Structural health monitoring meets data mining

Miao, S.

Author: Miao, Shengfa
Title: Structural health monitoring meets data mining
Issue Date: 2014-12-16

Chapter 1

Introduction

1.1 Background

Over the last decade, assessing the service-life of concrete civil structures is a theme that has gained a lot of interest. Despite concrete being a construction material that can last several decades to centuries, it has become clear that external influences may substantially (and often unexpectedly) shorten the service-life of concrete structures. More in detail, the factors that affect the service-life of civil structures have various origins, such as traffic load, varying climate conditions as well as the natural degradation of the material involved, notably the concrete and the reinforcement bars.

The traditional way to assess the actual condition of infrastructural assets is based on visual inspection or portable instruments, an approach which suffers from the following drawbacks:

- It is fairly subjective and difficult to quantify.
- It requires a lot of manpower, material and equipment.
- It may have blind spots, and completeness cannot be guaranteed.
- Its inspection period is long and inefficient.
- It interferes with the normal flow of traffic.

According to a recent survey from the US Federal Highway Commission [1], on average 56% of the assessments made by visual inspection are inappropriate. Driven by these drawbacks, the field of *Structural Health Monitoring* (SHM) is emerging, which is an interdisciplinary field, including civil engineering, signal processing, sensor technology, material sciences, data management and mining. The SHM process can be approached from a Statistical Pattern Recognition paradigm [2, 3], which employs an array of sensors to periodically collect the dynamic response of the monitored structure, and assesses the system's health, with damage-sensitive features extracted from these measurements.

In this thesis, we discuss results from a Dutch SHM project, the so-called *InfraWatch* project. The project is one of the key projects of a Dutch STW perspective program, called *Integral Solutions for Sustainable Construction* (IS2C). The IS2C program is composed of nine research projects, aiming to enforce new innovations in the state-of-the-art service-life assessment and to set a new standard for sustainable construction. The InfraWatch project covers the aspects of sensing, monitoring and degradation mechanisms, and is a joint research project between Leiden University and Delft University of Technology. The data used for this project is captured by a monitoring system that is installed at a major highway bridge in the Netherlands, called the *Hollandse Brug*. The computational data analysis and data mining has been conducted by researchers at Leiden University, while the physical interpretation and matching with structural analysis models was conducted at Delft University of Technology.

1.2 Objectives and Scope

According to Farrar and Sohn's approach [2, 3], the SHM process can be broken down into four parts:

- Part 1: Operational Evaluation: damage definition, life-safety, economic justification, operational and environmental conditions and limitations are considered in this step.

- Part 2: Data Acquisition, Fusion and Cleansing: excitation methods, structural response and data transmission are considered in this step.
- Part 3: Feature Extraction and Information Condensation: this step focuses on selecting features that indicate the health of the structure, such as natural frequencies, damping ratios and mode shapes.
- Part 4: Statistical Model Development for Feature Discrimination: this step aims to design algorithms to distinguish between features from the undamaged and damaged structures.

Part 1 and Part 4 are beyond the scope of this thesis. The bridge in our project was closed for renovation in 2007, and since then a sensor network has been installed on the bridge. InfraWatch started two years after the renovation and the sensor network installation, so the operational evaluation step is skipped in this work. Since the installation of the network, three years of data has been collected. During this period, it is reasonable to assume that the bridge has not suffered any major damage, so the damage identification in Part 4 is not covered in this thesis.

In this thesis, we focus on Part 2 and Part 3 in the above paradigm. We are interested in understanding the specifics of each sensor type and individual sensors, and the dependencies between sensors of different types. Each dataset collected with an individual sensor produces a time series, which is sensitive to several external factors, most notably daily traffic and temperature variations. Some of these factors are useful, like truck events, which help to excite (bring in motion) the bridge, while some of them are interference factors, like temperature influence, which hinder the proper analysis. To select and extract reliable features from massive datasets, we employ a number of signal processing and data mining techniques, which are briefly introduced in the following subsections.

1.2.1 Data Acquisition and Signal Processing

In SHM, one can distinguish two general excitation methods [3] for data acquisition: the *forced* and the *ambient* excitation method. With the forced excitation method, the input forces are controllable and measurable (such as with an impact hammer, a shaker, or a controlled load such as a weighed truck), so it is easy to obtain a clear and interpretable signal. This method is usually adopted in laboratory tests or to obtain short-term data from a real structure in the field. In contrast, with the ambient excitation method, it is hard to measure the input forces accurately, because these forces are usually varying and occur at random intervals. However, this method is suitable for long-term monitoring of structures.

Our InfraWatch project is based on the ambient excitation method for data acquisition. There are three sensor types involved in the sensor network, measuring strain, vibration and temperature. The strain sensors indirectly measure the load of the bridge by measuring strain experienced parallel to the bridge, in two horizontal directions. Strain measurements are sensitive not only to the experienced load of the bridge (which will make the structure bend), but also to a large extent to temperature effects. The vibration sensors measure vertical shaking of the bridge, caused by the impact and passing of traffic. Contrary to strain gauges, they are hardly sensitive to temperature changes. The temperature sensors measure the local temperature of the bridge, at the exact point where they are attached to the bridge. This temperature may vary a bit, depending on the location of the bridge.

On the Hollandse Brug, we can generally recognise three different phenomena at different time scales. This shows up as three different components in the strain signal: a low-frequency component, a medium-frequency component and a high-frequency component. The low-frequency component includes effects such as the daily temperature fluctuations or traffic jams. These effects will show up in the strain signal as a drifting baseline. Removing the low-frequency component from the signal is known as *baseline correction* [4]. A good baseline correction method helps to study useful patterns hidden in the raw time series. The medium-frequency component consists of normal traffic events, which are considered useful patterns in this work. The high-frequency component consists of noise (for example caused by the rolling tires of vehicles or measurement noise), which can be eliminated using smoothing methods.

Although the strain signal seems more informative than the vibration signal when considering the time domain, it turns out to be less useful in the frequency domain. Especially when considering details of the vibrations caused by (heavy) vehicles, which is captured to varying degrees by both sensors, the vibration signal is more clear, and not affected by the baseline drift. By combining the strain and the vibration signals, we succeeded in developing a method to select high-quality data for feature extraction.

1.2.2 Feature Extraction

The integrated performance of the bridge can be studied through so-called *modal parameters*: natural frequencies, damping ratios and mode shapes [5, 6, 7, 8]. In this thesis, we are interested in the following topics:

- The selection of high-quality datasets.
- Modal analysis methods.
- The influence of temperature on modal parameters.
- The influence of traffic mass on modal parameters.

Based on the understanding of the different behaviour of the strain and vibration sensors, we select the vibration sensor as our target sensor type for modal analysis. To get rid of the influence of traffic mass, we select datasets during so-called *free-vibration periods* as our target datasets. The free-vibration period is the period right after a vehicle has passed, and before a next vehicle appears on the bridge. The reason for choosing this period is that the bridge is put in motion by the vehicle, but the actual weight does not influence the frequency of vibration after the vehicles has disappeared, nor do any other vehicles.

A number of methods have been developed for modal analysis, such as structural calculation methods (Finite Element Method (FEM)), the *Peak-Picking* method (PP) [9, 10] and the *Stochastic Subspace Identification* method (SSI) [11, 12, 13, 14, 15]. We extract modal parameters by combing these modal analysis methods, and pay special interest to the relationship between natural frequencies and temperature.

1.3 Thesis Outline

This thesis is composed of eight chapters. Most of these chapters are based on published papers by the author. The following provides a brief description of each chapter.

Chapter 2 presents some basic concepts that play a role in the remainder of the thesis.

Chapter 3 presents an introduction to the InfraWatch project. In this chapter, we introduce the bridge and the sensor network in detail. Parts of the content in this chapter were previously published in the following paper:

> Veerman R., Miao S., Koenders E., and Knobbe A. *Data-Intensive Structural Health Monitoring in the InfraWatch Project*. In Proceedings of the 6th International Conference on Structural Health Monitoring of Intelligent Infrastructure (SHMII6), Hongkong, 2013.

Chapter 4 explores sensor dependencies among multiple sensor types. All the sensors in the sensor network are sensitive to related aspects of the measured system, that is to say there are certain dependencies. To gain insight into these dependencies, and how the placement and location of sensors influences them, we employ linear regression, convolution, envelope and band pass filters to model signals in both the time and the frequency domain, and then utilise Subgroup Discovery [16, 17, 18] to further analyse the obtained models. This work was published in the following paper:

> Miao S., Vespier U., Vanschoren J., Knobbe A., and Cachucho R. *Modeling Sensor Dependencies between Multiple Sensor Types*. In Proceedings of BeneLearn, Nijmegen, 2013.

Chapter 5 looks into the problem of baseline drift. To separate the influence of normal traffic events from other environmental factors, we propose a novel baseline correction method, the *Most-Crossing* method, which is a piece-wise method, based on probability theory. The method assumes that patterns of the same scale follow the same probability distribution, so that patterns of different scales can be distinguished based on their probability distributions. In strain signals of the sensor network, the probability distribution of environmental factors, which contribute to baseline, is different from that of traffic events. Based on this observation, we propose the Most-Crossing method to extract the baseline from strain signals. This work was previously published in the following paper:

> Miao S., Koenders E., and Knobbe A. *Automatic Baseline Correction of Strain Gauge Signals.* In Structural Control and Health Monitoring 22 (1), pp. 36-49, 2015.

Chapter 6 covers the topic of *predefined pattern detection*. Given a pattern (template), we can characterise it as a combination of landmarks and constraints. Landmarks are remarkable points in the pattern, e.g., local extrema. Constraints are composed of local constraints and global constraints. The former focus on properties of individual landmarks, and the latter focus on relationships between properties of different landmarks within the pattern. If the prior knowledge is given to us by domain experts, the pattern detection procedure can be addressed as a predefined pattern detection issue. Predefined pattern detection has its advantage in processing huge datasets collected from a specific domain. It will be extremely expensive to detect patterns with traditional pattern detection methods, which work through all possible pattern lengths. What's more, most of the existing pattern detection methods focus on full sequence matching, that is, sequences with clearly defined beginnings and endings, where all data points contribute to the match. These methods will become ineffective when deformations appear in both temporal and amplitude dimensions. This work has been submitted to the journal of Information Sciences:

> Miao S., Vespier U., Meeng M., Cachucho R., and Knobbe A. *Predefined Pattern Detection in Large Time Series.* revised to Information Sciences, 2014.

Chapter 7 presents modal analysis of the bridge. Changes in the integrity of the material and/or structural properties of structures are known to adversely affect their performance, which can be observed from structures' dynamic response. We propose a procedure to select high-quality datasets, and employ two modal analysis methods to extract modal parameters from them. We also look into the influence of environmental factors, such as traffic mass and temperature, on modal parameters. This work was published in the following papers:

> Miao S., Veerman R., Koenders E., and Knobbe A. *Modal Analysis of a Concrete Highway Bridge — Structure Calculations and Vibration-Based Results.* In Proceedings of the 6th International Conference on Structural Health Monitoring of Intelligent Infrastructure (SHMII6), Hongkong, 2013.
>
> Miao S., Knobbe A., Koenders E., and Bosma C. *Analysis of Traffic Effects on a Dutch Highway Bridge.* In Proceedings of the International Association for Bridge and Structural Engineering (IABSE), Rotterdam, 2013.

Chapter 8 concludes the research involved in this thesis, and presents a number of recommendations for further work.