	Missile	23.9%	Weapon-of-Mass-D	21.5%	Biological-Weapon	11.0%	anthrax	6.4%

System Architecture



Scalability of Solution

- High Scalability / Extensibility
  - Other anomaly detectors can be added to provide additional indicators
  - Risk Assessment Policy provides a means for writing new rules and sets of rules
- Generalizability
  - Methodology provides abstraction mechanisms for managing complexity
  - Approach can be generalized to other domains
- Reusability / Interoperability
  - Anomaly detectors can provide indicators to other types of systems
  - XML-based interfaces – provide "loose" couplings between modules

ARDA

Limitations/Vulnerabilities

❑Non-cyber activities

- Mitigation: Security Administrator Application for entering / managing non-cyber indicators

❑Undetected cyber observables:

- Most non-textual media (Images, Audio, Video)
  - » Example: Communications analyst inappropriately retrieving images unrelated to task
  - » Mitigation: Analyze image meta-data to provide basic analysis of image content
- Anonymous user behavior – Guest, and other potentially anonymous activities such as access through web-based applications
  - » Mitigation: Can still monitor to identify risk
- Account "masquerading"
  - » Mitigation: Focus on individual insiders; detect shifts in behavior

ARDA

Summary

- Currently under experimentation using controlled simulation with synthetic data sets (scenarios):
  - Baseline scenario – observables under normal conditions
  - "Threat" scenarios – baseline scenario with anomaly injection
  - Includes supporting UNCLASSIFIED document collections on a variety of topics (e.g., Terrorism/WMD)

- Preliminary results indicate
  - *Role-Goal-Task*-orientation of Expected Behavior Model provides a basis for modeling context-dependent behavior
  - Relational Matrix approach very well suited to anomaly detection in entity-to-entity interaction
  - Semantic Analysis approach works well to identify off-topic information access

ARDA

Acknowledgements

- Advanced Research and Development Activity (ARDA) Advanced Countermeasures for Insider Threat (ACIT) Program (*sponsor)*
- Other ARDA Programs
  - Cyber Indications & Warning (CIW) Workshop (MITRE, Aug 03)
  - Advanced Question & Answering for Intelligence (AQUAINT)
  - Novel Intelligence from Massive Data (NIMD)
- Mitigating the Insider Threat to Information Systems - #2; Workshop Proceedings (RAND, Aug 00)

ARDA