Aggregating Detectors for New Worm Identification

Eric Anderson and Jun Li
University of Oregon
{anderson, lijun}@cs.uoregon.edu

Internet worms have resulted in considerable disruption of our communications infrastructure [1] and could cause much more [2]. We propose a design for coordinating a widely distributed set of network monitors to detect the emergence of new high-speed worms, develop and validate signatures for their identification, and model their spreading dynamics in real time. The primary new contribution of our design is a mechanism for going from the observation that there is a possible worm to automatically validating that observation, developing a signature for the worm if one does not already exist, providing a predictive model for the worm's spreading, and statistically quantifying the level of confidence in these characterizations.

Our work presupposes the ability to detect worm-like behavior with moderate accuracy as it occurs. This supposition appears to be well-founded for a significant class of worms [3], [4], [5].

We represent worm signatures as small graphs annotated with constraints, consisting of event nodes and causation edges [6]. Signature nodes are labeled with constraints describing event characteristics, and the graph as a whole is annotated with inter-node relational constraints. An actual behavior is said to match a signature when the signature is isomorphic to an event history subgraph which represents that behavior, and the events described by the nodes in that subgraph satisfy the signature's constraints.

Whenever a monitor, using existing heuristics and some new ones, observes suspicious events which are not matched by any existing signature, our system performs signature extraction, optimization and validation:

1. A candidate signature is created based on the graph of those suspicious events.
2. The candidate signature is generalized to maximize the number of those events which are matched while minimizing the number of non-suspicious events matched. It is possible that a signature which appears optimal relative to the traffic observed at one site may be undesirable for the Internet as a whole. To remedy this, a monitor further disseminates its locally optimal candidate signature to a set of monitors which then report back both legitimate and false match counts to re-optimize the signature.
3. Signature validation is accomplished by gathering data on the number of observed matches over time from monitors and fitting that data to spreading models using regression analysis. The quality of the fit is indicative of the probability that the observed phenomenon is a worm.

We believe that this proposed system has the potential to detect and characterize any present or potential worm that has a spreading rate and victim host selection accuracy within knowable limits. We have completed an initial design. The next steps are to complete the prototype which is currently under development, evaluate the system, and make the results public. The primary component processes (suspicious traffic detection, signature generation and model fitting) will be studied independently to assess and improve their correctness and performance using actual benign and worm traffic traces, simulated worm behavior and analytical models of possible worms. In addition, we will test the system as a whole using both controlled experimental data and live traffic in our local test bed and then on a large-scale test system such as PlanetLab[7].

This project also motivates new fundamental research on network traffic characterization and measurement methodology. We are looking for characteristics which separate benign and worm traffic with sufficient confidence, either using low-complexity analysis over large sample sizes, or high-complexity analysis over small samples.

References

[1] David Moore, Vern Paxson, Stefan Savage, Colleen Shannon, Stuart Staniford, and Nicholas Weaver, "Inside the slammer worm," *IEEE Security and Privacy*, vol. 4, pp. 33–39, July 2003.

[2] Nicholas Weaver, Vern Paxson, and Stuart Staniford, "A worst-case worm," Tech. Rep., Silicon Defense, 2003.

[3] Matthew Williamson, "Throttling viruses: Restricting propagation to defeat malicious mobile code," Tech. Rep. HPL-2002-172, Hewlett Packard Laboratories, 2002.

[4] Thomas Toth and Christopher Kruegel, "Connection-history based anomaly detection," in *Proceedings of the 2002 IEEE Workshop on Information Assurance and Security*, June 2002, pp. 30–25.

[5] Stuart Staniford, "Containment of scanning worms in enterprise networks," *Journal of Computer Security*, to appear.

[6] S. Staniford-Chen, S. Cheung, R. Crawford, M. Dilger, J. Frank, J. Hoagland, K. Levitt, C. Wee, R. Yip, and D. Zerkle, "GrIDS – A graph-based intrusion detection system for large networks," in *Proceedings of the 19th National Information Systems Security Conference*, 1996.

[7] "PlanetLab web site," http://www.planet-lab.org/.