RELATIONSHIP BETWEEN CONGESTION AND TRAFFIC ACCIDENTS ON EXPRESSWAYS AN INVESTIGATION WITH BAYESIAN BELIEF NETWORKS

By Charitha Dias**, Marc Miska***, Masao Kuwahara****, and Hiroshi Warita*****

1. Introduction

Accidents and congestion are two frustrating events, which can be observed very frequently on roads. Accidents, especially on expressways, can trigger heavy traffic congestions imposing huge external costs and reducing the level of service. Therefore it is obvious that accidents clearly have an impact on congestion. But the opposite, i.e. the effect of congestion on occurrence of accidents, is less studied and still questionable [11]. One can argue that congestion can reduce the high speeds on expressways and as a result of that the accident rate is reduced. But in a congested road section vehicles are closely packed and as a result of that rear-end collisions, back-up collisions as well as side collisions can occur. Therefore it is important to analyze the impact on the accidents by congestion so that the policy makers can implement relevant measures to reduce the external costs of both accidents and congestion.

This paper investigates the effects of traffic congestion on the occurrence of accidents on 8 radial routes (inbound direction) of Metropolitan Expressway (MEX). Data were obtained from the International Traffic Database (ITDb) [6]. Two softwares, namely WinMine Toolkit [2] and MSBNx [5], which use the concept of Bayesian Belief Networks (BBN), were used to model the interrelationships among occurrence of accidents and other variables such as congestion index (CI), traffic density and volume.

2. Relationship between congestion and accidents

Very limited attempts have been made, in the past by several authors, to describe the relationship between accidents and congestion. Among those, Wang *et al*.[11] claimed that traffic congestion, controlling other factors such as flow, curvature, gradient, section length, no. of lanes etc., has little or no impact on frequency of accidents (fatal or non-fatal), using data for M25 highway. But the CI values in their data were relatively low, i.e. less than 0.5, for most of the cases. Therefore, it is questionable that those data really represented congested situations.

Noland and Quddus[8] used a series of negative binomial models to analyze the effect of congestion on road safety. Their results were not conclusive, suggesting that there is little effect of congestion on road safety. They suspected that this might be due to the weakness of proxies they used to represent congestion, plus might be due to the method they implemented to model relationships. While above mentioned studies claimed that there is no any significant relationship between accidents and congestion, Golob and Recker [4], using nonlinear multivariate statistical analysis, concluded that rear-end collisions are more likely to occur under heavily congested stop-and-go traffic. Though this is an indication that congestion has an effect on accidents, the

* Keywords: traffic accidents, congestion, Bayesian belief networks
** Non-member of JSCE, M. Sc., Chodai Co., Ltd.
(2-1-3 Higashi-Tabata, Kita-Ku, Tokyo, Japan 114-0013, Tel: +81-3-3894-3236, Fax: +81-3-3894-3265)
*** Member of JSCE, Lecturer, Institute of Industrial Science, University of Tokyo
(4-6-1 Komaba, Meguro-ku, Tokyo, Japan 153-8505, Tel: +81-3-5452-6419, Fax: +81-3-5452-6420)
**** Member of JSCE, Professor, Institute of Industrial Science, University of Tokyo
(4-6-1 Komaba, Meguro-ku, Tokyo, Japan 153-8505, Tel: +81-3-5452-6419, Fax: +81-3-5452-6420)
***** Member of JSCE, Planning and Environment Department, Metropolitan Expressway Co., Ltd.
(1-4-1 Kasumigaseki, Chiyoda-ku, Tokyo, Japan 100-8930, Tel: +81-3-3539-9389, Fax: +81-3-3502-2412)

link between accidents and congestion has not elaborated clearly.

Though some other authors have attempted to explore the relationship between accidents and traffic volume or flow [1)3)], these studies may not describe the relationship between accidents and congestion, because according to speed-flow characteristics not only flow but also speeds are also needed to determine whether the prevailing traffic condition is congested or un-congested.

These limited previous studies suggest us the need for more studies to better understand the impact of congestion on accidents. In this study BBNs, which can model and graphically interpret the relationships in any system, were used to model the interrelationships among CI, volume, density and occurrence of accidents.

3. Methodology

(1) Data description

Figure 1: Metropolitan expressway (MEX) network

ITDb datasheet

Data: Metropolitan Expressway, Route: 04, Section: 05, Direction: 01

Owner: Metropolitan Expressway Corp. Japan

Provider: International Traffic Database (ITDb) - www.trafficdata.info

Timestamp	Q [veh/5mins]	avg. Speed [km/h]	Occupancy	Event
200606050000	74	81.2	2.7	none
200606050005	75	80.1	2.8	none
200606050010	76	78.7	2.8	none
200606050015	71	77.0	2.6	none
200606081645	178	14.7	38.4	none
200606081650	188	15	35.6	none
200606081655	191	15.4	34.8	none
200606081700	171	13.6	37	none
200606081705	174	13	39	none
200606081710	182	14	35.2	none
200606081715	177	13.1	36.9	none
200606081720	149	11	41.5	none
200606081725	128	8.1	41.5	Accident
200606081730	167	13.3	36.7	Accident
200606081735	170	13.6	36.1	Accident
200606081740	153	12.6	37.7	Accident
200606081745	135	9.8	42.8	Accident
200606081750	136	10.2	39.1	Accident
200606081755	140	10.4	40.5	Accident
200606081800	134	10.2	39.9	Accident
200606081805	144	11.3	34.8	Accident
200606081810	128	9.3	43.8	Accident
200606081815	148	11.7	37.6	Accident
200606081820	240	33.9	21.8	none

Pre-accident period

Figure 2: Available data

Five-minutes aggregated volume, average speed, occupancy and incident data on inbound direction of 8 radial routes (Route numbers 1, 2, 3, 4, 5, 6, 7 and 9) of the MEX, collected over one week were used. One-week data (from 5^{th} June 2006 to 11^{th} June 2006) of ITDb on MEX is free for public and that free data was used for this analysis. Format of raw data is shown in Figure 2.

(2) Data arrangement

a) Pre-accident period.

First the other incidents such as car breakdowns, road works etc. were removed and the 15 minutes time period before an accident was defined as pre-accident period (Figure 2). Here, 15 minutes was selected just to represent the average traffic conditions on the road section before an accident.

b) Congestion Index (CI)

Congestion index expresses the congestion level of a given road section relative to the free-flow conditions. It is non-negative and dimensionless value. Average speed data was converted in to congestion data, i.e. congestion indexes (CI), with slightly modified version of Taylor's [9)] method, as follows.

$$CI = \begin{cases} (V_{FF} - V)/V_{FF} & ;\ if\ V \le V_{FF}\ \ and\ V_{FF} > 0 \\ 0 & ;\ if\ V > V_{FF} \end{cases}$$

Where;

V – *Average speed for the road section*

V_{FF} – *Free flow speed for the road section*

c) Volume classes, CI classes and density classes

Volumes, CI's and occupancies were categorized in to classes as shown in Table 1. This was done because categorized data were needed to learn BBNs.

Table 1: Categorized volumes, CIs and densities

Volume range (veh/5mins)	Volume class
Volume <= 100	Low
100<Volume<=200	LowerMedium
200<Volume<=300	UpperMedium
Volume>301	High

CI Range	CI Class
CI <= 0.2	FreeFlow
0.2 <CI <=0.4	Comfortable
0.4 <CI <=0.6	Medium
0.6 <CI <=0.8	Congested
CI >0.8	HeavilyCongested

Occupancy range	Density class
Occupancy <=15	Low
15<Occupancy<=35	Medium
Occupancy>=35	High

(3) Bayesian belief Networks (BBN)

BBNs are probabilistic graphical models that represent a set of variables and their conditional independencies via a directed acyclic graph [12]. BBN is an effective technique to understand the relationships among variables because BBNs can model the interrelationships among variables with their conditional probabilities in any kind of a system and represent them graphically.

Two applications namely WinMine Toolkit[2] and MSBNx[5], which use the BBN concept, were used to identify conditional independencies among variables and to perform inference, respectively. WinMine toolkit is set of tools, designed for Windows environment, that allow constructing statistical models from categorized data. WinMine has many advantages compared with other traditional methods such as bar graphs, contingency tables and odd ratios to mine for information in data because it can provide a better picture of the factor interrelationships among variables [7]. MSBNx is a component-based application, which can be used to create, assess and evaluate Bayesian networks [5]. It is an excellent tool for inference or updating probabilities.

Categorized volume, CI, density and event for each route were fed in to WinMine toolkit separately in order to obtain the dependency network, which is very similar to Bayesian network. For inference purposes the conditional probabilities, determined with WinMine, were fed in to MSBNx. Figure 3 depicts dependency network along with the conditional probability table (CPT) for the node "Event", constructed with WinMine, and Figure 4 depicts the BBN with un-conditional probability tables for all nodes, constructed with MSBNx, for the inbound direction of Route 4 of MEX.

At a glance some important information can be drawn from these figures. Figure 3 describes the dependencies among the variables in data fed in to WinMine. Figure 4 shows how the inbound direction of Route 8 functions. 15% of the time, within the considered one week time period, it is congested or heavily congested, 76% of the time medium volumes and 81% of the times low densities could be observed. And there is around 1% of chance of an accident.

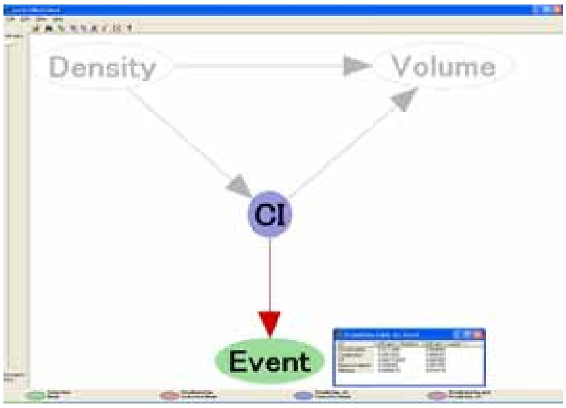

Figure 3: Dependency network constructed with WinMine

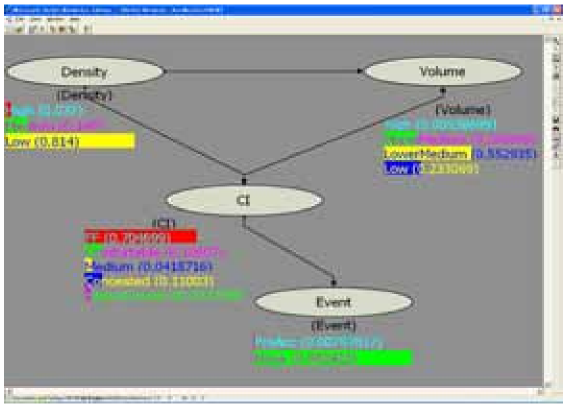

Figure 4: BBN constructed with MSBNx

Un-conditional probabilities, which are shown in Figure 4, are probabilities before any evidence is observed and those are called prior probabilities. When evidence is observed, for example when it is known that the CI reflects heavily congested situation, probabilities for all other nodes can be updated and those updated probabilities are called posterior probabilities. And this process of updating probabilities is called inference.

Inference was performed for BBNs for each considered routes with MSBNx and updated probabilities were noted. Results are described in Section 4.

4. Results

Probabilities of occurring a pre-accident period were plotted against congestion levels, as in Figure 5, and probability of occurring a pre-accident period was plotted against density levels, as in Figure 6. Figure 5 clearly shows that the chance of an accident is increased with the congestion level.

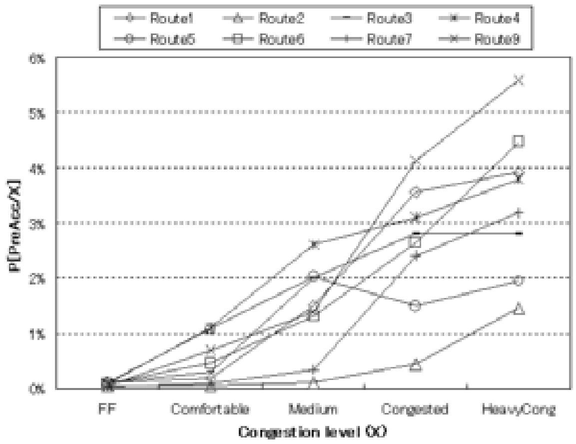

Figure 5: Probability of observing a pre-accident condition given the congestion level

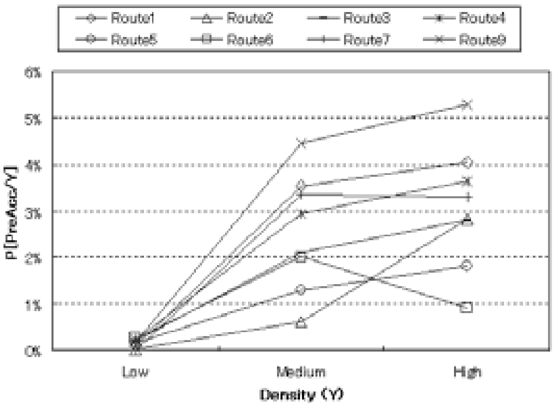

Figure 6: Probability of observing a pre-accident condition given the density level

High density reflects high congestion levels. When occurrence of accidents is increased with the increase of density levels, as shown in Figure 6, it reflects that probability of accident occurrence is increased with congestion levels. Therefore both Figure 5 and Figure 6 describe that the accident occurrence is increased with congestion levels.

Table 2 compares the congestion levels and density levels when a pre-accident condition is evidenced and non-pre-accident condition is evidenced.

Table 2: Conditional probabilities of CI and density if pre-accident condition is observed

	Route1	Route2	Route3	Route4	Route5	Route6	Route7	Route9
P(CI=Cong OR HeavyCong/PreAcc)	74%	64%	60%	64%	65%	71%	85%	87%
P(Density=High/PreAcc)	39%	33%	21%	18%	42%	23%	76%	70%
P(CI=Cong OR HeavyCong/None)	8%	8%	17%	14%	20%	10%	4%	19%
P(Density=High/None)	4%	1%	6%	4%	12%	12%	4%	15%

When a pre-accident situation is evidenced, the probability that the road section is in congested conditions is in the range of [60%, 87%]. And compared to that, when none (a non-pre-accident condition or any other incident) is observed the probability that the road section is congested is pretty low, i.e. in the range of [8%, 20%]. This reflects that when an accident occurred most of the time road section was congested.

These results probably describe the drivers' behavior at congested situations. On congested roads, where stop-and-go conditions are prevailed, frustrated drivers try to accelerate resulting rear-end collisions. And some drivers may try to change their lanes resulting side collisions plus back-up collisions. However, it should be kept in mind that fatalities are less in congested road sections compare to un-congested sections where operating speeds are high [10)].

5. Summary

(1) Conclusions

Flow, speed, occupancy and event data on inbound direction of 8 radial routes of MEX, collected over one week period, were used to explore the impact of congestion on occurrence of accidents. BBNs were used to model the relationships among pre-accident conditions and other variables. It was found that when a road section is getting congested, i.e. when the CI is increased and density is increased, the chance for an accident could be increased. Analysis of accident types will more clearly provide the answer for why probability of accident occurrence increases with congestion level.

(2) Further research

These results may not totally describe the mechanism of occurrence of accidents and congestion on MEX. Broad analysis is needed, which considers all factors that may cause accidents and congestion such as weather, road geometry, driver characteristics etc., to get a better picture of pre-accident conditions.

Although the probability of occurring a pre-accident situation is increased with the increase of congestion, accidents occur under congested situations are not fatal compared to accidents occur under high-speed situations. Here in this analysis, severity of accidents and types of accidents were not considered due to lack of data. Therefore, it is interesting to extend the analysis to consider the severity levels as well as types of accidents, too.

6. Acknowledgement

All traffic measurement data for this study was provided by the International Traffic Database (ITDb). Data available to the public can be accessed through www.trafficdata.info.

References

1) Ceder, A.: Relationships between road accidents and hourly traffic flow. II. Probabilistic approach, Accident Analysis and Prevention, 14 (1), pp.35–44, 1982.

2) Chickering, D.: The WinMine Toolkit. Redmond, WA., Microsoft technical report no. MSR-TR- 2002-103, 2002.

3) Dickerson, A., Peirson, J., Vickerman, R.: Road accidents and traffic flows: An econometric investigation, Economica, London School of Economics and Political Science, Vol. 67(265), pp. 101-21, 2000.

4) Golob, T. F. and Recker, W. W.: Relationships Among Urban Freeway Accidents, Traffic Flow, Weather and Lighting Conditions, Institute of Transportation Studies, University of California, Berkeley, 2001.

5) Kadie, C. M., Hovel, D., and Horvitz, H.: MSBNx: A Component-Centric Toolkit for Modeling and Inference with Bayesian Networks, Microsoft Research Technical Report MSR-TR-2001-67, 2001.

6) Miska, M., *et al.*: The International Traffic Database Project, 14th World congress on intelligent transport systems, CD-ROM, 2007.

7) Milburn, N. J. *et al*: Mining for information in accident data, Final report, Federal aviation administration, Civil Aerospace medical institute, Oklahoma city, 2006.

8) Noland, R.B., Quddus, M.A.: Congestion and safety: a spatial analysis of London, Transportation Research Part A: Policy and Practice 39 (7-9), pp. 737-754, 2005.

9) Taylor, M.A.P., Woolley, J.E., Zito, R.: Integration of the global positioning system and geographical information systems for traffic congestion studies. Transportation Research Part C: Emerging Technologies 8 (1-6), pp.257-285. 2000.

10) Taylor, M. C.: The speeds of vehicles which are involved in fatal accidents, Traffic Engineering and Control 42(2), 2001.

11) Wang, C., Quddus, M. A., Ison, S. G.: Impact of traffic congestion on road accidents: A spatial analysis of the M25motorway in England, Accident Analysis and Prevention, 2009.

12) Wikipedia contributors: Bayesian network, Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=Bayesian_network&oldid=293220919 (accessed May 30, 2009).