

An intelligent decision support system for bridge safety assessment based on Data Mining models

M. F. Santos¹, H. Quintela¹, P. Cortez¹ & J. O. Almeida²

¹*Department of Information Systems/Algoritmi R&D Centre, University of Minho, Portugal*

²*School of Technology and Management of the Polytechnic Institute of Viana do Castelo, Portugal*

Abstract

Bridges are one of the primary infrastructures in our society. During the life cycle of a bridge structure, the service conditions should be evaluated on a regular basis in order to assure the necessary levels of strength and durability. Taking into account (i) the social-economic importance of bridges' use, (ii) the necessary safety assurance and (iii) the high costs of any physical intervention, there is a need for continuous online bridge monitoring, for investment and use optimization. Recently, smart structures, which combine remote sensors (which send a stream of time series data) with intelligent information systems for real-time decision support through embedded Data Mining (DM) models, have been proposed to handle this task. Indeed, the application of DM techniques to analyse civil engineering data has gained an increasing interest in recent years, due to intrinsic characteristics such as the ability to deal with nonlinear relationships.

In this study, Artificial Neural Networks have been used to predict the following ratios: global efficiency, structural adequacy and safety, serviceability, essentiality for public use, and special reductions, using a ratio-based framework, and data collected during inspections of bridges in the north of Portugal. In particular, the global efficiency ratio is very useful to identify intervention priorities and to schedule the repair, strengthening and rehabilitation needs.

The obtained results are encouraging and the most accurate model for global efficiency presents a low error (Root Mean Squared Error of 0.149).

This approach opens room for the development of intelligent decision support systems for Bridge Management Systems. These systems are being recognized as a good way to systematize all the management process and to minimize the ratio cost/benefit during the bridge lifetime.

Keywords: Data Mining, knowledge discovery from databases, intelligent information systems, intelligent decision support systems, artificial neural networks, civil engineering structures.



1 Introduction

Recently, there were two major disasters concerning the fall of bridges in Portugal. Since then, structure experts are trying to develop a new Bridge Management System (BMS) to be applied by the authorities in order to assure the safety of bridges' use [2].

A BMS is a computerised system that supports the following activities: structures inventory information; inspection; assessment; maintenance bids, prioritisation and allocation; works data; network structures condition monitoring; planning and forecasting; databases [3]. The main objectives of a BMS are helping the administrations to assure public safety and risk reduction, and preservation of investment to produce budget, maintenance and program policies. One of the most important roles of BMS is to provide a systematic procedure for anticipating bridge maintenance, repair and rehabilitation needs and the prioritization of these activities.

BMS systems currently face the great challenge of balancing the available funds with the increasing needs for bridge maintenance, repair and rehabilitation.

In order to assure high levels of security, it is necessary to deploy systems that quickly gather the necessary structure related data, allowing the early identification of damage, in order to plan the necessary interventions [4].

The interest in the development of BMS is increasing and nowadays they are based on life cycle costs analysis to optimize maintenance costs.

The knowledge of how the bridges deteriorate is essential in those systems, and the forecasting models for structure evolution are being applied in a theoretical basis on several structural components, using probabilistic or deterministic approaches. However, there is no deterministic law to predict the deterioration of a bridge in general. The Markov's Model for deterioration forecasting is used in the most parts of BMS Systems (e.g. Pontis [5], Bridgit [6], Danbro + [7]).

A Markov model [8] assumes that a system may occupy one of a finite number of states ($State_1, State_2, \dots, State_n$) at different moments in time (T_1, T_2, \dots, T_n), being the probability of occupying a state determined solely by recent history. More specifically, a first-order Markov model assumes that these probabilities depend only on the preceding state. Markov analysing techniques are applied in a wide range of engineering applications. A Markov process has the property that the bridges state at the inspection t depending only on its previous state at the inspection $t-1$. With this assumption, the present state is the only one that is taken into account to determine the future of the bridge.

In this paper an alternative way to solve the elicited problems is proposed: an Intelligent Decision Support System (IDSS) for Bridge Safety Assessment using Artificial Neural Networks models and a ratio based framework. As noted by [9] the development of IDSS incorporating models generated by Knowledge Discovery in Databases (KDD) processes and implemented as Agent-Based Systems is an interesting way to solve complex and dynamic problems, such as bridge safety assessment.



The paper is organized as follows: first, the basic concepts are introduced and the data is presented and described; then the applied DM techniques are presented; next, the experiments are described, the results analysed in terms of several criteria; then, the IDSS architecture is proposed; finally, closing conclusions are drawn and future research activities are depicted.

2 Materials and methods

The dataset used in the experiments considers 194 inspections of bridges located in the North of Portugal, with different ages, materials, structural functioning, dimensional characteristics, and use demands. Each inspection registry includes 165 variables, such as: accessibility; location; historical/cultural and strategic importance; bridge length; number of traffic ways; average day traffic; bridge deck state; bridge's deck geometry evaluation; erosion effects; number of previous inspections; and conservation state. The following ratios were also recorded: global efficiency (GER), structural adequacy and safety (R_1), serviceability (R_2), essentiality for public use (R_3), and special reductions (R_4) (Table 1).

Structural safety tends to be higher when R_1 has a higher value, and that is explained by the relative importance of the ratio in the final value of GER . The GER value tends to a maximum safety as R_1 , R_2 and R_3 have higher values, and R_4 presents a lower value.

For example, a bridge that presents a null value for R_1 , has a lower value for R_3 and R_4 , and the value of GER will be around 30%.

Table 1: The bridge ratios variables and relative importance for Global Efficiency value.

Ratio	Domain	Importance (%)
Structural and adequacy safety	[0,...,0.98]	55%
Serviceability	[0,...,0.3]	30%
Essentiality for public use	[0,...,0.15]	25%
Special reductions	[0,...,0.08]	-

The GER can assume values between 0 and 1 (0% and 100%), where a lower value means a high level of risk and possible collapse of the structure, which implies an urgency in interventions (Figure 1). According to the data analysis and experts' experience, R_1 is the ratio that has more importance in the GER final value (Table 1).

In the experiments carried out, several approaches were followed for ratio prediction; the SAS Enterprise Miner tool [10] was adopted to conduct the whole KDD process. This tool provides an easy management of all KDD steps, including the use of Artificial Neural Networks.

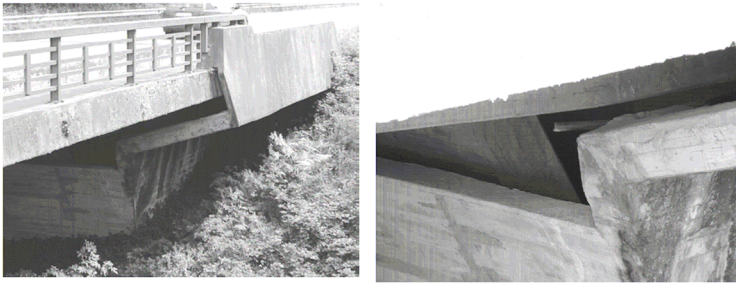


Figure 1: Bridge presenting a *GER* value of 48% [3].

2.1 KDD and Data Mining

2.1.1 Pre-processing

In the pre-processing phase of the KDD course, it was identified a total of 78 variables containing missing values in different rows and with different combinations. To solve this issue several approaches could be considered, such as filling values using expert advice, imputation with average or most frequent values or deletion of the rows containing the missing attributes.

Since the number of rows is not high, it was decided to use just the attributes related to deterioration. The final pre-processed dataset (for the DM phase) contains 138 rows and 20 columns, with no missing values.

2.1.2 Data Mining

The aim of the DM step is the induction of predictive models for the following ratios: global efficiency, structural adequacy and safety, serviceability, essentiality for public use, and special reductions, using the Artificial Neural Networks (ANN) technique. ANNs are connectionist models inspired in the behaviour of the human brain. In particular, the Multi-Layer Perceptron (MLP) is a popular architecture that will be adopted in this work. In MLPs, neurons are grouped in layers and only forward connections exist. This architecture is capable of nonlinear mappings and is a very powerful technique for dealing with complex regression problems [11].

3 Results

3.1 Framework

For the ratios prediction, it was followed a ratio-based framework. This framework includes a total of 6 distinct models, which differ in the feature selection approach. Regarding the later, three classes of models were used. The first one uses all the attributes for a prediction model with 4 outputs altogether: structural adequacy and safety, serviceability, essentiality for public use, and special reductions ratios. The second selection uses all the input attributes for the generation of 4 prediction models, each one for one output ratio. Finally, the last

one uses all the attributes for global efficiency ratio prediction. It should be stressed that in all the experiments, the ratios were not used as input attributes, because one the main objectives of this work was to develop an agile method for *Global Efficiency Ratio* prediction.

In the ANN modelling the number of hidden nodes in the hidden layer was automatically set by the SAS tool, and the accuracy was estimated using the Holdout method by randomly dividing the available data into three exclusive partitions: 40% for training, 30% for validation and 30% for testing.

3.2 Experimental results

Table 2 presents the predictive performance in validation and test sets for all the models proposed by the framework, using the Root Mean Squared Error (RMSE) measure (Equation 1). The RMSE is one of the most commonly used measures on regression tasks. This metric, which will be adopted in this work, has the property of being more sensitive to high error values.

$$\sqrt{\frac{(a_1 - c_1)^2 + (a_2 - c_2)^2 + \dots + (a_n - c_n)^2}{(a_1 - \bar{c})^2 + (a_2 - \bar{c})^2 + \dots + (a_n - \bar{c})^2}} \quad (1)$$

As the results suggest (Table 2) there are improvements in the results when each model predicts just one ratio, although this effect is not as impressive as it could be expected.

Table 2: The predictive test performance (RMSE values).

Approach	Ratio	Validation	Test
1st	Structural and adequacy safety	0,091	0,101
	Serviceability	0,057	0,052
	Essentiality for public use	0,057	0,055
	Special reductions	0,010	0,016
2nd	Structural and adequacy safety	0,075	0,115
	Serviceability	0,055	0,04
	Essentiality for public use	0.048	0.06
	Special reductions	0.099	0.01
3rd	Global Efficiency	0,122	0,149

In the first approach the global accuracy for the ANN based model with four outputs (one for each ratio) is 0.063, and the results analyses take us to try to develop one model for each ratio, trying to improve the accuracy. That approach conduces to a best RMSE value for Serviceability, Essentiality for public use and Special Reductions, but to a higher error value for Structural and adequacy safety

and Essentiality for public use ratios' prediction. However, the global improvements attained, lead us to recommend the ratios prediction with separate models, because that strategy implies that the models are more specialized, and consequently, as shown in Table 1, the results are globally better.

The moderate improvement in the results from the 1st to the 2nd approach is important, since when it is crucial to minimize the error due to safety issues. Regarding the global efficiency ratio, which is very useful to identify intervention priorities and to program the repairing, strengthening and rehabilitation needs, the results achieved are very encouraging, opening room for the use of DM based models for value prediction. The most important attributes are: inventory rating, construction material, deck condition, substructure condition, superstructure condition, and bridge roadway width.

4 IDSS system

Intelligent Decision Support Systems (IDSS) are computer based systems that support one or several phases of the individual, team organisation or inter-organisational decision making process [12]. While IDSS constitute one of the most popular areas of research in information systems in the past, a closer look seems to indicate that this interest is declining or not increasing as expected [13,14]. This is due to the requirements for decision support tools posed by new dynamic and complex environments, demanding for situated and active IDSS.

To improve the success of implementing an IDSS in a real environment for Bridge Safety Assessment based on a ratio framework, the following modules, presented in Figure 2, should be addressed:

Data Entry: capture all the data using remote sensors in the case of variables that could be measured automatically [15], and using mobile systems for inspections variables introduced by the specialists, and store them in a transactional database.

Intelligent Monitoring: this module uses the Data Mining module, because the predicting models should be invoked in a continuous way for online monitoring. In case of abnormal values for the safety ratios, a set of alerts events should be displayed, and the alert(s) should be delivered for the system managers by mobile systems. Each time the models are invoked, should also be created an entry in the knowledge database, with the values measured and the predicted values for the ratios.

Since the values monitored by the remote sensors could be measured in continuous or in short intervals of time and the intervals between inspections are medium or long, for each measure of the remote variables, it should be considered the last inspection values.

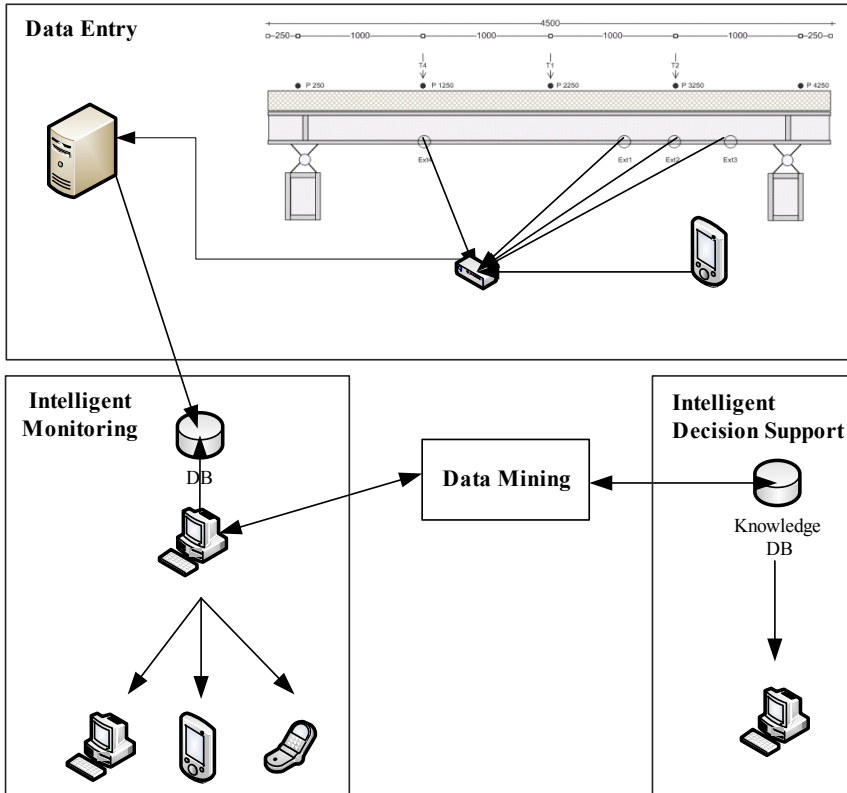


Figure 2: IDSS system architecture.

Intelligent Decision Support: this module implements the necessary functions to help managers to take decisions using the information available in the IDSS system, concerning the bridge security and interventions plans. It implements the tools for reporting and uses the Data Mining module for predicting models invocation and to overview their historic performance.

In a regular way, the Bridge Management System should deliver a report - available in the system and delivered by electronic mail for managers -, including the measured data, the predicted values, the alerts displayed, and the actions carried out, that are stored in the knowledge database. This data will also be used to monitor the history and evolution of the predicted models.

Data Mining: this module should implement mechanisms for models invocation, and functions to adapt the existing DM models with the new data collected. The results of invocation and evolution should be stored in the knowledge database, enabling the performance analysis.

5 Conclusions and further work

This paper presented a study for bridges safety ratios prediction based on a Data Mining (DM) approach. In particular, Artificial Neural Networks were used to model four safety ratios with different feature selection setups. The results obtained opening room for the automatic development of an Intelligent Decision Support System (IDSS) for bridge safety assessment.

In the future we intend to use data from multiple sources, such as traffic data, costs and predicted future conditions. It should be noted that in this work the data was collected from inspections carried out by specialists. Nevertheless, there is potential for automatic data collection by using bridge sensors. Indeed, the evolution from structures to smart structures opens room for cheap and real-time monitoring of more parameters that can be used to improve the ratios prediction and thus build more reliable Bridge Management Systems.

The use of IDSS in Bridge Management will help the specialists in day to day management. For instance, the need of interventions could be ranked in a proactive way. This would also help the authorities in adjusting the necessary funds.

Acknowledgements

This work was supported by the IN2TEC initiative of the School of Engineering of University of Minho. The authors thank Marta Vilas Boas for the work developed in the context of the SIISEC project.

References

- [1] Cruz, P., Lourenço, L., Santos, M., Quintela, H., Cortez, P., Strength of corroded tapered plate girders under pure shear, in Cruz, Frangopol & Neves (eds) Bridge Maintenance, Safety, Management, Life Cycle Performance and Cost, Taylor & Francis, 2006.
- [2] Almeida, J.O., Delgado, R., A bridge management system applied to a set of Portuguese bridges, Bridge Maintenance, Safety, Management, Life-Cycle Performance and Cost, Cruz, Frangopol & Neves (eds), Taylor & Francis Group, 2006.
- [3] Das, P.C., Maintenance Planning for Trunk Road Structures in England, http://onlinepubs.trb.org/onlinepubs/circulars/circ498/v2_H02.pdf.
- [4] Almeida, J., Gestão de Pontes Rodoviárias – Um modelo aplicável em Portugal, Msc Thesis, Faculdade de Engenharia da Universidade do Porto.
- [5] AASHTO – ASSHTOWare – Pontis, <http://aashtoware.org/>.
- [6] Hawk, H., Bridgit: user friendly approach to bridge management, Eighth International Bridge Management Conference, Transportation Research Circular, 498, <http://gulliver.trb.org/publications/circulars/circ498/circular498.pdf>.
- [7] Bjerrum, J., Jensen, F.M., Internet-based management of major bridges and tunnels using the Danbro+ system, Bridge Maintenance, Safety,



- Management, Life-Cycle Performance and Cost, Cruz, Frangopol & Neves (eds), Taylor & Francis Group, 2006.
- [8] Cappé, O., Moulines, E., Ryden, T., Inference in Hidden Markov Models, Springer Series in Statistics.
 - [9] Gago, P., Santos, M.F., Silva, A., Cortez, P., Neves, J., Gomes, L., INTCare: A Knowledge Discovery based Intelligent Decision Support System for Intensive Care Medicine, Journal of Decision Systems, Special issue on Design, Building and Evaluation of Intelligent DMSS, Lavoisier, 14(3): 241–259. 2005.
 - [10] Cerrito, P.B., Introduction to Data Mining Using SAS(R) Enterprise Miner (TM), SAS, 2007.
 - [11] Pompe, P., Feelders, A., Using Machine Learning, Neural Networks, and Statistics to Predict Corporate Bankruptcy, Microcomputers in Civil Engineering 12, 267–276, 1997.
 - [12] Forgione, G.A., Mora, M., Cervantes, F., Kohli, R., Development of integrated decision making support systems: a practical approach, in Proceedings of the AIS Conference, Long Beach, CA, 2132–2134, 2000.
 - [13] Claver E., Gonzalez, R., Llopes, J., An analysis of research in information systems (1981-1997), Information and Management, 37, 181–195, 2000.
 - [14] Vahidov, R., Kersten, G., Decision station: situating decision support systems, Decision Support Systems, 38, 283–303, 2004.
 - [15] Kwak, K.H., Hwang H.S., Sung, B.K., Development of measuring data system of bridges by wireless transmission using fiber Bragg Grating sensor, Bridge Maintenance, Safety, Management, Life-Cycle Performance and Cost, Cruz, Frangopol & Neves (eds), Taylor & Francis Group, 2006.