Computational Intelligence for Network Intrusion Detection: Recent Contributions

Asim Karim

Dept. of Computer Science, Lahore University of Management Sciences,
Opp. Sector U, DHA, Lahore 54792, Pakistan
akarim@lums.edu.pk

Abstract. Computational intelligence has figured prominently in many solutions to the network intrusion detection problem since the 1990s. This prominence and popularity has continued in the contributions of the recent past. These contributions present the success and potential of computational intelligence in network intrusion detection systems for tasks such as feature selection, signature generation, anomaly detection, classification, and clustering. This paper reviews these contributions categorized in the sub-areas of soft computing, machine learning, artificial immune systems, and agent-based systems.

1 Introduction

Network intrusion detection (NID) is essentially a pattern recognition problem in which network traffic patterns are classified as either ‘normal’ or ‘abnormal’. This is a difficult problem because of the wide diversity of traffic patterns and the need for accuracy in real-time operation. The NID problem has been tackled since the early days of computer networks but an efficient, effective, and practical solution is still being sought [30]. The incorporation of computational intelligence in network intrusion detection systems (NIDS) presents the greatest potential for an acceptable solution. Computational intelligence has yielded successful solutions to similar problems in other domains such as the highway incident detection problem [12]. Like the NID problem, the highway incident detection problem requires rapid and reliable identification of incidents (e.g. accidents) from raw traffic data obtained from sensors located at different points on the highway network.

In recent years significant contributions have been made towards the solution of the NID problem. Many of these contributions employ computational intelligence approaches derived and motivated from concepts of biological intelligence. This paper reviews recent computational intelligence contributions in the areas of soft computing (section 2), machine learning (section 3), artificial immune systems (section 4), and agent-based systems (section 5).

2 Soft Computing

Neural networks, fuzzy systems, and evolutionary algorithms have been used extensively for network intrusion detection. Neural networks are known to be

Y. Hao et al. (Eds.): CIS 2005, Part I, LNAI 3801, pp. 170 – 175, 2005.

effective classifiers and feature selectors / dimensionality reducers. As such, neural networks have been a component of several recent NIDSs. As part of the MAID intrusion detection system, a BP neural network is used for identifying DoS attacks [16]. The inputs to the neural network are statistically preprocessed SNMP MIB (management information base) data. Ng et al. use a stochastic radial-basis function neural network to extract and rank features for DoS attacks from the DARPA database [37]. NIDSs based on support vector machines (SVMs) and neural networks, and the study of various features for intrusion detection based on the performance of SVM and neural network classifiers is presented in [4, 34]. It has been found that SVMs perform better than neural networks for intrusion classifications. Liu et al. describe the design and performance of an anomaly detection system using the ART neural network [40]. Unlike the previously mentioned neural network based systems, this system can be trained in real-time operations. Valdes uses a competitive neural network to cluster unlabeled data into groups [3]. This overcomes the issue of having labeled normal and abnormal traffic for training. Sarasamma et al. propose a hierarchical Kohonen network model for detecting anomalies and outliers in connection events database [35]. Each level is a simple winner-take-all neural network that identifies features in the problem space. In this way, a clustering of the data is obtained which is then labeled using confidence measures.

Fuzzy sets and fuzzy logic have been used successfully for problems involving vagueness and imprecision including NID. The FIRE detection system, which is based on fuzzy techniques, uses the fuzzy cognitive map for decision making [17]. In the system design, several features in TCP, UDP, and ICMP protocol headers are evaluated by modeling some of them as fuzzy variables and visualizing them to ascertain their ‘fuzziness’. Lee and Mikhailov propose a fuzzy classification system for intrusion detection based on numeric features [19], while Shah et al. describe a fuzzy clustering approach applied to statistics obtained from low level kernel system and network data [13]. A hybrid neuro-fuzzy intrusion detection system is described in [32].

Evolutionary algorithms (EAs) provide robust solutions to many problems by adopting a process of selection and evolution similar to the natural process of evolution. For NID, evolutionary algorithms have proven useful for signature generation and feature selection. Pillai et al. describe a genetic algorithm (GA) for generating high quality rules from the DARPA intrusion detection database [23]. The generation of fuzzy rules or signatures from anomaly traffic is described in [9]. These rules serve as negative selectors in an immunity-based intrusion detection system. Florez et al. present a fuzzy association rule mining approach for generating fuzzy rule sets [11]. A threshold on the similarity between different (fuzzy) sets of rules is used to detect intrusions. Middlemiss and Dick use GA to rank features by evolving feature weights over multiple generations with a nearest neighbor rule for fitness function [22]. An evolutionary algorithm has also been used to optimally design a radial-basis function neural network for intrusion detection [1]. Song et al. present an efficient algorithm for training from large data sets using linear structured genetic programming (GP) [7]. The GP algorithm is able to find features from the KDDCup-99 data set that contains 0.5 million records in 15 minutes. The design, implementation, and evaluation of GA-based misuse detection system is described in [31]. A recent study of the suitability of a fitness function for NID is presented in [28].

3 Machine Learning

Several machine learning and data mining techniques have been proposed for NID. Markov models have been used for capturing the sequence of events in network traffic and determining the probabilities of significant changes from the norm. Ye et al. study the performance of a Markov-chain model of normal events for detecting abnormal ones by applying it to UNIX system logs [25], while Anming and Chunfu investigate hidden Markov models for network intrusion detection [39]. Zhang and Zhu integrate hidden Markov models and neural networks for NID [36]. In general, Markov based models are efficient but often less accurate because of the wide diversity of normal behaviors. Decision tree models have been employed to analyze protocol headers for feature selection and intrusion detection [33], while a comparison of naïve Bayes and decision trees is provided in [26]. Cho utilizes both soft and hard computing approaches by integrating a hidden Markov model of normal traffic with neural network and fuzzy logic [33]. Feature selection and dimensionality reduction is done with a SOM, construction of a database of normal sequences is performed with a Markov model, and a fuzzy inference system is used for decision making. A NIDS based on a probabilistic data mining approach for learning rules of normal and abnormal traffic is presented in [24].

4 Artificial Immune Systems

The human immune system is an example of an efficient and effective intrusion defense system. It is capable of identifying beneficial (self) and harmful (non-self) elements in the body and taking action to expel or eradicate unwanted elements. This defense system has a direct analogy to the network intrusion detection and response system. It is therefore not surprising that immunity-based NIDS have been proposed since the 1990s. The recent works in this area have focused on developing detectors for self and non-self. An investigation of the parameters involved in the creation of non-self detectors is presented in [15], while in [18] EA is employed to create hyper-ellipsoid detectors for negative selection. Gomez et al. describe the generation of fuzzy rules that characterize the non-self of an immunity-based NIDS [14]. Hang and Dai test an immunity-based NIDS with both positive and negative selection classes [38]. A study of anomaly detection using different features of traffic and protocol header in the context of an immunity-based NIDS is presented in [10]. Esponda et al. present a formal framework for positive and negative selection in immunity-based systems [8]. A survey of immunity-based computer defense systems and discussion of future development trends is presented in [20].

5 Agent-Based Systems

Computer networks are distributed in nature. As such, distributed agent-based systems are commonly proposed for their security including intrusion detection. Miller and Inoue describe the performances of distributed NIDSs in which agents perform local feature extraction using SOMs and global decision making [27].

Dasgupta and Brian present a distributed architecture for network security using packet, process, system, and user information [6]. It combines profile-based anomaly detection and parametric pattern matching in an agent-based system motivated from the human immune system [5]. Specific agents implement neural network classification and fuzzy inference for decision making at a central location. Siraj et al. describe their intelligent intrusion detection system which involves distributed information collection and central information processing and decision making using fuzzy cognitive maps [2]. An intelligent agent based distributed architecture is presented in [29], while a cooperative/collaborative architecture is presented in [41]. Zhou et al. propose a cooperative model for network security in which different elements of the security chain – firewalls, intrusion detection, VPNs – interact among themselves while maintaining a certain degree of independence [21].

6 Conclusion

Computational intelligence has figured prominently in many proposed solutions to the network intrusion detection problem in the recent past. In these solutions, computational intelligence provides functionalities such as feature selection, signature generation, anomaly detection, and decision making. This paper reviews these solutions categorized into the sub-areas of soft computing, machine learning and data mining, artificial immune systems, and agent-based systems. It is observed that computational intelligence holds great promise for an effective and practical solution to the network intrusion detection problem.

Acknowledgment

This work was supported by a research grant from PTCL R&D Fund, Pakistan.

References

1. A. Hofmann, T. Horeis, and B. Sick: Feature Selection for Intrusion Detection: An Evolutionary Wrapper Approach. Proc., International Joint Conference on Neural Networks (IJCNN ’04) (2004) 1563-1568
2. A. Siraj, R.B. Vaughn, and S.M. Bridges: Intrusion Sensor Data Fusion in an Intelligent Intrusion Detection System Architecture. Proc. Hawaii International Conference on System Sciences (2004) 902-911
3. A. Valdes: Detecting Novel Scans Through Pattern Anomaly Detection. Proc. DARPA Information Survivability Conference and Exhibition (DICEX ’03) (2003) 140-151
4. A.H. Sung and S. Mukkamala: Identifying Important Features for Intrusion Detection Using Support Vector Machines and Neural Networks. Proc. Symposium on Applications and the Internet (SAINT ’03) (2003) 209-216
5. D. Dasgupta and F. Gonzalez: An Immunity-based Technique to Characterize Intrusions in Computer Networks. IEEE Transactions on Evolutionary Computing, Vol. 6, No. 3 (2002) 281-291

6. D. Dasgupta and H. Brian: Mobile Security Agents for Network Traffic Analysis. Proc. DARPA Information Survivability Conference and Exhibition (2001) 332-340
7. D. Song, M.I. Haywood, A.N. Zincir-Heywood: Training Genetic Programming on Half a Million Patterns: An Example from Anomaly Detection. IEEE Transactions on Evolutionary Computation, Vol. 9, No. 3 (2005) 225-239
8. F. Esponda, S. Forrest, and P. Helman: A Formal Framework for Positive and Negative Decision Schemes. IEEE Transactions on Systems, Man, and Cybernetics – Part B (Cybernetics), Vol. 34, No. 1 (2004) 357-373
9. F. Gonzalez, J. Gomez, M. Kaniganti, and D. Dasgupta: An Evolutionary Approach to Generate Anomaly (Attack) Signatures. Proc. IEEE International Workshop on Information Assurance (IWIA '03) (2003) 251-259
10. F. Seredynski: Some Issues in Solving the Anomaly Detection Problem Using the Immunological Approach. Proc. IEEE International Parallel and Distributed Processing Symposium (IPDPS '05) (2005) 188-195
11. G. Florez, S.M. Bridges, and R.B. Vaughn: An Improved Algorithm for Fuzzy Data Mining for Intrusion Detection. Proc. North American Fuzzy Processing Society (2002) 457-462
12. H. Adeli and A. Karim: Wavelets in Intelligent Transportation Systems. John Wiley & Sons UK (2005)
13. H. Shah, J. Undercoffer, and A. Joshi: Fuzzy Clustering for Intrusion Detection. Proc. IEEE International Conference on Fuzzy Systems (2003) 1274-1278
14. J. Gomez, F. Gonzalez, and D. Dasgupta: An Immuno-Fuzzy Approach to Intrusion Detection. Proc. IEEE International Conference on Fuzzy Systems (2003) 1219-1224
15. J. Kim and P.J. Bentley: Towards an Artificial Immune System for Network Intrusion Detection: An Investigation of Clonal Selection with a Negative Selection Operator. Proc. Congress on Evolutionary Computing (2001) 1244-1252
16. J. Li and C. Manikopoulos. Early Statistical Anomaly Intrusion Detection of DoS Attacks Using MIB Traffic Parameters. Proc. IEEE International Workshop on Information Assurance (IWIA '03) (2003) 53-59
17. J. Xin, J.E. Dickerson, and J.A. Dickerson. Fuzzy Feature Extraction and Visualization for Intrusion Detection. Proc. IEEE International Conference on Fuzzy Systems (2003) 1249-1254
18. J.M. Shapiro, G.B. Lamont, and G.L. Peterson: An Evolutionary Algorithm to Generate Hyper-Ellipsoid Detectors for Negative Selection. Proc. GECCO '05 (2005) 337-344
19. K. Lee and L. Mikhailov: Intelligent Intrusion Detection System. Proc. IEEE International Conference on Intelligent Systems (2004) 497-502
20. K.P. Anchor, P.D. Williams, G.H. Gunsch, and G.B. Lamont: The Computer Defense Immune System: Current and Future Research in Intrusion Detection. Proc. Congress on Evolutionary Computing (2002) 1027-1032
21. L. Zhou, F. Liu, and J. Wu: Research on Co-operative Computer Network Security Technologies. Proc. IEEE International Conference on Systems, Man, and Cybernetics (2004) 1164-1168
22. M.J. Middlemiss and G. Dick: Weighted Feature Extraction Using a Genetic Algorithm for Intrusion Detection. Proc. Congress on Evolutionary Computing (2003) 1669-1675
23. M.M. Pillai, J.H.P. Eloff, and H.S. Venter: An Approach to Implement an Intrusion Detection System Using Genetic Algorithms. Proc. SAICSIT '04 (2004) 228-235
24. M.V. Mahoney and P.K. Chan: Learning Rules for Anomaly Detection of Hostile Network Traffic. Proc. IEEE International Conference on Data Mining (ICDM '03) (2003) 601-604

25. N. Ye, Y. Zhang, and C.M. Borror: Robustness of the Markov Chain Model for Cyber-Attack Detection. IEEE Transactions on Reliability, Vol. 53, No. 1 (2004) 116-123
26. N.B. Amor, S. Benferhat, and Z. Elouedi: Naïve Bayes vs Decision Trees in Intrusion Detectin Systems. Proc. SAC '04 (2004) 420-424
27. P. Miller and A. Inoue: Collaborative Intrusion Detection System. Proc. North American Fuzzy Information Processing Society (2003) 519-524
28. P.A. Diaz-Gomez and D.F. Hougen: Analysis and Mathematical Justification of a Fitness Function Used in an Intrusion Detection System. Proc. GECCO '05 (2005) 1591-1592
29. Q. Xue, L. Guo, and J. Sun: The Design of a Distributed Network Intrusion Detection System IA-NIDS. Proc. International Conference on Machine Learning and Cybernetics (2003) 2305-2308
30. R.A. Kemmerer and G. Vigna: Intrusion Detection: A Brief History and Overview. IEEE Computer (2002) 27-30
31. R.H. Gong, M. Zulkernine, and P. Abolmaesumi: A Software Implementation of a Genetic Algorithm Based Approach to Network Intrusion Detection. Proc. International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel Distributed Computing (2005) 246-253
32. S. Chavan, K. Shah, N. Dave, and S. Mukherjee: Adaptive Neuro-Fuzzy Intrusion Detection Systems. Proc. International Conference on Information Technology: Coding and Computing (ITCC '04) (2004) 70-74
33. S. Cho: Incorporating Soft Computing Techniques into a Probabilistic Intrusion Detection System. IEEE Transactions on Systems, Man, and Cybernetics – Part C (Applications and Reviews), Vol. 32, No. 2 (2002) 154-160
34. S. Makkamala and A.H. Sung: Detecting Denial of Service Attacks Using Support Vector Machines. Proc. IEEE International Conference on Fuzzy Systems (2003) 1231-1236
35. S.T. Sarasamma, Q.A. Zhu, and J. Huff: Hierarchical Kohonen Net for Anomaly Detection in Network Security. IEEE Transactions on Systems, Man, and Cybernetics – Part B (Cybernetics), Vol. 35, No. 2 (2005) 302-312
36. T. Abbes, A. A. Bouhoula, and M. Rusinowitch: Protocol Analysis in Intrusion Detection using Decision Tree. Proc. International Conference on Information Technology, Coding, and Computing (ITCC '04) (2004) 404-408
37. W. Ng, R. Chang, and D. Yeung: Dimensionality Reduction for Denial of Service Detection Problems Using RBFNN Output Sensitivity. Proc. International Conference on Machine Learning and Cybernetics (2003) 1293-1298
38. X. Hang and H. Dai: Applying Both Positive and Negative Selection to Supervised Learning for Anomaly Detection. Proc. GECCO '05 (2005) 345-352
39. X. Zhang and Z. Zhu: Combining the HMM and the Neural Network Models to Recognize Intrusions. Proc. International Conference on Machine Learning and Cybernetics (2004) 956-961
40. Y. Liu, D. Tian, and A. Wang: ANNIDS: Intrusion Detection System Based on Artificial Neural Network. Proc. International Conference on Machine Learning and Cybernetics (2003) 1337-1342
41. Y. Xiaoping and D. Yu: An Auto-Configuration Cooperative Distributed Intrusion Detection System. Proc. World Congress on Intelligent Control and Automation (2004) 279-283
42. Z. Anming and J. Chunfu: Study on the Applications of Hidden Markov Models to Computer Intrusion Detection. Proc. World Congress on Intelligent Control and Automation (2004) 256-260