A Study on Mining Fuzzy Association Rules and Episode Rules for Intrusion Detection

柯文元、曹偉駿

E-mail: 9422531@mail.dyu.edu.tw

ABSTRACT

At present, most of intrusion detection systems still existing excessive fault alerts problems, including high false positive and low detection rate. These erroneous alerts will let system managers take efforts to handle. Hence, how to enhance the detection ability is an important issue in computer security. The thesis proposes an intrusion detection system based on the combination of association rules and frequency episode rules. Firstly, we use the fuzzy clustering technique to cluster network packets. Secondly, we also mine the possible rules using the fuzzy association rule technology from each cluster in order to discover single intrusion attack. Besides, on the purpose of discovering the intrusion phase and order, we further mine fuzzy frequency episode rules to discover the multiple serial relationship from each cluster. Finally, we construct the rules into normal and abnormal rule database, respectively. The contribution of the proposed schemes is to accelerate the detection rate of the single intrusion or multiple intrusions by using the fuzzy data mining technique. In this thesis, we also implement the proposed schemes to validate their feasibility.

Keywords : Intrusion Detection System, Data Mining, Fuzzy Theory, Clustering Technology, Association Rules, Frequency Episode Rules

Table of Contents

封面內頁 簽名頁 授權書.....iii 中文摘要 iv
ABSTRACT.....v 誌謝.....vi 目錄.....vii 圖目錄.....ix 表目錄.....x 第一章 緒論.....1 1.1 研究背景與動機.....1 1.2 研究目的.....3 1.3 論文架構 .....4 第二章 文獻探討.....6 2.1 入侵偵測系統.....6 2.1.1 入侵偵測發展簡史.....7 2.1.2 入侵偵測系統分類方式.....9 2.1.3 網路攻擊類型分類.....13 2.1.4 DARPA Dataset.....14 2.2 模糊理論.....14 2.3 資料探勘.....18 2.3.1. 群集技術(Clustering Technology).....20 2.3.2. 關聯法則(Association Rules).....25 2.3.3. 情境法則(Episode Rules).....32 第三章 模糊關聯與情境法則探勘機制.....36 3.1. 研究流程.....36 3.2. 樣本特徵分析.....37 3.3. 研究方法.....41 第四章 實驗設計與分析.....46 4.1 實驗環境.....46 4.2. 系統實作.....47 4.3. 實驗結果與分析.....55 第五章 結論與未來發展方向.....58 5.1. 結論.....58 5.2. 後續發展建議.....59 參考文獻.....60

REFERENCES

[1] J. P. Anderson, “ Computer Security Threat Monitoring and Surveillance, ” Technical Report, James P. Anderson Co., Fort Washington, Pennsylvania, pp.76-82, 1980.

[2] R. Agrawal, T. Imielinski and A. Swami, “ Mining Association Rules between Sets of Items in Large Database, ” Proceedings of the ACM SIGMOD Conference on Management of Data, pp. 207-216, 1993.

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

[4] S. Axelsson, “ Intrusion Detection Systems: A Taxomomy and Survey, ” Technical Report, Dept. of Computer Engineering, Chalmers University of Technology, Goteborg, Sweden, pp. 99-15, 2000.

[5] J. C. Bezdek, “ Pattern Recognition with Fuzzy Objective Function Algorithms, ” Plenum, New York, 1981.

[6] D.E. Denning, “ An intrusion-detection model, ” The IEEE Transactions on Software Engineering, Vol. SE-13,pp. 222-232, 1987.

[7] D. E. Denning, D. Edwards, R. Jagannathan, T. Lunt and P. Neumann, “ A Prototype IDES: A Real-Time Intrusion Detection Expert System, ” Computer Science Laboratory, SRI International, 1987.

[8] H. Debar, M. Dacier, and A. Wespi, “ A Revised Taxonomy for IntrusionDetection Systems, ” Annales des Telecommunications, Vol. 55, No. 7/8, pp. 361-378, 2000 [9] T. Dobrowiecki, “ Episode Mining to Automatically Filter False Alarms, ” Proceedings of the 10th PhD Mini-Symposium on IEEE Hungary Section, pp. 44-45, 2003.

[10] J. E. Dickerson, J. Juslin, O. Koukousoula and J. A. Dickerson, “ Fuzzy intrusion detection, ” IFSA World Congress and 20th NAFIPS International Conference, Vol. 3, pp. 1506-1510, 2001.
[11] G. Florez, S.A. Bridges and 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.
[12] U. Fayyad, G. P. Shapiro, and P. Smyth, “ The KDD process for extracting useful knowledge from volumes of data, ” Communications of the ACM, pp.27-34, 1996.
[13] FBI/CSI, http://www.gocsi.com/press/20020407.html , 2002 [14] E. Forgy, “ Cluster analysis of multivariate data:efficiency versus interpreability of classifications, ” Biomertrics, Vol. 21, pp. 768, 1965.
[15] R. Graham, “ FAQ: Network Intrusion Detection System, ” http://www.robertgraham.com/pubs/network- intrusion-detection.html, 2000.
[16] H. Han, X. L. Lu, J. Lu, C. Bo and R. L. Yong, “ Data mining aided signature discovery in network-based intrusion detection system, ” Source ACM SIGOPS Operating Systems Review, Vol. 36 , Issue 4, pp. 7-13, 2002 [17] L. Heberlein, G. Dias, K. Levitt, B. Mukherjee, J. Wood and D. Wolber, “ A Network Security Monitor, ” Proceedings of the IEEE Symposium on Research in Security and Privacy, 1990.
[18] F. Hoppner, F. Klawonn, R. Kruse and T. Runkler, “ Fuzzy Cluster Analysis, ” WILEY, 1999.
[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] P. Innella, “ The Evolution of Intrusion Detection Systems, ” http://www.securityfocus.com/infocus/1514 , 2001 [21] M. F. Jiang, S. S. Tseng and C. M. Su., “ Two-phase clustering process for outliers detection, ” Pattern Recognition Letters, Vol. 22, pp. 691-700, 2001.
[22] 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.
[23] C. M. Kuok, A. Fu, and M. Wong, “ Mining Fuzzy Association Rules in Databases, ” SIGMOD record , Vol. 17, No.1, pp. 41-46, 1998.
[24] 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.
[25] W. Lee, S. J. Stolfo and K. W. Mok, “ A data mining framework for building intrusion detection models, ” Proceedings of the 1999 IEEE Symposium on Security and Privacy, pp. 120-132, 1999.
[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] Y. Li, N. Wu, X. S. Wang and S. Jajodia, “ Enhancing Profiles for Anomaly Detection Using Time Granularities, ” Journal of Computer Security, pp. 137-158, 2002.
[28] R. Lippmann, D. Fried, I. Graf, J. Haines, K. Kendall, D. McClung, D. Weber, S. Webster, D. Wyschogrod, R. Cunningham and M. Zissman, “ Evaluating intrusion detection systems: the 1998 DARPA off-line intrusion detection evaluation, ” Proceedings of the 2000 DARPA Information Survivability Conference and Exposition, Vol. 2, pp. 12-26, 2000.
[29] T. Lunt and R. Jagannathan, “ A Prototype Real-Time Intrusion Detection Expert System, ” Proceedings of the 1988 IEEE Symposium on Security and Privacy, Oakland, CA, 1988.
[30] T. Lunt, “ Detecting Intruders in Computer Systems, ” Proceedings of the 1993 Conference on Auditing and Computer Technology, 1993.
[31] 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.
[32] H. Mannila, H. Toivonen, and A. I. Verkamo, “ Discovery of Frequent Episodes in Event Sequences, ” Data Mining and Knowledge Discovery, Vol. 1, 1997.
[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] MIT Lincoln Laboratory – DARPA Intrusion Detection Evaluation, http://www.ll.mit.edu/IST?ideval/index.html, 2002.
[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] M. Qin, and K. Hwang ” , Frequent Episode Rules for Internet Anomaly Detection, ” Proceedings of The Third IEEE International Symposium on Network Computing and Applications, pp.161-168, 2004 [37] S. H. Rubin, “ A Fuzzy Approach Towards Inferential Datamining, ” Computers and Engine, Vol. 35, pp.267-270, 1998 [38] Symantec, “ Symantec Internet Security Threat Report, ” https://enterprisesecurity.symantec.com/Content/displaypdf.cfm?SSL=YES&EID=0&PDFID=665&promocode=ITR [39] Symantec Taiwan, “ Symantec network security threat research, ” Sep.2004. http://www.symantec.com.tw/region/tw/avcenter/10187537_SIST_es_CH.pdf [40] S. E. Smaha, “ Haystack: An Intrusion Detection System, ” Proceedings of Fourth Aerospace, Orlando, Florida, 1988.
[41] M. Sebring, E. Shellhouse, M. Hanna and R. Whitehurst, “ Expert Systems in Intrusion Detection: A Case Study, ” Proceedings of the 11th National Computer security Conference, 1988.
[42] R. Smith, A. Bivens and M. Embrechts, “ Clustering Approaches for Anomaly Based Intrusion Detection, ” Proceedings of the Walter Lincoln Hawkins Graduate Research Conference, 2002.

[43] R. Srikant and R. Agrawal, “ Mining Quantitative Association Rules in Large Relational Tables, ” Proceedings of ACM SIGMOD International Conference on Management Data, Montreal Canada, pp.1-12, 1996.
[44] W. J. Tsaur and I M. Fan, “ Anomaly Detection Mechanisms for Web Severs in Linux Environments, ” Communications of the CCISA, Vol. 8, No. 4, 2002.
[45] W. J. Tsaur and Y. C. Shieh, “ Constructing Fuzzy Association Rules for Intrusion Detection Systems, ” Proceedings of the 2003 National Computer Symposium, pp. 1256-1263, 2003.
[46] T. Verwoerd and R. Hunt, “ Intrusion Detection Techniques and Approaches, ” Computer Communications, Vol. 25, pp. 1356-1365, 2002.
[47] L. A. Zadeh, “ Fuzzy Sets, ” Information Control, Vol. 8, pp. 338-353, 1965.
[48] O. R. Zaiane, M. Xin and 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.