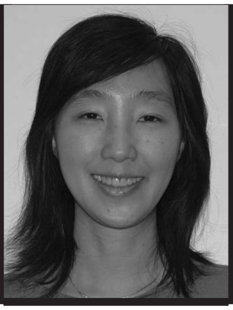
Recent Developments in Network Intrusion Detection

Thomas Chen



Zhi Fu

Liwen He

Tim Strayer

About 20 years ago in November 1988, the Morris worm spread through the Internet, taking down thousands of computers. The incident prompted the U.S. Defense Advanced Research Programs Agency to establish the CERT/CC to coordinate activities to defend against future Internet security problems, and was one of the first media stories to raise public awareness about network security. Security problems with the TCP/IP protocol suite were known (as noted by Steven Bellovin), but the Internet was a closed network for academics and researchers at the time. Spam and malware were minor problems, and the Web had not been invented. Security was understandably not one of the high priority concerns of the Internet designers 20 years ago, but the consequences of an open public Internet are now apparent.

Today network security has become an everyday problem with virtually all computers connected to the Internet. The average Internet user must be constantly vigilant against a number of network threats such as spam, worms, Trojan horses, bots, spyware, and phishing. Enterprises are forced to fortify their networks against remote intrusions into their servers and databases. Governments are concerned about espionage and the possibility of cyberwarfare.

Intrusion detection has been a critical component of network security since the 1980s. It is not realistic to expect that all attacks can be blocked by firewalls, access control lists, and other defenses. Intrusion detection serves the important function of identifying malicious activities and determining their nature, origin, and seriousness. Network-based and host-based intrusion detection methods commonly use a combination of signatures to recognize known attacks and anomaly detection to recognize suspicious behaviors.

This Special Issue is intended to present the state of the art in network-based intrusion detection. Although an enormous literature already exists, intrusion detection is a dynamic problem demanding constant research progress to keep up with new exploits, new evasion techniques, and increasing traffic rates. In response to the open call, we were pleased to receive 44 submissions from which six articles were accepted for this issue. The large number of submissions attests to the vitality of research efforts and high interest level in intrusion detection.

Intrusion detection is a problem well suited to intelligent sampling. Intrusion detection systems must observe a great amount of traffic (gigabits per second) looking for anomalies, but the vast majority of flows are normal and uninteresting. The first article in this special issue, "Network Anomaly Detection and Classification via Opportunistic Sampling" by G. Androulidakis, V. Chatziqiannakis, and S. Papavassiliou, argues that intelligent flow sampling can both reduce the data for processing and improve the effectiveness of anomaly detection. An entropy-based anomaly detection method is considered in combination with intelligent sampling. Two sampling algorithms are considered, one that favors large flows and another that favors small flows. Experiments are carried out to detect anomalies in traffic data collected from a university campus network.

The second article, "Self-Addressable Memory-Based FSM (SAM-FSM): A Scalable Intrusion Detection Engine" by Benfano Soewito, Lucas Vespa, Atul Mahajan, Ning Weng, and Haibo Wang, addresses the problem of high-speed string matching, commonly performed by signature-based network intrusion detection systems. A novel pattern matching engine is proposed to exploit a memory-based programmable finite state machine to achieve deterministic processing rates that are independent of packet and pattern characteristics. The engine is fully reconfigurable for new attack patterns and is storage-efficient. Memory space is saved by using a metapointer for multiple states and collapsing FSM states.

Accurate anomaly detection is a long-standing problem in intrusion detection. The third article, "Accurate Anomaly Detection through Parallelism" by Shashank Shanbhag and Tilman Wolf, proposes a parallel anomaly detection system. Instead of a single detection algorithm, the essential idea is to implement multiple anomaly detection algorithms and monitor multiple traffic subsets in parallel. This design approach has become practical only recently due to high-performance embedded processors (network processors). Each detection algorithm produces an anomaly metric, and all metric outputs are normalized and aggregated into an overall anomaly score reflecting the severity of the anomaly. This method increases detection accuracy by combining multiple anomaly detection algorithms (compared to any single algorithm) and increases sensitivity to anomalies specific to a particular traffic class.

Naturally, attackers are cognizant of intrusion detection systems and attempt numerous ways to avoid detection. One method is IP fragmentation, since it is problematic for intrusion detection systems to collect and reassemble all fragments. The fourth article, "Counting Bloom Filters for Pattern Matching and Anti-Evasion at the Wire Speed" by Gianni Antichi, Domenico Ficara, Stefano Giordano, Gregorio Procissi, and Fabio Vitucci, proposes to avoid fragment reassembly by means of counting bloom filters (CBFs), which are similar to bloom filters except for the use of fixed size counters (bins). This overcomes the problem that bloom filters cannot easily be emptied. CBFs have appealing features in terms of compactness, speed, update capability, and emptying feasibility. They are well suited to pattern matching because they are compact and quickly updatable for new detection signatures. Their ability to count occurrences of elements also makes them useful for anti-evasion if a CBF is set to represent the different substrings composing a string of interest to be matched. As occurrences of substrings are detected, the proper bins are decreased, and the entire string is detected when the filter is completely reset to zero.

Another long standing problem in intrusion detection is tracing the origin of attacks. Packet traceback has been difficult because IP source addresses can be spoofed, and IP routers are not designed to retain any memory of forwarded packets. This situation allows remote attacks to be carried out with anonymity and little risk. The fifth article, "An AS-Level Overlay Network for IP Traceback" by Andre Castelucio, Ronaldo M. Salles, and Artur Ziviani, addresses the problem of tracing the origin of distributed denial of service (DDoS) attacks and proposes a packet traceback system that can be partially deployed in the Internet. It operates on border routers of some autonomous systems that build an AS-level overlay network. BGP is used as a vehicle for information exchange between autonomous systems participating in packet traceback. The system requires a new extension to the BGP update message community attribute to allow information to be passed across autonomous systems, allowing autonomous systems willing to collaborate in traceback to join with participating autonomous systems. In the syste packets are marked by routers in a generalized bloom filter in the packet header. A new sequence marking process removes ambiguities in the traceback path. Results suggest that a relatively low number of autonomous systems can be sufficient for efficient packet traceback if they are strategically chosen.

The last article, "A Simple and Efficient Hidden Markov Model Scheme for Host-Based Anomaly Intrusion Detection" by Jiankun Hu, D. Qiu, Hsiao-Hwa Chen, and Xinghuo Yu, notes that widespread deployment of host-based anomaly detection has been hindered high false positive rates and computational demands. The authors propose an efficient hidden Markov model training scheme for system-call-based anomaly detection. Hidden Markov models are double stochastic processes that have been found to be useful for modeling nonobvious links in the context of different problems, including models of normal program behavior using system calls, but the training cost can be high. In the proposed method a long training data set is partitioned into subsequences used to train a submodel, which is incrementally merged into the final model using a weighted averaging algorithm. Highly similar subsequences detected by correlation analysis are skipped for efficiency.

We hope these articles will help give a snapshot of current research advances in intrusion detection. It should be evident from this special issue that a great deal of technological progress is being made, but at the same time challenging research issues remain open. In preparing this Special Issue, we wish to thank all the peer reviewers for their thorough and prompt reviews. We are grateful to Editor-in-Chief Ioanis Nikolaidis for his encouragement and constructive suggestions.

Biographies

THOMAS CHEN is a professor at the Institute of Advanced Telecommunications at Swansea University, Wales. He was formerly an associate professor in the Department of Electrical Engineering at Southern Methodist University, Dallas, Texas. Prior to joining SMU in 1997, he was a member of technical staff at GTE (now Verizon) Laboratories, Waltham, Massachusetts. He is a former Editor-in-Chief of *IEEE Communications Magazine* (2006–2007) and former founding Editor-in-Chief of *IEEE Communications Surveys*. He is a Senior Technical Editor of *IEEE Network*, Editor of *Journal on Security and Communication Networks*, and an Editor for IEEE Press. He is a co-editor of *Broadband Mobile Multimedia: Techniques and Applications* (CRC Press, 2008) and co-author of *ATM Switching Systems* (Artech House, 1995). He was co-recipient of the IEEE Communications Society's Fred W. Ellersick best paper award in 1996.

ZHI (JUDY) FU is a Principal Staff Researcher at Motorola Research Laboratories. She received her Ph.D. from the Computer Science Department of North Carolina State University in 2001 and since then joined Motorola Laboratories, focusing on wireless network security research. Her research interests include wireless AAA, protocol vulnerability analysis, security policy, and Web security and intrusion detection. She is an inventor on more than 10 pending patents, and has published over 20 papers in premier conferences and journals. She has been an invited expert reviewer for a number of conferences and journals, including *Journal of Network and Security Management* and *ACM Transactions on Autonomous and Adaptive Systems*.

LIWEN HE [SM] graduated from the University of Sheffield with a Ph.D. degree. In 1999 he joined BT Laboratories, researching optimal design, routing, capacity planning, performance analysis, and resilience in IP, optical and mobile networks. Now he is a principal security researcher at Security Research Centre, BT Group CTO. His main research interests are separation of IP control and data planes, MPLS security, routing protocol security and stability, IP traceback, gigabit passive optical network security, and next-generation router architecture. He has published more than 20 research papers in international conferences and technical journals, and has more than 10 international patents. He has organized and chaired a number of international conferences in network security and served as a technical program committee member. He is a member of the IEEE Communications and Information Security Technical Committee. He also participates in ITU-T Study Group 17 on telecom security standards. He was a guest editor for an *IEEE Communications Magazine* Special Issue on Security in Mobile Ad Hoc and Sensor Networks in February 2008.

TIM STRAYER [SM] is a division scientist in the BBN Technologies Network Technologies Business Unit. He received his Ph.D. from the University of Virginia in 1992, and joined BBN in 1997 from Sandia National Laboratories, California. While at BBN, he has worked on many DARPA and industry sponsored projects in the areas of active networking, satellite packet switching, mobile IP, virtual private networks, routing systems, and attack detection and tracing. His current research pursuits include many aspects of network security, especially IP traceback and botnet detection, where he has done work for DARPA, DHS, and IARPA. He has served as General Chair and a member of the Steering Committee for the IEEE Local Computer Networks Conference, and continues on the technical program committees of a dozen conferences and workshops. He has written over 30 journal and conference papers, several book chapters, and two Addison-Wesley books.