Extreme Events and Congestion on Urban Street Networks

A MS thesis

submitted by

Ajay Agarwal

(20202019)

under the supervision of

Dr. MS Santhanam

in partial fulfillment of
the requirements for the degree of
Master of Science in Physics(by Dissertation)

to the

Department of Physics

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, PUNE

April, 2023

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, PUNE
DEPARTMENT OF PHYSICS

DECLARATION

I hereby declare that the work reported in the project titled **"Extreme Events and Congestion on Urban Street Networks"** submitted for the partial fulfilment of M.S degree at the Department of Physics, IISER Pune, is a record of my work carried out under the supervision of *Dr. MS Santhanam*.

Ajay Agarwal

Ajay Agarwal
(20202019)
Department of Physics
IISER Pune
April 2023

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, PUNE
DEPARTMENT OF PHYSICS

CERTIFICATE

It is certified that **Ajay Agarwal** has submitted the project under my supervision for partial fulfilment of an M.S. degree in Physics on the topic **"Extreme Events and Congestion on Urban Street Networks"**. It is further certified that the above candidate has carried out the project work under my guidance during the academic session 2022-2023 at the Department of Physics, IISER Pune.

Dr. MS Santhanam
Professor
Department of Physics
IISER Pune-411008

ACKNOWLEDGEMENT

As a matter of first importance, I offer my genuine thanks to my supervisor Dr MS Santhanam, Professor, IISER Pune. I would like to express my deepest gratitude to him, for his invaluable guidance, support, and encouragement during the project work. His expertise, insights, and feedback have been instrumental in shaping the direction of my research, and I am deeply grateful for the time and energy he has invested in me.

I would also like to thank the faculty members of the physics department, especially Prof. Deepak Dhar for their mentorship, support, and valuable feedback throughout my MS journey. Their expertise, diverse perspectives, and intellectual curiosity have been a constant source of inspiration and motivation for me, and I am honoured to have had the opportunity to learn from them. I would like to thank Prof. Prasad Subramanian and Prof. Ramana Athreya for providing me with computer systems for my project work. Training under Prof. Ramana Athreya during my MS first-year project work in Radio Astronomy helped me a lot to complete this thesis.

I would like to express my gratitude to my labmates Aanjanye Kumar, Ritam Pal, J. Bharathi Kannan, Shaunak Roy, Dr S. Harshini Tekur, Dr P. G. Sreeram and Dr Vimal Kishore (during his visit for Nonlinear Dynamics Conference held at IISER Pune), for their support, and encouragement during MS project. Their friendship, collaboration, and intellectual curiosity have made my MS journey all the more enjoyable and meaningful.

Of course, my MS journey won't have been completed without healthy discussion and hanging out with my classmates Sandipan Manna, Rupesh Kumar, Sathyanarayanan Dhruva and others. I want to thank my family and loved ones, for their unwavering support, encouragement, and love throughout my academic journey. Their belief in me and my abilities has been a constant source of strength and motivation, and I am deeply grateful for their sacrifices and unwavering support.

The TAship for the ”Introductory Mechanics” course under Physics chair Prof. Sudherson Anant and Prof. Arka Banerjee was a constant source of motivation for completing this thesis. Also, TAship for the ”Non-Linear Dynamics” course under my guide itself gave me a wonderful experience. Lastly, I would like to express my gratitude to the participants of my research, who generously gave their time and expertise to help me advance my work. Without their contributions, this research would not have been possible.

Thank you all for your support, encouragement, and unwavering belief in me. I am honoured and humbled to have had the opportunity to pursue my MS thesis at this institution, and I look forward to continuing to contribute to the field of statistical physics, especially network theory, in the years to come.

Ajay Agarwal

(20202019)

ABSTRACT

Complex networks provide a common framework to study and understand dynamics on them, for instance, on internet, transportation networks, and protein interaction networks of biological systems. Though work on network science can be traced back to almost 200 years ago, there is considerable interest in the last two decades due to many interesting applications. Transport dynamics on complex networks, such as traffic on roads or information packets on network of routers, show many emergent phenomena, one of which is an extreme event, a rare event whose probability of occurrence is very low. An extreme event is said to occur if flux through a certain node goes beyond the prescribed threshold (may be related to its flux handling capacity). We use non-interacting degree-biased random walk routing (in real-time) on urban road transportation networks of four cities, namely, Mumbai, Delhi, Ahmedabad and New York. These are planar networks. We confirm the validity of a previously known result for planar networks as well – that small degree node are more prone to extreme events than hubs.

Another emergent phenomenon of interest is congestion arising due to walker interaction and finite handling capacity in the system. For example, road junctions can accommodate only a finite and small number of vehicles. We adopted a interacting random walk model for dynamics on city road transportation networks and studied the collective behaviour through phase transition. The congestion phase transition of real planar network shows similarity with that of 2D lattice network (a homogenous network), in spite of the fact that degree distribution of planar network is quite different from a 2D lattice network. Finally, we studied the extreme events using the generalized random walk model, a realistic transport model and showed that nodes with lower degree are more susceptible to encountering extreme events than the hubs.

Contents

Chapter 1

Introduction

1.1 Introduction

Network science is an interdisciplinary field that studies the structure, behaviour, and dynamics of complex networks, such as social networks, biological networks, and transportation networks. One of the key applications of network science is in the modelling and analysis of traffic flow in transportation networks. Traffic modelling is an important area of study, as it can help us to understand the behaviour of vehicles on the road, predict traffic congestion, and design more efficient transportation systems.

In network science, traffic modelling involves using mathematical and computational models to simulate the flow of vehicles through a network of roads, highways, and other transportation infrastructure. These models can be used to predict the behaviour of traffic under different conditions, such as changes in road design, traffic volume, and weather.

One of the central concepts in traffic modelling is the idea of network centrality, which refers to the importance of different nodes or edges in the network. Nodes or edges with high centrality are those that are most critical for the flow of traffic through the network. For example, in a road network, a major highway or intersection might have high centrality, as it connects many other roads and directs a large volume of traffic. At last, we assert that understanding and optimizing transportation systems have significant implications for improving the safety, efficiency, and sustainability of our cities and communities. These systems impact the daily lives of people and play a crucial role in enabling the mobility of goods and people. Efficient and well-planned transportation networks not only reduce travel time and costs but also minimize the negative impact on the environment and contribute to a better quality of life.

1.2 Networks

Any complex physical system can be abstracted as a graph, for instance, a city network consisting of roads and roads intersection, i.e. junction can be modelled as a graph whose nodes and edges or links depict the junction and road, respectively. Mathematically, a graph is represented as $G(N, E)$, where N and E are the numbers of nodes and edges in the graph. A graph $G(N, E)$ can be represented by an $N \times N$ matrix or adjacency list. The adjacency matrix of a graph is more informative, and any kind of graph, like directed, multigraph and weighted, can be represented easily.

1.2.1 Types of networks

- **Simple undirected graph**
  A network whose links do not have a defined direction. Examples: Internet, power grid, scientific collaboration networks.

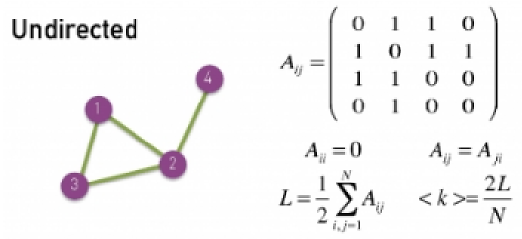
- **Simple Directed Network or Digraph**
  A network whose links have a defined direction.

- **Multi-Graph**
  A network with multiple links(or parallel links) between two pairs of nodes.

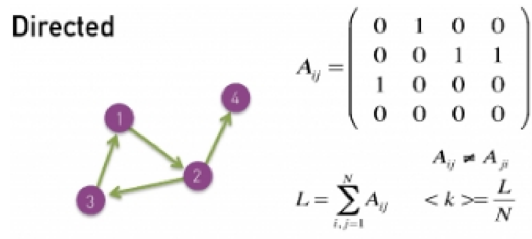
- **Multi-Digraph**
  A network with multiple directed links(or parallel links) between two pairs of nodes.

- **Weighted network**
  A network whose links have a defined weight, for example, road length, in the case of a city network.

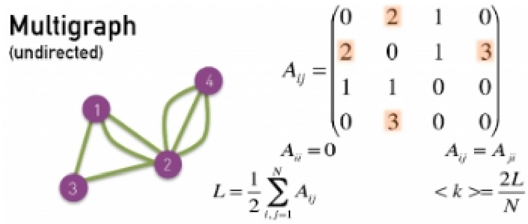
- **Planar network**
  A network that can be embedded in the plane, or in other words, it can be drawn on the plane in such a way that its edges intersect only at their endpoints.



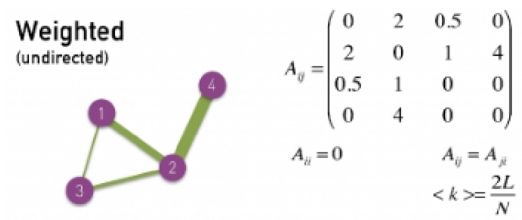
(a) Simple undirected graph



(b) Directed Graph



(c) Multi-Digraph



(d) Weighted Graph

Figure 1.1: Different graphs and its representation by adjacency matrix, taken from [4]

1.2.2 Basic characterization of networks

- **Degree, average degree and degree distribution**

  The degree of a node is the number of links to the nodes in the network. If k_i is the degree of i^{th} node, then the total number of links, E, in an undirected network is given by

$$E = \frac{1}{2}\sum_{i=1}^{N} k_i. \tag{1.1}$$

The average degree is given by,

$$\langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k_i = \frac{2E}{N}, \tag{1.2}$$

or

$$\langle k \rangle = \sum_{k=0}^{\infty} k p_k, \tag{1.3}$$

where p_k is the fraction of nodes with degree k or normalized degree distribution, which follows normalization condition $\sum_k^\infty p_k = 1$. The degree distribution plays a central role in network theory, for instance, for the discovery of scale-free networks, which assumes a power-law degree distribution $p(k) = k^{-\gamma}$, $2 \leq \gamma \leq 3$.

In a directed network, the total number of links E is given by

$$E = \sum_{i=1}^{N} k_i^{in} = \sum_{i=1}^{N} k_i^{out}, \tag{1.4}$$

where k_i^{in} and k_i^{out} represents number of links pointing inward and outward from a i^{th} node respectively. The total degree of node i is given by

$$k_i = k_i^{in} + k_i^{out}. \tag{1.5}$$

The average degree of a directed network is

$$\langle k_{in} \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i^{in} = \langle k_{out} \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i^{out} = \frac{E}{N}. \tag{1.6}$$

- **Paths and distances**

  A path is a sequence of nodes such that each node is connected to the next node successively by a link; for example, the path between nodes i_0 and i_n can be represented as an ordered list of n links $P = \{(i_0, i_1), (i_1, i_2), (i_2, i_3), ..., (i_{n-1}, i_n)\}$, comprising a total of $n + 1$ nodes. For a weighted network $A_{ij} = W_{ij}$, path length is $\sum_{(i,j)\in P} W_{ij}$. For an unweighted network, $W_{ij} = 1$, the length of a path is simply the number of links. There can be any number of paths between given two pairs of nodes, a path which corresponds to the least number of links (unweighted network) or the least distance (weighted network) is called the **shortest path**. Note that the shortest path does not need to be unique. The longest among all shortest paths in the network is called the network's **diameter**. The **average shortest path** is the average of the shortest paths between all pairs of nodes in the network.

- **Connectedness**

  A network is said to be connected if a path exists for every pair of nodes in the network. A network is disconnected if there is at least one pair of nodes with no path. A link which makes a disconnected network connected is called a **bridge**.

1.3 Extreme events and congestion phenomenon

1.3.1 Extreme events (EEs)

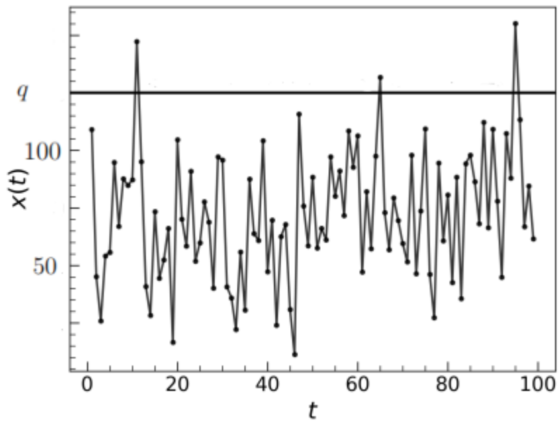
Extreme events are often associated with various types of crises, such as wildfires, heatwaves, hurricanes, volcanic eruptions, and stock market crashes. For example, when the temperature $T(t)$ of a certain location recorded in a particular season as a function of time t exceeds a predefined threshold T_q (which may be the normal temperature of that location in that particular season), due to some inherent fluctuations, it is considered an extreme event. It is worth noting the scalar variables such as temperature, wind speed, economic growth, and seismic activity. Many theoretical and empirical studies have been conducted on the statistics and dynamics of extreme events for such univariate scalar variables [2].
In a similar sense as of above examples, extreme events can also take place on complex networks (an abstract version of complex systems). In complex network settings, phenomenons like traffic jams in city roads, web servers not responding due to heavy load of web requests, floods in the network of rivers, and power blackouts due to tripping of power grids may be the result of some rare events on networks. For example, flux of vehicles on junction of roads goes beyond a certain threshold (related to its handling capacity) can cause traffic jam in the whole city. Real-world examples of extreme events include the China National Highway 110 traffic jam, which was a recurring traffic jam on China National Highway and Beijing–Tibet expressway, in Hebei and Inner Mongolia. The traffic jam slowed thousands of vehicles for more than 100 km and lasted for ten days. Many drivers were only able to move their vehicles 1 km per day, and some reported being stuck in the traffic jam for up to five days, making it one of the longest traffic jams on record [15]. Another notable example is the power blackout that occurred in the northeastern United States in 2003.

These events can be thought of as an emergent phenomenon due to the transport process on the networks and lead to losses ranging from financial and productivity to even of life and property, in order to handle them, it is important to estimate probabilities for the occurrence of EEs and, if possible, incorporate them to design resilience networks. Incorporating the estimation of probabilities of extreme events into the design of networks can help improve their resilience, to withstand and recover from extreme events, reducing the magnitude and duration of their impacts. This can involve various strategies, such as redundancy, robustness, and adaptability, to ensure that the network can continue to function even in the presence of extreme events. For example, in a city road network, incorporating alternative routes and transportation modes can help reduce the impact of traffic jams.

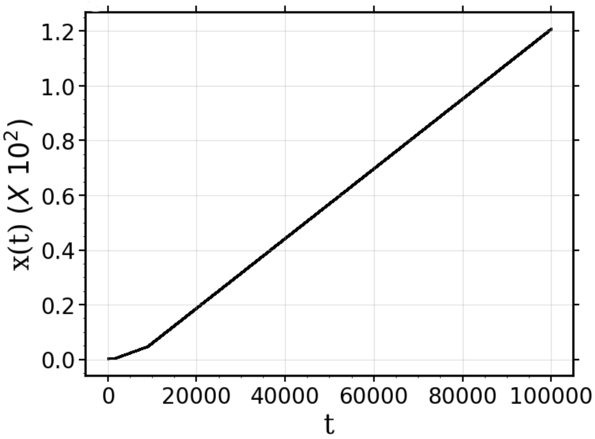
1.3.2 Congestion

Congestion is another important phenomenon on real transport networks such as the Internet, city roads and other queuing complex networks. It is considered to be another emergent phenomenon or collective behaviour that mainly arises due to interaction in the system. The first instance of congestion collapse on the Internet was recorded in October 1986, when the data throughput between Lawrence Berkeley National Laboratory and the University of California in Berkeley experienced a sharp decline from 32 Kbps to a mere 40 bps. Since then, many synthetic simulations have been done on real systems (like the Internet) to study the cause and design better routing protocols to mitigate congestion. The same phenomena can be thought to happen on city road networks as dynamics on the Internet are quite similar to that on city roads. People are now applying routing algorithms for traffic dynamics on roads which were developed for improving internet traffic flow [3].

Fig. 1.2, depicts how extreme events are different from congestion. It is possible for extreme events and congestion to occur concurrently within a transport system. It is imperative to conduct a thorough examination of these events in order to formulate theoretical forecasts and implement measures to mitigate the loss of property and life, while also saving energy and time.



(a) Extreme events is said to occur when $x(t) > q$, q is some threshold. It can be studied only for free flow regime or $x(t)$ fluctuates about some average value.



(b) Here, vehicles are accumulating on a node or junction, forming queue and hence slowing down the traffic flow. Thus, causing congestion.

Figure 1.2: Extreme event and congestion phenomenon.

Chapter 2

Statistical Features of Spatial Networks of Urban Streets

2.1 Introduction

A particular class of complex networks whose nodes occupy a precise position in two- or three-dimensional Euclidean space and whose edges are real physical connections referred to as spatial networks. Spatial networks of urban streets refer to the interconnected system of roads, streets, and pathways that comprise a city's transportation infrastructure. Understanding the properties and characteristics of these networks is essential to understanding how cities function and how they can be designed or improved.

One important aspect of spatial networks or any network is centrality. Centrality measures identify the most important nodes or intersections in a network, based on factors such as the number of connections they have to other nodes, the number of shortest paths through them, and the flow of traffic through them. By identifying these key locations, planners and policymakers can better understand the flow of vehicles, goods, and information through a city, and design interventions to improve accessibility, safety, and efficiency. There are different centrality measures, for instance, betweenness centrality, closeness centrality and degree centrality. Each of these measures provides a different perspective on the network, and can be used to identify different types of nodes that are important for different reasons.

2.2 Spatial networks of urban streets

A spatial network of urban streets refers to the interconnected system of roads and streets that make up a city's transportation infrastructure. These networks can vary in complexity, with some cities having a more grid-like pattern while others have a more organic and meandering layout.

New York City's street network is known for its grid-like pattern, which was established in the early 19th century as part of the Commissioners' Plan of 1811. The other cities that show grid-like pattern includes Portland, Oregon and San Francisco, California. Ahmedabad, Delhi and Mumbai, on the other hand, have more organic and unplanned street networks that have developed over time. In these cities, streets often follow the contours of the land, and intersections can be irregular and difficult to navigate.

In this thesis, we examine the urban street networks of major Indian cities, including Delhi, Mumbai, and Ahmedabad. To provide a comparative analysis, we also studied the planned city network of New York, USA. We obtained the spatial network data, denoted as $G(N, E)$, using OSMnx [5]. Here, N and E represent the number of nodes and edges in the network, respectively. To capture approximately 1000 nodes, we selected a region with a specific radius around the given coordinates. It is worth noting that the data structure of networks obtained using OSMnx is a multi-digraph. However, for simplicity, we converted it into a simple undirected network. Table 2.1 provides essential information regarding our networks, while fig. 2.1 showcases a visually appealing representation of the networks under consideration.

City	**Coordinates**	**Radius** (Km)	N	E	**Total road length** (Km)
Ahmedabad	23.03°N, 72.58°E	1.5	1064	1528	135.3
NewYork	40.71°N, 74.01°W	2	1117	1930	193.3
Delhi	28.70°N, 77.10°E	0.95	1092	1708	99.5
Mumbai	19.07°N, 72.87°E	2.15	1071	1522	169.7

Table 2.1: Network details

(a) NewYork city **(b)** New Delhi

(c) Mumbai **(d)** Ahmedabad

Figure 2.1: Spatial networks of urban streets, red dots represent nodes and grey lines edges

2.3 Degree distribution of networks

The degree distribution of a city's street network can provide insights into the city's transportation infrastructure, connectivity, and accessibility. In a city with a more grid-like pattern, like New York, the degree distribution is nearly uniform, with most intersections having a similar number of streets connected to them, for instance, there are equal numbers of nodes of degrees 3 and 4. In a city with a more organic street network, like Mumbai, Delhi and Ahmedabad, the degree distribution is more heterogeneous, peaking at degree 3, see fig. 2.2.

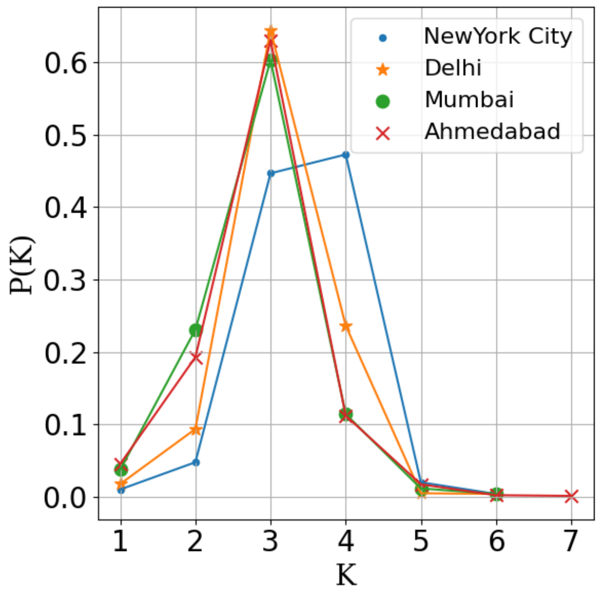


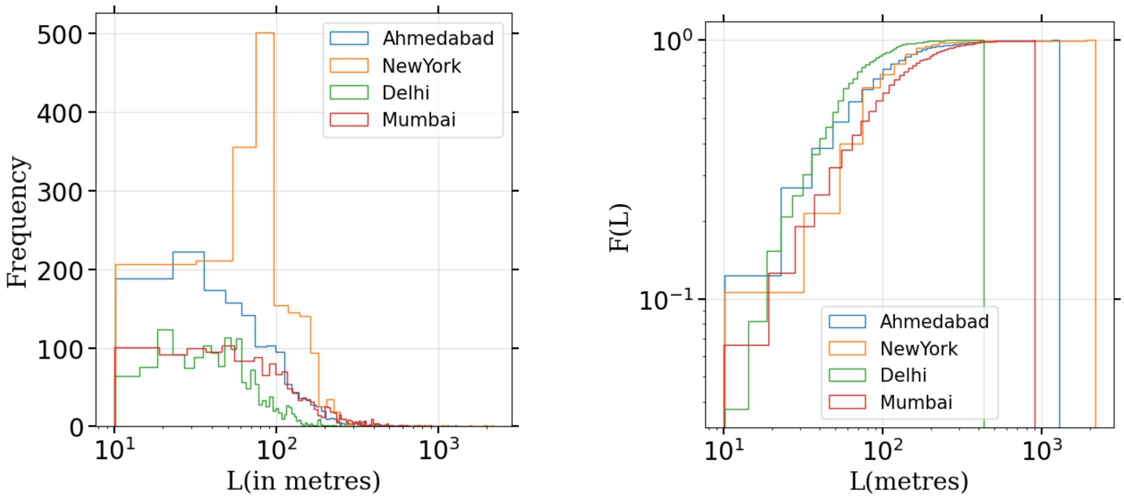
Figure 2.2: Degree distribution for different networks

2.4 Road length statistics

An important aspect of urban street networks is their street's road length distribution. In fig 2.3(a), we have shown the road length distribution. The cumulative distribution $F(L)$ is defined as $F(L) = \int_0^L \frac{E(L')}{E_T} dL'$ where E_T is the total number of edges and $E(L')dL'$ is the number edges with having edge length between L' and $L' + dL'$, see fig. 2.3(b). Table 2.2, shows statistics of road length of our considered networks. The minimum road length in all networks is $10m$.

City	$< L >$	σ_L	Maximum road length
Ahmedabad	88.4	95.5	1282.2
New York	100.0	117.2	2170.1
Delhi	58.3	39.5	432.9
Mumbai	111.4	101.5	903.6

Table 2.2: Statistics of road length in metres



(a) Histogram

(b) Cumulative distribution

Figure 2.3: Distribution of edge length in meters (m) for considered cities

2.5 Diameter of networks

The shortest distance between the two most distant nodes in the network is termed as diameter. It can be computed as, first compute all the shortest path lengths from every node to all other nodes, the maximum of all the paths is the diameter. The diameter for weighted and unweighted networks need not be the same. For an unweighted graph, diameter is the minimum number of links to be hopped between the two most distant nodes, while for a weighted graph, it is the minimum path length to be travelled. The table 2.3 shows the diameter of considered networks

City	Diameter (unweighted)	Diameter (weighted)
Ahmedabad	46	6.05Km
NewYork	44	6.15Km
Delhi	58	3.75Km
Mumbai	66	8Km

Table 2.3: Diameter of networks

2.6 Centrality

We have computed the two important centrality measures, namely,

- Betweenness centrality
- Closeness centrality

For any centrality C_x, we computed the cumulative distribution using

$$F(C_x) = \int_0^{C_x} \frac{N(C_x^{'})}{N_T} dC_x', \tag{2.1}$$

where N_T is the total number of nodes and $N(C_x')$ is the number of nodes having centrality value between $C_x^{'}$ and $C_x' + dC_x'$.

2.6.1 Betweenness centrality

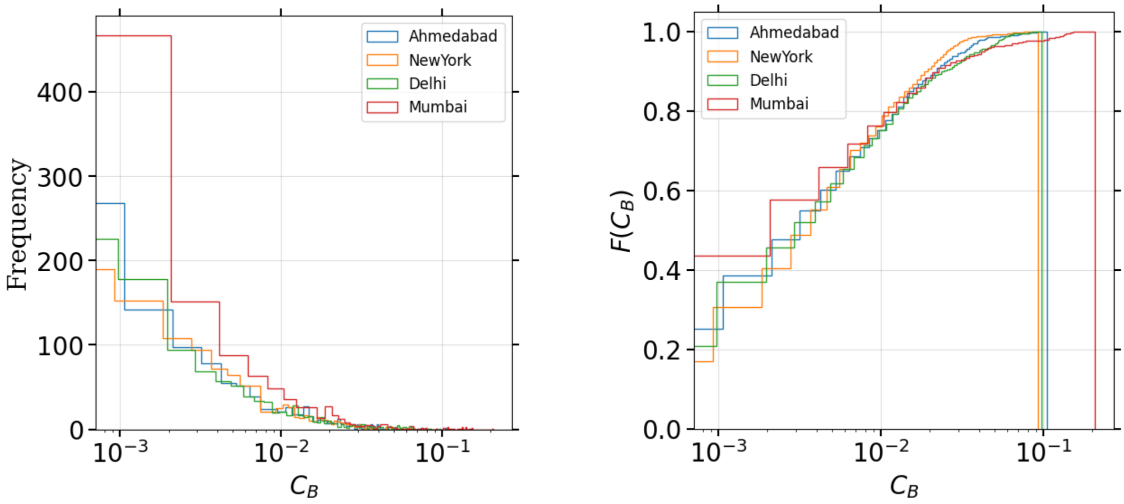
Betweenness centrality is a measure of a node's importance in a network. It quantifies the number of shortest paths that pass through a node, and it is often used to identify nodes that act as "bridges" or "gatekeepers" between different parts of a network. Betweenness centrality can be different for the network with unweighted and weighted edges. In an unweighted network, the shortest path for a pair of nodes is the least number of links connecting them. On the other hand, in a weighted network, the shortest path is the one for which the sum of the weight of edges is minimum. Mathematically, betweenness centrality of a node is given by

$$C_B(v) = \frac{1}{(N-1)(N-2)} \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}, \tag{2.2}$$

$\sigma(s,t)$ = number of shortest paths from node s to t.
$\sigma(s,t|v)$ = number of shortest paths from node s to t passing through node v.
And, N is the total number of nodes.

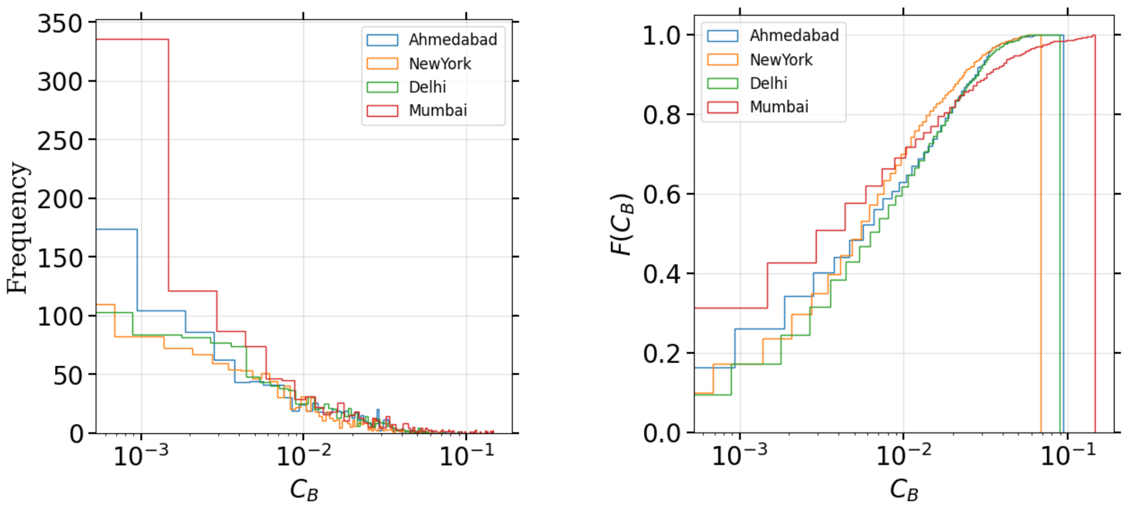
In 2.4(a) and 2.5(a), we can see the distribution of betweenness centrality for different real planar networks for unweighted and weighted, respectively.

In general, weighted betweenness centrality can be a more accurate measure of a node's importance in a network since it considers the actual distances and costs



(a) Histogram **(b)** Cumulative distribution

Figure 2.4: Distribution of betweenness centrality of nodes for unweighted networks



(a) Histogram **(b)** Cumulative distribution

Figure 2.5: Distribution of betweenness centrality of nodes for weighted networks

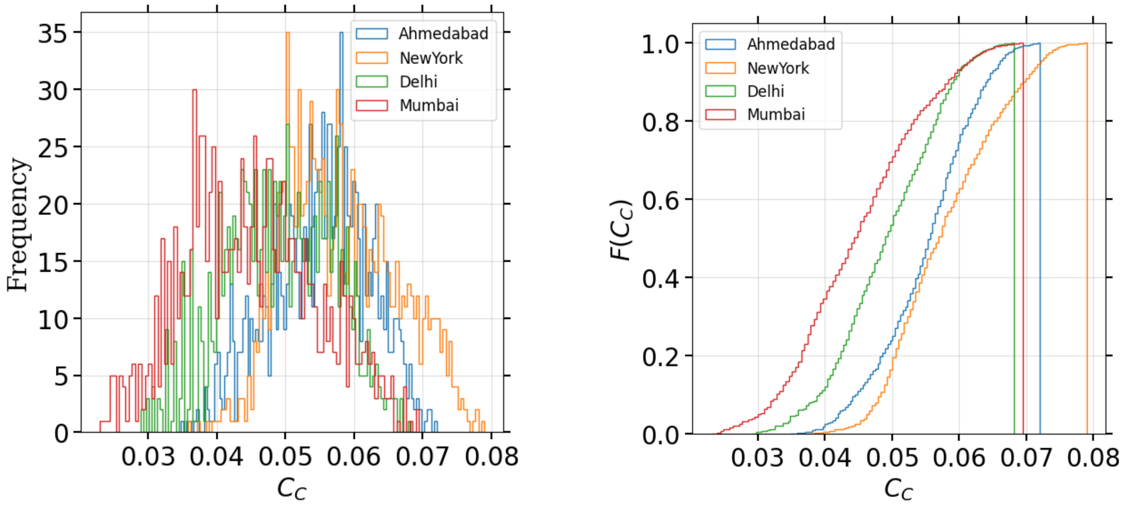
of paths between nodes. However, unweighted betweenness centrality can still be useful in some cases, especially when the weights of the edges in the network are not meaningful or when the focus is on the network's topology rather than the actual distances between nodes.

2.6.2 Closeness centrality

Closeness centrality is a measure of centrality in a network. It measures how easily a node can reach other nodes in the network. A node with high closeness centrality is considered to be "close" to other nodes in the network, in the sense that it can reach them quickly and efficiently. This can be an important characteristic in many real-world networks, such as social networks or transportation networks, where the ability to quickly and efficiently navigate the network is important. Closeness centrality of a node u is the reciprocal of the average shortest path distance to u over all $N-1$ reachable nodes.

$$C(u) = \frac{N-1}{\sum_{v=1}^{(n-1)} d(u,v)}, \tag{2.3}$$

where $d(u,v)$ = shortest path distance between u and v.
Fig. 2.6(a) and 2.7(a), show the distribution of closeness centrality for different planar networks for unweighted and weighted, respectively.



(a) Histogram **(b)** Cumulative distribution

Figure 2.6: Distribution of closeness centrality of nodes for different unweighted networks

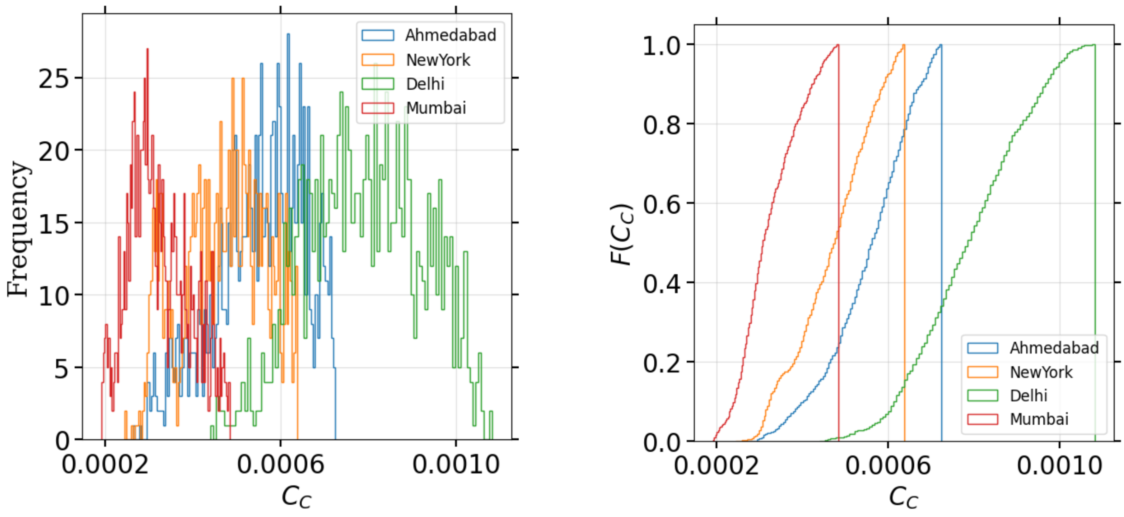
2.7 Clustering coefficient

The clustering coefficient indicates the degree of connectivity of neighbours of a given node. For a node i with degree K_i the local clustering coefficient is defined as

$$C_i = \frac{2L_i}{K_i(K_i-1)}, \tag{2.4}$$

where L_i represents the number of links between the K_i neighbours of node i. C_i for nodes of a network lies between 0 and 1.

- $C_i = 0$ if none of the neighbours of node i link to each other.



(a) Histrogram **(b)** Cumulative distribution

Figure 2.7: Distribution of closeness centrality of nodes for different weighted networks

- $C_i = 1$ if the neighbours of node i form a complete graph, i.e. they all link to each other.
- C_i can be considered as the probability of two neighbours of a node-link to each other. Consequently, $C = 0.5$ implies that there is a 50% chance that two neighbours of a node are linked.

In summary, C_i measures the network's local link density: The more densely interconnected the neighbourhood of node i, the higher its local clustering coefficient.

Average clustering coefficient

The degree of clustering of a whole network is captured by the average clustering coefficient, $\langle C \rangle$, representing the average C_i over all nodes $i = 1, ..., N$. Mathematically,

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i \tag{2.5}$$

City	$< C >$
Ahmedabad	0.0754
NewYork	0.0550
Delhi	0.0157
Mumbai	0.0810

Table 2.4: Average clustering values

2.8 Degree correlation

Consider a bunch of nodes, and let's assume that each node connects randomly to other nodes. Then, the probability that nodes with degrees K and K' link to each other can be given by

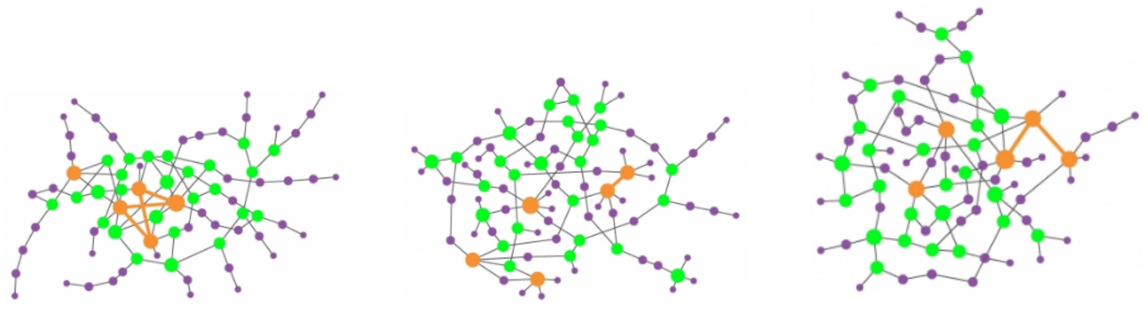
$$p_{K,K'} = \frac{KK'}{2L} \tag{2.6}$$

Where L is the total number of links in the network.
We can clearly see if K and K' are large so as $p_{K,K'}$, so we expect that $p_{1,2} << p_{13,56}$.
But, in general, we can find the networks in which hubs avoid hubs, falsifying the above.

2.8.1 Assortativity and disassortativity

The networks having approximately the same degree distribution may have different structures, giving rise to a different class of networks.

- Assortive
- Disassortive
- Neutral



(a) Assortive: Hubs connect hubs **(b)** Disassortive: Hubs avoid hubs **(c)** Neutral

Figure 2.8: Networks with precisely the same degree distribution, but displaying different degree correlations [4]

2.8.2 Measuring degree correlation: Degree correlation function

The degree correlation gives the relationship between the degrees of nodes that link to each other. We can quantify the degree correlations of a node by the average degree of its neighbours.

$$k_{nn,i} = \frac{1}{K_i} \sum_{j=1}^{K_i} K_j, \tag{2.7}$$

where K_i and K_j are the degree of node i and its neighbour j, respectively.

- Neutral Network

  For a neutral network, the degree correlation function is a constant given by

  $$k_{nn} = \frac{< K^2 >}{< K >} \tag{2.8}$$

  so a graph plotted between k_{nn} and degree K of nodes is a horizontal line. if we model the degree correlation function by an exponential function

  $$k_{nn} = AK^{\mu} \tag{2.9}$$

  where A is a constant, then μ would be zero for neutral networks.

- Assortive Network

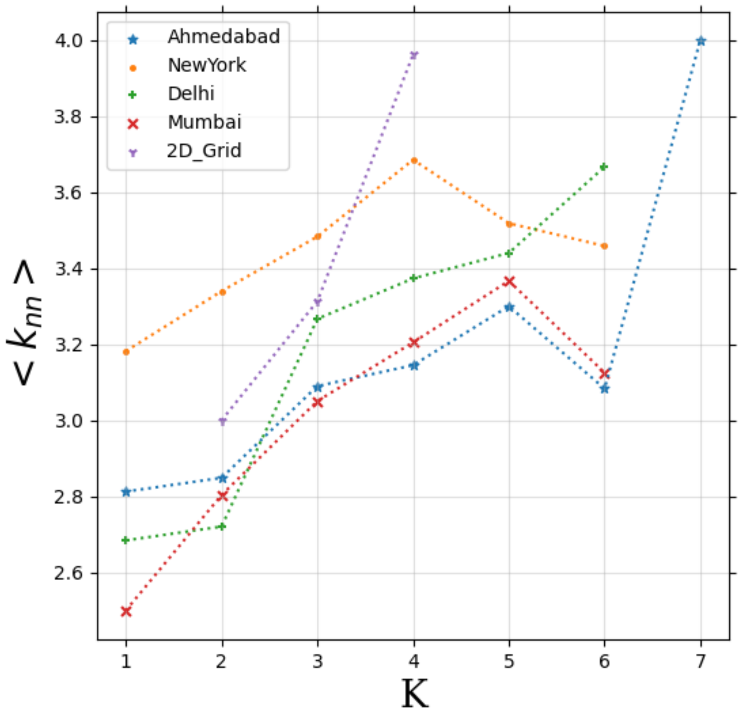
  For a assortive network, the degree correlation function k_{nn} is an increasing function of degree K, so if we model the degree correlation function by 2.9, then we have $\mu > 0$.

- Disassortive Network

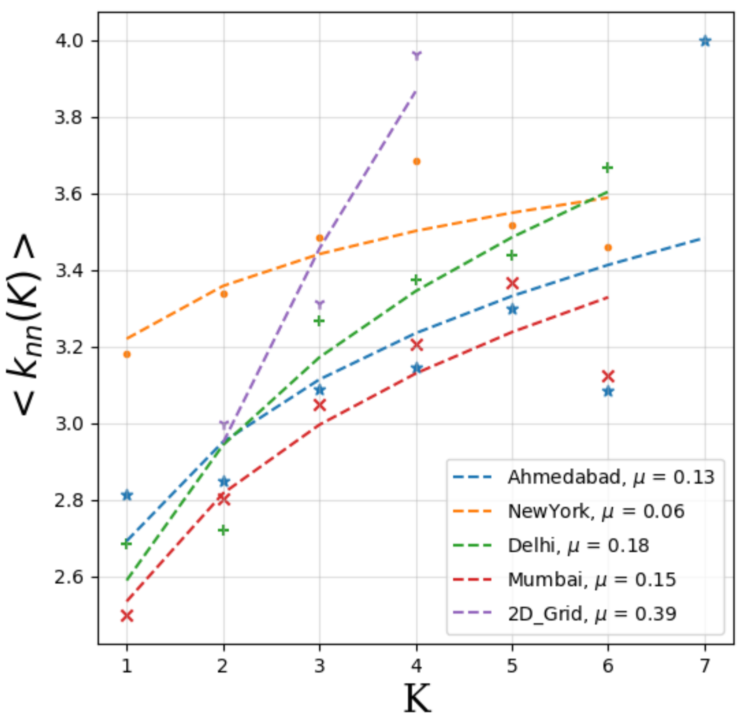
  For a disassortive network, degree correlation function k_{nn} is a decreasing function of degree K, so using equation 2.9, we have $\mu < 0$.

So correlation exponent μ acts as a single parameter for quantifying the type of network.

From the figure 2.9(b), $\mu > 0$ for all networks, hence all considered networks are assortative.



(a) Variation of average degree correlation function k_{nn} with degree K



(b) Curve fitting by Ak^{μ} model

Figure 2.9: Average degree correlation function $< k_{nn} >$ and curve fitting

Chapter 3

Degree Biased Routing Model and Extreme Events on Weighted Spatial Networks of Urban Streets

3.1 Introduction

For simulating transport dynamics on a real urban city road network, the random walk model is a simple and useful fundamental model to start with, compared to real complex transport dynamics. Assuming we disregard driver behaviour, vehicle interaction, and other intricate factors, we can consider the flow of vehicles through the road traffic network as a random walk. However, to create a more accurate model of this flow, we can use the generalized random walk. This approach incorporates a bias towards either high-degree nodes or low-degree nodes to better represent the actual flow of traffic [11], [13]. Imagine two distant road junctions that are not directly linked by a road. In such cases, these junctions are often connected through a significant hub in the city network. This scenario is one example of how traffic tends to be biased towards the hubs in practical situations. This tendency can be observed in other types of networks as well, such as telecommunication networks where phones connect to the nearest hubs, and railways that link remote areas to important or major junctions. The real-world instances we just mentioned inspired us to develop a model for transportation processes using topology-biased random walks. We then analyzed the likelihood of extreme events and other associated statistical measures using this model. In contrast to work in the paper [13], here we have adopted real transportation network of a city with edges depicting actual road length, and a vehicle is asked to take a speed from a normal distribution. On the other way, this can be seen as a continuous time random walk where a vehicle picks up a waiting time from a distribution of edge travel time shown in figure 3.1.

3.2 Degree biased routing model

We consider a weighted network consisting of a finite, undirected, and connected set of N nodes and E edges, with the weight of each edge representing the corresponding road length. There are W independent walkers performing biased random walks on this network in the sense explained below.

- A walker on a node i transits to node j with probability P_{ij} given by

$$P_{i \to j} = \frac{K_j^{\alpha}}{\sum_l K_l^{\alpha}} A_{ij} \tag{3.1}$$

  where α is a parameter, K_i is node degree and A_{ij} is adjacency matrix.

- A speed v is chosen from a normal distribution $N(\mu, \sigma)$, time taken to traverse the edge E_{ij} of length L_{ij} given by

$$t_{ij} = \frac{L_{ij}}{v} \tag{3.2}$$

- All walkers are allowed to traverse the whole network until total travel time reaches a maximum t_{max}.

- Suppose for node i, the walkers time series is $[(0,\ W_0), (t_{i1},\ W_1), (t_{i2},\ W_2)....(t_{im}, W_m)]$ such that $t_{im} < t_{max}$ where $W_j \in \mathbb{Z}^+$ is the number of walkers at time instance $t_{ij} \in \mathbb{R}^+$. If $\Delta\mathbb{N}$ is the window time (or time resolution), the interval in which walkers visited the node. So, our new time series is given as $[(0, W_0), (\Delta, W_1 + W_2 + ..W_k), (2\Delta, W_{k+1} + W_{k+2}..), ...(n\Delta, .. + W_m)]$, where n is some multiple of Δ such that $n\Delta < t_{im}$

3.3 Travel time distribution of walkers

The speed of a walker (or vehicle) is taken from a normal distribution $\mathcal{N}(\mu, \sigma)$ with mean $\mu = 40km/hr$ and $\sigma = 5km/hr$. For every edge E_{ij} with length L_{ij} of the network, possible travel time in seconds is given by

$$T_{edge} = 3.6\left(\frac{L_{ij}}{v}\right) \tag{3.3}$$

where 3.6 is a conversion factor and v $\in \mathcal{N}(40, 5)$ Km/hr.

Distribution of T_{edge} for all considered networks is shown in figure 3.1. Since the minimum edge length of all considered networks is $10m$, so minimum travel for each is $0.7s$. From the table 3.1, the average edge travel time is the smallest for Delhi's network. We used the Delhi network for random walk simulation, and optimal Δ was found to be $5s$, closer to the average edge travel time.

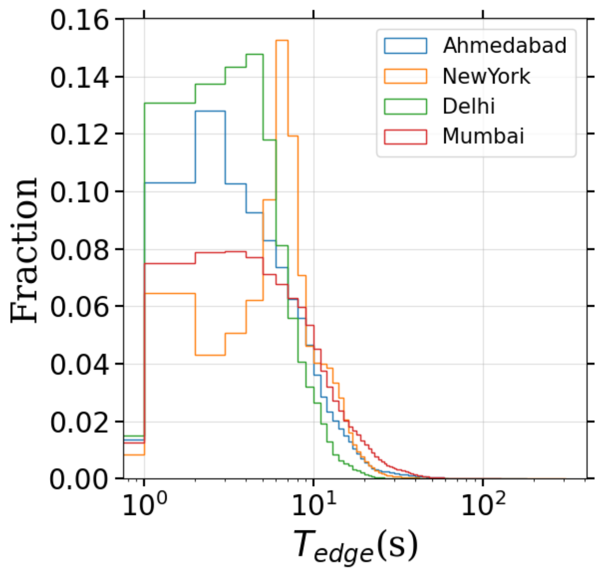


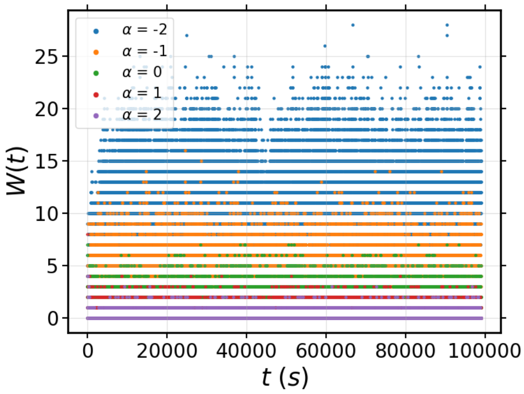
Figure 3.1: Distribution of edge travel time on different city roads in seconds s

City	Mean(s)	SD	Max	Min
Ahmedabad	8.1	8.8	181	0.7
NewYork	9.1	10.8	306.4	0.7
Delhi	5.3	3.7	61.1	0.7
Mumbai	10.2	9.4	127.6	0.7

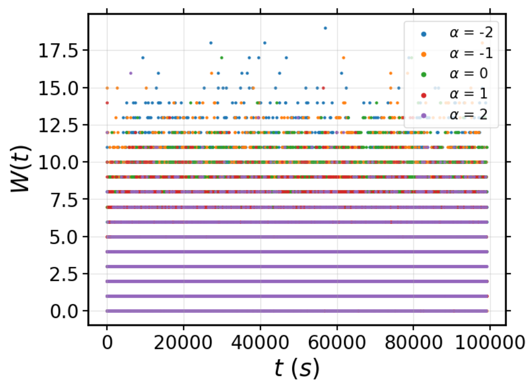
Table 3.1: Statistics of travel time on edges in seconds s

3.4 Analysis of dynamics

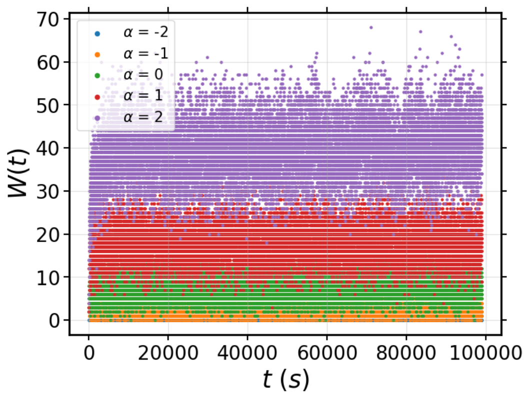
We took 5000 non-interacting walkers (or vehicles) on the Delhi city network with 1092 nodes and 1708 edges. All vehicles were allowed to traverse the network for the maximum time, $t_{max} = 120000s$. Since travel time at each edge is different so it is not possible that all walkers stopped at the same time. So, we chose $t_m = 100000s$. The walkers time series is shown in figure 3.2 for $\Delta = 5s$. The distribution of number of walkers, as shown in Figure 3.3, indicates that when $\alpha > 0$, the flux is inclined towards nodes with a higher degree, whereas for $\alpha < 0$, there is a greater number of walkers present on nodes with a smaller degree.



(a) K=1

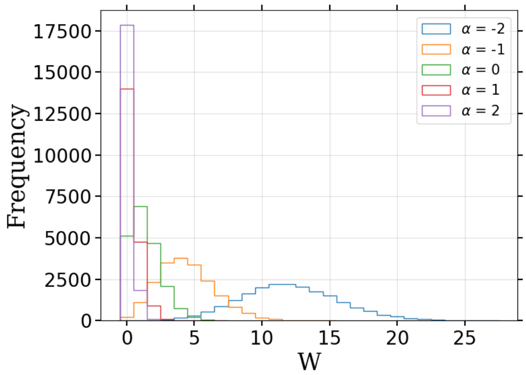


(b) K=3

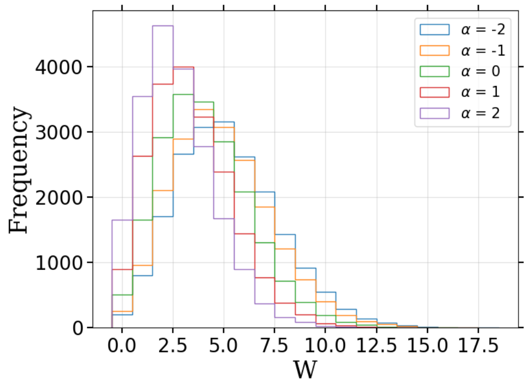


(c) K=6

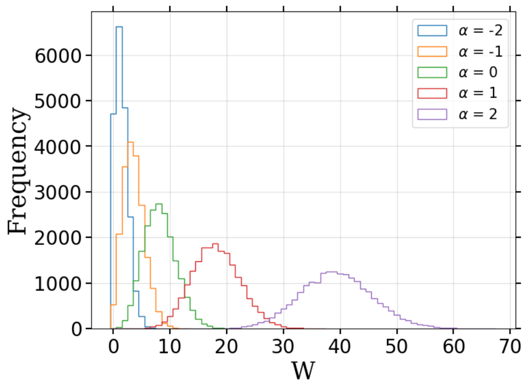
Figure 3.2: Walkers time series for node of different degree K



(a) K=1



(b) K=3



(c) K=6

Figure 3.3: Distribution of number of walkers W for nodes of different degree K

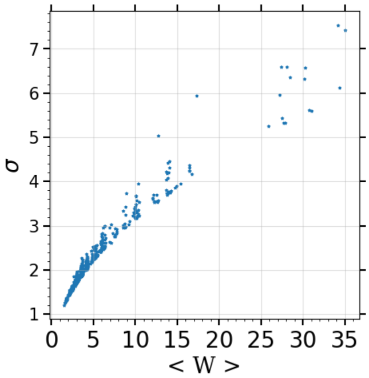
3.5 Statistics of the flux of walkers

In the paper [8], it was asserted that a flux fluctuation σ is related to average flux $< W >$ as

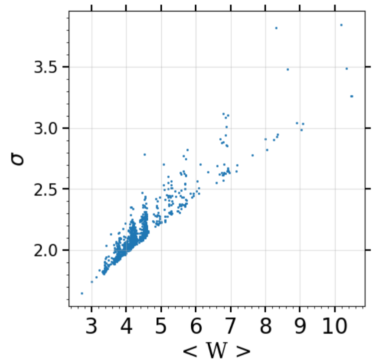
$$\sigma \sim < W >^{\beta} \tag{3.4}$$

where $\beta \in [0.5, 1]$.

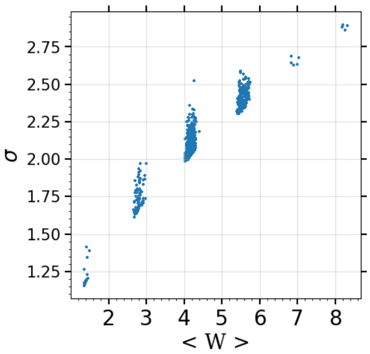
For standard discrete time random walk on the scale-free network, the average flux of walkers on a node is directly proportional to its degree K, $< W > \propto K$, while flux fluctuation varies as $\sigma \propto \sqrt{K}$, [12]. Thus, $\sigma \propto \sqrt{< W >}$, which implies $\beta = 0.5$ and it is evident from the figure 3.4.



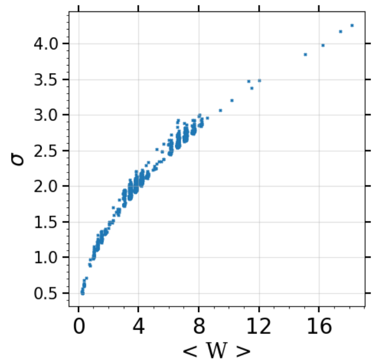
(a) α=-2



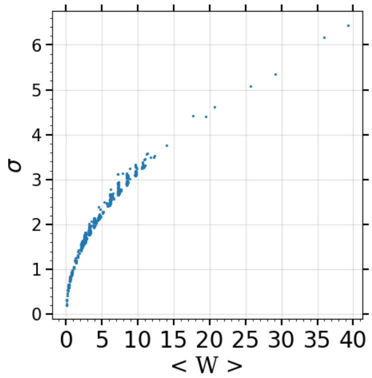
(b) α=-1



(c) α=0



(d) α=1



(e) α=2

Figure 3.4: Variation of flux fluctuation σ with average flux $< W >$, for all nodes of the network

3.5.1 Variation of flux statistics as a function of node parameter

The average flux and fluctuation passing through a node must depend on the degree of node, since higher-degree nodes are mostly visited by walkers. Figure 3.5 shows a plot of flux statistics with degree of nodes computed for all nodes. So, for a given biased parameter α, $< W_i >= \mathcal{F}(K_i)$ is not a good model. The flux through a node will also depend on the degree of neighbouring nodes. We introduced a new parameter [13] generalized strength ϕ, a measure of the ability of a node to attract walkers. The generalized strength is a function of the degree of the node and that of its nearest neighbours.

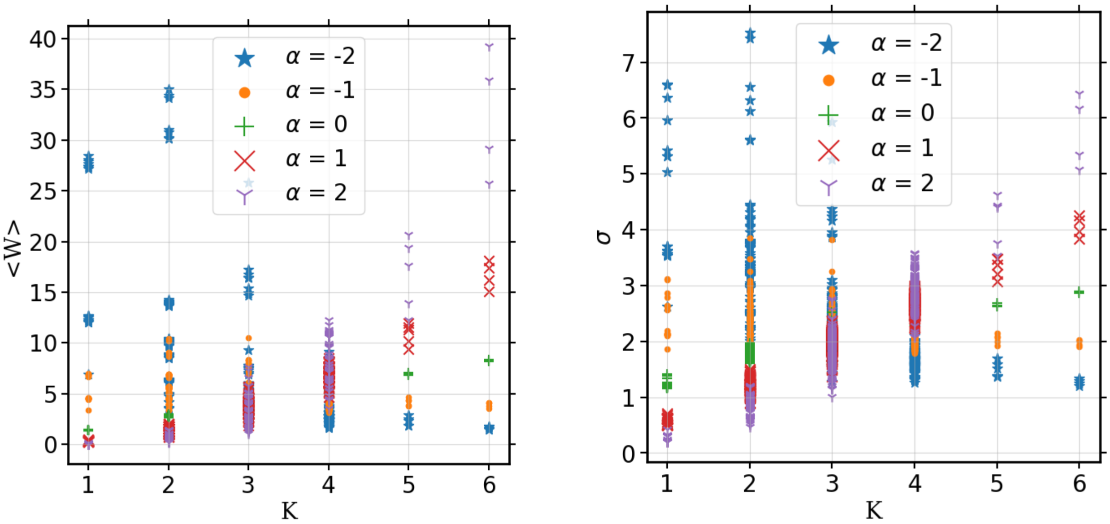
$$\tilde{\phi}_i = K_i^\alpha \sum_{j=1}^{K_i} K_j^\alpha \tag{3.5}$$

Note that ϕ_i depends on the bias parameter α and the degree of the nearest neighbours connected by an edge. So, nodes with the same degree can have different generalized strengths, which is evident from figure 3.6. The generalized strength of a node is determined by the local connectivity structure surrounding the node and is independent of the overall network structure. This differs from the behaviour of standard random walk on networks, where the larger-scale network topology does not play a significant role. In biased random walks on networks, however, the local network structure plays an important role.

To have generalized strength on the same scale for different values of α, we defined normalized generalized strength ϕ_i for every biased value α

$$\phi_i = \frac{\tilde{\phi}_i}{\sum_{j=1}^{N} \tilde{\phi}_j} \tag{3.6}$$

In Fig. 3.6, we show how the generalized strength ϕ depends on the degree of a node for several values of α in a planar network. For $\alpha = 0$, which is the standard random walk case, the generalized strength is the same for the same degree node. From the figure 3.7, average number of walkers shows a linear relationship with generalized strength(ϕ) irrespective of α, i.e $< W_i > \propto \phi$ and $\sigma_i \propto \phi^\beta$ thus holding universality of relation $\sigma \propto < W >^\beta$, $0.5 \leq \beta \leq 1$ see [8].



(a) Average flux $< W >$ vs K **(b)** Flux fluctuation σ vs K

Figure 3.5: Variation of flux statistics with degree K, computed for all nodes

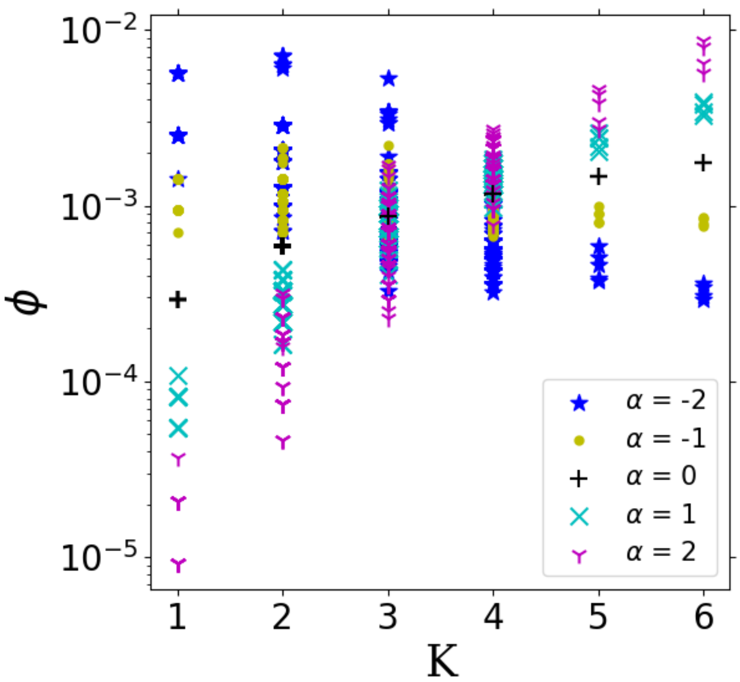
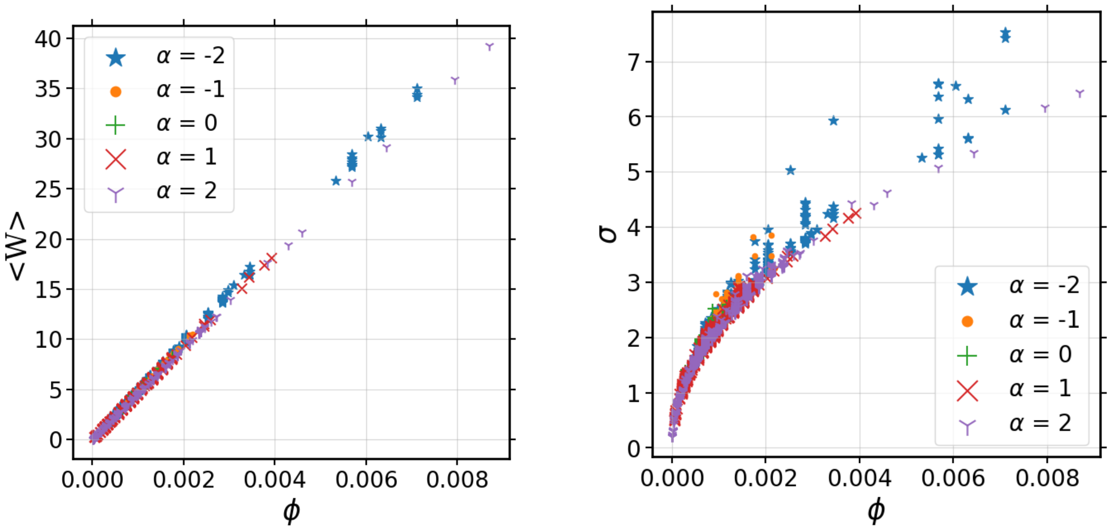


Figure 3.6: Generalized strength ϕ as function of degree K



(a) Average flux $< W >$ as a function of ϕ **(b)** Flux fluctuation σ as a function of ϕ

Figure 3.7: Flux statistics as a function of generalized strength ϕ

3.6 Extreme events (EE) statistics

In the context of extreme value statistics, an extreme event refers to a rare occurrence with a low probability of happening, usually located at the tail of the probability distribution function. In the context of a network, we will adopt the same concept for each node. Specifically, we will consider an event as extreme if the number of walkers passing through a particular node at any time exceeds a threshold value, q. We modelled threshold $q$for node i as

$$q_i \ = \ \langle W_i \rangle + \ m\sigma_i \tag{3.7}$$

where $\langle W_i \rangle$ is the average flux & σ_i is flux fluctuation through node i, and m is a real number.

We assert that threshold q should depend on node statistics rather than a constant for all nodes of the network because when a uniform threshold is applied regardless of the node, it can result in certain nodes consistently experiencing extreme events while others never encountering any extreme events at all, [12][13]. So, the probability of occurrence of an extreme event on node i is given by,

$$P_{EE}^{i} = \frac{1}{n}\sum_{t=0}^{n\Delta} \theta(W_i(t) - q_i) \tag{3.8}$$

where $W_i(t)$ is a time series of walkers on node i, n is the number of time steps in the time series, q_i = threshold for node i and $\theta(x)$ is a Heaviside function defined as

$$\theta(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} \tag{3.9}$$

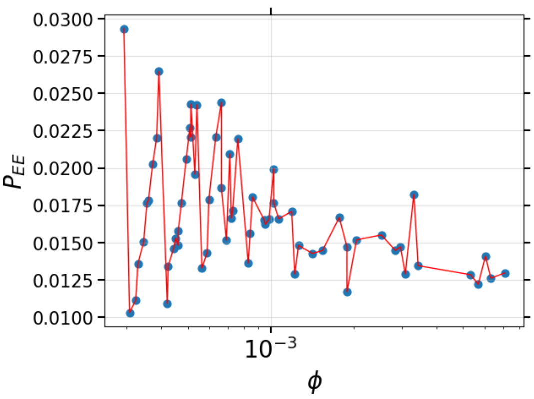
3.6.1 Probability of occurrence of extreme events as a function of generalized strength

From figure 3.7, we can clearly see that the flux statistics through a node are proportional to its generalized strength value. It is better to see the probability of occurrence of an extreme event as a function of generalized strength ϕ, averaged over nodes having the same value of ϕ_i.

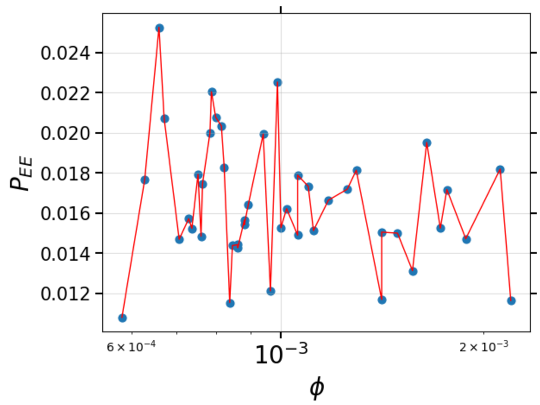
If there are N_i number of nodes having the same value of generalized strength ϕ_i, let $\{\phi_i\}$ denotes the set of nodes with the same value of generalized strength, then, the probability of occurrence of an extreme event is given by

$$P_{EE}(\phi_i) = \frac{1}{N_i}\sum_{j\in\{\phi_i\}} P_{EE}^{j} \tag{3.10}$$

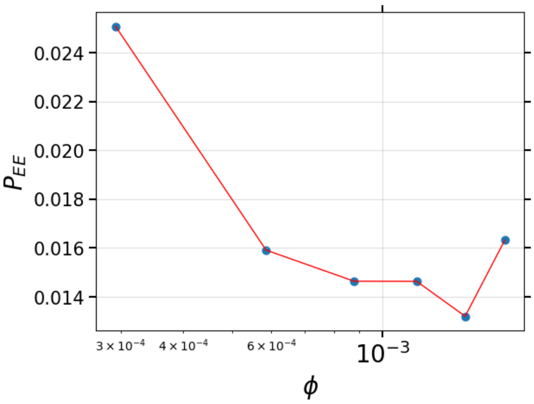
The simulation results depicted in Figure 3.8 display how the probability of extreme events varies as a function of ϕ for different values of α. Notably, the results indicate that nodes with lower generalized strength ϕ values tend to have a higher predicted probability of encountering extreme events, on average, than nodes with higher ϕ values. This is an unexpected observation. These results are in accordance with, that were obtained in the paper by Kishore et al., 2012 [13] for the discrete degree biased random walk on SF networks, which shows that extreme events are more probable for nodes which attract large flux (on average).



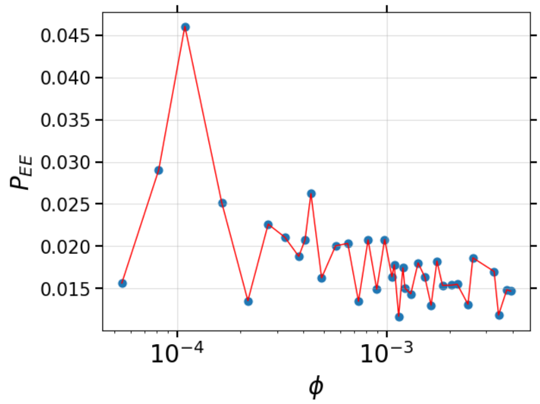
(a) α=-2



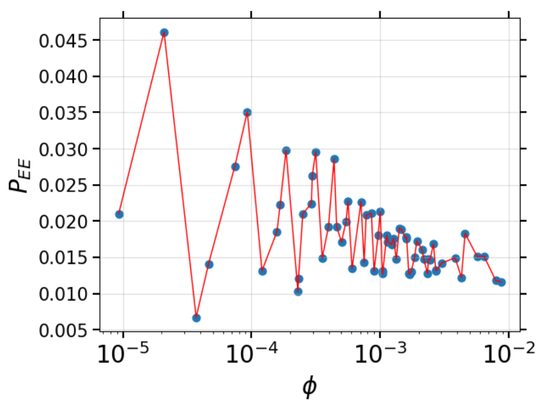
(b) α=-1



(c) α=0



(d) α=1



(e) α=2

Figure 3.8: Variation of probability of occurrence of extreme event P_{EE} as a function of generalized strength ϕ

Chapter 4

Interactive Random Walk Routing And Congestion Phenomenon on Unweighted Spatial Networks of Urban Streets

4.1 Introduction

Traffic congestion is a condition in transport that leads to slower speeds, longer trip times, and increased vehicular queuing. Traffic congestion in a city can arise due to its large population, limited road infrastructure, underlying road network and high number of vehicles on the road. Major cities of a country are more prone to congestion mainly due to a high number of automobiles and inefficient traffic routing strategy, resulting in large waiting hours, pollution and waste of energy and fuel. It is easier to execute soft strategies rather than hard strategies for traffic management. Soft strategies involve designing of efficient routing protocol, on the other hand, modifying the underlying network structure comes under hard strategies. Hard strategies are cost and time expensive and hence hard to implement.

Apart from a standard random walk, Computer scientists have proposed many routing algorithms for the efficient transport of data packets in the communication network, one of which is the shortest-path routing protocol (SPRP). The real transport networks are queuing networks, where each node generates packets or vehicles and delivers to their destination node via a suitable routing protocol. In [9] author proposes a routing strategy incorporating congestion level on neighbouring nodes with SPRP. If j is the destination node for a packet generated on node i, then neighbouring node l is chosen such that the quantity:

$$d_l^{eff} = hd_{ij} + (1 - h)c_l, \tag{4.1}$$

is minimum, where d_{ij} is the minimum number of hops one has to pass by in order to reach j, and h is a tunable parameter that accounts for the degree of traffic awareness incorporated in the delivery algorithm ($h = 1$, recovers SPRP, no congestion control). It was verified that algorithms which integrate topological and traffic information have been shown to perform better than the standard protocol (SPRP) in terms of efficiency and packet delivery. Later on, many other protocols were proposed, including probabilistic routing protocols [14]. But, there is a need for a parameter to quantify the effectiveness of the protocol in the sense that all packets generated with rate $\mathcal{P}$ reach their destination or, in other words, can the network process all the incoming data packets. If it can do it, the total number of packets in the system $A(t)$ will be stationary in time or in free-flow regime, if it cannot, $A(t)$ will grow in time or the network will be in congested phase. One useful metric for distinguishing between the two phases is to divide the average growth rate of the queues $\langle \dot{A} \rangle$ by the average rate $\bar{\mathcal{P}}$ at which packets are received. This yields the fraction of packets that are either stuck in queues or not delivered,

$$\rho = \frac{\langle \dot{A} \rangle}{\bar{\mathcal{P}}}. \tag{4.2}$$

In the paper, Echenique et al. [10], author came up with two scenarios, in one where purely topological routing, such as routing along the shortest paths, is employed, the emergence of congested phases occurs in a continuous manner. However, when a traffic-aware scheme is incorporated, the transition from uncongested to congested phases becomes discontinuous, shown in fig. 4.1.

Aiming to analyze a wide range of congestion phenomena from a statistical perspective showing continuous, discontinuous, and hybrid phase transitions between a free-flow and a congested phase, a generalized random walk routing protocol was de-

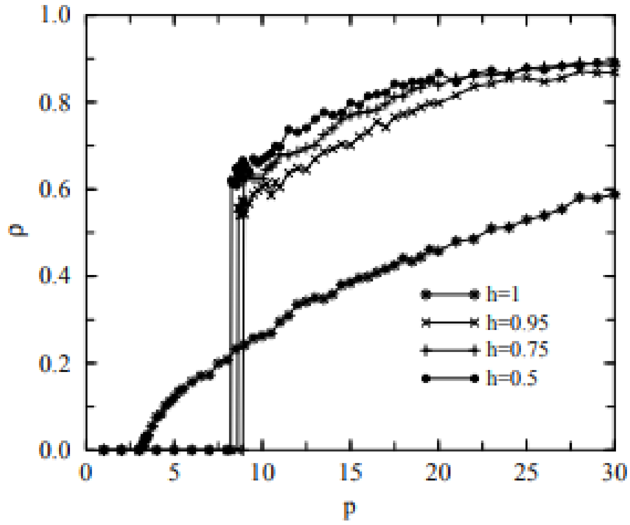


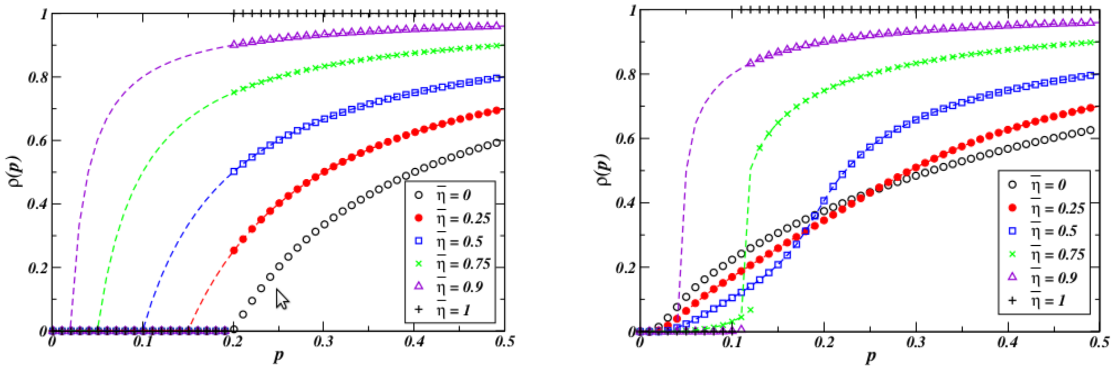
Figure 4.1: Phase transition diagram for traffic protocol with different congestion control parameter h, taken from[10]

veloped, simple enough to obtain analytical results. This protocol incorporates key elements from real-world routing schemes, including a balanced approach between topology-based and traffic-based routing strategies. This generalized routing protocol was introduced with the intention of studying congestion dynamics comprehensively[]. The model was capable of replicating the observed cross-over in the scaling of traffic fluctuations in the Internet, see Fig. 4.1, and furthermore, a modified version of the model can qualitatively replicate certain stylized facts of traffic behaviour in transportation networks.

Here, we adopted the model to study the congestion phenomenon on a real planar network (city road network). The congestion phenomenon on the scale-free network and grid network are well studied and shows the following phase transition behaviour, shown in fig. 4.2.

4.2 Interacting random walk routing model

Let us consider a network of N nodes and let $\{\Omega_i\}$ be the set of neighbours of node i. At each time step, particles are generated at each node i with probability p_i. Each node is endowed with a FIFO queue in which arriving particles are stored. Let n_i be the number of particles in the queue of node i. If $n_i > 0$, at most r_i topmost particles can leave the node and jumps in the queue of a randomly chosen neighbour $j \in$



(a) Random regular graph, degree $K = 4$ **(b)** Scale-free network, exponent $\gamma = 2.5$

Figure 4.2: Phase transition diagram for synthetic network, each with 1000 nodes. The symbols represent the behaviour for increasing values of birth probability p, $\mu = 0.2$, $n^* = 10$ and different congestion control parameter η. The dashed curve is obtained by decreasing the value of p, taken from [7]

$\{\Omega_i\}$. The arrival node rejects the particle with probability $\eta(n_j)$, and it stays on the departure node. If not rejected, the particle is either destroyed during the hopping with a probability μ_j or enters the queue on node j, schematic for the dynamics is shown below 4.3.

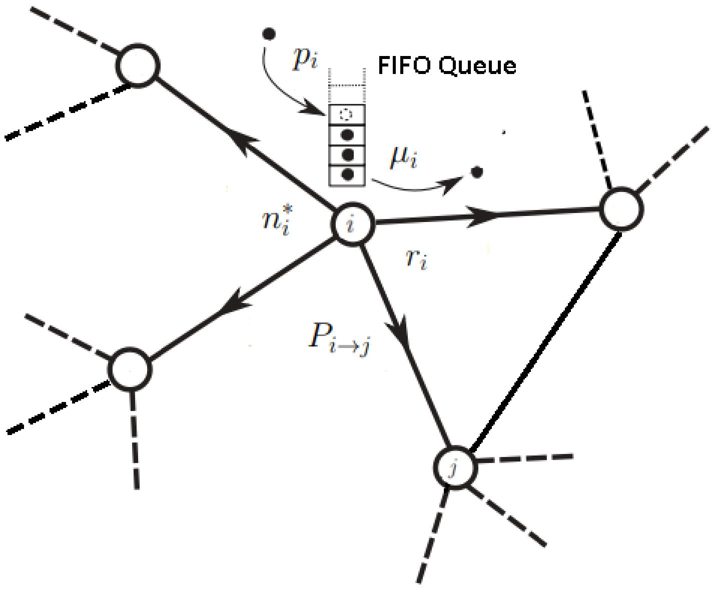


Figure 4.3: For a node i in queuing network, p_i (Creation rate), r_i (Service rate or outflux), μ_i (Absorbing rate), $P_{i\to j}$ (Transition probability)

For simplicity, we consider $p_i = p$, $\mu_i = \mu$ and $\eta(n_i) = \bar{\eta}\theta(n_i - n_i^*)$ where $\theta(x)$ indicates the Heaviside step function for all node i. Thus, as per the routing protocol, node i refuses to accept a walker with probability η if $n_i \geq n_i^*$. Thus, here n_i^* is the threshold queue length for i^{th} node.

The primary objective of this model is to examine the statistical properties of queue lengths, rather than individual particle trajectories. This approach enables us to focus on probabilistic events and overlook the fate of individual particles generated and routed to their destination. The steady state of the system is governed by the balance between the creation and absorption rates. When creation rates exceed absorption rates, queues become saturated with particles, leading to congestion and indefinite growth of the queues. However, as the external drive that creates particles remains constant over time, the system eventually reaches a non-equilibrium steady state where the queues grow at a constant rate. To characterize the phase transition from a free phase to a congested one, we use the order parameter ρ defined as

$$\rho = \lim_{t\to\infty} \frac{A(t+\tau) - A(t)}{Np\tau}, \tag{4.3}$$

where τ is the observation time, $A(t) = \sum_i^N n_i$, is the total number of particles in the system at time t, Np is the average number of particles created at each time step and τ is the observation time. Note that p $= \frac{\sum_i p_i}{N}$ is the average birth probability and N is the number of nodes. System is said to be congested in the steady state if the order parameter is greater than 0. The higher the value of ρ higher is the level of congestion, $\rho = 1$ corresponds to a highly congested state, while $\rho = 0$ means no congestion or free flow. We can determine the congestion state of any node i by computing its local order parameter just by replacing $A(t)$ by $n_i(t)$ and p by p_i. In this work, we considered two scenarios

- **Constant parameters values**: All parameters of the model are kept same for all the nodes of the network.
- **Degree dependent outflux and node capacity**: Only outflux r_i and node capacity n_i^* are made degree dependent.

4.3 Analysis of the model dynamics with constant parameter values

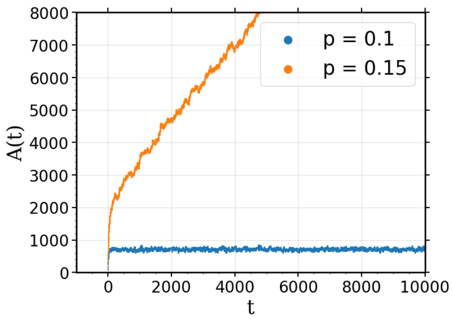
Here, we consider $p_i = p$, $\mu_i = \mu$, $r_i = r$, and $n_i^* = n^*$. The rejection probability is given by

$$\eta(n_i) = \bar{\eta}\theta(n_i - n^*), \tag{4.4}$$

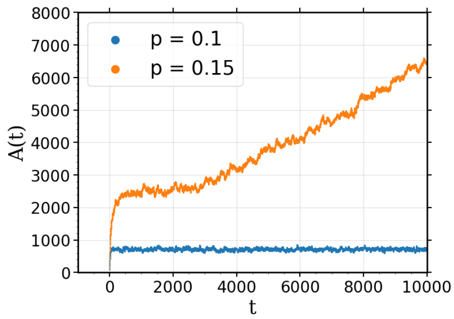
where $\bar{\eta} \in [0, 1]$ and $\theta(x)$ is a unit step function as defined in chapter 3, equation 3.9. The transition probability for a walker to go from node i to j, can be written as

$$P_{i \to j} = \frac{A_{ij}(1 - \eta(n_j))}{k_i}. \tag{4.5}$$

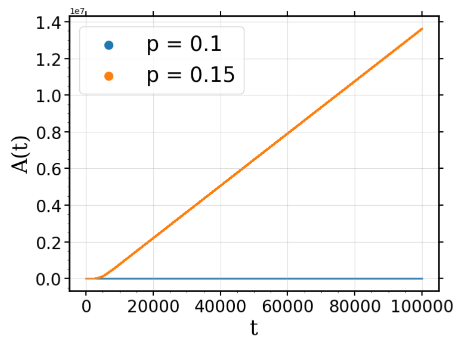
Note that for $\bar{\eta} = 0$, we recover the standard random walk with no congestion control.



(a) $\bar{\eta} = 0$

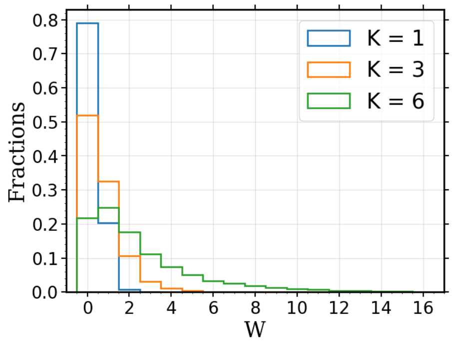


(b) $\bar{\eta} = 0.5$

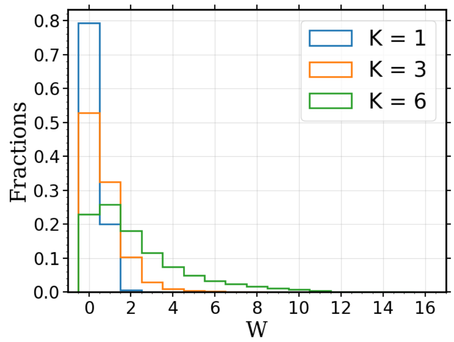


(c) $\bar{\eta} = 0.9$

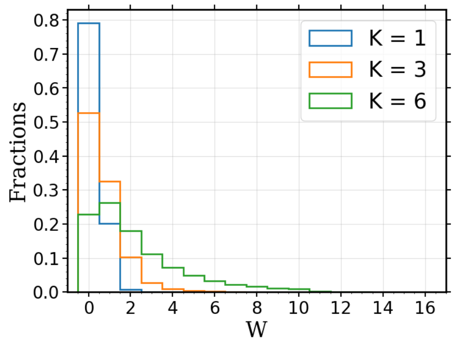
Figure 4.4: Variation of $A(t)$ with time t, $\mu = 0.2$, $n^* = 10$ and different congestion control parameter $\bar{\eta}$



(a) $\bar{\eta} = 0$

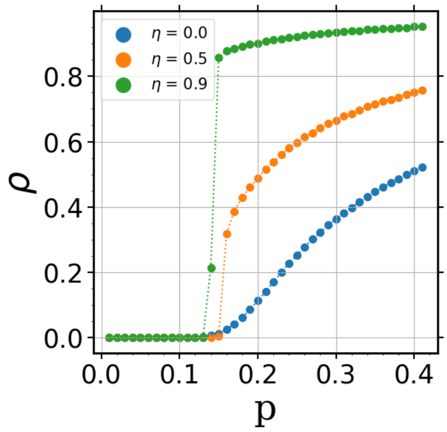


(b) $\bar{\eta} = 0.5$

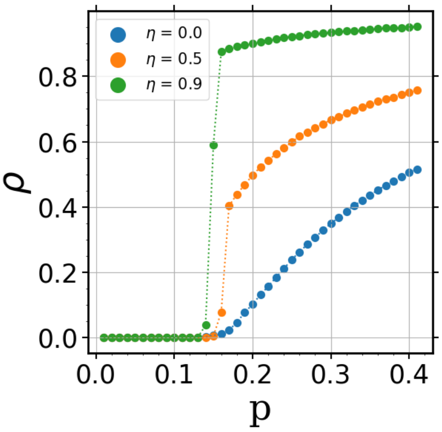


(c) $\bar{\eta} = 0.9$

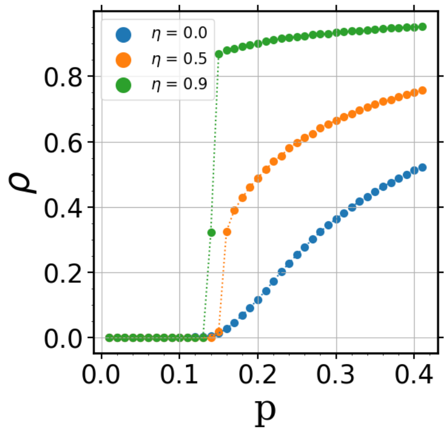
Figure 4.5: Distribution of number of walkers on nodes of different degree K, $p = 0.1$, $\mu = 0.2$, $n^* = 10$, and different congestion control parameter $\bar{\eta}$



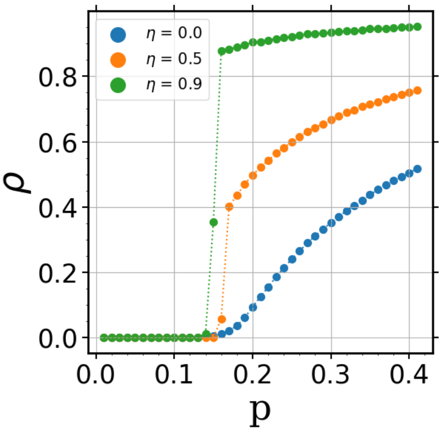
(a) Mumbai city network, 1071 nodes



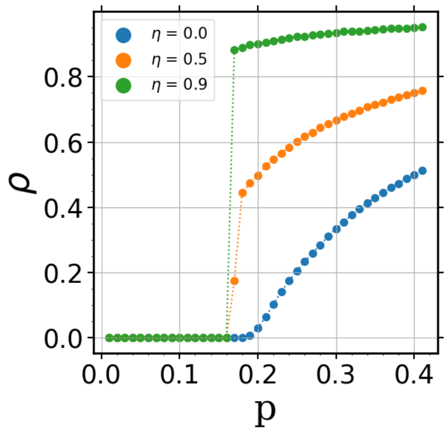
(b) Delhi city network, 1092 nodes



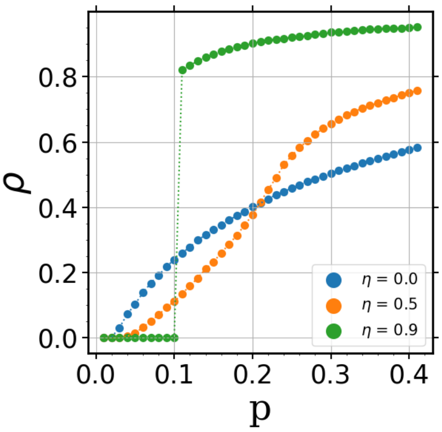
(c) Ahmedabad city network, 1064 nodes



(d) New York city network, 1117 nodes



(e) 2D grid network,1000 nodes



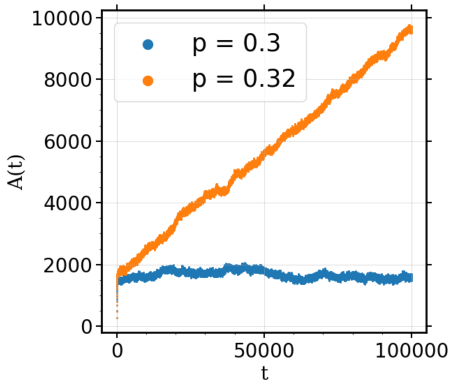
(f) Scale-free network,1000 nodes

Figure 4.6: Order parameter ρ as a function of birth probability p, $\mu = 0.2$, $n^* = 10$ and $r_i = 1$, and different rejection probability η

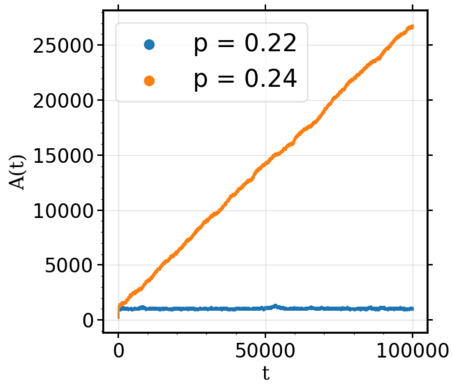
4.4 Analysis of the model dynamics when outflux r_i and node capacity n_i^* are degree dependent

Here, we consider $p_i = p$, $\mu_i = \mu$, $r_i = b_0 k_i$, and $n^* = a_0 k_i$, where k_i is degree of node i, a_0 and b_0 are nonzero-positive integers such that $b_0 \leq a_0$. So, the rejection probability is modified to

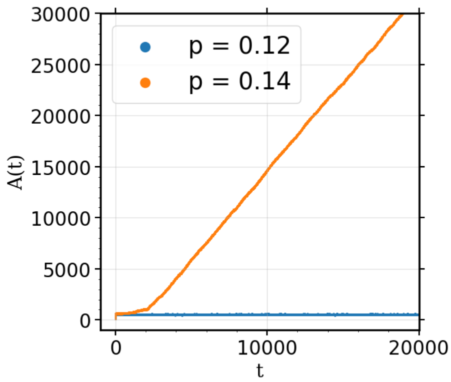
$$\eta(n_i, k_i) = \bar{\eta}\theta(n_i - n^*(k_i)), \tag{4.6}$$



(a) $\bar{\eta} = 0$



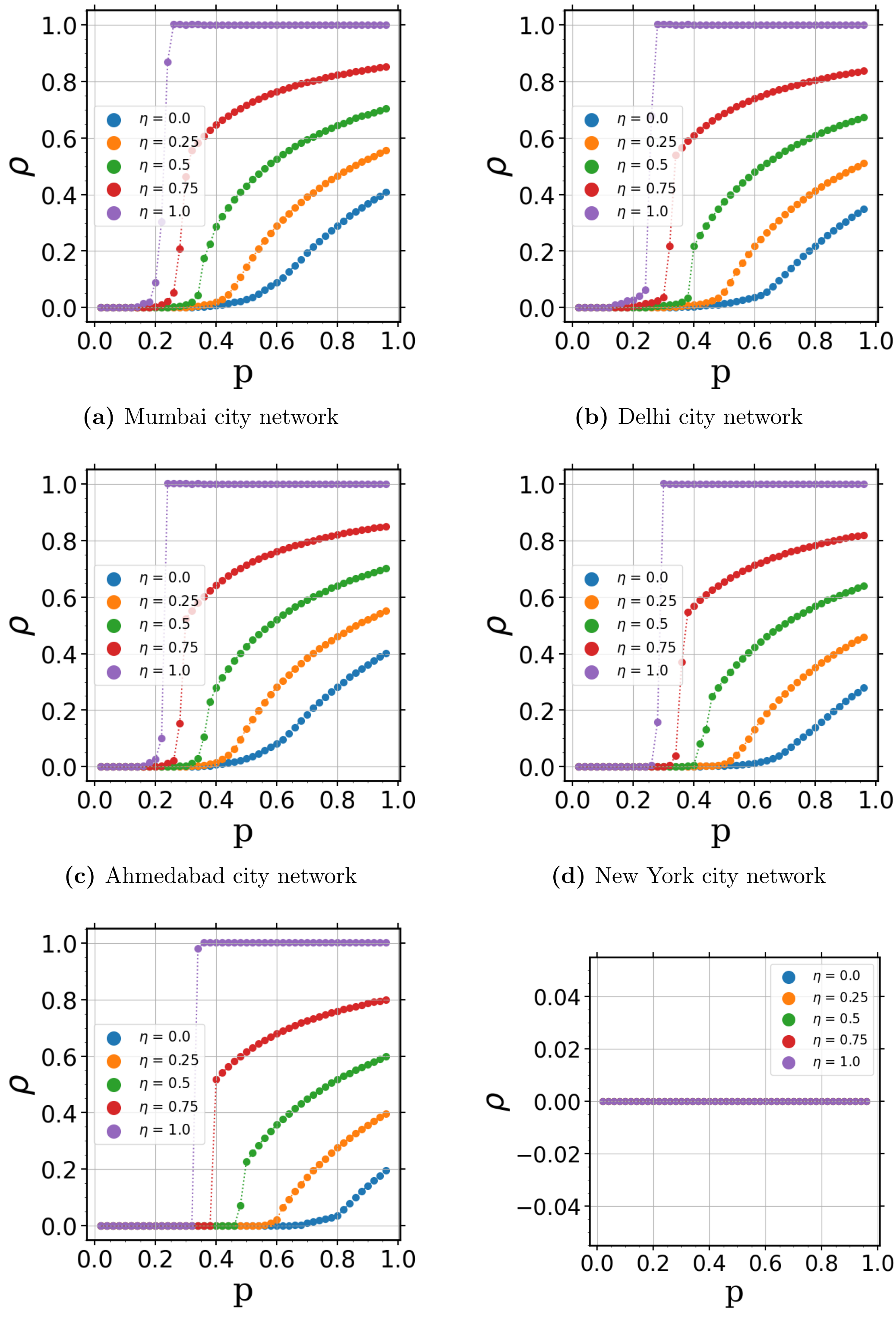
(b) $\bar{\eta} = 0.5$



(c) $\bar{\eta} = 1$

Figure 4.7: Variation of $A(t)$ with time t, $\mu = 0.2$, $r_i = k_i$, $n^* = k_i$ and different congestion control parameter $\bar{\eta}$

When node capacity n_i^* and outflux r_i are made degree dependent, congestion regime completely disappears for scale-free networks, while for planar networks, it still exists though critical birth probability (p_c) shifted to a higher value as shown in fig. 4.8.



(a) Mumbai city network

(b) Delhi city network

(c) Ahmedabad city network

(d) New York city network

(e) 2D grid network

(f) Scale-free network

Figure 4.8: Order parameter ρ as a function of birth probability p, $\mu = 0.2$, $n^* = k_i$ and $r_i = k_i$, for different rejection probability η

Considering parameter values $p = 1$, $\mu = 1$, $\bar{\eta} = 1$, $r_i = b_0 k_i$, and $n^* = a_0 k_i$, it would be interesting to see for what values of a_0 and b_0, we see no congestion on planar networks. As a matter of computational complexity, we took a_0 up to 9. For different combination of a_0 and b_0, we find mainly four behaviour,

1. **Linear variation**

   In this case, total number of walkers, $A(t)$ linearly increasing with time t, as shown in fig. 4.10(a).

2. **Piecewise linear**

   Linear variation with a kink at a certain time t_k, shown in fig. 4.10(b) for $b_0 = 4$ and $a_0 = 5$.

3. **Congested and free-flow phase coexist**

   In paper [7], it was observed that for high $\bar{\eta}$, there is the coexistence of two phases in a certain range of p, shown in fig. 4.9. We also found the coexistence of two phases, as shown in fig. 4.10(b) for $b_0 = 4$ and $a_0 = 7$, where system enters in congested state after critical time, t_c.

4. **Free flow**

   $A(t)$ fluctuates about some mean value with $\dot{A}(t) = 0$, after some small transient time.

We grouped these four behaviour of variation of $A(t)$ with time t into three cases,

- Highly congested (includes case 1)
- Weakly congested or two phases coexist (includes cases 2 and 3)
- Free flow (includes case 4)

We have shown three different cases by a heatmap for different combinations of a_0 and b_0, shown in fig. 4.11. New York shows similarity with 2D lattice network. It can be said that all cities show quite similar behaviour.

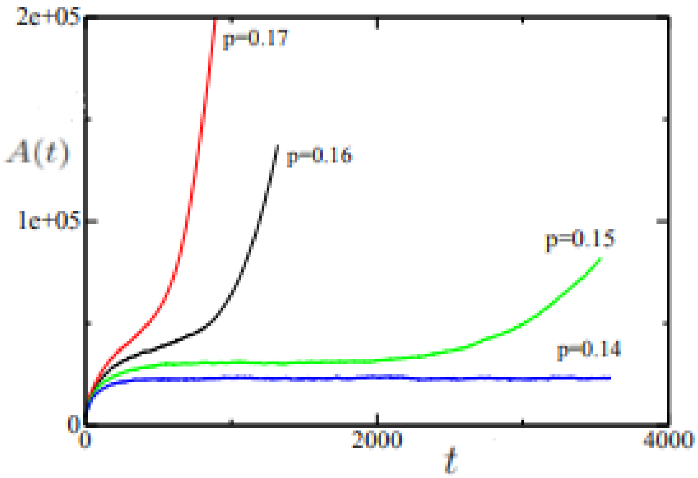
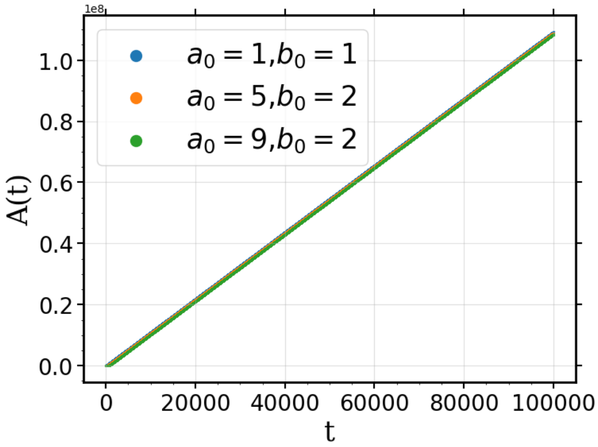
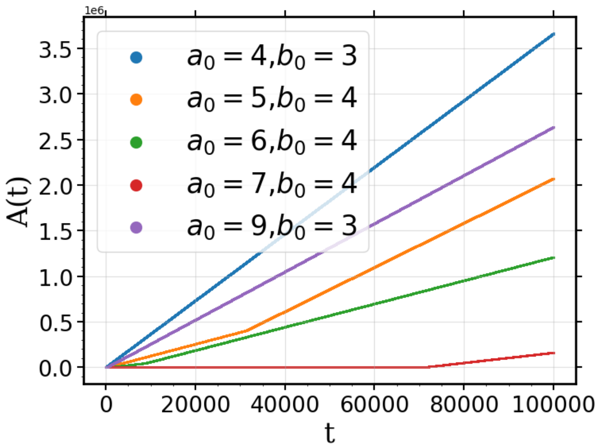


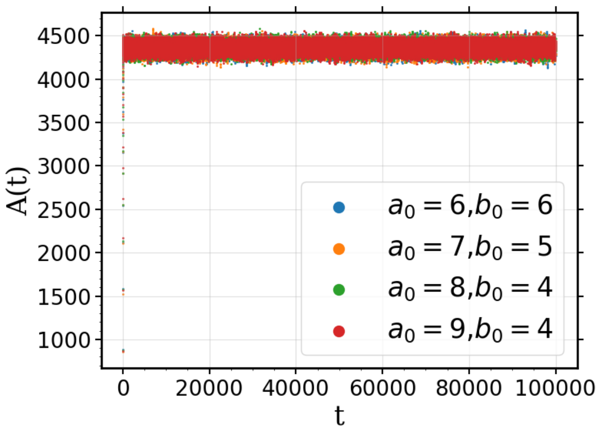
Figure 4.9: Variation of total number of particles $A(t)$ with time t for different birth probability p, $\mu = 0.2$, $r_i = 1$, and $\bar{\eta} = 0.9$, taken from [7]



(a) Highly congested

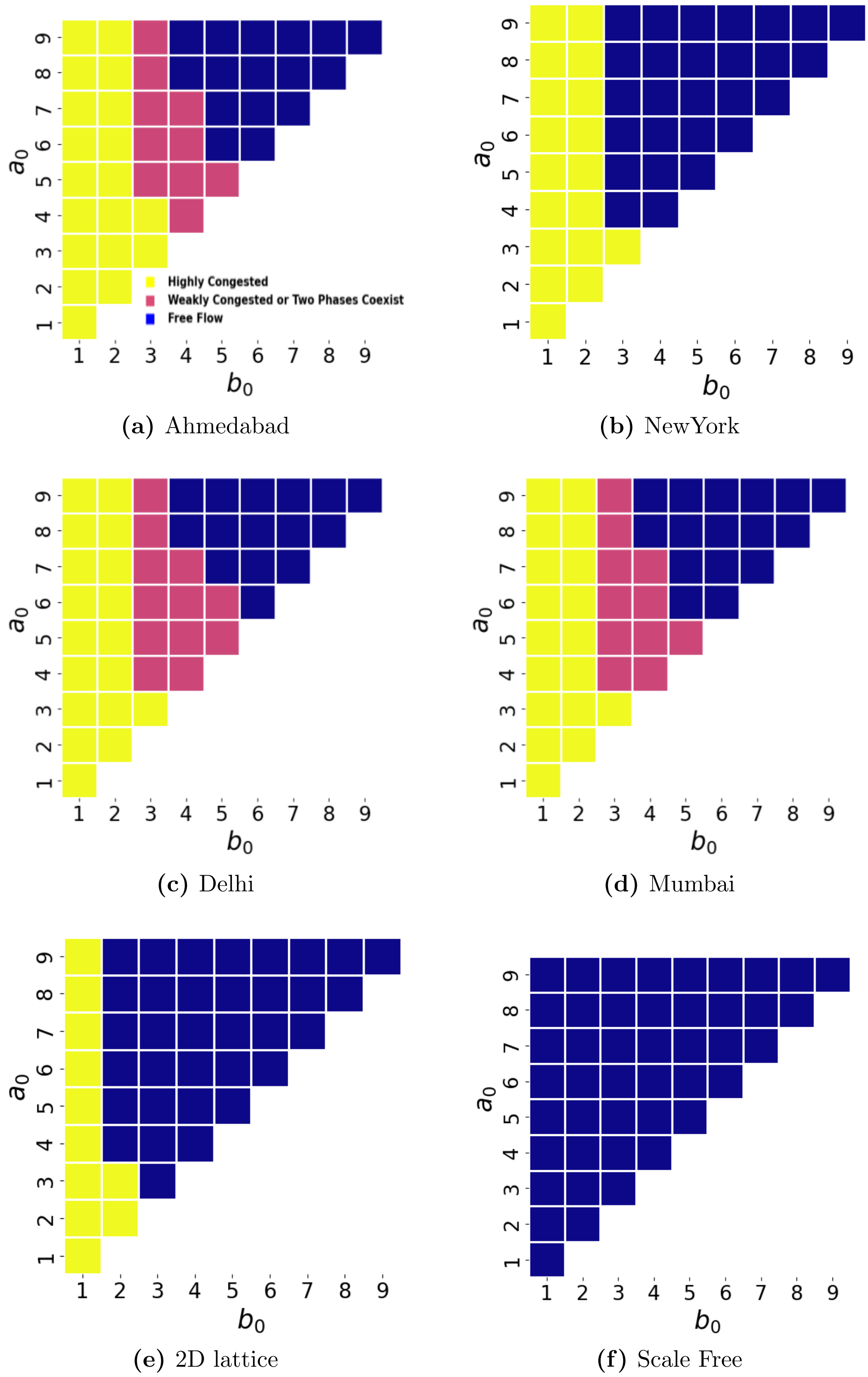


(b) Weakly congested or two phases coexist



(c) Free Flow

Figure 4.10: Variation of $A(t)$ with time t, $p = 1$, $\mu = 0.2$, $n^* = a_0 k_i$, $r_i = b_0 k_i$ and $\bar{\eta} = 1$



(a) Ahmedabad (b) NewYork (c) Delhi (d) Mumbai (e) 2D lattice (f) Scale Free

Figure 4.11: Heat map showing level of congestion for different values of a_0 and b_0, $p = 1$, $\mu = 0.2$, $n^* = a_0 k_i$, $r_i = b_0 k_i$, and $\bar{\eta}$ = 1

Chapter 5

Interactive Random Walk Routing and Extreme Events on Unweighted Spatial Networks of Urban Streets

5.1 Introduction

In the previous chapter, we studied congestion phenomenon on queuing networks. In this chapter, we will study extreme events adopting the model introduced in the previous chapter. As we discussed before, the extreme event is the other emergent phenomenon on networks. Extreme events, are defined as exceedance above a prescribed quantile. Unlike congestion, it is not necessary that extreme event emerges due finiteness of node size. EE arise from the natural fluctuation in the traffic passing through a node or a junction and not due to external constraints or the finiteness of the junction. People adopted the simple random walk model and its different versions to study extreme events, [12] [13]. In chapter 2, we studied extreme events by adopting non-interacting continuous time degree biased random walk model on city road transportation networks. Contrary to expectation, it was observed that the nodes which attract large flux experience less extreme events than those which attract small flux or vehicles. But, there is a need for a more realistic routing protocol for transport dynamics. It was found that even the intelligent routing protocol (SPRP) doesn't change the nature of the result [12]. The actual dynamics of transportation are interactive in nature and, therefore, require queuing networks. Hence, we have adopted the generalized random walk model on queuing networks from [6], which has already been described in chapter 4 of this thesis. The model is non-conservative in nature. Hence the total number of vehicles in the system does not remain constant over time. Instead, it may fluctuate about an average or linearly increase with time depending on the specific values assigned to the model parameters.

The presence of a stationary distribution is an essential requirement for determining the EE probabilities. Hence, if a system is in a non-equilibrium steady state, it is not possible to calculate the EE probabilities. It is possible to determine the EE probabilities for a system that exhibits only slight variations in the total number of walkers. In a free flow state, the system achieves a stationary state in which there is a small fluctuation of walkers about some average value. Consequently, our analysis of the extreme event probabilities focuses on the free-flow state of the system.

5.2 Interacting random walk model and analysis of dynamics

For transport dynamics, we adopt an interacting random walk model on four urban street networks of Ahmedabad, New York, Delhi and Mumbai. As a review, our model parameters are of the form $p_i = p$, $\mu_i = \mu$, $r_i = b_0 k_i$, and $n^* = a_0 k_i$, where k_i is the degree of node i, a_0 and b_0 are nonzero-positive integers such that $b_0 \leq a_0$. So, the rejection probability is modified to

$$\eta(n_i, k_i) = \bar{\eta}\theta(n_i - n^*(k_i)), \tag{5.1}$$

where $\bar{\eta} \in [0, 1]$.
We would like to assert that when road junction reached its capacity and becomes fully occupied, no other vehicle can enter it or there is a complete rejection i.e $\eta(n_i) = 1$ for $n_i > n_i^*$. so, to include this realistic feature, we will take $\bar{\eta} = 1$.

We fixed our parameter values $p = 1$ (generating vehicles at full capacity), $\mu = 0.2$, $r_i = b_0 k_i$, $n_i^* = a_0 k_i$ & $\bar{\eta} = 1$. In order to study EEs, the system should be in free flow regime. Hence, we look for values of a_0 and b_0, for which the system is in free-flow regime. For Ahmedabad, Delhi and Mumbai, the free-flow state is found at $a_0 = 6$ and $b_0 = 6$, while for New York, $a_0 = 4$ and $b_0 = 4$, see fig. 4.11

Fig. 5.1 shows the variation of the total number of walkers, $A(t)$, as a function of time for different cities. The distribution of number of walkers is shown in fig. 5.2. [1]

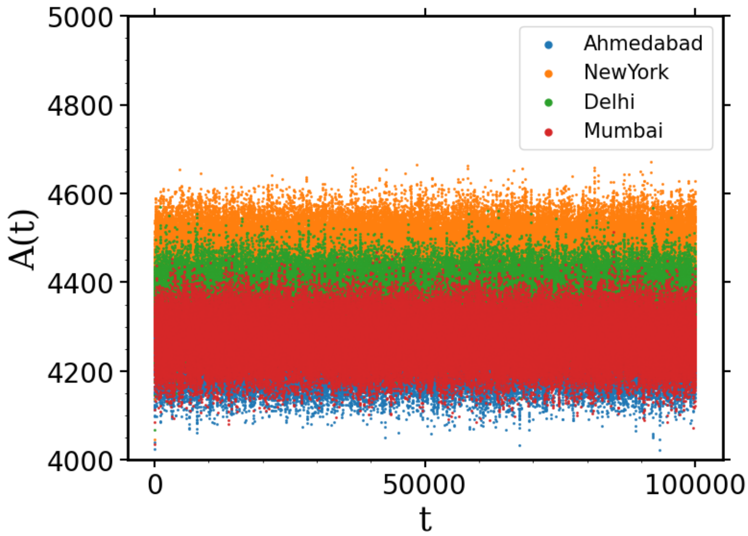


Figure 5.1: Variation of $A(t)$ with time t, p=1, $\mu = 0.2$, $n^* = a_0 k_i$ and $r_i = b_0 k_i$, and $\bar{\eta}$=1

5.3 Statistics of flux of walkers

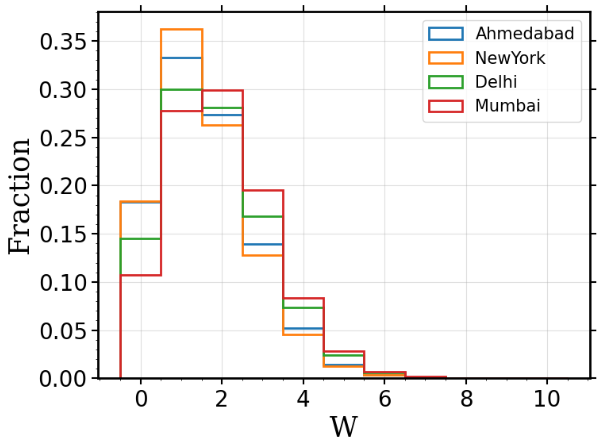
The variation flux fluctuations σ with average flux $\langle W \rangle$ depicts the behaviour of the system. Here, the behaviour is quite different and interesting, as shown in fig. 5.3

5.3.1 Variation of flux statistics as a function of node parameter

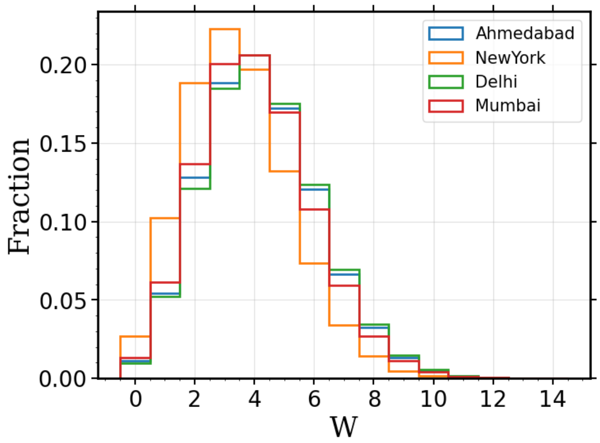
It is good to see how average flux and fluctuations vary as a node parameter. By intuition, a good node parameter is its degree K. In fig. 5.4, we plotted the average flux $\langle W \rangle$ and fluctuation as a function of the degree of nodes (average over nodes with the same degree).

Let us now see how fluctuations σ_K (average over nodes with the same degree) vary with average flux, $\langle W \rangle$. In fig. 5.5, curves coincide for all cities.

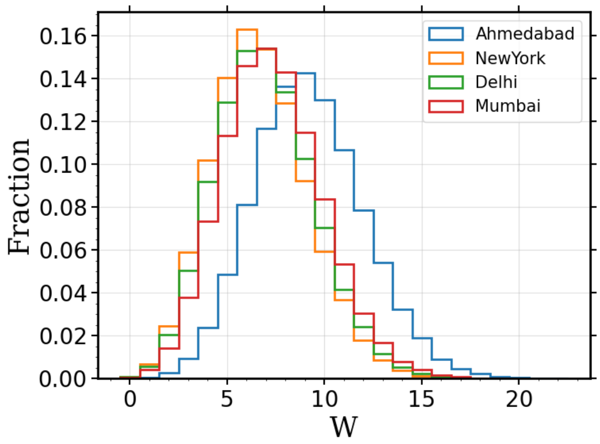
[1] $(a_0, b_0) = (6, 6)$ for Delhi, Mumbai and Ahmedabad and $(a_0, b_0) = (4, 4)$ for New York



(a) $K = 1$

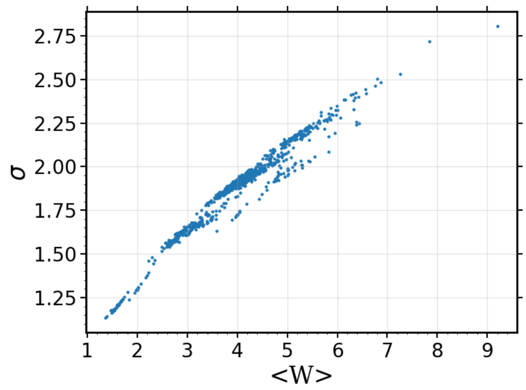


(b) $K = 3$

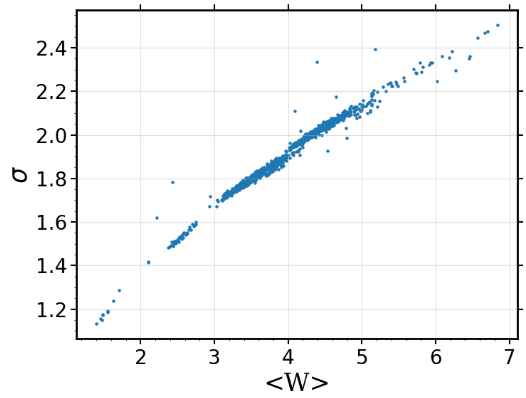


(c) $K = 6$

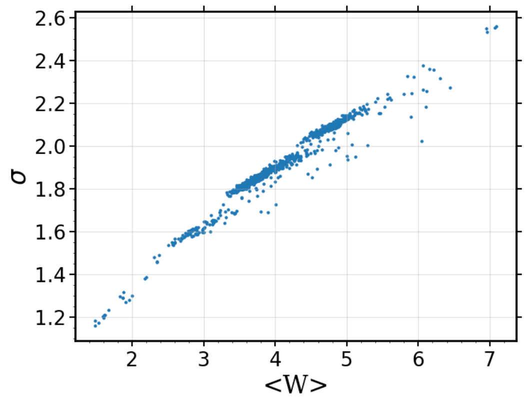
Figure 5.2: Distribution of number of walkers W on nodes of different degrees K



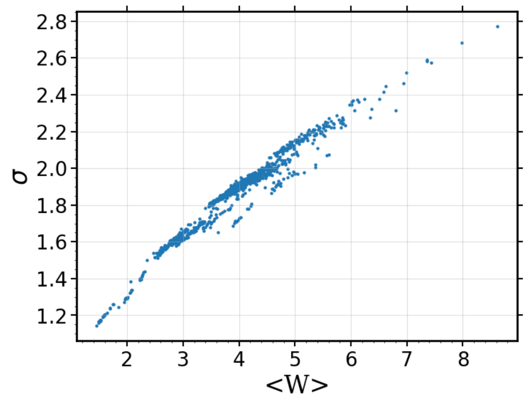
(a) Ahmedabad



(b) NewYork

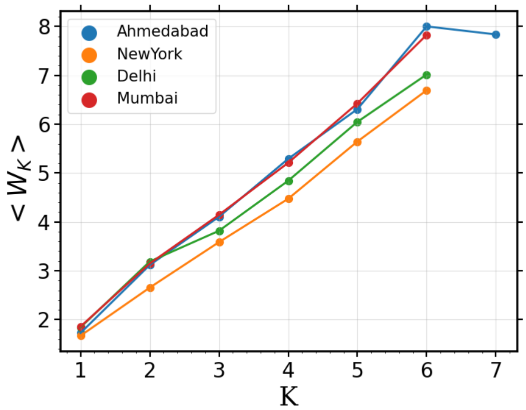


(c) Delhi

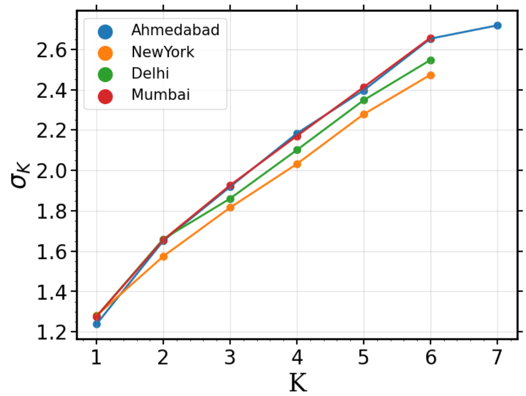


(d) Mumbai

Figure 5.3: Variation of standard deviation σ of walkers with average number of walkers $\langle W \rangle$, for all nodes of the network



(a) Average flux $\langle W \rangle$ as a function of degree K



(b) Flux fluctuation σ as a function of degree K

Figure 5.4: Statistics of flux of walkers as a function of degree K

5.4 Probability of occurrence of EEs

As discussed in chapter 3, an event as extreme if flux passing through a particular node at any time exceeds a threshold value, q. We modelled threshold q for node i as

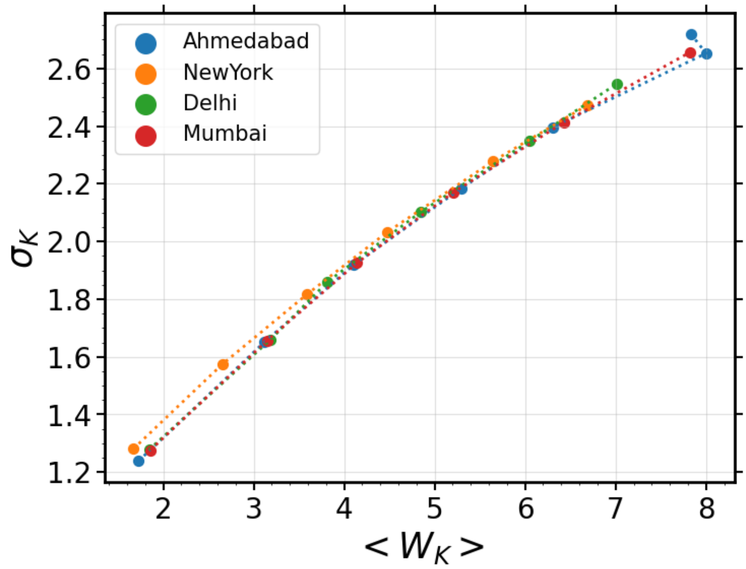


Figure 5.5: Variation of flux fluctuations σ_K with average flux $< W_K >$

$$q_i = \langle W_i \rangle + m\sigma_i \tag{5.2}$$

where $\langle W_i \rangle$ & σ_i are average flux and fluctuations through node i and m is a real number. Here, we chose the scaling parameter $m = 4$, fig 5.6 shows the probability of occurrence of EE as a function of degree K.

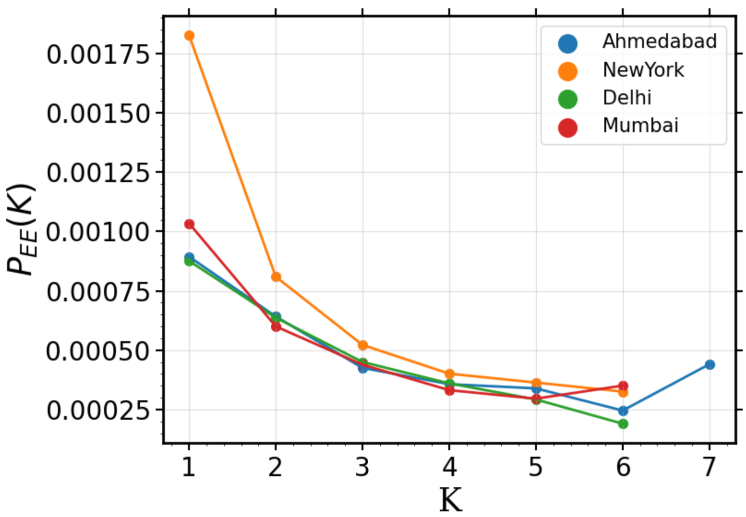


Figure 5.6: Variation of probability of occurrence of EE, P_{EE} with degree K

Chapter 6

Conclusion

In chapter 1, we introduced types of networks and their basic characteristics, and described extreme events and congestion phenomena.

In Chapter 2, we focused on examining the statistical features of the spatial network of urban streets. Our analysis included studying the degree distribution of the networks, and we observed that Mumbai, Delhi, and Ahmedabad exhibit quite similar distributions, with a large number of nodes having a degree of 3. In contrast, New York, being a planned city, has the highest number of nodes with degrees 3 and 4. We observed that the distribution of road length in most cities shows power law behaviour. Among the spatial networks we analyzed, Mumbai appears to have the largest diameter in both the weighted and unweighted versions. We computed the different centrality measures for weighted and unweighted networks, namely betweenness centrality and closeness centrality. From the degree correlation function, we can conclude that all the networks under consideration were assortative.

In Chapter 3, inspired by the real-world examples, we adopted degree biased routing model on spatial networks and found that on average, extreme events are more likely to occur at nodes with lower generalized strength ϕ values compared to those with higher generalized strength ϕ.

In Chapter 4, we studied the congestion phenomenon on unweighted spatial networks, in spite of heterogeneity in the degree distribution of city planar networks, the congestion diagram looks similar to that of a grid network with just a small shift in critical birth probability value(p_c) at which the network becomes congested. The congestion appears earlier in spatial networks of streets due to heterogeneity in its degree distribution. If the outflux r_i and node capacity n_i^* are made dependent on the degree, the scale-free network does not exhibit congestion even when the birth probability is $p = 1$. However, congestion persists in planar networks.

In chapter 5, the probability of occurrence of extreme events was studied using interacting random walk model on unweighted spatial networks where nodes form queues. Here, too we found that lower-degree nodes are more susceptible to extreme events.

During my MS project I came up with the idea of real traffic model see appendix[1] B. It is just the extension of the model introduced in chapter 3 of this thesis. Traffic dynamics actually happens on roads or links rather than nodes or junctions. It would be interesting to study the phase transition.

[1]Due to time constraint analysis wasn't completed, python code see [1]

Bibliography

[1] Ajay Agarwal. Python code for real traffic simulation. https://github.com/AJ-21cm/Real-Traffic-Simulation. Accessed: April 12, 2023.

[2] Sergio Albeverio, Volker Jentsch, and Holger Kantz. *Extreme events in nature and society*. Springer Science & Business Media, 2006.

[3] Julio Amézquita-López, Jorge Valdés-Atencio, and David Angulo-García. Understanding traffic congestion via network analysis, agent modeling, and the trajectory of urban expansion: a coastal city case. *Infrastructures*, 6(6):85, 2021.

[4] Albert-László Barabási. Network science book. *Network Science*, 625, 2014.

[5] Geoff Boeing. Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65:126–139, 2017.

[6] Daniele De Martino, Luca Dall'Asta, Ginestra Bianconi, and Matteo Marsili. Congestion phenomena on complex networks. *Physical Review E*, 79(1):015101, 2009.

[7] Daniele De Martino, Luca Dall'Asta, Ginestra Bianconi, and Matteo Marsili. A minimal model for congestion phenomena on complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(08):P08023, 2009.

[8] M Argollo De Menezes and A-L Barabási. Fluctuations in network dynamics. *Physical review letters*, 92(2):028701, 2004.

[9] Pablo Echenique, Jesús Gómez-Gardeñes, and Yamir Moreno. Improved routing strategies for internet traffic delivery. *Physical Review E*, 70(5):056105, 2004.

[10] Pablo Echenique, Jesús Gómez-Gardenes, and Yamir Moreno. Dynamics of jamming transitions in complex networks. *Europhysics letters*, 71(2):325, 2005.

[11] Agata Fronczak and Piotr Fronczak. Biased random walks in complex networks: The role of local navigation rules. *Physical Review E*, 80(1):016107, 2009.

[12] Vimal Kishore, MS Santhanam, and RE Amritkar. Extreme events on complex networks. *Physical review letters*, 106(18):188701, 2011.

[13] Vimal Kishore, MS Santhanam, and RE Amritkar. Extreme events and event size fluctuations in biased random walks on networks. *Physical Review E*, 85(5):056120, 2012.

[14] Wen-Xu Wang, Bing-Hong Wang, Chuan-Yang Yin, Yan-Bo Xie, and Tao Zhou. Traffic dynamics based on local routing protocol on a scale-free network. *Physical review E*, 73(2):026111, 2006.

[15] Wikipedia. Longest traffic jam. https://en.wikipedia.org/wiki/China_National_Highway_110_traffic_jam#cite_note-1. Accessed: April 12, 2023.

Appendix A

Traffic Speed Models

A.1 Introduction

The important parameters of traffic models are vehicular speed and vehicle density. Many people proposed models to give an exact relationship between speed and density. The following are the models

- Greenshield's macroscopic stream model
- Greenberg's logarithmic model
- Underwood's exponential model
- Pipes' generalized model

we will see the details about the models in the coming sections.

A.2 Greenshield's macroscopic stream model

Macroscopic stream models represent how the behaviour of one parameter of traffic flow changes concerning another. The most important among them is the relation between speed and density. Greenshield proposed the first and most simple relation between them. As illustrated by the equation below, he assumed a linear speed-density relationship to derive the model.

$$v = v_f - \left[\frac{v_f}{k_{jam}}\right]k, \tag{A.1}$$

where v_f is the free flow speed, k is vehicular density and k_{jam} is the vehicular density at jam condition. When density becomes zero, speed approaches free flow speed (i.e. v $\rightarrow v_f$ as k $\rightarrow$ 0).

A.2.1 Flow

- Flow of vehicles is given by

$$q = kv \tag{A.2}$$

- Relation between flow and density is given by

$$q = kv_f - \left[\frac{v_f}{k_{jam}}\right]k^2 \tag{A.3}$$

- Relation between flow and speed is given by

$$q = k_{jam}v - \left[\frac{k_{jam}}{v_f}\right]v^2 \tag{A.4}$$

A.2.2 Density corresponding to maximum flow

We can calculate the density at which flow is maximum. In equation A.3 put $\frac{dq}{dk} = 0$, we get $k_0 = \frac{k_{jam}}{2}$. So, maximum flow is $q_{max} = \frac{v_f k_{jam}}{4}$
velocity v_0 at maximum flow can be obtain by putting $k = k_0$, $v_0 = \frac{v_f}{2}$

A.2.3 Applying Greenshield model

For applying this model in my traffic simulation, k is the number of vehicles per edge. Every edge has some finite capacity for holding the vehicles. So speed taken by a vehicle will depend on the number of vehicles currently held by the edge.

$$v_{next} = \begin{cases} v_f\left[1 - \left(\frac{k_{current}}{k_{jam}}\right)\right], & k \leq k_{jam} \\ 0, & k_{current} > k_{jam} \end{cases} \tag{A.5}$$

Variation of v with k

If free speed is taken to be, v_f=80 km/h = 22.8 m/s

$$v = 22.8\left[1 - \left(\frac{k}{k_{jam}}\right)\right], 0 \leq k \leq k_{jam} \tag{A.6}$$

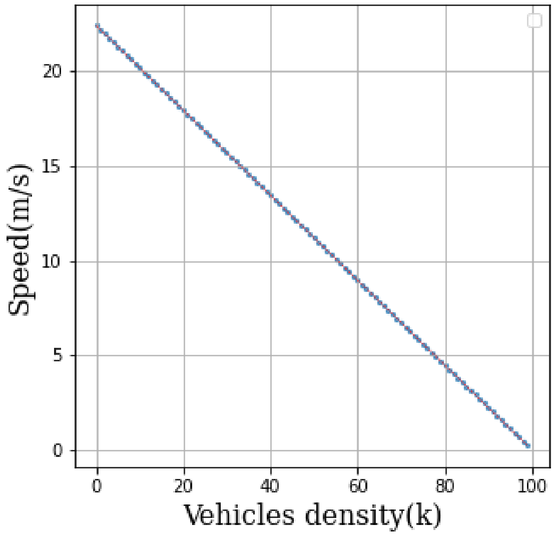
A.1

Distribution of edge travel time

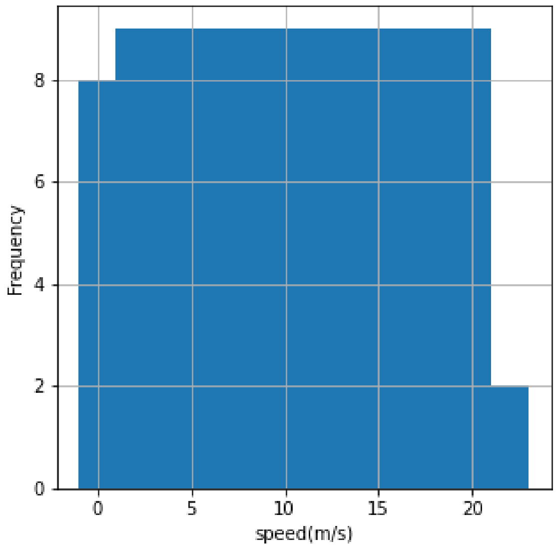
time of travel is given by

$$t = \frac{D}{v}, \tag{A.7}$$

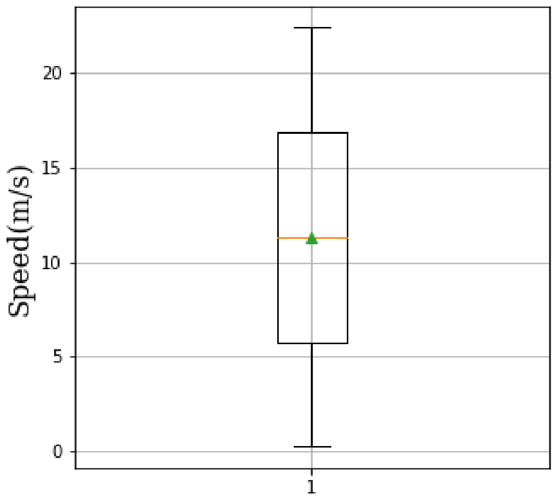
$$t = \frac{D}{22.8\left[1 - \left(\frac{k}{k_{jam}}\right)\right]}. \tag{A.8}$$



(a) Speed profile

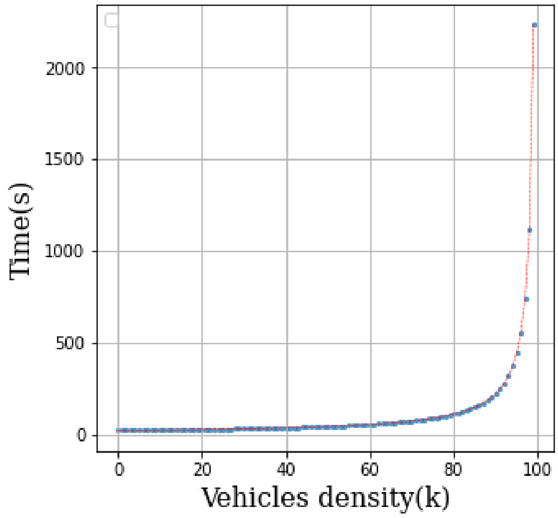


(b) Histogram

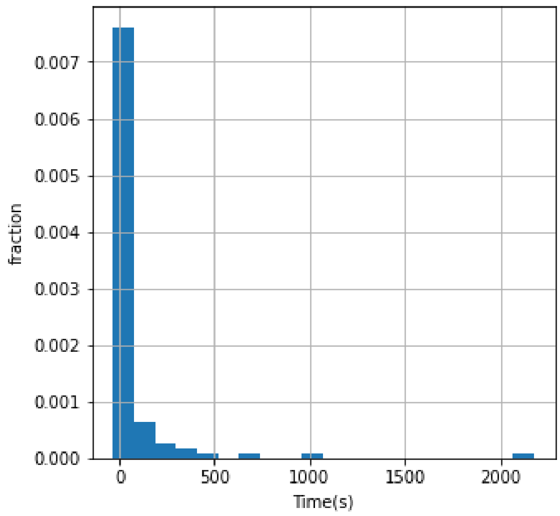


(c) Boxplot

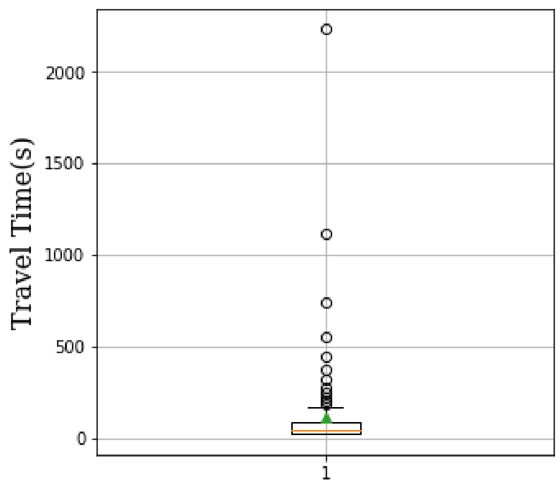
Figure A.1: Speed profile and distribution, Greenshield model



(a) Travel time profile



(b) Histogram



(c) Boxplot

Figure A.2: Travel time profile and distribution, D = 500m, Greenshield model

A.3 Greenberg's logarithmic Model

Greenberg assumed a logarithmic relation between speed and density.

$$v = v_0 \ln\left[\frac{k_{jam}}{k}\right] \tag{A.9}$$

The model has gained excellent popularity because this model can be derived analytically. The main drawback of this model is that as density tends to zero, the speed tends to infinity. This shows the inability of the model to predict the speeds at lower densities.

A.3.1 Applying Greenberg's logarithmic Model

For practical application, there is a need for modification to the Greenberg model since speed can not be infinite at high density.

$$v_{next} = \begin{cases} 1.01\ v_0 \ln(k_{jam}), & k_{current} = 0 \\ v_0 \ln\left[\frac{k_{jam}}{k_{current}}\right], & k_{current} \leq k_{jam} \\ 0, & k_{current} > k_{jam}. \end{cases} \tag{A.10}$$

Variation of v with k

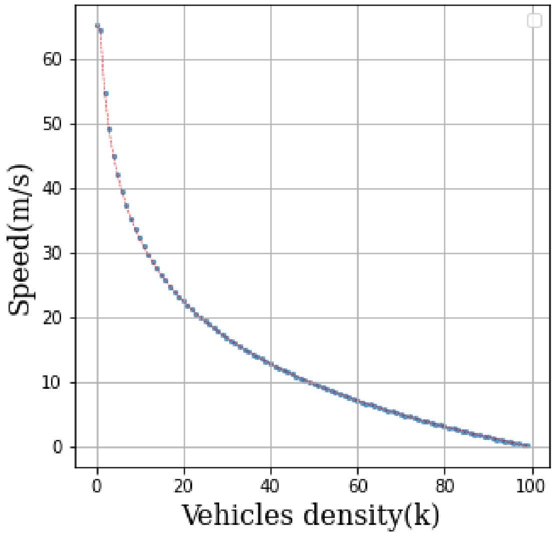
Taking $v_0 = 40km/h$, speed is given by

$$v = 11.4\ \ln\left[\frac{k_{jam}}{k}\right] m/s,\ 0 < k < k_{jam}. \tag{A.11}$$

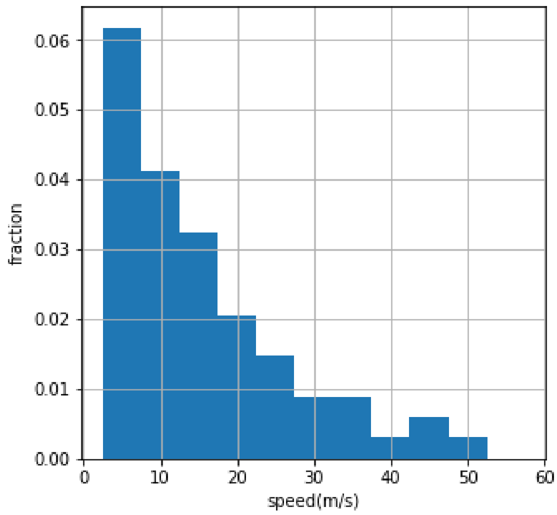
Distribution of time intervals

Time of travel is given by using the equation

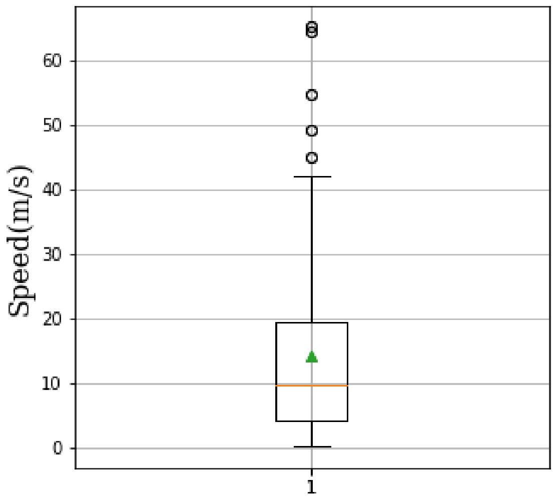
$$t = \frac{D}{11.4\ \ln\left[\frac{k_{jam}}{k}\right]} \tag{A.12}$$



(a) Speed profile

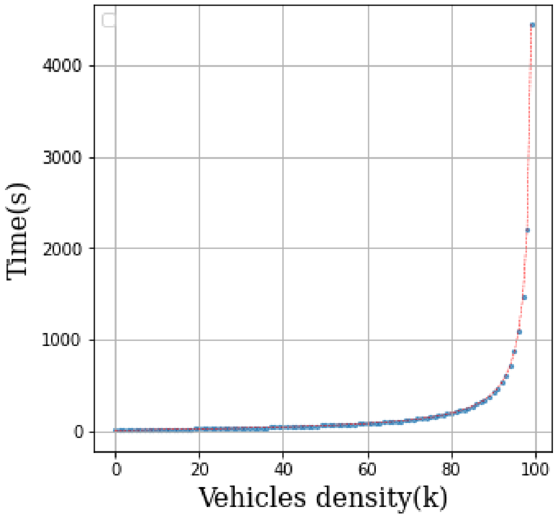


(b) Histogram

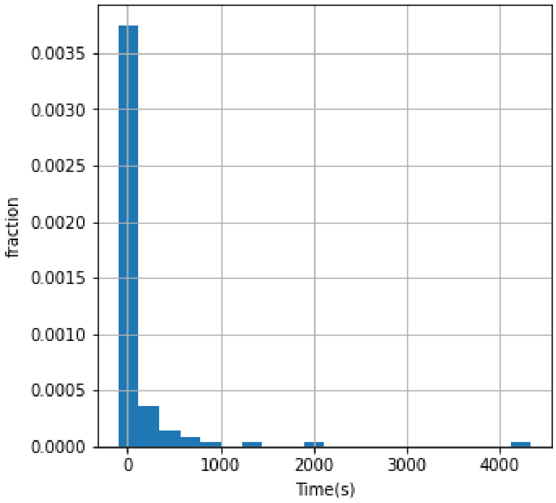


(c) Boxplot

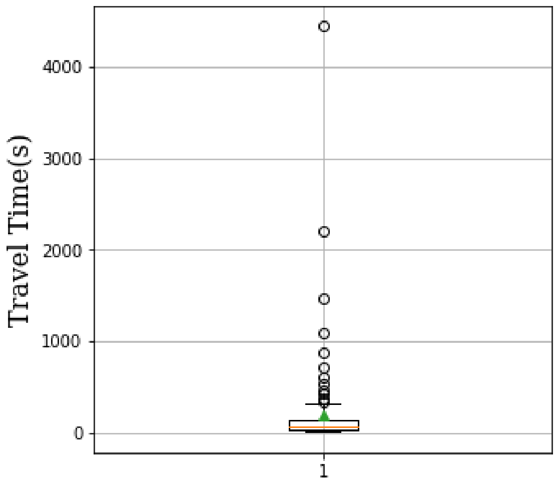
Figure A.3: Speed profile and its distribution, Greenberg's model



(a) Travel time profile



(b) Histogram



(c) Boxplot

Figure A.4: Travel time profile and distribution, D = 500m, Greenberg's model

A.4 Underwood's exponential model

Underwood proposed an exponential model to overcome the limitation of Greenberg's model, as shown below.

$$v = v_f e^{-\frac{k}{k_{jam}}}, \tag{A.13}$$

where k_{jam} is the optimal density, i.e. the density corresponding to the maximum flow. v_f is free to flow speed.
In this model, speed becomes zero only when the density reaches infinity which is the drawback of this model. Hence this cannot be used for predicting speeds at high densities.

A.4.1 Applying Underwood's exponential model

For practical application, I modified the exponential model.

$$v_{next} = \begin{cases} v_f e^{-\frac{k_{current}}{k_{jam}}} & , \quad k_{current} \leq k_{jam} \\ 0 & , \quad k_{current} > k_{jam} \end{cases} \tag{A.14}$$

Variation of v with k

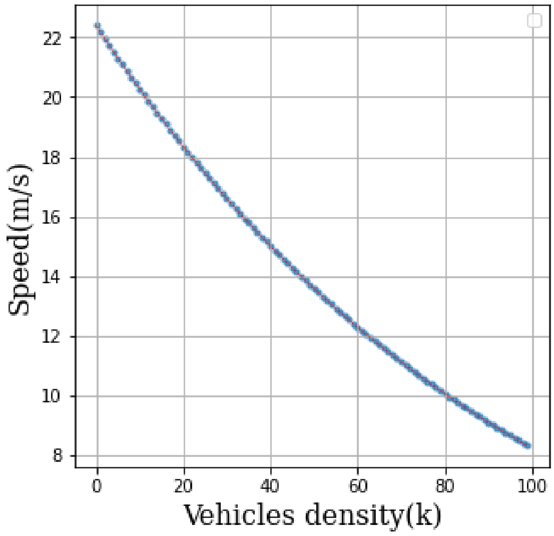
If we take free speed, $v_f = 22.8 m/s$ so, our equation of speed is given by

$$v = 22.8 e^{-\frac{k}{k_{jam}}} m/s \tag{A.15}$$

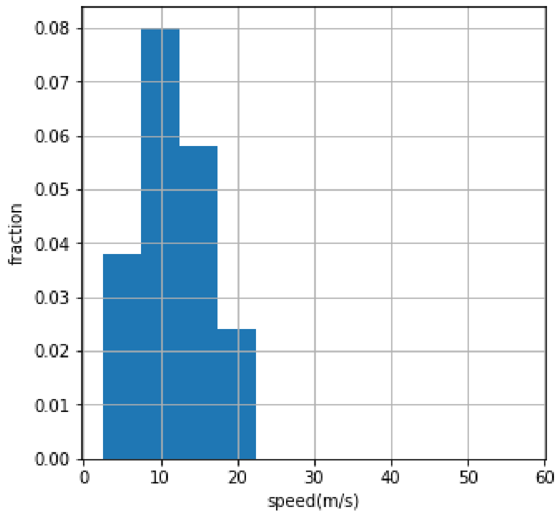
Distribution of time intervals

Time of travel is given by

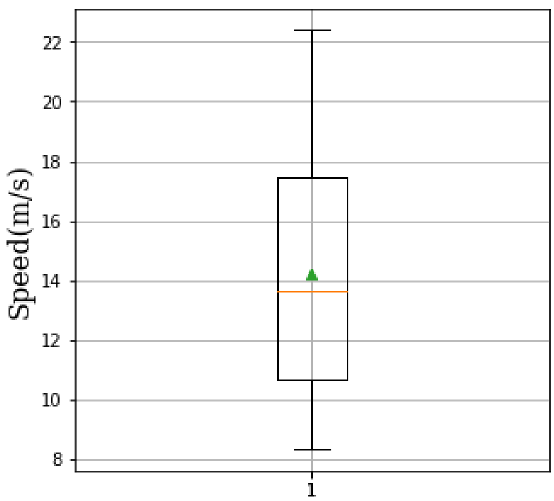
$$t = \frac{D}{22.8 e^{-\frac{k}{k_{jam}}}} \tag{A.16}$$



(a) Speed profile

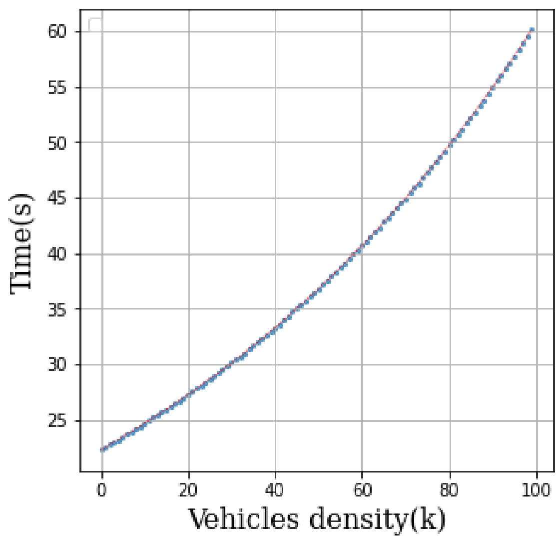


(b) Histogram

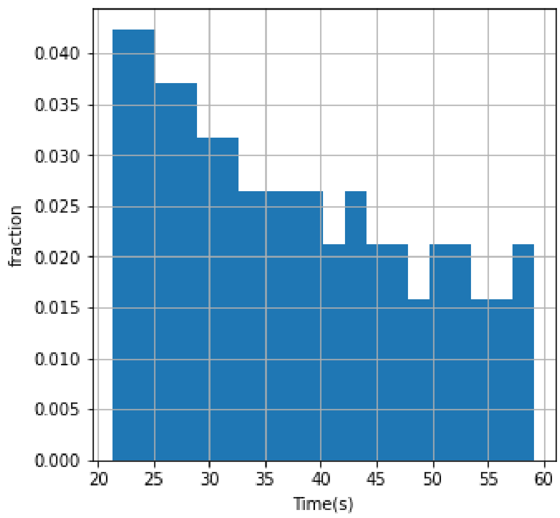


(c) Histogram

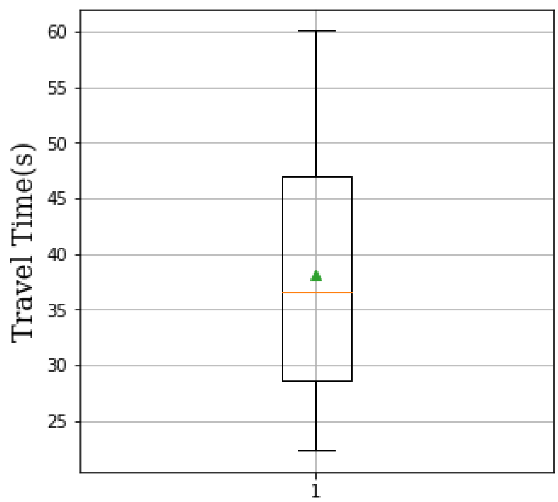
Figure A.5: Speed profile and distribution, Underwood's model



(a) Travel time profile



(b) Histogram



(c) Boxplot

Figure A.6: Travel time profile and distribution, D = 500m, Underwood's model

A.5 Pipes' Generalized model

Pipes proposed a model given by the following equation.

$$v = v_f \Big[1 - \Big(\frac{k}{k_{jam}} \Big)^n \Big] \tag{A.17}$$

where $n \in \mathbb{Z}^+$.

It is a more generalized modelling approach with introducing a new parameter n. When n is set to one, Pipe's model resembles Greenshield's model. Thus, by varying the values of n, a family of models can be developed.

A.5.1 Applying Pipes' Generalized model

Let n=2, so speed is given by

$$v_{next} = \begin{cases} v_f \Big[1 - \Big(\frac{k_{current}}{k_{jam}} \Big)^2 \Big], & k \leq k_{jam} \\ 0, & k_{current} > k_{jam} \end{cases} \tag{A.18}$$

Variation of v with k

If we take free speed, $v_f = 80 km/h = 22.8 m/s$,

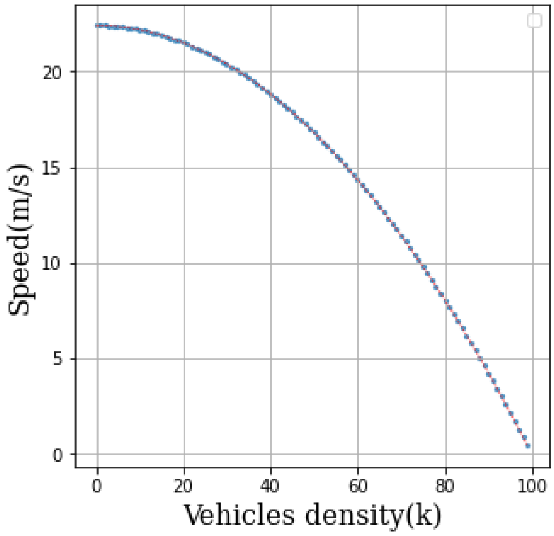
$$v = 22.8 \Big[1 - \Big(\frac{k}{k_{jam}} \Big)^2 \Big], \; 0 \leq k \leq k_{jam} \tag{A.19}$$

Distribution of time intervals

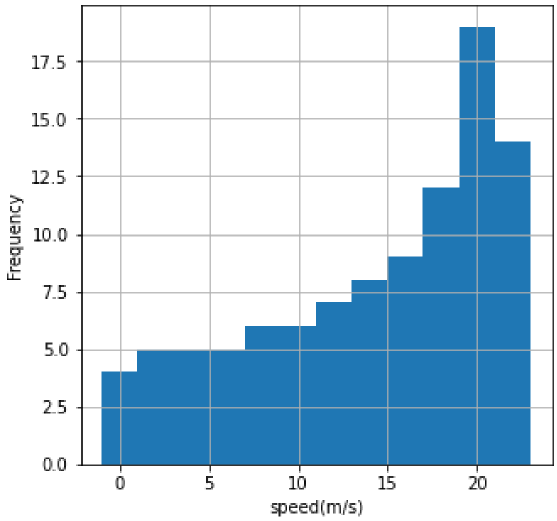
Time of travel is given by,

$$t = \frac{D}{v}, \tag{A.20}$$

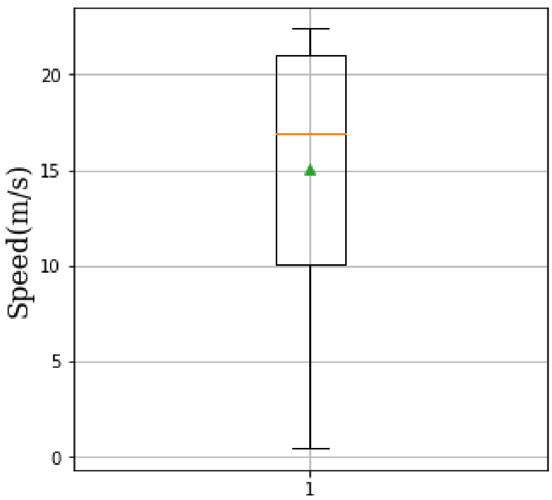
$$t = \frac{D}{22.8 \Big[1 - \Big(\frac{k}{k_{jam}} \Big)^2 \Big]}. \tag{A.21}$$



(a) Velocity profile

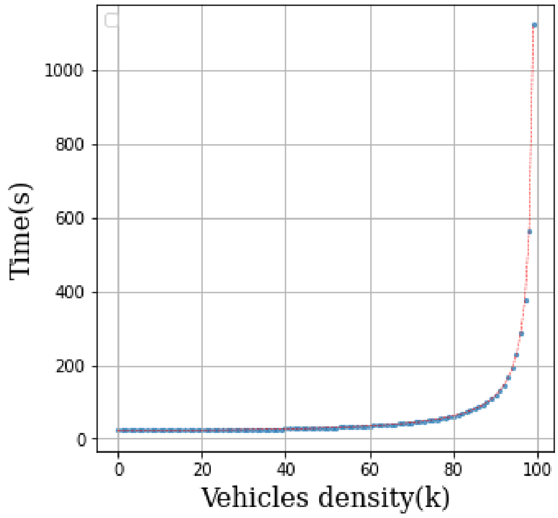


(b) Histogram

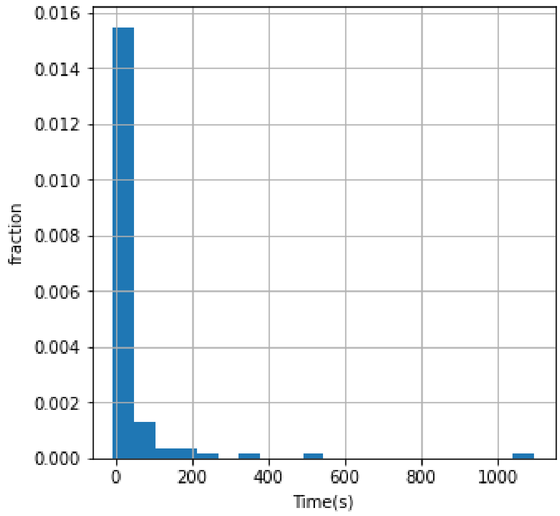


(c) Boxplot

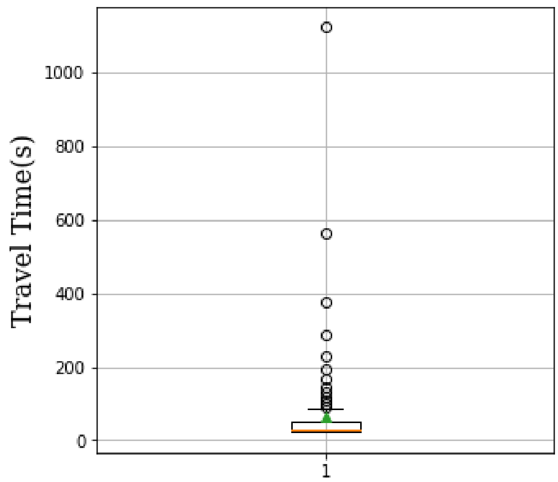
Figure A.7: Speed profile and distribution, Pipes' model



(a) Travel time profile



(b) Histogram



(c) Boxplot

Figure A.8: Travel time profile and distribution, $D = 500m$, Pipes' model

Appendix B

The Real Traffic Model

B.1 A link model to mimic real traffic

- Consider a real planar city-directed network consisting of N nodes and $2E$ edges, where E is the number of edges in the undirected version of the considered network. Each edge E_{ij} has a finite length $L_{ij} \in \mathbb{R}^+$ (actual road length of a city).

- Each node and edge is endowed with a FIFO finite queue.

- Vehicles are created only on nodes with probability p_i per time step (time resolution) and can only be destroyed on node with probability μ_i. No vehicle is destroyed on edges.

- Vehicle on node i enters the queue of edge E_{ij}, with probability $P_{i\to j}$ given by,

$$P_{i\to j} = \frac{1}{K_i}, \tag{B.1}$$

  where K_i is degree of node i.

- Speed acquired by a vehicle on node i going to node j is given by,

$$V_{i\to j} = \begin{cases} V_f \exp\left[-\frac{n_{(i,j)}}{n^*_{(i,j)}}\right], & n_{(i,j)} \leq n^*_{(i,j)} \\ 0, & n^*_{(i,j)} < n_{(i,j)} \end{cases} \tag{B.2}$$

  where V_f is free flow speed, $n_{(i,j)}$ is number of vehicles on edge E_{ij}, whose capacity is $n^*_{(i,j)}$ defined as,

$$n^*_{(i,j)} = b_0 \lfloor L_{ij} \rfloor, \tag{B.3}$$

where $b_0 \in \mathbb{Z}^+$ is some constant, L_{ij} is edge length and $\lfloor\rfloor$ represents the mathematical floor function.

- Node i can deliver at most r_i vehicle to connecting edges per time step (time resolution).