

**SAFETY ANALYSIS OF UNSIGNALIZED INTERSECTIONS UNDER  
ERRONEOUS ROAD CRASH DATA INFLUENCE**

**by**

**ASHAR AHMED**

**Thesis submitted in fulfillment of the  
requirements for the degree of  
Doctor of Philosophy**

**July 2016**

## TABLE OF CONTENTS

	<b>Page</b>
<b>ACKNOWLEDGMENTS</b>	ii
<b>LIST OF CONTENTS</b>	iv
<b>LIST OF TABLES</b>	ix
<b>LIST OF FIGURES</b>	xi
<b>LIST OF SYMBOLS</b>	xiv
<b>LIST OF ABBREVIATIONS</b>	xv
<b>ABSTRAK</b>	xvii
<b>ABSTRACT</b>	xix
<b>CHAPTER ONE: INTRODUCTION</b>	
1.1 Background	1
1.2 Problem Statement	3
1.3 Objectives	6
1.4 Scope of Study	6
1.5 Limitations of Study	8
1.6 Organization	9
<b>CHAPTER TWO: LITERATURE REVIEW</b>	
2.1 Introduction	11
2.2 Errors in Accident Data	11
2.2.1 Type of Errors	16
2.2.2 Causes of Errors	21

2.2.3	Methods of Data Rectification	23
2.3	Behavioral Analysis	31
2.4	Severity Analysis	34
2.5	Summary	38

### **CHAPTER THREE: METHODOLOGY**

3.1	Introduction	41
3.2	Data Rectification Procedure	43
3.2.1	Study Data	43
3.2.2	Processing of Raw Data	44
3.2.3	Selection of Attributes for Location Identification	50
3.2.4	Equipment Used for Primary Survey	51
3.2.5	Primary Survey Procedure	52
3.2.6	Formulation of Error Free Database	53
3.2.7	Algorithm	54
3.2.8	Introduction of Intersection Geometry as Additional Filtering Parameter	58
3.2.9	Detailed Data Collection	59
3.2.9.1	Equipment Used	60
3.2.9.2	Procedure	61
3.3	Behavioral Analysis	64
3.3.1	Behavioral Observations	64
3.3.2	Weaving Merging Right Turn ‘WMRT’	66
3.4	Binary Logistic Regression	68
3.4.1	Odds Ratio	70

3.4.2	Variable Significance	71
3.5	Data Division	71
3.6	Conditional Probability	73
3.7	Summary	76

#### **CHAPTER FOUR: DATA RECTIFICATION PROCEDURE**

4.1	Introduction	78
4.2	Descriptive Statistics	78
4.2.1	Concentration of Uncontrolled Unsignalized Intersections	79
4.2.2	General Description	83
4.2.2.1	Penang Mainland	83
4.2.2.2	Penang Island	97
4.2.3	Safety Analysis With Respect to Island versus Mainland	111
4.2.3.1	Effect of Geometric Parameters	111
4.2.3.2	Effect of Landuse	113
4.2.3.3	Effect of Unsignalized Intersections Density	115
4.3	Data Rectification	118
4.4	Summary	126

#### **CHAPTER FIVE: EXPLORATORY AND BEHAVIORAL ANALYSIS**

5.1	Introduction	128
5.2	Site Characteristics	128
5.3	Descriptive Analysis	130
5.3.1	Volume	130
5.3.2	Speed	132

5.3.3	Spacing between Vehicles	134
5.3.4	Descriptive Analysis of Single versus Multiple Accidents	136
5.3.5	Descriptive Analysis of Serious versus No Serious Conflicts	140
5.4	Right Turning Behavior of Vehicles Entering The Major Road	144
5.5	Traffic Conflicts with respect to Movement Type of Vehicles Entering the Major Road	144
5.6	Traffic Conflicts with respect to Major Road Width	146
5.7	Discussion	148
5.8	Summary	154

## **CHAPTER SIX: STATISTICAL MODELING AND PROBABILISTIC**

### **ANALYSIS**

6.1	Introduction	155
6.2	Statistical Modeling	155
6.2.1	Correlation Analysis	157
6.2.2	Fitting All Possible Regression Equation	159
6.2.3	Models with only One Independent Variable	160
6.2.4	Models with Two Independent Variables	161
6.2.5	Models with Three Independent Variables	163
6.2.6	Models with Four Independent Variables	167
6.2.7	Models with All Five Independent Variables	169
6.2.8	Final Model	170
6.2.9	Discussion	172
6.3	Safe Critical Gap	175
6.4	Probabilistic Analysis	179

6.4.1	Conditional Probability	180
6.5	Summary	184

## **CHAPTER SEVEN: CONCLUSION AND RECOMMENDATIONS**

7.1	Conclusion	185
7.2	Recommendations for Future Research	189
7.3	Recommendations for Safer Roads and Better Road Safety Planning	191

<b>REFERENCES</b>	<b>194</b>
-------------------	------------

## **APPENDICES**

APPENDIX A: Example of Raw Data Obtained from MIROS

APPENDIX B: Example of the Data Contained in Excel Sheet Named  
'PENANG'

APPENDIX C: Example Field Survey Sheet

APPENDIX D: Equipment Used in Detailed Survey

APPENDIX E: Maps of the 28 Roads Surveyed

APPENDIX F: Results of the Calculations Performed Based on the Model

APPENDIX G: Definitions

## **LIST OF PUBLICATIONS**

## LIST OF TABLES

	<b>Page</b>
Table 2.1 Studies with respect to year of publication and name of country or territory having error in accident data	30
Table 3.1 List of attributes	51
Table 3.2 Event description, coding and scenario	76
Table 4.1 Descriptive statistics of geometric and physical attributes of each road	80
Table 4.2 Geometric parameters with respect to geographical location	112
Table 4.3 Name of road, number of matched accidents and intersections by algorithm	121
Table 4.4 Site name, number and the number of accidents matched as a result of the rectification procedure	124
Table 5.1 Site characteristics	129
Table 5.2 Site name, number and percentage of motorcycles in the far side traffic	132
Table 5.3 Descriptive statistics of sites with single accidents	138
Table 5.4 Descriptive statistics of sites with multiple accidents	139
Table 5.5 Descriptive statistics of sites with serious conflicts	142
Table 5.6 Descriptive statistics of sites with no serious conflicts	143
Table 5.7 Percentage of vehicles performing WMRT and Non-WMRT maneuver	144
Table 5.8 Conflicts with respect to movement type	145
Table 5.9 Traffic parameters of conflicts with respect to movement type	146

Table 5.10	Conflicts with respect to road width	147
Table 5.11	Traffic parameters of conflicts with respect to road width	148
Table 5.12	Comparison of vehicles that performed or not performed WMRT and had a serious conflict	153
Table 6.1	Type, unit, category, design value and acronym of variables in SPSS	156
Table 6.2	Correlations of all independent variables	158
Table 6.3	All possible 31 models	159
Table 6.4	Models with one independent variable only	161
Table 6.5	Models with two independent variables	162
Table 6.6	Models with three independent variables	164
Table 6.7	Models with four independent variables	168
Table 6.8	Model with all five independent variables	170
Table 6.9	Model summary of the best models from each class	171
Table 6.10	Results of conditional probability with respect to accidents and conflicts	181



## LIST OF SYMBOLS

$A_i$	Number of terms with $i^{\text{th}}$ % of Motorcycles Far Side
$B_j$	Number of terms with $j^{\text{th}}$ Number of Accident(s)
$\beta$	Coefficient of an independent variable in the model
$\beta_i$	Coefficients of the independent variables $x_i$
$\beta_0$	Constant of the equation
C	Number of terms with near to far volume ratio $> 1$
$C_{IU}$	Concentration of uncontrolled unsignalized intersections
$l(\beta)$	Likelihood function
$L(\beta)$	Log of the likelihood function
Pr	Probability
$R^2$	Coefficient of correlation
S.D	Standard Deviation
$t_{sc}$	Safe critical gap
$T_{IU}$	Total number of uncontrolled intersections along a particular road
$x_i$	Series of independent variables $i$ ranging from 1 to $n$ with $n$ being the total number of variables
$z$	Latent variable

## LIST OF PUBLICATIONS

### Journals

Ahmed, A., Sadullah, A.F.M. and Yahya, A.S. (2015). Field study on the behavior of right-turning vehicles in Malaysia and their contribution on the safety of unsignalized intersections. *Transportation Research Part F*, DOI:10.1016/j.trf.2015.03.006. *In Press*.

Ahmed, A., Sadullah, A.F.M. and Yahya, A.S. (2015). Effect of roadside development on the safety of intersections without signals. *Applied Mechanics and Materials Vol. 802*, pp 393-398.

Ahmed, A., Sadullah, A.F.M. and Yahya, A.S. (2015). Difference in the behavior of right-turning vehicles and their effect on safety at unsignalized intersections in Malaysia. *Jurnal Teknologi, Vol 76, No. 14*, pp 37–41.

Ahmed, A., Sadullah, A.F.M. and Yahya, A.S. (2015). Evaluating the contribution of physical parameters on the safety of unsignalized intersections. *Journal of Engineering Science and Technology Vol. 10, No. 5 (2015)*, pp 654 – 666.

Ahmed, A., Sadullah, A.F.M. and Yahya, A.S. (2014). Accident analysis using count data for unsignalized intersections in Malaysia. *Procedia Engineering Vol. 77*, pp 45 – 52.

Ahmed, A., Sadullah, A.F.M. and Yahya, A.S. (2013). Contemporary developments in the safety analysis of unsignalized intersections. *International Journal of Advancements Civil Structural and Environmental Engineering, Vol. 1, issue 1*, pp 52 - 55.

### Conferences

Ahmed, A., Sadullah, A.F.M. and Yahya, A.S. (2015). Underreporting of serious and minor injury accidents in Malaysia. *MUTRFC 2015, Selangor, Malaysia*.

Ahmed, A., Sadullah, A.F.M. and Yahya, A.S. (2015). Critical assessment of the infrastructure and vehicle type contributing most in fatal accidents in Penang, Malaysia. *CARS 2015, Kuala Lumpur, Malaysia*.

Gazder, U., Ahmed, A., Sadullah, A.F.M. and Yahya, A.S. (2015). Accident analysis of unsignalized intersections. *7<sup>th</sup> International Civil Engineering Congress (ICEC 2015), June 12-13, Karachi, Pakistan, pp 161-167.*

Ahmed, A., Sadullah, A.F.M. and Yahya, A.S. (2013). Studying the effect of geometric characteristics on crash severity at unsignalized intersections in Penang. *MUTRFC, Kuala Lumpur, Malaysia.*

## LIST OF FIGURES

		<b>Page</b>
Figure 1.1	Vehicular density per square kilometer and per kilometer paved road in each state of Malaysia	8
Figure 2.1	Types of errors in accident data	18
Figure 3.1	Methodology framework	42
Figure 3.2	Selection of road name	45
Figure 3.3	Usage of 'Find and Replace' tool	46
Figure 3.4	Copying of road names	47
Figure 3.5	Type 'none' in case of no match found	48
Figure 3.6	Widths of the minor and major roads	53
Figure 3.7	Procedure of error free database formation	54
Figure 3.8	Pseudo code for the designed search technique	55
Figure 3.9	Flowchart of the proposed algorithm	56
Figure 3.10	Tree model of the proposed algorithm	57
Figure 3.11	Location of road tubes and camera for data recording	63
Figure 3.12	Weaving Merging Right Turn (WMRT) and conventional right turn	67
Figure 3.13	The probability of serious conflict versus the logit transformation	70
Figure 3.14	Division of data with respect to travel direction and vehicle type	72
Figure 4.1	Concentration of $I_U$ per kilometer and % $I_U$	82
Figure 4.2	Map of Jalan Raja Uda (Google Maps 2014)	84
Figure 4.3	Map of Jalan Bagan Ajam (Google Maps 2014)	85

Figure 4.4	Map of Jalan Pongsu Seribu (Google Maps 2014)	86
Figure 4.5	Map of Jalan Besar Nibong Tebal (Google Maps 2014)	87
Figure 4.6	Map of Lorong Mak Mandin 5 (Google Maps 2014)	88
Figure 4.7	Map of Jalan Bagan Lalang (Google Maps 2014)	89
Figure 4.8	Map of Jalan Pengkalan Macang (Google Maps 2014)	90
Figure 4.9	Map of KM22 Jalan Besar Simpang Ampat (Google Maps 2014)	91
Figure 4.10	Map of Jalan Bertam 1, Kepala Batas (Google Maps 2014)	92
Figure 4.11	Map of Jalan Dua Kepala Batas (Google Maps 2014)	93
Figure 4.12	Map of Jalan Merbau Kudung (Google Maps 2014)	94
Figure 4.13	Map of Jalan Telaga Air (Google Maps 2014)	95
Figure 4.14	Map of Jalan Padang Bengali (Google Maps 2014)	96
Figure 4.15	Map of Jalan Penaga P001 (Google Maps 2014)	97
Figure 4.16	Map of Lebuhraya Thean Teik (Google Maps 2014)	98
Figure 4.17	Map of Jalan Paya Terobong (Google Maps 2014)	99
Figure 4.18	Map of Jalan Persiaran Gurney (Google Maps 2014)	100
Figure 4.19	Map of Jalan Cantonment (Google Maps 2014)	101
Figure 4.20	Map of Jalan Air Itam (Google Maps 2014)	102
Figure 4.21	Map of Jalan Padang Tembak (Google Maps 2014)	103
Figure 4.22	Map of Jalan Burma (Google Maps 2014)	104
Figure 4.23	Map of Jalan Jelutong (Google Maps 2014)	105
Figure 4.24	Map of Jalan Bukit Gambir (Google Maps 2014)	106
Figure 4.25	Map of Jalan Tan Sri Teh Ewe Lim (Google Maps 2014)	107
Figure 4.26	Map of Jalan Tanjung Bungah (Google Maps 2014)	108
Figure 4.27	Map of Jalan Batu Gantung (Google Maps 2014)	109

Figure 4.28	Map of Jalan Penaga Z0372 (Google Maps 2014)	110
Figure 4.29	Map of Jalan Sultan Azlan Shah (Google Maps 2014)	111
Figure 4.30	Distribution of landuse for the roads surveyed with respect to geographical location	114
Figure 4.31	Average number of unsignalized intersections and accidents with respect to geographical location	115
Figure 4.32	Average number of accidents with respect to unsignalized intersection density per km	116
Figure 4.33	Typical rectification process	120
Figure 5.1	Total volume of motorcycle versus others	131
Figure 5.2	Total volume near side versus far side	131
Figure 5.3	Speed of motorcycle versus others	133
Figure 5.4	Speed near side versus far side	134
Figure 5.5	Gap size of motorcycle versus others	135
Figure 5.6	Gap size near side versus far side	135
Figure 6.1	Gap size (seconds) versus probability of serious conflict for each of the four combinations	177
Figure 6.2	Accident frequency, near to far volume ratio and percentage of motorcycles in the far side direction	180
Figure 6.3	Near and far side conflict areas	182

## LIST OF ABBREVIATIONS

AD	Accident Database
AES	Automated Enforcement System
ARF	Accident Reporting Form
AT	Area Type
CBP	Community-Based Programs
DRL	Day-Running Light
FD	Field Database
G	Gap acceptance
GPS	Global Positioning System
HCMM	Highway Capacity Manual Malaysia
IRTAD	International Traffic Safety Data and Analysis Group
L	Total length of road in kilometer
LM	Lane Marking
MC	Motorcycle
MIROS	Malaysian Institute of Road Safety Research
ML	Master Loop
MTW	Motorized Two Wheelers

NCAP	New Car Assessment Program
NFSA	Nested Filtered Search Algorithm
OR	Odds Ratio
R	Road width
RTI	Road Traffic Injuries
S	Average speed on road
SPF	Safety Performance Function
TC	Traffic Control
UI	Unsignalized Intersections
V	Vehicle type
W	WMRT
WHO	World Health Organization
WMRT	Weaving Merging Right Turn



**ANALISIS KESELAMATAN DI PERSIMPANGAN TIDAK BERLAMPU  
ISYARAT DI BAWAH PENGARUH DATA KEMALANGAN JALANRAYA  
YANG MEMPUNYAI RALAT**

**ABSTRAK**

Data kemalangan merupakan asas data pengangkutan yang sangat penting bagi sesebuah negeri, Pengecaman kawasan berisiko tinggi lazimnya menggunakan data kemalangan ini. Tanpa maklumat tepat dan yang boleh tentang lokasi kemalangan maka tiada kemalangan boleh dikenalpasti dengan tepat pada seluruh rangkaian pengangkutan. Walaupun kekurangan ini telah disebutkan dalam pelbagai dokumen ilmiah tetapi kajian untuk memperbaiki kekurangan ini masih tidak dapat dikenalpasti. Ini merupakan motivasi untuk penyelidikan ini. Tujuan penyelidikan ini adalah untuk membangunkan kaedah yang boleh digunakan untuk menganggar lokasi tepat bagi kemalangan yang telah berlaku, menjalankan analisis mikroskopik pada tapak yang dikenalpasti dan untuk merumuskan model anggaran risiko yang boleh merangkumi tingkah laku serta parameter trafik yang bersesuaian dengan keadaan di Malaysia. Kaedah menganggar lokasi memerlukan data daripada dua punca; iaitu data lapangan dicerap daripada tapak manakala data kemalangan diperolehi dari MIROS (Institut Penyelidikan Keselamatan Jalan Raya, Malaysia). Daripada keputusan proses pembaikan, setelah menggunakan Algoritma “Nested Filtered Search (NFSA)” untuk memadankan kemalangan dengan persimpangan di mana kemalangan itu mungkin berlaku, maka 16 persimpangan telah berjaya dipadankan dengan betul. Lanjutan dari langkah pertama, fasa kedua telah dimulakan. Dalam fasa ini, data telah dicerap di setiap lokasi menggunakan pencerap

data bersama rakaman video digunakan secara serentak untuk memperolehi data tingkahlaku. Hasil kajian telah mengenalpasti pergerakan belokan khusus yang diistilahkan sebagai perjalinan percantuman ke kanan (WMRT) yang dilakukan oleh kenderaan dari jalanraya minor telah dikenalpasti. Analisis selanjutnya, menggunakan teknik konflik, mendedahkan bahawa WMRT adalah lebih selamat daripada pembelokan ke kanan yang konvensional. Parameter ini telah digunakan dalam pembentukan model bersama dengan penerimaan sela dan jenis kenderaan untuk menganggarkan kebarangkalian konflik serius berlaku. Model ini telah berjaya menganggarkan sela genting selamat untuk motorsikal dan kenderaan lain yang membelok ke kanan daripada jalanraya minor ke major. Satu parameter baharu dikenali sebagai 'nisbah kepadatan sisi dekat kepada sisi jauh' telah diperkenalkan dan didapati mempunyai hubungan yang tinggi antara kemalangan dan konflik ditapak. Ini telah menghasilkan penemuan kepada hubungan bersyarat antara peratusan motorsikal, nisbah itu dan kemalangan serta konflik yang berlaku ditapak. Kebarangkalian kemalangan dan konflik meningkat dengan penurunan peratusan motorsikal yang bergerak pada arah sisi jauh dengan syarat nisbah kepadatan sisi dekat kepada sisi jauh adalah lebih daripada satu. Hasil penyelidikan ini telah berjaya menangani ralat daripada data kemalangan dan meningkatkan keselamatan pada persimpangan yang tidak mempunyai lampu isyarat.

# **SAFETY ANALYSIS OF UNSIGNALIZED INTERSECTIONS UNDER ERRONEOUS ROAD CRASH DATA INFLUENCE**

## **ABSTRACT**

Accident data is the backbone of any country's transportation database. Identification of black spots and high risk zones are based on it. Without reliable and accurate information pertinent to accident location not even a single accident could be identified on the entire transportation network. The primary flaw in the Malaysian accident data was the ambiguous location of accident sites. Although this shortcoming was highlighted in the literature by different researchers but none had taken up the task to develop a procedure for its rectification which provided the motivation for this study. The aims of this research were to develop a method that could be used for the estimation of correct location for the accidents that occurred in the past, to conduct microscopic analysis on the sites identified and to formulate risk estimation models that can incorporate the behavioral and traffic parameters relevant to Malaysian conditions. The rectification procedure required data from two sources; therefore, field data was collected from site while accident data was acquired from MIROS (Malaysian Institute of Road Safety Research). As a result of the rectification process, which involved the use of Nested Filtered Search Algorithm (NFSA) to match the accidents with the intersections where they might have occurred, 16 sites were correctly identified that initiated the second phase of the study. In this phase traffic data on each site was collected using data loggers while simultaneous video recording was made to extract behavioural data. A unique type of turning movement termed as the Weaving Merging Right Turn (WMRT) made by

the minor road vehicles was newly identified. Further analysis, using conflict technique, revealed that WMRT results into less number of serious conflicts on wider major roads than the conventional right-turn, thus reducing the chance of occurrence of a serious conflict for all types of motorized vehicles. This parameter was utilized in the model formation along with gap acceptance and vehicle type to estimate the probability of serious conflict. The model enabled the estimation of safe critical gap for motorcycles and other vehicles turning right from the minor to the major road. Another new parameter known as 'near side to far side volume ratio' was formulated and found to be highly related with the accidents and conflicts occurring on site. This led to the discovery of a conditional relationship between the percentage of motorcycles, the ratio and the accidents and conflicts that occurred on each site. The probability of accidents and conflicts increased with decrease in the percentage of motorcycles moving in the far side direction provided that the near side to far side volume ratio remained greater than one. The results of this study will help remove errors from the accident data and improve the safety at unsignalized intersections.

## CHAPTER ONE

### INTRODUCTION

#### 1.1 Background

High-income countries share only 16% of the world's population with 47% of the world's motorized vehicles registered in them but their share in the global road traffic deaths is only 8%. While middle-income countries contain 72% of the world's population and 52% of the world's motorized vehicles but their share in the global road traffic deaths is 80%. This shows that despite being highly motorized, the road safety environment in high-income countries is 10 times better than middle ones. Unfortunately, Malaysia falls into the category of middle-income countries (WHO, 2015) which have the second highest annual road traffic fatality rates in the world, which is 18.5 per 100,000 population. As per International Traffic Safety Data and Analysis Group-IRTAD (2014) Malaysia had the highest road traffic fatality rate per 100,000 population, which was 23.6, among all the countries evaluated in the report. It is more than double than that of Korea, which had 10.8, and approximately six times than that of Japan, which had 4.1.

In order to improve the situation, recent initiatives taken by the Malaysian Government bodies responsible for road safety include the formulation and implementation of the 'Road Safety Action Plan 2014-2020' (IRTAD, 2015), which is based on the UN road safety plan for the decade of action for road safety. The primary target of this plan is to reduce the forecasted number of road fatalities upto 2020 by 50%. This corresponds to fewer than 5368 deaths by 2020 which is a 22% reduction as compared to 2010. Other initiatives with respect to the three pillars of

road safety, which are the user/driver, the vehicle and the environment/infrastructure is as follows:

*User/Driver*

- Community-Based Programs (CBP)
- Automated Enforcement System (AES)
- Concentrated enforcement activity during festival period

*Vehicles*

- Day-Running Light (DRL)
- New Car Assessment Program (NCAP)
- Safety star grading for bus operators
- Performance indicators for periodic technical inspections

*Environment/Infrastructure*

- Authorized left turn
- Policy to enhance guardrail standard

The targets set by the governing agencies can not be achieved unless the measures taken are based on accurate investigation of the causes behind road accidents that lead to road fatalities. The primary requirement for the accuracy of the results of an accident investigation is the availability of an error free accident database. Mistakes in the accident database are common not only in low and middle-income countries but in a high-income and highly advanced country like USA as well (Tegge and Ouyang, 2009; Kim and Levine, 1996). Missing or incomplete information, regarding accident location along with other attributes such as traffic signal and road surface condition, recorded by police has been reported in a study in China (Wang et al., 2011). Mistakes such as duplication and missing data along with

under-reporting of traffic injury deaths have also been reported in the South African Police Data (Chokocho et al., 2013). It was found that the location of 88% of crashes reported in the data acquired for a study in a low and middle income country could not be associated geographically to a point or an area (Wang et al 2013). Although Saudi Arabia is a high income country but in terms of discrepancies in accident data it is equivalent to a low or middle-income country because of the errors due to incomplete filling of accident reports. The most important missing information in their data is the accurate identification of accident location (Al-Ghamdi, 2003). A recent study has acknowledged the discrepancies in the Malaysian national accident database (Abdul Manan et al., 2013). The most common mistake is the incorrect recording of the location where the accident occurred. Because of the mistakes that may occur in accident data it has been recommended in standard road safety manuals to perform accuracy checks on variables such as location, number of vehicles involved, time, accident severity, number of casualties and accident type (PIARC, 2003). Therefore, the need to conduct a road safety study that aims at developing a rectification procedure for Malaysia arises.

## **1.2 Problem Statement**

Descriptive statistics provides standalone numbers for individual parameters such as volume, speed, number of lanes, road width, curvature, vehicle type, vehicular distribution, directional distribution, etc. While a model provides a specific combination of the above parameters linked together through mathematical operators (such as +, -, ×, ÷, e, etc.) such that when values are plugged in for a specific facility (such as an intersection, roundabout, toll plaza, highway, etc.) risk estimates are obtained which are pertinent to the said facility. These values can be utilized by

engineers, transportation planners and law enforcers to improve safety. But the estimates can only be accurate if the values provided belong to the site under investigation. Erroneous values will result into faulty estimates, thereby making the entire road safety exercise ineffective. Therefore, it is extremely important to rectify the data before proceeding with the analysis so that those spots could be correctly identified where the problem exists. As mentioned in the previous section, errors in the Malaysian accident database had been reported in the literature, but no known initiative has been taken to establish a comprehensive procedure that can be used to rectify it. This argument became the initial motivation to conduct this study.

Risk estimation can not be accurate without incorporating behavioral parameters. At unsignalized intersections the behavior of vehicles making a right-turn onto the major road is often characterized by their compliance with the stop rule (Kodsi and Muttart, 2009; Kosaka, et al., 2007; Muttart, et al., 2011; Pradhan, et al., 2005), use of turning indicator (Abdul Manan, 2014a) and aggressive driving or force merging (Kaysi and Alam 2000; Kaysi and Abbany 2007). In low and middle income countries the drivers of minor road vehicles make indigenous maneuvers such as the Opposite Indirect Right Turn (Abdul Manan, 2014a). Their behavior is very different from the vehicles in high income countries. Only one previous study was conducted in Malaysia with respect to the behavior and was limited to motorcycles only and did not take into account other types of vehicles which include cars, vans, lorries, high occupancy vehicles, and heavy goods vehicles. Furthermore, such microscopic studies are not possible without a rectified accident data set. Therefore, unavailability of location based microscopic accident data resulted into