



Ain Shams University
Faculty of Science
Department of Mathematics

On Detecting Network Intrusion Using Computational Intelligence

A thesis

Submitted to the Department of Mathematics, Faculty
of Science, Ain Shams University, Cairo, Egypt.

For the Degree of M.Sc. In Science as a Partial Fulfillment For
Requirements of the Master Science
(Computer Science)

By

Ahmed Abd El-Rahman Mahmoud Elngar

Demonstrator, Faculty of Information Technology & Computer Science, Sinai University,
El-Arish, Egypt.

Supervisors

Prof. Fayed Fayek Mohamed Ghaleb

Prof. Emeritus of Mathematics

Faculty of Science

Ain Shams University

Dr. Dowlat Abd El Aziz Mohamed

Lecturer of Mathematics and Computer Science

Faculty of Science

Ain Shams University

2013

List of Publications

Journal Papers:

1. **Ahmed A. Elngar**, Dowlat A. El A. Mohamed , Fayed F. M. Ghaleb, "A Fast Accurate Network Intrusion Detection System", International Journal of Computer Science and Information Security, vol. 10, no. 9, pp. 29 - 35, 2012.
2. **Ahmed A. Elngar**, Dowlat A. El A. Mohamed , Fayed F. M. Ghaleb, "A Real-Time Anomaly Network Intrusion Detection System with High Accuracy", Information Sciences Letters International Journal, vol. 02, no. 2, pp. 49-56, 2013.

Peer Reviewed International Conference:

1. **Ahmed A. Elngar**, Dowlat A. El A. Mohamed , Fayed F. M. Ghaleb, "A Real-Time Network Intrusion Detection System with High Accuracy", The 6th International Conference on Mathematics, Trends and Development (ICMTD12) , 27 - 29 December, Cairo, Egypt, 2012.

Abstract

According to rapid development and popularity of Internet and online procedures, the potential of network attacks has increased substantially in recent years. Therefore, network security needs to be concerned to provide secure information channels. Intrusion Detection System (IDS) becomes an essential component of computer security. Network Intrusion Detection Systems (NIDS) aims to dynamically identify unusual access or attacks to secure the internal networks, by looking for potential malicious activities in network traffic. However, building a high-performance and fast NIDS is a major research problem in network security.

One of the important problems for NIDS is dealing with data containing high number of features. High dimensional data may leads to decrease the predictive accuracy and the speed of the NIDS. Therefore, Feature Selection (FS) is one of the key topics in building NIDS. (FS) can serve as a pre-processing tool for high dimensional data before solving the classification problems. The purpose of the feature selection is to reduce the number of irrelevant and redundant features. (FS) searches for a subset of features which improve the prediction accuracy and improves the NIDS speed.

This thesis is devoted to focuss on how to construct a fast accurate NIDS. The thesis propose two different hybrid NIDS, the proposed hybrid NIDS models involves data preprocessing, data reduction and intrusion classifica-

tion. Experiments and Analysis of the proposed hybrid NIDSs with other previous NIDSs demonstrated that; the two proposed hybrid NIDSs enhance the intrusion detection rate and decreasing the testing speed.

To

My wife Dr. Heba Eid, you supported me each step of my way.

**You have always been the first to back and strengthen me
whenever I felt exhausted and weak.**

**My baby girl Farida, the smile on your face brings a smile to my
heart.**

Acknowledgement

*Firstly, I would like to thank **ALLAH** for all his grace, mercy and strength that has sustained me throughout this time of my life.*

*I would like to express my deep thanks and gratitude to **Prof. Fayed M. Ghaleb** for his continuous support, valuable advices and abundant experience throughout this work. I have learnt many things about research working from him.*

*I'm indebted to **Dr.Dowlat Abd El Aziz** for her encouragement and helpful co-operation throughout this work. She has provided a lot of motivation.*

And I do not forget to thanks all my teachers, Department of Mathematics Faculty of Science Al-Azhar University ,Ain Shams University and Sinai University on the continued support and encouragement.

*Finally, I would also like to express special thanks to **my father, my mother, my brother, and my sisters** for their love and support during of this work.*

Contents

List of Publications	iii
Abstract	iv
Acknowledgement	vii
1 Introduction	1
1.1 Network security	2
1.1.1 Overview	2
1.2 Taxonomy of Cyber Attacks	3
1.2.1 Denial of Service (DoS) Attacks	3
1.2.2 Remote to Local (R2L) Attacks	4
1.2.3 User to Root (U2R) Attacks	4
1.2.4 Probing Attacks	4
1.2.5 Examples of Cyber Attacks	5
1.2.6 Denial of Service (DoS) Attack: The TCP-SYN Attack	8
1.3 Intrusion Detection System	9

1.3.1	Taxonomy of Intrusion Detection System (IDSs)	10
1.3.1.1	Depending on Location Techniques	11
1.3.1.2	Depending on Detection Techniques	14
16	Thesis Motivation	1.4
17	Thesis Contributions	1.5
19	Thesis Organization	1.6
2	Feature Selection in Computational Intelligence	21
2.1	Introduction	21
2.2	Computational Intelligence	21
2.3	Feature Selection (FS)	23
2.3.1	Introduction	23
2.4	A Feature Selection Approaches (FSAs)	25
2.4.1	Filter Approach	25
2.4.2	Wrapper Approach	26
2.5	A General Algorithm of Feature Selection	26
2.5.1	Subset Generation	28
2.5.2	Subset Evaluation	28
2.5.3	Stopping Criteria	28
2.5.4	Result Validation	28
2.6	Sequential Selection Algorithms	30
2.6.1	General Sequential Forward Selection (GSFS):	30

32	2.6.2 Sequential Forward Selection (SFS): .	
34	2.6.3 Sequential Backward Selection (SBS)	
34	Randomized Selection Algorithms(RSA) . . .	2.7
34	2.7.1 Information Gain (IG)	
34	2.7.1.1 Introduction	
37	2.7.2 Genetic Algorithm (GA)	
37	2.7.2.1 Introduction	
40	2.7.2.2 Chromosome Structure	
41	2.7.2.3 Genetic Algorithm Operation	
	2.7.2.4 Genetic algorithm-based Feature Selection Using	
51	Correlation	
52	Feature Fitness Function	2.7.2.5
53	Chromosome Fitness Function	2.7.2.6
57	Swarm Optimization(PSO)	2.7.3
57	Introduction	2.7.3.1
	2.7.3.2 Particle Swarm Optimization Operation	59
	2.7.3.3 Particle Swarm Optimization-based Feature Sub-	
63	set Selection	
64	Fitness Function	2.7.3.4
69	Simple Function Example Using PSO	2.7.3.5

3 Classification in Computational Intelligence Techniques 70

70	Introduction	3.1
70	Classification Overview	3.2
73	Classification Procedure	3.3
77	Classification Methods .	3.4
78	3.4.1 Decision Tree Method .	
85	3.4.1.1 ID3 Classifier .	
90	3.4.1.2 C4.5 Classifier .	
94	3.4.2 Bayesian Method . . .	
	3.4.2.1 Naïve Bayes Classifier	95
	3.4.2.2 Hidden Naïve Bayes Classifier	97

4 The Proposed Hybrid Network Intrusion Detection Models102

4.1 Introduction	102
4.2 Network intrusion Dataset	103
4.3 Performance Evaluation	108
4.4 Network Intrusion Detection Related Work	110
4.5 The Proposed Hybrid PSO-DT ID Model	113
4.5.1 Preprocessing Phase	115
4.5.2 PSO Feature Selection Phase	115
4.5.3 C4.5 Intrusion Detection Phase	116
4.5.4 Implementation Results and Analysis	118
4.5.4.1 Experiments and analysis	118

4.6 The Proposed Hybrid PSO-Discritize-HNB ID Model	120
121 IEM Discritization Phase	4.6.1
122 HNB Intrusion Detection Phase . . .	4.6.2
124 Implementation Results and Analysis	4.6.3
124 4.6.3.1 Experiments and analysis . .	
5 Conclusion and Future Directions	130
5.1 Conclusions	130
5.2 Future Directions	132

List of Tables

40	Features values differences	2.1
105 . . .	KDD'99 dataset Features and their data types	4.1
107 . . .	KDD'99 dataset reduction statistics	4.2
109 . . .	Confusion Matrix	4.3
118 . . .	C4.5 DT detection measurements (41-dimension feature)	4.4
119 . . .	PSO-DT detection measurements (11-dimension feature)	4.5
119 . . .	GA-DT detection measurements (12-dimension feature)	4.6
4.7	Testing accuracy, Features Number and Timing comparison .	120
4.8	HNB accuracy measurements (41-dimension feature)	124
4.9	The Proposed-PSO-Discritize-HNB accuracy measurements (11- dimension feature)	125
4.10	Classical PSO-Discritize-HNB accuracy measurements (19-dimension feature)	126
4.11	Testing accuracy, Features Number and Timing comparison .	127

4.12 IG-Discritize-HNB accuracy measurements (24-dimension feature)	128
4.13 Testing accuracy, Features Number and Timing comparison .	128

List of Figures

- 5 Remote to Local attacks. 1.1
- 6 User to Root attacks. . . 1.2
- 6 Probe attacks. 1.3
- 7 Denial of Service attacks . 1.4
- 8 Handshake of TCP-SYN. 1.5
- 9 SYN Flood Attacker. . . . 1.6
- 1.7 Intrusion Detection System Taxonomy 10
- 1.8 Host-Based Intrusion Detection Systems (HIDS) 12
- 1.9 Network-Based Intrusion Detection Systems (NIDS) 13
- 15 1.10 A Misuse-based IDS.
- 16 1.11 An anomaly-based IDS.
- 27 A key Steps Of Feature Selection. . 2.1
- 39 Genetic Algorithm flow chart. . . . 2.2
- 41 chromosomes representation. . . . 2.3