

**CHARACTERIZATION OF INTERNET TRAFFIC IN UUM  
WIRELESS NETWORKS**

**WISAM DAWOOD ABDULLAH**

**UNIVERSITI UTARA MALAYSIA**

**2012**

# CHARACTERIZATION OF INTERNET TRAFFIC IN UUM WIRELESS NETWORKS

A project submitted to Dean of Research and Postgraduate Studies Office in partial

Fulfillment of the requirement for the degree

Master of Science (Information Technology)

Universiti Utara Malaysia

By

Wisam Dawood Abdullah

## **PERMISSION TO USE**

In presenting this project in partial fulfillment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the Universiti Utara Malaysia the Library may make it freely available for inspection. I further agree that permission for copying of this project in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor or, in her absence by the dean of the Faculty of Information Technology. It is understood that any copying or publication or use of this project or parts thereof for financial gain should not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use may be made of any material from my project.

Request for permission to copy or to make use of material in this project, in whole or in part should be addressed to:

**Dean of the Awang Had Salleh Graduate School  
College of Arts and Sciences  
University Utara Malaysia  
06010 UUM Sintok  
Kedah Darul Aman**

## **ABSTRACT**

The development in communication technology and the propagation of mobile devices, lightweight, with built-in, high-speed radio access in wireless are making wireless access to the Internet the popular situation rather than a wire line. Whereas, the growth of the wireless network with additional mobile devices in the UUM and increasing number of users led to slow wireless connection. Therefore, understanding the behavior of traffic analysis helps us to develop, manage WLAN technology, and deploy. It help us to apply our workload analysis results to issues in wireless network deployment, such as capacity planning, and potential network optimizations, such as algorithms for load balancing across multiple Access Points (APs) in a wireless network. The trace composes of two parts: firstly, one that connects to the core switch in computer center which is connected with the distribution switches that link the Access Point (APs) with the wireless network at campus, and secondly, another one for the measurement of bulk data transfers and interactive data exchange between two nodes in UUM library, which had been initiated at that time. This thesis investigates the performance network and users' behavior in UUM wireless network.

## ACKNOWLEDGEMENTS

### **“In the name of Allah the Most Beneficent and Most Merciful”**

All praises and thanks to the Almighty, Allah (SWT), who helps me to finish this project, Allah gives me the opportunity, strength and the ability to complete my study for Master degree after a long time of continuous work.

No volume of words is enough to express my gratitude towards my guides, Dr. Adib M. Monzer Habbal and Dr. Masuddi Bin Mahmuddin, who have been very concerned and have aided for all the material essential for the preparation of this thesis report. They have helped me explore this vast topic in an organized manner and provided me with all the ideas on how to work towards a research-oriented venture.

I am thankful to Mr. Adi Affandi Ahmed, Mr. Mohd Samsu and Dr. Omer Abdullah, for the motivation and inspiration that triggered me for this thesis work.

I am also thankful to Prof. Dr. Zulkhairi Md Dahalim, Mr. Khairil Adli Bin Abdul Rahman, Madam Farah, and all staff in Computer Center of UUM, for to support me and help me in data capture. Finally, it would not been possible for me to complete the study and this project without the help by Allah and then supporting and encourage from my family and friends. First and foremost, my gratitude goes to my father and mother for motivation me and for their prayers for me, may Allah bless them. To my brothers (Salam, Hussam, and Bassam), nephew humam and sisters for supporting me and had a great influence to finish my master study. To My friends PhD (Student) Munadil K. Faaeq , Ahmed fareed, Ibrahim and Saifuddin Hatim thanks for standing beside me and giving support in all period of study. Thanks for all persons who helped or contributed to finish my Master program.

*Wisam*

# TABLE OF CONTENTS

PERMISSION TO USE.....	i
<b>ABSTRACT.....</b>	<b>ii</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>iii</b>
<b>TABLE OF CONTENTS .....</b>	<b>iv</b>
<b>LIST OF TABLE .....</b>	<b>vii</b>
<b>LIST OF FIGURE .....</b>	<b>viii</b>
<b>LIST OF ABBREVIATION.....</b>	<b>xi</b>
<b>CHAPTER ONE: INTRODUCTION</b>	
1.1 BACKGROUND .....	1
1.2 PROBLEM STATEMENT .....	4
1.3 RESEARCH QUESTION.....	5
1.4 PROJECT'S OBJECTIVE .....	5
1.5 SCOPE OF THE STUDY .....	5
1.6 SIGNIFICANCE OF THE STUDY.....	6
1.7 RESEARCH STRUCTURE .....	6
1.8 CONCLUSION.....	7
<b>CHAPTER TWO: LITERATURE REVIEW</b>	
2.1 INTRODUCTION .....	8
2.2 WIRELESS LOCAL AREA NETWORK (WLAN) TECHNOLOGY .....	8
2.2.1 MODULATION .....	11
2.2.1.1 ANALOG MODULATION.....	11
2.2.1.2 DIGITAL MODULATION .....	14
2.2.2 THE ELECTROMAGNETIC SPECTRUM.....	16
2.3 Wi-Fi ARCHITECTURES .....	20
2.3.1 BASIC SERVICE SET (BSS) .....	21
2.3.2 EXTERNAL SERVICE SET(ESS) .....	22
2.3.3 INDEPENDENT BASIC SERVICE SET(IBSS) .....	23
2.3.4 DIRECT-SEQUENCE SPREAD SPECTRUM (DSSS) .....	24
2.3.5 ORTHOGONAL FREQUENCY DIVISION MULTIPLE (OFDM) .....	25
2.3.6 ORTHOGONAL FREQUENCY DIVISION MULTIPLE ACCESSES (OFDMA) .....	27
2.4 OSI AND TCP/IP MODEL .....	28
2.4.1 APPLICATION LAYER .....	30
2.4.1.1 APPLICATION LAYER PROTOCOLS .....	33

2.4.2	TRANSPORT LAYER.....	40
2.4.2.1	TRANSPORT LAYER PROTOCOLS.....	41
2.5	THE IEEE 802.11 OPERATIONS .....	46
2.5.1	THE IEEE 802.11 FRAME .....	47
2.5.1.1	IEEE 802.11 FRAME TYPES .....	48
2.6	PERFORMANCE MEASUREMENT OF WIRELESS NETWORK .....	55
2.7	TYPES OF MEASUREMENT TOOLS .....	55
2.7.1	WIRELESS MONITORING .....	56
2.7.2	WIRELESS BENCHMARKING .....	57
2.8	HIGHER INSTITUTION LEARNING OF NETWORK .....	57
2.9	SUMMARY .....	60
<b>CHAPTER THREE: RESEARCH METHODOLOGY</b>		
3.1	INTRODUCTION .....	62
3.2	RESEARCH METHODOLOGY.....	62
3.3	NETWORK PHASE CONFIGURATION .....	63
3.3.1	NETWORK ENVIRONMENT .....	63
3.3.2	DATA CAPTURING.....	66
3.4	EVALUATION PHASE.....	77
3.4.1	ANALYSIS WIRELESS DATA .....	77
3.4.2	PRESENTATION PERFORMANCE METRICS AND INTERPRETATION... 94	
<b>CHAPTER FOUR: FINDINGS: PERFORMANCE</b>		
4.1	INTRODUCTION .....	96
4.2	UUM WIRELESS NETWORK MEASUREMENT .....	96
4.3	TRAFFIC COLLECTION AND ANALYSIS.....	97
4.4	USER DISTRIBUTION ACROSS THE APs .....	99
4.5	DAILY TRAFFIC PATTERN.....	100
4.5.1	AVERAGE PACKET SIZE VS. STANDARD DEVIATION OF THE SIZE OF EACH PACKET .....	105
4.5.2	TRAFFIC vs. NUMBER OF AUTHENTICATE USERS .....	105
4.5.3	THROUGHPUT VS. LOAD .....	107
4.6	BULK DATA TRANSFER AND INTERACTIVE DATA EXCHANGE.....	110
4.6.1	TCP_STREAM.....	111
4.6.2	UDP_STREAM .....	112
4.6.3	TCP_RR.....	112
4.6.4	TCP_CC.....	113
4.6.5	TCP_CRR .....	115

4.6.6	UDP-RR .....	116
4.7	SUMMARY .....	117
<b>CHAPTER FIVE: FINDINGS: USERS' BEHAVIOR</b>		
5.1	INTRODUCTION .....	119
5.2	USER'S BEHAVIOR IN UUM WIRELESS .....	119
5.3	NETWORK LAYER TRAFFICS .....	120
5.4	TRANSPORT LAYER TRAFFICS .....	124
5.5	APPLICATION LAYER TRAFFICS .....	128
5.6	APPLICATION TRAFFIC .....	139
5.7	SUMMARY .....	146
<b>CHAPTER SIX: CONCLUSION</b>		
6.1	INTRODUCTION .....	147
6.2	RESEARCH SUMMARY .....	147
6.3	PROBLEMS AND LIMITATION .....	149
6.4	CONTRIBUTION .....	150
6.5	FUTURE WORK .....	151
REFERENCES .....		152
APPENDIX .....		157



## LIST OF TABLE

Table 2.1: Radio Frequency Bands As Mentioned in .....	9
Table 2.2: The functions of each of the address fields for the four possible cases .....	51
Table 3.1: Presents the Switches supporting SPAN, RSPAN and ERSPAN .....	67
Table 3.2: The tcpdump Command-Line Options .....	74
Table 3.3: Command line options of tcpstat .....	80
Table 3.4: Substitution Strings .....	82
Table 3.5: Options for netperf .....	85
Table 3.6: Several of Wireshark filters .....	93
Table 4.1: Overall statistics for the capture .....	97
Table 4.2: The performance results per second to all week days.....	109
Table 4.3: TCP-STREAM .....	111
Table 4.4: UDP-STREAM.....	112
Table 4.5: TCP- Request/Response .....	113
Table 4.6: TCP- Connect/Close .....	114
Table 4.7: TCP-Connect/Request/Response .....	115
Table 4.8: UDP-Request/Response.....	116
Table 5.1: The top protocols responsible for the traffic.....	134
Table 5.2: The most popular applications (protocols) seen in the traffic.....	136
Table 5.3: The Application Category and Signatures .....	139

## LIST OF FIGURE

Figure 1.1: How's Wi-Fi work .....	3
Figure 1.2: Wireless networks classification .....	4
Figure 2.1: Amplitude modulation.....	12
Figure 2.2: Frequency Modulation (FM) .....	13
Figure 2.3: Phase modulation (PM) .....	13
Figure 2.4: Amplitude Shift Keying (ASK).....	15
Figure 2.5: Frequency Shift Keying (FSK).....	15
Figure 2.6: Phase Shift Keying (PSK) .....	16
Figure 2.7: The Electromagnetic Spectrum .....	17
Figure 2.8: Gain .....	18
Figure 2.9: Loss .....	19
Figure 2.10: BSS.....	22
Figure 2.11: ESS .....	23
Figure 2.12: IBSS .....	24
Figure 2.13: DSSS .....	25
Figure 2.14: OFDM modulation .....	27
Figure 2.15: Difference of OFDM and OFDMA.....	28
Figure 2.16: A comparison of the OSI and TCP/IP model .....	30
Figure 2.17: Application Layer .....	31
Figure 2.18: OSI Encapsulation Process.....	32
Figure 2.19: TCP/IP application layer protocols .....	33
Figure 2.20: Resolving DNS Addresses .....	36
Figure 2.21: Telnet Service.....	39
Figure 2.22: FTP Process.....	40
Figure 2.23: Transport layer .....	41

Figure 2.24: TCP - Segment format.....	42
Figure 2.25: UDP Datagram format.....	45
Figure 2.26: Two-Frame IEEE 802.11 Communication.....	47
Figure 2.27: The 802.11 frame .....	48
Figure 2.28: Data Frame .....	49
Figure 2.29: Control frame .....	53
Figure 2.30: Types of TCP/IP performance measurement tools.....	56
Figure 3.1: Phases of Methodology .....	63
Figure 3.2: UUM Wireless.....	65
Figure 3.3: Show the mirror port .....	69
Figure 3.4: Layers of Linux .....	71
Figure 3.5: Program Structure of tcpdump .....	73
Figure 3.6: Command line options for tcpdump .....	75
Figure 3.7: Functional diagram of tcpstat .....	79
Figure 3.8: Command line options of tcpstat.....	81
Figure 3.9: The interactions for client-server.....	85
Figure 3.10: Wireshark's works .....	89
Figure 3.11: Wireshark's GUI.....	91
Figure 3.12: Filter bar in wireshark .....	92
Figure 4.1: The traffic over the entire length of the trace .....	99
Figure 4.2: Number of authenticated users and number of active APs.....	100
Figure 4.3: High traffic Access Points/s .....	101
Figure 4.4: The high lose bandwidth during the week days .....	102
Figure 4.5: The low lose bandwidth during the week days.....	102
Figure 4.6: The packet lengths during the week days.....	103
Figure 4.7: The rates for packet length to each day of week .....	104
Figure 4.8: Average packet size vs. Standard deviation of size of packet .....	105
Figure 4.9: Number of authentication users.....	106

Figure 4.10: Total number of Packets/s .....	106
Figure 4.11: Throughput (kbps).....	108
Figure 4.12: Load.....	108
Figure 5.1: The classification of user traffic by network layer on Sunday .....	120
Figure 5.2: The classification of user traffic by network layer on Monday .....	121
Figure 5.3: The classification of user traffic by network layer on Tuesday .....	122
Figure 5.4: The classification of user traffic by network layer on Wednesday .....	122
Figure 5.5: The classification of user traffic by network layer on Thursday .....	123
Figure 5.6: The classification of user traffic by transport layer on Sunday .....	124
Figure 5.7: The classification of user traffic by transport layer on Monday .....	125
Figure 5.8: The classification of user traffic by transport layer on Tuesday .....	126
Figure 5.9: The classification of user traffic by transport layer on Wednesday .....	127
Figure 5.10: The classification of user traffic by transport layer on Thursday .....	128
Figure 5.11: The classification of user traffic by application on Sunday .....	129
Figure 5.12: The classification of user traffic by application on Monday .....	130
Figure 5.13: The classification of user traffic by application on Tuesday .....	131
Figure 5.14: The classification of user traffic by application on Wednesday .....	132
Figure 5.15: The classification of user traffic by application on Thursday .....	133
Figure 5.16: The proportion of application traffic on Sunday .....	140
Figure 5.17: The proportion of application traffic in Search Engines .....	141
Figure 5.18: The proportion of application traffic in Social Networks.....	142
Figure 5.19: The proportion of application traffic in Multimedia .....	143
Figure 5.20: The proportion of application traffic in Markets & News.....	144
Figure 5.21: The proportion of application traffic in Education .....	145

## LIST OF ABBREVIATION

AAA	Authentication, Authorization and Accounting Administrations.
AM	Amplitude Modulation.
ASK	Amplitude Shift Keying.
ASN	Access Service Network.
ASP	Application Service Provider.
AWGN	Additive White Gaussian Noise.
BER	Bit Error Rate.
BPSK	Binary Phase Shift Keying.
BS	Base Station.
DA	Destination Address
CC	Convolution Code.
CEPT	European Conference of Postal and Telecommunications.
CMIP	Common Management Information Protocol
CSN	Connectivity Service Network.
DAA	Detect and Avoid.
DKG	Dewan Kuliah Gugusan.
DL	Downlink.
DNS	Domain Name System.
DoS	Denial of Service.
DPSK	Differential Phase Shift Keying.
DPP	Dewan Penginapan Pelajar.
DPP YAB	Dewan Penginapan Pelajar Yagasan Al-Buqhari.
DSL	Digital Subscriber Line.
DSSS	Direct sequence Spread Spectrum.
DUR	Downlink to Uplink Ratio.
ECC	Electronic Communications Committee.
FCC	Federal Communications Commission.
FDM	Frequency Division Multiplexing.
FDMA	Frequency Division Multiple Access.
EDC	Executive Development Center.
FEC	Forward Error Correction.

FFT	Fast Fourier Transform.
FHSS	Frequency-hopping spread spectrum.
FTAM	File Transfer and Access Management Protocol
FTM	File Transfer Protocol
FTM	Fakulti Teknologi Maklumat
FM	Frequency Modulation.
FPAU	Fakulti Pengajian Antarabangsa dan Undang-Undang.
FSK	Frequency Shift Keying.
GW	Gateway.
HAP	High Altitude Platform.
HTTP	Hypertext Transfer Protocol
IEEE	Institute of Electrical and Electronic Engineers.
IFFT	Inverse Fast Fourier Transform.
IMS	IP Multimedia Subsystem.
IP	Internet Protocol.
ISI	Inter Symbol Interference.
ISP	Internet service provider.
ITU	International Telecommunication Union.
LAN	Local Area Network.
LOS	Line of Sight.
LTE	Long Term Evaluation.
MAC	Media Access Control.
MB-OFDM	Multiband OFDM.
Mbps	Mega bit per second.
MBWA	Mobile Broadband Wireless Access.
MFSK	Multiple Frequency Shift Keying.
MGF	Moment Generating Function.
MPSK	Multilevel Phase Shift Keying.
MS	Mobile Station.
MTRNG	Mersenne Twister Random Number Generator.
NLOS	None-Line of Sight.
NS	Network Simulator.
NSP	Network Service Provider.
NWG	Network Group.

OECD	Organization for Economic Co-operation and Development.
OFDM	Orthogonal Frequency Division Multiplexing.
OFDMA	Orthogonal Frequency Division Multiple Access.
PAPR	Peak-to-Average Power Ratio.
PE	Probability of Error.
PHY	Physical layer.
PK	Pusat Komputer.
PSD	Power Spectral Density.
PSK	Phase Shift Keying.
PSTN	Public Switched Telephone Network.
PUSC	Partially Used Sub-Carrier.
QAM	Quadrature Amplitude Modulation.
QoS	Quality of Service.
QPSK	Quadrature Phase Shift Keying
RA	Receiver Address
RNG	Random Number Generator.
RS	Reed-Solomon.
RSNA	Robust Security Network Association
SMTP	Simple Mail Transfer Protocol
SNR	Signal to Noise Ratio.
SS	Subscriber station.
STA	Station
TDMA	Time Division Multiple Access.
TA	Transmitter Address.
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UP	Uplink.
Wi-Fi	Wireless Fidelity.
WiMAX	Worldwide Interoperability for Microwave Access.
WLAN	Wireless Local Area Network.
WMAN	Wireless Metropolitan Area Network.

# **CHAPTER ONE**

## **INTRODUCTION**

In this chapter, wireless characterization is discussed in order to improve communication performance. This chapter highlights the concepts of Wi-Fi in term of standards and protocols. The attempt is to improve wireless services by study the characterization of UUM wireless network. The research problem, objectives and research questions together with significance of the study are included in this chapter.

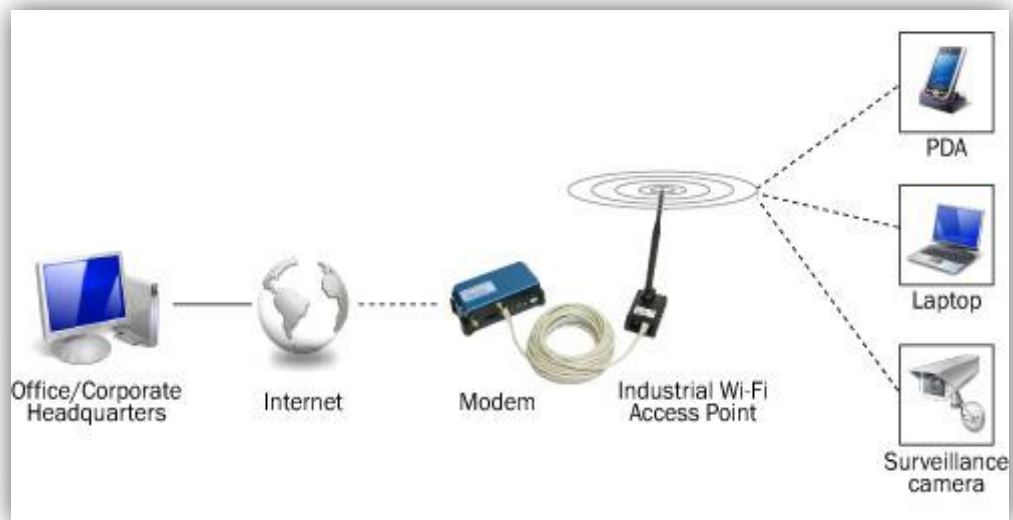
### **1.1 BACKGROUND**

The development in communication technology and the widespread use of Mobile devices that are lightweight, compact, high speed radio access in wireless technology are increasingly popularizing wireless access to the Internet. WLAN runs on IEEE 802.11 technology and are catering to connectivity in various places such as, universities, companies, corporation, and even in public places such as shopping malls, airports, lounges, and libraries, etc.; in other words, where personnel spend a considerable amount of time outside of work and home. In Malaysia, most of connections to the networks depend on wireless network and most of these rely on free frequency 2.4 and 5GHZ. The environment in this is study on the University Utara Malaysia (UUM). A few areas of information technology are developing so rapidly as that of the current Wireless - LAN (WLAN). Always, new Wireless – Standards are adopted by the demands for ever increasing data throughout and greater range [1, 2]. In 2005, there are ten completely new wireless technologies [3]. The needs for security requirements, so far doesn't indicate the signs of existence, and it is well known that wireless networking's update occurs most of the time and this includes the telecommunications field which in turn has many classifications or



categories. In the wireless networking industry environment, the manufacturers should follow regulations of the Institute of Electrical and Electronics Engineers (IEEE) and Federal Communications Commission (FCC). The expansion of the World Wide Web creates a heavier burden on the internet backbone because it had to get new technologies compatible with the regulations and meet the requirements for both. In order to be considered a Wireless producer or a marked competitor, a firm has to be familiar with requirements, and the responsibility entrusted to them, of the great challenges and convenience and flexibility of wireless location-independent connections. For that, LAN in networking became the focus of the users and manufacturers in term of flexibility, usability, bitrates throughput, frequency .etc. of performance factors, It became obligatory to reduce costs and increase efficiency to coincide with those challenges [4, 5]. There are many standard enacted networks by IEEE of Wi-Fi like 802.11a, 802.11b, 802.11g, 802.11n, each of which is covered under the Wi-Fi family and each version has its own specification which distinguishes it from the rest of the Wi-Fi family members [6]. Because of the forthcoming introduction of the Wi-Fi, the present study analyzes the wireless traffic model and its statistical parameters and investigates the traffic patterns to determine the user's behavior like number of users, protocols mix, and applications traffics, etc, in wireless network of UUM campus using traffic recorded in duration of one week. The trace composes of two parts: firstly, one that connects to the core switch in computer center which is connected with the distribution switches that link the Access Point (APs) with the wireless network at campus, and secondly, another one for the measurement of bulk data transfers and interactive data exchange between two nodes in UUM library. With that information in hand, the impression will be certainly the best clue about this field whether most areas are updated day by day or maybe less than a day.

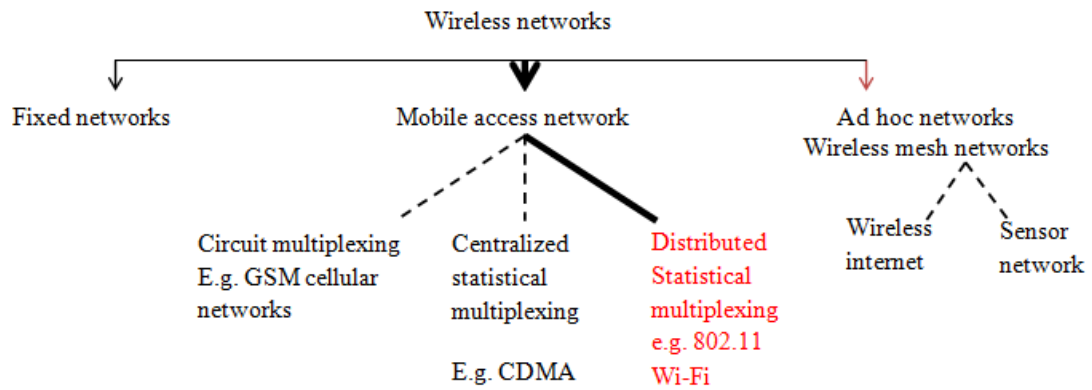
The revolution of networks lead to wireless network system which became the most important category of networking for the following reasons; availability, flexibility, mobility, etc. There are many features and criteria to make a successful Wi-Fi application that has to be taken into consideration by the developer, administrators and users as shows in Figure 1.1.



**Figure 1.1: how's Wi-Fi work [7]**

Wireless network are classified as three different types such as fixed networks, mobile access networks and ad-hoc networks "wireless mesh networks". The mobile access networks are divided into three types; circuit multiplexing (Global System for Mobile Communication (GSM) cellular networks), centralized statistical multiplexing (Code division multiple access (CDMA) cellular network). The earlier version of CDMA was Interim Standard (IS95) and is the first CDMA based digital cellular standard by Qualcomm. The brand name for IS-95 is CDMA One and known as TIA-EIA-95. The later generation is CDMA 2000 and 2- Wideband Code Division Multiple Access (WCDMA) IEEE 802.16 "WIMAX" networks.

The third type of mobile access network is distributed statistical multiplexing, while the example is IEEE 802.11 WLANs or Wi-Fi. It is also known as an ad-hoc network which consists of wireless internet and sensor network as shows in figure 1.2 [4].



**Figure 1.2:Wireless networks classification[6]**

## 1.2 PROBLEM STATEMENT

The number of hand-held devices in UUM wireless network and the amount of hand-held traffic are increasing rapidly. The lectures, reference materials' and other subjects (books) are currently and often online which increases the need for access to subjects [8, 9]. Hence, it is preferred that APs are installed in the FTM, library, EDC, DPP, DKG, and so on of buildings of the campus as most of the information invaluable to the students is currently online. But the growth of the wireless network with additional mobile devices in the UUM with increasing the number of users on wireless service or APs inhibits generic problem of wireless communication. Slow wireless connection is a popular complaint from students on this wireless network. Moreover, the uses of different web 2.0 application lead to the consumption more bandwidth.

### **1.3 RESEARCH QUESTION**

The two main questions in this study are:

- i. How does the performance of traffic wireless networks in higher institution learning perform?
- ii. How does users' behavior affect network performance?

### **1.4 PROJECT'S OBJECTIVE**

This study will be completed by achieving the following objectives:

- i. To investigate performance of network such as throughput, load, and end to end latency in UUM wireless data networks by performing real measurement.
- ii. To study the user's behavior by analysis the Internet traffic in UUM wireless network.

### **1.5 SCOPE OF THE STUDY**

Universiti Utara Malaysia is one of the universities in Malaysia considered as a government university and hence, new technology or approaches are first employed in the University for Trials by specialist teams to derive counseling immediately from experts. The data captures of 9 buildings (Fakulti Teknologi Maklumat (FTM), Pusat Komputer (PK), Dewan Kuliah Gugusan (DKG), Pusat Konvensyen, Dewan Penginapan Pelajar (DPP), Executive Development Center (EDC), Fakulti Pengajian Antarabangsa dan Undang-Undang (FPAU), Sultanah Bahiyah Library, and Dewan Penginapan Pelajar Yagasan Al-Buqhari (DPP YAB)) of UUM wireless network only because the time is not enough to study all UUM wireless network. The wireless network in UUM is operating as switched network, which means that all the devices connect to UUM campus network is

covered in the single subnet. The network at UUM is linked to the Internet through a Cisco router. Tcpdump tool was used to capture the network activity from the Access Point (APs) and at the computer center.

## **1.6 SIGNIFICANCE OF THE STUDY**

- i. The workload analysis findings can be applied to wireless network deployment like capacity planning, and potential network optimizations (e.g. algorithms for load balancing through multiple access points (APs) in a wireless network).
- ii. The results would help the team in computer center of UUM to carry out maintenance through the comprehension of usage patterns and students' behavior.
- iii. Analyzing wireless traffic is necessary to provide high quality wireless network services such as QoS management, traffic engineering, etc. The findings will assist in creating an effective model of network behavior as in reality; there is a lack of real characterization of user activity in a wireless setting.

## **1.7 RESEARCH STRUCTURE**

The study is organized as follows: Chapter One; wireless characterization is discussed in order to improve communication performance. The research problem, objectives and research questions together with significance of the study are included in this chapter. Chapter Two focuses on the wireless technologies that are used for communications which are discussed in depth for general understanding of the concept of this study. The issue regarding Wi-Fi standard, architectures for wireless transmissions and the measurement performance of wireless are discussed. Chapter

Three describes the methodology followed when gathering and analyzing the data, including steps taken to ensure user confidentiality. Chapter Four: contains the results of our analysis to UUM wireless performance .Chapter Five contains the results of our analysis to Users behavior. The study concludes in Chapter Six with summaries of our findings, problems and limitations, contribution, and future work.

## **1.8 CONCLUSION**

Era of ultra-speed which we contemporary accompanied occurrence many of the technology without a competitor are the telecommunications field and the most interesting part, Guess what? Wireless networking there is no scope to narrative it but no one still alive on this planet not use wireless tell now, the struggle and expansions vertically as well as horizontally with the most popular kind of wireless Wi-Fi, with few years after Wi-Fi established increased need for meet a requirement of developers, ISP and subscribers for many industry factors due to this. This study present usage traffics, and network performances in UUM wireless depended on a traffics captured at router in computer center for nine buildings of UUM wireless network. The goals of this study were to extend our understanding of wireless network performance and wireless behavior. In network performance to finds the throughput, load, average and end to end latency in UUM wireless and user's behavior to study web 2.0 applications. As well as learning which application and protocols consume the bandwidth, this helps us to understand if the wireless network on campus needs to improve the service to increase the amount of subscription.

## **CHAPTER TWO**

### **LITERATURE REVIEW**

#### **2.1 INTRODUCTION**

This chapter focuses on the wireless technologies that are used for communications which are discussed in depth for general understanding of the concept of this study. The issue regarding Wi-Fi standard, architectures for wireless transmissions and the measurement performance of wireless are discussed along with the issue regarding TCP/IP performance of measurement tools.

#### **2.2 WIRELESS LOCAL AREA NETWORK (WLAN) TECHNOLOGY**

A wireless LAN is term of network ability of interconnections within a limited area includes high rate of transfer data inside smaller or narrowed geographic places. WLAN interchange between two nodes or among nodes without cables and the link of connections that used by WLAN among the nodes is a special type of electromagnetic waves that known as radio waves [5, 10, 11]. Many applications use communications and wireless technology while there are no cable between nodes or base stations side and agents or subscribers sides also depends on radio waves in the exchange between senders and receivers transmissions. The radio waves transfer data through air which represent a media or the carrier. Wireless operates a radio frequency which modulates air by set of waves while this beam of connections between two nodes is not visible by human eyes [3]. The wireless technology might be considered as emerging branch of networking field due to its struggling past decades that witness important steps of development. The same principles and regulation which change during that period the transmission are depended on two major activities (send and receive) by using radio

waves or radio frequency that belong to spread spectrum of electromagnetic waves. The radio waves has a wavelength higher than infrared waves and uses an air as a media for transmission while sender and receiver have to use same channel to prescribe exchange process as implemented [5]. Radio waves applied within many fields and so many applications depend on it; TV, radio and cellular communications. Moreover, the navigation which also in contact astronauts, the wavelength of radio waves starts from little centimeters to few hundreds of meters, this variety in both wavelength and frequency give its own special features of spreading. Furthermore, many shapes of implementation of radio wave may be taken like amplitude modulation, frequency modulation and phase modulation AM/FM and PM. These modulations was still using analog till the early periods as another kind of digital implementation. AM is most familiar radio frequency often used by broadcast radio station. AM is unfortunately suffered from interference from outdoor spruces like lighting from thunderstorm, unlike FM which sometimes not generally used for data transfer and television video [12]. Radio broadcasting uses wireless signals that are scalable in frequency from 10 KHz to 30 GHz as shown in Table 2.1.

**Table 2.1:Radio Frequency Bands As Mentioned in[5].**

<b>Band</b>	<b>Frequency</b>	<b>Common uses</b>
<b>Very Low Frequency (VLF)</b>	10 KHz to 30 KHz	Maritime ship-to shore
<b>Low Frequency (LF)</b>	30 KHz to 300 KHz	Cordless telephones
<b>Medium Frequency (MF)</b>	300KHZ to 3 MHz	AM radio
<b>High Frequency (HF)</b>	3 MHz to 30 MHz	Short wave radio , CB radio
<b>Very High Frequency</b>	30 MHz to 144 MHz	TV stations 2.6,FM radio



<b>(VHF)</b>	144 MHz to 174 MHz	Taxi radio
	174 MHz to 328.6 MHz	TV stations 7-13
<b>Ultra High Frequency (UHF)</b>	328.6 MHz to 806 MHz	Public safety
	806 MHz to 960 MHz	Cellular telephones
	960 MHz to 2.3 GHz	Air traffic control radar
	2.3 GHz to 2.9 GHz	Wireless LANs *
<b>Super High Frequency (SHF)</b>	2.9GHz to 30 GHz	Wireless LANs
<b>External High frequency (EHF)</b>	30GHz and above	Radio astronomy

(\*) Wi-Fi operates free frequency such as 2.4. It takes IEEE standards 802.11b, 802.11g and 802.11n that can work on UHF as unlicensed frequency [3].

The components of radio communication consist of transducer, oscillator, antenna and modulator which are required for providing transmission process. The main function of transducer converts the information to be electronics as same as microphone in an audio communications system [12]. The oscillator is a signal producer with specific frequency which will be used for the information signal among transmission nodes. This signal has been embedded within carrier through modulation and come out from antenna as radiation signal export it for air as media for implementing the signal in space formed of electromagnetic spectrum waves [5, 13]. The receiver side has the same components to be communicated with sender in reverse order. The antenna detected signal from sender and capture it as radio signals which have been modulated from sender while waves received starting demodulated from radio signals into electrical forms. The information signal from modulate signal and demodulator

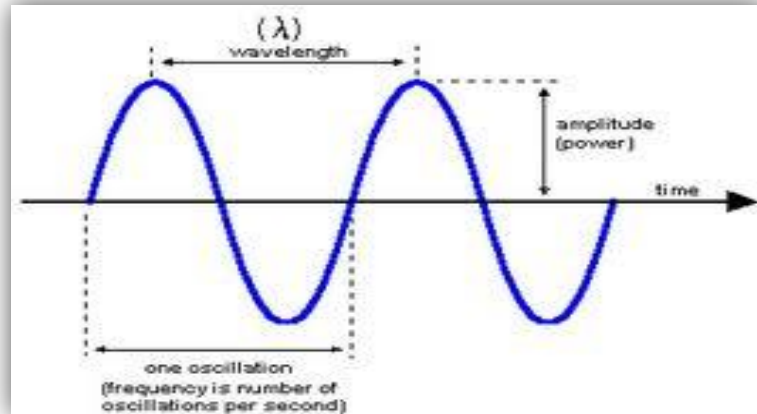
process is fetched by the oscillator process signal at carrier frequency, which is the source signal to extract the information signal [5, 13].

### **2.2.1 MODULATION**

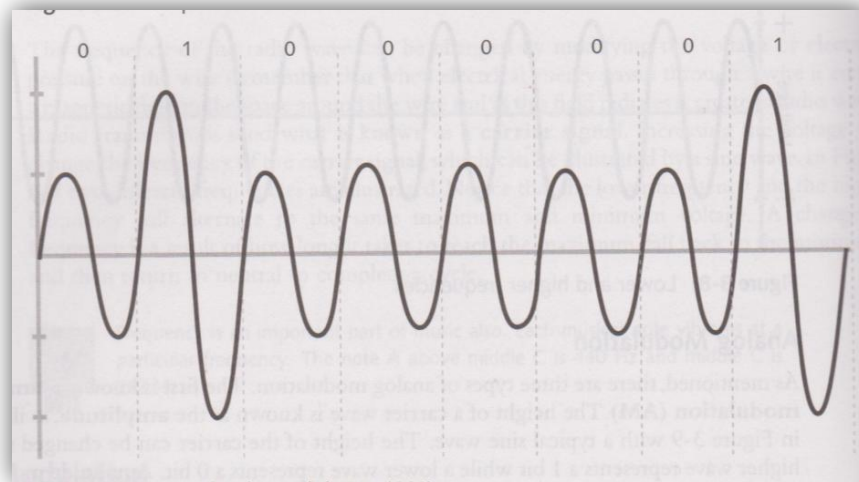
A carrier signal sent by radio transmissions carries no information. It is merely a continuous electrical signal. There are three types of modulations or changes to the signal that can be made to enable it to carry information as mentioned in section 2.1, the height, frequency and relative starting point are sometimes known as the “three degrees of freedom.” This modulation can be done in either analog or digital transmissions.

#### **2.2.1.1 ANALOG MODULATION**

This can be in diffluent type as AM, PM and FM. The Amplitude Modulation (AM) carries waves known as the amplitude as shown in Figure 2.1 with typical sine wave pattern. Figure 2.2 shows that the height of the carrier can be altered. By alteration, a higher wave is reflective of a 1 bit, while the lower wave is 0 bit. Broadcast radio stations use AM the most. Nonetheless, external sources for e.g., lightning during a thunderstorm often interfere with AM. Hence, it is not generally used for data transmissions [5, 13, 14].

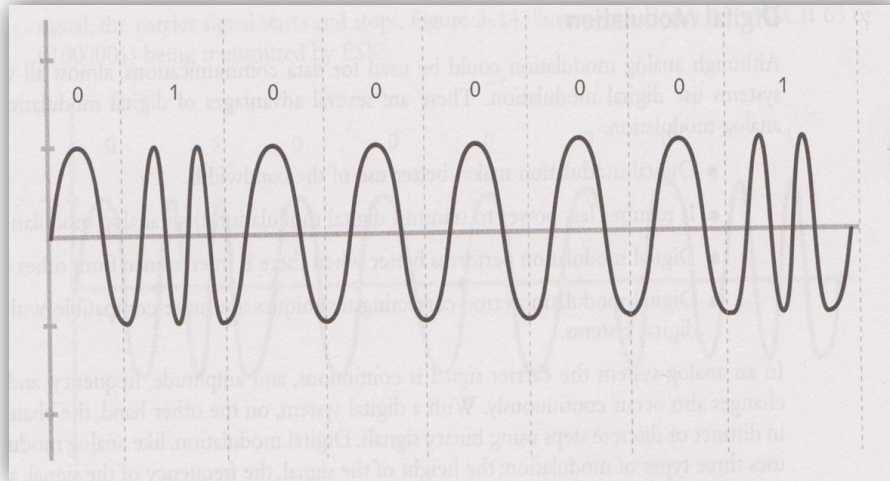


**Figure 2.1: Amplitude [5].**



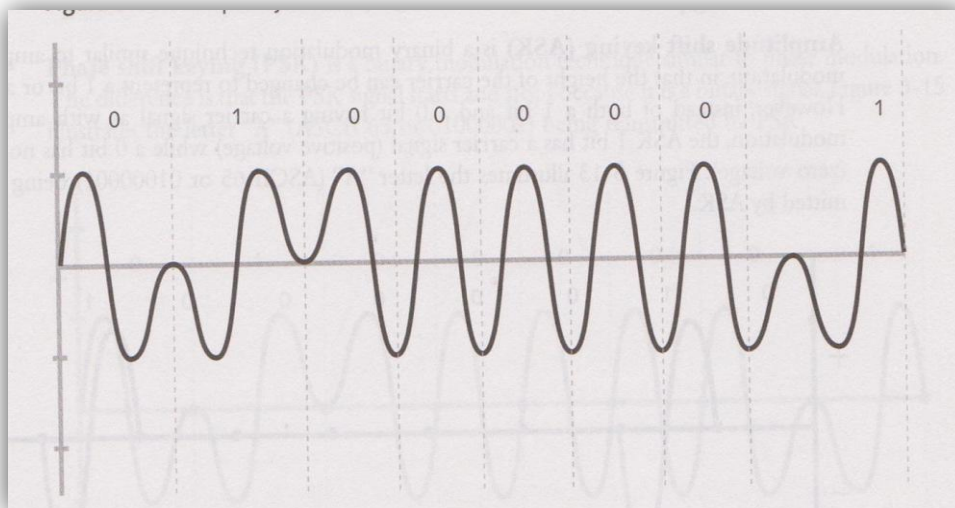
**Figure 2.2: Amplitude Modulation [5]**

In addition, AM changes with the height of the signal and the number of waves representing one cycle is altered by the frequency modulation (FM). When FM is used, the number of waves required to represent a 1-bit is greater than the number of waves that are required to represent a 0-bit. Figure 2.3, shows that similar to AM, broadcast radio station always uses FM. The difference is that, FM is not as susceptible to interference from external sources, like AM.



**Figure 2.3: Frequency Modulation (FM) [5]**

In terms of changes, AM changes the wave height and FM increases the number of waves per cycle. The starting point of the cycle is changed by phase modulation (PM). These changes only occur when the bits are transmitted from 1 bit to a 0 bit or vice versa. The change in starting point is indicative of a different bit being sent. Figure 2.4 illustrates the PM modulate.



**Figure 2.4: Phase modulation (PM) [5]**

Compared to radio broadcasts which use either AM or FM, AM, FM and PM are used by television broadcasts, whereby AM is used by the TV video, FM by the sound, and the color information by PM [5].

### **2.2.1.2 DIGITAL MODULATION**

Analog modulation could be used for data communications while almost all wireless systems use digital modulation. There are several advantages of digital modulation over analog modulation as it has been implemented in [10] and listed as:

- i Bandwidth is used better by digital modulation.
- ii Less power is required to transmit digital modulation compared to analog modulation.
- iii . When there is interference from other signals, the performance of digital modulation is better.
- iv The error-correcting techniques of digital modulation are more compatible with other digital systems.

In analog system, the carrier signal is continuous while amplitude, frequency and phase changes occur continuously. For the digital system, the changes are in distinct steps and binary signals are used. Like analog modulation, the digital modulation utilizes three types of modulation, i.e.

- i. the height of signal,
- ii. the frequency of the signal, and
- iii. the relative starting point [12, 14].

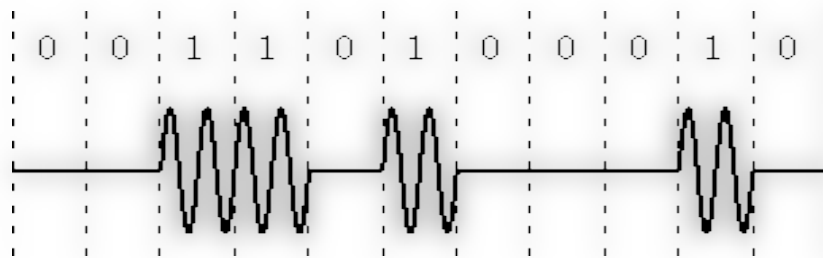
#### **• TYPES OF DIGITAL MODULATION**

The Amplitude Shift Key (ASK) is a binary modulation technique. ASK is similar to amplitude modulation. However, instead of having both a 1 bit and a 0 bit with

carrier signal (similar to amplitude modulation), the ASK 1 bit has a carrier signal which is a positive voltage, and a 0 bit has no signal which indicates Zero voltage.

Figure 2.5 shows process of transmission by ASK [15].

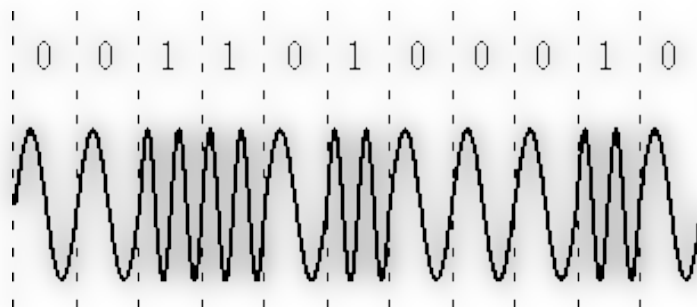
$$s(t) = \begin{cases} A \cos(2\pi f_c t) \\ 0 \end{cases}$$



**Figure 2.5: Amplitude Shift Keying (ASK) [5]**

Frequency shift keying (FSK) is similar to FM. FSK is a binary modulation technique which alters the frequency of the carrier signal. Since it transmits a binary signal, the carrier starts and stops, as illustrated in Figure 2.6 [15].

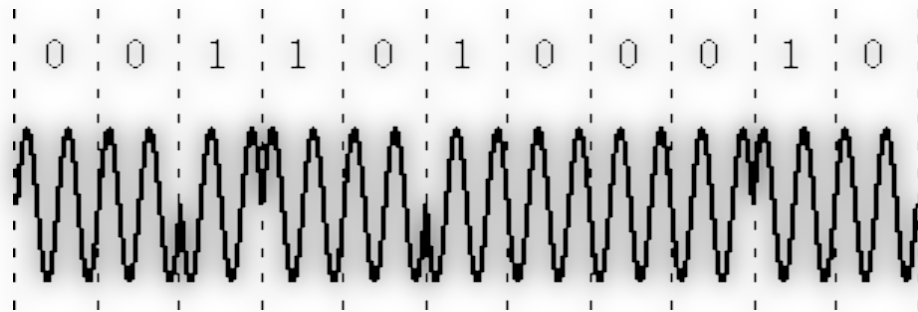
$$s(t) = \begin{cases} A \cos(2\pi (f_c + \delta) t) \\ A \cos(2\pi (f_c - \delta) t) \end{cases}$$



**Figure 2.6: Frequency Shift Keying (FSK) [5]**

Phase shift keying (PSK) is a binary modulation method like the Phase modulation in analog modulation. . The difference is that PSK signal, being a binary signal, starts and stops, as shown in Figure 2.7 [10].

$$s(t) = \begin{cases} A \cos(2\pi f_c t + 45^\circ) & 11 \\ A \cos(2\pi f_c t + 135^\circ) & 10 \\ A \cos(2\pi f_c t + 225^\circ) & 00 \\ A \cos(2\pi f_c t + 315^\circ) & 01 \end{cases}$$

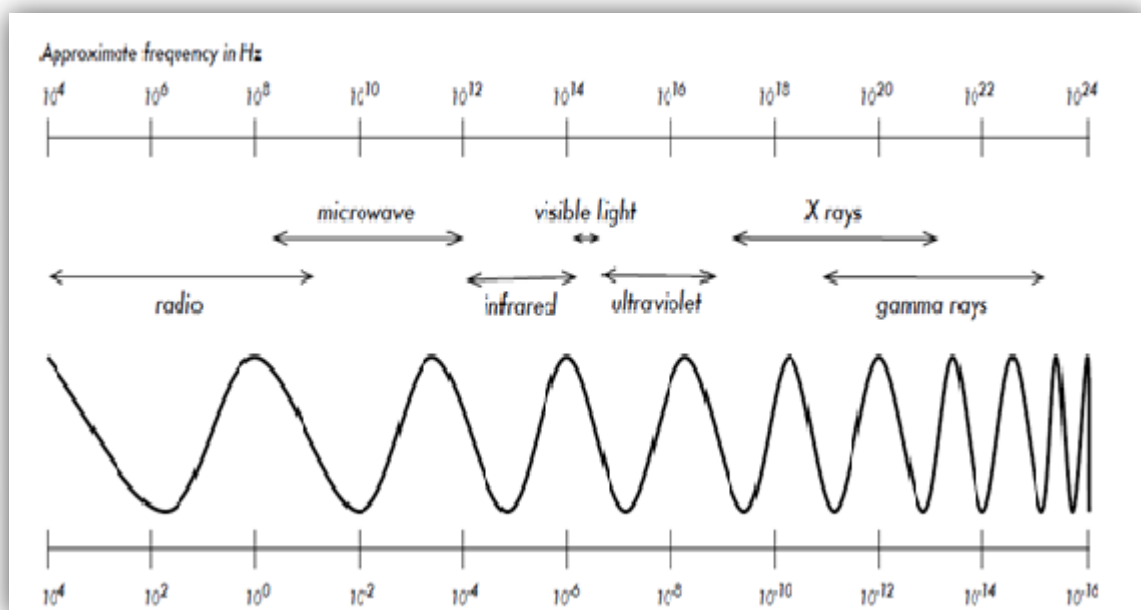


**Figure 2.7: Phase Shift Keying (PSK) [5]**

### 2.2.2 THE ELECTROMAGNETIC SPECTRUM

A wide range of frequencies (wavelengths) are spanned by electromagnetic waves. This range is known as the electromagnetic spectrum. The visible portion of the electromagnetic spectrum, or light, is the part of the spectrum that is most familiar to humans, lying roughly between the frequencies of  $7.5 \times 10^{14}$  Hz and  $3.8 \times 10^{14}$  Hz, and corresponding to wavelengths from circa 400 nm (violet/blue) to 800 nm (red). Other regions of the electromagnetic spectrum, are inclusive of alternating current (AC) or grid electricity at 50/60 Hz, Ultraviolet (on the higher frequencies side of visible light), infrared (on the lower frequencies side of visible light), X-Rays / Roentgen radiation, and many others [7, 14].

When waves are generated by the application of AC to an antenna, this is called Radio, which is the term used for the portion of the electromagnetic spectrum. This is applicable for the range from 3 Hz to 300 GHz. More specifically, the upper frequency limit will be 1 GHz. Many people tend to think of FM radio using a frequency of about 100 MHz, i.e., between radio and infrared subtend in the region of microwaves with frequencies ranging from about 1 GHz to 300 GHz, and wavelengths ranging from 30 cm to 1 mm. One of the most popular applications of microwaves is probably the microwave oven that works in the same region as the wireless standards, which lie within the bands that are kept open for general unlicensed use, called the ISM Industrial, Scientific, and Medical band. In addition, most parts of the electromagnetic spectrum are strictly monitored by licensing legislation, particularly for the parts of the spectrum which are appropriate for broadcast (TV, radio) and voice as well as data communication. The ISM bands, in many countries, are reserved for unlicensed use [1]. As shown in Figure 2.8.



**Figure 2.8: The Electromagnetic Spectrum [5]**

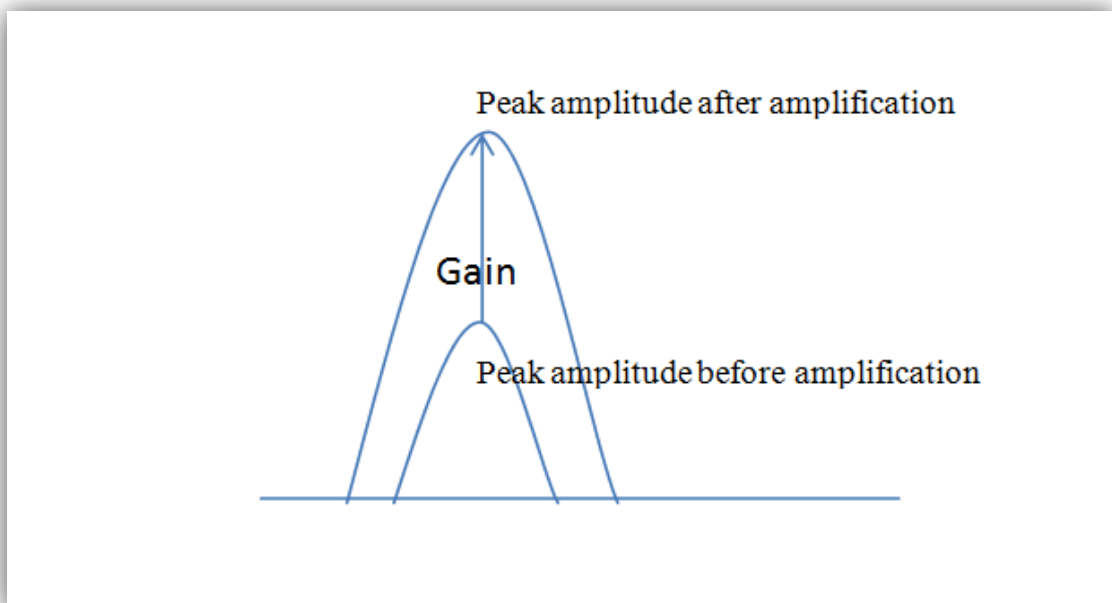


## RADIO FREQUENCY ATTITUDE

The conduct of an RF signal can be categorized by whether something adds power to the signal or takes power away from the signal known as gain and loss. These are now discussed in detail [13].

### GAIN

Gain is defined as the positive difference in amplitude between two signals. Gain is achieved by an amplification of the signal. Sometimes, gain is used synonymously with amplification. However, gain is technically the measure of amplification as shown in Figure 2.9. Gain can occur intentionally from an external power source that amplifies the signal, or unintentionally when an RF signal bounces off an object and combines with the original signal to amplify it as mentioned in [2]

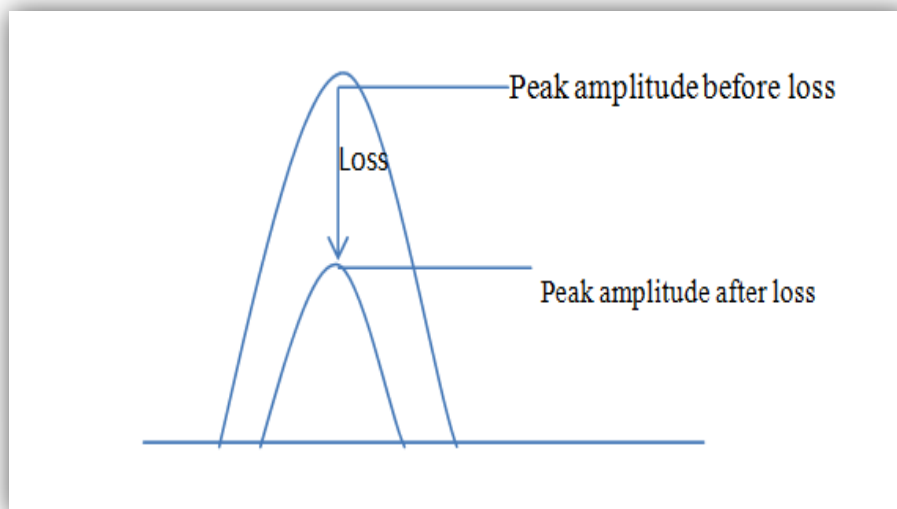


**Figure 2.9: Gain [10]**

## LOSS

Loss is also known as attenuation which is the negative difference in amplitude between signals, as seen in Figure 2.10. Like gain, loss can be either intentional or unintentional. Intentional loss may be necessary to decrease the strength of the signal to comply with standards or to prevent it from interfering with other RF signals [2]. More often, loss is unintentional. There are several factors that may result in RF loss. These include:

- **Absorption:** Certain type of materials can absorb the RF signal. This is known as absorption. The types of materials that will absorb an RF signal include concrete, wood, and asphalt.



**Figure 2.10: Loss as explained in [5]**

- **Reflection:** is the opposite of absorption. Instead of the signal being “soaked up” it is bounced back reflection which is caused by objects that are very large (in relation to the size of the wavelength of the signal) [13].

- **Scattering:** Where reflection caused by large and smooth objects, scattering is caused by small objects or rough surface. Objects that can cause scattering include foliage, rocks, and sand. Scattering can also occur when the RF signal comes in contact with elements in the air such as rain or heavy dust particles [2, 5, 10].
- **Refraction:** Over a long distance, an RF signal may move through different atmospheric conditions. For example, it may start out in relatively transparent condition, such as in bright sunshine, and then go through a much denser condition such as cold damp air. When an RF signal move from one medium to another of a different density the signal actually bends instead of traveling in a straight line.
- **Diffraction:** unlike refraction in which the medium through which the signal passes causes the RF signal to bend. Diffraction is bending caused by an object in the path of the transmission [2, 5, 10].

Refraction is the reason why swimming pool appears deeper than it actually is when you look into a pool, the light from the bottom is refracted away from the perpendicular because the index of refraction in air is less than that in water [2, 5, 10].

## 2.3 Wi-Fi ARCHITECTURES

Wireless connections refer to the way in which one or more unwired electronic devices are linked. These devices utilize special equipment to set up a connection to send and receive data through the air (media). Peripherals must be installed on both or among communication sides. This will be defined to the same behaviors between two nodes or among more nodes in order to be defined by all. The manufacturers,

developers and end-users must be legislated and measured. Therefore, manufactures and developers follow standards and regulations of the IEEE and FCC [5, 13, 16]. Wi-Fi wireless fidelity combines output between FCC and IEEE. The main purpose is to enable devices to make wireless connections. Many electronics devices are supported by this. Examples are, PC, cell phone, smart phone, digital mp3 and connections to internet via wireless providers' Wi-Fi with 802.11 is a standard enacted by IEEE, to be user-friendly. Wi-Fi uses radio waves for transmission, and the air uses it as an intermediary carrier. Three different wireless LAN configurations are defined by Wi-Fi as follows: Basic service set (BSS), External Service Set (ESS), and Independent Service Set (ISS)

### **2.3.1 BASIC SERVICE SET (BSS)**

Is a group of wireless devices served by a single access point (AP). It is also known as infrastructure mode. These devices transmit and receive transmissions to one access point. However, by definition, the AP need not be connected to the wired network. The BBS would have limited functionality if it were not connected. The mobile devices would only be able to communicate between themselves, to any other devices external to the BSS. The BSS must be given a unique identifier called the Service Set Identifier (SSID). The SSID serves as the Network name for the BSS [11, 17, 18]. The limitation or boundary of area for BSS is set by the Basic Service Area or BSA, as shown in Figure 2.11.

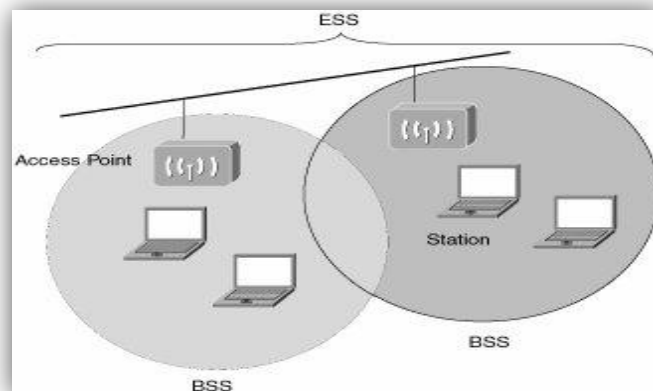


**Figure 2.11: BSS [16]**

### **2.3.2 EXTERNAL SERVICE SET (ESS)**

While BSS cannot support the required number of users for the BSA is too small. An ESS can be used as an ESS which comprises two or more BSS networks connected via a common distribution system, as figure 2.12. By utilizing multiple access points, additional users can be accommodated by ESS over wide area. ESS should be installed like cells in a cellular phone system. The APs can be positioned in such a way to enable the cells to overlap and to facilitate movement between cells (or roaming). When a mobile Wireless user carries a wireless laptop, TABs, or other devices, may enter into the range of more than one AP. The wireless device will select AP based on signal strength. Some devices also look at packet error rates [5, 12]. Once that device is accepted by the AP, the client device tunes to the radio channel at which the AP is set. The mobile device, at regular intervals, then carries onto look for appropriate radio frequency, to determine if a different AP can enhance service. If a better frequency is found, may be due to the user moving closer to the AP source,

then the device associates itself with the new AP (handoff), tuning to the radio frequency of the new AP to the user and seamless because the wireless device never has an interruption of service [2, 5, 18]. The only visible weaknesses of the Wi-Fi standard are: it does not stipulate how a handoff should occur. Since roaming between APs from different vendors can create problems, some industry experts suggest that all APs in an ESS come from the same vendor [10, 19].

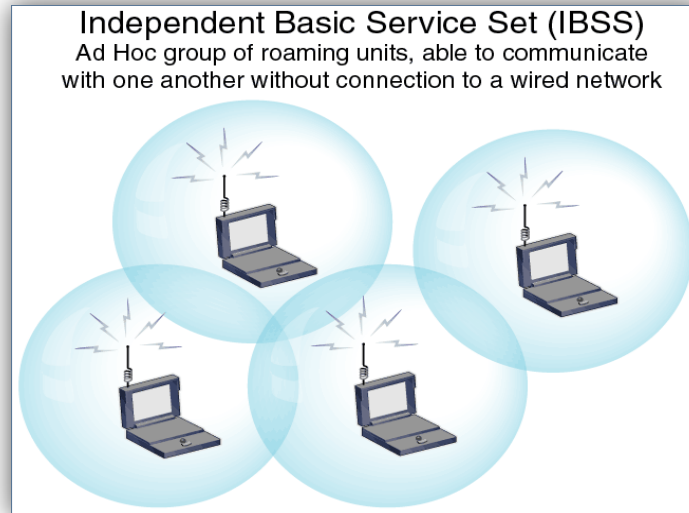


**Figure 2.12: ESS [16]**

### **2.3.3 INDEPENDENT BASIC SERVICE SET (IBSS)**

The IBSS or Independent Basic Service Set is a wireless network that does not utilize AP. It is also called peer-to-peer or ad hoc mode, whereby wireless devices communicate directly between them. While BSS has more flexibility in that it is able to connect to other wired or wireless networks, IBSS is useful for setting up to wireless networks anywhere quickly and easily, where users can share data between themselves, without the need to connect to the internet or an external network [7, 16, 20]. Wi-Fi has set of standard driver from 802.11 where each one distinguishes from others by some specification even they belong to same family. With Wi-Fi standards,

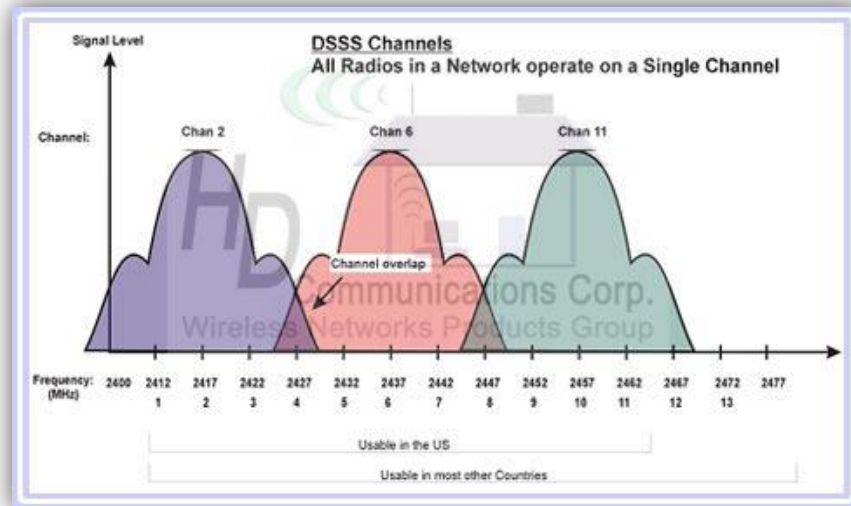
the differentiation from one to another is founded on the other hand the signal penetration less while signal face barriers [5, 15, 17, 21] as Figure 2.13: IBSS.



**Figure 2.13: IBSS [16]**

#### **2.3.4 DIRECT-SEQUENCE SPREAD SPECTRUM (DSSS)**

In Direct-sequence spread spectrum implying method as well as light spread spectrum mechanism, DSSS is a technology which has bit rate much higher data rates than the FHSS. The principle here is to spread an output signal using a predetermined bit sequence. In DSSS systems is to spread a pseudo-noise code (PN) is used the signal is converted directly from the PN code from a narrow band to a broadband signal, the spread signal is series sent to a broad band and not like FHSS staggered and spread over various channels. By the spreading of the signal intensity is greatly reduced. It achieves a reduction in the noise floor, thus minimizing interference with other systems, at the same time provides some protection against the spread of unauthorized interception. The signal is only detected by the receiver knows by the PN code [12, 15], as Figure 2.14.



**Figure 2.14: DSSS [15]**

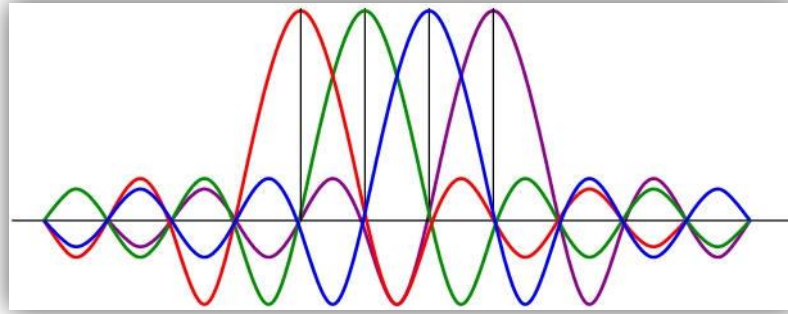
### 2.3.5 ORTHOGONAL FREQUENCY DIVISION MULTIPLE (OFDM)

Is based on parallel data transmission in frequency-division multiplexing, the Frequency Division Multiplexing (FDM), this is a symbol of several sub-channels in parallel sent at same time, these multiples transmission hold on multiples frequencies, each frequency band sub-carrier is modulated isolation with various data stream and a spacing guard band putted among the sub-carriers in order to overcome the signal overlap as attached in [12]. In addition, OFDM utilizes multiple sub-carriers. However, the sub-carriers are closely spaced, without causing interference, removing guard bands between adjacent sub-carriers, since the sub-carriers are orthogonal; i.e., the peak of one sub-carrier coincides with the null of an adjacent sub-carrier. A very high rate data stream is divided into multiple parallel low rate data streams in an OFDM system. Each smaller data stream is subsequently mapped to individual data sub-carriers and modulated using Phase Shift Keying (PSK) or Quadrature Amplitude Modulation (QAM) as indicated in [7, 22]. Compared to FDM, OFDM needs less bandwidth for similar amount of information, which translates to higher spectral



efficiency. Besides, the high spectral efficiency OFDM system (for e.g., WiMAX) is a flexible environment in NLOS. It can overcome interference and frequency-selective fading efficiently because of multipath propagation, as the tone is a subset of sub-carriers rather than the entire and a broader one. The effect is the suppression of Inter-symbol interference (ISI) as a symbol of the OFDM sub-carriers of the parallel bus system and the use of a single prefix Cyclic (CP).

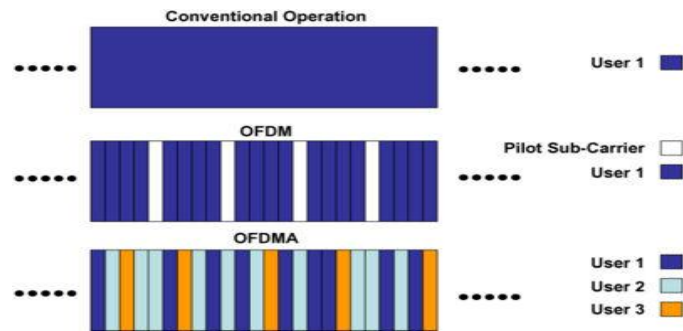
The OFDM theory has been for decades, but recently it has been implemented in real world application. Wired and Wireless chose to fixed and mobile telecommunications networks or OFDM technology to achieve the highest data rate (what is called broadband). Examples of these technologies are ADSL, Wi-Fi (802.11 a/g/n) and WiMAX. The big advantages of this technology are its insensitivity to disturbances and the high data rates. Disorders or unstable usually occur only in low frequency ranges. So the risk is high that when using only one channel, the data will be destroyed, spread across multiple sub-channels the data is lost and can only partially through the FEC method (Forward Error Correction) are corrected. The big disadvantage of the FDM process is very inefficient use of bandwidth. It has to run the sub-channel at certain distances from one another in order to avoid interference with each other. As a result of the large bandwidth overhead, OFDM systems achieve a reduction of the bandwidth by about 50%, it's allow the overlap of the transmission frequencies of the sub channels [2], as shows in Figure 2.15.



**Figure 2.15: OFDM modulation [23]**

### **2.3.6 ORTHOGONAL FREQUENCY DIVISION MULTIPLE ACCESSES (OFDMA)**

Like OFDM but the sub-carriers are separate into sets of sub-carriers, each of sets known by sub-channels. The sub-carriers that belong to sub-channels does not need an adjustment , regarding the downlink the sub-channels may be get various receivers , and uplink a transmitter may adjacent one or more sub-channels , as represent in Figure 2.16 which clarify it. The sub direction define sub channels which forwarded to the users stations “subscribers” regarding these channels states and data requirement, implemented sub channelization at the same time slot, a mobile WiMAX base station (BS) obtains more transmission power to the user device with lower levels of signal to noise ratio (SNR), as well as less power to devices of subscribers with higher SNR. Sub-channelization also affects the base stations to get more power to sub- channels that is allocated to indoor user stations which results in better coverage as illustrated in [24, 25].



**Figure 2.16: Difference of OFDM and OFDMA [24]**

In an OFDM only one users stations transmittable within same time per slot, unlike an OFMDA many users stations transmittable at same time slot not only that but OFMDA capable passing over many sub channels as explained in [16]. Sub-channelization at the uplink able to keep a user device transmits due to it able to focus power only on certain sub channels get to it. The reducing consuming power mechanism very useful particularly for devices which depending on battery as source power so working time within this feature will holder longer, like mobile WiMAX as referred in [26, 27].

## 2.4 OSI AND TCP/IP MODEL

The Internet can be accessed by using the World Wide Web, e-mail, and file-sharing programs, which provide the human interface to the network, and allowing the surfer to send and receive information easily. Most of the applications are intuitive and can be accessed without knowing how they work (Figure 2.17). It is important to know how an application can format, transmit, and interpret messages, sent and received across the network [28].

The Open Systems Interconnection (OSI) model is a product of the Open Systems and is a prescription of characterizing and standardizing the functions of

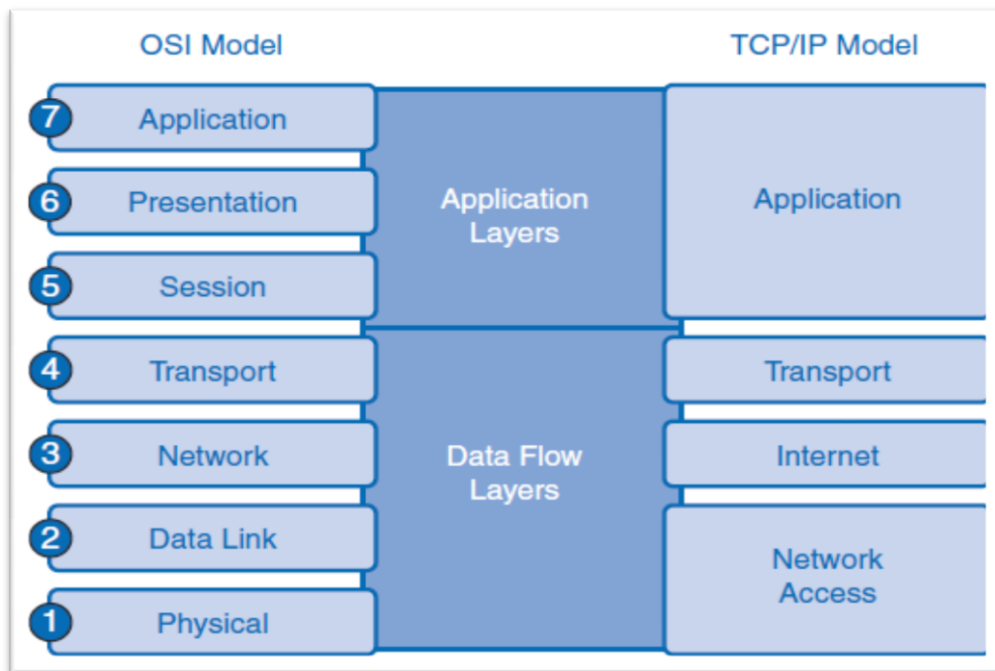
a communications system in terms of abstraction layers. In this model, similar communication functions are categorized into logical layers, whereby a layer serves the layer above and below it [28, 29]. The communication across the network is made easier if you use the layered framework of the (OSI) model. The OSI reference model is a layered, abstract representation created as a guideline for network protocol design and instruction. Under the OSI model, the networking process is divided into seven logical layers. Each layer has a unique functionality which is assigned specific services and protocols as figure 2.17. In the TCP/IP model, protocols are deliberately not as rigidly designed into strict layers compared to the OSI model. The TCP/IP recognizes four broad layers of functionality, derived from the operating scope of their contained protocols, i.e.:

- i. The scope of the software application,
- ii. The end-to-end transport connection,
- iii. The internetworking range, and
- iv. The scope of the direct links to other nodes on the local network.

Although the concept differs from the OSI model, these layers are nonetheless compared to the OSI layering scheme as follows:

- i. The Internet application layer includes the OSI application layer, presentation layer, and most of the session layer.
- ii. Its end-to-end transport layer includes the graceful close function of the OSI session layer and the OSI transport layer.
- iii. The internetworking layer (Internet layer) is a subset of the OSI network layer. The link layer includes the OSI data link and physical layers, and parts of OSI's network layer.

These comparisons are based on the original seven-layer protocol model according to ISO 7498.

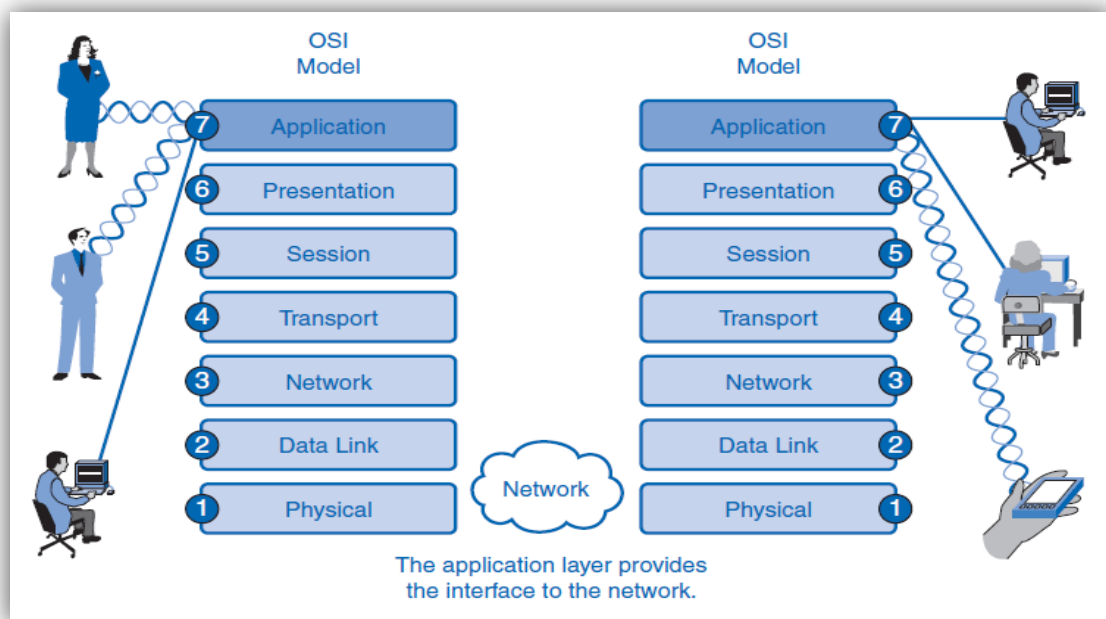


**Figure 2.17: Acomparison of the OSI and TCP/IP model [28]**

### 2.4.1 APPLICATION LAYER

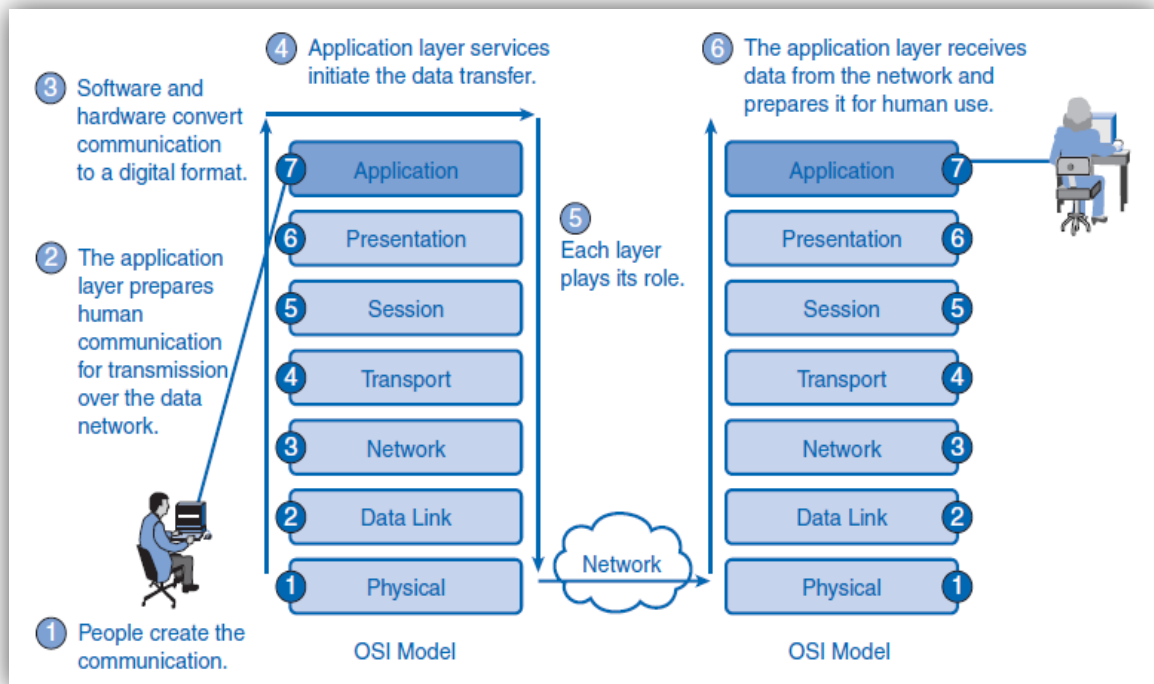
The application layer is the OSI layer closest to the end user. This means that both the OSI application layer and the user directly interact with the software application (Application software). Applications refer to the software programs people use to communicate over the network. Some examples of application software are HTTP, FTP, e-mail, etc.). This layer interacts with software applications, implementing a communicating component. These application programs fall outside the scope of the OSI model [30]. Application-layer functions normally, including the identification of communication partners, determination of the availability of resources, and

synchronization of communication. In the identification of communication partners, the application layer determines the identity and availability of communication partners, for an application to transmit data. In determining resource availability, the application layer must decide if sufficient network or the requested communications exists. All communication between applications requires cooperation managed by the application layer [28, 30, 31], as Figure 2.18.



**Figure 2.18: Application Layer [28]**

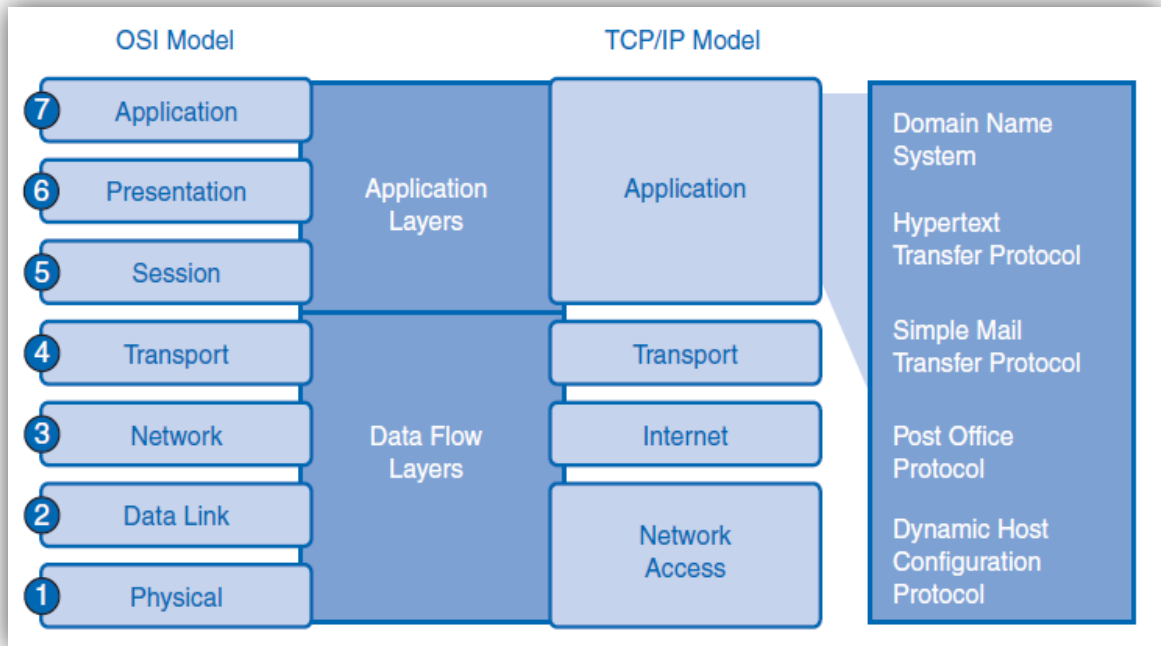
Information is passed from one layer to the next in the OSI model. This starts at the application layer on the transmitting host and proceeds down the hierarchy to the physical layer, and then passes over the communications channel to the destination host. Here, the information proceeds back up the hierarchy, ending at the application layer. Figure 2.19 depicts the steps in this process [28].



**Figure 2.19: OSI Encapsulation Process [28]**

The OSI layers encapsulate data down the stack. Encapsulated data travels across the media to the destination. OSI layers at the destination unencapsulate the data up the stack [31].

The application layer, Layer 7, is the top layer of both the OSI and TCP/IP models. Application layer protocols are used to exchange data between programs running on the source and destination hosts. Several application layer protocols are available (figure 2.20), and new protocols are constantly being developed. The TCP/IP protocol suite was developed before the definition of the OSI model; however, the functionality of the application layer protocols of TCP/IP fits roughly into the framework of the top three layers of the OSI model, i.e., application, presentation, and session [28]. Most applications, such as web browsers or e-mail clients, incorporate functionality of the OSI Layers 5, 6, and 7. A Figure 2.20: TCP/IP application layer protocols.



**Figure 2.20: TCP/IP application layer protocols [28]**

Most TCP/IP application layer protocols were developed before the emergence of personal computers, GUIs, and multimedia objects [28]. As a result, these protocols implement little of the functionality that is specified in the OSI model presentation and session layers.

#### 2.4.1.1 APPLICATION LAYER PROTOCOLS

During a communication session, both source and destination devices use application layer protocols. The source and destination host must match for the communications to be successful and the application layer protocols implemented [28]. The following tasks are performed by Protocols:

- i. Establish consistent rules for the exchange of data between applications and services loaded on the participating devices.
- ii. Specify how data inside the messages is structured and the types of messages that are sent between source and destination. These messages can



be requests for services, acknowledgments, data messages, status messages, or error messages [28, 30].

- iii. Define message dialogues, which ensure that a message that is sent elicits the expected response and the correct services are invoked when data transfer occurs.

Many different types of applications communicate across data networks. Hence, application layer services must implement multiple protocols for the desired range of communication experiences to be provided. Each protocol is specific in its purpose and comprises characteristics needed to meet that purpose. The right protocol details in each layer must be adhered to so that the functions at one layer interface correctly with the services in the lower layer.

#### **2.4.1.1.1 SOME OSI APPLICATION LAYER PROTOCOLS**

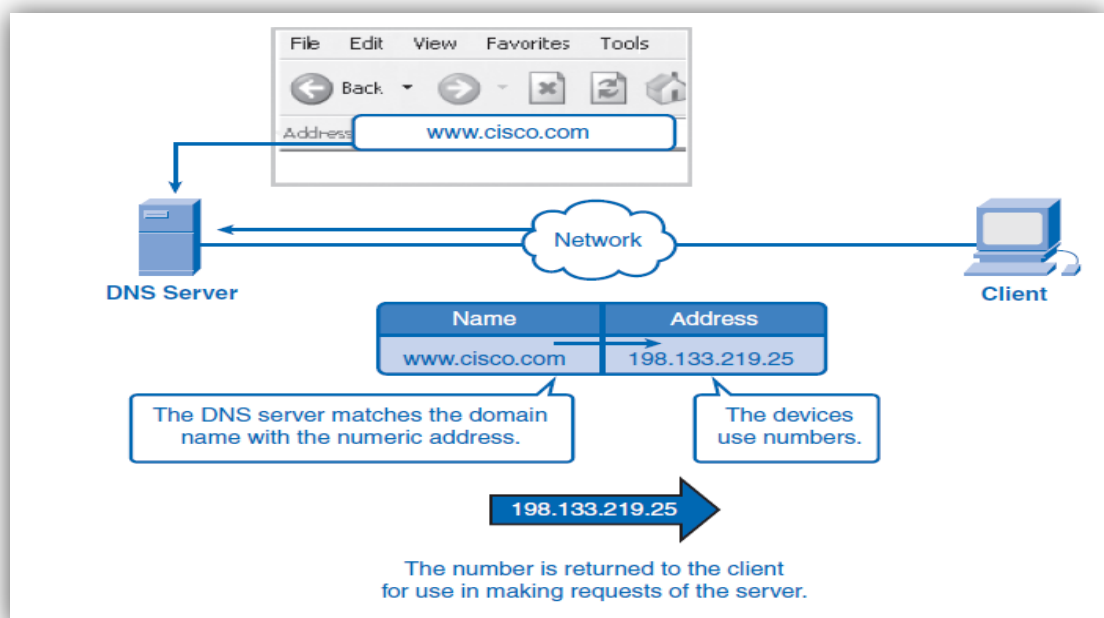
- **File Transfer and Access Management Protocol (FTAM)** is the OSI Application layer protocol for File Transfer Access and Management. FTAM aims to combine both file transfer, similar in concept to the Internet FTP, into a single protocol, as well as provide remote access to open files, like the NFS. However, like the other OSI protocols, FTAM has not been widely adopted and the TCP/IP based Internet has become the dominant global network [32, 33].
- **X.400 Mail:** is a suite of ITU-T Recommendations defining standards for Data Communication Networks for Message Handling systems (MHS). This is more commonly known as "email". Previously, X.400 was expected to be the main form of email; however, the SMTP-based Internet e-mail has taken over this role. In spite of this, it has been widely used in organizations, and was an

integral part of Microsoft Exchange Server up to 2006. Some variants continue to be used in the military and aviation sectors [34].

- **Common Management Information Protocol (CMIP)** is the OSI specified network management protocol. This is defined in the ITU-T Recommendation X.711, ISO/IEC International Standard 9596-1. It allows the implementation for the services defined by the Common Management Information Service (CMIS) specified in ITU-T Recommendation X.710, ISO/IEC International Standard 9595. This allows communication between network management applications and management agents. CMIS/CMIP is the network management protocol specified by the ISO/OSI Network management model, further defined by the ITU-T in the X.700 series of recommendations. CMIP models management information in terms of managed objects and allows both modification and performing actions on managed objects. Managed objects are described using GDMO (Guidelines for the Definition of Managed Objects), and can be identified by a distinguished name (DN), from the X.500 directory. CMIP also provides good security (support authorization, access control, and security logs) and flexible reporting of unusual network conditions [35].

#### 2.4.1.1.2 TCP/IP APPLICATION LAYER PROTOCOLS

- **Domain Name System (DNS)** (TCP/UDP port 53) is used to resolve Internet names to IP addresses [28]. It is a hierarchical distributed naming system for computers, services, or any resource connected to the Internet or a private network. Various information with domain names assigned to each of the participating entities can be associated with it. A service known as Domain Name Service, translates queries for domain names (meaningful to humans) into IP addresses. This allows the locating of computer services and devices worldwide. An often-used analogy to explain the Domain Name System is that it serves as the phone book for the Internet by translating human-friendly computer hostnames into IP addresses. For example, the domain name `www.cisco.com` translates to the addresses `198.133.219.25` (IPv4) as Figure 2.21 Resolving DNS Addresses [31].



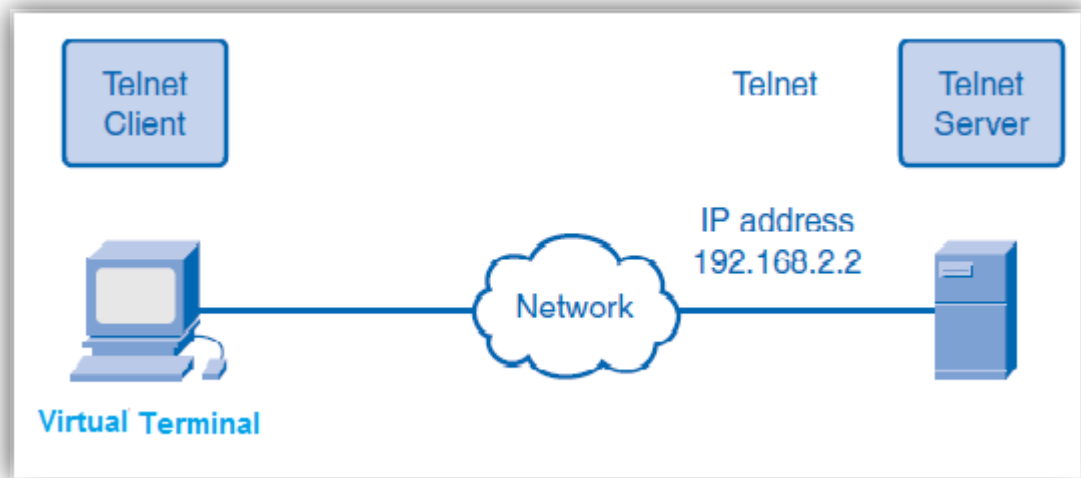
**Figure 2.21: Resolving DNS Addresses [28]**

- **Hypertext Transfer Protocol (HTTP)** (TCP port 80) is used to transfer files that comprise the web pages of the www. The HTTP is one of the protocols in the TCP/IP suite. It was initially developed to publish and retrieve HTML pages. However, now it is used for distributed, collaborative information systems. HTTP, which is one of the most used application protocols, is used across the www for data transfer [28]. It comprises two programs, i.e., a client program and a server program, both of which are executed on different end systems, and talk to each other by exchanging HTTP messages. The structure of these messages is defined by HTTP, as well as how the client and server exchange the messages [31]. HTTP specifies a request/response protocol. When a client, typically a web browser, sends a request message to a server, the HTTP protocol defines the message types the client uses to request the web page and the message types the server uses to respond. Users' browser the internet using the WWW. The actual protocol that allows one to download images or other object from another web site is called (HTTP) , it is uses the reliable service of TCP [17].
- **Simple Mail Transfer Protocol (SMTP)** is used for the transfer of mail messages and attachments [28]. SMTP is an Internet standard for electronic mail (e-mail) transmission across Internet Protocol (IP) networks. SMTP was first defined by RFC 821 (1982, eventually declared STD 10). It was last updated by RFC 5321 (2008), including the extended SMTP (ESMTP) additions, and is the protocol used widely today. SMTP is specified for outgoing mail transport, using TCP port 25. The protocol for new submissions is for all intents the same as SMTP, but it uses port 587 instead. SMTP connections secured by SSL are called shorthand SMTPS. Nonetheless, in its

own right, SMTPS is not a protocol [28, 31]. Electronic mail servers and other mail transfer agents use SMTP to send and receive mail messages. However, user-level client mail applications use only SMTP to relay messages to a mail server. To receive messages, client applications usually use Post Office Protocol (POP or POP3 port 110), Internet Message Access Protocol (IMAP) or a proprietary system (such as Microsoft Exchange or Lotus Notes/Domino) to access their mail box accounts on a mail server. To send and receive e-mail, one has to open an e-mail account with a network service provider [17].

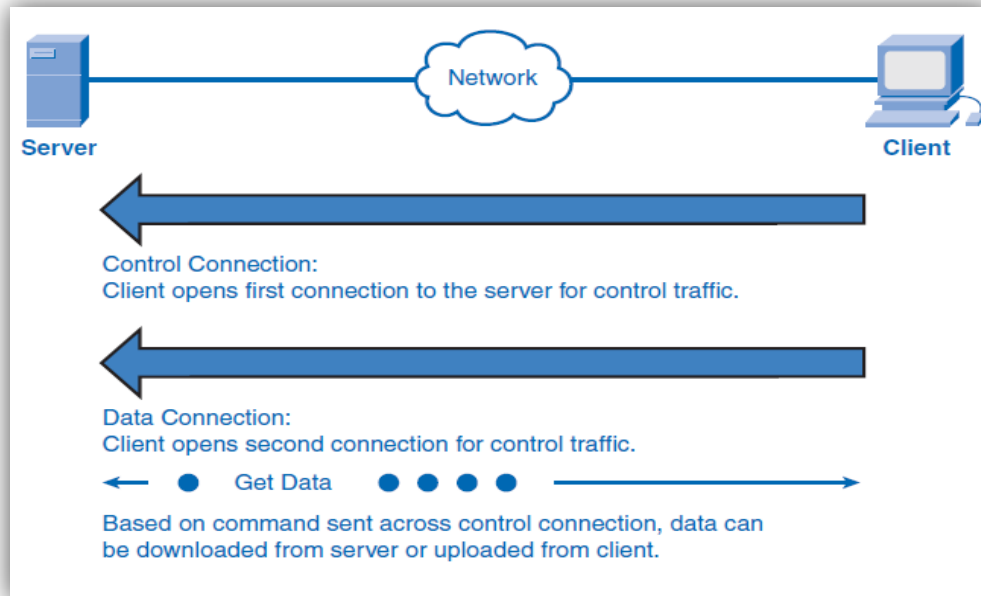
- **Telnet** (TCP port 23) is a terminal emulation protocol, and is used to provide remote access to servers and networking devices [28]. Telnet is a client/server protocol that provides a standard method of emulating text-based terminal devices over the data network [17, 28, 31]. Both the protocol and client software that implements the protocol are usually referred to as Telnet. It is used on Internet or LAN to provide a bidirectional interactive text-oriented communications facility by utilizing a virtual terminal connection. User data is interspersed in-band with Telnet control information in an 8-bit byte oriented data connection over the Transmission Control Protocol (TCP). It was developed in 1969 with RFC 15, extended in RFC 854, and standardized as Internet Engineering Task Force (IETF) Internet Standard STD 8 (one of the first Internet standards). Telnet provides access to a command-line interface (usually, of an operating system) on a remote host. Most network equipment and operating systems with a TCP/IP stack support a Telnet service for remote configuration (including systems based on Windows NT). Telnet

suffers from security issues; as such, its use for this purpose has declined in favor of SSH. Figure 2.22 shows the telnet service.



**Figure 2.22: Telnet Service [28]**

- **File Transfer Protocol (FTP):** FTP was developed to allow file transfers between a client and a server. An FTP client is an application that runs on a computer used to push and pull files from a server running the FTP daemon (FTPD)., FTP requires two connections between the client and the server, i.e., one for commands and replies, and the other for the actual file transfer, in order to successfully transfer files. The client establishes the first connection to the server on TCP port 21, which is used for controlling traffic. It consists of client commands and server replies. The client establishes the second connection to the server over TCP port 20, which is used for actual file transfer. It is created each time a file is transferred [28, 31, 36]. File transfer can occur in either direction, as depicted in Figure 2.23. The client can either download a file from the server or upload a file to the server.



**Figure 2.23: FTP Process [28]**

## 2.4.2 TRANSPORT LAYER

The transport layer provides transparent transfer of data between end users. It provides reliable data transfer services to the upper layers. The reliability of a given link is controlled by the transport layer through flow control, segmentation/desegmentation, and error control. Some protocols are state- and connection-oriented, which infers that the transport layer can keep track of the segments and retransmit those files. Besides this, the transport layer acknowledges successful data transmission and sends the next data if no errors occur [30]. The Transport layer provides for the segmentation of data and the control necessary to reassemble these pieces into the various communication streams [28]. Its primary responsibilities to accomplish this are:

- Tracking the individual communication between applications on the source and destination hosts.
- Segmenting data and managing each piece.

- Reassembling the segments into streams of application data.
- Identifying the different applications.

End-to-end transportation of packets across a network is managed by the transport layer. It connects application processes running on end hosts seamlessly, appearing as if the two end applications were connected by a reliable dedicated link. This makes the network to be “invisible” as illustrated in Figure 2.24.

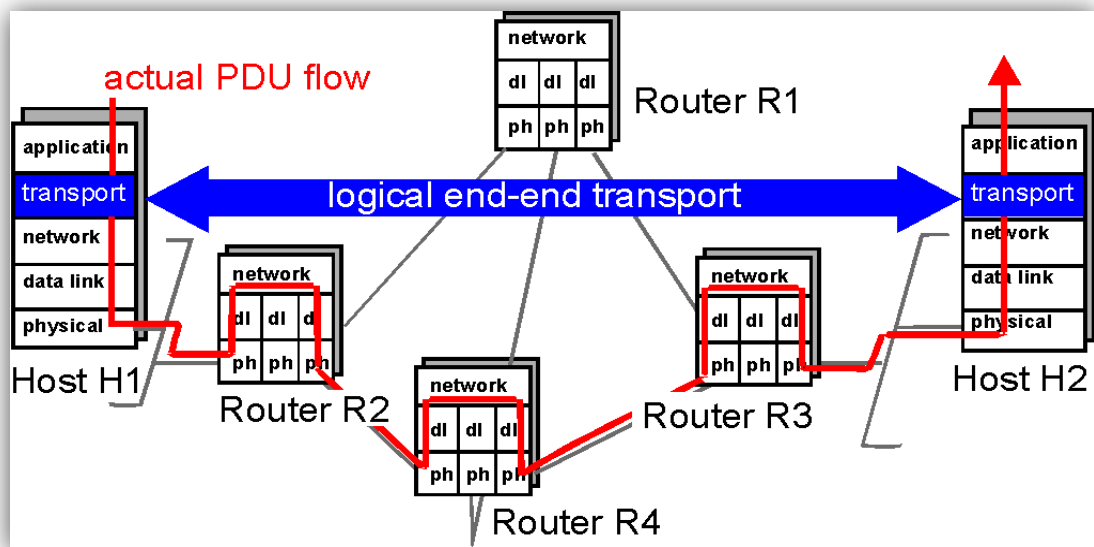


Figure 2.24: Transport layer [31]

### 2.4.2.1 TRANSPORT LAYER PROTOCOLS

TCP/IP uses two transport layer protocols: TCP and UDP.

- **Transmission Control Protocol (TCP)** TCP provides a connection oriented service. This means it sends data as a stream of related packets, making concepts such as the order of packets meaningful. Specifically, TCP provides reliable service to upper layer applications, so as to ensure that i) the packets are correctly received, and ii) the packets are received in the order in which they are sent. At the beginning of a connection, TCP uses a three-way



handshake to establish a connection between sender and receiver. Here, they agree on the protocol parameters to be used [17, 31, 37]. This process takes 1.5 round trip times (one side sends a Synchronize packet, the other replies with a SYN and an Acknowledge packet, and the first confirms with an ACK), which is an overhead that is avoided by UDP [37].

TCP receives data from the application as a single stream, e.g., a large file, and segments it into a sequence of packets. It tries to use large packets to minimize overhead, but there is a maximum size which the network can carry efficiently, called the MTU (maximum transfer unit). TCP is responsible for choosing the correct size, in a process called path MTU discovery. In contrast, UDP is given data already segmented into packets, and so it is the application's responsibility to observe MTU restrictions.

#### □ TCP segment format

Each TCP segment has two parts, a standard 20-byte header followed by a variable payload containing the application data [17, 38]. The TCP segment format is shown in Figure 2.25.

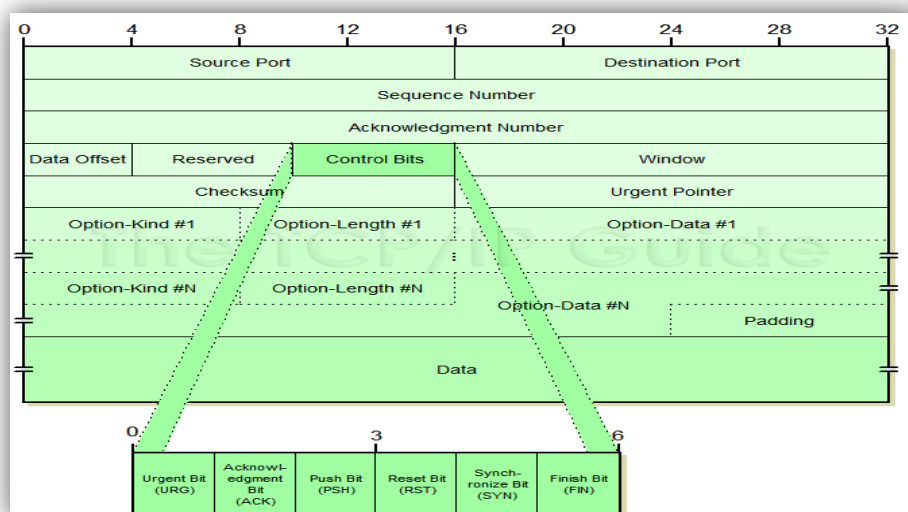


Figure 2.25: TCP Segment format [17]

Where:

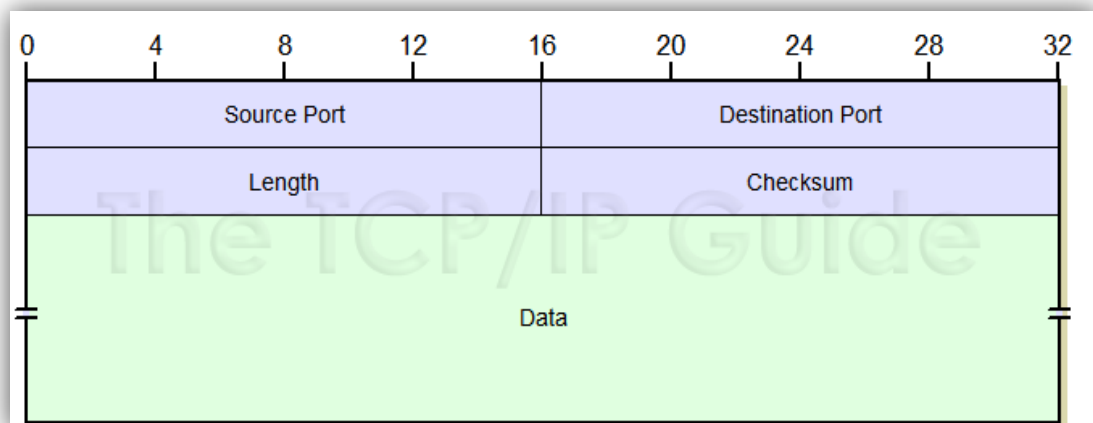
- i. **Source Port:** The 16-bit source port number, used by the receiver to reply.
- ii. **Destination Port:** The 16-bit destination port number.
- iii. **Sequence Number:** The 32-bits sequence number of the first data byte in this segment. If the SYN control bit is set, the sequence number is the initial sequence number (n) and the first data byte is n+1 [38].
- iv. **Acknowledgment Number** (32 bits): If the ACK control bit is set, this field contains the value of the next sequence number that the receiver is expecting to receive.
- v. **Data Offset:** The number of 32-bit words in the TCP header. It indicates where the data begins.
- vi. **Reserved:** The 6 bits reserved for future use; must be zero.
- vii. **URG:** Indicates that the urgent pointer field is significant in this segment.
- viii. **ACK:** Indicates that the acknowledgment field is significant in this segment.
- ix. **PSH:** Push function.
- x. **RST:** Resets the connection.
- xi. **SYN:** Synchronizes the sequence numbers.
- xii. **FIN:** No more data from sender.
- xiii. **Window:** Is the 16 bits used in ACK segments. The Window specifies the number of data bytes, starting with the one indicated in the acknowledgment number field that the receiver (= the sender of this segment) is willing to accept.

- xiv. **Checksum:** The 16-bit one's complement of the one's complement sum of all 16-bit words in a pseudo-header, the TCP header, and the TCP data. While computing the checksum, the checksum field itself is considered zero [38].
- xv. **Options (variable):** options are specified using multiples of bytes. There are two extra bytes preceding each option as follows i) the first byte indicates the option type, and ii) the second byte indicates the option in bytes. Examples of options :
- **Maximum Segment Size (MSS) (16 bits):** This option is used by the originating. TCP during connection establishment to negotiate the MSS to be used for the connection. The 16 bits used for this field limit the MSS to 64 KB [17].
  - **Timestamp (8 bytes):** The timestamp option is utilized for more accurate round-trip time (RTT) calculations. Two four-byte timestamp fields are used for this option. The sending TCP fills the first field with the current time. The receiver echoes back the timestamp value received in the second field in an ACK segment. This facilitates the sender for more accurate calculation of the RTT [17].
- **User Datagram Protocol (UDP):** is a very simple protocol, called 'connectionless' because all UDP packets are treated independently by the transport layer, and not as being part of an on-going flow. Other than minor error checking, UDP only does multiplexing and demultiplexing. UDP does not guarantee that packets will be received in the order that they are sent, or

that they will be received at all. Other than this, it also does not control its transmission rate. The rationale underlying the design of UDP is to allow applications to have more control over the data sending process and reduce the delay associated with setting up a connection, which are desirable features for certain delay-sensitive applications, for example, streaming video and Internet telephony. Generally, applications that can tolerate certain data loss/corruption (but which are sensitive to delay) prefer to use UDP [37].

#### □ UDP datagram format`

UDP datagram has a header and a payload. The payload carries the application message and the header carries the information necessary for the correct operation of the UDP protocol. The UDP header is very simple, only eight bytes long. Figure 2.26 shows UDP Datagram format [17, 38].



**Figure 2.26: UDP Datagram format [17]**

Where:

- i. **Source Port (16 bits):** Indicates the port of the sending process. It is the port to which replies should be addressed.

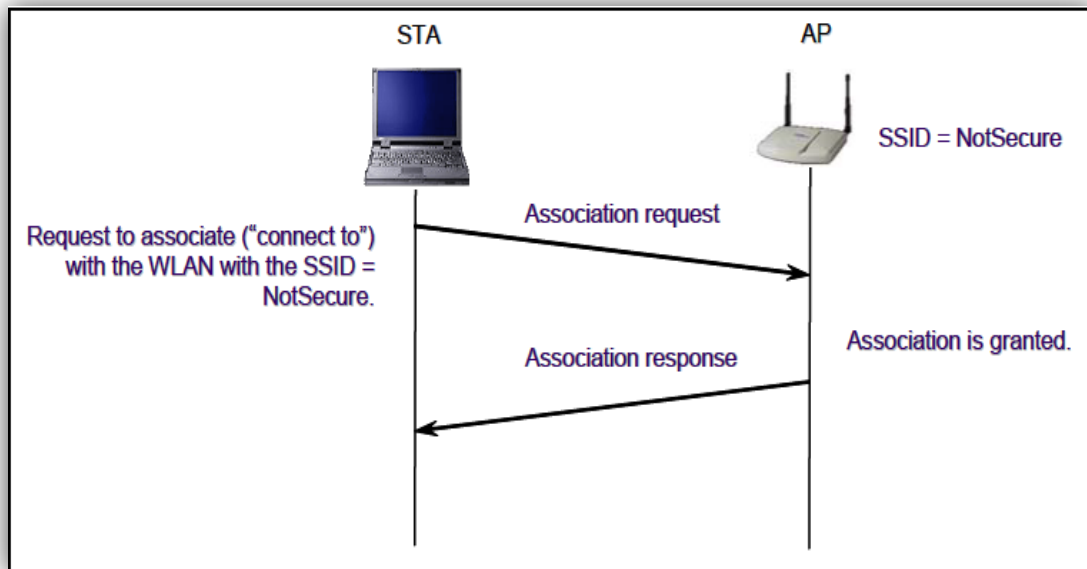
- ii. **Destination Port (16 bits):** Specifies the port of the destination process on the destination host.
- iii. **Length (16 bits):** The length (in bytes) of this user datagram, including the header.
- iv. **Checksum (16 bits):** An optional 16-bit one's complement of the one's complement sum of a pseudo-IP header, the UDP header, and the UDP data. The pseudo-IP header comprises the source and destination IP addresses, the protocol, and the UDP length [38].

## 2.5 THE IEEE 802.11 OPERATIONS

The IEEE 802.11 media access control (MAC) protocol supplies the functionality in WLANs. This is needed to provide reliable delivery of user data over the wireless media that is potentially noisy and unreliable. The IEEE 802.11 MAC protocol implements a frame exchange protocol. Under this exchange, the STA, which receives a frame, either returns an acknowledgement to the source of the frame that it was received correctly, or notifies the source if there is an error. Each STA executes the frame exchange protocol in the WLAN; and receives, decodes, and responds to information in the MAC header. This is done for every frame received, except for certain broadcast, multicast, and beacon frames [39, 40].

Figure 2.27 depicts a typical two-frame flow for IEEE 802.11 WLAN communication that illustrates an association request and response. First, the STA sends an Association Request frame to the AP, which is a request to connect to the WLAN with a service set identifier (SSID) of “NotSecure”. The SSID is a text name assigned to the WLAN. The AP, with matching SSID, responds to the STA successfully or otherwise. If the response is successful, the result is an association or a record-

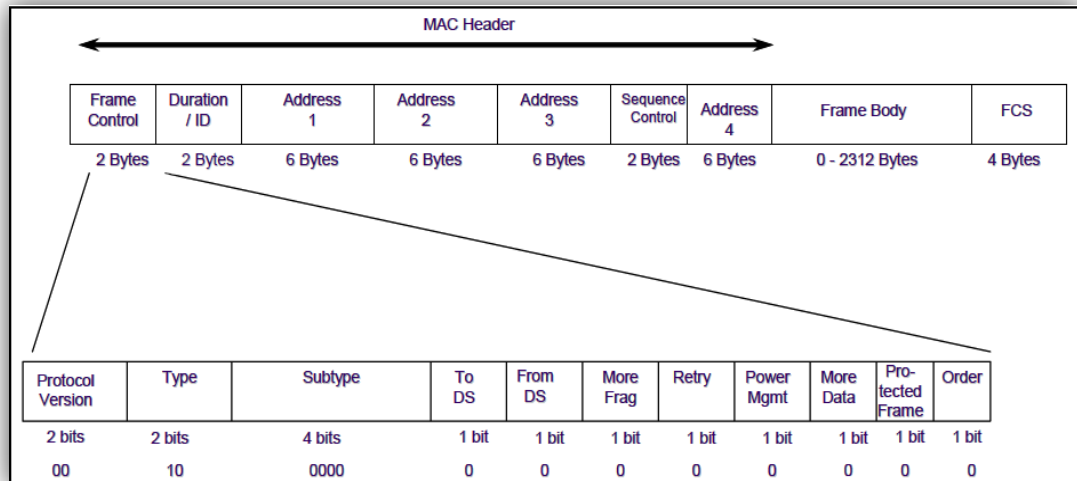
keeping procedure (not yet an RSNA) between the AP and STA. Association allows the DS to track STA location, so that DS frames are sent to the correct STAs [39, 41].



**Figure 2.27: Two-Frame IEEE 802.11 Communication [39]**

### **2.5.1 THE IEEE 802.11 FRAME**

The 802.11 frame have many similarities with Ethernet frame. It's composed of many of fields that are identified to its use for wireless link. The number over each of the fields in the frame represents the lengths of the fields in bytes [31]. But the number over each of the subfields in the frame control field of the frame represents the lengths in bits, as Figure 2.28.



**Figure 2.28: The 802.11 frame [31]**

The beginning of the data frame represents the MAC header. It is composed of several fields for transport of data in WLAN. The frame header provides the MAC addresses of the source and destination and transmitter address. This determines the address of the wireless interface card that it's responsible for transporting the frame onto the wireless medium, and receiver address, determining the group address or wireless station that should handle the frame [41].

#### **2.5.1.1 IEEE 802.11 FRAME TYPES**

The IEEE 802.11 frame exchange protocol involves three types of frames, as follows:

- i. **Data Frame:** Data frames compress packets from upper layer protocols, for example, IP. In turn, IP might contain application data (e.g., e-mail, Web pages). Data frames allow the sending of the upper layer protocol packets to a STA or AP. Robust Security Network Association (RSNA) security mechanisms protect these frames [41]. Data frames carry higher-level protocol data in the frame body. Some of the fields in the figure may not be used, depending on the particular type of data frame, [40]. The different types of data frame are categorized according to function, for example, data frames

used for contention-based service and that used for contention-free service. A frame that appears in the contention-free period only can never be utilized in an IBSS. Another example of a division is between frames that carry data and frames that perform management functions [40], show Figure 2.29.



**Figure 2.29: Data Frame [39]**

The data frame consists of several of fields, are:

- Frame Control:** This field defines a number of parameters for IEEE 802.11 operation. For example, it contains two bits used to identify the version of the IEEE 802.11 MAC. Another value within the field is the Protected Frame bit; if it is set to 1, the frame body is cryptographically protected using the negotiated cipher suite (e.g., CCMP, TKIP, WEP) [42]. Frame Control bits may affect the interpretation of other fields in the MAC header, though. Most notable are the address fields, which depend on the value of the To DS and From DS bits [39].
- Duration / ID:** This field is used by a STA to retrieve frames buffered at an AP. The field identifies the remaining duration in the frame exchange between a STA and AP [40]. Four rules specify the setting for the Duration field in data frames:
  - Any frames transmitted during the contention-free period set the Duration field to 32,768. Naturally, this applies to any data frames transmitted during this period.



- Frames transmitted to a broadcast or multicast destination (Address 1 has the group bit set) has duration of 0. Such frames are not part of an atomic exchange and are not acknowledged by receivers, so contention-based access to the medium can begin after the conclusion of a broadcast or multicast data frame. The NAV is used to protect access to the transmission medium for a frame exchange sequence. With no link-layer acknowledgment following the transmission of a broadcast or multicast frame, there is no need to lock access to the medium for subsequent frames [40, 41].
- If the more Fragments bit in the Frame Control field is 0, no more fragments remain in the frame. The final fragment need only reserve the medium for its own ACK, at which point contention-based access resumes. The Duration field is set to the amount of time required for one short inters frame space and the fragment acknowledgment.
- If the more Fragments bit in the Frame Control field is set to 1, more fragments remain. The Duration field is set to the amount of time required for transmission of two acknowledgments, plus three short interface spaces, plus the time required for the next fragment [40].
- **Address Fields:** The MAC header for a data frame contains four distinct address fields, although in some cases not all fields contain relevant addresses. The address fields identify the original Source Address (SA) and final Destination Address (DA) in a frame exchange, as well as the Receiver Address (RA). Depending on the function of the frame, the address fields also identify either the Transmitter Address (TA) or the BSS identifier BSSID, which is typically the address of the AP [41]. The sequence of the addresses in

the MAC header depends on two things: whether the transmitting station is in an IBSS or an infrastructure BSS, and whether the communicating stations are part of the DS. Table 2.2 identifies the functions of each of the address fields for the four possible cases, as defined by the values for the To DS and from DS subfields [40].

**Table 2.2: The functions of each of the address fields for four possible cases [40]**

Function	ToDS Subfields	From DS Subfields	Address1	Address 2	Address 3	Address 4
<b>IBSS</b>	0	0	RA=DA	SA	BSSID	N/A
<b>Infrastructure BSS: From the AP</b>	0	1	RA=DA	BSSID	SA	N/A
<b>Infrastructure BSS: To the AP</b>	1	0	RA = BSSID	SA	DA	N/A
<b>Infrastructure BSS: Wireless DS (AP to AP)</b>	1	1	RA	TA	DA	SA

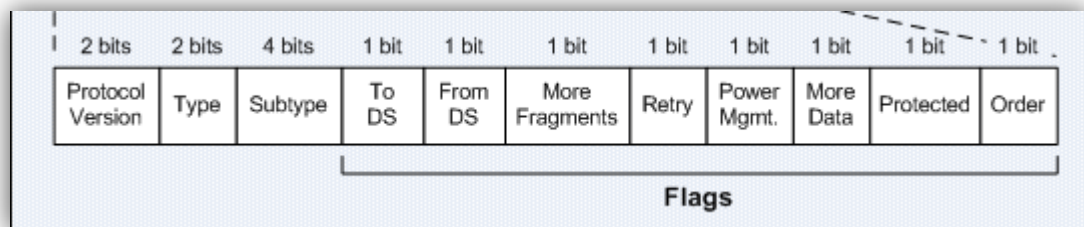
- **Sequence Control:** This field is used to allow a STA to identify received frames that are duplicates, and to assist it in reassembling fragmented frames [41].
- **FCS:** This field is used for error detection to detect random bit errors in the received frame. It contains the result of applying a 32-bit cyclic redundancy check (CRC-32) on the data. Because of this, the FCS is often called the CRC. The FCS calculation is performed on all data in the MAC header and frame body [40].

- **Frame Body:** This field, also called the Data field, holds a payload from a higher layer. The Frame Body field is variable in length, with a maximum size of 2312 octets [40, 41].
- ii. **Management Frame:** Management frames carry the information necessary for managing the MAC. They provide the ability to perform management functions such as authenticating or associating (the wireless equivalent to connecting or registering). These frames can easily be forged, since IEEE 802.11i does not protect management frames. IEEE 802.11w is working on a standard to protect some management frames [31, 40, 41]. Management is a large component of the 802.11 specification. Many types of management frames can be used to avail simple services on a wired network. It is easy to establish the identity of a network station on a wired network. This is because network connections need dragging wires from a central location to the new workstation. For several cases, patch panels in the wiring closet can be utilized to expedite installation. However, the crucial point is: new network connections can be authenticated by a personal visit when the new connection is brought up [40].

Management features must create wireless networks to provide similar functionality. The procedure is broken up into three components by 802.11. First, mobile stations searching for connectivity must locate a compatible wireless network to utilize for access. Typically, with wired networks, this involves finding of the appropriate data jack on the wall. Second, the network must authenticate mobile stations so that the authenticated identity is allowed to link to the network. The network itself provides the wired-network equivalent. If signals cannot leave the wire, obtaining physical access

somewhat represents an authentication process. Lastly, mobile stations must associate with an access point to gain access to the wired backbone. This final step is equivalent to plugging the cable into a wired network [40].

- iii. Control Frame:** Control frames are used for requesting and controlling access to the wireless media. An example of a control frame is the acknowledgement frame, which is used after data frames to ensure reliability. Its primary purpose is to alert the sender that the last frame was received correctly and there is no need to retransmit. This simple positive acknowledgement following each frame is expected, or the frame is considered lost. These frames can easily be forged, since IEEE 802.11i does not protect control frames [40]. Control frames assist in the delivery of data frames. They administer access to the wireless medium (but not the medium itself) and provide MAC-layer reliability functions [41], as shown in Figure 2.30.



**Figure 2.30: Control frame [31]**

All control frames use the same Frame Control field, which is shown in Figure 2.28.

- **Protocol version:** The protocol version is shown in Figure 2.30 has field to assign value depend on the version. Other versions may exist in the future [41].
- **Type:** Control frames are assigned the Type identifier 01. By definition, all control frames use this identifier.

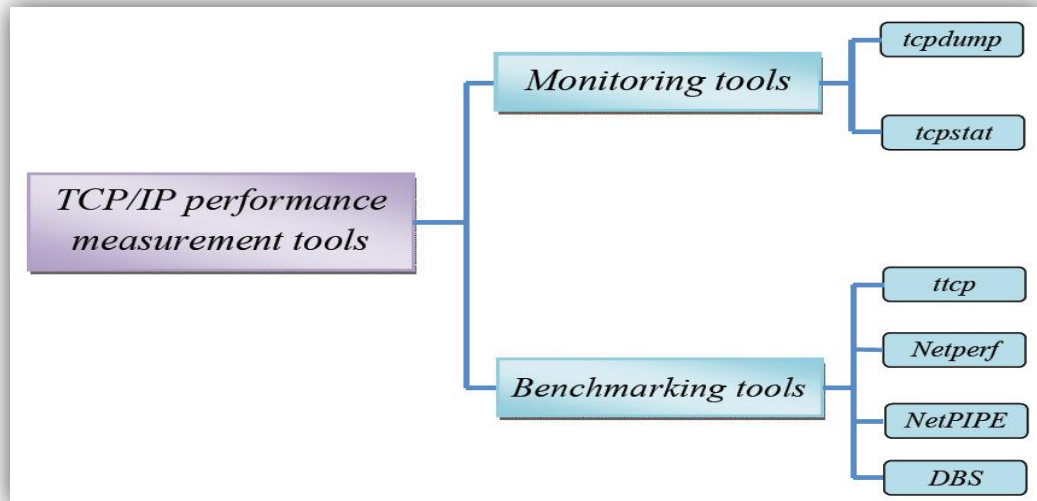
- **Subtype:** This field indicates the subtype of the control frame that is being transmitted.
- **To DS and from DS bits:** Control frames arbitrate access to the wireless medium. Hence, this can only originate from wireless stations. Since the distribution system does not transmit or receive control frames, these bits are always 0 [31, 40, 41].
- **More Fragments bit:** Control frames are not fragmented, so this bit is always 0.
- **Retry bit:** Control frames are not queued for retransmission like management or data frames, so this bit is always 0.
- **Power Management bit:** This bit is set to indicate the power management state of the sender after conclusion of the current frame exchange [40].
- **More Data bit:** The More Data bit is used only in management and data frames, so this bit is set to 0 in control frames.
- **WEP bit:** Control frames may not be encrypted by WEP, which may be used only for data frames and association requests. Hence, the WEP bit is always 0, for control frames [41].
- **Order bit:** Control frames are used as components of atomic frame exchange operations. Therefore, they cannot be transmitted out of order, and this bit is set to 0 [31].

## **2.6 PERFORMANCE MEASUREMENT OF WIRELESS NETWORK**

Network performance which is considered a complex issue has various independent variables impacting the clients' access to servers in a network. The majority of the elements that enables the performance of networks can be condensed into a few simple network principles which are measureable, controlled and monitored by the network administrator through simple and often free software [43]. The network's performance is affected by its protocol specification, internal algorithms, its implementation, memory management mechanism in its underlying operating system, system architecture, and the communication channels utilized. Networking measurement can enable the identification of possible bottlenecks, and lack of parameter settings. It can also examine the reliability of the component in complex TCP/IP networking environments [1]. It is a difficult task to measure network, whereby there are also some subtasks, such as Data collection, Analysis, Presentation and Interpretation.

## **2.7 TYPES OF MEASUREMENT TOOLS**

Several tools have been developed to measure and appraise the performance of the network, which are most valuable to assess, monitor or watch traffic, and benchmark several TCP implementations [17, 43]. Measurement tools can be classified based on a number of characteristics as in Figure 2.31.



**Figure 2.31: Types of TCP/IP performance measurement tools [17]**

### 2.7.1 WIRELESS MONITORING

The measurement of traffic from a wireless distant point (wireless monitoring) is popular in product development of both wireless research and commercial WLAN management. This provides comprehensive PHY/MAC information which is found on a wireless medium. To monitor security and diagnose network to detect anomalies, extensive wireless information is invaluable compared to that which is offered by SNMP or wired monitoring. The system consists of sniffers used to study traffic features upon the wireless medium. Hence, wireless monitoring enables the comprehension of traffic features and the detection of anomalies in a wireless network [44, 45]. In addition, wireless monitoring also highlights the features under study on the wireless medium so that PHY/MAC characteristics can be inferred. In sum, it allows the examination of the surface of the information which links factors such as signal strength, noise user arrivals and session durations with structured models. Some examples of monitoring tools are tcpdump and ttcp [46], this tools is been described in chapter three.

### **2.7.2 WIRELESS BENCHMARKING**

Wireless benchmarking carries out performance evaluation and provides performance information that is representative of the actual systems. This information can be used, either for the verification of performance modeling and simulation, or for performance changes detection [47]. It is utilized extensively for tracking changes in performance while carrying out software development, thus enabling feedback reception to developers, simultaneously while the work is going on. Besides this, it is very useful for testing traffic through the network and to measure varying performance indices, for e.g., throughput, delay, jitter, and approximated bandwidth. These measurements are often used in light of network management but they can also be useful for the implementation of TCP protocol or for the tuning of TCP implementations [1]. Examples of benchmarking tools are: `ttcp`, `Netperf` [48], this tools is been described in chapter three.

### **2.8 HIGHER INSTITUTION LEARNING OF NETWORK**

The analysis of network traffic and user behavior in different WLAN environments has been conducted by several studies [49-51]. Universities mostly use WLAN measurement [8, 9, 45, 52]. This study also examines the WLAN traffic in the Universiti Utara Malaysia, by analyzing the traffic over a period of three weeks, whereby data is captured using three different techniques (i.e., `tcpdump` traces, `wireshark` traces and `ntop`).

Tang and Baker's [52] study on the Stanford University Computer Science Department building was one of the pioneering studies on this issue. The researchers carried out an examination of the wired monitoring traces and the SNMP logs to try to analyze a twelve-week trace of a WLAN. In the public area wireless network,



Balachandran et al. [50] successfully collected the traces collected. The usage of the WLAN at the Saskatchewan University Campus was presented in [8]. The campus consists of 40 buildings encompassing public spaces (i.e., lounges, libraries, coffee shop, etc.), classrooms, laboratories and offices. Traffic trace was collected for a week, from in January 2003 by using Ether Peek, (software package allowing the recording of MAC addresses and traffic load information.

MAC addresses were matched with the authentication logs obtained from each of the 18 APs of the campus. In total, 134 unique users connected to the network. Individual users visited at most 8 different APs. Data recorded over three days at the ACM SIGCOMM 2001 conference were analyzed in [50] , the authors focused on modeling individual user bandwidth requirements and traffic loads on individual APs. They found that users distributed evenly across all APs.

The analysis of the Dartmouth College wireless network by Kotz and Essien [53] is most relevant to campus-wide networks. The wireless network in Dartmouth College consists of 476 APs offers covering in 161 buildings catering to 2000 users. The Dartmouth study used a combination of three forms of trace collecting: SNMP polling, packet header recording, and event-triggered log messages. The results show that network activity demonstrates clear patterns as follows: about fifty percent of the users were active on a typical day, and about one third of this number was mobile. It also illustrated a typical student's pattern of activity, whereby there was lower activity on Fridays and Saturdays, and accelerated activity on Sundays.

It is difficult to generalize the results in these works because of the low number of users observed (e.g., 74, 134 and 195, respectively). Besides, knowledge of the tracing may have upset user behavior in [52]. This study also indicates that no effort was made to ensure that the three weeks of analysis was representative of overall

usage patterns. Results in [50] are very specific to the conference settings. Hutchins et al. analyzed the WLAN at the Georgia Tech Campus over five months [54]. They extracted information about user behavior from the authentication logs at the firewall. The results present a strong diurnal cycle with peaks in the afternoon and higher activity during working days. They also collected mobility data through SNMP polling of AP association tables: 35% were static users, and 13% moved within one building, while the rest moved within the eighteen buildings of the campus.

In 2004, they revisited the WLAN [55] and found that, despite a drastic increase in traffic, users were mainly non-mobile. Similar user patterns were found in a corporate network from July 20 to August 17, 2002 [56]. Despite mobility results report higher mobility than on university campuses, users still spend a large fraction of time at one location. The results regarding the daily and weekly trends are similar to those observed on university campuses.

By studying multiple traces from different environments collected at different times, Hsu and Helmy [57] found that most traces display similar trends, but the details differ due to differences in population, environment, time and methodologies of trace collection. Their findings show that unrealistic assumptions are often undertaken in user modeling and computer simulation. One of the major problems for researchers who investigate the association patterns and session lengths of real WLAN users is the ability to separate from the traces those continuous associations and disassociations of the same user with several APs (the ping pong effect) because they can affect the correct interpretation of associations patterns. This problem has been addressed in many studies [8, 9, 58]. Our study investigates the usage patterns and performance for three weeks of the UUM WLAN .

## 2.9 SUMMARY

Wireless technology is operated by a radio waves or the radio frequency which its one member of spread spectrums. Modulation data from bits or electrical pulses to signals carried by the air as a media and demodulate it required achieved by special devices, this process represents the core tasks for any wireless devices, although despite the several of its type. Vary of modulation according to set of criteria such as first type of protocols e.g. Bluetooth, infrared, Wi-Fi, and cellular phone “this included it happens same protocol has a different way of modulation e.g. Wi-Fi standards 802.11b using DSSS while 802. 11n using OFDM, second type of used frequency e.g. free frequency or unique, third modulation schemes such as ASK, FSK and PSK which in turn divided for Quadrature phase shift keying QPSK, 16-level quad amplitude modulation 16-QAM and 64-QAM its consider effective element regarding power consuming, fourth the method of spread spectrum transmission is the most important factor that characterizes each standard of Wi-Fi, also consider major factor for develop and finding the new technology, finally coverage distance. Inside wireless itself as a technology there are many driven protocols like Wi-Fi, cellular phone, and satellite. Most popular of them and with wide using definitely the Wi-Fi introduces many of the standards since 1999 until this time couple of reason led to Wi-Fi consider a welcomed technology everywhere. Due to easy, flexibility, cheap and available beside rest features, although these features which stand with Wi-Fi occurring, important issue according to users” needs caused for looking new technology and main purpose of that enhanced performance. Regarding to the reasons which leading to resort a new technology, the lacks in Wi-Fi coverage area for vast place because of its own design as well as depends on free frequency and because of using in a large scale cause this to occur interference. Slow wireless connection is a popular complaint from students

on this kind of wireless network. In order to solve these problems must be identify of the layers OSI and TCP/IP. This study requires some tools for deals with traffic are monitoring tools and benchmarking tools.

## **CHAPTER THREE**

### **RESEARCH METHODOLOGY**

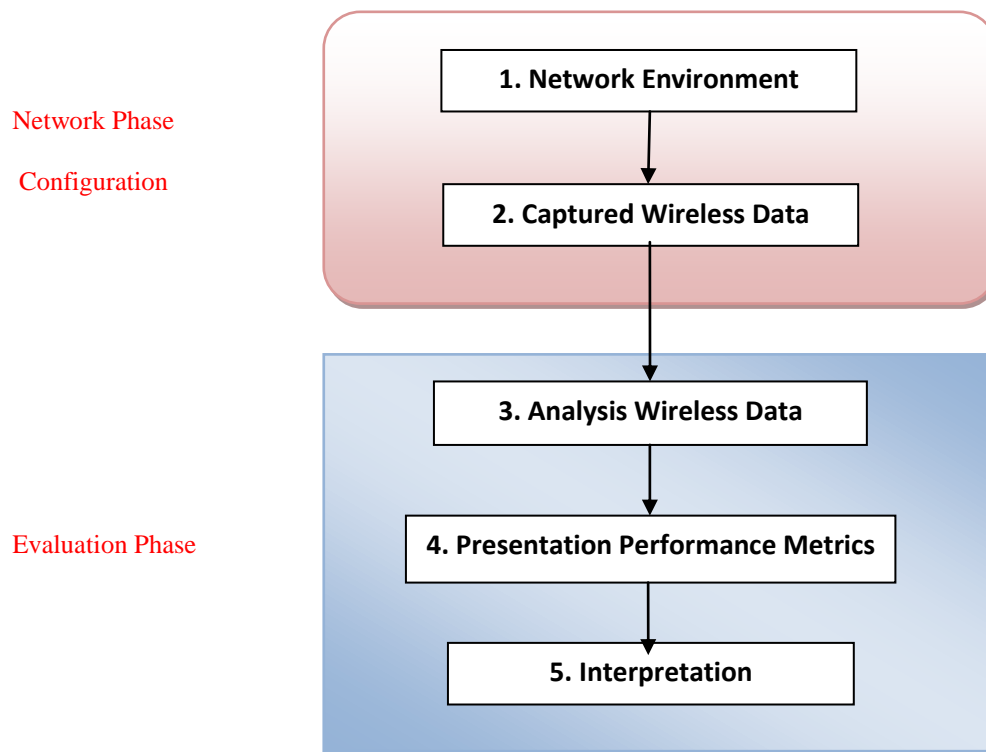
#### **3.1 INTRODUCTION**

Research methodology encompasses more than the methods for construction research and beyond what is known as a systematic approach to solve the research problem. A research methodology involves the combination of approaches that the researcher implemented in the research steps; for instance, data collection method, data processing method, instruments and evaluations as depicted in Figure 3.1. The research methodology for this study is described in detail in the following sections.

#### **3.2 RESEARCH METHODOLOGY**

A researcher utilizes suitable techniques and methods in an organized order in the research methodology because each step hinges on the one before it. Research methodology differs from one paradigm to another based on the style utilized. A systematic sequence is followed to the established set of events; steps that are akin to a waterfall system in software engineering as the next step adopts the output from the step before it as its own input. The group of blocks forming the methodology of the present study is clarified in Figure 3.1 [59]

The study basically hinges on the two levels with the first one being the network configuration phase while the second one is the evaluation phase. The former comprises of two steps which are network environment and captured wireless data phase. The latter on the other hand comprises of three steps which are the analysis of wireless data, the presentation of performance metrics and the interpretation. Therefore, a total of 5 levels are involved which are illustrated in Figure 3.1.



**Figure 3.1: Phases of Methodology [17]**

### **3.3 NETWORK PHASE CONFIGURATION**

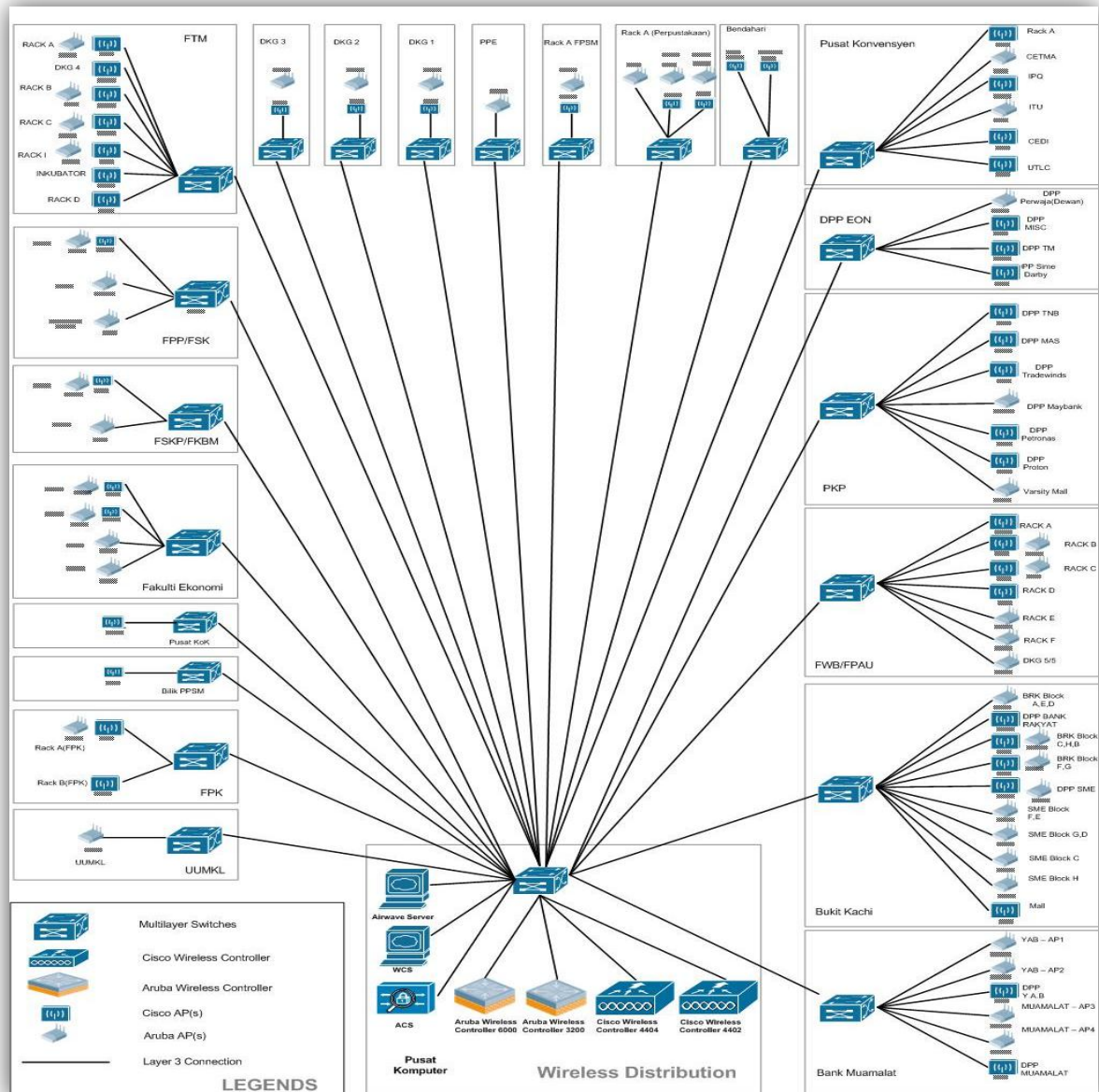
As mentioned, the network configuration phase comprises of two steps. These steps are described in detail in the following sub-sections.

#### **3.3.1 NETWORK ENVIRONMENT**

The wireless network in UUM operated as a switched network during data collection of the study implying that the University's Internet center was using a single subnet.

The UUM network links to the Internet by a Cisco router and it is made of a virtual network existing on a particular subnet distinct from the campus. The wireless network comprises of a total of four controller models including two Cisco Wireless Controller Model 4404 and two Aruba Wireless Controller models 6000 that are

linked to the Core multilayer switch in the computer center. All the Access Points (APs) in FTM, PK, DKG, Pusat Konversyen, DPP, EDC, FPAU, Sultanah Bahiyah Library, and DPPYAB of UUM totaling around 91 APs are managed by the Cisco wireless controller Model 4404. The distribution of these APs on nine buildings and another places of UUM is: 6 APs in Bank Muamalat, 13 APs in Bukit Kachi, 9 APs in FWB/FPAU, 7 APs in PKP, 4 APs in DPP EON, 6 APs in Pusat Konvensyen, 2 APs in Bendahari, 3-5 APs in Rack A (perpustakaan) /library, 2 APs in Rack A FPSM, 1 AP in PPE, 2 APs in DKG1, 2 APs in DKG2, 2 APs in DKG3, 11 APs in FTM, 4 APs in FPP/FSK, 3 APs FSKP/FKBM, 6 APs in Fakulti Ekonomi, 1 AP in Pusat KoK, 1 AP in Bilik PPSM, 3 APs in FPK, and 1 APs in UUMKL, as presented by Figure. 3.2.



**Figure 3.2: UUM Wireless**

Additionally, every building within the campus has APs linked with a multilayer Switch and these switches are linked to the Core switch in the Computer Center. The process can be described as – the wireless devices have packets sent to and from them and these packets travel on the same network as the UUM general traffic. Upon connection, the wireless devices are provided with IP addresses through a DHCP server upon which the wireless traffic is shifted to the UUM router. This is in turn



shifted to the internet and it returns to the campus network in the normal subnet. Hence, users who are using unauthorized wireless are blocked from linking to the UUM servers and internet in a direct manner.

### **3.3.2 DATA CAPTURING**

The wireless sniffer captures the first one thousand bytes of 802.11 frames and it proceeds to record the complete view (PHY/MAC/LLC/IP/Above-IP information) with the header containing useful PHY information like MAC Time, RSSI, SQ, Signal Strength, Noise, Signal Noise Ratio, and data rate. The entire information of signal and noise are placed in manufacture-specific units although they can also be utilized for comparisons [44]. The present study attempts to capture the IEEE 802.11 MAC frame structure which contains fields including, protocol version, frame type (management, data and control), Duration for Network Allocation Vector (NAV) calculation, BSS Id, Source and Destination MAC addresses, fragment, and sequence number etc [60]. Moreover it gathers information that is unprocessed depending upon the data utilization from the operational network. For example, the data needed to drive simulation may encompass the arrival time of individual packets as well as their sizes in bytes. The UUM wireless network's capture of data steps include:

#### **i. Port Mirroring**

A mirror port or what is commonly known as the Switched Port analyzer (SPAN) facilitates the duplication of all the traffic stemming from or ending at a single client device or access point to another port. It is considered invaluable in highlighting particular network issues. In addition, mirror mode should be utilized on the unused port as basically any connection made to this port shows

unresponsiveness [61]. The network traffic is selected by a network analyzer for the purpose of analysis. This network analyzer can be a Cisco Switch Probe device or other Remote Monitoring (RMON) probe. In the past, SPAN was considered as the general feature on the Cisco Series switches. Table 3.1 presents the Switches supporting SPAN, RSPAN and ERSPAN[10].

**Table 3.1: Presents the Switches supporting SPAN, RSPAN and ERSPAN [10]**

Switches	SPAN Support	RSPAN Support	ERSPAN Support
<b>Catalyst Express 500 / 520 Series</b>	Yes	No	No
<b>Catalyst 6500/6000 Series</b>	Yes	Yes	Yes Supervisor 720 with PFC3B or PFC3BXL running Cisco IOS Software Release 12.2(18)SXE or later. Supervisor 720 with PFC3A that has hardware version 3.2 or later and running Cisco IOS Software Release 12.2(18)SXE or later
<b>Catalyst 5500/5000 Series</b>	Yes	No	No
<b>Catalyst 4900 Series</b>	Yes	Yes	No
<b>Catalyst 4500/4000 Series (includes 4912G)</b>	Yes	Yes	No
<b>Catalyst 3750 Metro Series</b>	Yes	Yes	No
<b>Catalyst 3750 / 3750E Series</b>	Yes	Yes	No
<b>Catalyst 3560 / 3560E Series</b>	Yes	Yes	No
<b>Catalyst 3550 Series</b>	Yes	Yes	No
<b>Catalyst 3500 XL Series</b>	Yes	No	No
<b>Catalyst 2970 Series</b>	Yes	Yes	No
<b>Catalyst 2960 Series</b>	Yes	Yes	No
<b>Catalyst 2955 Series</b>	Yes	Yes	No
<b>Catalyst 2950</b>	Yes	Yes	No

Series			
Catalyst 2940 Series	Yes	No	No
Catalyst 2948G-L3	No	No	No
Catalyst 2948G-L2, 2948G-GE-TX, 2980G-A	Yes	Yes	No
Catalyst 2900XL Series	Yes	No	No
Catalyst 1900 Series	Yes	No	No

## ii. Configuring Port Mirroring

To enable port mirroring of controller page, following steps were done:

**Step 1** Choose **Controller > Ports** to open the Ports page

**Step 2** Click the number of the unused port for which the researcher want to enable mirror mode. The Port > Configure page appears.

**Step 3** Set the Mirror Mode parameter to **enable**.

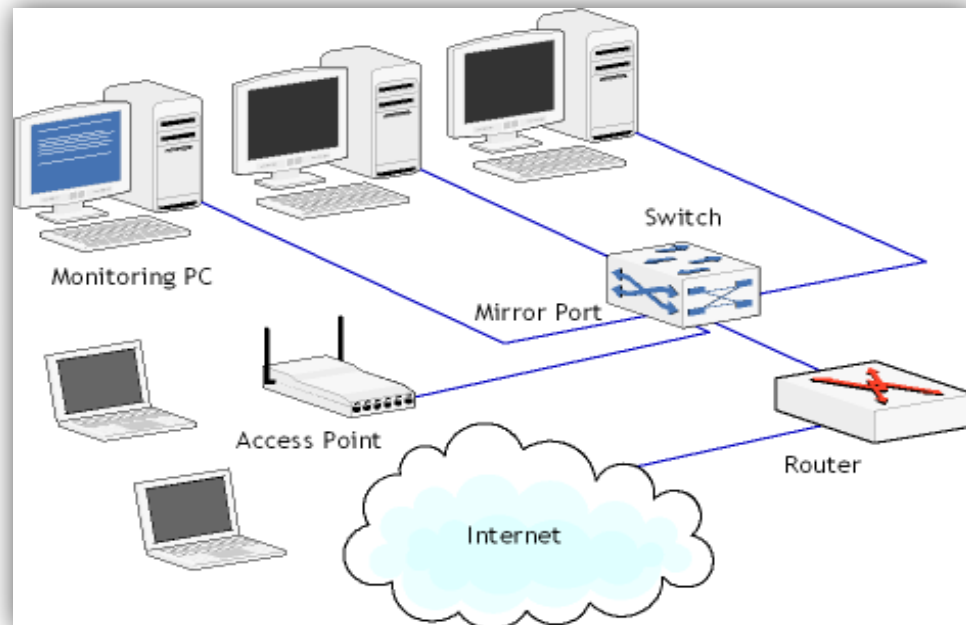
**Step 4** Click **Apply** to commit your changes.

**Step 5** Perform one of the following:

- To choose a specific client device that will mirror its traffic to the port, following steps were done:
  - ☐ Choose **Wireless > Clients** to open the Clients page.
  - ☐ Click the MAC address of the client for which the researcher want to enable mirror mode. The Clients > Detail page appears.
  - ☐ Under Client Details, set the Mirror Mode parameter to **enable**.
  - ☐ To choose which access point that will mirror its traffic to the port, following steps were done:
    - ☐ Choose **Wireless > Access Points > All APs** to open the All APs page.

- Click the name of the access point for which he want to enable mirror mode. The All APs > Details page appears.
- Choose the **Advanced** tab.
- Set the Mirror Mode parameter to **enable**.

**Step 6** Click **Save Configuration** to save your changes.



**Figure 3.3: The mirror port [61]**

### iii. Creating a Local Port Mirroring / APAN Session

Starting in privileged EXEC mode, for create a SPAN session and specify the source (monitored) ports or VLANs and the destination (monitoring) ports:

**Step 1** configure terminal.

**Step 2** no monitor session {*session number*| all |local | remote.

**Step 3** monitor session *session number* sources **Step 4**monitor session *session\_number* destination {**interface** *interface-id*

**Step 4** end.

#### **iv. The Operating System and Tools**

In the present study, Linux is the operating system used while tcpdump and Wireshark are utilized as tools for capturing data.

- **Linux Operating System**

Linux, can be described as a Unix-like operating system, hence the name, created for the provision of free or reasonable cost operating system to computer users and it is comparable to tradition systems [62]. It is known for being very efficient and fast-performing. Linux kernel, which is the core of the operating system, was created by Linus Torvaldsat of the University of Helsinki, Finland. In an attempt to complete the operating system, Torvalds and his colleagues utilized system components created by the members of the Free Software Foundation for the GNU Project [63]. Overall, Linux is considered as a multi-tasking, multi-user operating system enabling a number of people to run varying applications on one PC simultaneously. It is different from MS-DOS in a way that in the latter, only a single person is able to utilize the system at one time. In Linux, for the identification of an individual to the system, logging in is a must entailing the user to enter name and password – the personal key for logging into a personal account. The password is only known to the person and hence, the account is kept confidential [64].

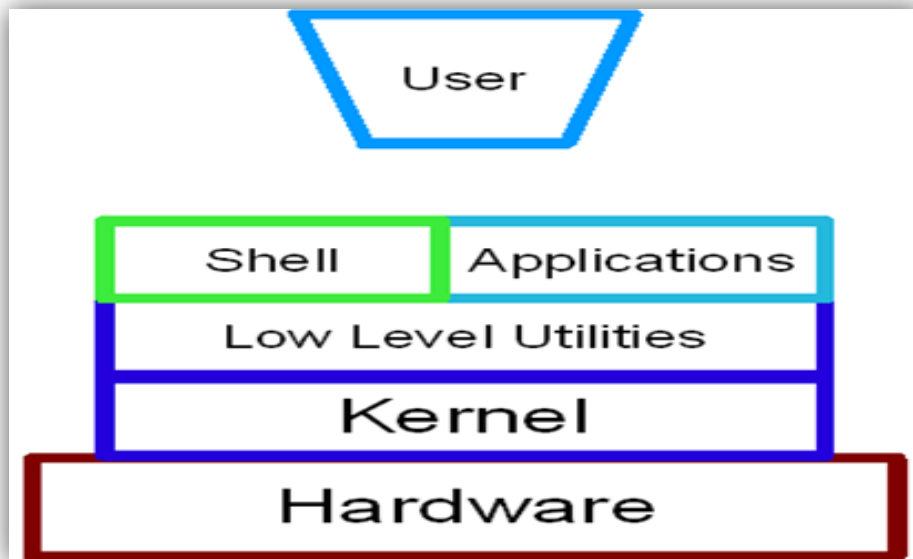
- **Linux Advantages**

- Linux is almost free system.
- It is portable and is compatible with any hardware platform [65].
- It is secured and flexible.
- It is scalable.

- Its Linux OS version and most of its applications are characterized by extremely short debug-times [65].

#### □ **Layers of Linux**

The Linux has three very impotent parts are: Kernel, Shell, and File system, as Figure 3.4.



**Figure 3.4: Layers of Linux [65]**

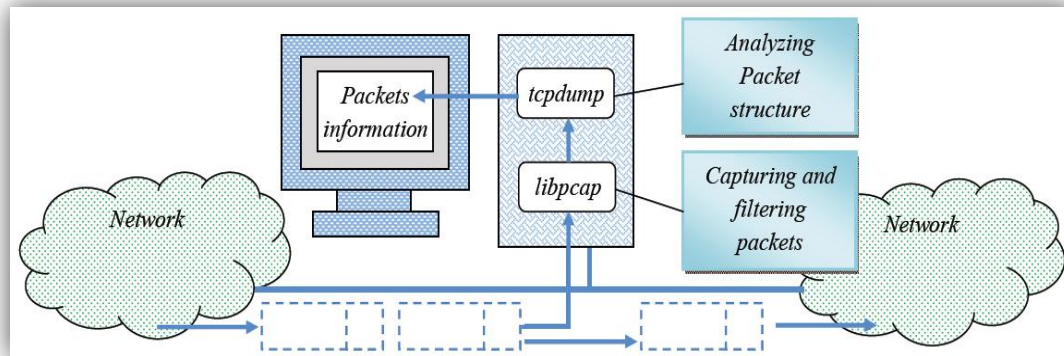
The kernel functions as the core of the operating systems and the interface that connects the applications and hardware. Shell is the interface between the user and the kernel. In Linux, everything is treated as a file even hardware device and it is arranged in a directory system [66].

- **Tcpdump**

Tcpdump is a tool that enables the interception and capture of packets going through a network interface, and thereby making it invaluable in understanding and solving network layer issues. It assists in monitoring of packet flow into the interface, the response of each packet, the packet drop and it provides information regarding ARP. This tool is not useful in the

identification and troubleshooting problems that are Application related prior to the understanding of the behavior of the problem. Tcpdump's main function [67] involves the monitoring of packets that are connected to the network, the dumping of headers, and the organization of payloads of packets into a readable format. Tcpdump's uses are listed in Figure 3.5. In monitoring traffic, it turns the network interface into promiscuous mode and thus, it accepts all the packets from the network and displays them in varying forms up on the console or stores them in files for analysis in a later stage [17, 43]. It makes use of the packet capture library (libpcap) whose development is attributed to Van Jacobson and its utilization has been prevalent in varying traffic monitoring applications. The library appropriates an Application Program Interface (API) to a system of packet capture that is developed in UNIX kernel and it facilitates packet filtering of packets gathered by the interface. The Tcpdump program acquires filtering specifications from command line arguments and passes them through in an organized way to the in-kernel packets capturing tool through the API facilitated by libpcap as depicted in Figure 3.5.

Majority of the UNIX distributions comprises of the tcpdump program but owing to security reasons, most of them are not installed by default. Those that are installed require root privileges for their running [43].



**Figure 3.5: Program Structure of Tcpdump [17]**

- **Using of Tcpdump With Option**

The first step in using Tcpdump is to know which interface is to be monitored. A lot of UNIX systems possess multiple network interfaces and PPP connections. The interface to be monitored must be determined and it can be shown with the help of programs such as UNIX ifconfig (Ethernet interface)/iwconfig (wireless interface). For the purpose of displaying the entire active network interfaces of the system, the `-a` option [17, 43, 67] has to be utilized. The Tcpdump is a simple and user-friendly tool and it supports many command lines of options for the user's convenience and versatility [17]. Tcpdump monitors are characterized as having the fewest number of active interfaces by default which is listed in the ifconfig output. In order to select a different interface, the `-i` option must be chosen [43]. There are varying command-line options that can be utilized for the modification of the Tcpdump program behavior. The table below Table 3.2 depicts the command-line options that are available. The program also has a filtering mechanism to store the packets chose. The filter is mentioned as an expression in the command line options [17].



**Table 3.2: The Tcpdump Command-Line Options [43]**

Option	Description
<b>-a</b>	Attempt to convert network and broadcast addresses to names.
<b>-c</b>	Exit after receiving count packets
<b>-C file size</b>	If the file is larger than file size, close the current save file and open a new one.
<b>-dd</b>	Dump packet-matching code as a C program fragment.
<b>-ddd</b>	Dump packet-matching code as decimal numbers (preceded with a count).
<b>-e</b>	Print the link-level header on each dump line.
<b>-E algo:secret</b>	Use algo: secret for decrypting IPsec ESP packets.
<b>-f</b>	Print foreign Internet addresses numerically.
<b>-F file</b>	Use file as input for the filter expression.
<b>-i interface</b>	Listen on interface.
<b>-l</b>	Make stdout line buffered.
<b>-m module</b>	Load SMI MIB module definitions from file module.
<b>-N</b>	Don't print domain name qualification of hostnames.
<b>-O</b>	Don't run the packet-matching code optimizer.
<b>-p</b>	Don't put the interface into promiscuous mode.
<b>-q</b>	Quick output. Fewer lines per packet are displayed.
<b>-R</b>	Assume ESP/AH packets to be based on old specification.
<b>-r file</b>	Read packets from file.
<b>-S</b>	Print absolute, rather than relative, TCP sequence numbers.
<b>-s snaplen</b>	Get snaplen bytes of data from each packet. The default is 68 bytes.
<b>-T type</b>	This option specifies the type of packet (rtp, snmp, rtcp, vat, or wb).
<b>-t</b>	Don't print a timestamp on each dump line.
<b>-tt</b>	Print an unformatted timestamp on each dump line.
<b>-ttt</b>	Print the delta time between packets.

<b>-tttt</b>	Print a timestamp in default format preceded by date on each dump line.
<b>-u</b>	Print undecoded NFS handles.
<b>-v</b>	Show verbose output.
<b>-vv</b>	Show more verbose output.
<b>-vvv</b>	Show even more verbose output.
<b>-w file</b>	Write the raw packets to file rather than printing them out.
<b>-x</b>	Print each packet in hex.
<b>-X</b>	When printing hex, print ASCII text as well.

Many command line options can be combined together and the program allows the user to determine the command line options that is required, by the separation of each option with a space as depicted in Figure 3.6 below:

```
tcpdump [-adeflnNOpqStvx] [-c count] [-F file]
        [- i interface] [-r file] [-s snaplen]
        [- T type] [-w file] [Expression]
```

**Figure 3.6: Command line Options for Tcpdump[17]**

For instance, in order to enable the capturing mechanism in the Cisco Wireless Controller 4404 through port mirroring Core Switch, the following command is provided:

```
# tcpdump- i wlan0 – C 1000 – s 1000 – w capture.dmp
```

The above command line is a combination of many options, - i wlan0 commands Tcpdump to capture packets passing through the wlan0 interface

while `-c 1000` commands Tcpdump to capture 1000 MB sized files. `-s` commands the program to capture the first 1000 bytes of the packet which means only the headers as opposed to the data. In addition, `-w` Tcpdump files/dsldump.dmp commands the program to set up a binary file for the dumping location. In case the study needs to capture the header with payload, the following default command line is used (65535 bytes):

```
#Tcpdump- i wlan0
```

- **Tcpdump Output Format**

The format of the Tcpdump output explains the information to TCP header in one line [17].

Src>dst: flags data-seqno ack window urgent options,

Where,

**Src and dst** are the source and destination IP address and TCP port numbers respectively

**Flags** are mixture of commands S(SYN), F(FIN), P (PUSH), or R(RST) with a single “.” (no flags).

**Data-seqno** is a description of the portion of the phase space covered by the data in a specific packet.

**Ack** is the number sequence of the next expected data byte from the other connection end.

**Windows** is the amount of bytes that are present in the receiver buffer.

**Urgent** shows the urgency of the data in the packet, and

**Options** are the options of the TCP that are inside the angle brackets.

### **3.4 EVALUATION PHASE**

As mentioned, the network configuration phase comprises of three steps. These steps are described in detail in the following sub-sections.

#### **3.4.1 ANALYSIS WIRELESS DATA**

Analysis wireless network is also referred to many terms including protocol analysis, sniffing, packet analysis, traffic analysis, eavesdropping and many other terms. It is described as the operation of gathering network traffic and conducting tests regarding its limitation and of determining what happens in the wireless network. A network analyzer basically decodes the data packets of common protocols and presents the network traffic in a format that is understandable. Sniffers that are unauthorized are considered threats to the network security owing to the difficulty of their detection and their insertion almost at any place; making them the hackers' weapon of choice. It comes either as a standalone hardware device having its own specialized software or software installed on a desktop or laptop computer. The difference between the types of network analyzers hinges on many features including the amount of supported protocols it is able to decode, the user interface, its graphing and statistical abilities. In addition, other differences may be in the form of inference capabilities, and quality of packet decodes [68]. Regardless of the fact that almost all network analyzers decode similar protocols, some works better than the rest according to the environment. A network analyzer is basically used to troubleshoot network problems, convert binary data in packets to formats that are understandable, to analyze the network performance, to discover bottlenecks, to detect network intrusion, log network traffic for the purpose of forensics and evidence, to analyze the application operations, to discover faulty network cards, to discover the source of virus outbreaks or denial of

service attacks, to detect spyware, to program network, to debug the system during the development stage, to detect a compromised computer to validate compliance with company policy, to be used as an educational resource when learning of protocols, to utilize in reverse-engineering protocols and to write clients and supporting programs [68]. In certain cases, the analysis of raw data is required to acquire certain features (for instance, throughput, latency) of data gathered [17]. The analysis of wireless data of UUM wireless network involves the following steps:

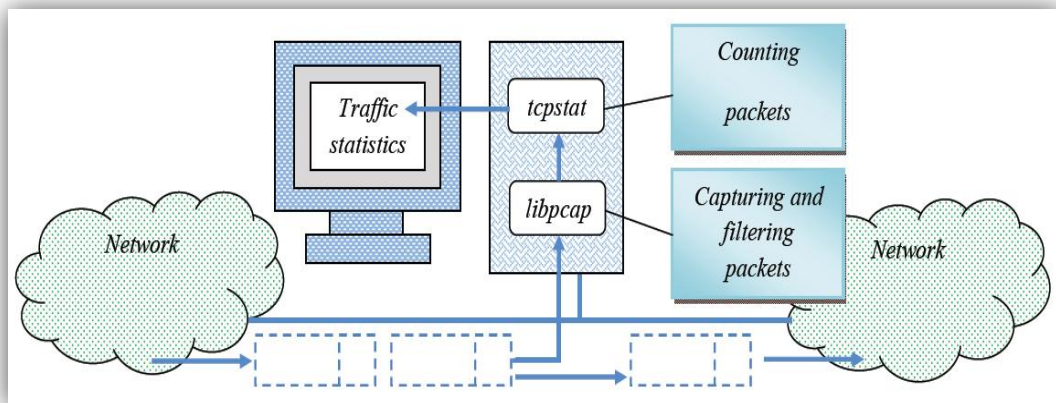
**Step1** – Following the capture of wireless traffic through the use of Tcpcap, a file is created; for instance, rawdata.cap or rawdata.dmp, and packets are stored continuously.

**Step2** – The tcpstat tool is utilized for the classification of the packets into different kinds and to appropriate statistics for each type. In the same step, several options can be made.

- i. The following can be measures: the number of ARP packets, the average packet size, the number of byte per second, the number of bits per second, the number of ICMP and ICMPv6 packet, the greatest packet size, the smallest packet size, the number of bytes, the number of packets, and the number of packets for every second.
- ii. The transport layer protocols can be investigated to know the number of TCP packets and the number of UDP packets.
- iii. The throughput and the load can be measured.

- **Tcpstat**

This is referred to as the network interface statistics reporting tool as it reports particular network interface statistics which is comparable to `vmstat` in its role in system statistics. The tool collects information through monitoring particular interface or through the reading of prior saved `Tcpdump` data gathered in a file. It reports TCP related statistics in a periodic manner on a particular network interface [17, 69]. Some statistics calculated through `tcpstat` include: the bandwidth used, the exchanged number of packets, the average packet size of a TCP stream, packets for every second, standard deviation of packet size, interface load and others. Its default action is the automatic search for a suitable interface and the presenting of current statistics on it. Similar to `Tcpdump`, `tcpstat` also utilizes the `libpcap` library. The functional structure of the program is presented in the figure 3.7 below Figure.



**Figure 3.7: Functional diagram of `tcpstat` [17]**

- **Usage of `Tcpstat` with Options**

Similar to `tcpdump`, the `tcpstat` also requires the setting of the wireless network interface into the promiscuous mode so that it can be considered in

the super user privilege. It should also be placed in a suitable network interface for monitoring and capturing of packets appearing on the wireless network interface and finally, it draws up a summary of the statistics after every five seconds. For the selection of the most suitable network interface, tcpstat needs to have list of the network interface and it can choose any active interface from with the exception of loopback interface [17, 69]. In the gateway multiple network, the interface has to be set at active mode to allow the monitoring of the interface and for the determination of option – in other words, to enable live capture as opposed to reading from a file on the interface provided by the command line. In case of automatic interfaces, the tcpstat program looks for the suitable interface on its own. A brief description of the options is depicted in the table 3.3 below.

**Table 3.3: Command Line Options of Tcpstat [17, 69]**

Options	Description
<b>a</b>	Accounting mode, displays the estimated number of bytes per second, minute, hour, day, and month.
<b>b bps</b>	Bandwidth mode. Displays the total number of seconds the data-throughput exceeded bps, and the percentage of total time this was, as if the interface were limited to bps bits per second. See the NOTES section below to see how the interval affects bandwidth calculation.
<b>B bps</b>	"Dumb" bandwidth mode. Displays the total number of seconds the data-throughput exceeded bps, and the percentage of total time this was. See the NOTES section below to see difference between "dumb" and normal bandwidth modes.
<b>e</b>	Suppresses the display of empty intervals.
<b>F</b>	Flush all output streams after printing each interval. Sometimes useful when redirecting output into a file, or piping tcpstat into another program like.
<b>f filter expr</b>	Filter the packets according the rules given by filter expr. For the syntax of these rules.
<b>h, -?</b>	Display version and a brief help message.
<b>i interface</b>	Do a live capture (rather than read from a file) on the interface interface given on the command line. If interface is "auto" then tcpstat tries to find an appropriate one by itself.

<b>l</b>	Include the size of the link-layer header when calculating statistics. (Ethernet only, right now. Usually 14 bytes per packet.).
<b>p</b>	Set the interface into non-promiscuous mode (promiscuous is the default) when doing live captures.
<b>o format</b>	Set the output format when displaying statistics. See the OUTPUT FORMAT section below for a description of the syntax.
<b>R seconds</b>	Show the timestamp relative to seconds. Avoid this option, because it will most likely go away in future versions.
<b>r filename</b>	Read all data from filename, which may be a regular file, a named pipe or "-" to read it's data from standard input. Acceptable file formats include pcap (tcpdump files) and "snoop" format files. filename is usually a file created by the tcpdump command using the "-w" option.
<b>s seconds</b>	When monitoring an interface, tcpstat runs for only seconds seconds, and then quits. When reading from a data file, tcpstat prints statistics for second's seconds relative to the first packet seen.

Many command line options can be a combination of two or more commands.

Tcpstat permits the user to determine the command line options as need be with each option separated by a space as illustrated in figure 3.8.

```
tcpstat [-? haeFlp] [-B bps] [-b bps] [-F filter expr ]

        [- i interface] [-o output]

        [-r filename] [-s seconds] [Interval]
```

**Figure 3.8: Command line options of tcpstat [69]**

For instance, in the present study, the timestamp, the number of packets, the average packet size, the standard deviation of the packet size and the bandwidth in bits per second is determined after every five seconds for the wireless interface referred to as wlan0;

```
# tcpstat -i wlan0
```

Where:



**-i** is a live capture as opposed to reading from the file of the interface provided on the command line

**wlan0** is the wireless network interface,

# Tcpstat is represented by **-r capture.cap**

Where,

**-r** is the command for read all data from file name and

**capture.cap** is the file created through the use of tcpdump to be read offline.

- **Tcpstat Output Format**

The output string is considered as any quoted string that tcpstat will write to the stdout. Moreover, tcpstat exchanges specific values in substrings which begin with % and most standard print “\” escape characters. Also, the string that follows option **-o** can be any string and tcpstat writes this particular string to the standard output. The table 3.4 below contains a list of the substitution strings [69].

**Table 3.4: Substitution Strings [69]**

Substitution String	Description
<b>%A</b>	The number of ARP packets.
<b>%a</b>	The average packet size in bytes.
<b>%B</b>	The number of bytes per second.
<b>%b</b>	The number of bits per second.
<b>%C</b>	The number of ICMP and ICMPv6 packets.
<b>%d</b>	The standard deviation of the size of each packet in bytes.
<b>%I</b>	The number of IPv4 packets
<b>%l</b>	The network "load" over the last minute, similar to uptime.

<b>%M</b>	The maximum packet size in bytes.
<b>%m</b>	The minimum packet size in bytes.
<b>%N</b>	The number of bytes.
<b>%n</b>	The number of packets.
<b>%p</b>	The number of packets per second.
<b>%R</b>	Same as %S, but relative to the first packet seen.
<b>%r</b>	Same as %s, but relative to the first packet seen.
<b>%S</b>	The timestamp for the interval in seconds after the "UNIX epoch"
<b>%s</b>	The timestamp for the interval in seconds. microseconds after the "UNIX epoch".
<b>%T</b>	The number of TCP packets.
<b>%U</b>	The number of UDP packets.
<b>%V</b>	The number of IPv6 packets.
<b>%number</b>	Switch the output to the file descriptor number at this point in the string.  All output for each interval before this parameter is by default the standard output (file descriptor 1). Useful when redirecting the output into more than one file (or fifo) for separate statistics. Be sure you know where they are going. Writing to "dangling" file descriptors (without directing them to a specific destination) may produce unexpected results.

The default format string for tcpstat is:

```
"Time:%S\tn=%n\tavg=%a\tstddev=%d\tbps=%b\n"
```

This will produce an output which would look similar to:

```
Time: 940948785 n=107avg=251.81stddev=422.45 bps=43110.40
```

Where:

**Time:** is timestamp.

**n:** is the number of packets passed through the interface.

**avg:** is average packet size.

**stddev:** is the standard deviation of the packet size.

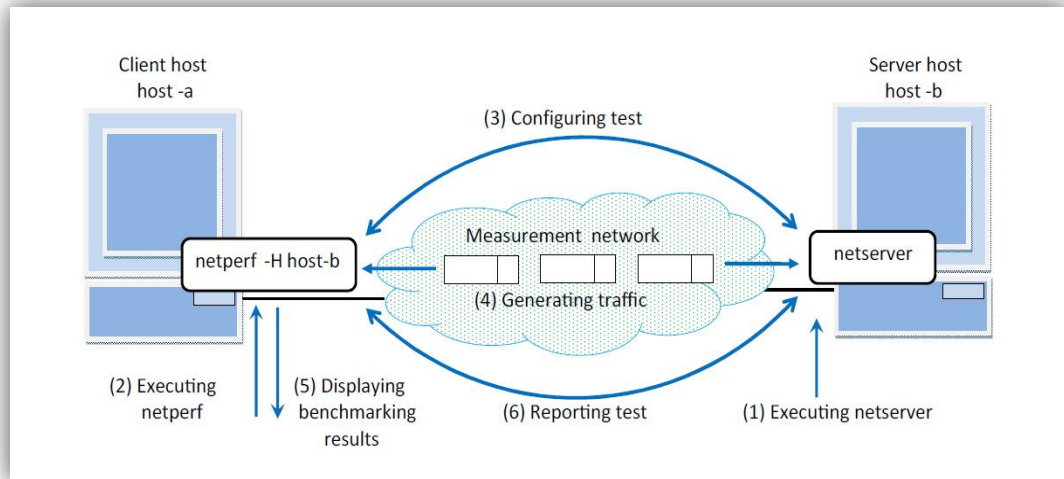
**bps:** bandwidth in bit per second.

**Step 3:** Make use of Netperf in the measurement of the present throughput between two particular hosts on library's wireless network through the active testing and the measurement of bulk data transference and interactive data exchange as well as latency between them. This is carried out through the use of:

- **Netperf**

Netperf is referred to as a network performance benchmark [70]. This can be utilized in the measurement of different aspects of networking performance. In the present study, the aim is the transference of bulk data and the request/response performance through the use of either TCP or UDP and the Berkeley Sockets interface. Moreover, other tests may be compiled in such as tests for DLPI, tests for Unix Domain Sockets, tests for IPv6.

The netperf is a tool developed by Rick Jones of Hewlett-Packard (HP) and its pioneering version was introduced in the mid-1990s. It comprises of two varying programs called the netserver and the netperf. The former acts as a server on the receiving host and is invoked by inetd super daemon while the latter is invoked on the host transmitting and operating as the client program for former [17]. When called for, netperf creates traffic to be sent to the netserver and carries out a measurement of performance. Figure 3.9 below shows an explanation of the interactions in the client server.



**Figure 3.9: The interactions for client-server [17]**

- **Usage of Netperf with Option**

At the onset, some rules have to be kept in mind regarding the netserver's role in the receiving host. No action is needed if the netserver's configuration is invoked from inetd super daemon; if not then the user has to initiate the server handle. While the netserver is running, the netperf has to be initiated at the transmitting host computer [17, 70]. In other words, both should be initiated and the option of performance measure is depicted in the following table 3.5.

**Table 3.5: Options for netperf [70]**

Option	Description
<b>-4</b>	Use AF_INET (aka IPv4) addressing for the control and possibly data connections.
<b>-6</b>	Use AF_INET6 (aka IPv6) addressing for the control and possibly data connections.
<b>-a sizespec</b>	Alter the send and receive buffer alignments on the local system. This defaults to 8 bytes.
<b>-A sizespec</b>	As -a, but for the remote system.

<b>-B brandstr</b>	Add brandstr to the output of a test with banners disabled.
<b>-c [rate]</b>	Request CPU utilization and service demand calculations for the local system. If the optional rate parameter is specified, netperf will use that instead of calculating the rate itself.
<b>-C [rate]</b>	As -c, but for the remote system.
<b>-d</b>	Increase the quantity of debugging output displayed during a test (possibly at the expense of performance).
<b>-D [secs,units]</b>	Display interim results at least every secs seconds using units as the initial guess for units per second. This is only available when netperf has been configured with --enable-demo.
<b>-f GMKgmK</b>	Change the units of measure for _STREAM tests. Capital letters are powers of two, lowercase are powers of ten.
<b>-F fill_file</b>	Pre-fill the send buffers with data from the named file. This is intended to provide a means for avoiding buffers that are filled with data which is trivially easy to compress. A good choice for a file that should be present on any system is this manpage - netperf.man. Other files may be provided as part of the distribution.
<b>-h</b>	Display a usage string, and exit.
<b>-H name[ip,family]</b>	Set the hostname (or IP address) and address family to use to establish the control connection to the remote system. Passing a single name with no comma will only set remote_host and will leave selection of address family for the control connection to the stack or by a -4 -r -6 command line option.
<b>-imax,min</b>	Set the maximum and minimum number of iterations when trying to reach certain confidence levels.
<b>-I lvl,[,intvl]</b>	Specify the confidence level (either 95 or 99 - 99 is the default) and the width of the confidence interval as a percentage (default 10).
<b>-l testlen</b>	Specify the length of the test (default 10 seconds). A negative value sets the number of request/response transactions, or the number of bytes for a stream test.
<b>-L name[ip,family]</b>	Set the local name IP and/or address family for the socket used for the control connection to the remote netserver.

<b>-n numcpus</b>	Specify the number of CPU's in the system on those systems for which netperf has no way to find the number of CPU's programmatically.
<b>-N</b>	This option will tell netperf to not establish a control connection to a remote netserver. Instead it will try to establish a data connection directly, using only the information supplied by the command line parameters and/or internal defaults. Unless other ports are provided by the command line, by default the data connection will be to the "discard" port for a "STREAM" or "SENDFILE" test, the "echo" port for an "RR" test or the "chargen" port for a "MAERTS" test.
<b>-o sizespec</b>	Set an offset from the alignment specified with -a.
<b>-O sizespec</b>	As -o, but for the remote system.
<b>-p portnum,localport</b>	Direct the control connection to a netserver listening on the specified port, rather than using a "netperf" entry in /etc/services or the internal default (port 12865). If ",localport" is specified the control connection will be established from that local port number. Specifying a single port number with no comma will specify only the remote netserver port number and will leave local port number selection to the stack.
<b>-P 0/1</b>	Show (1) or suppress (0) the test banner.
<b>-t testname</b>	Specify the test to perform. Valid test names include, but are not limited to, nor always compiled-in:  TCP_STREAMTCP_SENDFILETCP_MAERTSTCP_RRTCP_CRRUDP_STREAMUDP_RRDLCOSTREAMDLCO_RRDLCL_STREAMDLCL_RRSTREAM_STREAMSTREAM_RRDG_STREAMDG_RRLOC_CPUREM_CPU
<b>-T lcpu,remcpu</b>	Request that netperf be bound to CPU lcpu and/or netserver be bound to CPU rcpu.
<b>-v verbosity</b>	Set the verbosity level for the test (only with -P).

Steps to measure the throughput are:

- Start netperf as a daemon (Ubuntu): # Sudo /etc/initd/netperf/ start or stop.
- On a separate terminal: # ethstatus -I wlan0.

- Then start from the server: # netperf -H ip address of client.
- And vice versa on client: # netperf -H ip address of server.

After several seconds, the netperf shows the following performance:

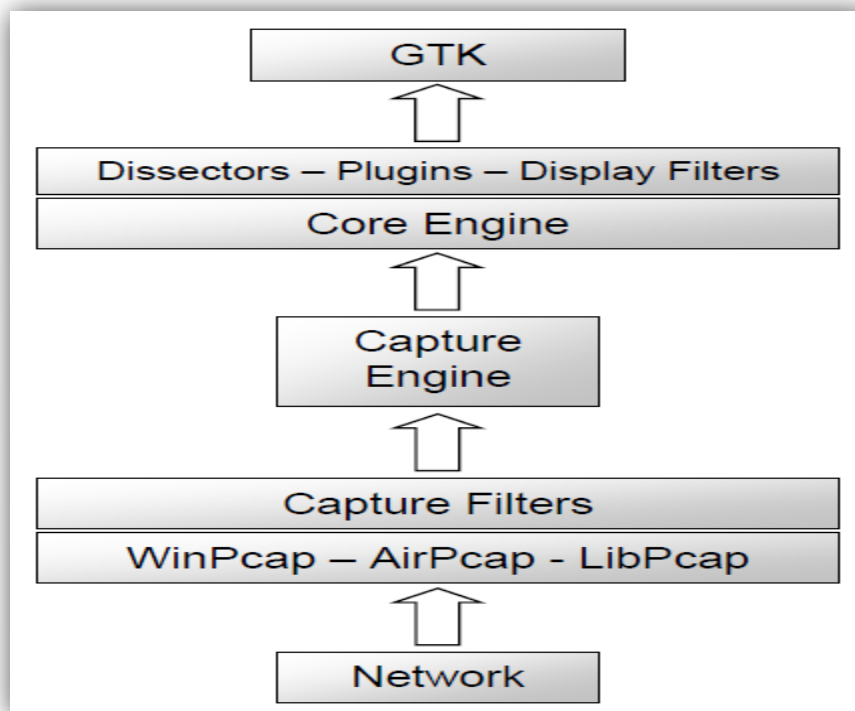
- Socket buffer size in bytes at the receiving host.
- Socket buffer size in bytes at the transmitting host.
- Message size in bytes.
- Elapsed time in seconds (length of this performance measurement).
- Effective (available) throughput between the two hosts

**Step 4:** The wireshark tool is utilized in the investigation of the user behavior -the investigates the traffic applications in an in-depth way involving different uses to web 2.0 applications and performs its distribution – these applications include facebook, YouTube, Google and others. Moreover, it facilitates the investigation into the usage pattern of protocols in applications such as, HTTP, FTP, SMTP, TELNET, and DNS and protocols present in the transport layer such as, TCP, UDP and others.

- **Wireshark**

Wireshark is a tool to analyze the network which was pioneered by Gerald Combs. The development and maintenance of Wireshark is conducted by its core developers, a group of individuals fixing bugs and providing novel functionality [71]. What it does is read packets from the network, decodes them and transforms them to an easy to read format. Wireshark is invaluable for its open source character, active maintenance, and cost-free element [68]. Wireshark is also considered as a powerful wireless security analysis tool as the user can easily go through large amounts of wireless traffic and identify security alerts in the wireless network through its filtering and protocol

decoders. These security threats include weak encryption or authentication mechanisms, and risks concerning information disclosure. Moreover, intrusion detection analysis can also be carried out through it to identify common attacks against wireless networks with the help of signal strength analysis. This identifies the station or the access point's location. Presently, wireshark operates in most UNIX platforms and different Windows platforms requiring GTK+, Glib, libpcap and other libraries for its operation [71]. The following Figure 3.10 explains the working of wireshark:



**Figure 3.10: Wireshark's works [68]**

Some other characteristics of Wireshark include: its distribution is carried through GNU's not UNIX, GPL, and it works under both promiscuous and non-promiscuous modes, it captures data from the network or it can read the data from the file captures, it has an easy to read format and can be configured to GUI, it has a veritable array of display filter capabilities, it supports



tcpdump format capturing filters, it can reconstruct a TCP session and display it in the American Standard Code for Information Interchange (ASCII), as well as in the Extended Binary Coded Decimal Interchange Code (EBCDIC), hexadecimal (hex) dump, or C arrays, it is present in pre-organized binaries and source code and it can run on more than 20 platforms with the inclusion of Uniplexed Information and Computing System (UNIX)-based operating systems (OSes) and Windows as well as third-party packages for Mac OS X. Moreover it supports more than 750 protocols and its open source character contributes to frequent new versions. It can also capture files from more than 25 varying products, save captured files in different formats and finally, it can capture data from different media forms [68].

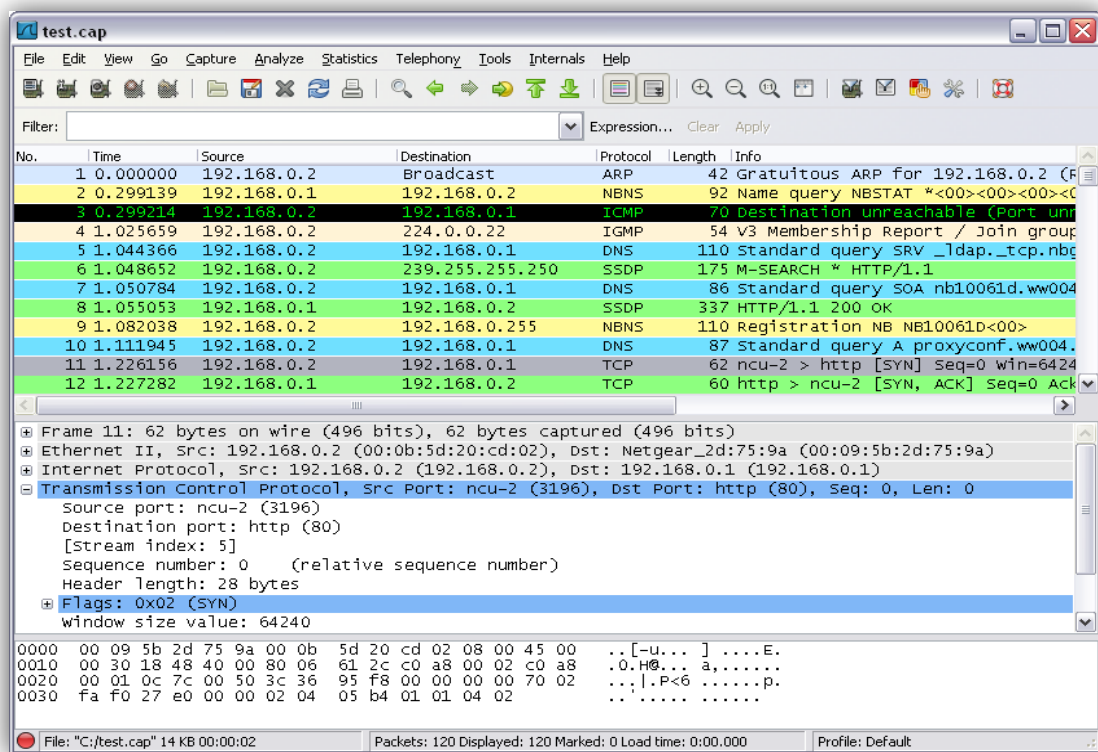
- **Wireshark's GUI and Output Format**

Wireshark possess a GUI that is user friendly and can be configured. The collected information is displayed in its three main panes [68, 71]. The capture function of wireshark looks alike for every capture. For every window, it has an adjustable application by clicking the row of dots between the panes and dragging it every which way. The highest pane is known as the summary pane and it presents a single line summary of the capture. The wireshark's GUI is presented in the figure below (Figure 3.11) and with its default fields including packet numbers, time, source address, destination address, and finally name and information regarding the highest layer protocol. Wireshark's default fields include:

- ☐ Packet number.
- ☐ Time.

- ☐ Source address.
- ☐ Destination address.
- ☐ Name and information about the highest-layer protocol.

The columns are easy to configure and sorted out.



**Figure 3.11: Wireshark's GUI [71]**

Where:

In the figure 3.11 above, shows that output format in the summary part of wireshark is:

Packet number: 11

Time: 1.226156,

Source Address: 192.168.0.2

Destination Address: 192.168.0.1

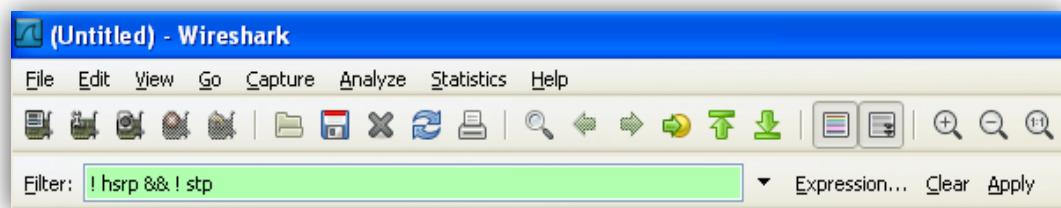
Protocol: TCP

Length: 62

Information: ncu-2 > http [SYN] seq=0 win 6424...

- **Wireshark' s Filters**

In the program, filters are made use of when the user is desirous of confining the number of packets to be captures. It assists in finding the desired packet without going through the entire packets. It can also utilize both capture and display filters. The capture filter syntax utilized in the libpcap library in tcpdump is similar to the one used her [68, 71-73]. The command line or the “Capture Filter’ dialog box is utilized to capture varying kinds of traffic. A more detailed description is presented in Figure 3.12.



**Figure 3.12: Filter bar in wireshark [71]**

There are two types of Wireshark filters:

- i Display filters: after you capture a lot of information, they help you to visualize only the packets that you are interested in, as table 3.6.
- ii Captures Filters: from the beginning you know what the interest is for you and capture only those packets.

**Table 3.6: Several of Wireshark filters [74]**

<b>FILTER</b>	<b>EXPLANATION</b>	<b>EXAMPLE</b>
<b>eth.addr</b>	source or destination mac-address	eth.addr == 00:1a:6b:ce:fc:bb
<b>eth.src</b>	source mac-address	eth.src == 00:1a:6b:ce:fc:bb
<b>eth.dst</b>	destination mac-address	eth.dst == 00:1a:6b:ce:fc:bb
<b>arp.dst.hw_mac</b>	target mac-address	arp.dst.hw_mac == 00:1a:6b:ce:fc:bb
<b>arp.dst.proto_ipv4</b>	target IPv4 address	arp.dst.proto_ipv4 == 10.10.10.10
<b>arp.src.hw_mac</b>	sender mac-address	arp.src.hw_mac == 00:1a:6b:ce:fc:bb
<b>arp.src.proto_ipv4</b>	sender IPv4 address	arp.src.proto_ipv4 == 10.10.10.10
<b>vlan.id</b>	vlan ID	vlan.id == 16
<b>ip.addr</b>	source or destination IPv4 address	ip.addr == 10.10.10.10
<b>ip.dst</b>	destination IPv4 address	ip.addr == 10.10.10.10
<b>ip.src</b>	source IPv4 address	ip.src == 10.10.10.10
<b>ip.proto</b>	IP protocol (decimal)	ip.proto == 1
<b>ipv6.addr</b>	source or destination IPv6 address	ipv6.addr == 2001::5
<b>ipv6.src</b>	source IPv6 address	ipv6.addr == 2001::5
<b>ipv6.dst</b>	destination IPv6 address	ipv6.dst == 2001::5
<b>tcp.port</b>	source or destination TCP port	tcp.port == 20
<b>tcp.dstport</b>	destination TCP port	tcp.dstport == 80
<b>tcp.srcport</b>	source TCP port	tcp.srcport == 60234

<b>udp.port</b>	source or destination UDP port	udp.port == 513
<b>udp.dstport</b>	destination UDP port	udp.dstport == 513
<b>udp.srcport</b>	source UDP port	udp.srcport == 40000
<b>fr.dlci</b>	Frame-Relay DLCI number	fr.dlci == 112
<b>wlan.sa</b>	source MAC address	wlan.sa == 00:1a:6b:ce:fc:bb
<b>wlan.da</b>	destination MAC address	wlan.da == 00:1a:6b:ce:fc:bb
<b>Dns.qry.name == Doman name</b>	Doman name	Dns.qry.name == www.google.com

For instance, in the study is an attempt to investigate how to filter Web 2.0 applications like facebook to determine the percentage of its use in UUM wireless network, the usage traffic it creates, the number of packets, the time between the packets and the average packets. Firstly, the filter (dns.qry.name) is inserted in the filter bar. Moreover, in order to determine the number of protocols and the request/response to them, the protocol is inserted in the filter.

### 3.4.2 PRESENTATION PERFORMANCE METRICS AND INTERPRETATION

Following the completion of the analysis, graphs and charts are required to present the performance metrics for visual use [75]. These graphs and charts should be clear and to the point as the interpretation of the meaning of the data presented is the ultimate goal of the performance practice [17]. In the present study, the Gnuplot tool is used to present the analysis of traffics and the performance. The throughput, load, and latency are been described in chapter four.

- **Gnuplot**

This is utilized in Linux, OS/2, MS Windows, OSX, VMS and other platforms as a portable command-line driven graphing utility. The rationale behind its creation was to enable scientists and students alike to visualize mathematical functions and data in an interactive manner. Since its creation, it has developed to support a variety of non-interactive uses including web scripting. It also helps in plotting engine by third-party applications like Octave and it explores data graphically. It has varying purposes; it generates plots and graphs from data or functions and it produces highly polished graphs suitable for publication. It is driven by command line whereby the user issues commands at a prompt and the gnuplot redraws the present plot in reaction to the command [75]. It has an interactive characteristic whereby the output is generated and displayed right after the command in the output window. Regardless of its other uses, it is mostly used as an interactive tool where the primary user interaction is made through a command language as opposed to a point-and-click GUI interface. Other reasons for choosing the Gnuplot include; its stability and active maintenance, its capability in handling huge data sets and quick reaction, its ability to read regular text files input and its tolerance for specific input file format, its support of all common graphic formats, its ability to generate polished, publication-quality graphs and to offer detailed control over the outcome of plots.

# **CHAPTER FOUR**

## **FINDINGS**

### **PERFORMANCE ANALYSIS**

#### **4.1 INTRODUCTION**

This chapter will present the results regarding to investigate network performance in UUM wireless by performing real measurement. After each tools executed like tcpstat, Netperf, wireshark, ntop, and the various options setting which have been used inside each command line in purpose to know and understand the performance of UUM wireless network, it will help us to improve the UUM wireless network.

This chapter is organized as follows: brief explanation about the wireless network measurement, performance results, user's distribution across the APs, daily traffic pattern, average packet size vs. standard deviation of the size of each packet, traffic vs. number of authenticate users, throughput vs. load, and bulk data transfer and interactive data exchange, and summary.

#### **4.2 UUM WIRELESS NETWORK MEASUREMENT**

Understanding network performance in wireless local-area networks (WLANs) is critical for those who develop, deploy, and manage WLAN technology as well as those who develop systems and application software for wireless networks. The performance of a UUM wireless network is considered as a complex issue having various independent variables impacting the clients' access to servers throughout a network. Nevertheless, majority of the elements that facilitates the performance of wireless networks can be summarized into a few simple network principles that are measureable, controlled and monitored by the network administrator through software

that is simple and often free [43]. Wireless Networking measurement can, therefore, facilitate the revelation of possible bottlenecks, lack of parameter settings, and examine the reliability of the component in a complex TCP/IP networking environments [1]. In conclusion, to investigate the UUM wireless helps staff in computer center to detect the weakness points in the UUM wireless network, moreover, to detect any application that takes most of the bandwidth, the throughput and load, the relationship between them, and interaction among nodes by sending the data between the two devices to know the latency, throughput, and so on.

### 4.3 TRAFFIC COLLECTION AND ANALYSIS

Table 4.1 summarized the high level characteristics of data captured. The researcher collected the captured data from 91 APs distributed on 9 buildings in UUM campus, one week after the mid of semester break. Trace collection started on Sunday, April 15, 2012 at 11:00 am local time. Each packet sent from and to the wireless network was mirrored to our trace gathering computer. Wireshark analyzed each packet individually and recorded information such as the date, time, origin, destination, and protocol. Trace collection stopped at the weekend on Thursday, April 19 at 12:00 pm.

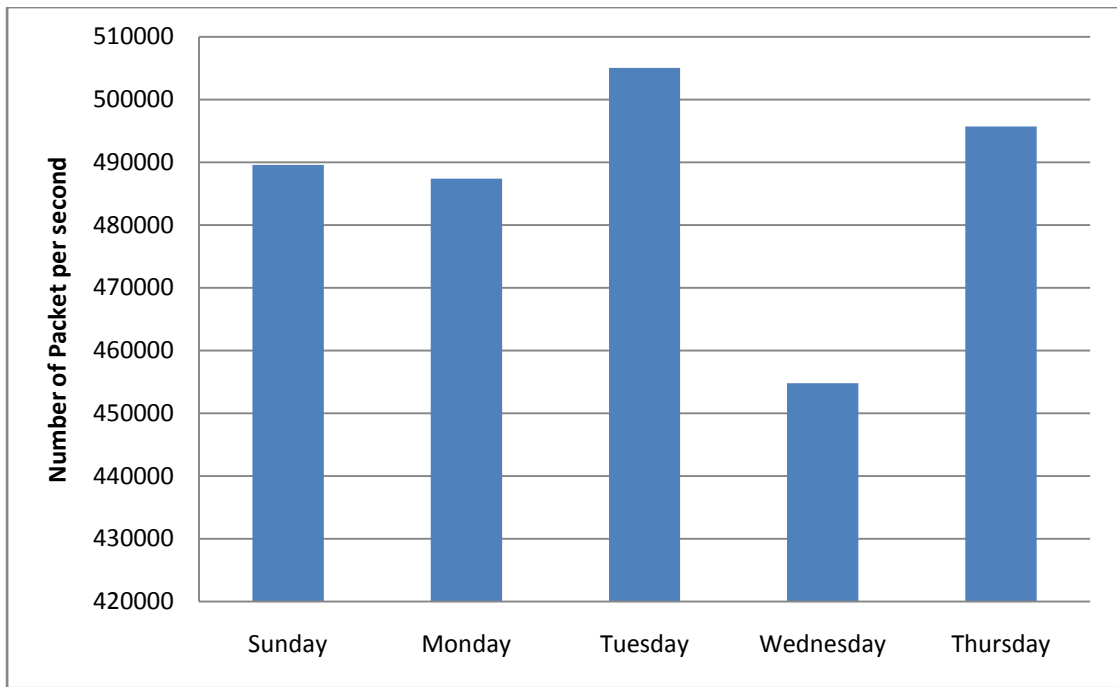
**Table 4.1: Overall statistics for the capture**

Attribute	Values
The number of APs	91 APs
Number of building	9 buildings
Peak throughput at an AP with 32768 packet size	10 <sup>6</sup> bits/sec 2.48 Mbps
Period of capture	One hour daily for one week



<b>The maximum authenticate users at an APs per second</b>	477
<b>Total of authenticate user's during the week per second</b>	2039
<b>Average of authenticate user's during the week per second</b>	407.8
<b>Average number of Packets per second of trace.</b>	486510.44
<b>Average ratio length packet during week days (ms)</b>	48.14921

The data captured consisted of aggregate packets level statistics of all traffic through the APs from nine buildings of UUM wireless network, including the information at network layer, transport layer, application layer, and application traffic. Whenever the study reach 100 bytes data captured, therefore, it receive the packets but can't it go dig deep to application traffic only at application layer protocols. So, the study increased the captured data to 1000 bytes. As a result, the study has characterized the wireless traffic in terms of the number of packets per second. Figure 4.1 shows the traffic over the entire length of the trace. The traffic level rises each day at around 11:00 am and remains high until the lunch time, and after that the traffic level is not stable.

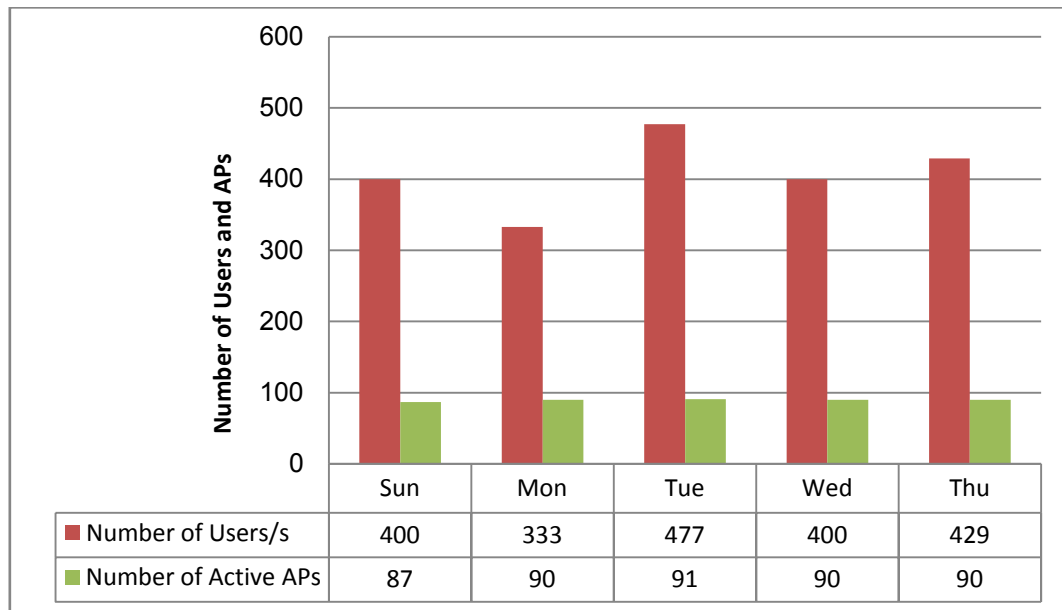


**Figure 4.1: The traffic over the entire length of the trace**

#### **4.4 USER DISTRIBUTION ACROSS THE APs**

First the study characterizes the distributions of users associated with APs. Through this investigation the study concludes that the amount of data sometime depends on the number of users. But the number of users are vary from one day to another depending on the lecture time, the opening time of the library during the week and another halls of the campus, and the type of work in this buildings of UUM wireless network. The number of active APs is also varying every day. Figure 4.2 shows the number of authorized users and number of active APs within data collection days. This investigation noticed that the number of users on Tuesday (477 users) of that week were the highest rate of users, but it observed that some of the users connected to the Internet without using it or without doing anything. Moreover, the investigation also observed the other characterizations of data and the wireless controller. This pattern reflects expected network performance were users socializing

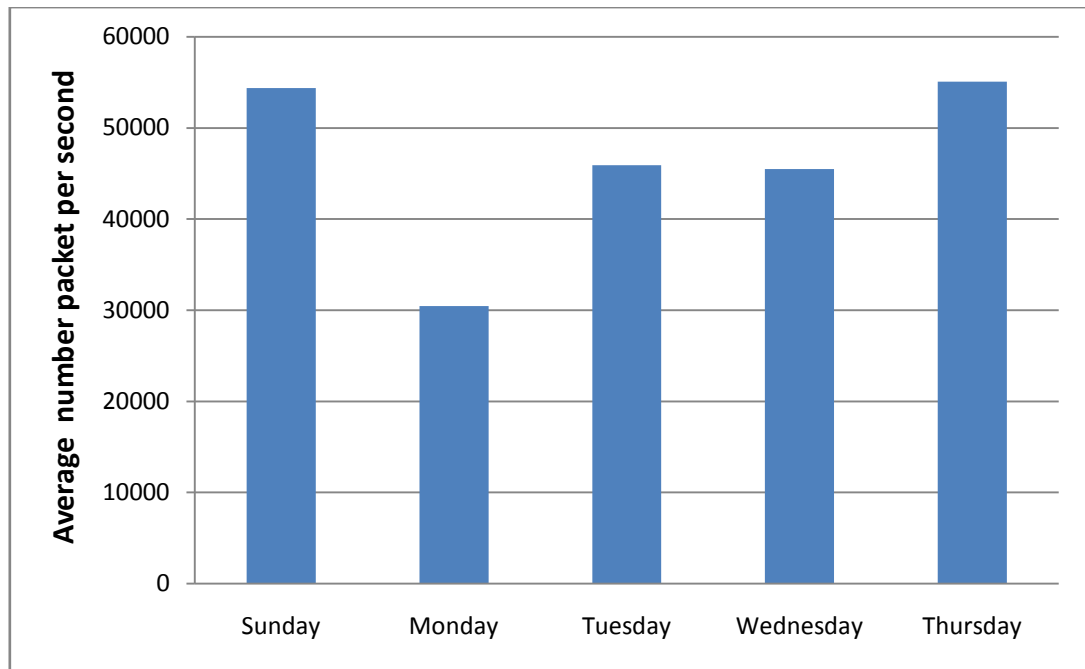
and enjoying refreshments during connected sessions. The study observed that the users' distribution follows the same pattern at all APs.



**Figure 4.2: Number of authenticated users and number of active APs**

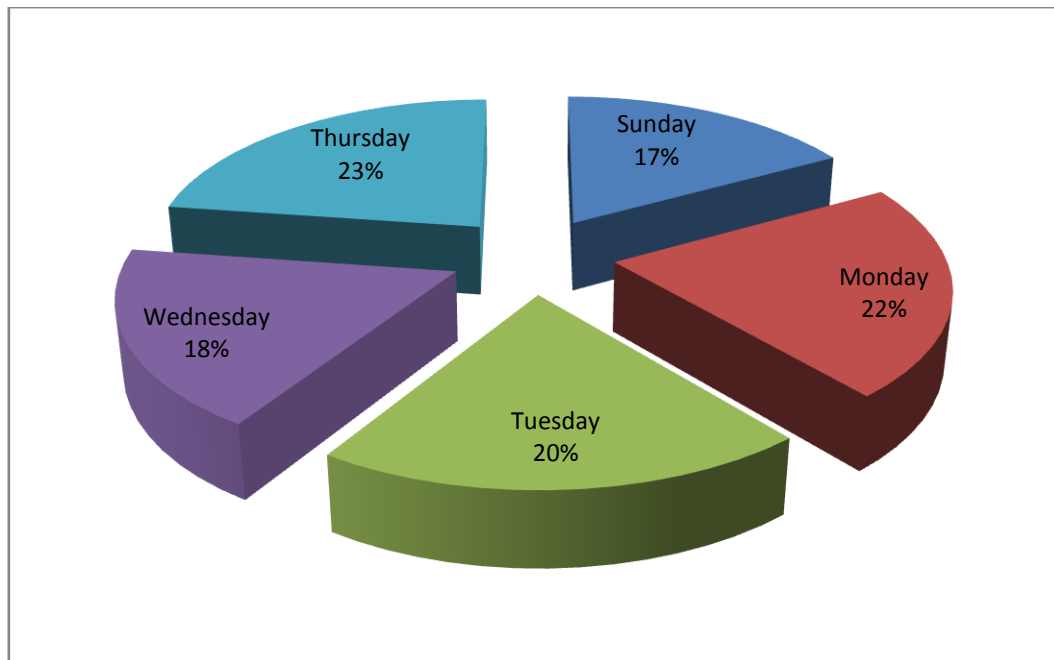
#### 4.5 DAILY TRAFFIC PATTERN

Figure 4.3 shows high traffics per second at all APs during the week. The study noticed variations from the average of packets every day adopted on the number of user with the type of data, though there are users connect to the Internet during the lecture time. Several distinctive features can be observed; the average packets was a little bit lower on Monday (333 users) and the reason behind that was the usage pattern of data and the type of data, although some of the students connect to the Internet only to check mails. The average is starting to increase again after Monday till it reaches the highest rate on Thursday. In addition, the researcher observed the number of users on Tuesday reached up to 477 users but the average packet didn't increase because of some students connect with the Internet without using it or because of the signal weakness.

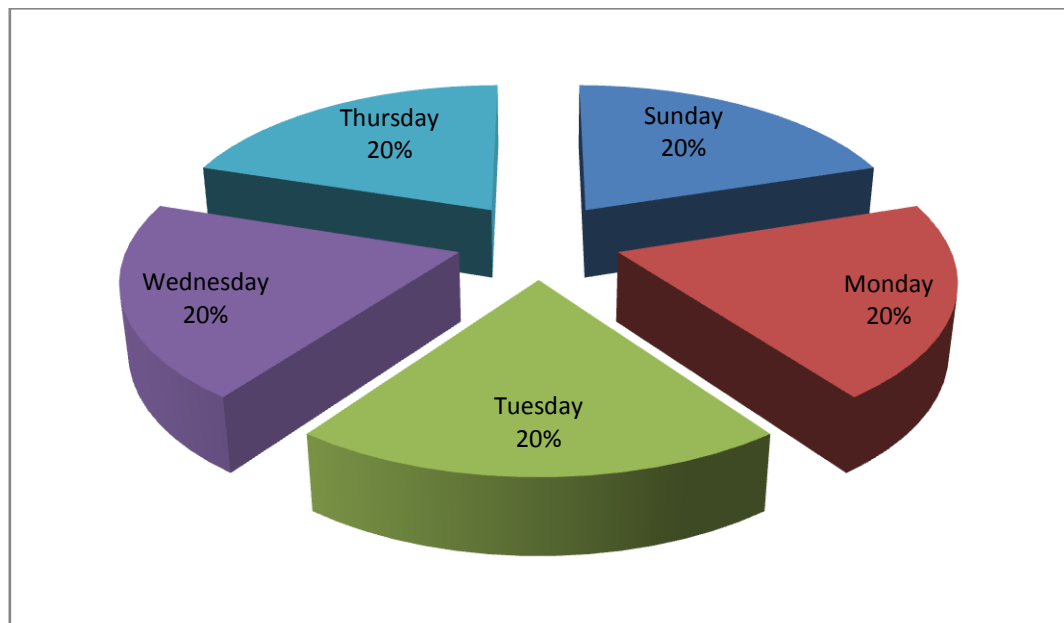


**Figure 4.3: High traffic Access Points/s**

Figure 4.4 shows the high lose of bandwidth during the week days starting of Thursday (23 %), Monday (22%), Tuesday (20%), Sunday (17%), and Wednesday (18%). This means that the bit rate and lose overall the bandwidth are increase on Thursday because the max packets size is high and the number of users is 429. The lose bandwidth on Monday is also high (22% and 333 users) because there are a lot of lectures. On Tuesday, the percent of lose bandwidth (20%) is not high although the number of users is 477, this could be because some users are not directly related to the packet rate. The difference between Thursday and other days could be attributed to the difference in the usage patterns of our students. It could also be due to poor signal strength or high levels of interference in the UUM Buildings causing repeated disconnections and reconnections. Figure 4.5 shows the low lose bandwidth during week days, where the observation is equal for all days of week.



**Figure 4.4: The high lose bandwidth during the week days**

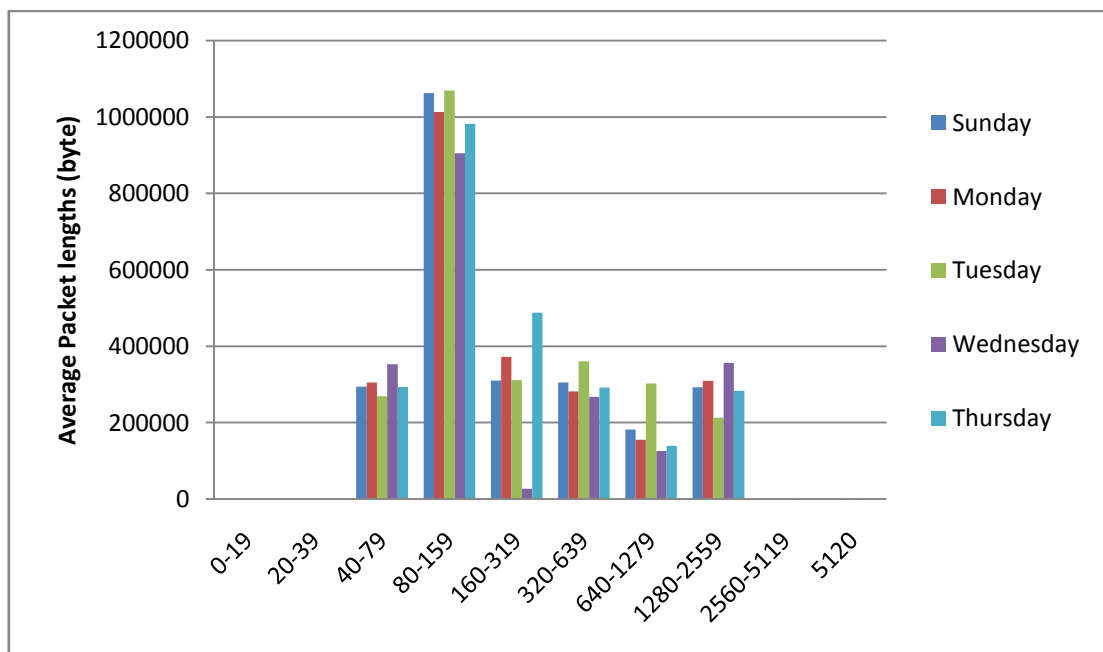


**Figure 4.5: The low lose bandwidth during the week days**

Popular public areas for socializing and studying such as DKG halls (student centre), the learning commons in the main library and a lounge outside it, and the computer labs for undergraduate and postgraduate students were used on every day of the trace. FTM, another student computing laboratory, situated in the classroom wing near large

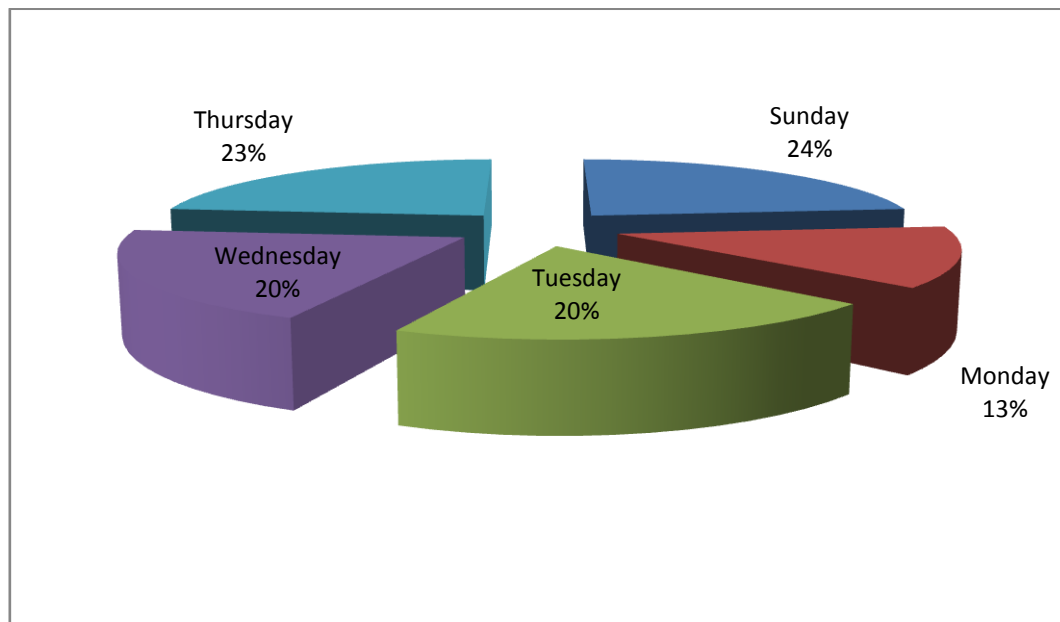
computer labs. While some other buildings appeared to be a logical place to install wireless access, but because of the lack of any non-classroom tables and chairs that allow laptop users to comfortably connect while they are not in classes, causing the minimal usage observed.

Figure 4.6 shows the packet lengths during the week days. The study observes that the length is very high (80-159) to all days of week. This means when the researcher is starting with the length 100 bytes to capture, a lot of packet the researcher will lose. The difference between lengths 80-159 and other lengths could be attributed to the difference in the usage patterns of our students. It could also be due to poor signal strength or high levels of interference in the buildings causing repeated disconnections and reconnections. Also this means that more bandwidth is consumed during the lengths 80-159 of UUM wireless network.



**Figure 4.6: The Average packet lengths during the week days**

Figure 4.7 shows the rate for the packet length. The study notices that the rate on Sunday is high. This means that the users are dealing with more lengths of packets so that the average size of the packet will increase on Sunday. This also leads to increase the load on the wireless network with increase the throughput and lose bandwidth. But the rate on Monday is low (13%) because the number of users is few, only 333 and this means the use for applications that need high lengths like the YouTube and so on also few. In additional, the study notices that the rate on Thursday is high because the number of users about 429 users and few lectures in weekend so that the students are using the web 2.0 most of the time like the YouTube , Facebook, Twitter , and Skype.

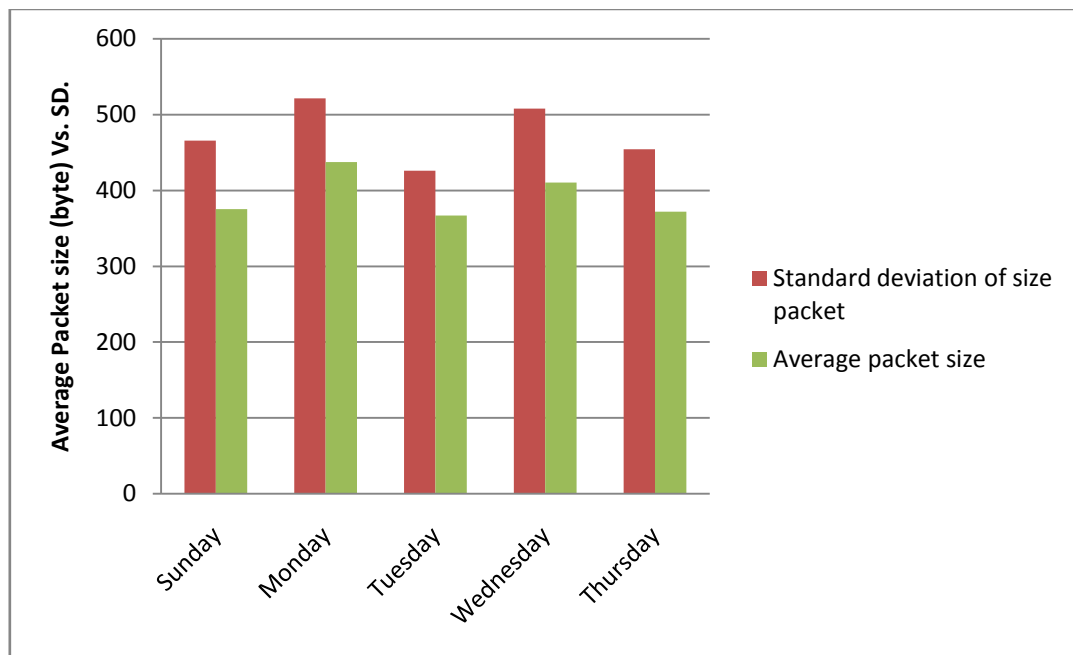


**Figure 4.7: The rates for packet length to each day of week**

But the rate on Tuesday and Wednesday is equal although of the number of user on Tuesday is 477 and number of user on Wednesday is 429. This means that some of students' connection is not directly related to the packet rate or they may connect without any interactive.

#### 4.5.1 AVERAGE PACKET SIZE VS. STANDARD DEVIATION OF THE SIZE OF EACH PACKET

Figure 4.8 compares the average packet size of every day of the week to the standard deviation size of each packet.



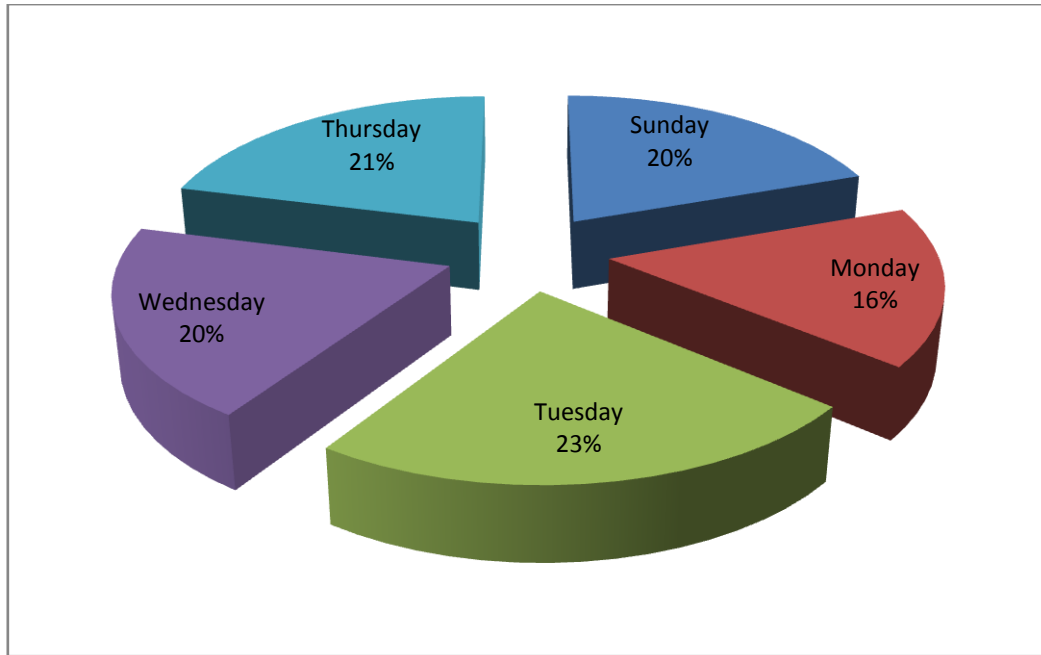
**Figure 4.8: Average packet size vs. Standard deviation of size of packet**

The Figure 4.8 above shows that the samples are near to the reality, where the researcher notes the different values of the standard deviation of the average, this depends on its size and the type of data. Wherever the standard deviation is close to the average, this denotes that the sample is closer to reality.

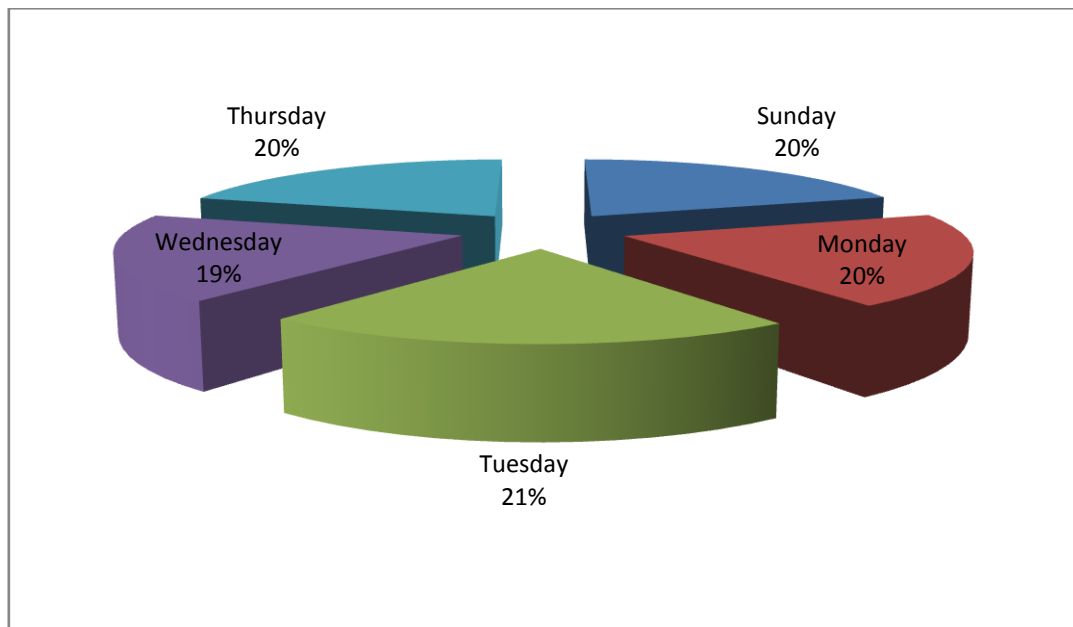
#### 4.5.2 TRAFFIC vs. NUMBER OF AUTHENTICATE USERS

Figure 4.9 compares between the number of authenticate users (a) at all access points and the total number of packets per second (b); send to and from users authenticated at all access points over the week of the captured.





**Figure 4.9: Number of authenticated users/s (a)**



**Figure 4.10: Total number of Packets/s (b)**

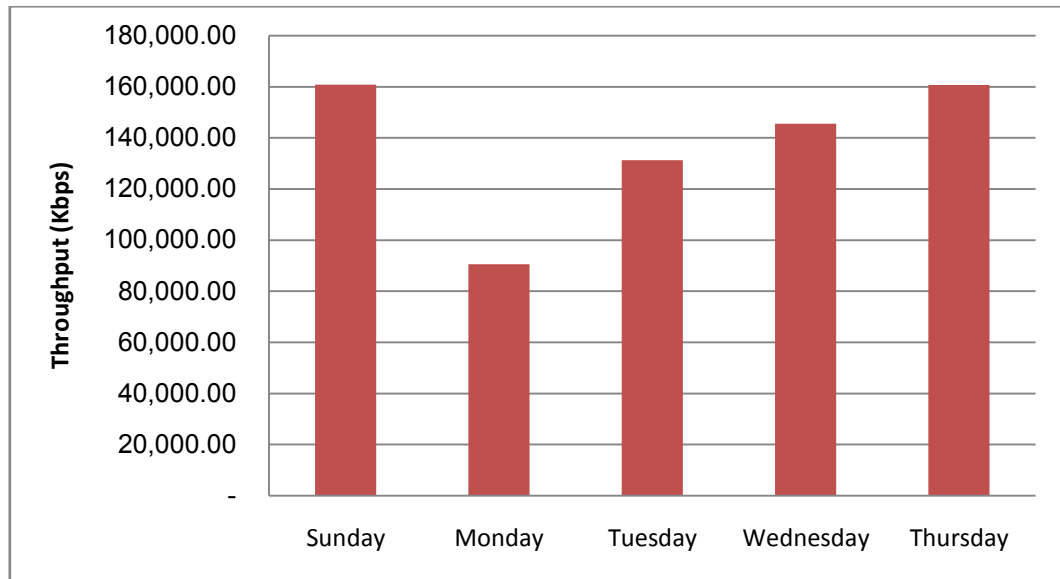
Although the access points on Tuesday generated the most authentications, the average number of packets associated with each of those authentications is lower than the number of users. This means that some of authentications users are not directly related to the packet rate. The access points on Monday generated most average

number of packets associated with each of those authentications, the number of authentications users at all access point is much lower than the number of packets. This means that the number of authentications is directly related to the packet rate and the type of data is different. The difference between Tuesday traffic and other traffic could be attributed to a difference in the usage patterns of our students. It could also be due to poor signal strength or high levels of interference in the Buildings of UUM campus causing repeated disconnections and reconnections.

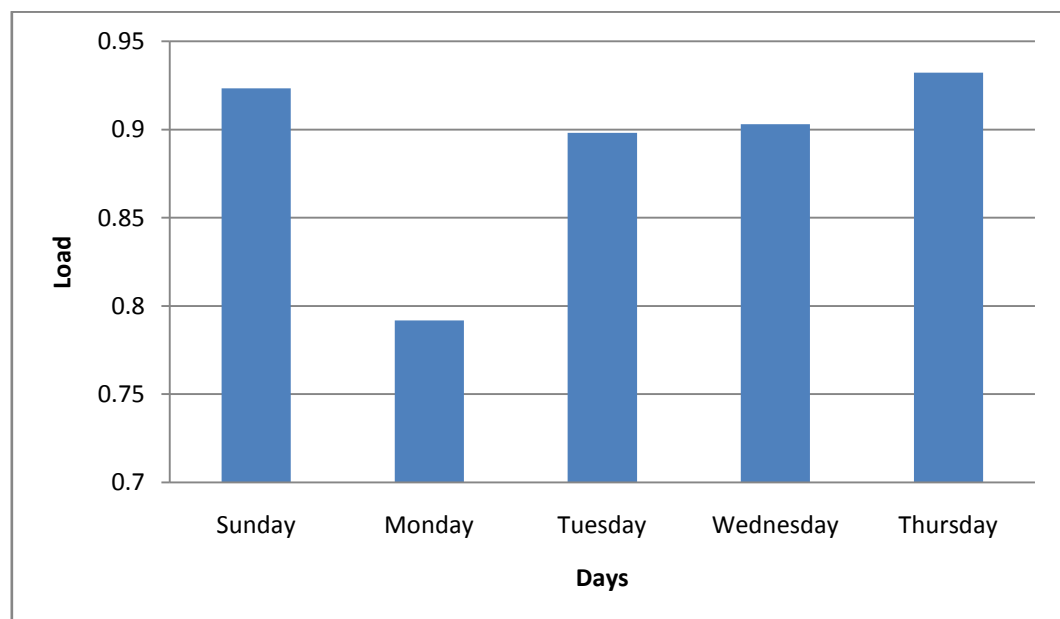
### **4.5.3 THROUGHPUT VS. LOAD**

Throughput is the total number of data which actually received from sender according to the time of sending and receiving process. The medium access control protocol should make as efficient use as possible of the wireless medium to maximize capacity. Also, it refers to how much data can be transferred from one location to another in a given amount of time. The throughput is usually measured in bits per second (bit/s or bps), and sometimes in data packets per second or data packets per time. In addition, when the throughput increases that means the load will increase. Figure 4.10 shows the comparison between the throughput and the load in UUM wireless network during one week and by take sample between 11:00 am-12:00 pm every day. As a result, the study observes that the throughput on Sunday arrives (24%). This means that the load will increase. Moreover, this indicates that there are much of data sent during Sunday because the number of users is 400, and also this is the first day of week where there are a lot of lectures and the use of Internet to make online reservation of library. But the study observes low rate on Monday (13%). This means a lot of students may use wireless connection but without doing any action and some time the connection is not directly related to the packet rate. Therefore, the load decreased during that day. Figure 4.10 shows that the loads during Tuesday and Wednesday are close, although

the number of users on Tuesday is 477 users and the number of users on Wednesday is 400. This means that there are number of users are not directly related to packet rate, or because of the weakness of the signals when the user is connecting to the Internet. Meanwhile, the connection is recorded in the wireless controller as current user but without any affect.



**Figure 4.11: Throughput (kbps) (a).**



**Figure 4.12: Load (b)**

**Table 4.2: The performance results per second to all week days**

<b>Days Parameter</b>	<b>Sunday</b>	<b>Monday</b>	<b>Tuesday</b>	<b>Wednesday</b>	<b>Thursday</b>
<b>Network Load</b>	0.923	0.791	0.898	0.903	0.932
<b>Max. packet size (Bytes)</b>	7656.44 bytes	9846.37 bytes	9097.27 bytes	8199.20 bytes	10242.66 bytes
<b>Min. packet size (Bytes)</b>	46 bytes	46 bytes	46 bytes	46 bytes	46 bytes
<b>Number of bit per second (bps)</b>	164722531.4 0 bits/s	92689062.3 0 bits/s	134434954. 80 bits/s	149085336.3 0 bits/s	164558154.30 bits/s
<b>Throughput (kbps)</b>	160861.85	90516.66	131284.14	145591.15	160701.32
<b>Standard deviation of size of each packet</b>	465.903	521.523	426.181	507.983	454.523
<b>Number of packet per second</b>	54396.911	30462.312	45915.018	45478.140	55081.822
<b>Number of Active Access Points (APs)</b>	87 APs	90 APs	91 APs	90 APs	90 APs
<b>Average number of authentication users/s</b>	400 users	333 users	477 users	400 users	429 users

#### **4.6 BULK DATA TRANSFER AND INTERACTIVE DATA EXCHANGE**

The bulk data transfer model is utilized in the measurement of the bandwidth or the effective throughput that the wireless network has between the transmitter and the receiving hosts. In the UUM wireless network, the netperf tool is used by the present study to measure the bulk data transfer between two computers. The researchers chose the library for the measurement of the bulk data as it presents the place where students congregate and it is the considered scope of the study. For the measurement of bulk data transfer and interactive data, the server with IP Address 10.6.3.54 is placed in level five of the library and the client's IP Address is placed as 10.19.82.180.161, 10, 19, 82, 187 and 10, 19, 69, 176 in the level four of the library. The study then initiated the measurement of the bulk data transfer and the interactive data existing between them. While carrying out this process, the researcher is faced by many challenges including system interruption, between the two and poor signals of the Internet in some areas of the library. Thus, either TCP or UDP can be utilized for this purpose. The examination assists in knowing the performance between the two notes for comparison and to provide knowledge regarding the bandwidth amount, the effective throughput and the latency in the wireless network of UUM. For the purpose of the present study, the researcher is required to employ two programs on the server and client or the netperf and netserver. Therefore, the researcher enabled data flow from the server to the client and vice versa. The tables (Table 4.3, 4.4, 4.5, 4.6, 4.7, and 4.8) present the performance measurement outcome.

#### 4.6.1 TCP\_STREAM

High-bandwidth TCP streaming was tested from client to server and the results are measured in Mbit/sec.

```
#Netperf -t TCP_STREAM -H 10.19.69.176 -- -m 1024
```

Where:

-m: Set the size of the buffer passed in- to the "send" calls of a\_STREAM test.

-t: testname like (TCP\_STREAM, UDP\_STREAM, TCP\_RR, and so on).

**Table 4.3: TCP\_STREAM**

<b>Recv Socket Size (Bytes)</b>	<b>Send Socket Size (Bytes)</b>	<b>Send Message Size (Bytes)</b>	<b>Elapsed (latency) Time (Secs.)</b>	<b>Throughput 10^6bits/sec</b>
<b>87380</b>	16384	1024	10.44	0.12

The research found the default receives socket buffer size for the purpose of the receiver (Ubuntu 11.10) is equal to 87380 bytes and for the sender (Ubuntu 11.4 kernel) is equal to 16384 bytes. In addition, throughput (0.12) is presented as 10^6 bits per second, with the test ran (latency) at 10.44 seconds. Moreover, IPv4 addresses (AF\_INET) were used. The study found that the default TCP\_STREAM is 16384 bytes and the data transfer was noted from the system running netperf to the system running netserver. The time consumed in establishing the connection is not included in the throughput calculation but the time spent flushing the final data to the remote side of the test is included.

### 4.6.2 UDP\_STREAM

High-bandwidth UDP streaming test from client to server.

```
# Netperf -t UDP_STREAM -H 10.19.69.176 -- -m 1024
```

Where:

-m: Set the size of the buffer passed in- to the "send" calls of a\_STREAM test.

-t: testname like (TCP\_STREAM, UDP\_STREAM, TCP\_RR, and so on).

**Table 4.4: UDP\_STREAM**

Socket Size (Bytes)	Message Size (Bytes)	Elapsed (latency) Time (Secs.)	Messages Okay #	Errors #	Throughput 10^6bits/sec
114688	1024	10.00	4787 3053	0	3.92
114688		10.00			2.50

The initial line of numbers is the statistical data from the netperf side while the second line of numbers comes from the netserver side. It was revealed that 4787 minus 3053 messages or 1734 messages failed to flow all the way to the netserver process.

### 4.6.3 TCP\_RR

A TCP\_RR (TCP Request/Response) is considered as a test carried out by passing the value of "TCP\_RR" to the global -t command-line option and the rate of transaction is considered as the number of completed transactions that are exchanged over the length of time consumed in performing the transactions.

```
# Netperf -t TCP_RR -H 10.19.180.161
```

**Table 4.5: TCP\_Request/Response**

<b>Local Socket Send (Bytes)</b>	<b>Remote Size Recv (Bytes)</b>	<b>Request Size (Bytes)</b>	<b>Resp. Size (Bytes)</b>	<b>Elapsed (Latency) Time (Secs.)</b>	<b>Trans. Rate Per Secs.</b>
<b>16384</b>	87380	1	1	10.00	62.60
<b>16384</b>	87380				

In this instance, both request and response sizes were one byte, the socket buffers were left as defaults and the test duration was noted at 10.00 seconds. The rate of transaction per second was not so good and as such, the UUM wireless requires additional APs in its library.

#### **4.6.4 TCP\_CC**

A TCP\_CC (TCP Connect/Close) test is a test conducted by passing a value of “TCP\_CC” to the global -t option and it is a test that measures the extent that the pair of systems can open and close connections quickly between the client and the server simultaneously. Although this is referred to as an RR (Request/Response) test, request or response is not exchanged over the connection.

```
# Netperf -t TCP_CC -H 10.19.180.161
```



**Table 4.7: TCP\_Connect/Close**

<b>Local Socket Send (Bytes)</b>	<b>Remote Size Recv (Bytes)</b>	<b>Request Size (Bytes)</b>	<b>Resp. Size (Bytes)</b>	<b>Elapsed (Latency) Time (Secs.)</b>	<b>Trans. Rate Per Secs.</b>
<b>16384</b>	87380	1	1	10.00	48.50
<b>16384</b>	87380				

The TIME\_WAIT reuse issue is a significant one for a test of TCP\_CC. Primarily, the TIME\_WAIT reuse occurs when a pair of systems churn through connections in a quick way that they end up wrapping the 16-bit port number space in a time less than the length of the TIME\_WAIT state. Theoretically, it is possible to reuse a connection in TIME\_WAIT, in rare and possible conditions. Attempt at reusing a connection in TIME\_WAIT can lead to a significant connection delay. Generally, any time the connection churn rate approaches is represented by:

$$\text{Sizeof(clientportspace)} / \text{Lengthof(TIME\_WAIT)}$$

Moreover, there is a risk connected with TIME\_WAIT reuse and as such to decrease the chances of this happening, netperf by default chooses its own client port numbers from 5000 to 65535. Systems with a 60 second TIME\_WAIT state, should allow approximately 1000 transactions per second. Also, the client port space size utilized by netperf can be manipulated through the test-specific -p option which adopts a sizespec as its value setting as its first value minimum and second value maximum port numbers utilized by netperf at the client's end.

#### 4.6.5 TCP\_CRR

The TCP Connect/Request/Response (TCP\_CRR) test is conducted by passing a value of “TCP\_CRR” to the global -t command-line option. This type of test is akin to a merger between TCP\_RR and TCP\_CC test measuring the performance of connection establishment, and exchanging one request/response transaction and cutting the connection. This is similar to what occurs in an HTTP 1.0 or HTTP 1.1 connection as opposed to using HTTP Keepalives.

```
# Netperf -t TCP_CRR -H 10.19.180.161
```

**Table 4.7: TCP\_Connect/Request/Response**

<b>Local Socket Send (Bytes)</b>	<b>Remote Size Recv (Bytes)</b>	<b>Request Size (Bytes)</b>	<b>Resp. Size (Bytes)</b>	<b>Elapsed (Latency) Time (Secs)</b>	<b>Trans. Rate Per Secs.</b>
<b>16384</b>	87380	1	1	10.00	4.20
<b>16384</b>	87380				

The TIME\_WAIT reuse issue occurs in the TCP\_CRR test like it does in the TCP\_CC test. Also, as the establishment of connection and tear-down is not symmetric, a TCP\_CRR test is asymmetrical even with similar request and response sizes.

#### 4.6.6 UDP-RR

A UDP Request/Response (UDP\_RR) test is done through passing a value of “UDP\_RR” to a global -t option and is used akin go to a TCP\_RR test with the exception that in the UDP\_RR, UDP is utilized as opposed to a TCP. UDP does not facilitate retransmission of lost UDP datagram, and netperf does not add anything to it implying that if \_any\_ request or response is lost, the request and response exchanges will halt at that point until the expiry of test timer. Netperf is not aware of this happening and the only sign provided is the low rate of transaction per second.

```
#Netperf -t UDP_RR -H 10.19.82.187
```

**Table 4.8: UDP\_Request/Response**

<b>Local Socket Send (Bytes)</b>	<b>Remote Size Recv (Bytes)</b>	<b>Request Size (Bytes)</b>	<b>Resp. Size (Bytes)</b>	<b>Elapsed (Latency) Time (Secs.)</b>	<b>Trans. Rate Per Secs.</b>
<b>114688</b>	114688	1	1	10.00	0.40
<b>114688</b>	114688				

## 4.7 SUMMARY

The study summarizes the conclusions of this chapter by giving a high-level characterization of a wireless in a UUM setting. The observation of this study is that the users distributed across all APs and user arrivals are correlated in various times and space. This is a direct consequence of the UUM wireless setting. The number of authenticated users sometimes depends on the time of lectures. On Tuesday, for example, the researcher observed that the number of users is reaching too high level. This means that most of the students do not have lectures in that day, which means that most of them are connecting to the Internet. Also the numbers of users are sometimes depends on the comfortable place and the number of chairs because some of buildings have APs but they do not have the best place to connect to the Internet. The traffic over the entire length of the trace is somehow good although the traffic is decreased during the Wednesday. The number of active APs is not stable during the week per second. This means that some of users will lose their Internet connection if the AP stops transmission. The rate of losing bandwidth depends on the number of traffic and the packet length. For example, the length of packet (80-159) consumes most of the bandwidth in UUM wirelesses. The low lose of the bandwidth is the same in all days per second. The bandwidth distribution across APs is highly uneven and does not directly correlate to the number of users at APs. Rather, the load at APs is determined more by individual user workload behavior. The rate for packet length reaches high level on Sunday although the loss bandwidth is low. This means that the students deal with a lot of usage pattern. Not surprisingly, the offered load on the network directly correlates with increase throughput. It is highest during the Sunday and lowest during the Monday. But some time the increase of the load does not affect on the throughput. The positive dealing with the wireless through sending the data

between two nodes to measure the latency, throughput, and transactions are not suitable in UDP-STREAM because the researcher sent 4787 messages and received 3053 messages. This means that 1734 messages did not make it at all the way to the remote netserver process; moreover, the latency time is not good. The TCP-STREAM is consider adequate when the researcher send message size 1024 bytes, where was latency 10.44 sec and throughput is 0.12 ( $10^6$ bits/sec). The transaction to TCP Request/Response is not suitable because the rate is 62.60 sec to send one byte. The transaction to TCP Connect/Close is not good. The UDP Request/Response is not good because the transaction is 0.40 sec and latency is 10.00 to response one byte. Finally, some buildings in UUM need to add more APs, like the library, FTM, and EDC. Also, some buildings need to provide comfortable places for students and add more chairs in APs places.

## **CHAPTER FIVE**

### **FINDINGS**

#### **USERS BEHAVIOR ANALYSIS**

##### **5.1 INTRODUCTION**

This chapter will present the results regarding to study users' behavior by analysis the Internet traffic in UUM wireless network. After each tools executed like tcpstat, Netperf, Wireshark, ntop, and the various options setting which have been used inside each command line in purpose to know and understand the user's behavior of UUM wireless network, it will help us to improve the UUM wireless network. This chapter is organized as follows: users' behavior in UUM wireless, user's application popularity, application traffic measurement, and summary.

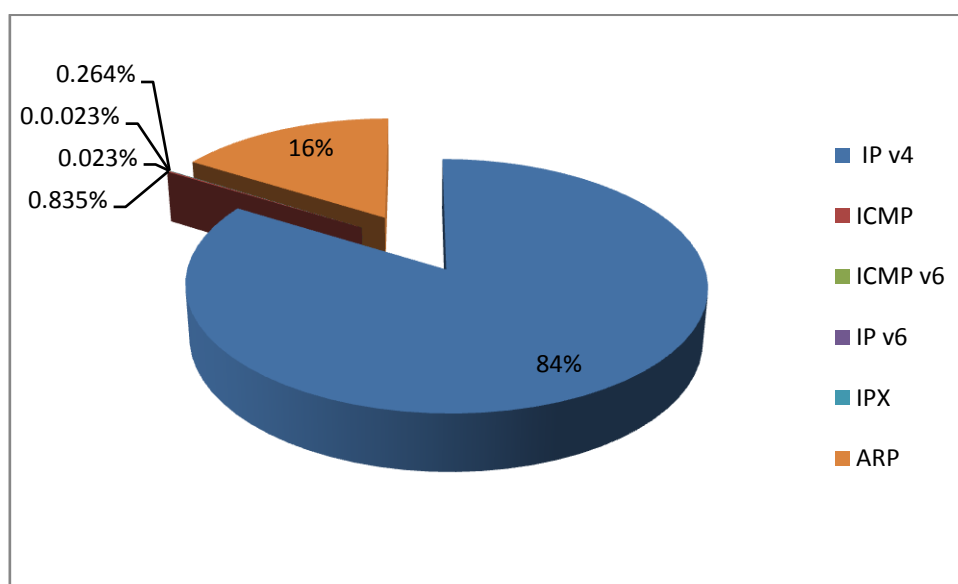
##### **5.2 USER'S BEHAVIOR IN UUM WIRELESS**

In this part, the study analyzes the characteristics of user behavior in UUM wireless network. Because the data was huge, therefore, a sample was selected to represent the population of the study. For clarity, the results were presented for just one hour of each day of 11:00 am to 12:00 pm which is considered the peak hour. Traffic collection started on Sunday, April 15, 2012 at 11:00 am. Each packet sent from and to the wireless network was mirrored to our traffic gathering computer. Wireshark analyzed each packet individually and recorded information such as the date, time, origin, destination, and protocol. Traffic collection stopped at the weekend on Thursday, April 19 at 12:00 pm. The data captured of 91 APs are distributed on 9 buildings of UUM campus (FTM, PK, DKG, Pusat Konvensyen, DPP, EDC, FPAU,

Sultanah Bahiyah Library, and DPPYAB). The data captured consists of the general statistics of all traffic through access points in the nine buildings of UUM wireless network. The data include the information at network layer, transport layer, and application layer. In this chapter we focus on the user's behavior by go dig deep to applications traffic like web 2.0 applications. In addition, to network layer, transport layer, and application layer Protocols. Investigate the applications traffic is very important in the network world because this study help us to know which applications traffic consume overall the bandwidth, and what are the user behavior in UUM wireless? As have been illustrated it in the chapter 4 paragraph 4.3 in relation to users distributions across APs.

### 5.3 NETWORK LAYER TRAFFICS

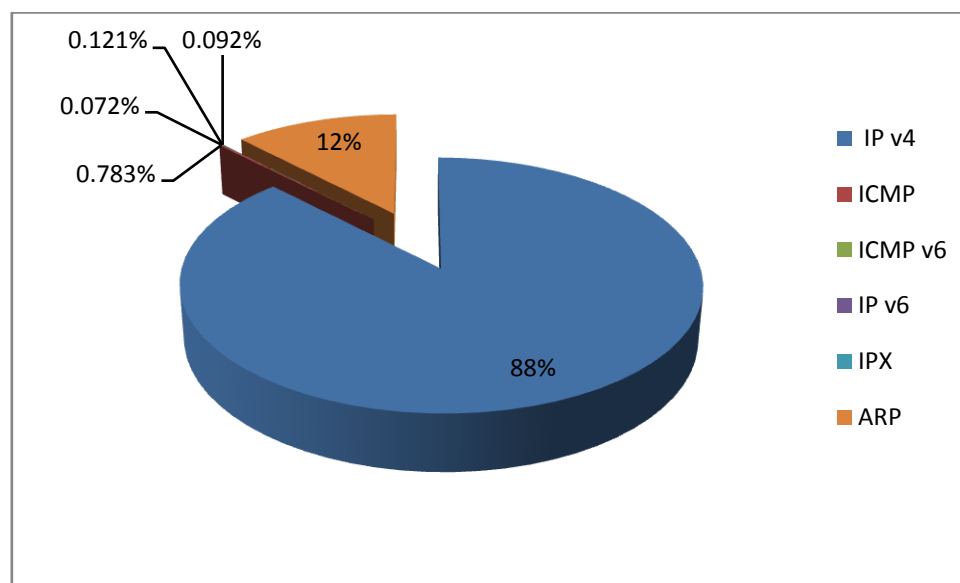
This section reports the analysis of the packet header traffic. Figures 5.1, 5.2, 5.3, 5.4, and 5.5 shows the classification of user traffic on Sunday, Monday, Tuesday, Wednesday, and Thursday by network layer traffics. Figure 5.1 shows the users' traffic on Sunday.



**Figure 5.1: The classification of user traffic by network layer on Sunday**

Figure 5.1 shows the network layer protocols responsible for the traffic on Sunday. It is cleared that most traffic is generated by IP v4 (84%). The remaining 16% of the traffic is accounted by ARP. Although some IPv6, IPX, ICMP, and ICMPv6 traffics are shown, but they accounted for less than 1% of the total traffic transferred.

As a result, the study observes that most bandwidth is consumed on Sunday with high traffic of IPv4, then ARP protocol. Figure 5.2 shows the classification of user traffic by network layer traffics on Monday.

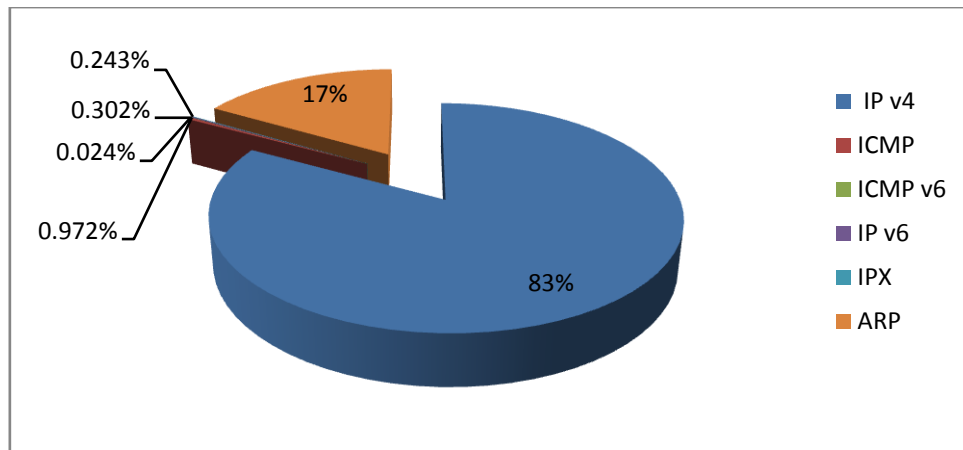


**Figure 5.2: The classification of user traffic by network layer on Monday**

Figure 5.2 shows the network layer protocols responsible for the traffic on Monday. It is cleared that most traffic is generated by IP v4 (88%). The remaining 12% of the traffic is accounted by ARP protocol. Although some IPv6, IPX, ICMP, and ICMPv6 traffics were observed, but it accounted for less than 1% of the total traffic transferred.

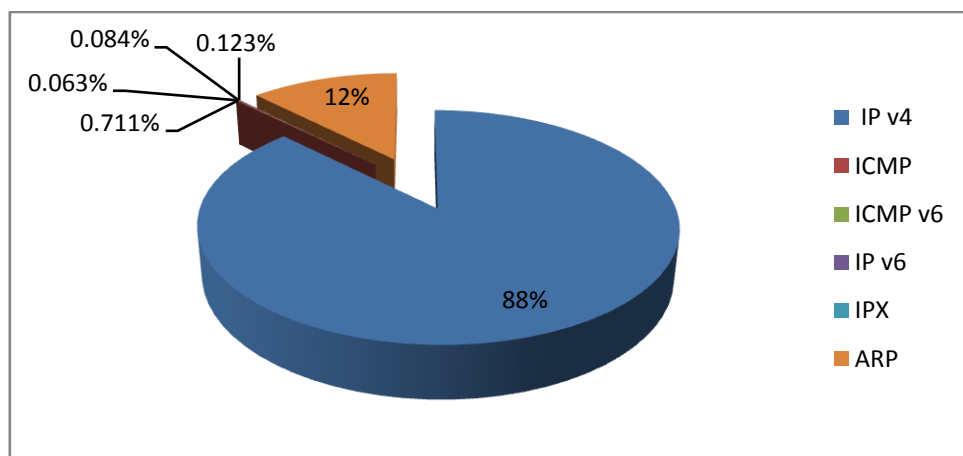
As a result, the study observes that most bandwidth is consumed on Monday by IPv4 traffics. This means that most users are directly related to packet rate traffics. Wherefore, the study noted that in section 4.4 of 4 the percent of lose bandwidth was high (22%) of another days. Figure 5.3 shows the classification of user traffic by network layer traffic on Tuesday.





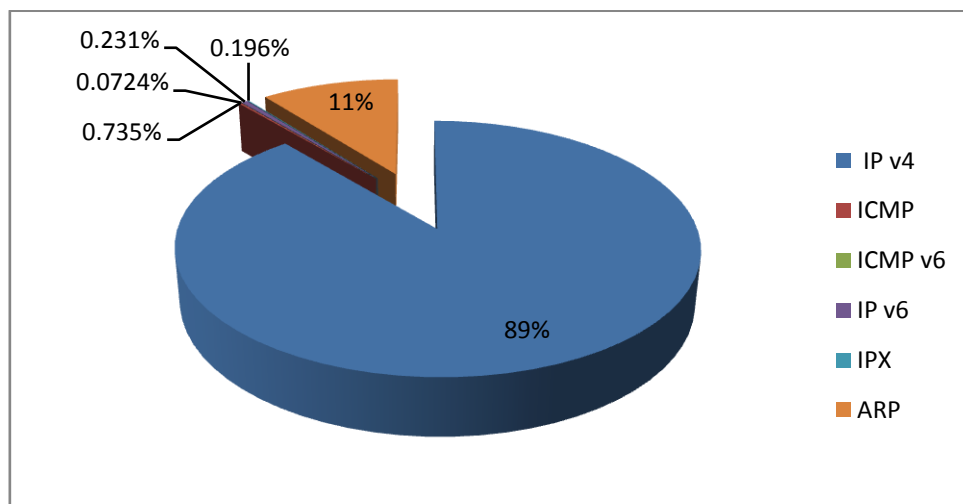
**Figure 5.3: The classification of user traffic by network layer on Tuesday**

Figure 5.3 shows the network layer protocols responsible for the traffic on Tuesday. The study sees that most traffic is generated by IP v4 (83%), 17% of the traffic is created by ARP. The study notes that the ARP traffic is increased 1% comparing with Sunday and 5% comparing with Monday. In addition, it is observed that the IP v4 traffic is decreased 1% comparing with Sunday and 5% comparing with Monday. Also some IPv6, IPX, ICMP, and ICMPv6 traffics were observed, they accounted for less than 1% of the total traffic transferred. As a result, it is observed that most bandwidth consumes by IP v4 traffic, then ARP traffic. In addition, the number of authenticated users on Tuesday is high (477 users) per second. Figure 5.4 shows the classification of user traffic by network layer on Wednesday.



**Figure 4.4: The classification of user traffic by network layer on Wednesday**

Figure 5.4 shows the network layer protocols responsible for the traffic on Wednesday. The study sees that most traffic is generated by IP v4 (88%), 12% of the traffic is created by ARP. The study notes that the ARP traffic is decreased 5% comparing with Tuesday. In addition, it is observed that the IP v4 traffic is increased to 5% comparing with Tuesday. Also some IPv6, IPX, ICMP, and ICMPv6 traffics were observed, they accounted for less than 1% of the total traffic transferred. As result, the study observes that the traffic that creates by IP v4 is still consumes most bandwidth. Figure 5.5 shows the classification of user traffic by network layer on Thursday.



**Figure 5.5: The classification of user traffic by network layer on Thursday**

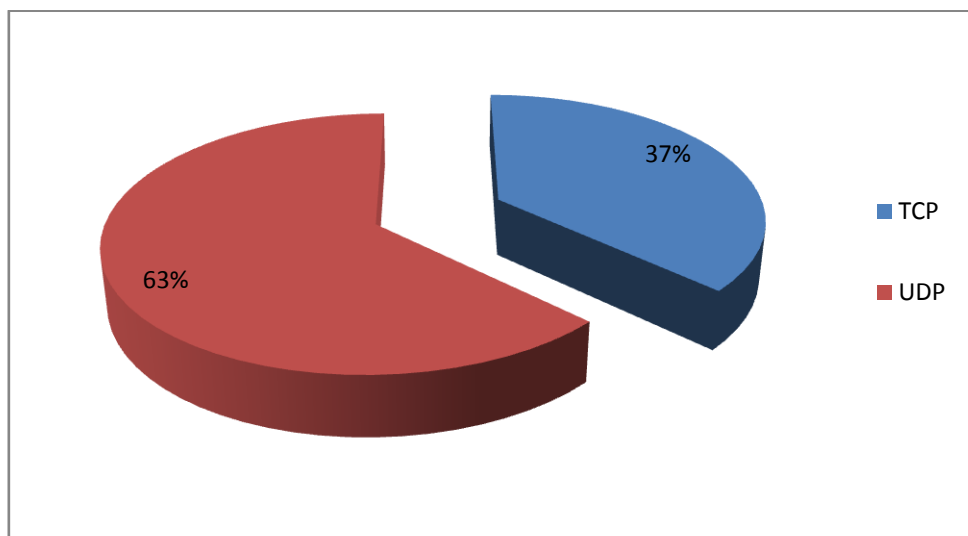
Figure 5.5 shows the network layer protocols responsible for the traffic on Thursday. The study sees that most traffic is generated by IP v4 (89%), 11% of the traffic is created by ARP. The study notes that the IP v4 traffic is increased 6% comparing with the Tuesday and 1% comparing with Wednesday. In addition, it is observed that the ARP traffic is decreased to 6% comparing with Tuesday and 1% comparing with Wednesday. Also some IPv6, IPX, ICMP, and ICMPv6 traffics were observed, they accounted for less than 1% of the total traffic transferred. In section 4.4 of chapter 4, it is noted that the percent of lose bandwidth was highest (23%) on Thursday, and the

number of users is 429 with high related to packet rate. This means that most traffic of this percent is by IP V4 then by ARP.

In conclusion, the IP v4 traffic is between 83% - 89% through days of the week. The ARP traffic is between 11% - 17% through the days of week. While IPv6, IPX, ICMP, and ICMPv6 traffics are less of 1% through all the days of week.

#### 5.4 TRANSPORT LAYER TRAFFICS

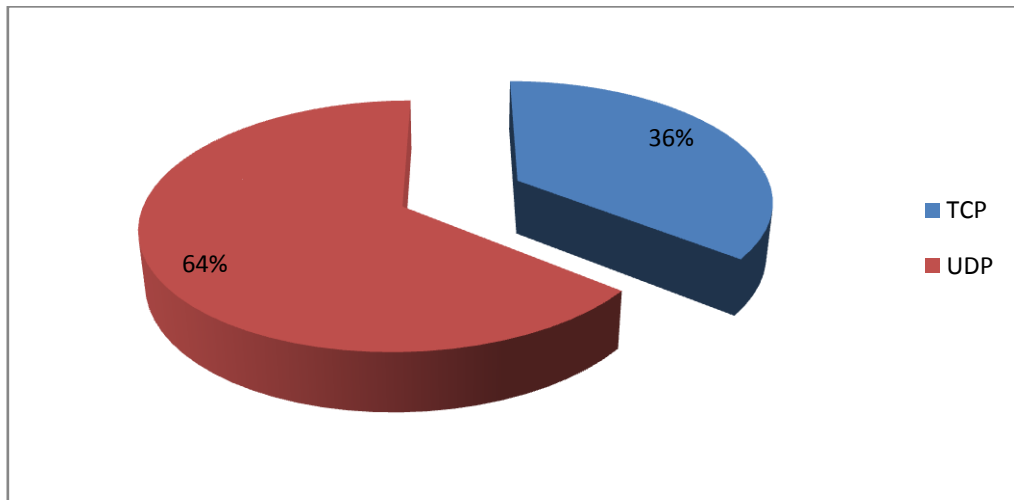
This section reports the analysis of the packet header traffic. Figures 5.6, 5.7, 5.8, 5.9, and 5.10 shows the classification of user traffic on Sunday, Monday, Tuesday, Wednesday, and Thursday by transport layer protocols. Figure 5.6 shows the users' traffic by the transport layer protocols on Sunday.



**Figure 5.6: The classification of user traffic by transport layer on Sunday**

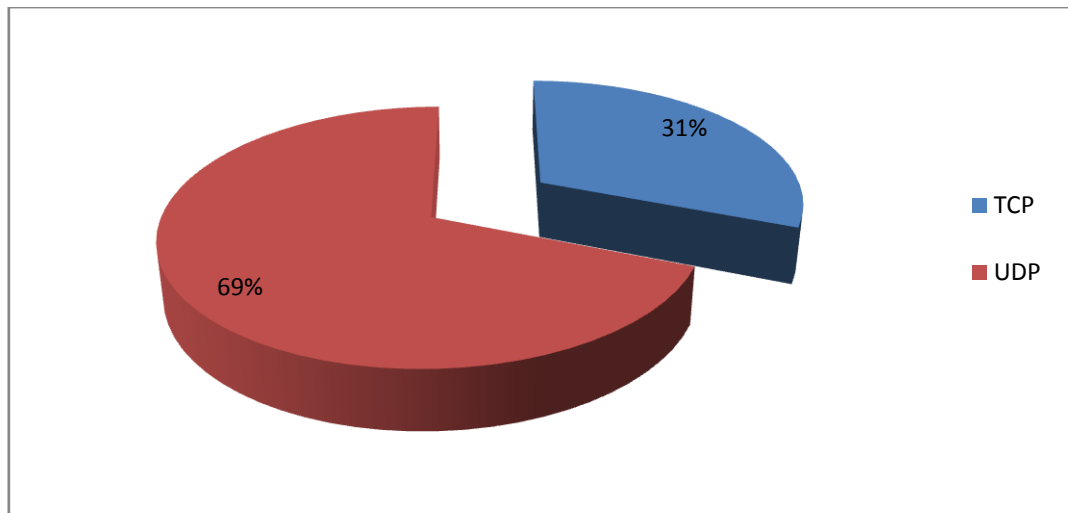
Figure 5.6 shows the top protocols responsible for the traffic on Sunday. It is cleared that most traffic is generated by UDP (63%); this means that a lot of bandwidth is consumed on Sunday by UDP. The remaining 37% of the traffic is accounted by TCP; although it mostly deals with HTTP, SMTP, FTP, and Telnet but the percent by UDP is high. As a result, the study observes that most of users are dealing with video and

audio streams. This means that most users are directly related to UDP traffics. Figure 5.7 shows the classification of user traffic by transport layer protocols on Monday.



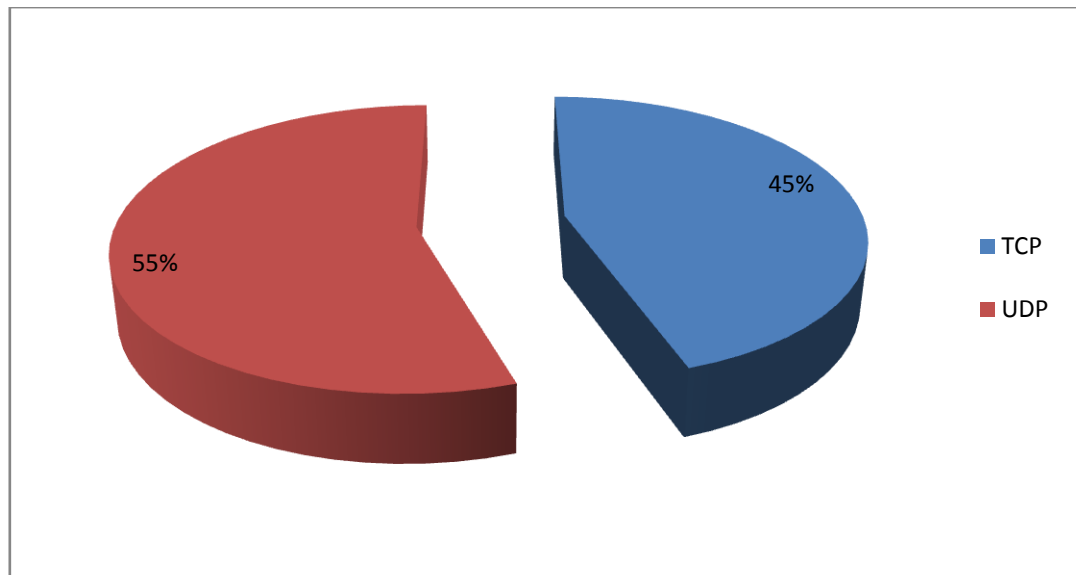
**Figure 5.7: The classification of user traffic by transport layer on Monday**

Figure 5.7 shows the top protocols responsible for the traffic on Monday. It is cleared that most traffic is generated by UDP (64%); this means that a lot of bandwidth is consumed on Monday by UDP. The remaining 36% of the traffic is accounted by TCP. As a result, the study observes that most of users are dealing with video and audio streams. This means that most users are directly related to UDP traffics. Although, the average number of authentication users on Monday is not more 333 users per second comparing with another days of week. Figure 5.8 shows the classification of user traffic by transport layer protocols on Tuesday.



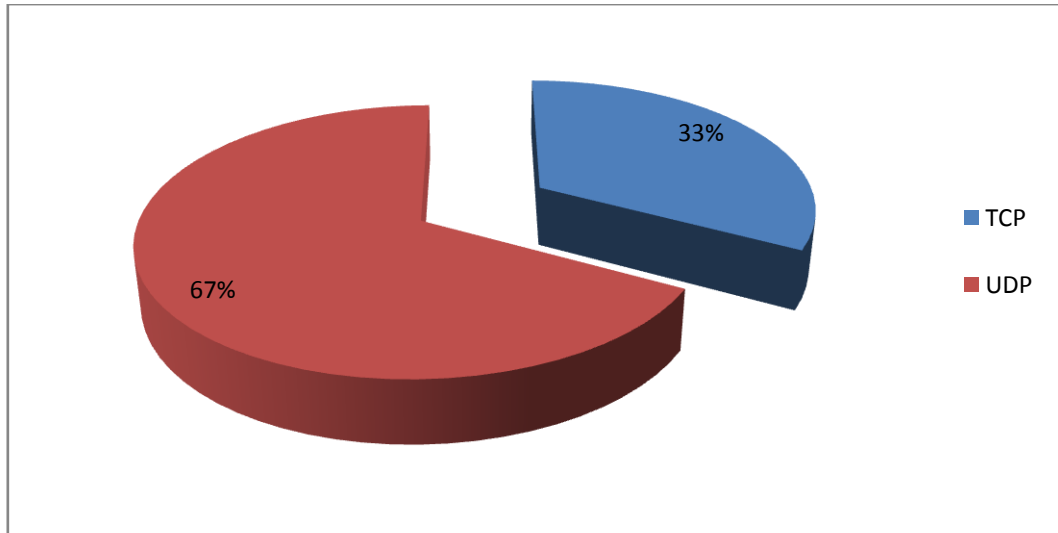
**Figure 5.8: The classification of user traffic by transport layer on Tuesday**

Figure 5.8 shows the transport layer protocols responsible for the traffic on Tuesday. The study sees that most traffic is generated by UDP (69%), 31% of the traffic is created by TCP. The study notes that the UDP traffic is increased 6.5% comparing with the first days of the week. In addition, it is observed that the TCP traffic is decreased about 5.5% comparing with Sunday and Monday. As a result, when the study sees that the number of authentication users on Tuesday is arrive to 477 users per second, this means that most of users are dealing with video and audio stream. Figure 5.9 shows the classification of user traffic by transport layer protocols on Wednesday.



**Figure 5.9: The classification of user traffic by transport layer on Wednesday**

Figure 5.9 shows the transport layer protocols responsible for the traffic on Wednesday. The study observes that most traffic is generated by UDP (55%), 45% of the traffic is created by TCP. But the study notes that the UDP traffic is decreased 14% comparing with Tuesday and 8.5% comparing with Sunday and Monday. In addition, it is observed that the TCP traffic is increased 14% comparing with Tuesday and 8% comparing with Sunday and Monday. As a result, the study observes that the users deal with video and audio stream is decrease comparing with the first days of week. Figure 5.10 shows the classification of user traffic by transport layer protocols on Thursday.



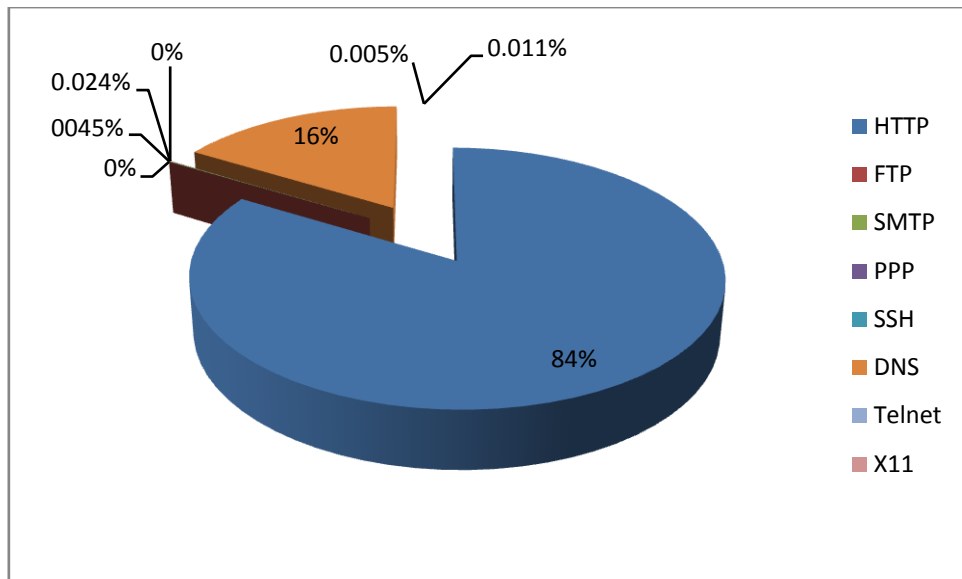
**Figure 5.10: The classification of user traffic by transport layer on Thursday**

Figure 5.10 shows the transport layer protocols responsible for the traffic on Thursday. The study sees that most traffic is generated by UDP (67%), 33% of the traffic is created by TCP. The study notes that the UDP traffic is increased 12% comparing with Wednesday. In addition, it is observed that the TCP traffic is decreased 12% comparing with Wednesday. As a result, the study observes that a lot of bandwidth is consumed on Thursday by UDP traffics, then TCP traffics.

In conclusion, the UDP traffic is between 55% - 69% through days of the week. The TCP traffic is between 31% - 45% through the days of week.

## **5.5 APPLICATION LAYER TRAFFICS**

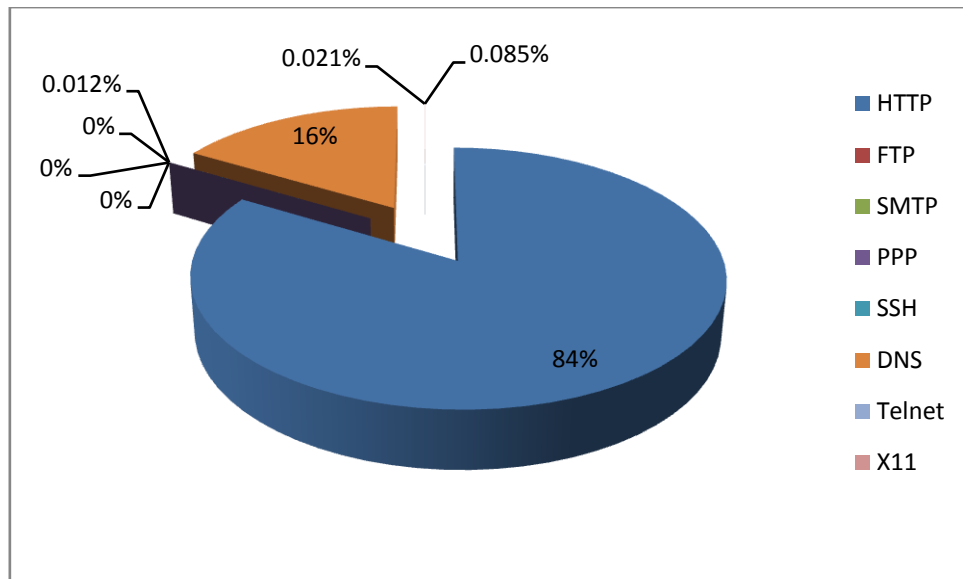
This section reports the analysis of the packet header traffic. Figures 5.11, 5.12, 5.13, 5.14, and 5.15 shows the classification of user traffic on Sunday, Monday, Tuesday, Wednesday, and Thursday by application layer protocols. Figure 5.11 shows the users' traffic by the application layer protocols on Sunday.



**Figure 5.11: The classification of user traffic by application on Sunday**

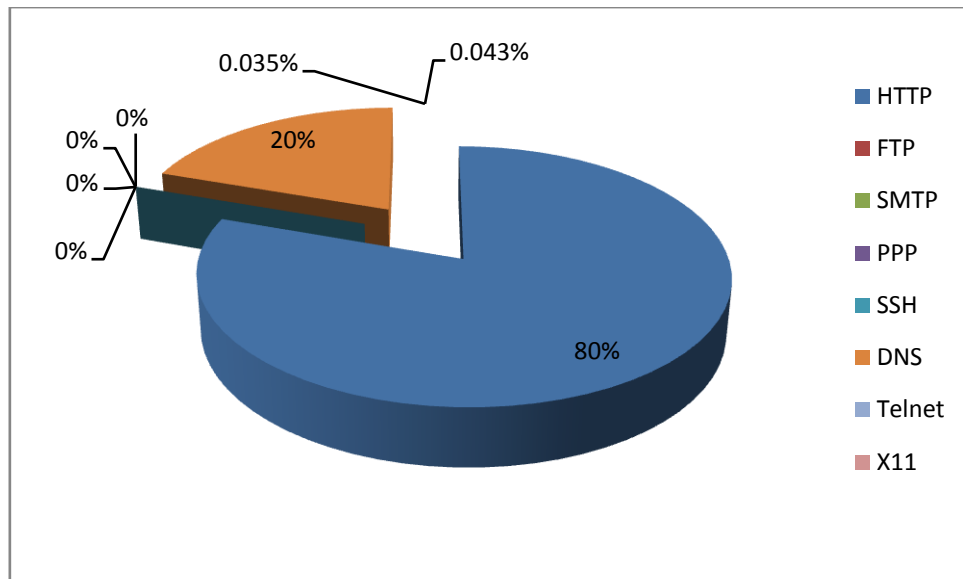
Figure 5.11 shows that web browsing (HTTP) is by far the most popular application protocol; contributing 84% of the total traffic transferred, followed by Domain Name System (DNS) (16%), and some SMTP, PPP, Telnet, and X11 traffics were observed, although they accounted for less than 1% of the total traffics transferred. As a result, the study notices that the popularity of web applications explain the peak users' bandwidth requirement. It is also observes that HTTP protocol consumes most the bandwidth. This means that most users' deal with web browsing under UDP protocol (like YouTube) because the UDP take 63% of remain traffic on Sunday. Figure 5.12 shows the classification of user traffic by Application on Monday.





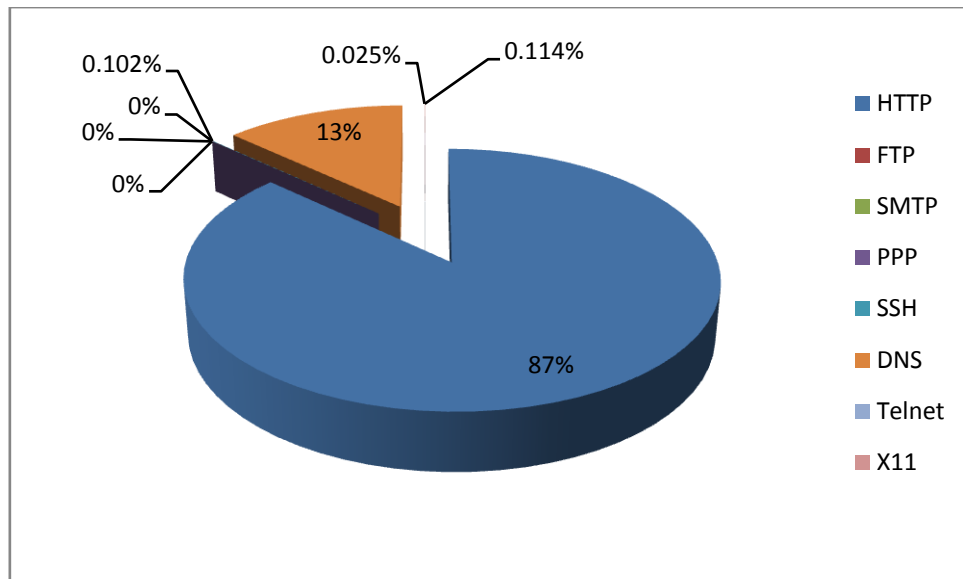
**Figure 5.12: The classification of user traffic by application on Monday**

Figure 5.12 shows that web browsing (HTTP) is by far the most popular application, contributing 84% (like Sunday) of the total traffic transferred, followed by Domain Name System (DNS) (16%), and some SSH, Telnet, and X11 traffics, although they accounted for less than 1% of the total traffics transferred, while the File Transfer Protocol (FTP), Simple Mail Transfer Protocol (SMTP), and Point-to-Point Protocol (PPP) are 0%. As a result, the study observes that the popularity of web applications explain the peak user bandwidth requirement. This means that most users deal with web browsing under UDP protocol because the UDP take 64% of remain traffic on Monday. Figure 5.13 shows the classification of user traffic by Application on Tuesday.



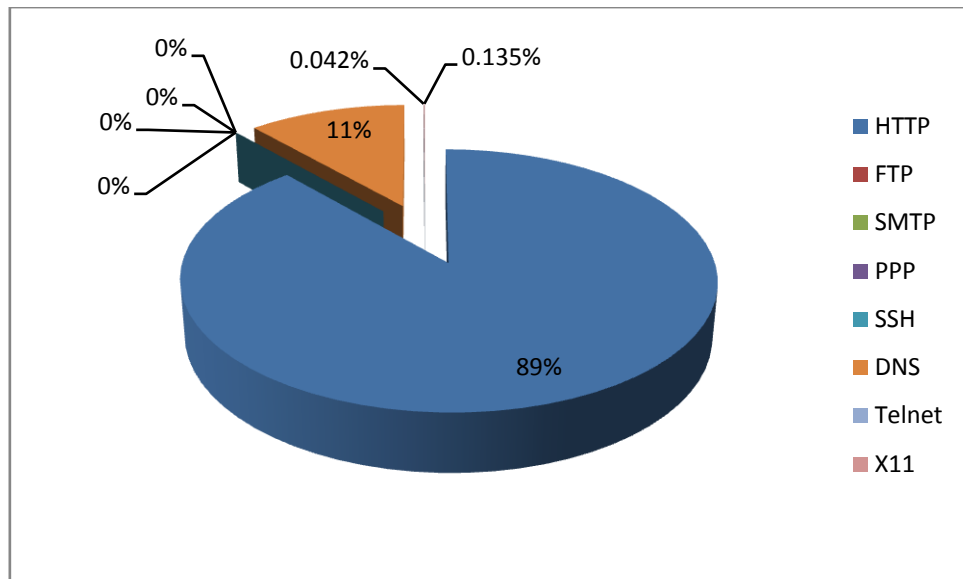
**Figure 5.13: The classification of user traffic by application on Tuesday**

Figure 5.13 shows that web browsing (HTTP) consume most of application traffic, contributing 80% of the total traffic transferred. The study observes that the percent is decreasing comparing with the first days of the week. The DNS contributes 20% of the traffic. Thus, the percent of DNS traffic is increasing on Tuesday 4% comparing with first days of week, and some Telnet, and X11 traffics were observed, although they accounted for less than 1% of the total traffics transferred. It is also sees that File Transfer Protocol (FTP), Simple Mail Transfer Protocol (SMTP), Point-to-Point Protocol (PPP) and Secure Shell are 0%. As a result, the web browsing is consuming most the application traffic during this day. Then it is also notices that the users increase their deals with the DNS during this day. Figure 5.14 shows the classification of user traffic by Application on Wednesday.



**Figure 5.14: The classification of user traffic by application on Wednesday**

Figure 5.14 shows that web browsing (HTTP) consumes most of application traffic; contributing 87% of the total traffic transferred, followed by Domain Name System (DNS) (13%), and although some SSH, Telnet, and X11 traffics were observed, they accounted for less than 1% of the total traffics transferred. It is also seen that File Transfer Protocol (FTP), Simple Mail Transfer Protocol (SMTP), and Point-to-Point Protocol (PPP) are 0%. As a result, the study observes that the users are dealing with a lot of web browsing in this day because the number of users is 400 and also the number of traffic is high, but the DNS traffic decreases. Figure 5.15 shows the classification of user traffic by Application on Thursday.



**Figure 5.15: The classification of user traffic by application on Thursday**

Figure 5.15 shows that web browsing (HTTP) is by far the most popular application; contributing 89% of the total traffic transferred, followed by Domain Name System (DNS) (11%). Although some Telnet and X11 traffics were observed, they accounted for less than 1% of the total traffics transferred. The study sees that File Transfer Protocol (FTP), Simple Mail Transfer Protocol (SMTP), Point-to-Point Protocol (PPP) and Secure Shell (SSH) are 0%.

In conclusion, the study sees that the web browsing traffic is between 80% – 89%. This means that most the application traffic in UUM wireless by HTTP. It is also observes that DNS traffic is between 11% - 20%, where arrive to the peak on Tuesday. Tables 5.1 and 5.2 of results after have been performed with wireshark, and tcpstat tools.

**Table 5.1: The top protocols responsible for the traffic**

		Sunday				
<b>Protocols</b>  <b>Parameter</b>	<b>TCP</b>	<b>UDP</b>	<b>ICMP</b>	<b>IP v4</b>	<b>IP v6</b>	<b>ARP</b>
<b>Avg. packets/sec</b>	18592.262	31696.836	99.193	56028.008	2.590	10743.397
<b>Avg. packet size (Bytes)</b>	649.933	407.147	125.893	393.746	150.086	107.397
<b>Avg. bytes/sec (Bps)</b>	12083732.562	12905256.199	12487.708	22060814.220	388.653	1153806.885
<b>Avg. Mbit/sec (Mbps)</b>	96.670	103.142	0.100	176.487	0.003	9.230
		Monday				
<b>Avg. packets/sec</b>	10611.107	19145.853	63.306	31590.124	8.848	4360.805
<b>Avg. packet size (Bytes)</b>	677.097	356.253	116.735	396.874	132.857	105.677
<b>Avg. bytes/sec (Bps)</b>	7184750.566	6820758.13	7390.083	12537310.497	1175.516	460837.595
<b>Avg. Mbit/sec (Mbps)</b>	57.478	54.566	0.059	100.298	0.009	3.687
		Tuesday				
<b>Avg. packets/sec</b>	12702.479	28366.607	173.588	47445.838	68.851	9477.070
<b>Avg. packet size (Bytes)</b>	599.906	421.185	125.157	382.457	134.964	106.130

<b>Avg. bytes/sec (Bps)</b>	7620289.624	11947592.86	21725.78	18145986.96	9292.38	1005798.47
<b>Avg. Mbit/sec (Mbps)</b>	60.962	95.581	0.174	145.168	0.074	8.046
		<b>Wednesday</b>				
<b>Avg. packets/sec</b>	20662.793	25436.896	97.762	48504.924	13.222	6822.955
<b>Avg. packet size (Bytes)</b>	651.997	400.228	124.466	424.751	132.624	107.500
<b>Avg. bytes/sec (Bps)</b>	13472072.244	10180556.29	12168.02	20602495.545	1753.511	73346.948
<b>Avg. Mbit/sec (Mbps)</b>	107.777	81.444	0.097	164.820	0.014	5.868
		<b>Thursday</b>				
<b>Avg. packet/sec</b>	18181.239	37206.566	111.439	56683.886	279.516	6667.881
<b>Avg. packet size (Bytes)</b>	649.737	373.009	120.007	388.603	137.417	106.701
<b>Avg. bytes/sec (Bps)</b>	11813014.651	13878401.318	13373.395	22027544.997	38410.279	711468.431
<b>Avg. Mbit/sec (Mbps)</b>	94.504	111.027	0.107	176.220	0.307	5.692

**Table 5.2: The most popular applications (protocols) seen in the traffic**

	Sunday									
<b>Protocols</b>	<b>HTTP</b>	<b>ICMP</b>	<b>ICMPV6</b>	<b>SMTP</b>	<b>IPX</b>	<b>PPP</b>	<b>SSH</b>	<b>DNS</b>	<b>Telnet</b>	<b>X11</b>
<b>Parameter</b>										
<b>Avg. packets/sec</b>	12013.644	99.193	2.516	10.988	37.459	3.168	0	2342.781	0.180	1.336
<b>Avg. packet size (Bytes)</b>	644.934	125.893	151.216	1232.910	153.045	359.077	0	182.334	89.500	1081.722
<b>Avg. bytes/sec (Bps)</b>	7748013.476	12487.708	380.391	13546.689	5732.919	1137.445	0	427169.558	16.123	1444.999
<b>Avg. Mbit/sec (Mbps)</b>	61.954	0.100	0.003	0.108	0.046	0.009	0	3.417	0.000	0.012
	Monday									
<b>Avg. packets/sec</b>	6500.857	63.306	1.504	0	8.327	0	0.428	1280.855	0.207	1.700
<b>Avg. packet</b>	698.140	116.735	142.114	0	219.768	0	335.200	188.484	113.000	1211.237

<b>size (Bytes)</b>										
<b>Avg. bytes/sec (Bps)</b>	4538510.254	7390.083	213.782	0	1829.990	0	143.533	241420.673	23.429	2085.727
<b>Avg. Mbit/sec (Mbps)</b>	36.308	0.059	0.002	0	0.015	0	0.001	1.931	0.000	0.016
	<b>Tuesday</b>									
<b>Avg. packets/sec</b>	10513.354	173.588	5.233	0	42.729	0	0	2579.750	0.472	0.263
<b>Avg. packet size (Bytes)</b>	574.550	125.157	146.303	0	242.359	0	0	154.033	280.882	1362.769
<b>Avg. bytes/sec (Bps)</b>	6040448.521	21725.781	765.610	0	10355.743	0	0	397366.184	132.612	358.667
<b>Avg. Mbit/sec (Mbps)</b>	48.324	0.174	0.006	0	0.083	0	0	3.179	0.001	0.003
	<b>Wednesday</b>									
<b>Avg. packets/sec</b>	10993.295	97.762	5.618	0	26.467	0	13.298	1654.211	0.770	3.464



<b>Avg. packet size (Bytes)</b>	723.257	124.466	142.280	0	183.353	0	210.286	169.108	190.065	1122.301
<b>Avg. bytes/sec (Bps)</b>	7950983.049	12168.092	799.390	0	4852.809	0	2796.418	279233.391	146.397	3887.535
<b>Avg. Mbit/sec (Mbps)</b>	63.608	0.097	0.006	0	0.039	0	0.022	2.234	0.001	0.031
	<b>Thursday</b>									
<b>Avg. packets/sec</b>	11071.613	111.439	7.663	0	48.690	0	0	1423.918	0.142	9.717
<b>Avg. packet size (Bytes)</b>	644.066	120.007	149.047	0	215.982	0	0	178.467	642.333	1134.387
<b>Avg. bytes/sec (Bps)</b>	7130851.683	13373.395	1142.101	0	10516.187	0	0	254114.596	91.124	11022.498
<b>Avg. Mbit/sec (Mbps)</b>	57.047	0.107	0.009	0	0.084	0	0	2.033	0.001	0.088

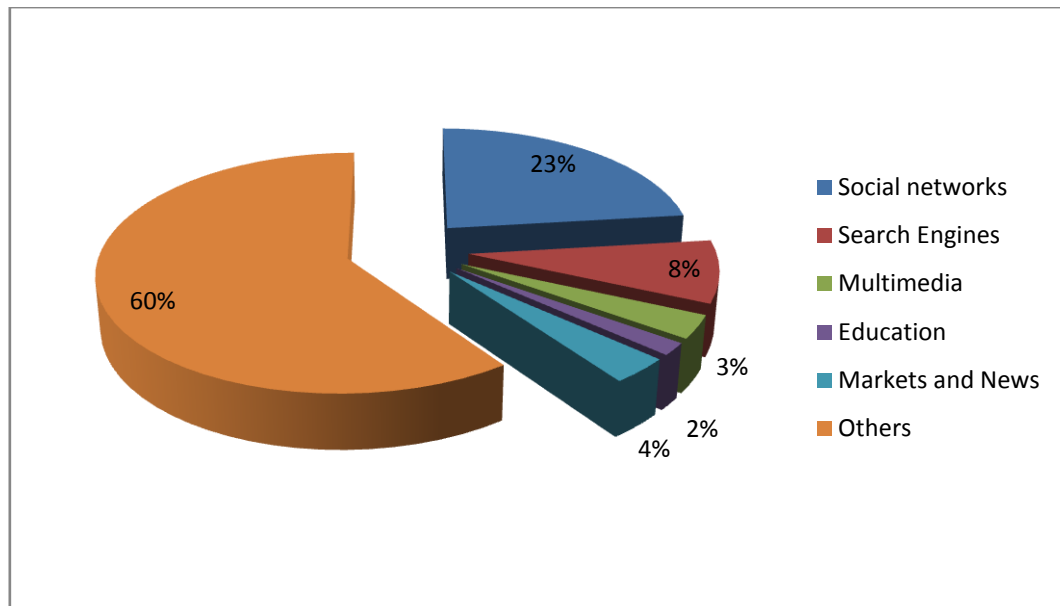
## 5.6 APPLICATION TRAFFIC

In this section, the study reports our analysis of the packet header traffic. This study goes dig deep to the application traffic. Table 5.3 shows the application categories and signatures (domains). The researcher select one day (Sunday) in this section because the time is not enough to select all the days of the week and the data is very huge. Moreover, this study needs more time, where each category of this study needs to classify the packet header based on the domain name and calculate the results to each application then combines it to represent the category. This investigation helps us to know the proportion between the application traffic, any applications take more time of users, and any applications take more bandwidth of wireless network.

**Table 5.3: The Application Category and Signatures**

Category	Signatures
<b>Social Networks</b>	Facebook; Twitter; Tagged; Digg; Flickr; Bing
<b>Search Engines</b>	Google; Yahoo; msn; Cari; Blogger
<b>Education</b>	cmslib.uum.edu.my (library); learningzone.uum.edu.my; umis.uum.edu.my (Portal); uum.org.my (UUM site); IEEE; ACM; Hub.sciverse
<b>Market &amp; News</b>	bbc; Apple; Mediafire; Airasia; Mudah
<b>Multimedia</b>	YouTube; Radioactive; Stumbleupon; Disqus; Reddit
<b>Others</b>	Unknown;

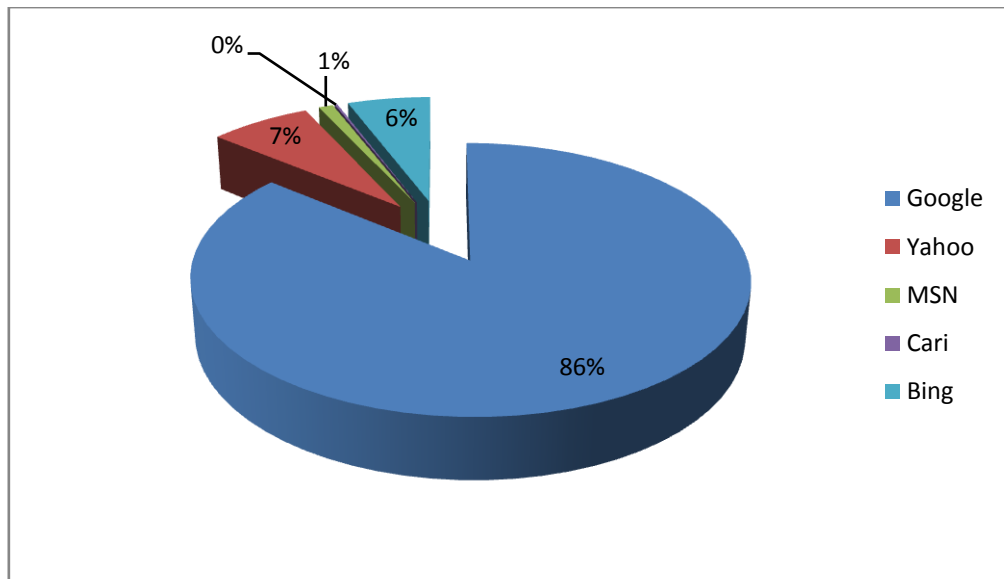
Figure 5.16 shows the classification of application traffic on Sunday as Social networks, Search Engines, Education, Markets & News, and Multimedia.



**Figure 5.16: The proportion of application traffic on Sunday**

Figure 5.16 shows a proportion of application traffic on Sunday. This study observes that most application traffic is generated by others 60% (unknown traffics), 21% of the application traffic created by Social networks, 10% of the application traffic created by Search Engines, 3 % of application traffic is generated by Multimedia, 2% of application traffic is created by Education, and 4% of application traffic is generated by News and Markets. As a result, the study observes that most of users deal with Social networks, then Search engines, Multimedia, Markets & News, and finally Education. This means that the social networks in UUM wireless take more bandwidth than Search Engines, Multimedia, Markets & News, and Education. Figures 5.17, 5.18, 5.19, 5.20, and 5.21 show the applications inside each category.

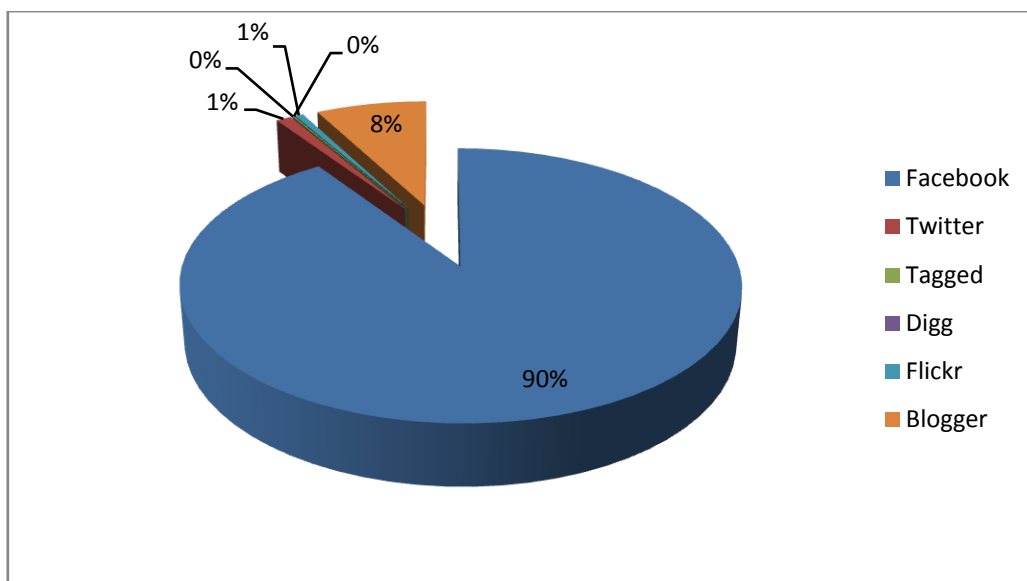
Figure 5.17 shows the percent to each application inside the Search Engines, likes Google, yahoo, Cari, Bing, and MSN. This helps us to know which application of this category takes high percent.



**Figure 5.17: The proportion of Search Engines application traffic**

Figure 5.17 shows a proportion of the most popular applications seen in the traffic in Search Engines. It is observed that Google Search Engine consumes most of application traffic, contributing 86% of the total traffic transferred, followed by Yahoo Search Engines 7%, Bing Search Engine (6%), MSN Search Engine (1%), and Cari Search Engine which accounted for less than 1% of the total social traffics. As a result, it is observed that most of users deal with Google Engine. This means that most bandwidth in Search Engines is consumed by the Google, then Yahoo, Bing, MSN, and Cari. This could be that the Google Engine takes high percent because it very popular, very advanced, and easy search.

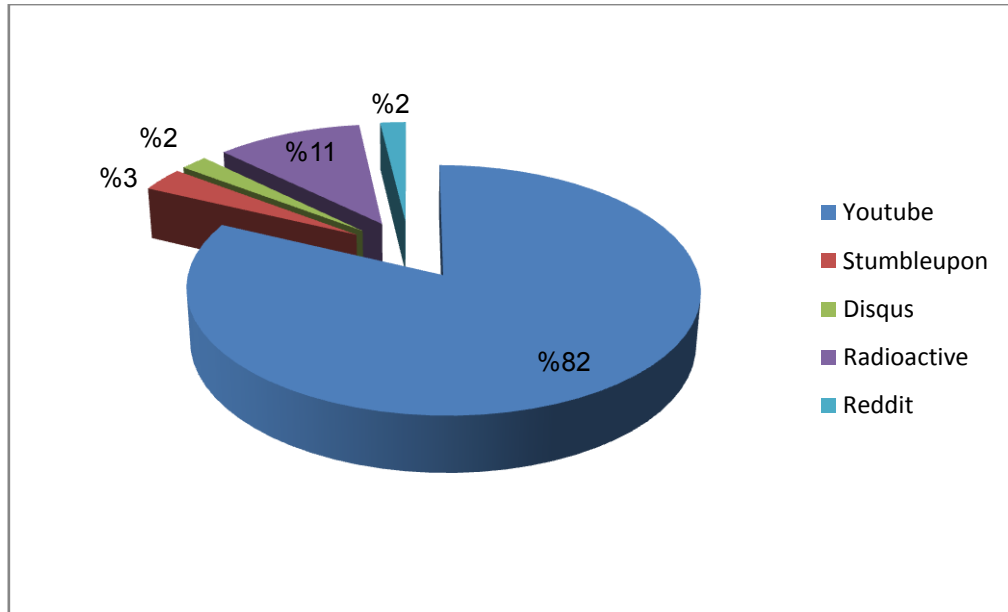
Figure 5.18 shows the percent to each application inside the Social Networks, likes Facebook, Twitter, Tagged, Blogger, Digg and Flickr. This helps us to know which application of this category takes the high percent.



**Figure 5.18: The proportion of Social Networks application traffic**

Figure 5.18 shows a proportion of the most popular applications seen in the traffic in Social Networks. It is observed that Facebook site consumes most of application traffic, contributing 90% of the total Social traffic transferred. This means that the facebook site is more widespread and common among the users; followed by Blogger site 8%, Twitter site (1%) Flickr site (1%), Tagged site accounted for less than 1% of the total social traffics, and Digg site has (0%). The results show that most of users deal with Facebook Social site. This means that most bandwidth in Social Networks is consumed by the Facebook, then Blogger, Twitter, Flickr, Tagged, and Digg. It is not surprising that the Facebook Social site takes more percent because it very popular and free to use.

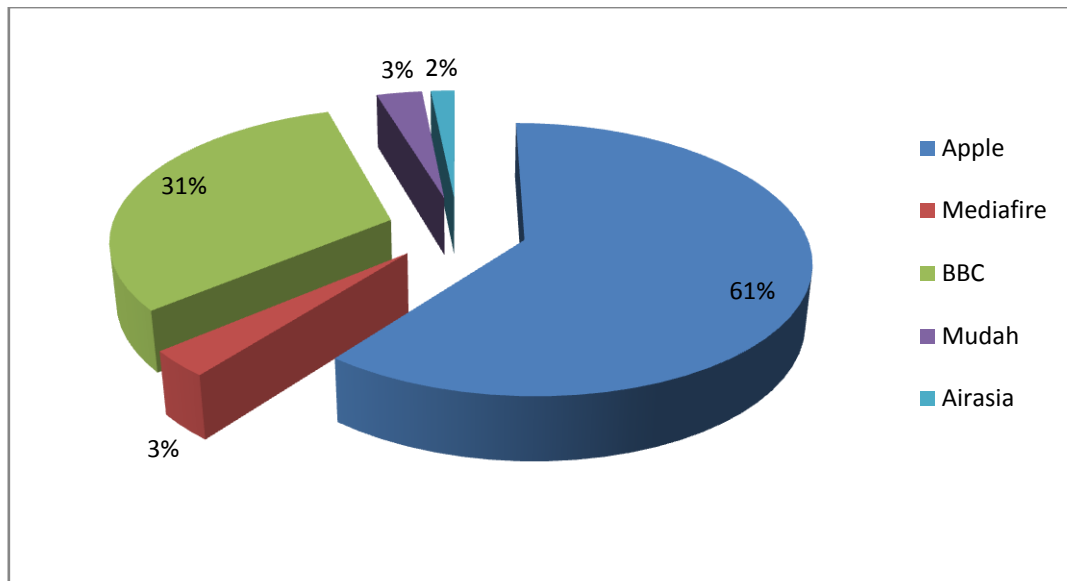
Figure 5.19 shows the proportion to each application inside the Multimedia, likes YouTube, Stumbleupon, Disqus, Radioactive and Reddit. This study helps us to know which application of this category takes high percent.



**Figure 5.19: The proportion of Multimedia application traffic**

Figure 5.19 shows a proportion of the most popular applications seen in the traffic in Multimedia. It is observed that YouTube site consumes most of application traffic, contributing 82% of the total multimedia traffic, followed by Radioactive site 11%, Stumble upon site (3%), Disqus site (2%), and Reddit site which has (2%) of the total traffics. As a result, it is observed that most of the users deal with YouTube. This means that most bandwidth in Multimedia category consumes by the YouTube, then Radioactive, Stumble upon, Disqus, and Reddit. Of course, the YouTube site takes high percent because it very popular and it provides more space to upload.

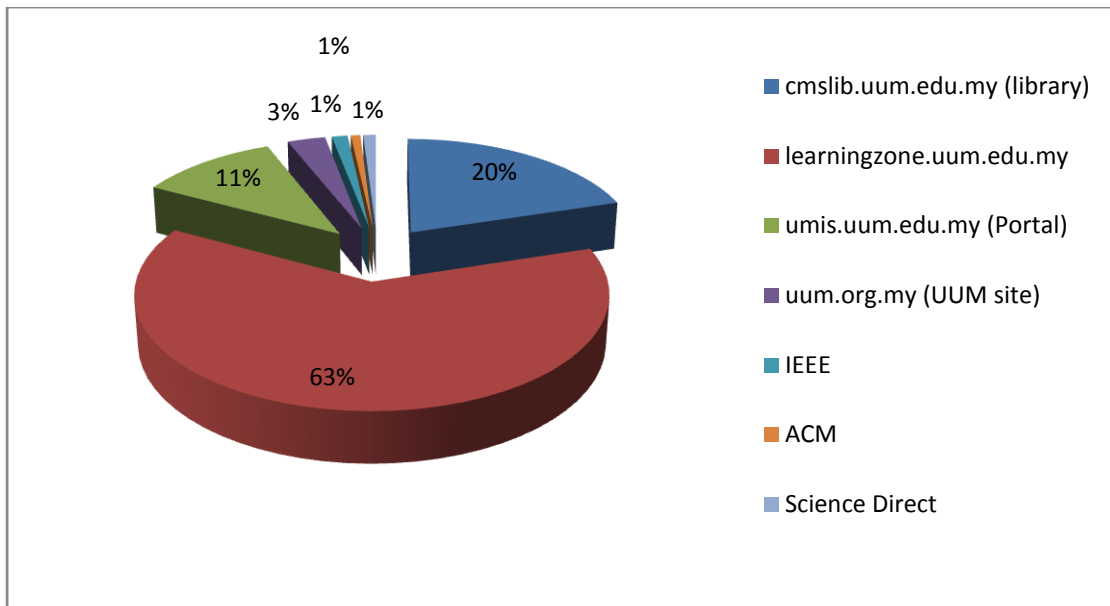
Figure 5.20 shows the proportion to each application inside the Markets & News, likes Apple, Mediafire, Mudah, BBC and Airasia. This investigation helps us to know which application of this category takes high percent.



**Figure 5.20: The proportion of Markets & News application traffic**

Figure 5.20 shows a proportion of the most popular applications observed in the traffic in Markets & News. It is noticed that Apple site consumes most of application traffic, contributing 61% of the total Markets & News traffic, followed by BBC site 31 %, Mediafire site (3%), Mudah site (3%), and Airasia site (2%). As a result, it is observed that most of users deal with Apple markets and BBC news. This means that most bandwidth in Markets & News is consumed by the Apple, then BBC, Mediafire, Mudah, and Airasia. Of course, the Apple site takes more percent because it very popular and display the news about each new product. Moreover, the political tension in the Middle East helped to increase follow-up news. Therefore, the percent to BBC is very high (31%) of total Markets & News traffics.

Figure 5.21 shows the proportion to each application inside the Education, likes Library site, learning zone site, Portal site, UUM site, IEEE site, ACM site, and Science Direct site. This study helps us to know which application of this category takes high percent.



**Figure 5.21: The proportion of Education application traffic**

Figure 5.21 shows a proportion of the most popular applications seen in the traffic in Education. It is observed that learning zone site consumes most of application traffic, contributing 63% of the total Education traffic transferred. Of course, the learningzone takes more percent because all the lectures and exams' news are updating in this site, followed by Library site (20 %), Portal site (11%), UUM site (3%), IEEE site (1%), Science Direct site (1%), and ACM site (1%). As a result, it is noticed that most of the users deal with Apple markets and BBC news. This means that most bandwidth in Education is consumed by the Learning zone, then Library, Portal, UUM site, IEEE, Science Direct, and ACM. It is not surprising that library takes high percent after learningzone because all the users need to book the resources online.



## 5.7 SUMMARY

The results summarize the conclusions of this chapter by giving a high-level characterization of wireless traffics in a Universiti Utara Malaysia wireless network. The high-level characterization of traffics is described in four parts are: network layer traffics, transport layer traffics, application layer traffics, and application traffic measurement. The first part shows the characterization of the network layer protocols like IP v4, ICMP v4, ICMP v6, IPX, IP v6, and ARP. The study observes that the percent of IP v4 protocols arrived to a high level, then ARP. This means that a lot of bandwidth consumes by IP v4, then ARP in the network layer traffic. The second part shows the characterization of the transport layer protocols like TCP and UDP. The study observes that the percent of UDP protocols arrived to a high level, then TCP. This means that a lot of bandwidth consumes by UDP, then TCP in the transport layer traffic. The third parts shows the characterization of the application protocols like HTTP, DNS, FTP, X11, PPP, SMTP, Telnet, and SSH. The study notes that web browsing (HTTP) are by far the most popular application; consumes most the application traffics, then DNS. Finally, the fourth parts shows the application traffics measurement like Social networks, Search Engines, Multimedia, Markets & News, and Education. The study in this part observe that the Social Network consumes more the bandwidth, then Search Engines, Markets & News, Multimedia, and Education.

## **CHAPTER SIX**

### **CONCLUSION**

#### **6.1 INTRODUCTION**

This chapter consists of four parts as follows: The first part presents a shorting summary of the five chapters, problems and limitations of the study are discussed in the second part, the contribution of the study is offered in the third part, and the proposal for future works is presented in the fourth part.

#### **6.2 RESEARCH SUMMARY**

This research study has the opportunity to introduce an investigation about the performance and users' behavior to UUM wireless network. Unlike previous studies of wireless networks, our traffics were captured of wireless controller in a computer center, where the researcher captured the first 1000 bytes of the packet for security reasons only. Our centralized network design made this study possible and more adequate as well as reduced the error-free traffic comparing with those analyzed in previous projects. The goal of our analysis is to determine where, when, how much, and for what our campus wireless network is being used.

Reviewing previous studies, there is no investigating performance of network such as throughput, load, and latency in UUM wireless data networks by performing real measurement as well as there is no study on users' behavior which focused on the Internet traffic analysis. Studying the performance and users' behavior of UUM wireless network help us to know where the loss of bandwidth, network load, throughput, and end to end latency that will make detection of bottle-neck area easier on the communication network, assist us in building the capacity of the network, and

evaluate if the wireless network on campus needs to improve the service in order to increase the amount of subscription. In addition, understanding the behavior of traffic analysis helps us to develop, manage WLAN technology, and also to apply our workload analysis results on issues in wireless network deployment such as capacity planning, and potential network optimizations such as algorithms for load balancing across multiple Access Points (APs) in a UUM wireless network.

While a single week might not be representative of overall usage patterns, the researcher feels that the results presented here do offer some real insights into how UUM wireless networks are used. The study observed that the average of authenticated wireless users on our campus connection is unstable because the number of active APs is unstable during the week. The popularity of a given access point was largely determined by its accessibility and familiarity to users. In our results, a small number of authenticated users that may generate high level of traffics and reverse is not true because this based on directly related with the packet rate. The number of traffic on Tuesday arrives to high level because the number of authenticate users on Tuesday 477 and the number of active APs 91. While the number of traffic on Wednesday has low level although of the number of users is high because some of users do not have directly related to the packet rate. But the percent is close between Sunday and Monday although the number of users is unequal between them.

The high lose of bandwidth on Thursday arrives a high level than on Monday, Tuesday, Wednesday, and Sunday, while the low loss of bandwidth is equal on all days. The packet lengths 80-159 consume most of the bandwidth. The throughput on Sunday arrives to a high level 24% than on Thursday 23%, Wednesday 21%, Tuesday 19%, and Monday 13%. The study observes the increase in throughput leads to increase in load.

In the top protocols, the study notes that the UDP protocol has the highest level during the week days followed by the TCP. While in the application protocol, HTTP protocol arrives to a high level during the week days, then the DNS protocol. This means that high consumption of the bandwidth by the UDP protocol, then TCP in the top protocols, and high consumption of the bandwidth by HTTP and DNS in the application protocols.

The study classified the application traffics into five categories as follows: Social Networks, Search Engines, Multimedia, Markets and News, and Education. Throughout the study on application traffics, the study observes that the Social Network arrives to a high level than Search Engines, Markets and News, Multimedia, and Education. But when the study goes dig deep to each category, it observes that the Facebook reaches a very high level in the Social Networks, then Google Search Engine reaches a high level in the Search Engines, YouTube site arrives a high level in the Multimedia, Apple markets and BBC news arrives a high level in the Markets and News, and Learningzone site arrives a high level in the Education.

### **6.3 PROBLEMS AND LIMITATION**

The major challenge encountered throughout the semester of this study was on data capture. In order to complete a careful work, specific data is required. The study classifies the challenges during the data capture to three parts: Operating system challenges, amount of data, and the tools. In first part, the researcher installed the Red Hat in the virtual box on the desktop of Windows 7. The researcher encountered a lot of problems during the installation such as the impossibility of installing some of the packages such as libpcap library to tcpdump, netperf, ntop, and tcpstat. This obstacle leads to another problem which is also the impossibility of making the sniffer or

monitoring by use this method of installation (Red Hat). Therefore, the researcher is not recommended to use this method of installation to Linux inside the virtual box, but he is recommended to use the Ubuntu in this study. Also, the researcher is recommended to install the operating system in a large part of hard disk. In the second part, because the amount of the received data during the capture is very large which is received from the mirror port on the Switch which was sometimes 1 GB of data by tcpdump tool every minute and some time every second, the researcher lost the operating system (Ubuntu 11.04). In the third part, the researcher encountered another problem with the tool that should be used to capture data. So, the researcher in this study used tcpdump tool to make data capture or data collection. The researcher is not recommended to use the Wireshark to capture the data because determining the promiscuous mode is not easy and it depends on NIC type. Finally, the limitation of the time is considered another problem.

## **6.4 CONTRIBUTION**

Users in UUM need high speed access to wireless connections that lead to increase the access to the online subjects and lectures. This study highlights the UUM wireless characterization of Wi-Fi. Due to that, Wi-Fi seems to be a good solution inside wireless networks. Thus, this study investigates the behavior and performance of UUM wireless network that lead to improve the speed of wireless to be suitable to user's needs. In addition, the occurrence of Wi-Fi could not pass some limitations or attract a comprehensive way in wireless networking behaviors.

## **6.5 FUTURE WORK**

The vision of our hope is that others will perform future studies of user behavior and network performance in similar and different wireless networks so that general characteristics and trends can be identified over time. The study also suggests other researchers to study the wireless in all buildings of UUM. In addition, this study would like to further study the geographic patterns of mobility. Seemingly, the majority users have regular habits as they move from their dorm to the class to the dining halls. The study was unable to distinguish between the users or types of users (students, faculty, and staff). It may be possible to infer the type of users from their behavior (for example, students are seen frequently in dorms), or to use clustering techniques. The study was also unable to distinguish the mobile host hardware (PDA, laptop, or desktop) or operating system, but for those seen in a tcpdump trace the study may be able to learn something from the protocols they use. Thus, the study suggests for future researchers to distinguish between mobile host hardware.

## REFERENCES

- [1] D. M. Anurag Kumar, Joy Kuri, *Wireless networking 2009*.
- [2] J. Feng, *Wireless networks*, 2011.
- [3] Alberto Escudero P. Sebastian Buettlich, "*Basic Wireless Infrastructure and Topologies*," 2009.
- [4] Y. Miyahara, "Next-generation wireless technologies trends for ultra low energy," in *Low Power Electronics and Design (ISLPED) 2011 International Symposium on*, 2011, pp. 345-345.
- [5] M. Ciampa, *CWNA Guide to Wireless LANS. Networking*. Thomson Publish, 2006.
- [6] A. Kumar, *et al.*, *Wireless networking*: Morgan Kaufmann, 2008.
- [7] R. Price, *Fundamentals of wireless networking*: McGraw-Hill, Inc., 2006.
- [8] D. Schwab and R. Bunt, "Characterising the use of a campus wireless network," in *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, 2004, pp. 862-870 vol.2.
- [9] D. Kotz and K. Essien, "Analysis of a campus-wide wireless network," *Wireless Networks*, vol. 11, pp. 115-133, 2005.
- [10] D. Hucaby and S. McQuerry, *Cisco field manual: catalyst switch configuration*: Cisco Systems, 2003.
- [11] T. Cooklev, "Wireless Communication Standards: A Study of IEEE 802.11, 802.15, and 802.16," *IEEE*, 2004.
- [12] A. Durresi and M. Denko, "Advances in wireless networks," *Mobile Information Systems*, vol. 5, pp. 1-3, 2009.
- [13] T. Cornelsen, "*WiFi- implementation Fundamentals and piloting*," *PHD Thesis, Computer science University for Applied Science of Regensburg*, 2010.
- [14] W. Gardner, "Spectral correlation of modulated signals: Part I--analog modulation," *Communications, IEEE Transactions on*, vol. 35, pp. 584-594, 1987.
- [15] V. K. Garg, *Wireless communications and networking*: Morgan Kaufmann, 2007.
- [16] P. Rengaraju, *et al.*, "Measuring and Analyzing WiMAX Security and QoS in Testbed Experiments," in *Communications (ICC), 2011 IEEE International Conference on*, 2011, pp. 1-5.
- [17] M. Hassan and R. Jain, *High performance TCP/IP networking*: Pearson Prentice Hall, 2004.

- [18] E. Kartsakli, *et al.*, "Multiuser MAC Protocols for 802.11n Wireless Networks," in *Communications, 2009. ICC '09. IEEE International Conference on*, 2009, pp. 1-5.
- [19] Y. Kawasumi, "Deployment of WiFi for rural communities in Japan and ITU's initiative for pilot projects," in *Enterprise Networking and Computing in Healthcare Industry, 2004. HEALTHCOM 2004. Proceedings. 6th International Workshop on*, 2004, pp. 200-207.
- [20] A. A. Khan and N. Zaman, "Comparative analysis of broadband wireless access from Wi-Fi to WiMax," in *Applied Sciences and Technology (IBCAST), 2009 6th International Bhurban Conference on*, 2009, pp. 8-14.
- [21] B. S. C. Choi and M. Gerla, "Wireless Interrupt: Inter-Device Signaling in Next Generation Wireless Networks," in *INFOCOM IEEE Conference on Computer Communications Workshops*, 2010, 2010, pp. 1-5.
- [22] Y. M. Li and J. H. Jhang-Li, "Integration of WiMAX and WiFi Services: Bandwidth Sharing and Channel Collaboration," *Scopus*, 2010.
- [23] K. G. Paterson, "Generalized Reed-Muller codes and power control in OFDM modulation," *Information Theory, IEEE Transactions on*, vol. 46, pp. 104-120, 2000.
- [24] F. Ohrtman, *WiMAX handbook*: McGraw-Hill, 2005.
- [25] Z. Zeyu, *et al.*, "ONU Placement in Fiber-Wireless (FiWi) Networks Considering Peer-to-Peer Communications," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, 2009, pp. 1-7.
- [26] S. Ahson and M. Ilyas, *The WiMAX handbook*: CRC Press, 2008.
- [27] J. G. Andrews, *et al.*, *Fundamentals of WiMAX: understanding broadband wireless networking*: Prentice Hall PTR, 2007.
- [28] M. A. Dye, *et al.*, *Network Fundamentals: CCNA Exploration Companion Guide*: Cisco Press, 2007.
- [29] H. Zimmermann, "OSI reference model--The ISO model of architecture for open systems interconnection," *Communications, IEEE Transactions on*, vol. 28, pp. 425-432, 1980.
- [30] L. L. Peterson and B. S. Davie, *Computer networks: a systems approach*: Morgan Kaufmann, 2003.
- [31] J. F. Kurose and K. W. Ross, *Computer networking*: Pearson/Addison Wesley, 2010
- [32] G. S. Poo and B. P. Chai, "ISO FTAM protocol performance," *Computer Communications*, vol. 14, pp. 413-422, 1991.
- [33] R. Lai, *et al.*, *On using PROTEAN to verify ISO FTAM protocol*: Springer, 1991.
- [34] S. Radicati, *Electronic mail: an introduction to the X-400 message handling standards*: McGraw-Hill, Inc., 1992.



- [35] U. Warrier, *et al.*, "Common management information services and protocols for the internet (CMOT and CMIP)," *RFC1189*, 1990.
- [36] J. Postel and J. Reynolds, "Rfc 959: File transfer protocol (ftp)," *InterNet Network Working Group*, 1985.
- [37] A. Tang, *et al.*, "Transport layer," *Wiley Encyclopedia of Computer Science and Engineering*, 2009.
- [38] L. Parziale, *et al.*, *TCP/IP Tutorial and Technical Overview*: IBM International Technical Support Organization, 2006.
- [39] J. F. Kurose and K. W. Ross, *Computer networking*: Pearson/Addison Wesley, 2011.
- [40] M. Gast, *802.11 wireless networks: the definitive guide*: O'Reilly Media, 2005.
- [41] S. Frankel, *et al.*, "Establishing wireless robust security networks: a guide to IEEE 802.11 i," *National Institute of Standards and Technology*, 2007.
- [42] D. Skordoulis, *et al.*, "IEEE 802.11 n MAC frame aggregation mechanisms for next-generation high-throughput WLANs," *Wireless Communications, IEEE*, vol. 15, pp. 40-47, 2008.
- [43] R. Blum, *Network Performance Open Source Toolkit Using Netperf, tcptrace, NISTnet, and SSFNet*: John Wiley & Sons, Inc., 2003.
- [44] J. Yeo, *et al.*, "A framework for wireless LAN monitoring and its applications," in *WiSe '04 Proceedings of the 3rd ACM workshop on Wireless security*, 2004, pp. 70-79.
- [45] J. Yeo, "Measuring traffic on the wireless medium: Experience and pitfalls," DTIC Document 2002.
- [46] P. Orosz and T. Skopko, "Software-Based Packet Capturing with High Precision Timestamping for Linux," in *Systems and Networks Communications (ICSNC), 2010 Fifth International Conference on*, 2010, pp. 381-386.
- [47] T. Kalibera, *et al.*, "Automated benchmarking and analysis tool," in *'06 Proceedings of the 1st international conference on Performance evaluation methodologies and tools*, 2006, p. 5.
- [48] S. S. Kolahi, *et al.*, "Performance Monitoring of Various Network Traffic Generators," in *Computer Modelling and Simulation (UKSim), 2011 UkSim 13th International Conference on*, 2011, pp. 501-506.
- [49] H. Asai, *et al.*, "Towards characterization of wireless traffic in coexisting 802.11 a/g and 802.11 n network," in *CoNEXT '10 Student Workshop Proceedings of the ACM CoNEXT Student Workshop*, 2010, p. 1.
- [50] A. Balachandran, *et al.*, "Characterizing user behavior and network performance in a public wireless LAN," in *SIGMETRICS '02 Proceedings of the 2002 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, 2002, pp. 195-205.

- [51] A. Gember, *et al.*, "A Comparative Study of Handheld and Non-handheld Traffic in Campus Wi-Fi Networks Passive and Active Measurement." vol. 6579, N. Spring and G. Riley, Eds., ed: Springer Berlin / Heidelberg, 2011, pp. 173-183.
- [52] D. Tang and M. Baker, "Analysis of a local-area wireless network," in *MobiCom '00 Proceedings of the 6th annual international conference on Mobile computing and networking*, 2000, pp. 1-10.
- [53] D. Kotz and K. Essien, "Characterizing usage of a campus-wide wireless network," Technical Report TR2002-423, Dartmouth College, March 2002, pp. 107-118.
- [54] R. Hutchins and E. W. Zegura, "Measurements from a campus wireless network," in *Communications, 2002. ICC 2002. IEEE International Conference on*, 2002, pp. 3161-3167 vol.5.
- [55] T. Henderson, *et al.*, "The changing usage of a mature campus-wide wireless network," in *MobiCom '04 Proceedings of the 10th annual international conference on Mobile computing and networking*, 2004, pp. 187-201.
- [56] M. Balazinska and P. Castro, "Characterizing mobility and network usage in a corporate wireless local-area network," in *MobiSys '03 Proceedings of the 1st international conference on Mobile systems, applications and services*, 2003, pp. 303-316.
- [57] H. Wei-jen and A. Helmy, "On Modeling User Associations in Wireless LAN Traces on University Campuses," in *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, 2006 4th International Symposium on*, 2006, pp. 1-9.
- [58] E. Zola, *et al.*, *User behaviour in a WLAN campus: a real case study*, 2009.
- [59] D. Niyato and E. Hossain, "Wireless broadband access: Wimax and beyond-integration of wimax and wifi: Optimal pricing for bandwidth sharing," *Communications Magazine, IEEE*, vol. 45, pp. 140-146, 2007.
- [60] I. L. M. S. Committee, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," *IEEE Standard*, vol. 802, p. 999, 1999.
- [61] M. L. Gress and L. Johnson, *Deploying and troubleshooting Cisco wireless LAN controllers*: Cisco Systems, 2009.
- [62] D. Phillips, "Computer music and the linux operating system: A report from the front," *Computer Music Journal*, vol. 27, pp. 27-42, 2003.
- [63] A. M. I. McHoes and I. M. Flynn, *Understanding operating systems*: Course Technology Ptr, 2010.
- [64] M. Welsh and S. S. Consultants, *Linux Installation and Getting Started*: Specialized Systems Consultants, 1998.
- [65] M. Garrels, *Introduction to Linux*: Fultus Publishing, 2010.
- [66] A. Petersson, "Operating systems," 2000.

- [67] C. L. Van Jacobon, and Steven McCanne. (1997). *tcpdump, manual page*. Available: <http://www.tcpdump.org>
- [68] A. Orebaugh, *et al.*, *Wireshark & Ethereal network protocol analyzer toolkit*: Syngress Media Inc, 2007.
- [69] P. Herman. (2000, 4/3/2012). *tcpstat* , *manual page*. Available: <http://frenchfries.net/paul/tcpstat/>
- [70] K. C. Rick Jones, Dave Shield. (1996). *Netperf. Manual Page*. Available: <http://www.netperf.org>
- [71] U. Lamping, "Wireshark Developer's Guide," 2004.
- [72] G. Combs, "Wireshark-network protocol analyzer," *Version 0.99*, vol. 5, 1998.
- [73] V. Y. Hnatyshin and A. F. Lobo, "Undergraduate data communications and networking projects using opnet and wireshark software," in *SIGCSE '08 Proceedings of the 39th SIGCSE technical symposium on Computer science education*, 2008, pp. 241-245.
- [74] CALIN. (2009). *Wireshark's most useful display filters*. Available: <http://www.firstdigest.com/2009/05/wiresharks-most-useful-display-filters/>
- [75] P. K. Janert, *Gnuplot in action: understanding data with graphs*: Manning Publications Co., 2009.

## APPENDIX



**UUM**  
Universiti Utara Malaysia

NetS / UUM / 12 / 012

College of Arts and Sciences  
Universiti Utara Malaysia  
06010 UUM Sintok  
Kedah Darul Aman, Malaysia  
Tel: (604) 928 6777  
Faks: (604) 928 6783  
<http://www.cas.uum.edu.my>

6 March 2012

Prof. Dr. Zulkhairi Md Dahalin  
Director of Computer Center  
Universiti Utara Malaysia

Assalamualaikum wr. wbkth.  
Dear Prof,


### DATA COLLECTION PHASE

Regarding on the above matter, **Mr. Wisam Dawood (808266)** and **Mr. Mustafa M. H. Ibrahim (808988)** are our Msc.IT students. They are currently doing their Master project on network performance measurements. They need to capture real data of our UUM network.

In accordance with that, I apply to seek your support and help to ease their works.

Consideration and The Honourable's support, I precede with a thousand thanks.

Yours Faithfully,

  
(Dr. Mohd Hasbullah Omar)  
Head of Computer Science Department  
School of Computing  
UUM College of Arts and Sciences

TERIMA

14 MAR 2012

NetS  
Pusat Komputer  
Universiti Utara Malaysia

En. Khalil  
7/3

En. Ali  
Tg. En. Lank  
Jilid 1/2012

TERIMA

- 7 MAR 2012

Pengarah  
Pusat Komputer  
Universiti Utara Malaysia



Universiti Pengurusan Terkemuka