Towards Anomaly/Intrusion Detection and Mitigation on High-Speed Networks

Yan Gao, Zhichun Li, Yan Chen

Northwestern Lab for Internet and Security Technology (LIST)

Department of Computer Science

Northwestern University

http://list.cs.northwestern.edu

Current Intrusion Detection Systems (IDS)

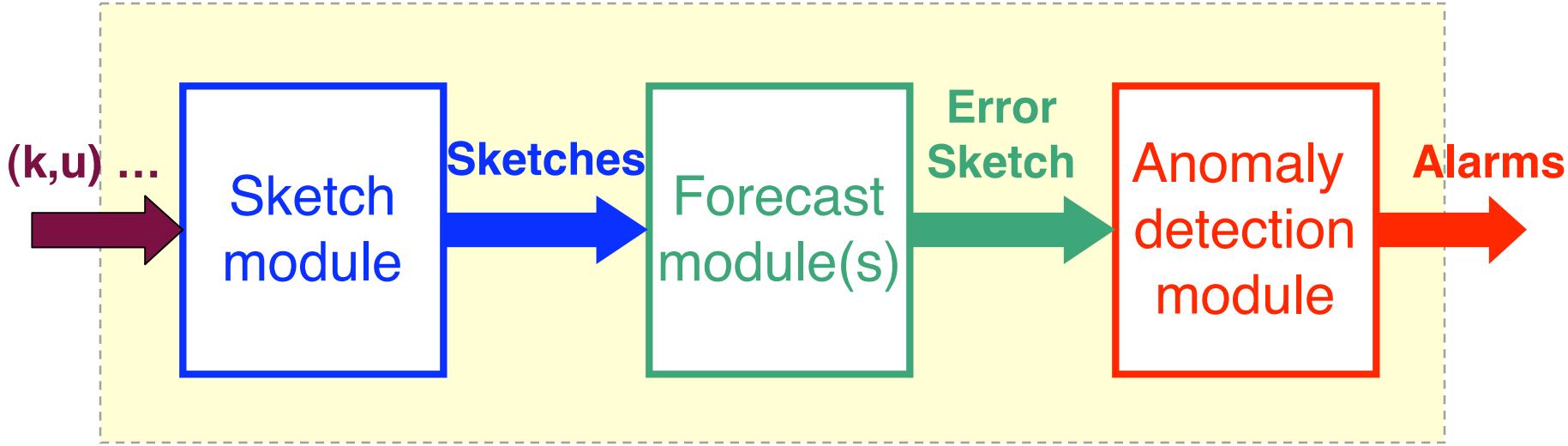
- Mostly host-based and not scalable to high-speed networks
  - Slammer worm infected 75,000 machines in <10 mins
  - Flash worm can take less than 1 second to compromise 1M vulnerable machines in the Internet [Staniford04]
  - Host-based schemes inefficient and user dependent
    - » Have to install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

Router-based Anomaly/Intrusion Detection and Mitigation System (RAIDM)

- Online traffic recording
  - Design reversible sketch for data streaming computation
  - Record millions of flows (GB traffic) in a few hundred KB
- Online flow-level anomaly/intrusion detection & mitigation
  - As a first step, detect TCP SYN flooding, horizontal and vertical scans even when mixed
    - » Existing schemes like TRW/AC, CPM will have high false positives
  - Infer key characteristics of malicious flows for mitigation
- Attach to routers as a black box

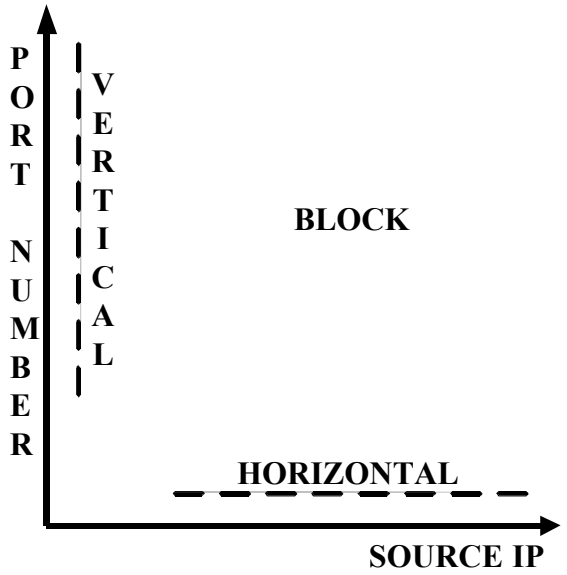
RAIDM: First flow-level intrusion detection that can sustain 10s Gbps bandwidth even for worst case traffic of 40-byte packet streams

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

Intrusion Detection and Mitigation



- 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

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 pc
    - » 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

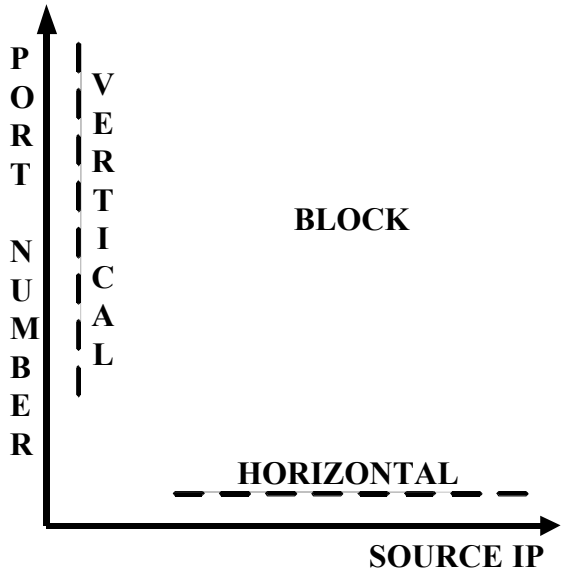
Description	Dport	count
Remote desktop scan	3389	1
SQLSnake	1433	3
W32.Rahack	4899	2
unknown scan	3632	1
Scan SSH	22	1
unknown scan	10202	1
Proxy scan	8118	1

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

Backup Slides

Intrusion Detection and 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
Spywares	Warn the end users being spied