Rand Waltzman
Program Manager, Information Innovation Office

Anomaly Detection at Multiple Scales

DARPA Cyber Colloquium
Arlington, VA

November 7, 2011

The problem...

- Why didn’t we see it coming?
  - Robert Hanssen
  - Aldrich Ames
  - Ana Belen Montes
- The trail of evidence was obvious *after the fact*
- Why is it so hard to pick up the trail *before the fact*?
- Answer:
  - Difficult to characterize anomalous v normal behaviors
  - Malicious activities distributed over time and cyberspace
  - Weak signal in a noisy background
  - Enormous amount of data

How much data?

- Find evidence of an insider at Fort Hood:
  - 65,000 soldiers at Fort Hood
  - Represent the e-mail and text message traffic as a graph
    - Nodes represent persons
    - Links represent e-mail or text messages
  - Analyze 47,201,879,000 links between 2,336,726 nodes over one year

- Find evidence against one person over the entire DoD:
  - E-mail and text message traffic only
  - Analyze 755,230,064,000 links between 37,387,616 nodes over one year

- And this does not include web-searches, file accesses, applications run, and many other forms of cyber observable behavioral data.

Anomaly Detection at Multiple Scales (ADAMS)

- Focus on malevolent insiders that started out as good guys
- Research organized into four coordinated thrusts
  - Topic analysis
    - Develop signatures for areas of responsibility
    - Detect straying from tasked topic areas, or produces unexpected content
  - System use
    - Temporal sequences of system and file accesses
    - Patterns of behavior
  - Social interactions and networks
    - Indicators
    - Social exchanges
  - Psychological state
    - Personal temperament and mental health
    - Distress, instability, or other vulnerability

Detect the signs that they are turning
before or shortly after they turn

Example: Insider Threat Scenarios in StackOverflow.com

- Data set with 645 thousand users, 5.5 million question posts, and 12 million responses
- Use of human controlled alias accounts (aka Sock Puppets) for voting fraud
- 9 inserted sock puppet voting fraud schemes
- Oregon State University graph analytics detected 7 out of 9 schemes and 310 out of 535 sock puppets.

Contact Information

- If you would like to pursue topics discussed in this presentation, please send your ideas to
- rand.waltzman@darpa.mil