An Intelligent Transportation System for Accident Risk Index Quantification

Andreas Gregoriades[1], Kyriacos Mouskos[2] and Harris Michail[3]
[1] Department of Computer Science and Engineering, European University Cyprus, Nicosia, Cyprus
[2] Cyprus Transport and Logistics Ltd, Nicosia, Cyprus
[3] Department of Electrical Engineering and Information Technology, Cyprus University of Technology, Limassol, Cyprus

Keywords: Bayesian Networks, Dynamic Traffic Assignment, Road Safety.

Abstract: Traffic phenomena are characterized by complexity and uncertainty, hence require sophisticated information management to identify patterns relevant to safety and reliability. Traffic information systems have emerged with the aim to ease traffic congestion and improve road safety. However, assessment of traffic safety and congestion requires significant amount of data which in most cases is not available. This work illustrates an approach that aims to alleviate this problem through the integration of two mature technologies namely, simulation-based Dynamic Traffic Assignment (DTA) and Bayesian Networks (BN). The former generates traffic flow data, utilised by a BN model that quantifies accident risk. Traffic flow data is used to assess the accident risk index per road section and hence, escape from the limitation of traditional approaches that use only accident frequencies to quantify accident risk. The development of the BN model combines historical accident records obtained from the Cyprus police and domain knowledge from road safety.

1 INTRODUCTION

Road safety constitutes a problem of paramount importance worldwide (Bartley, 2008). To deal with this problem, intelligent transportation systems (ITS) have emerged. ITS are also used in the following areas: congestion control, mobility enhancement, delivering environmental benefits, and boosting productivity and expanding economic and employment growth. The work presented herein describes a novel approach and tool for assessing the accident risk index of road networks. This prerequisites the assessment of accident risk. According to (Zheng, 2009), accident risk models are divided into two categories: social risk models, that measure probabilistic (frequentist) collective damage, and individual risk models, that measure probabilistic (frequentist) individual damage. These are categorized into aggregate and disaggregate methods. The former, use global statistics and the former specific events (Bartley 2008). However, predicting accident risk requires not only frequencies of crashes per road section but also traffic flow data. However, in most cases traffic flow and accident data cannot be found together. To that end authorities perform safety analysis using only crash data which is an approximate approach to accident risk estimation. This paper aims to address this problem through the development of a novel Intelligent Traffic Information System (ITIS) that leverages the capabilities of two mature methodologies namely simulation-based Dynamic Traffic Assignment (DTA) embedded in the VISTA simulator (Ziliaskopoulos et al., 1996) and Bayesian Belief Networks (BN). The former is widely used in transportation planning and operations to predict drivers' decisions (where and when to travel on the road network), and in work was used to estimate traffic flow conditions for each road section. The latter is a powerful uncertainty modelling technique used for the quantification of accident risk under varying conditions.

The paper is organised as follows. Next section describes the methodology. Subsequent sections concentrate on data pre-processing and BN model development. The integration of VISTA with the BN along with the results that emerge from the amalgamation of the two technologies in an ITS, is described next. The paper finishes with conclusions.

Gregoriades A., Mouskos K. and Michail H..
An Intelligent Transportation System for Accident Risk Index Quantification.
DOI: 10.5220/0003989203180321
In *Proceedings of the 14th International Conference on Enterprise Information Systems* (ICEIS-2012), pages 318-321
ISBN: 978-989-8565-10-5

2 METHODOLOGY

The Road Safety Assessor, ITS system proposed herein is the amalgamation of probabilistic risk assessment with a mesoscopic traffic simulation, namely VISTA. The need for this integration boils down to the limitations of traditional traffic information systems that mainly concentrate of data warehousing. The methodology proposed utilises data marts to generate projections of future system behaviour. To that end, intelligent information management techniques are employed to distil knowledge used to develop models that enable the prospective system behaviour. The two models that emerged from this process are the accident risk assessment model and the traffic simulation model. The accident risk assessment employed is causality-based and uses BN. In BN each node is used to represent a random variable that has been identified to have a causal influence on accident risk. Each directed edge represents an immediate dependence or direct influence between parent and child variables (Jensen, 2001). Evidence is entered in the model through instantiation of leaf node on the model. Inference is achieved by belief propagation through the models topology. BN technology is used to model how traffic and infrastructural factors influence accident risk. The second component of the approach is a road traffic simulator based on DTA. The DTA model is used in VISTA through the Dynamic User Equilibrium (DUE) model (Peeta et al., 2000). The use of DTA model enhances the limitations of existing practices by providing a consistent way of producing estimates of traffic flow conditions of road networks using limited information from traffic flow detectors. Moreover, it produces timely and complete traffic volume estimates for all sections of a road network and hence, can be used to assess accident risk using time varying conditions. The integration of BN with VISTA in the proposed traffic information system enables the dynamic assessment of accident risk using simulated traffic conditions and prior knowledge embedded in the BN. A pilot study conducted with the system aimed to assess the safety performance of the Nicosia road network in Cyprus and to investigate how it will behave under different scenarios.

Initially the road traffic model of Nicosia was specified, implemented, verified and validated in VISTA. Models in VISTA are represented by nodes connected by unidirectional links that represent flow of traffic in one direction. It is possible to have more than one link between two nodes to indicate separate lanes and lane direction. The completed VISTA simulation model was integrated with an accident risk assessor implemented in Java. The simulator provided the risk assessor with the traffic volumes of all road sections of the network for every 15 min interval. Traffic volumes along with infrastructural properties of the network were used by the BN to assess accident risk on a simulation step basis. For the development of the BN topology and the parameterization of its prior knowledge, historical road accident data were utilized.

3 ARCHITECTURE OF THE ITS

The Road Safety Assessor tool emerged from the integration of VISTA with BN technologies. The main components of the tool are: the BN engine, the accident risk assessor, the VISTA simulator, the data pre-processor that incorporates the scenario generator, the results analyzer and the visualizer. The tool was developed using a component-based software engineering methodology. With the initial specification of the system requirements captured, we proceeded in the identification of suitable software components that matched the initial system requirements. These components were subsequently integrated to implement parts of the system's functionality. In particular the Bayesian inference engine and the visualization components were selected after thorough investigation. The glue-code that enabled components integration was implemented in Java. The risk assessor quantifies accident risk using a Bayesian inference engine that utilizes the probabilistic model of accident risks. Input to the BN assessor is categorized into static and dynamic. The former is obtained from the VISTA database and the latter is the output of the VISTA simulation.

Input to the accident risk assessor is organized in the form of scenarios. An input scenario to the BN assessor is defined by the static and dynamic properties of each road section. Static information is obtained from the VISTA database and in combination with the dynamic input from the simulator. This provides the baseline for generating a number of plausible test scenario variations for each road section. Generated scenarios are executed by the risk assessor to quantify the probability of accident. The scenario generator is responsible for generating plausible scenario variations to stress-test the safety performance of each road section. The visualizer processes the results and depicts these to the user graphically. Input scenarios are executed by the BN model. Each scenario evidence is propagated down the BN topology to produce the posterior probability of accident risk per scenario.

The integration of the VISTA with the BN model was realized through asynchronous data interchange. To establish communication between VISTA and the risk assessor it was imperative to pre-process VISTA's output data prior to being utilized by the BN in the risk assessor. Specifically, VISTA variables are continuous by nature, hence, had to be converted into categorical/discrete to be processed by the BN model, since it uses only discrete nodes. Hence, it was necessary to discretize the output from VISTA prior to instantiating the BN model. For the discretization process it was necessary to refer to domain experts that specified the cut-off values for each variable. Specifically, for traffic volume three states were defined, namely, low, average and high. The first corresponding to less than 100 vehicles per 15 time interval, the second to less than 350 and the last to greater than 350.

4 BN MODEL DEVELOPMENT

Development of BNs requires the specification of the topology and the conditional probability tables. To that end historical accident records were obtained from the traffic safety department of the Cyprus Police. Preliminary compilation of the data was performed with the SPSS statistical package to reduce the dimensionality of the data. The accident dataset covered all accidents occurred in the Nicosia area from 2002 until 2008 and comprised over 9000 records. Each record consisted of 43 (six continuous and 37 categorical) input parameters covering global, local, temporal, accident, driver and car characteristics collected at the site of the accident by the police officers, eye witnesses and the involved parties. Each record was associated with a single categorical output parameter pertaining to accident severity, namely light, severe and fatal, as evaluated by the police officer at the site of the accident.

However, for the development of the BN model topology it was imperative to enhance the dataset with additional information regarding the traffic conditions of each accident record from VISTA simulation. Therefore each accident record was mapped on a geospatial GIS platform and subsequently import on VISTA to obtain the dynamic information of each accident location at different time intervals. This yielded an enhanced dataset of accident records.

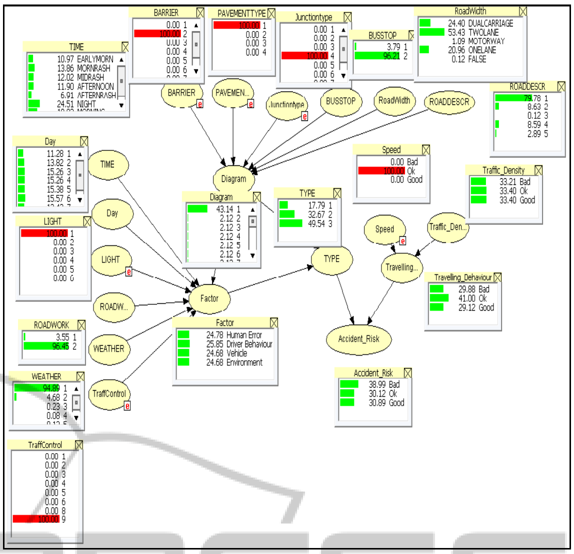

Figure 1: Data inferred BN topology at execution time.

A preliminary analysis of the dataset provided a generic indication of the influence of each variable to road accident risk. Data pre-processing was performed in two steps (a) replacement of missing and erroneous parameter values by the mean value, and (b) grouping related values of multi-valued categorical parameters so as to have a manageable number of states per parameter. Next, to reduce the dimensionality of the dataset, Principal Component Analysis (PCA) was used. This helped to identify the core variables of the model. Results from the dimensionality reduction using PCA, yielded 19 variables for the BN topology. The topology depicted in Figure 1, was learned from processed dataset using the Expectation Maximisation algorithm (Jensen, 2001). Figure 1 also shows an instantiation of the BN model in Hugin researcher tool. The developed ITS utilises the Hugin engine using its API. Each variable in this figure is accompanied by a monitor window that shows its states. The input evidence is showed as a solid bar in the monitor window of each variable. Collectively all variable instantiations correspond to one scenario variation that is provided by the scenario generator component of the tool that uses input from VISTA. In each scenario variation variables that are not instantiated using input from VISTA is varied systematically to produce additional scenario permutations that instantiates the BN model.

To estimate the accuracy of the developed BN model, validation was performed using the accident dataset obtained from the police. The dataset was utilised to identify locations on the network with high accident frequency. These are the networks black spots. These points were used to validate the model after it was implemented. Specifically, a subset of the accident dataset was used to validate the system. Black-spots that were identified using the dataset, were used to test the BN accuracy under varying conditions of traffic flow data.

5 RESULTS

Results from the accident risk assessor were used to calculate the accident risk index (ARI) of each road section. BN scenarios for each road segment were labeled accident prone if the BN accident risk probability was above a pre-specified threshold value. BN scenarios that fell below the threshold value were ignored. Scenarios were defined on the fly by the scenario generator component. Each segment is evaluated against scenarios that describe traffic condition at different time intervals and driver profiles. To assess the ARI it was imperative to normalize the number of accidents that were predicted by the BN with the traffic volume per time interval, for each road section. To that end, the developed system uses a systematic approach that utilizes the traffic volume estimates from the VISTA simulation and the accidents predicted using the BN risk assessor. Traffic volume acts as a normalizing factor for the number of accidents predicted using the BN risk assessor. In this study, the ARI is defined as:

> Accident Risk Index (ARI) = Number of accidents predicted by the BN/estimated traffic flow rate per time period of the day, from DTA

ARI results gave rise to road sections that inherently have safety issues. These are the network's black spots. An illustration of the preliminary results produced by the method is depicted in Figure 2. This figure illustrates a subset of the results and indicates that sections with IDs, 3, 21 and 47 have the highest ARI.

6 CONCLUSIONS

The ITS system described herein illustrates a novel approach to quantifying road safety using probabilistic inference expressed in causal relationships between factors leading to accidents with DTA simulation. The method escapes from the problem of traffic data shortage through the use of DTA simulation. VISTA provides complete traffic volume data estimates for all road sections of the network on a 24 hour basis. This constitutes advancement over existing methods that base their analysis on limited data obtained from a scarce number of traffic sensors on the network.

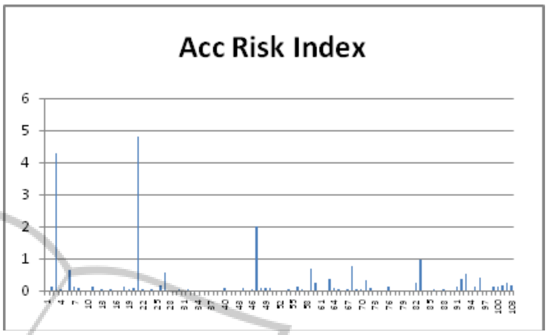

Figure 2: All road section with their ARI values (Y axis).

REFERENCES

Bartley, P., 2008. Traffic Accidents: Causes and Outcomes. *Nova*.

Florian, M., Mahut, M., Tremblay, N., 2008. Application of a simulation-based dynamic traffic assignment model. *European Journal of Operational Research*, 189, 1381–1392.

Jensen, F., 2001. Bayesian Networks and Decision Graphs. *Springer*.

Peeta, S., Ziliaskopoulos, A., 2001. Foundations of Dynamic Traffic Assignment: The Past, the Present and the Future. *Networks and Spatial Economics*, 1 (3/4), 233-65.

Zheng, X., Liu, M., 2009. An overview of accident forecasting methodologies. *Journal of Loss Prevention in the Process Industries*, 22(4), 484-491.

Ziliaskopoulos, A., Lee, S., 1996. A Cell Transmission Based Assignment-simulation Model for Integrated Freeway/Surface Street Systems. *Proc., 75th Transportation Research Board*, Annual Meeting, Washington, DC.