**Understanding Real Time Traffic Characteristics of Urban Zones Using GPS Data: A Computational Study on Dhaka City**

# Abstract

Dhaka, the capital of Bangladesh, is badly affected by traffic congestion and the situation is deteriorating day by day. In order to minimize traffic congestion, we need to investigate traffic pattern of this mega city and identify the factors that are responsible for high traffic intensity. However, only a few researches have been conducted to find out the reasons behind Dhaka city’s acute traffic congestion. Most of those researches were survey based, thus, prone to perception and human intervention. In this research, we have used Global Positioning System (GPS) data in order to analyze Dhaka city’s traffic pattern. We have considered 13 DPZ (Detailed Planning Zone) zones proposed by RAJUK, the city development authority of Bangladesh for traffic modeling. Computational methods are employed to find the similarities among clusters. We consider both intra zonal road segments in a zone and inter zonal road segments that connect two or more different zones. Land use pattern is also investigated to find out the traffic variability among zones. A graph is produced to visualize overall traffic scenario of Dhaka city, where the traffic intensity is measured by GPS dataset.

*Keywords*—Traffic pattern; Road segments; Global Positioning System (GPS); Traffic congestion; DPZ zones.

Introduction

Dhaka is one of the fastest growing cities in the world with an approximate population of 17 million [ (Placeholder1)1] and about 300,000 to 400,000 migrants coming into the city every year. Projected population of Dhaka city in 2020 is 20 million making it one of the most populated mega cities in the world [2]. Dhaka city is the center of the Bangladesh’s economic growth. One of the major and most alarming problems of Dhaka city is its traffic congestion. Traffic congestion costs 5 million working hour daily [3]. The approximate yearly loss due to traffic congestion is about 4.46 billion Dollar [4]. If proper steps are not taken, the average vehicle speed will decrease to less than 5 Kilometers per hour in the following years [5].

Japan International Cooperation Agency (JICA) [6] con- ducted a survey in order to address the existing problems. The main goal was to provide an overview of traffic pattern of Dhaka city and propose future planning for city development. In Dhaka Structural Plan (2016-2035) [7], published by Rajd- hani Unnayan Kartripakkha (RAJUK), the Dhaka Metropolitan Region (DMR) was divided into six zones and further de-

velopment was planned according to this zones. RAJUK has published the future planning of transportation based on these zones. Each zone has its own land use pattern, thus, those are considered differently in order to make improvement over overall traffic condition.

But the problem is that all the reports were mainly based on observational field study and interview, therefore, prone to human error. Moreover, such study is expensive and almost not feasible to carry out in regular basis to monitor traffic pattern. For a large scale research, some form of automatic data collection process is required. We have used Global Position- ing System (GPS) record obtained from GoBD which is a company that provides real time route suggestion. The dataset contains the traffic intensity information in a period of 15 days (from 1st September, 2015 to 15th September, 2015). The intensity level is measured in a scale from 0 to 1 where

1 being the highest traffic intensity.

In our previous works [8], [9] based on this GPS data and other secondary dataset, we have found out the impact of traffic congestion of one point on its neighbor points which follows the Gravity Model [8]. We also found out traffic pattern for different cases (e.g. overall, day-wise) based on dif- ferent factors (e.g. road segment intersections, marketplaces, rickshaws, land use pattern, public transportation) [9].

In this paper we have analyzed traffic pattern at both macroscopic and microscopic level. For this, we have used data from RAJUK which is the authoritative corporation for the development of Dhaka city. RAJUK divided the whole Dhaka city into 13 different zones. We use the land use information of those DPZ zones for macroscopic analysis. For microscopic modelling, We have determined both inter and intra road segments’1 of each zone and also found out the traffic pattern. Our contribution in this paper can be summarized as followed:

Generating intra and inter road segments among zones.

Intra and inter road traffic pattern for 13 different zones.

Finding the effect of road segments ratio and land use pattern on traffic congestion.

Determining factors that are responsible for traffic inten- sity pattern variation among same type of zones.

1In the remaining part of the paper, we will call intra zonal road segment as intra road and inter zonal road segment as intra road.

978-1-5386-5163-6/18/$31.00 ©2018 IEEE

Generating a graph that represents the overall traffic scenario of Dhaka city.

In section II, we have described related works on traffic modeling. Section III describes the form of our obtained data and the pre-processing method. In section IV, we have discussed about the methods of our works. We have concluded our paper in section V with a summary of our findings.

Related Work

A number of works have been done on traffic modeling to reduce congestion. In order to provide an accurate and precise result, some form of automatic methods need to be implemented.

Zhang et al. [10] demonstrated that using taxi GPS data for estimating travel time and speed conditions is practical by performing an error analysis of taking simple averages over historical data. The calculated intensity level of the traffic from given dataset is also based on vehicle GPS data. Obuhuma, J et al [11] proposed a real time system that provides the ability for monitoring and analyzing traffic scenarios. They emphasized on the necessity of real time traffic pattern analysis and use of high-tech computerized solutions. They have used GPS data collected by GPS server system. The server was specially developed in order to complete the study. Castro el al [12] proposed a method to construct density model. They have used large scale taxi traces. Their proposed model can also be used for the prediction of future traffic congestion.

Watling [13] evaluate the various models with regard to their suitability and compared various traffic models like TRAFFICQ [14] which has been designed for studying a detailed section of road network, DYNEMO [15] which is a mixed microscopic and macroscopic model.

Liu et al [16] analyzed GPS data from more than 6600 taxis. They used traffic pattern to describe intra urban land use variations. In our dataset, traffic intensity was measured using the GPS data from number of taxi cabs. We analyzed this dataset to determine different factors behind traffic congestion. Giuliano et al [17] have given an overview of the relation- ships between transportation and land use. It examines the impact of land use on the transportation. They have given a brief description of factors to consider while examining land use pattern and the impacts of it on transportation investments. The JICA study team and Dhaka Transportation Coordina- tion Board have been working together in order to formulate the basic concept of urban development for Dhaka Metropoli- tan Area in 2025 [6]. They have divided the study area into six zones and conducted household interviews, Cordon Line and Screen Line Survey. The study also investigated the traffic pattern and bus speed for morning peak, office hour and afternoon off-peak hour for different routes. From that, we can find out the routes which are more congested considering the bus speed. The study helps us to understand the current condition of traffic systems. But the main problem is that, this most of the analysis was completed manually. So the system

will be prone to human error and also expensive.

In our previous work [9], we analyzed provided data and found out traffic intensity pattern over hours of the day and days of the week. We clustered the DPZ zones based on their land use pattern. We have determined that zones in same cluster do not have same traffic pattern. Rickshaw which is a three wheeler slow moving vehicle and a popular mode of transportation in Bangladesh, is responsible for enhancing traffic congestion. Market places also have effect on traffic congestion. In our research, we have found out that public transportation has a negative relation with traffic congestion. We have also calculated the intersection density based on the latitude longitude information of the road segments and found out that traffic intensity increases with the number of road intersections. We also searched for the factors (e.g. land use pattern, road-intersections, bus route) that may cause conges- tion. We also did a regression analysis based on percentage of road network area, percentage of mixed land area, number of intersections/KM and number of bus routes.

We have also examined traffic expansion on a corridor using gravity model. From that analysis, we found out that, as the distance increases, the intensity impact of a point on neighboring points decreases [8]. For better city planning, land use pattern analysis is very important. The relationship between land use pattern and traffic intensity can help more efficient city planning. In this research, we have made some connections between these two.

Data Description

In this section, we have described our data, pre-processing method and secondary dataset. We have obtained our dataset from GoBD.co. It is a private company which is currently developing a route suggestion system. They use some propri- etary algorithms and use the GPS traces for calculating the traffic intensity. The data is from 1st September, 2015 to 15th September, 2015. The data has the traffic intensity information which is measured based on the road class, altitude of the GPS data using OSM and speed. A single road is divided into multiple road segments and we have total 11,769 road segments traffic data. The total number of record is 1,623,280. Table I describes the type and the attributes of our dataset.

TABLE I: Description of the variables of the Data.

|  |  |  |
| --- | --- | --- |
| Attribute | Representation | Type |

|  |  |  |
| --- | --- | --- |
| Identifier | Unique number for each record. | Object |

|  |  |  |
| --- | --- | --- |
| TimeStamp | Precious moment when the record was  created. | DateTime |

|  |  |  |
| --- | --- | --- |
| OsmID | Open Street Map identifier for the road  segment. | Integer |

|  |  |  |
| --- | --- | --- |
| Congestion | Traffic intensity which is in 0 to 1  range. | Float |

|  |  |  |
| --- | --- | --- |
| RoadLength | Road segment length in kilometer. | Float |

|  |  |  |
| --- | --- | --- |
| Coordinate | Latitude and longitude points of road  segment. | Lat-Long |

|  |  |  |
| --- | --- | --- |
| Utility | The cost of the road segment as a route. | Float |

*Data Pre-processing*

Our first step was to remove insufficient data from the dataset. In order to do that, we have divided our data into 30 minutes segment starting from 00:00:00. The whole day is divided into 48 segments. If a road segment has less than 5 records in 30 minutes slot from 1st September, 2015 to 15th September, 2015, we have removed them from the dataset. After cleaning the data we have 8,966 road segments traffic intensity information.

Secondly, we have found and removed outlier. We consider a record as an outlier if the traffic intensity is not within the 2-standard deviation from the arithmetic mean for that road segment in that particular time window (30 minutes).

We have plotted our data on map using QGIS and Google map. We have generated a KML file using all the latitude and longitude information of road segments from our dataset. We have drawn a polygon on Google map and converted it to KML file from map options. Then we load both of our KML files in QGIS as vector layer and export them as a shape file. Figure 1 shows road segments in the map which we have produced using QGIS and our GPS dataset.

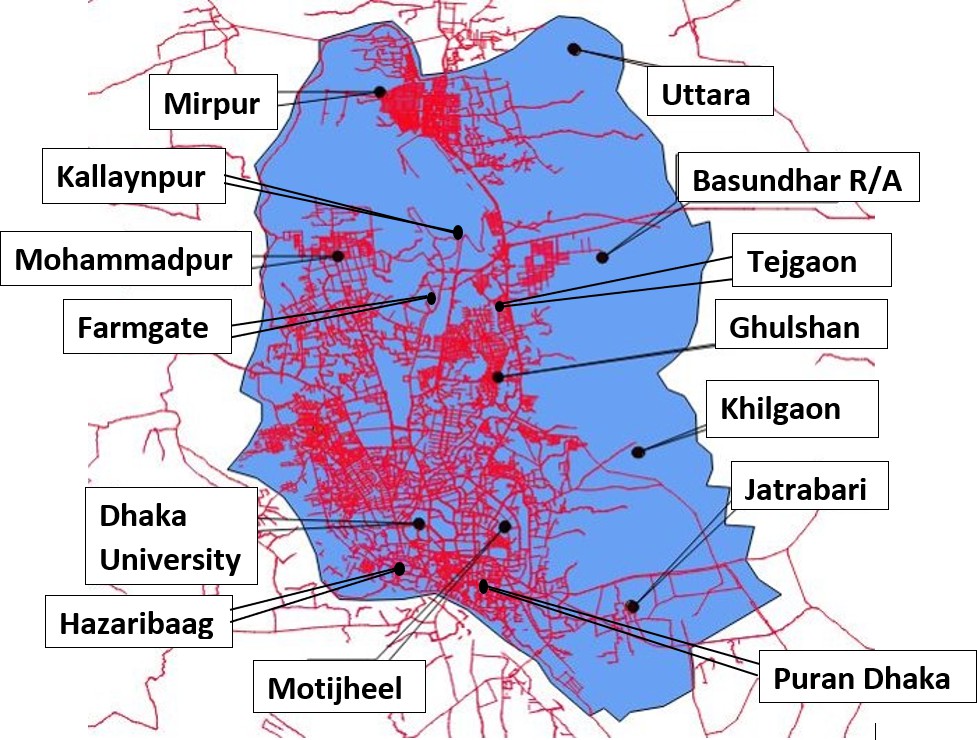


Fig. 1: Road segments and study area (blue colored area). There are total 11,769 road segments in our dataset. After preprocessing step we considered 8,966 road segments.

Figure 1 shows our study area and road segments. The total land area is 310 square kilometers.

*Secondary Dataset*

RAJUK have published a report [18] containing urban planning details of Dhaka city. The report has identified sub areas as DPZ, which contains detail land use proposals including social infrastructures information. We use that land use information for further traffic pattern modeling and analyze the modeling macroscopically. RAJUK also published the information about wards of DCC (Dhaka City Corporation) for each zone. They have divided the wards among different zones according to land use and homogeneity of the wards. Table II shows the information about all the zones and corresponding DCC (Dhaka City Corporation) wards.

TABLE II: DPZ zones with corresponding DCC wards, type and name.

|  |  |  |  |
| --- | --- | --- | --- |
| Zone | Dhaka City Corpo-  ration Ward | Zone Type | Informal Name |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-1 | 59, 60, 61, 62, 63, 64,  65, 66, 67, 68, 69, 70,  71, 72, 73 | Mixed | Puran Dhaka |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-2 | 74, 75, 76, 77,78, 79,  80, 81, 82, 83, 90 | Mixed | Jatrabari |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-3 | 56, 57 | Mixed | Dhaka University |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-4 | 32, 33, 36, 53, 54 | Commercial | Motijheel |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-5 | 22, 23, 24, 25, 27, 28,  29, 30, 31, 34, 35, 55,  84, 85, 86, 87, 88, 89 | Residential | Matuail |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-6 | 37 | Industrial | Tejgaon |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-7 | 38, 39, 40, 49, 50, 51 | Commercial | Farmgate |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-8 | 47, 48, 52, 58 | Suburb | Hagaribag |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-9 | 42, 43, 44, 45, 46 | Suburb | Mohammadpur |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-10 | 9, 10, 11, 12, 13, 14,  16, 41 | Suburb | Kallaynpur |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-11 | 19, 20 | Suburb | Gulshan |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-12 | 2, 3, 4, 5, 15 | Residential | West Mirpur |

|  |  |  |  |
| --- | --- | --- | --- |
| DPZ-13 | 6, 7, 8 | Residential | East Mirpur |

Traffic Analysis Methods

In this section we have described the method of our work. We have distributed the road segments into two categories: intra road segments and inter road segments.

*Distributing road segments among zones*

As we have plotted the DPZ zones into map, we consider each DPZ zone as a polygon. We have the latitude longitude points of each DPZ zones and latitude longitude points of each road segments. We checked if a road segment’s latitude and longitude points are distributed among multiple zones. If yes, the road segment is labeled as inter road segment, otherwise it is considered as intra road segment.

We have divided the labeled road segment DPZ wise. So now we have the information of each DPZ zone with the number of intra and inter road segments of that zone. Figure 2 shows the definition of intra and inter road segments.

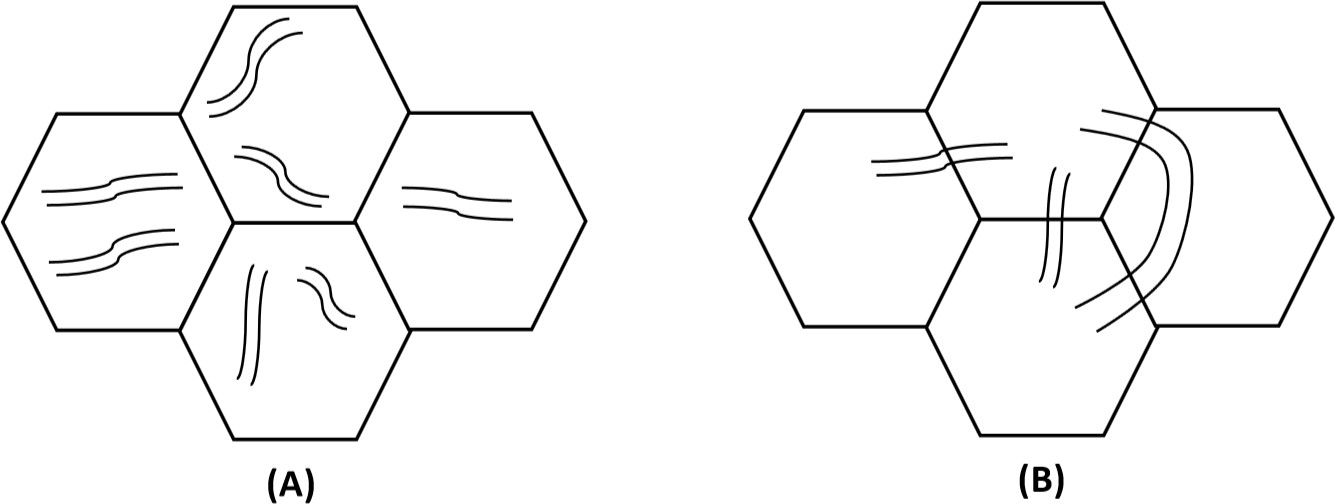


Fig. 2: Visual definition of intra (A) and inter (B) road segments. The polygon represents a zone where as the curve inside the polygon represents road segments.

Here, a road segment is considered as inter road segment if more than one zones has the coordinates of that particular road segment.

*Traffic intensity pattern between inter and intra road seg- ments*

As we have found out inter and intra road segments, we check their traffic pattern. In order to do that, we have put all the intra road segments in one class and inter road segments in another class. Then we have calculated average traffic intensity of these classes from 15 days data. Figure 3 shows the comparison of average traffic intensity between intra and inter road segments.

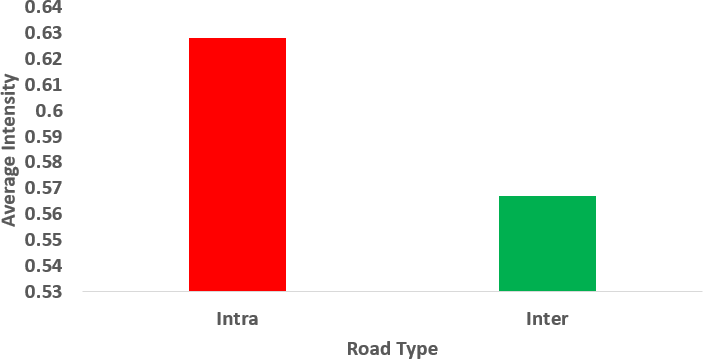


Fig. 3: Average intensity of intra road segments (red bar) and inter road segments (green bar). The intensity is calculated between 0 and 1 with 1 being the most intense.

Figure 3 represents that, intra road segments have higher traffic intensity than inter road segments.

We have also found out the connectivity between zones. We generate a graph consisting of nodes and edges. Here, the number on each edge is the average traffic intensity of the road segments that connect the two zones. Number inside a circle represents the average traffic intensity of the intra road segments of that zone. Figure 4 contains the graph that describes the traffic scenario among different zones.

In figure 4, value inside the vertex is this zone’s intra road segments intensity and value on the edge is the intensity of the inter road segments between connected zones. Bigger circle means more intensity on intra road segments of that zone. We have determined that the average intensity is more than 0.60 in 9 out of 13 zones.

*Finding Zone Cluster from Intra and Inter Road Segments*

As mentioned earlier, we have calculated the number of intra and inter road segments in zones. We calculate the ratio of intra and inter road segments using following formula.

*Numberof IntraRoadSegments*

*Ratio* =

*Numberof InterRoadSegments*

With the found ratio, we have used agglomerative clustering using Euclidean distance between two objects as distance metric. Figure 5 shows the dendrogram.

We have considered DPZ-7 (Farmgate) and DPZ-9 (Mo- hammadpur) zones as outlier as their road segment ratio is

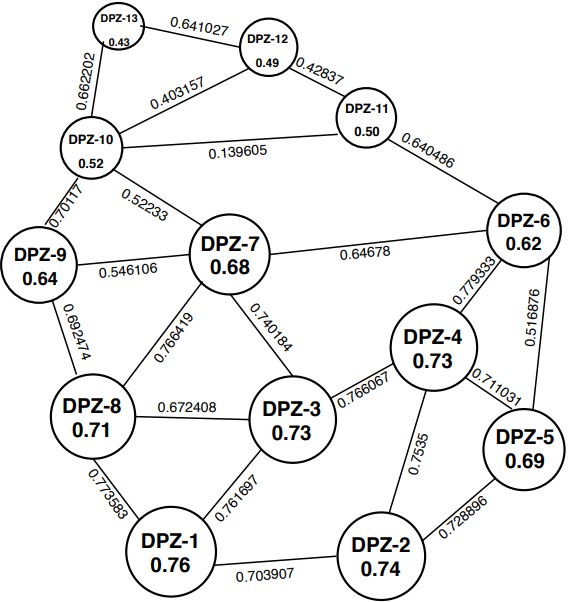


Fig. 4: Graphical representation of connected DPZ zones. Size of the vertex is proportional to its intra road segments average intensity.

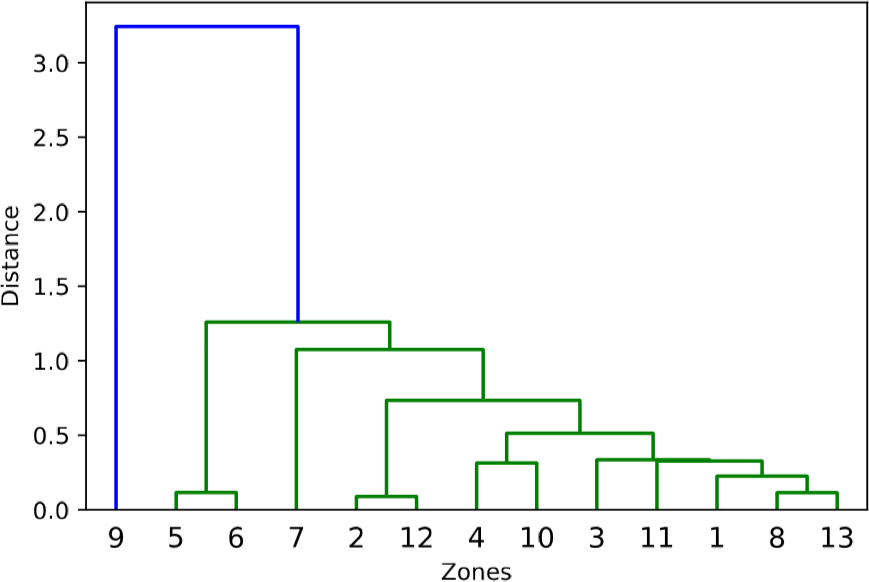


Fig. 5: Dendrogram of clusters based on ratio between intra and inter road segments.

different from all other zones. Excluding the outliers, we have considered 5 clusters. Table III shows our clustering result.

Each member of a cluster has similar ratio of intra and inter road segments. As the ratio is almost identical, the traffic intensity should be similar. But in reality, they vary in wide range. In our previous work, we clustered our data based on land use pattern. Despite having similar ratio, different land use pattern causes different traffic intensity in the same cluster.

*Traffic Variability in Intra and Inter Road Segments*

In order to find out inter road segments influence in overall traffic intensity of a zone, we filtered our data based on inter road segments traffic intensity. For this, we considered zones with inter intensity higher than 0.65. This left us with DPZ-1 (Puran Dhaka), 2 (Jatrabari), 3 (Dhaka University), 4

(Motijheel), 5 (Matuail) and 8 (Hazaribag).

TABLE III: Definition of different cluster based on intra and inter road segments ratio. Zones in the same cluster are sorted based on their ratio (lower to higher)

|  |  |  |
| --- | --- | --- |
| Cluster No | Zones in the Cluster | Ratio Range |

|  |  |  |
| --- | --- | --- |
| 1 | 6, 5 | 1.0 *<* ratio *<*=2.0 |

|  |  |  |
| --- | --- | --- |
| 2 | 3, 1 | 2.0 *<* ratio *<*=3.0 |

|  |  |  |
| --- | --- | --- |
| 3 | 13, 8, 11 | 3.0 *<* ratio *<*= 4.0 |

|  |  |  |
| --- | --- | --- |
| 4 | 4, 10 | 4.0 *<* ratio *<*= 5.0 |

|  |  |  |
| --- | --- | --- |
| 5 | 2, 12 | 5.0 *<* ratio *<*= 6.0 |

Then we analyzed these zones’ connected zones. We have found out that, each of this zone has at least one commercial area connected to it. So the mostly commercial zones are causing extra inter intensity.

As the ratio of intra vs. inter road segments increases, in most cases, the average intensity also increases except in cases of suburb and residential area. This tells us that, these suburb and residential area is less dependent on inter traffic movement for basic needs.

We have also calculated numerical difference between over- all average intensity of a zone and intra and inter intensity of a zone using the following formulas.

*OverallIntensity*(*avg*) *− InterIntensity*(*avg*)

As shown in the figure 6, intra road segments traffic intensity has less difference with overall intensity. The overall traffic intensity of a zone is mostly affected by intra road segments traffic intensity.

Conclusion

Accurate traffic intensity pattern analysis is very important in order to help proper city planning. This can also help to generate smart route suggestion which can eventually save time in transportation. We have analyzed traffic intensity pattern both macroscopically and microscopically. We have found that, intra road segments traffic intensity is much more than inter road segments traffic intensity. We have described the effect of commercial zones influence over neighboring zones. We have also found that, residential and suburb areas have less inter zones traffic intensity. Data shows us that, intra road segments traffic intensity is more responsible for overall traffic intensity.

Further research should focus more on microscopic analysis. Using this work and previous work [8], we can find the socio- economic impact of one zone on other connected zones. We can also compare the level of impact of a zone on its connected zones. Width of individual road can give us more insight about intensity on a particular road segment. Moving vehicles direction can help to find more accurate traffic pattern. Also, data about the condition of a road can be used as an additional

*Dinter* =

*Dintra* =

*OverallIntensity*(*avg*)

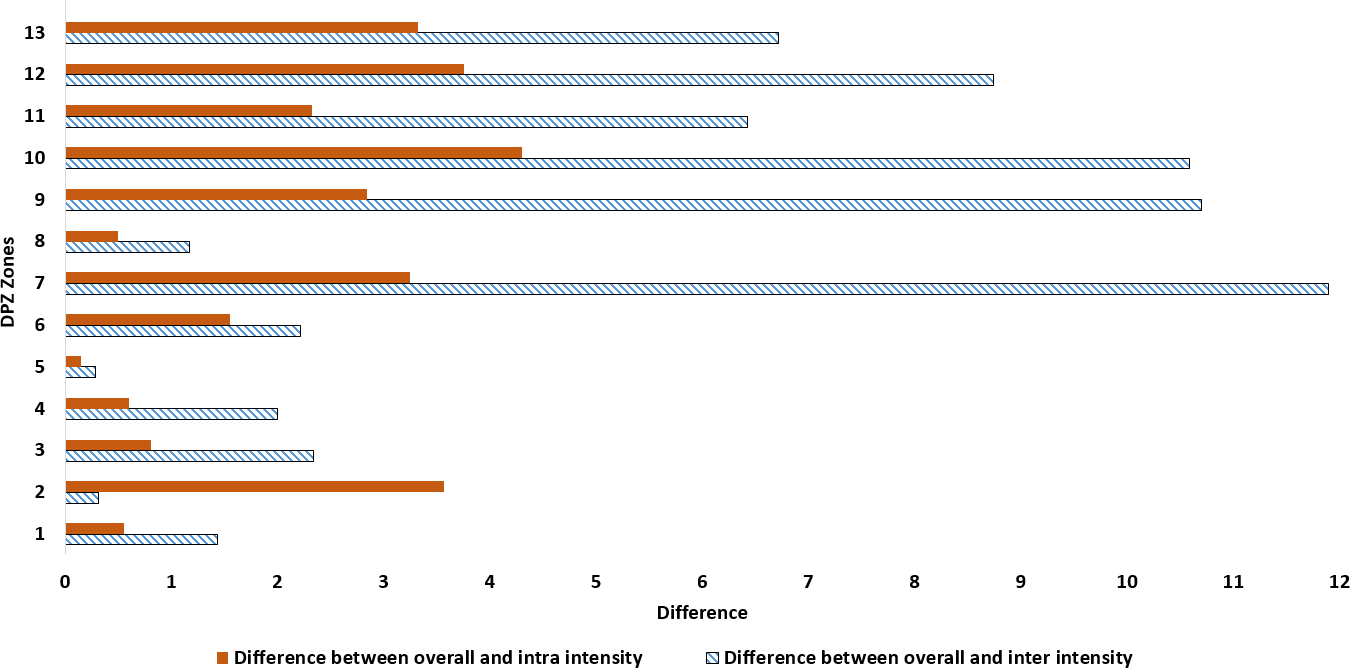
*OverallIntensity*(*avg*) *− IntraIntensity*(*avg*) *OverallIntensity*(*avg*)

parameter to analyze traffic intensity pattern. Finally, we have only 15 days GPS data. More data can be helpful for more accurate analysis. This can eventually provide a smart route suggestion system.

We plotted the result in a graph for each zone. For scaling, we have multiplied for result by 100. Figure 6 shows the result.

Acknowledgement

The authors would like to thank Gobd for providing GPS traffic dataset and LIRNEasia, Sri Lanka for providing logis- tics support.

Fig. 6: Comparative graph of difference between inter (tex- tured bar) and intra (solid bar) road segments average intensity with overall average intensity.

References

Asian Development Bank (2011), “Preparing the greater Dhaka sustain- able urban transport corridor project”, consultants report.

Roads and Highways Department (RHD) for the Peoples Republic of Bangladesh and the Asian Development Bank, “Greater Dhaka Sustain- able Urban Transport Project Prepared”, December 2015.

“Colossal loss” The Daily Star (Newspaper), March 25, 2018. Available online at [https://www.thedailystar.net/frontpage/colossal-loss-1553002,](http://www.thedailystar.net/frontpage/colossal-loss-1553002) Accessed on : 27th March, 2018.

Md. Saidur Rahman, “Dhaka’s traffic jam causes Tk 370b annual loss” The Daily Prothom Alo (Newspaper), March 27, 2018. Avail- able online at <http://en.prothomalo.com/bangladesh/news/173182/Dhaka-> s-traffic-jam-causes-Tk-370b-annual-loss, Accessed on : 4th April, 2018.

Dr. Bjorn Lomborg, The smartest ways to deal with traffic congestion in dhaka.“The Daily Star(Newspaper), May 09, 2016.

Japan International Cooperation Agency, “Preparatory Survey Report on Dhaka Urban Transport Network Development Study (DHUTS) in bangladesh. 2010.”

RAJUK (Rajdhani Unnayan Kartripakkha). “Draft Dhaka Structure Plan Report 2016-2035.”

Sayed, M. A., Rahman, M. M., Zaber, M. I., & Ali, A. A. “Understanding Dhaka City Traffic Intensity and Traffic Expansion Using Gravity Model”, 20th International Conference of Computer and Information Technology (ICCIT), 22-24 December 2017, Dhaka, Bangladesh.

Rahman, M. M., Shuvo, M.M.M., Zaber, M. I., & Ali, A. A. “Traffic Pat- tern Analysis from GPS Data: A Case Study of Dhaka City.” 2018 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), 16th-17th March, 2018, Bangalore, India.

ZHANG, W., XU, J., AND WANG, H. 2007. “Urban traffic situation calculation methods based on probe vehicle data.” J. Transport. Syst. Eng. Inform. Technol. 7, 1, 4349.

Obuhuma, J. I., Moturi C.A. , “Use of GPS With Road Mapping For Traffic Analysis”, International Journal Of Scientific Technology Research, 1(10), November 2012.

Castro, Pablo, Daqing Zhang, and Shijian Li. “Urban traffic modelling and prediction using large scale taxi GPS traces.” Pervasive Computing (2012): 57-72.

Watling, David. “A review of models of urban traffic networks (with particular reference to the requirements for modelling dynamic route guidance systems).” (1991).

Logie, M., “TRAFFICQ: A Comprehensive Model for Traffic Manage- ment Schemes”, Traffic Engineering and Control 20 (111, 516-518 (1979).

Leutzbach, W. Wiedemann, R., “Development and applications of traffic simulation models at the Karlsruhe Institut fu¨r Verkehrswesen”, Traffic Engineering & Control 22 (5), 270-278 (1986).

Liu, Yu, et al. “Urban land uses and traffic source-sink areas: Evidence from GPS-enabled taxi data in Shanghai.” Landscape and Urban Planning

106.1 (2012): 73-87.

Giuliano, Genevieve. “Land use impacts of transportation investments.” The geography of urban transportation 3 (2004): 237-273.

RAJUK (Rajdhani Unnayan Kartripakkha) , “Preparation of detailed area plan (dap) for dmdp area: Group-c, rajuk.”