WiFi Traffic Analysis

Akshitha Muthireddy
Department of Computer Science
Stony Brook University
Stony Brook, New York, USA

Divya Akuthota
Department of Computer Science
Stony Brook University
Stony Brook, New York, USA

Abstract — **We present the results of analysis of our campus-wide wireless network. We have analyzed a three-day trace of Stony Brook University's mobile network, WolfieNet, to analyze how the people make use of it. While collecting the trace, address and protocol information of each packet sent and received on the network is recorded. The campus-wide wireless network had four sub-divisions, WolfieNet Secure, WolfieNet-Get-Connected, WolfieNet-Open and WolfieNet-Guest. We collected the trace at three locations in the University, Chapin apartments, Frank Melville Jr. Memorial Library and Health Science Center Library. The trace collected had packets being sent and received on the above four networks. The trace was analyzed to find out about where, when and how much the wireless network is being used. The performance related issues are also addressed in the paper.**

Keywords-Wireless LANs, Traffic Analysis, 802.11

I. INTRODUCTION

Wireless network refers to any type of computer network that is not connected by cables of any kind. It is a method by which homes, telecommunications networks and enterprise installations avoid the costly process of introducing cables into a building, or as a connection between various equipment locations. Wireless telecommunications networks are generally implemented and administered using radio communication. This implementation takes place at the physical layer of the OSI model network structure.

WiFi stands for wireless frequencies / fidelity which enable the interconnectivity of many computers hence a way to connect internet from the access point to the computer or laptop. For enabling WiFi you need to use a device access point or technically named router having wireless facility to receive or transfer the signal.

IEEE 802.11 is a set of standards for implementing wireless local area network computer communication in the 2.4, 3.6 and 5 GHz frequency bands. These standards provide the basis for wireless network products using the Wi-Fi brand. Current 802.11 standards define frame types for use in transmission of data as well as management and control of wireless links. Frames are divided into very specific and standardized sections. Each frame consists of a MAC header, payload and frame check sequence (FCS).

In a wireless network, packets are continuously exchanged between various entities in the network infrastructure. These may be beacons, request packets, acknowledgement packets, data packets etc. These packets provide the essence of network activity.

In this project, we have analyzed a network trace of WolfieNet-Secure, WolfieNet-Get-Connected, WolfieNet-Open and WolfieNet-Guest at three different places in Stony Brook University like residence apartments and libraries during different times in three days. We chose residences and libraries as that is where we can expect busy environments as they are crowded with students. The main goal is to find out answers to questions like when was the network busiest, how was its performance during heavy traffic, etc.

This paper is organized as follows. Section II reviews the related work in this area. Section III describes the nature of environment in which we have carried out our analysis. Section IV details the approach we have used to perform to gather the data and analyze it. The results of the analysis have been listed in Section V. We conclude in Section VI with the summary of our work.

II. RELATED WORK

Tang and Baker [2] analyzed a 12-week trace collected from the wireless network used by the Stanford Computer Science department. This study is built on earlier work involving fewer users and a shorter duration. They analyzed traces of 74 computer-science students in one building. Their study provides a good qualitative description of how mobile users take advantage of a wireless network. Earlier, Tang and Baker [3] also characterized user behavior in a metropolitan area network, focusing mainly on user mobility. Furthermore, the network was spread over a larger geographical area and had very different performance characteristics.

Kotz and Essien [1] traced and characterized the Dartmouth College campus-wide wireless network during their Fall 2001(11 weeks) term. Their workload is quite extensive, both in scope (1706 users across 476 access points) and duration (12 weeks). Kotz and Essien focus on large-scale characteristics of the campus, such as overall application mix, overall traffic per building and AP, mobility patterns, etc. In terms of application mix, their network carries a richer set of applications that reflects the nature of campus-wide applications.

Schwab and Bunt [6] performed a week-long wireless traffic trace in January 2003 at University of Saskatchewan across 18 access points which were spread out in the university’s 40 buildings. A centralized authentication log was used to match packets with wireless access points. Their aim was to analyze the usage, such as the roaming patterns of the users and the purposes for which wireless networks were actually used, of the campus wireless network.

III. TESTING ENVIRONMENT

A. *Nature of Enviroment*

The traces were collected in three different settings which had different kinds of network traffic.

a) Chapin Apartments

These are the place where most of the graduate students reside on-campus.

b) Melville Library

The main library in the campus which has two different reading rooms, Central Reading Room where all kinds of books are available and North Reading Room where mainly science related books are available

c) Health Sciences Center Library

This library is mostly populated by the medical students.

B. *Nature of Network*

There are four different kinds of networks from which packets were collected. These networks are various WolfieNet networks.

a) WolfieNet-Secure

This is a secure network for students, faculty and staff. It utilizes WPA2 Enterprise authentication using NetID and NetID password. It is located in campus buildings, common areas and residence halls.

b) WolfieNet-Get-Connected

This needs to be used the first time a person is trying to connect to campus WiFi. XpressConnect technology is used to properly configure settings on the person's device to make for a quicker and easier connection to WolfieNet-Secure.

c) WolfieNet-Open

This network is unsecured and is present only in residence halls.

d) WolfieNet-Guest

There is no authentication requirement for this network and is meant for the visitors to the university on campus.

C. *Software and Hardware Specifications*

For collecting the network trace the device and software specifications are as follows:

a) Software Specifications

Wire Shark was used to filter and analyze the data that has been collected. It is a free and open-source packet analyzer. It is mainly used for network troubleshooting, analysis, software and communications protocol development, and education.

Data was collected in Ubuntu, an operating system based on Debian Linux distribution which is distributed as open source software.

b) Hardware Specifications

HP Pavilion dv6 laptop with Intel i7 processor acts as the sniffer.

IV. METHODOLGY

The data was first collected using tcpdump, stored in a .cap file for analysis in Wire Shark. We collected the data on 13th December, 2012, 14th December, 2012 and 15th December 2012.

A. *Trace Collection*

For collecting the trace, a series of steps were followed.

- Initially, we disable the wireless network. The Enable Networking is deselected.
- The wireless LAN interface is turned OFF using the command:

```
$ sudo ifconfig wlan0 down
```

- The wireless LAN interface is put in monitor mode, its status is checked and then it is turned ON again using the commands:

```
$ sudo iwconfig wlan0 mode monitor
$ iwconfig wlan0
$ sudo ifconfig wlan0 up
```

- Packet capture software, tcpdump, is used and the wireless LAN card captures the data which is not only directed to our laptop. This is stored in a .cap file which can be opened in Wire Shark for analysis. This is done using the command:

```
$ sudo tcpdump -i wlan0 -n -s 1000 -w file.cap
```

In the monitor mode, the wireless interface can only receive the packets but, cannot transmit the packets.

After collection of data, the mode has to be changed to managed from monitor for normal running of the wireless interface.

B. Analysis

The data collected is analyzed. The number of different types and also subtypes of 802.11 frames are filtered. The three different types of frames are:

i. Management Frames: The filtering equation used is:

wlan.fc.type eq 0

The different types of management frames are:

a.) Association request frame:
- Filtering equation is: wlan.fc.type_subtype eq 0

b.) Association response frame:
- Filtering equation is: wlan.fc.type_subtype eq 1

c.) Reassociation request frame:
- Filtering equation is: wlan.fc.type_subtype eq 2

d.) Reassociation response frame:
- Filtering equation is: wlan.fc.type_subtype eq 3

e.) Probe request frame:
- Filtering equation is: wlan.fc.type_subtype eq 4

f.) Probe response frame:
- Filtering equation is: wlan.fc.type_subtype eq 5

g.) Beacon frame:
- Filtering equation is: wlan.fc.type_subtype eq 8

h.) Authentication frame:
- Filtering equation is: wlan.fc.type_subtype eq 11

i.) Deauthentication frame:
- Filtering equation is: wlan.fc.type_subtype eq 12

ii. Control Frames: The filtering equation used is:

wlan.fc.type eq 1

The different types of control frames are:

a.) Acknowledgement frame (ACK):
- Filtering equation is: wlan.fc.type_subtype eq 29

b.) Request to Send frame (RTS):
-Filtering equation is: wlan.fc.type_subtype eq 27

c.) Clear to Send frame (CTS):
- Filtering equation is:
wlan.fc.type_subtype eq 28

iii. Data Frames: The filtering equation used to find the data frame is:

wlan.fc.type eq 2

We consider NULL data frames and the filtering equation used is:
wlan.fc.type_subtype eq 36

We have also filtered out the malformed and retransmitted packets. The malformed packets were found using the filter ‘malformed’ and retransmitted packets were found using the filter equation wlan.fc.retry == 1.

V. RESULTS

After collecting the data from different places like Chapin Apartments, Melville Library, Health Sciences Center Library, the data consists of packets of different types like Management frames, Control Frames, Data Frames. The number of Packet types differs with the traffic Diversity in that place. Given below are the statistics of the various kinds of frames and also malformed and retransmitted packets.

We plot various graphs for different scenarios. First, we plot a graph showing the different types of frames i.e., management, control and data at the three places i.e., Chapin Apartments, Melville Library and Health Sciences Center Library over the three days.

Plots that represent the relation between captured packets, malformed packets and retransmitted packets are also drawn. These are plotted using Wire Shark. Three different graphs are plotted for the three different locations on the same day i.e., 14th December, 2012.

Statistics for different types of 802.11 Frames and retransmitted and malformed packets are below:

Management Frames			
Subtypes	Chapin Apartments	Melville Library	HSC Library
Association request	234	290	301
Association response	182	202	225
Reassociation request	225	191	145
Reassociation response	185	154	71
Probe request	18991	12611	28117
Probe response	109663	51240	66389
Beacon	2107161	2175908	1318102
Authentication	1451	1281	1218
Deauthentication	150	126	51
Control Frames			
Request to Send	1165327	1773373	408803
Clear to Send	360743	7234890	279332
Acknowledgement	662061	959300	436307
Data Frames			
NULL data	664138	45259	55299

Table 1: Data related to different types and sub-types of frames

Packet Type	Chapin Apartments	Melville Library	HSC Library
Retransmitted Packets	424640	280014	1031827
Malformed Packets	52929	96434	38623

Table 2: Data related to retransmitted and malformed packets that were captured at the three locations

GRAPHS:

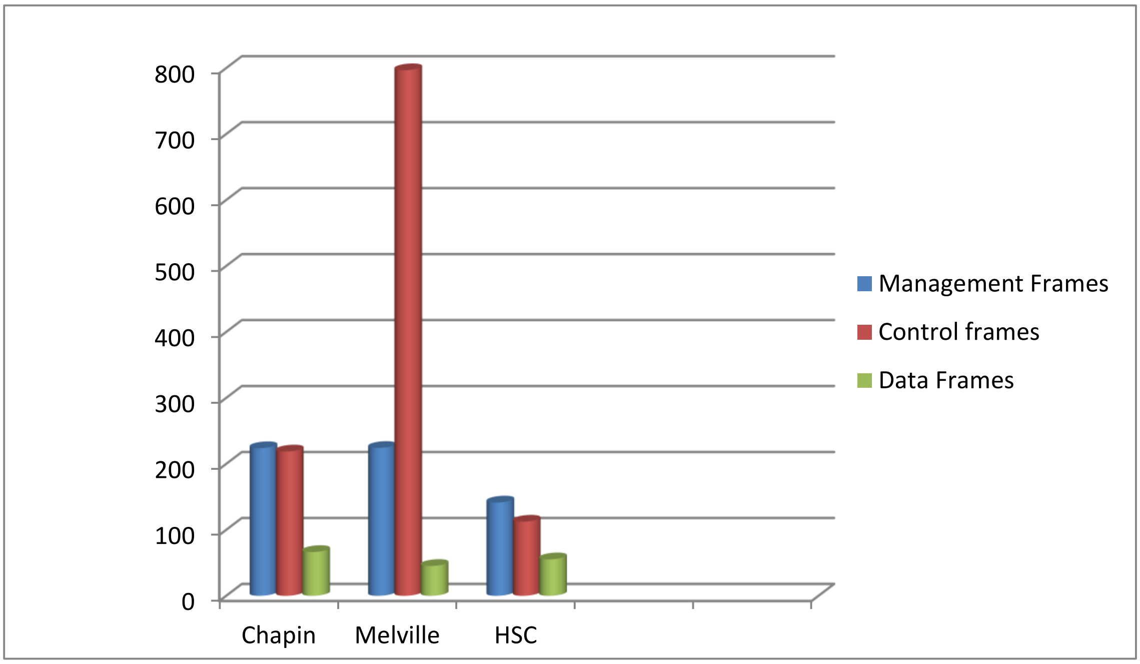

Figure 1: Comparison of the number of Management, Control and Data frames at 3 different places – Chapin Apartment, Melville Library, Health Science Center Library(number of frames are divided by 10000 for all categories)

For data collected at Chapin, the number of management frames and control frames are almost equal in number.
For the data collected at Melville, the number of control frames is very large when compared to Management Frames and Data Frames.
For the data collected at HSC, the number of frames decreased from Management to control to Data.

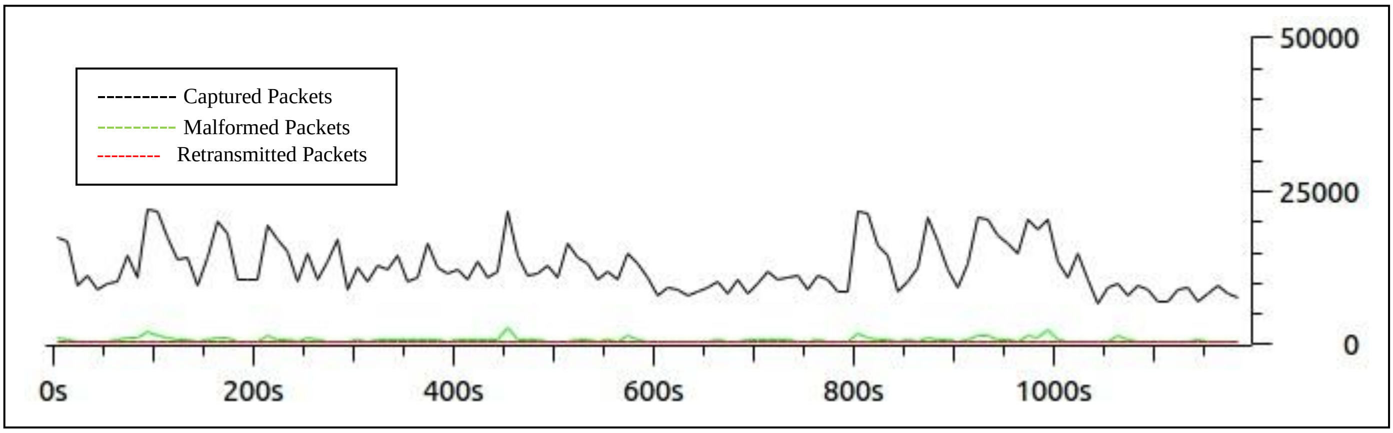

Figure 2: Relation between captured packets, malformed packets and retransmitted packets in HSC on Date 14th December 2012.

The number of malformed packets is very less when compared to the captured packets. The number of retransmitted packets is also less.

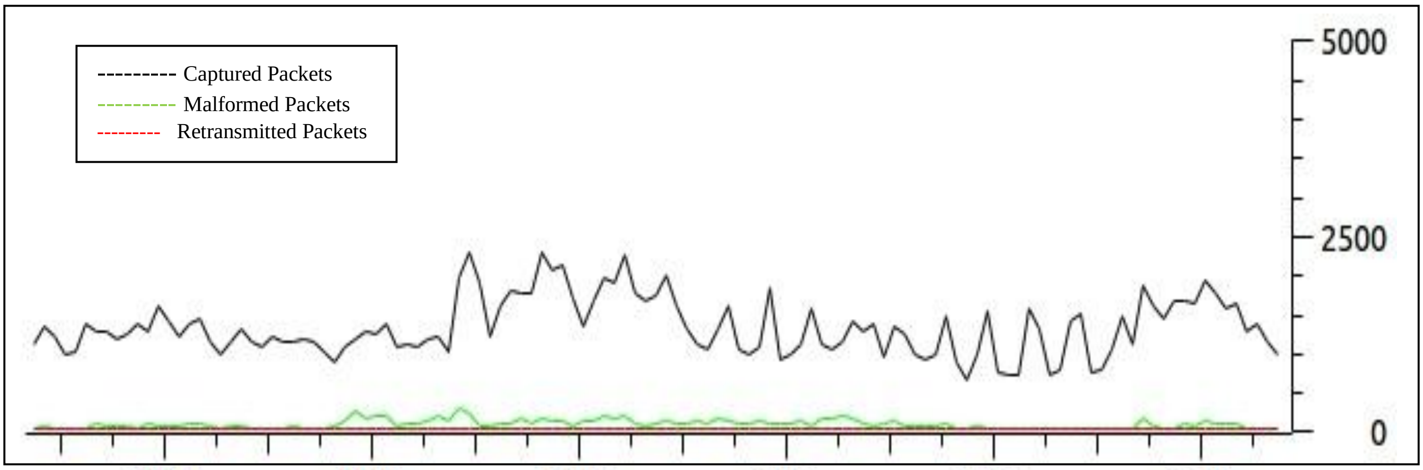

Figure 3: Relation between captured packets, malformed packets and retransmitted packets in Chapin Apartments on 14th December, 2012.

The number of malformed packets is almost negligible and the number of retransmitted packets is also less. The number of retransmitted packets in Chapin Apartments is more than those in Health Science Center Library.

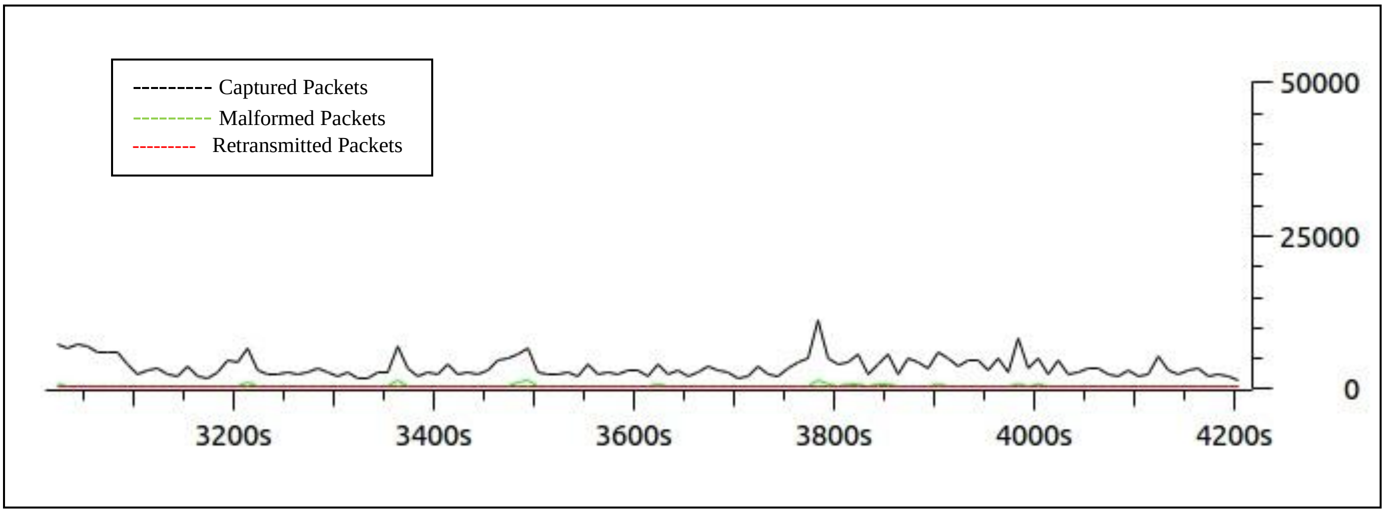

Figure 4: Relation between captured packets, malformed packets and retransmitted packets in Melville Library on 14th December, 2012

When compared to the number of packets captured, both retransmitted and malformed packets are negligible in Melville Library. This is a good performance with respect to retransmissions and packet loss.

CONCLUSION

In this experiment, we illustrate the performance of the network in different settings and on different days. The traffic is more in libraries (Melville Library and HSC Library) on weekdays when compared to that on weekends. This is due the increase in number of students in libraries on weekdays. When residence halls (Chapin Apartments) are considered, traffic is more on weekend than on weekdays. The traffic is relatively more in libraries on the days the data was collected as it is exam time. Although our findings are logically accurate, there could be some discrepancies owing to the distance from the access point which varied as we did not sniff packets from the exact same position on all three days that we collected the data. We collected data in these three different settings as each had different traffic diversities.

References

[1] David Kotz and Kobby Essien, "Analysis of Campus-wide Wireless Network," MobiCom'02, Proceedings of the 8th annual international conference on Mobile computing and Networking, pp. 107-118,September 23–28, 2002, Atlanta, Georgia, USA.

[2] D. Tang and M. Baker, "Analysis of a Local-Area Wireless Network," Proceedings of ACM MobiCom '00, pp. 1-10, Boston, MA, August 2000.

[3] D. Tang and M. Baker, "Analysis of a Metropolitan-Area Wireless Network", Proceedings of ACM MobiCom'99, pages 13–23, August 1999.

[4] K. Lai, M. Roussopoulos, D. Tang, X. Zhao, and M. Baker, "Experiences with a Mobile Testbed," Worldwide Computing and Its Applications, Lectures notes in Computer Science (1368). Berlin: Springer, 1998, 222-237.

[5] B. J. Bennington and C. R. Bartel, "Wireless Andrew: Experience building a high speed, Campus-Wide Wireless Data Network," Proceedings of ACM MobiCom'97, pages 55–65, August 1997.

[6] D. Schwab and R. Bunt, "Characterizing the use of a Campus Wireless Network," IEEE Infocom 2004.

[7] A. Balachandran, et. al., "Characterizing User Behaviour and Network Performance in a Public Wireless LAN," Proceedings of ACM SIGMETRICS '02, pp. 195-205, Los Angeles, CA, June 2002.

[8] B. Noble, M. Satyanarayanan, G. Nguyen, and R. Katz, "Trace-based Mobile Network Emulation," Proceedings of ACM SIGCOMM'97, pages 51–61, September 1997.