

Segmentation de séries temporelles par modèles de mélange avec contraintes sur les instants d'apparition des classes

Emmanuel Ramasso¹

Thierry Denœux²

¹ FEMTO-ST, Département Mécanique Appliquée

² Université de technologie de Compiègne, CNRS, Heudiasyc

Univ. Bourgogne Franche-Comté, 15B Av. des Montboucons, 25000 Besançon, emmanuel.ramasso@femto-st.fr
Université de technologie de Compiègne, Rue Roger Couitolenc, 60203 Compiègne Cedex, thierry.denoeux@utc.fr

Résumé :

La surveillance des structures est un domaine en plein essor faisant appel à certaines technologies de capteurs qui génèrent des données dont les spécificités nécessitent de développer des algorithmes d'analyse de données dédiés. Dans cette courte communication, nous nous intéressons à la technique de l'émission acoustique dont les capteurs fournissent des données temporelles massives, non régulièrement espacées, en temps quasi-continu et sans a priori sur le lien entre les signaux d'émissions acoustiques et l'endommagement dans les matériaux. Pour cela, après une description de la technique, de l'état de l'art et de la problématique, nous proposons une esquisse d'un modèle de mélange permettant de tenir compte de l'ordre des observations en temps continu avec possibilité de déterminer et de contraindre les instants d'apparition des classes en lien avec les endommagements.

Mots-clés :

Surveillance de structures, émissions acoustiques, partitionnement.

Abstract:

Structural Health Monitoring is an active field of research relying on various sensors technologies to collect data with strong specificities requiring the development of dedicated algorithms. In this short communication, we discuss about the acoustic emission sensing technique that is based sensors providing massive time-series made of irregularly spaced in time events occurring in almost continuous time and without prior knowledge about the link between the signals collected and the real damages occurring in the material. For that, after a description of the technique, the literature and the problem, we propose the main lines of a method which relies on a modification of Gaussian mixture models in order to explicitly take the timestamp of acoustic emission signals into account. In that way, it is possible to determine the sequence of damage-related clusters.

Keywords:

Structural Health Monitoring, acoustic emission, clustering.

1 Introduction

Les moyens de transport (avions, trains, automobiles, bateaux) et les constructions et infrastructures civiles et industrielles (entrepôts, réservoirs, éoliennes, conduites pour le transport de matières, routes, voies ferrées, tunnels, barrages, digues, etc.) sont directement liés à la croissance démographique et économique globale de tout pays. Prévenir l'apparition d'une défaillance sur ces structures est donc un enjeu majeur, d'autant plus pour des structures vieillissantes qu'il est nécessaire de maintenir pour des raisons évidentes liées aux coûts de maintenance, à leur disponibilité et la sécurité des utilisateurs.

Pour détecter au plus tôt un dysfonctionnement lié à une usure ou un dommage, il est nécessaire de comprendre et modéliser ces structures complexes. Cela passe par leur instrumentation à l'aide de capteurs et par l'analyse des données.

L'émission acoustique est une technique de mesures non destructive utilisée dans divers domaines depuis plusieurs décennies. Lorsqu'apparaît un changement dans l'intégrité d'une structure lié aux sollicitations lors de l'utilisation en service, une partie de l'énergie libérée prend la forme d'ondes élastiques transitoires se propageant depuis la source de l'émission (l'endommagement) jusqu'à la surface du matériau et que l'on peut détecter par des capteurs piézo-électriques. Cette technique est dite passive dans le sens où la source d'émission est liée

à l'endommagement de la matière qui n'est pas contrôlée, ce qui diffère des techniques actives reposant sur une source externe utilisée généralement comme référence.

Le contenu fréquentiel des ondes est compris entre 20 kHz et 1 MHz selon les capteurs utilisés. Les émissions acoustiques sont ainsi des séries temporelles collectées en continu avec des fréquences d'échantillonnage autour de 5 MHz. L'autre caractéristique des émissions acoustiques concerne le fait que les endommagements apparaissent au gré des sollicitations et des évolutions des paramètres des matériaux. Ainsi, les signaux transitoires liés à des endommagements ne sont pas espacés de manière régulière. Etant donnée la fréquence d'échantillonnage élevée, ces signaux transitoires peuvent apparaître pratiquement à n'importe quel instant. Enfin, même si certaines sources d'émissions acoustiques sont connues a priori, il n'existe à l'heure actuelle aucun moyen d'associer avec certitude une source à tous les signaux collectés pendant des essais de fatigue (les plus intéressants en pratique car représentatifs de la réalité) et sur des matériaux innovants tolérants aux dommages comme les composites. Ces spécificités (données temporelles massives, non régulièrement espacées, temps quasi-continu et sans a priori) rendent les données d'émissions acoustiques complexes à analyser.

2 Etat de l'art

La technique de l'émission acoustique est très utilisée depuis des décennies pour les essais de caractérisation de matériaux, l'objectif étant d'identifier différentes familles d'endommagement qui apparaissent au cours des sollicitations. Les instants d'apparition de certains événements sont aussi étudiés pour la surveillance. L'émission acoustique est cependant assez peu utilisée de cette manière pour des structures en service du fait de la complexité des données. Les très nombreuses applications industrielles exploitent en effet cette technique

comme détecteur d'événement : dès qu'un signal avec une certaine énergie est détectée, la structure est inspectée. Il n'y a donc pas de suivi de la progression des dommages comme cela peut être fait a posteriori sur des essais en laboratoire ou en utilisant d'autres techniques (notamment actives).

L'une des raisons de la faible utilisation des techniques l'émission acoustique pour la surveillance des structures en service vient du fait que la compréhension et la modélisation des sources et des émissions acoustiques représentent un sujet de recherche assez peu exploré. Par ailleurs, lorsque plusieurs sources sont actives à des instants proches (ce qui est majoritairement le cas sur les matériaux innovants), les signaux se convoluent et la résultante dépend des paramètres des matériaux, qui évoluent en fonction des sollicitations, ainsi que des capteurs. L'analyse de données, qui doit permettre d'aider à améliorer l'interprétation des signaux, reste confinée à l'utilisation de méthodes de partitionnement assez simples telles que l'algorithme des centres mobiles ou plus récemment les modèles de mélanges gaussiens, généralement disponibles dans les logiciels commerciaux accompagnant les capteurs et les systèmes d'acquisition.

L'algorithme de référence pour l'analyse des données d'émissions acoustiques est le suivant. La première phase consiste à extraire des descripteurs des signaux d'émissions acoustiques. (Nous passons les détails sur la phase de détection des transitoires d'intérêt dans les flux continus d'émissions [5]). Etant donnée la méconnaissance des sources, une vingtaine de descripteurs sont généralement extraits [4]. Ensuite, soit une analyse en composantes principales est employée pour réduire la dimension [3], soit tous les ensembles de n descripteurs sont considérés pour le partitionnement des données à l'aide d'un algorithme de classification automatique [10]. Pour chaque partition, un indice de validité (comme celui de Davies-Bouldin le plus souvent, ou une combinaison d'indices) est utilisé pour déduire le meilleur

sous-ensemble, ainsi que le nombre de classes. Les classes représentent les sources d'émissions acoustiques. Une analyse de corrélation entre quelques points dans certaines classes et les endommagements réellement en présence est réalisée avec d'autres mesures expérimentales [11].

Cette approche de référence présente certaines limites. D'une part, l'ordre des données n'est pas pris en compte, ce qui suppose que l'ordre d'apparition des endommagements n'est pas pertinent dans l'analyse; or, ceci est peu probable en termes de mécanique de la rupture. D'autre part, les indices de validité sont tous liés à la forme des classes dans l'espace des descripteurs et privilégient les classes compactes et bien séparées; or, il a été montré expérimentalement que ces critères ne donnaient pas toujours satisfaction [8], c'est pourquoi d'autres pistes ont été explorées [9, 1].

La méthode décrite par Ramasso et al. dans [9] vise à sélectionner les paramètres de la chaîne de reconnaissance de formes (notamment le nombre de classes et le sous-ensemble de descripteurs) de façon à mettre en évidence une chronologie dans l'apparition des classes. Sous l'hypothèse que les classes représentent des endommagements, il est peu probable qu'elles apparaissent toutes aux mêmes instants ou de manière très rapprochée dans le temps, en particulier en début d'essai [8]. Cela est cependant très souvent le cas avec l'algorithme standard.

L'algorithme proposé dans [9] suit l'algorithme standard excepté que l'évaluation de la pertinence des partitions ne repose plus sur les critères de forme des classes mais sur le nombre d'événements dans chaque classe. L'algorithme a été développé spécifiquement pour une structure composite particulière où le lien entre ces proportions et la chronologie des instants d'apparition des classes a été montré expérimentalement. Le travail proposé dans [7, 1] a permis de développer un critère directement lié à la chronologie d'apparition des classes.

Dans tous les travaux précédents, soit l'ordre des données n'est pas du tout pris en compte, soit il l'est seulement dans le critère d'évaluation. Dans la plupart des cas, la méthode de partitionnement utilisée donne les mêmes résultats quel que soit l'ordre des données. Dans le présent article nous présentons l'esquisse d'un modèle visant à créer une partition des données d'émissions acoustiques avec, à la fois, la prise en compte du temps et de la forme des classes de manière explicite dans la méthode de partitionnement et la possibilité d'ajouter des contraintes sur la séquence des instants de première apparition des classes (pour en prendre en compte d'éventuels a priori sur les déclenchements de certains endommagements).

3 Esquisse de la méthode

Un des modèles utilisés en analyse des émissions acoustiques est le modèle de mélange Gaussien

$$p(\mathbf{x}_1, \dots, \mathbf{x}_T; \boldsymbol{\theta}) = \prod_{t=1}^T \sum_{k=1}^K \pi_k \phi(\mathbf{x}_t; \mu_k, \Sigma_k)$$

où \mathbf{x}_t est un vecteur de descripteurs observés à l'instant t pour une séquence de longueur T , $\phi(\cdot; \mu, \Sigma)$ est la fonction de densité de la loi normale multivariée de moyenne μ et de variance Σ , π_k est la proportion de la composante k dans le mélange, K est le nombre de composantes et $\boldsymbol{\theta}$ est le vecteur des paramètres. L'échantillon $\mathbf{x}_1, \dots, \mathbf{x}_T$ est supposé iid. Les descripteurs communément utilisés représentent le contenu fréquentiel des signaux (puissance partielle, fréquence maximum ou moyenne...), l'énergie, la durée, le temps de montée, l'amplitude, etc. Les articles mentionnés en référence, par exemple [4, 5, 11] décrivent plus formellement ces descripteurs et en proposent d'autres.

Pour prendre en compte l'apparition progressive des classes dans le temps, nous proposons de rendre les proportions π_k croissantes en

fonction du temps, selon le modèle suivant :

$$p(\mathbf{x}_1, \dots, \mathbf{x}_T; \theta) = \prod_{t=1}^T \sum_{k=1}^K \pi_{tk} \phi(\mathbf{x}_t; \mu_k, \Sigma_k),$$

avec

$$\pi_{tk} = \frac{\alpha_{tk}}{\sum_{\ell=1}^K \alpha_{t\ell}}, \quad k = 1, \dots, K,$$

où $\alpha_{t1} = 1$ pour $t = 1, \dots, T$ et

$$\alpha_{tk} = \frac{\beta_k}{1 + \exp[-\gamma(t - \tau_k)]},$$

pour $k = 2, \dots, K, t = 1, \dots, T$ où

$$0 \leq \tau_2 \leq \tau_3 \leq \dots \leq \tau_K \leq T$$

$$\beta_k \geq 0, \quad \gamma \geq 0.$$

Le degré d'activation de la k^e classe est ainsi dépendant du temps au travers de la sigmoïde retardée de τ_k , d'amplitude β_k et de pente contrôlée par γ (supposée identique dans toutes les classes).

Pour t dans $] -\infty, +\infty[$, α_{tk} parcourt $]0, \beta_k[$, les π_{tk} restent entre 0 et 1 et leur somme vaut toujours 1.

Il est possible d'optimiser les paramètres de ce modèle avec une approche reposant sur l'algorithme EM [2] pour les paramètres μ et Σ comme dans le modèle de mélange standard, et d'optimiser les autres paramètres par maximisation sous contraintes de la fonction auxiliaire.

Nous travaillons actuellement sur le positionnement de cette approche par rapport aux modèles de Markov à temps continu [6] et sur sa validation avec des données réelles.

Remerciements :

Ce travail a été financé par l'EIPHI Graduate school (contrat ANR-17-EURE-0002) et le projet RESEM-COALESCENCE géré par l'Institut de Recherche Technologique Matériaux Métallurgie Procédés (IRT M2P) supporté par le Plan d'Investissement d'Avenir.

Références

[1] N. Chandarana, E. Ramasso, C. Soutis, and M. Gresil. Damage identification in composites through acoustic emission monitoring. In *9th Int. Conf. on Acoustic Emission (ICAE-9) & 61st Acoustic Emission Working Group Meeting (AEWG-61)*, Chicago, IL, USA, June 2019.

[2] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1) :1-38, 1977.

[3] N. Godin, P. Reynaud, and G. Fantozzi. Challenges and limitations in the identification of acoustic emission signature of damage mechanisms in composites materials. *Applied Sciences*, 8(8) :1267, 2018.

[4] S. Kattis. Noesis : Advanced data analysis, pattern recognition & neural networks software for acoustic emission applications. In *Kolloquium Schallemission, Statusberichte zur Entwicklung und Anwendung der Schallemissionsanalyse*, Fulda, 9-10 March 2017.

[5] M. Kharrat, E. Ramasso, V. Placet, and M.L. Boubakar. 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.

[6] Y.-Y. Liu, S. Li, F. Li, L. Song, and J.M. Rehg. Efficient learning of continuous-time hidden markov models for disease progression. In C. Cortes, N.D. Lawrence, D.D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 3600-3608. Curran Associates, Inc., 2015.

[7] E. Ramasso, P. Butaud, T. Jeannin, F. Sarasini, V. Placet, N. Godin, J. Tirillò, and X. Gabrion. Learning the representation of raw acoustic emission signals by direct generative modeling and its use in natural clusters identification. *Engineering Applications of Artificial Intelligence*, 2019. Submitted.

[8] E. Ramasso, X. Gabrion, DD. Doan, M. Kharrat, V. Placet, and L. Boubakar. Reconnaissance des sources acoustiques dans les composites à matrice organique : quel(s) critère(s) utiliser pour une classification non-supervisée des signaux? In *Congrès Français d'Acoustique*, pages 1611-1617. 2014.

[9] E. Ramasso, V. Placet, and M.L. Boubakar. 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] M. Sause, A. Gribov, A. Unwin, and S. Horn. Pattern recognition approach to identify natural clusters of acoustic emission signals. *Pattern Reco. Lett.*, 33 :17-23, 2012.

[11] Markus G.R. Sause. *In Situ Monitoring of Fiber-Reinforced Composites : Theory, basic concepts, methods, and applications*, volume 242. Springer Int. Publishing, springer series in materials science edition, 2016.