



NAIF ARAB UNIVERSITY FOR SECURITY SCIENCE
COLLEGE OF COMPUTER INFORMATION SECURITY
DEPARTMENT

USING ENSEMBLE MACHINE LEARNING FOR INTRUSION DETECTION USING
MULTIPLE MODELS WITH SAME ADAPTIVE AI CLASSIFICATION
ALGORITHMS ON WEKA

BY:

MUTEB AL-YOUSEF

A THESIS

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF:

MASTER OF SCIENCE IN INFORMATION SECURITY

TO:

INFORMATION SECURITY DEPARTMENT

THE COLLEGE OF COMPUTER AND INFORMATION SECURITY

NAIF ARAB UNIVERSITY FOR SECURITY SCIENCES

MAR 2018

SUPERVISED BY:

DR. NABIH T.ARAR

NAIF ARAB UNIVERSITY FOR SECURITY SCIENCES

Table of Contents

Dedication	iv
Acknowledgment	v
List of Figures	V
List of Tables.....	VII
List of Abbreviations.....	VIII
1 Introduction	1
1.1 Introduction	1
1.2 Study Problem	3
1.3 Study Hypotheses	4
1.4 Aims and Objectives	4
1.5 Importance of the Study	5
1.6 Motivations.....	6
1.6.1 First Motivation.....	7
1.6.2 Second Motivation	7
1.6.3 Third Motivation	7
1.7 Research Scope	7
1.7.1 Intrusions Signature.....	7
1.7.2 Signature Database Update	8
1.7.3 Detecting Unknown intrusions Dynamically	8
1.8 Contributions	8
1.8.1 First Contribution	9
1.8.1 Second Contribution.....	9
1.8.1 Third Contribution.....	9
1.8.1 Fourth Contribution.....	9
1.9 Publications	10
1.10 Conclusion.....	10
2 Background	12
2.1 Introduction	12
2.2 Intrusions, Intrusion Detection and Intrusion Detection System	13

2.3	IDS Overview	14
2.4	IDS History	15
2.5	Architecture of IDS	17
2.6	Characteristics of IDS	18
2.7	Importance of the IDS	20
2.8	IDS Types.....	20
2.9	NIDS vs HIDS.....	21
2.9.1	Network-Based Intrusion Detection System	21
2.9.1.1	Advantages of Network based Intrusion Detection Systems	21
2.9.1.2	Disadvantages of Network based Intrusion Detection Systems.....	22
2.9.2	Host-Based Intrusion Detection System	23
2.9.2.1	Advantages of Host based Intrusion Detection Systems.....	23
2.9.2.2	Disadvantages of Host based Intrusion Detection Systems	24
2.10	Intrusion Detection Categories.....	25
2.10.1	Signature-based or Misuse Detection.....	25
2.10.1.1	Advantages of Signature-based Detection	25
2.10.1.2	Disadvantages of Signature-based Detection.....	26
2.10.2	Behaviour-based or Anomaly Detection	26
2.10.2.1	Advantages of Behaviour-based Detection	27
2.10.2.2	Disadvantages of Behaviour-based Detection.....	27
2.10.3	Hybrid-Based Detection.....	27
2.11	Challenges of Intrusion Detection systems	28
2.12	IDS vs Firewall.....	29
2.13	IDS vs. IPS	31
2.14	Protecting the IDS itself	31
2.15	IDS's Packet, Signature and Alerts	32
2.15.1	Packets.....	32
2.15.2	Signatures	33
2.15.3	Alerts	33
2.16	Main types of Security Threats and Attacks	33

2.17	Open Source IDS tool	35
2.17.1	Components of Snort.....	35
3	Literature Review.....	37
3.1	Related Work.....	37
3.2	Conclusion.....	52
4	Research Methodology	54
4.1	Introduction	54
4.2	Methodology	54
4.3	Study Population	55
4.4	Study Samples	55
4.5	Study Tools	56
4.5.1	IDS Tool.....	57
4.5.2	Attacking Tool.....	57
4.6	Samples Source	57
4.7	Conclusion.....	58
5	Proposed Model (Dynamic Detecting and Updating the database of Signature-Based IDS with new attacks)	59
5.1	Introduction	59
5.2	Challenges	60
5.3	Proposed Solution (Detecting and Updating the database of Signature-Based IDS with new attacks automatically).....	61
5.3.1	IDS Engine Stages.....	63
5.3.2	Filtering Engine.....	64
5.3.3	Updating Engine.....	67
5.4	Conclusion.....	68
6	Results and Discussion.....	69
6.1	Introduction	69
6.2	Dataset.....	69
6.3	Installed Programes	70
6.4	Implementation and Testing.....	70

6.4.1	Implementation of the Proposed Model	71
6.4.2	Testing.....	74
6.4.2.1	First Test.....	74
6.4.2.2	Second Test	75
6.5	Proposed Model vs Other Related Models.....	76
6.6	Result Discussions	77
6.6.1	Result of First Test	77
6.6.2	Result of Second Test.....	77
6.7	Conclusion.....	78
7.	Conclusion and Future Works.....	79
7.1	Introduction	79
7.2	Contributions.....	80
7.3	Recommendations and Future Work.....	81
7.4	Conclusion.....	81

List of Figures

Figure 2.1. General Structure of IDS	17
Figure 2.2. Characteristics of IDS based on 6 criteria	18
Figure 2.3. Categories of IDS.....	25
Figure 2.4. Places where the firewall and IDS can be within the network	30
Figure 2.5. Components of Snort	36
Figure 3.1. Automatic update of signatures on secondary IDS.....	38
Figure 3.2. Automatic update of signatures on primary IDS	38
Figure 3.3 Algorithm to generate and update signature databases.....	39
Figure 3.4. Multiple site's IDS architecture.....	41
Figure 3.5 Algorithm of updating signature database	43
Figure 3.6 Proposed Model of the IDS.....	44
Figure 3.7 Structure of the small database's mobile agent	45
Figure 3.8 Flowchart for CA-NIDS	47
Figure 3.9 Algorithm 2 CA-NIDS.....	48
Figure 3.10 Structure of proposed model	49
Figure 3.11 The order of updating process	50
Figure 3.12 algorithm for updating the small frequency databases from the complementary database.	50
Figure 4.1 Interface of the attacking tool	57
Figure 5.1 General Structure of the proposed model.....	62
Figure 5.2 The flow work of the similarity factor.....	65
Figure 5.3 The priorities among the Similarity and Black List of IPs factors	66
Figure 5.4 Updating process of both CDB and FSDB automatically	67
Figure 6.1 The website where the dataset were collected	71
Figure 6.2 Part of RuleMaker's code for reading and reformatting the collected rules	72
Figure 6.3 The formate of the collected rules according to SNORT's rules formate	72
Figure 6.4 Part of the RuleClassifier's code for distributing the rules into smaller databases	73

Figure 6.5 Part of the RuleClassifier's code responsible for distributing the rules based on protocol
type 73
Figure 6.6 TCP Listener 74

List of Tables

Table 1.1: Some statistics about intrusions such are security breaches, email threats, malware,web attacks and ransomware over 2015-2016	2
Table 2.1: Pros and cons of intrusion detection methodologies.....	27
Table 2.2: Network Packet Structure	32
Table 3.1: Testing the performance of IDS using both small and big signatures databases	39
Table 3.2: Performance rate of the proposed IDS model	48
Table 3.3: Pre-defined priorities by network administrator	51
Table 6.1: Attacker system specifications.....	75
Table 6.2: IDS system specifications	75
Table 6.3: Test outputs of the proposed model with one large database	75
Table 6.4: Attacker system specifications.....	77
Table 6.5: IDS system specifications	76
Table 6.6: The way where the signatures were distributed based on protocol type.....	76
Table 6.7: Test outputs of the proposed model with multiple smaller databases.....	76
Table 6.8 Comparing between the outputs of both proposed model and [27] 's model in terms of accuracy.....	77
Table 6.9 Comparing between the outputs of both proposed model and [27] 's model in terms of performance.....	77
Table 6.10 Comparison between both proposed model's tests in terms of accuracy and performance.....	78

List of Acronyms

Abbreviations	Explanations
BIDS	Behavior-Based Intrusion Detection System.
CDB	Complementary Database.
DDOS	Distributed Denial of Service.
DOS	Denial of Service.
FP	False Positive.
FSDB	Frequent Signatures Database.
HIDS	Host-Based Intrusion Detection System.
ICMP	Internet Control Message Protocol.
IDS	Intrusion Detection System.
IPS	Intrusion Prevention System.
NIDS	Network-Based Intrusion Detection System.
SIDS	Signature-Based Intrusion Detection System.
TCP	Transmission Control Protocol.
UDP	User Datagram Protocol.