ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/284736478

Data mining techniques in DoS/DDoS attack detection: A literature review

Article *in* INFORMATION, Japan · August 2015

CITATIONS
10

2 authors:

Bayu Adhi Tama
Ulsan National Institute of Science and Technology
31 PUBLICATIONS **105** CITATIONS
SEE PROFILE

Kyung Hyune Rhee
Pukyong National University
168 PUBLICATIONS **626** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Conference Paper View project

All content following this page was uploaded by Bayu Adhi Tama on 08 December 2017.

The user has requested enhancement of the downloaded file.

Data Mining Techniques in DoS/DDoS Attack Detection: A Literature Review

Bayu Adhi Tama, Kyung-Hyune Rhee

Department of IT Convergence and Application Engineering, Pukyong National University
Busan 48513, South Korea
E-mail: {bayu, khrhee}@pknu.ac.kr

Abstract

This paper attempts to classify papers concerning DoS/DDoS attack detection using data mining techniques. Thirty five papers were selected and carefully reviewed by authors from two online journal databases. Each of selected paper was classified based on the function of data mining such as association, classification, clustering, and hybrid methods. The findings of this work indicate that classification and hybrid techniques received a great deal of attention from researchers. Our literature review provides a state of the art analysis concerning DoS/DDoS attack detection using data mining techniques.

Key Words: —DoS/DDoS attack, data mining, survey, classification.

1. Introduction

With the development of Internet, many businesses are now shifting their operation to online. Internet become comfortable place for attackers to run their misbehave activities. So as, critical security mechanism such as attack detection and prevention is necessary. Denial of Service (DoS) or distributed DoS (DDoS) are aimed at making network resource unavailable to its legitimate users [1]. In the past few years, DoS/DDoS attack has grown significantly both in size and frequency. According to Arbor Network, 90% respondents cited flood attack as the biggest threat in 2004, and 90% of them experienced application-layer attack in 2015 [2].

Nowadays, DoS/DDoS attack detection has attracted many researchers worldwide. Attack detection techniques have been developed in order to protect network against misbehaving users. Such techniques have been continually improved in order to boost their detection capability (measured by high detection and low false positive rate). Data-centric approach such as knowledge discovery or data mining is one of popular method which has gained a lot of attention in many areas.

Data mining is a process of extracting or detecting hidden pattern knowledge from large databases using statistical, mathematical, artificial intelligence, and machine learning techniques [3], [4]. Among data mining techniques, support vector machine (SVM) has been reported as the most successful classification algorithms in the data mining area. For instance, it provides good performance on traffic flooding attack detection [5].

This paper present a comprehensive literature review related to DoS/DDoS attack detection using data mining techniques published in academic journals between 2007 and 2015. The rest of the paper is organized as follows: at first, research methodology which includes classification process and classification framework is discussed in Section II; Section III describes classification of DoS/DDoS attack based on the framework. Finally, we draw a conclusion, limitation, and future work in Section IV.

2. Research Methodology

In this section, research methodology such as classification process and classification framework is described. An illustrative diagram of this section is depicted in Fig. 1

2.1 Classification Process

The literature were searched based on the keyword data mining and DoS or DDoS attack detection which produced approximately 170 papers. Two online digital library databases such as IEEE Xplore and Science Direct were chosen to provide a comprehensive bibliography of academic literature on data mining and DoS/DDoS attack detection. As such research papers published in journals represent the highest level of research [6], master and doctoral dissertations, conference papers, textbooks, reports, and unpublished working papers were excluded. All papers with published and in-press status were considered. Of 170 papers, only 35 papers were selected for classification.

2.2 Classification Framework

Selected papers were categorized by the function of data mining used. Data mining model for detecting DoS/DDoS attack generally include association, classification, clustering, and hybrid technique. Association or correlation analysis enables the identification of frequently occurring features of network pattern. By characterizing network pattern, attack detection can be performed. Common tools for association analysis are frequent pattern growth and fuzzy association rules.

Classification is one of common supervised learning models in data mining. Classification aims at building classification model to predict future network behavior through classifying dataset into a predefined class. C4.5, Naive Bayes are common tools for classification analysis.

Clustering is an unsupervised learning aims to develop unknown clusters. Unlike classification, it has no predefined clusters. Common tools for cluster analysis include neural network (NN) and SVM. Hybrid techniques employ more than one algorithms in order to increase its performance rather than single classifier. To date, combining more than one algorithms is challenging task.

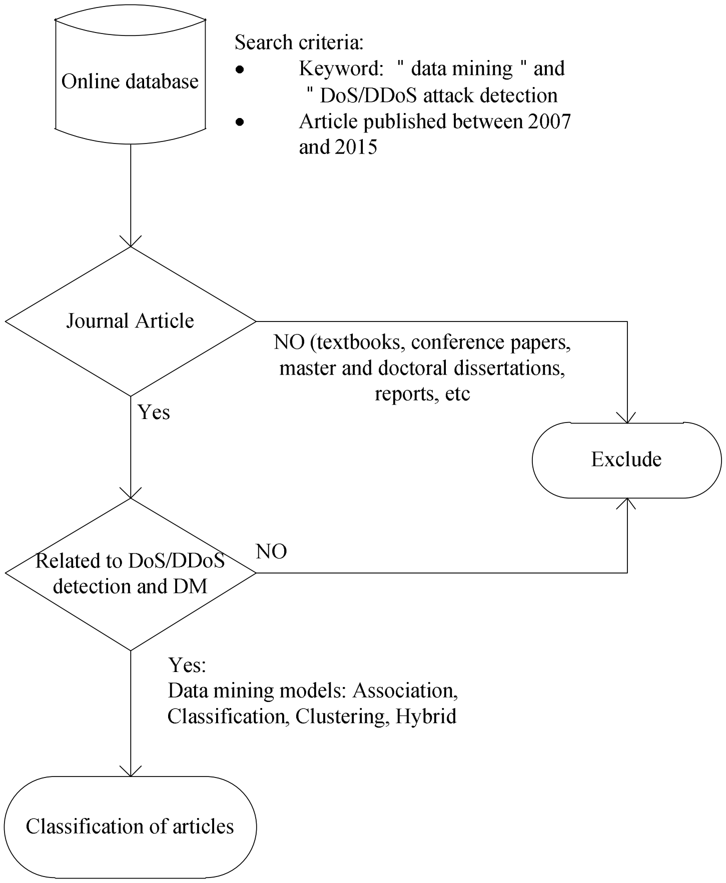
Online database
Search criteria:
• Keyword: " data mining " and " DoS/DDoS attack detection
• Article published between 2007 and 2015
Journal Article
NO (textbooks, conference papers, master and doctoral dissertations, reports, etc
Yes
Exclude
Related to DoS/DDoS detection and DM
NO
Yes:
Data mining models: Association, Classification, Clustering, Hybrid
Classification of articles

Fig. 1. Classification Process and Classification Framework

3. Classification of DoS/DDoS Attack

In this section, a detailed distribution of 35 papers are categorized.

3.1 Papers Distribution by Data Mining Techniques

Table I shows detail classification of the papers based on data mining techniques. It is noted that each paper may have employed more than one data mining techniques. Among 35 data mining techniques which have been applied in DoS/DDoS attack detection, SVM is the most commonly used technique. It has been mentioned in 6 (17.14%) papers out of 35 papers.

Following is fuzzy systems which has been described in 5 (14.29%) papers out of 35 papers.

Table 1. Classification by Data Mining Techniques

Data Mining Functions	Data Mining Techniques	Reference
Association	Fuzzy association rules	[7], [8]
	Frequent structure mining	[9]
	Multivariate correlation analysis	[10]
	Sequence analysis	[11]
	Apriori	[12], [13]
Classification	SVM	[14], [15], [16]
	Class construction	[12]
	Classification tree	[15]
	Multiagent pattern recognition	[17]
	Entropy-based	[18], [19]
	Ensemble neural	[20]
	Case-based reasoning	[21]
	Genetic algorithm	[22]
	Decision tree (DT)	[16], [13]
	Naïve Bayes	[16]
	Bayesian Network	[16]
	k-Nearest Neighbor	[16]
	Ripper	[16]
	Neural network (NN)	[16]
	Fuzzy estimators	[23]
	Particle swarm optimization (PSO)	[24]
	Extreme learning machine (ELM)	[25]
Clustering	Hierarchical clustering	[26]
	Outlier detection	[27]
	k-Means	[28]
Hybrid	DT + SVM	[29]
	Wavelet + SVD	[30]
	Fuzzy Association rule + genetic optimization	[31]
	Hierarchical clustering + SVM	[32]
	Genetic algorithm + k-NN	[33], [34]
	SOM + k-Means	[35]
	Clustering + Ant-Colony + SVM	[36]
	ensemble of adaptive + hybrid neuro-fuzzy	[37]
	Hybrid PSO + DT	[38]
	RBF + PSO	[39]
	genetic fuzzy systems + pairwise learning	[40]

3.2 Papers Distribution by Year

Fig. 2 shows the distribution of papers by year of publication. From the figure, it can be shown that the number of publications related to DoS/DDoS attack detection using data mining.

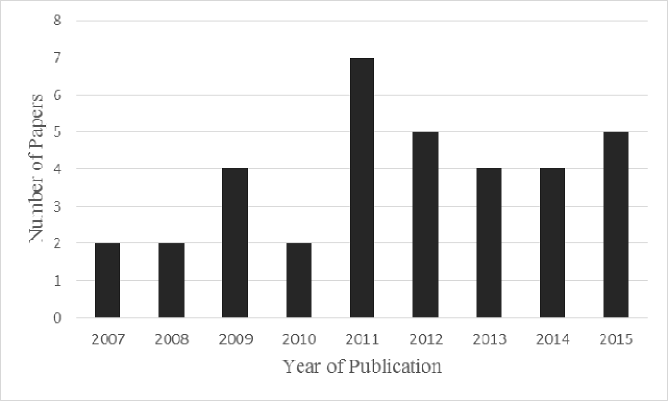

Fig. 2. Papers Distribution by Year

3.3 Papers Distribution by Journal’s Title

Table II presents distribution of papers by journals title in which the papers were published. It is obvious that ”Expert System with Applications” which focuses on information relating to expert and intelligent systems applied in industry, government, and universities worldwide, contains 10 (28.6%) papers out of the total papers published. Following are the journal namely, ”Computer Communications” and ”Computers & Security” which share the same amount.

Table 2. Paper Distribution by Journal’s Name

Journal Title	Amount
Expert Systems with Applications	10
Computer Communications	5
Computers & Security	5
Journal of Network and Computer Applications	2
IET Information Security	1
Computer Networks	1
IEEE Communications Letters	1
Applied Soft Computing	1
Computers and Mathematics with Applications	1
Computers and Electrical Engineering	1
Computer Applications	1
Information Sciences	1
Journal of Communications And Networks	1

Egyptian Informatics Journal	1
Neurocomputing	1
IEEE Transaction on Parallel and Distributed Systems	1
Journal of Systems Architecture	1

4. Conclusion

We presents a comprehensive literature review concerning DoS/DDoS attack detection using data mining technique. It is not surprising that SVM plays dominant role in many researches because of its accuracy and robustness. This work has limitation. This study only contains papers were published from 2007-2015, and only two popular online databases were used. We focus to gather more articles from other online databases.

Acknowledgement

This research was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (MSIP) (No. NRF-2014R1A2A1A11052981)

References

[1] M. H. Bhuyan, H. J. Kashyap, D. K. Bhattacharyya, and J. Kalita, "Detecting distributed denial of service attacks: Methods, tools, and future directions," The Computer Journal, 2013.

[2] A. Network, "Worldwide infrastructure security report," April 2015, http://www.arbornetworks.com/images/documents/infographics/AttackTimeline Final.pdf.

[3] J. Han and M. Kamber, Data Mining: Concepts and Techniques 2nd Edition. San Fransisco: Morgan Kaufmann, 2006.

[4] E. Turban, J. E. Aronson, T. P. Liang, and R. Sharda, Decision support and Business Intelligence Systems (Eighth ed.). Pearson Education, 2007.

[5] J. Yu, H. Lee, M.-S. Kim, and D. Park, "Traffic flooding attack detection with SNMP MIB using SVM," Computer Communications, vol. 31, no. 17, p. 42124219, 2008.

[6] J. H. Nord and G. Nordb, "MIS research: Journal status assessment and analysis," Information & Management, vol. 29, no. 1, p. 2942, 1995.

[7] M.-Y. Su, G.-J. Yu, and C.-Y. Lin, "A real-time network intrusion detection system for large-scale attacks based on an incremental mining approach," Computers & security, vol. 28, no. 5, pp. 301–309, 2009.

[8] G.-Y. Chan, C.-S. Lee, and S.-H. Heng, "Discovering fuzzy association rule patterns and increasing sensitivity analysis of XML-related attacks," Journal of Network and

Computer Applications, vol. 36, no. 2, pp. 829–842, 2013.

[9] R. Sadoddin and A. A. Ghorbani, “An incremental frequent structure mining framework for real-time alert correlation,” computers & security, vol. 28, no. 3, pp. 153–173, 2009.

[10] Z. Tan, A. Jamdagni, X. He, P. Nanda, and R. P. Liu, “A system for denial-of-service attack detection based on multivariate correlation analysis,” Parallel and Distributed Systems, IEEE Transactions on, vol. 25, no. 2, pp. 447–456, 2014.

[11] A. A. Ramaki, M. Amini, and R. E. Atani, “RTECA: Real time episode correlation algorithm for multi-step attack scenarios detection,” Computers & Security, 2014.

[12] N.-Y. Jan, S.-C. Lin, S.-S. Tseng, and N. P. Lin, “A decision support system for constructing an alert classification model,” Expert Systems with Applications, vol. 36, no. 8, pp. 11 145–11 155, 2009.

[13] J. Yu, H. Kang, D. Park, H.-C. Bang, and D. W. Kang, “An in-depth analysis on traffic flooding attacks detection and system using data mining techniques,” Journal of Systems Architecture, vol. 59, no. 10, pp. 1005–1012, 2013.

[14] J. Yu, H. Lee, M.-S. Kim, and D. Park, “Traffic flooding attack detection with SNMP MIB using SVM,” Computer Communications, vol. 31, no. 17, pp. 4212–4219, 2008.

[15] S.-Y. Wu and E. Yen, “Data mining-based intrusion detectors,” Expert Systems with Applications, vol. 36, no. 3, pp. 5605–5612, 2009.

[16] D. Stevanovic, A. An, and N. Vlajic, “Feature evaluation for web crawler detection with data mining techniques,” Expert Systems with Applications, vol. 39, no. 10, pp. 8707–8717, 2012.

[17] Z. A. Baig and K. Salah, “Multi-agent pattern recognition mechanism for detecting distributed denial of service attacks,” IET information security, vol. 4, no. 4, pp. 333–343, 2010.

[18] X. Ma and Y. Chen, “DDoS detection method based on chaos analysis of network traffic entropy,” Communications Letters, IEEE, vol. 18, no. 1, pp. 114–117, 2014.

[19] B. Tellenbach, M. Burkhart, D. Schatzmann, D. Gugelmann, and D. Sornette, “Accurate network anomaly classification with generalized entropy metrics,” Computer Networks, vol. 55, no. 15, pp. 3485–3502, 2011.

[20] P. A. R. Kumar and S. Selvakumar, “Distributed denial of service attack detection using an ensemble of neural classifier,” Computer Communications, vol. 34, no. 11, pp. 1328–1341, 2011.

[21] C. I. Pinz´on, J. F. De Paz, M. Navarro, J. Bajo, V. Juli´an, and J. M. Corchado, “Real-time CBR-agent with a mixture of experts in the reuse stage to classify and detect DoS attacks,” Applied Soft Computing, vol. 11, no. 7, pp. 4384–4398, 2011.

[22] S. M. Lee, D. S. Kim, J. H. Lee, and J. S. Park, “Detection of DDoS attacks using

optimized traffic matrix," Computers & Mathematics with Applications, vol. 63, no. 2, pp. 501–510, 2012.

[23] S. N. Shiaeles, V. Katos, A. S. Karakos, and B. K. Papadopoulos, "Real time DDoS detection using fuzzy estimators," computers & security, vol. 31, no. 6, pp. 782–790, 2012.

[24] S. Jamali and V. Shaker, "Defense against SYN flooding attacks: A particle swarm optimization approach," Computers & Electrical Engineering, vol. 40, no. 6, pp. 2013–2025, 2014.

[25] J. M. Fossaceca, T. A. Mazzuchi, and S. Sarkani, "MARK-ELM: Application of a novel multiple kernel learning framework for improving the robustness of network intrusion detection," Expert Systems with Applications, 2014.

[26] K. Lee, J. Kim, K. H. Kwon, Y. Han, and S. Kim, "DDoS attack detection method using cluster analysis," Expert Systems with Applications, vol. 34, no. 3, pp. 1659–1665, 2008.

[27] P. Casas, J. Mazel, and P. Owezarski, "Unsupervised network intrusion detection systems: Detecting the unknown without knowledge," Computer Communications, vol. 35, no. 7, pp. 772–783, 2012.

[28] W. Cerroni, G. Moro, R. Pasolini, and M. Ramilli, "Decentralized detection of network attacks through P2P data clustering of SNMP data," Computers & Security, vol. 52, pp. 1–16, 2015.

[29] S. Peddabachigari, A. Abraham, C. Grosan, and J. Thomas, "Modeling intrusion detection system using hybrid intelligent systems," Journal of network and computer applications, vol. 30, no. 1, pp. 114–132, 2007.

[30] C. S. Sastry, S. Rawat, A. K. Pujari, and V. P. Gulati, "Network traffic analysis using singular value decomposition and multiscale transforms," Information Sciences, vol. 177, no. 23, pp. 5275–5291, 2007.

[31] M.-Y. Su and S.-C. Yeh, "An online response system for anomaly traffic by incremental mining with genetic optimization," Communications and Networks, Journal of, vol. 12, no. 4, pp. 375–381, 2010.

[32] S.-J. Horng, M.-Y. Su, Y.-H. Chen, T.-W. Kao, R.-J. Chen, J.-L. Lai, and C. D. Perkasa, "A novel intrusion detection system based on hierarchical clustering and support vector machines," Expert systems with Applications, vol. 38, no. 1, pp. 306–313, 2011.

[33] M.-Y. Su, "Using clustering to improve the kNN-based classifiers for online anomaly network traffic identification," Journal of Network and Computer Applications, vol. 34, no. 2, pp. 722–730, 2011.

[34] ——, "Real-time anomaly detection systems for Denial-of-Service attacks by weighted k-nearest-neighbor classifiers," Expert Systems with Applications, vol. 38, no. 4, pp.

3492–3498, 2011.

[35] S. Lee, G. Kim, and S. Kim, “Self-adaptive and dynamic clustering for online anomaly detection,” Expert Systems with Applications, vol. 38, no. 12, pp. 14 891–14 898, 2011.

[36] Y. Li, J. Xia, S. Zhang, J. Yan, X. Ai, and K. Dai, “An efficient intrusion detection system based on support vector machines and gradually feature removal method,” Expert Systems with Applications, vol. 39, no. 1, pp. 424–430, 2012.

[37] P. A. R. Kumar and S. Selvakumar, “Detection of distributed denial of service attacks using an ensemble of adaptive and hybrid neuro-fuzzy systems,” Computer Communications, vol. 36, no. 3, pp. 303–319, 2013.

[38] G. Nadiammai and M. Hemalatha, “Effective approach toward intrusion detection system using data mining techniques,” Egyptian Informatics Journal, vol. 15, no. 1, pp. 37–50, 2014.

[39] A. Karami and M. Guerrero-Zapata, “A hybrid multiobjective RBFPSO method for mitigating DoS attacks in named data networking,” Neurocomputing, vol. 151, pp. 1262–1282, 2015.

[40] S. Elhag, A. Fern´andez, A. Bawakid, S. Alshomrani, and F. Herrera, “On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on intrusion detection systems,” Expert Systems with Applications, vol. 42, no. 1, pp. 193–202, 2015.

View publication stats