Tunnelling and Underground Space Technology 81 (2018) 68–82

Contents lists available at ScienceDirect

Tunnelling and Underground Space Technology

journal homepage: www.elsevier.com/locate/tust

An ensemble-based change-point detection method for identifying unexpected behaviour of railway tunnel infrastructures

Matteo Vagnoli*, Rasa Remenyte-Prescott

Resilience Engineering Research Group, The University of Nottingham, Science Road, University Park, NG7 2RD Nottingham, United Kingdom

ARTICLE INFO

Keywords:
Structural Health Monitoring (SHM)
Data mining
Change-point detection
Ensemble of change-point detection methods

ABSTRACT

A large amount of data is generated by Structural Health Monitoring (SHM) systems and, as a consequence, processing and interpreting this data can be difficult and time consuming. Particularly, if work activities such as maintenance or modernization are carried out on a bridge or tunnel infrastructure, a robust data analysis is needed, in order to accurately and quickly process the data and provide reliable information to decision makers. In this way the service disruption can be minimized and the safety of the asset and the workforce guaranteed.

In this paper a data mining method for detecting critical behaviour of a railway tunnel is presented. The method starts with a pre-processing step that aims to remove the noise in the recorded data. A feature definition and selection step is then performed to identify the most critical area of the tunnel. An ensemble of change-point detection algorithms is proposed, in order to analyse the critical area of the tunnel and point out the time when unexpected behaviour occurs, as well as its duration and location. The work activities, which are carried out at the time of occurrence of the critical behaviour and have caused this behaviour, are finally identified from a database of the work schedule and used for the validation of the results. Using the proposed method, fast and reliable information about infrastructure condition is provided to decision makers.

1. Introduction

The size of the European railway network is expected to continuously increase in order to transport most of the long-distance passengers and freight by 2030 (IRA, 2015). Railways are, indeed, among the most emission-efficient transportation systems, and electric trains can offer a carbon-free journey (if they are powered using nuclear or renewable power sources). However, the European railway network is ageing, and its materials are degrading due to environmental threats (such as traffic, wind and temperature changes (Lee et al., 2005; Chattopadhyay and Reddy, 2007; Boller et al., 2015). The railway infrastructure is thus calling for: *i)* a real-time condition monitoring of its health state over time; *ii)* maintenance activities to restore the health state to a new safe condition; and *iii)* renewal activities to enhance the performance of the railway network, in terms of comfort to passengers, speed, safety and availability of the service (Baxter, 2015). Structural Health Monitoring (SHM) strategies are needed to guarantee the safety, reliability and availability of the infrastructure both during normal operation of the railways and during maintenance and modernization activities (Lee et al., 2010; Chen and Wang, 2017). Particularly, the former requires an SHM method in order to accurately monitor the infrastructure behaviour over time, by identifying ongoing degradation mechanisms of the materials. The latter requires an SHM strategy in order to monitor the infrastructure behaviour in the short-period, by ensuring the safety of the workforce and understanding whether the infrastructure behaviour during the work activities is within the predicted safety limits (Brownjohn, 2007; Rajabi et al., 2017).

SHM strategies are especially necessary for underground infrastructure during their maintenance and modernization activities. In fact, the uniqueness of each underground infrastructure can lead to unexpected behaviour of the in-field structure during the work activities (Bhalla et al., 2005). For example, the UK railway network is subject to an electrification process, which aims to provide cleaner, quicker and more comfortable trains. The old infrastructure of the UK railway network does not usually have the clearance necessary to install the Overhead Line Equipment (OLE) system, which consists of 25 kV AC live conducting wires, insulators and supporting equipment (Kilsby et al., 2017). Hence, maintenance and modernization activities of railway infrastructures are planned, in order to provide the necessary clearance for the OLE system. In the case of tunnels, the way chosen to provide the necessary clearance for the OLE is: *i)* to remove the track, sleepers and ballast; *ii)* to excavate into the sub-formation in order to obtain a new lowered ground; *iii)* to re-establish the ballast, sleepers and track to the new lower level that provides the necessary clearance

* Corresponding author.
E-mail address: Matteo.vagnoli@nottingham.ac.uk (M. Vagnoli).

https://doi.org/10.1016/j.tust.2018.07.013
Received 13 July 2017; Received in revised form 13 December 2017; Accepted 7 July 2018
0886-7798/

for the OLE system. During these works, a real-time SHM system is required with the aim of monitoring the tunnel behaviour. The safety of the workforce is then guaranteed by verifying whether the predicted safety limits are respected by the actual behaviour of the tunnel (Brownjohn, 2007).

Several long-term SHM systems for tunnel infrastructure are presented in literature (Bhalla et al., 2005; Li et al., 2015; Rajabi et al., 2017; Wang et al., 2017), with the aim of developing and studying an optimal measurement system to monitor and predict the behaviour of the tunnel. For example, (Bhalla et al., 2005) present a comprehensive review of measurement strategies for a tunnel, by analysing and discussing the positive effect of the SHM on the tunnel life cycle. (Li et al., 2015) introduce a new wired sensor strategy to monitor the convergence of a tunnel under work activities; (Rajabi et al., 2017) discuss an Artificial Neural Network (ANN)-based method in order to predict horizontal displacements of a tunnel, by relying on the data provided by a 2D Finite Element Model (FEM) of the infrastructure; (Wang et al., 2017) present a long-term SHM system in order to assess the safety of a number of tunnels in China. A large amount of data is however generated by these SHM methods continuously. Data mining techniques are therefore required to analyse this data automatically, accurately and rapidly (Duan and Zhang, 2006). Indeed, data mining techniques are able to transform the recorded data into valuable information for decision makers, by pointing out vulnerabilities of the tunnel. As a consequence, the safety of the asset and the workforce can be improved by taking rapid informed decision (Alves et al., 2015; Zhou et al., 2015; Li et al., 2016).

In this paper, an ensemble-based data mining method is proposed in order to detect the unexpected behaviour of a railway tunnel. The tunnel is subjected to enhancement works, which are necessary for the installation the OLE system during the electrification process of the UK railway network. An SHM system is installed on the tunnel for monitoring the displacements of the tunnel during the works. An FEM of the tunnel has been developed by the contractor of the works in order to predict the displacement of the tunnel during the works. The FEM showed that the expected displacement of the tunnel should remain around the value of zero during each phase of the works, and, if the tunnel would converge, the displacement would increase at a rate of 0.001 mm/h. A data mining analysis is proposed in order to analyse measurements over one month, and identify unexpected behaviour of the tunnel, i.e. the behaviour that does not agree with the FE predictions. The works that are carried out on the tunnel at the time of the unexpected behaviour are also investigated. The recorded data are firstly analysed by using a pre-processing step, which removes the noise of the data. Then, a feature definition and selection process is adopted in order to identify the critical locations of the tunnel, by the means of a K-means algorithm (Jain, 2010). The unexpected behaviours, which are measured at the critical location on the tunnel, are evaluated by using a novel ensemble-based change-point detection method. The proposed method is used to identify the time when the tunnel starts to converge rapidly and point out the duration of the unexpected behaviour. The performance of four individual change-point detection algorithms is merged to detect and diagnose the most critical behaviour of the tunnel (Lavielle, 2005; Carslaw et al., 2006; Killick et al., 2012; Liu et al., 2013). The works at the tunnel site that are carried out at that time are investigated, with the aim of pointing out the causes of the unexpected behaviour. The ensemble-based change-point detection method is needed in order to identify the most critical change-point of the monitored tunnel behaviour. Indeed, single change-point methods, such as Cumulative Sum (CUSUM)-based (Carslaw et al., 2006) or probability distributions-based (Liu et al., 2013) methods, are able to detect only abrupt changes in the data, without pointing out the most severe changes. Furthermore, the longer the duration of the monitored behaviour of the system, the higher the number of the abrupt changes, which are identified by an individual change-point method. Thus, the most severe change in the data can be lost among all the change-points (Killick et al., 2012). Individual change-point methods are also usually unable to identify the duration of the most critical system behaviour, as their objective is to point out the moment when the data deviates from the average behaviour (Maleki et al., 2016). For these reasons, individual change point methods can lead to identifying the incorrect works that might have caused the unexpected behaviour of the tunnel. Conversely, the proposed ensemble-based of change-point methods is able to identify the most critical change in the data, by highlighting its start and end time. In this way, only the information regarding the most critical behaviour of the tunnel is provided to the decision maker. At the same time, the corresponding works at the worksite at the time of occurrence of the most critical behaviour are provided to the decision maker. The performance of the proposed ensemble-based method is compared with the results of each individual change-point algorithm. In this way, the superior ability of the ensemble-based change point method in identifying the most critical change-point is demonstrated.

The remaining of the paper is organized as follow: Section 2 introduces the change-points algorithms and the proposed ensemble-based method; Section 3 describes the tunnel case study and the SHM system and the results; conclusion and remarks are provided in Section 4.

2. The proposed ensemble-based change-point detection method

Change-point detection methods are developed in order to detect changes of the monitored behaviour of a system efficiently and reliably (Tartakovsky et al., 2014). Change-point methods aim to identify the exact moment when the monitored variable of the system starts to deviate abruptly from an equilibrium level. Change-point detection algorithms are adopted in several frameworks, such as SHM and prognosis of gas turbines (Lipowsky et al., 2010; Maleki et al., 2016), variation of air pollution concentration (Carslaw et al., 2006), variation of climate parameters in order to monitor climate change characteristics (Reeves et al., 2007), failure of pipes in chemical industries (Tickle et al., 2010). However, change-point methods are usually unable to point out the most critical change-point clearly, i.e. the change-point where the monitored variable experiences the highest variation. In fact, the most critical change point is identified among all the change-points of the system. The duration of the unexpected changes is also not assessed. Furthermore, the choice of an individual change-point algorithm can jeopardize the reliability and robustness of the data analysis, due to different results of the individual change-point detection methods. For these reasons, in this paper we propose a novel ensemble-based change-point method to analyse a large database of displacements of a railway tunnel, by coupling the performance of four individual change-points methods. In this way, the proposed ensemble-based method is able to point out the most critical change-point of the system, providing its duration and possible causes. As a consequence, the reliability and robustness of the data mining analysis are expected to improve accordingly. Decision makers can then schedule future work activities by using the results of the ensemble-based change-point method directly. Indeed, the most critical behaviour of the tunnel is pointed out clearly, in terms of both duration and corresponding works at the worksite.

In what follows, the theoretical background of the four individual change-point algorithms is presented briefly, and then the proposed ensemble-based method is introduced.

2.1. Change-point methods: theoretical background

The change-point analysis can be divided in two groups - real-time and retrospective detection: the former aims to identify a change-point of system behaviour as soon as it occurs; the latter aims to identify a change-point of system behaviour by analysing the history of the monitored parameter. The focus in this paper is on the retrospective change-point analysis, which provides more robust and accurate