Mining Patterns for Structural Health Monitoring

Ugo Vespier, Arne Koopman, Arno Knobbe, Marvin Meeng
LIACS, Universiteit Leiden

Since the last decades, numerous application domains produce an explosion of digital data, together with an increase in its consumption. To make sense out of this data flood, the research community continuously devises new techniques to manage and analyse these massive data collections. Much of this data is captured in so called data streams, which are large sequences of elements of a given data type. These data types can vary widely and can be derived from various applications; from sensor networks to website logs and financial transactions.

Sensor networks, in particular, are applied to measure the structural characteristics of large infrastructures such as buildings and bridges. The novel field of Structural Health Monitoring (SHM) uses techniques to monitor their status by analysing the data derived from these sensors. This is a challenging and critical problem that can positively affect the manner how we manage and maintain public infrastructures.

Our work is part of a Structural Health Monitoring project called *InfraWatch* that investigates new data management and analysis techniques. In particular, it focusses on one of the largest highway bridge in the Netherlands: the "Hollandse Brug". This bridge is equipped with a network of 145 sensors that measure stress, vibration and temperature on several locations. Moreover, the sensor network is coupled with a video data stream, and produces more than 5GB of data every day. This architecture serves as a perfect real-world evaluation platform for SHM algorithms that operate on large scale data.

We present our preliminary findings within the context of InfraWatch and introduce a technique to model aspects of the bridge's behavior via equations. We exploit the continuous nature of the data by using the Lagramge equation discovery system to generate various types of equations that model relations between sensors. Due to the abundance and redundancy of generated equations, a subset selection algorithm is used to select those equations that best model and cover the bridge's behavior. Our experiments show that relationships between sensors can be effectively modelled by equations and that we can select from this abundance a small set based on their relevancy.