

AN ADAPTIVE AND DISTRIBUTED INTRUSION DETECTION SCHEME FOR  
CLOUD COMPUTING

NURUDEEN MAHMUD IBRAHIM

A thesis submitted in fulfilment of the  
requirements for the award of the degree of  
Doctor of Philosophy (Computer Science)

School of Computing  
Faculty of Engineering  
Universiti Teknologi Malaysia

SEPTEMBER 2019

## **DEDICATION**

To Almighty Allah (SWT) be the Glory

## **ACKNOWLEDGEMENT**

First I wish to express immense gratitude to Almighty Allah (SWT) for keeping me in good health and granting me the courage and ability to conduct the research.

Also my profound gratitude to my parents Alhaji Ibrahim Mahmud and Hauwa Salihu and all family members for their moral support during difficult times in the course of my Ph.D study.

My immense gratitude also goes to my supervisor Dr. Anazida Zainal for the guidance, helpful advices and critics. Your comments and constructive criticism have gone a long way in shaping my research and providing useful insights towards the realization of the research works. May Allah reward you for your effort.

Furthermore, I wish to express profound gratitude to Professor Richard Jensen, Department of Computer Science, Aberystwyth University, UK and Tian Huang, Shanghai Jiao Tong University, China for useful assistance towards realization of some of my research objective. Finally I wish to express my gratitude to all those whose assistance either consciously or unconsciously have contributed to the actualization of the research, may Allah shower His blessings on you all.

## ABSTRACT

Cloud computing has enormous potentials but still suffers from numerous security issues. Hence, there is a need to safeguard the cloud resources to ensure the security of clients' data in the cloud. Existing cloud Intrusion Detection System (IDS) suffers from poor detection accuracy due to the dynamic nature of cloud as well as frequent Virtual Machine (VM) migration causing network traffic pattern to undergo changes. This necessitates an adaptive IDS capable of coping with the dynamic network traffic pattern. Therefore, the research developed an adaptive cloud intrusion detection scheme that uses Binary Segmentation change point detection algorithm to track the changes in the normal profile of cloud network traffic and updates the IDS Reference Model when change is detected. Besides, the research addressed the issue of poor detection accuracy due to insignificant features and coordinated attacks such as Distributed Denial of Service (DDoS). The insignificant feature was addressed using feature selection while coordinated attack was addressed using distributed IDS. Ant Colony Optimization and correlation based feature selection were used for feature selection. Meanwhile, distributed Stochastic Gradient Decent and Support Vector Machine (SGD-SVM) were used for the distributed IDS. The distributed IDS comprised detection units and aggregation unit. The detection units detected the attacks using distributed SGD-SVM to create Local Reference Model (LRM) on various computer nodes. Then, the LRM was sent to aggregation units to create a Global Reference Model. This Adaptive and Distributed scheme was evaluated using two datasets: a simulated datasets collected using Virtual Machine Ware (VMWare) hypervisor and Network Security Laboratory–Knowledge Discovery Database (NSL-KDD) benchmark intrusion detection datasets. To ensure that the scheme can cope with the dynamic nature of VM migration in cloud, performance evaluation was performed before and during the VM migration scenario. The evaluation results of the adaptive and distributed scheme on simulated datasets showed that before VM migration, an overall classification accuracy of 99.4% was achieved by the scheme while a related scheme achieved an accuracy of 83.4%. During VM migration scenario, classification accuracy of 99.1% was achieved by the scheme while the related scheme achieved an accuracy of 85%. The scheme achieved an accuracy of 99.6% when it was applied to NSL-KDD dataset while the related scheme achieved an accuracy of 83%. The performance comparisons with a related scheme showed that the developed adaptive and distributed scheme achieved superior performance.

## ABSTRAK

Pengkomputeran awan mempunyai potensi besar, namun masih mengalami banyak masalah keselamatan. Oleh itu, terdapat keperluan dalam sistem perlindungan sumber awan untuk memastikan keselamatan data pelanggan di awan. Sistem Pengesanan Pencerobohan (IDS) awan sedia ada mengalami ketepatan pengesanan yang lemah disebabkan sifat awan yang dinamik serta penghijrahan Mesin Maya (VM) yang menyebabkan pola trafik rangkaian mengalami perubahan. Ini memerlukan IDS adaptif yang mampu mengendalikan corak trafik rangkaian yang dinamik. Oleh itu, penyelidikan ini membangunkan skim pengesanan pencerobohan awan adaptif yang menggunakan algoritma pengesanan titik perubahan Pensegmenan Binari untuk mengesan perubahan dalam profil normal trafik rangkaian awan dan mengemas kini Model Rujukan IDS apabila terdapat perubahan. Selain itu, penyelidikan ini membincangkan isu ketepatan pengesanan lemah yang disebabkan oleh ciri-ciri yang tidak penting dan serangan terancang seperti Perkhidmatan Penafian Teragih (DDoS). Ciri tidak penting ditangani menggunakan pemilihan ciri manakala serangan terancang ditangani menggunakan IDS teragih. Pengoptimuman Koloni Semut dan pemilihan ciri berasaskan korelasi digunakan untuk proses pemilihan ciri. Selain itu, Penurunan Cerun Stokastik teragih dan Mesin Vektor Sokongan (SGD-SVM) telah digunakan untuk IDS teragih. IDS teragih terdiri daripada unit pengesanan dan unit pengagregatan. Unit pengesanan mengesan serangan menggunakan SGD-SVM teragih untuk mencipta Model Rujukan Tempatan (LRM) pada sebilangan nod komputer. Kemudian LRM dihantar ke unit pengagregatan untuk penciptaan Model Rujukan Global. Skim Adaptif dan Teragih telah dinilai menggunakan dua dataset: dataset simulasi yang dikumpulkan menggunakan dataset pengesanan pencerobohan penanda aras pengkomputeran makmal (VMWare) dan *hypervisor* serta *Database Keselamatan Makmal-Pangkalan Data Pengetahuan (NSL-KDD)*. Untuk memastikan kaedah ini dapat menangani sifat dinamik penghijrahan VM di awan, penilaian prestasi dilakukan sebelum dan semasa senario penghijrahan VM. Hasil penilaian skim adaptif dan teragih pada dataset simulasi menunjukkan sebelum penghijrahan VM, ketepatan klasifikasi keseluruhan sebanyak 99.4% dicapai oleh skim adaptif dan teragih manakala skim yang berkaitan mencapai ketepatan 83.4%. Semasa senario migrasi VM, ketepatan pengelasan 99.1% dicapai oleh skim cadangan manakala skim berkaitan mencapai ketepatan 85%. Skim ini mencapai ketepatan 99.6% apabila ia digunakan untuk dataset NSL-KDD manakala skim yang berkaitan mencapai ketepatan 83%. Perbandingan yang dibuat menunjukkan bahawa prestasi skim cadangan adalah lebih hebat daripada skim berkaitan.

## TABLE OF CONTENTS

	<b>TITLE</b>	<b>PAGE</b>
	<b>DECLARATION</b>	<b>ii</b>
	<b>DEDICATION</b>	<b>iii</b>
	<b>ACKNOWLEDGEMENT</b>	<b>iv</b>
	<b>ABSTRACT</b>	<b>v</b>
	<b>ABSTRAK</b>	<b>vi</b>
	<b>TABLE OF CONTENTS</b>	<b>vii</b>
	<b>LIST OF TABLES</b>	<b>xii</b>
	<b>LIST OF FIGURES</b>	<b>xiv</b>
	<b>LIST OF ABBREVIATIONS</b>	<b>xviii</b>
	<b>LIST OF SYMBOLS</b>	<b>xxii</b>
<b>CHAPTER 1</b>	<b>INTRODUCTION</b>	<b>1</b>
1.1	Overview	1
1.2	Problem Background	3
1.3	Problem Statement	10
1.4	Research Aim	11
1.5	Research Objectives	11
1.6	Scope of Study	12
1.7	Significance of the Research	12
1.8	Research Contributions	13
1.9	Thesis Organization	15
<b>CHAPTER 2</b>	<b>LITERATURE REVIEW</b>	<b>17</b>
2.1	Introduction	17
2.2	Cloud Computing	17
2.2.1	Cloud Architecture	18
2.2.2	Key Cloud Features	19
2.2.3	Core Technologies in Cloud Computing	20

2.2.4	Cloud Service Delivery Model	22
2.2.5	Cloud Deployment Models	23
2.3	Security Issues in Cloud Computing	24
2.3.1	Attacks in Cloud Computing	26
2.4	Intrusion Detection System	28
2.5	Taxonomy of IDS in Cloud Computing	29
2.5.1	Detection Domain	30
2.5.2	Detection Technique	32
2.5.2.1	Signature-based Technique	32
2.5.2.2	Anomaly Detection	33
2.5.2.3	Hybrid Technique	38
2.5.3	Architectural Structure	39
2.5.3.1	Standalone	39
2.5.3.2	Distributed/Collaborative	39
2.5.4	Detection Time	40
2.6	Change point Detection	40
2.6.1	Binary Segmentation	41
2.7	Research Issues Hampering Detection Accuracy of Cloud IDS	42
2.7.1	Redundant and Insignificant Features	43
2.7.1.1	Feature Selection	43
2.7.1.2	Filter Approach	43
2.7.1.3	Related Works on Filter-based Feature Selection	45
2.7.1.4	Wrapper-based Approach	47
2.7.1.5	Embedded Approach	51
2.7.2	Adaptive Cloud-based IDS	53
2.7.2.1	Related works for adaptive cloud IDS	54
2.7.2.2	Existing adaptive cloud intrusion detection schemes	55
2.7.3	Distributed Attacks	58

2.7.3.1	Existing distributed approach for cloud intrusion detection	59
2.7.3.2	Other distributed IDS approach for non-cloud based applications	64
2.8	Validation Dataset	66
2.9	Summary of Research works in Cloud IDS	72
2.10	Research Direction	79
2.11	Definition of Terms	80
2.12	Summary	81
<b>CHAPTER 3</b>	<b>RESEARCH METHODOLOGY</b>	<b>83</b>
3.1	Introduction	83
3.2	Overview of Research Framework	83
3.3	Benchmarking	87
3.4	Details of the Research Framework	87
3.4.1	Phase 1: Data pre-processing	87
3.4.2	Phase 2: Design of the Adaptive Cloud Intrusion Detection Scheme	89
3.4.3	Phase 3: Design of the Adaptive and Distributed Cloud Intrusion Detection Scheme	90
3.5	Datasets	91
3.6	Experimental Set-up	96
3.7	Evaluation Metrics	96
3.8	Result Validation and Analysis	97
3.9	Summary	98
<b>CHAPTER 4</b>	<b>ANT COLONY OPTIMIZATION AND CORRELATION-BASED FEATURE SELECTION</b>	<b>99</b>
4.1	Introduction	99
4.2	Design and Implementation of the Proposed ACO-CFS	99
4.2.1	The Proposed Feature Selection using ACO and CFS	100
4.2.2	Evaluation of Selected Features using Random Forest	106
4.2.3	Parameter Settings	108



4.2.4	Dataset Organization	110
4.3	Result and Analysis	111
4.3.1	Result and Analysis of the Proposed ACO-CFS on Simulated Dataset	111
4.3.2	Result and Analysis of the Proposed ACO-CFS on NSL-KDD	116
4.4	Significance Test	120
4.5	Summary	121
<b>CHAPTER 5</b>	<b>ADAPTIVE CLOUD INTRUSION DETECTION SCHEME</b>	<b>123</b>
5.1	Introduction	123
5.2	Design and Implementation of the Proposed Adaptive Cloud Intrusion Detection Scheme	123
5.2.1	Dataset Organization	133
5.2.2	Parameter Settings	134
5.3	Result and Analysis	135
5.3.1	Comparison with Related Works	143
5.3.2	Significance Test	150
5.4	Summary	150
<b>CHAPTER 6</b>	<b>ADAPTIVE AND DISTRIBUTED CLOUD INTRUSION DETECTION SCHEME (A-D-CIDS)</b>	<b>153</b>
6.1	Introduction	153
6.2	Design and Implementation of the Proposed Adaptive and Distributed Cloud Intrusion Detection Scheme (A-D-CIDS)	154
6.2.1	Dataset Organization	161
6.2.2	Parameter Settings	161
6.3	Result and Analysis	162
6.3.1	Comparison with Related Works	171
6.3.2	Significance Test	173
6.4	Summary	174
<b>CHAPTER 7</b>	<b>CONCLUSION</b>	<b>175</b>
7.1	Concluding Remarks	175

7.2	Research Contributions and Future Work	178
	<b>REFERENCES</b>	<b>179</b>
	<b>LIST OF PUBLICATIONS</b>	<b>196</b>

## LIST OF TABLES

<b>TABLE NO.</b>	<b>TITLE</b>	<b>PAGE</b>
Table 2.1	Related Works on Feature Selection Techniques	52
Table 2.2	Related Works on Adaptive Cloud IDS	55
Table 2.3	Related Works on Distributed Cloud IDS	60
Table 2.4	Summary of Related Works in Cloud IDS	74
Table 3.1	General Research Framework	86
Table 3.2	Description of Simulated Dataset	94
Table 3.3	Features of Simulated Dataset	94
Table 3.4	Features of NSL-KDD	95
Table 3.5	IDS Evaluation Metrics	97
Table 4.1	Parameter settings for ACO	109
Table 4.2	Parameter settings for Random Forest	109
Table 4.3	Parameter settings for GA	109
Table 4.4	Training and Test Samples for Simulated Datasets	110
Table 4.5	Training and Test Samples for NSL-KDD	110
Table 4.6	Features selected by the proposed ACO-CFS, EMFFS and GA on simulated datasets	113
Table 4.7	Result of accuracy, DR and FPR of the proposed ACO-CFS on simulated datasets	114
Table 4.8	Comparison of accuracy, DR and FPR of the proposed ACO-CFS with EMFFS and GA on simulated datasets	114
Table 4.9	Features selected by ACO-CFS, GR and IG on NSL-KDD	117
Table 5.1	Train/Test Split for Simulated Dataset	133
Table 5.2	Train/Test Split for NSL-KDD	134
Table 5.3	Parameter settings for SGD-SVM	135
Table 5.4	Parameter settings for Binary Segmentation	135
Table 5.5	Parameter settings for PCA, K-Means and LOF	135
Table 5.6	Result of change point on all samples of dataset 1 and 2	137

Table 5.7	Result of change point on NSL-KDD	137
Table 5.8	Result for A-CIDS on Simulated Datasets	142
Table 5.9	Comparison between A-CIDS vs K-Means, PCA and A-DR-LOF-SAX on Simulated Datasets	145
Table 5.10	Comparison between A-CIDS vs K-Means, PCA and A-DR-LOF-SAX on NSL-KDD	145
Table 5.11	Result of t-test for A-CIDS against K-Mean, PCA and A-DR-LOF-SAX	150
Table 6.1	Result for A-D-CIDS on the simulated datasets	168
Table 6.2	Comparison between A-CIDS and A-D-CIDS on simulated dataset	168
Table 6.3	Comparison between A-D-CIDS and CIDS on NSL-KDD	172
Table 6.4	Result of t-test for A-D-CIDS vs A-CIDS	173

## LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
Figure 1.1	Anomaly detection process (Mishra <i>et al.</i> , 2017)	2
Figure 1.2	Scenario leading to the problem	5
Figure 1.3	Phases in the design of the adaptive and distributed cloud intrusion detection scheme	13
Figure 1.4	Research contribution	14
Figure 2.1	Cloud Architecture (Bohn <i>et al.</i> , 2011)	19
Figure 2.2	Virtual Machine Migration (Vmware Inc, 2009a)	22
Figure 2.3	Security issues in cloud computing	25
Figure 2.4	DDoS attack (Beitollahi and Deconinck, 2012)	28
Figure 2.5	Taxonomy of Cloud IDS	29
Figure 2.6	Host-based/Network-based IDS (Poston, 2017)	31
Figure 2.7	Cause effect diagram for issues hampering accuracy of cloud IDS	42
Figure 2.8	Adaptive anomaly detection scheme proposed by Huang <i>et al.</i> , (2013)	56
Figure 2.9	Adaptive anomaly detection scheme proposed by Huang <i>et al.</i> , (2016)	58
Figure 2.10	Distributed IDS proposed by Singh <i>et al.</i> , (2016)	61
Figure 2.11	Distributed IDS proposed by Idhammad <i>et al.</i> , (2018)	62
Figure 2.12	Distributed IDS proposed by Alhaddad <i>et al.</i> , (2016)	63
Figure 2.13	Distribution of datasets used for research in cloud IDS	67
Figure 2.14	Xen architecture (Barham <i>et al.</i> , 2003)	70
Figure 2.15	KVM architecture (Binu and Kumar, 2011)	70
Figure 2.16	VSphere components (VMware, Inc., 2009b)	71
Figure 3.1	Framework for the Adaptive and Distributed Cloud Intrusion Detection Scheme	85
Figure 3.2	Data simulation set-up (VMware Inc., 2009a)	93
Figure 4.1	Feature selection using ACO-CFS	102

Figure 4.2	Pseudo code for the proposed ACO-CFS feature Selection	103
Figure 4.3	ACO graph representation	104
Figure 4.4	Pseudocode for Random Forest classifier (Kouzani and Nasireding, 2009).	108
Figure 4.5	Comparison of accuracy of ACO-CFS vs EMFFS and GA on simulated dataset	114
Figure 4.6	Comparison of DR of ACO-CFS vs EMFFS and GA on simulated datasets	115
Figure 4.7	Comparison of FPR of ACO-CFS vs EMFFS and GA on simulated datasets	115
Figure 4.8	Comparison of accuracy of ACO-CFS vs IG and GR on NSL-KDD	118
Figure 4.9	Comparison of DR of ACO-CFS vs IG and GR on NSL-KDD	119
Figure 4.10	Comparison of FPR of ACO-CFS vs IG and GR on NSL-KDD	119
Figure 5.1	Framework for the Adaptive Cloud Intrusion Detection Scheme	125
Figure 5.2	Activity diagram for the Adaptive Cloud Intrusion Detection Scheme	126
Figure 5.3	Pseudo code for the Adaptive Cloud Intrusion Detection Scheme	127
Figure 5.4	Pseudo code for Binary Segmentation	130
Figure 5.5	Pseudo code for SGD-SVM	132
Figure 5.6	Change points on Sample 1 of simulated Dataset 1 and 2	138
Figure 5.7	Change points on Sample 2 of simulated Dataset 1 and 2	138
Figure 5.8	Change points on Sample 3 of simulated Dataset 1 and 2	138
Figure 5.9	Change points on Sample 4 of simulated Dataset 1 and 2	139
Figure 5.10	Change points on Sample 5 of simulated Dataset 1 and 2	139
Figure 5.11	Change points on Sample 6 of simulated Dataset 1 and 2	139
Figure 5.12	Change points on Sample 7 of simulated Dataset 1 and 2	140
Figure 5.13	Change points on NSL-KDD dataset	140

Figure 5.14	Comparison of detection rate on simulated dataset between the proposed A-CIDS vs K-Means, PCA and A-DR-LOF-SAX before VM migration	145
Figure 5.15	Comparison of detection rate on simulated dataset between the proposed A-CIDS vs K-Means, PCA and A-DR-LOF-SAX during VM migration	146
Figure 5.16	Comparison of FPR before VM migration between the proposed A-CIDS vs K-Means, PCA and A-DR-LOF-SAX on simulated dataset	146
Figure 5.17	Comparison of FPR during VM migration between the proposed A-CIDS vs K-Means, PCA and A-DR-LOF-SAX on simulated dataset	147
Figure 5.18	Comparison of accuracy before VM migration between the proposed A-CIDS vs K-Means, PCA and A-DR-LOF-SAX on simulated datasets	147
Figure 5.19	Comparison of accuracy during VM migration between the proposed A-CIDS vs K-Means, PCA and A-DR-LOF-SAX on simulated datasets	148
Figure 5.20	Comparison of accuracy between A-CIDS vs PCA, K-Means and A-DR-LOF-SAX on NSL-KDD	148
Figure 5.21	Comparison of DR between A-CIDS vs PCA, K-Means and A-DR- LOF-SAX on NSL-KDD	149
Figure 5.22	Comparison of false positive rate between A-CIDS vs PCA, K-Means and A-DR-LOF-SAX on NSL-KDD	149
Figure 6.1	Overview of the Adaptive and Distributed Cloud Intrusion Detection Scheme	155
Figure 6.2	Framework for the Proposed Adaptive and Distributed Cloud Intrusion Detection Scheme	156
Figure 6.3	Activity Diagram for the Adaptive and Distributed Cloud Intrusion Detection Scheme	156
Figure 6.4	Pseudo code for the Adaptive and Distributed Cloud Intrusion Detection Scheme	157
Figure 6.5	Destination_IP count during Normal period of simulated datasets	160
Figure 6.6	Destination host count during Normal period on NSL-KDD	160
Figure 6.7	Destination IP count during DDoS on Sample 1 of Dataset 1 and 2	163

Figure 6.8	Destination IP count during DDoS on Sample 2 of Dataset 1 and 2	163
Figure 6.9	Destination IP count during DDoS on Sample 3 of Dataset 1 and 2	164
Figure 6.10	Destination IP count during DDoS on Sample 4 of Dataset 1 and 2	164
Figure 6.11	Destination IP count during DDoS on Sample 5 of Dataset 1 and 2	164
Figure 6.12	Destination IP count during DDoS on Sample 6 of Dataset 1 and 2	165
Figure 6.13	Destination IP count during DDoS on Sample 7 of Dataset 1 and 2	165
Figure 6.14	Destination host count during DDoS on NSL_KDD	165
Figure 6.15	Comparison of accuracy before VM migration between A-D-CIDS and A-CIDS on simulated datasets	168
Figure 6.16	Comparison of accuracy during VM migration between A-D-CIDS and A-CIDS on simulated data	169
Figure 6.17	Comparison of detection rate before VM migration between A-D-CIDS and A-CIDS on simulated data	169
Figure 6.18	Comparison of detection rate during VM migration between A-D-CIDS and A-CIDS on simulated data	170
Figure 6.19	Comparison of false positive rate before VM migration between A-D-CIDS and A-CIDS on simulated dataset	170
Figure 6.20	Comparison of false positive rate during VM migration between A-D-CIDS and A-CIDS on simulated dataset	171
Figure 6.21	Comparison of accuracy between A-D-CIDS and CIDS on NSL-KDD	172
Figure 6.22	Comparison of DR between A-D-CIDS and CIDS on NSL-KDD	172
Figure 6.23	Comparison of FPR between A-D-CIDS and CIDS on NSL-KDD	173



## LIST OF ABBREVIATIONS

A-CIDS	-	Adaptive Cloud Intrusion Detection Scheme
ACO	-	Ant Colony Optimization
A-D-CIDS	-	Adaptive and Distributed Cloud Intrusion Detection Scheme
A-DR- LOF-SAX	-	Adaptive Dimension Reasoning Local Outlier Factor Symbolic Aggregate Approximation
AMD-V	-	Advanced Micro Dynamics Virtualization
ANN	-	Artificial Neural Network
AWS	-	Amazon Web Services
API	-	Application Programming Interface
CFS	-	Correlation Based Feature Selection
CIDS	-	Cloud Intrusion Detection System
CPU	-	Central Processing Unit
CSP	-	Cloud Service Provider
CSM	-	Cloud Service Management
CVE	-	Common Vulnerabilities and Exposures
DoS	-	Denial of Service
VMM DDoS	-	Virtual Machine Monitor Distributed Denial of Service
DMA	-	Direct Memory Access
DITG	-	Distributed Internet Traffic Generator
DR	-	Detection Rate
EC2	-	Elastic Compute Cloud
EMFFS	-	Ensemble Multi Filter Feature Selection
EM	-	Expectation Maximization

FNR	-	False Negative Rate
FPR	-	False Positive Rate
FSM	-	Finite State Machine
GA	-	Genetic Algorithm
GAE	-	Google Application Engine
GRM	-	Global Reference Model
KDD	-	Knowledge Discovery in Databases
LIDS	-	Local Intrusion Detection System
GA	-	Genetic Algorithm
GR	-	Gain Ratio
GAE	-	Google Application Engine
HTTP	-	Hyper Text Transfer Protocol
IaaS	-	Infrastructure-as-a-Service
IDM	-	Identity Management
IDS	-	Intrusion Detection System
IMP	-	Infrastructure Monitoring Probe
IG	-	Information Gain
IoT	-	Internet of Everything
IP	-	Internet Protocol
I/O	-	Input Output
IT	-	Information Technology
ITOC	-	Information Technology Operation Center
KVM	-	Kernel Virtual Machine
LAN	-	Local Area Network
LRM	-	Local Reference Model
LDAP	-	Light Weight Directory Access Protocol
MIMT	-	Man-in-the-Middle-Attack

NAS	-	Network Attached Storage
NFIS	-	Neuro Fuzzy Inference System
NSL- KDD	-	Network Security Laboratory Knowledge Discovery in Databases
OSLR	-	Open System Library Repository
OS	-	Operating System
PaaS	-	Platform-as-a-Service
PCA	-	Principal Component Analysis
PSO	-	Particle Swarm Optimization
R2L	-	Remote to Local
RST	-	Rough Set Theory
SaaS	-	Software-as-a Service
SGD	-	Stochastic Gradient Descent
SLA	-	Service Level Agreement
SMB	-	Small Medium Business
SMM	-	System Management Mode
SMRAM	-	System Management Random Access Memory
SNMP	-	Simple Network Management Protocol
SQL	-	Structured Query Language
SSL	-	Secure Socket Layer
SVM	-	Support Vector Machine
TCP/IP	-	Transmission Control Protocol/Internet Protocol
TNR	-	True Negative Rate
TPR	-	True Positive Rate
TSL	-	Transport Layer Security
U2R	-	User to Root
UDP	-	User Datagram Protocol

UCLA	-	University of Californian Los Angeles
UML	-	Unified Modelling language
VM	-	Virtual Machine
VMI	-	Virtual Machine Instrospection
VMM	-	Virtual Machine Monitor
VT	-	Virtualization Technology
WS	-	Web Sevices
XML	-	Extensible Mark-up Language
XSS	-	Cross Site Scripting

## LIST OF SYMBOLS

$Bf(m)$	- Penalty to guard against over fitting
$C$	- Cost function
$E_n(f_w)$	- Empirical risk
$f_w(x)$	- A function that maps value of $x$ to an output
$j_i^k$	- Set of features unvisited by ant $k$
$l$	- Loss function
$m$	- Number of change points
$ML(T_{1:m})$	- Maximum likelihood of change point occurrence
$N_{ij}$	- Heuristic desirability of choosing feature $j$ when at feature $i$
$\overline{r_{cf}}$	- Average feature class correlation
$\overline{r_{ff}}$	- Average feature-feature intercorrelation
$P(Y)$	- The marginal probability density function for an attribute $Y$
$P(Y X)$	- The conditional probability of $Y$ given $X$ .
$S^k$	- Feature subset found by ant $k$
$ S^k $	- Subset Size
$T_{1:m}$	- Number of change points and their position in Segment Neighborhood
$T_{ij}$	- Amount of virtual pheromone on edge $(i, j)$
$v$	- Value of feature
$v'$	- Scaled value
$w$	- Weight vector for SGD-SVM
$\alpha$	- A parameter that determines the relative importance of pheromone
$\beta$	- A parameter that determines the relative importance of heuristic desirability

- $\rho$  - A constant used to simulate pheromone evaporation
- $\gamma$  - Fitness function

# CHAPTER 1

## INTRODUCTION

### 1.1 Overview

Cloud computing is a computing paradigm that offers computing resources as a service via the internet (Xiong *et al.*, 2014). It has revolutionized the conventional usage of hardware and software resources as organizations can cut the cost of purchasing and maintaining expensive hardware and software by subscribing for it on a pay-per-use basis. Cloud computing is a promising and emerging IT technology with enormous potentials and benefits to customers; however it has underlying security issues and vulnerabilities (Khorshed *et al.*, 2012). Examples of security threats capable of compromising the cloud security are Virtual Machine Monitor (VMM) DoS, port scanning and Man-in-the-Middle-Attack (Mishra *et al.*, 2017). Also, new features of cloud computing such as virtualization and VM migration introduces additional challenges to cloud security as studies report that the detection accuracy of anomaly detection system is degraded during the migration of VM from one host to another (Adamova *et al.*, 2014; Shirazi *et al.*, 2014). Performance degradation during VM migration can result because the cloud behaviour constantly changes during VM migration (Huang *et al.*, 2013), thereby making it difficult to maintain a consistent normal profile for anomaly detection. Therefore, providing effective security is crucial to the quality of service in cloud computing.

Intrusion detection is the process of monitoring events occurring in a system or network and analysing it for evidences of security incidents that breaches or presents impending threat of breach of system security policy or standard security practice (Scarfone and Mell, 2007). IDS can be classified into signature-based and anomaly detection depending on whether the kind of attack to be detected is known beforehand or unknown. The signature-based detection process captures activities in a network and compare them with a collection of attack signatures (Liao *et al.*, 2013). Anomaly

detection is concerned with the identification of events that appears to be anomalous with respect to normal system behaviour. Figure 1.1 shows the anomaly detection process. Anomaly detection has been well researched as a classical issue in the domain of intrusion detection and machine learning. Due to the recent advent of cloud computing with its new operational and technical features the problem of anomaly detection has risen again though well-established in classical computer system (Huang *et al.*, 2016). Anomaly detection techniques can be used for cloud to detect both known and unknown attacks at different levels such as IaaS, PaaS and SaaS (Modi *et al.*, 2013). The three major categories of anomaly-based IDS are: statistical-based, knowledge-based and machine learning (Garcia-Teodoro *et al.*, 2009). Anomaly detection using statistical technique involves observing the data of the current network profile and comparing it against the statistical profile previously created (Denning and Neumann, 1985). Knowledge-based techniques uses expert system for anomaly detection by employing a set of rules to classify a set of data (Anderson *et al.*, 1995). Machine learning techniques create a model that is used to classify the pattern analysed. Various anomaly detection techniques have been used for cloud based IDS such as Local Outlier Factor (Huang, *et al.*, 2013), PCA and K-Means clustering (Shirazi, *et al.*, 2014), Naïve Bayes and Random Forest (Idhammad, *et al.* 2018), Fuzzy C-Means clustering (Mehibs and Hashim (2018).

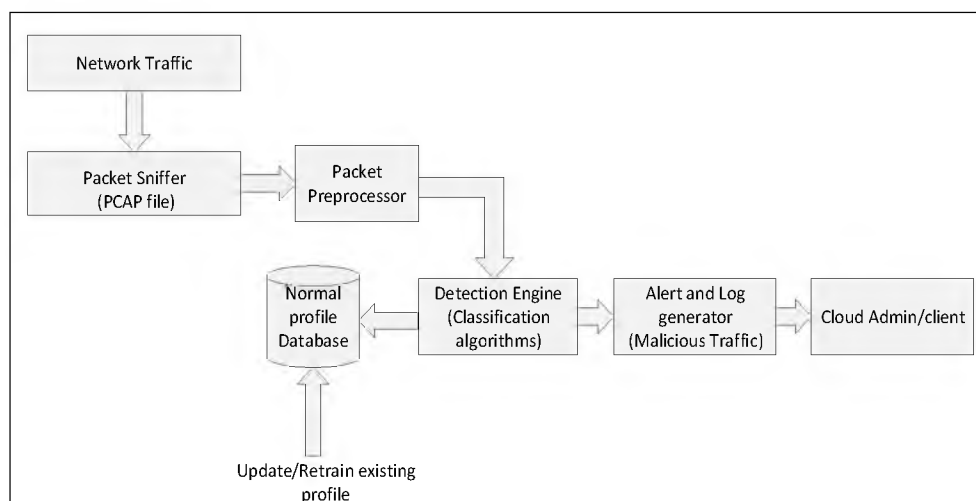


Figure 1.1 Anomaly detection process (Mishra *et al.*, 2017)



## 1.2 Problem Background

The distributed and multi-tenant nature of cloud computing makes it vulnerable to security threats. Cloud computing systems can be exposed to threats to its availability, data and the virtualized infrastructure which can be used as a launching pad for new attacks (Patel *et al.*, 2013). Cloud computing resources have always been primary target for DoS attack. Data of all costumers are kept at one geographical location and the Cloud service providers offers its services through the Internet. This makes Cloud data centres more vulnerable to attack. According to Cloud Security Alliance report Ko and Lee (2013), the number of incidents on Cloud environment has risen over the years. In fact, from 2009 to 2011, the number of Cloud vulnerability incidents has increased from 33 to 71, most likely due to the phenomenal growth in the Cloud services. There are three types of DoS attacks which are: volume-based attack, protocol attack, and application layer attack. According to Arbor Networks (2014), 61% of the organizations surveyed have faced volume-based attack, 24% faced protocol attack and 20 % faced application layer attack. Combination of more than one type of DDoS attack (multi-vector attack) is becoming a new trend among the attackers. According to Incapsula report Incapsula Inc. (2014), 81 % of the attacks are multi-vector attacks. Also, according to Incapsula Inc. (2014) SYN flood DDoS attack is the most common form of DDoS attack against the cloud infrastructure. Therefore it is essential to safeguard the cloud resources against DDoS attacks.

A number of research works have been conducted in cloud IDS both on host-based ( Kwon *et al.*, 2011; Alarifi and Wolthusen, 2013) and network based (Modi *et al.*, 2012; Xiong, *et al.*, 2014). On the detection methodology numerous research works have been conducted on anomaly detection such as (Shamsolmoali and Zareapoor, 2014; Shirazi, *et al.*, 2014) and the signature-based technique (Ficco *et al.*, 2012; Gupta *et al.*, 2013). The signature based detection approach is known for its accuracy in detecting known attack signature as long as the database is always up-to-date. The major drawback is its inability to detect unknown attacks or variation of known attack signatures (Osanaiye, *et al.*, 2016b). Anomaly detection on the other hand is more suited for detection of unknown attack but it suffers from high false alarm (Garcia-Teodoro, *et al.*, 2009; Singh *et al.*, 2016). The false alarm can be attributed to redundant and noisy features because they can have negative impact on the accuracy

of IDS (Aghdam and Kabiri, 2016). Also, the behaviour of the cloud network rapidly changes due to the heterogeneity of the clients using the service, the elastic nature of the services delivered (Dalmazo *et al.*, 2014; Xiong, *et al.*, 2014) and the dynamic nature of VM migration (Huang, *et al.*, 2013; Nagarajan and Perumal, 2015; Huang, *et al.*, 2016) which results in load fluctuation which affects the ability of the security monitoring system to detect attacks (Giannakou, *et al.*, 2015). VM migration adds difficulty to anomaly detection since it is based on large number of memory copy operations which may result in anomaly (Zhang *et al.*, 2013). Furthermore, coordinated attacks such as DDoS attacks which simultaneously occur in many network results in difficulties in detection of this attack (Zhou *et al.*, 2009). This difficulty is due to the coordinated nature of the attacks where attack are spread over multiple network. Therefore a collaborative effort is required to tackle the attack. For instance Smurf based DDoS uses a spoofed IP address to send ICMP request to large number of reflector host when the reflector host receives the request, they reply to the spoofed IP address thereby flooding it (Bhuyan *et al.*, 2015). The overall problem situation leading to detection inaccuracy in cloud IDS as shown in Figure 1.2 can be summarized into three points namely: redundant and insignificant features, dynamic cloud nature and distributed attacks.

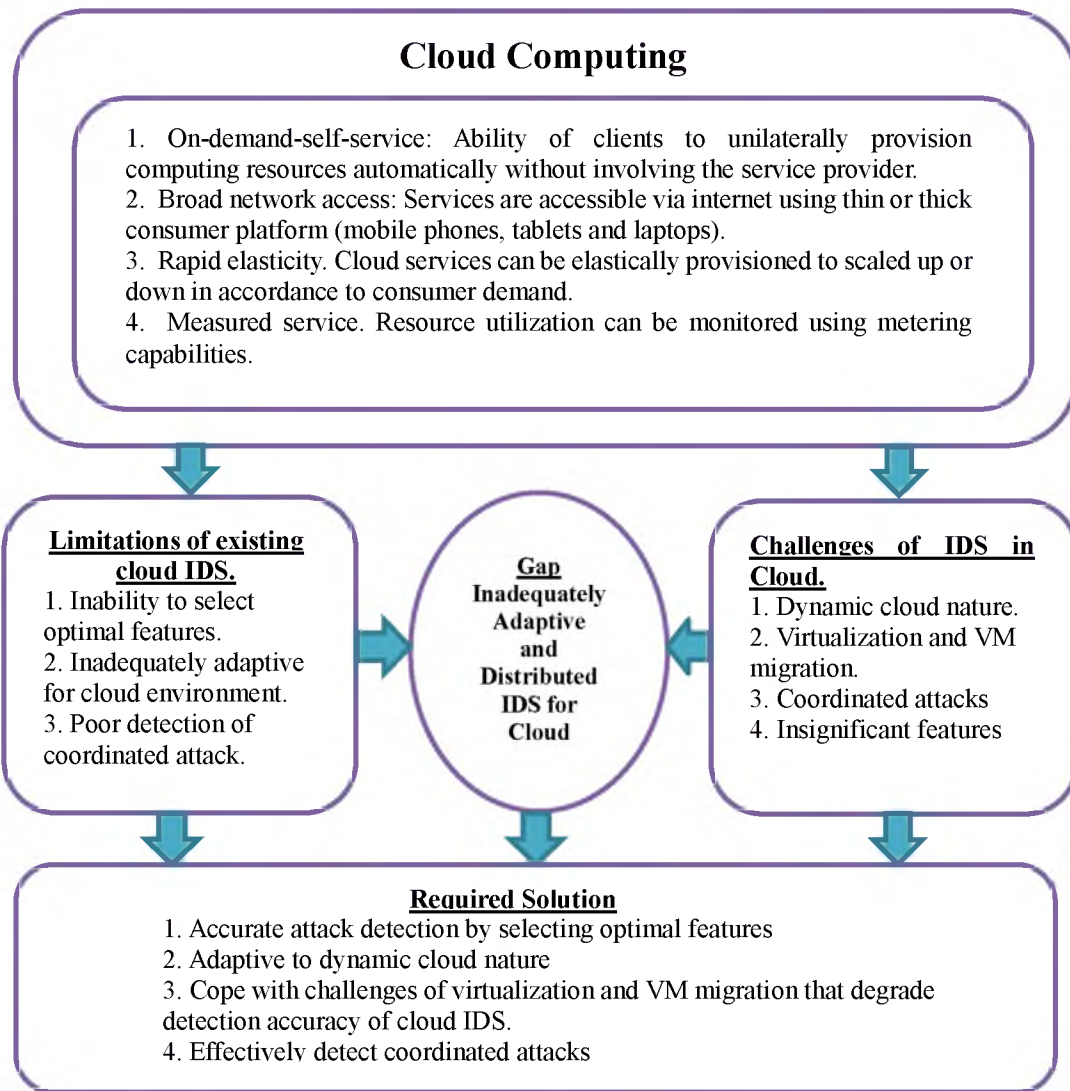


Figure 1.2 Scenario leading to the problem

**a. Insignificant and Redundant Features**

Insignificant and redundant features can have a negative impact on the accuracy of IDS hence it is necessary to remove the insignificant features to improve performance accuracy (Aghdam and Kabiri, 2016). A pre-processing component for choosing only significant features is an essential component for an effective IDS (Kannan *et al.*, 2012). The accuracy and efficiency of a machine learning based IDS is hinged on the features selected. Data that is explained with fewer features offers a better explanation of the processes underlying the data and therefore simplify the process of knowledge extraction (Kang and Kim, 2016). Feature selection is the process of eliminating redundant features in a dataset in order to improve classification

accuracy. In cloud IDS various feature selection techniques have been proposed as follows.

Osanaiye, *et al.* (2016a) used a combination of four filters (Information Gain, Gain Ratio, ReliefF and Chi-Square) to select features from NSL-KDD intrusion detection dataset. However, the filter approach may discard important features that are less informative on their own but more informative when combined with others (Chandrashekar and Sahin, 2014). Muthurajkumar *et al.* (2013) proposed a technique for feature selection using Rough Set Theory. Zhou *et al.* (2011) proposed a feature selection technique using multi-objective Particle Swarm Optimization. Kannan *et al.* (2012) proposed a technique for selecting significant features for intrusion detection using Genetic Algorithm. However these approaches are based on the heuristic search and the heuristic techniques cannot guide to optimal subset every time (Jensen and Shen, 2005). Besides, the selected features are based on traditional network datasets that may not capture and represent the cloud peculiarities.

#### **b. Dynamic Cloud Nature**

The behaviour of the cloud network changes due to the heterogeneity of the clients using the services, the elastic nature of the services delivered (Xiong, *et al.*, 2014) and dynamic nature of VM migration (Huang, *et al.*, 2013; Huang, *et al.*, 2016). Cloud computing enables virtual machines to be migrated from one node to another in order to provide efficient elasticity, load balancing and fault tolerance (Huang *et al.*, 2016). Despite being a key feature in cloud computing, VM migration poses security challenge to anomaly detection system. For instance legitimate migration can be misclassified as anomaly (Shirazi, *et al.*, 2014), since the cloud infrastructure settings may change a lot during migration (Huang, *et al.*, 2013). The normal behaviour of cloud applications may change owing to technical and non-technical reasons. Changes due to technical reasons involve cloud migrations and software/hardware upgrade while non-technical aspect could be due to seasonal events. Moreover, updating of IDS model is even more important during migration process since the infrastructure settings may change a lot during migration (Huang, *et al.*, 2013). In addition, VM migration adds difficulty to anomaly detection since it is based on large number of memory copy operations which may result in anomaly. Also anomaly detection

becomes challenging when VMs' are migrated to destination with different infrastructural settings such as network conditions, memory size and workload making the applications behave differently (Zhang, *et al.*, 2013). Due to the changing behaviour of cloud environment there is a need for cloud anomaly detection system to be adaptive.

Anomaly based intrusion detection creates a normal usage profile and a deviation from this profile is flagged as anomaly (Tsai, *et al.*, 2009). The normal profile creation can be static or adaptively updated in order to prevent false alarm caused by changing network pattern. The static anomaly-based IDS performs one-time training at the beginning of the IDS development to obtain a reference model which is subsequently used during detection stage to predict network behaviour while the adaptive IDS adopt a dynamic strategy to update the normal reference model (Zainal, 2011). According to Krishnan and Chatterjee (2012) an anomaly-based adaptive IDS should have a crucial surveillance component that monitors the normal profile for changes in order to update the normal profile when a change is observed. This surveillance component can help in reducing false alarm by adaptively updating the behavioural parameter. A number of research works have been proposed for adaptive cloud IDS as discussed in the following paragraph.

To address the performance degradation due to VM migration in cloud anomaly detection, an adaptive scheme for anomaly detection using Dimension-Reasoning Local Outlier Factor and Symbiotic Aggregate Approximation (DR-LOF-SAX) was proposed by Huang, *et al.*, (2016). DR-LOF was used to identify the data dimensions with significant impact on the anomaly. To further validate the result obtained from the Local Outlier Factor (LOF), Symbolic Aggregate Approximation (SAX) algorithm was used for comparison of the symbolic distance before and after migration. Small distance that is below the threshold will be dismissed as a false alarm meanwhile; a large distance indicates that the behaviour is an anomaly. Huang, *et al.* (2013) proposed an LOF based adaptive anomaly detection scheme that update the Reference Model each time test data is collected. However the limitations of these schemes is that they lack change tracking mechanism to determine when the changes in the normal profile is occurring and update the IDS model accordingly. In addition both schemes are host-based which will have a low visibility of the cloud network

activities. Krishnan and Chatterjee (2012) proposed an adaptive IDS framework for cloud computing that incorporates signature based and anomaly detection and a component for surveillance of normal behaviour changes in order to update the IDS Reference Model. Giannakou, *et al.* (2015) proposed an adaptive IDS that uses two component namely infrastructure monitoring probe and adaptation manager to monitor change and perform update. However these approaches are based on theoretical framework with no algorithmic technique been proposed for the monitoring component nor performing experimental test to validate the efficacy of the technique. Other related adaptive IDS proposed for cloud are the work of (Meng *et al.*, 2013; Chou and Wang, 2015; Toumi *et al.*, 2015; Chouhan and Hasbullah, 2016; Wahab *et al.*, 2017) . However these approaches are not suitable for the cloud environment as they do consider the effect of the critical cloud features such as VM migration which is reported to cause poor detection accuracy of cloud anomaly detection. Hence it is essential for cloud IDS to be able to cope with the challenges of VM migration for effective anomaly detection. In addition, the techniques to track changes prior to performing adaptive detection are not clearly specified.

In summary the limitations of the existing adaptive cloud-based IDS can be summarized as follows: no algorithmic technique has been proposed to track change in the normal profile of the data so as to update the IDS model accordingly. As earlier stated it is essential for an adaptive IDS to have change monitoring component to determine when the change in the normal profile is occurring in order to update the IDS model. This component can aid in reducing false alarm by updating the anomaly detection model parameters (Krishnan and Chatterjee 2012). Furthermore, most of the adaptive IDS proposed for cloud (Krishnan and Chatterjee, 2012; Huang, *et al.*, 2013; Meng, *et al.*, 2013; Chou and Wang, 2015; Toumi, *et al.*, 2015; Chouhan and Hasbullah, 2016; Wahab, *et al.*, 2017) are not adequate for the cloud environment as they do not consider the cloud peculiarities such as VM migration. The adaptive IDS (Huang, *et al.*, 2016) proposed to address VM migration issues is limited to host based which will have a low visibility of the cloud network.

### c. **Detection of distributed attacks**

The proliferation of distributed attacks such as DDoS and distributed port scan has brought forth challenges to centralized cloud-based IDS. A single IDS only monitors a single sub-network. Hence, it is unable to detect distributed attack accurately as it lacks the ability to link attack information from various sub-network (Zhou, *et al.*, 2009). To tackle the distributed nature of this attack, a collaborative defence mechanism is required.

In collaborative IDS, participating agents collaborate to detect distributed attacks by sharing attack information among themselves (Pérez *et al.*, 2013). In cloud computing both standalone and distributed approach have been adopted to detect distributed attacks. Under the standalone category a number of research works have been conducted (Bakshi and Dujodwala, 2010; Sahi *et al.*, 2017) however an isolated IDS cannot accurately detect coordinated attacks like DDoS (Singh *et al.*, 2016; Al Haddad *et al.*, 2016). Therefore this research focused on distributed approach. A distributed or collaborative IDS is comprised of many IDS over different sub networks or host that share alerts among each other to detect coordinated attacks. A collaborative IDS have the potentials of detecting attacks shared over several host or networks by linking attack evidence across several sub networks (Elshoush and Osman, 2011). Collaborative or distributed IDS consists of detection units and aggregation units. The detection units detect attacks and send to aggregation unit for aggregation (Patel *et al.*, 2013). The research work in distributed cloud IDS aimed at detecting distributed DDoS can be classified as signature-based proposed by Gul and Hussain, (2011) and Lo, *et al.*, (2010), however the limitation of the signature-based approach is that the attack information is sent from detection units to aggregation unit whenever new signature is found and the limitation of this is that zero-day attacks will not be detected. The anomaly detection approach proposed by Badis *et al.*, (2015) and Bharajwaja *et al.*, (2011) sends attack information from detection units to aggregation unit whenever anomaly is detected and this could lead to high false alarm. Because the anomaly detection approach is prone to false alarm (Garcia-Teodoro, *et al.*, 2009; Singh *et al.*, 2016) while the hybrid techniques (Man and Huh 2012; Singh *et al.*, 2016; Al Haddad *et al.*, 2016) send alert from detection units to aggregation units whenever attack signature is found or an anomaly is detected and this technique also inherits the limitations of both signature-based and anomaly detection. Further limitations of these

works is that they are not adequate for the cloud environment as the effect of certain cloud peculiarities such as VM migration which is reported to cause false alarm in cloud IDS (Shirazi *et al.*, 2014) is not investigated. Hence it is essential for cloud IDS to be able to cope with the challenges of VM migration for effective anomaly detection.

### **1.3 Problem Statement**

The behaviour of the cloud network rapidly changes due to the heterogeneity of the clients using the services, the elastic nature of the services delivered and the migration of VM from one host to another makes it difficult to create a consistent normal profile for anomaly detection. Existing research works on adaptive approach proposed to address performance degradation due to VM migration is limited to host-based which does not cover the effect of the entire migration picture in the cloud network. Furthermore an approach to monitor the changes in the normal profile of the data to determine when changes occur in order to update the IDS reference model accordingly has not been investigated using practical algorithmic approach.

In addition, the widespread of coordinated attacks such as DDoS has introduced challenge to centralized cloud-based IDS, hence a distributed IDS is required to tackle coordinated attacks. However, existing distributed cloud IDS do not address the appropriate time to share attack information among the nodes in the distributed IDS. In addition, they are not adequate for the cloud environment, as they do not address the peculiarities of cloud computing such as the issue of VM migration.

Furthermore, an IDS requires a pre-processing component for choosing only significant features. However, existing feature selection technique proposed for cloud IDS are based on heuristic search techniques which cannot guarantee optimal features. Besides the selected features are based on traditional network datasets that may not capture and represent the cloud peculiarities.



The Research Hypothesis is:

*Poor accuracy of cloud-based IDS due to dynamic nature of VM migration and distributed attacks can be improved using an adaptive and distributed cloud-based IDS. Detection accuracy for cloud-based IDS can be improved using feature selection technique.*

The research aims to address the following research questions:

- i. How to select optimal feature subset using feature selection technique in order to improve detection accuracy of cloud IDS.
- ii. How to determine the change pattern in the normal profile of the data and update IDS Reference Model according to change pattern.
- iii. How to determine when to share Reference Model for detecting attack among distributed IDS to improve detection accuracy of distributed attacks.

#### **1.4 Research Aim**

The aim of this research is to propose an adaptive and distributed cloud intrusion detection scheme that uses change point detection to determine when to update the IDS reference model and when to share attack information among nodes in the distributed IDS to improve detection accuracy of cloud IDS.

#### **1.5 Research Objectives**

The research aims to achieve the following objectives:

- i. To propose an enhanced hybrid Ant Colony Optimization and Correlation-based feature selection technique capable of selecting optimal feature set to improve accuracy of cloud IDS.

- ii. To design an adaptive cloud intrusion scheme that can improve detection accuracy of cloud IDS by monitoring change pattern in the normal profiles and update the IDS Reference Model accordingly.
- iii. To propose an enhanced adaptive and distributed cloud intrusion detection scheme that monitors the traffic volume of destination IP address from the various computing nodes in the distributed IDS to determine when to aggregate the reference models from the various computing nodes in order to improve detection accuracy of coordinated attacks.

## **1.6 Scope of Study**

The research is limited to the following:

- i. The effect of VM migration was only investigated on attacks such as port scanning and Distributed Denial of Service (DDoS). These attacks were also considered by other cloud researchers such as (Adamova, *et al.*, 2014; Shirazi, *et al.*, 2014). These attacks are considered because the cloud has suffered from several outages due to DDoS attacks. Therefore it is imperative to safeguard the cloud from such attack.
- ii. The study is limited to attack detection and does not consider pre-emptive actions.
- iii. Three machine learning techniques such as Stochastic Gradient Descent, Support Vector Machine, and Random Forest were investigated as a proof of concept for the proposed scheme.

## **1.7 Significance of the Research**

The research is significant from a theoretical and practical perspective. The motivation and the rationale for the research are:

- i. Cloud computing is confronted by an increasing number of cyber-attacks. Various security measures have been proposed to enable the detection of attacks in cloud computing. To enable cloud-based IDS effectively detect attacks when the normal reference model is frequently changing due to

dynamic cloud characteristics. It is essential to design adaptive IDS that can cope with the dynamic cloud nature.

- ii. The research findings is expected to offer better insight and contribute to the robustness of the cloud security.
- iii. Both practitioners and researchers can benefit from the research. As more data and applications from various sectors such as academia, government and industries are being migrated to the cloud therefore providing security measures to allay consumers worry about the security issues in cloud is crucial.

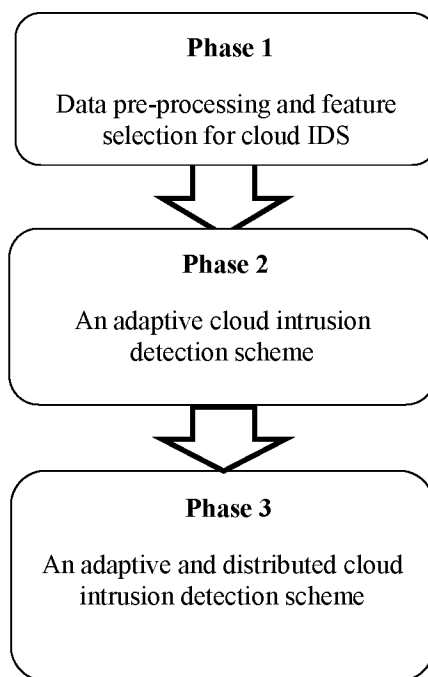


Figure 1.3 Phases in the design of the adaptive and distributed cloud intrusion detection scheme

## 1.8 Research Contributions

In this section, the contribution of the research is discussed. The research has three contributions as shown in Figure 1.4.

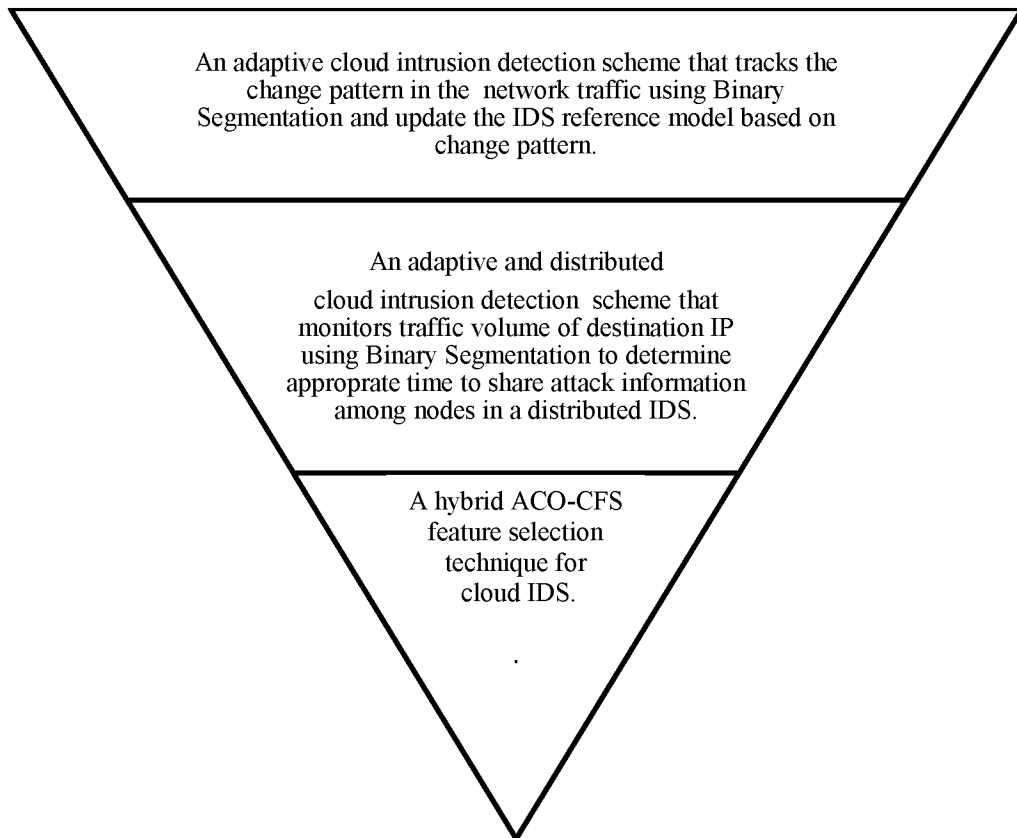


Figure 1.4 Research contribution

- i. The first contribution is an adaptive cloud intrusion detection scheme that uses change points detection to track the change pattern in the cloud data and perform Reference Model update according to the change pattern. The design is based on the philosophy of tracking change in statistical property of the data using Binary Segmentation change point detection algorithms and updating the IDS Reference Model periodically based on change pattern.
- ii. The second contribution of the research is an adaptive and distributed cloud intrusion detection scheme that monitors the traffic volume of destination IP using Binary Segmentation to determine the appropriate period to share attack information among nodes in the distributed IDS.
- iii. The third contribution is a hybrid Ant Colony Optimization and Correlation-based Feature Selection (ACO-CFS) technique for cloud IDS.

## **1.9 Thesis Organization**

The Thesis is comprised of seven chapters. Chapter 1 introduces the research. Chapter 2 provides a review on the current IDS in cloud computing and the issues that need to be addressed. Chapter 3 presents the research methodology. Chapter 4 discusses on feature selection and data pre-processing, Chapter 5 presents the adaptive cloud intrusion detection scheme, Chapter 6 presents the adaptive and distributed cloud intrusion detection scheme and chapter seven concludes the research.

## REFERENCES

- Abels, T., Dhawan, P. and Chandrasekaran, B. (2005). An overview of xen virtualization, *Dell Power Solutions*, 8, 109-111.
- Adamova, K., Schatzmann, D., Plattner, B. and Smith, P. (2014). Network anomaly detection in the cloud: The challenges of virtual service migration. *Proceedings of the IEEE International Conference on Communications (ICC)*. 10-14 June. Sydney, Australia : IEEE, 3770-3775.
- Aghdam, M. H., Ghasem-Aghaei, N. and Basiri, M. E. (2009) . Text feature selection using ant colony optimization. *Expert systems with applications*, 36(3), 6843-6853.
- Aghdam, M. H. and Kabiri, P. (2016) . Feature Selection for Intrusion Detection System Using Ant Colony Optimization. *International Journal of Network Security*, 18(3), 420-432.
- Al Haddad, Z., Hanoune, M. and Mamouni, A. (2016). A Collaborative Network Intrusion Detection System (C-NIDS) in Cloud Computing. *International Journal of Communication Networks and Information Security*, 8(3), 130-135.
- Alarifi, S. and Wolthusen, S. (2013) Anomaly Detection for Ephemeral Cloud IaaS Virtual Machines. *Proceedings of the International Conference on Network and System Security*. 3-4 June. Madrid, Spain: Springer, 321-335.
- Aljawarneh, S., Aldwairi, M. and Yassein, M. B. (2018). Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model. *Journal of Computational Science*, 25, 152-160.
- Amazon, (2017). AmazonElastic Compute Cloud (AmazonEC2), Available at :<https://aws.amazon.com/ec2/> (Accessed: July 2017).
- Amiri, F., Yousefi, M. R., Lucas, C., Shakery, A. and Yazdani, N. (2011). Mutual information-based feature selection for intrusion detection systems. *Journal of Network and Computer Applications*, 34(4), 1184-1199.
- Anderson, D., Lunt, T. F., Javitz, H., Tamaru, A. and Valdes, A. (1995). *Detecting unusual program behavior using the statistical component of the Next-generation Intrusion Detection Expert System (NIDES)*: SRI International, Computer Science Laboratory. Available at <http://www.csl.sri.com/papers/5sri/5sri.pdf>. (Accessed: June 2017).

- Arbor Networks, Inc (2014). Worldwide infrastructure security report volume IX [Online]. Available at <http://pages.arbornetworks.com/rs/arbor/images/WISR2014.pdf> (Accessed April 2019).
- Aslahi-Shahri, B., Rahmani, R., Chizari, M., Maralani, A., Eslami, M., Golkar, M., (2016). A hybrid method consisting of GA and SVM for intrusion detection system. *Neural computing and applications*, 27(6), 1669-1676.
- Avallone, S., Guadagno, S., Emma, D., Pescape, A. and Ventre, G. (2004). D-ITG Distributed Internet Traffic Generator. *Proceedings of the First International Conference on the Quantitative Evaluation of Systems, QEST 2004*. 27-30 September. Enschede, Holland: IEEE, 316-317.
- Badis, H., Doyen, G. and Khatoun, R. (2015). A Collaborative Approach for a Source based Detection of Botclouds. *Proceedings of the 2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)*. 11-15-May. Ottawa, Canada: IEEE, 906-909.
- Bakshi, A. and Dujodwala, Y. B. (2010). Securing Cloud from DDoS Attacks using Intrusion Detection System in Virtual Machine. *Proceedings of the 2010 Second International Conference on Communication Software and Networks, ICCSN'10*. 26-28 February. Singapore: IEEE, 260-264.
- Barham, P., Dragovic, B., Fraser, K., Hand, S., Harris, T., Ho, A., et al. (2003). Xen and the Art of Virtualization. *Proceedings of the nineteenth ACM symposium on Operating Systems Principles*. 19-22 October. Bolton Landing, New York: ACM, 164-177.
- Beitollahi, H. and Deconinck, G. (2012). Analyzing well-known countermeasures against distributed denial of service attacks. *Computer Communications*, 35(11), 1312-1332.
- Bharadwaja, S., Sun, W., Niamat, M. and Shen, F. (2011). Collabra: A Xen Hypervisor based Collaborative Intrusion Detection System. *Proceedings of the 2011 Eighth International Conference on Information Technology: New Generations*. 11-13 April. Las Vegas, Nevada: IEEE, 695-700.
- Bhuyan, M. H., Bhattacharyya, D. and Kalita, J. K. (2015). An empirical evaluation of information metrics for low-rate and high-rate DDoS attack detection. *Pattern Recognition Letters*, 51, 1-7.
- Binu, A. and Kumar, G. S. (2011). Virtualization Techniques: A Methodical Review of XEN and KVM. *Proceedings of the 2011 International Conference on Advances in Computing and Communications*. 22-24 July. Kochi, India: Springer, 399-410.

- Binu, S. and Misbahuddin, M. (2013). A survey of traditional and cloud specific security issues. *Proceedings of the 2013 International Symposium on Security in Computing and Communication*. 19-22 September. Bangalore, India: Springer, 110-129.
- Bohn, R. B., Messina, J., Liu, F., Tong, J. and Mao, J. (2011). NIST Cloud Computing Reference Architecture. *Proceedings of the 2011 IEEE World Congress on Services (SERVICES)*. 4-9 July. Washington, DC: IEEE, 594-596.
- Bonabeau, E., Dorigo, M. and Theraulaz, G. (1999). *Swarm intelligence: from natural to artificial systems*: Oxford university press.
- Borah, P. and Gupta, D. (2017). On Lagrangian Twin Parametric-Margin Support Vector Machine. *Proceedings of the 2017 International Conference on Next Generation Computing Technologies*. 30-31 October. Dehradun, India: Springer, 474-487.
- Bottou, L. (1998). Online learning and stochastic approximations. *On-line learning in neural networks*, 17(9), 142.
- Bottou, L. (2010). Large-Scale Machine learning with Stochastic Gradient Descent *Proceedings of the International Conference on Computational Statistics, COMPSTAT'2010*. 22-27 August. Paris, France: Springer, 177-186.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- Breunig, M. M., Kriegel, H.-P., Ng, R. T. and Sander, J. (2000). LOF: identifying density-based local outliers. *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*. 15-18 May. Dallas, Texas: ACM, 93-104.
- Chandrashekar, G. and Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1), 16-28.
- Chen, B., Chen, L. and Chen, Y. (2013). Efficient ant colony optimization for image feature selection. *Signal processing*, 93(6), 1566-1576.
- Chen, Y., Li, Y., Cheng, X.-Q. and Guo, L. (2006). Survey and Taxonomy of Feature Selection Algorithms in Intrusion Detection System. *Proceedings of the 2006 International Conference on Information Security and Cryptology*. 29 November-1st December. Beijing, China: Springer, 153-167.
- Chen, Y., Miao, D. and Wang, R. (2010). A rough set approach to feature selection based on ant colony optimization. *Pattern Recognition Letters*, 31(3), 226-233.



- Chen, X., and Yu, X. (2015). A collaborative intrusion prevention architecture for programmable network and SDN. *Computers and Security*, 58, 1-19.
- Chen, Z., Xu, G., Mahalingam, V., Ge, L., Nguyen, J., Yu, W., et al. (2016). A cloud computing based network monitoring and threat detection system for critical infrastructures. *Big Data Research*, 3, 10-23.
- Cheng, J., Yin, J., Liu, Y., Cai, Z. and Li, M. (2009). DDoS Attack Detection Algorithm using IP Address Features. *Proceedings of the 2009 International Workshop on Frontiers in Algorithmics*. 20-23 June. Hefei, China: Springer, 207-215.
- Chhikara, R. R., Sharma, P. and Singh, L. (2016). A hybrid feature selection approach based on improved PSO and filter approaches for image steganalysis. *International Journal of Machine Learning and Cybernetics*, 7(6), 1195-1206.
- Chiba, Z., Abghour, N., Moussaid, K. and Rida, M. (2016). A cooperative and hybrid network intrusion detection framework in cloud computing based on snort and optimized back propagation neural network. *Procedia Computer Science*, 83, 1200-1206.
- Chou, H.-H. and Wang, S.-D. (2015). An Adaptive Network Intrusion Detection Approach for the Cloud Environment. *Proceedings of the 2015 International Carnahan Conference on Security Technology (ICCST)*. 21-24 September. Taipei, Taiwan: Springer, 1-6.
- Chouhan, M. and Hasbullah, H. (2016). Adaptive Detection Technique for Cache-based Side Channel Attack using Bloom Filter for secure cloud. *Proceedings of the 2016 3rd International Conference on Computer and Information Sciences (ICCOINS)*. 15-17 August. Kuala-lumpur, Malaysia: IEEE, 293-297.
- Cisco. (2013 ). Cisco Annual Security Report Available at [https://www.cisco.com/web/offer/gist\\_ty2\\_asset/Cisco\\_2013\\_ASR.pdf](https://www.cisco.com/web/offer/gist_ty2_asset/Cisco_2013_ASR.pdf) (Accessed: August 2017).
- Clark, C., Fraser, K., Hand, S., Hansen, J. G., Jul, E., Limpach, C., et al. (2005). Live Migration of Virtual Machines. *Proceedings of the 2005 2nd Conference on Symposium on Networked Systems Design & Implementation*. 2-4 May. Boston, Massachusetts: Usenix , 273-286.
- Dahliyusmanto D., Abdul Hanan A. (2014). *Enhancing Anomalies Traffic Detection in Grid using Support Vector Machine*. PhD Thesis, Universiti Teknologi Malaysia, Skudai.
- Dalmazo, B. L., Vilela, J. P. and Curado, M. (2014). Online Traffic Prediction in the Cloud: A Dynamic Window Approach. *Proceedings of the 2014 International*

*Conference on Future Internet of Things and Cloud (FiCloud)*. 27-29 August. Barcelona, Spain: IEEE, 9-14.

- Denning, D. E. and Neumann, P. G. (1985). Requirements and model for IDES—a real-time intrusion detection expert system. *Document A005, SRI International*, 333.
- Dillon, T., Wu, C. and Chang, E. (2010). Cloud Computing: Issues and Challenges. *Proceedings of the 2010 24th IEEE International Conference on Advanced Information Networking and Applications (AINA)*. 20-23 April. Perth, West Australia: IEEE, 27-33.
- Dinesh Singh, D. P., Bhavesh Borisananiya, Chirag Modi. (2016). Collaborative IDS Framework for Cloud Computing. *International Journal of Network Security*, 18( 4), 669-709.
- Diro, A. A. and Naveen, C. (2018). Distributed attack detection scheme using deep learning approach for Internet of Things. *Future Generation Computer Systems*, 82, 761-768.
- Dorigo, M., Di Caro, G. and Gambardella, L. M. (1999). Ant algorithms for discrete optimization. *Artificial life*, 5(2), 137-172.
- Dorigo, M., Maniezzo, V. and Colomi, A. (1996). Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 26(1), 29-41.
- Eberhart, R. C. and Kennedy, J. (1995). A New Optimizer using Particle Swarm Theory. *Proceedings of the sixth international symposium on micro machine and human science*. 4-6 October. Nagoya, Japan: IEEE, 39-43.
- Elshoush, H. T. and Osman, I. M. (2011). Alert correlation in collaborative intelligent intrusion detection systems—A survey. *Applied Soft Computing*, 11(7), 4349-4365.
- Erdman, C. and Emerson, J. W. (2007). bcp: an R package for performing a Bayesian analysis of change point problems. *Journal of Statistical Software*, 23(3), 1-13.
- Estevez-Tapiador, J. M., Garcia-Teodoro, P. and Diaz-Verdejo, J. E. (2003). Stochastic Protocol Modeling for Anomaly based Network Intrusion Detection. *Proceedings of the 2003 First IEEE International Workshop on Information Assurance, IWIAS 2003*. 24 March. Darmstadt, Germany: IEEE 3-12.
- Fernandes, D. A., Soares, L. F., Gomes, J. V., Freire, M. M. and Inácio, P. R. (2014). Security issues in cloud environments: a survey. *International Journal of Information Security*, 13(2), 113-170.

- Ficco, M., Tasquier, L. and Aversa, R. (2013). Intrusion Detection in Cloud Computing. *Proceedings of the 2013 Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC)*. 28-30 October. Compiegne, France: IEEE, 276-283.
- Ficco, M., Venticinque, S. and Di Martino, B. (2012). Mosaic-based intrusion detection framework for cloud computing. *Proceeding of the OTM On the Move to Meaningful Internet Systems Confederated International Conferences: OTM 2012*. 10-14 September. Rome, Italy: Springer, 628-644.
- Fonseca, N. and Boutaba, R. (2015). *Cloud services, networking, and management*: John Wiley & Sons.
- Foster, I., Zhao, Y., Raicu, I. and Lu, S. (2008). Cloud Computing and Grid Computing 360-degree Compared. *Proceedings of the 2008 Grid Computing Environments Workshop, 2008. GCE'08*. 16 November. Austin, Texas: IEEE, 1-10.
- Fraiwan, L., Lweesy, K., Khasawneh, N., Wenz, H. and Dickhaus, H. (2012). Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier. *Computer methods and programs in biomedicine*, 108(1), 10-19.
- Garcia-Teodoro, P., Diaz-Verdejo, J., Maciá-Fernández, G. and Vázquez, E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. *computers & security*, 28(1), 18-28.
- Ghareb, A. S., Bakar, A. A. and Hamdan, A. R. (2016). Hybrid feature selection based on enhanced genetic algorithm for text categorization. *Expert Systems with Applications*, 49, 31-47.
- Genge, B., Haller, P. and Kiss, I. (2016). A framework for designing a resilient distributed intrusion detection systems for critical infrastructures. *International Journal of Critical Infrastructures*. 15, 3-11.
- Giannakou, A., Rilling, L., Pazat, J.-L., Majoreczyk, F. and Morin, C. (2015). Towards Self Adaptable Security Monitoring in IaaS Clouds. *Proceedings of the 2015 15th IEEE/ACM International Symposium on, Cluster, Cloud and Grid Computing (CCGrid)*. 4-7 May. Shenzhen, Hong Kong: IEEE, 737-740.
- Goldberg, D. E. and Holland, J. H. (1988). Genetic algorithms and machine learning. *Machine learning*, 3(2), 95-99.
- Google. (2017a). Google App Engine Documentation. Available at: <https://cloud.google.com/appengine/docs/> (Accessed: November 2017).

- Google. (2017b). Google Docs. Available at: <https://www.google.com/docs/about/> (Accessed: November 2017).
- Grobauer, B., Walloschek, T. and Stöcker, E. (2011). Understanding cloud computing vulnerabilities. *Security & privacy, IEEE*, 9(2), 50-57.
- Guha, S., Yau, S. S. and Buduru, A. B. (2016). Attack Detection in Cloud Infrastructures using Artificial Neural Network with Genetic Feature Selection. *Proceedings of the 2016 Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*. 8-12 August. Auckland, New Zealand: IEEE, 414-419.
- Gul, I. and Hussain, M. (2011). Distributed cloud intrusion detection model. *International Journal of Advanced Science and Technology*, 34(38), 71-81.
- Guller, M. (2015). *Big data analytics with Spark: A practitioner's guide to using Spark for large scale data analysis*: Springer.
- Gupta, B. and Badve, O. P. (2017). Taxonomy of DoS and DDoS attacks and desirable defense mechanism in a cloud computing environment. *Neural Computing and Applications*, 28(12), 3655-3682.
- Gupta, S. and Kumar, P. (2013). VM Profile based Optimized Network Attack Pattern Detection Scheme for DDoS Attacks in Cloud. *Proceedings of the 2013 International Symposium on Security in Computing and Communication*. 22-24 August. Mysore, India: Springer, 255-261.
- Hall, M. A. and Smith, L. A. (1997). Feature Subset Selection: A Correlation Based Filter Approach. *Proceedings of the 1997 International Conference on Neural Information Processing and Intelligent Information Systems*. Berlin: Springer, 855-858.
- Hall, M. A. and Smith, L. A. (1999). Feature selection for machine learning: comparing a correlation-based filter approach to the wrapper. *Proceedings of the 1999 FLAIRS conference*, 235-239.
- Hira, Z. M. and Gillies, D. F. (2015). A review of feature selection and feature extraction methods applied on microarray data. *Advances in bioinformatics*. Available at: <https://www.hindawi.com/journals/abi/2015/198363/abs/>. (Accessed: January 2018).
- Huang, T., Zhu, Y., Wu, Y., Bressan, S. and Dobbie, G. (2016). Anomaly detection and identification scheme for VM live migration in cloud infrastructure. *Future Generation Computer Systems*, 56, 736-745.

- Huang, T., Zhu, Y., Zhang, Q., Zhu, Y., Wang, D., Qiu, M., et al. (2013). An LOF-based Adaptive Anomaly Detection Scheme for Cloud Computing. *Proceedings of the 2013 IEEE 37th Annual Computer Software and Applications Conference Workshops (COMPSACW)*. 22-26 July. Japan: IEEE, 206-211.
- Hutchens, J. (2014). *Kali Linux network scanning cookbook*. Packt Publishing Ltd.
- Hwang, T., Shin, Y., Son, K. and Park, H. (2013). Design of a Hypervisor-based Rootkit Detection Method for Virtualized Systems in Cloud Computing Environments. *Proceedings of the 2013 AASRI Winter International Conference on Engineering and Technology*. 28-29 December. Saipan, USA: Atlantis Press, 27-32.
- Idhammad, M., Afdel, K. and Belouch, M. (2018). Distributed Intrusion Detection System for Cloud Environments based on Data Mining techniques. *Procedia Computer Science*, 127, 35-41.
- Incapsula Inc. (2014). 2013–2014 DoS threat landscape report [Online]. Available at <http://www.incapsula.com/blog/ddosthreat-landscape-report-2014.html>. (Accessed 22, April 2019).
- Javadpour, A., Abharian, S. K. and Wang, G. (2017). Feature Selection and Intrusion Detection in Cloud Environment Based on Machine Learning Algorithms. *Proceedings of 2017 IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC)* . 12-15 December, Guangzhou, China: IEEE, 1417-1421.
- Jensen, R. (2006). Performing feature selection with ACO *Swarm Intelligence in Data Mining* (pp. 45-73): Springer.
- Jensen, R. and Shen, Q. (2003). Finding Rough Set Reducts with Ant Colony Optimization. *Proceedings of the 2003 UK workshop on computational intelligence*, 15-22.
- Jensen, R. and Shen, Q. (2005). Fuzzy-rough data reduction with ant colony optimization. *Fuzzy sets and systems*, 149(1), 5-20.
- Jin, J., Fu, K. and Zhang, C. (2014). Traffic sign recognition with hinge loss trained convolutional neural networks. *IEEE Transactions on Intelligent Transportation Systems*, 15(5), 1991-2000.
- Igbe, O., Darwish, I. and Saadawi, T. (2016). Distributed Network Intrusion Detection System: An Artificial Immune System Approach. *IEEE First International Conference on Connected Health: Applications, Systems and*

- Engineering Technologies (CHASE)*. 27-29 June. Washington, DC, USA: IEEE, 101-106.
- Kanan, H. R., Faez, K. and Hosseinzadeh, M. (2007). Face Recognition System using Ant Colony Optimization-based Selected Features. *Proceedings of the 2007 IEEE Symposium on Computational Intelligence in Security and Defense Applications*. 12-15 December. Honolulu, Hawaii: IEEE, 57-62.
- Kang, S.-H. and Kim, K. J. (2016). A feature selection approach to find optimal feature subsets for the network intrusion detection system. *Cluster Computing*, 1-9.
- Kannan, A., Maguire, G. Q., Sharma, A. and Schoo, P. (2012). Genetic algorithm based feature selection algorithm for effective intrusion detection in cloud networks. *Proceedings of the 2012 IEEE 12th International Conference on Data Mining Workshops (ICDMW)*. 10 December. Brussels, Belgium: IEEE, 416-423.
- Kapil, D., Pilli, E. S. and Joshi, R. C. (2013). Live Virtual Machine Migration Techniques: Survey and Research Challenges. *Proceedings of the 2013 IEEE 3rd International Advance Computing Conference (IACC)*. 22-23 February. Ghaziabad, India: IEEE, 963-969.
- Kene, S. G. and Theng, D. P. (2015). A Review on Intrusion Detection Techniques for Cloud Computing and Security Challenges. *Proceedings of the 2015 2nd International Conference on Electronics and Communication Systems (ICECS)*. 26-27 February. Coimbatore, India: IEEE, 227-232.
- Kholidy, H. A. and Baiardi, F. (2012). CIDD: A Cloud Intrusion Detection Dataset for Cloud Computing and Masquerade Attacks. *Proceedings of the 2012 Ninth International Conference on Information Technology-New Generations*. 16-18 April. Las Vegas, Nevada: IEEE, 397-402.
- Khorshed, M. T., Ali, A. S. and Wasimi, S. A. (2012). A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing. *Future Generation computer systems*, 28(6), 833-851.
- Killick, R. and Eckley, I. (2014). changepoint: An R package for changepoint analysis. *Journal of Statistical Software*, 58(3), 1-19.
- Killick, R., Fearnhead, P. and Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, 107(500), 1590-1598.
- Kira, K. and Rendell, L. A. (1992). The Feature Selection Problem: Traditional methods and a new algorithm. *Proceedings of the 1992 Aaai*, 129-134.

- Ko R, Lee SSG (2013). Cloud computing vulnerability incidents: a statistical overview [Online] Available at [https://downloads.Cloud security alliance.org/initiatives/cvwwg/CSA\\_Whitepaper\\_Cloud\\_Computing\\_Vulnerability\\_Incidents.zip](https://downloads.Cloud%20security%20alliance.org/initiatives/cvwwg/CSA_Whitepaper_Cloud_Computing_Vulnerability_Incidents.zip). (Accessed 22 April, 2019).
- Kouzani, A. Z. and Nasireding, G. (2009). Multilabel classification by bch code and random forests. *International journal of recent trends in engineering*, 2(1), 113-116.
- Krishnan, D. and Chatterjee, M. (2012). An adaptive distributed intrusion detection system for cloud computing framework *Recent Trends in Computer Networks and Distributed Systems Security* (pp. 466-473): Springer.
- Kruegel, C. and Toth, T. (2000). A Survey on Intrusion Detection Systems. *Proceedings of the 2000 TU Vienna, Austria*.
- Kumar, K. and Stankowic, C. (2015). *VMware vSphere Essentials*: Packt Publishing Ltd.
- Kwon, H., Kim, T., Yu, S. J. and Kim, H. K. (2011). Self-similarity based lightweight intrusion detection method for cloud computing *Intelligent Information and Database Systems* (pp. 353-362): Springer.
- Lazarevic, A., Kumar, V. and Srivastava, J. (2005). Intrusion detection: A survey *Managing Cyber Threats* (pp. 19-78): Springer.
- Lee, K., Kim, J., Kwon, K. H., Han, Y. and Kim, S. (2008). DDoS attack detection method using cluster analysis. *Expert Systems with Applications*, 34(3), 1659-1665.
- Liao, H.-J., Lin, C.-H. R., Lin, Y.-C. and Tung, K.-Y. (2013). Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications*, 36(1), 16-24.
- Ließ, M., Glaser, B. and Huwe, B. (2012). Uncertainty in the spatial prediction of soil texture: Comparison of regression tree and Random Forest models. *Geoderma*, 170, 70-79.
- Lin, J., Keogh, E., Wei, L. and Lonardi, S. (2007). Experiencing SAX: a novel symbolic representation of time series. *Data Mining and knowledge discovery*, 15(2), 107-144.
- Liu, H. and Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on Knowledge and Data Engineering*, 17(4), 491-502.

- Lo, C.-C., Huang, C.-C. and Ku, J. (2010). A Cooperative Intrusion Detection System Framework for Cloud Computing Networks. *Proceedings of the 2010 39th international conference on Parallel processing workshops (ICPPW)*. 13-16 September. San Diego, California: IEEE, 280-284.
- Lunt, T. F. (1989). *Automated audit trail analysis and intrusion detection: A survey*: SRI International, Business Intelligence Program.
- Lyon, G. F. (2009). *Nmap network scanning: The official Nmap project guide to network discovery and security scanning*: Insecure.
- Maiti, S., Garai, C. and Dasgupta, R. (2015). A Detection Mechanism of DoS Attack using Adaptive NSA Algorithm in Cloud Environment. *Proceedings of the 2015 International Conference on Computing, Communication and Security (ICCCS)*. 4-5 December. Pamplousses, Mauritius: IEEE, 1-7.
- Man, N. D. and Huh, E.-N. (2012). A Collaborative Intrusion Detection System Framework for Cloud Computing. *Proceedings of the International Conference on IT Convergence and Security*. 91-109.
- Mehibs, S. M. and Hashim, S. H. (2018). Proposed Network Intrusion Detection System Based on Fuzzy c Mean Algorithm in Cloud Computing Environment. *Journal of University of Babylon*, 26(2), 27-35.
- Mell, P. and Grance, T. (2011). The NIST definition of cloud computing. Available at <http://faculty.winthrop.edu/domanm/csci411/Handouts/NIST.pdf> (Accessed: February, 2016).
- Meng, Y., Li, W. and Kwok, L.-F. (2013). Towards Adaptive False Alarm Reduction using Cloud as a Service. *Proceedings of the 2013 8th International Conference on Communications and Networking in China (CHINACOM)*. 14-16 August. Guilin, China: IEEE, 420-425.
- Meng, X., Bradley, J., Yavuz, B., Sparks, E., Venkataraman, S., Liu, D., et al. (2016). Mllib: Machine learning in apache spark. *The Journal of Machine Learning Research*, 17(1), 1235-1241.
- Mishra, P., Pilli, E. S., Varadharajan, V. and Tupakula, U. (2017). Intrusion detection techniques in cloud environment: A survey. *Journal of Network and Computer Applications*, 77, 18-47.
- Modi, C., Patel, D., Borisaniya, B., Patel, H., Patel, A. and Rajarajan, M. (2013). A survey of intrusion detection techniques in cloud. *Journal of Network and Computer Applications*, 36(1), 42-57.



- Modi, C., Patel, D., Borisanya, B., Patel, A. and Rajarajan, M. (2012). A Novel Framework for Intrusion Detection in Cloud. *Proceedings of the Fifth International Conference on Security of Information and Networks*. 25-27 October. Jaipur, India.: IEEE, 67-74.
- Muthurajkumar, S., Kulothungan, K., Vijayalakshmi, M., Jaisankar, N. and Kannan, A. (2013). A Rough Set based Feature Selection Algorithm for Effective Intrusion Detection in Cloud Model. *Proceedings of the International Conference on Advances in Communication, Network and Computing*. 8-13
- Nagarajan, P. and Perumal, G. (2015). A Neuro Fuzzy Based Intrusion Detection System for a Cloud Data Center Using Adaptive Learning. *Cybernetics and Information Technologies*, 15(3), 88-103.
- Nam, C. F., Aston, J. A. and Johansen, A. M. (2012). Quantifying the uncertainty in change points. *Journal of Time Series Analysis*, 33(5), 807-823.
- Novakovic, J. (2009). Using Information Gain Attribute Evaluation to Classify Sonar Targets. *Proceedings of the 2009 17th Telecommunications forum TELFOR*, 1351-1354.
- Osanaiye, O., Cai, H., Choo, K.-K. R., Dehghantanha, A., Xu, Z. and Dlodlo, M. (2016a). Ensemble-based multi-filter feature selection method for DDoS detection in cloud computing. *EURASIP Journal on Wireless Communications and Networking*, 2016(1), 130.
- Osanaiye, O., Choo, K.-K. R. and Dlodlo, M. (2016b). Distributed Denial of Service (DDoS) Resilience in Cloud: Review and Conceptual Cloud DDoS Mitigation Framework. *Journal of Network and Computer Applications*. 67(1), 147-165.
- Patel, A., Taghavi, M., Bakhtiyari, K. and JúNior, J. C. (2013). An intrusion detection and prevention system in cloud computing: A systematic review. *Journal of network and computer applications*, 36(1), 25-41.
- Pawlak, Z. (1998). Rough set theory and its applications to data analysis. *Cybernetics & Systems*, 29(7), 661-688.
- Pearce, M., Zeadally, S. and Hunt, R. (2013). Virtualization: Issues, security threats, and solutions. *ACM Computing Surveys (CSUR)*, 45(2), 17.
- Pérez, M. G., Mármol, F. G., Pérez, G. M. and Gómez, A. F. S. (2013). RepCIDN: A reputation-based collaborative intrusion detection network to lessen the impact of malicious alarms. *Journal of network and systems management*, 21(1), 128-167.
- Pham, L. H., Albanese, M. and Venkatesan, S. (2016). A quantitative risk assessment framework for adaptive Intrusion Detection in the cloud. *Proceedings of the*

2016 IEEE Conference on Communications and Network Security (CNS). 17-19 October. Philadelphia, Pennsylvania: IEEE, 489-497.

- Pham, N. T., Foo, E., Suriadi, S., Jeffrey, H. and Lahza, H. F. M. (2018). Improving performance of intrusion detection system using ensemble methods and feature selection. *Proceedings of the Australasian Computer Science Week Multiconference*. 29 January- 2 February. Brisbane, Australia: ACM, 2.
- Poston, A., (2017). The ultimate cheat sheet on IDS, IPS and HIDS. Available at <https://www.uzado.com/blog/the-ultimate-cheat-sheet-on-ids-ips-and-hids>. [Accessed: 21 June, 2017].
- Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1(1), 81-106.
- Rashedi, E., Nezamabadi-Pour, H. and Saryazdi, S. (2009). GSA: a gravitational search algorithm. *Information sciences*, 179(13), 2232-2248.
- Reeves, J., Chen, J., Wang, X. L., Lund, R. and Lu, Q. Q. (2007). A review and comparison of changepoint detection techniques for climate data. *Journal of Applied Meteorology and Climatology*, 46(6), 900-915.
- Ringberg, H., Soule, A., Rexford, J. and Diot, C. (2007). Sensitivity of PCA for traffic anomaly detection. *ACM SIGMETRICS Performance Evaluation Review*, 35(1), 109-120.
- Robnik-Šikonja, M. and Kononenko, I. (2003). Theoretical and empirical analysis of ReliefF and RReliefF. *Machine learning*, 53(1-2), 23-69.
- Rokach, L. and Maimon, O. (2005). Decision trees *Data mining and knowledge discovery handbook* (pp. 165-192): Springer.
- Sadeghzadeh, M. and Teshnehlab, M. (2010). Correlation-based Feature Selection using Ant Colony Optimization. *World Academy of Science, Engineering and Technology*, 64, 497-502.
- Sahi, A., Lai, D., Li, Y. and Diykh, M. (2017). An efficient DDoS TCP flood attack detection and prevention system in a cloud environment. *IEEE Access*, 5, 6036-6048.
- Salesforce. (2017). Available at <https://www.salesforce.com/my/> (Accessed: August 2017).
- Sanflippo. (2005). HPING3 Available at: <http://www.hping.org/hping3.html>. (Accessed: September 2017)
- Scarfone, K. and Mell, P. (2007). Guide to intrusion detection and prevention systems (idps). *NIST special publication*, 800(2007), 94.

- Sekar, R., Gupta, A., Frullo, J., Shanbhag, T., Tiwari, A., Yang, H., et al. (2002). Specification-based anomaly detection: a new approach for detecting network intrusions. *Proceedings of the 9th ACM conference on Computer and communications security*. 18-22 November. Washington, DC: ACM, 265-274.
- Shamsolmoali, P. and Zareapoor, M. (2014). Statistical-based Filtering System against DDOS Attacks in Cloud Computing. *Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. 24-27 September. New Delhi, India: IEEE, 1234-1239.
- Shang, W., Huang, H., Zhu, H., Lin, Y., Qu, Y. and Wang, Z. (2007). A novel feature selection algorithm for text categorization. *Expert Systems with Applications*, 33(1), 1-5.
- Sharma, P., Sood, S. K. and Kaur, S. (2011). Security issues in cloud computing. *High Performance Architecture and Grid Computing*, 36-45.
- Shirazi, N.-u.-H., Simpson, S., Marnerides, A. K., Watson, M., Mauthe, A. and Hutchison, D. (2014). Assessing the Impact of Intra-cloud Live Migration on Anomaly Detection. *Proceedings of the 2014 IEEE 3rd International Conference on Cloud Networking (CloudNet)*. 8-10 October. Luxembourg: IEEE, 52-57.
- Singh, A. and Chatterjee, K. (2017). Cloud security issues and challenges: A survey. *Journal of Network and Computer Applications*, 79, 88-115.
- Singh, D., Patel, D., Borisaniya, B. and Modi, C. (2016). Collaborative ids framework for cloud. *International Journal of Network Security*, 18(4), 699-709.
- Singh, K., Guntuku, S. C., Thakur, A. and Hota, C. (2014). Big data analytics framework for peer-to-peer botnet detection using random forests. *Information Sciences*, 278, 488-497.
- Stiitzle, T. and Hoos, H., (1997). The MAX-MIN ant system and local search for the traveling salesman problem. *Proceedings of 1997 IEEE International Conference on Evolutionary Computation (ICEC '97)*.13-16 April. Indianapolis, USA: IEEE, 309-314.
- Subashini, S. and Kavitha, V. (2011). A survey on security issues in service delivery models of cloud computing. *Journal of network and computer applications*, 34(1), 1-11.
- Tavallaee, M., Bagheri, E., Lu, W. and Ghorbani, A. A. (2009). A detailed analysis of the KDD CUP 99 data set. *Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications. CISDA*. 8-10 July. Ottawa, Canada: IEEE, 1-6.

- Toumi, H., Talea, A., Marzak, B., Eddaoui, A. and Talea, M. (2015). Cooperative trust framework for cloud computing based on mobile agents. *International Journal of Communication Networks and Information Security*, 7(2), 106.
- Tsai, C.-F., Hsu, Y.-F., Lin, C.-Y. and Lin, W.-Y. (2009). Intrusion detection by machine learning: A review. *Expert Systems with Applications*, 36(10), 11994-12000.
- Vapnik, V. N. and Vapnik, V. (1998). *Statistical learning theory* (Vol. 1): Wiley New York.
- Verma, A. and Kaushal, S. (2011). Cloud computing security issues and challenges: a survey. *Advances in Computing and Communications*, 445-454.
- Vieira, S. M., Mendonça, L. F., Farinha, G. J. and Sousa, J. M. (2013). Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients. *Applied Soft Computing*, 13(8), 3494-3504.
- VMware, Inc (2009a). Introduction to VMware vSphere. Available at [https://www.vmware.com/pdf/vsphere4/r40/vsp\\_40\\_intro\\_vs.pdf](https://www.vmware.com/pdf/vsphere4/r40/vsp_40_intro_vs.pdf). (Accessed: 27 Aug, 2016).
- VMware, Inc.(2009b). vSphere Basic System Administration. Available at [https://www.vmware.com/pdf/vsphere4/r40/vsp\\_40\\_admin\\_guide.pdf](https://www.vmware.com/pdf/vsphere4/r40/vsp_40_admin_guide.pdf) (Accessed 27 August, 2016).
- Wahab, O. A., Bentahar, J., Otrok, H. and Mourad, A. (2017). I Know You Are Watching Me: Stackelberg-Based Adaptive Intrusion Detection Strategy for Insider Attacks in the Cloud. *Proceedings of the 2017 IEEE International Conference on Web Services (ICWS)*. 25-30 June. Honolulu, Hawaii: IEEE, 728-735.
- Wang, L., Zhu, J. and Zou, H. (2006). The doubly regularized support vector machine. *Statistica Sinica*, 16(2), 589.
- Wu, Q. and Shao, Z. (2005). Network Anomaly Detection using Time Series Analysis. *Proceedings of the 2005 Joint International Conference on Autonomic and Autonomous Systems and International Conference on Networking and Services - (icas-isns'05)*. 23-28 October. Papeete, French Polynesia: IEEE, 42.
- Wyld, D. C. (2010). Risk in the clouds?: Security issues facing government use of cloud computing. *Innovations in computing sciences and software engineering*, 7-12.

- Xiang, C., Yong, P. C. and Meng, L. S. (2008). Design of multiple-level hybrid classifier for intrusion detection system using Bayesian clustering and decision trees. *Pattern Recognition Letters*, 29(7), 918-924.
- Xiong, W., Hu, H., Xiong, N., Yang, L. T., Peng, W.-C., Wang, X., (2014). Anomaly secure detection methods by analyzing dynamic characteristics of the network traffic in cloud communications. *Information Sciences*, 258, 403-415.
- Xu, J., Yan, J., He, L., Su, P. and Feng, D. (2010). CloudSEC: A cloud architecture for composing collaborative security services. *Proceedings of the 2010 IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom)*. 30 November- 3 December. Indianapolis, Indiana: IEEE, 703-711.
- Yang, Y. and Pedersen, J. O. (1997). A comparative study on feature selection in text categorization. *Proceedings of the 1997 Icml*, 412-420.
- Ye, N., Emran, S. M., Chen, Q. and Vilbert, S. (2002). Multivariate statistical analysis of audit trails for host-based intrusion detection. *IEEE Transactions on Computers*, 51(7), 810-820.
- Yeung, D.-Y. and Ding, Y. (2003). Host-based intrusion detection using dynamic and static behavioral models. *Pattern recognition*, 36(1), 229-243.
- Yiakopoulos, C. T., Gryllias, K. C. and Antoniadis, I. A. (2011). Rolling element bearing fault detection in industrial environments based on a K-means clustering approach. *Expert Systems with Applications*, 38(3), 2888-2911.
- Zainal, A. (2011). *An Adaptive Intrusion Detection Model for Dynamic Network Traffic Patterns Using Machine Learning Techniques*. PhD Thesis, Universiti Teknologi Malaysia, Skudai.
- Zainal, A., Maarof, M. A., Shamsuddin, S. M. and Abraham, A. (2012). Design of Adaptive IDS with Regulated Retraining Approach *Advanced Machine Learning Technologies and Applications* (pp. 590-600): Springer.
- Zaharia, M., Xin, R. S., Wendell, P., Das, T., Armbrust, M., Dave, A., (2016). Apache spark: a unified engine for big data processing. *Communications of the ACM*, 59(11), 56-65.
- Zhang, J. (2013). Advancements of outlier detection: A survey. *ICST Transactions on Scalable Information Systems*, 13(1), 1-26.
- Zhang, Q., Wu, Y., Huang, T. and Zhu, Y. (2013). An intelligent anomaly detection and reasoning scheme for VM live migration via cloud data mining. *Proceedings of the 2013 IEEE 25th International Conference on Tools with Artificial Intelligence (ICTAI)*. 4-6 November. Herndon, Virginia: IEEE, 412-419.

- Zhang, T. (2004). Solving large scale linear prediction problems using stochastic gradient descent algorithms. *Proceedings of the twenty-first international conference on Machine learning*, 116.
- Zhao, Y. and Zhang, Y. (2008). Comparison of decision tree methods for finding active objects. *Advances in Space Research*, 41(12), 1955-1959.
- Zhou, C. V., Leckie, C. and Karunasekera, S. (2009). Decentralized multi-dimensional alert correlation for collaborative intrusion detection. *Journal of Network and Computer Applications*, 32(5), 1106-1123.
- Zhou, L.-H., Liu, Y.-H. and Chen, G.-L. (2011). A feature selection algorithm to intrusion detection based on cloud model and multi-objective particle swarm optimization. *Proceedings of the 2011 Fourth International Symposium on Computational Intelligence and Design (ISCID)*. 28-30. Hangzhou, China: IEEE, 182-185.
- Zhu, B. and Ghorbani, A. A. (2006). Alert correlation for extracting attack strategies. *International Journal of Network Security*, 3(3), 244-258.
- Zinkevich, M., Weimer, M., Li, L. and Smola, A. J. (2010). Parallelized stochastic gradient descent. *Proceedings of the 2010 Advances in neural information processing systems*, 2595-2603.

## LIST OF PUBLICATIONS

### Journal with Impact Factor

1. **Ibrahim, N. M.,** Zainal, A. (2017). A feature selection technique for Cloud IDS using Ant Colony Optimization and Decision Tree. *Advanced Science Letters*, 23(9), pp. 9163-9169, <https://doi.org/10.1166/asl.2017.10045> (**Q2, IF: 1.25**).

Status: Published.

### Indexed Journal

1. **Ibrahim, N. M.,** Zainal, A. (2018). Adaptive cloud intrusion detection model. *International Journal of Swarm Intelligence Research*, 5(5) (**Indexed by ISI Web of Science**).

Status: Published.

2. **Ibrahim, N. M.,** Zainal, A. Distributed Cloud Intrusion Detection Scheme. *International Journal of Distributed Systems and Technologies* (**Indexed by ISI Web of Science, Q3**)

Status: Accepted with revisions.

- 3 **Ibrahim, N. M.,** Zainal, A. (2018). Intrusion detection technique in Cloud Computing: A review. *International Journal of Computer Applications*, 179(12), pp. 0975-8887, (**Indexed by EBSCO**).

Status: Published.

### Indexed Conference Proceedings

1. **Ibrahim, N. M.,** Zainal, A. (2018). A Model for Adaptive and Distributed Intrusion Detection for Cloud Computing. Presented in 2018 *IEEE 7<sup>th</sup> ICT International Student Project Conference*, Mahidol, Thailand.