A Study on Constructing Fuzzy Association Rules for Intrusion Detection Systems

施衣喬、曹偉駿

E-mail: 9222463@mail.dyu.edu.tw

ABSTRACT

The category of the intrusion detection system (IDS) which detects and guards intruders includes host-based and network-based protection. Regardless of which system is adopted, both of them detect the intrusion by matching the behavior pattern. Moreover, the methods used by the IDS to detect the intrusion are divided into two parts: anomaly and misuse detection. If we use only one intrusion detection mechanism, the false positive rate or false acceptance rate is usually higher. Therefore, how to reduce the false rate is an attractive topic on the intrusion detection mechanism. Beside, since the network utilization is more and more high, how to find out the useful user behavior pattern precisely and efficiently from a large amount of log data is also a major issue in intrusion detection research. In the situation which the intrusive methods are renovated unceasingly, to detect the unknown intrusion is a key point of developing the intrusion detection system. In this thesis, we propose a fuzzy association rules mechanism for host-based intrusion detection systems by using the system log file as data source. First, we cluster the data from system log files with the fuzzy clustering technology. Next, we further derive association rules from the clustering results. Finally, we construct the anomaly and misuse rule database, respectively. Moreover, once there are new data added into log file, we use the fuzzy incremental data mining scheme to derive newer association rules for these data. Our proposed method can not only build the user behavior patterns efficiently from a large amount of data, but also detect the intrusion precisely after combining misuse and anomaly intrusion detection mechanisms. Furthermore, this approach can detect the unknown intrusion. Finally, we also implement an intrusion detection system for validating the proposed mechanism.

Keywords : Intrusion Detection System, Data Mining, Fuzzy Theory, Clustering Technology, Association Rules, Incremental Data Mining.

Table of Contents

REFERENCES

[1] An Introduction to Intrusion Detection, http://www.acm.org/crossroads/xrds2-4/intrus.html [2] FBI/CSI, http://www.gocsi.com/press/20020407.html [3] R. Agrawal, T. Imielinski, and A. Swami, "Mining Association Rules between Sets of Items in Large Database," Proceedings of the ACM SIGMOD Confe- rence on Management of Data, pp. 207-216, 1993.

[4] R. Agrawal, and R. Srikant, "Fast Algorithms for Mining Association Rules in Large Databases," Proceedings 20th International Conference Very Large Data Bases, pp. 478-499, 1994.

[5] N. Ayan, A. Tansel, and E. Arkun, "An Efficient Algorithm to Update Large Itemsets with Early Pruning, " Proceedings of the 5th ACM International Conference on Knowledge Discovery and Data Mining, pp. 287-291, 1999.

[6] J. C. Bezdek, "Pattern Recognition with Fuzzy Objec- tive Function Algorithms," Plenum, New York, 1981.

[7] E. Biermann, E. Cloete, L.M. Venter, "A Comparison of Intrusion Detection systems, Computers and Security," Vol. 20, Issue8, pp. 676-683, 2001.

[8] S.M. Bridges, R.B. Vaughn, "Fuzzy Data Mining and Genetic Algorithms Applied to Intrusion Detection," National Information Systems Security Conference, pp. 13-31, 2000.

[9] D.W. Cheung, S.D. Lee, and B. Kao, "A general incremental technique for maintaining discovered association rules," Proceedings of Database Systems for Advanced Applications, pp. 185-194, 1997.
[10] D.E. Denning, "An intrusion-detection model", IEEE Transactions on Software Engineering, Vol. SE-13, pp. 222-232, 1987.
[11] J.E. Dickerson, J. Juslin, O. Koukousoula, J.A. Dickerson, "Fuzzy intrusion detection," IFSA World Congress and 20th NAFIPS International Conference, Vol. 3, pp. 1506-1510, 2001.
[12] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "The KDD process for extracting useful knowledge from volumes of data." Communications of the ACM, pp. 27-34, 1996.
[13] G. Florez, S.A. Bridges, R.B. Vaughn, "An Improved Algorithm for Fuzzy Data Mining for Intrusion Detection," Proceedings of the North American Fuzzy Information Processing Society Conference (NAFIPS- 2002), pp. 457-462, 2002.
[14] E. Forgy, "Cluster analysis of multivariate data: efficiency versus interpreability of classifications," Biomertrics, Vol. 21, pp. 768, 1965.
[15] G. Helmer, J. Wong, V. Honavar, L. Miller,"Automated Discovery of Concise Predictive Rules for Intrusion Detection," Journal of Systems and Software, Vol. 60, Issue3, pp. 165-175, 2002.
[16] R. Hart, D. Morgan and H. Tran, "An Introduction to Automated Intrusion Detection Approaches," Infor- mation Management and Computer Security 7, pp. 76-82, 1999.
[17] F. Hoppner, Frank Klawonn, Rudolf Kruse, Thomas Runkler, "Fuzzy Cluster Analysis," WILEY, 1999.
[18] Y.C. Hu, R.S. Chen, G.H. Tzeng, "Discovering fuzzy association rules using fuzzy partition methods," Knowledge Based Systems, Vol. 16, pp. 147-147, 2003.
[19] M. Hossain, "Integrating Association Rule Mining and Decision Tree Learning for Network Intrusion Detection: A Preliminary Investigation," International Conference on Information Systems, Analysis and Synthesis, Vol. 11, pp. 65-70, 2002 [20] James and P. Anderson, "Computer Security Threat Monitoring and Surveillance," Technical report, James P. Anderson Co., Fort Washington, PA., 1980.
[21] M.F. Jiang, S.S. Tseng, C.M. Su., "Two-phase clustering process for outliers detection," Pattern Recognition Letters, Vol. 22, pp. 691-700, 2001.
[22] L. Kaufman and P.J. Rousseeuw, "Finding Groups in Data: an Introduction to Cluster Analysis, " John Wiley & Sons, 1990.
[23] C. S. Kuo, T. P. Hong and S. C. Chi, "A study of fuzzy data mining algorithms for quantitative values," Graduate school of Management Science I-Shou University, Thesis, 1999.
[24] C.M. Kuok, A. Fu, and M. Wong, "Mining Fuzzy Association Rules in Databases," SIGMOD record , Vol. 17, No.1, pp. 41-46, 1998.
[25] C.H. Lee, C.R. Lin, and M.S. Chen, "Sliding-Window Filtering: An Efficient Algorithm for Incremental Mining," Proceedings of the ACM 10th International Conference on Information and Knowledge Management, pp. 263-270, 2001.
[26] W. Lee, S.J. Stolfo, P. Chan, E. Eskin, W. Fan, M. Miller, S. Hershkop and J. Zhang, "Real Time Data Mining-based Intrusion Detection," Proceedings of the 2001 DARPA Information Survivability Conference and Exposition (DISCEX II), pp. 89-100, 2001.
[27] W. Lee, S.J. Stolfo, "Data Mining Approaches for Intrusion Detection," Proceedings of the 7th USENIX Security Symposium, 1998.
[28] W. Lee, S.J. Stolfo, and K.W. Mok, "Mining audit data to build intrusion detection models," In 4th International Conference on Knowledge Discovery and Data Mining, pp. 66-72, 1998.
[29] W. Lee, S.J. Stolfo and K.W. Mok, "A Data Mining Framework for Building Intrusion Detection Models," Proceedings of the IEEE Symposium on Security and Privacy, pp. 120-132, 1999.
[30] Y. Li, N. Wu, X.S. Wang ,S. Jajodia, "Enhancing Pro- files for Anomaly Detection Using Time Granul- arities," Journal of Computer Security, pp. 137-158, 2002.
[31] T. Lunt, " Detecting Intruders in Computer Systems," Proceedings of the 1993 Conference on Auditing and Computer Technology, 1993.
[32] J. Luo and Susan M. Bridges, "Mining Fuzzy Association Rules and Fuzzy Frequency Episodes for Intrusion Detection," International Journal of Intelligent Systems, Vol. 15, pp. 687-703 ,2000.
[33] J. A. Marin, J., D. J. Ragsdale, and J. R. Surdu, "A Hybrid Approach to Profile Creation and Intrusion Detection," Proceedings of the DARPA Information Survivability Conference and Exposition - DISCEX , pp. 69-76 ,2001.
[34] B. Ozden, S. Ramaswamy and A. Silberschatz, "Cyclic association rules," Proceedings of the 14th Interna- tional Conference on Data Engineering, pp. 412-421, 1998.
[35] L. Portnoy, E. Eskin, and S.J. Stolfo, "Intrusion Detection with Unlabeled Data Using Clustering," Proceedings of the ACM CCS Workshop on Data Mining for Security Applications, 2001.
[36] S. Ramaswamy, S. Mahajan, and A. Silberschatz, "On the discovery of interesting patterns in association rules," Proceedings of the International Conference on Very Large Data Bases, pp. 368-379, 1998.
[37] S. H. Rubin, "A Fuzzy Approach Towards Inferential Data mining," Computers and Engine, Vol. 35, pp. 267-270, 1998 [38] R. Smith, A. Bivens, M. Embrechts, "Clustering Approaches for Anomaly Based Intrusion Detection," Walter Lincoln Hawkins Graduate Research Confe- rence, 2002.
[39] R. Srikant and R. Agrawal, "Mining Quantitative Association Rules in Large Relational Tables," Proceedings of ACM SIGMOD International Confe- rence on Management Data, Montreal Canada, pp. 1-12, 1996.

[40] W.J. Tsaur and I-Ming Fan, "Anomaly Detection Mechanisms for Web Severs in Linux Environments," Communications of the CCISA, Vol. 8, No. 4, 2002.
[41] T. Verwoerd and R. Hunt, "Intrusion Detection Techniques and Approaches,"Computer Communications, Vol. 25, pp. 1356-1365, 2002.
[42] L.A. Zadeh. Fuzzy Sets, Information Control, Vol. 8, pp. 338-353, 1965.
[43] O.R. Zaiane, M. Xin, J. Han, "Discovering Web Access Patterns and Trends by Applying OLAP and Data Mining Technology on Web Logs," Proceedings of Advances in Digital Libraries Conference (ADL- 98), pp. 19-29, 1998.