Safety analysis for expressway based on Bayesian network: a case study in China

Ling Wang[1,3], Hua-pu Lu[1], Yi Zheng[2,3]*, Zhijun Qian[4]

[1] *Institute of Transportation Engineering, Tsinghua University, Beijing, China*

[2] *School of civil Engineering, Tianjin University, Tianjin, china*

[3] *National Defense Transportation Department，Military Transportation University，Tianjin, China*

[4] *Logistics Department of Air Force, Beijing, China,*

Received 23 November 2014, www.cmnt.lv

Abstract

The accident rates and the mortality rates of expressway in china show an obviously opposite trend compared with most countries, and they are significantly higher than other classes of highways. Moreover, the rate of devastating accidents on expressway in china is higher, and road safety on expressway in china is still serious. The aim of this research is to attempt to develop the causation of traffic accidents occurred on expressway and seek the accidents mechanisms. This paper presents a safety analysis for expressway in China, analyzes the accidents occurred on one expressway of Shanxi Province in China, and selects 8 variables from four influence factors including driver characteristics, highway characteristics, vehicle characteristics, and atmospheric characteristics. The authors consider the relationship of these variables and use the Netica Software to develop BN model involved 8 nodes. Then, the sensitivity analysis is processed for each variable. The research draws a conclusion that four variables including cause (the driver's illegal behavior), experience, weather and lighting were the main cause of the occurrence of accidents on expressway in China.

Keywords: Safety analysis, Expressway, Bayesian network

1 Introduction

On account of full of control of access, separating the traffic flow in opposite directions, all or part of the interchange and the perfect traffic ancillary facilities, the accident rates and the mortality rates of expressway is significantly lower than ordinary highway. Nonetheless, the accident rates and the mortality rates of expressway in china showed an obviously opposite trend. In 2011, 9583 accidents were found on expressways, accounted for 4.55%, killing 6448 and injuring 13007 in China. And the number of accidents per one hundred kilometers on expressway is 11.29, while the number of accidents per one hundred kilometers on highway is 5.13. The death doll per one hundred kilometers on expressway is 7.59, while the death doll per one hundred kilometers on highway is only 1.52. Furthermore, in 9583 accidents occurred on expressway in china in 2011, the number of accidents with 3 fatalities in one accident was 275, the number of accidents with over 5 fatalities in one accident was 72, and devastating accidents with over 10 fatalities in one accident were 6, taking up 22.90%, 24.32% and 22.22% of the total similar accidents. In other words, the rate of devastating accidents on expressway in china is higher, and road safety on expressway in china is still serious.

Traffic accidents are caused by several different factors, such as driver characteristics, highway characteristics, vehicle characteristics, and atmospheric characteristics. One of the principal objectives of traffic accident analyses is to identify key factors that affect the severity of an accident[1]. The researchers in the domain of traffic accident injury severity focused on attempting to identify the significant variables that contribute to the occurrence of an injury severity in a traffic accident and some researchers initiated to use data mining techniques such as artificial neural networks, regression trees, and Bayesian networks to study the issue mentioned above.

For example, Abdelwahab and Abdel-Aty (2001) used two neural network paradigms, the multilayer perceptron (MLP) and fuzzy adaptive resonance theory (ART) neural networks to model the relationship between driver injury severity and crash factors related to driver, vehicle, roadway, and environment characteristics[2]. Delen (2006) classified the injury severity of a traffic accident into five categories (no injury, possible injury, minor non-incapacitating injury, incapacitating, and fatality) and conducted sensitivity analysis on the trained neural network models to identify the prioritized importance of crash-related factors as they apply to different injury severity levels[3].

Some researchers used classification tree techniques to model injury severity in traffic accidents[4]. They developed a Classification and Regression Tree (CART) model to establish the relationship between injury severity and twenty explanatory variables which included driver/vehicle characteristics, highway/environmental variables and accident variables.

There are quite limited researchers that use Bayesian networks to analyze crash-related injury severity. A Bay-

* ***Corresponding author's*** e-mail: rainvgyi@163.com

esian network represents a set of variables together with a joint probability distribution with explicit independency assumptions. Bayesian networks employ techniques from probability and graph theory to model complex systems with interrelated components[5].

Bayesian networks are increasingly used for understanding and simulating computational models in many domains, including medical diagnostic system[6], student modeling[7][8], and intelligent help assistant in Microsoft Office[9]. Recently, Bayesian networks are used in traffic research domain. Sun (2006) used Bayesian network for traffic flow forecasting based on comprehensive experiments on urban vehicular traffic flow data of Beijing, then included that the Bayesian network is a effective approach for traffic flow modeling and forecasting, both for complete data and incomplete data[10].

Gregoriades (2007) highlighted the interest of using Bayesian Networks (BNs) to model traffic accidents and discussed the need to not consider traffic accidents as a deterministic assessment problem[11]. Juan (2011) pointed out that BNs could describe accidents that involve many interdependent variables; the relationship and structure of the variables could be studied and trained from accident data; they do not need to know any pre-defined relationships between dependent and independent variables[12]. A two-car accidents injury severity model was constructed using BNs based on the data on two-car accidents for Slovenia in 1998[5]. Juan (2011) analyzed 1536 accidents on rural highways in Spain, and built 3 different BNs to classify the severity of accidents into slightly injured and killed or severely injured[12]. Juan (2013) used latent class cluster (LCC) as a preliminary tool for segmentation of 3229 accidents on rural highways in Granada (Spain) between 2005 and 2008 and used Bayesian networks to identify the main factors involved in accident severity[1]. Randa (2011) used Bayesian networks to present a simplifying analysis of traffic accidents injury severity on two-lane way[13]. Randa's study provided a methodology that could be used to minimize the number of variables used in order to determine efficiently the injury severity of traffic accidents without reducing the performance of the model [13].

This research put the safety of expressway as the research object because of the special case in China that the accident rates and the mortality rates of expressway in china showed an obviously opposite trend Compared with many other countries. The primary objective of this paper is to use Bayesian Networks (BNs) to analyze the safety of expressway system based on limited statistic data in order to develop the causation of traffic accidents and seek the accidents mechanisms.

2 Methodology

2.1 BAYESIAN NETWORK DEFINITION

A BN is a Directed Acyclic Graph (DAG) consisting of a set of nodes, representing variables with a finite set of states, and edges, representing the probabilistic causal dependence among the variables. The DAG represents the structure of causal dependence between nodes and gives the qualitative part of causal reasoning in a BN, thus the relations between variables and the corresponding states give the quantitative part, consisting of a Conditional Probabilistic Table (CPT) attached to each node with parents. If U is a universe of variables:

$$U = \{X_1, X_2, \cdots, X_n\}, \quad (1)$$

the joint probability of U is then:

$$P(U) = \prod_{i=1}^{n-1} P(X_i | X_{i+1} \cdots, X_n), \quad (2)$$

from the joint probability distribution $P(U)$, various marginal and conditional probabilities can be computed, e.g. $P(X_i)$, $P(X_i|X_j)$ or $P(X_i|e)$ where , in general, e is an evidence: $e = (e_1, e_2, \cdots, e_m)$ .That is information received from external sources about the possible states/values of a subset of the network. For a set of discrete variables, X_i , the evidence appears in the form of a likelihood distribution over the states of X_i: if an observation is given over some variables if the network, the probability of occurrence of some events can be calculated given the evidence:

$$P(U|e) = \frac{P(U, \mathrm{e})}{P(e)}, \quad (3)$$

$$B_P = \{P(X_i | Pa(X_i), X_i \in U)\}. \quad (4)$$

That is to say that a BN over a set of variables U is a network structure, which is a DAG over U and a set of probability tables $B_P = \{P(X_i | Pa(X_i), X_i \in U)\}$ where $Pa(X_i)$ is the set of parents of is X_i in BN and $i = (1, 2, 3, \cdots, n)$. A BN represents joint probability distributions $P(U) = \prod_{X_i \in U} P(X_i | Pa(X_i))$ [13].

2.2 BUILDING BAYESIAN NETWORK

The construction of BNs consist of a few procedures, the first step is identifying and defining the research domain, identifying the relevant variables. Once the variables are identified, variable states and their relevant initial probabilities are assigned [14]. That is to say, the values of probability normally estimated based on certain sources of evidence (empirical data, expert's belief, literature review, or intuition). The second step is to determine the relationships among the variables and establish the graphical structure of the model. The third step is to apply Bayesian rules to compute conditional probability values for each of variables in the model. The fourth step requires development of scenarios to update and train the model. Once a model is updated, the fifth phase is to run sensitivity analysis to assess the performance of the model against its parameters. The sixth phase of the model development is model validation [15].

3 Data

Traffic accident data were obtained from the public security department traffic management bureau of Shanxi Province in China for a section of freeway in mountain areas for four years (2009-2012). The total number of accidents obtained for this period was 949 involved of 127 injury and casualty accidents and 822 property accidents. Table 1 provides information on the data used for this study.

TABLE 1 Data on expressway accidents

Variable	Values	Frequency	Relative Frequency(%)
accident type	fatality	51	5.37
	injury	76	8.01
	property	822	86.62
lighting	dusk	273	28.77
	daylight	676	71.23
weather	bright	492	51.84
	cloudy	134	14.12
	fog	7	0.74
	rainy	141	14.86
	snow	169	17.81
	other	6	0.63
age	18-24	75	7.9
	25-64	874	92.1
	65-inf	0	0
experience	0-1	95	10.01
	2-5	295	31.09
	6-10	274	28.87
	11-inf	285	30.03
sex	male	910	96.0
	female	39	4.0
cause	fatigue driving	139	14.65
	not according to stipulations	285	30.03
	unsuitable safety distance	140	14.75
	overspeed	12	1.26
	wrong overtaking	9	0.95
	improper operation	16	1.69
	others	348	36.67
vehicle type	large cars	382	40.25
	small cars	567	59.75

Eight variables will be used to build BN model. The data included variables describing the conditions that contributed to the accident and injury severity. The variables include accidents severity variables (fatality, injury, property), weather conditions when the accidents occurred (e.g. bright, cloudy, fog, rainy, snow), accident information (e.g. time of accident, vehicles involved in the accident), characteristics of the driver (such as age or gender, experience of driving), the cause for accidents.

Table 1 presents data for traffic accidents on expressway for selected variables. There are eight variables related to the accident in table 1 including accident type, lighting, weather, age, experience, sex, cause and vehicle type. The share of accidents that resulted in a fatality of at least one person is 5.37%. Over 70% of accidents on expressway occurred at night. More than half happened in good weather and more than 30% of accidents happened in rainy or snow weather(14.86% of accidents happened in rainy and 17.81% of accidents happened in snow weather). Among participants, most of share corresponded to drivers 25 to 64 years old and reached 92.1%, while the share of drivers under 25 years of age was only 7.9%, and the share of drivers over 64 years of age was 0. It also showed that the number of older drivers was relatively less than other age groups of drivers. For drivers involved in accidents, the share of one year of driving experience was significant and reached 10.01%. The vast majority of drivers involved in accidents were male (96%).

4 Modeling with bayesian networks

4.1 BAYESIAN NETWORKS CONSTRUCTION

One can identify three main methods for constructing Bayesian networks when trying to model a particular situation. The first method is largely subjective, one reflects on their own knowledge or the knowledge of others and then captures it into a Bayesian network. The second is the method which one automatically synthesizes the Bayesian network from some other type of formal knowledge. And the above two methods are sometimes known as the knowledge representation (KR) approach for constructing Bayesian networks. The third method for constructing Bayesian networks is based on learning them from data, Here either the structure, the probabilities, or both can be learned from the given data set. Since learning is an inductive process, one needs a principle of induction to guide the construction process according to this machine learning (ML) approach.

This research uses the Netica software to build the BN [16]. The concept of causality was used as the guiding principle whereby the causalities amongst the variables, as determined from the literature review, domain knowledge and views from local fire management practitioners, were graphically expressed by the direction of causation for directing arcs [17]. To avoid over-fitting, it was necessary to construct a network with a simple and shallow topology but not compromising on the ability to infer accidents occurrence with an acceptable error rate as suggested by Marcot et al. (2006) [18]. This required developing a model which best fits the data yet contains the fewest total parameters and therefore, the interest was not in a complex model that did not estimate a different topology or have significant effects on posterior probability estimates.

In machine learning, variable selection is a process that is used to select a subset of variables and to remove variables that do not contribute to the performance of the machine learning technique used. This research basically does not consider the variables selection because of data deficiency. The primary cause of accidents data deficiency in china is that road traffic accident information collection system is imperfect. In addition, another reason is the lack of mass historical data. The background of this research is limited data, and this research shall use the existing accident data adequately. So, Eight variables will be used to build BN model, including accidents severity variables (fatality, injury, property), weather conditions when the accidents occurred (e.g. bright, cloudy, fog, rainy, snow), accident information (e.g. time of accident, vehicles involved in the accident), characteristics of the driver(such as age or gender, experience of driving), and the cause for accidents.

4.2 BAYESIAN NETWORK ESTIMATION

A Bayesian network for a given domain can be estimated using different approaches. This paper uses a template model that should not vary from one problem to another. Our purpose here is to estimate a fixed Bayesian network over a given set of variables, obtained by a combination of expert judgment and empirical data. Specifications for some alternative possibilities for estimating a Bayesian network are presented below.

A difficult part of building a Bayesian network is quantifying probabilities, which can be derived from various sources: from domain experts (subjective probabilities); from published statistical studies; derived analytically, or learned directly from raw data. This research uses the last option, mainly because of the limited data.

According to the structure of Bayesian networks, models can be classified as those with a known structure or those with an unknown structure. There are basically two different approaches to learning the structure of a Bayesian network from data: 1) search and scoring methods and 2) dependency analysis methods. Cooper and Herskovits (1992) presented the Bayesian scoring method [19], and Lam and Bacchus (1994) proposed the minimum descriptption length method. The above mentioned approaches are the two of well-known scoring criteria approaches [20].

All these methods can be expected to find the correct structure only when the probability distribution of the data satisfies certain assumptions. But both methods find only approximations for the true structure. According to the available data, models for learning Bayesian networks can be classified into those with complete data available or those with incomplete data available. In the first case, all variables are observed for all instances in the database while, in the second case, values for some variables may be missing or some variables may not even be observed (hidden variables). Because the available database used for this paper contains complete data, the first possibility is relevant.

4.3 VARIABLES CONSIDERATION

Some conditions of an accident may be called exogenous. They are tied to the accident and happen without the volition or action of the drivers involved [5]. Variables in table 1 in this category are: 1) lighting, 2) weather, 3) vehicle type. Besides these external and objective conditions, there are also internal subjective conditions that relate to the drivers involved: 1) age, 2) experience, 3) sex, 4) cause.

Based on the limited data in this research, objective and subjective internal conditions influence the occurrence of an accident. In this research, three types of accident outcomes are considered: fatality, injury, or property. The fatality accident refers to an accident occurred with at least one person killed. The injury accident refers to an accident occurred with some injuries and some property loss. Then, the fatality accident and the injury accident are defined road traffic accident or ordinary accident in statistic data for road traffic accidents in China. Whereas, the property accident refers to an accident with property loss only, and it is defined simple accident and in general, the property accident is not considered in some researches.

This research used the Netica software to implement the BNs; however, the approach described here should extend to any system allowing equations to specify the probability distributions for each node. Fig.1 and Fig.2 show the initial BN model and the compiled BN model for accident analysis on expressway respectively. The compiled model is shown in Fig. 2 and the beliefs are shown for each node in the form of the belief bars. These represent the initial beliefs (presented as probabilities) about the determinants of the severity of accident (presented as accident type) in Shanxi Province in China as evidenced by the data used.

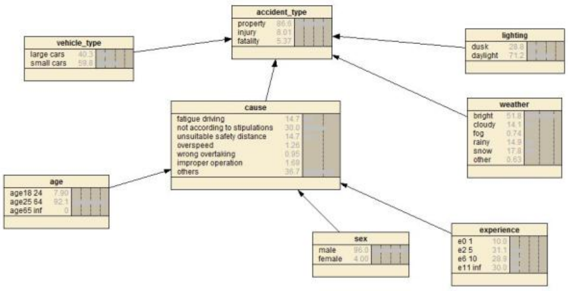

FIGURE 1The initial BN model for accident analysis on expressway

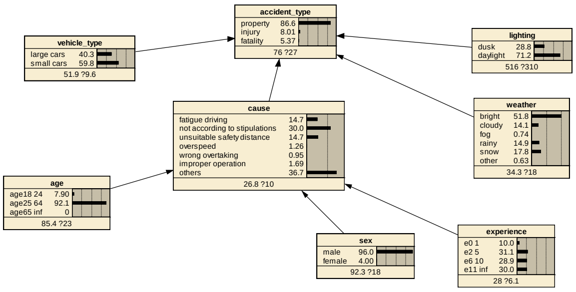

FIGURE 2 The compiled BN model for accident analysis on expressway

4.4 ANALYSIS OF MODEL PARAMETERS

In the above model, two variables including age and experience are continuous variables. Table2 emphasizes the discretization levels we choose.

TABLE 2 Discretization levels for age and experience

Variable	Ranges
age	[18,24],[25,64],[65,∞)
experience	(0,1],[2,5],[6,10],[11,∞)

Other discrete variables process different values as follows.

(1) Accident type. From the above mentioned in this research, accident types are divided three types, namely, fatality, injury and property.

(2) Weather condition. There are six values, namely, bright, cloudy, fog, rainy, snow and other. The definition is based on the statistic data for road traffic accident in China.

(3) Cause. This research analyzes 949 accidents occurred on one expressway in Shanxi province in China and summarizes seven causes related to road accident, including fatigue driving, not according to stipulations, unsuitable safety distance, overspeed, wrong overtaking, improper operation and others.

5 Sensitivity analysis

Sensitivity analysis plays a role to know how sensitive is our belief in this node's value to the findings of other nodes. If it is very sensitive, we may want and it is important for us to know the state of that node, then we may want to invest the effort in determining the values of all the nodes that substantially influence it. The sensitivity analysis was used to determine which variables 'drive' the model and answers the question 'What matters in this decision?' The value of information analysis is the model that determines how much money one is willing to pay to obtain more information.

TABLE 3 The sensitivity of 'vehicle type'

Probability Ranges	Min	Current	Max	RMS Change
large_cars	0	0.4025	1	0.4904
small_cars	0	0.5975	1	0.4904
Mean of Real Value	40.25	51.9	59.75	9.563
Variance reduction = 91.45 (100 %); Entropy reduction = 0.9724 (100 %); Belief Variance = 0.2405 (100 %).				

Table 3 is the Sensitivity of 'vehicle type'. In table 3, range refers to the minimum and maximum values that this measure can take on. RMS refers to "root mean square," which is the square root of the average of the values squared. "Belief" means posterior probability, that is to say, conditioned on all findings currently entered. The definition of RMS Change of Belief is that the square root of the expected change squared of the belief of state f the query variable, due to a finding at the varying variable, with the finding at the varying variable distributed by P(f), and f is a state of the varying variable. Entropy Reduction refers to the mutual information between the query variable and the varying variable.

TABLE 4 The sensitivity of 'lighting'

Probability Ranges	Min	Current	Max	RMS Change
dusk	0	0.2877	1	0.4527
daylight	0	0.7123	1	0.4527
Mean of Real Value	28.77	516.3	713.2	309.8
Variance reduction = 9.6e+004 (100 %); Entropy reduction = 0.8657 (100 %); Belief Variance = 0.2049 (100 %).				

TABLE 5 The sensitivity of 'weather'

Probability Ranges	Min	Current	Max	RMS Change
bright	0	0.5184	1	0.4997
cloudy	0	0.1412	1	0.3482
fog	0	0.0074	1	0.0857
rainy	0	0.1486	1	0.3557
snow	0	0.1781	1	0.3826
other	0	0.0063	1	0.07912
Mean of Real Value	0.63	34.26	51.84	18.36
Variance reduction = 337(100 %); Entropy reduction = 1.841 (100 %); Belief Variance = 0.4659 (100 %).				

TABLE 6 The sensitivity of 'age'

Probability Ranges	Min	Current	Max	RMS Change
age18_24	0	0.079	1	0.2697
age25_64	0	0.921	1	0.2697
age65_inf	0	0	0	0
Mean of Real Value	7.9	85.45	92.1	22.71
Variance reduction = 515.8 (100 %); Entropy reduction = 0.3986 (100 %); Belief Variance = 0.07276 (100 %).				

TABLE 7 The sensitivity of 'sex'

Probability Ranges	Min	Current	Max	RMS Change
male	0	0.96	1	0.196
female	0	0.04	1	0.196
Mean of Real Value	4	92.32	96	18.03
Variance reduction =325 (100 %); Entropy reduction = 0.2423 (100 %); Belief Variance =0.0384 (100 %).				

TABLE 8 The sensitivity of ' experience '

Probability Ranges	Min	Current	Max	RMS Change
e0_1	0	0.1001	1	0.3001
e2_5	0	0.3109	1	0.4629
e6_10	0	0.2887	1	0.4532
e11_inf	0	0.3003	1	0.4584
Mean of Real Value	10.01	28.02	31.09	6.068
Variance reduction =36.82 (100 %); Entropy reduction = 1.895 (100 %); Belief Variance =0.5218 (100 %).				

TABLE 9 The sensitivity of 'cause'

Probability Ranges	Min	Current	Max	RMS Change
Fatigue driving	0	0.1465	1	0.3536
Not according to stipulations	0	0.3003	1	0.4584
Unsuitable safety distance	0	0.1475	1	0.3546
overspeed	0	0.0126	1	0.1115
Wrong overtaking	0	0.0095	1	0.097
Improper operation	0	0.0169	1	0.1289
others	0	0.3667	1	0.4819
Mean of Real Value	0.95	26.84	36.67	10.35
Variance reduction =107.1 (100 %); Entropy reduction = 2.108 (100 %); Belief Variance = 0.5459 (100 %).				

We conduct the sensitivity analysis for each variable in the BN models. The sensitivity analysis results of this model are presented from table 3 to table 9. And the nodes are ranked in according to the degree of influence of their findings on the outcomes of accident node calculated as a measure of mutual information or variance reduction (expressed as a percentage). Cause (i.e. the driver's illegal behaviour) is the most significant factor causing the largest entropy reduction in accident occurrence. Experience, weather, and lighting also show a strong conclusive influence on accident occurrence with more than 0.5 entropy reduction values each. So, we can conclude that some variables 'drive' the model such as cause, experience, weather, and lighting. That is to say that these four factors are the main cause of accidents on expressway in China based on the research data. Similarly, the beliefs for the first four important factors when the finding on states of each model variable is entered are given in table 3 to table 9, and indicate the change in beliefs on the presence or occurrence of fires. The changes in beliefs reveal that the occurrence of accident is highly sensitive and increased in cause (0.5459), experience (0.5218), weather (0.4659) and lighting (0.2049). There are also some changes in beliefs and the changes in beliefs reveal that the occurrence of accident is low sensitive and increased in vehicle type(0.2405), age (0.07276) and sex(0.0384).

6 Conclusion

This study presents a Safety analysis for expressway in China based on Bayesian belief network. Firstly, this paper perform a literature review for identifying the significant variables that contribute to the occurrence of an injury severity in a traffic accident, and the Bayesian Networks application in some domain, especially, in traffic safety research domain. Then, the authors analyze the 949 accident records which occurred on one mountain expressway of Shanxi Province in China. The four kinds of factors influence the traffic safety, including driver characteristics, highway characteristics, vehicle characteristics, accidents characteristics, and atmospheric characteristics. Despite the loss of accident data, this research selects 8 variables, considers the relationship of these variables and use the Netica Software to develop BN model involved 8 nodes based on limited research data, in order to attempt to develop the causation of traffic accidents and seek the accidents mechanisms. In this paper, the authors process the sensitivity analysis for each variable and draw a conclusion that four variables including cause, experience, weather and lighting are the main causes of the occurrence of accidents on expressway in China. These four variables consist of both external factors and internal factors.

Acknowledgments

This work was partially supported by state technology supporting program of China (2012BAG06B01) and road traffic safety technology action and plan of China (2009-2011). The authors also thank the public security department traffic management bureau of Shanxi Province in China for the accident records used in this study.

References

[1] Juan, D. O., G. Lopez, R. Mujalli, F. J. Calvo. Analysis of Traffic Accidents on Rural Highways Using Latent Class Clustering and Bayesian Networks. *Accident Analysis and Prevention*, Vol. 51, 2013, pp. 1–10.

[2] Abdelwahab, H. T., and M. A. Abdel-Aty. Development of Artificial Neural Network Models to Predict Driver Injury Severity in Traffic Accidents at Signalized Intersections. *In Transportation Research Record: Journal of the Transportation Research Board, No. 1746, Transportation Research Board of the National Academies*, Washington, D.C., 2001, pp.6-13.

[3] Delen, D., R. Sharda, and M. Bessonov. Identifying Significant Predictors of Injury Severity in Traffic Accidents Using a Series of Artificial Neural Networks. *Accident Analysis and Prevention*, Vol. 38, 2006, pp. 434-444.

[4] Chang, L.Y., H.W. Wang. Analysis of Traffic Injury Severity: an Application of Non-parametric Classification Tree Techniques. *Accident Analysis and Prevention*, Vol.38, 2006, pp. 1019–1027.

[5] Simoncic, M., A Bayesian Network Model of Two-car Accidents. *In: Proceedings of the 22th International Conference on Mathematical Methods in Economics*. 2004, pp, 282–287.

[6] Niedermayer,D. *An Introduction to Bayesian networks and their contemporary applications*. 1998. Retrieved December 16, 2006, from http://www.neidermayer.ca/papers/bayesian/.

[7] Reye, J. Student Modeling Based on Belief Networks. *International Journal of Artificial Intelligence in Education*, Vol. 14, 2004, 63-96

[8] Vomlel, J. Bayesian Networks in Educational Testing. *International Journal of Uncertainty, Fuzziness and Knowledge Based Systems*, Vol. 12, 2004, pp, 83-100.

[9] Heckerman, D., E. Horwotz. Inferring Informational Goals from Free-text Queries: A Bayesian Approach. *In: Proceedings of the 14th Conference on Uncertainty in AI*, 1998, pp. 230-238.

[10] Sun S. L., C. S. Zhang, and G. Q. Yu. A Bayesian Network Approach to Traffic Flow Forecasting. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 7, 2006, pp, 124-132

[11] Gregoriades, A., Towards a User-centred Road Safety Management Method Based on Road Traffic Simulation. *In: Proceedings of the 39th Conference on Winter Simulation*, Washington, DC, 2007. pp. 1905–1914.

[12] Juan, D. O., R. Mujalli, F. J. Calvo. Analysis of Traffic Accident Injury Severity on Spanish Rural Highways Using Bayesian Networks. *Accident Analysis and Prevention*, Vol. 43, 2011, pp. 402–411.

[13] Randa O. M., D. O. Juan. A Method for Simplifying the Analysis of Traffic Accidents Injury Severity on Two-lane Highways Using Bayesian Networks. *Journal of Safety Research*, Vol. 42, 2011, pp, 317-326.

[14] Druzdzel, M.J., L. C. Gaag. Building Probabilistic Networks: Where Do the Numbers Come From? *Data Engineering*, Vol.12, 2000, pp. 481-486.

[15] Mittal, A., A. Kassim. *Bayesian Network Technologies: Applications and Graphical Models*. IGI Publishing. New York. 2007.

[16] Norsys Software Corporation, 2010. Netica version 4.16 [Online URL]. *http://www.nosys.com*.

[17] Dlamini, W. M. A Bayesian Belief Network Analysis of Factors Influencing Wildfire Occurrence in Swaziland. *Environmental Modelling & Software*, Vol. 25, 2010, pp, 199-208

[18] Marcot, B. G., J. D. Steventon, G. D. Sutherland, R. K. Mccann. Guidelines for Developing and Updating Bayesian Belief Networks Applied to Ecological Modeling and Conservation. *Canadian Journal of Forest Research-Revue Comedienne de Recherché Forestiere*. Vol. 36, 2006, pp, 3063-3074.

[19] Cooper, G.F., E. Herskovits. A Bayesian Method for the Induction of Probabilistic Networks from Data. *Machine Learning*, Vol. 9, 1992, pp, 309–347.

[20] Lam, W. and F. Bacchus. Learning Bayesian Belief Networks: An Approach Based on the MDL Principle. *Computational Intelligence*, Vol.10, 1994, 269-293.

Authors

Ling Wang, 1980. 08, Xinyang City, Henan Province, China

Current position, grades: a Ph.D. candidate, the Institute of Transportation Engineering, Tsinghua University
University studies: master's degree from Academy of Military Transportation in 2006; civil engineering
Scientific interest: transportation planning and management, traffic safety
Publications: more than 10 papers

Huapu Lu, 1957. 02, Tieling City, Liaoning Province, China

Current position, grades: a professor and a Ph.D. student supervisor in the Institute of Transporta- tion Engineering, Department of Civil Engineering, Tsinghua University
University studies: Ph.D. degree from Nagoya University, Japan, in 1993.
Scientific interest: transportation planning and management, intelligent transportation system, traffic flow theory and public traffic
Publications: more than 180 papers

Yi Zheng, 1973.03, Shenyang City, Liaoning Province, China

Current position, grades: a Ph.D. candidate, School of civil Engineering, Tianjin University
University studies: master's degree from Chongqing Jiaotong University in 2007.
Scientific interest: national defense transportation
Publications: more than 10 papers

Zhijun Qian, 1978.11, Xuzhou City, Jiangsu Province, China

University studies: bachelor degree from Academy of Military Transportation in 1999.
Scientific interest: national defense transportation
Publications: more than 5 papers