Results of the DARPA 1998 Offline Intrusion Detection Evaluation

Richard P. Lippmann, Robert K. Cunningham, David J. Fried, Isaac Graf,
Kris R. Kendall, Seth E. Webster, Marc A. Zissman
rpl@sst.ll.mit.edu

MIT Lincoln Laboratory
Room S4-121
244 Wood Street
Lexington, MA 02173-0073

Abstract

DARPA sponsored the first realistic and systematic evaluation of research intrusion detection systems in 1998. As part of this evaluation, MIT Lincoln Laboratory developed a test network which simulated a medium-size government site. Background traffic was generated over two months using custom traffic generators which looked like 100's of users on 1000's of hosts performing a wide variety of tasks and generating a rich mixture of network traffic. While this background traffic was being generated, automated attacks were launched against three UNIX victim machines (SunOS, Solaris, Linux) located on the inside of this simulated government site behind a router. More than 300 instances of 38 different attacks were embedded in roughly two months of training data and two weeks of test data. Six DARPA research sites participated in a blind evaluation where test data was provided without specifying the location of embedded attacks.

Results were analyzed by generating receiver operating characteristic curves (ROCs) to determine the attack detection rate as a function of the false alarm rate. Performance was evaluated for old attacks included in the training data, new attacks which only occurred in the test data, and novel new never-before-seen attacks developed specifically for this evaluation. Detection performance for the best systems was reasonable (above 60% correct) at a false alarm rate of 10 false alarms per day for both old and new probe attacks and attacks where a local user illegally becomes root (u2r). Intrusion detection systems trained on old probe or u2r attacks generalized well to other attacks in these same categories. Detection rates were worse, especially for new and novel denial of service (dos) attacks and attacks where a remote user illegally accesses a local host (r2l). Although detection accuracy for old attacks in these two categories was roughly 80%, detection accuracy for new and novel attacks was below 25% even at high false alarm rates. An intrusion detection system formed from the best components of the submitted systems performed much better than a baseline keyword spotting system that is similar to many commercial and government systems. Across all 120 attacks in the test data, it reduced the false alarm rate by more than two orders of magnitude (from roughly 600 false alarms per day to 6) and it also increased the detection accuracy (from roughly 20% detections to 60%). These results suggest that future intrusion detection research should move towards developing algorithms that find new attacks and away from older approaches that focus on creating rules to find attack signatures.

The current 1999 DARPA evaluation is extending the 1998 evaluation by adding Windows/NT victims, including new Windows/NT attacks, including insider attacks,

and adding attacks designed to measure the ability of intrusion detection systems to detect new, previously unseen, attacks.

Further information concerning the 1998 evaluation, including instructions for obtaining training and test data, is available at http://www.ll.mit.edu/IST/ideval.

Acknowledgements

This work was co-sponsored by the Defense Advanced Research Projects Agency and the Dept. of the Air Force. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Air Force.

Author List (Senior Author)

Dr. Richard P. Lippmann received a B.S. degree in Electrical Engineering from the Polytechnic Institute of Brooklyn, in 1970 and a Ph.D. degree in Electrical Engineering from the Massachusetts Institute of Technology, in 1978. From 1978 to 1981 he was Director of the Communications Engineering Laboratory of the Boys Town Institute for Communication Disorders in Children, Omaha, NE. He worked on speech perception, speech training aids for deaf children, sound alerting aids for the deaf, and signal processing for hearing aids. In 1981 he joined Massachusetts Institute of Technology, Lincoln Laboratory, and is currently a Senior Staff Member in the Information Systems Technology Group. Research interests include speech recognition, developing improved neural network and statistical pattern classifiers, and the application of neural networks and statistics to problems in computer intrusion detection. Dr Lippmann is currently a Distinguished Lecturer for the IEEE Signal Processing Society and he received the first IEEE Signal Processing Magazine award for an article entitled "An Introduction to Computing with Neural Nets" published in April 1987. Abstracts of some publications and an extended biography are available at http://www.ll.mit.edu/IST/pubs.