DETECT 802.11 MAC LAYER SPOOF BASED DENIAL OF SERVICE ATTACKS AT SERVICE PROVIDER USING RECEIVED SIGNAL STRENGTH

A dissertation submitted for the Degree of Master of Science in Information Security

G.A.R ELALASINGHAM

University of Colombo School of Computing

2016

UCSC

DECLARATION

The thesis is my original work and has not been submitted previously for a degree at this or any other university/institute.

To the best of my knowledge it does not contain any material published or written by another person, except as acknowledged in the text.

Students Name: G.A.R Elalasingham

____

Signature: Date:

This is to certify that this thesis is based on the work of Mr. G.A.R Elalasingham under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Certified by:

Supervisor Name: Dr.D.A.S Atthukorala

____

Signature: Date:

ABSTRACT

Service Providers today are increasing their dependence on wireless networks in order to operate and maintain a cost effective and competitive advantage as well as offer mobility among corporate users to physically move about whilst maintaining a connection to the corporate wireless network.

However, service providers need to control and prevent their network and systems from being exposed to wireless attacks. Service providers overlook the potential impact of a Denial of Service (DoS) attack against their wireless networks as wireless networks can be very vulnerable to DoS attacks as these attacks could be possibly launched by competitors, for political reasons, as part of a combined attack or just frustration on an attacker's part of not being able to break into service providers' network. However, the results can be anything from degradation of the wireless network to a complete loss of availability and the reputation of the provider.

On an 802.11 wireless network, an attacker can transmit packets using a spoofed source MAC addressees as the activity does not require much expertise and expensive equipment as a result the recipient of spoofed frames has no way of finding if they are legitimate or illegitimate requests and will process and allows MAC layer DoS attacks to take place against the service provider.

A solely RSSI based 802.11 MAC spoof detection and device classification approach is proposed for service providers to detect and classify legitimate or illegitimate requests in the event of MAC layer DoS attacks against their wireless networks. When the attacker is fixed at a location and either surrounded by some stationary devices or none and executes the MAC layer DoS attacks using randomly generated MAC addresses, this approach will guide the examiner to follow the Sampling, Quantization, Pattern Recognition and clustering phases to inspect the location, classify devices and identify such misbehaviors .

This approach was evaluated using real experiments under different criteria and given a confident that the proposed approach can be used to distinguish devices between different model devices, same model devices, different models of devices from one vendor and applicable on any indoor environments.

ACKNOWLEDGEMENTS

I would like to express profound gratitude to my supervisor, Dr. Ajantha Athukorala, for his invaluable motivation, encouragement and support by providing relevant knowledge, advices, supervision and provide valuable feedbacks, appreciations, suggestions throughout this research work.

Meanwhile, I am grateful for Dr. Chamath Keppitiyagama, Mr. Kenath, Mr.Charith Madushanka for their valuable inputs and feedbacks which enabled me to focus the research from different perspective and extend my research work.

Further I would like to thank all my colleagues and friends who motivated and helped me on finding relevant research materials, gather knowledge and the encouragement to manage my MSc research work successfully.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

3G - Third Generation.

LAN - Local Area Network.

Wi-Fi - Wireless Fidelity.

WLAN - Wireless Local Area Network.

PIN - Personal Identification Number.

DHCP - Dynamic Host Configuration Protocol.

DoS - Denial of Service.

RSS - Received Signal Strength.

RSSI - Received Signal Strength Indicator.

MAC - Media Access Control.

IEEE - Institute of Electrical and Electronics Engineers.

ARP - Address Resolution Protocol.

RF - Radio Frequency.

PC - Personal Computer.

FFT - Fast Furious Transfer.

PHY - Physical.

dBm - decibel-milliwatts.

AP - Access Point.

IP - Internet Protocol.

NIC - Network Interface Card.

|RSSI| - Mod of Received Signal Strength Indicator.

QRSSI - Quantized Received Signal Strength Indicator.

Q3 - Quantized interval 3.

Q4 - Quantized interval 4.

Q5 - Quantized interval 5.

CHAPTER 1 INTRODUCTION

1.1 Growing Demand for Wi-Fi Access

Wi-Fi has become a platform for service innovations, revenue growth and experience transformations. As a result, the demand for and availability of Wi-Fi access as a substitute for traditional cellular voice and mobile data has been growing, especially among both mobile and cable operators as it incorporates a number of key benefits to expand their traditional voice and data services like providing high speed data access for smart phone users to surf web sites, download movies, view video clips, share files, access location based services, get navigations and do video conferencing etc. And also it enables service providers to enable new business opportunities like location analytics, 3G offload, mange wireless LANs and the Internet of everything applications that consumers want today as Wi-Fi is enabled to the thermostats, the light bulbs, home security, monitoring and control systems, the appliances, and the automotive products. For example, Dialog has launched 4000 Wi-Fi hotpots, provided 88 high end WLAN solutions across large and medium business organizations and commissioned 114 Public Sites across major bus stands, railway stations, theatres and private institutes especially over the past 18 months.

1.2 Importance of Wireless Network Security

A wired connection can give users a sense of security as the data travelling over a wire is probably not going to be overheard by others. A wireless connection is inherently different; data travelling over the air can be overheard by anyone within range and let the security concerns continue to grow. From the service provider perspective securing a wireless network becomes just as important as any other aspect. Users should be authenticated by some means before they can become functioning members of the wireless network. The identification process is discussed among the operators in the context of identifying the endpoints of a wireless connection and Identifying the end user. The task of managing authentication between client devices and the networks often involve multiple login names and passwords and a safeguard often implemented by a classic login screen.

1.3 Motivation and Objective of the Project

Generally, continuous user authentication can be seen in security policies that lock the user's workstation after a period of inactivity and demanding that users to enter their PIN when user want to continue the services. These traditional security measures annoy people and leave much room for innovation for seamless authentication techniques which allow service providers to identify the end points and the users without hassle the users.

DHCP Option 82 is a technique was designed to insert circuit specific information into a DHCP request that is being forwarded by an end point to a centralized DHCP server. This requests can be used to apply security and service policies based on which endpoint and the user making the request. Dialog specifically uses this option to identify both the cooperate and public end points and the respective users to provide different class and level of services and support assurance.

From user perspective this is quite desirable as the user attachment to the network services has become automatic and even allows the users to hop freely from one network to another without any sign on. However, Dialog has to control and prevent their network and systems from being exposed to wireless attacks. Dialog has to overlook the potential impact of a Denial of Service (DoS) attack against their wireless networks as wireless networks can be very vulnerable to DoS attacks as these attacks could be possibly launched by competitors, for political reasons or as part of a combined attack. However the results can be anything from degradation of the wireless network to a complete loss of availability and ultimately the reputation of Dialog.

For example, On an 802.11 Dialog wireless network, an attacker can transmit packets using a spoofed source MAC addressees without much expertise and expensive equipment and fully utilize the provided centralized DHCP service, as Dialog has no way of finding if they are legitimate or illegitimate requests and will process the requests and allows MAC layer DoS attacks to take place against Dialog and deny or degrade the provided network services for its loyal genuine customers.

1.4 Scope

Though several mitigation approaches [1]-[5] have been proposed to tackle the problem of mac spoof attacks in a wireless networks, this research provides an approach to detect 802.11 MAC spoof activities through observing a device classification solely based on a device's receive signal strength (RSS), which enables service providers to detect and classify legitimate or illegitimate requests in the event of MAC layer DoS attacks against their wireless networks.

This study will be focused to detect and classify the devices or MAC addresses in the event of MAC layer DoS attack against a DHCP service; where the attacker is fixed at an indoor location surrounded either by stationary genuine devices or no devices and sent the illegitimate requests via a single access point (AP) using randomly generated fake MAC addresses.

However, after the classification and detection of devices or suspicious MAC addresses this study will not talk about the further classification of identified MAC addresses as genuine or spoofed MAC addresses and will not leads towards an approach to the prevention of MAC layer DoS attack executed against the DHCP service.

Meantime, this study will not examine or evaluate the functionality, accuracy and the integrity of the functions, tools and devices being used.

1.5 Structure of the Dissertation

The rest of the dissertation is organized as follows.

Chapter 2 –Literature Review

The related work chapter identifies similar solutions and research work done in the area of study and will help identify the requirements and the limitations in detail.

Chapter 3 - Design Preparation

The design chapter describes the design decisions and the design techniques, which were used for the same.

Chapter 4 - Implementation

The implementation chapter discusses about the implementation of the method. All components will be identified and described. Further, the implementation environment, the tools and the data structures used will be discussed.

Chapter 5 - Results &Evaluation

The results and evaluation chapter will describe the methodologies used to verify and validate the approach and explain how all aspects of the objective were tested.

Chapter 6 - Conclusion& Future Work

The conclusion and future work chapter concludes the dissertation by discussing the final results and the future works.

CHAPTER 2 LITERATURE REVIEW

The following sections are classified to identify similar researches and solutions available related to the detection of 802.11 MAC spoofing from device classification based on device's RSS. The sections are organized to describe the availability to change the 802.11 MAC addresses of the devices, explain already existing 802.11 MAC spoofing detection methods ,approaches and the limitations, the findings and the limitations from existing transceiver based spoof detection methods and the RSS based classification approaches followed by exploring a device's energy impact on received RSS.

2.1 802.11 MAC Spoofing

Nearly all 802.11 cards in use permit their MAC addresses to be altered, often with full support and drivers from the manufacturer. Both Linux and Windows users are commonly permitted to change their MAC address using open-source tools and changing network card properties respectively. An attack execute via a brute-force attack script with a random MAC address for each successive connection attempt go undetected by network activity analysis applications [1].

2.2 802.11 MAC Spoof Detection

The IEEE makes the list of prefix allocations and the assigned company information available to the public, largely for users to match a piece of equipment with a MAC address to its manufacturer. We can use this list to evaluate all source MAC addresses on the network to determine if the prefix is allocated by the IEEE. MAC addresses that appear on the network using a prefix as yet unallocated by the IEEE can be flagged as misbehavior activity [1].However still sophisticated attackers may deliberately forge to evade the detection.

As part of the 802.11 specifications, the IEEE included a method to accommodate fragmentation for large management and data frames through the use of frame sequencing. Wright [1] also proposes to use these sequence number gaps as the detection clue. If the gap exceeds a certain threshold, a spoofing alert is raised. This method, however, may raise false alerts in the presence of lost or duplicated frames, which are common in practice.

Guo and Chiueh extend this method to use ARP to confirm the current sequence number from the genuine station [2]. This approach, however, does not work when there is no genuine station in the network.

2.3 Transceiver based Detection and Classification

Hall [3] proposes a hardware based approach to detect transceiver-print anomaly. A transceiver-print is extracted from the turn-on transient portion of a signal. It reflects the unique hardware characteristics of a transceiver. This RF pattern cannot be manipulated at the software level, and is hard to forge by even using software radio. Thus this approach is potentially the most reliable method for detecting spoofing attacks. However, though this method potentially reliable, it has a higher cost in both measurement and analysis devices, and limits its use practical at service providers.

Chen [4]. Propose a method for detecting spoofing attacks and locating the adversary, in 802.11WLANs assuming that received signal strength (RSS) values follow a Gaussian distribution with a uniform derivation. Their work does show that per-frame RSS analysis and multiple coordination of access points are Promising for spoofing detection. However In a realistic deployment their algorithm proven may not work well, due to non-Gaussian RSS distribution and missing RSS measurements.[5]

Yong Sheng and his Team [5] propose an approach based on Gaussian mixture models, building RSS profiles for spoofing detection assuming the attacker and the victim are separated by a reasonable distance. They claim that the RSS is a measurement that is hard to forge arbitrarily and it is highly correlated to the transmitter’s location and can be used to differentiate devices to detect MAC spoofing as proposed by several other researches. They have demonstrated using experiments on a building-scale wireless test bed, for detecting moving attackers who spoof the MAC addresses of stationary devices. However, there approach will not detect attackers when there is no genuine station available and the attacker also not moving

2.4 RSSI models for classification

Chalampos, Franco and Robert [7] have formulated a probability distribution of the position of a sensor node based on one reading produced with the RSSI model. They have evaluated the probability that a sensor node lies within a certain region, given that the power received from the beacons is modeled with RSSI.

They show that the actual position is lognormal and given an evidence to the role of the parameters standard deviation and the path loss component in the probability distribution of the actual distance. They also present a method for estimating the location of a node from multiple sample power readings computes the expected value of the received power and combines it with the mean and the standard deviation of the sample readings. Their algorithm for location estimation based on several samples does not involve any complex calculations (such as square roots) which is very important to consider when we develop algorithms to be executed.

Venkatraman, Frederik, Thiemo [6] have presented a detection and classification scheme for concurrent multi-source interference affecting wireless sensor networks by accounts collective emissions from the interferers by sampling the received signal strength (RSSI), and perform interference detection and classification in sequence. The detection method uses an RSSI sampler that captures the emissions from all interferers as a series of RSSI bursts involves an unsupervised learning approach, i. e., clustering, to distinguish the bursts from the different Interferers. The output of the interference detection component tis then passed to a classification component that inspects each cluster for periodicity. The classification relies on the temporal patterns of an interferer’s emissions. Though this algorithm is designed to be executed on sensor nodes, this method is potentially suitable for this study.

Similarly, Zacharias etal. [9] Classify interference based on a fixed set of simple conditions. In contrast to approach in [6], this classification method includes processing of computationally expensive tasks such as FFTs and execution on a PC rather than on sensors.

2.5 Energy impact on RSSI

A Wi-Fi radio can be in several operating modes, known as power states for that device, each draining a different amount of power. The power states and transitions for Wi-Fi are simpler, as shown in Figure 1.

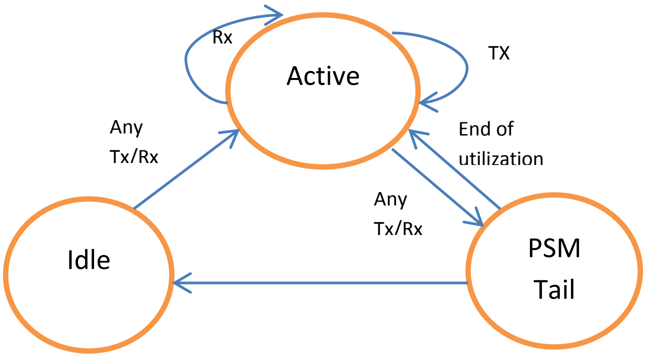


Figure 1: Wi-Fi Power State Machine

The above power state behavior of Wi-Fi suggests that the total energy drain of Wi-Fi in carrying out a fixed amount of traffic equals the sum of the energy consumed while in productive power states and while in the tail state, which is determined by the total duration they stay in these states. The total time spent in a productive state is affected by the data rate and retransmissions. The total time spent in the tail state depends on the total number of times the device enters the tail state and the duration of each episode. However, conceptually, the tail power state phenomenon is independent of the signal strength [11].

Ning Ding, Daniel Wagner, Xiaomeng Chen [11] contended the quantitative impact of poor signal strength on the device battery drain. The poor signal strength can significantly inflate the actual energy consumption by the wireless interface to be much higher than under good signal strength. The more obvious impact on energy drain is that reduced signal strength triggers rate adaptation at the PHY layer to lower the data rate which elongates packet transmission, result in retransmissions, disassociation and re-association with the access point and hence increases power consumption by the radio.

CHAPTER 3 DESIGN PREPARARTION

This section presents the preliminary decisions and actions taken to identify a suitable model to detect MAC spoofing attacks solely based on RSSI measurements obtained from stationary devices fixed at a location and surrounded by some stationary devices or none.

3.1 Reference Model

The underlying objective of the reference model as shown in below figure 2 is to illustrate the methodology used to reads RSSI values and performs a detection step prior to explicitly classifying sources of spoofing attack.

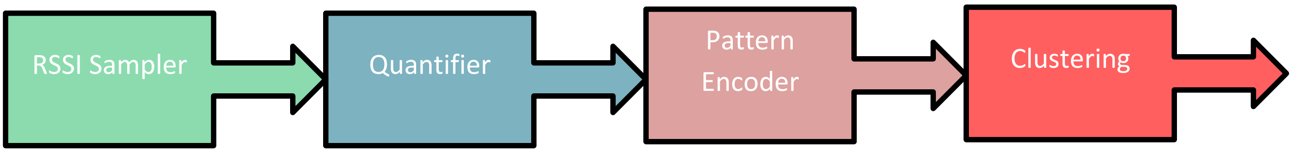


Figure 2: Reference Model

RSSI Sampler

The RSSI sampler captures the emissions from all devices as a series of samples at a fixed defined interval.

Quantifier

Quantization is motivated by the emissions from a given device may vary slightly over time in their strength. These minor variations are not relevant to detecting the spoofing attack, and hence they can be abstracted away by quantizing the RSSI reading. For example, establish power level 1 for RSSI values below -90 dBm, and divide the RSSI range above > -90 dBm evenly over the remaining number of levels like power level 2 (-90 dBm < RSSI <= - 60 dBm), power level 3 (-60 dBm < RSSI <= -30 dBm),and power level 4 (-30 dBm < RSSI).Then the following sequence of RSSI samples from a device -85,-86,-56,-84, -56,-54,-53 can be interpreted as 2,2,3,2,3,3,3. Higher number of intervals allows capturing more distinct values however, this quantization levels will be concluded at the implementation steps.

Pattern Encoder

The encoder simply counting the number of occurrences of quantified power levels and provides a pattern based on power levels used by a device at most during its transmission. For example the following sequence 2,2,3,2,3,3,3 will be encoded as (3, 2) as the pattern of transmission used by the device the most in highest order.

Cluster

The clustering component arranges the RSSI samples into groups such that samples that are belong to particular transmission pattern are assigned to the same group. The underlying intuition is that similar bursts reference to particular transmission pattern is likely to come from the same device.

3.2 Test Bed Preparation

The test execution environment setup to test the concept is shown in Figure 3.

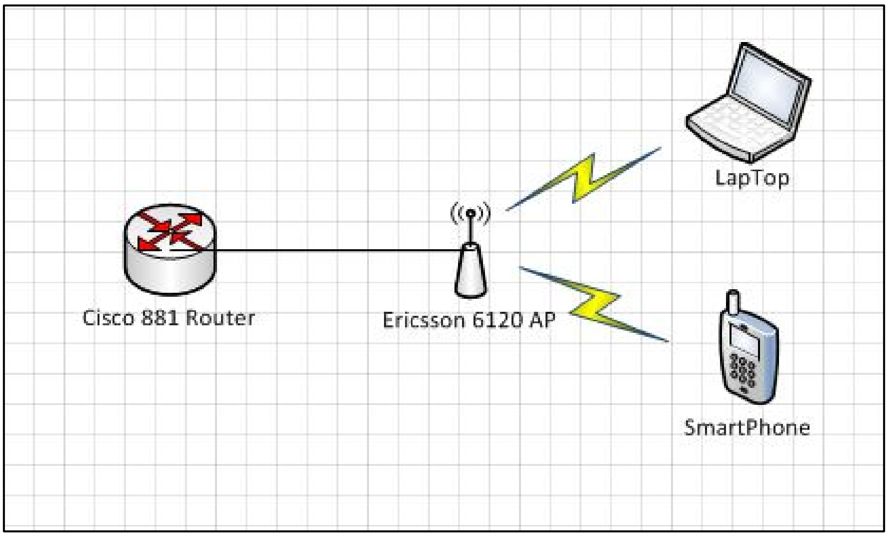


Figure 3: Test Bed

The experimental setup comprises an Ericsson omni directional Wi-Fi Access Point 6120, a Cisco 881 Router, a Samsung S3 smart phone and an AcerE5471 laptop. The Wi-Fi Access Point is used to provide a omni directional Wi-Fi Coverage in 2.4 GHz band. In order to mitigate the impact on device battery drain due to low data rate and retransmissions the transmission power of the AP is fixed to 17dbm and the low transmission rates like 1, 2, 5.5,6,11 Mbps are disabled.

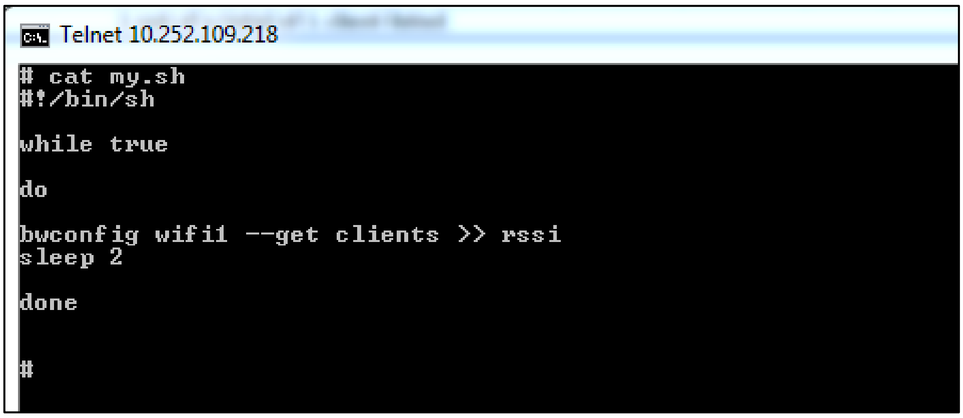
The router is used to offer a DHCP service for the laptop and the smartphone. The DHCP pool is configured inside the router with a 255 IPs with a lease time of one hour.

The lap top is used to execute the spoofing attack at a fixed distance from the AP. The underlying objective of using the lap top and the smart phone is to observe the devices' transmission pattern while they are fixed at a distance from the AP and one of its MAC address get changed using random MAC addresses, detect MAC spoofing activities based on the observations.

3.3 Tools

Bash Script

A script as shown in figure was written inside the AP to record the RSSI values of the devices connected on its 2.4 GHz radio at every 2 seconds and redirects the out to a file named as rssi inside the AP. The obtained data set in the file rssi will be processed by the quantifier.

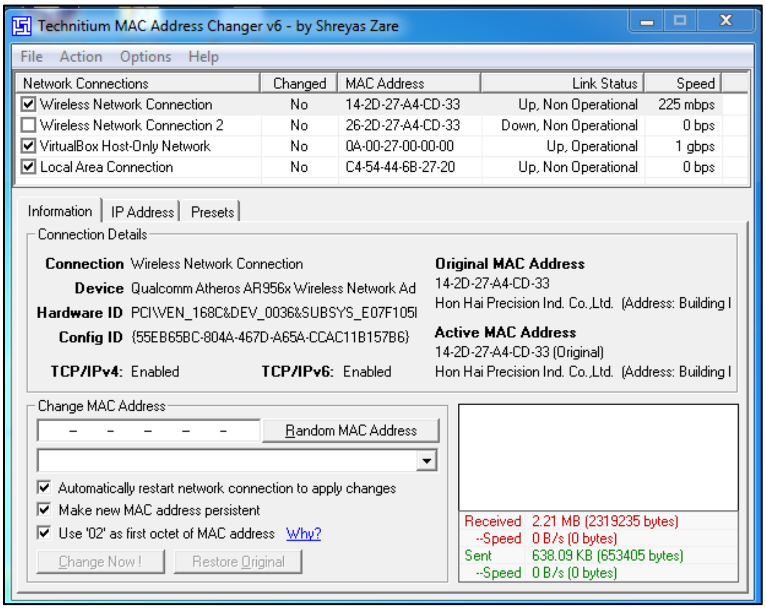


```
# cat my.sh
#!/bin/sh

while true

do

bwconfig wifi1 --get clients >> rssi
sleep 2

done


#
```

Figure 4: Bash script to get rssi samples

TMAC V6

A freeware utility allows spoofing MAC Address of Network Interface Card instantly on machines works on Windows 10, 8 & 7 for both 32-bit and 64-bit. It has a very simple user interface to set a new MAC address to the NIC, bypassing the original hard coded MAC address and provides information regarding each NIC in the machine as shown in Figure 5.



Network Connections	Changed	MAC Address	Link Status	Speed
☑ Wireless Network Connection	No	14-2D-27-A4-CD-33	Up, Non Operational	225 mbps
□ Wireless Network Connection 2	No	26-2D-27-A4-CD-33	Down, Non Operational	0 bps
☑ VirtualBox Host-Only Network	No	0A-00-27-00-00-00	Up, Operational	1 gbps
☑ Local Area Connection	No	C4-54-44-6B-27-20	Up, Non Operational	0 bps

Connection Wireless Network Connection
Device Qualcomm Atheros AR956x Wireless Network Ad
Hardware ID PCI\VEN_168C&DEV_0036&SUBSYS_E07F105I
Config ID {55EB65BC-804A-467D-A65A-CCAC11B157B6}
TCP/IPv4: Enabled **TCP/IPv6:** Enabled

Original MAC Address
14-2D-27-A4-CD-33
Hon Hai Precision Ind. Co.,Ltd. (Address: Building I

Active MAC Address
14-2D-27-A4-CD-33 (Original)
Hon Hai Precision Ind. Co.,Ltd. (Address: Building I

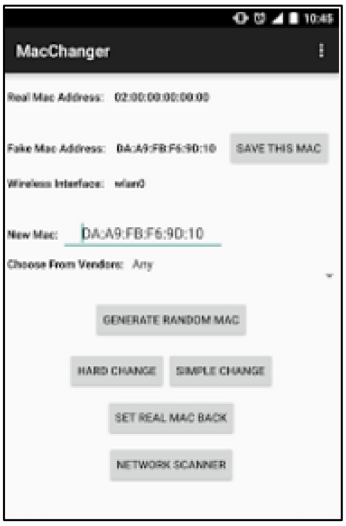
☑ Automatically restart network connection to apply changes
☑ Make new MAC address persistent
☑ Use '02' as first octet of MAC address

Received 2.21 MB (2319235 bytes)
--Speed 0 B/s (0 bytes)
Sent 638.09 KB (653405 bytes)
--Speed 0 B/s (0 bytes)

Figure 5: TMAC

Wifi Mac Changer

This Application changes and spoof wireless Interface MAC address to any wanted MAC addresses on android devices. However it requires rooted devices and BusyBox to be installed.



MacChanger

Real Mac Address: 02:00:00:00:00:00

Fake Mac Address: DA:A9:FB:F6:9D:10

Wireless Interface: wlan0

New Mac: DA:A9:FB:F6:9D:10

Choose From Vendors: Any

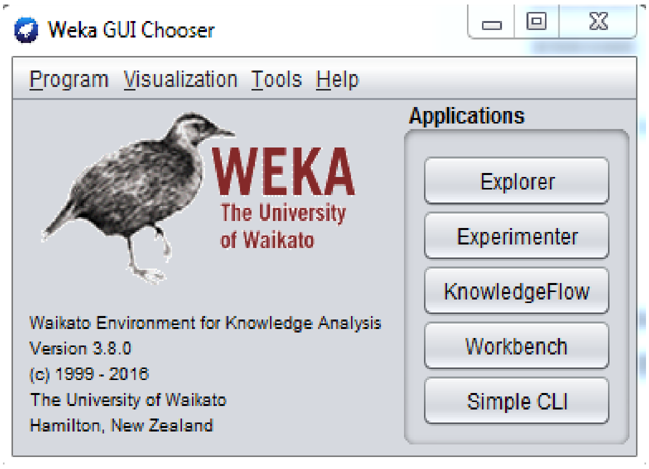
Figure 6: Wifi Mac changer

Microsoft Excel 2010

The familiar excel spread sheet was used to quantify the RSSI samples from the sample data and obtain the device emission pattern using excel functions like “IF” and “pivot table” to populate a large, detailed data set to feed into the clustering component.

Weka

Weka is a popular suite of machine learning open source software. It has a collection of machine learning algorithms for data mining tasks. The machine learning algorithms can be applied directly to a dataset for data pre-processing, classification, regression, clustering, association rules, and for visualization. Weka is also well-suited for developing new machine learning schemes. Weka is used to cluster the samples from the dataset and classify the devices for inspection through visualization.



Weka GUI Chooser

Waikato Environment for Knowledge Analysis
Version 3.8.0
(c) 1999 - 2016
The University of Waikato
Hamilton, New Zealand

Figure 7: Tool Weka

CHAPTER 4 IMPLEMENTATION

The test bed in figure 3 was implemented inside a 10x15 feet closed room as shown in below figure 8.

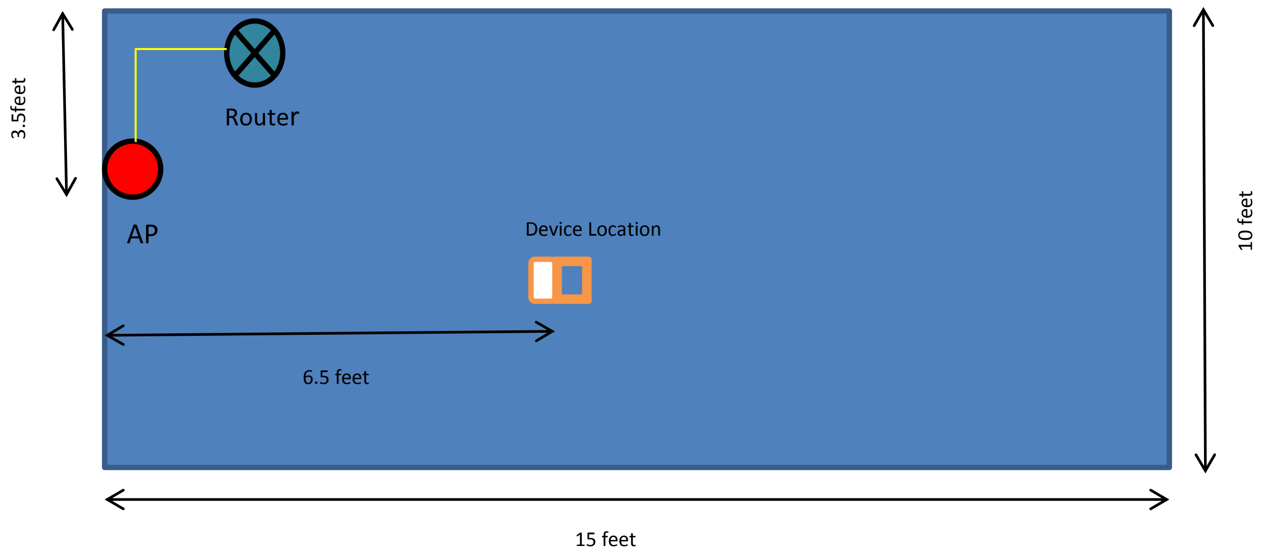


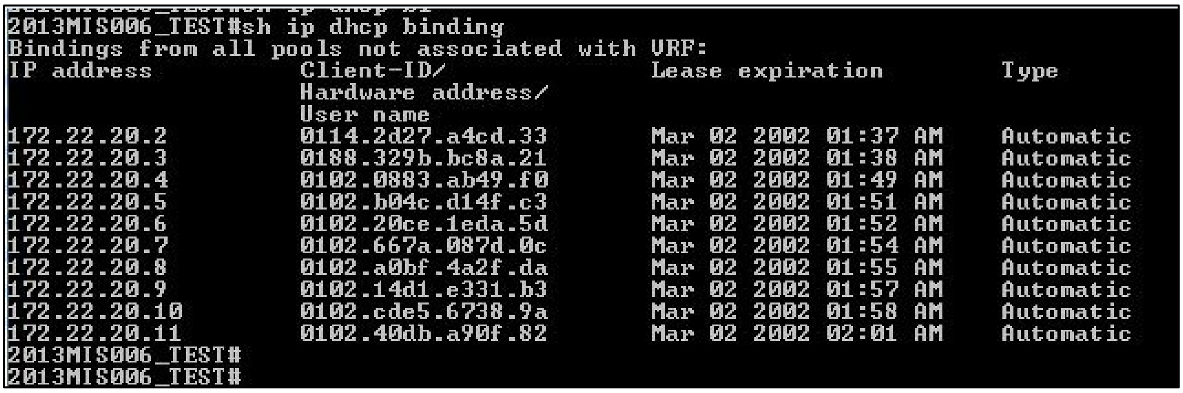
Figure 8: Test Bed Implementation

The AcerE14 laptop with the original mac addresses of 14:2d:27:a4: cd:33 and the Samsung Galaxy S3 smartphone with the original mac address of 88:32:9b:bc:8a:21 were kept at a fixed distance of 6.5 feet from the AP. The former device was used to perform the spoofing activity whereas the later one was used for a genuine user reference. The underlying objective of the attempt is to detect spoofing activities based on the observations over both devices' transmission pattern while they are fixed at the distance from the AP and changes the MAC addresses of the device AcerE14 using random MAC address values.

Initially, the laptop and the smartphone are attached to the access point and the script shown in the Figure 4 was executed inside the AP to record the RSSI samples.

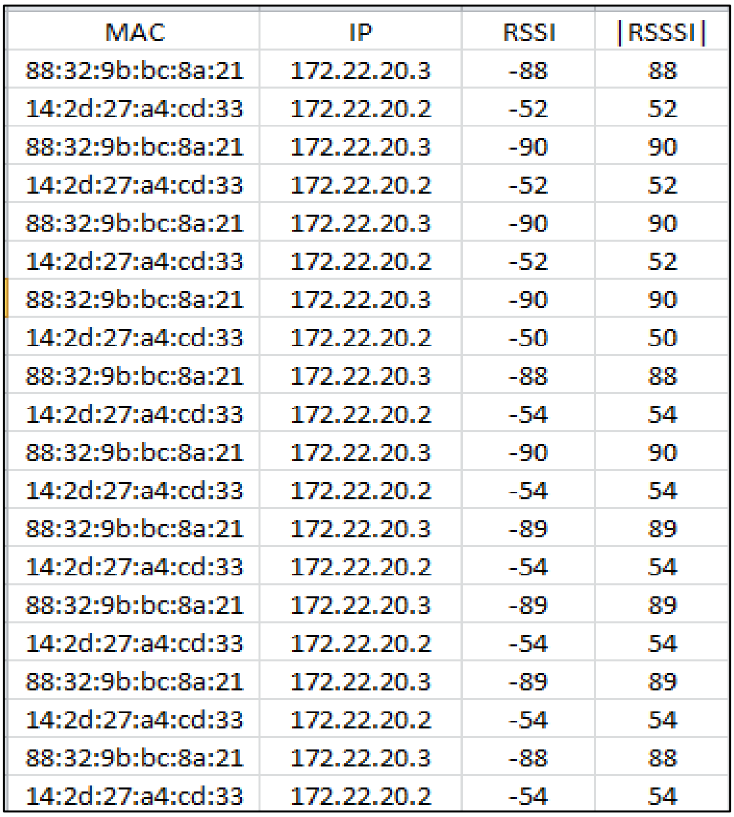
During the sampling period , the MAC address of the laptop is randomly changed to the value 02:08:83:ab:49:f0 using TMAC v6 shown in Figure 5 and ensured that the laptop with the new mac address has attached to the AP and obtained a new IP from the DHCP server. After repeating this activity on the laptop for 8 times at one minute interval the script shown in Figure 4 running inside the AP manually got stopped. The changed MAC address values and the obtained IP addresses during the activity are shown in Figure 9. The date and the time stamp inside the router is not synced properly as a result the lease expiration stamp indicate incorrect date and time. It can be ignored as it does not have any influences over upcoming activities and impact on the results.

The recorded sample values inside the AP under the file "rssi" as explained in
Figure 4, are populated into an excel sheet as a part of the sheet is shown in Figure 10. To avoid dealing values with minus signs, an additional column | RSSI | is introduced in the sheet.



```
2013MIS006_TEST#sh ip dhcp binding
Bindings from all pools not associated with VRF:
IP address          Client-ID/              Lease expiration        Type
                    Hardware address/
                    User name
172.22.20.2         0114.2d27.a4cd.33       Mar 02 2002 01:37 AM    Automatic
172.22.20.3         0188.329b.bc8a.21       Mar 02 2002 01:38 AM    Automatic
172.22.20.4         0102.0883.ab49.f0       Mar 02 2002 01:49 AM    Automatic
172.22.20.5         0102.b04c.d14f.c3       Mar 02 2002 01:51 AM    Automatic
172.22.20.6         0102.20ce.1eda.5d       Mar 02 2002 01:52 AM    Automatic
172.22.20.7         0102.667a.087d.0c       Mar 02 2002 01:54 AM    Automatic
172.22.20.8         0102.a0bf.4a2f.da       Mar 02 2002 01:55 AM    Automatic
172.22.20.9         0102.14d1.e331.b3       Mar 02 2002 01:57 AM    Automatic
172.22.20.10        0102.cde5.6738.9a       Mar 02 2002 01:58 AM    Automatic
172.22.20.11        0102.40db.a90f.82       Mar 02 2002 02:01 AM    Automatic
2013MIS006_TEST#
2013MIS006_TEST#
```

Figure 9: DHCP leased addresses

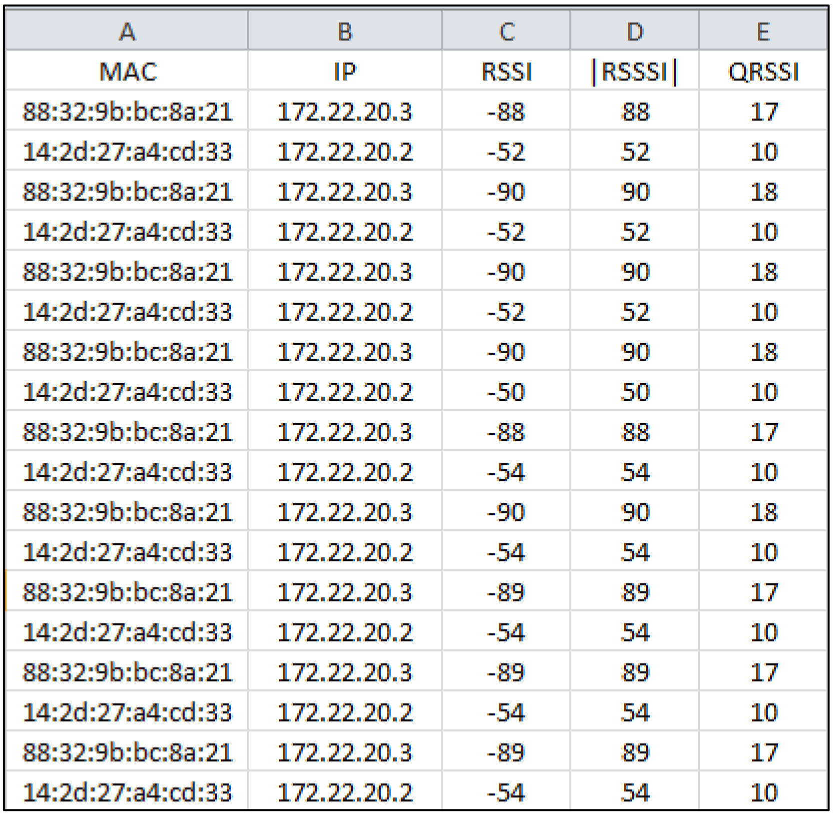


MAC	IP	RSSI	|RSSSI|
88:32:9b:bc:8a:21	172.22.20.3	-88	88
14:2d:27:a4:cd:33	172.22.20.2	-52	52
88:32:9b:bc:8a:21	172.22.20.3	-90	90
14:2d:27:a4:cd:33	172.22.20.2	-52	52
88:32:9b:bc:8a:21	172.22.20.3	-90	90
14:2d:27:a4:cd:33	172.22.20.2	-52	52
88:32:9b:bc:8a:21	172.22.20.3	-90	90
14:2d:27:a4:cd:33	172.22.20.2	-50	50
88:32:9b:bc:8a:21	172.22.20.3	-88	88
14:2d:27:a4:cd:33	172.22.20.2	-54	54
88:32:9b:bc:8a:21	172.22.20.3	-90	90
14:2d:27:a4:cd:33	172.22.20.2	-54	54
88:32:9b:bc:8a:21	172.22.20.3	-89	89
14:2d:27:a4:cd:33	172.22.20.2	-54	54
88:32:9b:bc:8a:21	172.22.20.3	-89	89
14:2d:27:a4:cd:33	172.22.20.2	-54	54
88:32:9b:bc:8a:21	172.22.20.3	-89	89
14:2d:27:a4:cd:33	172.22.20.2	-54	54
88:32:9b:bc:8a:21	172.22.20.3	-88	88
14:2d:27:a4:cd:33	172.22.20.2	-54	54

Figure 10 : A Part of RSSI Samples with MAC Addresses and IP Addresses

Then an additional column QRSSI is introduced to the data sheet in Figure 10 to include the quantified RSSI sample values obtained from defined quantization interval as shown in Figure 11. In order to observe a considerable variation in the transmission pattern of the devices, quantification interval is started from the interval 5. The following excel function is included in the data sheet in Figure 10 to obtain the quantized sample values of the samples at the interval of 5.

```
=IF(D2<10,1,IF(D2<15,2,IF(D2<20,3,IF(D2<25,4,IF(D2<30,5,IF(D2<35,6,IF(D2<40,7,IF(D2<45,8,IF(D2<50,9,IF(D2<55,10,IF(D2<60,11,IF(D2<65,12,IF(D2<70,13,IF(D2<75,14,IF(D2<80,15,IF(D2<85,16,IF(D2<90,17,IF(D2<95,18,19))))))))))))))))))
```



A	B	C	D	E
MAC	IP	RSSI	|RSSSI|	QRSSI
88:32:9b:bc:8a:21	172.22.20.3	-88	88	17
14:2d:27:a4:cd:33	172.22.20.2	-52	52	10
88:32:9b:bc:8a:21	172.22.20.3	-90	90	18
14:2d:27:a4:cd:33	172.22.20.2	-52	52	10
88:32:9b:bc:8a:21	172.22.20.3	-90	90	18
14:2d:27:a4:cd:33	172.22.20.2	-52	52	10
88:32:9b:bc:8a:21	172.22.20.3	-90	90	18
14:2d:27:a4:cd:33	172.22.20.2	-50	50	10
88:32:9b:bc:8a:21	172.22.20.3	-88	88	17
14:2d:27:a4:cd:33	172.22.20.2	-54	54	10
88:32:9b:bc:8a:21	172.22.20.3	-90	90	18
14:2d:27:a4:cd:33	172.22.20.2	-54	54	10
88:32:9b:bc:8a:21	172.22.20.3	-89	89	17
14:2d:27:a4:cd:33	172.22.20.2	-54	54	10
88:32:9b:bc:8a:21	172.22.20.3	-89	89	17
14:2d:27:a4:cd:33	172.22.20.2	-54	54	10
88:32:9b:bc:8a:21	172.22.20.3	-89	89	17
14:2d:27:a4:cd:33	172.22.20.2	-54	54	10

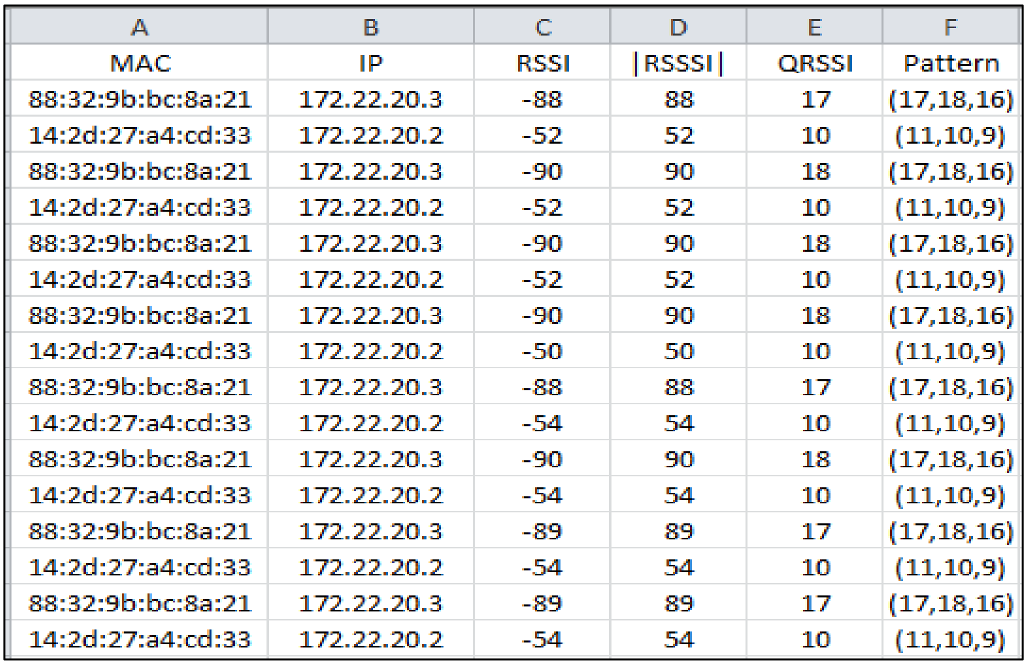
Figure 11: A Part of RSSI Samples with Quantified Values

Then, using Excel pivot table, the data stored in the sheet shown in Figure 11 is summarized by dragging and dropping fields graphically to get a summary on count of QRSSI value under each MAC addresses and discover the transmission pattern of the MAC address based on the highest order of quantified values used by each MAC addresses as summarized in below Table 1.

MAC Addresses	QRSSI	Count of QRSSI	Transmission Pattern
02:08:83:ab:49:f0	10	102	(10,11,9)
	11	8	
	9	2	
02:14:d1:e3:31:b3	10	114	(10,11,9)
	11	12	
	9	2	
02:20:ce:1e:da:5d	10	88	(10,11,9)
	11	30	
	9	2	
02:40:db:a9:0f:82	11	61	(11,10,9)
	10	37	
	9	15	
02:66:7a:08:7d:0c	10	116	(10,11,9)
	11	5	
	9	2	
02:a0:bf:4a:2f:da	11	89	(11,10,9)
	10	35	
	9	2	
02:b0:4c:d1:4f:c3	10	108	(10,11,9)
	11	17	
	9	4	
02:cd:e5:67:38:9a	10	84	(10,11,9)
	11	53	
	9	24	
14:2d:27:a4:cd:33	11	71	(11,10,9)
	10	39	
	9	5	
88:32:9b:bc:8a:21	17	401	(17,18,16)
	18	94	
	16	1	

Table 1 : Summary on QRSSI and the discovered transmission pattern of each MAC addresses.

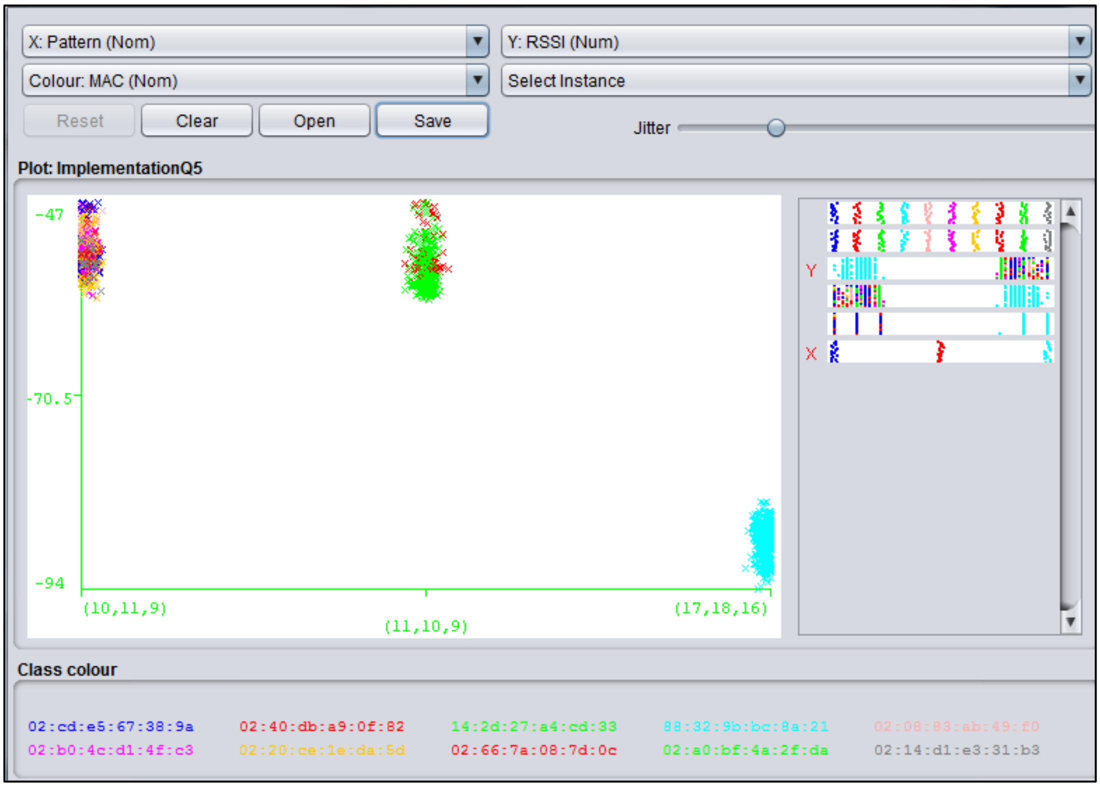
The discovered transmission patterns in Table 1 is then appended to the excel sheet in Figure 11 subjected to MAC addresses under a new column named as “Pattern”. A part of the sheet after the appendance is shown in below Figure 12.



A	B	C	D	E	F
MAC	IP	RSSI	|RSSSI|	QRSSI	Pattern
88:32:9b:bc:8a:21	172.22.20.3	-88	88	17	(17,18,16)
14:2d:27:a4:cd:33	172.22.20.2	-52	52	10	(11,10,9)
88:32:9b:bc:8a:21	172.22.20.3	-90	90	18	(17,18,16)
14:2d:27:a4:cd:33	172.22.20.2	-52	52	10	(11,10,9)
88:32:9b:bc:8a:21	172.22.20.3	-90	90	18	(17,18,16)
14:2d:27:a4:cd:33	172.22.20.2	-52	52	10	(11,10,9)
88:32:9b:bc:8a:21	172.22.20.3	-90	90	18	(17,18,16)
14:2d:27:a4:cd:33	172.22.20.2	-50	50	10	(11,10,9)
88:32:9b:bc:8a:21	172.22.20.3	-88	88	17	(17,18,16)
14:2d:27:a4:cd:33	172.22.20.2	-54	54	10	(11,10,9)
88:32:9b:bc:8a:21	172.22.20.3	-90	90	18	(17,18,16)
14:2d:27:a4:cd:33	172.22.20.2	-54	54	10	(11,10,9)
88:32:9b:bc:8a:21	172.22.20.3	-89	89	17	(17,18,16)
14:2d:27:a4:cd:33	172.22.20.2	-54	54	10	(11,10,9)
88:32:9b:bc:8a:21	172.22.20.3	-89	89	17	(17,18,16)
14:2d:27:a4:cd:33	172.22.20.2	-54	54	10	(11,10,9)

Figure 12: A Part of Spreadsheet updated with the column "Pattern"

Finally, the data sheet in Figure 12 is uploaded into the tool WEKA shown in Figure 7 to arrange the RSSI samples into clusters such that samples that are belong to particular transmission pattern are assigned to the same cluster as shown in Figure 13.



X: Pattern (Nom)
Y: RSSI (Num)
Colour: MAC (Nom)
Select Instance

Plot: ImplementationQ5

Figure 13: RSSI clusters belong to transmission pattern with quantization level 5

In the Figure 13, the x-axis represents the transmission pattern of the MAC addresses; the y-axis represents the RSSI values recorded against the MAC addresses. Each MAC addresses is represented by different colours.

The Figure 13 exhibits three clusters formed by three unique transmission patterns. The clusters with more than one colour gives a hint that the cluster is formed by RSSI samples subjected the transmission pattern but it has involved more than one MAC addresses which greatly helped to classify the devices and detect the 802.11 MAC spoof activity held during the process by concluding that either there is a possible presence of three devices each with the transmission pattern subjected to x-axis, two of them are suspicious devices. Or there are 10 MAC addresses exist subjected to the transmission patterns in the x-axis where the MAC address 88:32:9b: bc: 8a:21 belongs to a genuine device and the other 9 MAC addresses belongs to some non-genuine devices or spoofed MAC addresses. However, the former conclusion reveals more about the device classification with a 50% deviation from the actual expected count where the latter conclusion gives information about both device classification and the detection of suspicious MAC addresses.

A similar approach repeated on the dataset shown in Figure 10 under the quantization intervals 4 and 3. The functions included in the dataset in Figure 10 to obtain the quantized values of the samples under intervals 4 and 3 are shown below respectively.

=IF(D2<=10,1,IF(D2<=14,2,IF(D2<=18,3,IF(D2<=22,4,IF(D2<=26,5,IF(D2<=30,6,IF(D2<=34,7,IF(D2<=38,8,IF(D2<=42,9,IF(D2<=46,10,IF(D2<=50,11,IF(D2<=54,12,IF(D2<=58,13,IF(D2<=62,14,IF(D2<=66,15,IF(D2<=70,16,IF(D2<=74,17,IF(D2<=78,18,IF(D2<=82,19,IF(D2<=86,20,IF(D2<=90,21,IF(D2<=94,22,IF(D2<=98,23,0)))))))))))))))))))))))

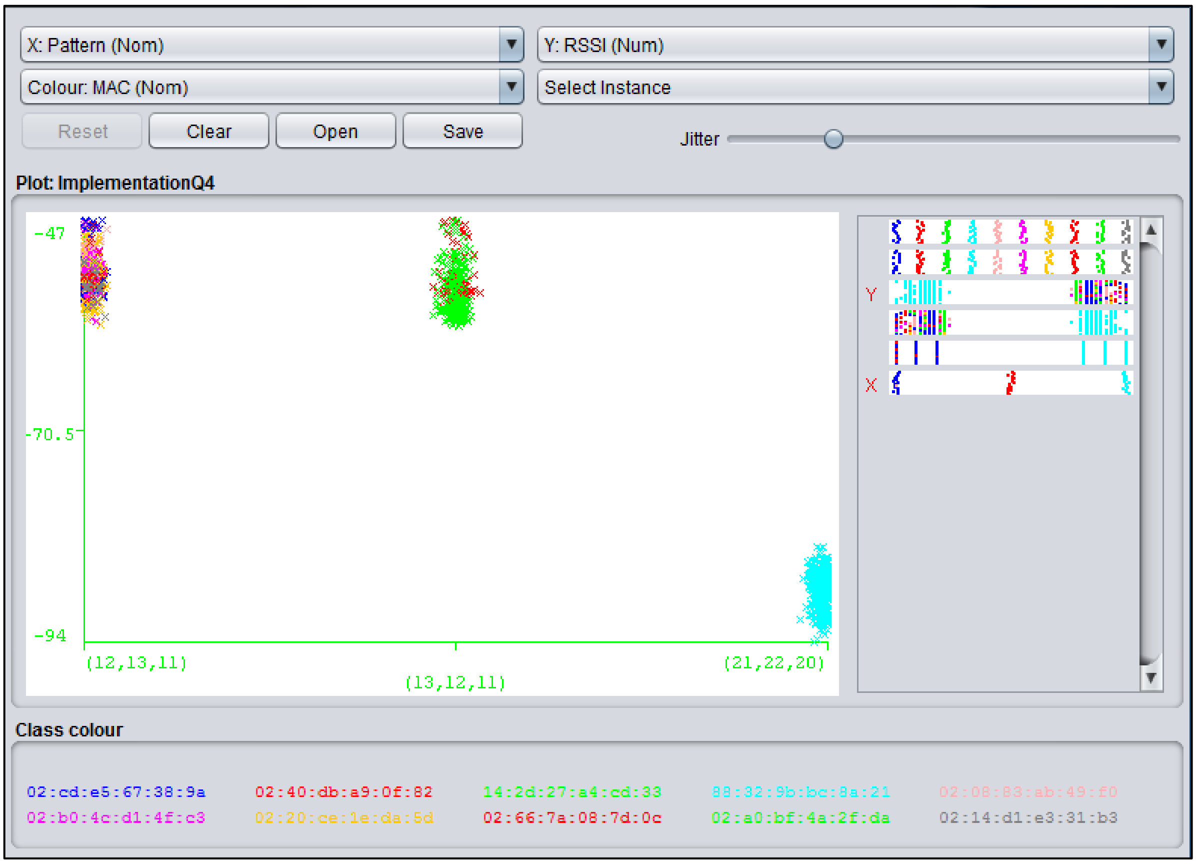
=IF(D2<=10,1,IF(D2<=13,2,IF(D2<=16,3,IF(D2<=19,4,IF(D2<=22,5,IF(D2<=25,6,IF(D2<=28,7,IF(D2<=31,8,IF(D2<=34,9,IF(D2<=37,10,IF(D2<=40,11,IF(D2<=43,12,IF(D2<=46,13,IF(D2<=49,14,IF(D2<=52,15,IF(D2<=55,16,IF(D2<=58,17,IF(D2<=61,18,IF(D2<=64,19,IF(D2<=67,20,IF(D2<=70,21,IF(D2<=73,22,IF(D2<=76,23,IF(D2<=79,24,IF(D2<=82,25,IF(D2<=85,26,IF(D2<=88,27,IF(D2<=91,28,IF(D2<=94,29,IF(D2<=97,30,0))))))))))))))))))))))))))))))

The identified transmission patterns of the MAC addresses subjected to quantization intervals 4 and 3 are tabulated in Table 2.

MAC	Transmission pattern under different quantization level		
	Q5	Q4	Q3
02:08:83:ab:49:f0	10,11,9	12,13,11	15,16,17,14
02:14:d1:e3:31:b3	10,11,9	12,13,11	16,15,17,14
02:20:ce:1e:da:5d	10,11,9	12,13,11	16,15,17,14
02:40:db:a9:0f:82	11,10,9	13,12,11	16,14,17,15
02:66:7a:08:7d:0c	10,11,9	12,13,11	16,15,14,17
02:a0:bf:4a:2f:da	11,10,9	13,12,11	17,16,15,14
02:b0:4c:d1:4f:c3	10,11,9	12,13,11	15,16,17,14
02:cd:e5:67:38:9a	10,11,9	12,13,11	16,14,15,17
14:2d:27:a4:cd:33	11,10,9	13,12,11	17,16,15,14
88:32:9b:bc:8a:21	17,18,16	21,22,20	27,28,29,26

Table 2: Transmission patterns under different quantization levels

The WEKA clustering output of the RSSI samples subjected to quantization intervals 4 and 3 are shown in figure 14 and 15 respectively.



X: Pattern (Nom)

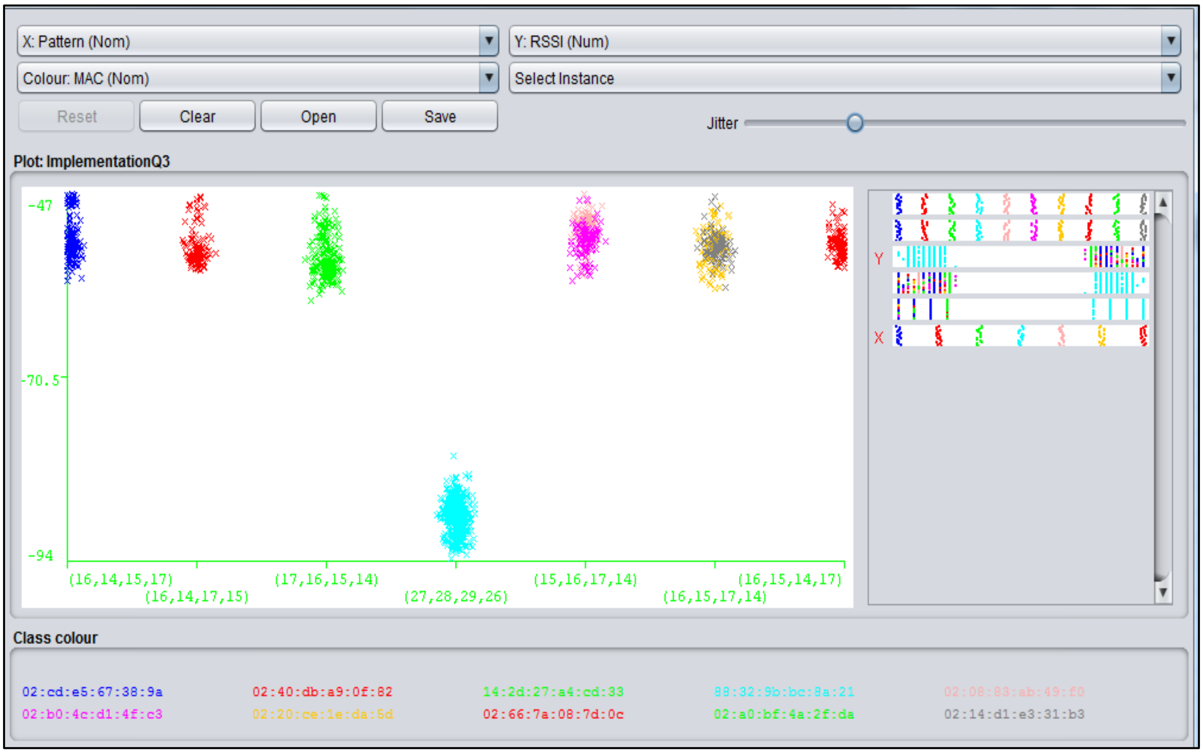
Y: RSSI (Num)

Colour: MAC (Nom)

Select Instance

02:cd:e5:67:38:9a 02:40:db:a9:0f:82 14:2d:27:a4:cd:33 88:32:9b:bc:8a:21 02:08:83:ab:49:f0
02:b0:4c:d1:4f:c3 02:20:ce:1e:da:5d 02:66:7a:08:7d:0c 02:a0:bf:4a:2f:da 02:14:d1:e3:31:b3

Figure 14: RSSI clusters belong to transmission pattern with quantization level 4



X: Pattern (Nom)

Y: RSSI (Num)

Colour: MAC (Nom)

Select Instance

02:cd:e5:67:38:9a 02:40:db:a9:0f:82 14:2d:27:a4:cd:33 88:32:9b:bc:8a:21 02:08:83:ab:49:f0
02:b0:4c:d1:4f:c3 02:20:ce:1e:da:5d 02:66:7a:08:7d:0c 02:a0:bf:4a:2f:da 02:14:d1:e3:31:b3

Figure 15: RSSI clusters belong to transmission pattern with quantization level 3

In the Figure 14 and 15, the x-axis represents the transmission pattern of the MAC addresses; and the y-axis represents the RSSI values recorded against the MAC addresses. MAC addresses are represented using different colors. Figure 14 illustrates that there are 3 devices involved in the experiment; each subjected to the transmission pattern marked in the x-axis. The clusters formed by the samples subjected to the transmission patterns “(12, 13, 11)” and (13, 12, 11) had involved more than one MAC addresses which is suspicious. Any MAC addresses subjected to those two transmission patterns are classified as suspicious or spoofed addresses. Which concludes the same as concluded in the approach using the quantization interval 5.

However , Figure 15 concludes that there are 7 devices present at the location, each with the transmission pattern marked in the x-axis, the transmission patterns “(15, 16,17,14)” and “(16,15,17,14)” are suspicious as the clusters formed by the both patterns reflects more than one MAC addresses. Thus all the 6 MAC addresses subjected to the mentioned transmission pattern are suspicious where all other 4 MAC addresses are genuine which increases the error possibility in the conclusion subjected to device classification and MAC detection counts compare to other two experiments conducted under quantization intervals 5 and 4.

Thus, quantization interval level 3 will be eliminated from quantization phase and quantization intervals 5 and 4 are considered during rest of the chapters.

CHAPTER 5 EVALUATION

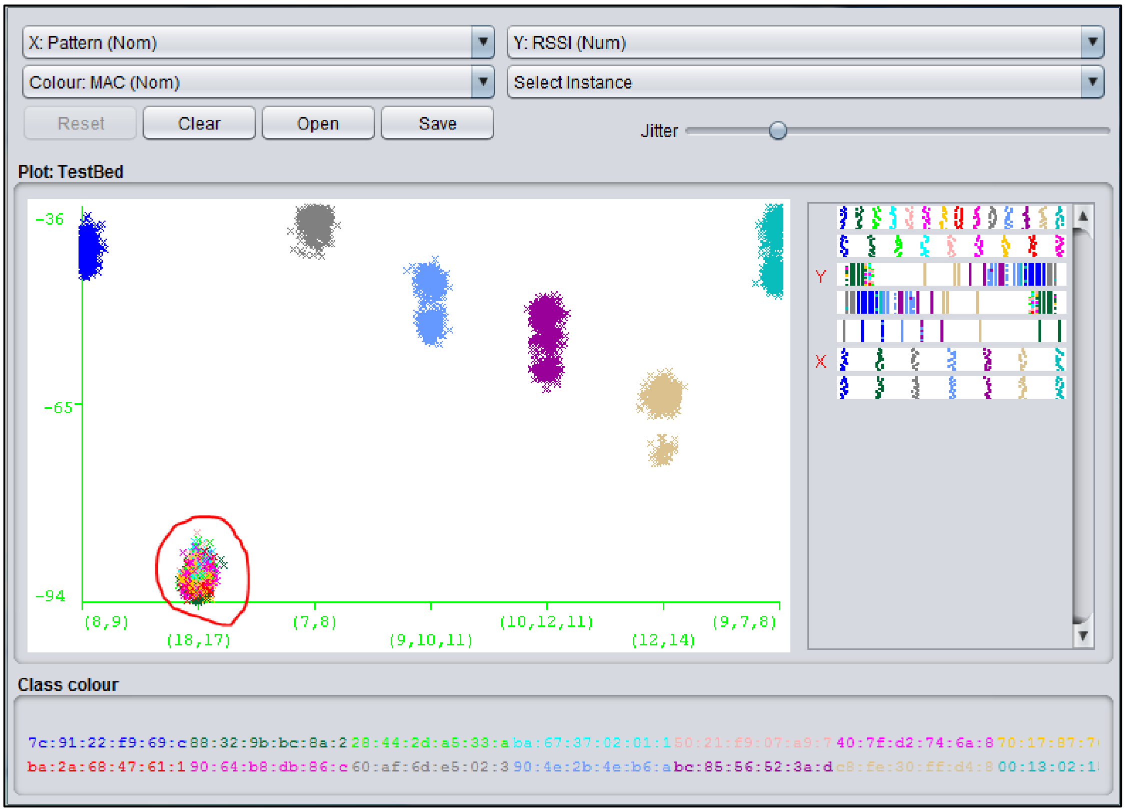
The following sections describe the different evaluation criteria and the evaluation results of the proposed RSSI based 802.11 MAC based spoof activity detection approach. The experiments are conducted to evaluate the ability of the proposed detection approach to distinguish devices between different model devices, same model devices, same vendor devices and the results at different indoor location and setup.

5.1 Experiment using more devices

A Samsung J2, a HP 2000 laptop, a HuaweiG5100200 smart phone, an AcerE554 laptop, a Samsung A5 smartphone, an EX553 TAB and the Samsung S3 are involved to the experiment setup in figure 8. The samples from devices are recorded inside the AP for 20 minutes using the bash script shown in Figure 4. The tool Wifi Mac Changer shown in Figure 6 is used on Samsung S3 to change its original MAC addresses 88:32:9b: bc: 8a:21 and conduct the spoofing activity during the sampling process. The MAC addresses of the devices and the transmission patterns obtained for each MAC addresses via the pattern identification process under quantization intervals 5 and 4 are listed in table 3. The clustering of RSSI samples obtained from WEKA for the transmission pattern subjected to quantization intervals 5 and 4 are also shown in Figure 16 and Figure 17 respectively.

Devices	MAC Addresses	Transmission pattern under different quantization level	
		Q5	Q4
Samsung S3	28:44:2d:a5:33:a0	18,17	21,22,20
	40:7f:d2:74:6a:8a	18,17	21,22,20
	50:21:f9:07:a9:77	18,17	21,22,20
	70:17:87:76:36:40	18,17	21,22,20
	88:32:9b:bc:8a:21	18,17	21,22
	90:64:b8:db:86:ce	18,17	21,22
	ba:2a:68:47:61:10	18,17	21,22,20
	ba:67:37:02:01:11	18,17	21,22
Samsung A5	60:af:6d:e5:02:3c	7,8	8,9
Huawei G5100200	90:4e:2b:4e:b6:aa	9,10,11	11,12,13
HP2000	bc:85:56:52:3a:d7	10,12,11	12,14,13
EX 553 TAB	c8:fe:30:ff:d4:85	12,14	15,17
AcerE554	00:13:02:15:b5:31	9,7,8	10,9,11,8
Samsung J2	7c:91:22:f9:69:cf	8,9	9,10

Table 3: Transmission patterns of multiple devices from different quantization interval



X: Pattern (Nom)

Y: RSSI (Num)

Colour: MAC (Nom)

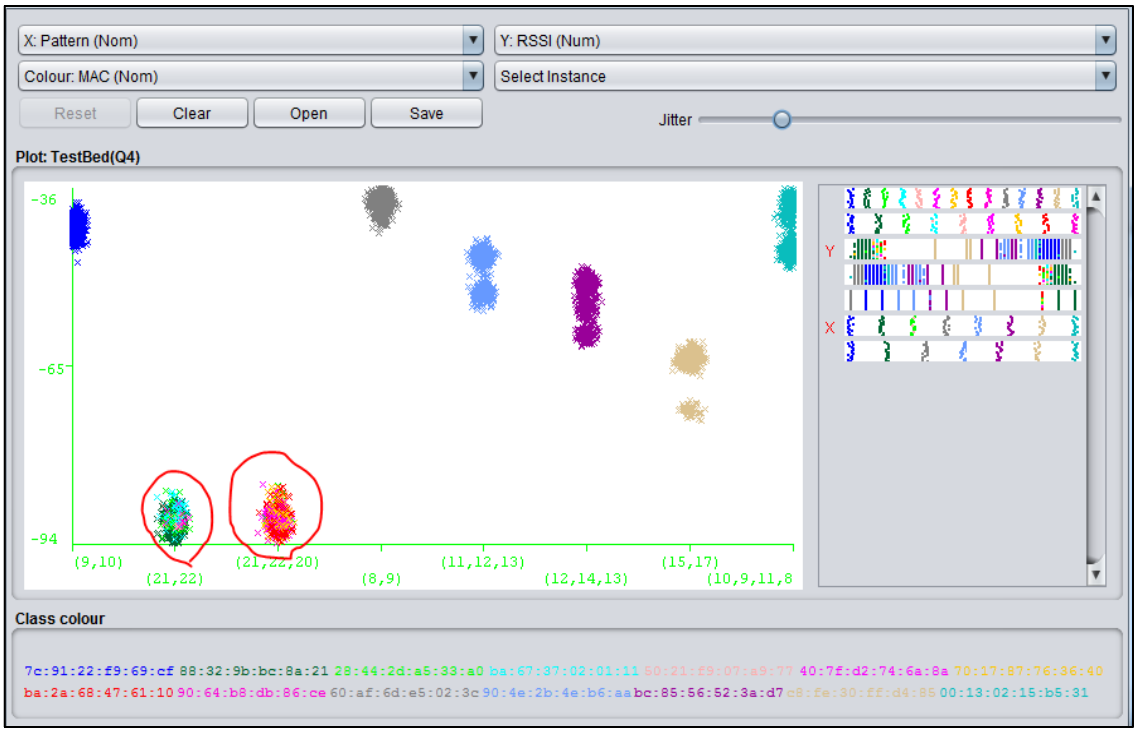
Select Instance

Plot: TestBed

Class colour

7c:91:22:f9:69:c 88:32:9b:bc:8a:2 28:44:2d:a5:33:a ba:67:37:02:01:1 50:21:f9:07:a9:7 40:7f:d2:74:6a:8 70:17:87:7
ba:2a:68:47:61:1 90:64:b8:db:86:c 60:af:6d:e5:02:3 90:4e:2b:4e:b6:a bc:85:56:52:3a:d c8:fe:30:ff:d4:8 00:13:02:1

Figure 16: RSSI clusters belong to multiple devices under quantization interval 5



X: Pattern (Nom)

Y: RSSI (Num)

Colour: MAC (Nom)

Select Instance

Plot: TestBed(Q4)

Class colour

7c:91:22:f9:69:cf 88:32:9b:bc:8a:21 28:44:2d:a5:33:a0 ba:67:37:02:01:11 50:21:f9:07:a9:77 40:7f:d2:74:6a:8a 70:17:87:76:36:40
ba:2a:68:47:61:10 90:64:b8:db:86:ce 60:af:6d:e5:02:3c 90:4e:2b:4e:b6:aa bc:85:56:52:3a:d7 c8:fe:30:ff:d4:85 00:13:02:15:b5:31

Figure 17: RSSI clusters belong to multiple devices under quantization interval 4

In the Figure 16 and Figure 17, the x- axis represents the obtained transmission patterns of the devices subjected to the quantization intervals and the y-axis represents the RSSI sample values recorded from the devices. Each MAC address is represented using different colours; the colours in the clusters highlight the MAC addresses accountable for the samples in that clusters. Any clusters having more than one color noted suspicious and circled by the red colour.

Figure 16 illustrates that there are 7 devices involved in the experiment; each with the transmission pattern marked in the x-axis. The device with the transmission pattern of "(18, 17)" had used more than one MAC addresses where the other devices had used only one MAC addresses. Meanwhile Figure 17 explains that there are 8 devices involved in the experiments, each with the transmission pattern marked in the x-axis, the devices with the transmission patterns "(21, 22)" and "(21, 22, 20)" are suspicious as both the devices reflects more than one MAC addresses within the cluster while other devices exhibits only one MAC addresses.

As a conclusion Figure 16 reveals the truth of expected device classification count 7, identification of suspicious device count 1 and the detection of suspicious mac address 8. Whereas Figure 17 reveals that counts as 8, 2 and 8 which creates a 14.9%, 50% and 0% deviation from the actuals. However when observe the transmission patterns (21, 22) and (21, 22, 20), both patterns have the value (21, 22) in common which gives a hint that both transmission patterns could possibly belongs to one device. Having said that the figure 17 also gives the correct expected device classification count 7, identification of suspicious device count 1 and the detection of suspicious mac addresses 8.

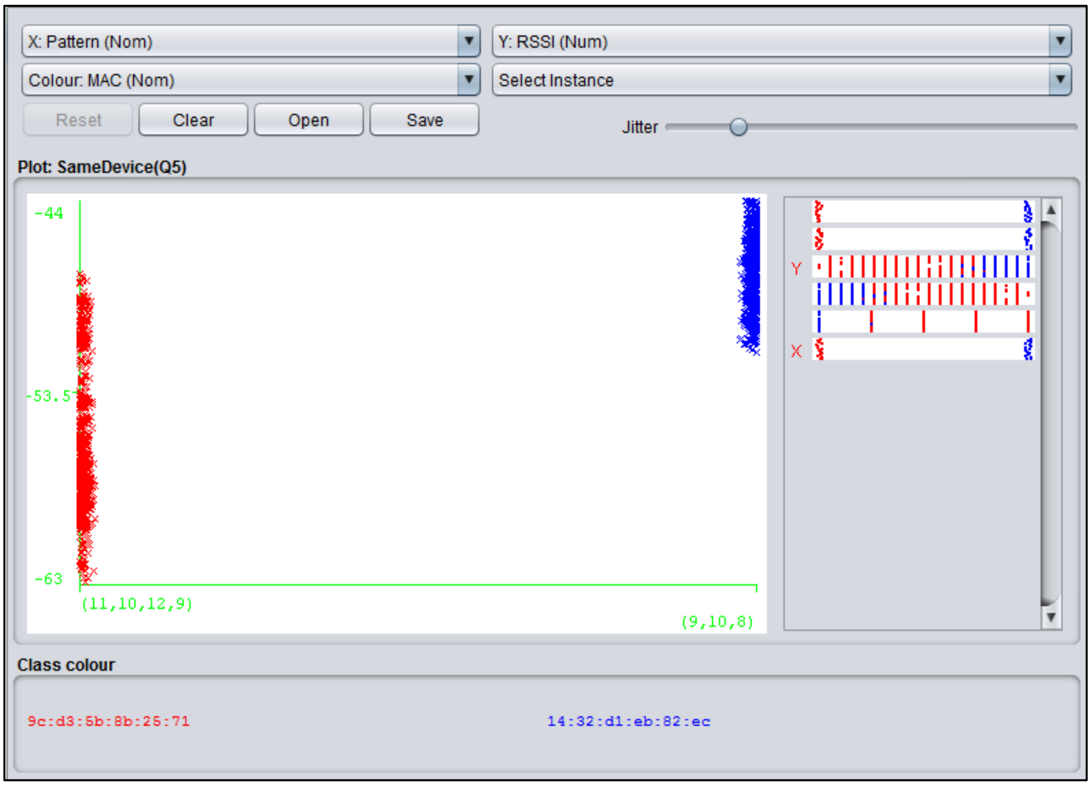
5.2 Experiment using same model devices

The underlying objective of this evaluation part is to examine the ability of the proposed approach to separate two same model stationary devices presence at a location. Two Samsung Galaxy Grand Prime smartphones have same firmware installed are involved to the experiment setup in Figure 8. Both devices are kept close enough to each other on the device location shown in Figure 8 and recorded the RSSI samples from the devices inside the AP for 30 minutes using the script shown in Figure 4. The owner of the devices did not allow installing applications or changing the device MAC addresses. Thus, the spoofing activity is not carried out from either device during the sampling process. The recorded samples are then passed into the quantization and pattern identification processes and followed by the clustering component. The transmission pattern of the devices obtained at pattern identification phase under quantization intervals 5 and 4 are tabulated in Table 4.

Devices	MAC Addresses	Transmission pattern under different quantization levels	
		Q5	Q4
GalaxyGrandPrime 1	14:32:d1:eb:82:ec	9,10,8	11,10,12
GalaxyGrandPrime 2	9c:d3:5b:8b:25:71	11,10,12,9	13,14,12,11,15

Table 4: Transmission patterns of same devices under different quantization interval

The final clustering output of the samples obtained via WEKA under quantization intervals 5 and 4 are shown in Figure 18 and 19 respectively.



X: Pattern (Nom)

Y: RSSI (Num)

Colour: MAC (Nom)

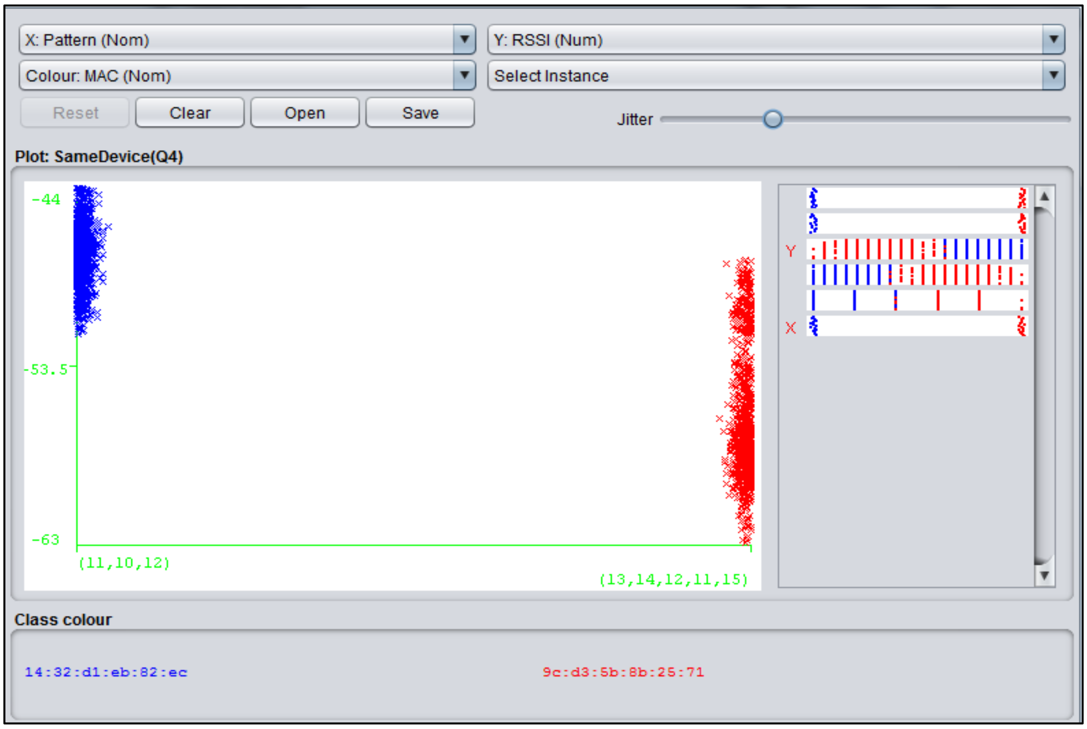
Plot: SameDevice(Q5)

Class colour

9c:d3:5b:8b:25:71

14:32:d1:eb:82:ec

Figure 18: Clusters from same devices under quantization interval 5



X: Pattern (Nom)

Y: RSSI (Num)

Colour: MAC (Nom)

Plot: SameDevice(Q4)

Class colour

14:32:d1:eb:82:ec

9c:d3:5b:8b:25:71

Figure 19: Clusters from same devices under quantization interval 4

In the Figure 18 and Figure 19, the x- axis represents the obtained transmission patterns of the devices subjected to the quantization intervals and the y-axis represents the RSSI sample values recorded from the devices. The red color and blue color clusters represent the group of samples from the MAC addresses 9c:d3:5b:8b:25:71 and 14:32:d1:eb:82:ec respectively.

This experiment concludes that regardless of the quantization intervals both devices can be differentiated from each other using the proposed detection approach.

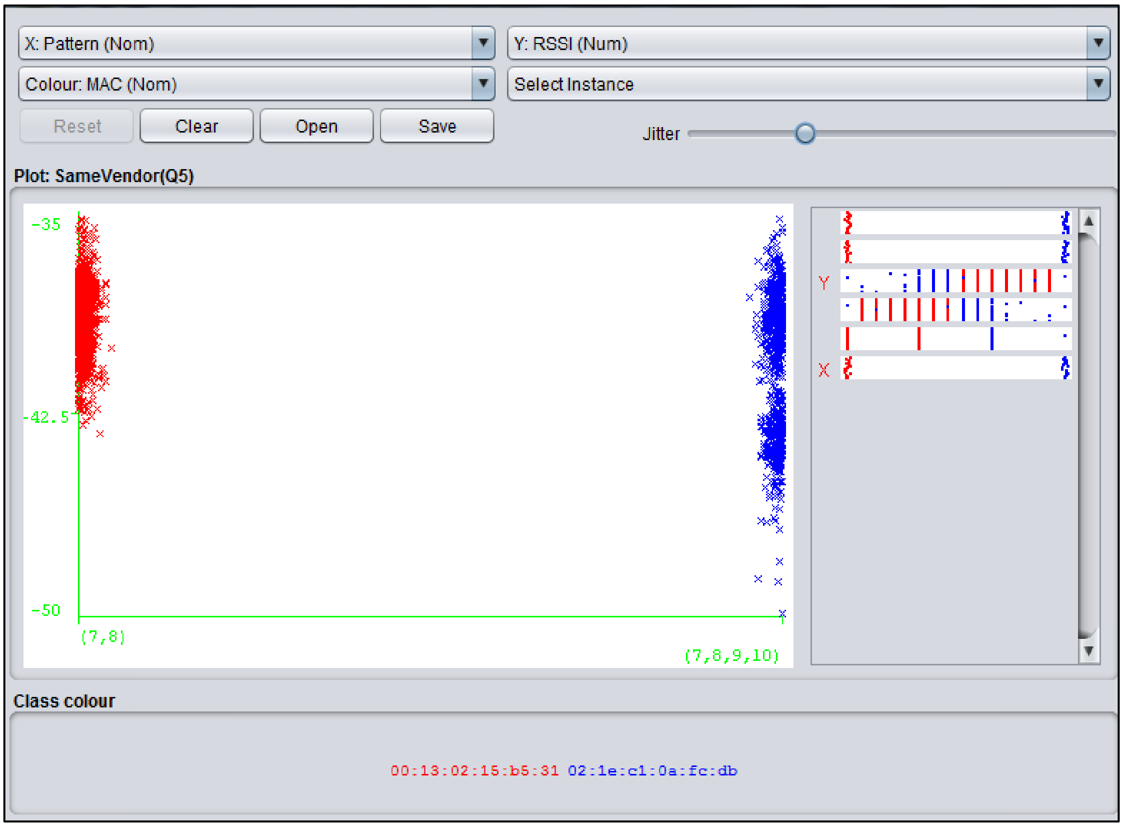
5.3 Experiment using different models of devices from a vendor

This experiment is carried out to examine the ability of the proposed detection approach to differentiate two different models of devices from one vendor. The Acer laptops E5471 and E554 are selected for the experiments and placed close enough to each other in the experiment setup shown in Figure 8. The RSSI samples of the devices are recorded for 30 minutes inside the AP using the script shown in Figure 4. Since chapter 4 has already discussed the ability of the approach to detect and classify the device AcerE5471 regardless of its spoofed MAC addresses, spoofing is not carried out during the sampling process. The recorded samples are then passed into the quantization and pattern identification phases of the proposed detection approach to obtain the transmission patterns of the devices subjected to the quantization intervals 5 and 4. Finally the populated dataset is processed by the clustering component. The transmission patterns of the devices obtained at pattern identification phase under quantization intervals 5 and 4 are tabulated in Table 5. The clusters subjected to each transmission pattern under quantization intervals 5 and 4 are shown in figure 20 and 21 respectively.

Devices	MAC Addresses	Transmission pattern under different quantization levels	
		Q5	Q4
Acer554	00:13:02:15:b5:31	(7,8)	(9,8,10)
AcerE5471	02:1e:c1:0a:fc:db	(7,8,9,10)	(9,10,8,11)

Table 5: Transmission Pattern of two different models from the vendor Acer

In the Figure 20 and Figure 21, the x- axis represents the obtained transmission patterns of the devices subjected to the quantization intervals and the y-axis represents the RSSI sample values recorded from the devices. The red color and blue color clusters represent the group of samples from the devices with the MAC addresses 00:13:02:15:b5:31 and 02:1e:c1:0a:fc:db respectively.



X: Pattern (Nom)

Y: RSSI (Num)

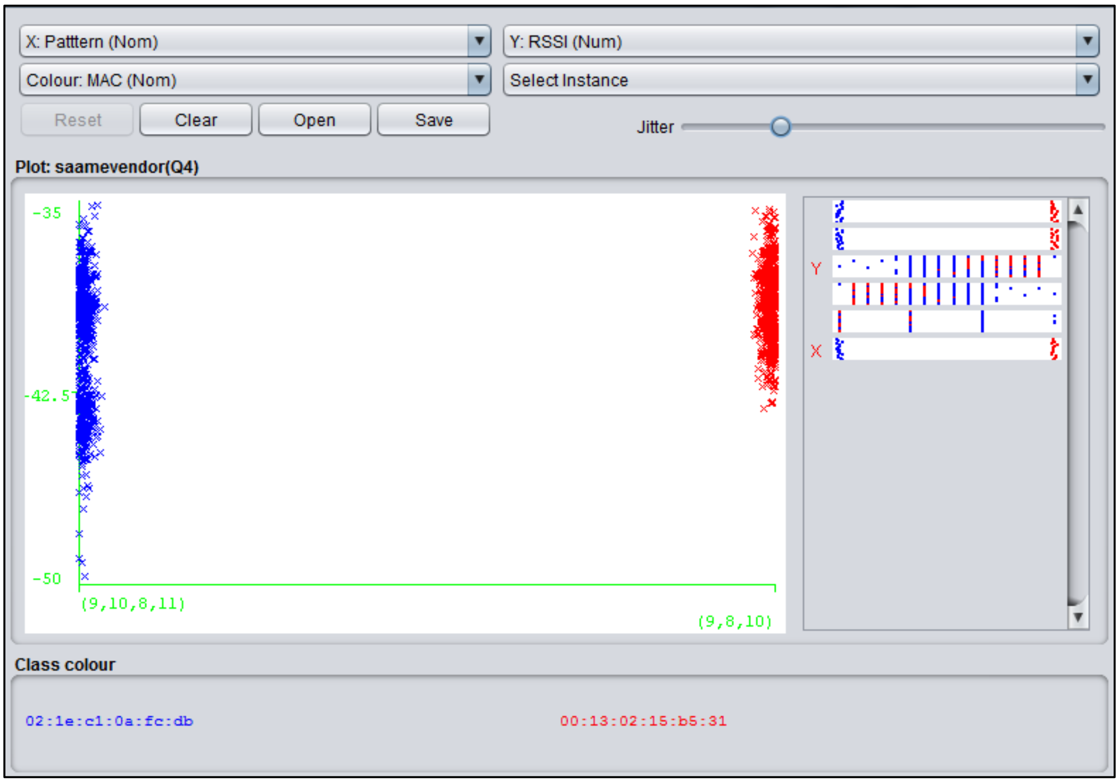
Colour: MAC (Nom)

Select Instance

Plot: SameVendor(Q5)

Class colour

Figure 20: Same vendor different model devices under quantization level 5



X: Patttern (Nom)

Y: RSSI (Num)

Colour: MAC (Nom)

Select Instance

Plot: saamevendor(Q4)

Class colour

Figure 21: Same vendor different model devices under quantization level 4

The Figure 20 and Figure 21 concludes that the proposed approach is capable of differentiate the samples from two different model of devices from a vendor regardless of the quantization intervals proposed. However when observe the transmission patterns (7,8) and (7,8,9,10) obtained under quantization interval 5, both patterns start using (7,8). If the patterns are interpreted using a line chart as shown in Figure 22,it gives a hint that when both the devices started following the same transmission pattern at some extend as circled in Figure 22 the output will confuse the examiner. Having said that, this evaluation identifies a better output from the proposed detection approach under quantization interval 4.

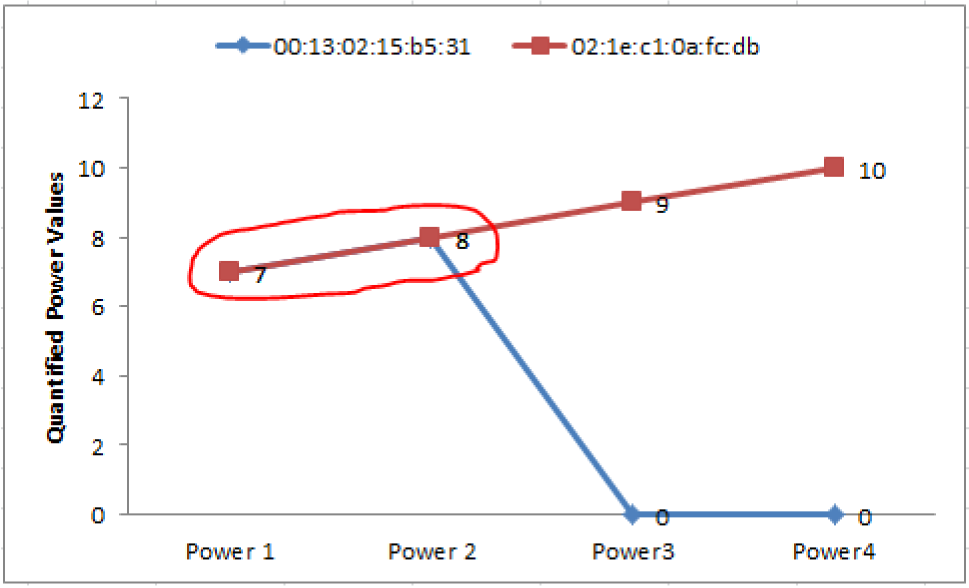


Figure 22: Line chart representation of patterns (7, 8) and (7, 8, 9, 10)

5.4 Experiment at different indoor location

Indoor office environments represent challenging conditions due to non-line of sight between devices and the AP that causes multipath fading effects. The extent of these effects may also vary over time, e.g., due to people moving, thereby increasing a considerable variance in received signal strength of the devices on the AP. This study is solely relies on RSSI observations to detect spoofing activity, thus, it is important to characterize its output in such office environment.

The setup implemented at an office cubical cabin floor to continue this evaluation is shown in Figure 23. The AP is placed at 9.4 feet on the ceiling to provide the Omni directional coverage to the devices. An Apple 6S, a Windows Lumina, an Apple 5S, a Huawei Corner 4X, the Samsung S3and the AcerE5471 laptop are involved in the evaluation process. The specification of the location and the placements of the AP, Router, and the devices are shown in Figure 23.

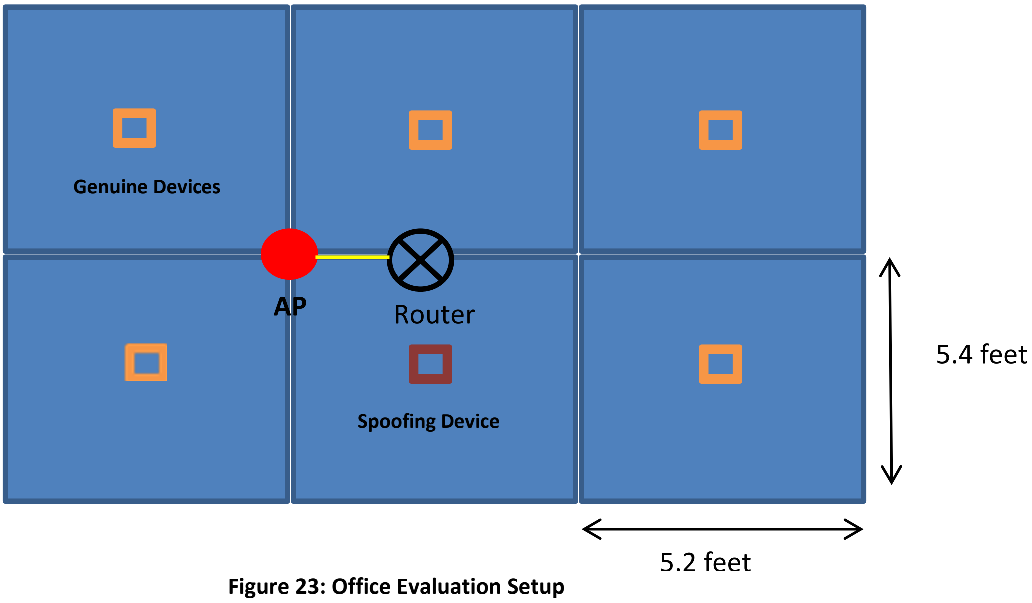


Figure 23: Office Evaluation Setup

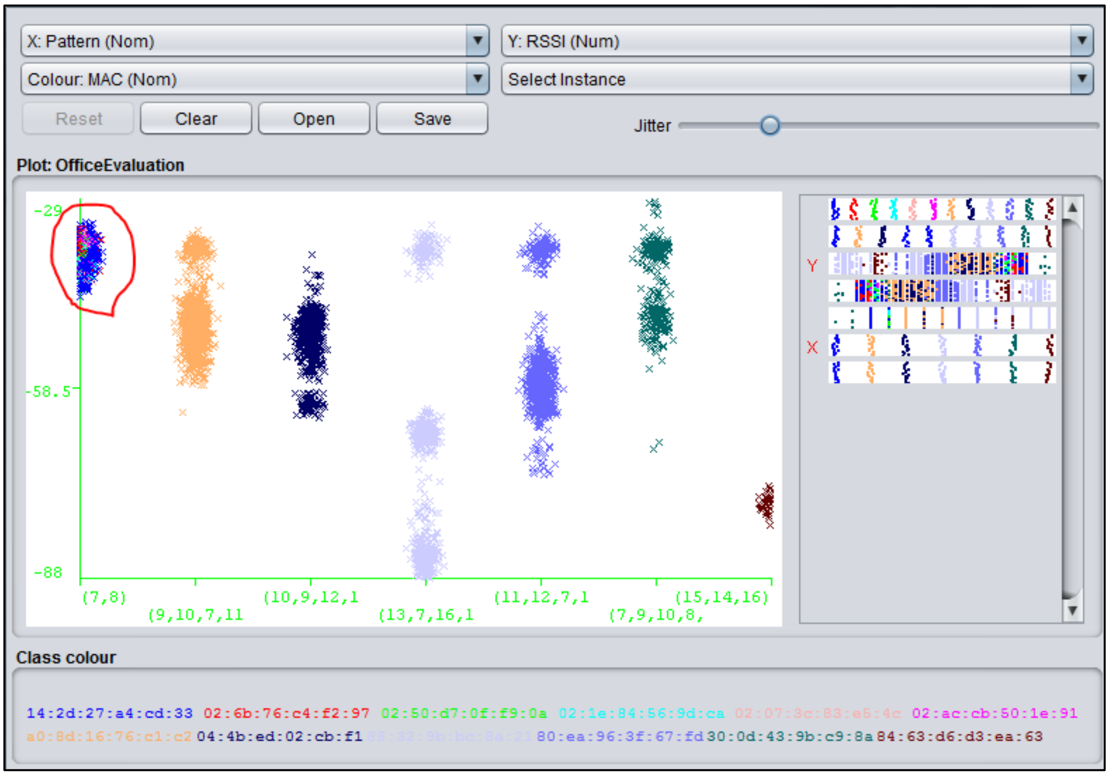
The RSSI samples of the devices are recorded for 20 minutes inside the AP using the script shown in Figure 4. During the sampling period the AcerE5471 is used to conduct the MAC spoofing activity at random time intervals using the tool TMAC v6 shown in Figure 5.

The recorded samples are then processed by the quantization and pattern identification phases in the proposed detection approach to identify the transmission pattern of each device. During this processing an unknown MAC address belongs to a Microsoft device also identified additionally connected to the setup .Thus it is also considered as an additional device for this evaluation process. The transmission patterns of the devices obtained through pattern identification phases subjected to quantization intervals 5 and 4 are shown in Table 6.

Devices	MAC Addresses	Transmission pattern under different quantization levels	
		Q5	Q4
Apple 6S	04:4b:ed:02:cb:f1	(10,9,12,11,8,7)	(11,12,14,13,10,9,8)
Windows Lumina	30:0d:43:9b:c9:8a	(7,9,10,8,6,11,13,5)	(8,11,10,12,9,6,7,16,13)
Apple 5S	80:ea:96:3f:67:fd	(11,12,7,10,13,14,8)	(13,14,8,12,15,16,17,9,11)
An Unknown MS Device	84:63:d6:d3:ea:63	(15,14,16)	(18,17,19)
S3	88:32:9b:bc:8a:21	(13,7,16,17,12,15,8,14)	(15,20,8,16,19,18,21,9,17)
Huawei Corner 4X	a0:8d:16:76:c1:c2	(9,10,7,11,8,12)	(11,12,10,8,13,9,15)
AcerE5471	14:2d:27:a4:cd:33	(7,8)	(8,9)
	02:07:3c:83:e5:4c	(7,8)	(8,9)
	02:1e:84:56:9d:ca	(7,8)	(8,9)
	02:50:d7:0f:f9:0a	(7,8)	(8,9)
	02:6b:76:c4:f2:97	(7,8)	(8,9)
	02:ac:cb:50:1e:91	(7,8)	(8,9)

Table 6: Transmission Pattern of devices at office under different quantization levels

Finally the populated dataset is fed into the clustering component to analyze the data and get the final output of the process. The clustering output of the processed data provided by the WEKA subjected to quantization intervals 5 and 4 are shown in Figure 24 and 25 respectively.



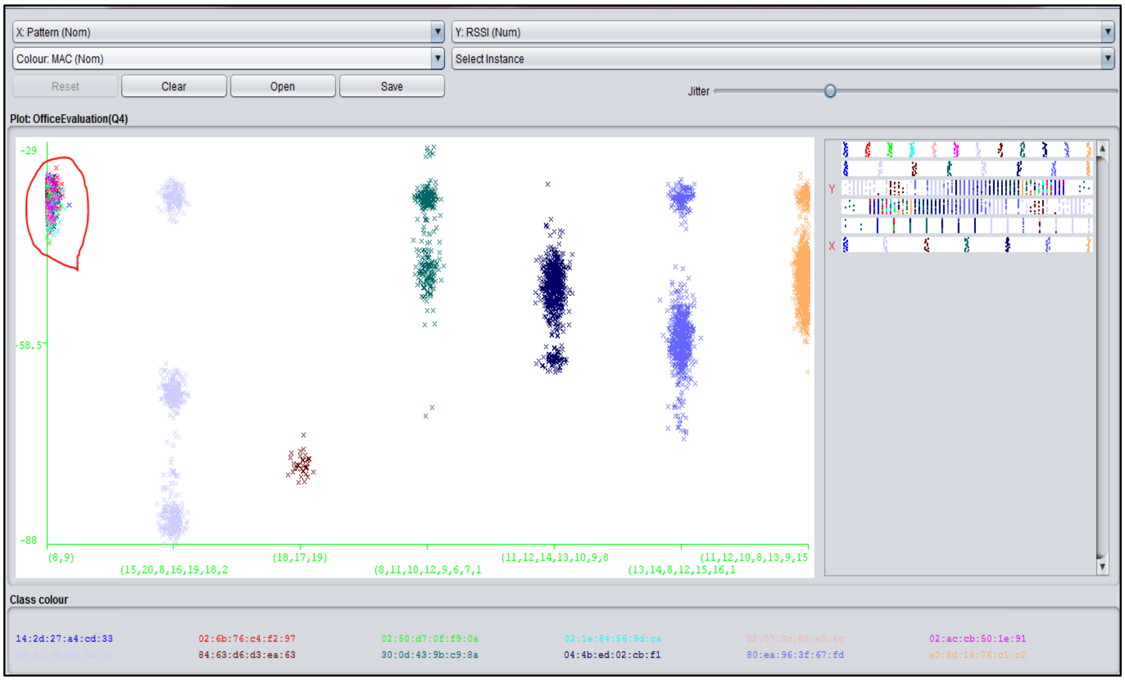
X: Pattern (Nom)
Y: RSSI (Num)
Colour: MAC (Nom)
Select Instance

Plot: OfficeEvaluation

Class colour

14:2d:27:a4:cd:33 02:6b:76:c4:f2:97 02:50:d7:0f:f9:0a 02:1e:84:56:9d:ca 02:07:3c:83:e5:4c 02:ac:cb:50:1e:91
a0:8d:16:76:c1:c2 04:4b:ed:02:cb:f1 88:32:9b:bc:8a:21 80:ea:96:3f:67:fd 30:0d:43:9b:c9:8a 84:63:d6:d3:ea:63

Figure 24: Device clusters at office under quantization level 5



X: Pattern (Nom)
Y: RSSI (Num)
Colour: MAC (Nom)
Select Instance

Plot: OfficeEvaluation(Q4)

Class colour

14:2d:27:a4:cd:33 02:6b:76:c4:f2:97 02:50:d7:0f:f9:0a 02:1e:84:56:9d:ca 02:07:3c:83:e5:4c 02:ac:cb:50:1e:91
88:32:9b:bc:8a:21 84:63:d6:d3:ea:63 30:0d:43:9b:c9:8a 04:4b:ed:02:cb:f1 80:ea:96:3f:67:fd a0:8d:16:76:c1:c2

Figure 25: Device clusters at office under quantization level 4

In Figure 23 and Figure 24, the x- axis represents the obtained transmission patterns of the devices subjected to the quantization intervals and the y-axis represents the RSSI sample values recorded from the devices. Each MAC address is represented using different colours; the colours in the clusters highlight the MAC addresses accountable for the samples in that clusters. Thus, the clusters having more than one colour noted suspicious and circled by the red color.

Both Figures 23 and 24 illustrates that there are 7 devices present at the location. Each with the transmission pattern marked in the x-axis. The device with the transmission pattern (7,8) and (8,9) subjected to quantization intervals 5 and 4 had used more than one MAC addresses inside the location where the other devices had used only one MAC addresses. Thus, referring to the Table 6, the identified suspicious device represented by the transmission pattern (7, 8) and (8, 9) subjected to Q5 and Q4 is AcerE5471, which is true. However referring Q4 column of the Table 6, the Apple 6s and the Huawei Corner 4X devices have transmission patterns started from (11,12). Consider the line chart interpretation of the two transmission patterns shown in Figure 26. When both devices' transmission pattern relies inside the area as circled in the Figure 26, the detection and the classification output under quantization interval 4 gives false outputs after some extent.

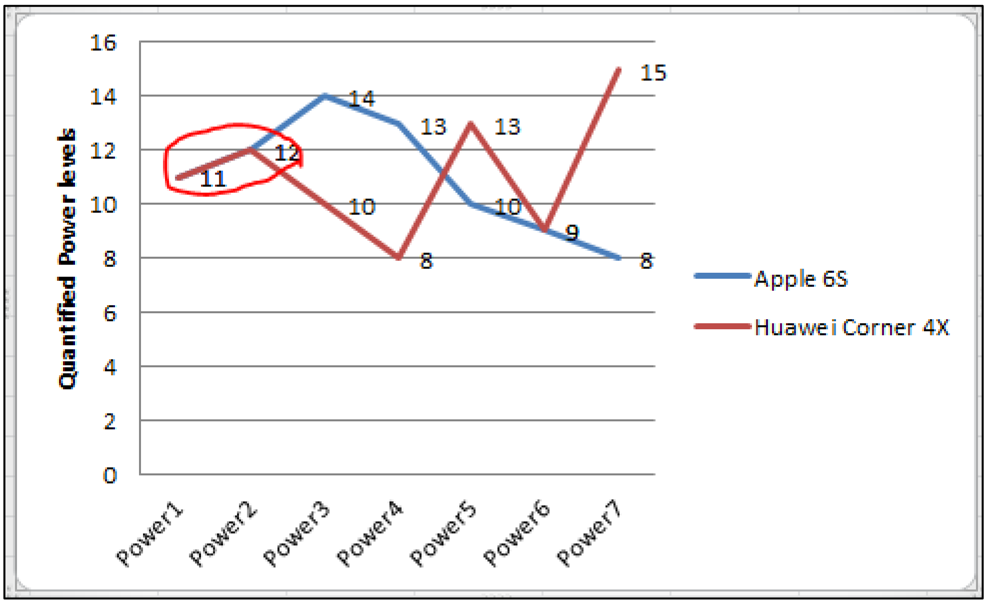


Figure 26: Transmission Pattern of Apple6s and Huawei 4X under Q4

Thus, this evaluation activity concludes that the output is better under quantization interval 5 when the proposed approach is used to classify the devices and detect spoofing activities solely based on the RSSI samples obtained from the devices stationary and resides inside an equally separated cubical cabin indoor environment.

CHAPTER 6 CONCLUTION & FUTURE WORK

6.1 Conclusion

This study has gone through a comprehensive literature view on existing MAC spoof detection methods, transceiver based spoof detection methods, existing classification models, approaches and identified the existing limitations and understands the impact of the energy in the device on the RSSI and ultimately has proposed a suitable model to identify 802.11 MAC spoof activity solely based on devices' RSS values. However, the proposed model is applicable on circumstance where the attacker is stationary at an indoor environment, surrounded by either stationary devices or none and sent legitimate requests towards a victim target using randomly generated MAC addresses.

The study has adopted the steps sampling, quantization, pattern recognition and clustering as the major components of the model and tested initially on the test bed implemented inside an indoor room consists of an AP, a router with a DHCP service as the target, a laptop to conduct the spoofing activity and a stationary smartphone as a reference.
The samples from the devices were recorded inside the AP at every 2 seconds using a bash script while spoofing the MAC address of the laptop. Then, the obtained data set is fed into an excel sheet to quantify the samples and identify the devices' transmission pattern using pivot table feature in the Excel. To obtain an initial classification between the transmission patterns of the two devices, the study has started the initial quantization interval from level 5 and adjusted to levels 4 and 3 to identify the optimal interval for better classification. Finally the updated dataset was fed into WEKA to examine the clustering output.
Both clustering outputs under quantization intervals 5 and 4 had a 50% deviation from the expected classification count and 87.5 % correct detection of suspicious MAC addresses. However, the output under quantization interval 3 had further deviated the expected classification count by 200 % and reduced the correct number if suspicious MAC addresses by 50%. Thus, the quantization interval 3 was ignored at the later part of the research.

However, at the phase of evaluation, the experiments in real, using 15 different devices have proven the capability of the proposed approach to distinguish the RSS samples from multiple model devices, same model devices, and different model devices provide by same vendor and able to classify the devices which are stationary at indoor locations like office cabin. The evaluation part concludes that regardless of the quantization intervals same model device can be differentiated where quantization interval 4 is highly recommended to differentiate different model of devices from the same vendor. However, both quantization intervals 5 and 4 shall be used in the approach to differentiate multiple devices stationary at any indoor location but quantization interval 5 gives a better output compare to Q4.

6.2 Future Work

The proposed approach also can be used to classify stationary devices when the attacker conducts a spoofing activity using another existing stationary 802.11 MAC addresses from the location.

Meantime the proposed approach can be examined on circumstances where

- The attacker is alone but not stationary.
- The attacker is not stationary among other stationary devices.
- The attacker is stationary among non-stationary devices.

However, the proposed approach can be taken further through a development with real time pattern recognition and device classification.

REFERENCES

[1]. Joshua Wright, J. Wright, “Detecting wireless LAN MAC address spoofing,” 2003, technical document.

[2].F. Guo and T. ckerChiueh, “Sequence number-based MAC address spoof detection,” in Proceedings of the 8th International Symposium on Recent Advances in Intrusion Detection, Seattle, WA, USA, Sept. 2005.

[3].J. Hall, M. Barbeau and E. Kranakis. Using Transceiver prints for Anomaly Based Intrusion Detection. In Proceedings of 3rd IASTED, CIIT 2004, November 22-24,2004, St. Thomas, US Virgin Islands.

[4] Y. Chen, W. Trappe, and R. P. Martin, “Detecting and localizing wireless spoofing attacks,” in SECON’07: Proceedings of the 4th Annual IEEE Conference on Sensor, Mesh and Ad Hoc Communications and Networks, June 2007.

[5].Yong Sheng, Keren Tan, Guanling Chen, David Kotz, Andrew Campbell. ”Detecting 802.11 MAC Layer Spoofing Using Received Signal Strength”.

[6]. VenkatRaman,Fredrick,Thiemo “Detecting and Avoiding Multiple Sources of Interference in the 2.4 GHz Spectrum”. Uppsala University and SICS, Sweden

[7]. CharalamposPapamanthou, Franco P. Preparata, and Roberto Tamassia Department of Computer Science and Center for Geometric Computing Brown University. “Algorithms for Location Estimation
based on RSSI Sampling”.

[8]. S. Zacharias, T. Newe, S. O’Keeffe, and E. Lewis, “A light weight classification algorithm for external sources of interference in ieee802.15.4-based wireless sensor networks operating at the 2.4 GHz,”IJDSN, 2014.

[9] S. Zacharias, T. Newe, S. O’Keeffe, and E. Lewis, “Identifying sources of interference in RSSI traces of a single IEEE 802.15. 4 channel,” in ICWMC, 2012

[10]ShravanRayanchu, AshishPatro, Suman Banerjee, “Detecting Non-WiFi RF Devices using Commodity WiFi Hardware”

[11] Ning Ding, Daniel Wagner, Xiaomeng Chen,” Characterizing and Modeling the Impact of Wireless Signal Strength on Smartphone Battery Drain”.