

Signal-level clustering of acoustic emission streaming

Pauline Butaud¹, Emmanuel Ramasso¹, Hanaa Ahayan¹, Thomas Jeannin¹, Nathalie Godin², Vincent Placet¹

1 Univ. Bourgogne Franche-Comté, FEMTO-ST, Dep. of Applied Mechanics,
Besançon, France. emmanuel.ramasso@univ-fcomte.fr
2 INSA Lyon, MATEIS, Villeurbanne cedex, nathalie.godin@insa-lyon.fr

Abstract

Clustering of acoustic emission signals aims at interpreting the dynamical materials behaviour through solicitations. It is usually posed as problem of finding clusters with good shape and well separated in the feature space, where features are extracted from AE signals. We here propose an alternative which does not use such features. The methodology relies on anomaly detection and includes a criterion that optimises the spread of onsets of clusters in place of shape.

1. Introduction

The Acoustic Emission (AE) technique is widely used for material characterization and structural health monitoring. Once collected, the standard approach to process and interpret the potentially large amount of AE streaming data is generally based on four main phases: 1) wave picking, 2) feature extraction, 3) clustering, 4) evaluation of clusters with, if possible, correlation with AE sources.

The first phase, wave picking, isolates relevant AE signals from noise within the streaming. It is generally based on a set of thresholds to ensure that AE signals have a sufficiently high amplitude during a certain amount of time. Several wave picking methods have been proposed in the past. A general methodology have been proposed by Unnpórrsson (10) using peak picking of a detection function. The detection function may be based on frequency content analysis using Fourier transform or wavelets (9).

In the second phase, features are extracted from AE signals. The use of features is principally motivated by representing AE signals in the same dimensional space. Indeed, AE signals have generally different lengths which cannot be managed by standard clustering methods. Numerous features have been proposed in the past, a compilation has been for example made in (6).

The potentially high number of features and the unknowns concerning the correct features to use for given damage families, justify, in some way, the use of data reduction methods, such as the principal component analysis (4).

Clustering then aims at finding out the data structure to infer the behavior of the material during solicitations. A “good” set of clusters has to optimize some predefined criterion which is used in the evaluation step. Features, clustering and evaluation steps are thus closely related and all require to select properly the good set of parameters:



which features to be extracted? which clustering algorithm (and its parameters)? and which evaluation criterion? (3).

Being used as inputs of clustering and of evaluation criterion, as well as highly influenced by the wave picking method, features are thus of key importance. Finding a consensus for their interpretation is also a critical issue (9).

In this paper, we propose to work at the AE signal level, without feature extraction in the usual sense, and in an unsupervised manner. While the supervised case has already been tackled in the past, unsupervised learning at the signal (or streaming) level for AE interpretation is a novel approach.

The proposed method has three main advantages:

- 1) it can be used for supervised or unsupervised learning;
- 2) it can work directly on streaming without wave picking. This particularly simplifies the processing by reducing the number of algorithms and parameters to be used;
- 3) if wave picking is used to extract AE hits, then the method does not require feature extraction from AE signals. It simplifies the processing by avoiding the heavy task of feature selection.

In the sequel, we first describe the methodology. We then show a practical example.

2. Methodology

The proposed method relies on a statistical model of the temporal evolution of the streaming data or of an AE signal. The generalization of the proposed methodology to AE streaming is simple, however we restrict the description to AE signals in the sequel for the sake of convenience.

The idea relies on novelty detection. Novelty detection has been widely used in various applications and implemented in many publications dedicated to SHM using vibration data (2). The principle is to build a mathematical model representing how “normal” data are distributed. This model is then applied online, during test, and generates a similarity between the current data compared to the normal ones. If a drift occurs a modification in the data is likely to occur. Exploiting those modifications for AE signal clustering is the main purpose of this work.

In (2), an Autoregressive model (AR) is used to represent vibration data collected by accelerometers for SHM purposes in real-world situations. In this paper, the methodology relies on a new probabilistic model called Autoregressive Weakly-Hidden Markov Model. We have developed the theoretical part for prognostics (not for AE) in a previous conference paper (5) where details can be found on the model.

The main idea to keep in mind is that an AE signal called reference is used as an input of this model which estimates a set of parameters to fit the signal. The parameters are then used in inference during test: for each AE signal tested (not used during the

learning phase), a set of novelty indicators are computed. Those indicators reflect how similar is the tested signal compared to the model.

We finally obtained a matrix of novelty indicators, with as many lines as the number of AE signals (or pieces of streaming) and as many columns as the number of indicators. We then apply a clustering method to find the structure of the indicators. For that we have used the method proposed in (9) which has the following specificities:

- It uses the Gustafson-Kessel algorithm in place of Kmeans for clustering the indicators by considering general hyper-ellipsoids to define clusters shape.
- It uses multiple subsets of indicators to perform clustering: in place of using one unique subset, the algorithm makes all combinations of n among all indicators and perform clustering. It then selects a subset of subsets according to an information theoretic criterion onsets or clusters proportions. This specificity allows representing multimodal clusters.
- It uses an unsupervised fusion method to combine the partitions obtained by considering multiple subsets. The fusion relies on reordering the clusters according to their order of occurrence: This specificity allow to consider the sequence of clusters as a criterion while previous methods, though plotting sequences, actually optimizes clusters according to their shape in the feature space and not according to sequences or onsets.
- It estimates in an unsupervised manner the uncertainty around clusters estimated after fusion.
- It estimates the robustness of the results and optimizes the number of clusters by maximizing robustness. This optimization does not rely on features, only on partitions, and makes use of an information theoretic criterion.

The only critical parameter of this methodology is the reference AE signal, as for most anomaly detection methods. We propose to perform bootstrapping to solve this issue: Sample AE signals randomly, learn a model for each, apply the models on the remaining AE signals, and repeat this process several times. Then we apply the previous clustering methodology using all obtained novelty indicators. By doing so, uncertainty due to the reference is quantified and robustness maximized against its selection.

3. Results

The proposed method can be applied on any existing datasets, provided AE waveforms have been recorded (for example in DTA files of Mistras' AEWin software). We propose below an illustration on a unidirectional flax-epoxy composite coupon.

2.1 Materials

Specimen preparation: Unidirectional flax-epoxy composite plates were fabricated by thermo-compression of the Flaxpreg T-UD material provided by LINEO®. It is a range of pre-impregnated material based on an epoxy resin system and a unidirectional flax fibres reinforcement. The flax fabric has an areal weight of 110 g/m². The resin that is used is the Huntsman XB3515® and the hardener is the Aradur5021®. The fibre fraction is approximately 50% in weight before curing.

Composite fabrication: Plates were fabricated using a thermo-compression machine (AGILA Presse 100KN). These plates consist in 8 elementary unidirectional plies. The layers are compressed in a Teflon coated aluminium mould and cured for one hour at the cure temperature of 120 °C. A pressure equal to 3 bars is applied for a period of 1 h after the temperature of the sample has reached 60°C. The mould was under vacuum during the fabrication.

Sample preparation: Specimens were cut in the plates using a Trotec Speed300® laser device in the fibre direction at the following dimensions: 240mm x 15mm x 1mm. Each specimen was characterized in terms of fibre content as proposed in (1). The fibre volume fraction was measured to be equal to $61 \pm 2\%$.

Environmental conditioning: After fabrication, samples were stored in a climatic chamber at 23°C and 50% RH during at least 20 days before mechanical testing. During this period, relative weight uptake and dimensions were monitored. A Kern 770 balance was used for weight monitoring (precision of 0.001 g), a calliper with a precision of 10^{-2} mm was used to evaluate the swelling.

2.2 Estimation of damage kinetics

The tensile properties were determined using an Instron Electropuls E10000 machine equipped with a 10 kN load sensor. The sample was subjected to a displacement control ($2 \text{ mm} \cdot \text{min}^{-1}$) up to failure. A total of about 33000 AE signals was recorded in the test, using a similar configuration of the acquisition and signal processing as in (7) using streaming data.

The aforementioned methodology was applied using 4 reference AE signals randomly selected in the first 2000 AE signals recorded. Six indicators were computed for each AE signal and all combinations of 2 indicators have been considered for a number of clusters ranging from 3 to 8.

Damage onsets with sequence of clusters (also called cumulated number of hits per cluster, chronology or time-dependent behaviour) are represented in Figure 1. In figure 2 we represent the total number of times a particular novelty indicator has been used by the consensus fusion method to estimate the damage sequence.

From figure 1 we can observe that the sequence of AE signals has been decomposed into 5 steps by the algorithm:

1. a first onset at the beginning of the test
2. a second onset around 24 s.
3. a third onset around 42 s.
4. a fourth onset around 72 s.
5. and the fifth onset around 95 s.

The uncertainty envelop is the largest for cluster 3. It means that varying the inputs (the indicators) leads to different results. For the other clusters, results are accompanied with small uncertainty. This chronology can easily be used for prognostics since it decomposes the behaviour into useful steps.

As already observed in (9), the proportions of AE hits in each cluster is quite different and the sooner the onset, the higher the proportion. For fibre reinforced composites, this behaviour was observed multiple times in our tests. However, this observation is material-dependent.

Figure 2 shows how many times indicators have been used to estimate the sequence. 6 indicators were computed for each of the 4 reference AE signal leading to 24 possible inputs of the clustering method. All combinations of 2 indicators were considered during clustering. About 20 combinations were kept at the end in order to perform the fusion of the partitions. This figure shows for example that the reference AE signal number 4 has been considered as the most relevant to estimate the sequence. For this reference, the indicator number 5 has been considered as the most relevant.

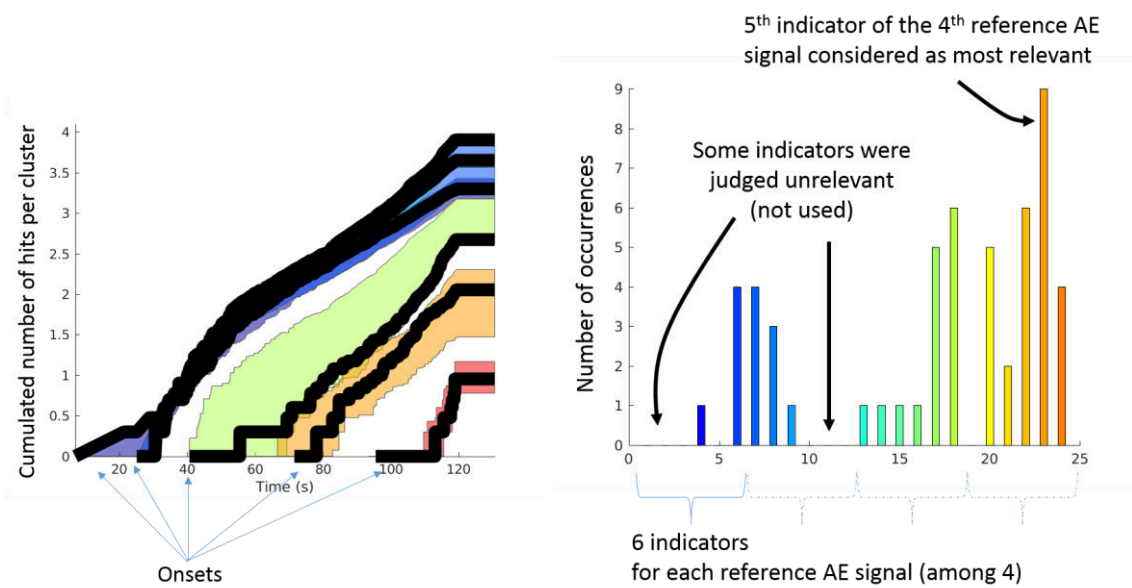


Figure 1. Estimation of the damage sequence and its uncertainty.

Figure 2. Number of times a particular indicator has been used for estimating the sequence.

4. Conclusions

A new methodology has been proposed for the clustering of AE signals. It makes use of anomaly detection and optimises spread of onsets. It is a novel methodology in the sense that it does not require feature extraction as done in common approaches. It also quantifies uncertainty and includes robustness to select the best parameterization. Results show the interest of this approach for prognostics.

Acknowledgements

This work has been supported by the Labex ACTION project (contract "ANR-11-LABX-0001-01").

References and footnotes

1. Amiri A, Triplett Z, Moreira A, Brezinka N, Alcock M, Ulven CA., “Standard density measurement method development for flax fiber”. *Industrial Crops and Products*, 96(Supplement C):196-202, 2017.
2. Farrar, C. R., & Worden, K., “Structural health monitoring: a machine learning perspective”. John Wiley & Sons. 2012.
3. Godin, N., Huguet, S., Gaertner, R., & Salmon, L., “Clustering of acoustic emission signals collected during tensile tests on unidirectional glass/polyester composite using supervised and unsupervised classifiers”. *Ndt & E International*, 37(4), 253-264, 2004.
4. Johnson, M., “Waveform based clustering and classification of AE transients in composite laminates using principal component analysis”. *NDT & E International*, 35(6), 367-376, 2002.
5. Juesas, P., Ramasso, E., Drujont, S., & Placet, V. “On partially supervised learning and inference in dynamic Bayesian networks for prognostics with uncertain factual evidence: Illustration with Markov switching models”. *Conference on Prognostics and Health Management, Bilbao*, 2016.
6. Kattis S., “Noesis – Advanced Data Analysis, Pattern Recognition & Neural Networks Software for Acoustic Emission Applications”, *Kolloquium Schallemission, Statusberichte zur Entwicklung und Anwendung der Schallemissionsanalyse*, 9-10 March, Fulda, 2017.
7. Kharrat, M., Ramasso, E., Placet, V., & Boubakar, M. L., “A signal processing approach for enhanced acoustic emission data analysis in high activity systems: application to organic matrix composites”. *Mechanical Systems and Signal Processing*, 70, 1038-1055, 2016.
8. Kurz, J. H., Grosse, C. U., & Reinhardt, H. W. “Strategies for reliable automatic onset time picking of acoustic emissions and of ultrasound signals in concrete”, *Ultrasonics*, 43(7), 538-546, 2005.
9. Ramasso, E., Placet, V., & Boubakar, M. L., “Unsupervised consensus clustering of acoustic emission time-series for robust damage sequence estimation in composites”, *IEEE Transactions on Instrumentation and Measurement*, 64(12), 3297-3307, 2015.
10. Unnþórsson, R., “Hit detection and determination in AE bursts”, *Acoustic Emission-Research and Applications*, Chapter 1, InTech, 2013.