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 penderia 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 pengsahihan 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

LIST OF TABLES

LIST OF FIGURES

LIST OF ALGORITHMS

LIST OF APPENDICES

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.