

# Software Defined Networking: Network Intrusion Detection System



Tuan Anh Tang

Submitted in accordance with the requirements for the degree of

*Doctor of Philosophy*

The University of Leeds

Department of Electronic and Electrical Engineering

June 2019

## Declaration

The work in this thesis is based on research carried out at the Institute of Robotics, Autonomous Systems and Sensing (IRASS), School of Electronic and Electrical Engineering, Leeds University, UK. The candidate confirms that no part of this thesis has been submitted elsewhere for any other degree or qualification and it is all his own work except where work which has formed part of jointly authored publications. The contribution of the candidate and other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others. It is to assert that the candidate has contributed solely to the technical part of the joint publication under the guidance of his academic supervisors. The detailed contributions of the authors are as the following:

**Author:** Tuan Anh Tang

**Contributions:** Proposed and implemented the SDN-based intrusion detection models. Wrote first and final drafts of the manuscripts.

**Co-Author:** Dr. Des McLernon

**Contributions:** Provided overall supervision on the whole project. Provided feedback on technical analyses and final drafts of the manuscripts.

**Co-Author:** Dr. Lotfi Mhamdi

**Contributions:** Provided feedback on technical analyses related to SDN aspects and drafts of the manuscripts.

**Co-Author:** Dr. Syed Ali Raza Zaidi

**Contributions:** Provided feedback on technical analyses and drafts of the manuscripts.

**Co-Author:** Professor Mounir Ghogho

**Contributions:** Provided feedback on technical analyses and drafts of the manuscripts.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Tuan Anh Tang to be identified as the author of this work has been asserted by himself in accordance with the Copyright, Designs and Patents Act 1988.

© 2019 Tuan Anh Tang and The University of Leeds

This thesis is dedicated to my family.

## Acknowledgements

First and foremost, I would like to express my sincere gratitude to Dr. Des McLernon for all of his support in my PhD life from the beginning. His guidance, feedback and encouragement have inspired and motivated me a lot in my PhD journey. But above everything, thank you very much for your caring and kindness. I wholeheartedly appreciate your kindness.

I would like to express my gratitude to Dr Lotfi Mhamdi for guiding me to the field of SDN and continuously supporting me during my research.

I would like to thank Dr. Syed Ali Raza Zaidi for all of his advice and feedback. Thank you very much for your guidance since the beginning.

My appreciation also goes to Professor Mounir Ghogho for his support and guidance. Thank you very much for the constructive talks at the early stage of my research that gave me the right research direction.

I thank my fellow lab-mates in rooms 3.62 and 2.59, School of Electronic and Electrical, for all the fruitful discussions and for all the fun we have had in the last four years. I would like to especially thank Yen, Bao, Asma, Ali, Mohanad, Edmond, Bao and Naveed for all the precious things that we have shared.

Last but not least, I would like to thank my family: my parents, and my sisters for supporting and encouraging me endlessly. Thank you very much for being with me through all the ups and downs. Without you, I would not be here now.

## **Abstract**

Software Defined Networking (SDN) is developing as a new solution for the development and innovation of the Internet. SDN is expected to be the ideal future for the Internet since it can provide controllable, dynamic and cost-effective networking. The emergence of SDN provides a unique opportunity to achieve network security in a more efficient and flexible manner. One key advantage of SDN, as compared to traditional networks, is that by virtue of centralized control, it allows better provisioning of network security. Nevertheless, the flexibility provided by the SDN architecture manifests several new network security issues that must be addressed to strengthen SDN security. The SDN has original structural vulnerabilities, which are the centralized controller, the control-data interface and the control-application interfaces. These vulnerabilities can be exploited by intruders to conduct several types of attacks.

Network Intrusion Detection System (NIDS), which is an important part of network architecture, is used to detect network intrusions and secure the whole network. In this thesis, we propose an SDN-based NIDS (DeepIDS) using Deep Learning (DL) algorithms to detect anomalies in the SDN architecture. Firstly, we evaluate the potential of DL for flow-based anomaly detection with different flow features.

Through experiments, we confirm that the DL approach has the potential for flow-based anomaly detection in the SDN environment. Secondly, we propose a Gated Recurrent Unit Recurrent Neural Network (GRU-RNN) to improve the detection rate of the DeepIDS. Our experimental results show that the proposed GRU-RNN model improves the detection rate significantly without deteriorating network performance. The performance of our system in terms of accuracy, throughput, latency and resource utilization shows that DeepIDS does not affect the performance of the OpenFlow controller, and so is a feasible approach.

Finally, we introduce an unsupervised approach (SAE-1SVM) to solve an unlabeled and imbalanced dataset problem. This approach yields a high detection rate while maintaining a significantly low processing time. Through extensive experimental evaluations, we conclude that our proposed approach exhibits a strong potential for intrusion detection in the SDN environments.

## Abbreviations

<i>ACC</i>	Accuracy
<i>AE</i>	Autoencoder
<i>AUC</i>	Area Under the Curve
<i>DDoS</i>	Distributed Denial of Service Attack
<i>DoS</i>	Denial of Service Attack
<i>DL</i>	Deep Learning
<i>DNN</i>	Deep Neural Network
<i>F1</i>	F1-measure
<i>FP</i>	False Positive
<i>FPR</i>	False Positive Rate
<i>FN</i>	False Negative
<i>GRU</i>	Gated Recurrent Unit
<i>GRU – RNN</i>	Gated Recurrent Neural Network
<i>R</i>	Recall
<i>ROC</i>	Receiver Operating Characteristic Curve
<i>RNN</i>	Recurrent Neural Network
<i>OC – SVM</i>	One-class Support Vector Machine
<i>P</i>	Precision
<i>IDS</i>	Intrusion Detection System
<i>LSTM</i>	Long Short-Term Memory
<i>ML</i>	Machine Learning
<i>NN</i>	Neural Network
<i>NIDS</i>	Network Intrusion Detection System
<i>SAE</i>	Stacked Autoencoder
<i>SDN</i>	Software Defined Networking
<i>SOM</i>	Self-Organizing Map
<i>SVM</i>	Support Vector Machine
<i>TP</i>	True Positive
<i>TPR</i>	True Positive Rate
<i>TN</i>	True Negative



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.2	Challenges . . . . .	4
1.3	Objective and Scope of the Thesis . . . . .	5
1.4	Limitations and Constraints . . . . .	5
1.5	Thesis Outline and Contributions . . . . .	6
1.6	Publications . . . . .	9
<b>2</b>	<b>Software Defined Networking</b>	<b>11</b>
2.1	Software Defined Networking . . . . .	11
2.1.1	Definition . . . . .	11
2.1.2	OpenFlow Protocol . . . . .	18
2.1.3	Security in SDN . . . . .	25
2.2	Network Intrusion Detection System . . . . .	28
2.2.1	Signature-based Detection . . . . .	30
2.2.2	Anomaly-based Detection . . . . .	32

2.3	Intrusion Detection in the SDN . . . . .	36
2.4	Intrusion Detection Performance Evaluation . . . . .	39
2.5	SDN-based NIDS: An Example . . . . .	41
2.5.1	Experimental Setup . . . . .	43
2.5.2	Simulation Results . . . . .	44
2.6	Conclusion . . . . .	46
<b>3</b>	<b>Deep Learning</b>	<b>47</b>
3.1	Notations . . . . .	47
3.2	Introduction . . . . .	48
3.3	Advantages and Disadvantages of DL . . . . .	53
3.4	Styles of Learning . . . . .	54
3.5	Training and Testing Neural Networks . . . . .	54
3.6	Network Datasets . . . . .	57
3.6.1	NSL-KDD Dataset . . . . .	58
3.6.2	CICIDS2017 Dataset . . . . .	63
3.7	Conclusion . . . . .	65
<b>4</b>	<b>DeepIDS: Deep Learning Approach for Intrusion Detection in Software Defined Networking</b>	<b>66</b>
4.1	Introduction . . . . .	67
4.1.1	Motivation . . . . .	67
4.1.2	Contribution . . . . .	67
4.1.3	Chapter Organization . . . . .	68
4.2	DeepIDS System Architecture . . . . .	69
4.2.1	The Collector . . . . .	71

4.2.2	The Anomaly Detector . . . . .	73
4.2.3	The Counter Measure Deployment . . . . .	73
4.3	Experimental Methodology . . . . .	74
4.3.1	DL Experimental Setup . . . . .	74
4.3.2	Network Performance Evaluation Setup . . . . .	77
4.4	Detection Performance Evaluation . . . . .	79
4.4.1	Effect of The Learning Rate . . . . .	79
4.4.2	Effect of The Feature Set . . . . .	81
4.4.3	Efficiency Evaluation . . . . .	88
4.5	Network Performance Evaluation . . . . .	89
4.5.1	Throughput Evaluation . . . . .	89
4.5.2	Latency Evaluation . . . . .	91
4.5.3	Resource Utilization . . . . .	92
4.6	Conclusion . . . . .	93
<b>5</b>	<b>Deep Recurrent Neural Networks for SDN-based Intrusion De-</b>	
	<b>tection Systems</b>	<b>94</b>
5.1	Introduction . . . . .	95
5.1.1	Motivation . . . . .	95
5.1.2	Contribution . . . . .	95
5.1.3	Chapter Organization . . . . .	96
5.2	Recurrent Neural Networks . . . . .	96
5.3	Detection Performance Evaluation . . . . .	99
5.3.1	Experimental Methodology . . . . .	99
5.3.2	Experimental Result Analysis . . . . .	101

5.4	Network Performance Evaluation . . . . .	107
5.4.1	Throughput Evaluation . . . . .	108
5.4.2	Latency Evaluation . . . . .	109
5.5	Conclusion . . . . .	110
<b>6</b>	<b>Deep Learning Approach Combining Stacked Autoencoder with One-class SVM for DDoS Attack Detection in SDNs</b>	<b>112</b>
6.1	Introduction . . . . .	113
6.1.1	Motivation . . . . .	113
6.1.2	Contribution . . . . .	114
6.1.3	Chapter Organization . . . . .	114
6.2	Denial-of-Service Attacks . . . . .	115
6.3	SAE-1SVM for DDoS Attack Detection . . . . .	118
6.4	Detection Performance Evaluation . . . . .	121
6.4.1	Experimental Setup . . . . .	121
6.4.2	DDoS Attack Detection with a Hard Threshold . . . . .	123
6.4.3	DDoS Attack Detection with the SAE-1SVM . . . . .	127
6.5	Conclusion . . . . .	129
<b>7</b>	<b>Conclusions and Future Work</b>	<b>130</b>
7.1	Conclusions . . . . .	130
7.2	Limitations . . . . .	132
7.3	Future Work . . . . .	133
7.4	Final Remarks . . . . .	135
<b>A</b>	<b>Appendix</b>	<b>137</b>

## CONTENTS

---

A.1	Preparing Data . . . . .	137
A.2	Performance Evaluation Processes . . . . .	137
A.3	Experiment Testbed Setup . . . . .	138
	<b>References</b>	<b>155</b>

# List of Figures

1.1	The SDN Market Size Prediction [1] . . . . .	2
1.2	The Overview of the NIDS . . . . .	5
2.1	A three-layer SDN Architecture [2] . . . . .	16
2.2	OpenFlow Switch [3] . . . . .	19
2.3	OpenFlow Packet Processing Pipeline [3] . . . . .	22
2.4	OpenFlow Packet Matching Process [3] . . . . .	24
2.5	Overview of Intrusion Detection System . . . . .	29
2.6	Commonly Used Machine Learning Techniques [4] . . . . .	35
2.7	The ROC Curve Example. The closer the ROC curve to the top left corner, the better the result is . . . . .	41
2.8	The Emulated Network Topology . . . . .	43
2.9	Entropy Comparision for Legitimate, 25% Rate DDoS Attack and 75% Rate Attack Traffic . . . . .	45
3.1	The AI, ML and DL Relationships [5] . . . . .	48
3.2	A Drawing of a Biological Neuron [6] . . . . .	49

**LIST OF FIGURES**

---

3.3	The Single Artificial Neuron Structure . . . . .	49
3.4	The NN Architecture with an Input Layer, a Hidden Layer and an Output Layer . . . . .	51
3.5	Overview of The DL Algorithm . . . . .	52
3.6	The Difference Between Traditional ML and DL Algorithms . . . .	53
3.7	The green line represents an overfitted model, and the black line represents a regularized model. While the green line best follows the training data, it is too dependent on that data, and it is likely to have a higher error rate on new unseen data, compared to the black line [7] . . . . .	58
3.8	ML/DL Training and Testing Phase . . . . .	59
4.1	The DeepIDS Architecture . . . . .	69
4.2	Network Intrusion Detection Framework . . . . .	70
4.3	The DNN Structure . . . . .	75
4.4	The Network Evaluation Process . . . . .	78
4.5	ROC curve comparison for different feature sets. The Mixed Fea- ture Set achieves the best result with highest AUC and lowest FPR	82
4.6	ROC Curve Comparison of Different Algorithms for Basic Feature Set . . . . .	84
4.7	ROC Curve Comparison of Different Algorithms for Traffic Feature Set . . . . .	85
4.8	ROC Curve Comparison of Different Algorithms for Mixed Feature Set . . . . .	86
4.9	Throughput Evaluation (log scale on x-axis) . . . . .	90

## LIST OF FIGURES

---

4.10 Latency Evaluation (log scale on x-axis) . . . . .	92
5.1 The RNN Unfolded in Time . . . . .	97
5.2 Gated Recurrent Unit Structure [8] . . . . .	98
5.3 The RNN Structure . . . . .	101
5.4 Traing and Testing Phase Evaluation . . . . .	102
5.5 Performance Metric Comparison . . . . .	104
5.6 ROC Curve Comparison for Different Algorithms . . . . .	105
5.7 P vs R Curves . . . . .	106
5.8 Throughput Evaluation . . . . .	108
5.9 Latency Evaluation . . . . .	110
6.1 DDoS Attack Topology . . . . .	116
6.2 A General Structure of an AE . . . . .	118
6.3 SAE-1SVM System Detail . . . . .	121
6.4 Reconstruction Error Rate Comparision . . . . .	126
A.1 The Emulated Network Topology . . . . .	139



# List of Tables

2.1	SDN Controllers . . . . .	18
2.2	Main Components of a Flow Entry . . . . .	20
2.3	Match Fields in OpenFlow . . . . .	21
2.4	SDN Threat Vectors . . . . .	26
2.5	SDN Attack Summary . . . . .	26
2.6	A Binary Confusion Matrix . . . . .	39
3.1	Definitions of Activation Function . . . . .	50
3.2	The NSL-KDD Dataset Distribution . . . . .	59
3.3	The NSL-KDD Dataset Features . . . . .	60
3.4	Attacks in The NSL-KDD Dataset . . . . .	61
3.5	The NSL-KDD Dataset Examples . . . . .	62
3.6	The CICIDS2017 Dataset Description . . . . .	63
3.7	The CICIDS2017 Dataset Examples . . . . .	64
4.1	Flow Feature Examples . . . . .	71
4.2	The DNN Model Structure . . . . .	75

## LIST OF TABLES

---

4.3	Feature Set Description . . . . .	76
4.4	Network Parameters . . . . .	78
4.5	Loss and Accuracy Evaluation for Different Learning Rates . . . . .	79
4.6	Accuracy Metrics for Different Learning Rates . . . . .	80
4.7	Accuracy Evaluation for Different Feature Set . . . . .	81
4.8	Performance Metric Evaluation of Three Feature Sets . . . . .	82
4.9	Performance Metric Evaluation of the Mixed Feature Set . . . . .	85
4.10	Accuracy Comparison for Different Algorithms Using the Mixed Feature Set . . . . .	87
4.11	Accuracy comparison of different algorithms. The DNN uses the Mixed Feature Set with 6 features. The others use 41 features . . .	87
4.12	Running Time Evaluation . . . . .	88
4.13	Resource Utilization Evaluation . . . . .	93
5.1	The CICIDS2017's Feature Description . . . . .	100
5.2	Neural Network Model Structures . . . . .	100
5.3	Accuracy Comparison with Other Algorithms . . . . .	103
5.4	Performance Metric Evaluation for the NSL-KDD Dataset . . . . .	103
5.5	Accuracy comparison with previous studies. The GRU-RNN and DNN use the Mixed Feature Set with 6 features. The SVM and NB Tree use the NSL-KDD dataset with 41 features . . . . .	107
5.6	Performance Metric Evaluation for the CICIDS2017 dataset . . . . .	107
6.1	The SAE Architecture . . . . .	122
6.2	The CICIDS2017's Feature Description . . . . .	122
6.3	Network Architecture Details . . . . .	123

## LIST OF TABLES

---

6.4	Reconstruction ACC Comparison . . . . .	124
6.5	Accuracy Metrics for Different Thresholds . . . . .	125
6.6	The Evaluation Metric Comparison . . . . .	128
6.7	The Detection Performane Results with the Friday Dataset . . . . .	128
6.8	The Training and Testing Time Comparison . . . . .	129