OMAR RASHEED
01-134161-051

Mobility Analytics Application For Intelligent Transportation System Using Blockchain,IOT and Machine Learning

Bachelor of Science in Computer Science

Supervisor: Dr. MUHAMMAD MUZAMMAL

Department of Computer Science
Bahria University, Islamabad

December 2019

Certificate

We accept the work contained in the report titled "Mobility Analytics Application For Intelligent Transportation System Using Blockchain,IOT and Machine Learning", written by Mr. Omar Rasheed as a confirmation to the required standard for the partial fulfillment of the degree of Bachelor of Science in Computer Science.

Approved by . . . :

Supervisor: Dr Muhammad Muzammal

_____

Internal Examiner: Sir Ali Irfan and Dr Moazam Ali

_____

External Examiner:

_____

Project Coordinator: Miss Zubaria Inayat

_____

Head of the Department: Dr Muhammad Muzammal

_____

November 22^{nd}, 2019

Abstract

Intelligent Transportation System is one of the component of Smart City development and it has played an important role in providing sustainable environment to the citizen.Traffic is a phenomenon that we all have to deal with,due to urbanization,the trend of traffic congestion has increased a lot,which causes multiple problems to all the people travelling,In this project I have analysed the Geospatial data to understand traffic congestion trends and to design a solution to traffic congestion problem.The project focus is mainly on finding the best solution for the issues that are faced on the road,that is high traffic congestion mainly because of Urbanization.There are two modules that I have developed in the project,one is Hazard Warning System that is based on IOTA TANGLE technology and the other one is Traffic congestion analysis and prediction,using Machine Learning.

Acknowledgments

First of all,I really want to thanks Allah Almighty,for giving me the strength,motivation and encouragement to develop the final year project in time. I really want to thanks my parents,whose love,support and continuous encouragement during all ups and downs really inspired me to work on the final year project with determination and love.

I submit my heartiest gratitude to my Supervisor Dr. Muhammad Muzammal for his sincere guidance and help for completing this project and I also want to specially thanks Dr. Sumaira Kausar who really guided me a lot during the project. I would also like to thank all concerned persons who helped me in this regard,who helped me to inject quality in the final year project **"Mobility Analytics Application for Intelligent Transportation System Using Blockchain,IOT and Machine Learning "**.

OMAR RASHEED
Islamabad, Pakistan

November 2019

"A father gives his child,
Nothing better
then a Good Education"

Prophet Hazrat Muhammad (S.A.W)

Contents

Abstract i

1 Introduction 1
- 1.1 Overview . . . 1
- 1.2 Problem Statement . . . 3
- 1.3 Problem Description . . . 3
- 1.4 Objective . . . 4
- 1.5 Project Scope . . . 4

2 Literature Review 7
- 2.1 GPS Trajectories using Visualisation and Clustering: . . . 7
- 2.2 Method for assessment of Traffic Congestion of Urban road networks based on Speed . . . 8

3 Requirement Specifications 11
- 3.1 Purpose . . . 11
- 3.2 Product Scope . . . 11
  - 3.2.1 Hazard Warning System . . . 11
  - 3.2.2 Traffic Congestion Analysis . . . 12
- 3.3 Overall Description . . . 12
  - 3.3.1 Product Perspective . . . 12
  - 3.3.2 Product Functions . . . 12
  - 3.3.3 User Interface . . . 13
  - 3.3.4 Hardware Interface . . . 13
- 3.4 Use case . . . 13
  - 3.4.1 User Login Use case . . . 14
  - 3.4.2 User Logout Use case . . . 14
  - 3.4.3 User Registration Use case . . . 15
  - 3.4.4 Input Message . . . 15
  - 3.4.5 Spatial Data Analysis . . . 16
  - 3.4.6 Perform System Testing . . . 16
  - 3.4.7 Machine2Machine Transaction Use case . . . 16
  - 3.4.8 Input Message . . . 18
  - 3.4.9 Data Loading Usecase for Spatial Analysis . . . 18
  - 3.4.10 Data Visualization Use case . . . 19

4 Design **21**
4.1 System Architecture . . . 21
4.1.1 Traffic Congestion analysis and prediction module Methodology,Analysis and Results . . . 21
4.1.2 Hazard Warning System Module . . . 29
4.2 Design Constraints . . . 30
4.3 Design Methodology . . . 31
4.3.1 Process Model . . . 31
4.4 High Level Design . . . 31
4.4.1 Hazard Warning System Process . . . 31
4.4.2 Registration Process . . . 32
4.4.3 System Login Process . . . 32
4.4.4 System Logout Process . . . 33
4.4.5 IOTA TANGLE Process . . . 33
4.5 Database Design . . . 34
4.6 GUI Design . . . 34
4.6.1 Registration Page . . . 34
4.6.2 Login Page . . . 35
4.6.3 Home Page . . . 36
4.6.4 Messages Page . . . 36

5 System Implementation **39**
5.1 System Architecture . . . 39
5.1.1 Development Environment and Tools . . . 40
5.1.2 Implementation Techniques . . . 41
5.2 Conclusion . . . 42

6 System Testing and Evaluation **43**
6.1 Deriving Specification . . . 43
6.2 Testing Environment . . . 43
6.3 Test Cases . . . 43
6.3.1 Smoke Testing . . . 44
6.3.2 Sanity Testing . . . 44
6.3.3 Graphical User Interface and Performance Testing . . . 44
6.3.4 Login Test Case . . . 45
6.3.5 Logout Test Case . . . 45
6.3.6 Registration Test Case . . . 45
6.3.7 IOTA MAM M2M Transaction Test Case . . . 46
6.3.8 Geospatial Data Loading Test Case . . . 46
6.4 Azure Devops Environment . . . 46

7 Conclusions **49**
7.1 Discussion . . . 49
7.2 Conclusion . . . 49
7.3 Limitations . . . 50
7.4 Future Work . . . 50

A User Manual **53**

References **55**

List of Figures

1.1 Urbanization Trends . . . 2
1.2 Traffic Congestion Reason . . . 3

2.1 Spatial Trajectory Visualization . . . 8
2.2 Grouping of speed for traffic congestion threshold . . . 8

3.1 Structure Of Masked Authenticated Messaging . . . 12
3.2 Use Case Diagram . . . 13

4.1 Shape of the dataset . . . 23
4.2 dataset head . . . 23
4.3 dataset head . . . 23
4.4 Dataset Target value/Class Label . . . 24
4.5 dataset head . . . 24
4.6 Traffic Congestion Trends . . . 25
4.7 Shape files used for rendering maps . . . 26
4.8 Original Trajectory of cleaned spatial data . . . 26
4.9 Heatmap representing traffic congestion at different places . . . 27
4.10 Heatmap Analysis using Identify Feature option in QGIS . . . 28
4.11 Heatmap Results using feature selection tool . . . 28
4.12 System Architecture for Hazard Warning System Module . . . 30
4.13 MAM Bundle Structure . . . 30
4.14 Iterative Model . . . 31
4.15 Hazard warning system sequence diagram . . . 32
4.16 Registration sequence diagram . . . 32
4.17 Login Sequence diagram . . . 33
4.18 Logout Sequence diagram . . . 33
4.19 IOTA TANGLE MAM Sequence diagram . . . 34
4.20 Hazard Warning System database diagram . . . 34
4.21 Registration Page of Hazard Warning System Module . . . 35
4.22 Login Page of Hazard Warning System Module . . . 35
4.23 Home Page of Hazard Warning System Module . . . 36
4.24 Messages Page of Hazard Warning System Module . . . 37

5.1 MAM Architecture . . . 40

6.1 Devops Main Dashboard . . . 47
6.2 Creating work item . . . 47

6.3 Work item dashboard page 48

List of Tables

3.1 User login use case 14
3.2 User logout use case 14
3.3 User Registration use case 15
3.4 Input Message Use case 15
3.5 Spatial Data Analysis Use case 16
3.6 Input Message Use case 16
3.7 M2M Transaction processing Use case 17
3.8 Input Message Use case 18
3.9 Data Loading Use case for Geospatial Analysis 18
3.10 Data Visualization Use case 19

6.1 Test Case 1:Smoke Testing 44
6.2 Test Case 2:Sanity Testing 44
6.3 Test Case 3:GUI and Performance testing 45
6.4 Test Case 4:Login Test Case 45
6.5 Test Case 5:Logout Test case 45
6.6 Test Case 6:Registration Test case 45
6.7 Test Case 7:IOTA MAM M2M Transaction test case 46
6.8 Test Case 8:Geospatial Data loading Test case 46

Acronyms and Abbreviations

M2M	Machine2Machine
DLT	Digital Distributed Ledger
DAG	Direct Acyclic Graph
AI	Artificial Intelligent
GPS	Global Positioning System
ML	Machine Learning
ANN	Artificial Neural Network
MAM	Masked Authenticated Messaging
DS	Data Science
IOT	Internet of Things
AV	Autonomous Vehicles
DBSCAN	Density-based spatial clustering of applications with noise
USLA	Unsupervised Learning Algorithms
DDM	Digital Data Marketplace

Chapter 1

Introduction

In this chapter,I will be giving a brief introduction about my Final year project.The main idea of this final year project is actually to look for all the possible solution,tools,techniques and methods that I can use in order to solve the problem of high traffic congestion on the roads due to urbanization.
We are currently living in 2019 and we are blessed with so many tools,technologies that we can use in our benefit,with right guidance,and right method,we can definitely solve many problems that we are facing in our society,because nowadays we have high computational systems available to us that can handle and process large amount of data as well as provide us with analytics,data is indeed the most important asset of current century that is why data is called as new oil.Data is so handful and useful for us that just by using data,we can make decisions for better future.Data when properly cleaned,preprocessed and analysed,gives valuable information to us that we can use in making right decisions,nowadays almost every device is getting smarter,we are getting data from many devices,from cars,lights,bulbs,AC,heater,not only the home appliances,we are getting the data from Camera sensors as well that are installed in different smart cities,why all the data is being collected,the reason is that we want to find useful information that data can provide to us.
Data that I will use in this project is Geo-spatial data that will contain the GPS Coordinates.complete information regarding the dataset is provided in chapter 4 of the Final Year Project report.

1.1 Overview

“Intelligent Transportation System” is a component of smart city development,with the help of which,the transportation of the city can be improved a lot,which results in many benefits like less fuel consumption,less traffic jams and provides a better living environment for

the residents of the city. In the recent decades,the rate of urbanization is really fast and drastic,with the cities being more developed and people having more oppurtunities to work,there are better school facilities and medical facilities ,alongside many other such facilities that makes it very obvious to understand the reason for such increase in the urbanization,which is expected to even move further as the time progresses due to the advancements in the technology,and IOT (Internet Of Things) where internet connectivity plays a crucial role to help citizens performs the important tasks which were previously performed in quit tediously and in difficult manner,For example in past,and is still in the trend that if someone wants to perform some bank transaction,they go to the bank via some traffic source like cars,vehicles etc which result in unnecessary fuel consumption,traffic jams and all sorts of environmental pollution issues arise from these problems,so by the use of Intelligent Transportation System,a clean,productive environment can be provided to the citizen which will help them to improve their life standards and be productive for the development of the country.

According to a famous Statistics website,it was described that in the year 2017,the percentage of urbanization was 36.44 PC ,on the average if we consider this percentage,it is quite significant,in literal meanings it means that in every year,out of 100 people in a village/rural areas,almost 35 out of them settles in the Towns and urban places,which is quite a large number.

Below is the chart that shows the rapid increase in the Urbanization in Pakistan from Year 2007 to Year 2017 (10 years span).

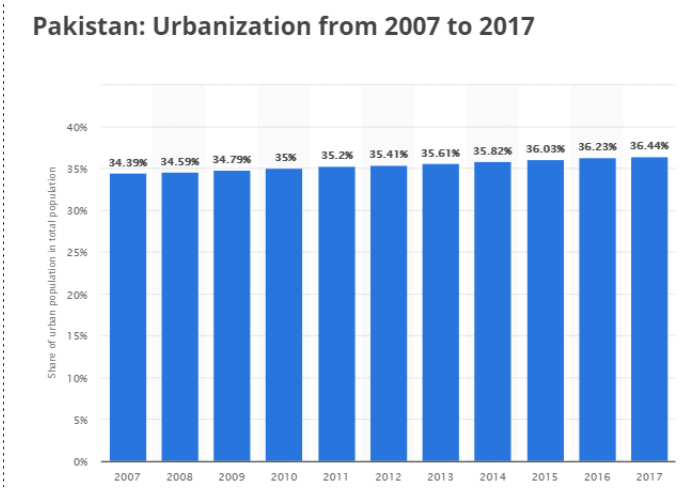


Figure 1.1: Urbanization Trends

Blockchain,Machine Learning and IOT are the latest technologies that are making their way to provide better solutions to the people in different walks of life. The main objective of this project is to prepare a framework,tool that provides a platform to reduce problems faced in traffic,to provide a solution for sustainable development. This application will consist

of an application that will help the user to understand the patterns of traffic congestion at different times of the day,that will be the Data Analysis and Machine Learning Module of the project whereas IOTA TANGLE will be used as a tool to increase the communication and trust levels among the drivers on the road,for example a driver is facing some problem on the road like he has gone through some accident and he is injured,so he will send a message via IOTA TANGLE and it will then be broadcasted to all other drivers that are subscribed to that particular MAM (Masked Authenticated Messaging) channel. Traffic Congestion and Trust among the vehicles on the road is very important for the whole society because of high rate of urbanization.So the main purpose of this Analytical and Dapp is to provide the user of the urban cities,a better,and improved experience on the road,to reduce the number of traffic jams by developing smart cities,to reduce number of accidents that happen every week and to provide a better experience to the people on the road which will be very beneficial for the environment.

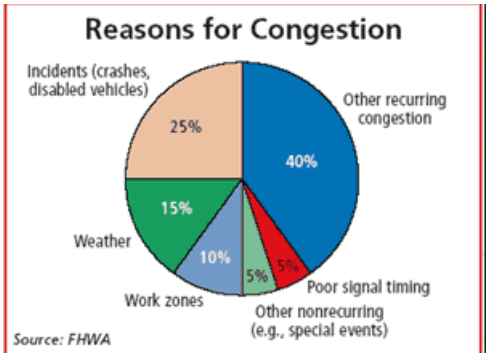


Figure 1.2: Traffic Congestion Reason

1.2 Problem Statement

To develop an application that would :

- Predict the level of traffic congestion at different times of the day.
- Allow the drivers on road to broadcast emergency messages on the road to the network of subscribed drivers.

1.3 Problem Description

Traffic congestion, air pollution and fuel wastage are few major issues that are caused by urbanization which requires efficient and smart solutions to deal appropriately. Traffic congestion has a number of negative impacts on the whole social environment, reduction of regional economic health, unnecessary fuel wastage, air pollution, delays (which may result in late arrival for employment, meetings, and education, resulting in lost business,

disciplinary action or other personal losses),inefficient traffic predictions, frustrated and stressed motorists, higher chance of accidents ,all of these are few of the major problems caused by traffic congestion, and if efficient data analysis of mobility data is carried out, different traffic congestion patterns can be detected which can prove to be a catalyst for the solution of traffic congestion problem and that is one of my major module in this project as I will be conducting data analysis of the geospatial data as well to find out the trends of traffic congestion.

Lack of direct communication between the vehicles is another huge factor that plays an important role in less efficient communication and understanding between the vehicles hence causing accidents and many other catastrophes. Also lack of communication and lack of trustful social environment of vehicles on the road can become a hindrance for the emergency requirement of vehicles or drivers like for example if a vehicle is having accident and the driver is injured, then rather calling the emergency or some hospital, efficient solution will be to inform the ongoing trip peers(i.e. other vehicles) so the necessary help can be provided in less time as opposed to ambulance or any other service which might take long time to reach that location and all of this can only be possible by developing a trustful, social network among the vehicles.

1.4 Objective

There are two main objectives of developing this Final Year Project :

- To develop a running prototype of Hazard Warning System.
- To analyze the geospatial data to predict the traffic congestion patterns.

1.5 Project Scope

This application as earlier stated would be designed to Predict the traffic congestion level at different times of the day for development of smart cities and a Hazard warning system by which driver will be able to broadcast the message when he is facing any issue on the road like accident and congestion. The project's main idea is to provide an interactive environment for user to practice driving on the road that will help the drivers/users themselves ,will avoid unnecessary fuel consumption and saving environment to get polluted . Intelligent transportation system consists of different units, that combines to for the complete ITS ,in this project ,the main focus and emphasis will be on the development of hazard and danger warning system, along with the use of mobility analytics for descriptive analysis of the geospatial data ,that is the basic scope of this project, as IOTA TANGLE is a developing architecture and technology in which there are a lot

on innovations and changings occurring on weekly and monthly basis, it provides the opportunity to use the latest structure for the implementation of the project.

Chapter 2

Literature Review

In this new age of technology, Datascience,IOT and Blockchain jhas proved to be much more than just a means of entertainment. The quantity and quality of spatial data are increasing rapidly. This is particularly evident in the case of movement data. Devices capable of accurately recording the position of moving entities have become ubiquitous and created an abundance of movement data. Valuable knowledge concerning processes occurring in the physical world can be extracted from these large movement data sets. Geovisual analytics offers powerful techniques to achieve this.

2.1 GPS Trajectories using Visualisation and Clustering:

This article describes a new geovisual analytics tool specifically designed for movement data. The tool features the classic space-time cube augmented with a novel clustering approach to identify common behaviour. Datascience,Machine Learning are not new technologies,it was actually started alot earlier,it is just because now we have alot,alot of data that we find it useful to detect the patterns from it,data is like new oil,now its upto us how we can utilize it to get the best from it,there are many techniques and methodologies that we can apply to get the data to find the hidden patterns inside the data,this is called data mining,to find the meaning,hidden meaning,hidden information from the data,and there are many techniques which help us to find that.

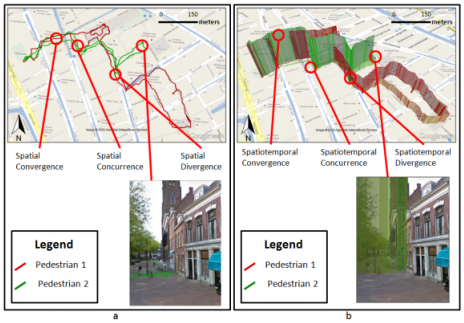


Fig. 5. Using 2D (a) and 3D (b) visualisations to display the path taken by 2 pedestrians in a city.

Figure 2.1: Spatial Trajectory Visualization

In general, determining the validity and quality of the clusters produced by any clustering technique is subjective and requires domain knowledge. While ground truth data can be used to assess this, in exploratory tasks, such data is often unavailable. In the case study presented here, the ground truth was of coarser detail than the categories produced by the clustering algorithm and so were not directly comparable. In future studies, we will collect such supplementary data alongside the raw movement data. This will enable the validity of the clustering results to be assessed. Additionally, we will assess the benefits of using this approach with movement data collected in other environments where speed and acceleration also act as a proxy for behaviour.

2.2 Method for assessment of Traffic Congestion of Urban road networks based on Speed

This is one of the most relevant and interesting research paper that I found after reading many research papers regarding Spatial Data Mining,in this research paper,the author have used speed attribute to analyse the traffic congestion patterns,following is the figure in which the divisions are shown for speed according to their dataset.

$$R_v = \frac{v}{V_{max}} \times 100 \quad (1)$$

where,
R_v denotes the speed performance index;
v denotes the average travel speed, km/h;
V_{max} denotes the maximum permissible road speed, km/h.

Table 1. The evaluation criterion of Speed Performance Index on expressway.

Speed Performance Index	Traffic State Level	Description of Traffic State
[0,25]	Heavy Congestion	The average speed is low, road traffic state poor.
(25,50]	Mild Congestion	The average speed is lower, road traffic state bit weak.
(50,75]	Smooth	The average speed is higher, road traffic state better.
(75,100]	Very Smooth	The average speed is high, road traffic state good.

Figure 2.2: Grouping of speed for traffic congestion threshold

As it can be clearly seen that the minimum speed they have is of 0 Kph while the maximum one is 100 Kph,while the dataset I am working on has 0 Kph as minimum speed and 160 Kph as maximum speed unit. To find the extent of traffic congestion

on the road,they have used average road segment state and the time of non congestion state in the observation period to find out the traffic congestion trends. I have taken this case study and I have build my own research and development based on this research with different threshold values based on the dataset that I have,and I have used Machine Learning algorithm to cluster the dataset based on the attributes of time(hour) and the speed attribute.

Chapter 3

Requirement Specifications

3.1 Purpose

The purpose of the project is to use the data,the mobility analytics to analyse the road traffic patterns for finding the reason for traffic congestion,so that it can be used for development a sustainable society,which can look after the high urbanization rate,to provide better transportation facilities to the common citizen by developing an Intelligent Transportation System.

3.2 Product Scope

Intelligent transportation system consists of different units, that combines to for the complete ITS ,in this project ,the main focus and emphasis will be on the development of hazard and danger warning system, along with the use of mobility analytics for descriptive analysis of the geospatial data ,that is the basic scope of this project, as IOTA TANGLE is a developing architecture and technology in which there are a lot on innovations and changings occurring on weekly and monthly basis, it provides the opportunity to use the latest structure for the implementation of the project. The details of the two module that I am trying to develop in this project are :

3.2.1 Hazard Warning System

The reason for developing hazard warning system is that when we look at current traffic system,there are few issues that need urgent attention,for example,if you imagine that a car is having accident,now the driver is injured,but the problem is that it will take quite some time to reach the target place and help the driver ,instead to make this time duration short and increase the probability of reducing the time period for the help reaching to the

driver,I will be developing a prototype,that will consist of an Arduino Uno,and GPS NEO 6m,IOTA TANGLE MAM channel,GPS MODULE will be used to send the information like (Latitude,Longitude,timestamp) and a message will be written from Arduino and it will be concatenated to the GPS location,and it will be broadcasted to the MAM channel where all the drivers subscribed to the channel will be able to read that message.

3.2.2 Traffic Congestion Analysis

The second module of the project is the Machine learning part that will be doing the Data Analysis and prediction of traffic congestion at different times of the day.

The strategy I will use in this module is that I am having the dataset of Beijing Bicycle,the dataset is of 182 users,and it is provided by Microsoft Geolife. I am doing the data analysis of the Spatial data,which will be both descriptive and visual analysis of the data.Then I will perform Feature Engineering.This project module can be further divided into few parts for the sake of better understanding.

3.3 Overall Description

3.3.1 Product Perspective

The main components of the systems can be clearly distinguished in the following block diagram. User would be able send the message to the IOTA TANGLE MAM,which can be seen by all the drivers that are subscribed to that MAM channel.

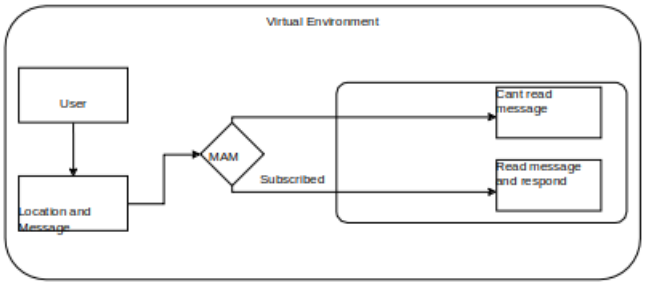


Figure 3.1: Structure Of Masked Authenticated Messaging

Once the subscribed drivers read the message ,they will be able to know the location of the driver who is facing a problem and also the problem that he is facing.

3.3.2 Product Functions

The application prototype would provide a virtual platform that can be served as bases for practicing transportation system. The system would use the message generated by the

driver that consists of the problem he is facing and the location of the driver. And the second module will analyse the dataset, for understanding the spatial trajectory patterns.

3.3.3 User Interface

The User interface would be kept as simple as possible. Since there would be no complex setting that are required to be done by the user, the user interface would simply consists of a menu that would guide user to proceed to send the message that consist of the GPS location and the message from the driver.

3.3.4 Hardware Interface

The application would provide a virtual platform that can be run on a website. Following are the hardware specifications required for the project.

- GPS Module Neo6m to send the GPS coordinates.
- Arduino Uno

3.4 Use case

Following is the usecase Figure:

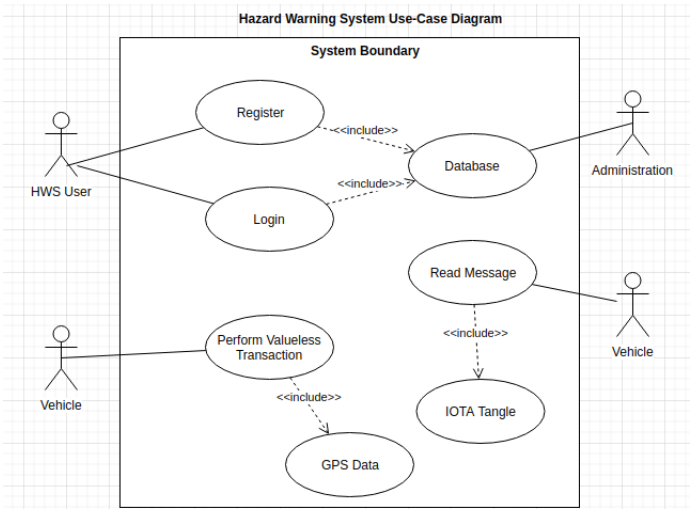


Figure 3.2: Use Case Diagram

3.4.1 User Login Use case

Following is the use case for user login usecase

User Login Use case	
Actors	IOTA TANGLE user,MongoDB,Passport.js
Description	To use the System,the user must be login first so the application can be used.
Flow of events	If the user is not registered,he should first register,then open the login page,and then enter the credentials and click on login.
Pre Condition	The user must be registered and the credentials entered in the login page must be valid.
Post Condition	The user is logged in to the system successfully.
Data	Username and password
Stimulus	User wants to use the application.
Response	Login successful
Comments	It is important for a user to be registered first to use the system

Table 3.1: User login use case

3.4.2 User Logout Use case

Following is the use case for User Logout

User Logout Use case Message	
Actors	IOTA TANGLE user,Passport.js
Description	Once the user has performed the desired action,then he has to logout as well from the system.
Flow of events	User first registers,after registration,the user login into the system,after login,the user perform desired activities,then user clicks on log out to get logout from the system
Pre Condition	User is logged in
Post Condition	User is logged out.
Data	Session cookies
Stimulus	User wants to logout from the system
Response	Logout successful
Comments	User has to logout after performing the activities.

Table 3.2: User logout use case

3.4.3 User Registration Use case

Following is the use case for user registration

User Registration Use case	
Actors	IOTA TANGLE Non-user,MongoDB
Description	To use the System,the user must first registered in the database.
Flow of events	If the user is not registered,he should first register by providing valid information.
Pre Condition	The user wants to use the system and provide valid credentials while registering.
Post Condition	The user is registered successfully.
Data	username,password,confirm password,email.
Stimulus	User wants to use the system so he has to register first.
Response	Registration successful
Comments	It is important for a user to be registered first.

Table 3.3: User Registration use case

3.4.4 Input Message

Following is the use case for input message

Input Message	
Actors	Driver/User,Subscribed users
Description	Whenever the driver see some threat on the road or face some emergency,he will broadcast the message with the location.
Flow of events	User will open the application,then send the message,the message is encrypted,which only can be read by users having the side key
Pre Condition	The user has already subscribed to MAM channel The user faces some problem like accident on road,or traffic congestion
Post Condition	The message is successfully delievered to MAM channel
Data	Text Message,GPS location
Stimulus	User wants to broadcast the message to subscribed users.
Response	Message is broadcasted
Comments	Only those users can read the message who are subscribed to the MAM channel

Table 3.4: Input Message Use case

3.4.5 Spatial Data Analysis

Following is the use case for spatial data analysis

Input Message	
Actors	Vehicles,Spacial Data
Description	Data is saved ,feature engineering is performed,and traffic congestion prediction is made
Flow of events	Data is collected Data Preprocessing is performed Once the data is cleaned,It is passed to the model for prediction
Pre Condition	Data should be available Model should be trained
Post Condition	Data is cleaned Traffic congestion prediction is made
Data	Geospatial Data
Stimulus	User wants to analyse traffic congestion
Response	Traffic congestion prediction is made
Comments	Data should have Geospatial points (Latitude and Longitude)

Table 3.5: Spatial Data Analysis Use case

3.4.6 Perform System Testing

Usecase for performance testing

Input Message	
Actors	Fullstack Developer and Software Quality Assurance (Myself)
Description	All modules and sub modules to be tested
Flow of events	WhiteBox testing,of every unit from start
Pre Condition	Input with both valid and invalid data
Post Condition	Result/output
Data	Text Message
Stimulus	Validation
Response	Confirmation
Comments	Every module to be tested seperately

Table 3.6: Input Message Use case

3.4.7 Machine2Machine Transaction Use case

Following is the use case for M2M(Machine2Machine Transaction usecase)

M2M transaction use case	
Actors	IOTA TANGLE User,Transaction(can be valueless as well as contain value)
Description	User has logged in to the system and perform the transaction,now proof of work has to be performed for the transaction to be approved.
Flow of events	User has registered,logged in to the system and now wants perform transaction,transaction is launched,now proof of work is performed,2 previously unattended transaction has to be first approved,it acts as an incentive for the user
Pre Condition	User is logged in and transaction is performed
Post Condition	Two previously unattended transactions are approved,so the transaction is being processed for approval
Data	transaction,can be valued or valueless,user credentials.
Stimulus	User wants perform a transaction.
Response	Transaction is processed successfully.
Comments	DAG is really a good choice for fast systems in which transactions are performed within miliseconds by the sensors,and incentive is provided to the user to keep approving the transactions,hence removing the role of miners as seen in other distributed ledger technologies such Bitcoin,Etherium etc.

Table 3.7: M2M Transaction processing Use case

3.4.8 Input Message

Following is the use case for input message

Input Message	
Actors	Driver/User,Subscribed users
Description	Whenever the driver see some threat on the road or face some emergency,he will broadcast the message with the location.
Flow of events	User will open the application,then send the message,the message is encrypted,which only can be read by users having the side key
Pre Condition	The user has already subscribed to MAM channel The user faces some problem like accident on road,or traffic congestion
Post Condition	The message is successfully delievered to MAM channel
Data	Text Message,GPS location
Stimulus	User wants to broadcast the message to subscribed users.
Response	Message is broadcasted
Comments	Only those users can read the message who are subscribed to the MAM channel

Table 3.8: Input Message Use case

3.4.9 Data Loading Usecase for Spatial Analysis

Following is the use case for Data Loading use case

Data Loading use case	
Actors	Geospatial Data,Data Analyst(Myself)
Description	To analyze the Traffic Congestion trends,the Geospatial data has to be first loaded into the system
Flow of events	Data is selected first and then loaded into the system for analysis.
Pre Condition	Geospatial data is successfully loaded into the system
Post Condition	Traffic Congestion Analysis and prediction is performed
Data	Geospatial Data
Stimulus	Traffic Congestion to be analysed.
Response	Geospatial data is analysed
Comments	The data should be large in size and must be clean(processed) so better analysis can be performed.

Table 3.9: Data Loading Use case for Geospatial Analysis

3.4.10 Data Visualization Use case

Following is the use case for Data Visualization

Data Visualization Use case	
Actors	Qgis(Data visualization tool specially for Geospatial data)
Description	Once the Geospatial data is analyzed,then it indicates the times of the day when the traffic congestion is more and when it is less, i.e freeflowing,so as it is now clear from the analysis that congestion is more,the next task is to see from visualization that at that places we have the most congestion,it can only be seen from visualization (heatmaps,dotplots) as I have the traffic congestion target value now.
Flow of events	Data is loaded in the system,data is cleaned after the data preprocessing is performed,traffic analysis is done,Selected data is used for visualization
Pre Condition	Data is clean and Data Analysis is performed.
Post Condition	Real world Data visualization is created showing the traffic congestion trends visually
Data	Geospatial Data
Stimulus	Performing the visualization for understanding the trends visually
Response	Data Visualization is done successfully showing the traffic congestion trends.
Comments	Data Visualization is such a beautiful and informative technique that it explains each and everything which the data analysis reflects.

Table 3.10: Data Visualization Use case

Chapter 4

Design

Systems design is the process of defining the architecture, components, modules, interfaces, and data for a system to satisfy specified requirements. This chapter should have the following sections:

4.1 System Architecture

This application is a dual module application which is intended to run on the Ubuntu Operating system.The module of hazard warning System to use IOTA TANGLE for the Peer To Peer Network,MAM (Masked Authenticated Messaging) for the message broadcasting and subscription of drivers. Following is the basic interaction of this application.

4.1.1 Traffic Congestion analysis and prediction module Methodology,Analysis and Results

Nowadays we have alot of data,is data valueable?can we achieve something from data?well ,these are some of the questions that comes into our minds whenever we hear the words datascience,artificial intelligence and such words,if we analyse Machine Learning algorithms we find out that actually data is one of the most important asset we have these day,why?because we can do thousands of different useful things using data,on of which is data mining,by applying certain principles,we can find out all the trends and hidden insights from the data in less time,which we can use for many purposes.The reason I have performed this research is because I strongly believe that we can use the hidden information that we find out using data mining techniques,we can use them to tack the problems and issues caused by the urbanization,urbanization is at peak these days due to more opportunities for people in cities then rural areas,now since more and more people are moving towards cities,that also means that more traffic on the streets,which indirectly means that more

noise,air pollution,traffic jams,different diseases,high stress level,poor affect on the minds of people,frustrations,all of these are few of many problems that are caused by traffic congestion. Now as a Computer Scienctist,a DataScience enthusiasts and a Machine learning student,its my our responsibility to use the technology as a tool to fight with this problem which is growing day by day.Following is the information regarding the dataset I have used and the methods,techniques that I have applied to achieve my desired results in this research and development module of my final year project.

Dataset Information:

Following is the information regarding the dataset used for the Traffic Congestion Analysis and prediction module.
Source: www.figshare.com
Country: China
City: Beijing
Location: Ring Road
Year:2011
Dataset Posted By: Dr. Wang, Jingyuan,He is also the leader of the Big Data Intelligence for SmartCity development and research group in University of Beihand,China
. **Number of days:**30
Number of trajectories per day:288
Unique Taxi's: Yes
Time Duration:24 Hours for 30 days
Dataset File Type for analysis: Comma Seperated Value File(csv)
Target value/Class Label: None (it will be Traffic-Congestion-Level) assigned after applying unsupervised algorithm
Dataset file for visualization in QGIS: Vector Layer:Shape File (.shp extension)
Features: Following are the features :

- Longitude
- Latitude
- Speed
- Timestamp
- Direction

Exploratory Data Analysis(EDA) and Data cleaning:

Previous Subsection was about the brief overview of the dataset used. Due to hardware limitation and processing issues,I have used 70 taxi's trajectory data for the traffic congestion analysis module so as to work with the data efficiently and in best possible way.

Shape of the dataset : Shape of the dataset shows the total number of records and features in the dataset.

```
In [9]: df.shape
Out[9]: (157266, 5)
```

Figure 4.1: Shape of the dataset

Head of dataset : head is a Pandas function that is used to have a look of first 5 rows in the dataset,it gives us a rough idea of the features and the values.

```
In [8]: df.head()
Out[8]:
```

	longitude	latitude	time	speed	direction
0	116.303913	39.939207	20111101010630	45	178
1	116.303926	39.938604	20111101010731	47	180
2	116.267906	39.906812	20111101010955	73	178
3	116.268172	39.907342	20111101010606	75	0
4	116.269191	39.919079	20111101010711	67	2

Figure 4.2: dataset head

Preprocessing Timestamp: In the dataset,the timestamp is in this format yyyymmddhhmmss,in which I just want the Hours portion for the unsupervised learning algorithm,so after preprocessing,i get the hour.

Data for Kmean Clustering : For Kmean clustering,I have used the hour and speed feature.

```
In [23]: x_feature.head()
Out[23]:
```

	speed	hour
0	45	1
1	47	1
2	73	1
3	75	1
4	67	1

Figure 4.3: dataset head

Getting the target value : After performing the kmean algorithm on the dataset,I get the points that represent my target label,which I have named it as Traffic-Congestion-level as seen in this figure.

```
In [78]: classification_dataset.head()
```

Out[78]:

	traffic_congestion_level	latitude	longitude	direction	hour	speed
0	0	39.939207	116.303913	178	1	45
1	0	39.938604	116.303926	180	1	47
2	0	39.906812	116.267906	178	1	73
3	0	39.907342	116.268172	0	1	75
4	0	39.919079	116.269191	2	1	67

Figure 4.4: Dataset Target value/Class Label

The figure shows the target label with values.

Criteria for Speed : The minimum speed in the dataset is 0Kph,while the maximum speed is 160Kph.I have used the following criteria for the speed grouping.

- 0 To 65 (Heavy Traffic congestion)
- 65 To 90 (Medium Traffic congestion)
- 90 To 160 (No traffic congestion)

Criteria for Time(Hour) : There are total 8 ring roads in China,which are quite vast and they connect different cities of China.Following is the criteria that I have set for the time classification for clustering

- 04:00 To 10:00 (Morning Time)
- 10:00 To 16:00 (Noon and Afternoon Time)
- 16:00 To 22:00 (Evening Time)
- 20:00 To 24:00 Or 00:00 To 05:00 (Night Time)

Traffic Congestion Prediction Results : After performing unsupervised learning K-Mean,following results were achieved

Out[52]:

		traffic_congestion_level
hour	speed	
1	1	54.094337
	2	8.930093
	3	0.080119
2	1	0.000000
	2	0.000000
	3	0.000000
3	1	19.926112
	2	2.329811
	3	0.019076
4	1	6.921394
	2	5.315834
	3	0.102374

Figure 4.5: dataset head

It is clearly visible from the results that the most heavy traffic jam is in the early morning and in the afternoon,now I have found out the time of the day when the congestion

is most,the next task is to find at what places the congestion is alot,and their is heavy traffic flowing,and also the reasons for that,for this purpose I have taken use of an opensource sofftware called QGIS which is specially developed for analyzing spatial data and visualizing it.

Data Visualization:

Once the exploratory data analysis(EDA) and data preprocessing is performed,and the results are obtained,then the next task is to visualize the dataset.Different python libraries are available for that like matplotlib and many other tools,like ARCGIS,D3.JS,Geopandas and many more,but after trying different tools,i found the QGIS software as a very nice tool for spatial data analysis due to its simple interface and easy to use functions,and tools inside it.Also it has package manager so if I want some package for analysis or visualization,I can directly install that using package downloader.

Traffic Congestion Trends based on results : One of the best advantage of using data for finding the information is that it provides us with information of alot of trends,we can find many trends using data. It can be clearly seen in the plot that traffic congestion is quite high in the morning,then in the afternoon its less,and then in the evening,their is a peak in the traffic congestion.X axis represents the Hour(1,2,3,4) where 1 represents morning,2 represents afternoon,3 represents evening and 4 represents night time.on Y-axis is traffic congestion level.

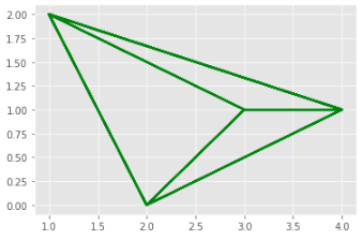


Figure 4.6: Traffic Congestion Trends

Shape Files : Shape file is a file that is used to store vector data storage format for storing the attributes,shape and location for drawing on map.Previously I had CSV file which i exported using pandas function to-csv(directory),and from qgis i exported them as shape file so i can use them to make visualizations on map using bing map. Originally I had the main file,then I created three different layers from the main trajectory using filtering based on traffic-congestion feature.the values with 0 were named as heavy-congestion,values with 1 were named medium congestion and values with 2 were named no-congestion.

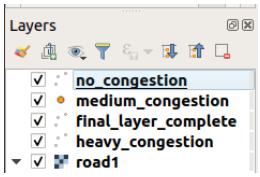


Figure 4.7: Shape files used for rendering maps

Trajectory of Preprocessed Spatial data : Trajectory is a set of coordinated that represents a collection of spatial(longitude and latitude) data points with respect to changing time.Following is the trajectory for spatial data.

Figure 4.8: Original Trajectory of cleaned spatial data

Heatmap showing the traffic congestion : When we want to analyse the spatial data,heatmap is one of the best tool we have for that purpose,it helps us to analyse every single part of the trajectory on map,I have used the opactity property to 50pc so as to analyse the heatmap in best possible way.Light areas in the heatmap represents less congestion whereas the dark areas represent high traffic congestion.

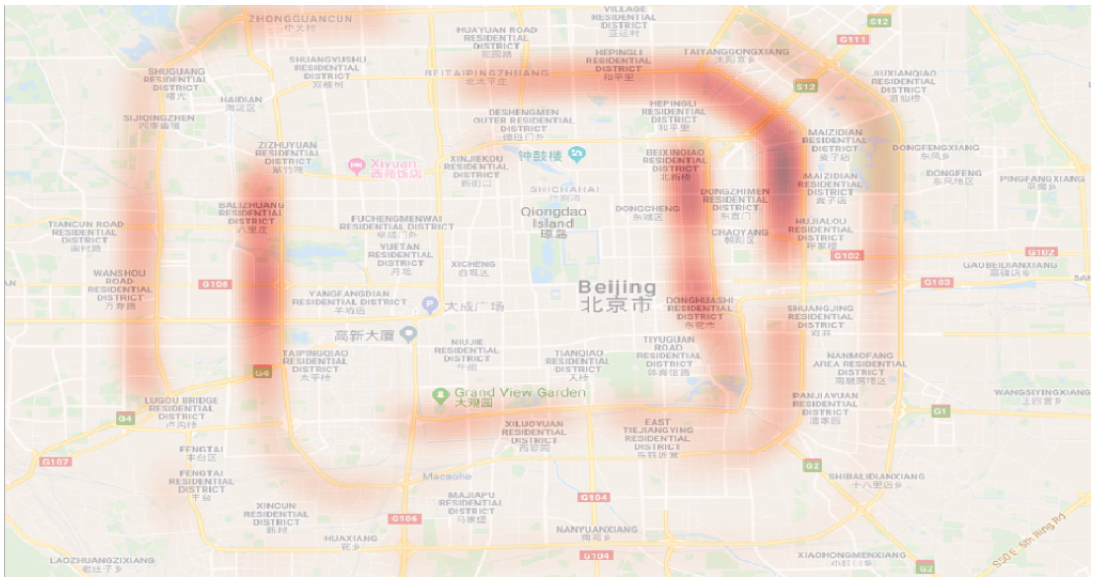


Figure 4.9: Heatmap representing traffic congestion at different places

Where is the maximum traffic congestion and why?

From my analysis,i have analysed and predicted that at what time of the day is most traffic congestion and also how much is the traffic congestion in terms of pc(percentage).The next task is to find out the location where the maximum traffic congestion is. As it is quite clearly seen from the heatmap that in the top right corner/side we have a quite darker area/spot in the heatmap,it represents that this has higher congestion rate,now how can i find the coordinates of this place?for that I have 2 options,one is to see the coordinates in status bar down the screen or I can also select that feature using identify features option in the menu bar,which is much better option as it also give alot of information and details about that point,as it can be seen in the figure.

Result based on heatmap analysis

Now that I have the coordinates of place I was interested in to find information about,I have got that point using identify feature tool in QGIS.I have used that tool to find information about the place where the traffic congestion maximum,one of many spots in the trajectory where the traffic congestion is maximum. Following are the places that lie in that area :

- Beijing Zeizhong Elementary School
- China Minsheng Bank 24 hour self service
- Beijing Post Office

Now it can be quite easily compared and seen that why traffic congestion is high at longitude 116.4 and latitude 39.9,the reason is there are important public offices,schools and

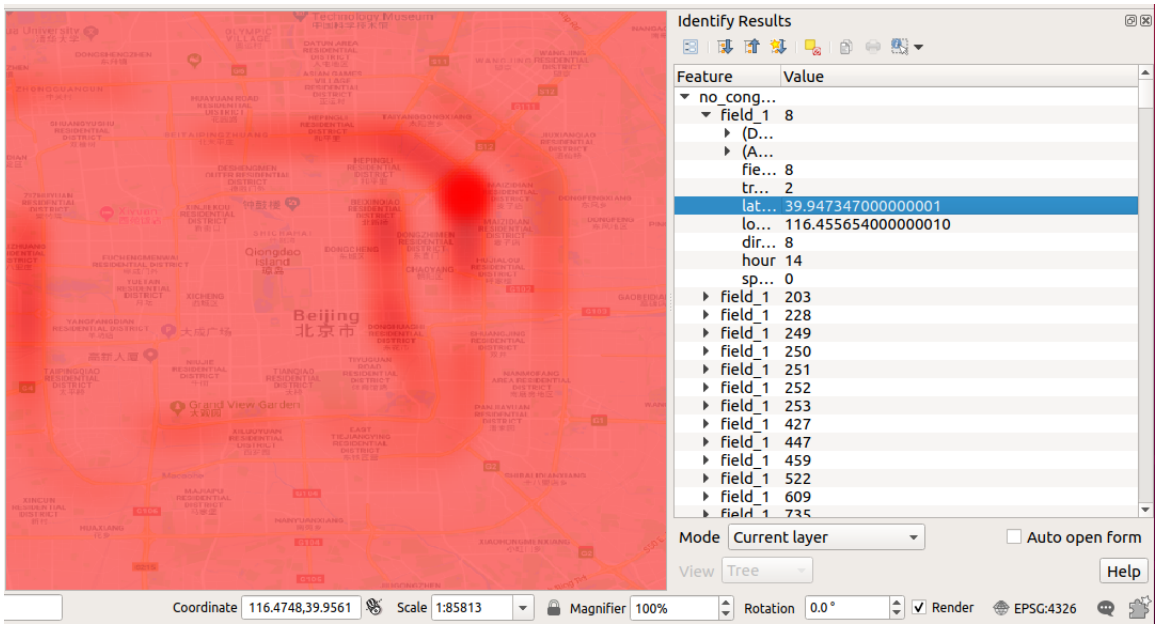


Figure 4.10: Heatmap Analysis using Identify Feature option in QGIS

post office at that location,thats why the traffic congestion is high their in the morning and evening mostly because of heavy traffic,movement on this route.Kids,teachers,professor,post office staff and many bank workers,all are using that ring route to reach to their destinations.

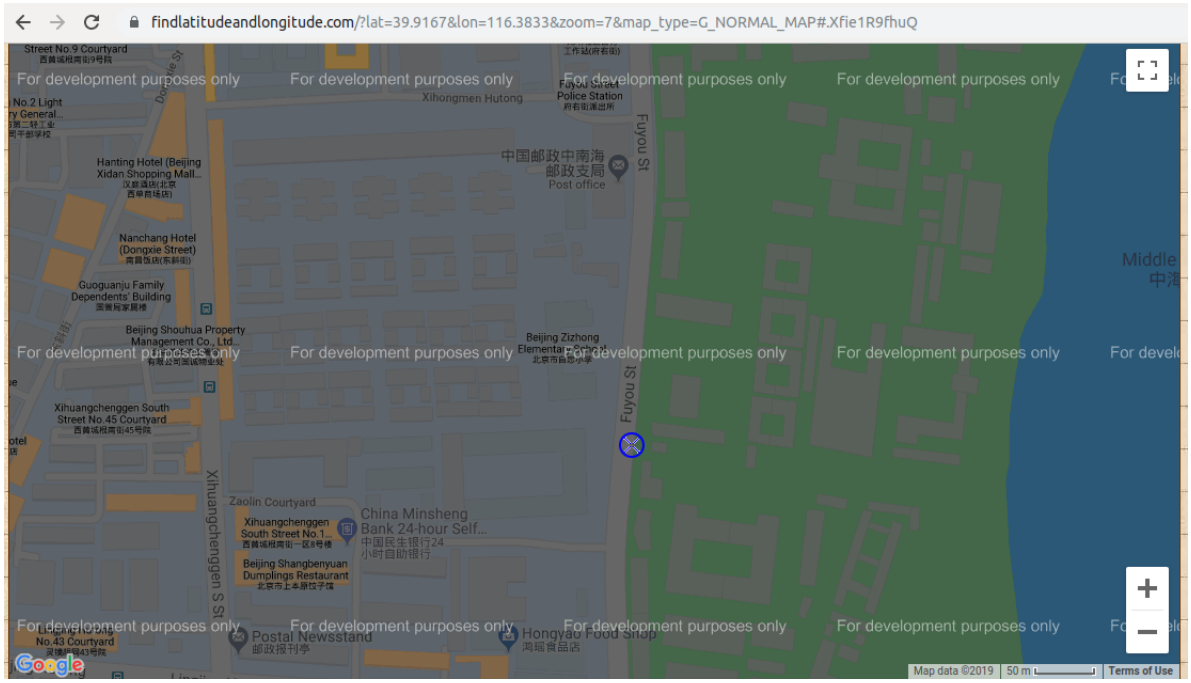


Figure 4.11: Heatmap Results using feature selection tool

Solution for the traffic congestion

Now I have all the data,analysis,visualization and and information regarding the traffic congestion,it can be broken down into small parts in trying to achieve the best possible solutions to counter this issue.

- In the ring road,provide Uturns where their is too much of congestion,this can be achieved by monitoring the realtime traffic at rush hours,which is in morning period as proven in this project(Traffic Congestion Analysis Module)
- Automate the street signals,normally the traffic light signals are not automated,which results in poor coordination on traffic sometimes resulting in chaos and frustation for traffic users on roads,automating will help optimize regulating the whole traffic,for example,when the traffic is heavy,congestion is high,the traffic light signals can be set for high time,and when the traffic is low,the congestion is low,maybe the traffic signals can be automated to like 30 seconds for saving people time and keep traffic moving on the road.
- Provide alternative routes for important destination near this spot/part of heatmap. These are few of many solution that can be used in order to deal with the issue of traffic congestion.

Conclusion:Traffic Congestion Analysis,prediction,findings and Possible Solutions

This was about my finding,analysis,and results i achieved based on Beijing Ring road dataset.It can be clearly seen how important the role of data is in order to find the root of the problem,in software development root cause analysis is one of the most important method used for finding the defect,issue and using that for solving it. This analysis can be applied on Pakistan Traffic as well,and this is one of my future goal as well to use latest scientific approaches in solving the largescale issues and problems faced by Pakistan due to Urbanization.There are many solutions to every problem and they can always be used to fix that problem.

4.1.2 Hazard Warning System Module

Hazard Warning System module is the IOTA TANGLE based module.I have tried to develop a prototype that resemble the functional behaviour of autonomous vehicles that are currently under test,Autonomous vehicles uses censors to detect the environment and the scenarios and they react accordingly.It all works on Machine2Machine transaction,IOTA Tangle provides solution for that as it allows valueless machine2machine (M2M) transactions as well. Following is the flow for the module

- User will open the interface and write the text information that the user want to pass on to the subscribed drivers.
- GPS MODULE NEO 6m, installed on the Arduino Uno will be used to deliever the message of the user to MAM.
- The users subscribed on the MAM (Masked Authenticated Message) channel will be able to read the message and act accordingly.
- Those users will be able to read the message who are subscribed to the MAM channel and have the side key of the channel as the mode used is restricted one.
- Once the users read the message they can respond to the message accordingly and a trustful environment can be developed on the road.

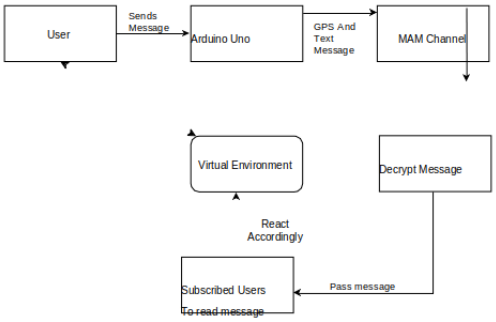
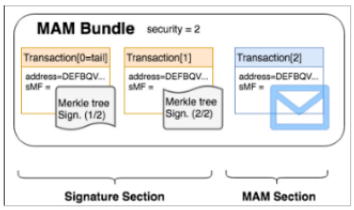


Figure 4.1: System Architecture

Figure 4.12: System Architecture for Hazard Warning System Module



*sMF = signatureMessageFragment

*[ ]: current index

*signature: security*2187 [trytes]

* MAM section gets longer, if message gets longer than 2187 trytes.

* Merkle tree is explained later

Figure 4.13: MAM Bundle Structure

4.2 Design Constraints

Following are the design constraints

- The system is not designed for Android Application

- The system requires internet connectivity for message delievery
- Subscribed Users on MAM Channel can read the information
- GPS Module is required to transfer the location information
- SerialConnection is used to send data from Arduino Uno to IOTA Tangle

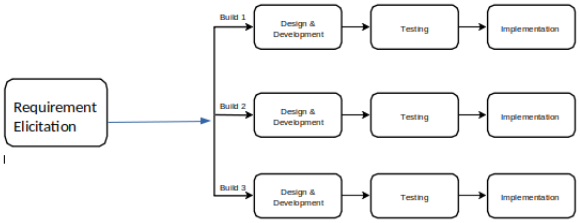


Figure 4.14: Iterative Model

This section describes any constraints in the system design (reference any trade-off analyses conducted such, as resource use versus productivity, or conflicts with other systems) and includes any assumptions made during the developing the system design.

4.3 Design Methodology

Design methodology of this application is described below

4.3.1 Process Model

I will use Iterative model to create this application. In Iterative model,we do not start with complete specification,rather the development is started with simple set of software requirement, which is iteratively updated until the complete requirements are fulfilled by the system and the system is ready to use. The best feature of Iterative development is that,using this model,I will be able to analyse my application much better and understand what improvement I need to make in the current version based on my final product requirements.

4.4 High Level Design

In this section,the system is viewed as high level design.

4.4.1 Hazard Warning System Process

Following is the complete process diagram of the system

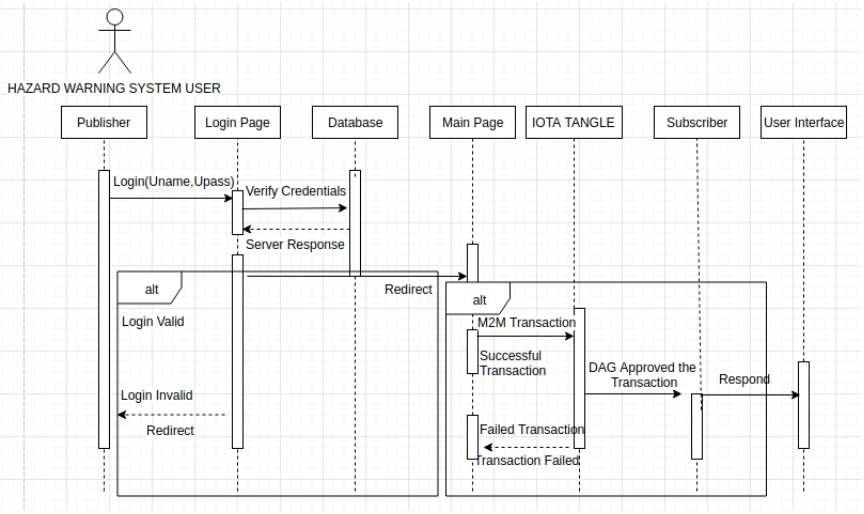


Figure 4.15: Hazard warning system sequence diagram

4.4.2 Registration Process

Following is the process diagram for registration

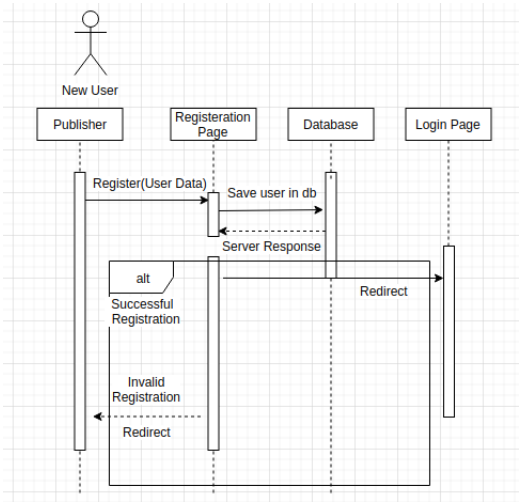


Figure 4.16: Registration sequence diagram

4.4.3 System Login Process

Following is the complete process diagram of the Login

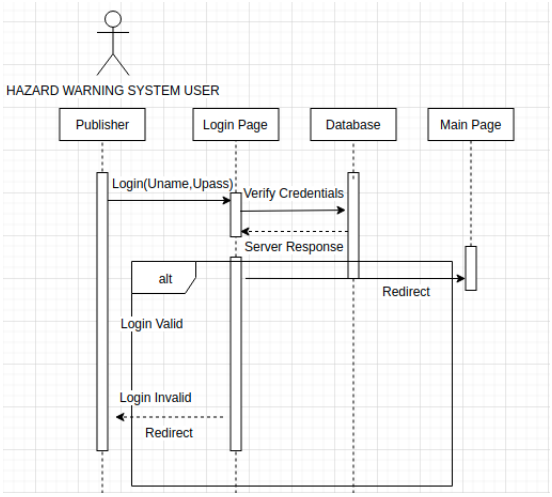


Figure 4.17: Login Sequence diagram

4.4.4 System Logout Process

Following is the complete process diagram of the Logout

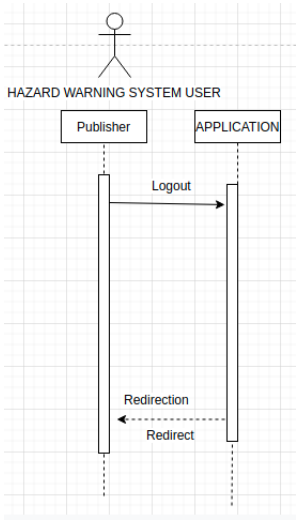


Figure 4.18: Logout Sequence diagram

4.4.5 IOTA TANGLE Process

Following is the complete process diagram of the IOTA TANGLE MAM transaction

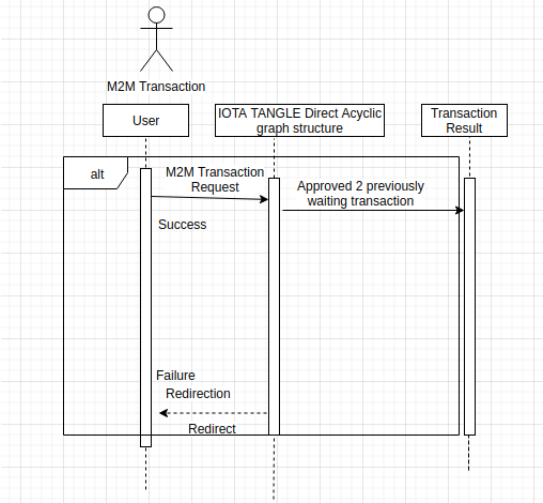


Figure 4.19: IOTA TANGLE MAM Sequence diagram

4.5 Database Design

Database consists of the relations and their relationships with each other.UserId is the primary key for both the tables and it is also a foreign key for Users table,and there is a composition between the 2 tables,its a 'has a' relationship,one to many.One user can send multiple messages across the network.

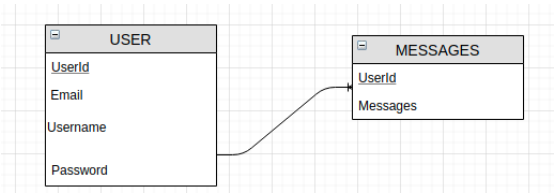


Figure 4.20: Hazard Warning System database diagram

4.6 GUI Design

Graphical User Interface is one of the most important aspect of any software application,there is a complete field that is Human Computer Interaction that has been designed specifically to design the best interface for the application.I have tried my level best to design the best Graphical User Interface for my project.

4.6.1 Registration Page

The first step in any application is that,the user must register first to use that application,following is the registration page in which user will enter their valid information to get registered in the system.

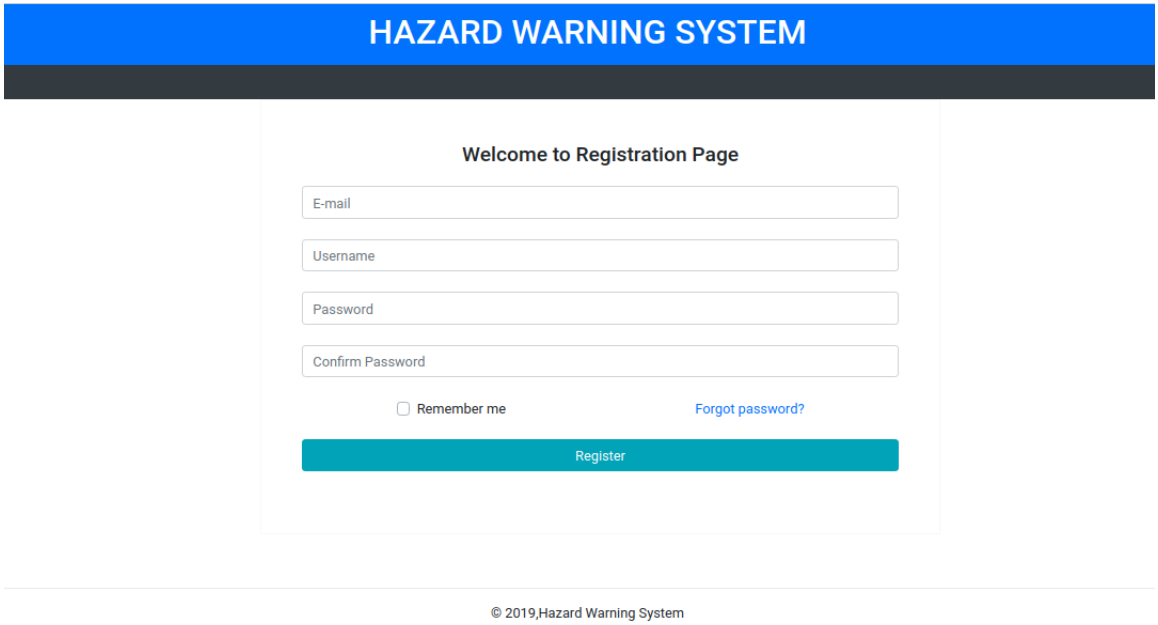


Figure 4.21: Registration Page of Hazard Warning System Module

4.6.2 Login Page

Once the user is registered,then they can login and use the system.Following is the figure of login page.

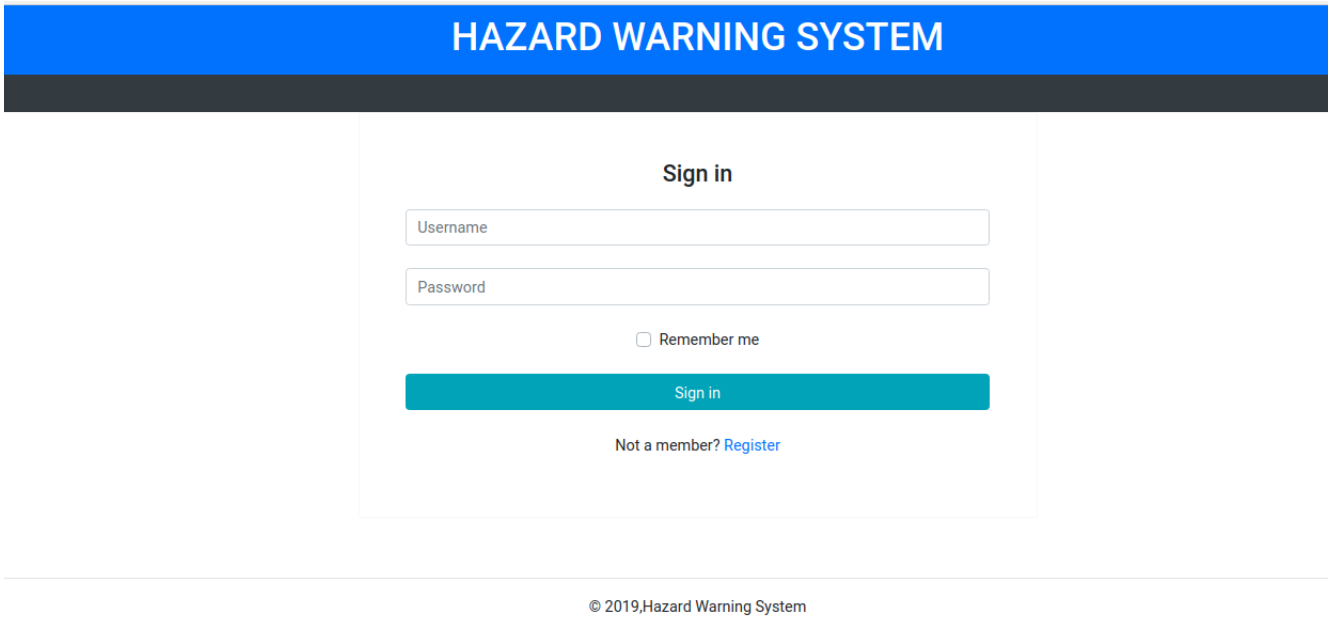


Figure 4.22: Login Page of Hazard Warning System Module

4.6.3 Home Page

When the credentials of user are verified in the login page,the user is redirected to the Home Page.

Figure 4.23: Home Page of Hazard Warning System Module

4.6.4 Messages Page

All the users who are registered in the system can view the messages.

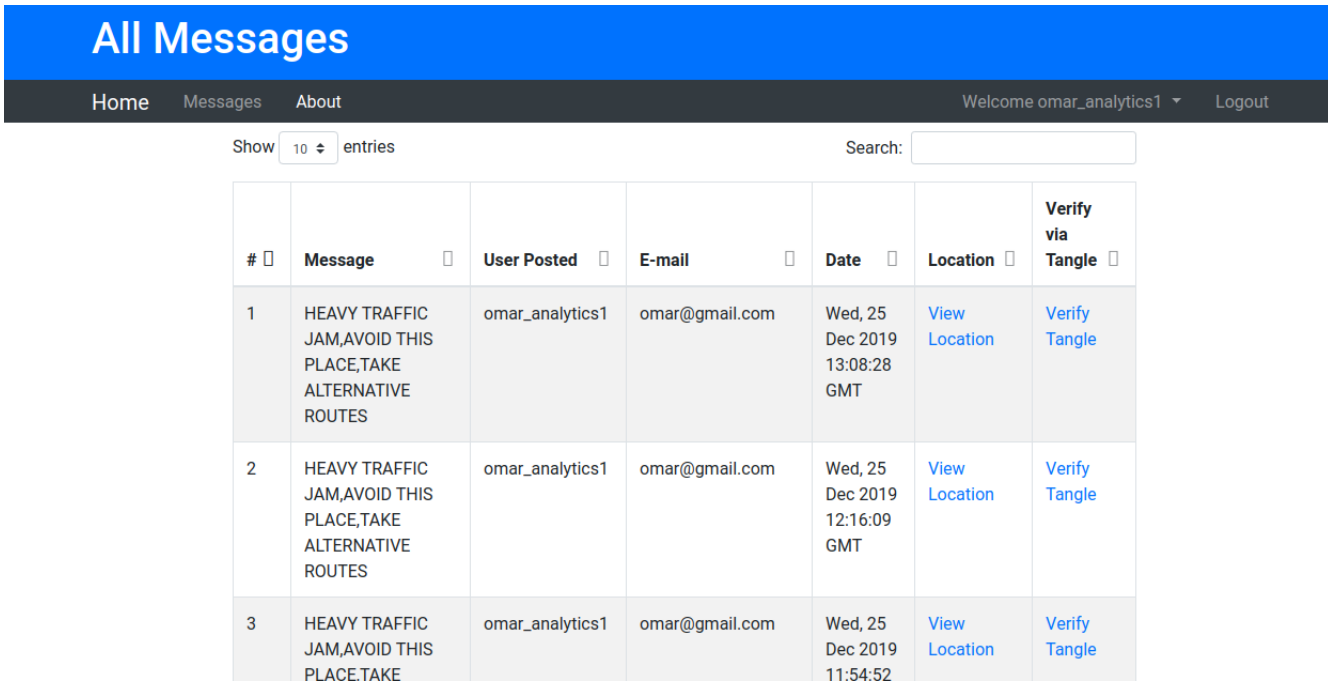


Figure 4.24: Messages Page of Hazard Warning System Module

Chapter 5

System Implementation

Implementation is the process of moving an idea from concept to reality. The System implementation is a realization of a technical specification or algorithm as a program, software component, or other computer system through programming and deployment.

5.1 System Architecture

The system takes input in the form of message (string) that contains the combination of text and the GPS Coordinates so all the users(drivers) nearby can receive the message,and respond to it.The architecture followed in the Hazard Warning System is Publish/Subscribe architect pattern.The publisher publishes a message via Masked Authenticated Messaging Restful API and the mode of communication is restricted so that only those drivers subscribed to that channel will get the notification. The structure of MAM is such that once a message is passed to the tangle,it can never be edited or deleted,this property of IOTA TANGLE ensures transparency and security.Another main quality of the IOTA TANGLE is that it is very flexible,it allows zero value transactions as well,which are called valueless transactions,in which zero value is used.These are really helpful in developing the networks where data is recieved from the sensors,it is recieved in huge numbers,which normal database can not handle and gets crashed.If we take example of Bitcoin for example,it is also very fast that is one of the reason that IOTA TANGLE is preferred for IOT related projects,where data is recieved from censors.One of the main reason that IOTA TANGLE is fast is that there are no miners involved in the process like there are in other Distributed Ledger Technologies like bitcoin,etherium etc,now the question arises,if there are no miners,then how are transactions verified?where is the incentive,actually the incentive lies in the beautiful structure of DAG,DAG stands for Direct Acyclic Graph,the incentive is that for a transaction to be completed,it has to first confirm previous 2 transactions only

then it will get processed.That is the main incentive and it promotes honesty among the users. There are three modes of communication possible via MAM,these are :

- Private Mode
- Public Mode
- Restricted Mode

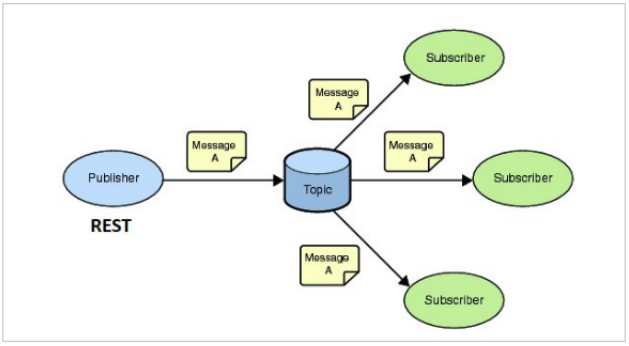


Figure 5.1: MAM Architecture

5.1.1 Development Environment and Tools

I have used following tools and technologies for my Final Year Project.

5.1.1.1 Tools and Technologies used for Traffic Congestion Analysis and Prediction Module

I have used following tools and technologies for the Traffic Congestion Analysis and Prediction module:

1. QGIS
2. Anaconda Navigator
3. Jupyter Notebook
4. Pandas
5. Numpy
6. Matplotlib
7. Seaborn

5.1.1.2 Tools and Technologies used for Hazard Warning System Module

I have used following tools and technologies for the Hazard Warning System Module:

1. Visual Studio Code
2. Arduino IDE
3. Mlab
4. MongoDB
5. Node.js
6. IOTA TANGLE MAM(Masked Authenticated Messaging)
7. Passport
8. React
9. Express

5.1.1.3 Tool for Documentation

There are different options available for documenting the project,I have used **TeXstudio** for documention due to its precise and controlled environment for documentation.My experience with TeXstudio has been really nice because of the Html like syntax as tags are used to perform the documentation.For providing references I have used Online website **www.citethisforme.com** as I was having an issue with providing refence using TeXstudio,it was a nice tool for providing the references.

5.1.1.4 Tool for Testing

In development,it is really important to test to code while it is being developed,if it is left at the end to be performed,then it makes it really difficult sometimes to do debugging.For testing purpose I used **Postman** and **Azure Devops** tool.

5.1.2 Implementation Techniques

The main idea of the project is to use Machine Learning algorithms to understand the traffic congestion trends and to use Blockchain for Hazard Warning System module.For traffic congestion analysis,preprocessing must be first performed in order to apply the algorithms on that data,otherwise the algorithm will not perform efficiently.

5.2 Conclusion

The analysis is performed much better when the data is sufficient and complete Exploratory Data Analysis is performed.

Chapter 6

System Testing and Evaluation

In this chapter I test my application to give the stakeholder details of the quality of our project. By Testing I will be able to know the degree to which my project obtained its goals.

6.1 Deriving Specification

I used Software Requirement Specification document for developing my test cases. I tested my projects to some limits and checked whether it is working or not. The following Test were done on "Mobility Analytics Application For Intelligent Transportation System using Blockchain,IOT and Machine Learning"

6.2 Testing Environment

The application consists of two modules,Hazard Warning System and Traffic Congestion Analysis and prediction,Following tools were used for testing purpose:

- Visual Studio Code
- Jupyter Notebook
- QGIS
- Postman

6.3 Test Cases

The following test cases have been performed on the system:

6.3.1 Smoke Testing

Smoke Testing is a test that is conducted as a preliminary/prerequisite step for regression testing,in smoke test,the basic critical functionalies of the application are tested,if the smoke test is pass,then the build is sent for regression testing.

Steps	Tasks	Expected Outcome	Actual Outcome
1	GPS data is loaded for analysis	Pass/Fail	Pass
2	Libraries are imported for preprocessing of data	Pass/Fail	Pass
3	Both Express and React(Client) servers are running concurrently	Pass/Fail	Pass
4	Application is loading properly on the browser	Pass/Fail	Pass

Table 6.1: Test Case 1:Smoke Testing

6.3.2 Sanity Testing

Sanity Testing is also a prerequisite test for regression testing,but it is performed when the application has matured and many builds have been developed.

Steps	Tasks	Expected Outcome	Actual Outcome
1	Users can send and receive messages(Transactions)	Pass/Fail	Pass
2	User data (Valueless transaction) is successfully passed to IOTA TANGLE	Pass/Fail	Pass
3	User data (Valueless) transaction is successfully fetched from IOTA TANGLE	Pass/Fail	Pass
4	Data is successfully sent from Arduino Uno to IOTA TANGLE	Pass/Fail	Pass

Table 6.2: Test Case 2:Sanity Testing

6.3.3 Graphical User Interface and Performance Testing

Graphical User Interface testing is a testing technique that is used to test the front end functionalities,interaction between user and the system is one of the key element in order to achieve the best user experience. Performance Testing is used to check whether the system produces the desired output or not,in case the desired output is not produced,then possible solutions are designed to solve that issue.

Steps	Tasks	Expected Outcome	Actual Outcome
1	Front end is responsive	Pass/Fail	Pass
2	System doesnt take alot of time to load	Pass/Fail	Pass
3	System handles the exceptions well	Pass/Fail	Pass
4	Data is encrypted	Pass/Fail	Pass

Table 6.3: Test Case 3:GUI and Performance testing

6.3.4 Login Test Case

Steps	Tasks	Expected Outcome	Actual Outcome
1	System accepts the valid credentials	Pass/Fail	Pass
2	System verify the credential after checking the database	Pass/Fail	Pass
3	User is not logged in when wrong username or password is entered	Pass/Fail	Pass
4	User is logged in when the credentials are valid	Pass/Fail	Pass

Table 6.4: Test Case 4:Login Test Case

6.3.5 Logout Test Case

Steps	Tasks	Expected Outcome	Actual Outcome
1	User is successfully logged in	Pass/Fail	Pass
2	User session is activated	Pass/Fail	Pass
3	User database is maintained	Pass/Fail	Pass
4	User gets logged out when he clicks on logout button	Pass/Fail	Pass

Table 6.5: Test Case 5:Logout Test case

6.3.6 Registration Test Case

Steps	Tasks	Expected Outcome	Actual Outcome
1	Username is unique for every user	Pass/Fail	Pass
2	Registration page is loading correctly	Pass/Fail	Pass
3	User record is stored in the database	Pass/Fail	Pass
4	User is registered into the system upon providing valid credentials	Pass/Fail	Pass

Table 6.6: Test Case 6:Registration Test case

6.3.7 IOTA MAM M2M Transaction Test Case

Steps	Tasks	Expected Outcome	Actual Outcome
1	User is logged in the system	Pass/Fail	Pass
2	User successfully creates a transaction	Pass/Fail	Pass
3	IOTA TANGLE verifies the transaction	Pass/Fail	Pass
4	The message is delievered to the IOTA TANGLE network,and can be verified as well	Pass/Fail	Pass

Table 6.7: Test Case 7:IOTA MAM M2M Transaction test case

6.3.8 Geospatial Data Loading Test Case

Steps	Tasks	Expected Outcome	Actual Outcome
1	Dataset is loaded into the system	Pass/Fail	Pass
2	Dataframe is created successfully	Pass/Fail	Pass
3	Dataset is preprocessed successfully	Pass/Fail	Pass
4	Traffic Congestion is predicted at different times of the day	Pass/Fail	Pass

Table 6.8: Test Case 8:Geospatial Data loading Test case

6.4 Azure Devops Environment

I have used devops environment for documenting the test cases,bugs,epics and features. Azure Devops provide an interactive and efficient way of collaboration for the developers.I can plan test cases,upload my code for version controlling,see bug reports,check which bugs are resolved and which bugs are still not resolved,hence Azure Devops provide us a dashboard sort of tool which is really useful in Software Project Development.There are different packages available to use Azure Devops,I have used Basic version of the Azure Devops for my Final Year Project.

Following are the screenshots showing the work done on Azure Devops

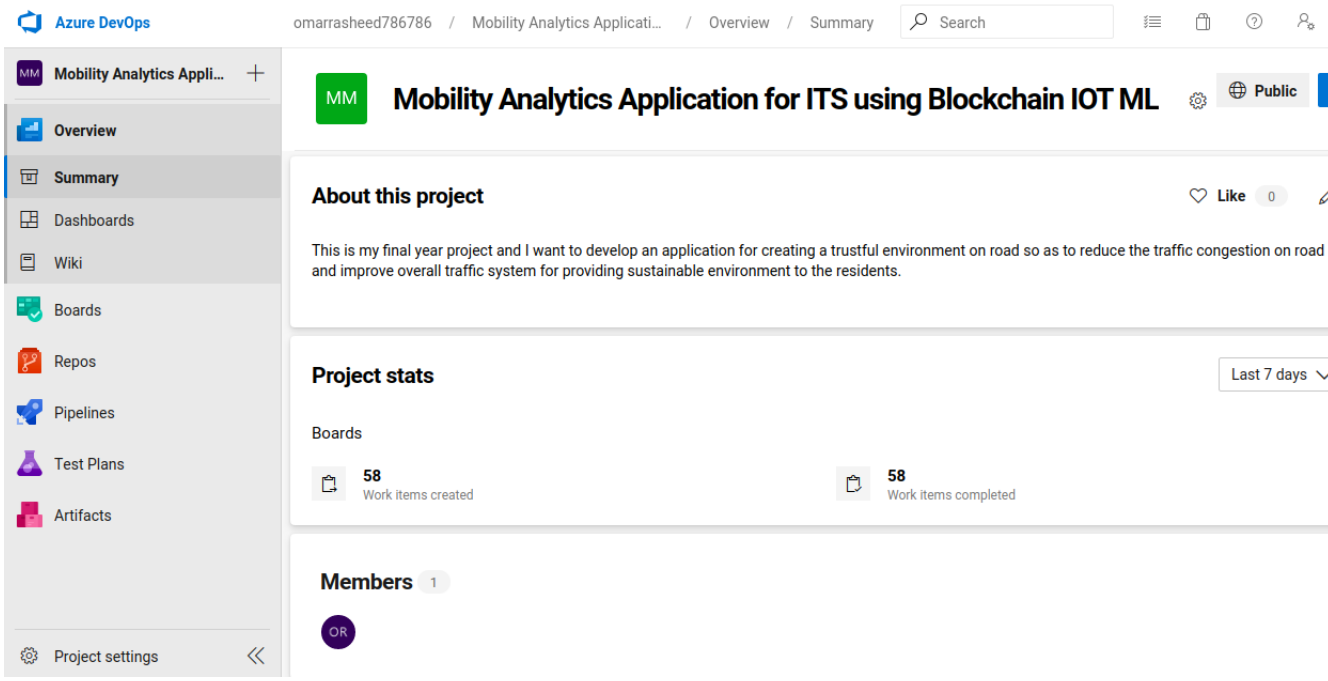


Figure 6.1: Devops Main Dashboard

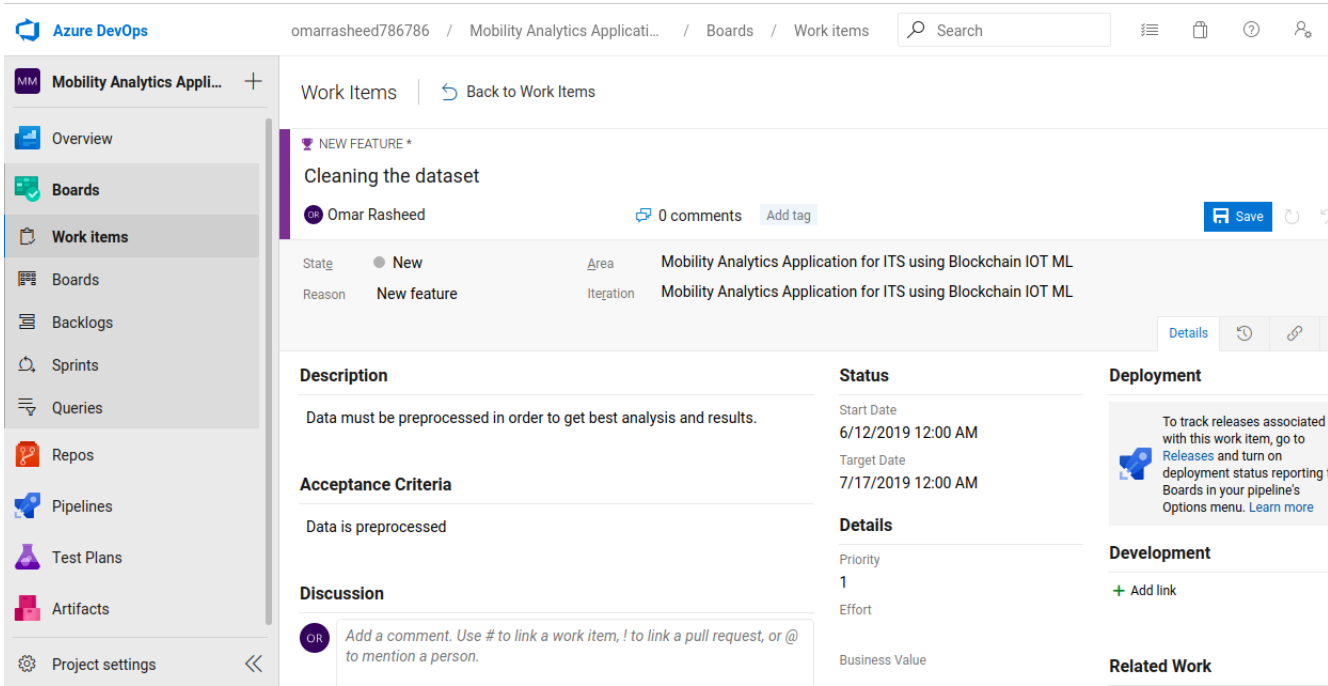


Figure 6.2: Creating work item

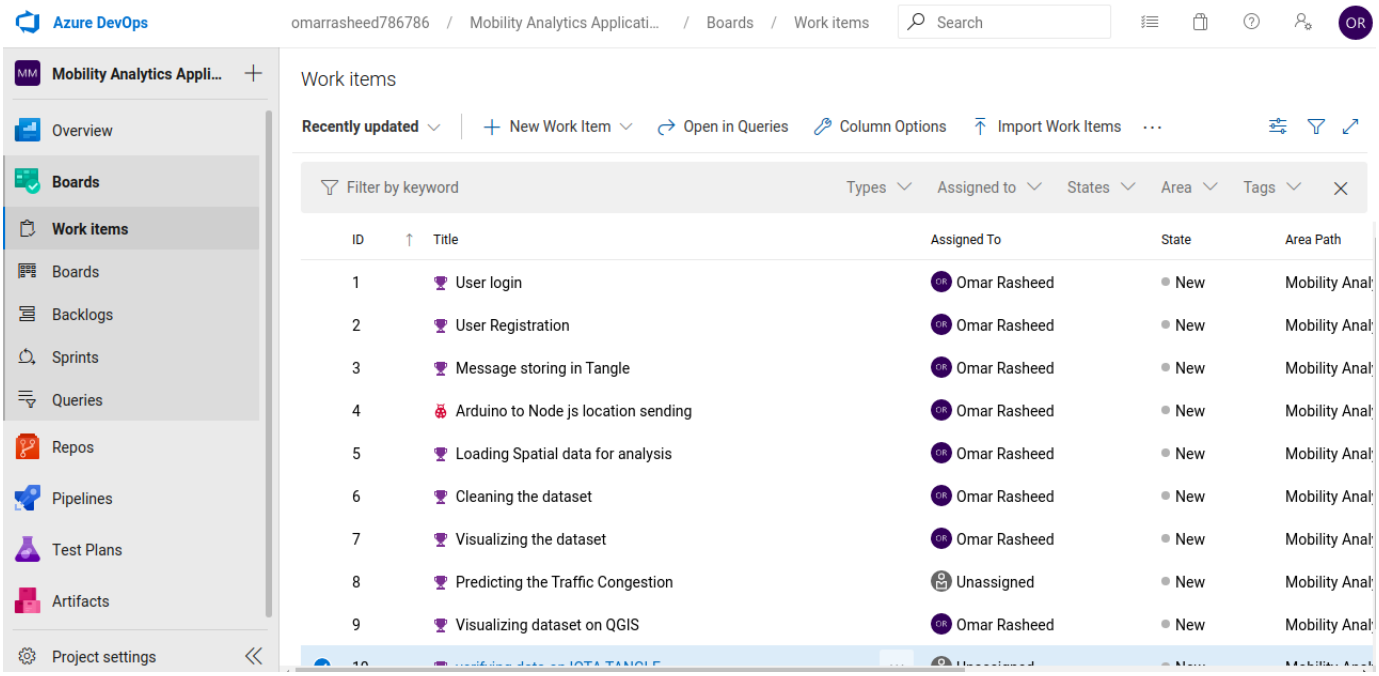


Figure 6.3: Work item dashboard page

It was indeed really nice to work with Azure Devops,it provide us really nice option to control our project and perform the task in efficient way and also help in keeping good coordination among the team members in software development project.

Chapter 7

Conclusions

7.1 Discussion

In this project “Mobility Analytics Application for Intelligent Transportation System Using Blockchain,IOT and Machine Learning” I have created an application which provide an estimate of traffic congestion levels at different times of the day and a trustful environment on roads between the vehicles using IOTA TANGLE Masked Authenticated Messaging so that the user can have best experience on the road free of any miscommunication and any traffic congestion.Traffic Congestion and problems arising from this issue have been witnessed for long,specialy in the recent years,due to Urbanization,this trend has increased a lot,and I strongly believe that with the help of Machine Learning and Blockchain technology we can easily tackle this issue and find out a solution for this problem.My heart cries out whenever I see people getting late due to traffic congestion,the mother nature getting filthy due to dangerous gases released by traffic congestion,patients getting late to get the treatment or the first aid,many die on their way to hospital in traffic these and many other problems caused by traffic congestion are really painful to even think about,and then when I look at the tools I have,technologies that are currently available to me,I strongly and firmly believe that I can play my role to solve this problem using my skill,knowledge and with the help of my teachers,elders and every other stackholder.

7.2 Conclusion

By the Blessing of Allah Almighty,I am very happy that I have completed my Final Year Project and with my application I will be able to help people from all walks of life. This project would be impossible without the help,guidance and Supervision of my supervisor

"Dr. Muhammad Muzammal" as he helped me whenever I wanted any guidance. Some of the main learning benefits from my projects are as follow:

- Improved Knowledge in the field of Machine Learning
- Improved Learning in the field of Datascience
- Improved Knowledge in the field of Blockchain(Distributed Ledger Technology)
- GPS Datasets and visualization,mapping of the GPS data
- Communicating with IOTA TANGLE DAG
- New tools and technologies
- Testing and reviewing
- Documentation

7.3 Limitations

Due to time and other constraints there are some limitations of my projects which are as follows:

- Limited to laptop only and not integrating the system on some car,as I wanted to develop dummy prototype like autonomous vehicle
- Limited to website supported devices
- Not utilizing all the dataset of 30 days because of hardware limitations of my laptop.

7.4 Future Work

This is a demo project of what I am trying to accomplish and would require a good deal of future work.Trajectory Mining has really fascinated me,I really find it interesting in doing research in the field of trajectory mining,yes it has its own challenges but I believe that if the intentions are clear and will is strong,every obstacle,issue can be resolved,and traffic issues are no strange to this rule,they can be also resolved if proper planning is done to solve them.Few of the main improvements that can be made to this project are as follow:

- Support Multiple Platforms
- Work on live input data
- Give direct real time response as MAM subscriber

- Remove as much Limitations as possible
- Get Spatial Dataset of any Pakistan City and apply the same technique to provide solution for traffic congestion
- Apply other Unsupervised Learning algorithms like DBSCAN and compare the results.
- Integrate the hazard warning system to android application.
- Introduce incentives option as well for honest publishers of data
- In this project,due to time constraint,I have applied Kmean Unsupervised Learning algorithm to analyse the traffic congestion cluster,but I will continue this project and apply other Unsupervised Learning algorithms like Genetic Algorithm,DBSCAN(Density-based spatial clustering of applications with noise) etc to have comparative analysis and to check which algorithm works best in analyzing and predicting traffic congestion.

Appendix A

User Manual

For feasibility analysis of Hazard Warning System,I have conducted an offline and online survey.The sample size is 50 as I recieved 50 responces,and most of the feedback showed the dire and immediate need of such an application that can help create a trustful environment on the road and help reduce traffic congestion and help the peer drivers, autonomous vehicles in future as much as possible as many respondents answers reflected their lack of confidence in ongoing traffic management systems and they want a change,the data analysis reflects.

References

[1] Peter J. Denning. Is computer science science? *Commun. ACM*, 48(4):27–31, April 2005. No Citations.