Design and Development of a Probabilistic Framework for Traffic Management

MAJOR PROJECT REPORT

SEMESTER- I

FOUR-YEAR UNDERGRADUATE PROGRAMME
(DESIGN YOUR DEGREE)

SUBMITTED TO UNIVERSITY OF JAMMU, JAMMU

SUBMITTED BY THE TEAM: **PASCAL GROUP**

TEAM	ROLL NO
Aditya Gupta	DYD-24-01
Muneeb Ajmer	DYD-24-24
Somil Narang	DYD-24-36
Vaishnavi Pole	DYD-24-38
Vishnu Uphadhya	DYD-24-40

UNDER THE MENTORSHIP OF

Prof. K.S Charak
(Dept of Mathematics)

Dr. Sunil Kumar
(Dept of Statistics)

Dr. Jatinder Manhas
(Dept of IT)

CERTIFICATE

The work embodied in this report entitled "**Design and Development of a Probabilistic Framework for Traffic Management"** has been done by Group Pascal including group members- Aditya Gupta, Muneeb Ajmer, Somil Narang, Vaishnavi Pole, Vishnu Uphadhya as a Major Project for Semester I of Four-Year Undergraduate Programme (Design Your Degree). This work was carried out under the guidance of Mentors Prof. K.S Charak, Dr. Sunil Bhougal and Dr. Jatinder Manhas partial fulfilment of the award of the Design Your Degree, Four Year Undergraduate Programme, University of Jammu, Jammu & Kashmir. This project report has not been submitted anywhere else.

Signature of Mentors

Prof. K.S Charak

(Dept of Mathematics)

Dr. Sunil Kumar

(Dept of Statistics)

Dr. Jatinder Manhas

(Dept of IT)

Prof. Alka Sharma

Director

SIIDEC, University of Jammu

ACKNOWLEDGMENT

We avail this opportunity to acknowledge all those who helped and guided us during the course of our work. First and foremost, we thank Almighty God from the depth of our heart for generating enthusiasm and granting our spiritual strength to successfully pass through this challenge. Words are too meager to express our esteem indebtedness and whole hearted sense of gratitude towards our mentors. It is our pleasure beyond words to express our deep sense of feelings for their inspiring guidance, generous encouragement and well-versed advice. They provided us undaunted encouragement and support in spite of their busy schedules. Really, fortunate we are and we feel extremely honored for the opportunity conferred upon us to work under their perpetual motivation.

We are extremely fortunate for having an opportunity to express our heartiest gratitude to Prof. Alka Sharma, Director of SIIEDC (Design your degree), University of Jammu, Jammu and member of advisory committee for her valuable advice and generous help and providing us all the necessary facilities from the department. We are highly thankful to all respected mentors for their immense help during the conduct of this major project. The words are small trophies to express our feelings of affection and indebtedness to our friends who helped us throughout the major project, whose excellent company, affection and co-operation helped us in carrying out our research work with joy and happiness.

We acknowledge all the people, mentioned or not mentioned here, who have silently wished and gave fruitful suggestions and helped us in achieving the present goal.

Abstract

Traffic congestion remains a persistent issue in urban environments, causing delays, economic losses, increased fuel consumption, and environmental pollution. Effective traffic management solutions require the ability to accurately predict vehicle arrivals and flows at critical points, such as intersections and highways, to optimize traffic control measures. This project, titled "Design and Development of a Probabilistic Framework for Traffic Management," addresses this challenge by introducing a predictive model based on the Poisson distribution method.

The Poisson distribution is well-suited for modelling random, discrete events, such as vehicle arrivals, over fixed intervals of time. By analysing historical traffic data, the framework identifies patterns in vehicle arrivals and establishes a probabilistic model that predicts traffic flow under varying conditions. This predictive capability can be integrated with existing traffic management systems to dynamically adjust signal timings, reroute vehicles, and prioritize emergency responses, thereby minimizing congestion and enhancing road efficiency.

The project involves the development of a computational algorithm to process traffic data, apply the Poisson distribution, and generate real-time predictions. The framework's accuracy was tested on simulated traffic scenarios, demonstrating its effectiveness in forecasting vehicle arrival rates and adapting to fluctuations in traffic density. The results highlight the potential of probabilistic methods in addressing complex traffic management problems.

TABLE OF CONTENT

Chapter 1

Introduction

In today's world traffic congestion is one of the most time-consuming problems and there are many different solutions we are introduced with but most of them are based on real-time data and can be a costly method to be applied everywhere in Jammu. Our project**, Design and Development of a Probabilistic Framework for Traffic Management**, proposes a solution using **historical traffic data** and the **Poisson distribution** [2] to simulate traffic flow. This approach enables us to predict vehicle arrivals at various checkpoints without relying on real-time data, offering a cost-effective and scalable model for traffic management.

The Poisson distribution method principle is:

This method is used to calculate the likelihood of a specific number of events occurring in a fixed time interval, space, area etc (See [2]). This method finds the probability of events which are independent of each other and the events should occur at a known constant rate.

We encountered many methods to approach the problem of traffic congestion in Jammu and Poisson distribution is the method we chose as it describes the likelihood of independent events occurring and since arrival of one vehicle does not always influence the arrival of another vehicle in a fixed interval of time (See [5]). We used this formula to find the vehicle arrivals in the intervals of 5 minute in the form of a simulation while finding the appropriate statistical measures like mean, variance, standard deviation and threshold.

Overview of the Poisson distribution:

$$\boldsymbol{P(X = k) = \frac{(\lambda^k e^{-\lambda})}{k!}}$$

Where:

P(X=k) is the probability of k events occurring.

λ is the average vehicle arrivals in one minute in a location.

k is the actual number of events.

e is the exponential constant.

1.1 OBJECTIVES

1.Create a model which offers a **cost-effective and reliable solution** to improve traffic flow

2. Predict the traffic flow and **give insights** about the traffic conditions in the future.

3. To enhance the efficiency of **traffic signal optimization**.

4. To make an easy-to-use environment for future predictions.

5. Using Poisson Distribution method for more accurate predictions.

Key steps involved in the working of Poisson distribution:

- **Initialization**: We have to find the value of **λ** in a location, we can do it by manually counting the vehicles.
- **Find the Probability:** We will find the probability of **k** events occurring (P(X=k)) for the values of **k** ranging from 0 to ∞ since theoretically **k** has an infinite range (See [2]).
- **Random number generation:** A random number is generated ranging from 0 to 1 to check whether it is greater than or smaller than the P(X=k) for every value of k and this process will be followed continuously till the random number generated is smaller than P(X=k) and the corresponding **k** value is the number of vehicles (See[4]) that can arrive at the location for the 1st 5-minute interval.
- **Calculating Statistical measures:** By calculating the statistical measures like mean, variance and standard deviation we can find the threshold value (mean ± (n* Standard Deviation)) where n* standard deviation is any multiple of standard deviation.
- **Traffic Light Timings:** After calculating the threshold value, if the calculated value of **k** for different 5-minute intervals exceeds the threshold value, the traffic light timings change itself according to the high traffic congestion and the green-light duration increases and red-light duration decreases for the traffic to flow seamlessly. If the value of **k** does not exceed the threshold value, then it is considered a normal congestion and the green-light duration is decreased and red-light duration is increased.

This method is a well-suited method to get better insights about the upcoming traffic and controlling it by adjusting the traffic lights timings based on the results of statistical measures.

Benefits of using Poisson distribution:

- This method is a simple and efficient method to implement and it only requires two parameters **λ** and **k.**
- This method is most suitable for this type of problem as the arrival of vehicles are independent events to each other and this method describes the probability of independent events occurring in a fixed interval of time.
- This is a flexible method which can be used in various situations of traffic congestion.
- This method calculates probability for events that are random and since vehicle arrivals are random in a location this becomes another reason to use this method.

Chapter 2

METHODOLOGY

We used Poisson distribution data since it is a well-suited method for these types of problems where the events occur independently. The method is used by calculating the average number of vehicle arrivals at that place (historical data). The statistical measures calculated are mean, variance, standard deviation, and threshold.

DATA COLLECTION

Historical data collection is done by visiting some of the major locations prone to heavy traffic congestion and calculated the average number of vehicle arrival in a minute.

TOOLS AND TECHNOLOGY USED

1. Programming Language: C

- **High Performance**: Ensures fast computation and efficient handling of large datasets.
- **Low-Level Control**: Offers better memory management and system resource control.

2. Statistical Method: Poisson Distribution

- Used to model vehicle arrivals at traffic checkpoints, based on historical data.
- Helps in simulating realistic traffic patterns and predicting traffic flow during different time intervals.

3. Tools: GCC Compiler

- **Cross-Platform Compatibility:** Ensures the code runs across different operating systems.
- **Debugging and Optimization:** Provides powerful debugging tools and optimizations to enhance code performance.

4. Data Sources

- **Historical Traffic Data:**

    1. Collected from traffic studies and local data sources to simulate realistic traffic conditions and patterns.
    2. Collected primary data from some major sites during field visits.

LIMITATIONS OF THE MODEL

- Poisson distribution is meant for independent events but in **reality, vehicles not always move independently.**
- It works for only Single stream of events and **not for multi-lane roads.**
- Poisson allows for an infinite number of events but there is a **physical upper limit of road capacity.**
- This formula predicts randomness but sometimes **traffic might become deterministic.**

Chapter 3

CODING

```
# include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <time.h>

// Function to calculate factorial (used in Poisson probability)
unsigned long long factorial(int n) {
    if (n == 0 || n == 1) return 1;
    return n * factorial(n - 1);
}

}

// Function to calculate mean
double calculate_mean(int data[], int size) {
    double sum = 0.0;
    for (int i = 0; i < size; i++) {
        sum += data[i];
    }
    return sum / size;
}

// Function to calculate variance
        k++;
        p *= (double)rand() / RAND_MAX;
    }
    return k - 1;

double calculate_variance(int data[], int size, double mean) {
    double sum = 0.0;
    for (int i = 0; i < size; i++) {
        sum += pow(data[i] - mean, 2);
    }
    return sum / size;
}

// Function to calculate standard deviation
double calculate_std_dev(double variance) {
    return sqrt(variance);
}

// Function to get traffic light timings as a string
const char* get_traffic_light_timings(int vehicles, double threshold) {
    if (vehicles > threshold) {
        return "High Congestion: Green=90s, Yellow=10s, Red=30s";
    } else {
        return "Normal Traffic: Green=50s, Yellow=10s, Red=60s";
    }
}

int main() {
    double lambda;   // Average arrival rate (vehicles per minute)
    int total_time;  // Total time to simulate (in minutes)
    int time_step = 1;  // Time interval for each simulation step (in minutes)

    // Initialize random number generator
    srand(time(NULL));

    // Input historical data
    printf("Enter average arrival rate (vehicles per minute): ");
    scanf("%lf", &lambda);
    printf("Enter total simulation time (in minutes): ");
    scanf("%d", &total_time);

```

```
    int interval_duration = 5; // 5-minute intervals
    int num_intervals = total_time / interval_duration; // Total number of 5-minute intervals
    int vehicle_data[num_intervals]; // Array to store traffic data for each interval
    for (int i = 0; i < num_intervals; i++) vehicle_data[i] = 0; // Initialize to 0

    // Simulate traffic
    printf("\nSimulating traffic for %d minutes (%d intervals of %d minutes each):\n", total_time, num_intervals, interval_duration);

    for (int t = 0; t < total_time; t += time_step) {
        int vehicles = simulate_poisson(lambda); // Simulate vehicle arrivals
        int interval_index = t / interval_duration; // Determine current 5-minute interval index
        vehicle_data[interval_index] += vehicles;  // Accumulate vehicles for the interval
    }

    // Calculate statistical measures
    double mean = calculate_mean(vehicle_data, num_intervals);
    double variance = calculate_variance(vehicle_data, num_intervals, mean);
    double std_dev = calculate_std_dev(variance);

    // Set congestion threshold

// Function to calculate Poisson probability
double poisson_probability(int k, double lambda) {
    return (pow(lambda, k) * exp(-lambda)) / factorial(k);
}

// Simulate vehicle arrivals using Poisson distribution
int simulate_poisson(double lambda) {
    double L = exp(-lambda);
    double p = 1.0;
    int k = 0;

    while (p > L) {
    double congestion_threshold = mean + std_dev;

    // Identify peak interval
    int peak_interval = 0;
    int max_vehicles = vehicle_data[0];
    for (int i = 1; i < num_intervals; i++) {
        if (vehicle_data[i] > max_vehicles) {
            max_vehicles = vehicle_data[i];
            peak_interval = i;
        }
    }

    // Display 5-minute interval data with traffic light timings
    printf("\nTraffic Data (5-minute Intervals):\n");
    printf("Interval\tVehicles\tTraffic Light Timings\n");
    for (int i = 0; i < num_intervals; i++) {
        printf("%d-%d\t%d\t%s\n",
               i * interval_duration,
               (i + 1) * interval_duration,
               vehicle_data[i],
               get_traffic_light_timings(vehicle_data[i], congestion_threshold));
    }

    // Print statistical measures
    printf("\nStatistical Measures:\n");
    printf("Mean: %.2f vehicles/interval\n", mean);
    printf("Variance: %.2f\n", variance);
    printf("Standard Deviation: %.2f vehicles/interval\n", std_dev);
    printf("Congestion Threshold: %.2f vehicles/interval\n", congestion_threshold);

    // Print peak interval
    printf("\nPeak Interval Analysis:\n");
    printf("Peak Interval: %d-%d minutes with %d vehicles\n",
```

```
            peak_interval * interval_duration,
            (peak_interval + 1) * interval_duration,
            max_vehicles);

    return 0;
}

```

Chapter 4

RESULTS:

```
Enter average arrival rate (vehicles per minute, must be > 0): 25
Enter total simulation time (in minutes, must be > 0): 60

Simulating traffic for 60 minutes (12 intervals of 5 minutes each)

Traffic Data (5-minute Intervals):
Interval        Vehicles        Traffic Light Timings
0-5     128     Normal Traffic: Green=50s, Yellow=10s, Red=60s
5-10    136     High Congestion: Green=90s, Yellow=10s, Red=30s
10-15   112     Normal Traffic: Green=50s, Yellow=10s, Red=60s
15-20   144     High Congestion: Green=90s, Yellow=10s, Red=30s
20-25   113     Normal Traffic: Green=50s, Yellow=10s, Red=60s
25-30   105     Normal Traffic: Green=50s, Yellow=10s, Red=60s
30-35   120     Normal Traffic: Green=50s, Yellow=10s, Red=60s
35-40   128     Normal Traffic: Green=50s, Yellow=10s, Red=60s
40-45   120     Normal Traffic: Green=50s, Yellow=10s, Red=60s
45-50   138     High Congestion: Green=90s, Yellow=10s, Red=30s
50-55   123     Normal Traffic: Green=50s, Yellow=10s, Red=60s
55-60   111     Normal Traffic: Green=50s, Yellow=10s, Red=60s

Statistical Measures:
Mean: 123.17 vehicles/interval
Variance: 132.64
Standard Deviation: 11.52 vehicles/interval
Congestion Threshold: 134.68 vehicles/interval

Peak Interval Analysis:
Peak Interval: 15-20 minutes with 144 vehicles
```

4.1 WORKING OF THE CODE:

Step 1: Initialization:

- The user inputs the value of **λ** (average arrival rate of vehicles).
- The user inputs the desirable duration of simulation in minutes.
- The compiler ensures that the values inputted are meeting the requirements and it proceeds.

Step 2: Poisson Distribution

- Firstly, the factorial of k is calculated.
- The Poisson distribution formula is then used with the by using the inputted values in the formula $\boldsymbol{P(X=k)=\frac{(\lambda^{k}e^{-\lambda})}{k!}}$.
- After finding the probability of k events occurring then a random number is generated between 0 and 1 by using rand () /RAND_MAX function and time(null) seed is used to make sure that the random numbers generate do not follow any pattern and then it will be compared with the probabilities till we find the random number to be smaller than the probability and we will declare the corresponding k value as the no of vehicle arrival in the 1st 5- minute interval and this process is repeated till the simulation ends.

Step 3: Calculating Statistical measures

- Statistical measures like mean, variance, and standard deviation are found and from mean and standard deviation the threshold value is calculated
- The threshold value is then used to determine the traffic light timings, if the vehicle arrivals is more than the threshold value the traffic light timing changes making green-light duration longer and red-light duration shorter and vice versa.

Step 4: Finding the peak interval

- The peak interval is founded by storing the vehicle count of first interval in the max_vehicle and if any other interval have a value greater than the first interval than the max vehicle number is changed into the new number and after checking the whole list the final value max_vehicle is printed as the peak interval.

Chapter 4

Pilot Study

A pilot study is a preliminary investigation to test and evaluate the feasibility, methodology, and design of the main traffic management project. It serves as a trial run to identify potential issues, refine the approach, and ensure the effectiveness of data collection, analysis, and implementation strategies.

Objective of the Pilot Study

To conduct a small-scale evaluation of traffic conditions and public transportation at Bikram Chowk and Panama Chowk, Jammu, and test the viability of a probabilistic traffic simulation model using historical data. The goal is to refine methodologies for the larger study.

Scope of the Pilot Study

Geographic Focus: Bikram Chowk and Panama Chowk Jammu.

Key Stakeholders: Commuters, public transport operators, and traffic authorities.

Study Duration: 2 weeks.

Core Areas of Focus:

- Data collection techniques.
- Simulation model validation.
- Traffic flow analysis.

Pilot Study Design and Methodology:

1. Preliminary Planning

Define Key Parameters:

- Traffic volume, peak and off-peak hours, vehicle types.
- Public transport frequency and capacity.

Set Objectives:

- Assess current congestion levels.
- Validate simulation accuracy with real-world data.

2. Area Selection and Preparation

Location: Bikram Chowk and Panama Chowk, a high-traffic areas.

Site Analysis:

- Identify key bottlenecks and intersections.
- Identifying problems related to Traffic signals.

3. Data Collection

- **Primary Data:**
- Manual traffic counts (vehicle types, direction, and frequency).

Secondary Data:

- Historical traffic data from Jammu Traffic Police (e.g., average daily traffic).

4. Development of Simulation Model

- **Programming Approach:**
  - Use C programming for a simple probabilistic traffic simulation model.
  - Input Variables:
    - Number of vehicle arrivals per minute.
  - Output:
    - Predicted congestion levels and provide adequate traffic light timings.

5.Validation and Testing

- Conduct real-time traffic observation and compare results with simulation output.
- Identify discrepancies and refine the algorithm.

Key Metrics for Pilot Study Success

- **Accuracy:** The degree to which the simulation model predicts real traffic patterns.
- **Feasibility:** Ease of data collection and adaptability of the model.
- **Relevance:** The insights gained from the pilot study that apply to the full-scale project.

Deliverables of the Pilot study

- **Traffic Flow Report:**
  - Current congestion patterns, peak hours, and problem areas.
- **Simulation Output:**
  - Validation of the probabilistic traffic simulation model.
- **Feasibility Report:**
  - Assessment of the methodologies and resources required for the full-scale study.
- **Recommendations:**
  - Modifications to improve data collection and algorithms

Challenges Faced

Data Collection Difficulties:

- Inconsistent manual counts during peak hours.
- Gaps in historical data availability.

Model Accuracy:

- Accounting for unexpected factors like weather or accidents.

Outcomes of the Pilot Study

- Refined methodology for data collection.
- Improved traffic simulation model ready for full-scale implementation.
- Actionable insights into traffic challenges at Bikram Chowk.

- **SUCCESS RATE:**
- Our Model gave predictions with 76% Success rate (Tested on Field visit).
- Calculation of success rate is done by using the formula (predicted vehicle counts/actual vehicle counts) *100. This is calculated on a field visit and the same outcome has been observed after many tries.

Chapter 5

CONCLUSION

This project of ours "**Design and Development of a Probabilistic Framework for Traffic Management"** provides a simple and effective implementation of Poisson distribution in high traffic congestion prone areas and offering a cost-effective and a scalable method to avoid traffic congestion.

With a Success rate of 76%, This model can be considered a successful model taking in account that there is no highly advanced language and concepts of machine learning used in this.

Future Scope

- Incorporate **variable λ values** for time-dependent simulations (e.g., peak vs. off-peak hours).
- Extend the model to simulate **multi-lane roads** and **networked intersections**.
- Add **visualization tools** for better data presentation and analysis.
- Expand to include **pedestrian movement** and **public transport arrivals** for broader application.

Chapter 7

REFRENCES

[1] J. Adachi and A. Takasu. Traffic Incident Detection Using Probabilistic Topic Model. *CEUR Workshop Proceedings*, March 2014.

[2] Daniel L. Gerlough and Andre' Schuhl. Use of Poisson Distribution in Highway Traffic. The Probability Theory Applied to Distribution of Vehicles on Two-Lane Highways. *Eno Foundation for Highway Traffic Control*, 1955.

[3] R. Sathiyaraj, A. Bharathi, Sikandar Khan, Tayybah Kiren, Inam Ullah Khan, Muhammad Fayaz. A Genetic Predictive Model Approach for Smart Traffic Prediction and Congestion. *Hindawi*, 2022.

[4] Tom V. Mathew. Vehicle arrival model. *IIT Bombay*, 2023

[5] Caleb Okeyo Oyala and Edgar Ouko Otumba. Modelling of Distribution of the "Matatu" Traffic Flow Using Poisson Distribution in a Highway in Kenya. *Hikari Ltd,* 2018.

[6] Ninad Lanke and Sheetal Koul. Smart Traffic Management System. *Foundation of Computer Science (FCS),* 2013.