Final Report for DOE Early CAREER Award on "HPNAIDM: The High-Performance Network Anomaly/Intrusion Detection and Mitigation System"

Award #: DE-FG02-05ER25692
Name of Recipient: Office for Sponsored Research,
Northwestern University,
633 Clark Street, Evanston, IL 60208-1108

PI: Yan Chen
2005-2008

Executive Summary

Identifying traffic anomalies and attacks rapidly and accurately is critical for large network operators. With the rapid growth of network bandwidth, such as the next generation DOE UltraScience Network, and fast emergence of new attacks/virus/worms, existing network intrusion detection systems (IDS) are insufficient because they:
- Are mostly host-based and not scalable to high-performance networks;
- Are mostly signature-based and unable to adaptively recognize flow-level unknown attacks;
- Cannot differentiate malicious events from the unintentional anomalies.

To address these challenges, we proposed and developed a new paradigm called high-performance network anomaly/intrustion detection and mitigation (HPNAIDM) system. The new paradigm is significantly different from existing IDSes with the following features (research thrusts).
- Online traffic recording and analysis on high-speed networks;
- Online adaptive flow-level anomaly/intrusion detection and mitigation;
- Integrated approach for false positive reduction.

Our research prototype and evaluation demonstrate that the HPNAIDM system is highly effective and economically feasible. Beyond satisfying the pre-set goals, we even exceed that significantly (see more details in the next section). Overall, our project harvested 23 publications (2 book chapters, 6 journal papers and 15 peer-reviewed conference/workshop papers). Besides, we built a website for technique dissemination, which hosts two system prototype release to the research community. We also filed a patent application and developed strong international and domestic collaborations which span both academia and industry.

Comparison of Actual Accomplishments with the Objectives and Goals of the Project

In this section, we compare our project achievement with the goals set at the proposal. Basically, we used less than 2 years to finish the goals set up in the proposal. After that, we significantly expand the goal of our research with the following additional achievement beyond those specified in the proposal.

1. Correlation of attacks observed from multiple sites for a global intrusion alert
2. Network situational-aware analysis, and intrusion forensics
3. Botnet detection in high-speed network systems.

Please refer to the next section on the rationale behind the expanded goal.

Project Activities

In this project, we successfully verify and evaluate our design (including the original hypothesis and approaches used) with a system prototype tested on real world traffic. During the evaluation, we found that we need to further expand our system to really understand the attacks and anomalies in real world traffic. That is why we added the first two components listed above. Furthermore, with emergence of botnet threat, we added the botnet detection system. Next, we will explain the system architecture and how we achieved these goals in each year.

The HPNAIDM system has four major components: high-speed network monitoring using sketches, statistical anomaly/intrusion detection, network diagnosis and polymorphic worm signature generation. I attach the system architecture graph as below to show the four components in our proposal.

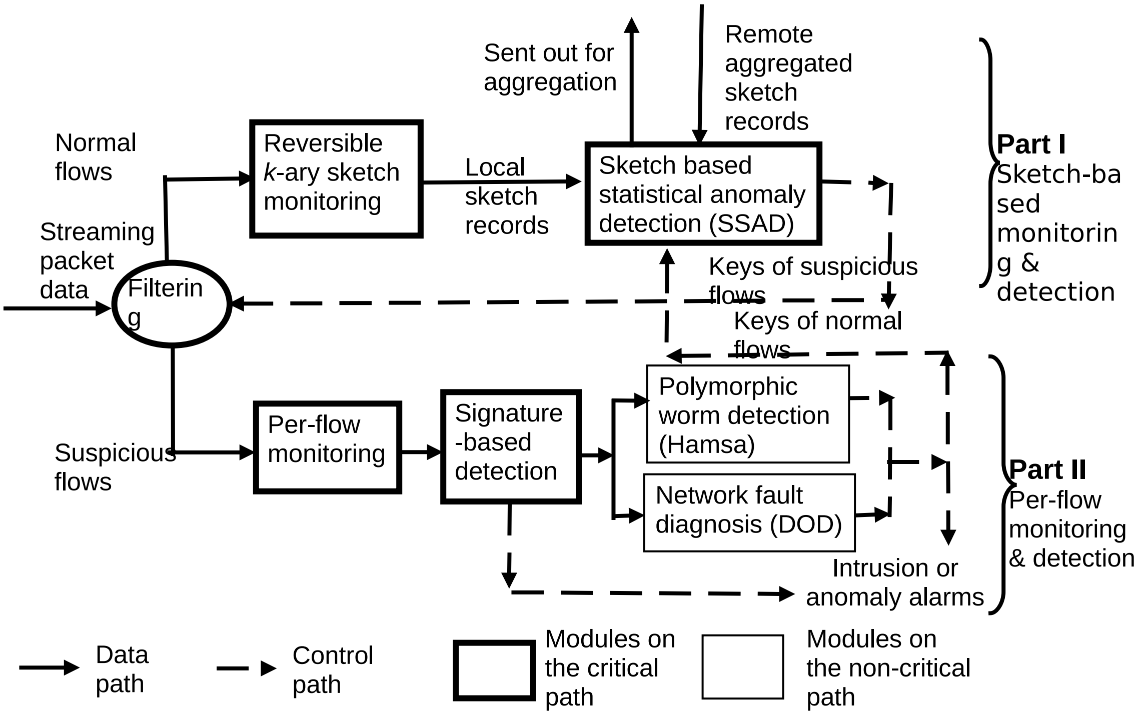
Sent out for aggregation
Remote aggregated sketch records
Normal flows
Reversible *k*-ary sketch monitoring
Local sketch records
Sketch based statistical anomaly detection (SSAD)
Part I
Sketch-based monitoring & detection
Streaming packet data
Filtering
Keys of suspicious flows
Keys of normal flows
Polymorphic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

Figure 1. The system architecture graph of HPNAIDM

In the first year of this project (2005-2006), we have finished the design, and first stage evaluation of the first two components: high-speed network monitoring and intrusion detection. For the statistical anomaly detection, we combine the visualization techniques for effective detection. This will be published as one conference and one journal paper in security visualization community as below.

In addition, we made significant progress for the third and forth components. Basically, we finished the first iteration of design and evaluation, and were working on improving the mechanisms to detect more stealthy polymorphic worms and to build more flexible tool for finer level Internet diagnosis.

In the second year of this project (2006-2007), we have made very good progress on the other two components: polymorphic worm signature generation and network fault diagnosis.

In addition, we are expanding the scope of the research by studying two additional directions: correlation of attacks observed from multiple sites for a global intrusion alert system, and intrusion forensics. In particular, we are conducting research on the state-of-the-art botnet systems. Currently, Botnets dominate the Internet attack landscape and most of the attacks are launched with botnets. To understand the behavior and intention of the botnets are of crucial importance for securing the emerging high-speed network.

In the final year (2007-2008), for proposed area, we mainly finished up the network diagnosis. In addition, we worked on the expanded area: network situational-aware analysis, intrusion forensics, and botnet detection.

Products

Publications:

Book Chapter:

1. Zhichun Li, Anup Goyal, and Yan Chen, Honeynet-based Botnet Scan Traffic Analysis, invited book chapter for Botnet Detection: Countering the Largest Security Threat, Springer, 2007.
2. Ehab Al-Shaer and Yan Chen, Integrated Fault and Security Management, invited book chapter for Information Assurance: Dependability and Security in Networked Systems, Morgan Kaufmann Publishers, 2007.

Journal papers (Some of the journal papers takes a couple years after the end of project to be finally published.):

3. Zhichun Li, Yan Gao, and Yan Chen, HiFIND: A high-speed flow-level intrusion detection approach with DoS resiliency, Journal of Computer Networks, Volume 54, Issue 8, June 2010.

4. Lanjia Wang, Zhichun Li, Yan Chen, Zhi (Judy) Fu, and Xing Li, Thwarting Zero-Day Polymorphic Worms With Network-Level Length-Based Signature Generation, in ACM/IEEE Transaction on Networking (ToN), Volume 18, Issue 1, 2010.
5. Yao Zhao, Yan Chen, and David Bindel, Towards Unbiased End-to-End Network Diagnosis, in ACM/IEEE Transaction on Networking (ToN), Volume 17, Number 6, Dec. 2009.
6. Yao Zhao and Yan Chen, FAD and SPA: End-to-end Link-level Loss Rate Inference without Infrastructure, in the Journal of Computer Networks, 53(9): 1303-1318, 2009.
7. Robert Schweller, Zhichun Li, Yan Chen, Yan Gao, Anup Gupta, Elliot Pearson, Yin Zhang, Peter Dinda, Ming-Yang Kao, and Gokan Memik, Reversible Sketches: Enabling Monitoring and Analysis over High-speed Data Streams, in ACM/IEEE Transaction on Networking, Volume 15, Issue 5, Oct. 2007.
8. P. Ren, Y, Gao, Z. Li, Y. Chen and B. Watson, "IDGraphs: Intrusion Detection and Analysis Using Stream Compositing", *IEEE Computer Graphics & Applications, special issue on Visualization for Cyber Security*, Volume 26, Number 2, March/April, 2006.

Conference papers:

9. Zhichun Li, Anup Goyal, Yan Chen, and Vern Paxson, Automating Analysis of Large-Scale Botnet Probing Events, in the Proc. of ACM Symposium on Information, Computer and Communications Security (ASIACCS), 2009 (33/147=22.4%).
10. Yan Gao, Yao Zhao, Robert Schweller, Shobha Venkataraman, Yan Chen, Dawn Song, and Ming-Yang Kao, "Detecting Stealthy Spreaders Using Online Outdegree Histograms", in the Proc. of *the 15th IEEE International Workshop on Quality of Service (IWQoS), 2007* (17/64=26.6%).
11. Zhichun Li, Lanjia Wang, Yan Chen and Zhi Fu, Network-based and Attack-resilient Length Signature Generation for Zero-day Polymorphic Worms, in the Proc. of the 15th IEEE International Conference on Network Protocols (ICNP), 2007(32/220=14%).
12. Yao Zhao and Yan Chen, A Suite of Schemes for User-level Network Diagnosis without Infrastructure, in the Proc. of *IEEE INFOCOM, 2007* (252/1400=18%).
13. Guohan Lu, Yan Chen, Stefan Birrer, Fabian E. Bustamante, Chi Yin Cheung and Xing Li, End-to-end Inference of Router Packet Forwarding Priority, in the Proc. of *IEEE INFOCOM, 2007* (252/1400=18%).
14. Yan Gao, Leweng Deng, Aleksandar Kuzmanovic, and Yan Chen, Internet Cache Pollution Attacks and Countermeasures, in Proc. of *the 14th IEEE International Conference on Network Protocols (ICNP)*, Nov. 2006
15. Yao Zhao, Yan Chen, and David Bindel, Towards Unbiased End-to-End Network Diagnosis, in Proc. of *ACM SIGCOMM 2006* (37/340=10%).
16. Zhichun Li, Yan Chen, and Aaron Beach, Towards Scalable and Robust Distributed Intrusion Alert Fusion with Good Load Balancing, in Proc. of *ACM SIGCOMM Workshop on Large-Scale Attack Defense 2006*(11/33=33%).

17. R. Schweller, Z. Li, Y. Chen, Y. Gao, A. Gupta, E. Pearson, Y. Zhang, P. Dinda, M. Kao, and G. Memik, Reversible Sketches: Enabling Monitoring and Analysis over High-speed Data Streams, to appear in ACM/IEEE Transaction on Networking.
18. Y. Gao, Z. Li and Y. Chen, "A DoS Resilient Flow-level Intrusion Detection Approach for High-speed Networks", accepted by *IEEE International Conference on Distributed Computing Systems (ICDCS)*, 2006. (75/536=14%)
19. R. Schweller, Z. Li, Y. Chen, Y. Gao, A. Gupta, Y. Zhang, P. Dinda, M. Kao, and G. Memik, "Reverse Hashing for High-speed Network Monitoring: Algorithms, Evaluation, and Applications", in the Proc. of *IEEE INFOCOM*, April 2006. (252/1800=18%)
20. Z. Li, M. Sanghi, Y. Chen, M. Kao, and B. Chavez, "Hamsa: Fast Signature Generation for Zero-day Polymorphic Worms with Provable Attack Resilience", accepted by *IEEE Symposium on Security and Privacy*, 2006. (about 8%)
21. Y. Zhao, Y. Chen, and D. Bindel, "Deterministic Overlay Diagnosis", accepted as a poster by *ACM SIGMETRICS*, 2006. (30 full + 17 poster papers out of 217 (14-22%))
22. 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)*, Oct. 2005.
23. Z. Li, Y. Gao, and Y. Chen, Towards a High-speed Router-based Anomaly/Intrusion Detection System, poster in *ACM SIGCOMM*, Aug., 2005.

Website:

We built a website to disseminate the research results, such as publications and tools to release.
http://list.cs.northwestern.edu/midf.html

Networks or Collaboration Fostered:

In this project, we foster strong collaborations with Prof. Bin Liu of Tsinghua University China. Besides Prof. Liu, there were two students Chengchen Hu and Gao Xia from their lab coming to Northwestern for half a year each as visiting PhD students. There are also several joint publications results from the collaboration.

We also widely collaboration with several NU colleagues, Profs. MingYang Kao, Peter Dinda, Gokan Memik. In addition, we collaborate with Prof. Vern Paxson from UC Berkeley. Besides the academia collaboration, we also worked with industry labs (e.g., Motorola Labs and Microsoft Research).

Software Release:

There are two software component release we hosted on the aforementioned project website:

1. Hamsa polymorphic worm signature generator

2. Reversible Sketch Code

No Computer Modeling involved.

For this project, we mainly design algorithms to detect intrusions with high speed, based on real world attacks. So there is no computer modeling involved.