

Intelligent Machine Learning Architecture for Detecting DDoS attacks in IoT networks

by

Yahya Sulaiman Al-hadhrami

Thesis submitted in fulfilment of
the requirements for the degree of

Doctor of Philosophy

Under the supervision of **Dr. Farookh Khadeer Hussain**

University of Technology Sydney
Faculty of Engineering and Information Technology

September-2020

Certificate of Authorship/Originality

I, Yahya Al-Hadhrami declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philanthropy (C02029), in the computer science school/faculty of engineering and information technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Production Note:

Signature: Signature removed prior to publication.

Date: 16/09/2020

ABSTRACT

Intelligent Machine Learning Architecture for Detecting DDoS attacks in IoT networks

by

Yahya Sulaiman Al-hadhrami

The Internet of Things (IoT) is growing rapidly across a wide range of applications; one example of such an application is the smart city, in which a city's infrastructure, such as road management, building automation, and people and crowd surveillance, is connected to the Internet. Such applications are being extended to factories, smart agriculture, and even smart devices, which are becoming very common. The rapid growth in the IoT has driven other technologies, such as 5G networks, to grow rapidly to adjust to the sheer number of devices connected to the Internet, and these technologies are expected to further expand the spread of the IoT. However, the existing IoT deployment does not come without challenges, including the large number of connected devices, security issues, and a variety of new standards. From a security perspective, IoT faces a growing threat when it comes to the availability of networks. Distributed denial of service (DDoS) attacks are one well-known threat. However, investigation of the literature shows a lack of solutions with which to tackle DDoS attacks in the IoT.

To address this gap in the literature, this thesis proposes an intelligent machine-learning-based platform that can detect denial-of-service attacks, termed IDD-IoT. The proposed platform consists of several components, including building a real-time dataset generation framework to generate IoT-based datasets (IoT-DDoS) to detect malicious attacks in the IoT, allowing scientists and researchers in the field to further enhance intrusion detection systems with an up-to-date dataset. The

platform then builds on the dataset generation framework, developing an intelligent machine-learning-based framework for detecting three kinds of IoT-DDoS attacks: blackhole, selective forwarding, and flooding attacks. We utilize this framework to build a novel advanced intrusion detection system (IDS) for IoT networks capable of analyzing and detecting DDoS attacks. The IDS consists of a real-time monitoring and analysis unit capable of monitoring traffic in real time with the assistance of an IDS agent that works as a communication link between the IDS and IoT network. We show that our proposed intelligent framework can efficiently detect malicious attacks in respect to security goals such as confidentiality, privacy, and availability, by building an emulated smart IoT environment using the Cooja simulation platform, and we evaluate its performance. Finally, we present the simulation and evolution results to highlight the proposed platform's efficiency, taking into consideration the limitations associated with resource-constrained devices.

thesis directed by Associate Professor Farookh Hussain

School of Computer Science

Faculty of Engineering and Information Technology (FEIT)

Dedication

I dedicate this thesis to my family, friends and all who supported me through this journey.

Acknowledgements

I would like to humbly extend my utmost gratitude and thankfulness to the hands that helped craft my academic endeavours throughout my PhD journey. This journey has been blessed by the Almighty Allah and such blessings were bestowed upon me in the form of people guiding me to the light. One person to whom I must express my sincere appreciation is my superior and principal advisor, Associate Professor Farookh Khadeer Hussain for his endless support. His critical vision and leadership contributed greatly to my work ethic and the outlet it took. You shaped my academic character which will be of great value to me as I begin my career. My journey was also blessed by the encouragement of my family, friends and colleagues to whom I owe my deep gratitude. May Allah give you back the light you brought into my journey. Thank you.

Yahya Al-Hadhrami
Sydney, Australia, 2020.

List of Publications

The following is a list of my research papers produced during my PhD study.

Journal Papers

- J-1. **Yahya Al-hadhrami**, and Farookh Khadeer Hussain, “Real time dataset generation framework for intrusion detection systems in IoT,” *Future Generation Computer Systems*, 08:414 – 423, 2020, doi:”doi.org/10.1016/j.future.2020.02.051”.
- J-2. **Yahya Al-hadhrami**, and Farookh Khadeer Hussain, “DDoS attacks in IoT Networks: A Comprehensive Systematic Literature Review,” *World Wide Web Journal*, 2020 (Minor Revision).

Conference Papers

- C-1. **Yahya Al-hadhrami**, Nassser Al-hadhrami and Farookh Khadeer Hussain: Data Exportation Framework for IoT Simulation Based Devices. *International Conference on Network-Based Information Systems*, pp.212-222 , Springer, 2019
- C-2. **Yahya Al-hadhrami**, and Farookh Khadeer Hussain :A Machine Learning Architecture Towards Detecting Denial of Service Attack in IoT. *Conference on Complex, Intelligent, and Software Intensive Systems*, pp.417-429, Springer, 2019

Contents

Certificate	ii
Abstract	iv
Dedication	vi
Acknowledgments	vii
List of Publications	viii
List of Figures	xvi
List of Tables	xix
1 Introduction	1
1.1 Background	2
1.1.1 Security Requirements and Goals	2
1.1.2 Constraints and Limitations	4
1.1.3 IoT Architecture	6
1.1.4 RPL Protocol	7
1.1.5 IoT Security	8
1.2 SIGNIFICANCE OF THE THESIS	9
1.2.1 Scientific Significance	10
1.2.2 Social Significance	10
1.3 Structure of the Thesis	10
1.4 Conclusion	12

2 Literature Review	13
2.1 Introduction	13
2.2 Research Strategies	13
2.2.1 Keywords	15
2.2.2 Research Questions	15
2.2.3 Research Filtration Process	16
2.3 Security Issues and Attack Classifications	17
2.3.1 Perception Layer Attack	18
2.3.2 Network	20
2.4 Literature Review	22
2.4.1 Protocol-Based Solutions	23
2.4.2 Trust Based Solution	26
2.4.3 Intrusion Detection Based Solutions	29
2.4.4 Others	34
2.5 Comprehensive Discussion	36
2.5.1 Summary of the Literature Shortcoming	38
2.6 Conclusion	39
3 Problem Definition	40
3.1 Introduction	40
3.2 Terms	41
3.2.1 Internet of Things	41
3.2.2 IoT Security	41
3.2.3 DDoS Attacks	41
3.2.4 Selective Forwarding Attacks	41

3.2.5	Blackhole Attack	41
3.2.6	Flooding Attack	41
3.2.7	6LoWPAN Protocol	42
3.2.8	RPL Protocol	42
3.2.9	SVM	42
3.2.10	Neural Networks	42
3.2.11	Decision Tree	42
3.2.12	Node	42
3.2.13	Sink/Root Node	43
3.2.14	Sniffer Node	43
3.2.15	Attacker Node	43
3.2.16	6BR Router	43
3.2.17	IDS	43
3.2.18	IDS Agent	43
3.2.19	Data Extraction	43
3.2.20	Feature Selection	43
3.2.21	DODAG	44
3.2.22	DIS	44
3.2.23	DIO	44
3.2.24	Route	44
3.3	Problem Overview	44
3.4	Research Issues	45
3.5	Research Questions	46
3.6	Research Objectives	47

3.7	Research Approach to Problem-Solving	47
3.7.1	Design Science Research Methodology	47
3.7.2	Social Science Method	49
3.7.3	The Choice of the Science and Engineering-Based Research Method	50
3.8	Conclusion	51
4	Solution Overview	52
4.1	Introduction	52
4.2	Overview of IDD-IoT Framework Solution for DDoS Detection in IoT	52
4.3	Overview for the Data Collection and Dataset Creation Framework . .	54
4.3.1	Data Collection Module	54
4.4	Overview of the Machine Learning Methodology and Evaluation . . .	55
4.4.1	Artificial Neural Networks	56
4.4.2	Support Vector Machine	56
4.4.3	Decision Tree (Random Forest)	56
4.4.4	Detection Methods	56
4.4.5	Machine Learning Validation	57
4.5	Overview of the ML IDS Implementation in the IoT Network	60
4.5.1	Proposed Network Diagram:	61
4.5.2	IDS Agent Module:	61
4.5.3	IDS Evaluation Overview	62
4.6	Conclusion	62
5	A Machine Learning Architecture for Detecting Denial- of-Service Attacks in IoT	63

5.1	Introduction	63
5.2	Framework	63
5.2.1	Data Collection Module	65
5.2.2	Feature Selection	67
5.2.3	Data Classification Module	68
5.2.4	IDS Agent Module	75
5.3	Conclusion	76
6	Real-time Dataset Generation Framework for Intrusion Detection Systems in IoT	77
6.1	Introduction	77
6.2	Data Exportation Framework	78
6.2.1	Data Handler Module (DHM)	79
6.2.2	Data Visualization Module (DVM)	80
6.2.3	Database Management Module (DMM)	81
6.2.4	API Module (APIM)	81
6.2.5	DEF Evaluation	82
6.2.6	Simulation Results and Discussion	83
6.3	Real-Time Data Collection Framework	85
6.3.1	Network and Attack Scenarios	86
6.3.2	Traffic Generation	90
6.3.3	Capturing the Data	91
6.3.4	Data Aggregation	93
6.3.5	Queue	94
6.3.6	Feature Extraction Unit	95

6.3.7	Data Labelling	99
6.3.8	Quantitative Description of IoT-DDoS	99
6.4	Analysis of the Dataset	101
6.5	Comparison With Other Datasets	103
6.6	Conclusion	104
7	A Machine Learning Platform for Detecting DDoS Attacks in IoT Based Networks	106
7.1	Introduction	106
7.2	Machine Learning Algorithms for Attack Detection	107
7.2.1	Support Vector Machine Testing	107
7.2.2	Artificial Neural Networks	112
7.2.3	Random Forest	115
7.3	Machine Learning Pre-Processing	116
7.4	Machine Learning Detection Evaluation	123
7.4.1	SVM Experiment	123
7.4.2	ANN Experiment	126
7.4.3	Random Forest Experiment	129
7.5	Discussion and Evaluation	131
7.6	Conclusion	133
8	A Machine Learning IDS Implementation in IoT Networks	134
8.1	Introduction	134
8.2	IDS Implementation in the IoT Network	134
8.2.1	Real-Time Data Monitoring and Aggregation Unit	135

8.2.2	Attack Detection Unit	136
8.2.3	IDS Agent Implementation	138
8.3	Deployment of the Machine Learning Model	140
8.4	Evaluation	145
8.4.1	Network Performance Evaluation:	148
8.4.2	Signal Node Evaluation:	150
8.5	Web Interface to Monitor IDS-Performance	151
8.6	Conclusion	152
9	Recapitulation and Future Research Directions	153
9.1	Introduction	153
9.2	Problems Addressed in This Thesis	154
9.3	Contribution of This Thesis to the Existing Literature	154
9.3.1	Contribution 1: Comprehensive State-of-the-art Survey of the Existing Literature	154
9.3.2	Contribution 2: A Framework for Attack Detection in IoT Using Intelligent Machine learning Methods	155
9.3.3	Contribution 3: A real-time Framework for Dataset Generation in IoT Environments	156
9.3.4	Contribution 4: Intelligent Machine Learning Methodology for DDoS Attack Detection in IoT	157
9.3.5	Contribution 5: Implementation and Evaluation of the Proposed IDS Framework in a Semi-Real IoT Environment	158
9.4	Conclusion and Future Work	159
	Bibliography	161

List of Figures

1.1	IoT three-layer architecture [9]	6
2.1	Literature review filtration process	14
2.2	Paper distribution based on database	17
2.3	Distribution of papers by keyword	18
3.1	Overview of the design science research methodology	48
4.1	IDD-IoT framework	53
4.2	Detection method	57
4.3	Example of a confusion matrix	58
4.4	Post learning phase	60
4.5	Proposed network diagram	61
5.1	Intelligent DDoS detection framework (IDD-IoT)	64
5.2	Framework phases	65
5.3	Data collection module	67
5.4	Example of training ML model	68
5.5	Pre-learning workflow	69
5.6	Data detection model	71

5.7	Workflow of the GET operation	74
5.8	Example IoT network	75
6.1	DEF tool	78
6.2	Screenshot of the DEF tool	80
6.3	Data exportation framework API	81
6.4	Example of the JSON output	82
6.5	IoT simulation network	84
6.6	The proposed IDS network architecture and placement	85
6.7	Data collection module	87
6.8	Pictorial visualization of attacks	87
6.9	Attack network topology	92
6.10	UDP statistics	93
6.11	Data collection model	96
6.12	UDP payload	99
6.13	Protocol distribution	100
6.14	Network flow	101
6.15	Rank change over time	101
6.16	DIO comparison	102
7.1	SVM diagram	108
7.2	Multi-layer perceptron	113
7.3	Random forest example	115
7.4	IoT-DDoS subsets	117
7.5	The ML framework used in this thesis	118

7.6	IoT-DDoS heatmap	119
7.7	SVM confusion matrix BH	124
7.9	SVM hyperplane separation SF	125
7.11	Confusion matrix for artificial neural network	127
7.12	Random forest confusion matrix for blackhole attack	129
7.14	Recall chart comparing SVM, ANN and RF	131
7.15	Precision chart comparing SVM, ANN and RF	131
7.16	F1-score chart comparing SVM, ANN and RF	132
8.1	Overview of the IDS framework	135
8.2	Pipeline for the real-time data monitoring and aggregation unit . . .	135
8.3	Sequence diagram for the IDS framework	137
8.4	Alert packet	138
8.5	The IDS framework REST-API unit used for third party plugins . . .	141
8.6	Blackhole attack implementation	145
8.7	IDS in operation	147
8.8	Network throughput	148
8.9	packet loss chart	149
8.10	Rank change over time (blackhole)	150
8.11	Power consumption of node 10	150
8.12	Power consumption of node 2	151
8.13	Web monitor	151

List of Tables

2.1	Survey paper comparison	15
2.2	Databases	16
2.3	Attack summary	17
2.4	Protocol-based solutions	22
2.5	Trust-based solutions summary	27
2.6	Intrusion detection solution summary	30
2.7	Overview of IoT security approaches	36
5.1	Data collection features	66
6.1	Workstation setup	83
6.2	DEF Tool Performance	85
6.3	Simulation parameters	86
6.4	Sensor node parameters	88
6.5	Data collection features	98
6.6	Protocol distribution	100
6.7	Sample of the dataset	103
6.8	Dataset comparison	103
7.1	Dataset samples	119

7.2	Main matrices used	121
7.3	SVM confusion matrix comparison	123
7.4	Example of one of the K-fold tests	126
7.5	Example of one of the K-fold tests	128
7.6	RF hyperparameter tuning	130
7.7	Final results	133
8.1	Detection rate table	146
8.2	Detection rate table	146
8.3	Simulation parameters	148