SELF SIMILARITY AND QUEUING ANALYSIS OF LTE SYSTEMS

A Project Report Submitted

in Partial Fulfilment of the Requirements

for the Degree of

BACHELOR OF TECHNOLOGY

in

Mathematics and Computing

by

Prem Sujan Kotta (170123027)

and

Garvit Mehta (170123018)

to the

DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI - 781039, INDIA

April 2021

CERTIFICATE

This is to certify that the work contained in this project report entitled "SELF SIMILARITY AND QUEUING ANALYSIS OF LTE SYSTEMS" submitted by Prem Sujan Kotta (170123027) and Garvit Mehta (170123018) to the Department of Mathematics, Indian Institute of Technology Guwahati towards partial requirement of Bachelor of Technology in Mathematics and Computing has been carried out by them under my supervision.

It is also certified that this report is a survey work based on the references in the bibliography.

Turnitin Similarity: 21 %

Guwahati - 781 039 (Prof. Selvaraju N.)

April 2021 Project Supervisor

ABSTRACT

The main aim of the project is to analyse self-similar property of LTE network traffic to understand burstiness of the network. We also provide a Queuing Model in the context of LTE Networks and establish various performance metrics for them. We then calculate the hurst parameter for traffic accumulated at several base stations. We will also simulate M/M/1 queuing system using NS3 and derive performance metrics experimentally.

Contents

List of Figures

Chapter 1

Introduction

1.1 LTE systems

LTE stands for Long-Term Evolution and it is a standard wireless data communication technology for mobile devices and data terminals. The main aim of LTE is to facilitate low response time and high throughput supporting flexible bandwidth deployments.

The main components of LTE architecture are as follows:

UE: User Equipment represent the terminal devices which deal with modules like data communication and terminating data channels.

E-UTRAN: Evolved UMTS Terrestrial Radio Access Network are evolved base stations(eNB) which send and receive radio transmissions from the user equipment using the digital and analogue signal processing.

EPC: Evolved Packet Core deals with network management and gateway servicing modules.

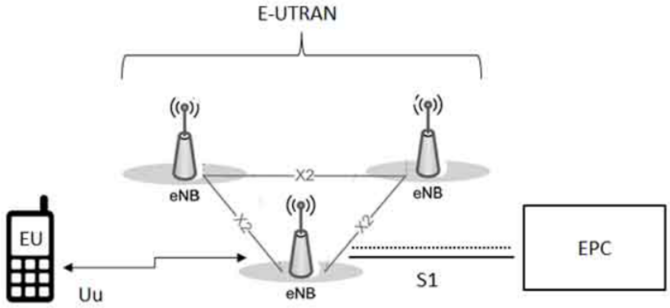


Figure 1.1: High-level architecture of LTE systems

1.2 Queuing systems

Queuing systems deal with the mathematical study of congestion. Generally, the study of servicing requests from customers/entities arriving in a queue fashion at a facility is called queuing theory. Examples: Ticket booking, customer service centers.

1.3 Importance of queuing theory in LTE systems

With growing modern internet applications and mobile devices, there is a spike in the usage of cellular networks. So it has become very important to investigate and analyze the performance of LTE networks.

Here eNB acts as a server and requests from wireless devices arrive at eNB. So we can treat the LTE network as a queuing system. The terminal wireless devices use protocols like up-link and down-link for data transfer in LTE networks.

Some important metrics of LTE networks such as waiting delay and block probability are analyzed using queuing theory.

Thus it is very important to study Queuing systems to understand LTE systems and how applications with different burst patterns effect LTE network performance.

Definition 1.3.1. A process is memory-less if the system forgets the state constantly i.e. probabilities are not influenced by the history of the process. A random variable G is memory-less if the distribution satisfies the property:

$$P(G > a + b|G > a) = P(G > b)$$

This property is also referred to as Markov property.

Definition 1.3.2. A stochastic counting process S(t) with rate $\lambda > 0$ is a Poisson process if it satisfies with the following properties:
a. $S(0) = 0$.
b. $S(t)$ has independent increments
c. The count of arrivals in any time span of length $t > 0$ follows the distribution Poisson(λt).

Definition 1.3.3. Birth death Process: When a service request arrives at a queue system, it is allocated a resource and eventually, the customer leaves the system after the service. This is called a birth-death process. Each request arrival is considered as birth and each served customer is considered as death.

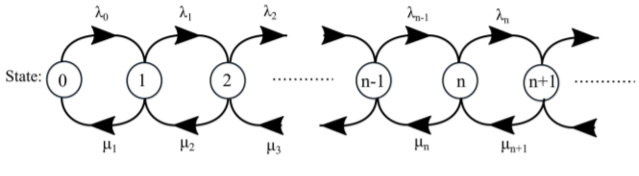


Figure 1.2: Figure showing the transitions in birth-death process

The arrival rate of customers(λ) and service rate(μ) are constant or state-dependent. Mathematical models of queuing systems are studied using the birth-death framework.

Definition 1.3.4. Little's Law: The fundamental relationship between the three parameters λ(arrival rate), L(length of the system), W(waiting time in the system)

$$L = \lambda W$$

1.4 Queuing models

Queuing systems are mainly characterized by the arrival of requests, the service mechanism, and queuing discipline

Definition 1.4.1. Kendall notation is a standard system used in queuing systems to describe a queuing model with three factors A/S/c. Sometimes also referred to as A/S/c/K/N/D.
A: Arrival process.
S: Servicing process.
c: Number of service channels.
K: Length of the queue.
N: Customer population
D: Service discipline

1.4.1 Simple Markovian queues

Single server Queues:
M/M/1 model: M/M/1 is a simple markovian birth-death process consisting a single server. Arrival and service rates are state independent. Inter-

arrival time and service times for M/M/1 queues follow exponential distributions.

$$d_{arr}(t) = \lambda e^{-\lambda t}$$

$$d_{ser}(t) = \mu e^{-\mu t}$$

λ: The rate of arrival.

μ: The rate of service.

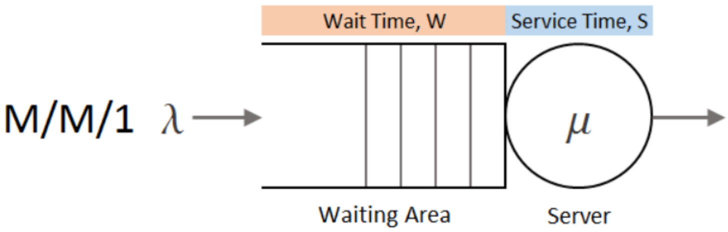


Figure 1.3: M/M/1 queue

Theorem 1.4.2. *An exponential random variable S has the memory-less property.*

Proof. From Bayes theorem,

$$P(S > a+b|S > b) = \frac{P(S > a+b, S > b)}{P(S > b)} = \frac{P(S > a+b)}{P(S > b)} = \frac{e^{-\lambda(a+b)}}{e^{-\lambda b}} = e^{-\lambda a}.$$

$$P(S > a+b|S > b) = P(S > a)$$

Implies exponential random variable is memoryless ☐

Let P_n be the probability for the system in long-term to stay in state n as shown in Figure 1.2.

For a small period of time h, The steady state probability of the system to stay in state n at time t+h (infinitesimally small h) from the birth-death state figure 1.2 gives 3 cases:

- From state n-1 with one request and no service
- From state n+1 with no request and one service
- From state n with no request and no service

Mathematically written,

$$P_n(t+h) = P_{n-1}(t)\lambda h(1-\mu h) + P_{n+1}(t)(1-\lambda h)\mu h + P_n(t)(1-\lambda h)(1-\mu h)$$

For steady state,

$$\frac{P_n(t+h) - P_n(t)}{h} = 0$$

$$P_{n-1}(t)\lambda + P_{n+1}(t)\mu - P_{n-1}(t)(\lambda + \mu) = 0$$

$$(\lambda + \mu)P_n = P_{n+1}\mu + P_{n-1}\lambda(n \geq 1) \tag{1}$$

$$P_0(t+h) = P_1(t)(1-\lambda h)\mu h + P_0(t)(1-\lambda h)$$

$$\mu P_0 = \lambda P_1 \tag{2}$$

From (1) and (2),

$$P_n = \left(\frac{\lambda}{\mu}\right)^n P_0$$

Sum of probabilities of steady states

$$P_0 + P_1 + P2 \quad ... \quad P_\infty = 1$$

$$P_0(1 + (\lambda/\mu) + (\lambda/\mu)^2 + \quad ... \quad + (\lambda/\mu)^\infty) = 1$$

An important parameter of the queueing system is **traffic rate** $\rho = \frac{\lambda}{\mu}$

$$P_0\left(1 + \rho + \rho^2 + \quad ... \quad + \rho^\infty\right) = 1$$

$$P_0\left(\frac{1}{1-\rho}\right)=1$$

$$P_0=1-\rho$$

$$P_n=(1-\rho)^n\rho$$

To analyze a queuing model we study the following parameters

Average number of entities in the system(L_s)

Average number of entities waiting in the queue(L_q)

Average time spent by an entity in the system(W_s)

Average time spent by an entity waiting in the queue(W_q)

L_s is the average number of customers in the system either waiting in the queue or being serviced.

$$\begin{aligned}
L_s &= \sum_{i=0}^{\infty} iP_i \\
&= \sum_{i=0}^{\infty} i\rho^i(1-\rho) \\
&= \sum_{i=0}^{\infty} i(\rho^i-\rho^{i+1}) \\
&= 1(\rho-\rho^2)+2(\rho^2-\rho^3)+.... \\
&= \rho+\rho^2+\rho^3+... \\
&= \rho(1+\rho+\rho^2+\rho^3+...) \\
&= \frac{\rho}{1-\rho}
\end{aligned}$$

$$\implies L_s=\frac{\lambda}{\mu-\lambda} \tag{1.1}$$

L_q is the average number of customers in the system waiting in the queue

$$\begin{aligned}
L_q &= \sum_{i=1}^{\infty}(i-1)P_i \\
&= \sum_{i=0}^{\infty} iP_n - \sum_{i=1}^{\infty} P_n \\
&= L_s - (1 - P_0) \\
&= \frac{\rho}{1-\rho} - (1 - (1-\rho)) = \frac{\rho}{1-\rho} - \rho \\
&= \frac{\rho^2}{1-\rho}
\end{aligned}$$

$$\implies L_q = L_s\rho \tag{1.2}$$

W_s is the average time spent by a customer in the queue system waiting or being serviced.

From Little's law can be written as

$$W_s = \frac{L_s}{\lambda} = \left(\frac{\lambda}{\mu-\lambda}\right)\frac{1}{\lambda} = \frac{1}{\mu-\lambda} \tag{1.3}$$

W_q is the average waiting time spent by a customer in the queue.

From Little's law can be written as

$$W_q = \frac{L_q}{\lambda} = \left(\frac{\rho}{1-\rho}\right)\frac{\rho}{\lambda} = \frac{\lambda}{\mu(\mu-\lambda)} \tag{1.4}$$

From the above equations,

$$W_q = W_s\rho \tag{1.5}$$

Now, we study queuing systems for LTE networks.

Chapter 2

N-Burst Model for LTE Networks

After many years of research it has become clear that usual models of queuing theory are not sufficient for the modelling many telecommunication services. Maybe one of the reasons for this is that data is not transmitted continuously rather it is transmitted in bursts of packets.But this is not at all enough for explaining the unpredictable performance of routers. This traffic must also be self similar along with bursty. The self-similar property can be explained and modeled by using power-tailed distributions of burst-sizes.

2.1 Some Key Words

2.1.1 Telecommunication Network

Group of nodes which are joined by links which are used to exchange messages between them. To exchange messages, these links may use packet switching, circuit switching, and many other methods.

2.1.2 Self Similarity

The bursty traffic can be described using the notion of Self-Similarity. This property is often used to describe the infinite complex patterns. However in our case ,this term is used to describe the distribution of object i.e. The distribution of object remains similar even if it is viewed at varying scales.

2.1.3 Reliability Function

It is also called Complementary cumulative distribution function. If we have a non negative variable Z having mean E(Z). Then it can be described as

$$R_Z(x) = Pr(Z > x) \tag{2.1}$$

2.1.4 Bursts

Burst is a group of consecutive packets which have shorter inter-packet gaps than the packets which arrive after or before the burst of packets.

2.1.5 Packet Delay

Time taken by packets to move from source to destination, or rather from one source to another.

2.1.6 Buffer Overflow

When unexpected number of entities arrive at the router the requests are lost, this is called buffer overflow.

2.1.7 Semi Markov process

If we have a process that can be in N states 1,2...N and the system stays in i^{th} state for a random interval of time having mean μ_i and makes a transition to state j with a probability P_{ij} then such a process is called semi-Markov process. Also abbreviated as SM.

2.1.8 Power tail distribution

Power-Tail distributions with exponent β follow the properties

$$E[X^l] = \int_0^\infty x^l f(x)\, dx \begin{cases} = \infty & \text{for } l \geq \beta \\ < \infty & \text{for } l < \beta \end{cases}$$

Similarly,

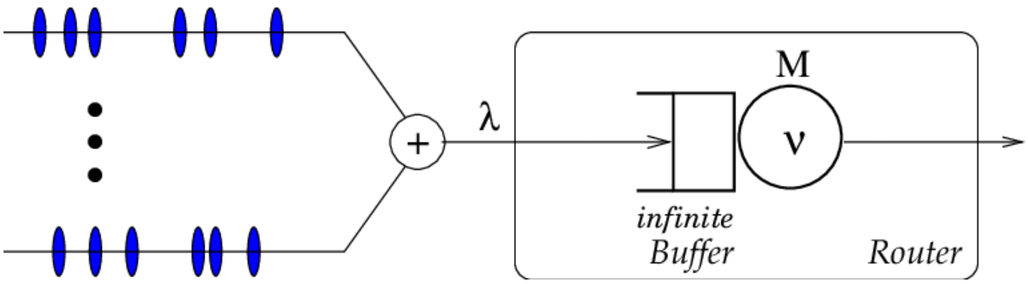
$$\lim_{x \to \infty} x^l R(x) = \begin{cases} = \infty \text{ for } l > \beta \\ = 0 \text{ for } l < \beta \end{cases}$$

Density function c/x^β is a simple form having the above property.

Truncated Power-Tail distribution is when $x^\beta R(x)$ remains almost a constant for large order of magnitudes and drops to 0 eventually

2.2 LTE Network model

A simple LTE network consists of a evolved nodeB, which provides service to a set of N wireless devices. These devices use the LTE technology for downloading data (down-link) and upload and request data (up-link). The mobile devices use several applications to access information using down-link.This



In this figure the packet generated by each user arrive at a single queue which is maintained by a single router or server

Figure 2.1: NBurst/M/1 queue

results in variable traffic for requests in the LTE system. Down-link communication will be our main focus.

The number of packets requested by a mobile device is considered as a random variable. The properties and distribution of this random variable is purely based on the mobile application used to request the data. Which is why standard queuing models are not appropriate to understand telecommunication systems.

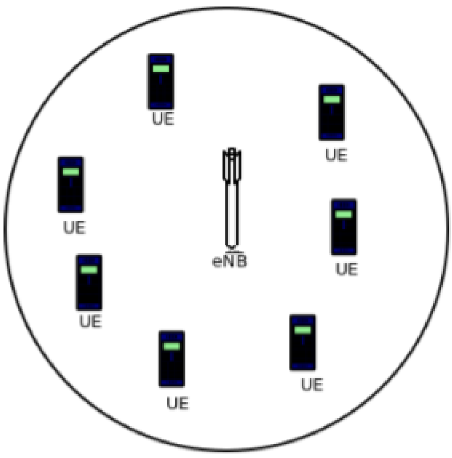


Figure 2.2: LTE network topology

2.2.1 Requirements for an appropriate network model

2.3 Description of Model

N Burst Model is one of the variants of ON-OFF models. The arrival process of N-Burst is the superposition of N ON-OFF type identical and independently distributed source's traffic streams. Let (λ_p) be the peak transmission rate during a burst i.e. at ON time. No packets are transmitted during OFF period.

Given below are the essential parameters of the model :-

κ	: average rate of packet arrival for each source.
λ	: The overall arrival rate that is produced by N sources i.e. $\lambda = \kappa N$
N_b	: Average number of packets during a burst (at ON period)
λ_p	: The peak transmission rate of a node sending packets during a burst.
λ_b	: Mean burst arrival rate $= \lambda/N_b$
$\overline{ON}$	: the average time during which the node is active
$\overline{OFF}$	: the average time during which the node transmission is OFF.
μ	: the mean service rate of the router
$\rho = \frac{\lambda}{\mu}$	: Load utilisation of the router

We have another parameter which is of prime importance called burstiness parameter b defined below:

$$b = \frac{\overline{OFF}}{\overline{ON} + \overline{OFF}} = 1 - \frac{\kappa}{\lambda_p} = \frac{\lambda}{N\lambda_p} \tag{2.2}$$

Since ρ (the amount of data sent per unit time), and ρ can be held constant as b is varied over 0 to 1. So the parameter b can be considered as a shape parameter.

Case 1: b=0

When b=0, it means that average time during which the node transmission is OFF is 0. In this case all the bursts come after one another and ON/OFF process reduces to a simple one.

Case 2: b=1

In this case, the average time during which burst comes is zero. It also means that all the packets in the burst arrive simultaneously i.e. They arrive as a bulk process.

All other cases of b are not simple and require a careful analysis.

2.4 Distribution of Sub Processes

For simplicity let us consider N=1 i.e 1-Burst model.

This model depends on four separate distributions , with random variables denoted by X_{ON}, X_{OFF}, X_{SV} and X_{INP}.

ON: **On time distribution** have mean $\overline{ON}$. Therefore the mean number of packets during a burst is given by $N_b = \overline{ON}.\lambda_p$

OFF: **OFF time distribution** have mean $\overline{OFF}$ (It will depend on how often the bursts are generated)

SV: **Service time distribution** has mean $1/\mu$ (It will depend on distribution of packet sizes, service rate μ depends on packet size and router speed)

INP: **Inter packet time distribution** is the distribution during a burst whose mean is $1/\lambda_p$

The parameters to be varied must have to be selected carefully in order to have comparison study of the model useful. We can have two distributions whose means are different but have same shape, whose higher moments will be scaled proportionately. The description of this can be given as:
Let X and E(X) be the random variable and its mean respectively, then its Complementary Cumulative Distributive Function be given by

$$R_X(x) = Probability(X > x) \tag{2.3}$$

Let Z is the non negative random variable with mean E(Z) and let its distribution has same shape as $R_X(.)$.Then,

$$R_Z(x) = R_X(rx), \text{ where } r = E(X)/E(Z), \text{ for}$$

$$\begin{aligned} E(Z) &= \int_0^\infty R_Z(x)dx \\ &= \int_0^\infty R_X(rx)dx \\ &= \int_0^\infty R_Z(u)du/r \\ &= E(X)/r \\ &= E(Z) \end{aligned}$$

Remark 2.4.1. Moments of functions which have same shape scale according to the below formula:

$$E(Z^n) = r^n E(X^n) \tag{2.4}$$

or

$$\frac{E(Z^n)}{[E(Z)]^n} = \frac{E(X^n)}{[E(X)]^n} \tag{2.5}$$

Using this we can vary the distributions of all the above methods and keeping their shapes same.

2.5 Limiting cases for $b \to 0$ and $b \to 1$

We consider this for 1-Burst process i.e. for N=1.
From (2.2) we can conclude that the the transmission of packets can occur at the slowest rate at $\overline{ON}$ times when $\overline{OFF} = 0$ or thus b=0 ($\lambda_p = \kappa$). In this case the distribution of $\overline{OFF}$ time is irrelevant.
It can be said that since there is no halt in between two bursts , $\overline{ON}$ distribution also does not have an impact on the system.

Remark 2.5.1. Let G_{INP} represents a general distribution which have Inter arrival times as governed by X_{INP}. Interpret the similar symbols accordingly.SM denotes Semi Markov.

Thus for $b \to 0$ the 1-Burst process reduces from $SM/G_{SV}/1$ queue (X_{OFF},X_{INP} and X_{ON} are the distributions on which SM depends) to a $G_{INP}/G_{SV}/1$ queue.
On the other hand if we increase λ_p indefinitely, then according to (2.2), we get $b \to 1$. This is the case when $\overline{ON}$ tends to 0. It also means that all the packets in the burst arrive almost simultaneously i.e. They arrive as a bulk process The bulk size is distributed according to ON-time distribution whose mean is N_b. The time in between two bursts is distributed according to $\overline{OFF}$. Now the 1-Burst process reduces to a ${G_{(OFF)}}^{[ON]}/G_{SV}/1$ queue.

Remark 2.5.2. The above notation of ${G_{(OFF)}}^{[ON]}$ represents the bulk arrival process whose Inter Arrival times X_{OFF} and bulk sizes proportional to X_{ON}.

Our base model should be simple, so we take X_{OFF}, X_{SV} and X_{INP} as exponential distributions. The ON distribution will be varied to obtain different results. Our model have parameters from each of these models. For these cases,

$$b \to 0 \text{ the system becomes an } M_\lambda/M_\mu/1 \text{ queue}$$

$$b \to 1 \text{ the system becomes an } {M_{\lambda_b}}^{[(ON)]}/M_\mu/1 \text{ queue}$$

Remark 2.5.3. M_λ stands for exponential inter arrival times having rate λ.

2.6 LTE Performance Metrics

Remark 2.6.1. Although, the analytic N-Burst model can be evaluated for all possible distributions,the calculations are not trivial.But it is possible to gain insight by looking at the limiting cases of b=0 and b=1.In the description below, we will hold ρ constant, while varying b for various ON time distributions. That is the average load is held constant, while the packets in a burst are bunched up or spread out as much as possible according to the value of b.

There are many modes to use this model. First one being when the idle time approaches zero which leads to continuous flow (This is described by b=0 and no bursts). So, in this case the model can be reduced to Poisson arrival which can also be represented as ($M_\lambda/M_\mu/1$ queuing model) and thus the mean delay

$$mdb0 = \frac{1/\mu}{1-\rho} \tag{2.6}$$

where ρ is as described above (as given in [1]).

The other mode of use is when the active time is reduced to almost 0 (This is described by b=1). Now in this case , there is a simultaneous arrival of all packets and thus mean delay can be written as

$$mdb1 = C.(\frac{1/\mu}{1-\rho}) \tag{2.7}$$

(as given in [1]), where $C = \frac{E(N(N+1)/2)}{E(N)}$ random variable N represent the count of packets in a single burst.

Remark 2.6.2. :b=1 is the bulk arrival process given by queue $M_{\lambda_b}{}^{[(ON)]}/M_\mu/1$. This behaviour is well known and can give the above result (2.7).

If we assume that load utilisation ρ of router is constant, then the best performance of nodes occurs in the case when b=0 and the worst performance for b=1. The size of the burst depends on the burstiness parameter b (each b value represents a unique type of traffic). Now we will write without proof that the queuing model can be represented as matrix exponentiation approach and steady state solution for the system can given as in [1]:
The end to end packet delay using little's formula can be given as:

$$\overline{DEL} = \frac{1}{\kappa}.R(I-R)^{-1}\overline{\epsilon} \tag{2.8}$$

Block probability is given by:

$$Pr(Block) = \frac{1}{\kappa}\pi(R^B L)\overline{\epsilon} \tag{2.9}$$

where R can be calculated by solving system as Quasi-Birth-Death process,I is the identity matrix,κ is the average arrival rate, and ϵ is the unity vector.

Chapter 3

Self-similarity

3.1 What is self similarity?

LTE is a recent advancement in telecommunication systems, and the design of its traffic is of particular interest. Since LTE technology is rapidly expanding in terms of coverage and user base, it is essential to investigate its network traffic. Since data traffic is the most common form of traffic in LTE, this study focuses on the evidence that LTE data traffic has a clear Self-Similarity property. Based on the given parameters, the intensity of this Self-Similarity property is evaluated in LTE networks.

If the network traffic was essentially a Poisson or Markovian arrival process, the burst arrivals should have been smoothed out over a long time period.But, measuring of real traffic data indicates that there is a significant burstiness (traffic variance) present even in a wide time scale.

Self-Similarity is the property we can associate with an object whose appearance remains same irrespective of the scale at which it is viewed. Self-Similarity is used in the distributional sense in the case of stochastic objects like time series: the object's correlational structure remains unchanged when presented at different scales. As a consequence, at a variety of time scales, such a time series experiences bursts.

A bursty traffic can be described statistically using self-similarity. Since bursts are observed on all time scales, traffic at certain time are generally correlated with traffic at a future time.

It has the property that when a time series is aggregated by summing the original data points the resulting time series has the same auto-correlation function as the original time series.

Mathematically, for a time series $X = (X_t : t = 0, 1, 2, ...)$, the m aggregated series, $X^{(m)} = (X_k^{(m)} : k = 0, 1, 2, ...)$ by summing the original series X over non-overlapping blocks of size m. If X is self similar, it has the same auto-correlation function $r(k) = E[(X_t - \mu)(X_{t+k} - \mu)]$ as $X^{(m)}$ for all m. This is referred as distributionally self-similar. Self-similar time series shows long-range dependence, $r(k) \sim k^{-\beta}$ as $k \to \infty$ where $0 < \beta < 1$.

Self similarity is expressed using a single variable, representing the speed of decay of autocorrelation function known as Hurst parameter $H = 1 - \beta/2$.

For self similar series $1/2 < H < 1$, when $H \to 1$, degree of self-similarity increase.

3.1.1 Existence of self similarity

There are several arguments made by researchers about why the Internet traffic is self-similar ranging from file size distribution on web servers, ON/OFF models of heavily tailed distribution, user behavior, network protocols, buffer in routers and the TCP congestion avoidance algorithms.

Extensive statistical analysis shows that the data at the level of user equipment or source-destination pairs are self-similar and exhibit high variability.

3.1.2 Effects of self similarity

Self-Similarity in network traffic results to packet loss. When traffic increases to threshold the bandwidth and the router buffer sizes can't handle the bursts, resulting in packet lost. Packet lost is money lost for network operators. In certain situations the lost packets are sent again which again leads to congestion and wastage of resources. Some methods are used to control traffic.

Predictive feedback control method uses dynamic traffic flow control, by adjusting congestion based on either nodes have on-set of concentrated periods of high or low activity.

Error correction method uses re-transmission of non viable data like streaming audio or video. The level of redundancy is adjusted according to the congestion level. This method has the risk of damaging the congestion level due to high traffic from these nodes.

3.2 Hurst Parameter

The Hurst parameter, also known as the Self-Similarity parameter, is a measure of time series long-term memory. It has to do with time series autocorrelation and the rate at which it decreases as the lag between pairs of values increases.It denotes a time series's proclivity to regress strongly to the mean or cluster in a particular direction.

A time series with a value H in the range of 0.5-1 has long-term positive autocorrelation, which means that a high value in the series will almost certainly be accompanied by another high value, and that the values in the future will also appear to be high.A value in the range of 0 to 0.5 means a time series of long-term swapping between high and low values in adjacent pairs, implying that a single high value will most likely be replaced by a low value, and the value after that will appear to be high, with the propensity to transition between high and low values lasting a long time. H=0.5 may seem to be the value for a fully uncorrelated series, but it is actually the value for series in which the autocorrelation at small time lags may be positive or negative, but the absolute values of the autocorrelation decay exponentially to zero.

There are a number of methods to calculate hurst parameter like, Variance-time plot, R/S plot, Whittle's Estimator, we used Variance-time plot to estimate H,

For a self-similar process, the variance of the aggregated time series follows ,

$$Var(X^{(m)}) \approx \frac{Var(X)}{m^{\beta}}$$

Taking logarithm on both sides give us,

$$log[Var(X^{(m)})] \approx log[Var(X)] - \beta log(m)$$

Since $Var(X)$ is constant, if we plot $Var(X^{(m)})$ and m on log-log plot, we should get a straight line with slope $-\beta$

Chapter 4

Traffic analysis

4.1 Data collection

We used traffic of 4G cell towers traffic data to study self-similarity of the network.These cell towers serve user equipment in their vicinity. When a user makes a data service request, that device will be served by a 4G cell closest to the user. The traffic of a cell within an hour is given by the data capacity of all devices served by the station.
Example: Cell X is serving 30 subscribers, assuming if a customer on average uses 20Mb per hour. Traffic of cell X that hour = 30 * 20 = 600Mb.

Data consists of 50 cells. Approximately the data is collected over 1 year x 24 hours x 57 cells. We will see that the nature of traffic varies from time to time, there is peak traffic around 10-12 AM and 11-12PM, and low traffic is observed during early hours of the day.

4.2 Traffic representation

A sample real world LTE traffic data of one of the cells is represented below. The y-axis represents traffic accumulated at a eNode base station in Megabytes per hour over 1 year time frame.

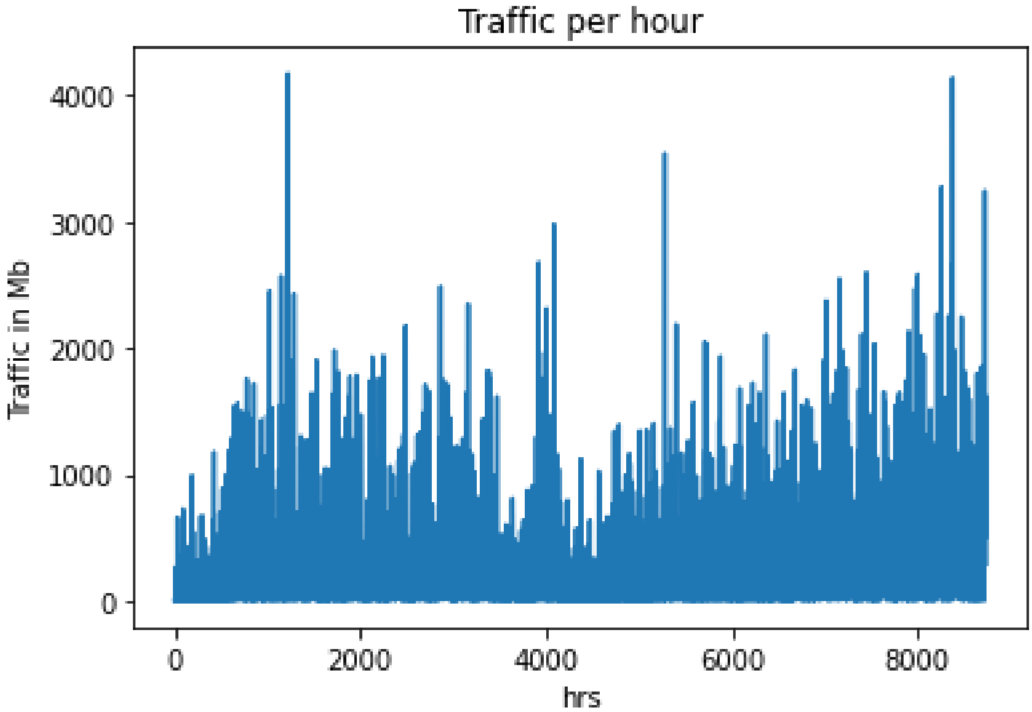


Figure 4.1: Traffic/hour

Next, we aggregate the traffic accumulated per 10 hours by summing the 10 non-overlapping consecutive traffic per hour.

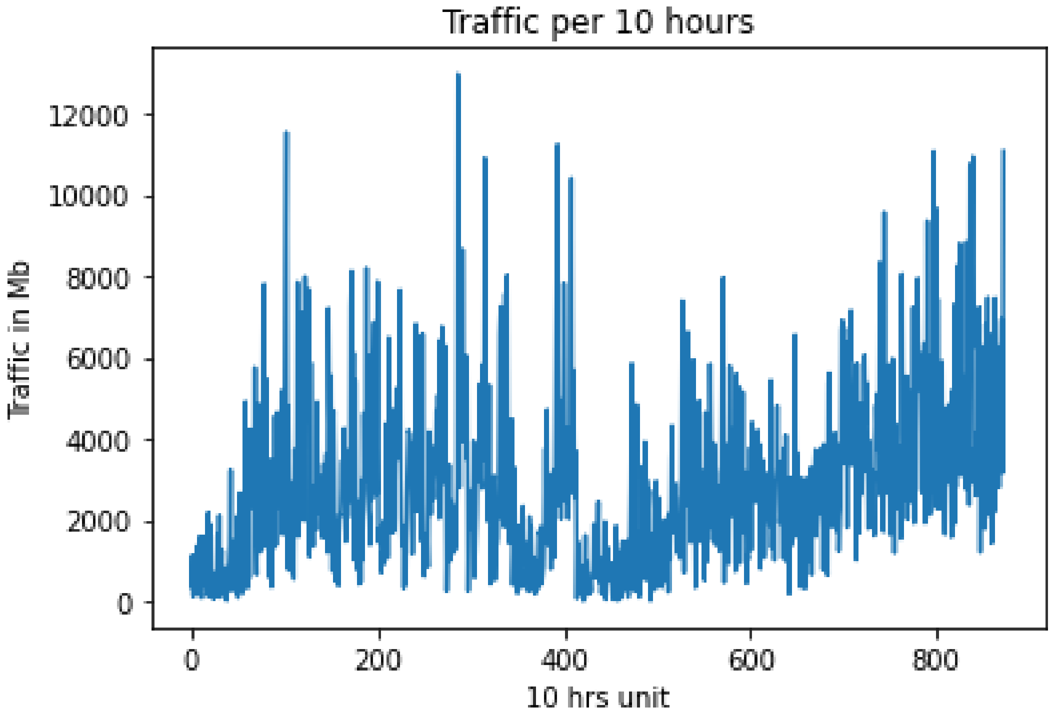


Figure 4.2: Traffic/10 hours

4.3 Variance-time plots

For the data represented above, the $log(X^m)$ vs $log(m)$ is plotted below,the blue cross points represent the log of variance of aggregated time series and the log of aggregation size, the red line represents a line with slope -1. We can clearly see that the points follow along a line with slope between 0 and -1.

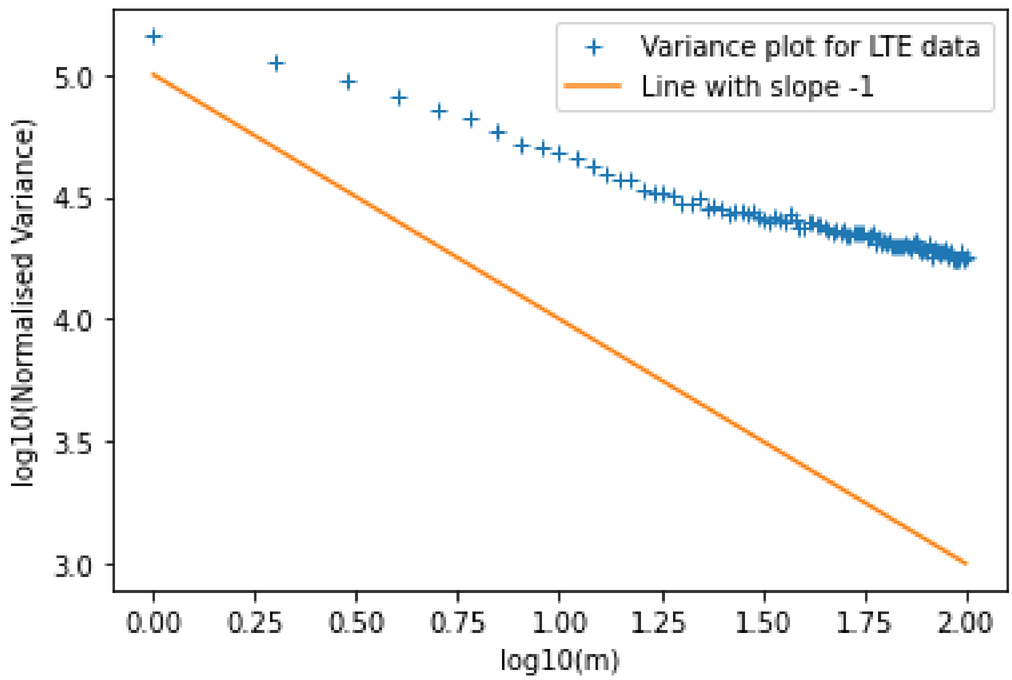


Figure 4.3: log Variance plot

To calculate this slope we use linear regression to fit them along a line, The plotted curve is fitted with a Y=mX+C curve with best fit slope of -0.44535 and y-intercept of 5.11624

$\implies -\beta = -0.44535 \implies H = 1 - \beta/2 = 1 - (0.44535)/2 = 0.777325$

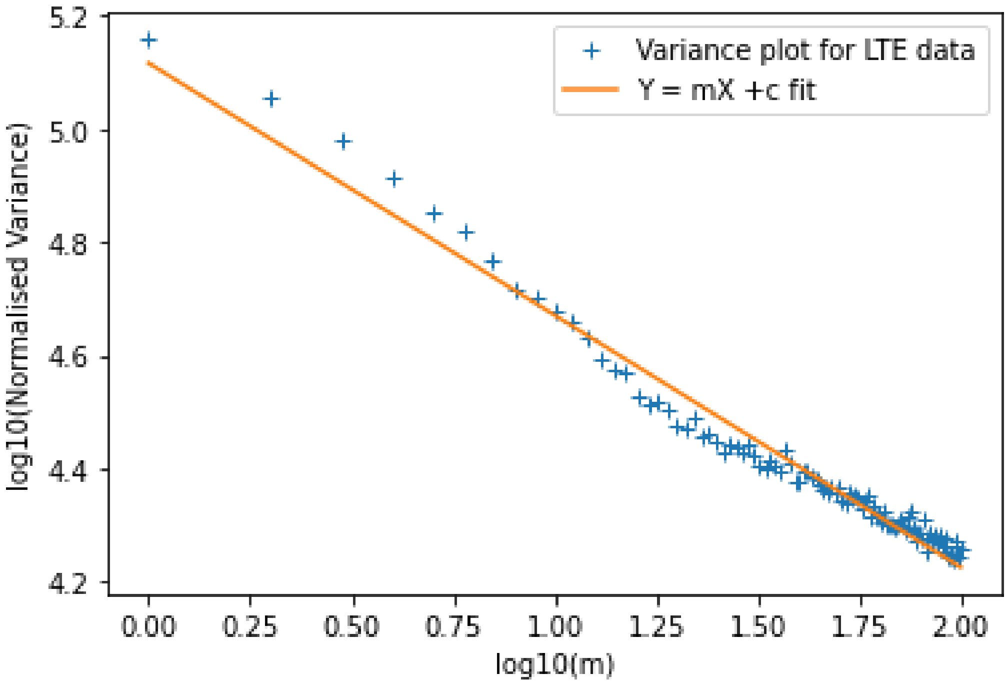


Figure 4.4: Fitting data on a straight line

We take data from more eNode base stations and verify the self-similarity of the traffic using the same method as above, the results of the fit and hurst-parameter are given below

Cell Id	A*x + B	beta	Hurst parameter
Cell_000111	-0.5048886657050763*x + 5.881310697038597	0.504889	0.747556
Cell_000112	-0.47420761455870564*x + 5.3577196417669635	0.474208	0.762896
Cell_000113	-0.34961805216086006*x + 5.453306975104272	0.349618	0.825191
Cell_000231	-0.4732641697597102*x + 5.559633607880568	0.473264	0.763368
Cell_000232	-0.4518799424290425*x + 5.462535811927402	0.45188	0.77406
Cell_000233	-0.3898929645820931*x + 5.66215612546308	0.389893	0.805054
Cell_000461	-0.41319447360042366*x + 5.575030923528314	0.413194	0.793403
Cell_000462	-0.6007279557227103*x + 5.026070152619324	0.600728	0.699636
Cell_000463	-0.44186054206982117*x + 5.606661506167345	0.441861	0.77907
Cell_001912	-0.4453502764645639*x + 5.116624398168283	0.44535	0.777325

Figure 4.5: Table of Hurst parameter for different cells

4.4 Hourly traffic in a week

Traffic of all hours in a week are averaged in the data and is plotted below

The data points start from Wednesday hour 0, and include traffic for each hour upto an year. We can observe that the traffic in the early hours of the day is much less compared to the peak traffic observed within a day.

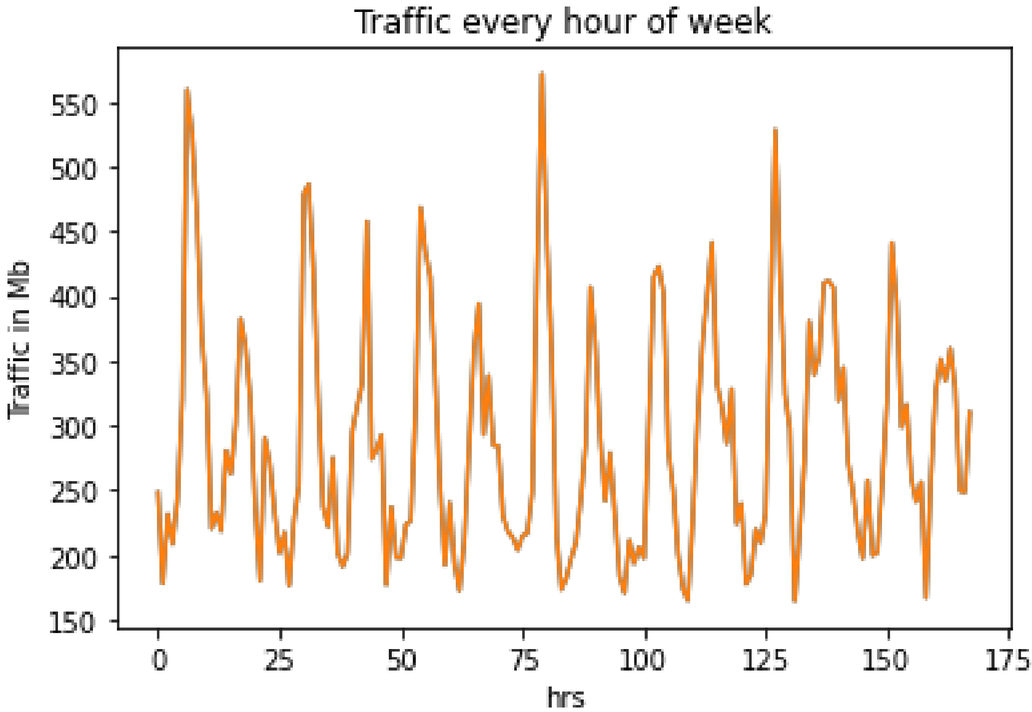


Figure 4.6: Traffic per hour in a week

4.5 Daily traffic in a week

Traffic of all days in a week are averaged in the data and is plotted below

We can observe that the traffic is more during Wednesdays and Mondays, which is generally true for internet traffic as well.

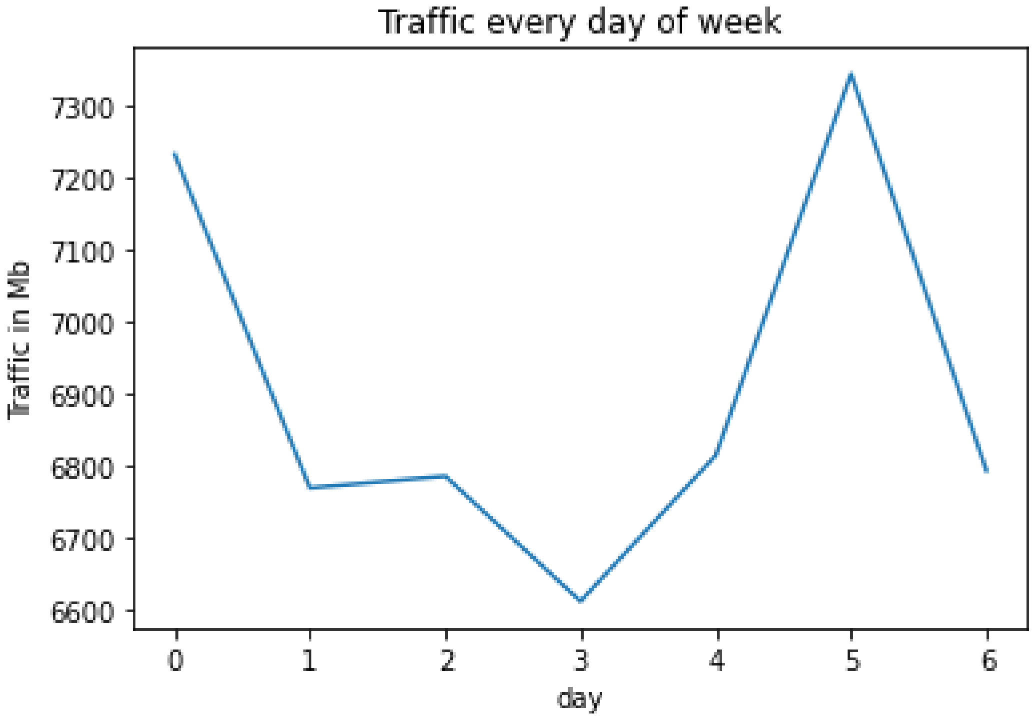


Figure 4.7: Traffic per day in a week

Chapter 5

M/M/1 simulation

5.1 NS-3

NS-3 is a discrete-event network simulator, targeted primarily for research and educational use. NS3 is used to simulate modern network systems with real time scheduler many used for academic and research purposes.

5.2 Model implementation

We implemented M/M/1 queuing model in ns3 which can take inputs like lambda referring to arrival rate, mu referring to service rate, initialpackets referring to number of packets in queue initially, numpackets referring to number of packets to en-queue, queue limit referring to size of the queue. The packets are scheduled to be en-queued according to the arrival rate, and the packets are dequeued according to the service rate.

We log all the activity happening in the queue system and collect traces and store them in a .dat file.

5.3 Simulation results

We study the idle time of the queue, mean queue length, and average packet delay, while varying the arrival rate from $\lambda = 1$ to 9 keeping service rate constant at $\mu = 10$. We send 1,00,000 packets and keep queue limit a relatively large value to avoid packet drops

We plot the results below,

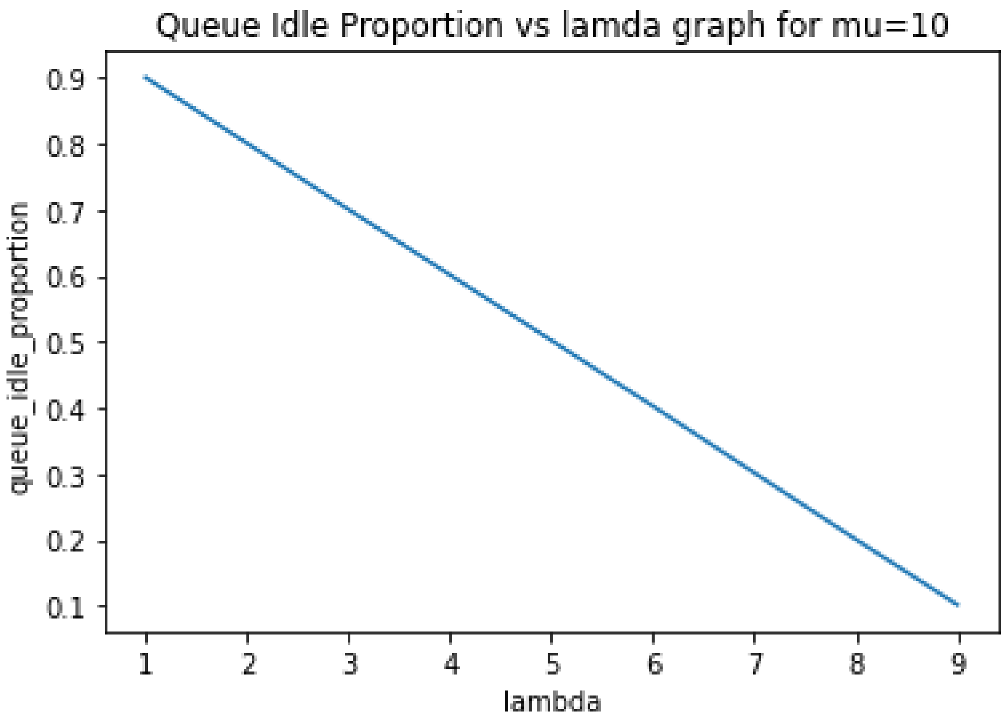


Figure 5.1: Idle time of queue

Increasing lambda linearly decrease the idle proportion of the queue linearly following the formula $P(idle) = 1 - \rho$

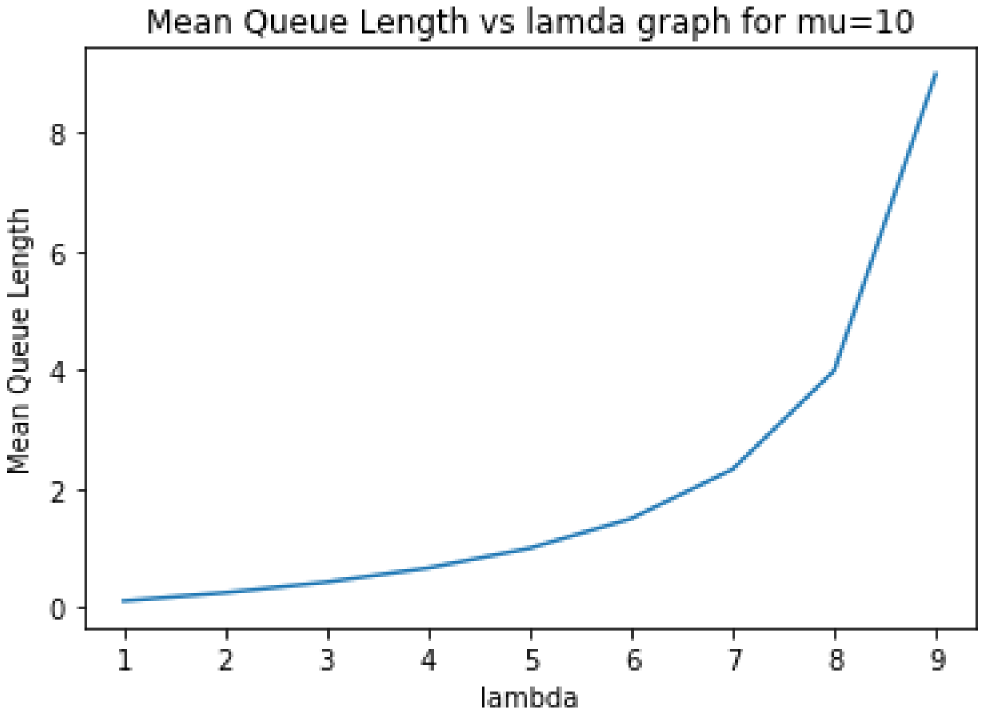


Figure 5.2: Avg length of queue

Length of the system in steady state is consistent with the given by the formula, $L_s = \frac{\rho}{1-\rho}$

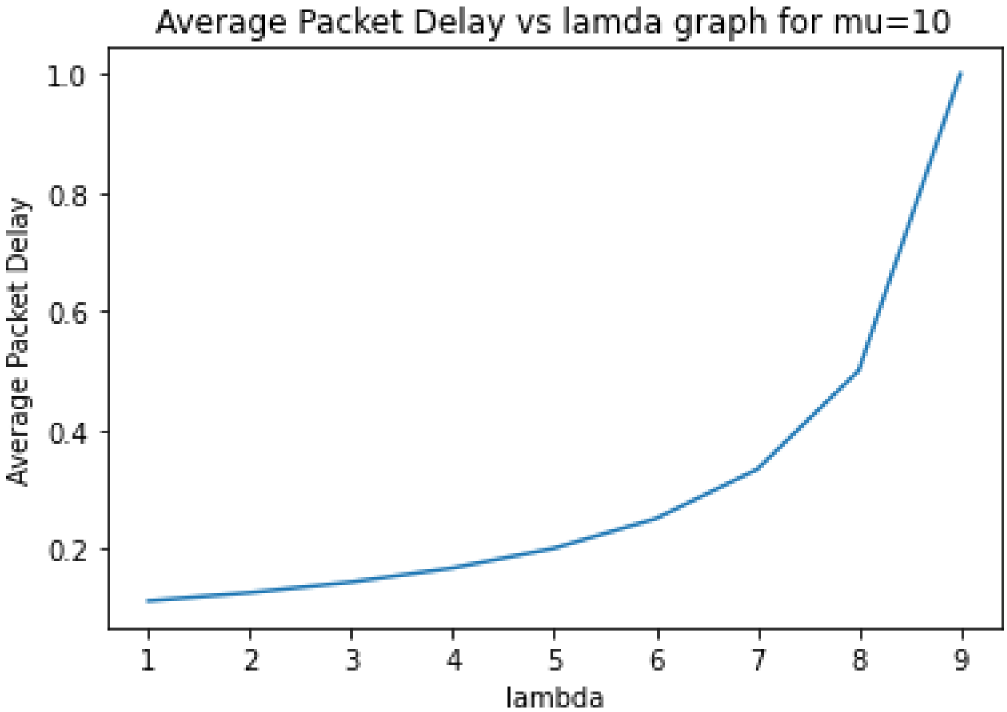


Figure 5.3: Packet delay

Mean Delay of packets is consistent with the formula $W_s = L_s/\lambda = \frac{1}{\mu - \lambda}$

Next we experiment with the buffer overflow times for sending 1,00,000 packets in the M/M/1 queue starting with no packets initially in the queue with the buffer queue size of 100 packets. We fix µ = 10 and run the simulation varying lambda values as follows λ = 9.5, 9.7, and 9.9. We measure how long it takes for the queue to overflow.

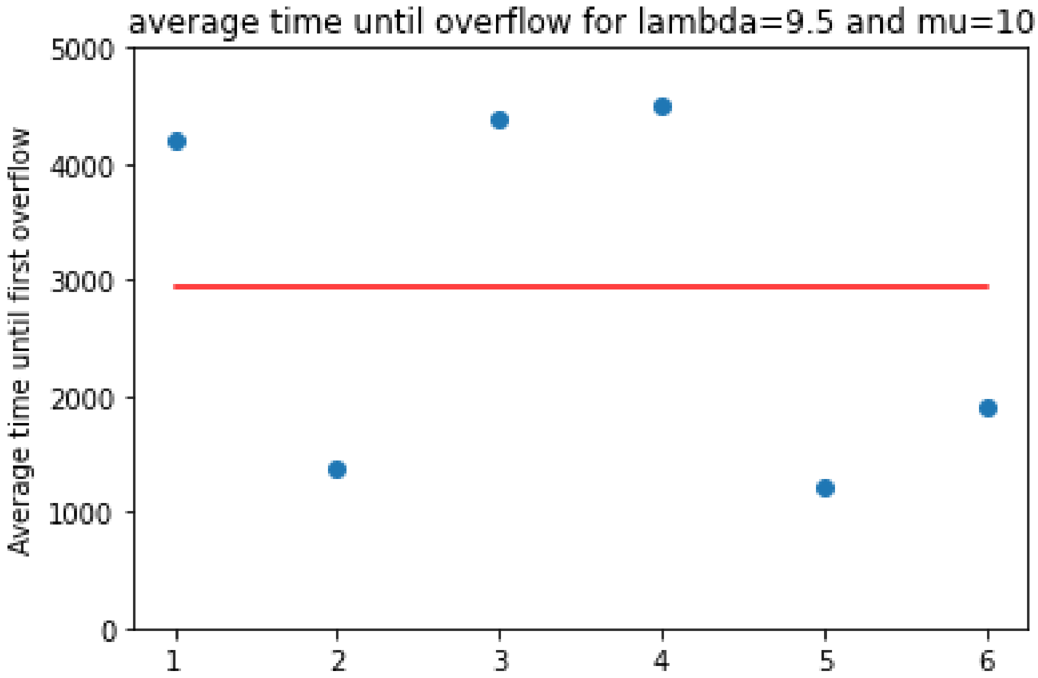


Figure 5.4: Overflow times for lambda=9.5

The individual points represent the times at which the queue overflowed for the first time for different simulation runs with $\lambda = 9.5$. The horizontal line represents the average of the overflow times of these simulations

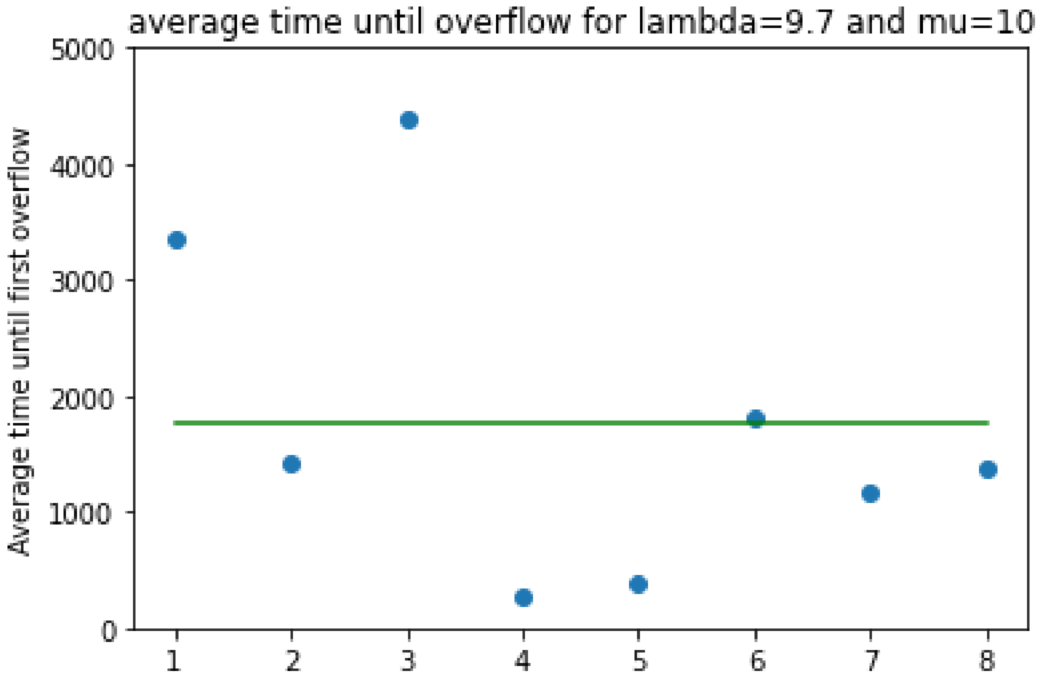


Figure 5.5: Overflow times for lambda=9.7

The individual points represent the times at which the queue overflowed for the first time for different simulation runs with $\lambda = 9.7$. The horizontal line represents the average of the overflow times of these simulations

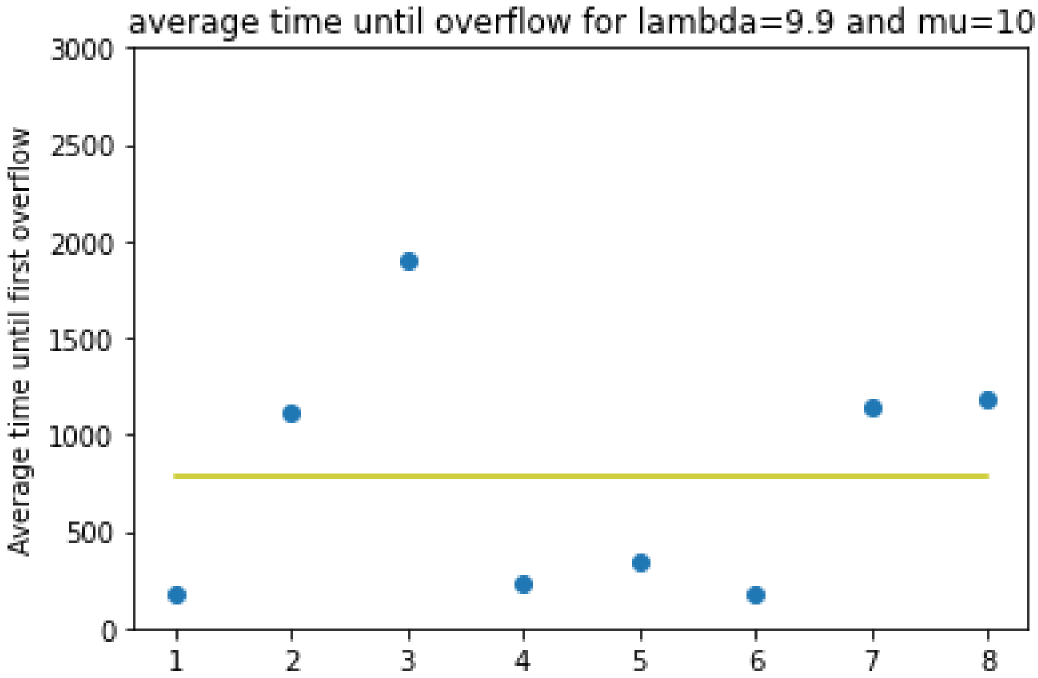


Figure 5.6: Overflow times for lambda=9.9

The individual points represent the times at which the queue overflowed for the first time for different simulation runs with $\lambda = 9.9$. The horizontal line represents the average of the overflow times of these simulations

We can see that the average time until first overflow decreases with lambda value coming close to mu, this behaviour is expected as the increased arrival rate congest the queue faster and probability of a packet dropped from the queue increase at lower times.

Chapter 6

Conclusions

This thesis work started with understanding the basics of queuing systems like queue notation, Markovian processes, performance metrics like mean length of queue/system, mean waiting time in a queue/system and the relation between LTE networks and queuing systems. Next we saw that the basic queue model M/M/1 doesn't take care of the important aspects of a traffic in a LTE network. We study what a queue model must consist to behave like an LTE network. So we study Nburst/M/1 model which takes into consideration of the burstiness of the packet arrivals in the network and look into the performance metrics of the system analytically.

Next we study about self-similarity and see why an LTE network has self-similarity.We also look into the effects of self-similarity. We study different ways to calculate Hurst parameter to find the degree of self-similarity.

We analyse a data set consisting of network traffic at 4G cell stations which is collected over a certain time period and we use Variance-time plot to determine the Hurst parameter of the network, this is done for several datasets. Further we look into the traffic analysis of network by plotting average traffic per hour in a week and average traffic per day in a week.

Next we implemented M/M/1 queue model in ns3 simulator and verified the previously studied results by plotting idle time proportion of the queue, mean length of the queue, mean delay of the packet for varying arrival rates.

Since the calculated Hurst parameters for the analysed LTE dataset are close to 0.8 which is greater than 0.5, we proved that the LTE network traffic is self-similar ,and hence bursty in nature. So we can conclude that Nburst/M/1 model is a suitable model to analyse LTE netwroks.

Bibliography

[1] Adel Agamy and Ahmed Mohamed. Influence of various application types on the performance of lte mobile networks. *International Journal of Electrical and Computer Engineering*, 10:5420–5429, 10 2020.

[2] Jan Beran. *Statistics for long-memory processes*, volume 61. CRC press, 1994.

[3] Mark E. Crovella and Azer Bestavros. Self-similarity in World Wide Web traffic: Evidence and possible causes. *IEEE/ACM Transactions on Networking*, 5(6):835–846, December 1997. Revised and substantially corrected version of [Crovella and Bestavros, 1996].

[4] John F. Shortle, James M. Thompson, Donald Gross, Carl M. Harris. *Fundamentals of Queueing theory.* Fourth Edition. Wiley Publications, 2008.

[5] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1–15, 1994.

[6] Lester Lipsky, Manfred Jobmann, Michael Greiner, and Hans-Peter Schwefel. Comparison of the analytic n-burst model with other approxi-

mations to telecommunications traffic. *International Journal of Electrical and Computer Engineering*, pages 122–132, 02 2001.

[7] R. Sherman D. V. Wilson W. Willinger, M. S. Taqqu. Self-similarity through high-variability: statistical analysis of ethernet lan traffic at the source level. *IEEE/ACM Transactions on Networking*, pages 71–86, 1997.

[8] Walter Willinger, Murad S. Taqqu, Will E. Leland, and Daniel V. Wilson. Self-Similarity in High-Speed Packet Traffic: Analysis and Modeling of Ethernet Traffic Measurements. *Statistical Science*, 10(1):67 – 85, 1995.