Mechanical Systems and Signal Processing 62-63 (2015) 381–394

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Data quality enhancement and knowledge discovery from relevant signals in acoustic emission

Felipe Mejia [a,*], Mei-Ling Shyu [b], Antonio Nanni [a]

[a] Department of Civil Architectural and Environmental Engineering, University of Miami, Coral Gables, FL 33146, United States
[b] Department of Electrical and Computer Engineering, University of Miami, Coral Gables, FL 33146, United States

A R T I C L E I N F O

Article history:
Received 29 November 2012
Accepted 17 February 2015
Available online 17 March 2015

Keywords:
Acoustic emission
Bridges
Data mining
Non-destructive testing
Structural health monitoring

A B S T R A C T

The increasing popularity of structural health monitoring has brought with it a growing need for automated data management and data analysis tools. Of great importance are filters that can systematically detect unwanted signals in acoustic emission datasets. This study presents a semi-supervised data mining scheme that detects data belonging to unfamiliar distributions. This type of outlier detection scheme is useful detecting the presence of new acoustic emission sources, given a training dataset of unwanted signals. In addition to classifying new observations (herein referred to as "outliers") within a dataset, the scheme generates a decision tree that classifies sub-clusters within the outlier context set. The obtained tree can be interpreted as a series of characterization rules for newly-observed data, and they can potentially describe the basic structure of different modes within the outlier distribution. The data mining scheme is first validated on a synthetic dataset, and an attempt is made to confirm the algorithms' ability to discriminate outlier acoustic emission sources from a controlled pencil-lead-break experiment. Finally, the scheme is applied to data from two fatigue crack-growth steel specimens, where it is shown that extracted rules can adequately describe crack-growth related acoustic emission sources while filtering out background "noise." Results show promising performance in filter generation, thereby allowing analysts to extract, characterize, and focus only on meaningful signals.

1. Introduction

Acoustic emission (AE) testing—a relatively new structural health monitoring (SHM) technology—has enjoyed high growth and focus in both the private and research industries [1–4]. However, even though AE has established itself as an inexpensive damage detection technology, it suffers from shortcomings that are still the focus of current research. More specifically, the presence of unwanted signals (also loosely referred to as "noise") has made the analysis of AE datasets a notoriously difficult process.

Historically, AE unwanted signal removal techniques have been classified into two types: spatial techniques and parametric techniques. Spatial techniques focus on the time and the order of arrival of waves detected by two or more sensors. On the

* Correspondence to: 1251 Memorial Drive, McArthur Engineering building, Room 301, Coral Gables, FL 33146-0630, United States. Tel: +1 305 284 1299; fax: +1 305 284 3492.
E-mail address: f.mejia2@umiami.edu (F. Mejia).

http://dx.doi.org/10.1016/j.ymssp.2015.02.017
0888-3270/

other hand, parametric techniques require the identification and classification of unwanted signals, followed by the creation of pre-processing or post-processing filtering rules. Both types of signal removal techniques are compatible and can be applied during testing or as part of post-processing analysis. Spatial techniques suffer from requiring more than one sensor and knowledge of wave properties inside the material, and parametric techniques require a manual interpretation of unwanted signals. Therefore, there is a clear need for automated "noise" removal approaches that can be used on single or sparse sensor setups, which are common in many SHM applications.

Fortunately, advances in data mining and pattern recognition techniques have given AE signal characterization a renewed interest, although studies have been centered on characterization of material failure mechanisms instead of systematic identification and removal of unwanted signals. This study implements a data mining scheme for the identification and characterization of data belonging to unfamiliar distributions (herein referred to as "outliers"). By further clustering the outlier and subsequently extracting classification rules, statistically similar emissions can be efficiently filtered and, by nature of the extracted rules, the underlying structure of the AE source can be characterized with a certain degree of accuracy.

The proposed data mining scheme is validated in three steps. After an overview of the involved algorithms, the algorithms are applied to a synthetic dataset, where the outlier classification accuracy and performance is examined. Then the algorithm is fit to work with AE data arising from pencil lead breaks under background "noise" conditions in order to prove that AE hits from different distributions can be successfully segregated. Finally, the algorithm is applied to single-edge cracked tension (SE(T)) steel specimens, where the procedure is used as a semi-supervised background "noise" removal filter and as a tool for crack-growth signal characterization. Lastly, a discussion of extracted rules and recommendations for improved rule generalization are given.

2. Background

2.1. SHM and AE

SHM is usually defined as the implementation of a damage detection strategy for aerospace, civil, and mechanical engineering infrastructures. This process involves the observation of a structure or mechanical system over time using periodically spaced measurements, the extraction of damage-sensitive features from these measurements, and the analysis of these features to determine the current state of system health [5]. An SHM system may consist of sensors, data acquisition and transmission systems, databases for effective data management, and health diagnosis methodologies (including data processing, data mining, damage detection, model updating, safety evaluation, and reliability analysis). The number of SHM technologies is continuously growing, but research efforts have mainly been concentrated on the areas of (1) acoustic signals, (2) electromagnetic, (3) radiography, (4) fiber optics, (5) radar and radio frequency, (6) optics, and (7) piezoelectric ceramic [6]. A more detailed description of SHM can be found in [7].

Among all the different SHM techniques, AE is particularly attractive since it allows analysts to observe the dynamics of material performance in real time. The process of AE occurs when a material suddenly releases localized stress energy, thereby causing a transient elastic wave that propagates through the material. AE is usually associated with irreversible dynamic processes such as friction, fracture, impacts, crack growth, corrosion, and other types of damage [8]. AE is considered a "passive" method since it does not itself input any energy into the observed system and, as such, allows analysts to test under typical operating conditions [9].

AE has been shown to occur during testing of metals in fatigue crack-growth experiments. It is produced not only by the onset of yielding at the crack tip or crack extension, but also by the rubbing of fatigue crack surfaces due to closure [10]. Abrading surfaces produce frequent emissions, which have a slow rise time and low amplitude. AE from crack closure can occur even during tension-tension cyclic loading [11].

2.2. Parametric analysis in AE

Parametric analysis of AE in metals is not a new concept. Initial analyses of AE were based on emission rates, or "hit count" [12]. Shortly after, as more powerful circuitry became available in AE systems, basic waveform parameters such as amplitude were able to be computed. The potential of amplitude analysis as a viable characterization method of emission signals was recognized in [13] and became the dominant practice for several years. More recently, with the advent of waveform recording capabilities, feature extraction (e.g., frequency, time–frequency, and wavelet transforms) in post processing has been used extensively [14–17].

2.2.1. Pattern recognition, machine learning, and data mining in AE

Farrar et al. [5] described SHM primarily as a problem of pattern recognition, and the AE research community has since adopted this paradigm. Pattern recognition has been used in AE mainly as a technique for characterizing the structure and natural "signatures" in AE datasets [15,18–20].

More recently, pattern recognition in AE has evolved towards machine learning and data mining, which have been used mainly for classification, regression, and prediction. Current implementations are implementing data mining tools for hypothesis searching, rule extraction, and decision-making. For example, artificial neural networks (ANN) have been used to classify and predict background "noise" in aerospace composites [21]; waveform classification in FRP monitoring [22]; and