Modeling lane-changing behavior based on queue length at an urban signalized intersection

Amiruddin Ismail, Shahrum Abdullah, Azami Zaharim, and Ibrahim Ahmad

***Abstract*—**This research aims to study and develop models for driver's lane-changing behavior in urban areas using logistic regression method. Initially a pilot study was conducted using a videotape recording technique to film an approach road leading to a signalized intersection in an urban road during the morning off-peak period. Inter related coding methods were designed to described and verify the driver's lane-changing maneuver. Later more video-taping studies were done to develop part of the questionnaires. A questionnaires study to analyze the driver's background, experience, attitudes, lane-changing practices and their driving behavior on the road was carried out in order to develop lane-changing behavior models using the logistic regression method. 14 models of lane and non-lane changing were developed and validated statistically. The statistical validations were based on parameters such as Omnibus test of model coefficients, - 2 Log likelihood, Cox and Snell R square, Nagelkerke R square, Hosmer and Lemeshow test, Chi-square, classification table, standard error, wald statistic, degree of freedom, test for significance, odds ratio and histogram of estimated behavioral probabilities.

***Keywords*—**Lane-changing, pass, queue length, racei.

I. Introduction

The consequences of traffic jam issues had become a big concern for community especially in a big city. However, it is impossible to identify the whole story about it without identifying on the specific factor that contributes to the traffic jam. One of the many factors that have been identified in contributing to traffic jam is the driver's behavior. Thus, the purpose of this research was one way to bring up the issues of driver's lane-changing behavior in an urban area.

Previously, numerous researches have been conducted to investigate the driver's behavior in an urban area. Ikenouek et al. [1] and Botma [2] had developed a model based on the mechanic lane-changing maneuver aspects in urban road. In their studies, the main aspects is the related forces in lane-changing maneuvers happened such as acceleration, reaction times, break forces, vehicle signal aspect and others without decision making aspect to make any lane-changing maneuver. A comparison study between a car-following model and lane-changing model had been made by Gipps in 1986 [3]. From his study, he concluded that lane-changing is more complex than car-following because lane-changing decision depends on the identified objectives that sometimes can contrary on each other.

Actually, the drivers are not good enough in lane-changing making decision [4] since less of lane-changing discipline was ranked at the third position as the main cause of an accident happened in highway. In another context, one researcher [5] explained that the traffic jam problem caused by the stopping bus in the bus stand. A related study by [6] stated that it is important to make an explanation about the cause of lane-changing. Meanwhile, another study [7] at a traffic light intersection shows many drivers changed to the shortest queue lane at started red. A study made in United States by Mason et. al [8] stated that careless in lane-changing was listed in 10^{th} position out of 25 driver's behavior as the cause of accident happened in highway.

II. Background

In this study, initially pilot study was carried out in the city of Birmingham, United Kingdom and followed by several pilot studies in the city of Kuala Lumpur, Malaysia. The method used was by video filming several chosen intersections to gather traffic data and information related to the lane and non-lane changing behavior. For the pilot studies to be effective, the main criteria for having "good" intersections is to have intersections with at least two lanes, preferably five or more lanes when approaching the stop line. With only one lane, a driver will tend to follow the preceding vehicle. With two lanes, a driver will tend to change to the next lane when able, possible and necessary. But if there were more than two lanes, the option to change lane is likely to be more whenever able, possible and necessary.

Manuscript received April 11, 2008: Revised version received, Nov. 1, 2008.

Amiruddin Ismail is with Engineering Faculty, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, MALAYSIA (phone: 603-8921-6203; fax: 603-8921-6147; e-mail: abim@vlsi.eng.ukm.my).

Shahrum Abdullah is with Engineering Faculty, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, MALAYSIA (e-mail: shahrum@eng.ukm.my).

Azami Zaharim is with Engineering Faculty, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, MALAYSIA (e-mail: azami@eng.ukm.my).

Ibrahim Ahmad is with Engineering Faculty, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, MALAYSIA (e-mail: ibrahim@eng.ukm.my).

The location of the intersections for several pilot studies are as shown in Fig. 1, Plates 1 & 2 respectively, and the details layout one of the approach road is as shown in Fig. 2.

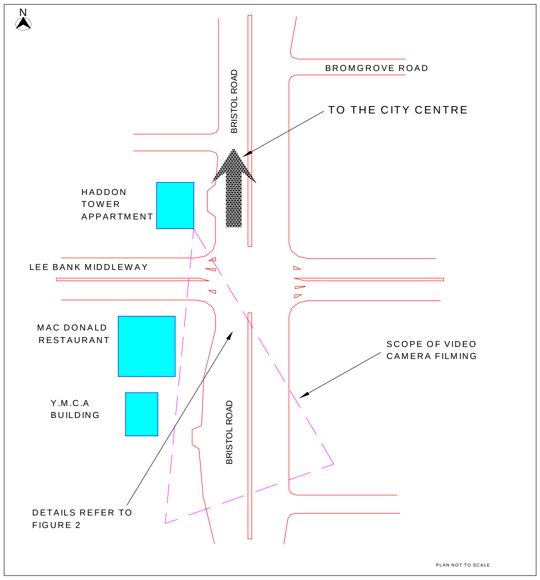


Fig. 1 location of the pilot study in Birmingham, U.K

Plate 1 location one of the pilot study in Kuala Lumpur, Malaysia

Plate 2 location of another study area in Kuala Lumpur, Malaysia

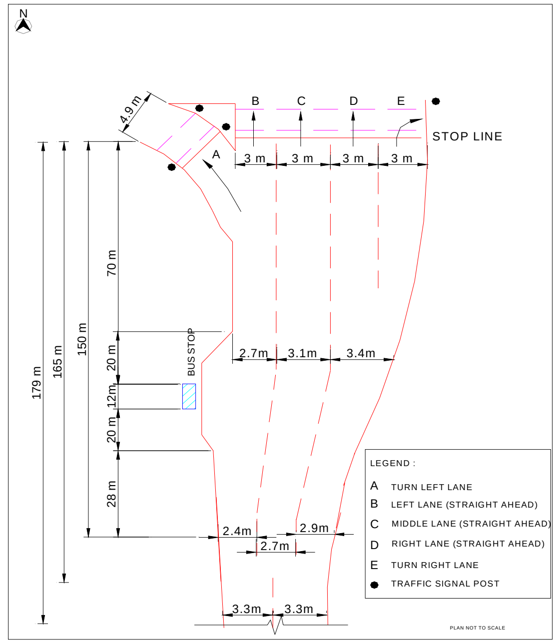


Fig. 2 details layout one of the approach road leading to the signalized intersection

III. The Questionnaires Study

Next, the planning of this study can be shown in the following flow chart as depicted in Fig. 3. After transcribing all the video tapes captured from the pilot studies in Birmingham and Kuala Lumpur, followed by reading the related materials (documents, books, articles) and having discussions with traffic study experts, a set of questionnaires was designed and formulated. Initially the questionnaires were pre-test with several respondents to observe their acceptance or rejection to the questions rose. The final sample of the questionnaires form (in Malay language) is available in appendix C of Amiruddin [9] that comprises of 83 questions with some of these questions are assisted with diagrams and digital photographs.

The questionnaires form is divided into five sections comprises of the number of questions in brackets, namely:

i) Respondent background (9 questions)
ii) Driving experience (8 questions)
iii) Driving attitude (22 questions)
iv) Driver practices of lane changing at an intersection (21 questions)
v) Driver behavior on road (23 questions)

In this study, 2230 sets of questionnaires were carried out to the drivers randomly. 384 or 17.22 % sets of questionnaires were returned back by the respondents. The results of the returning questionnaires were recorded to make an analysis. These results were useful in developing lane and non lane-changing models by using computer software named as Statistical Package of Social Science (SPSS) version 10.0. This study convergent on Klang Valley or *Lembah Klang* which is the most developed place in Malaysia which have many big cities such as Kuala Lumpur, Petaling Jaya, Shah

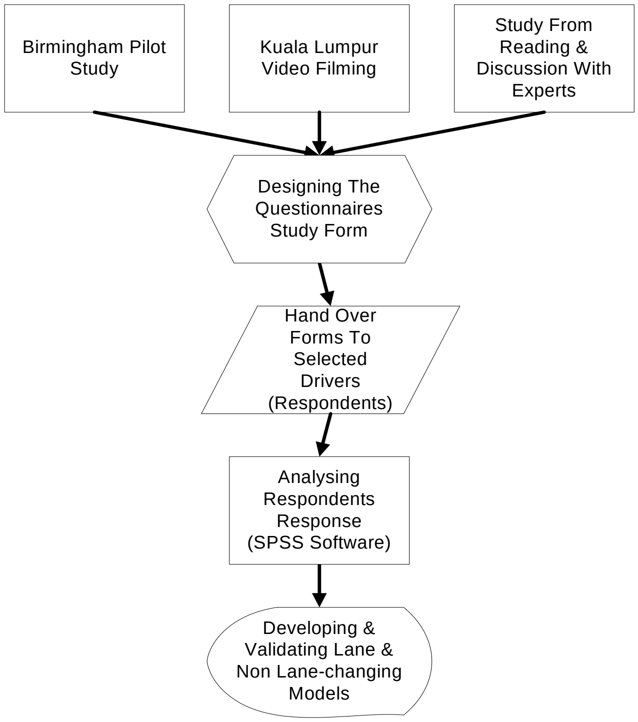


Fig. 3 the study planning flow chart

Alam, Klang, Ampang Jaya, Gombak, Selayang and Kajang. Earlier study by Ismail et al. [10] has developed and validated statistically 14 models of lane and non lane-changing in the same area. This paper will describe only one of the models developed from this study to suit to the title of the paper.

A. Logistic Regression Model

Based on SPSS version 10.0 regression model for one independent variable, logistic regression model is written as:

$$\text{Probability (event)} = \frac{e^{B_0 + B_1 X}}{1 + e^{B_0 + B_1 X}}$$

or the same as

$$= \frac{1}{1 + e^{-(B_0 + B_1 X)}} \qquad (1)$$

Where B_0 and B_1 are the coefficients or estimated parameters from the data which is a scalar, X is the independent variable which is a vector, and e is the basic logarithm value the same with 2.7182818.

For two or more independent variable, logistic regression model is written as:

$$\text{Probability (event)} = \frac{e^{Z}}{1 + e^{Z}}$$

or the same as

$$= \frac{1}{1 + e^{-Z}} \qquad (2)$$

and

$$Z = B_0 + B_1 X_1 + B_2 X_2 + \ldots + B_p X_p \qquad (3)$$

where $B_0, B_1, \ldots B_p$ are the coefficients or estimated parameters from the data which is a scalar, $X_1, X_2, \ldots X_p$ are independent variables which is a vector, and p is the number of independent variables.

If the value of the Probability (event) is less than 0.5, hence can conclude that the event was not occurred but if the probability value is larger than 0.5, hence can conclude that the event was occurred.

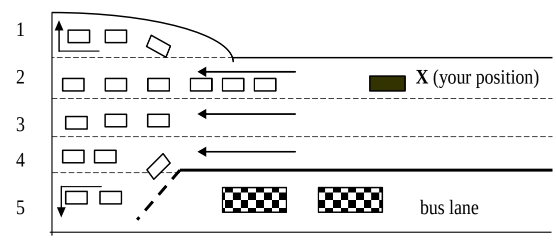
Finally, probability of the event not occurred can be estimate as:

Probability (not occurred) = 1 – Probability (event) (4)

IV. Results And Discussions

A. Lane-Changing Model Based on Queue Length

The lane-changing model based on queue was developed by using a dependent variable named as Action. This model refers to a diagram posed on question 55 of the questionnaire as shown in Fig. 4. This question related to the driver's action whether to choose lane 2, 3 or 4 when the driver drives on lane 2 and facing on various kind of queue length at lane 2, 3 and 4 at the signalized intersection. The driver direction is straight.



Vehicles stop at the stop line due to red light signal. You are in the 2nd lane and still moving, and perceived lane 4 has the shortest queue. What is your Action? (You are going straight ahead).

[ ] Proceed to lane 2.
[ ] Change lane to lane 3 only.
[ ] Change to lane 3 and then change to lane 4.

Fig. 4 question 55 about the respondent (driver) action when facing with various queue lengths approaching the stop line

Based on the SPSS result, four independent variables were significant in developing this model. These variables were selected refers on their significance level which is less or equal to 5 percent (0.05). The independent variables named as:

Racei = refers to a driver group

Fast = refers on drives fast because late to attend an appointment

Queue = refers on selecting a lane with the shortest queue at the signalized intersection

Pass = refers on given passage to another driver

Table 1 refers to the dependent variable that is constant (Action) included in the model. Table shows that no respondent were predicted true on not changing lane when facing on this scenario (Question 55), meanwhile 269 respondents were predicted true on changing lane. Otherwise, 113 respondents have been misclassified in not changing lane and no one else respondent have been misclassified in changing lane. On the whole, 70.4% (> 50%) respondents were predicted true in making decision.

Table 1 classification table (constant in the model)

Observation			Prediction		
			Go straight Changing lane		Right Percentage
			Not changing lane	Changing lane	
Step 0	Go straight Changing lane	Not changing lane	0	113	0.0
		Changing lane	0	269	100.0
	Overall percentage				70.4

Table 2 presents the dependent variable included in the equation. Otherwise, Table 3 shows the independent variables not included in the equation. Both of the tables display all the independent variables and constant are significant when their significance value are less than five percent.

Table 2 dependent variable included in the equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	0.867	0.112	59.859	1	0.000	2.381

Table 3 independent variables not included in the equation

			Score	Df.	Sig.
Step 0	variables	Racei	9.294	1	0.002
		Fast	9.788	1	0.002
		Queue	23.645	1	0.000
		Pass	6.414	1	0.011
	Overall Percentage		41.858	4	0.000

Table 4 demonstrates the result of Hosmer and Lemeshow test. From the results obtained, it was found that the significance value is more than five percent which is 0.666 or 67%. Meanwhile, a Chi-square and degree of freedom values are 5.833 and 8 respectively. The null hypothesis which stated that there is no significant difference between the observed and expected value cannot be rejected. So, this model was very suitable with the collected data.

Table 4 Hosmer and Lemeshow test

Step	Chi-square	Degrees of Freedom	Sig.
1	5.833	8	0.666

Table 5 displays the overall percentage of the true predicted have increased about 3.7% from 70.4% to 74.1% when all the significance independent variables included in the model. Table shows that 36 or 31.9% respondents have been classified true in not changing lane (more than 31.9% from Table 1) and 247 or 91.8% respondents were classified true in changing lane (decrease 8.2% from 100% in Table 1). Besides that, 77 (from 113 in Table 1) respondents were misclassified as not changing lane, meanwhile an addition of 22 respondents from nil have been misclassified in changing lane. On the whole, we can use this changing model.

Table 5 classification table

Observation			Predicted		
			Go straight Changing lane		Right Percentage
			Not changing lane	Changing lane	
Step 1	Go straight Changing lane	Not changing lane	36	77	31.9
		Changing lane	22	247	91.8
	Overall percentage				74.1

Table 6 arrays the variables include in the equation. The Table states that all the independent variables and constant are significant since their significance value is less than five percent except Pass (6 percent).

Table 6 variables include in equation

		B	S.E	Wald	Df	Sig.	Exp(B)	95.0% C.I for Exp(B)	
								Lower	Upper
Step 1	Racei	-1.219	0.394	9.564	1	0.002	0.296	0.136	0.640
	Fast	-0.235	0.106	4.940	1	0.026	0.791	0.643	0.973
	Queue	1.037	0.240	18.612	1	0.000	2.820	1.761	4.516
	Pass	-0.457	0.243	3.527	1	0.060	0.633	0.363	1.020
	Constant	1.295	0.388	11.134	1	0.001	3.653		

Based on the results above, the produced lane-changing model is:

$$\text{Probability (Lane-Changing Based on Queue Length (Action))} = \frac{1}{1+e^{-Z}}$$

Refer to the equation (2),

$$Z = 1.295 - 1.219\ (\text{Racei}) - 0.235\ (\text{Fast}) + 1.037\ (\text{Queue}) - 0.457\ (\text{Pass}) \quad (5)$$

B. Value of Z

The value of Z was identified based on how the variables are recorded in SPSS. Actually, there were two different range used in recoding the respondent's answer such as positive and negative value or yes and no or strongly agree and strongly disagree. For positive range in this case, substituted one (yes) in Racei, Queue and Pass, meanwhile substituted five (strongly agree) to Fast. Hence,

$$Z = 1.295 - 1.219\ (1) - 0.235\ (5) + 1.037\ (1) - 0.457\ (1) = 0.519$$

So,

$$\text{Prob(Lane-Changing Based on Queue Lenght(Action))} = \frac{1}{1+e^{-(-0.519)}} = 0.3731 \quad (6)$$

Based on the result given in the equation, since the probability value is less than 0.5, hence concluded that the lane-changing does not occurred. In other word, this type of driver group was predicted not changing lane when facing on various kind of queue length.

C. Probability Sensitivity

The probability sensitivity was tested to prove its ability in another day practices. This test was made by including the negative range value in the equation. So, substituted zero on Racei, Queue and Pass meanwhile one to Fast. Hence,

Z =1.295 - 1.219 (0) – 0.235 (1) + 1.037 (0) - 0.457 (0) = 1.06

So,

$$\text{Prob(Lane - Changing Based on Queue Lenght (Action))} = \frac{1}{1+e^{-(1.06)}} = 0.7427 \quad (7)$$

Since the probability value is more than 5 percent (>0.5), hence we concluded that the event (lane-changing) was occurred. It means that the driver is predicted to change lane when facing on the differences between queue lengths. So, this model is sensitive with the value of the independent variables.

D. Predicted Probability Histogram Graph

The scatter probability plot of the event occurred or not is represented by Predicted Probability Histogram Graph in Fig. 5. It shows that many groups of event (Lane-changing) occurred were skewed to the right of 0.5 which is 245 cases compared with 80 cases not lane-changing. There were a few groups of event does not occurred were skewed to the left of 0.5 which is 30 cases (Not Lane-changing) and 20 cases (Lane-changing). The result produced by the graph was similar with the equation (7).

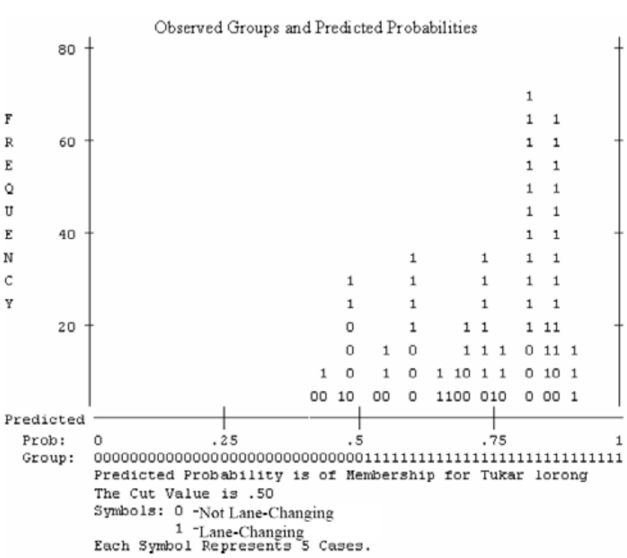


Fig. 5 predicted probability histogram graph (data analysis on SPSS 2001)

The suggested Lane-Changing Model Based on Queue Length is:

$$\frac{1}{1+e^{-[1.295-1.219\,(\text{Racei}) - 0.235\,(\text{fast}) + 1.037(\text{Queue}) - 0.457\,(\text{Pass})]}}$$

V. Conclusion

In this study, initially the method of video filming was used to study the driver behavior on lane-changing when approaching a signalized intersection. Then, the questionnaires study which was partly derived from the video-taping transcriptions was used to study why driver change and not to change lane when approaching a signalized intersection in urban areas. Consequently, models were then developed from these questionnaires and verify statistically by the SPSS logistic regression technique. This study limited itself to only one scenario (Question 55) faced by the drivers around the capital city of Malaysia. The suggested model above only gives an estimation of the probability whether the driver was changed lane or not. Since the lane-changing model based on queue length is sensitive, it can predict whether the driver changed lane or not when facing on the scenario stated in the questionnaires. It can be done by substituted the value of independent variables with the collected range value that have from the current record.

References

[1] K. Ikenouek, T. Saito, H. Hanado, Simulation model of lane-changing on streets. *Reps. of the Nat. Res. Inst. of Police Sci.*, Tokyo, Vol. 14, No. 2, 1973, pp. 1-23.

[2] H. Botma, State of the art report traffic flow models. *Part viii: microscopic traffic flow models*. Stichting Wetinschappiligk Undesgack Veskussveiligheid, No. R-78-42, 1978.

[3] PG. Gipps, Computer program MULTSIM for simulating output from vehicle detectors on a multi lane signal controlled road. *TORG, Dept. of Civil Engrg., Univ. of Newcastle Upon Tyne, U.K.*, Working Paper No. 20, 1976.

[4] C. Seton, Police forces unite in campaign against m-way offenders. *The Times*, Wednesday, 22 January 1992, p. 5.

[5] AF. Sadullah, Bus stop congestion and its contribution to delay. *Proc. U.T.S.G. Annual Conf., Traff. Management*, 1990 (not published).

[6] S. Yousif, J. Hunt, Modelling lane utilisation on British dual-carriageway roads: effects on lane-changing. *Traff. Engrg. & Control*, Vol. 36, No. 12, 1995, pp. 680-687.

[7] Ismail, A, Drivers' lane-changing behaviour in Kuala Lumpur. *Proc. of the 4th. Malaysian Road Conference Part 5-6*, 2000, pp. 1-8.

[8] JM. Mason, K. Fitzpatrick Jr., DL. Seneca, TB. Davinroy, Identification of inappropriate driving behaviors. *Journal of Transpn. Engrg*, ASCE, Vol. 118, No. 2, 1992, pp. 281-298.

[9] Amiruddin Ismail, *Kelakuan pemandu menukar lorong di persimpangan jalan raya berlampu isyarat di kawasan bandar* (lane-changing behavior at signalized intersections in urban areas). Ph.D. thesis (in Malay), Dept. of Civil & Structural Engrg., Fac. of Engrg., Universiti Kebangsaan Malaysia, 2002, pp. 1-472.

[10] A, Ismail, K, Jumari, A.F, Mohd. Sadullah, Developing lane & non lane changing models in urban areas. *Proc. Conf. of the Australian Road Research Board, Vol. 21*, 2003, pp 1523-1533.

Associate Professor Dr Amiruddin Bin Ismail is currently an engineer and a lecturer at the Department of Civil and Structural Engineering, Faculty of Engineering, Universiti Kebangsaan Malaysia (UKM). He was born in Kuala Lumpur in 1956. He earned a Diploma in Civil Engineering from the MARA Institute of Technology (ITM) in 1979. After working for a while he furthered his study in the USA and earned a BSCE from the Point Park College in 1983 and an MSCE from the University of Pittsburgh, Pittsburgh in 1984. Later he joined UKM as a lecturer in 1986. In early 1990 he went to the University of Birmingham, UK to complete a research program. He then furthered his studies at UKM as a part-time student and was awarded a Doctoral Degree in

Civil and Structural Engineering in 2002 for his research in lane-changing behavior in urban areas. His specialty areas are traffic engineering specializing in drivers' lane-changing behavior on urban roads, transport planning and modeling, traffic control and management and pavement engineering. He is a member of the UK Institution of Highways & Transportation (IHT) Malaysian Branch, Road Engineering Association of Malaysia (REAM), Road Engineering Association of Asia & Australasia (REAAA), Transportation Science Society of Malaysia (TSSM) and The Malaysian Academy of Islamic Science (ASASI).

Currently he is the Chairman of REAM Traffic Control and Management Technical Committee 7 (TC07), was Deputy Chairman of the Institution of Engineers' Malaysia (IEM) Highway and Transportation Engineering Technical Division until September 2006, Vice President 1 of the TSSM and a member of the REAM Technical Standing Committee on Technology and Road Management.

Dr. Ismail has published numeral publications in journals and proceedings at international and national arenas.