Received 12 October 2011
Accepted 1 December 2011

6. Conclusion

This paper presents two novel approaches to anomaly detection. These approaches make use of data mining techniques to learn normal pattern bases. Model *FP* uses frequency pattern mining technology to quickly find out frequent system call sequences, so it has efficient advantages in computing and memory usage. On the other hand, Model *TP* makes use of tree-like pattern mining methods, that could get a good detection accuracy. We also thoroughly evaluated our models in false positive rates and false negative rates by different control parameters. The experiments showed and validated their advantages disadvantages.

Future work is going to apply more data mining methods to anomaly detection, and to compare their performances each other including of the models in this paper.

Acknowledgments

This work is supported by grants from the National Science Foundation in China (60873145) and the Discipline Construction Foundation of CUFE.

References

1. S. Forrest, S. A. Hofmeyr, A. Somayaji and T. A. Longstaff, A sense of self for UNIX processes, in *Proc. 1996 IEEE Symposium on Security and Privacy*, IEEE Computer Society Press (Los Alamitos, CA, 1996), pp. 120–128
2. S. A. Hofmeyr, S. Forrest and A. Somayaji, Intrusion detection using sequences of system calls, *J. Computer Security*, **6**(3) (1998) 151-180.
3. W. Lee, S. J. Stolfo and M. Chan, Learning patterns from Unix process execution traces for intrusion detection, in *Proc. AAAI Workshop: AI Approaches to Fraud Detection and Risk Management*, AAAI Press (Menlo Park, 1997), pp.50–56.
4. P. Helman and J.Bhangoo, A statistically based system for prioritizing information exploration under uncertainty, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, **27**(4)(1997) 449–466.
5. W. Lee and S. J. Stolfo, Data mining approaches for intrusion detection, in *Proc. 7th USENIX Security Symposium*, Usenix Association (San Antonio, TX, 1998), pp. 79-94.
6. S. T. Brugger, Data mining methods for network intrusion detection, *Ph. Dissertation, University of California* (Davis, CA, USA, 2004).
7. C. Warrender, S. Forrest and B.Pearlmutter, Detecting intrusions using system calls: alternative data models, in *Proc. 1999 IEEE Symposium on Security and Privacy*, IEEE Computer Society Press (Oakland, CA, USA, 1999), pp.133-145.
8. J. Z.Mohammed, Efficiently mining frequent trees in a forest, in *Proc. 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM Press (Alberta, Canada, 2002), pp. 71-80.
9. W. Lee and S. Stolfo, A data mining framework for Building intrusion detection models, *ACM Transactions on Information and System Security*, **3**(4) (2000) 227-261.
10. W. Lee and X. Dong, Information-theoretic measures for anomaly detection, in *Proc. 2001 IEEE Symp. on Security and Privacy*, IEEE Computer Society Press (2001), pp.130−143.
11. R. Lippmann, D. Fried and I. Graf, Evaluating intrusion detection systems: the 1998 DARPA off-line intrusion detection evaluation, in *Proc. 2000 DARPA Information Survivability Conference and Exposition* (2000), pp. 12-26.
12. J. W. Haines, D. J. Fried and J. Korba, Analysis and results of the 1999 DARPA off-line intrusion detection evaluation, in *Proc. Intl. Symposium on Recent Advances in Intrusion Detection*, Springer-Verlag (2000), pp. 162-182.
13. X. Li, and N. Ye, Decision tree classifiers for computer intrusion detection. *Journal of Parallel and Distributed Computing Practices*, **4**(2) (2001) 179−190.
14. N. Ye, X. Li, Q. Chen, Probabilistic techniques for intrusion detection based on computer audit data. *IEEE Trans. Syst. Man Cybern*., **31**(4) (2001) 266–274.
15. G. Singh, F. Masseglia1, C. Fiot and A. Marascu1, Data mining for intrusion detection: from outliers to true intrusions, in *Proc. the 13th Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Springer (Bangkok, Thailand, 2009), pp. 891-898.
16. A .Patcha and J. M. Park, An overview of anomaly detection techniques: existing solutions and latest technological trends. *Computer Networks*, **51**(7) (2007) 3448-3470.
17. V. Chandola, S. Boriah and V. Kumar, A reference based analysis framework for analyzing system call traces, in *Proc. the ACM Sixth Annual Workshop on Cyber Security and Information Intelligence Research* (2010). Retrieved from http://dx.doi.org/10.1145/1852666.1852703.