Adaptive Anomaly/Intrusion Detection and Mitigation Systems for High-speed Wireless Networks

Yan Chen, Department of Computer Science, Northwestern University
ychen@cs.northwestern.edu http://list.cs.northwestern.edu

I. Motivation

Traffic anomalies and attacks are commonplace in today's networks, and identifying them rapidly and accurately is critical for large network operators. It was estimated that malicious code (viruses, worms and Trojan horses) caused over $28 billion in economic losses in 2003, and will grow to over $75 billion in economic losses by 2007 [11]. Meanwhile, the broadband wireless networks, represented by the emerged IEEE 802.16 standards, is gaining its popularity and will be seamlessly integrated into the current Internet. Such high-speed wireless network may even be connected to the regional backbone directly to construct the wireless metropolitan-area network (MAN). Thus such wireless network will be exposed to all the attacks, viruses and worms in the current Internet.

Like the 802.11 wireless LAN, 802.16 network is facing scrutiny on its security mechanisms. Driven by customers' security awareness and demand, intrusion detection and security management tools become essential part of current 802.11 product offering suite. Although there are a lot of existing 802.11 intrusion detection products, they mostly target to detect denial-of-service attacks caused by the WEP authentication vulnerability, *e.g.*, Airespace [6]. Existing IDSs for wired network have various shortcomings when they are applied to broadband wireless MAN as discussed below [5, 10, 13–15, 17, 20]. Furthermore, there is little intrusion detection research done tailored to 802.16 and beyond 3G.

First, they are mostly host-based and not scalable to high-speed networks. However, nowadays rapid propagation of viruses/worms can infect most vulnerable machines in the Internet in only ten minutes [12], and even 30 seconds with advanced techniques [18]. Thus it is crucial to identify such outbreaks in their early phases, which can only be possibly achieve d by detection at the base stations or routers provided by the network infrastructures instead of at end hosts [23]. In fact, it is very hard to implement certain detection techniques, such as those for port scanning, at end hosts. The end hosts, especially for wireless devices, may have various weak computational/power limits to do advanced detection. Thus to augment the wireless MAN infrastructure with accurate intrusion detection and mitigation system can significantly attract the subscribers. However, the existing schemes are not scalable to the link speeds and number of flows for high-speed 802.16 wireless MAN as illustrated below.

Second, they are mostly signature-based and unable to adaptively recognize flow-level unknown attacks. Most of them can only detect known attacks with signatures, but not unknown new attacks. Also, attackers can easily evade such IDSs by garbling signatures. Statistical IDSs are therefore proposed to detect anomalous behaviors [2, 7, 8, 21, 22]. To enable accurate detection and attack mitigation, the detection needs to be executed at flow level. Given a spoofed TCP SYN flooding attack sending 40-byte packet streams to a 802.16 network which can provide up to 134Mbps bandwidth, each packet may be considered as a flow, and even to record 10-minute traffic in memory for analysis will take more than 4GB. Thus existing flow-level schemes themselves are vulnerable to attacks [7, 13, 14].

In addition, the statistical parameters are often manually set, and hard to adapt to the traffic pattern changes. However, wireless networks often have transient connections which makes it a big challenge to differentiate collisions, interference, and real attacks.

Third, they cannot differentiate malicious events from the unintentional anomalies. Such unintentional anomalies may be caused by network element (*e.g.*base stations, routers or links) faults. The statistical IDS cannot tell it from malicious attacks, and thus have high false positive rates.

II. Proposed Solution

A. System Design and Features

To address these challenges, we propose to break the limitations of existing IDSs, and develop a new paradigm called *adaptive Anomaly/Intrustion Detection and Mitigation system for high-speed Wireless networks* (WAIDM). The new paradigm is significantly different from existing IDSs as illustrated with the following features (research thrusts). We believe this intrusion detection technology for 802.16 will bring Motorola to a uniquely strong marketing position ahead of competition in 802.16 portfolio products offered to customers. And we have received strong support from our liaisons in Motorola: Gregory W. Cox, Judy Z. Fu, and Philip Roberts in the Networks and Systems Group of Motorola Research Center.

First, online traffic recording and analysis on high-speed wireless MAN We propose to leverage the PI's previous work on *k-ary sketch* [9], an efficient tool for data streaming computation, to record flow-level traffic as the basis for statistical anomaly detection. We will also investigate reversible sketches to infer the characteristics (*e.g.*, source IP) of culprit flows when detected, and then mitigate them. With sketches, we can record millions of simultaneous flows with even Gigabtye traffic with only a few hundred

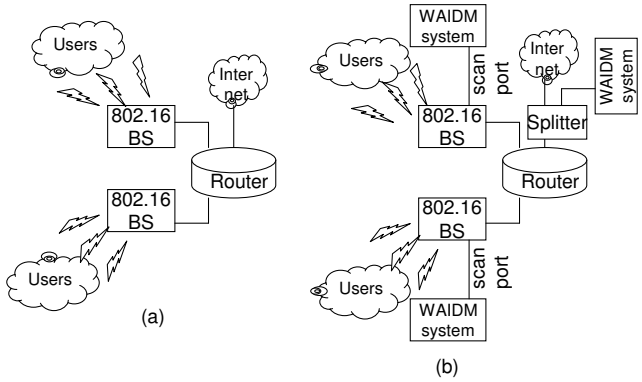

Fig. 1. Attaching the WAIDM systems to high-speed 802.16 base stations. (a) original configuration, (b) two methods to deploy WAIDM for monitoring: 1) connect WAIDM to a scan port and copy the traffic of a base station; 2) connect WAIDM to a splitter which aggregates the traffic from all the ports of a router where multiple base stations are connected to.

kilobytes.

Second, online adaptive flow-level anomaly/intrusion detection and mitigation We will design sketch-based algorithms to detect and mitigate various flow-level intrusions *online*, leveraging the emerging statistical learning theory (SLT) which is able to adaptively learn the traffic pattern changes with even *unsupervised* learning.

In particular, we will use statistics from Management Information Base (MIB) of the base station (or similar management information obtained through management interface or mobile switch center) to understand the current wireless network status. We will consider metrics such as capacity, transmission fail count, multiple retry count, duplicate count, received fragment count, *etc.*. Then we can infer the wireless network status whether it is congested or Interfered. Based on such inference, we will automatically adapt to different learned profiles on observing status changes. Given different attacks detected, various mitigation schemes will be investigated and applied to protect the infrastructure and customers. We believe that such scheme is applicable to both 802.16 wireless MAN and celluar network infrastructure protection.

Third, integrated approach for false positive reduction We plan to bridge the traditional gap between the network measurement/trouble shooting research and the intrusion detection research by *integrating the signature-based detection, and network element fault diagnostics* to analyze the traffic anomalies discovered by the statistical methods, and further reduce the false positives.

We will particularly explore the 802.16 specific attacks. The current 802.16 standard has a "security sublayer" which mostly focus on packet data encryption, authentication and authorization, using IPSec [19]-like schemes. We will investigate the vulnerability of these schemes, and design a signature library for attacks that are specific to 802.16, in addition to the existing signatures of Internet attacks.

As shown in Figure 1, WAIDM detection systems will be implemented as black boxes to attach to the broadband 802.16 base stations without affecting the normal operations of the base stations. It can enable the early detection of global scale attacks, which is crucial because such attacks can subvert the wireless MAN infrastructure and many hosts instantly with exponential growth.

Detection on edge networks (in the case of 802.16 wireless MAN, the base station) is particularly critical, powerful and efficient (without deploying IDSs on all the edge hosts), according to a recent research agenda for large scale malicious code by the famous Defense Advanced Research Projects Agency (DARPA) [23]. For instance, an infected machine that generates a large number of scans is detectable on its associated network links, as is a large amount of aggregate scanning. Since most normal traffic receives positive responses, while scans are met primarily with negative responses or non-responses, this anomalous pattern of behavior can be easily detected when there exists symmetric routing. Furthermore, a wireless MAN operator knows its unused address space, and can raise a red flag if there is a high amount of traffic into that space. Other propagation strategies may also show clearly anomalous behaviors.

Our base station based anomaly/intrusion detection scheme can be naturally extended for responses. That is, when alerted, filters can be automatically constructed to filter classes of traffic, and isolate infected machines. Such a scheme is highly ranked as "powerful and flexible" by the DARPA research agenda [23]. For instance, for distributed SYN flooding, we can start the *SYN defender* [1], *SYN proxy* and/or *SYN cookies* [3] for the particular victim machines and applications to alleviate the DoS effects. For port scans and point-to-point SYN flooding, we use ingress filter to block the traffic from the attacker IP. For vertical scans, we can quarantine the victim machine for further inspection.

B. System Architecture

As illustrated in Figure 2, the WAIDM system can be divided into two parts: k-ary sketch-based monitoring and anomaly detection (part I), and per-flow monitoring and anomaly/intrusion detection (part II). Part I filters out the major portion of legitimate traffic so that we can afford to examine suspicious flows in full detail. That is, part I provides scalability while part II offers accuracy.

III. Project Deliverables and Evaluation

We will build a prototype of the WAIDM system with the following expected properties:

- It can record traffic and detect intrusions at flow-level online for high-speed wireless MAN, with small memory consumption, small amount of memory access, and scalable to large IP address space, like IPv6.
- It can effectively distinguish collisions, interference and intrusions attacks. We will benchmark the false alarm rates and compare with existing wired and wireless IDS products.
- It can automatically response to attacks through intrusion mitigation. We will evaluate the response effectiveness in terms of the percentage of intrusion

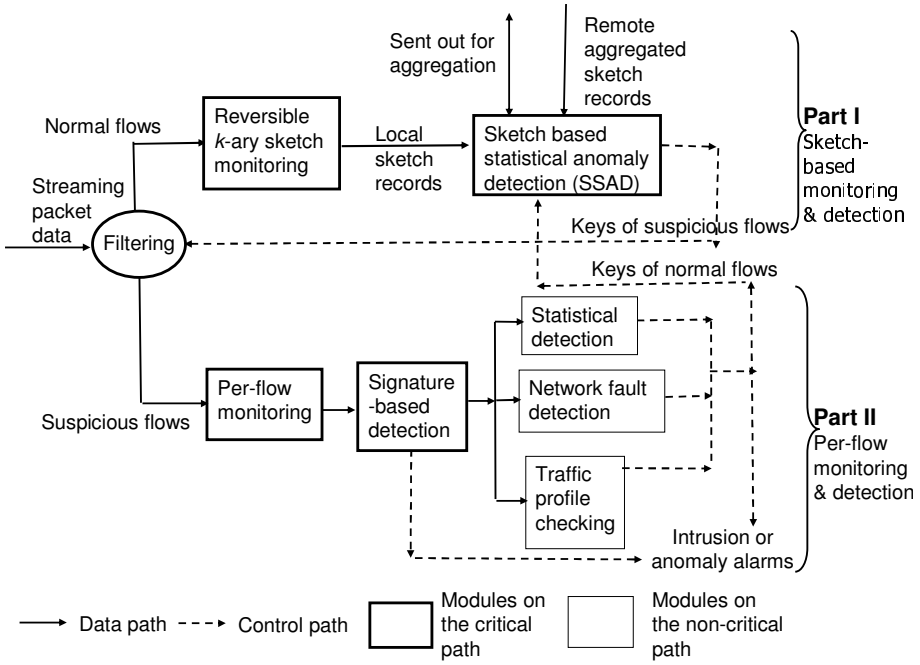

Fig. 2. Architecture of a WAIDM system.

traffic automatically blocked.

We will build prototypes for the WAIDM system. We have set up collaborations with Motorola and will use their wireless access network data obtained from its partners for evaluation. Furthermore, We will apply network and attack traces from other various sources for evaluation: Northwestern University (NU), Fermi National Lab, AT&T, CERNET, and DShield (the largest worldwide intrusion log repository). Note that these organizations like NU and Fermi Lab can be potential customers of the 802.16 wireless MAN. This puts the PI in a very unique position to combine data from both a core network provider (AT&T) and numerous access networks (Motorola, NU, and Fermi) for studying global anomalies/intrusions and how it will affect the wireless MAN if deployed. We also plan to test our system on site at NU and Fermi Lab.

IV. Preliminary Results

We have successfully design and implemented sketches for network monitoring on high-speed traffic, both in software and hardware. When implemented on a single FPGA board, we can sustain more than 16Gbps even for a stream of 40-byte-packets (the worst case traffic). With a software implementation, for inferring even 1,000 heavy change keys out of millions of flows, we find more than 99% of the heavy change keys with less than a 0.1% false positive rate within 13 seconds on a P4 PC. Typically, the anomaly detection examines much less number of heavy changes, like the top 100, which takes only 0.34 seconds. These work are published in the premier ACM SIGCOMM Internet Measurement Conference (IMC) [9, 16].

We have successfully designed and implemented sketch-based flow-level intrusion detection system, and applied it on several router-level traffic traces from Northwestern University, Fermi National Lab and Lawrence Berkeley National Lab. We further compared the results with other state-of-the-art detection methods, like TRW for port scan detection [7] and CPM for SYN flooding detection [21, 22]. In comparison to TRW, we can detect vertical scans which TRW cannot, and the memory consumption is much less than that of TRW which will easily run out of memory when there is an intensive spoofed SYN flooding attacks. For CPM, it can only detect attacks for overall aggregated traffic instead of at per-flow level. We found our method are much more accurate than CPM, and unlike CPM, we can mitigate attacks once detected.

V. Qualification of PI

PI Yan Chen is an Assistant Professor of Computer Science at Northwestern University. He has eight years of experience on distributed systems and networking research from both academia and industry, especially in the area of network security, network measurement and P2P systems. For instance, he designed the first simulation-based network DoS attack resilience benchmark, and applied it for evaluating and comparing the centralized, replicated, and emerging distributed object location services [4]. He has widely collaborated with both industry and academia researchers on numerous projects with publications in ACM SIGCOMM, SIGCOMM IMC, IEEE ICNP, IEEE JSAC, *etc.*. In 2004, he won the Microsoft Trustworthy Computing Grant and the Murphy Society Grant of Northwestern University.

REFERENCES

[1] TCP Flooding Attack and Firewall-1 SYNDefender. Checkpoint Software Technologies.

[2] BARFORD, P., KLINE, J., PLONKA, D., AND RON, A. A signal analysis of network traffic anomalies. In *ACM SIGCOMM Internet Measurement Workshop* (November 2002).

[3] BERNSTEIN, D. Syn cookies. http://cr.yp.to/syncookies.html.

[4] CHEN, Y., BARGTEIL, A., BINDEL, D., KATZ, R., AND KUBIATOWICZ, J. Quantifying network denial of service: A location service case study. In *Proc. of the Third International Conference on Information and Communications Security (ICICS)* (2001).

[5] HOFMEYR, S., AND FORREST, S. Intrusion detection using sequences of system calls. *Journal of Computer Security 6* (1998).

[6] INC., A. Wireless intrusion detection and prevention. `http://www.airespace.com/technology/whitepapers.php`.

[7] JUNG, J., PAXSON, V., BERGER, A. W., AND BALAKRISHNAN, H. Fast portscan detection using sequential hypothesis testing. In *Proceedings of the IEEE Symposium on Security and Privacy* (2004).

[8] KOMPELLA, R. R., SINGH, S., AND VARGHESE, G. On scalable attack detection in the network. In *Proc. of ACM/USENIX IMC* (2004).

[9] KRISHNAMURTHY, B., SEN, S., ZHANG, Y., AND CHEN, Y. Sketch-based change detection: Methods, evaluation, and applications. In *Proc. of ACM SIGCOMM Internet Measurement Conference (IMC)* (2003).

[10] LOYALL, J. P., PAL, P. P., SCHANTZ, R. E., AND WEBBER, F. Building adaptive and agile applications using intrusion detection and response. In *Proc. of the Network and Distributed Systems Security Conference* (2000).

[11] MARS, E., AND JANSKY, J. D. Email defense industry statistics. `http://www.mxlogic.com/PDFs/IndustryStats.2.28.04.pdf`.

[12] MOORE, D., PAXSON, V., SAVAGE, S., SHANNON, C., STANIFORD, S., AND WEAVER, N. The spread of the Sapphire/Slammer worm. http://www.caida.org, 2003.

[13] PAXSON, V. Bro: A system for detecting network intruders in real-time. *Computer Networks 31*, 23-24 (1999), 2435–2463.

[14] ROESCH, M. Snort: The lightweight network intrusion detection system, 2001. `http://www.snort.org/`.

[15] RYUTOV, T., NEUMAN, C., KIM, D., AND ZHOU, L. Integrated access control and intrusion detection for web servers. *IEEE Trans. on Parallel and Distributed Systems 14*, 9 (2003), 841–850.

[16] SCHWELLER, R., GUPTA, A., PARSONS, E., AND CHEN, Y. Reversible sketches for efficient and accurate change detection over network data streams. In *ACM SIGCOMM Internet Measurement Conference (IMC)* (2004).

[17] SEKAR, R., BENDRE, M., DHURJATI, D., AND BOLLINENI, P. A fast automaton-based method for detecting anomalous program behaviors. In *Proc. of the IEEE Symposium on Security and Privacy* (2001).

[18] STANIFORD, S., PAXSON, V., AND WEAVER, N. How to own the Internet in your spare time. In *Proceedings of the 11th USENIX Security Symposium* (2002).

[19] THAYER, R., DORASWAMY, N., AND GLENN, R. RFC 2411 - IP Security Document Roadmap. `http://www.faqs.org/rfcs/rfc2411.html`.

[20] WAGNER, D., AND DEAN, D. Intrusion detection via static analysis. In *Proceedings of the IEEE Symposium on Security and Privacy* (2001).

[21] WANG, H., ZHANG, D., AND SHIN, K. G. Detecting SYN flooding attacks. In *Proc. of IEEE INFOCOM* (2002).

[22] WANG, H., ZHANG, D., AND SHIN, K. G. Change-point monitoring for detection of DoS attacks. *IEEE Transactions on Dependable and Secure Computing 1*, 4 (2004).

[23] WEAVER, N., PAXSON, V., STANIFORD, S., AND CUNNINGHAM, R. Large scale malicious code: A research agenda. Tech. Rep. DARPA-sponsored report, 2003.