IMPROVING PERFOMANCE OF 3G MOBILE DATA OFFLOADING THROUGH WIFI NETWORKS BY DATA CACHING

By

CECILIA WAMBUI GICHUHI

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE IN DATA COMMUNICATION IN THE FACULTY OF COMPUTING AND INFORMATION MANAGEMENT AT KCA UNIVERSITY

AUGUST, 2013

DECLARATION

I declare that this Research project is my original work and has not been previously published or submitted elsewhere for award of a degree. I also declare that this Research project contains no material written or published by other people except where due reference is made and author duly acknowledged.

Student Name: ______________________ Reg. No. ____________

Sign: ______________________ Date: ____________

I do hereby confirm that I have examined the master's Research project of

Cecilia Gichuhi

AND have certified that all revisions that the Research project panel and examiners recommended have been adequately addressed.

Sign: ______________________ Date: ____________

Prof. Ddembe Williams

IMPROVING PERFOMANCE OF 3G MOBILE DATA OFFLOADING THROUGH WIFI NETWORKS BY DATA CACHING

ABSTRACT

In the recent years there has been a skyrocket rise in mobile data usage especially due to the large number of smartphones in use. This has led to a rigorous traffic overloading in cellular networks and the trend is expected to continue. It is imperative that architectures are put in place to handle this data. Use of Wifi to offload data has been considered as one of the immediate solutions.

In this paper I propose Wifi caching to offload mobile data. Tests were carried out using OMNet++ a discreet event simulator. From the experiments done it can be seen that the throughput is best when the number of hosts are about 3 and 10. Further increase of hosts above 10 hosts then the throughput and efficiency begin to go down. However when LRU caching is implemented the general throughput is slightly lower than when caching is not implemented which is a concern, and was not the expected result. The user request latency is slightly higher after caching.

The AP caching model can be tested using other different caching algorithms to see which one would give best results.

Key words: Data offloading, Wifi offloading, Wifi caching, Caching model, OMNet++, AP caching

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First and foremost my gratitude goes to the almighty God whose gift of life and grace; without which I would not reach this far.

Thanks to my research project supervisor Dr. Ddembe Williams, for the invaluable supervision and time spent discussing the progress of the project.

I am especially thankful to my husband and family for their understanding, encouragement and morale they gave me during my research study.

Thank you to my friends and classmates at KCA University for the continuous support and encouragement.

ABBREVIATIONS

3G	Third Generation
4G	Fourth Generation
AP	Access point
CAGR	Compounded Annual Growth rate
CCK	Complementary Code Keying
EB	Exabyte
GUI	Graphical User Interface
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IEEE	Institute of Electrical and Electronics Engineers
LFU	Least Frequently Used
LTE	Long Term Evolution
LRU	Least Recently Used
MB	Megabyte
MDO	Mobile data offloading
MoSoNets	Mobile Social Networks
NED	Network description language
OFDM	Orthogonal Frequency Division Multiplexing
OMNeT++	Objective Modular Network Testbed in C++
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RF	Radio Frequency

SS	Spread Spectrum
STBC	Alamouti Space-Time Block Coding
VANET	Vehicular Adhoc Network
WLAN	Wireless Local Area Network
Wi-Fi	Wireless Fidelity

LIST OF FIGURES

LIST OF TABLES

DEFINITION OF TERMS

3G - Is short for third Generation, it is a word used to signify the 3rd generation of mobile telecommunications technology. It is also named Tri-Band 3G.

4G- 4G is the fourth generation of mobile phone mobile communication technology standards. It is a successor of the 3G standards.

AP – A wireless access point is a device that allows wireless devices to connect to a wired network using Wi-Fi, or related standards.

Caching- Caching is the temporary storage of content, such as HTML pages, images, files and web objects, for the benefit of offloading the origin web or application server and reducing the time required to deliver content to a client.

CAGR- Compounded Annual Growth rate is a business and investing specific term for the smoothed annualized gain of an investment over a given time period.

Exabyte- A unit of information or computer storage equal to one quintillion. The unit symbol for the exabyte is EB. 1 EB = 1,000,000,000 gigabytes = 1,000,000terabytes = 1,000petabytes

Femtocells- In telecommunications, a femtocell is a small, low-power cellular base station, typically designed for use in a home or small business.

LTE-Long Term Evolution refers to a standard for smooth and efficient transition toward more advanced leading-edge technologies to increase the capacity and speed of wireless data networks. It is a 4G Technology.

MB- The term MB stands "megabyte" is commonly used to mean either 1000^2 bytes or 1024^2 bytes.

MDO - Mobile data offloading, also called data offloading is the use of complementary network technologies for delivering data originally targeted for cellular networks.

MoSoNets - Mobile Social Networks is social networking where individuals with similar interests converse and connect with one another through their mobile phone and/or tablet.

Smart phones - A smartphone is a mobile phone built on a mobile operating system, with more advanced computing capability and connectivity than a feature phone.

VANETs - A Vehicular Ad-Hoc Network, is a technology that uses moving cars as nodes in a network to create a mobile network. VANET turns every participating car into a wireless router or node, allowing cars approximately 100 to 300 metres of each other to connect and, in turn, create a network with a wide range.

Wi-Fi – Wi-Fi short for Wireless Fidelity is a popular technology that allows an electronic device to exchange data wirelessly (using radio waves) over a computer network, including high-speed Internet connections.

CHAPTER 1: INTRODUCTION

1.1. Background of study

Subsequent to the increase of internet use via smart phones, there has been a severe traffic overloading of cellular networks, leading to deteriorating 3G quality especially in metropolitan areas. “The increase of smart phone devices has resulted in a skyrocketing growth of mobile traffic, which in 2011 grew 2.3-fold, more than doubling for the fourth year in a row, and is expected to grow 18 times from 2011 until 2016” (V. Siris & D. Kalyvas, 2012). The difference between a regular mobile phone and a smart phones is that a smart phone has superior computing ability and more connecting options than a regular cellphone, meaning the user can use applications like media players, camera, video streaming, messaging and so on, leading to increased network traffic. The Cisco Visual Networking index ,2011-2016 reports that Smartphones represent only 12 percent of total global handsets in use today, but they represent over 82 percent of total global handset traffic. In 2011, the typical smartphone generated 35 times more mobile data traffic (150 MB per month) than the typical basic-feature cell phone (which generated only 4.3 MB per month of mobile data traffic). “According to Cisco forecasts and practical experiences of mobile operators, we are now facing the “mobile data apocalypse”. Mobile data traffic grows at a compound annual growth rate (CAGR) of 131 percent between 2008 and 2013, and will exceed two exabyte per month in 2013” (Dimatteo et al, 2011)

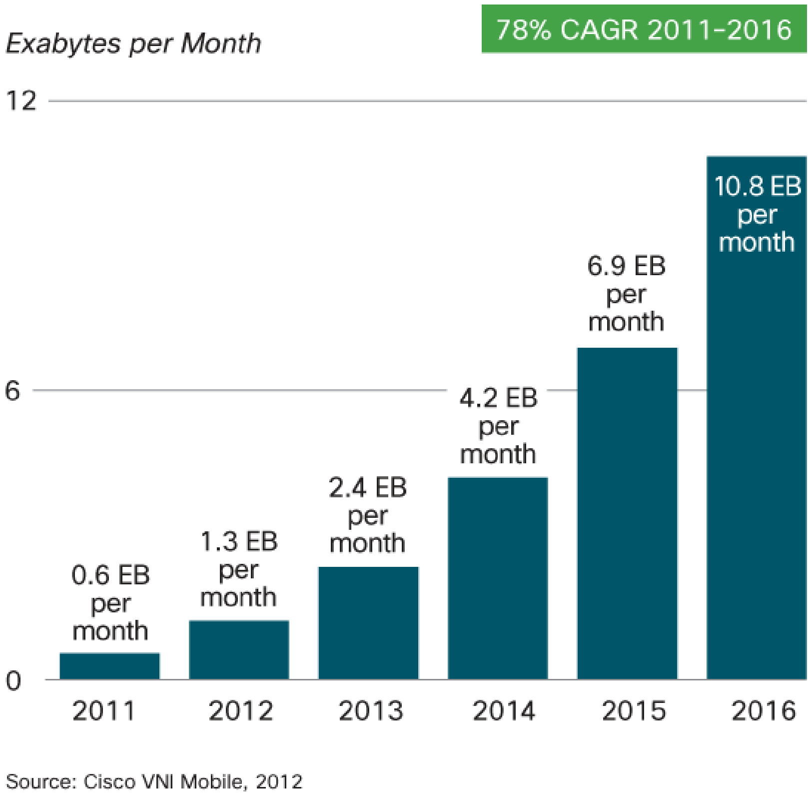

Figure 1.1 Cisco Forecasts 10.8 Exabytes per month of Mobile Data Traffic by 2016

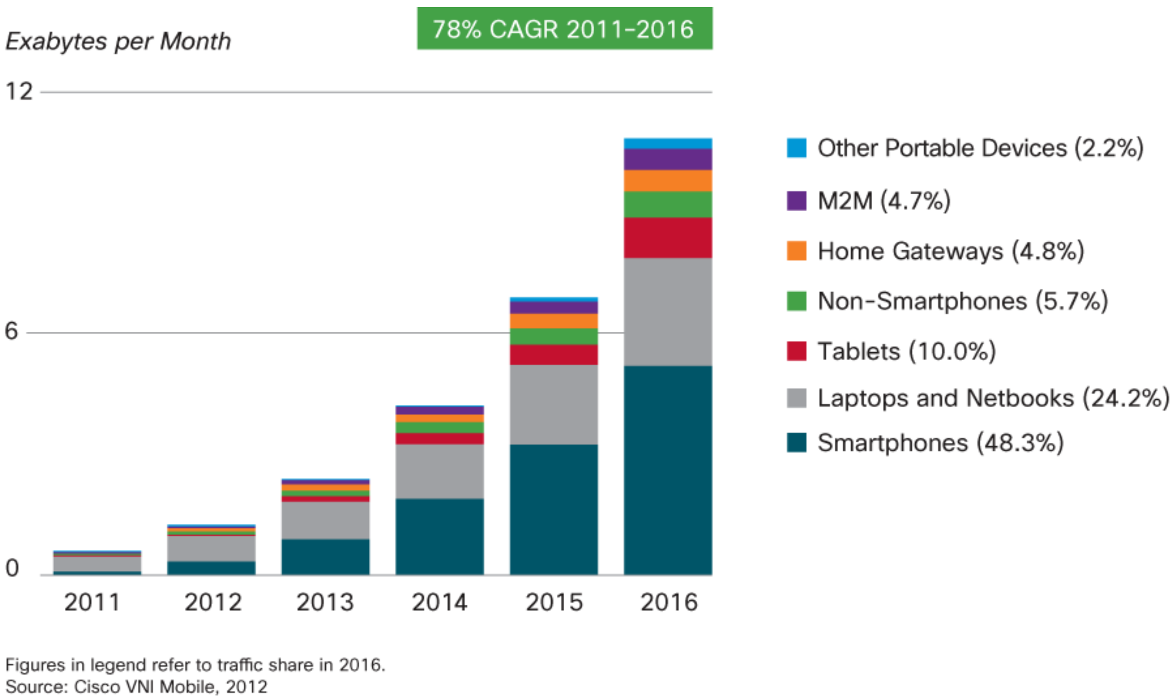

Figure 1.2 Laptops and Smartphones Lead Traffic Growth

This creates a problem where 3G networks are overloaded especially in metropolitan areas or areas where there is a large concentration of smart phones in use and hence degrading the experience of the smartphone users. Mobile operators on the other hand are also not recouping anywhere near as much in revenue as they are being forced to provide in network capacity. Though operators continue to roll out higher bandwidth e.g. Long Term Evolution (LTE), increasing speed may not always be cost effective; and even with 4G the bandwidth may not be enough. The typical user given higher bandwidth tends to do even more downloads and uploads. Cisco Visual Networking Index shows that In 2011, a fourth-generation (4G) connection generated 28 times more traffic on average than a non-4G connection. Although 4G connections represent only 0.2 percent of mobile connections today, they already account for 6 percent of mobile data traffic.

Offloading the data via small cell technologies like Wi-Fi and Femtocell technologies have been considered as an immediate resolution for the problem. Thus in this paper I propose the use of Wifi caching to offload mobile data. The Wifi in this case is a cellular operator assisted solution.

1.2. Problem Statement

The main problems facing 3G networks today are:

1) The Increasing network traffic – There is a tremendous surge in network traffic. Rich data services and bandwidth usage offered by 3G is not anticipated to decrease V. Gupta and M. Rohil, (2012). According to Bulut & Szymanski (2012) with the mobile data explosion

already being experienced, it is inevitable that in the near future cellular networks will be overloaded and congested.

2) The Increased use of the smart phone for Internet – According to Cisco, (2012) in the year 2012 the average use of smart phones rose to 81 percent. Lee et al (2010) observes that the main reason behind this growth is the increase in smart mobile devices that offer ubiquitous internet access and varied multimedia authoring and playback capabilities. Dimatteo, Hui, Han & Li (2011) also observes that mobile networks are presently encountering the challenges of sudden increase in data from mobiles. It is of the essence that freshly designed architectures be designed to deal with large amounts of mobile data. Though operators continue to roll out higher bandwidth e.g. Long Term Evolution (LTE), increasing speed may not always be cost effective; and even with 4G the bandwidth may not be enough.
3) The Spectrum is costly and scarce – V. Gupta & M. Rohil, (2012) observes that Systems for telecommunication all need a particular amount of electromagnetic bandwidth to run. As the world becomes increasingly wireless, it becomes increasingly difficult to distribute available spectrum to various technologies. Wireless equipment manufacturers want additional bandwidth to enable them sell more equipment. Because of this shortage, the institutions allocating the spectrum to buyers sell to them at very high prices.

Some of the solutions to the above problems are:

1. Scaling: This involves building additional base stations and/or increasing the backhaul capacity. This requires more infrastructure and thus more investment, which directly results in high cost per MB. However research shows that the typical user given higher bandwidth tends to do even more downloads and uploads. Cisco Visual Networking Index reports that In 2011, a 4G connection generated 28 times more traffic on average than a non-4G connection.

Although 4G connections represent only 0.2 percent of mobile connections today, they already account for 6 percent of mobile data traffic.

2. Optimization: This refers to optimizing the radio and backhaul usage. “Starting from 1G to 3G many technological upgrades have happened for the optimized usage of radio but there is certainly a limit on the number of bits which can be packed onto radio waves” (V. Gupta and M. Rohil, 2012). Caching of frequently used data is one way of optimizing the backhaul use. Because it assists in control of flow it is a solution that is quite promising. However it poses a few shortcomings:

a) Intensive inspection of packets may cause the network to slow down.

b) Users do not like to be policed due to invasion of privacy (V. Gupta and M. Rohil, 2012)

3. Mobile Data Offloading (MDO): Also called data offloading is the use of complementary network technologies for delivering data originally targeted for cellular networks. Rules triggering the mobile offloading action can either be set by an end-user (mobile subscriber) or an operator. The benefit of mobile data offloading for the end users is getting service cost control and getting readily available high bandwidth. The operators main aim for offloading on the other hand is to ease congestion of their mobile networks. MDO has emerged as a promising solution. Offloading helps in reducing the cost of telecom operators as a good amount of their revenue is used to pay for the exclusive use of spectrum, without the services being affected.

Primary offloading technologies in use by the industry include Wi-fi and femtocells.

Thus in this paper I am proposing the use of Wifi caching to offload mobile data. The Wifi in this case is a cellular operator assisted solution.

1.3. Aims/Purpose

The main aim of this study is to propose the use of Wi-fi caching to improve performance of 3G mobile data offloading.

1.4. Specific objectives

To improve performance of 3G mobile data offloading through Wi-Fi caching by:

a) Identifying a caching scheme model
b) Defining the parameters to be measured
    i) Network throughput
    ii) Cache hit ratio
    iii) User request latency
c) Designing the caching scheme model
d) Implementing the caching scheme model using simulation
e) Testing the cache scheme model for improved performance

1.5. Justification of Project

1.5.1. Significance

Traffic on the mobile network is increasing rapidly; networks must be managed efficiently by the service providers to match customer demands. According to Cisco (2012) the technology evolution of radio access networks is limited by the laws of physics, and significant growth in radio frequency (RF) efficiency can no longer be expected. LTE radio access is reaching the limits of Shannon's law, the spectrum available for mobile data applications is limited.

Accuris Networks (2010) Observe that Network Operators who provide a flawless mobile Wi-Fi offload experience for their consumers will create a substantial opportunity to save costs by diverting low value, high volume data traffic to an alternative access network and by creating a powerful, mass market engine that differentiates their offers and drives top-line growth.

By use of Caching data in the Wi-fi data access, performance and availability are largely improved and load on cellular networks is reduced. Addressing the problem will be a win- win situation for both the cellular network providers and the mobile phone users.

Other key application areas that will benefit from these is in high speed mobility i.e. trains and buses and VANETs (Vehicular Ad-hoc Networks).

1.5.2. Scope and Limitations

The scope of this study was to develop a Wi-fi caching scheme model for 3G mobile data offload. The study has used simulation to test the improvement of 3G mobile data offload performance using Wi-Fi data caching. In this paper the issues of security of caching using Wi-fi are not addressed. Similarly issues of policy handling and how charging would be controlled or setup are not discussed in this paper.

1.6. Assumptions of the study

This study assumes that the solution is an operator assisted solution, thus Wi-fi's are linked to a particular operator.

CHAPTER 2: LITERATURE REVIEW

2.1 State of the Art: Introduction to Mobile Data offloading

Mobile data offloading, also called data offloading is the use of complementary network technologies for delivering data originally targeted for cellular networks Wikipedia, 2013. The requirement for offloading mobile data has been brought about by the sudden increase of internet data traffic particularly data going through the mobile networks coming from smart phones. According to forecast reports by Cisco average smartphone usage nearly tripled in 2011. The average amount of traffic per smartphone in 2011 was 150 MB per month, up from 55 MB per month in 2010. Global mobile data traffic will increase 18-fold between 2011 and 2016. Mobile data traffic will grow at a compound annual growth rate (CAGR) of 78 percent from 2011 to 2016, reaching 10.8 exabytes per month by 2016. This usage is expected to keep on rising and as such then; the Mobile network providers to avoid congestion and poor quality of service have to look for alternative methods of dealing with this influx of data in their networks.

2.1.1 Offloading via Wi-fi

Offloading via Wi-fi is an economical and easily deployable solution that can assist operators deal with data traffic increase at minimized cost. Wi-fi operates over unlicensed spectrum and is easier to build than larger network deployments and upgrades D. Srinivsan et al (2012). Wi-Fi enables devices to electronically transmit information or connect to the internet wirelessly using radio waves. Wifi makes use of the unlicensed radio spectrums of 2.4Ghz and 5Ghz. Coverage for Wifi usually extends over a 50 to 150meters radius of the Access point, and speeds range from 11 Mbps to 54 Mbps but with new Wifi speed are expected to go higher to 100Mbps.

Figure 2.1 Smart phone, tablet, laptop and desktop connecting to a Wireless access point

Various Wifi IEEE standards exist and include, 802.11b, 802.11a, 802.11g, and 802.11n. The IEEE 802.11 specifications are wireless standards that specify an “over-the-air” interface between a wireless client and a base station or access point, as well as among wireless client.

The IEEE 802.11 deals with:

1. Type of modulation, Frequency, power levels i.e. The physical network layer
2. Software based protocols that enable equipment to communicate i.e.The media access control layer (MAC)

Standard	**802.11a**	**802.11b**	**802.11g**	**802.11n**
Frequency	5GHz	2.4 GHz	2.4 GHz	5GHz and/or 2.4GHz
Typical Data Rate	23 Mbit/s	4.5 Mbit/s	19 Mbit/s	74 Mbit/s
Max Data rate	54 Mbit/s	11 Mbit/s	54 Mbit/s	300 Mbit/s
Range	115 feet	115 feet	125 feet	230 feet
Modulation and	PSK, QAM,	PSK, CCK,SS	PSK, QAM,	QAM, Alamouti,

coding technique	OFDM		OFDM	OFDM

Table 2.1 Comparison of IEEE 802.11 standards

2.1.2 Offloading via femtocells

Femtocells are small, inexpensive, low-power base stations that are generally consumer deployed and connected to their own wired backhaul connection. In these respects, they resemble WiFi access points, but instead they utilize one or more commercial cellular standards and licensed spectrum. Andrews et al (2012). According to Andrews et al (2012) Perhaps the most significant and widely-discussed challenge for femtocell deployments is the possibility of stronger, less predictable, and more varied interference and the lack of a low delay connection to the core network which can result in significant handover signaling delays.

2.2 State of the Practices: Mobile Data offloading application areas

Some of the key areas where mobile data offloading will be of great benefit include in areas of:

- High speed mobility in trains and buses.

According to Paul Kapustka (2012) Availing internet on buses and trains is very appealing to a lot of people, as it bring better communication almost any time and additionally the idea of working as your travel becomes possible. Traditionally reading the newspaper was regarded as the only productive work that could be done while commuting. Just like adding Wi-fi in the schools and libraries, availing internet to buses and trains seems like the subsequent step to take in transport services.

From the business angle, there is a rising interest in the concept of sponsors advertising their products online to the commuting audience in transit.

According to Michael Kennedy (2012), an analysis of 3G smart phone use with WiFi offloading in South Korea sponsored by the U.S. National Science Foundation monitored the usage of 100 iPhone subscribers for 24 hours per day for 2.5 weeks. This research provides data on the timing and mix of service availability for urban subscribers including access to home WiFi, 3G mobile, WLAN at work and public WiFi in retail areas and transportation facilities. The study found that about 65 percent of the total traffic was offloaded using on-the-spot WiFi access. It found, further, that an additional 29 percent of the total data load would be offloaded if offload delays of at least an hour were permitted. From the subscriber's perspective, mobile offloading will go a long way in keeping 3G data usage below mobile operators' monthly data usage caps.

According to Kevin Fitchard (2011) “AT&T now has 460 hotspots in the NYC metro area that AT&T mobility can seamlessly access, potentially relieving the wider area cellular network of an enormous data load. That’s roughly 2% of AT&T’s total hotspots in the U.S. Most of those hotspots are in Starbucks coffee shops, other cafes and restaurants and hotels, but AT&T has been deploying outdoor Wi-Fi in key locations, using high-capacity long-range wireless technologies that greatly expand the reach and load of a typical access point. In addition to the public parts, Times Square is now blanketed with Wi-Fi (though it is only available to AT&T customers). AT&T has extended that hotzone program to other parts of the country, for instance covering San Francisco’s Embarcadero Center and Austin’s Sixth Street”.

2.3 Review of Technologies & Technological Capabilities

The general solutions for mobile traffic offloading are: Femtocells, Peer-to-Peer opportunistic approaches and Wi-fi.

Dimateo et al (2011) state that Femtocells are smaller cellular base stations, that connect to the service providers network using broadband; they are usually designed for using indoors. They use a similar spectrum that is licensed, like that for macrocells in the mobile networks and thus no additional hardware is required to use mobile phone. Femtocells allow service providers to expand their service coverage on the indoors or at the cell edge, mostly where access would otherwise be limited or unavailable. According to Kumar et al (2011) Femtocell transmission power is very small and these devices are usually incorporated to small plastic desktop and wall mounts cases that are powered from the consumers power sockets. Where femtocells are available mobile phone user connects to the femtocell and does not search for a base station that is external. A handoff occurs between the cells inside and the cells outside when the user moves. However some challenges with Femtocells are that there may be conflicts in spectrum between the major network and femtocells, and between femtocells that are next to each other. According to Gupta et al (2012) The other challenge with femtocells is their availability is not yet widespread and they have not yet matured. The demand for the femtocells is small and they still utilize the spectrum that is licensed; including FCC rules. Finally also deploying femtocells is complicated as compared to deploying of Wi-fi.

Peer-to-Peer opportunistic approaches have also been considered as an option for mobile data offloading especially since there is minimal monetary cost to it. Mobile Social Networks (MoSoNets) are exploited as opportunistic communications channels to facilitate the information dissemination and thus reduce the amount of cellular traffic. “Target-users help to further propagate the information among all the subscribed users through their social participation, when their mobile phones are within the transmission range of each other and can communicate opportunistically”.(B. Han, 2012)

Wi-Fi is a commonly used technology that enables an electronic devices to communicate wirelessly (using radio waves) over a computer network, including high-speed Internet connections. The Wi-Fi Alliance defines Wi-Fi as any "wireless local area network (WLAN) products that are based on the Institute of Electrical and Electronics Engineers' (IEEE) 802.11 standards". Wi-fi wireless networking technology uses radio waves to provide wireless high-speed Internet and network connections. It uses the unlicensed spectrum; and offers a much faster speed compared to the 3G service (Wikipedia, accessed 2013). According to Gupta (2012) Wi-Fi operates in unlicensed ISM 2.4 GHz and 5 GHz frequency bands. The spectrum availability in the two respective bands is 83 MHz and 505 MHz. This means that regulator approval is not required for individual deployments, and Wi-Fi has a larger “free” spectrum available to cater to any size of network deployment. The equipment cost for Wi-fi deployment is low and it is easy to deploy. Additionally Wifi has advanced security with WPA and WPA2 encryptions and also has support for quality of service (QoS). Wi-fi is in widespread use especially in metropolitan areas either rolled out by operators as business hotspots, shared as community network or installed by users for residential use.

2.4 Critique of Literature

Opportunistic peer-to-peer offloading has been proposed as a method of mobile data offloading. Han et al*., 2012* proposes using mobile network users to push the content by identifying MoSoNets. They use three algorithms for their problem and test the results through a simulation using traces. The algorithms are Random, heuristic and greedy. Their preliminary experimental outcome demonstrates that when they are briefly in contact cell phones can be able to swap around 1.48 Megabytes of data. In this scenario peers co-operate and share data.

While peer-to-peer offloading can be used by groups with similar interests, major challenge occurs in that mobile phone can only store data to a very limited capacity.

Various solutions of offloading via Wi-fi have been proposed. Some in quantitative studies and others in caching. Several quantitative studies using traces have been carried out by Lee et al, 2010, Dimatteo et al, 2011, Bulut & Szymanski, 2012. Lee et al.,2010 carried out a quantitative study which showed that there was a reduction in traffic by 65% achieved using on-the-spot offloading onto Wi-fi. They recruited 100 phone users from Metropolitan areas and collected statistics on their Wi-fi connectivity during a period of about two and a half weeks. Dimatteo et al., 2011 presents an architecture (MADNet) for the integration of Wifi networks and mobile-to-mobile Pocket switched networks with cellular networks. They look at transfer of files which are bulk and streaming of video over 3G. Real mobility data set of 500 taxis in urban areas study give the number of AP's needed for different requirement of QoS for data delivery in metropolitan areas. They show that only few hundreds of Wi-fi APs deployed in an area of 313.83km^2 can offload half of the mobile data from the 3G network in their scenario for both downloads and uploads. Bulut and Szymanski, 2012 studied Wi-fi Access points deployment in a metropolitan area. They analyze a large scale real user mobility traces and propose a deployment algorithm based on the density of user data request frequency. They measure how much offloading can be done with different no. of APs. In the three above studies concentration was given in trying to evaluate the optimal quantities of APs required in different scenarios. The issue of caching of data to offload data is not addressed.

Balasubramanian et al, 2012 designed a system called Wiffler to augment mobile 3G capacity using Wi-fi. It uses leveraging delay tolerance and fast switching in vehicular networks. Wiffler uses a simple model of the environment to predict WiFi connectivity. It uses these predictions to delays transfers to offload more data on WiFi, but only if delaying reduces 3G usage and the transfers can be completed within the application's tolerance threshold. Small scale analysis is used. Wiffler considerably leads to a reduction in 3G usage. For a realistic workload, the reduction is 45% for a delay tolerance of 60 seconds.

Some studies have been carried out in the area of caching by Goemans et al, 2006 and Mashadi & Hui, 2012. Goemans et al., 2006 proposes an architectural and protocol framework that allows 3G service providers to offload content distribution to an ad-hoc network. Resident subscribers are provided incentives to cache popular data items while mobile users access this data from the resident subscribers through adhoc networks. However the main aim of the paper is not on the results of the effects of how mobile the nodes are but rather on the theoretical aspects of the game and user cooperation. Mashadi & Hui, 2010 propose an opportunistic proactive caching strategy which exploits available access points to proactively push contents to nodes through Wi-fi. They also incorporate the use of peer-to-peer sharing from their local cache. Using the push method can overload the mobile nodes unnecessarily a pull method would be preferable where a request is done when needed.

In this paper I propose the use of Wi-fi caching to offload mobile data. Mobile users requests for data are serviced by the nearest AP that has the data already cached. The Wi-fi in this case is a cellular operator assisted solution.

CHAPTER 3: METHODOLOGY APPROACH

To improve performance of 3G mobile data offloading through Wi-Fi caching by:

a) Identifying a caching scheme model

b) Defining the parameters to be measured

   i) Network throughput

   ii) Cache hit ratio

   iii) User request latency

c) Designing the caching scheme model

d) Implementing the caching scheme model using simulation

e) Testing the cache scheme model for improved performance

3.1 Current measures/methodologies used

Caching is the temporary storage of content, such as HTML pages, images, files and web objects, for the benefit of offloading the origin web or application server and reducing the time required to deliver content to a client. Web caching comprise of storing frequently requested objects on a caching server instead of the original server, so that web servers can make better use of network bandwidth, reduce the workload on servers and improver the response time for users. If an HTTP request is made in a cache networks ,the cache checks whether it has a local copy before forwarding a request to an original server *(Chan et al., 2009).* A cache hit is a request that is able to be satisfied from the data from the cache. A cache miss is a request that could not be satisfied by the data that is in the cache.

3.1.1 Types of Caching

Various methods are in use for caching, though most of them are targeted on to web proxies rather than mobile networks. According to Kumar et al the two main kinds of cache sharing techniques are:

1. Push based - In the push based cache sharing, when a node caches a new data item, it will proactively broadcast the caching updates to it's neighboring nodes. The neighbouring nodes update the caching information for the future use.
2. Pull based - In the pull based cache scheme when a node wants to access a data item, it broadcasts a request packet to all its neighbors' node.

3.1.2 Key Caching Aspects

Kumar et al (2011) states that some of the key issues to put into consideration when caching are:

1. A cache discovery algorithm that is able to discover and return items that have been requested.
2. A cache replacement algorithm that replaces old data that is cached with new ones when there is no space.
3. A cache consistency algorithm to ensure that the cached data items are up to date and not stale.

3.1.3 Caching Replacement Strategies

Podlipnig et al (2003) groups caching replacement strategies into 3 main categories:

1. Recency based - In recency based strategies time is the main factor used to decide which cached data should be discarded first. The most known strategy of this type is LRU or Least Recently Used, which removes the least recently referenced object

2. Frequency based - In frequency based replacement strategy the data least requested for is the one removed first. The most known of these strategies is LFU or Least Frequently Used.
3. Recency and Frequency based - Recency and frequency based take both the frequency and recency of use into account.

3.1.4 Methodologies in use for mobile caching

Various mobile caching methods have been implemented as below:

i. M. Akon et al (2012) use an optimal cache replacement policy for data access applications in wireless networks where data updates are injected from all the clients. Two cache access policies are used Proactive Access Policy (PAP) and Reactive Access Policy(RAP). An update-oriented replacement policy (URP) is used.
ii. T. Wang et al (2012) design a network model to study this in a setting with two characteristics: 1. delay tolerant; 2. lack of infrastructure. Mobile users generate requests and opportunistically download from other users they meet, via Bluetooth or WiFi. The difference in popularity of different web content induces a non-uniform request distribution, which is usually a Zipf's law distribution.
iii. J. Wang et al (2006) implement Caching popular web documents in wireless base stations to reduce network traffic between base stations and web servers, and reduce user web request latencies. A cache line migration (CLM) strategy is proposed to flow and replicate the heads of network cache lines of web documents associated with a moving mobile host to the new base station during the mobile host handoff. These replicated cache line heads provide direct links to the cached web documents accessed by the moving mobile hosts in the previous base station, thus improving the mobile web caching performance. (Wang et al, 2006)

iv. A. Mashhadi et al (2010) use an opportunistic proactive caching strategy which exploits available Access Point to proactively push contents to nodes through Wi-fi. They model access points as fixed nodes, monitoring network requests and downloading missing content from their gateways to the internet. The Access points then proactively push popular content to users in their Wi-fi range. Two basic selection approaches are used. Bulky selection – access points proactively push the bigger popular files to adjacent nodes. Random selection – any popular content can be pushed by access points and proactively cached by adjacent nodes.

3.2 Evaluation of methodology approaches

The proposed methods by M. Akon et al, T. Wang et al and J. Wang use cooperate caching methods where the mobile nodes have to cooperate within themselves. Caching is done on the mobile phone which again might be a limitation because of the available size.

The proposed method by A. Mashhadi et al uses a push based method. According to P. Kumar et al, 2010 Schemes which are push based improve the availability of data, but there is an overhead cost on communication. The disadvantage of the push based scheme is that an advertisement may become useless if no demand for the cached data items occurs in the vicinity. One more problem with the push based scheme is that caching information may not be longer used if the node moves out from the zone or due to the cache replacement. Also there is chance that the pushed content will not be used and duplication of the same content will be in the users. Thus a pull

Author	Pull data from (Access Point, Base Station, Server)	Cooperative	Push/Pull	Replacement Policy
M. Akon	Originating Server	✓	Pull	Recency and

				Frequency based
T. Wang	Originating Server	✓	Push	Frequency
J. Wang	Base Station	✓	Pull	Cache line migration
A. Mashhadi	Access Point	X	Push	Recency and Frequency based

Table 3.1 Comparison of different mobile caching methodologies

3.3 Proposed methodological approach

When a mobile node requires to download content, it sends a query to the base station presently in range. The base station in turn does a search for the nearest AP to the mobile node that has the content already cached. If found the AP feeds the mobile node with the content. If no AP has the required content the base station retrieves data from the original server and serves the mobile node with the content, which is also cached onto the nearest AP. Caching replacement techniques to be used in this method will be the LRU (Least Recently Used) based strategy. Also popular objects are distributed within the various APs. Thus when the cache is full then the least recently used object is deleted from the cache.

User request latency is measured from the time the user sends a request to the time they get a response. Network throughput which is the average rate of successful message delivery over a communication channel is measured. Efficiency is measured basically by network mean throughput divided by the data rate multiplied by 100.

3.4 Characteristics of proposed (tools, model, artefacts etc)

3.4.1 Simulation

Simulation is the reproduction of an operation in the real-world or system over time (Wikipedia, 2013). Simulation has various functions which include Prediction,Theory discovery, Perfomance, Training, Education and Proof. There are three different types of simulations which are System dynamics simulation, discreet event simulation and Agent-based simulation.

Simulation was chosen as the preferred method to test the model, as it would be very expensive and time consuming to evaluate on real-systems.

3.4.1 OMNeT++ Simulator

The selected simulator is the OMNeT++ as it had the required attributes. OMNeT++ simulator is a discrete event object oriented network simulator programmed by Andras Varga (University of Budapest) and its initial release was in 1992 (Varga, 2008). The OMNeT++ was written using C++ language and distributed as an open source program that can be used for non-profit and academic purposes for free. In order to use it commercially, a license for another version called OMNEST is needed. It supports various packages like MiXim, INET and INETmanet which enable networks both adhoc and wireless to be simulated.

The OMNET++ Programming model consists of simulated objects which are represented by modules. Modules can be either simple or compound. A network simulation in OMNET++ is a compound module comprising of other Compound modules. Simple modules are used to define algorithms and are active components of OMNeT++ in which events take place and the behaviour of the model is defined (generation of events, reaction on events) Sarkar et al,(2010) . Modules communicate by messages (sent directly or via gates). One module description consists of:

- (.NED file) which are Interface description
- (C++ class) which are Behavior description
- Modules
- Links and Gates which can be created statically at the beginning of the simulation (NED file) or dynamically during the simulation.

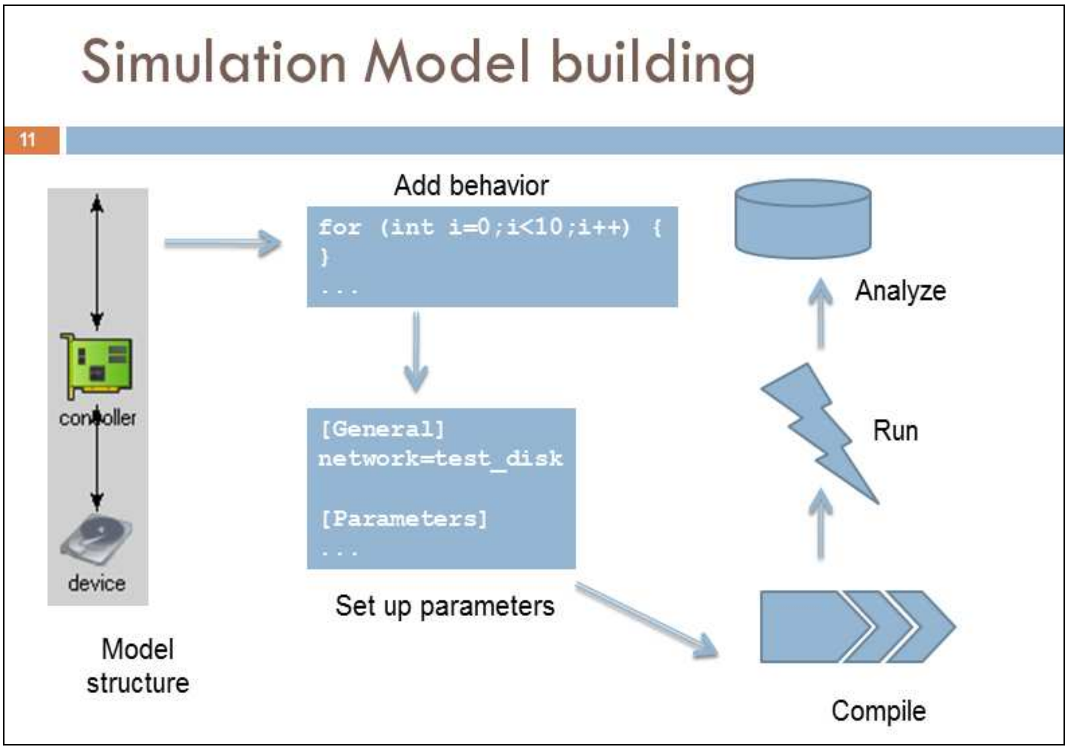

Figure 3.1 Simulation model building using OMNET++

3.4.2 Comparison with other Simulators

While some simulators like OPNET (Optimized Network Engineering Tool) were considered for use; the Academic version does not provide mobility and telecommunication features. Another simulator ns2 does not have a graphical presentation of simulation output data, and it is not user friendly because of it's text-based interface. OMNet++ has a good interactive interface for users that is friendly, and according to Weingartner et al, 2009 who carried out experiments to compare performance of five networking simulators that is ns2, ns-3, OMNetimp++, jiST/SWANS and Simpy. OMNeT++ is a good choice if scalability is a main concern. OMNeT++ also has the advantage of being able to run on most commonly used operating systems like Microsoft Windows, Linux, UNIX and Macintosh.

CHAPTER 4: CONCEPTUAL DESIGN

4.1 Conceptual Design

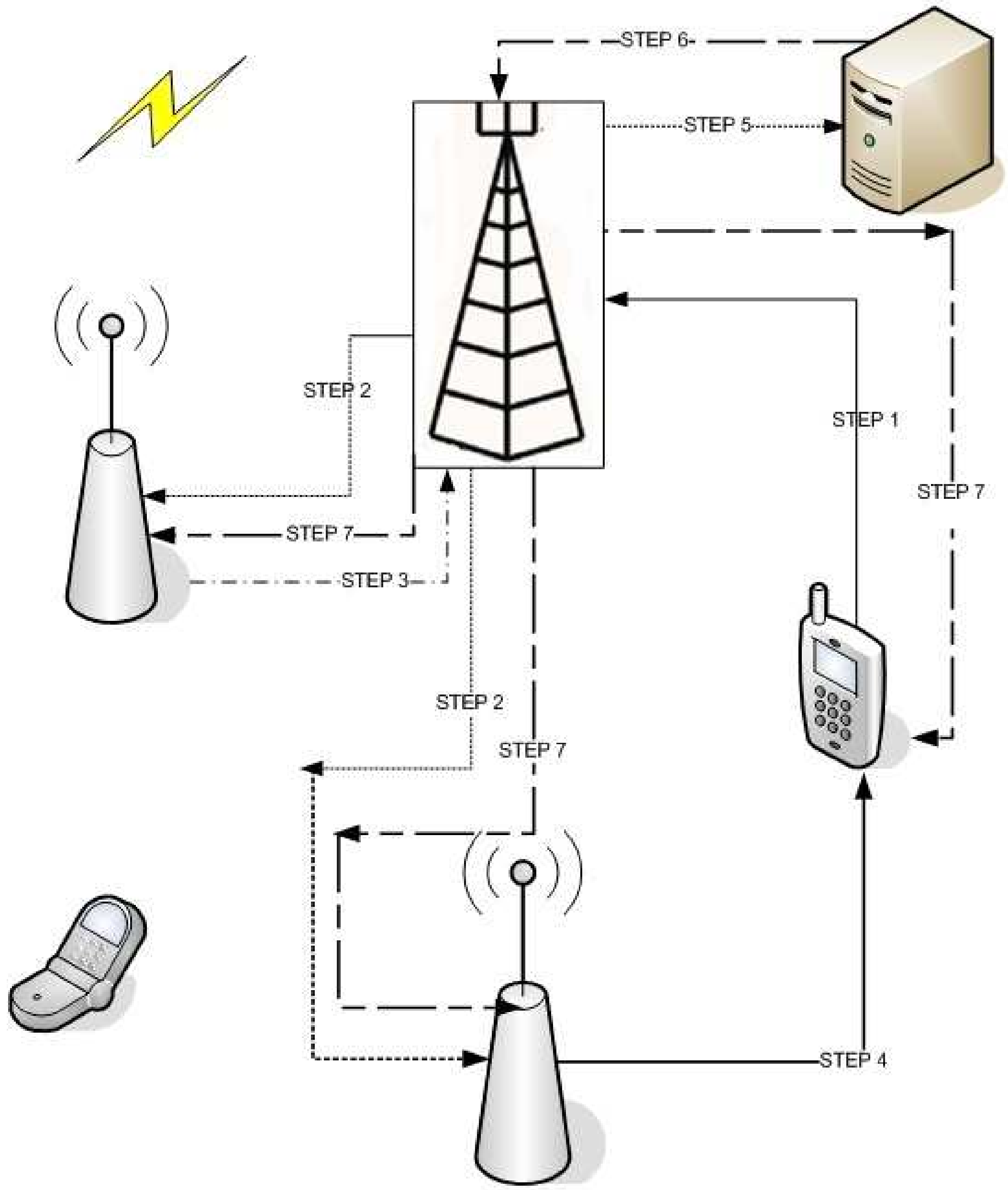

Figure 4.1 The conceptual design

STEP 1: Mobile phone requests data from the Base Station

STEP 2: Base station ask for cache data from nearest Access Points

STEP 3: Access Points gives response to the Base Station

STEP 4: Access Point with cached data sends data to Mobile phone

STEP 5: If access points which are near do not have the data required then the base station request for the data from original server, the data is then cached onto the APs and the mobile phone receives the data as well from the base station.

STEP 6: Originating server feeds Base Station with the data requested

STEP 7: Base Station feeds the mobile with the requested data, similarly the Base Station feeds the nearest Access Points with the same data.

When a Mobile phone user requests for internet data, the base station will send a request to the nearest Access points for cached data as requested by the mobile phone user. The nearest Access points will then give a response to the base station confirming if they have the requested data cached or not. If one of the Access points has the cached data then that Access point will send data to the mobile phone user. In a scenario where more than one Access point has the data available the first one to give response to the base station is allocated to give the feedback to the user.

If none of the Access points has the data cached then the base station requests data from the originating server. The originating server feeds the data to the base station; the data is then cached onto the nearest Access points and the mobile phone user receives the data as well from the base station.

During the caching on the Access points Least recently used strategy is used to decide which cached data should be discarded first; basically Least Recently Used (LRU), which removes the least recently referenced object

4.2 Field studies

The Omnet++ simulator with Inet and MiXim frameworks are used to design and test the model. Number of mobile nodes are varied together with different no. of Access points.

4.3 Data/ Input Data (expected outputs)

The key inputs to be measured will be:

- The Network throughput which is the average rate of successful message delivered over a communication channel.
- Cache hit ratio which is a measure of how often a searched-for item is actually found in the cache.
- User request latency which is the time the user sends a request to the time they get a response.

Data Collection is done by:

- Specifying during the simulation that the result of the simulations that are run are recorded as scalar values, vector values or histograms.
- An analysis file(*.anf) with raw data is created.
- The relevant information is extracted and then analysed using statistical methods to draw a conclusion.

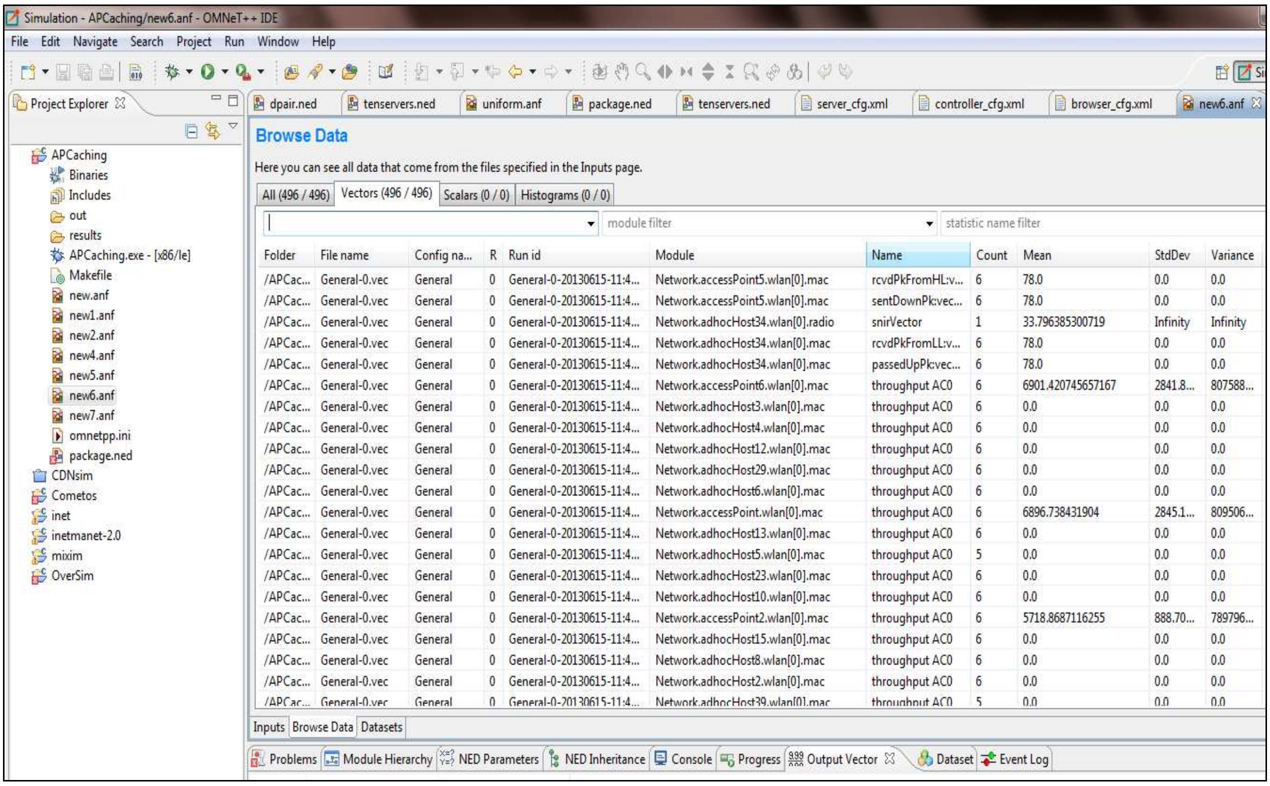

Figure 4.2 Sample raw output data from simulation

CHAPTER 5: IMPLEMENTATION AND RECOMMENDATIONS

5.1 Implementation

A laptop with intel atom is used to install the OMNet++ simulator. Various frameworks are then added namely Inet and InetMANet as they contain the modules for wireless and mobile simulation. OMNet++ is made up of simple modules and compound modules. Simple modules represent the active components of OMNeT++ where events occur and model behaviors are defined. Simple modules are implemented in C++. Compound modules comprise of simple modules or other compound modules; see figure 5.1 An AP is a compound module comprising of several simple modules. A network description files (.ned) file is then created to setup the network required. See below figure 5.2 package .ned with various adhoc hosts, access point and channel control.

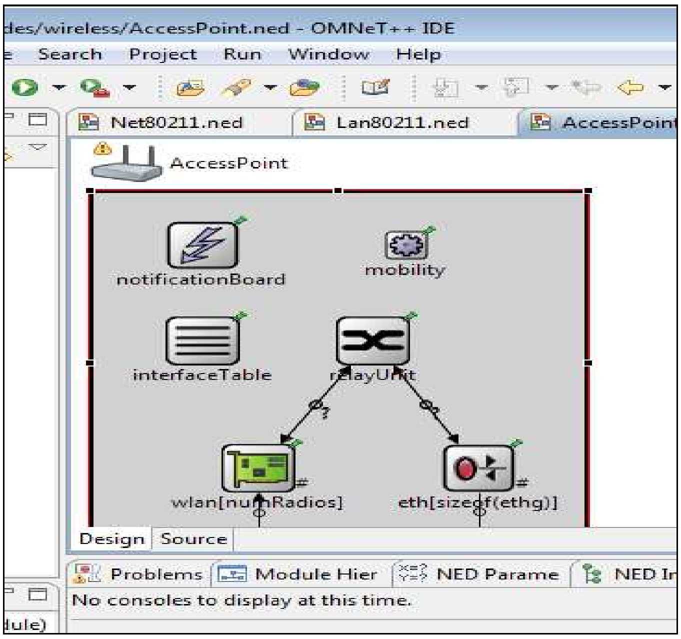

Figure 5.1 An Access point compound module comprising of several simple modules

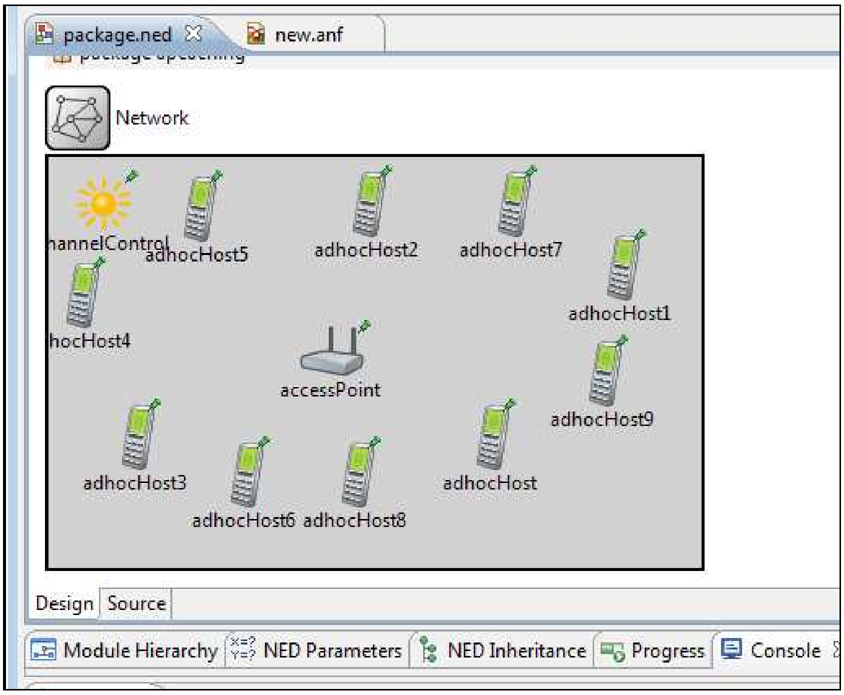

Figure 5.2 Omnet project window displaying 10 hosts and 1 access point

To configure how the network is required to run, parameters are specified in the .ned file and in the omnet.ini configuration file.

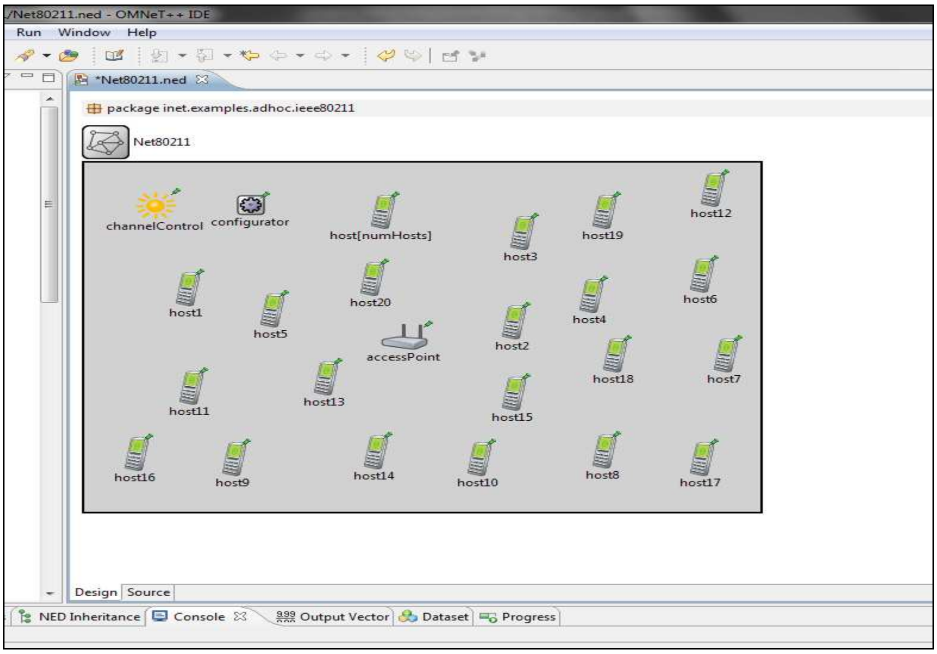

Figure 5.3 Omnet project window displaying 20 hosts and 1 access point

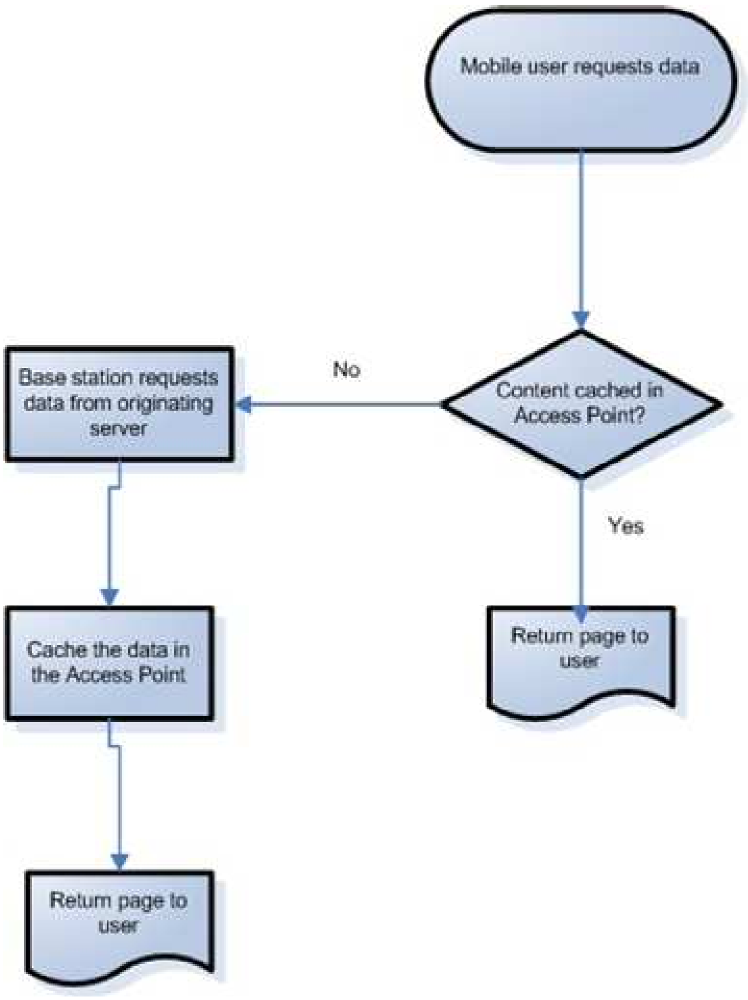

Fig 5.4 Implementation model

5.2 Testing (Evaluate Framework)

OMNeT provides two environments for simulation Tkenv (GUI) (Fig 5.7) and Cmdenv (non GUI) environment. Simulation modules are setup as per figure 5.2. The configuration file omnet.ini is then checked for correct settings.To run the simulation the run button is pressed and the simulation runs in the Tkenv window. Each simulation experiment was run for approximately 15 minutes. The number of mobile hosts were varied and the results of the run done. Simulation parameter settings are as below figure 5.6. Results of the simulation are then either stored as a vector file or a scalar file. Several simulations are run with varying no. of hosts i.e. 3 , 5, 7, 10, 20, 30

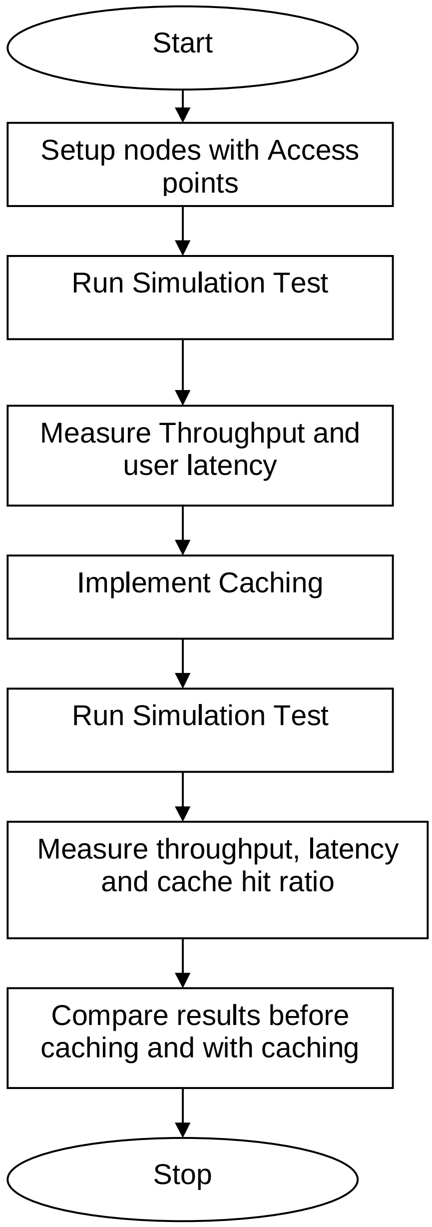

Figure 5.5 Implementation Steps

```
**.wlan[*].radio.snirThreshold = 4dB



[General]
network = Net80211
#record-eventlog = true
#eventlog-message-detail-pattern = *:(not declaredOn(cMessage) and not declaredOn(cNamedObject) and not declaredOn(cObject))
num-rngs = 3
**.mobility.rng-0 = 1
**.wlan[*].mac.rng-0 = 2
#debug-on-errors = true
tkenv-plugin-path = ../../../etc/plugins
**.constraintAreaMinX = 0m
**.constraintAreaMinY = 0m
**.constraintAreaMinZ = 0m
**.constraintAreaMaxX = 600m
**.constraintAreaMaxY = 400m
**.constraintAreaMaxZ = 0m
**.debug = true
**.coreDebug = false
**.host*.**.channelNumber = 0
# channel physical parameters
*.channelControl.carrierFrequency = 2.4GHz
*.channelControl.pMax = 2.0mW
*.channelControl.sat = -110dBm
*.channelControl.alpha = 2
*.channelControl.numChannels = 1
# mobility
**.host*.mobilityType = "MassMobility"
**.host*.mobility.initFromDisplayString = false
**.host*.mobility.changeInterval = truncnormal(2s, 0.5s)
**.host*.mobility.changeAngleBy = normal(0deg, 30deg)
**.host*.mobility.speed = truncnormal(20mps, 8mps)
**.host*.mobility.updateInterval = 100ms
# tcp apps
**.host[0].numTcpApps = 1
**.host[0].tcpApp[*].typename = "TCPSinkApp"
**.host[0].tcpApp[0].localPort = 1000
**.host*.numTcpApps = 1
**.host*.tcpApp[*].typename = "TCPSessionApp"  # ftp
**.host*.tcpApp[0].active = true
**.host*.tcpApp[0].connectAddress = "host[0]"
**.host*.tcpApp[0].connectPort = 1000
**.host*.tcpApp[0].tOpen = 0
**.host*.tcpApp[0].tSend = 0
**.host*.tcpApp[0].sendBytes = 100KiB
**.host*.tcpApp[0].tClose = 0
**.tcpApp[*].localAddress = ""
**.tcpApp[*].localPort = -1
**.tcpApp[*].sendScript = ""
**.cacheTimeout = 120s
**.queryInterval = 120s
**.queryResponseInterval = 120s
**.wlan[*].bitrate = 54Mbps
**.wlan[*].mgmt.frameCapacity = 10
**.wlan[*].mac.address = "auto"
**.wlan[*].mac.maxQueueSize = 14
**.wlan[*].mac.rtsThresholdBytes = 3000B
**.wlan[*].mac.retryLimit = 7
**.wlan[*].mac.cwMinData = 7
**.wlan[*].mac.cwMinBroadcast = 31
**.wlan[*].radio.transmitterPower = 2mW
**.wlan[*].radio.thermalNoise = -110dBm
**.wlan[*].radio.sensitivity = -85dBm
**.wlan[*].radio.pathLossAlpha = 2
```

Figure 5.6 Simulation parameter setting

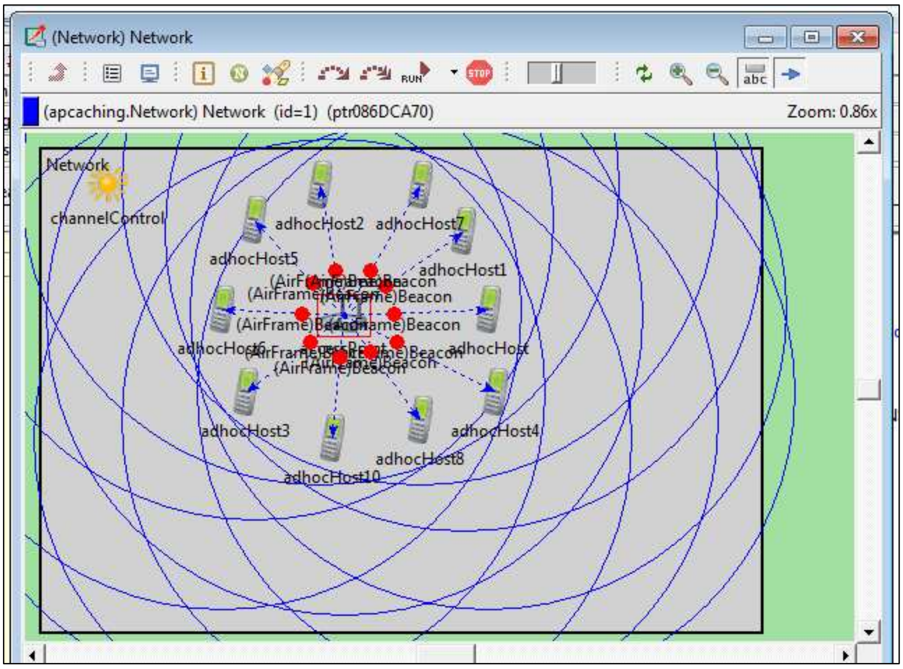

Figure 5.7 OMNeT++ gned

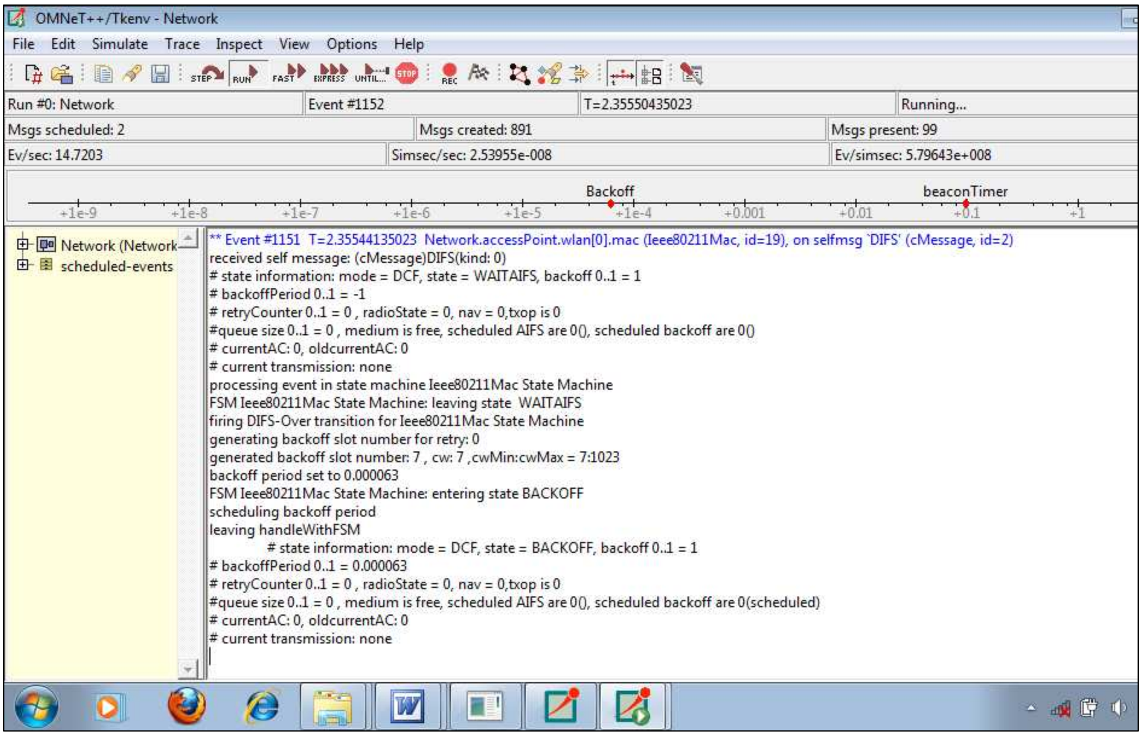

Figure 5.8 OMNeT++ Tkenv

After the simulations are run data is collected from the vector files, and then it is analyzed. From the experiments done it can be seen that the throughput is best when the number of hosts are about 3 and 10. Further increase of hosts above 10 hosts then the throughput and efficiency begin to go down. The results are in line with the work of research done by (Pelleta E., 2004).

No. of Hosts	**Throughput Mbps**	**Efficiency %**	**Request Latency**
3	6.79	12.57	0.0678
5	6.52	12.07	0.0677
7	6.57	12.17	0.0712
10	6.81	12.61	0.0692
20	6.54	12.11	0.0722
30	6.33	11.72	0.0785

Table 5.1 Throughput measurement, efficiency and request latency on varying no. of hosts

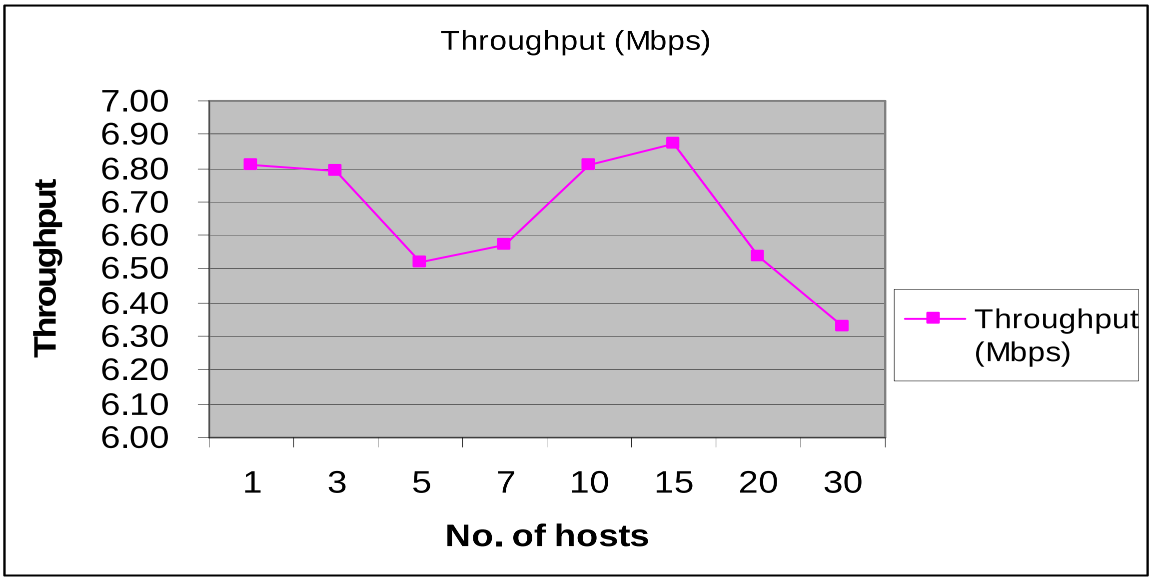

Figure 5.9 Impact of changing no. of hosts on throughput.

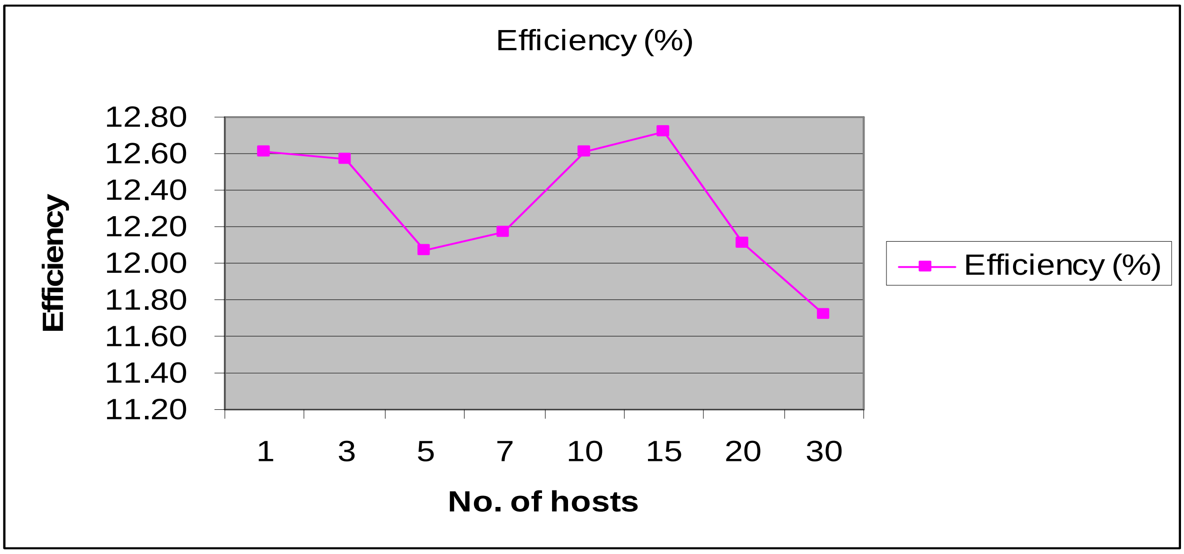

Figure 5.10 Impact of varying no. of hosts on efficiency

Further testing was then done with the implementation of caching on the access point. The results are as below. Access point was test with varying no. of hosts. The cached data in all scenarios is 1024kb. After caching the throughput is highest when the no. of hosts are 10. Cache hit ratio is highest when the hosts are around 7. Request latency is also best when the hosts are 10.

No. of Hosts	**Throughput Mbps**	**Request latency**	**Cache hit ratio**
3	4.277	0.0899	0.82
5	4.280	0.0896	0.911
7	4.276	0.0898	0.922
10	4.288	0.0890	0.876
20	4.268	0.0904	0.83
30	4.283	0.0894	0.773

Table 5.2 Throughput, Request Latency and Cache hit ratio on varying no. of hosts with cache

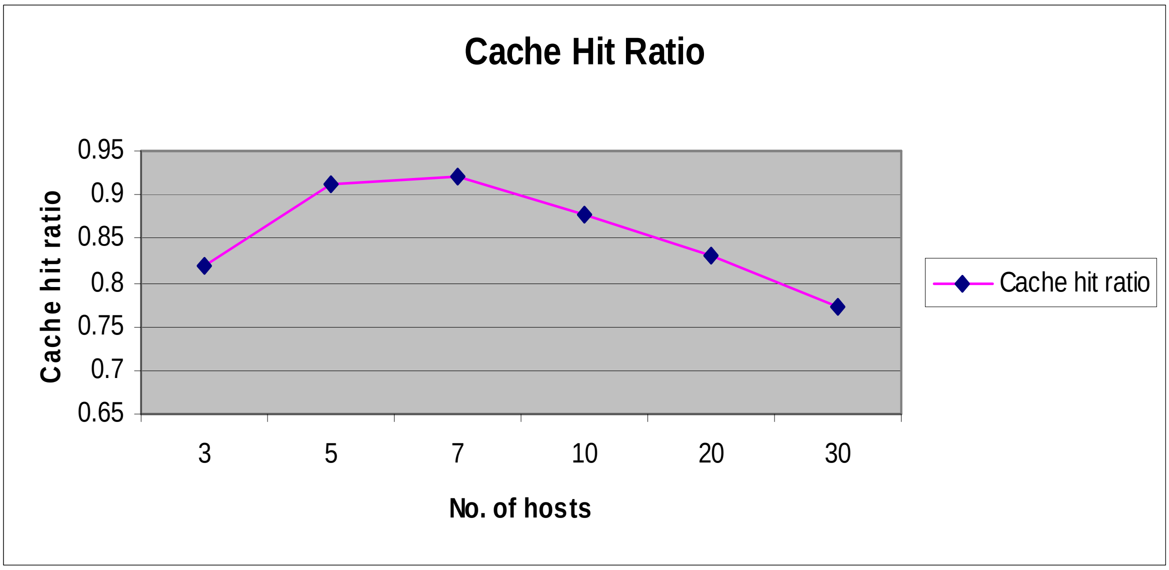

Figure 5.11 Cache hit ratio on varying no. of hosts

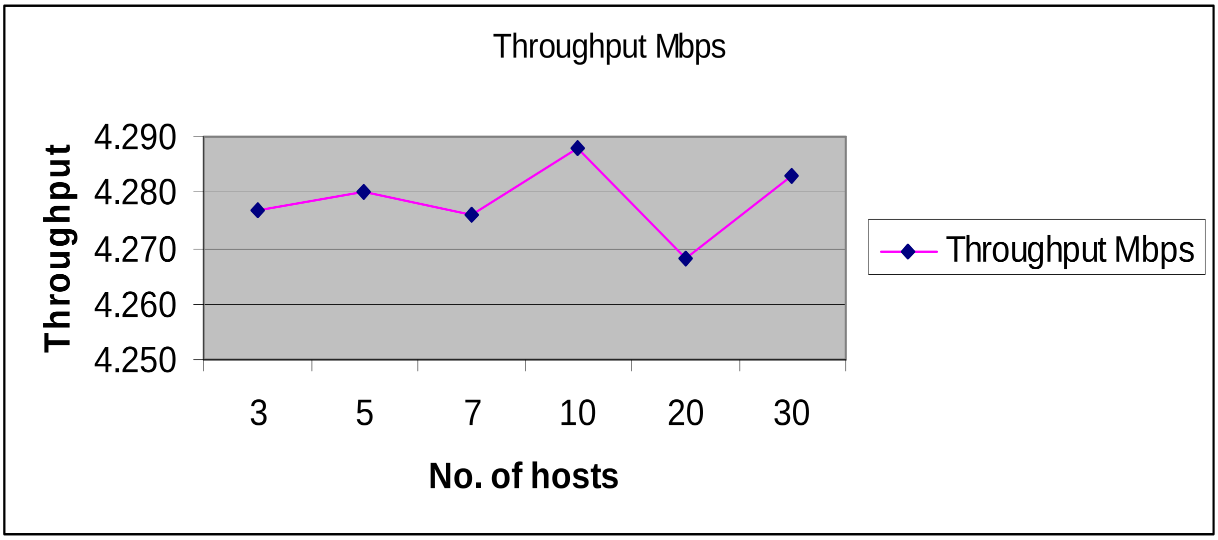

Figure 5.12 Throughput measurement on varying no. of hosts

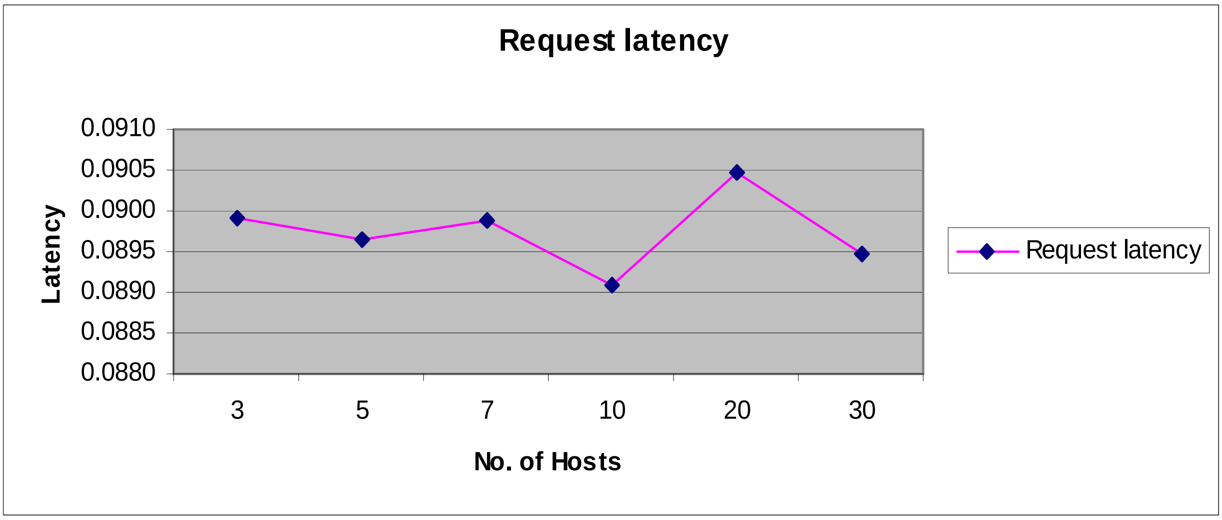

Figure 5.13 Request latency on varying no. of hosts

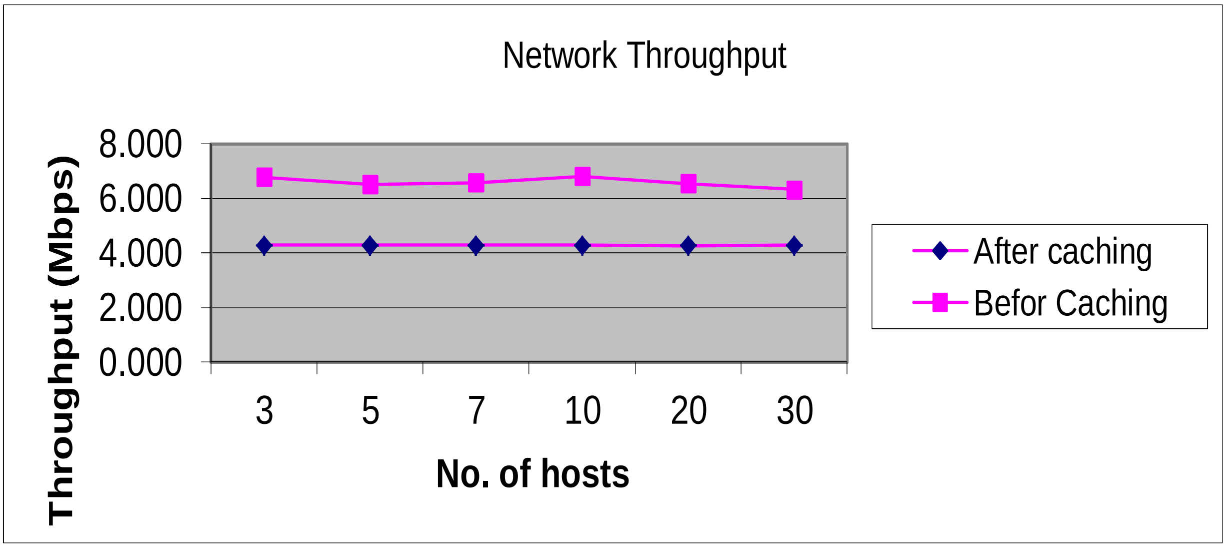

Figure 5.14 Comparison of Network Throughput before and after caching

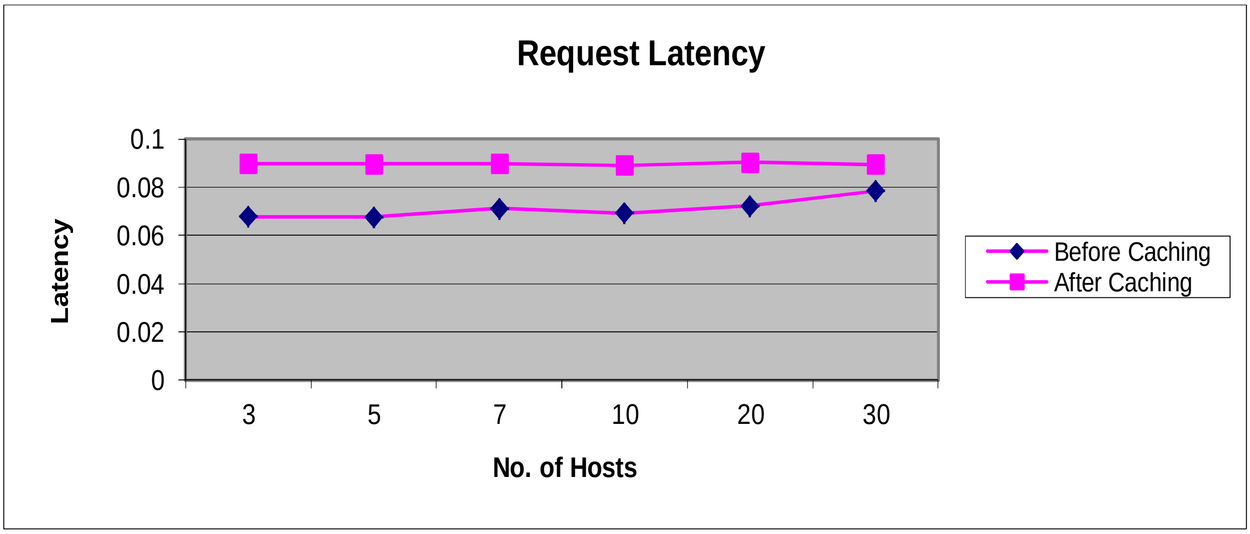

Figure 5.15 Comparison of request latency before and after caching

CHAPTER 6: DISCUSSIONS OF FINDINGS

6.1 Discussions of Findings

From the experiments done it can be seen that the throughput is best when the number of hosts are about 3 and also when they are 10. Further increase of hosts above 10 hosts then the throughput and efficiency begin to go down. Similar trend is seen even with caching implemented highest throughput is achieved when the hosts are about 10 after that there's a downward trend. However when caching is implemented the general throughput is lower than when caching is not implemented which is a concern, and was not the expected result. Similarly the user request latency is higher after the caching has been implemented.

When it comes to the cache hit ratio best hits are when hosts are between 5-8 and then from there; there is a drop in the hit ratio.

6.2 Conclusions

3G Cellular networks will continue to experience capacity constraints despite implementing new technology like LTE. AP Caching could be used to offload the Wifi data enhancing the user experience with faster downloads. Possibly with different placement and replacement algorithms in the cache more hits can be enhanced compared to the misses; in additional better throughput can also be attained.

While deploying AP's will go along way in offloading mobile data, implementing of caching as part of the solution will be of great benefit to both the service providers and the users. Caching will minimize network latency since the cache is located closer to users, and all requests need not traverse through the network.

6.3 Critical review and reflections

Various challenges were faced in using the simulation tool, as it has numerous functions and learning it was a challenge though beneficial in the long run.

6.4 Future

The Access Point caching model can be tested using different caching algorithms to see which one would give best results or better results. Caching of different sizes of data can be done to give a more realistic approach.

References

1. V . Siris and D. Kalyvas, "Enhancing Mobile Data Offloading with Mobility Prediction and Prefetching", MobiArch '12 Proceedings of the seventh ACM international workshop on Mobility in the evolving internet architecture Pages 17-22
2. S. Dimatteo, P. Hui, B. Han and V. Li, "Cellular Traffic Offloading through Wifi Networks", In proceeding of: IEEE 8th International Conference on Mobile Adhoc and Sensor Systems, MASS 2011, Valencia, Spain, October 17-22, 2011
3. http://en.wikipedia.org/wiki/Mobile_data_offloading, accessed on 14/01/2013
4. B. Han, P. Hui, V. Kumar, M.V. Marathe, G. Pei, J. Shao and A. Srinivasan, "Data Offloading through Opportunistic Communications and Social Participation", IEEE Transactions On Mobile Computing, Vol. 11, No. 5, May 2012
5. V. Gupta and M. Rohil, "Enhancing wi-fi with IEEE 802.11u for Mobile data offloading", International Journal of Mobile Network Communications & Telematics (IJMNCT) Vol.2, No.4, August 2012
6. K. Lee, I. Rhee, J. Lee, S.Cong and Y. Yi, "Mobile Data Offloading: How Much Can WiFi Deliver?", Proceedings of the ACM SIGCOMM 2010 conference Pages 425-426
7. A.Balasubramanian, R. Mahajan and A.Ventakataramani, "Augmenting Mobile 3G Using WiFi", MobiSys '10 Proceedings of the 8th international conference on Mobile systems, applications, and services Pages 209-222 , June 2012
8. E. Bulut and B. Szymanski, "WiFi Access Point Deployment for Efficient Mobile Data Offloading", Proc. ACM Workshop PINGEN at mobicom 12, Istanbul, Turkey, August 26, 2012
9. M. Goemans, L. Li, V. Mirrokni and M. Thottan, "Market Sharing Games Applied to Content Distribution in Ad-Hoc Networks", IEEE Journal on Selected Areas in Communications, Volume:24,I ssue:5,P ages 1020 - 1033, May 2006

10. E. Weingartner, H. Lehn and K. Wehrle, “A perfomance comparison of recent network simulators”, Communications, ICC '09 IEEE International Conference, pages 1-5, June 2009
11. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011–2016
12. Cisco White paper: “Architecture for Mobile Data Offload over Wi-Fi Access Networks”, 2012
13. A. Kumar, Divya, T. Singh, Dr. S. Khurmi**, “**Femtocell – A Review”, IJCST VOL. 2, ISSUE 3, September 2011
14. http://en.wikipedia.org/wiki/Wi-Fi, accessed on 14/01/2013
15. A. Mashhadi, P. Hui, “Proactive Caching for Hybrid Urban mobile Networks”, UCL Department of Computer Science, Research note RN/.10./05, October 2010
16. S. Podlipnig and L. Boszormenyi, “A survey of web cache replacement strategies”, ACM Computing Surveys, p. 398, 2003
17. P. Kumar, N. Chauhan, LK Awasthi and N. Chand, “Proactive approach for Cooperative caching in Mobile Adhoc Networks”, International Journal of computer Science issues, Vol. 7 issue 3 no. 8, May 2010
18. N. Sarkar and S. Halim, “A review of Simulation of telecommunication Networks: Simulators, Classification, comparison, Methodologies, and Recommendations”, ch.3.6, 2010
19. E. Pelletta, “Maximum Throughput of IEEE 802.11 Access Points: Test Procedure and Measurements”, August 2004
20. O. H. Sabrie and R. Salleh, “Enhancing Handoff Process using IAPP with Caching Techniques”, Journal of Applied Sciences, volume 9 (22), 2009
21. Chan, E., W. Li and D. Chen, (2009). Energy saving strategies for cooperative cache replacement in mobile ad hoc networks. Pervasive Mobile Comput., 5: 77-92.

22. M. Akon, M. T. Islam, X. Shen and A. Singh, “A bandwidth and effective hit optimal cache scheme for wireless data access networks with client injected updates”, Computer Networks 56, pp. 2080–2095, 2012

23. D. Srinivasan, J Dey, R.Mukherjee, .S Kumar, “Data Offload Approaches For Mobile Operators Improving Network Efficiency And Strengthening Quality Of Service”, 2012

24. http://www.fiercetelecom.com/story/making-business-case-mobile-data-offloading-service-no-downside/2012-01-23, Accessed 19/08/2013

25. http://connectedplanetonline.com/3g4g/news/att-now-has-460-mobile-data-offload-points-in-nyc-0610/, Accessed 19/08/2013

26. J. Wang, Z.Du and P.Srimani, “Cooperative Proxy Caching For Wireless Base Stations”, Mobile Information Systems, Volume 3 Issue 1, January 2007

27. T. Wang and P Hui, “Cooperative Caching based on File Popularity Ranking in Delay Tolerant Networks”, Proc. Of ExtremeCom , 2012

28. Jeffrey G. Andrews, Holger Claussen, Mischa Dohler, Sundeep Rangan and Mark C. Reed, “Femtocells: Past, Present, and Future”, IEEE Journal on selected areas in communication, Vol. 30, No. 3, April 2012