Integration of Spatial Models for Web-based Risk Assessment of Road Accident

Sippakarn KASSAWAT[1], Sunya SARAPIROME[1,*] and Vatanavongs RATANAVARAHA[2]

[1]*School of Remote Sensing, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand*
[2]*Department of Transportation Engineering, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand*

([*]Corresponding author's e-mail: sunyas@sut.ac.th)

Received: 21 August 2014, Revised: 8 January 2015, Accepted: 22 February 2015

Abstract

This study is a combination of the web-based system of the Poisson regression model and the Decision expert (DEX) approach to assess the risk of traffic accidents on each segment of Highway 304 in the province of Nakhon Ratchasima, Thailand. The variables of the Poisson model include average daily traffic (ADT), road geometric and environmental parameters. Geometric parameters were used in a factor analysis to the high accident segment portion of the road. The DEX was used as a tool to determine environmental parameters derived from environmental conditions potentially promoting road accidents. The system developed allows users' interaction to vary environmental conditions subject to change with different times of a day and weather. The system can provide the analytical results to identify potential positions at risk of accidents on the highway based on individual users' situations. The system developed can be used as a guide for planning and managing to reduce the number of accidents on the highway. Additionally, the system can provide warning information of road segments for highway users.

Keywords: Spatial analysis, road accident risk assessment, spatial web-based application, Poisson regression model, Decision expert

Introduction

In both developed and developing countries, traffic accidents are one of the main causes of human and economic losses. According to the World Health Organization report in 2010, road accidents were an increasing worldwide problem. There were around 1.24 million deaths and over 23 million injuries per year and around 85 % of these fatalities occurred in developing countries [1]. Road accidents have been a major problem in Thailand for years. According to a 2010 Royal Thai Police report there have been more than 71,000 road collisions that have taken approximately 10,000 lives at an economy expense of more than 232,000 million baht [2].

Accident prediction models (APMs) are generally applied to estimate the expected accident frequencies from a number of roadway entities (highways, intersections, interstates, etc.) and also identify geometric, environmental and operational factors that are related to the occurrence of accidents. It is important to examine the nature of relationships between roadway, environmental and operational factors, and accidents, to understand the specific mechanisms involved in accidents to better predict their occurrence [3].

Current literature dealing with road accident analysis has found that there are diverse models that can be used to analyze the subject. The Poisson regression model is one of the original methods used. Also, it is a simple method to understand and gives the highest accuracy for forecasting [4,5].

The multi-attribute decision making (MADM) is considered a beneficial tool in the scientific and professional contexts of complex problem solving [6]. The basic rule is a decomposition of the decision problem into smaller, less complex sub-problems [7,8].

Combining the MADM and the DEX approaches has created a multi-attribute decision support system [9]. It is designed as an interactive system shell that offers tools for building and verifying a knowledge base, evaluating options and explaining the results. The structure of the knowledge base and evaluation procedures nearly comply with the multi-attribute decision making paradigm and therefore makes a specialized system for decision support [10].

This research focuses on a web-based application on road accident assessment using an integrated Poisson regression model and MADM decision tree (DEX approach). Highway 304 in Nakhon Ratchasima, Thailand was selected as a case study. To supply useful information of road accidents, a web-based application was used. It supports the changes of different factors that affect road accidents and shows the simulation of changing circumstances in various environmental conditions. This application may help in making decisions of improvement or additional safety measures to decrease accidents.

Study area

Nakhon Ratchasima is 259 kilometers from Bangkok with area of around 20,494 square kilometers, making it the biggest province in Thailand, and acts as a gateway to other provinces in the Northeast region of Thailand. Nakhon Ratchasima has a population of 2,582,089 making it the second most populated province in Thailand behind Bangkok.

Highway 304 is a major arterial highway heading from Nakhon Ratchasima towards Eastern Thailand. A critical section of this road (about 78 km) also includes a number of tourist attractions which creates an enormous amount of traffic congestion. This section of road includes a complicated slope and numerous turns, which causes an increment in traffic volume as well as increment in the risk of accidents. According to the data from the Department of Highways, from 2008 to 2011 there were more than 100 accidents, 15 deaths and 174 injures. For these reasons, Highway 304 is an interesting candidate for spatial analysis to assess the risk of accidents.

Materials and methods

The procedure of research regarding the development of a web-based spatial decision support system for the risk assessment of road accidents is displayed as a flowchart in **Figure 1**. The architecture of the proposed system developed consists of a knowledge base that houses the database and the rule base, the Poisson regression model, the DEX approach, and a World Wide Web component allowing the system to be accessible over the Internet.

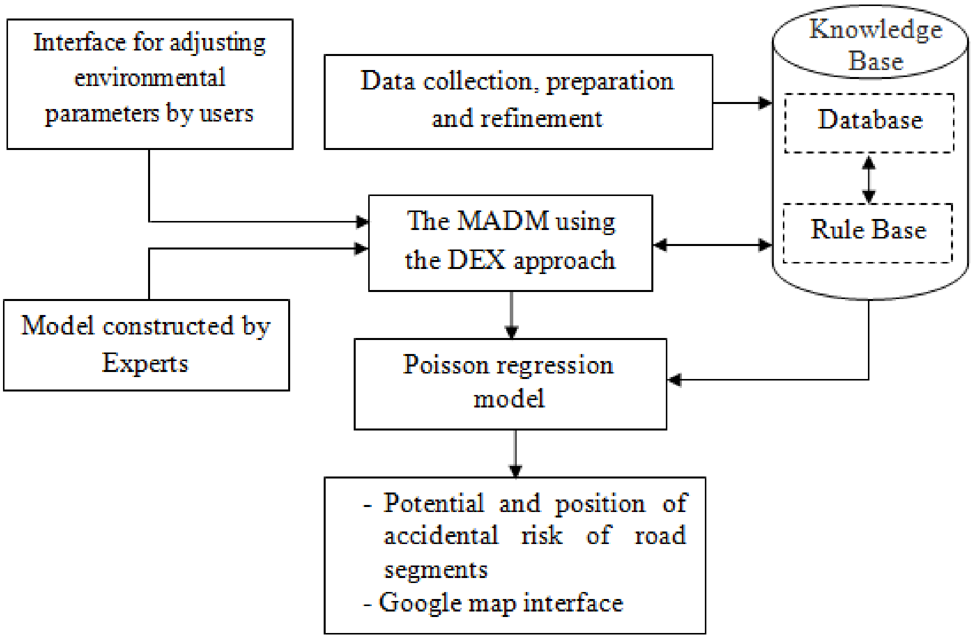

Figure 1 Research methodology.

Road database construction

From previous studies, factors relevant to evaluating the risk of accidents can be listed as shown in **Table 1**. All these attributes were attached to each 100-m road segment in the database of Highway 304. The data on road accidents (2009 - 2012) describe detailed conditions of accidents; such as date and time, road surface conditions, location, etc. The road geometry data were obtained from road design blue prints. ADTs of each 1-km road segment were compiled by the Highways Statistics Unit, 8 Highways Bureau, Nakhon Ratchasima supplied raw data based on 6 survey points (4 inside and 2 outside the study area).

Table 1 Summary of variables affecting traffic accident.

Factors	Accidental causatives	Direction*
ADT	The ADT showed a positive association with crashes [11]. The ADT showed a higher risk for pedestrian accidents [12].	+
Shoulder width	Several studies point to the fact that shoulder width is more beneficial to safety at higher traffic volumes than at lower ones [13].	-
Pavement width	General tendencies for increasing the lane width leading to a decrease in accident rate [13].	-
Degree of horizontal curve	The average crash rate is increased to be 3 times higher on curves compared to tangents [14]. Degree of horizontal curve	+
Number of intersections	The complexity of junctions and the method of junction control seem to show an association with the crash rate [15].	+
Percentage of vertical grade	The elements of vertical alignment that are believed to influence safety show an equivalent increase in the accident frequency of 1.16 percent per percent grade [16].	+
U-turn	The models also show that providing U-turns will result in more crashes at weaving sections [17].	+
Road light Day light (time)	The risk of injury accidents was found to increase in darkness [18]. For crashes of all types, night-time crash risk is only slightly higher overall than daytime risk [19].	+
Weather	An increased risk for having a motor vehicle accident under adverse weather conditions was compared to not driving under adverse weather conditions in a case control study [20].	+

*Positive (+) or negative (-) direction means that the higher value of the factor indicates the higher or lower risk of accident.

MADM using DEX approach for environment criteria

The qualitative hierarchical decision model is based on a selected list of variables or factors, parameters, criteria, and process that are monitored in decision making [8]. As for the knowledge representation in DEX, the methodology of hierarchical decision models is composed of attributes X_i and utility functions F (**Figure 2**). Attributes are variables that represent decision sub-problems. Utility functions or decision rules are designed to incorporate multi-attributes to provide an overall evaluation result. The structure of the knowledge base and evaluation procedure developed are compiled to form a multi-attribute decision making paradigm. A particular knowledge base of DEX includes a tree of attributes and utility functions [7,9].

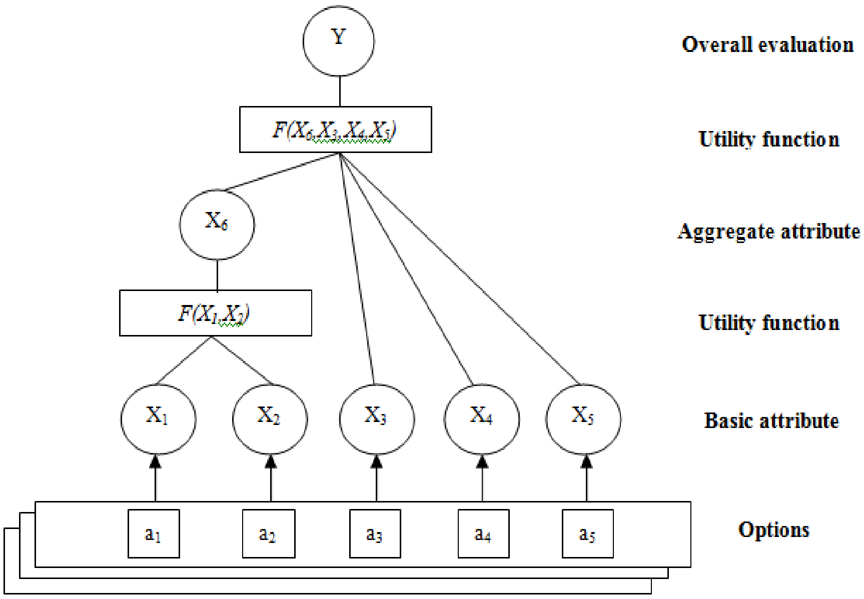

Figure 2 Components of a hierarchical decision model [6,8].

Following the model developed, a questionnaire was constructed to illicit preferences and values for criteria by surveying experts in the Department of Highway and Department of Transportation faculties.

Poisson regression model

A binomial trial is an experiment with 2 possible outcomes, success or failure. The probability of success is the same in each trial and each trial is independent of the other trials [4]. However, accidental occurrence fundamentally follows a Bernoulli trial with unequal probability of independent events that is also known as the Poisson trial.

For typical accident cases, the number of trials (n) is very large and success is very low and it is not the same in each trial. Let X be the number of events in a given interval and the expected number of accident occurrence is λ or the mean of the Poisson distribution. The Poisson distribution, $P = \lambda/n$, where n is a number of trials, y is a number of accidents at a road segment (0, 1, 2, 3, …), e is a mathematical constant ≈ 2.718282. The Poisson distribution is given as [21];

$$P(X = y) \cong \frac{\lambda^y}{y!} e^{-\lambda} \tag{1}$$

The model parameters were estimated using the maximum likelihood method. The values of the parameter vectors (β_x) can be obtained by maximizing the likelihood function $L(\lambda|y)$. The function can be expressed as;

$$L(\lambda|y) = \prod_{i=1}^{n} \frac{\lambda^{y_i}}{y_i!} e^{-\lambda} \tag{2}$$

The result provides coefficients of each factor and a constant, as expressed in Eq. (3), for finding λ to construct and calculate the model.

$$\lambda = EXP(\beta_1 F_1 + \beta_2 F_2 + \beta_3 F_3 + \cdots \beta_x F_x) \tag{3}$$

Model validation

The results as risk of accident of each road segment obtained from the model developed are divided into 3 levels: low, moderate, and high, and are scored as 1, 2, and 3, respectively. These scores of segments were used to validate by comparing with the data of actual incidents. Based on data of 2009 - 2012, the highest number of actual annual incidents in a segment is 11 cases while the lowest is 0. To be comparable, the number of incidents in a year to be validated should be classified and scored as follows:

0 - 3 accidents represent low risk scored as 1.

4 - 7 accidents represent moderate risk scored as 2.

> 7 accidents represent high risk scored as 3.

Then, an evaluation was conducted using the Root Mean Square Error (RMSE) to determine the difference between the actual and the estimated scores of the model. The data from 2009 - 2011 are for model construction while the data from 2012 are for model validation.

Due to the limited number of accidents in a validating year, the environment scenarios cannot be separated into poor, moderate, and good. Therefore, the environment scenarios for model input were setup to be poor, moderate, and good instead. Then, the RMSEs of the result from each scenario level and the levels of accident for validating were estimated.

Web-based development

The system was developed as a web-based application working with imported attribute data and users' requirements on variation of environmental attributes, i.e. time, weather and light conditions. The simulation processing includes the performance of the model on road segment attributed data and scenarios of environmental attributes varying according to the user requirements. The result is displayed on Google Maps for user interaction. The system was designed and developed using a relation database. The software used MySQL as a database management system, PHP as an interface between the database server and users, and Apache as a web server.

Results and discussion

The results obtained from the model, which are analyzed based on road segment data, are displayed for each 1-km segment on Google Maps (**Figure 3**). The resulting 3 levels of risk of accidents are colored as green for low risk, yellow for moderate risk and red for high risk.

Dealing with environmental attributes, weather conditions (clear, fog, rain) and the time period of a day are specified and analyzed in the hierarchical decision model. The system then retrieves the geometry and ADT data according to the time specified from the database and analyzes in the Poisson regression model.

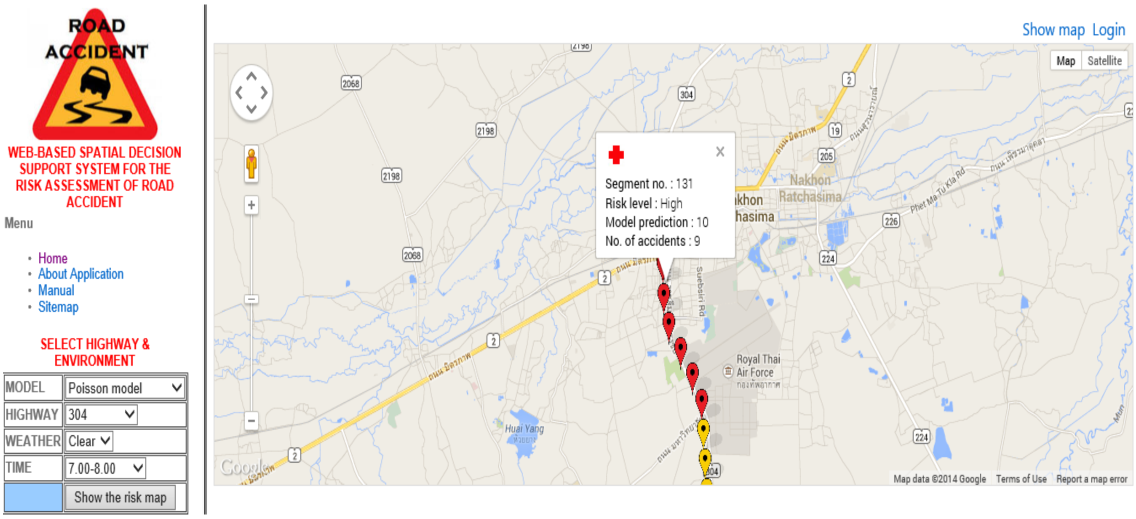

Figure 3 Web interface shows result from the model.

The geometry data are constant according to the actual road conditions, regardless of the change in time or weather. Each segment of the road has a number of fixed factors that have a positive impact on the risk of accidents (higher value represents higher risk), including the degree of horizontal curve, percentage of vertical grade, the number of road connections, the number of U-turns and the number of intersections. A number of fixed factors that have a negative impact on the risk of accidents (higher value represents lower risk) includes the pavement width and shoulder width.

The ADT data will change based on the time specified in the system by the user and vary according to the number of kilometers which is obtained from estimating the raw data from various survey points on Highway 304 and the land use. In this study, it is found that the risk of accidents is higher in areas with a high ADT value.

In addition to the ADT, the road light is taken into consideration based on the period of time. During daylight hours (6.00 am - 6.00 pm), there is no difficulty in light conditions for every segment. During nighttime (6.00 pm - 6.00 am), the system considers the segments with or without road lights. The segments without road lights are more at risk of accidents.

The weather is a factor that changes according to the input of a user which can affect the occurrence of accidents. When the user specifies the weather conditions in the system, it will apply this to every segment. It can be found that the rainy weather is more likely to increase the risk of accidents than the misty and the clear weather, respectively.

The results presented show the risk of accidents will vary according to the data on road geometry, ADT, weather and time at each segment. The model indicates that the road segments with high risk of accidents are in the community areas of Nakhon Ratchasima, where there is the traffic congestion and there are a number of connections to various blocks, or in the kilometers of 128 to 132. In the study area, the risk is reduced when the distance is away from the community areas. This means that the road segments in community areas that have a high traffic volume are more at risk of accidents. The results were consistent with Lord [22], Martin [23] and Pande and Abdel-Aty [24]. In addition, the weather and time that changes according to the requirements of users can affect the change in the risk of accidents in each segment.

A validation of the system results was performed using estimated results of each road segment from 2009 - 2011, under 3 environmental scenarios, and incident levels for 2012. The validation resulted in RMSEs of 0.546, 0.395, and 0.426 for poor, moderate, and good environment scenarios, respectively.

This means that the results obtained from the model are close to the level of actual accidents. The highest error is approximately 25 % for the poor environment scenario. This indicates that the spatial accuracy in terms of identifying highly accidental potential of road segments is considered acceptable for the study of road accident assessment. Therefore, it can be concluded that the model developed in the system is reliable.

Conclusions

The system developed integrated the MADM using the DEX approach and Poisson regression model to assess the risk of road accidents. The system allows database and rule base to be updatable. It also allows users to interact with the system in terms of varying environment scenarios with the resulting observation displayed on Google Maps. The research used data of 2009 - 2011 for model construction and 2012 for validation.

The validation results, RMSEs, based on the comparison of levels of estimated Poisson distribution from the model and actual incidents of road segments in the validating year were acceptable. This leads to the conclusion that the identification of high accident potential of road segments is adequate for planning and management in road accident reduction. As a result, the system developed could serve the objective of the study.

Acknowledgements

The authors are grateful to the Prince of Songkla University for providing financial support of the research. Special thanks is extended to the Bureau of Highway 8 Nakhon Ratchasima, Department of Highway for providing data of Highway 304. The authors also would like to acknowledge the reviewers for kindly improving our manuscript.

References

[1] World Health Organization. *Global Status Report on Road Safety 2013: Supporting a Decade of Action*. WHO Press, Geneva, 2013, p. 1-4.
[2] A Leelakajonjit. Comparison of international road accident database and Thailand's. *J. Soc. Transport. Traffic Stud*. 2011; **3**, 39-46.
[3] B Nambuusi, T Brijs and E Hermans. *A Review of Accident Prediction Models for Road Intersections*. Policy Research Centre for Mobility and Public Works, Hasselt, 2008, p. 1-9.
[4] RR Dinu and A Veeraragavan. Random parameter models for accident prediction on two-lane undivided highways in India. *J. Saf. Res.* 2011; **42**, 39-42.
[5] SH Kim, SB Chong, KH Song and KS Chon. Development of an accident prediction model using GLIM (Generalized Log-linear Model) and EB method: A case of Seoul. *J. Eastern Asia Soc. Transport. Stud*. 2005; **6**, 3669-82.
[6] T Arh and BJ Blazic. Application of multi-attribute decision making approach to learning management systems evaluation. *J. Comput*. 2007; **2**, 28-37.
[7] M Bohanec and V Rajkovic. Multi-attribute decision modeling: Industrial applications of DEX. *Informatica* 1999; **23**, 487-91.
[8] J Abellan, G Lopez and J Ona. Analysis of traffic accident severity using Decision Rules via Decision Trees. *Expert Syst. Appl.* 2013; **40**, 6047-54.
[9] M Bohanec, B Zupan and V Rajkovic. Applications of qualitative multi-attribute decision models in health care. *Int. J. Med. Informat*. 2000; **59**, 191-205.
[10] M Pipan, T Arh and BJ Blazic. Evaluation cycle management - model for selection of the most applicable learning management system. *WSEAS Trans. Adv. Eng. Edu*. 2008; **5**, 129-36.
[11] DG Kang, S Washington and J Oh. Modeling crash types: New insights into the effects of covariates on crashes on rural intersections. *J. Transport. Eng*. 2006; **132**, 282-92.
[12] C Lee and M Abdel-Aty. Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. *Accid. Prev.* 2005; **37**, 775-86.

[13] N Garber and A Ehrhart. *The Effect of Speed, Flow, and Geometric Characteristic on Crash Rates for Different Types of Virginnia Highways*. Virginia Transpotation Research Council, Virginia, 2000, p. 1-4.
[14] R Lyles and W Taylor. Communicating changes in horizontal alignment, technical report Michigan State University. *Transport. Res. Board* 2006; **559**, 60-3.
[15] EH Choi. *Crash Factors in Intersection-Related Crashes: An On-Scene Perspective*. NHTSA's National Center for Statistics and Analysis, Washington DC, 2010, p. 30-1.
[16] J Ma and K Kockelman. Crash frequency and severity modeling using clustered data from Washington State. *In*: Proceedings of the 2006 IEEE Intelligent Transportation Systems Conference, Toronto, 2006, p. 1621-6.
[17] P Liu, T Pan, JJ Lu and B Cao. Estimating capacity of u-turns at unsignalized intersection: Conflicting traffic volume, impedance effects, and left-turn lane capacity. *Transport. Res. Rec.* 2008; **32**, 44-51.
[18] P O Wanvik. Effects of road lighting: An analysis based on Dutch accident statistics 1987-2006. *Accid. Anal. Prev*. 2009; **41**, 123-8.
[19] A Williams. Teenage drivers: Patterns of risk. *J. Saf. Res.* 2003; **34**, 5-15.
[20] M Hijar, C Carrillo, M Flores, R Anaya and V Lopez. Risk factors in highway traffic accidents: A case control study. *Accid. Anal. Prev*. 2000; **32**, 703-9.
[21] D Lord, SP Washington and JN Ivan. Poisson, Poisson-gamma and zero-inflated regression models of motorvehicle crashes: Balancing statistical fit and theory. *Accid. Anal. Prev*. 2005; **37**, 35-46.
[22] D Lord. Application of accident prediction models for computation of accident risk on Transportation networks. *Transport. Res. Board* 2002; **1784**, 17-26.
[23] JL Martin. Relationship between crash-rate and hourly traffic flow on interurban motorways. *Accid. Anal. Prev*. 2002; **34**, 19-29.
[24] A Pande and M Abdel-Aty. Assessment of freeway traffic parameters leading to lane-change related collisions. *Accid. Anal. Prev*. 2006; **38**, 936-48.