

Bayesian Networks Probabilistic Safety Analysis of Highways and Roads

Elena Mora^(✉), Zacarías Grande, and Enrique Castillo

Department of Applied Mathematics and Computational Sciences,
University of Cantabria, 39005 Santander, Spain
elenamoravil@gmail.com, zgandrade@gmail.com, castie@unican.es

Abstract. A probabilistic safety analysis methodology based on Bayesian networks models for the probabilistic safety assessment (PSA) of highways and roads is presented. The main idea consists of (a) identifying all the elements encountered when travelling the road, (b) reproducing these elements by sets of variables, (c) identifying the direct dependencies among variables, (d) building a directed acyclic graph to reproduce the qualitative structure of the Bayesian network, and (e) building the conditional probability tables for each variable conditioned on its parent nodes. Since human error is the most important cause of accidents, driver's tiredness and attention are used to model how the driver's behaviour evolves with driving and how it is affected by the environment, signs and other factors. A computer program developed in Matlab implements the Bayesian network model from the list of road items and a set of parameter values given by a group of experts. In this way, the most critical elements can be identified and sorted by importance, thus, an improvement of the global safety of the road can be done savings time and money. The proposed methodology is illustrated in real examples of a Spanish highway and a conventional road.

Keywords: Bayesian networks · Road safety
Probabilistic safety analysis

1 Introduction

One of the main concerns of our society today is the great number of accidents that occur on roads. Thus, the improvements on safety traffic analysis on roads are considered very useful. The study of this work focuses on Probabilistic safety analysis (PSA) since it is well-known and the best technique to asses the safety level of a system. It has been widely used for nuclear power plants and in recent years it has been also employed for railway lines (see for example [1, 2]).

The analysis presented here is especially based on Bayesian networks (BN) because they have a great power to reproduce multidimensional random variables and present numerous advantages against other existing methods such as fault or event trees. However, there are few analysis of traffic infrastructures based on them. Some of the studies applying BN are for example [3–6].

In this work a Bayesian network for probabilistic safety assessment of highways and roads is proposed and described.

2 Bayesian Network Proposed Model

In this section the Bayesian network proposed model is introduced and the steps of its building are detailed. A Bayesian network has two main components: (a) an acyclic graph which define the qualitative structure of the multidimensional variables and (b) the conditional probability tables which quantify the Bayesian network. In order to define the qualitative structure here, is used the subjective method, based on our knowledge of the problem since the existing data are not adequate to provide this information.

The first steps to build the Bayesian network are the selection of the most relevant items that have influence on traffic safety being likely to cause incidents, and the identification of the variables related with these items. In order to identify all possible items, such as curves, stop signals, intersections, roundabouts, tunnels, acceleration or deceleration lines, traffic light signals, and pavement failures, a video is recorded from the start to the end of the road (see [7]). The list of the chosen variables for this model with their definitions and their possible values are shown in Tables 1 and 2 depending on they are related to the driver, infrastructure or the incident respectively.

Table 1. List of variables related to the driver and the infrastructure with their possible definitions and values.

V. Related to the driver	Definition	Values
D: Driver’s attention	This variable represents the driver’s attention	Distracted, Attentive, Alert
T: Driver fatigue	Measures driver fatigue	A positive value that increases with driving time
Sd: Driver decision on speed	Represents the action of the driver in cases where he must adjust the speed	Correct, Error I, Error II
Dri: Driver type	This variable reflects the quality of the driver	Professional, Experienced, Standard, Bad
AS: Decision of the driver at a traffic light	Represent the decision of the driver at a traffic light	Correct, Error I, Error II
DS: Decision of the driver in a signal	Represents the decision of the driver in a signal, such as Stop, yields or speed limit signals	Correct, Error
V. Related to the infrastructure	Definition	Values
It: Traffic intensity	This variable measures traffic intensity	Light, Medium or Dense
Vis: Visibility	This variable measures the visibility existing at the point considered	Good, Average and Poor
Vt: Type of vehicle	Refers to the type of vehicle	Heavy Vehicle, Automobile and Moto
S: Speed	Is the circulation speed at the point considered	Set of positive values
SS: Signal status	Represents the status of the signal	Green, Yellow, Red
TF: Technical failure	It represents the possible failure of a vehicle, signal, etc.	Yes, No
E: Failure in the environment	It represents the possibility of unwanted events, such as obstructions of the road by stones, trees or other materials, animals as well as defects of the road, clearings, embankments, etc	No fault, Minor fault, Medium fault or Serious fault

Table 2. List of variables related to the incident with their possible definitions and values.

V. Related to the incident	Definition	Values
V: Vehicle failure	Consider the possibility of a vehicle failure	No fault, Minor fault, Medium fault or Serious fault
P: Pavement condition	Represents the state of the pavement	Good condition, Mild failure, Medium failure or Serious fault
Co: Collision	Represents the possibilities of collision with other vehicles that circulate on the road in the same or opposite direction	No Collision, Mild Collision, Medium Collision, or Severe Collision
W: Weather	Represents the type of climate	Fair, Rain/Snow, Wind, Fog, Snow/Ice
CF: Crossover Frequency	Represents the frequency of crossing between two elements of brakes such as vehicles and trains or pedestrians and vehicles at a specific point on the road, such as pedestrian crossings or level crossings	Yes, No
I: Incident	Represents possible incidents that may occur at a particular location on the road or in a no-signal range	Incident, Mild Incident, Medium Incident, and Serious Incident

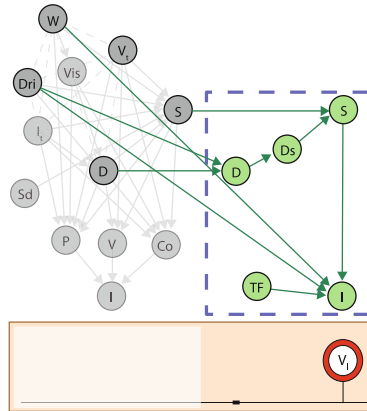


Fig. 1. Sign sub-Bayesian network as example of the different sub-Bayesian networks.

The identification of the direct dependencies among the variables involved is the next step. In order to obtain the acyclic graph different sub-Bayesian networks with a particular structure and associated variables with their corresponding conditional probabilities have been modelled. As a representative example, the sign sub-Bayesian network which is used at locations where some action subject to error is required, is shown in Fig. 1.

Finally, the conditional probability tables are defined using closed formulas.

Once these four steps are covered, all the information needed to calculate the marginal probabilities of the incident types nodes and the ENSI¹ values for each location become available.

¹ ENSI refers to the expected number of equivalent severe incidents, where 6.4 medium incidents and 230 light incidents are considered equivalent to one severe incident.

3 Software Development and Outcomes

A computer software written in Matlab with calls to Latex, JavaBayes and BNT software has been developed by our group to obtain the results for this model. As data input it is used a sequential description of all the items encountered when travelling along the highway. The computer program checks the given data for errors, builds the acyclic graph of the Bayesian network, builds the conditional probability tables, calculates the incident probabilities, evaluates the ENSI of incidence nodes, provides a table of the ENSI frequencies for all items, sorts ENSI values by importance, provides the expected number of incidents of each item type, plots the segments Bayesian subnetworks, gives the ‘JavaBayes’ code, gives the ‘BN’ code, and provides a report file.

In order to display the results in a useful way, the following elements, shown in Fig. 2 have been included divided by segments: (A) the segment acyclic graph of the Bayesian network, (B) the graphical representation of traffic signs and track elements, (C) the segment characteristics, and (D) a cumulated risk chart where the reader can easily identify the relative importance of the different items by comparing the discontinuities (jumps) of the graph. The whole line representation has been divided into short segments, with the aim of obtaining a more detailed information. In order to obtain a global view of the roads assessment some plant sites, where the riskiest points are allocated and differentiated by severity level, are represented, as well as, different types of tables to facilitate the understanding of the results of the probabilistic safety analysis. In the following Section, representative examples of this graphical information provided are illustrated.

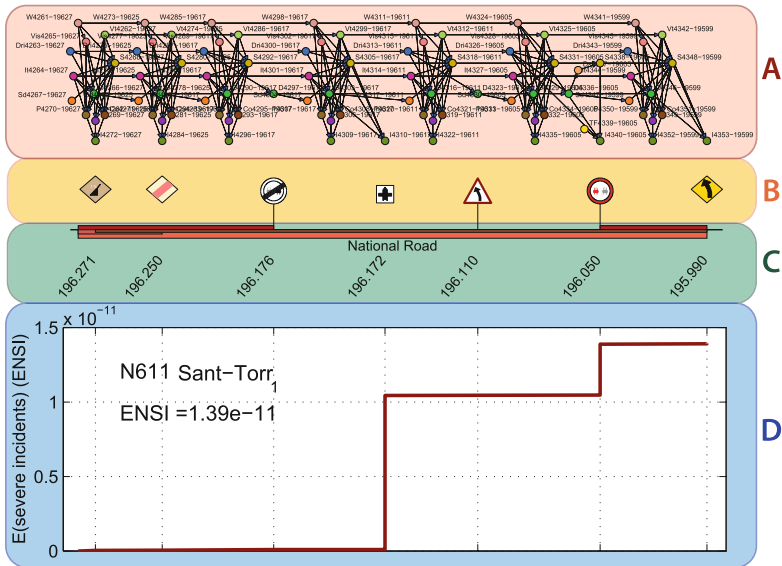


Fig. 2. Example of the information supplied by the computer program.

4 Examples of Application

The usefulness of the model can be appreciated in the next examples. In the first, the National N-634 road from (Kilometer Point, KP) 260.945 to KP 264.870 is considered, with an Annual Average Daily Traffic (AADT) of 2145 vehicles. In Fig. 3 the acyclic graph of the segment between KP 262.180 and KP 262.276 is represented. As commented before, the relative importance of the different

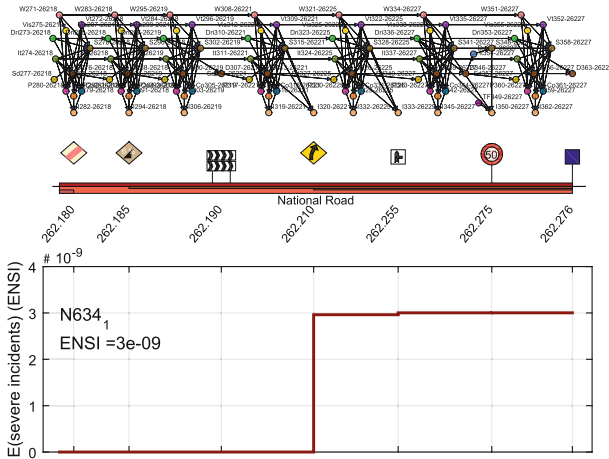


Fig. 3. Example of the segment between KP 260.945 to KP 264.870 of the N-634 national road without improvements.

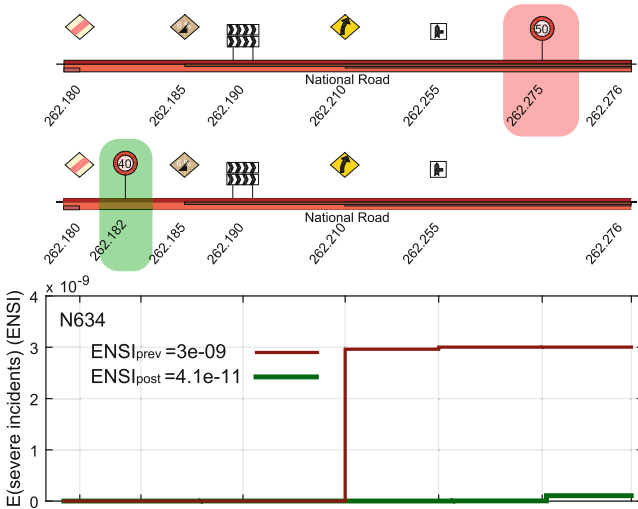


Fig. 4. Example of the segment between KP 260.945 to KP 264.870 of the N-634 national road with proposed improvements.

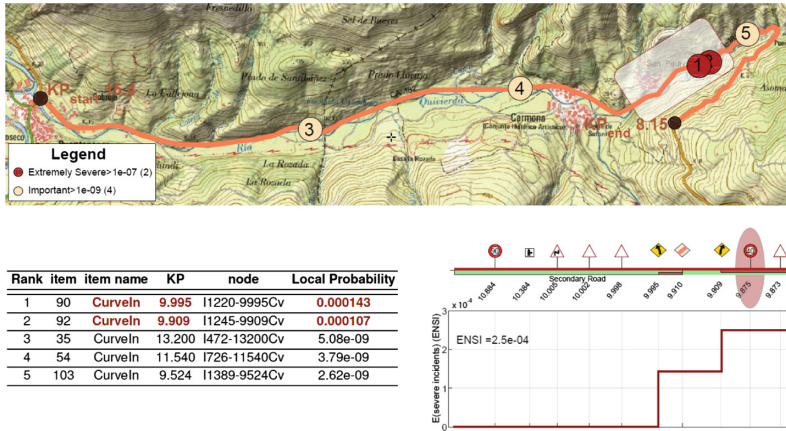


Fig. 5. Example of the CA-182 secondary road initially.

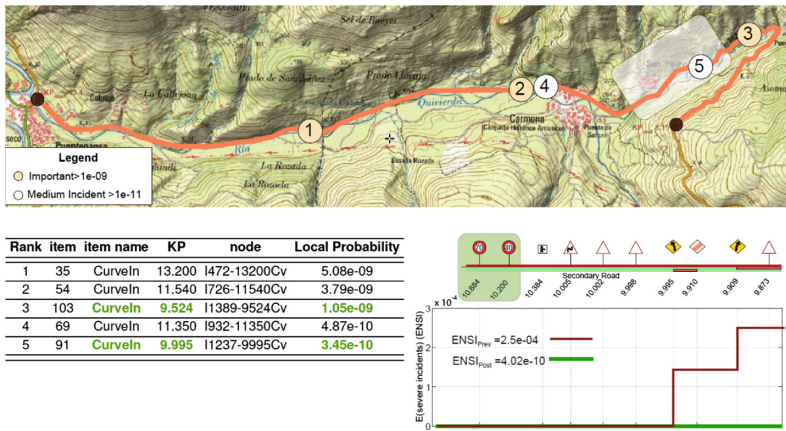


Fig. 6. Example of the CA-182 secondary road with proposed improvement.

items can be identified by comparing the discontinuities in the graph. It can be easily seen that curves and lateral entries are the most critical items. The worst segment of the line is placed between KP 262.210 and KP 262.416, thus, the speed limit sign (50 Km/h) of KP 262.275 have been reduced to 40 km/h and have been moved to KP 262.184. In Fig. 4 how this change improves significantly the resulting ENSI can be seen.

In the second example, shown in Fig. 5, the CA-182 secondary road trace from KP 9.801 to KP 15.250 with an AADT of 558 vehicles is represented. Here, the solution for the misplacement of a sign (speed limit signal 40 km/h at KP 9.875), which generated two items with a significant high incident probability, was to move the sign from its original location to KP 10.200. In Fig. 6 how the risk decrease significantly and the reduction of the ENSI can be appreciated.

5 Conclusions

Based on the model and outcomes presented in this paper the following conclusions can be drawn:

1. Bayesian network models are very useful to reproduce and to perform a probabilistic safety assessment of highways and roads.
2. This model allows to identify the most dangerous points of any road.
3. It is possible to determine the most frequent events that produce accidents.
4. The real examples commented in this paper show that the method can identify relevant incidents and quantify their probabilities of occurrence.
5. The proposed methodology allows to take corrective measures to the safety problems of the road and to predict accident concentration zones before they occur.
6. Applying this model resources for improving safety and maintenance could be optimized.
7. A lot of work still needs to be done in the future in this direction above all about the parameter estimation. The collaboration of diversity groups of experts would improve the power and the efficiency of the method.

References

1. Castillo, E., Grande, Z., Calviño, A.: Bayesian networks based probabilistic risk analysis for railway lines. *Comput. Aided Civil Infrastruct. Eng.* **31**, 681–700 (2016)
2. Castillo, E., Grande, Z., Calviño, A.: A Markovian-Bayesian network for risk analysis of high speed and conventional railway lines integrating human errors. *Comput. Aided Civil Infrastruct. Eng.* **31**, 193–218 (2016)
3. Pawlovich, M., Li, W., Carriquiry, A., Welch, T.: Experience with road diet measures: use of Bayesian approach to assess impacts on crash frequencies and crash rates. *TRR: J. Transp. Res. Board* **1953**, 163–171 (2006)
4. Persaud, B., Lyon, C.: Empirical Bayes before-after safety studies: lessons learned from two decades of experience and future directions. *Accid. Anal. Prev.* **39**, 546–555 (2007)
5. Deublein, M.K.: Roadway accident risk prediction based on Bayesian probabilistic networks. ETH Zürich. Ph.D (2013)
6. Deublein, M., Schubert, M., Adey, B.T., de Soto García, B.: A Bayesian network model to predict accidents on Swiss highways. *Infrastruct. Asset Manag.* **2**, 145–158 (2015)
7. Grande, Z., Castillo, E., Mora, E.: Highway and road probabilistic safety assessment based on Bayesian network model. *Computer Aided Civil and Infrastructure Engineering* (2017, to appear)