HPNAIDM: the High-Performance Network Anomaly/Intrusion Detection and Mitigation System

Yan Chen

Lab for Internet & Security Technology (LIST)
Department of EECS
Northwestern University
http://list.cs.northwestern.edu

Battling Hackers is a Growth Industry!

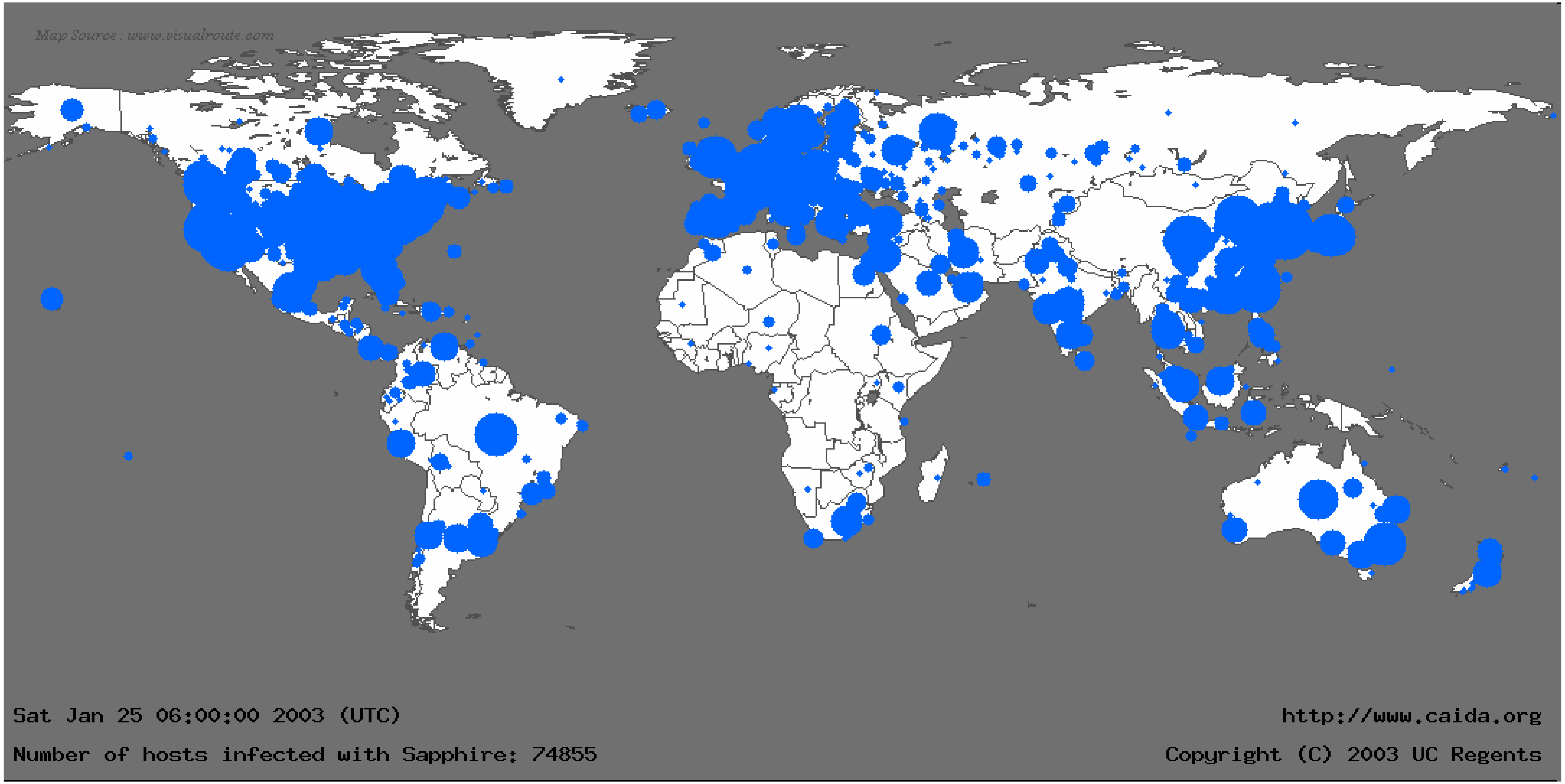
--Wall Street Journal (11/10/2004)

- The past decade has seen an explosion in the concern for the security of information
- Denial of service (DoS) attacks
  - Cost $1.2 billion in 2000
  - Thousands of attacks per week in 2001
  - Yahoo, Amazon, eBay, Microsoft, White House, etc., attacked
- Virus and worms faster and powerful
  - Melissa, Nimda, Code Red, Code Red II, Slammer …
  - Cause over $28 billion in economic losses in 2003, growing to over $75 billion in economic losses by 2007

Current Intrusion Detection Systems (IDS)

- Mostly host-based and not scalable to high-speed networks
  - Slammer worm infected 75,000 machines in <15 mins
  - Flash worm can take less than 1 second to compromise 1M vulnerable machines in the Internet [Staniford04]
  - Host-based schemes inefficient and user dependent
    - » Have to install IDS on all user machines !
  - Existing network IDS unscalable: In a 10Gbps link, each 40-byte packet only has 10ns for processing !
  - Many DOE national labs have over 10Gbps high-performance networks

The Spread of Sapphire/Slammer Worms


Map Source : www.visualroute.com
Sat Jan 25 06:00:00 2003 (UTC)
Number of hosts infected with Sapphire: 74855
http://www.caida.org
Copyright (C) 2003 UC Regents

Current Intrusion Detection Systems (II)

- Mostly signature-based
  - Cannot recognize unknown anomalies/intrusions
  - New viruses/worms, polymorphism
- Statistical detection
  - Hard to adapt to traffic pattern changes
  - Unscalable for flow-level detection
    - » IDS vulnerable to DoS attacks
  - Overall traffic based: inaccurate, high false positives
- Cannot differentiate malicious events with unintentional anomalies
  - Anomalies can be caused by network element faults
  - E.g., router misconfiguration

High-Performance Network Anomaly/Intrusion Detection and Mitigation System (HPNAIDM)

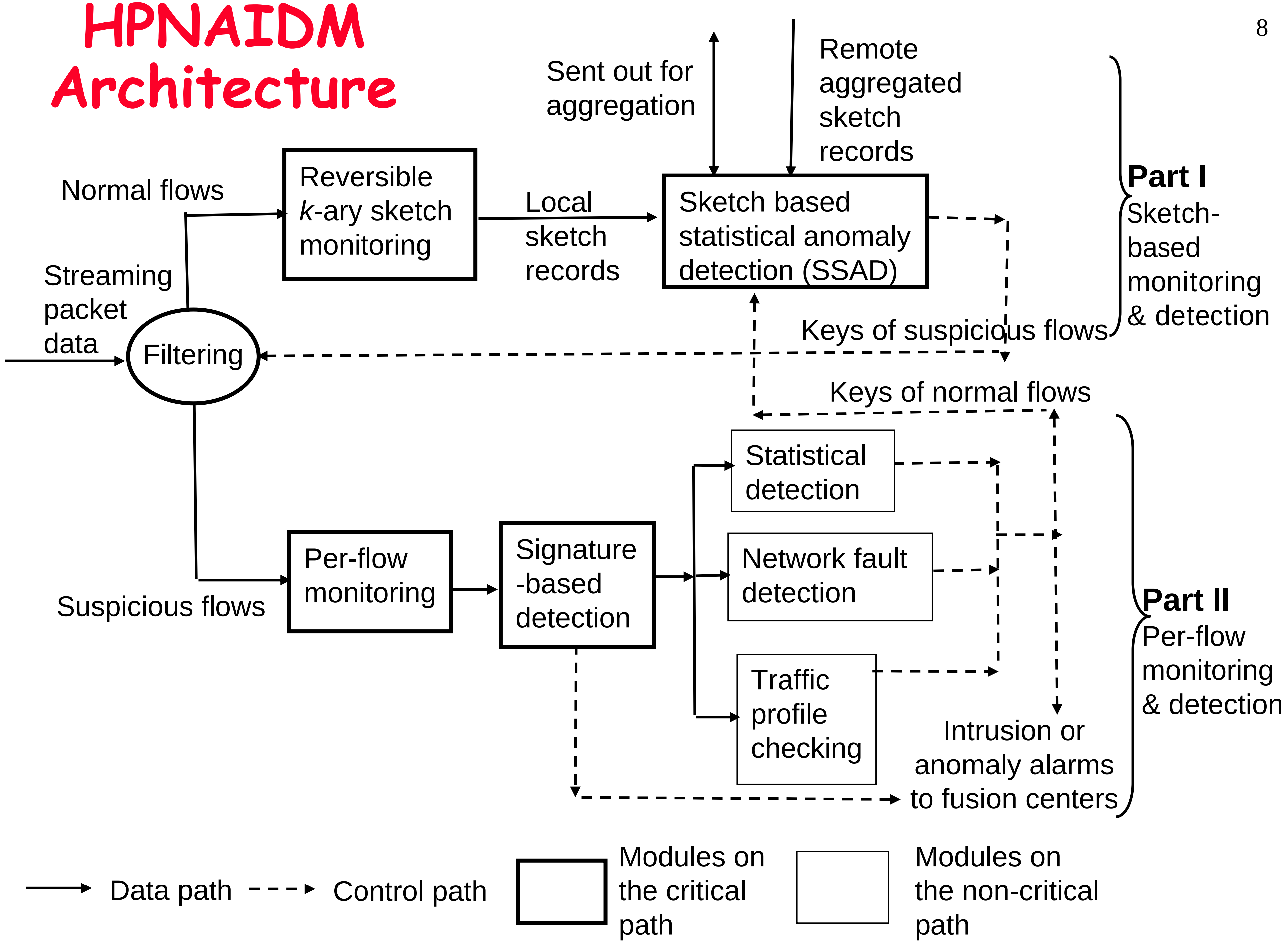
- Online traffic recording
  - Design reversible sketch for data streaming computation
  - Record millions of flows (GB traffic) in a few hundred KB
  - Small # of memory access per packet
  - Scalable to large key space size (2^{32} or 2^{64})
- Online sketch-based flow-level anomaly/intrusion detection
  - Leverage statistical learning theory (SLT) adaptively learn the traffic pattern changes
  - As a first step, detect TCP SYN flooding, horizontal and vertical scans even when mixed

HPNAIDM (II)

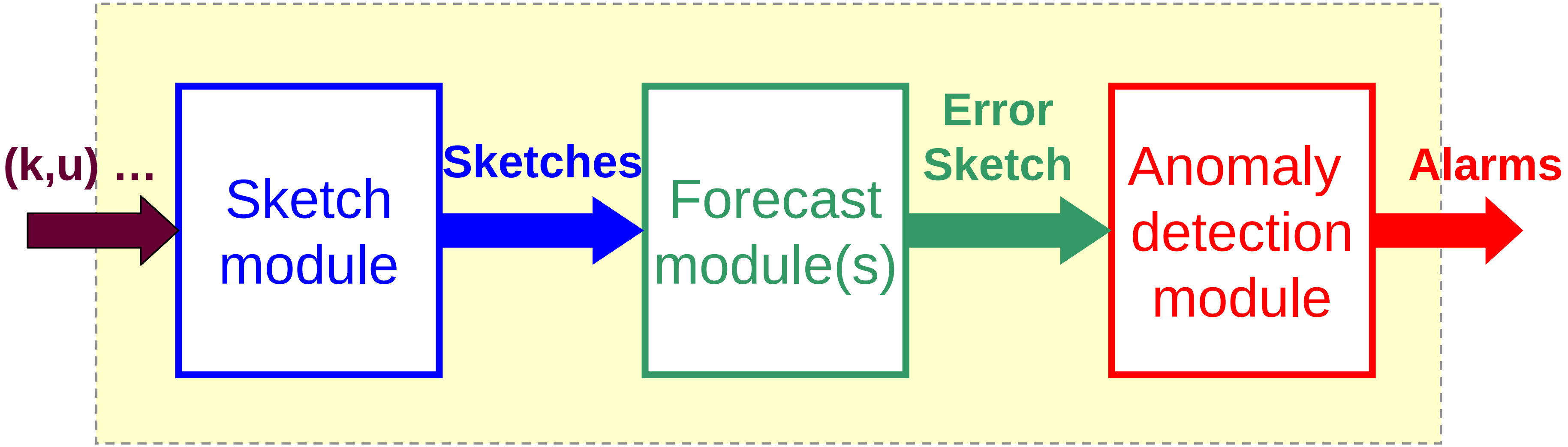
- Integrated approach for false positive reduction
  - Signature-based detection
  - Network element fault diagnostics
  - Traffic signature matching of emerging applications
- Infer key characteristics of malicious flows for mitigation

HPNAIDM: First flow-level intrusion detection that can sustain 10s Gbps bandwidth even for worst case traffic of 40-byte packet streams

HPNAIDM Architecture

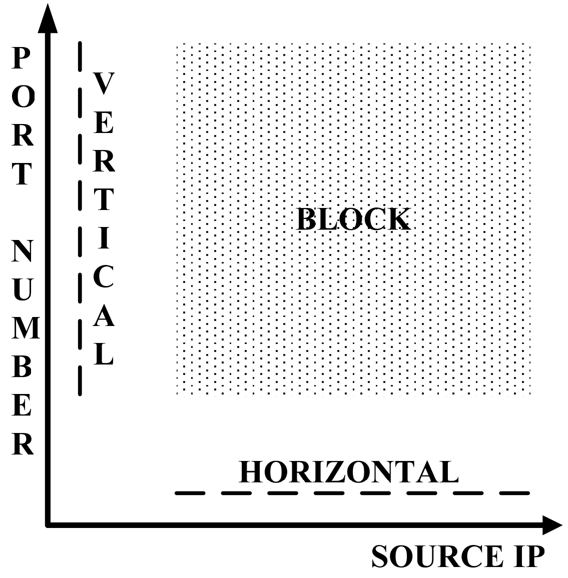

Sent out for aggregation
Remote aggregated sketch records
Normal flows
Reversible *k*-ary sketch monitoring
Local sketch records
Sketch based statistical anomaly detection (SSAD)
Part I
Sketch-based monitoring & detection
Streaming packet data
Filtering
Keys of suspicious flows
Keys of normal flows
Statistical detection
Network fault detection
Per-flow monitoring
Signature -based detection
Suspicious flows
Traffic profile checking
Intrusion or anomaly alarms to fusion centers
Part II
Per-flow monitoring & detection
Data path
Control path
Modules on the critical path
Modules on the non-critical path

Reversible Sketch Based Anomaly Detection



- Input stream: (key, update) (e.g., SIP, SYN-SYN/ACK)
- Summarize input stream using sketches
- Build forecast models on top of sketches
- Report flows with large forecast errors
- Infer the (characteristics) key for mitigation

Sketch-based Intrusion Detection



- RS((DIP, Dport), SYN-SYN/ACK)
- RS((SIP, DIP), SYN-SYN/ACK)
- RS((SIP, Dport), SYN-SYN/ACK)

Attack types	RS((DIP, Dport), SYN-SYN/ACK)	RS((SIP, DIP), SYN-SYN/ACK)	RS((SIP, Dport), SYN-SYN/ACK)
SYN flooding	Yes	Yes	Yes
Vertical scans	No	Yes	No
Horizontal scans	No	No	Yes

Intrusion Mitigation

Attacks detected	Mitigation
Denial of Service (DoS), e.g., TCP SYN flooding	SYN defender, SYN proxy, or SYN cookie for victim
Port Scan and worms	Ingress filtering with attacker IP
Vertical port scan	Quarantine the victim machine
Horizontal port scan	Monitor traffic with the same port # for compromised machine

Preliminary Evaluation

- Evaluated with NU traces (239M flows, 1.8TB traffic/day)
- Scalable
  - Can handle hundreds of millions of time series
- Accurate Anomaly Detection w/ Reversible Sketch
  - Compared with detection using complete flow-level logs
  - Provable probabilistic accuracy guarantees
  - Even more accurate on real Internet traces
- Efficient
  - For the worst case traffic, all 40 byte packets
    - » 16 Gbps on a single FPGA board
    - » 526 Mbps on a Pentium-IV 2.4GHz PC
  - Only less than 3MB memory used

Preliminary Evaluation (cont'd)

- 25 SYN flooding, 936 horizontal scans and 19 vertical scans detected (after sketch-based false positive reduction)
- 17 out of 25 SYN flooding verified w/ backscatter
  - Complete flow-level connection info used for backscatter
- Scans verified (all for vscan, top and bottom 10 for hscan)
  - Unknown scans also found in DShield and other alert reports

Top 10 horizontal scans

Description	Dport	count
Remote desktop scan	3389	1
SQLSnake	1433	3
W32.Rahack	4899	2
unknown scan	3632	1
Scan SSH	22	1
unknown scan	10202	1
Proxy scan	8118	1

Bottom 10 horizontal scans

Description	Dport	count
W32.Sasser.B.Worm	5554	1
Backdoor.CrashCool	9898	2
Unknown scan	42	1
VNC scan	5900	3
Unknown scan	6101	2
Scan SSH	22	1

Activities

- Publications
  - Z. Li, Y. Gao, and **Y. Chen**, "Towards a High-speed Router-based Anomaly/Intrusion Detection System", poster in *ACM SIGCOMM*, 2005. Also, Work in Progress talk with the same title at USENIX Security Symposium, Aug. 2005.
  - P. Ren, Y, Gao, Z. Li, **Y. Chen** and B. Watson, "IDGraphs: Intrusion Detection and Analysis Using Histographs", Proc. of *the IEEE Workshop on Visualization for Computer Security (VizSEC), 2005*
  - R. Schweller, A. Gupta, E. Parsons, and **Y. Chen**, "Reversible Sketches for Efficient and Accurate Change Detection over Network Data Streams", in *ACM SIGCOMM Internet Measurement Conference (IMC)*, 2004
  - **Y. Chen**, D. Bindel, H. Song, and R. H. Katz, "An Algebraic Approach to Practical and Scalable Overlay Network Monitoring", in Proceedings of *ACM SIGCOMM*, 2004
  - B. Krishnamurthy, S. Sen, Y. Zhang, and **Y. Chen**, "Sketch-based Change Detection: Methods, Evaluation, and Applications", Proceedings of *ACM SIGCOMM Internet Measurement Conference (IMC)*, 2003
  - **Y. Chen**, D. Bindel, and R. H. Katz, "Tomography-based Overlay Network Monitoring", Proceedings of *ACM SIGCOMM Internet Measurement Conference (IMC)*, 2003
- Invited talk
  - Y. Chen, "Adaptive Intrusion Detection and Mitigation Systems for WiMAX Networks", Motorola Research Lab, 2005

Potential Collaboration with DOE National Labs

- Dr. Wu-chun Feng
  - Integrate w/ RADIANT

Los Alamos
NATIONAL LABORATORY
EST.1943

- Dr. Don Petravick and Dr. Matt Crawford
  - Collaborated on a NSF proposal

Fermilab

- Dr. Nageswara Rao (UltraScience testbed)

OAK RIDGE NATIONAL LABORATORY
Managed by UT Battelle for the Department of Energy

  - Traffic data collection, intrusion detection and analysis
  - On-site testing