Modeling TCP/IP network traffic for intrusion detection by genetic evolution

Filippo Neri
DSTA - University of Piemonte Orientale
Corso Borsalino 54, 15100 Alessandria (AL), Italy
email: neri@di.unito.it, neri@mfn.unipmn.it
Phone:+39 011 6706783, Fax:+39 011 751603

Abstract

The detection of intrusions over computer networks (i.e., network access by non-authorized users) can be cast to the task of detecting anomalous patterns of network traffic. Learning systems based on Genetic Algorithms can contribute powerful search techniques for the acquisition of patterns of the network traffic from the large amount of data made available by audit tools.

1 Introduction

The raise in the number of computer intrusion, virtually occurring at any site, determines a strong request for computer security techniques to protect the site data. One kind of such techniques tries to detect when a non-authorized user has gained access (i.e., intrusion) to the computer site. A variety of approaches to intrusion detection do exist [Denning, 1987]. One of this approach tries to characterize the normal usage of the resources under monitoring. An intrusion is then suspected when a significant shift from the resource's normal usage is detected. This approach seems to be more promising because of its potential ability to detect unknown intrusions [Forrest et al., 1996, Lee et al., 1999, Neri, 2000].

We concentrate our research on the impact of different learning methods and of alternative data representation, with respect to the ones used in previous works, on the detection performances. As learning methods, we exploited two rule based systems: a heuristic one, RIPPER [Cohen, 1995], and a genetic based one, REGAL [Neri and Saitta, 1996]. And, as network data made available from the Information Exploration Shootout project and the 1998 DARPA Intrusion Detection Evaluation have been chosen as experimental testbed.

The experimental results support the use of compressed feature values as a promising method to increase detection performances.

References

[Cohen, 1995] Cohen, W. (1995). Fast effective rule induction. In *Proceedings of International Machine Learning Conference 1995*, Lake Tahoe, CA. Morgan Kaufmann.

[Denning, 1987] Denning, D. (1987). An intrusion detection model. *IEEE Transaction on Software Engineering*, SE-13(2):222–232.

[Forrest et al., 1996] Forrest, S., Hofmeyr, S. A., Somayaji, A., and Longstaff, T. A. (1996). A sense of self for unix processes. In *Proceedings of 1996 IEEE Symposium on Computer Security and Privacy.*

[Lee et al., 1999] Lee, W., Stolfo, S., and Mok, K. (1999). Mining in a data-flow environment: experience in network intrusion detection. In *Knowledge Discovery and Data Mining KDD'99*, pages 114–124. ACM Press.

[Neri, 2000] Neri, F. (2000). Mining tcp/ip traffic for network intrusion detection by using a distributed genetic algorithm. In *Proc. of European Conference on Machine Learning*, page In press, Barcelona, Spain. Springer-Verlag.

[Neri and Saitta, 1996] Neri, F. and Saitta, L. (1996). Exploring the power of genetic search in learning symbolic classifiers. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-18:1135–1142.