Situational Awareness for Cybersecurity

Tim Finin and Anupam Joshi

Computer Science and Electrical Engineering

University of Maryland, Baltimore County

21 October 2011
Situational Awareness

• Awareness of what’s happening around you to understand how information, events, and actions will impact your goals & objectives, now and in future

• A common theme in as we become more *instrumented* and *interconnected*
  
  Cybersecurity, cyber-physical systems, hot conflicts, homeland security, disaster relief, health-care, IT services, network operations & management ...

• Applies to people, smart interfaces, sensors, AI, wireless networks, embedded systems, streaming data, image processing, SIGINT, HUMINT, smartphones, etc.

• Challenges for distributed, dynamic & interconnected systems
Current work at UMBC

• We’re exploring building situationally aware systems at UMBC

• I’ll briefly touch on examples for social media and mobile computing

• Then describe some ongoing work for cybersecurity supported by Northrop Grumman
E.g.: Response System for Gulf Oil Spill (NSF)

• Mine data from social media to improve oil spill trajectory model used by NOAA Emergency Response Division

• Used surface winds, currents, oil spill rates, boundary locations, oil dispersion, type of oil, diffusion coefficients, drift velocities, etc.

• Coupled with atmospheric, hydrologic and storm surge models

• Performed regressions, fishery impacts, animations
E.g.: Smartphones sharing context

• Platys is an $1.8M NSF project with Duke & NCSU

• Sensor-rich android phones learn to recognize their user’s context: what, who, where, when, how ...

• Information is shared securely and with appropriate detail following user specified privacy policies and context

• The shared information helps other devices learn faster and provide better services
Pre-facto threat/vulnerability detection and monitoring
Our Approach

1. Detect potential new vulnerabilities from Web descriptions and discussions, extract information and map to IDS KB [ebiq.org/p/540]

2. Recognize potential attacks and intrusions in data from low level intrusion detection systems and map to IDS KB [ebiq.org/p/63]

3. Integrate and reason over results of (1) and (2) to identify actual attacks [ongoing]
Tracking Security Vulnerability Info

- Working with Northrop Grumman on system to discover new software vulnerabilities and track their spread and evolution
- Use human language technology, machine learning and cybersecurity knowledge bases to extract, evaluate and fuse structured information from Web, chat rooms, and social media
- Prototype automatically adds to, updates and maintains a structured knowledge base
Microsoft Windows Knowledge Base Article 2507618 update is not installed on the system, which could allow an attacker to exploit the following vulnerability:

Microsoft Windows is vulnerable to a stack-based buffer overflow, caused by improper bounds checking when validating the parameter values of specially-crafted OpenType fonts by the OpenType Compact Font Format (CFF) driver. By persuading a victim to visit a specially-crafted Web site containing a malicious OpenType font, a remote attacker could overflow a buffer and execute arbitrary code on the system with the privileges of the victim.
Buffer overflow in Fax4Decode in LibTIFF 3.9.4 and possibly other versions, as used in ImageIO in Apple iTunes before 10.2 on Windows and other products, allows remote attackers to execute arbitrary code or cause a denial of service (application crash) via a crafted TIFF Internet Fax image file that has been compressed using CCITT Group 4 encoding, related to the EXPAND2D macro in libtiff/tif_fax3.h.
Linked Open Data Cloud 2007-11

Uses Semantic Web Technology to publish shared data & knowledge

LOD is the new Cyc: a common source of background knowledge

Data is interlinked to support integration and fusion of knowledge

2011: 31B facts in 295 datasets interlinked by 504M assertions on ckan.net
Knowledge-based intrusion detection

- Low level system data is compared to a model of the quiescent state
- Nonconforming data modeled in an IDS ontology and reasoned over to predict potential attack type
- Undercoffer, 2003, 2004
3 Integration, reasoning & prediction

• Update the IDS ontology & populate with data from NVD, CERT and NGC resources
• Improve and deploy cybersecurity information extraction prototype
• Integrate into the Linked Open Data cloud to exploit background knowledge
• Ingest low-level data from IBM Proventia Network Enterprise Scanner ES 750
• Evaluate in our UMBC lab environment to identify potential attacks
for info, contact finin@umbc.edu or joshi@umbc.edu or see http://ebiq.org/