FACTORS AFFECTING THE REAL-TIME TRAFFIC INFORMATION DEMAND IN BANGKOK

Malai Potipun, Pairoj Raothanachonkun Nakorn Indra-Payoong
Faculty of Logistics, Burapha University, Chonburi 20131
Tel: 0-3810-2222 ext. 3100-3103, Fax: 0-3839-3231
E-mail: {mpotipun, pairoj.iang, nakorn.ii}@gmail.com

ABSTRACT
Purpose – To study the factors affecting the demand of real time traffic information service and classify travelers based on the travel behavior.

Design/methodology/approach – Data from respondent are analyzed and form the basis of the generated Binary logistic model. Binary logistic regression is used for analysis used for predicting the traveler's decision that purchases the real-traffic information service or not and the equation derived from the analysis shows that the factors influencing the decision of the traveler. In addition to, the Hierarchical Cluster Analysis and the K-Mean Cluster Analysis are used for grouping of travelers.

Findings: The results will show that 5 factors affecting the traffic information demand such as traveler' needs, traffic jam experience, distance, willingness to pay for information service, and cost, respectively. The factors that influence the decision are different according to the groups (commuter and recreation). The commuter group, the traffic jam experience influences the traveler's decision. Meanwhile, the traveler in recreation group emphasized the factors of distance. In addition, the travelers are classified into three groups according to the travel behavior. Each of the group is divided into the difference of time, costs, distance and willingness to pay for service.

Research limitations/implications - This survey is specification among travelers who travel in Bangkok regularly so limits its usefulness elsewhere.

Originality/value - The results of this paper will be beneficial to the government and business in order to develop the traffic information systems for the travelers' needs.

Key words: Traveler Behavior, Real Time Traffic Information, Binary Logistic Regression

Article Type: Research paper

Real time traffic information system is the advancement of technology that facilitates information technology for solving traffic congestion (Wei-Hsun, 2008). Today, information technology and communication tools are used for traffic information systems in most of the large cities around the world. The increase of real time traffic information systems has become prevalent in many research fields related to the use of information systems such as traveler behavior (Katrin, 2007), information performance (Minoru, 2009), solving the traffic congestions and willingness to pay for traffic information service (Amalia,1997). This study seeks to investigate whether information technology can be used to benefit traveling and transportation. In addition, information system has become an attractive commercial product and enjoyed growing demand. The information service has grown together with the communication business (Rong and Erik, 2006), such as internet services and mobile phones (Lisa, et.al.2007). The purpose of this paper is to find factors that influence the selection of real time traffic information service and to create a Binary Logistic Regression Model that use forecast traveler's demand.

Travel and Real time traffic information in Bangkok
Bangkok is the capital, where the most population of Thailand and is the center of government, education, communication transportation, financial, commercial and other aspects of Thai life. Bangkok has land border contact with the other provinces. Travelling in Bangkok can be done in several ways, by express boat, automobile, motorcycle, bus, taxi, BTS sky train and by MRT subway. In Bangkok begin the main streets of Thailand's infrastructure, Phaholyothin Road leading all the way to the north of the country, Sukhumvit road which is the route to the east and Phet Kasem road going to the south. Bangkok is the center of trade and transportation, which also means it also has the worst

traffic jams in Thailand. People's expectation that the traffic jams would lessen in Bangkok once the underground and sky train opened also did not take into account the Thai obsession with the car. Owning a car has become an enormous status symbol and even the lower middle-class, who can barely afford to get through to the end of each month, are spending all their money on buying a car. This is causing an even larger congestion of Bangkok's roads, and it's expected to continue. Every year, more cars are added to the capital cities roads and less people use public transportation because of bus users took longer time than users of other transportation (Sureeporn, 2001).

The traffic report system are soon launch a web site that provides traffic information and projections to help busy Bangkok motorists save time on their daily commutes and is reaching out to smart phone users to contribute to the project to increase its accuracy. The web site, traffy.nectec.or.th, will be fully operational in the next four months and will provide a service that predicts the travel time between two points based on traffic conditions. It can currently predict journey times up to 30 minutes in advance, but that will be extended to 24 hours by launch.

A preliminary survey of travelers' behavior in Bangkok (Malai, et.al, 2010) found that motorists spend 90 minutes 56 kilometers and 150 Baht for travel expenses each day. The obtained results are rather similar to the survey by Krungthep Thanakom Co., Ltd. in 2003. In 2010, the National Electronics and Computer Technology Centre report traffic conditions on the main road and highway around Bangkok consisted of multiple channels, such as radio station, internet, call center and traffic variable messages signs or intelligent traffic signs. This system displayed information for the driver in real time to those who were actually traveling at that time by organizing a system that displays traffic conditions on the main routes in Bangkok, offering routing information for the drivers on the road. The Intelligent traffic signs are display traffic on each route at the current time and collect data processing through the detector camera over 150 units that are installed on the main roads around Bangkok measuring the traffic volume and display three line color The color indicate various level of traffic flow with green indicating traffic mobility, yellow showing slow movement but no standstill and red for heavy traffic jam and the need to be avoided (NECTEC, 2010). There are any researcher currently done that studied about the traveler's behavior and traffic information (Karin,2006), (Waadt, 2009), (David,2002).

Willingness to pay for Real time traffic information

In economics, the willingness to pay (WTP) means the maximum amount a person would be willing to pay, sacrifice or exchange in order to receive a good or to avoid something undesired. The studied the willingness to pay for advanced traffic information system or SmarTraveler, an ATIS that provides, via telephone, real-time location-specific traffic and transit information in the greater Boston area. The model is an integrated system of discrete choice and latent variable models. It predicts travelers' frequency of use and subscription under varying pricing scenarios. Two models are presented: one for current SmarTraveler users, and one for non-users. The SmarTraveler usage rate is modeled as a function of payment method and pricing, travelers' travel and socioeconomic characteristics, and their attitudes and perceptions toward ATIS. The data used in model estimation included willingness-to-pay scenarios involving two methods of payment: a flat monthly fee, and a charge per call (Amalia,1997). Similarly, the analyzed the economic valuation of traffic congestion costs in Bangkok and found that the impact of traffic congestion on travelers and their Willingness-To-Pay for improving Bangkok traffic condition was 37.64 baths per trip for auto travelers and 21.37 baths per trip for public transport users (Thanit,2001).

Binary Logistic Regressions

Logistic regression is a statistical method for analyzing a dataset in which there are one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable. The goal of logistic regression is to find the best fitting model to describe the relationship between the dichotomous characteristic of dependent variable and a set of explanatory variables. Logistic regression generates the coefficients and its standard errors and significance levels of a formula to predict a logit transformation of the probability of presence of the characteristic of interest.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_p X_p + \varepsilon \qquad \text{---------- (1)}$$

where Y is the categorical variable
$X_1, X_2, \ldots, X_p$ are independent variables or categorical variable ; p ≥ 1
β_0-β_p are coefficients of independent variables
ε is random Variable

Factors that make people decide in the event of any one or more alterative choice. We consider a rational basis for decision making by economic concepts called "Utility Theory" which indicate that people choose the alternative which promises greater satisfaction. The idea is to write the equation.

$$U_{in} = V_{in} + \varepsilon_{in} \qquad \text{---------- (2)}$$

By U_{in} is the function of satisfaction for person who prefer alternative i
V_{in} is the independent variables that affect satisfaction
ε_{in} is the value of the variable is not observed

The people choose alternative i instead of alternative j in the all option (Cn), when alternative i promises greater satisfaction or maximum utility. In the event that consists of two alternatives (event and no event) we are interested in the opportunity to study the event that the logit model of the basic functions follows.

$$P_{(event)} = \frac{e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p)}} \qquad \text{---------- (3)}$$

By $P_{(event)}$ is the probability that the event occurred
$\beta_1, \beta_2, \ldots, \beta_p$ are parameters on the estimated value
$X_1, X_2, \ldots, X_p$ are the independent variables
e is natural Logarithm (mathematical value =2.71828)

When we know the probability of occurrence, subsequently we can calculate the probability of no event that is:

$$P_{(no-event)} = 1 - P_{(event)} \qquad \text{---------- (4)}$$

Or

$$P_{(no-event)} = 1 - \left[\frac{e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p)}} \right] \qquad \text{---------- (5)}$$

From equation (4) and (5), we can study the ratio between the event and no event is called "Odds Ratio". This ratio can help explain changes in factor X_1 unit will change the Odds Value which requires the interpretation of the marks and the value were calculated. The Odds Ratio is calculated from:

$$Odds\ Ratio = \frac{P_{(event)}}{1 - P_{(event)}} \qquad \text{---------- (6)}$$

Cluster Analyses

Cluster analysis is a class of statistical techniques that can be applied to data that exhibit "natural" groupings. Cluster analysis sorts through the raw data and groups them into clusters. A cluster is a group of relatively homogeneous cases or observations. Therefore, the selected characteristic or or variable that is used to divide the case is very important. In addition, the case must be in only one group. Cluster Analysis is divided into several methods, but the most widely used techniques are the Hierarchical Cluster Analysis and the K-Mean Cluster Analysis. Hierarchical Cluster Analysis is a technique often applied in the cluster variables. This technique deals with a small amount of data or N fewer than 200 and does not know the number of group before. The K-Mean Cluster Analysis is a technique classification into subgroups based on the number of large cases. K-Mean technique uses iteration method that is included in any group of each round case. Then calculates the new median and continues to do so until the median is unchanged or completed (Kanlaya, 2010).

This research data was collected from motorists who travel in Bangkok. The questionnaire was administered to 447 respondents. The questionnaire is divided into three parts: Part 1 is general information of the respondents, part 2 is traveler's behavior and part 3 is the willingness to pay for real time traffic information. The questions consist of both open-ended and closed questions. Data from 447 respondents are divided into two parts, part1 (412 respondents) for modeling and part 2 (65 respondents) for checking the accuracy of models. Creating Binary Logistic model in order to study the relationship between the dependent variables (Y) and independent variable (X) in this study considered variables as Y indicate the purchase of information service or not and the independent variable (X) , composed of distance, travel time, congestion time, attitude and willingness to pay for service. The statistical program is used for calculated parameters and determines the parameters in the equation and validation of the model that consists of the coefficients checking, the significance checking, and the accuracy checking. The percentage accuracy of prediction is calculated from

$$\%correct = \frac{\sum_{n=1}^{N} An}{N} x100 \qquad \text{---------- (7)}$$

Classification of respondents is based on purchase information services decision. The binary Logistic model analysis shows the influence purchase decisions in real time traffic service and the clustering technique is used to divide the group of respondents.

The result of this paper

Factors that affect decision to purchase real time traffic information service.

Factors	All motorist			Commuter group			Recreation group		
	β	S.E.	Sig.	β	S.E.	Sig.	β	S.E.	Sig.
Distance	-.051	.023	.025*	-.056	.025	.027*	.005	.061	.935
Travel Time	.062	.016	.000**	.060	.019	.001**	.076	.033	.022*
Travel cost	.012	.004	.003**	.012	.005	.008**	.006	.011	.614
Congestion Time	.066	.015	.000**	.063	.016	.000**	.091	.041	.026*
Needs	2.063	.302	.000**	2.027	.348	.000**	2.561	.769	.001*
WTP per month	.036	.010	.000**	.040	.012	.001**	.031	.017	.062

Table 1: Variable in the Equation

$$Y = \beta_0 + \beta_1(DIS) + \beta_2(TT) + \beta_3(EX) + \beta_4(CT) + \beta_5(NE) + \beta_6(WTP)$$

Y = Result of the decision (1= purchase, 0 = not purchase)
DIS = Distance of traveling per day
TT = Time of traveling per day
EX = Expenses of traveling per day
CT = Congestion time
NE = Traveler' needs
WTP = Willingness to pay for service

The estimation of parameters in the function found that:

- the estimated coefficients for distance for all motorists were negative (-.051) and significant at the 95% confidence level. Besides distance, travel time, travel cost, congestion time, needs and willingness to pay for service were found to be positive and significant at the 99% confidence level.
- in the commuter group the coefficients for distance were found to be negative (-.056) and significant at the 95% confidence level. Besides distance, travel time, travel cost, congestion time, needs and willingness to pay for service were found to be positive and significant at the 99% confidence level.
- in the recreation group, factors influencing the decision were travel time, congestion time, needs and willingness to pay for service which these factors have the positive coefficients but there were differences significant level, as follow travel time factor was significant at the 95% confidence level, congestion time factor was significant at the 95% confidence level, needs factor was significant at the 99% confidence level and willingness to pay for service factor was significant at the 90% confidence level.

The study of factors influencing the decision to purchase the real time traffic information service among 412 respondents then analysis with computer software, the results show a picture choice and equation following.

$$\mathrm{Y} = -16.075 - .051(DIS) + .062(TT) + .012(EX) + .066(CT) + 2.063(NE) + .036(WTP)$$

The model above shows the most important factors for the real time traffic information purchasing to be the real-time information facilitated for travelers in BKK and its use while on the road. The next factors are the congestion time, travel time per day, distance, willingness to pay for service and travel cost per day.

Form 412 respondents are divided into 2 group according to the purpose of the trip, such as commuter groups who travel for work or study (N=318) and recreation groups who travel for entertainment or other activities (N=94). As commuters, the models show that factors influencing the real-time traffic information services, (sort descending) such as needs, followed by travel time, the congestion time, distance, willingness to pay for service and travel cost per day.

$$\mathrm{Y} = -15.453 - .056(DIS) + .060(TT) + .012(EX) + .063(CT) + 2.027(NE) + .040(WTP)$$

In the recreation groups, the model shows that factors influencing the real-time traffic information services are needs, followed by congestion time, travel time and willingness to pay for information service.

$$\mathrm{Y} = -21.151 + .076(TT) + .091(CT) + 2.561(NE) + .031(WTP)$$

Accuracy and reliability of the model

The validity check for the sign of the coefficients in the model shows the direction of the relationship variables that influence the decision choice any way that sign (+) of the positive coefficient shows that factors correlated to the same direction that we interested in alternative. On the other hand, if the sign (-) of the coefficient show that factor is in opposite relationship with alternative that we are interested.

Factors	The sign of the coefficients		
	All motorist	Commuter group	Recreation group
Distance	-	-	+
Travel Time	+	+	+
Travel Cost	+	+	+
Congestion Time	+	+	+
Needs	+	+	+
Willingness to pay per month	+	+	+

Table 2: Marks show the relationship between the independent samples.

In this study, found that coefficient for distance variable is negative (-); it would mean travelers who journey much more likely to not buy the real-time traffic information services. The sign (+) before coefficients for distance variable coefficient is positive, which means if the traveler go far, they would have increasing opportunity to purchase the information services. The sign before coefficients of other variable is positive for the event which interested. There are factors that positive for purchasing traffic information service. Travel time; if takes more travel, opportunity to purchase services to increased more. Travel costs; if higher the cost to travel more, opportunity to purchase services to increased more. The congestion time; if longer traffic jam, opportunity to purchase information services to increased more. The importance level that when the data show that real-time traffic is very important the opportunity to purchase services to increased accordingly. The willingness to pay factor show that when willingness to pay is expensive, the opportunity to purchase services increased accordingly.
Check the accuracy of the prediction from the sample.

Models	**N**	**forecasting**			
		Frequency Correct	Frequency Missing	% Correct	% Missing
All motorist $Y = -16.075 - .051(DIS) + .062(TT) + .012(EX) + .066(CT) + 2.063(NE) + .036(WTP)$	**412**	**372**	**40**	**90.29**	**9.71**
Commuter group $Y = -15.453 - .056(DIS) + .060(TT) + .012(EX) + .063(CT) + 2.027(NE) + .040(WTP)$	**318**	**289**	**29**	**90.88**	**9.12**
Recreation group $Y = -21.151 + .076(TT) + .091(CT) + 2.561(NE) + .031(WTP)$	**94**	**81**	**13**	**86.17**	**13.83**

Table 3: The accuracy of the prediction samples

Check the accuracy of the prediction from samples collected showed that the group who traveled all (412), the model can predict the decision to buy information services can be 372 people correct and 40 people be wrong that represent the accuracy per error about 90: 10. In the commuter group (318 people), the model can predict the decision to buy the information services can be 289 correct and 29 people wrong that represent accuracy per error about 91: 9. In the recreation group (94 people), the model can predict the decision to buy information services can be 81 people correct and 13 people wrong that represent the accuracy per error about 86:14.

Check the accuracy of the prediction from outside the sample population.
The most appropriate model be tested the accuracy of prediction. The questionnaire divided into 3 groups to analyze the model results in Table 4

Models	**N**	**forecasting**			
		Frequency Correct	Frequency Missing	% Correct	% Missing
All motorist $Y = -16.075 - .051(DIS) + .062(TT) + .012(EX) + .066(CT) + 2.063(NE) + .036(WTP)$	65	47	18	72.31	27.69
Commuter group $Y = -15.453 - .056(DIS) + .060(TT) + .012(EX) + .063(CT) + 2.027(NE) + .040(WTP)$	49	38	11	77.55	22.45
Recreation group $Y = -21.151 + .076(TT) + .091(CT) + 2.561(NE) + .031(WTP)$	16	10	6	62.5	37.50

Table 4: Accuracy of forecasting based on objective for travel

This tables presents the accuracy of the prediction with the total of people traveling (65 people), the model predicts demand of real time traffic information. In the commuter group, the model can predict the decision to buy information services are accuracy 77.55% and errors 27.69 %. In the recreation group, the model can predict the decision to buy information services are accuracy 62.5% and errors 37.50%.

Classification of the travelers
Clustering methods commonly used consists of Hierarchical Clustering and K-Mean clustering because the two methods have different limitations. To study the effect of the best both methods are applied using.

Factors	Group	Mean	Estimate	Meaning
Distance (Kilometers per day)	1	29.88	30	short distance
	2	54.97	55	middle-distance
	3	86.75	85	long distance
Travel Time (Minute per day)	1	51.49	50	short-time
	2	86.93	90	middle-time
	3	121.79	120	long time
Expense (Baht per day)	1	142.83	140	low expense
	2	277.98	280	Average expense
	3	507.58	500	High expense
Congestion Time (Minute per day)	1	49.23	50	Little traffic jam
	2	95.61	90	Heavy traffic
	3	184.80	180	The worst traffic jam
Traveler's needs (scale 1-10)	1	3.82	3	Little needs
	2	5.52	5	Medium needs
	3	7.00	7	The most needs
Willingness To Pay (Bath per month)	1	24.86	25	Minimum willingness to pay
	2	246.38	250	Medium willingness to pay
	3	468.18	470	Maximum willingness to pay

Table 5: The average and means of the various factors

Conclusion
Adoption of modern technology to help manage traffic and transport system has been accepted and has been developed for high performance. The next step is the integration into more commercial services. This study was to analyze travel behavior on Bangkok and to study the real time traffic information service demand for decision making on the road, including the various factors affecting the selection of information services. The results are 3 models that show the factors affecting the real time traffic information demand.

Model 1 $\mathrm{Y} = -16.075 - .051(DIS) + .062(TT) + .012(EX) + .066(CT) + 2.063(NE) + .036(WTP)$
Model 2 $\mathrm{Y} = -15.453 - .056(DIS) + .060(TT) + .012(EX) + .063(CT) + 2.027(NE) + .040(WTP)$
Model 3 $\mathrm{Y} = -21.151 + .076(TT) + .091(CT) + 2.561(NE) + .031(WTP)$

The analysis of three models concluded that the factors impacting the demand for real time traffic information are the factors that most effective is the traveler's needs. The second factor is the congestion time and other factors, such as, travel time, distance, willingness to pay and expense, respectively. For the recreation group travel time and congestion time are prevalent in the traveler's needs and impact their willingness to pay more than commuter group. In addition, the result of clustering group travelers found that the average travel time of the short distance about 30 km with 50 minutes and lowest cost that about 140 bahts, while the travelers who travel a long haul with an average distance of 85 km and a travel time of 120 minutes, are willing to spend about 500 bahts. Finally, in regard to travelers and their willingness to pay for the real-time traffic information service the amount varies between 25 bahts per month as the lowest and 470 bahts per month as the highest to be paid.

References:

Burgess, L., Toppen, A., & Pretorius, P. (2007), Real-time Traveler Information Services Business Model, Washington, DC, USA.

Dziekan, K. & Kottenhoff, K. (2007), "Dynamic at stop real-time information displays for public transport: effect on customers", Transportation Research Part A: Policy and Practice, Elsevier, vol. 41 (6), pp 489-501.

Fukui, M., Nishinari, K., Yokoya, Y., & Ishibashi, Y. 2009, "Effect of real-time traffic information upon traffic flows on crossing roads", Physical A: Statiscal Mechanics and its Applications, Elsevier, vol. 388(7), pp.1207-1212.

Gechan, T. and Suen, L. (1993), "User acceptance of advanced traveler information system for elderly & disabled travelers in Canada", Proceeding of The Vehicle Navigation & Information System Conference in Ottawa,1993, pp 389-392.

Krungthep Thanakom Co., Ltd., 2005 ,"A study on perception of people who live in Bangkok on Bus

Rapid Transit (BRT), Bangkok, Thailand.

Levinson, D. (2003), "The value of advanced traveler information system for route choice", Transportation Research Part C: Emerging Technologies, Elsevier, vol. 11(1), pp.75-87

Luis, A. H., Armijaya, H., Takanashi, I., & Siswosoebrotho, I. B. (2003), "The development of traffic information system and electronic toll collection in Jabotabek toll road network", Proceeding of The Eastern Asia Society for Transportation Studies, Vol.4, pp 995-1010.

Nachaiwieng, T. (2001), "Economic valuation of traffic congestion costs in Bangkok" (Master thesis), Chulalongkorn University, Bangkok, Thailand.

National Electronics and Computer Technology Center: NECTEC, ITS (2007) "ITS Book," Available at: http://wiki.nectec.or.th/its/Cluster/ITSBook (accessed 10 October 2011). September 20, 2011)

Polydoropoulou, A. Gopinath, D. A., & Moshe, B. (1997) "Willingness to pay for advanced travel

information system: SMARTRAVELER case study", Transportation Research PartRecord 1588, Available at: http://www.angelfire.com/ultra/atis0/pagar.pdf (accessed 10 October 2011).

Potipun, M., Raothanachonkun, P., & Indra-Payoong, N., (2010), "The demand of traffic information service in Bangkok: Pilot survey", Proceeding of Seventh National Transport Conference (NTC7) in Bangkok, Thailand.

Punpuing, S., and Ross, H., (2001), "The human side of Bangkok's transport problems", cities, Vol. 18 No1, pp. 43-50.

Vanichbuncha, K. (2010), Advance Statistical Analysis with SPSS for Window, Bangkok, Thailand.

Waadt, A. , Wang, S., Bruck,G. H., & Jung, P. (2009) " Traffic congestion estimation service exploiting

mobile assisted position schemes in GSM networks. Procedia Earth and Planetary Science, Elsevier, vol. 1 (1), pp 1385-1392.

Wallacc, R. R., and Streff, M. F. (1993), "Traveler information in support of driver's diversion decision: A survey of driver's preference", Proceeding of The Vehicle Navigation & Information System Conference in Ottawa, 1993, pp 242-246.

Zhang, R. and Verhoef, E.T., 2006. "A monopolistic market for advanced traveler information systems and road use efficiency," Transportation Research Part A: Policy and Practice, Elsevier, vol. 40(5), pp. 424-443.