

# Study on the Effect of Road Infrastructure, Socio-Economic and Demographic Features on Road Crashes in Bangladesh

Shakil M. Rifaat, Md. H. Rahman, Mohammed, Mosabbir Pasha

**Abstract**—Road crashes not only claim lives and inflict injuries but also create economic burden to the society due to loss of productivity. The problem of deaths and injuries as a result of road traffic crashes is now acknowledged to be a global phenomenon with authorities in virtually all countries of the world concerned about the growth in the number of people killed and seriously injured on their roads. However, the road crash scenario of a developing country like Bangladesh is much worse comparing with this of developed countries. For developing proper countermeasures it is necessary to identify the factors affecting crash occurrences. The objectives of the study is to examine the effect of district wise road infrastructure, socioeconomic and demographic features on crash occurrence. The unit of analysis will be taken as individual district which has not been explored much in the past. Reported crash data obtained from Bangladesh Road Transport Authority (BRTA) from the year 2004 to 2010 are utilized to develop negative binomial model. The model result will reveal the effect of road length (both paved and unpaved), road infrastructure and several socio economic characteristics on district level crash frequency in Bangladesh.

**Keywords**—Demographic, Negative Binomial Model, Road Infrastructure, Socio-economic, Traffic Safety.

## I. INTRODUCTION

AROUND the world, an estimated 1.2 million people are killed in road crashes each year and as many as 50 million are injured. Moreover, motor vehicle collision injuries will move from ninth place in 1990 to third place by 2020 as the leading causes of disease burden [1]. Injuries and mortalities from road crashes are problems for both developed and developing countries. For example, recent Canadian crash data in 2009 indicates that there were 2,209 fatalities and 11,451 serious injuries occurred in that particular year [2]. However, the road crash scenario of developing country is much worse comparing with this of developed countries. Over 80% of traffic fatalities occur in so called developing and emerging countries, even though these countries account only about one third of the total motor vehicle fleet. Accident rates in developing countries are often 10-70 times higher than in developed countries [3].

Shakil. M. Rifaat is with Islamic University of Technology, Board Bazar, Gazipur- 1704, Bangladesh (Phone: +8801928029882; Fax: +88029291260; e-mail: smrifaat@iut-dhaka.edu).

Md. H. Rahman, Mohammed, and Mosabbir Pasha are with Islamic University of Technology, Board Bazar, Gazipur- 1704, Bangladesh (e-mail: hisham\_xndc@yahoo.com, mohammediut@gmail.com, mosabbir@iut-dhaka.edu).

As a developing country, Bangladesh is not out of this context. Fatality rate in Bangladesh is almost 25 times higher comparing with that of most of the developed countries, 8 times higher than that of Thailand and 3 times higher than that of India [4]. Each year nearly 5000 accidents as reported by police occur in Bangladesh [5]. About 60% of the accident occurs on national and regional highways and 40% on urban roads [6].

One way of improving road safety is to reduce crash occurrences by implementing crash reduction countermeasures. Another way is to reduce the severity of crashes with safer vehicle and road user behavior modification programs. However, these two methods can only be successfully applied if the relevant factors those contribute toward the occurrence or increased the severity of crashes are known. The injury risks of individuals in traffic crashes are influenced by a multitude of factors, including vehicle features, roadway design and operation, driver characteristics, type of collisions and environmental conditions. In addition to identify the significant factors, it is essential to quantify the relative magnitudes of the impact of these factors on collision frequency and severity so that countermeasures to prevent collisions can be prioritized and implemented.

While identifying factors affecting crashes or related severities, several studies have concentrated on the collisions at an intersection or a particular road segment. In these studies, a wide range of factors were examined, including driver attributes, vehicles features, environmental condition and crash characteristics [7], [8].

However, when crashes are grouped in a larger spatial scale such as county, ward or state rather than on a particular road segment or at an intersection, many factors related to land-use policy, road infrastructure and socio-economic characteristics can be examined. Recently, a number of studies have been done where these factors are identified as playing important roles on road related crashes [9], [10]. These macro level studies provide valuable insights to transportation planners on how crashes are influenced by spatial factors.

The key aims of our research are: to identify the magnitude of traffic accidents and casualties and to examine the effect of district wise road infrastructure, land use & demographic features on crash occurrence. Previously several micro level studies have been carried out at some specific zones in Bangladesh, but macro level studies have been observed very rarely, if any [11], [12]. In our study, district level data has

been taken as the unit of this macro level analysis. It is to be noted that, Bangladesh is divided into 64 districts under 7 divisions. Fig. 1 shows the map of 64 districts in Bangladesh.

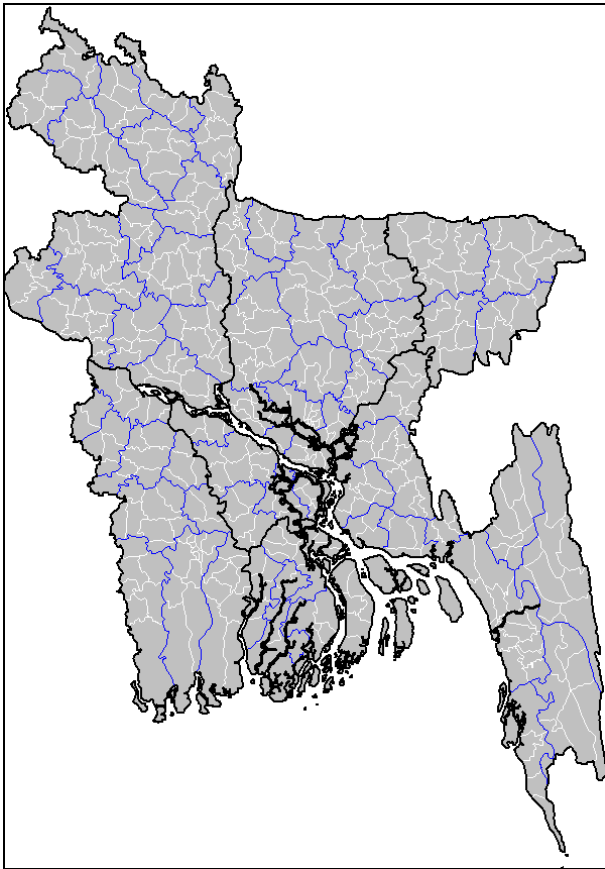


Fig. 1 Sixty four districts in Bangladesh

## II. METHODOLOGY AND DATA

To achieve the objectives of the study, suitable statistical model needs to be selected. The model will be developed using actual accident data that relates crashes with road infrastructure, socioeconomic and demographic factors. The methodology is divided into several steps:

1. Statistical model selection to correlate number of crashes as a function of different factors.
2. Collection and processing of data to establish the model.
3. Analysis and interpretation of model result.

Several sources of data were used in this study. The organizations considered for data collections are: Roads & Highways Department (RHD), Bangladesh Statistical bureau (BSB), Bangladesh Road Transport Authority (BRTA) and Local Government Engineering Department (LGED). Bangladesh Road Transport Authority (BRTA) provided road traffic accident reports for 64 districts in the country. The reported crashes for the years 2004-2010 are used in this study. During the period analyzed, 21448 crashes occurred in

the 64 districts in Bangladesh, giving an average of 48 crashes per district each year. To understand the accident trend in Bangladesh, the crash statistics for the year 2004-2010 is shown in Fig. 2.

Demographic and socioeconomic characteristics for each district were collected from the census data published by Bangladesh Statistical Bureau (BSB). Information such as average dwelling household size, household living structure, household economic condition, number of population, birth & death by sex, literacy rate, economically active population by labor force participation rates, etc. are extracted from the census database. Several district wise socioeconomic data were found missing from the census data such as: education level, income level, age distribution etc.

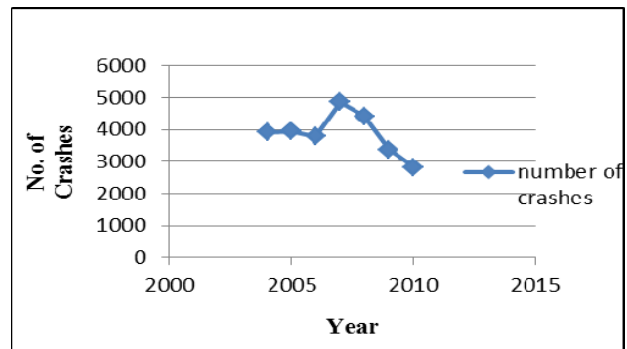


Fig. 2 Year wise crash occurrence in Bangladesh

District wise number and span of structure such as bridge, culvert etc. were collected from the Local Government Engineering Department (LGED). From Roads and Highways Department (RHD), the total length of road in each district, lengths of paved & unpaved road etc. were collected. From all these information variables are formed for model development. The variables are either continuous or categorical in nature. For example, the length of paved road in each district is categorized into 0-40km, 41- 80km, 81- 120km and above 120km. The road network of Bangladesh in different district is shown below.

## III. MODEL SELECTION

The road accidents are influenced by several factors, including vehicle characteristics, driver behavior, pedestrian characteristics, type of collisions, road and traffic characteristics etc. It is essential to quantify the relative magnitudes of the impacts of these factors on collision occurrence. Various studies have been done to identify those factors related to crash frequency. These studies use statistical models either to determine collision frequencies, collision involvement rates, collision probabilities and collision severities.

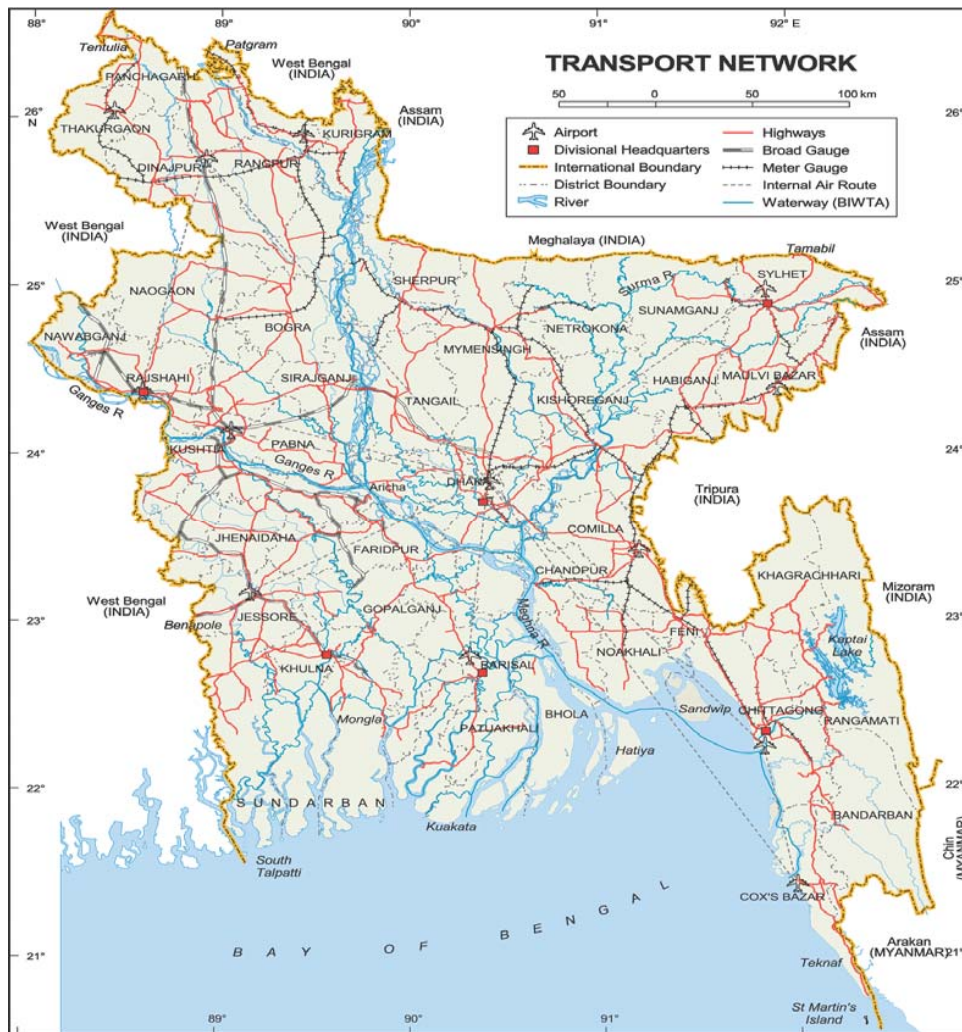


Fig. 3 Transport Network in Bangladesh

The principal objective of a statistical model is to identify a probabilistic system of the form:

$$Y = f(x) \quad (1)$$

where, the dependent variable  $Y$  is a function of a set of independent variables of  $X$ .

The most widely used regression model to analyze collision frequency data is the multiple linear regression model which is given by:

$$y = X\beta + \varepsilon \quad (2)$$

where  $y$  is the number of collisions,  $X$  is a vector of explanatory variables;  $\beta$  is a vector of parameters to be estimated and  $\varepsilon$  is the error term.

However, since accident data are usually random, discrete, nonnegative, sporadic, and count data, there are a lot of undesirable properties in MLR such as assumption of normality and common variance as well as the possibility of

negative outcomes that results in misinterpretation of count data [13]. To overcome the problems associated with MLR models, it was proposed that the Poisson regression should be instead used for modeling accident frequencies [14].

If event ' $n$ ' occurs according to a Poisson process with parameter  $\mu$ , then the Poisson distribution can be written as:

$$\Pr(n_{it} | \mu_{it}) = \{\exp(-\mu_{it}) * \mu_{it}^{n_{it}}\} / n_{it}! \quad (3)$$

where  $(n_{it})$  is the probability of ' $n$ ' accidents occurring on roadway section ' $i$ ' in time ' $t$ ',  $\mu_{it}$  is the expected number of accidents on roadway section ' $i$ ' in time ' $t$ '. The Poisson distribution has the limitation that the variance and mean should be approximately equal i.e.

$$\text{Var}(n_{it}) = E(n_{it}) = \mu_{it} \quad (4)$$

In the case of accident frequencies, the variance is generally much larger than the mean (described as the over-dispersion phenomenon) at which point the Poisson model becomes

inappropriate. To deal with the limitations of the Poisson model, a negative binomial based on a gamma-distributed error term is commonly used. The equation can be rewritten as follows:

$$\ln \mu_{it} = \beta X_{it} + \varepsilon_{it} \quad (5)$$

where  $\exp(\varepsilon_{it})$  is a gamma distributed error term with a mean one and variance  $k$ . The resulting probability distribution under the negative binomial assumption is:

$$\Pr(n_{it} | \mu_{it}, k) = [\Gamma\{(n_{it}+1)/k\} / \Gamma(1/k) n_{it}!] * \{k\mu_{it} / (1+k\mu_{it})\}^n * \{1/(1+k\mu_{it})\}^{1/k} \quad (6)$$

In (6),  $k > 0$  is often referred as over-dispersion parameter. If  $k$  reduces to zero, then the Negative Binomial model reduces to the Poisson regression model. In this way, the Poisson regression model is nested within the Negative Binomial model and a t-test for  $k = 0$  can be used to evaluate the significant presence of over-dispersion in the data. The statistical models are evaluated to select the best model from the competitive set of models. The evaluation will be done with the help of two statistics: Likelihood Ratio Test and Log-likelihood Ratio Index ( $\rho^2$ ).

$$X^2 = -2 [(\beta_R) - (\beta_U)] \quad (7)$$

where,  $(\beta_R)$  is the log likelihood at convergence of the restricted model &  $(\beta_U)$  is the log likelihood at convergence of the unrestricted model. The test statistic is  $X^2$  distributed with the degrees of freedom equal to the difference in the numbers of parameters in the restricted and unrestricted model. To measure the overall goodness of the models, the log-likelihood ratio index will be calculated which is shown below:

$$\rho^2 = 1 - L(\beta)/L(0) \quad (8)$$

where,  $L(\beta)$  is the log-likelihood value of the fitted model.  $L(0)$  is log-likelihood value of the model only with constant term.

The elasticity of accident frequency  $\mu_{it}$  with respect to  $x_{itk}$  is defined as:

$$E = (d\mu_{it} / \mu_{it}) * (x_{itk} / dx_{itk}) \quad (9)$$

where,  $x_{itk}$  is the  $k$  th independent variable for section  $i$  in year  $t$ . The pseudo-elasticity for indicator variable is computed as:

$$E = \{\exp(\beta) - 1\} / \exp(\beta) \quad (10)$$

In our study, as the over-dispersion parameters ( $k$ ) was statistically significant ( $p$ -value  $< 0.0001$ ) which shows the validity of using the Negative Binomial model instead of the Poisson model. Any variable with a ninety percent confidence level was considered to be marginally significant and retained in the final model.

#### IV. RESULTS AND DISCUSSION

TABLE I  
ESTIMATION RESULTS FOR NEGATIVE BINOMIAL MODEL

<b>Number of observation</b>		64	
<b>Chi-square</b>		90.98	
<b>Alpha</b>		0.1482	
<b>P-value (Alpha)</b>		<0.0001	
<b>Road characteristics</b>	<b>Variables</b>	<b>Coefficient</b>	<b>Elasticity</b>
	<b>National road (km)</b>	<b>Categories:</b> Total,0-40, 41-80, 81-120, 120>	
	81-120	0.5556**	0.4263
	120>	0.8726**	0.5821
	<b>National road</b>	<b>Categories:</b> Total,0-5, Above five	
	0-5	-0.7971**	-1.2191
	<b>Regional road(km)</b>	<b>Categories:</b> Total,0-60, 61-120, 121-180,Above 180	
	61-120	0.3837**	0.3187
	121-180	0.6178**	-1.2191
	<b>Zilla road(km)</b>	<b>Categories:</b> Total,0-30,31-60,Above 90	
	0-30	-0.3331*	-0.3953
	31-60	-0.4905**	-0.6331
	<b>Union zilla road(km)</b>	<b>Categories:</b> Total,0-400, 400-800	
	700-1200	0.2327*	0.2076
	<b>Number of gap</b>	<b>Categories:</b> Total,0-300, 301-600	
<b>Socio-economic and Demographic Characteristics</b>	Total	-0.0007**	-0.1308
	<b>Total household population</b>	<b>Categories:</b> Male% 15+, Female% 15+, Economically active male % 15+, Economically active female % 15+	
	Male% 15+	-0.0762*	-4.0354
	Econ. active male % 15+	0.0076*	0.5666
	<b>Temporary insolvency%</b>	<b>Categories:</b> Total%	
	Total%	-0.0250**	-0.5779
	<b>Saving%</b>	<b>Categories:</b> Below 28%, Above 28%	
	Above 28%	-0.9511**	-1.5887
	<b>Household size</b>	<b>Categories:</b> Below 5, Above 5,	
	Above 5	-0.2452*	-0.2778
<b>Sex ratio</b>	<b>Categories:</b> Male: Female (<100:100, 100-120:100, Above 120:100)		
	Above 120	-0.9873**	-0.6274

Note: \* & \*\* denote statistically significant at  $\alpha = 0.10$  &  $0.05$  respectively

To find out significant variables, district wise road characteristics, socioeconomic and demographic characteristics are included in the model in order to determine their complex effects. Discussions of the significant control variables in the final models are presented in two categories: road characteristics, and socio-economic and demographic characteristics.

Several road characteristics were found as significant variables and their effects would be briefly discussed. A number of studies have shown the effect of these variables on the expected crash rate [7], [15]. The expected effect of length of a road on crashes may vary depending on the type of road. For example, higher functional classes of roadways such as national and regional road have better design standard and thus may have fewer number of crashes. On the contrary,

these roads are designed for higher speed which may be associated with more casualties. During model development, two categories are considered for each functional classification (i.e., paved road and unpaved road). Being consistent with other studies, our study has found that increasing the length of paved road per year has an adverse effect on the number of crashes [16]. This result is consistent when variables are entered in the model either continuous or categorical form. Particularly all classes of paved road such as Regional paved road, Union zilla paved road and National paved road are associated to increase accidents. Perhaps better road management and access control measures lead to increase speed of the vehicles on roads. While considering elasticity it is observed that districts having more than 120 km National paved road, the chance of crashes increases 0.5821%.

On the other hand, according to the model result, decrease in paved road length as well as increase in unpaved road length (i.e., National unpaved road, Zilla unpaved road) causes a reduction in road accident. Probably reduced traffic volume in unpaved road may be the reason of this finding. The elasticity for the districts having 0- 5 km unpaved road is -1.2191 which means the reduction of the chance of crash occurrence for that particular category is 1.2191%.

Another statistically significant factor in the model result is the existing gap which means the gap in the continuous road system. These gaps decrease the traffic volume by providing alternative transportation route and mode such as ferries etc. By minimizing traffic exposure these gaps may reduce the road accidents.

In parallel with road characteristics, various socio-economic and demographic factors are also examined in this study. Here different aspects of socio demographic features such as: total number of household, percentage of male and female, sex ratio etc. were considered. Several economic features such as: economically active population, percentage of economically active male and female, different earning group (permanent solvency, temporary solvency, saving, solvent etc.) were also examined. In the model development process both continuous and categorical variables are formed from these socioeconomic and demographic characteristics. From a previous study it was found that demographic changes seemed to be more associated with fatality reductions [9]. In another study it was observed that, counties with a higher percentage of the population under poverty level had significantly increased crash risk [17].

It is shown from the Table I that increase in overall male population reduce the number of crashes. A similar result is echoed by the variable sex ratio as well. On the contrary, economically active male population escalates the crash occurrence. Perhaps economically active male are more exposed to traffic because of job purpose comparing with the economically inactive male group. The elasticity of this variable is 0.5666 which represents one percent increase in active male population of a district will increase the likelihood of crash frequency by 0.5666%.

Model results show that if the percentage of insolvent people increases in a district, the chances of crashes decrease. Reduced mobility of the insolvent people might be the reason of this finding. Moreover, it is found that if the percentage saving by the people in a district increases, then the likelihood of crashes decrease. More percentage savings in a district reflects economic affluence and perhaps the overall education level of this district is much higher comparing with that of other districts.

The economic groups who save their earning after all expenditures are relatively educated population and thus are more conscious about traffic safety. On the other hand, insolvent population perhaps is less concerned about their safety. It is noticed from the model result that if the household size increases, the chances of crash decrease. Perhaps the large household size restricts the movement of the family members, thus reduces the exposure of traffic.

#### V.CONCLUSION

It is clear that factors related to road infrastructure, socioeconomic and demographic features have significant impact on crash occurrence. The model out comes reveal which factors increase or decrease the district level crash frequency. Our result indicates that when the length of the paved road increases, it enhances the chances of crashes. On the contrary, when the length of the unpaved road increases, it reduces the likelihood of crashes.

Socioeconomic and demographic characteristics have significant role on influencing the quantity of accident. From the study it was found that increase in household population will result in the increase of accident number. We also obtained that male population decreases the crash number, while female population shows vice versa. But this situation changes when the effect of economic condition on gender is incorporated. It is observed that increase in active male population increases the likelihood of crashes. From the model results we have also seen that different economic conditions such as: permanent insolvency, temporary insolvency, saving amount etc. play an important role on the outcome of district level crash number.

This study will provide information to the policy makers as well as transportation engineers or planners about the factors affecting crash occurrence at district level. Moreover, this study reveals which factors will increase the crashes and which factors will decrease the crashes. Based on these information proper countermeasures can be developed for the reduction of crashes. This study is beneficial for a developing country like Bangladesh, which is experiencing a significant number of road casualties in recent times because of increased population, increased traffic, less development in roads and highways sector.

#### ACKNOWLEDGMENT

The authors would like to thank Roads & Highways Department (RHD), Bangladesh Statistical bureau (BSB),

Bangladesh Road Transport Authority (BRTA) and Local Government Engineering Department (LGED) for their assistance in this research.

## REFERENCES

- [1] M. Peden, R. Scurfield, and D. Sleet, E. Jarawan, and C. Mathers, *World Report on Road Traffic Injury Prevention*, World Health Organization, Geneva, 2004.
- [2] Transport Canada, *Road Safety in Canada*, Transport Canada, Ottawa, TP 15145E, ISBN 978-1-100-18621-4, 2011.
- [3] M. M. R. Sheikh, "A Statistical Analysis of Road Traffic Accidents and Casualties in Bangladesh," M.S. thesis, Dept. Civil Eng., Napier Univ., Germany, 2004.
- [4] A. A. Jahan, "Road safety countermeasures and traffic calming practices in Roads and Highways Department in Bangladesh," in *Proc. International Conference on Road Safety*, Dhaka, Bangladesh, 2006, pp.120-123.
- [5] M. M. Hoque, M. F. Rahman, S. Sarkar, and M. Ashrafuzzaman, "Hazardous road locations on national highways in Bangladesh," in *Proc. International Conference on Road Safety*, Dhaka, Bangladesh, 2006, pp.229-237.
- [6] M. M. Hoque, "The road to road safety: Issues and initiatives in Bangladesh," in *Proc. 2nd Asian Regional Conference on Safe Community and 1st Bangladesh Conference on Injury Prevention*, Dhaka, 2004.
- [7] R. Tay, and S. M. Rifaat, "Factors contributing to the severity of crashes at intersections," *Journal of Advanced Transportation*, Vol. 41, No. 3, pp.245-265, 2007.
- [8] X. Wang, and M. Abdel-Aty, "Analysis of left-turn crash injury severity by conflicting pattern using partial proportional odds models," *Accident Analysis and Prevention*, Vol. 40, pp.1674- 1682, 2008.
- [9] R. B. Noland, and L. Oh, "The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois country-level data," *Accident Analysis Prevention*, 36, pp.525-532, 2004.
- [10] D. J. Graham, and S. Glaister, "Spatial variation in road pedestrian casualties: the role of urban scale, density and land use mix," *Urban Studies*, Vol. 40, No. 8, pp.1591-1607, 2003.
- [11] S. Rahman, S., "Development of an Accident Prediction Model for Intersections of Dhaka City, Bangladesh," *International Journal of Computer Applications*, Vol. 47, No. 16, pp.10-16, 2012.
- [12] M. S. Haque., and K. G. Meraz, "Study of hazard index on Ambarkhana to Bandar, Under Sylhet City Corporation," *Journal of civil engineering, The institution of engineers Bangladesh*, Vol. 36, pp.15-23, 2008.
- [13] S. Joshua, and N. Garber, "Estimating truck accident rate and involvement using linear and Poisson regression model," *Transportation Planning and Technology*, Vol. 15, pp.41-58, 1990.
- [14] S. P. Miaou, "The relationship between truck accidents and geometric design of road section: poisson versus negative binomial regression," *Accident Analysis and Prevention*, Vol. 26, pp.471-482, 1994.
- [15] E. Amoros, J. L. Martin., and B. Laumon, "Comparison of road crashes incidence and severity between some French counties," *Accident Analysis and Prevention*, 35, pp.537-547, 2003.
- [16] A. Hedayeghi, A. S. Shalaby, B. N. Persaud, B. N. and C. Cheung. C., "Temporal transferability and updating of zonal level accident prediction model," *Accident Analysis and Prevention*, 38, pp.579-589, 2006.
- [17] J. A. Valverde, and P. P. Jovanis, "Spatial analysis of fatal and injury crashes in Pennsylvania," *Accident Analysis and Prevention*, 38, pp.618-625, 2006.



**Dr. Shakil Mohammad Rifaat** was born in Dhaka, Bangladesh, in 1976. He received his PhD degree in transportation engineering from the University of Calgary, Alberta, Canada, in 2010 and Masters degree (Specialization: Transportation Engineering) from the National University of Singapore. He did his B.Sc. in Civil Engineering degree from Bangladesh University of Engineering and Technology in 2000.

He joined at the Islamic University of Technology in 2010 and is working as an Associate Professor in the Department of Civil & Environmental Engineering. His main research interests are Traffic Safety, Transportation Data Analysis, Econometric Modeling.

Dr. Rifaat is the author of several articles published in the premium journals.



**Md. Hishamur Rahman** was born in Dhaka, Bangladesh, in 1990. He received his B.S. degrees in civil engineering from the Islamic University of Technology, Board Bazar, Gazipur, in 2012.

From 2013, he is doing his Masters degree in civil engineering. His research interests include Transportation Infrastructure, Resource Allocation and Management, Smart Infrastructure System.

Mr. Rahman was awarded Merit Scholarship for his outstanding result in the board examinations.



**Mohammed** was born in Munshiganj, Dhaka, Bangladesh, in 1989. He received his B.S. degrees in civil engineering from the Islamic University of Technology, Board Bazar, Gazipur, in 2012.

From 2013, he is working as a design engineer with the Development Design and Management. His research interests include Traffic Safety, Smart Infrastructure System, Handling Iteration and Feedback Loop in design process.

Mr. Mohammed was awarded Merit Scholarship for his outstanding result in the board examinations.



**Mosabbir Pasha** was born in Chittagong, Bangladesh, in 1990. He received his B.S. degrees in civil engineering from the Islamic University of Technology, Board Bazar, Gazipur, in 2012.

He is working as a Lecturer in the Department of Civil & Environment engineering with the Islamic University of Technology. His main research interests are Integrated Transportation System, Traffic and Pedestrian Safety and Pavement Design.

Mr. Pasha received "IUT Gold Medal Award" on 2012 for academic excellence.