# 網路入侵偵測之調適性模糊資料探勘


## 鍾居璋、陳鴻文

E-mail: 9607617@mail.dyu.edu.tw



**摘 要**

由於計算機與電腦網路的普及，網路攻擊行為隨之增加，攻擊手法也日趨複雜，因此入侵偵測系統IDS應運而生；然而傳統IDS普遍存在著偵測率過低或誤報率過高的問題。本研究基於資料探勘與模糊理論的技術，以DARPA 集合的封包作為處理資料。首先針對訓練用的攻擊型網路封包，我們使用了C5.0演算法來建構出相對的誤用入侵偵測規則庫。另一方面，則將所有訓練用的網路封包，利用ISODATA演算法進行適當的分群後，再使用C5.0演算法來建構出這67個群集的決策樹，以形成龐大的類異常偵測規則庫。此外，藉由序列相關分析(sequential pattern)技術，本研究從同一來源及目的之封包序列中，找出入侵者的三種循序式入侵行為。若發現使用者正在依循某種特定入侵序式模式時，判斷出極有可能後續會有入侵行為的發生，以期降低入侵所造成的傷害。 本研究提出之模糊分類法及異常行為序列封包分析法，可達到平均86%及60%的辨識準確率，同時亦降低了誤報率到平均2%的程度；故可明確且有效地判斷使用者是否具有入侵之行為，大幅改善了傳統IDS所面臨的偵測率過低或誤報率過高的問題。



關鍵詞 : 資料探勘、模糊理論、入侵偵測系統、決策樹、次序相關分析


**目錄**



**參考文獻**


一、中文部份 台灣網路資訊中心(2007)，[線上資料]，來源: http://www.myhome.net.tw/2007_03/main03_1.htm [2007, September]. 邱美珍(1996)，決策樹學習法中連續屬性之分類研究，中原大學資訊工程研究所未出版碩士論文。 柯文元(2005)，模擬關聯與情境法則探勘於入侵偵測之研究，大葉大學資訊管理學系碩士班未出版碩士論文。 二、英文部份 Adler, P. S. (1993). Time-and-motion regained. Harvard Business Review, 71 (1), 97-108. Agrawal, R. & Srikant, R. (1994). Mining Sequential Patterns. Proceedings of the International Conference on Data Engineering. Berry, M. J. A. & Linoff, G.. (1997). Data Mining Technique for Marketing Sale and Customer Support. Wiley Computer, New York, NY. Berry, M., & Linoff, G.. (1997). Data Mining Techniques for marketing, sales, and Customer Support. New York. Wiley Computer Publishing. Bezdek, J. C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum, New York. C5.0. Available: http://www.rulequest.com/ [No date]. Debar, H., Becker, M., & Siboni, D. (1992). A Neural Network Component for an Intrusion Detection System. IEEE Security and Privacy, 10(2), 155-169. Dickerson, J. E., Juslin, J., Koukousoula, O., & Dickerson, J. A. (2001). Fuzzy intrusion detection. IFSA World Congress and 20th NAFIPS International Conference, 3, 1506-1510. Fayyad, U., Shapiro, G. P. , & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. Communications of the ACM, 31(7), 27-34. Han, J. (1999). Data Mining. Encyclopedia of Distributed Computing, Kluwer Academic Publishers, 1-7. Helman, P., & Liepins, G. (1993). Statistical Foundations of Audit Trail

Analysis for the Detection of Computer Misuse. IEEE Software Engineering. 14(5), September – October. Kumar, S., & Spafford, E. (1994). A Pattern Matching Model for Misuse Intrusion Detection. the 17th National Computer Security Conference. Marin, J. A., Ragsdale, D. J., & Surdu, J. R. (2001). A Hybrid Approach to Profile Creation and Intrusion Detection. Proceedings of the DARPA Information Survivability Conference and Exposition - DISCEX, 69-76. MIT Lincoln Laboratory - DARPA Intrusion Detection Evaluation. Available: http://www.ll.mit.edu/IST/ideval/index.html [1997]. Piatetsky-Shapiro, G., & Frawley, W. J. (1991). Knowledge Discovery in Databases. AAAI/MIT Press. Portnoy, L., Eskin, E., & Stolfo, S. J.(2001). Intrusion Detection with Unlabeled Data Using Clustering. Proceedings of the ACM CCS Workshop on Data Mining for Security Applications. Smith, R., Bivens, A., & Embrechts, M.(2002). Clustering Approaches for Anomaly Based Intrusion Detection. Proceedings of the Walter Lincoln Hawkins Graduate Research Conference. Sundaram, A. An Introduction to Intrusion Detection, ACM Crossroads Student Magazine. Available: http://www.acm.org/ crossroads/xrds2-4/intrus.html [No date]. Symantec. Available: http://www.symantec.com/index.htm [2007]. Tsaur ,W. J., & Fan, I M. (2002). Anomaly Detection Mechanisms for Web Severs in Linux Environments. Communications of the CCISA, 8(4). Qin, M, & Hwang, K, (2004). Frequent Episode Rules for Internet Anomaly Detection. Proceedings of The Third IEEE International Symposium on Network Computing and Applications, 161-168. Quinlan, J.R., (1993). C4.5 Programs for machine learning. Morgan Kaufmann Publishers, San Mateo, California. Zadeh, L. A., (1965). Fuzzy sets. Information and Control , 8, 338-353. Zadeh, L. A., (1975). The concept of a linguistic variable and its application to approximate reasoning I, II, III. Information Science , 8 , 199-251 , 301-357 ; 9, 43-80. Zaiane, O. R., Xin, M., & Han, J. (1998). 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), 19-29.