Cyber Security @CEISARE

Professor S. Upadhyaya
Department of Computer Science and Engineering
SUNY at Buffalo

CSE 501 Talk
November 7, 2013

Outline

- Overview of Cyber Security Center at UB
  - Research and Education
- Motivation – Why Cyber Security?
- Types of Research Projects

CEISARE

- CEISARE designated as a National Center of Excellence in 2002 by NSA, DHS
  - Through a competitive process
  - We were one of 13 centers designated that year (36 across the country)
  - Today, there are 100+ centers

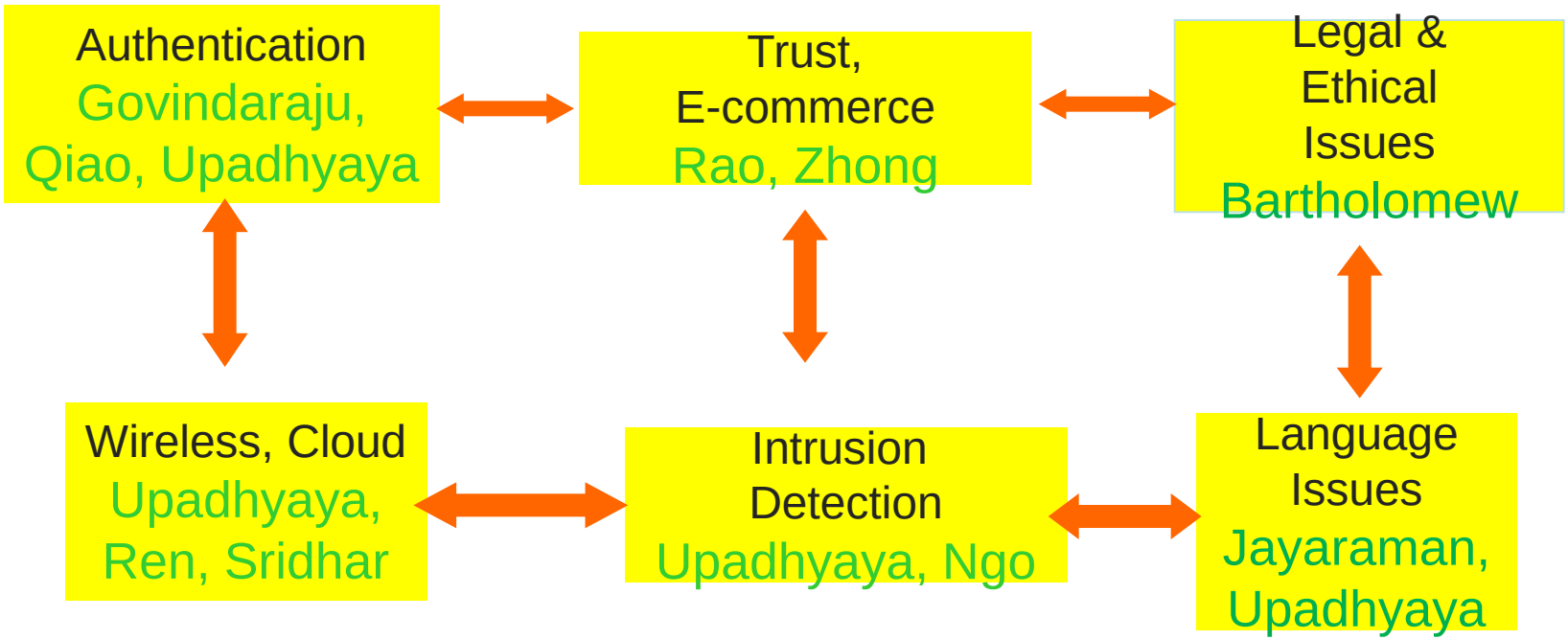
Graduate Certificate in IA

- Effort started with funds from DoD, 2003
  - Funding was to create a new integrative course in IA
- Two tracks – technical and managerial
- Requirements
  - 6 credits of core courses in the track
  - 5-6 credits of elective in the dept.
  - 3 credits of required integrative course
- Technical track
  - Core – Intro. to crypto, Computer security, Applied crypto, Wireless networks security (choose two)
- Managerial track
  - Core – Network management, E-Commerce security

CEISARE Courses

- Courses with IA Content
  - CSE 565 Computer Security (currently teaching in fall 2013)
  - CSE 566 Wireless Networks Security (spring 2015)
  - CSE 664 Applied Crypto and Computer Security
  - CSE 671 Security in Wireless Ad Hoc and Sensor Networks
  - LAW 629 Computers, Law, Technology and Society
  - LAW 645 Copyright
  - Law 956 E-Commerce Law
  - MGA 615 Fraud Examination
  - MGS610 Digital Forensics
  - MGS 650 Information Assurance
  - MGS 651 Network Management
  - MGS 659 E-Commerce Security
  - MGT 681 Intellectual Property
  - MHI 512 Ethical, Social & Human Factors in Medical/Health Informatics
  - MTH 529/530 Introduction to the Theory of Numbers I/II
  - MTH 535 Introduction to Cryptography
  - MTH 567 Stream Ciphers
- Other Technical Electives
  - http://www.cse.buffalo.edu/caeiae/advanced_certificate_program.htm

Computer Security & IA Faculty at UB



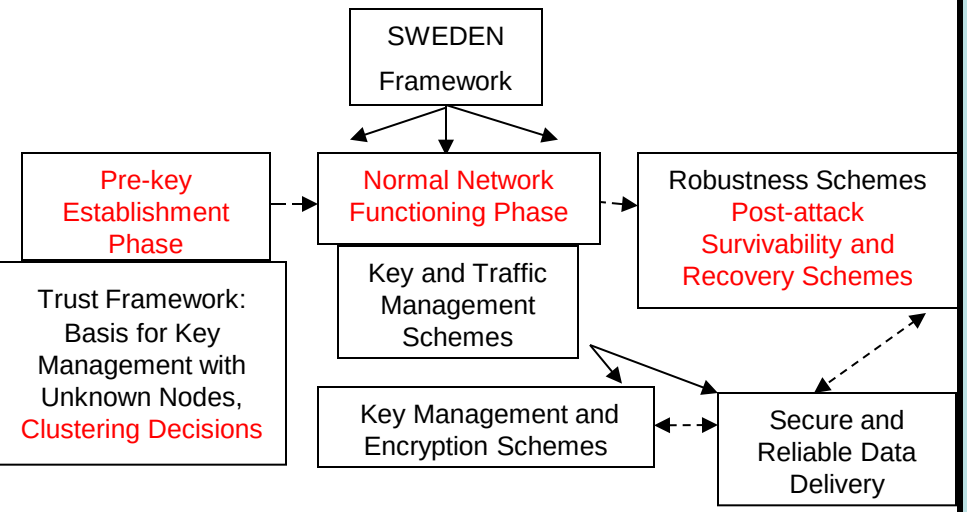
Types of Attacks

- Simple script kiddie attacks
  - Worms, viruses, DoS attacks, SQL injection
  - Examples: Sasser worm (2004), SoBig.F virus (2003),Yahoo (2000), UN Website defacing (2007)
- High profile attacks
  - Targeted at Intelligence communities, industrial plants, corporate networks
  - Examples: Hactivism/Anonymous at FBI (2011), Stuxnet at Iran's nuclear facility (2010), Operation Aurora at Google (2010)
- Threats to national security
  - Insider attacks
  - Examples: Bradley Manning leaked classified documents to WikiLeaks (2009), Edward Snowden leaked NSA documents to the media (2013)
- Nuisance attacks such as phishing

Research Projects

- Most federally funded
- Some industry funded
- Disciplines ranging from
  - Systems security
  - Networks security
  - Social networks

SWEDEN: A New Framework for Secure and Trusted Communications in Wireless Data Networks, Shambhu Upadhyaya, Funded by AFRL, NSF/Cisco, 2004-09



Goals

- Design decision making framework for nodes to establish keys with other unknown nodes
- Use this framework for cluster forming decisions in ad-hoc networks
- Improve on existing key management schemes and design secure data delivery schemes for enhanced reliability in data transfer
- Provide schemes for resiliency against attacks and post-failure recovery

Novel Ideas

- **Trust between the nodes used as a metric for decision making**
- **Differential encryption (header and payload differently) scheme for ad-hoc networks, and hashing based lightweight techniques for sensor networks**
- **Evaluating security of paths and nodes based on their relative position in the network**
- **Building in survivability in the network architecture proactively for surviving potential attacks**
- **Robustness, Recovery and Survivability Schemes**

Accomplishments

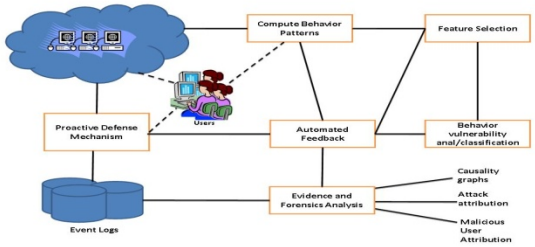
- **NSF/Cisco sponsored Wireless Security Lab**
- **Representative Publications:**
  - **IEEE Conference on Local Computer Networks (LCN), Tampa, FL, Nov 2004**
  - **IEEE ACM IWIA, College Park, MD, Mar 2005**
  - **IEEE Conference on Knowledge Intensive Multi-agent Systems (KIMAS), Boston, MA, Apr 2005**
  - **Secure Knowledge Management (SKM), Sep 2004**
  - **MMM 2007, St. Petersburg, 2007**
  - **IEEE Trans. Computers (forthcoming, 2014)**
- **Future Plans**
  - **Security Schemes for mesh networks**
  - **Security of smartphones**
  - **Privacy issues in smartphone applications**

Impact

- Graduate Students
  - M. Virendra (Ph.D., June 2008), M. Jadliwala (Sept. 2009), Ameya Sanzgiri (Dec. 2013 expected), Steve Salerno (M.S., June 2011), Sam Pezzino (M.S., June 2014)
- Publications
  - KIMAS 2005, SKM 2006, IEEE ICC 2007, MMM-ACNS 2007, IEEE SRDS 2007, Infocom 2008, WiSec 2009, IEEE TC 2014
- Funding Agency
  - Air Force Research Laboratory (2007-09)

Objective

- A unified behavior based framework for mitigating threats and damaging attacks on the Internet
- Address phishing, zero-day exploits, spyware, email authorship attribution, information leak in documents
- Hardware acceleration to support scalability



The approach

- Behavior Capture and Analysis (feature selection, simulated annealing)
- Behavior Based Monitoring and Detection (support vector machines)
- Attack Attribution and Forensics (causality graphs)
- Attack-Agnostic Framework (component based approach, implementing theories in hardware on modern CPUs)
- Validation (user studies)

State of the art in the area

- Malware on the Internet is rampant
- Behavior-based defense used successfully in real-world
- Extended to cyber-world by researchers at Columbia U.
- Behavior capture and correlation of applications using programming languages
- Behavior based monitoring for attack detection using statistical and rule-based algorithms
- Behavior based techniques for network forensics using causality graphs
- Designing new hardware for content processing and cryptography

Novel ideas

- Attack-agnostic framework to address all facets of security – attack protection, detection, response and forensics
- A holistic approach
- Proof-of-concept prototypes for anti-phishing, handling zero-day exploits, malicious email attribution, anti-spyware,
- information leak detection
- Hardware acceleration techniques to handle “pump and dump” malware
- – "Spycon: Emulating user activities to detect evasive Spyware”, IEEE Malware 2007 (Best paper award)

Impact

- Graduate Students
  - M. Chandrasekaran (Ph.D., June 2009), N. Pulera (M.S., June 2008), H. Alkebulan, (M.S., Dec. 2008), N. Campbell (B.S., Dec. 2008), V. Keshavamurthy (M.S., 2012)
- Publications
  - Ubisafe 2006, Malware 2007 (Best Paper Award), Albany IA Conference 2007, 2008, DNCMS 2011
- Funding Agency
  - DoD (2007-08), Intel Corporation (2008-10)

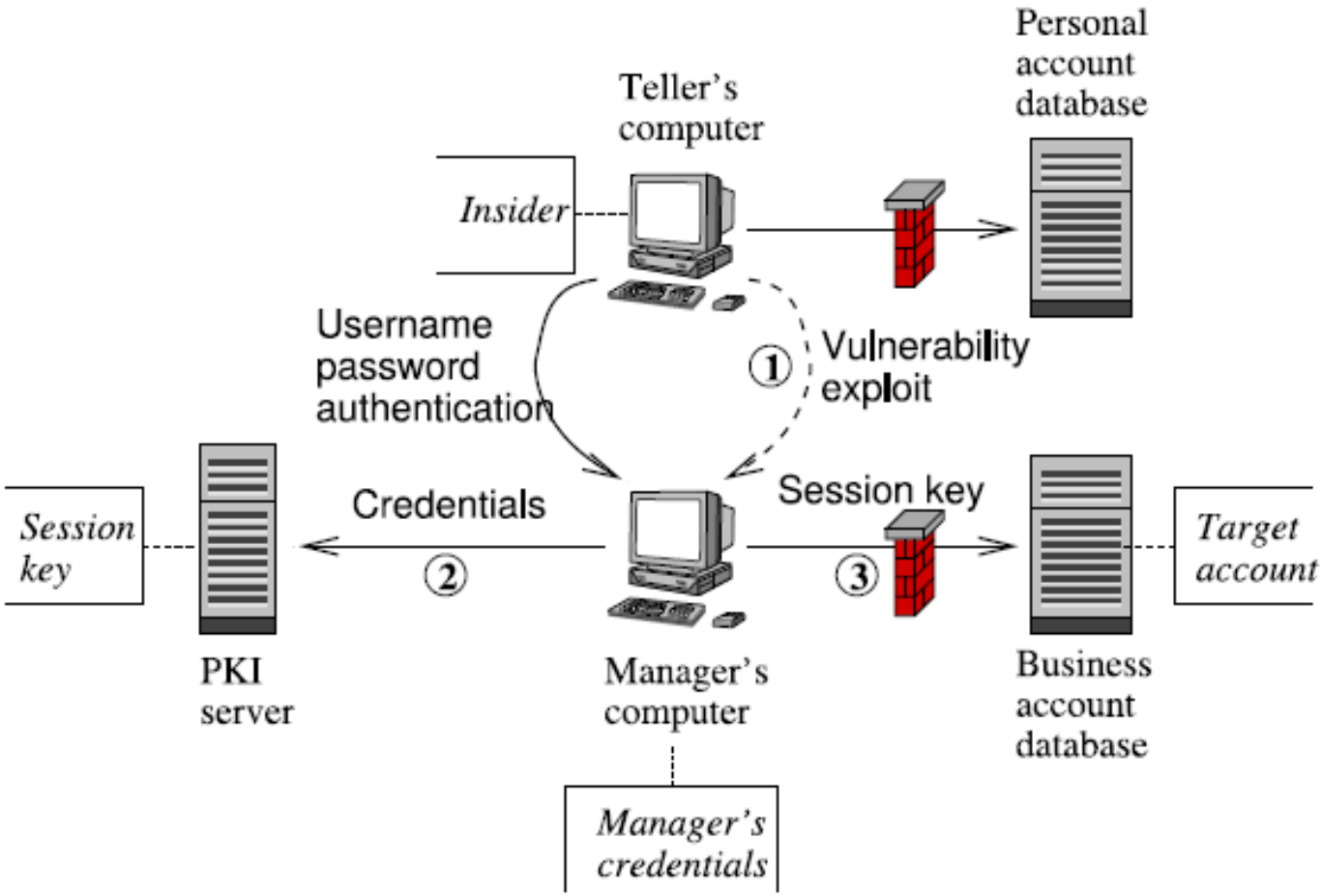
Ongoing Projects

- Insider threat assessment
- Cyber situation awareness
- Malware propagation in social networks
  - Twitter is vulnerable
  - Use epidemic theory to study malware diffusion
- Active authentication (recently funded by NSF)
  - One-time authentication is not sufficient
  - Use behavioral biometrics for continual authentication
  - Examples: keystroke dynamics, mouse movements
  - Extension of work to network of smartphones

Insider Threat – A Financial Institution Example

- Scenario
  - Every teller performs sundry personal accounting tasks
  - Manager endorses large transactions and also performs business transactions
  - The two databases are separated
  - All transactions to the DB are encrypted
  - Teller to personal accounts DB uses lower strength encryption
  - Business transactions require the manager to refer to a PKI server and get a session key
  - Both DBs are protected behind a firewall
- Attack
  - Teller knows the manager doesn't apply security patches regularly
  - Rogue teller exploits some vulnerability to compromise manager's account

Modeling the Attack (Physical Graph)



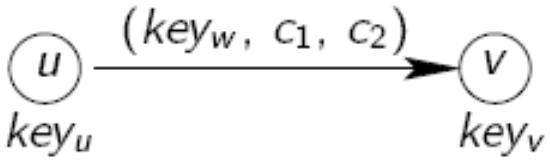
ICMAP (Info-Centric Modeler and Auditor)

- Information-centric modeling concept
- A Capability Acquisition Graph (CAG) generation for insider threat assessment
- Part of a DARPA initiative
- Ideas published in ACSAC 2004, IEEE DSN 2005, JCO 2005, IEEE ICC 2006, IFIP 11.9 Digital Forensics Conference 2007, Springer 2010, RAID 2010
- DOE SBIR (technology transfer in 2010-11)

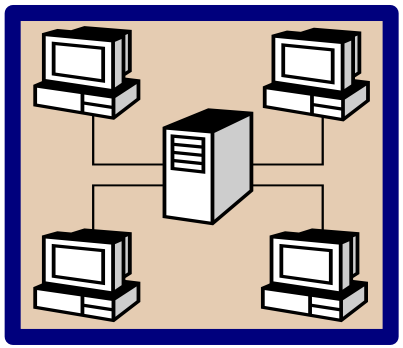
Basic CAG Model

Focus on an insider's view of an organization such as Hosts, Reachability, and Access Control

Model Component	Abstraction
Vertex	Hosts, People
Edge	Connectivity, Reachability
Key	Information, Capability
Key Challenge	Access Control
Starting Vertex	Location of insider
Target Vertex	Actual target
Cost of Attack	Threat analysis metric



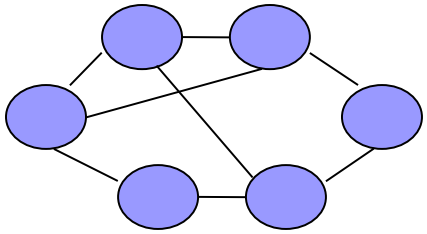
ICMAP Overview



Network topology

Network entity rules

ICMAP Engine



Cap. acquisition graph

vulnerabilities

Authentication mechanism

Social Eng. Awareness

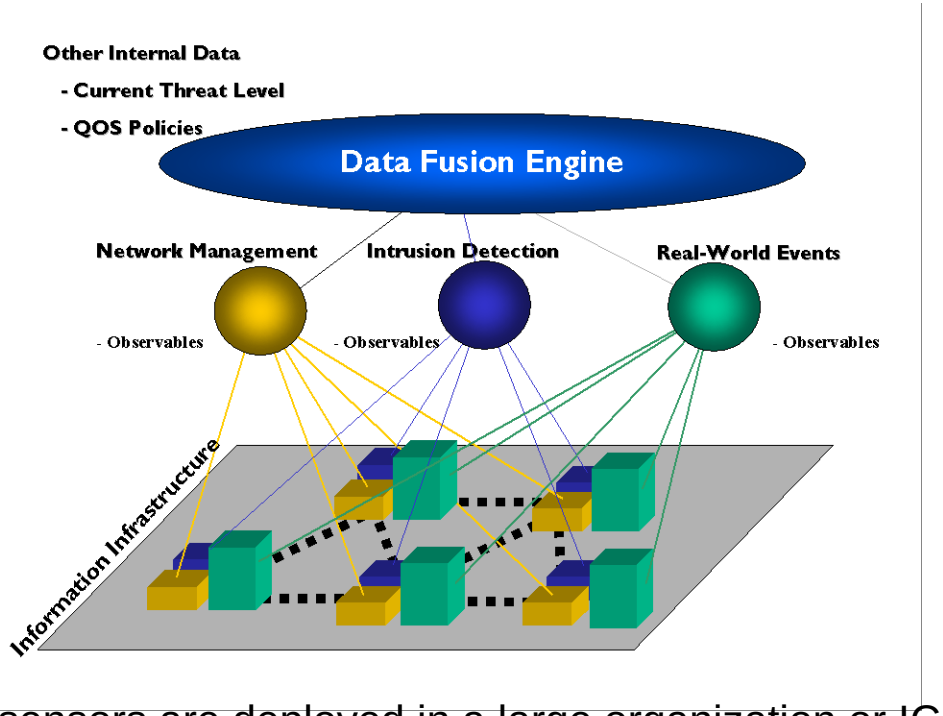
Cost Rules

Defense centric approach

Perform sensitivity analysis

feedback

High Level View of Situation Awareness



- Scenario
  - 100's of IDS sensors are deployed in a large organization or IC network
  - Sys logs, web logs, etc. are being collected
  - Alerts are streaming at high rate
- How would a security analyst make sense out of it
- Basically, you need a dash board display in simple terms

From Causation to Correlation

- Fareed Zakaria, Time Magazine contributor, July 8-15, 2013
  - (article on big data, big brother)

  - Today, instead of figuring out why a crime happens, they look at crimes and notice what events or behavior seem to precede them. The tricky work of turning info into knowledge has shifted from causation to correlation
- Situation Awareness
  - Forensics
  - Attack prediction/cyber warning

Solution Overview

- Generic attack model (derived from military model)
  - *Reconnaissance→Intrusion→ Privilege Escalation→Goal*
- Define semantic classes that convey stages in multistage attack progression
- Use semantic classes as basic constructs in high-level attack model (Guidance Template)
  - Reflects topology of network (cf. sentence construction)
- Map entire sensor signature set to semantic classes
- Correlate events based on IP address relatedness and attack-model to generate multistage attack hypotheses (Attack tracks) with *criticality values*
- Test, Validate and Analyze on realistic datasets

Data Dimensionality Challenges and Solutions

- Challenges for practical attack scenario detection
  - 10-20K alerts each day in organizational networks!
  - Huge log data
  - A good part of data is redundant (sensors may record redundant information)
  - Heterogeneous state vector with high dimensionality
- Mitigation techniques
  - Transform from high dimension to low dimension → enable analyst to handle and visualize scenario
    - Principal component analysis
  - De-duplication of alerts prior to correlation
    - Fingerprinting algorithms (e.g., Rabin fingerprinting)

Principal Component Analysis

- PCA
  - Widely used in many domains to address unsupervised dimension/noise reduction
  - Uses singular value decomposition (SVD) to discover “true” dimensionality, obtain nice geometrical structure
  - Good for visualization
  - Selecting the right features is important
  - E.g., Alert count, distinct event type count, etc.
- Ref:
  - Mathew S. and S. Upadhyaya, “Attack Scenario Recognition through Heterogeneous Event Stream Analysis”, *IEEE MILCOM 2009, Boston, MA, October 2009*

Data De-duplication of Alerts

- Pre-process alerts to bring dataset to manageable size where correlation algorithms can process and obtain real-time responses
- Use Rabin fingerprinting for pattern matching and data dedup
- Rabin fingerprinting has nice properties
  - Linear time complexity for string matching
  - Collision probability is extremely low
- De-duplication can reduce alert set by more than an order of magnitude
- Ref:
  - Nagarajaiah H., S. Upadhyaya and V. Gopal, “Data De-duplication and Event Processing for Security Applications on an Embedded Processor”, *5th International Workshop on Dependable Network Computing and Mobile Systems (DNCMS 2012), in conjunction with IEEE SRDS 2012, Irvine, CA, October 2012*

Q&A

- shambhu@buffalo.edu
- http://www.cse.buffalo.edu/~shambhu