

Contents

1. INTRUCTION

1.1 Network anomaly detection and its significance.....	2
1.2 Usefulness of dataset preparation.....	3
1.3 Motivation.....	3
1.4 Proposed approach and its significance	3
1.5 Problem definition.....	3
1.6 Our contributions.....	4
1.7 Organization of dissertation.....	5

2. ANOMALY DETECTION AN OVERVIEW

2.1 What are anomalies.....	6
2.2 Different aspects of anomaly detections.....	7
2.2.1 Nature of input data.....	7
2.2.2 Types of anomalies.....	7
2.2.3 Data labels.....	8
2.2.4 Output of anomaly detection.....	9
2.3 Discussion.....	9

3. RELATED WORKS

3.1 Statistical approach.....	10
3.2 Machine learning.....	11
3.2.1. Bayesian networks.....	11
3.2.2. Markov models.....	11
3.2.3. Fuzzy logic techniques.....	11
3.3 Neural networks.....	12
3.4 Data mining based.....	12
3.4.1 Clustering and outlier detection.....	12
3.5 knowledge based.....	12

4. DATASETS PREPARATION: CAPTURING, PREPROCESSING, FEATURE EXTRACTION AND CORRELATION

4.1 Testbed setup.....	13
4.2 Launching of real life attacks.....	13
4.2.1 Port scanning	13
4.2.2 Nmap : network mapper.....	16
4.2.3 DoS attacks generation.....	19
4.2.4 Other attacks.....	22
4.3 Packet level network traffic capturing and dataset preparation.....	23

4.3.1	Packet level network traffic capturing.....	23
4.3.2	Preprocessing.....	25
4.3.3	Feature extraction.....	26
4.3.4	Correlation of various features.....	30
4.4	Flow level network traffic capturing and dataset preparation.....	31
4.4.1	Network flow.....	31
4.4.2	Properties of network flow.....	32
4.4.3	Network flow capturing.....	34
4.4.4	Tool used for capturing the NetFlow traffic.....	34
4.4.5	Netflow processing.....	36
4.4.6	Aggregating flows.....	38
4.4.7	Filter syntax.....	38
4.4.8	Top N statistics.....	40
4.4.9	Prerequisites for Nfsen.....	41
4.4.10	Working with Nfsen.....	41
4.5	Preprocessing, feature extraction and correlation.....	42
4.5.1	Preprocessing.....	42
4.5.2	Feature extraction.....	43
4.5.3	Correlation of various features.....	45
4.6	Discussion.....	46
5.	PROPOSED APPROACH	
5.1	Proposed algorithm.....	48
5.2	Hybridization of packet based and flow based.....	49
5.3	Complexity analysis.....	51
5.4	Dataset used.....	51
5.5	Results and performance evaluation.....	51
5.6	Discussion.....	52
6.	CONCLUSION AND FUTURE WORK	
6.1	Conclusion.....	53
6.2	Future Work	53
8	BIBLIOGRAPHY	54

List of figures:

Sl no.	Table name	Page No.
1	Similarity graph based approach	4
2	Example of anomalies in a 2-dimensional dataset	6
3	Architecture of the testbed setup	14
4	One to one attack	16
5	One to many attack	16
6	Many to one	16
7	Many to many	16
8	Syn attack launched by using nmap tool	17
9	TCP window Scan on a closed port	18
10	Tcp window Scan on an open port	18
11	XMAS Scan on a closed port	18
12	XMAS Scan on a open port	18
13	SYN Scan on a closed port	19
14	SYN Scan on an open port	19
15	Gulp captured data in expanded format	25
16	Extracted basic features (using c routine)	27
17	Tcptrace detailed summery	28
18	Packet level dataset	31
19	Netflow parameters	33
20	Operation of Nfdump	35
21	Netflow processing	36
22	Nfdump Extended format	37
23	Nfdump Filtering Outputs	40
24	Nfsen graphical view of UDP traffic	42
25	Filtered Flow data	43
26	Basic features of flow level data	44
27	The flow level dataset	46
28	Proposed framework of our Hybrid IDS	50

List of tables:

Table No.	Table name	Page No.
1	Different types of attack generated by targa	20
2	Different options of Gulp	24
3	Different capture file options of Gulp	25
4	Descriptions of basic feature	26
5	Different content based features	29
6	Different time-based features	29
7	Different window-based features	30
8	Unique keys of Netflow	33
9	Different versions of Netflow	33
10	Different options of Nfdump tool to read nfcapd files	36
11	Description of custom options	38
12	Descriptions of top N statistics	40
13	List of parameters we considered	42
14	Llist of basic features	44
15	Descriptions of time-window based features	45
16	Description of connection based features	45
17	Distributions of normal and attack connections	51
18	Calculation results of the experiment	51