



**PASSIVE CLIENT-CENTRIC ROGUE ACCESS POINT
DETECTION FRAMEWORK FOR WI-FI HOTSPOTS**

NAZRUL MUHAIMIN BIN AHMAD

DOCTOR OF PHILOSOPHY

2018



Faculty of Information and Communication Technology

**PASSIVE CLIENT-CENTRIC ROGUE ACCESS POINT DETECTION
FRAMEWORK FOR WI-FI HOTSPOTS**

Nazrul Muhaimin bin Ahmad

Doctor of Philosophy

2018

**PASSIVE CLIENT-CENTRIC ROGUE ACCESS POINT DETECTION
FRAMEWORK FOR WI-FI HOTSPOTS**

NAZRUL MUHAIMIN BIN AHMAD

**A thesis submitted
in fulfillment of the requirements for the degree of Doctor of Philosophy**

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2018

DECLARATION

I declare that this thesis entitled “Passive Client-centric Rogue Access Point Detection Framework for Wi-Fi Hotspots” is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature :

Name : Nazrul Muhaimin Bin Ahmad

Date :

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in term of scope and quality for the award of Doctor of Philosophy.

Signature :

Supervisor Name : Assoc Prof. Dr Mohd Faizal Bin Abdollah

Date :

DEDICATION

I dedicate my thesis to my beloved family, especially

To my lovely wife:
Azwina Binti Ibrahim

To my wonderful daughters:
Adlina and Hannah Mariah

To my late mother:
Fatimah Binti Abdullah

To my father:
Ahmad Bin Abdul Aziz

ABSTRACT

The proliferation of Wi-Fi hotspots in public places provides seamless Internet connectivity anywhere at any time to the wireless clients. Although many hotspots are often unprotected, unmanaged and unencrypted, this does not prevent the clients from actively connecting to the network. The underlying problem is that the network Access Point (AP) is always trusted. The adversary can impersonate a legitimate AP by setting up a rogue AP to commit espionage and to launch evil-twin attack, session hijacking, and eavesdropping. To aggravate the threats, existing detection solutions are ill-equipped to safeguard the client against rogue AP. Infrastructure-centric solutions are heavily relied on the deployment of sensors or centralized server for rogue AP detection, which are limited, expensive and rarely to be implemented in hotspots. Even though client-centric solutions offer threat-aware protection for the client, but the dependency of the existing solutions on the spoofable contextual network information and the necessity to be associated with the network makes those solutions are not viable for the hotspot's client. Hence, this work proposes a framework of passive client-centric rogue AP detection for hotspots. Unlike existing solutions, the key idea is to piggyback AP-specific and network-specific information in IEEE 802.11 beacon frame that enables the client to perform the detection without authentication and association to any AP. Based on the spatial fingerprints included in the broadcasted information from the APs in the vicinity of the client, this work discloses a novel concept that enables the rogue AP detection via the client's ability to self-colocalize and self-validate its own position in the hotspot. The legitimacy of the APs in the hotspot, in this view, lies in the fact that the correct matching between the Received Signal Strength Indicator (RSSI) measurements at the client and pre-recorded fingerprints is attainable when the beacons are transmitted only from the legitimate APs. Hence, any anomalousness in AP's beacon frame or any attempt to replay the legitimate AP's beacon frame from different location can be detected and classified as rogue AP threats. Through experiments in real environment, the results demonstrate that with proper algorithm selection and parameters tuning, the rogue AP detection framework can achieve over 90% detection accuracy in classifying the absence and presence of rogue AP threats in the hotspot.

ABSTRAK

Kepesatan pertumbuhan kawasan khas Wi-Fi di tempat awam menyediakan sambungan Internet tanpa sempadan di mana sahaja pada bila-bila masa kepada pelanggan wayarles. Walaupun banyak kawasan khas sering tidak dilindungi, tidak terurus dan tidak dienkripsi, ini tidak menghalang pelanggan dari menyambung secara aktif kepada rangkaian. Masalah asas adalah bahawa titik capaian sentiasa dipercayai. Pihak musuh boleh memasang titik capaian palsu untuk menyamar sebagai titik capaian yang sah bagi melakukan pengintipan dan untuk melancarkan serangan kembar-jelik, sesi rampasan, dan mencuri dengar. Lebih memperburukkan lagi ancaman, penyelesaian pengesanan yang sedia ada tidak dilengkapi untuk melindungi pelanggan daripada titik capaian palsu. Penyelesaian berpaksikan infrastruktur yang sangat bergantung pada penggunaan penerima atau pelayan terpusat untuk pengesanan titik capaian palsu adalah terbatas, mahal dan jarang untuk dilaksanakan di kawasan khas. Walaupun penyelesaian berpaksikan pelanggan menawarkan perlindungan ancaman terhadap pelanggan, tetapi kebergantungan penyelesaian yang sedia ada pada konteks maklumat rangkaian yang dapat diserang dan keperluan untuk bersekutu dengan rangkaian membuat penyelesaian tersebut tidak berdaya maju untuk pelanggan kawasan khas. Oleh itu, kajian ini mencadangkan satu rangka kerja pengesanan titik capaian palsu yang pasif serta berpaksikan pelanggan untuk kawasan khas. Tidak seperti penyelesaian yang sedia ada, idea utama adalah untuk menggendong maklumat khusus titik capaian dan rangkaian di dalam IEEE 802.11 bingkai suar yang membolehkan pelanggan untuk melaksanakan pengesanan tanpa pengesahan dan penyekutuan kepada mana-mana titik capaian. Berdasarkan cap jari ruang yang terdapat di dalam maklumat yang disiarkan dari titik-titik capaian di dalam persekitaran pelanggan, kajian ini mendedahkan satu konsep baru yang membolehkan pengesanan titik capaian palsu melalui keupayaan pelanggan untuk melakukan swa-penempatan bersama dan swa-pengesahsahihan kedudukan diri sendiri di dalam kawasan khas. Kesahihan titik capaian di dalam kawasan khas, dalam pandangan ini, terletak pada hakikat bahawa padanan yang betul boleh dicapai antara ukuran penunjuk kekuatan isyarat yang diterima (RSSI) pada pelanggan dan cap jari yang telah dirakam apabila suar dipancarkan hanya dari titik-titik capaian yang sah sahaja. Oleh itu, mana-mana ciri yang beranomali di dalam bingkai suar titik capaian atau apa-apa percubaan untuk memainkan semula bingkai suar titik capaian yang sah dari lokasi yang berbeza boleh dapat dikesan dan diklasifikasikan sebagai ancaman titik capaian palsu. Melalui eksperimen dalam persekitaran yang sebenar, keputusan menunjukkan bahawa dengan pilihan algoritma yang betul dan penalaan parameter yang tepat, rangka kerja pengesanan titik capaian palsu ini boleh mencapai lebih 90% ketepatan pengesanan didalam mengklasifikasikan ketiadaan dan kehadiran ancaman titik capaian palsu dalam kawasan khas.

ACKNOWLEDGEMENTS

In the name of Allah, the Most Gracious and the Most Merciful. First and foremost, I would like to thank Allah Almighty, who made me capable to complete this thesis throughout those difficult years. I am indebted to my supervisors, Assoc. Prof. Mohd Faizal Abdollah and Dr Robiah Yusof for their excellent supervision, guidance, endless patient and encouragement throughout my PhD journey. I would like to express gratitude to my colleagues Dr Anang Hudaya Muhamad Amin and Dr Subarmaniam for their insightful comments on the thesis. I also would like to extend my thanks to the members of Thundercloud Research Lab, Faculty of Information Science & Technology (FIST), Multimedia University (MMU) for their time, guidance, and kind support during my studies. I would like to express greatest gratitude to Ministry of Education (MoE) for sponsoring my studies under MyBrain Scheme and for funding this research under Fundamental Research Grant Scheme (FRGS).

I would like to express my profound gratitude to my parent, abah and arwah mak who always provide love and moral support during my studies. I am deeply indebted to my dear wife Azwina for her love and understanding through my years as a graduate student. Especially, her much needed emotional support on that particular moment when I lost my mother upon completion of my thesis. Thanks also to our daughters Adlina and Hannah Mariah for the joy and the happiness they bring to us during our many moments together. Last, but certainly not least, I must acknowledge with tremendous and deep thanks to my family, friends, neighbours, and students.

TABLE OF CONTENTS

	PAGE
DECLARATION	
APPROVAL	
DEDICATION	
ABSTRACT	i
ABSTRAK	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	ix
LIST OF ALGORITHMS	xiii
LIST OF APPENDICES	xiv
LIST OF ABBREVIATIONS	xv
LIST OF SYMBOLS	xvii
LIST OF PUBLICATIONS	xviii
CHAPTER	
1. INTRODUCTION	1
1.1 Introduction	1
1.2 Threats in Public Wi-Fi Hotspot	2
1.3 Common Countermeasures in Wi-Fi Hotspot	6
1.4 Research Problem	10
1.5 Research Questions	15
1.6 Research Aim and Objectives	16
1.7 Research Approach	18
1.8 Research Operational Framework	20
1.9 Research Contributions	22
1.10 Thesis Organization	26
2. LITERATURE REVIEW	28
2.1 Introduction	28
2.2 Rogue Access Point (RAP)	29
2.3 A Taxonomy of RAP Detection Mechanisms	33
2.4 Existing RAP Detection Mechanisms	38
2.4.1 Infrastructure-centric RAP Detection	40
2.4.1.1 Wired-side Fingerprinting	40
2.4.1.2 Wireless-side Sniffing	46
2.4.2 Client-centric RAP Detection	56
2.5 RAP Detection Mechanisms for Wi-Fi Hotspot	68
2.5.1 Related Work on RAP Detection Framework	68
2.5.2 Characteristics of Proposed RAP Detection Framework	72
2.6 Wi-Fi Localization	77
2.6.1 Localization Algorithms	78
2.6.2 Localization Infrastructure	83
2.6.3 Related Issues in Fingerprint-based Localization	85
2.6.3.1 Grid-based AP Fingerprint Collection	86
2.6.3.2 The Availability of AP Fingerprints during Online Detection	87

2.6.3.3	Radio Map Construction	90
2.6.3.4	Spatial Neighbourhood Formation	93
2.6.3.5	Hierarchical Agglomerative Clustering	95
2.7	Summary	99
3.	RESEARCH METHODOLOGY	102
3.1	Introduction	102
3.2	Research Methodology	103
3.2.1	Preliminary Study Phase	105
3.2.2	Framework Formulation Phase	110
3.2.3	Development Phase	111
3.2.4	Testing and Validation Phase	113
3.2.4.1	Adversary Models	114
3.2.4.2	Procedure for Testing and Validation	119
A.	Validation Datasets I	120
B.	Validation Datasets II	123
C.	Signal Strength Attack Model	124
3.2.5	RAP Classification	126
3.3	Summary	127
4.	THE CONCEPTUAL DESIGN OF PASSIVE CLIENT-CENTRIC RAP DETECTION FRAMEWORK	128
4.1	Introduction	128
4.2	Formulation of RAP Detection Framework	130
4.2.1	Fingerprint-based Localization Classifier Algorithms	131
4.2.1.1	Point-based Localization	131
4.2.1.2	Area-based Localization	132
A.	Simple Point Matching (SPM)	132
B.	Area Based Probability (ABP)	133
4.2.2	Formulation of Self-colocalization Concept	135
4.3	RAP Detection Framework	138
4.4	Development of RAP Detection Framework	146
4.4.1	Offline Deployment Phase: Piggybacking AP Profile in Beacon Frame	147
4.4.1.1	IEEE 802.11 Beacon Frame	149
4.4.1.2	RAPDet-Fi: An Enhanced Association-less Broadcast System	151
4.4.2	Online Detection Phase: Dynamic Radio Map Construction	155
4.4.3	Online Detection Phase: Coarse-grained Colocalization Stage	158
4.4.3.1	The Preliminaries in Colocalization	161
4.4.3.2	New Self-colocalization Classifier Algorithms	163
4.4.4	Online Detection Phase: Fine-grained Colocalization Stage	166
4.4.4.1	Manifold Spatial Neighbourhood Relationships	166
4.4.4.2	A New Location Refinement via Spatial Neighbourhood Relationship	168
4.5	RAP Detection Performance Metrics	173
4.6	Summary	175
5.	EVALUATION OF RAP DETECTION FRAMEWORK	179
5.1	Introduction	179

5.2	Evaluation of Framework Design and Implementation	181
5.2.1	Evaluation of Offline Calibration Phase	181
5.2.1.1	RSSI Distributions	182
5.2.2	Evaluation of Online Detection Phase: Dynamic Radio Map Construction	185
5.2.3	Evaluation of Online Detection Phase: Coarse-grained Colocalization Stage	187
5.2.3.1	Accuracy-Precision Trade-off	187
5.2.3.2	Evaluation of Localization Classifier Algorithms	189
A.	Classifier Accuracy	189
B.	Classifier Precision	193
C.	Comparison of Point-based Classifiers and Area-based Classifiers	198
5.2.3.3	Evaluation of Self-colocalization Algorithm	200
5.2.4	Evaluation of Online Detection Phase: Fine-grained Colocalization Stage	203
5.2.4.1	The Formation of the Clusters	203
5.2.4.2	Effect of Distance Thresholds	205
5.2.5	Recommended Values for RAP Detection Parameters	208
5.3	Evaluation of RAP Detection Framework	210
5.3.1	Detection Accuracy Rate	211
5.3.2	TPR-FPR Trade-off	216
5.3.3	Classification of RAP	221
5.3.4	RAP Detection Framework Responses against Threat Model II	225
5.4	Expert Review on the Framework	230
5.4.1	Industry Expert Background Information	231
5.4.2	Validation on the Key Concepts of the Proposed Framework	232
5.4.3	General Validation on the Overall Framework	233
5.5	Summary	235
6.	CONCLUSION AND FUTURE WORK	237
6.1	Introduction	237
6.2	Concluding Remarks	238
6.3	Summary of Contribution	243
6.4	Future Work	246
	REFERENCES	248
	APPENDICES	278

LIST OF TABLES

TABLE	TITLE	PAGE
1.1	Summary of Research Problems (RPs)	15
1.2	Research Sub-Questions	15
1.3	Mapping of ROs to RPs and RQs	18
1.4	Summary of the Research Contributions	25
2.1	Classification of RAP	31
2.2	Related Literature in RAP Detection Mechanisms	38
2.3	Key Characteristics of RAP Detection	39
2.4	Detection Characteristics of Wired-side Fingerprinting	44
2.5	Summary of Wired-side Fingerprinting Mechanisms	46
2.6	Detection Characteristics of Wireless-side Sniffing	53
2.7	Summary of Wireless-side Sniffing Mechanisms	55
2.8	Detection Characteristics of Client-centric Mechanisms	64
2.9	Summary of Client-centric Detection Mechanisms	67
2.10	Detection Characteristics of Existing RAP Detection Framework	70
2.11	The Evaluation of Client-centric Mechanism Against the Essential RAP Detection Characteristics for Wi-Fi Hotspot	73
2.12	The Summary of the Desirable Characteristics for the Proposed RAP Detection Framework	77
2.13	Summary of Wireless Localization Algorithms	83

3.1	Attack Scenario	118
3.2	Classification of the AP	127
4.1	RAP Detection Decision Rule	174
5.1	Sample of Wi-Fi Hotspot Radio Map	187
5.2	Precision of the Point-based Localization Classifier Algorithm	194
5.3	Highest σ Values for Area-based Classifier Algorithms for Achieving Precision Below 10%	199
5.4	Recommended σ Values for Area-based Classifier Algorithms	209
5.5	RAP Detection Performances when FPR is 10%	220
6.1	Sub-contributions of Offline Calibration Phase	244
6.2	Sub-contributions of Online Detection Phase	245

LIST OF FIGURES

FIGURE	TITLE	PAGE
1.1	The Forecast of Global Public Hotspots from 2009 to 2015	2
1.2	Potential Issues when Using Public Wi-Fi to Access and Transmit Confidential Information	4
1.3	Sample of Captive Portals	9
1.4	Conceptual Framework of the Proposed Rogue AP Detection	19
1.5	Operational Framework of the Research	21
2.1	The Structure of Chapter 2	30
2.2	Phishing AP Attack (a) Evil Twin AP is connected to LAP for Internet connectivity (b) Evil Twin AP is connected to Internet via mobile data connection	32
2.3	Overview of the RAP Detection Mechanisms According to the Proposed Taxonomy	37
2.4	Generic Topology of Infrastructure-centric Detection Mechanism	40
2.5	MITM Attack	59
2.6	Distributed Monitoring Module (DMM)	69
2.7	Graphical Representation of Wireless Localization Algorithms	78
2.8	Topology for Localization Infrastructure	84
3.1	Research Methodology Diagram	104
3.2	Preliminary Study Phase	107

3.3	Classification of RAP Detection Mechanisms	108
3.4	RAP Detection Parameters from Various Wireless Protocol Stack	109
3.5	Framework Formulation Phase	110
3.6	Development Phase of the Research	112
3.7	Floor Plan and RSSI Measurement Points	114
3.8	AP Phishing Attack Operation	116
3.9	Partition of the Dataset into K Folds for the CV Technique	120
3.10	Partition of the Measurement Datasets via CV Technique	121
3.11	Evaluation of RAP Framework Design	123
3.12	Sample of Signal Strength Attack Variation for Random Locations	126
4.1	The Structure of Chapter 4	130
4.2	Self-colocalization of RAP Detector via 3-tuple Reference APs	136
4.3	RAP Detection Framework – Offline Calibration Phase	140
4.4	RAP Detection Framework – Online Detection Phase	141
4.5	The Fluctuation of RSSI Values from LAP and RAP	144
4.6	The Development of the Modules in the Proposed RAP Detection Framework	148
4.7	Generic IE Fields in Beacon Frame	150
4.8	Proposed New RAPDet-Fi IE Field in the Beacon Frame	153
4.9	Hostapd Files Modification for AP Profile Embedding	156
4.10	Beacon Frame Scanning and Radio Map Construction	157
4.11	Beacon Replay Attack	160
4.12	Coarse-grained Colocalization Stage	162
4.13	Sample Dendrogram based on Single-linkage Criterion	168
4.14	Fine-grained Colocalization Stage	170

5.1	The Structure of Chapter 5	180
5.2	RSSI Measurements at Each Location from 4 Different APs	183
5.3	RSSI Measurements in Two Adjacent Locations over Time	184
5.4	RSSI Distributions of APs for Two Adjacent Locations	185
5.5	Signal Strength Contour Map from One of Fingerprint Dataset	186
5.6	Trade-off between Accuracy and Precision	188
5.7	Accuracy of Point-based Localization Classifier Algorithms	190
5.8	Accuracy of Area-based Localization Classifier Algorithms	193
5.9	Precision CDF of Area-based Algorithms over 2-tuple Reference APs	195
5.10	Precision CDF of Area-based Algorithms over 3-tuple Reference APs	196
5.11	Precision CDF of Area-based Algorithms over 4-tuple Reference APs	197
5.12	Accuracy-precision Relationships for Area-based Classifier Algorithms	199
5.13	The Evaluation of Self-colocalization Process	201
5.14	The Formation of Clusters via Single-linkage under Various Thresholds	204
5.15	The Formation of Clusters via Average-linkage under Various Thresholds	205
5.16	Effect of Single-linkage Distance Threshold on Cluster Size and Density	206
5.17	Effect of Average-linkage Distance Threshold on Cluster Size and Density	207
5.18	The Accuracy of RAP Detection Framework with Single-Linkage Clusters	212
5.19	The Accuracy of RAP Detection Framework with Average-Linkage Clusters	215
5.20	TPR-FPR Trade-off Performances under Single-linkage Clusters	217
5.21	TPR-FPR Trade-off Performances under Average-linkage Clusters	219

5.22	RAP Classification under Single-linkage Clusters	223
5.23	RAP Classification under Average-linkage Clusters	224
5.24	RAP Detection Framework Responses I	227
5.25	RAP Detection Framework Responses II	230

LIST OF ALGORITHMS

ALGORITHM	TITLE	PAGE
2.1	Hierarchical Agglomerative Clustering	96
4.1	RAP Detector Self-colocalization	137
4.2	Enhanced AP Profile Anomaly Detection Pseudo Code	143
4.3	Beacon.c	155
4.4	Radio Map Construction	158
4.5	New KNN Classifier for Colocalization	163
4.6	New SPM Classifier for Colocalization	164
4.7	New ABP- α Classifier for Colocalization	165
4.8	Code Snippet of Fine-grained Colocalization Stage	173

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	Scapy Sniffer Script for RSSI Collection	278
B	Reviewer I	280
C	Reviewer II	286

LIST OF ABBREVIATIONS

A	Active
ABP	Area Based Probability
AP	Access Point
BSSID	Basic Service Set Identifier
CC	Client-centric
CDF	Cumulative Density Function
CS	Crowd-sourced
CV	Cross Validation
DoS	Denial of Service
ET	Evil Twin
FPR	False Positive Rate
IAT	Inter-packet Arrival Time
IC	Infrastructure-centric
IC-W	IC-wired
IC-WiFi	IC-wireless
ISP	Internet Service Provider
KNN or K-NN	K-Nearest Neighbour
LAP	Legitimate AP
LBS	Location Based Services
MAC	Medium Access Control

MITM	Man-in-the-middle
NUL	Network and Upper Layer
O	Off the Shelf Hardware
P	Passive
PHY	Physical
RAP	Rogue AP
RC	Research Contribution
RO	Research Objective
RP	Research Problem
RQ	Research Question
RSSI	Received Signal Strength Indicator
RTT	Round Trip Time
S	Specialized Hardware
SN	Sequence Number
SNR	Spatial Neighbourhood Relationship
SPM	Simple Point Matching
SS	Single-sourced
SSID	Service Set Identifier
TCP	Transmission Control Protocol
TOA	Time of Arrival
TPR	True Positive Rate
UDP	User Datagram Protocol
WLAN	Wireless Local Area Network

LIST OF SYMBOLS

C	Dendrograms' Clusters
CL_I	Colocalization I (Coarse-grained Colocalization Stage)
CL_{II}	Colocalization II (Fine-grained Colocalization Stage)
D	Dendrogram
$D(x, y)$	Similarity or Dissimilarity Function
F'	Radio Map
$G(x)$	Partitioning Function
h	Dendrogram Height or Distance Threshold or Detection Threshold
L	Location
P	Total Number of Q -tuple Reference APs in Self-colocalization
Q	Number of APs Used in the Localization
RL	Returned Location or Area
RP	Returned Partition
SP	Cluster of Online RSSI Values according to Q -tuple Reference APs
SS	Set of Online RSSI Values
ss	Location Fingerprint
T	Number of Active APs
α	ABP's confidence level
σ	Standard Deviation in SPM and ABP algorithms

LIST OF PUBLICATIONS

Ahmad, N. M., Abdollah, M. F., Yusof, R., Muhamad Amin, A. H., and Kannan, S., 2014. A RSSI-based Rogue Access Point Detection Framework for Wi-Fi Hotspots. *IEEE 2nd International Symposium on Telecommunication Technologies*, pp. 150-155.

Ahmad, N. M., Abdollah, M. F., Yusof, R., Muhamad Amin, A. H., and Kannan, S., 2014. Detecting Access Point Spoofing Attack using Partitioning-based Clustering Methods, *Journal of Networks*, 9(12), pp. 3470-3477.

Ahmad, N. M., Muhamad Amin, A. H., Abdollah, M. F., and Yusof, R., 2015. An Empirical Investigation of RSSI-based Distance Estimation for Wireless Indoor Positioning System, *International Journal of Wireless and Mobile Computing*, 8(2), pp. 206-212.

Ahmad N. M., Muhamad Amin, A. H., Kannan, S., Ali, A. M. M., Abdollah, M. F., and Yusof, R., 2016. A Passive and Privacy-friendly Area Based Localization for Wireless Indoor Networks. *IEEE Region 10 Symposium (TENSymp)*, pp. 213-218.