CRACK GROWTH MONITORING WITH HIERARCHICAL CLUSTERING OF AE

N. F. INCE¹, CHU-SHU KAO², M. KAVEH¹, A. TEWFIK¹ and J. F. LABUZ²
¹) Department of Electrical and Computer Engineering, ²) Department of Civil Engineering,
University of Minnesota, Minneapolis, MN, USA

Abstract

This paper presents a sequence of signal processing and hierarchical clustering techniques utilized to process signals with low signal-to-noise ratio (SNR) measured by multiple AE sensors. Noise and other extraneous events present major challenges for the detection and monitoring of AE signals generated by the inception of microcracks and their growth. Characteristics of AE waveforms released during controlled mode-I fracture are explored, and these characteristics are used for clustering AE and locating the fracture. With hierarchical clustering and signal “denoising” techniques, it is possible to locate the position of the crack plane with high accuracy using AE signals of poor quality, wherein the spatial distribution of clusters is indicative of crack propagation.

Keywords: Crack propagation, Signal processing, Hierarchical clustering, Denoising.

Introduction

Broadband acoustic emission (AE) events are widely used for characterizing damage, particularly as generated from the initiation and propagation of a fracture [1 - 3]. The waveform history of an AE signal is frequently used to investigate the health of a structure [2]. For example, the cumulative count of AE from different sensors has been used as a marker for damage quantification of structures under overload conditions.

Besides investigation of waveform history, as a more advanced application, AE events can be used to locate the damage. The spatial distribution of AE events estimated with techniques based on time of arrival can be used for determining position of a fracture. However, the use of AE events for locating damage is often obscured by noise and spurious events, which may cause a misinterpretation of the data. Even in controlled laboratory settings, it is difficult to account for all the sources of noise. Therefore, an AE system that automatically “learns” crucial patterns from data may provide clues for distinguishing between real events and extraneous signals, thus improving the spatial accuracy of localized AE and reduce false alarms. Consequently, the separation of AE events from noise is an important challenge.

Other investigators have employed signal processing techniques to improve the interpretation of AE data. A subspace approach based on principal component analysis and a self-organizing map was proposed to discriminate AE events due to crack propagation from noise of various origins [3]. In particular, their proposed algorithm consisted of two steps. In the first step, a combination of principal component analysis and differential time estimates were used to separate the noise from actual AE. A self-organizing map from a neural network was used to cluster the noise and AE signals into separate neurons. The algorithm was verified with two sets of data, and a correct classification ratio over 95% was achieved.
Accordingly, semi-supervised or unsupervised algorithms [3] are quite important in real life scenarios for distinguishing between signals of interest and noise. In this scheme, continuously stored events can be evaluated by their similarities in signal characteristics and spatial origin. For example, cluster identification plays a critical role in petroleum geomechanics, by computing cross-correlation of AE waveforms in both time and frequency domains, in order to precisely trace the structures activated during oil or gas stimulation [4]. Therefore, clustering methods may be used as an effective tool for the exploration and discrimination of AE events, where minimum or no prior knowledge is available.

In this paper, a sequence of signal processing and clustering techniques is presented, which can be used to locate damage in the form of a propagating fracture. Specifically, this work focuses on adaptive autoregressive modeling to reduce certain imperfections from the AE signals, and on the implementation of hierarchical clustering from the cross correlation across different events. The feasibility of the proposed techniques in determining the location of a fracture is presented by examining AE events recorded by eight sensors attached to a structure where microcracks have localized and crack propagation is occurring.

The experiments and the AE data sets recorded from two specimens during controlled crack propagation trials are described. Next, the signal preprocessing techniques used for enhancing the measured AE signals in the presence of noise and data acquisition imperfections are presented. This is followed by a description of a novel clustering technique to group the AE events. Finally, the spatial distribution of localized events is related to the results of the experiments in terms of crack monitoring.

**Fracture Testing**

Mode-I fracture tests using three-point bend specimens were conducted in a closed-loop, servo-hydraulic load frame with crack-mouth-opening displacement (CMOD) as the feedback signal, which provided the opportunity to monitor crack propagation during unstable growth (Fig. 1). The beams were composed of a brittle material called St. Cloud (Charcoal) granite, which produces significant AE before and during crack propagation. The rock has a density of 2.70 Mg/m³, P- and S- wave velocities of 5.4 and 3.2 km/s, Young’s modulus of 68 GPa, Poisson’s ratio of 0.23, and tensile strength of 13 – 14 MPa. Test 1, as a reference test, was conducted on a specimen with dimensions of 220 × 73 × 32 mm and a 4-mm notch cut in the bottom surface of the beam to induce AE events within a known region; i.e., the fracture initiated and propagated from the tip of the notch. Test 2 was performed on a specimen with a similar size (220 × 71.5 × 32 mm), but this beam was not notched; the boundary was smooth and the location of the eventual fracture was unknown, although the loading configuration promoted crack development at the center of the beam. The distance between the supports for both beams was the same, 169 mm, and both beam were loaded at a CMOD rate of 5×10⁻² mm/s.

Eight AE sensors were coupled to the front and back surfaces of the beams using a cyanoacrylate adhesive. The sensors surrounded a region approximately 50 mm in diameter adjacent to the lower-center region of beam (Fig. 1). The AE data were collected with high-speed, CAMAC-based data acquisition equipment, consisting of four two-channel modular transient recorders (LeCroy model 6840) with 8-bit analog-to-digital converter (ADC) resolution and a sampling rate of 20 MHz. The data acquisition system was interfaced with eight piezoelectric transducers (Physical Acoustics model S9225), and eight preamplifiers with bandpass filters from 0.1 to 1.2 MHz and 40 dB gain were used for conditioning the raw AE signals. The
The frequency response of these transducers ranged from 0.1 to 1 MHz, with a diameter of approximately 3 mm. All channels were triggered when the signal amplitude exceeded a certain threshold on the first sensor. This sensor is referred to as the “anchor” sensor. The recording window was 100 µs, with a 50-µs pre-trigger. In each test, a total of approximately 1600 events were recorded, with 62% and 45% of the events for Test 1 and Test 2, respectively, being located within an error of 5 mm using standard techniques [5].

**Signal Enhancement**

The AE data acquisition system was set up for targeting the P-wave arrivals. Due to high amplifier gain, a number of AE time histories in both tests were “saturated” (Fig. 2), and these events were uniformly distributed during the tests. Although the signals were saturated, the P-wave amplitudes remained in the linear range of the measurement system. Because the energy released was greater for the saturated events (the amplitudes was larger), the signal-to-noise ratio (SNR) of the P-waves of saturated AE events was 5.2 dB (Test 1) and 7.12 dB (Test 2) higher than the unsaturated AE events, indicating that the saturated events have P-wave amplitudes around two times larger than the unsaturated cases. Consequently, the arrival time can be determined with greater accuracy for the saturated signals, which means that the AE location have smaller error for the saturated events.

In order to extract information from signals with low SNR, the unsaturated events, a clustering procedure was used. Basically, the procedure captures those AE events that are correlated and then averages the time histories to improve the SNR. Approximately 82% (Test 1, notched) and 48% (Test 2, unnotched) of the total events recorded by the anchor sensor were selected for further analysis, as these were unsaturated signals. Within this group, several events contained spikes (Fig. 3), which probably originated from ADC sign errors. To remove these spikes, the prediction error of an adaptive autoregressive (AAR) model was employed. The autoregressive model is a stochastic signal modeling technique, which represents the current sample as a weighted linear combination of past samples and a white noise error [6]. In the adaptive version, the model parameters (weights) are time-varying, which can be adjusted to the non-stationary characteristics of the signal.
\[ y[n] = a_{1,n}y[n-1] + a_{2,n}y[n-2] + \ldots + a_{p,n}y[n-p] + e[n] \]  

(1)

Indeed, the AE events are also non-stationary, and the AAR model is expected to adapt to the time-varying signal characteristics. In this study, each AE was assumed as an AAR model of order \( p = 2 \). The model parameters \( a_{p,n} \) were estimated with the energy normalized least mean squares algorithm [7]:

\[
a_{p,n+1} = a_{p,n} + \frac{2\mu e[n]y[n]}{\lambda + y^T[n]y[n]} 
\]

(2)

where \( \lambda \) is a regularization constant for those values where the local energy of the signal is close or equal to zero. A learning rate of \( \mu = 0.005 \) and a regularization parameter of \( \lambda = 1 \) was used in this study. The model parameters and prediction error \( e[n] \) were computed sample by sample on normalized AE events, where the sample mean was removed (i.e. zero mean) and then normalized to the sample standard deviation (i.e., unit variance). When the absolute value of the prediction error is larger than a pre-defined threshold (8), related sample points were replaced with the predicted values of the AAR model. Two representative corrupted AE events and related prediction errors and corrected signals are given in Fig. 3. In some cases, where several consecutive spikes occurred, the AAR-based method failed to enhance the signal. These events were associated with large final prediction error values. A simple search algorithm was developed to find and remove those events with maximum absolute error values greater than eight.

Fig. 2. Time histories of (a) unsaturated and (b) saturated events. The P-wave amplitude of the saturated events was much higher, but still remained in the linear range of the system.

### Clustering of AE Events

Once the locations are computed from the time-of-arrival (TOA) information, the damage features (e.g. distributed or localized microcracking) are inspected visually by projecting the AE locations on a 2D surface [8]. The TOA is typically calculated by comparing the signal amplitude to a predefined threshold, where the earliest arrival is associated with the P-wave, as shown in Fig. 2. This type of method may produce misleading TOA information if the signal is noisy. Moreover, the spurious events recorded by the system may be included during the inspection of spatial distribution of cracks, and this may cause misinterpretation of the damage features. Therefore, before applying the amplitude threshold for TOA estimation, correlated events generated from microcracking were grouped and averaged to increase the SNR. A hierarchical clustering approach that uses the cross-correlation function computed between different events was applied.
As a first step, the normalized cross-correlation function $R_x[k]$ was computed for only 256 shifts between any two events on the signals $x[n]$ and $y[n]$ acquired at the anchor sensors:

$$R_x[k] = \frac{1}{(N-k)\sigma_x\sigma_y} \sum_n x[n]y[n+k], \quad |k| \leq 256$$  \hspace{1cm} (3)

Then, a correlation matrix was constructed that keeps the maximum value of the absolute cross-correlation function between all event pairs. This correlation matrix was used to build a hierarchical cluster dendrogram [2]. The correlation matrices of the two datasets of the fracture tests are shown in Fig. 4.
The average linkage method was employed to construct the dendrogram, and then it was cut at level 0.2 in order to cluster those events that have average cross correlations equal or larger than 0.8. At this level, 212 and 82 clusters were obtained with two or more members for the data of Test 1 (notched beam) and Test 2 (unnotched beam), respectively. Two examples of the overlap plot of clusters with different numbers of AE events are shown in Fig. 5, where the average correlation values are 0.875 and 0.810, respectively.

This step was followed by computing the averages of each cluster. Indeed, due to averaging, the random components in the data were suppressed, where repetitive components remained the same. Therefore, the average acoustics computed from each cluster have higher SNR than individual AE events (Fig. 6). In this scheme, averaging will cancel out the uncorrelated noise in the signal and the repetitive components will remain the same, which improves the SNR of the signals in their amplitude with a factor of $\sqrt{N}$, where $N$ is the number of correlated events in a
cluster. In order to improve the SNR by a factor of more than two, clusters with at least five members were used. Furthermore, the reliability increases since it may be difficult to observe repeated recording by chance in real life situations.

In order to locate the “source” of the clustered AE waveforms (actually several sources of similar character), the envelope of the signal was computed with a Hilbert transform (Fig. 6). The TOA is then computed on the envelope of the signal if it exceeds a threshold of 0.5. After calculating the arrival information on each sensor location, a location algorithm [5] was used to estimate the 3D position of the “source.”

Results

Figure 7a shows the estimated positions of all the unsaturated AE events, regardless of the location error. For the notched beam in Test 1, the estimated AE locations clustered around the notch area and extended along the Y-axis. For the unnotched beam in Test 2, however, the AE locations were more scattered on one side of the centerline (mid-span), and it is more difficult to pinpoint the fracture location. To verify the AE locations, upon the completion of each test, the front (Z = 32 mm) and back (Z = 0) surfaces of the specimens were scanned with 1200 dpi resolution and 24 bit color. After enhancing the contrast level, the crack trajectories were carefully mapped onto the image (Fig. 8).

Using hierarchical clustering on both specimens, the clustered events (Fig. 7b) and the larger clusters with at least five members (Fig. 7c) accurately find the fracture position. Most of the clustered events in Test 1 are localized around the notch tip, with a vertical extension expected in the mode I fracture test. Similarly, in Test 2, although the individual events were dispersed on the 2D projection plane, the hierarchical clustering analysis extracted the fracture location from a cloud of AE events (Figs. 7b, 7c) and clearly indicated the fracture path. In Fig. 9, the position of AE clusters in Test 2 (Fig. 7c) was overlapped by the mapped crack trajectory, indicating that the AE clusters with more than five events were highly correlated with the actual fracture.

Plots of event indices of the largest clusters with decreasing number of members are shown in Fig. 10. It is observed that the clusters were formed from consecutive or closely spaced events in time. Note that the AE recorded in both tests were mostly after reaching peak load (failure). For the notched beam (Test 1), larger clusters tended to occur well after failure, where the peak load was reached at the time of event number 42. This observation is reasonable, since the notch promoted the fracture earlier in loading, and the well-localized fracture created more space- and time- correlated events. For the unnotched beam (Test 2), the trend was not obvious due to

(a) the cluster analysis being performed on only 48% of the events (i.e. unsaturated events only) with low SNR, and
(b) the number of events in each cluster was fewer.

If AE with higher SNR were used, a similar result would be expected from the hierarchical clustering analysis. Note that none of the clusters span the entire duration of the experiments, which is reasonable for a mode-I fracture reaching a critical opening with a traction-free region, where no AE is expected.
**Conclusions**

The use of low signal-to-noise (SNR) ratio events for damage detection was investigated by introducing various approaches for clustering AE signals. Specifically, adaptive autoregressive modeling was used to remove spikes from the recorded waveforms as they cause poor estimation of correlations between events and the P-wave arrivals using a simple threshold. This was followed by the calculation of the cross-correlation between events to build clusters of AE. By computing the averages of each cluster, the SNR of the signals was improved by a factor of the square root of the number of correlated events. In particular, this step suppressed the random fluctuations in the baseline segment before the P-wave arrival. As a result, the time of arrival was detected with higher accuracies with the low-amplitude AE signals. This was accomplished by computing the envelopes of averaged events in each cluster and then comparing its amplitude to a predefined threshold. It was observed that the clustered events can accurately locate the position of a fracture, even with AE of low SNR. As the cluster size increases the spatial location of them were highly correlated with the location and propagation direction of the cracks. Further-
more, a well-localized fracture exhibits both space- and time- correlated events. These results suggest that the proposed signal processing algorithm can be used to build a damage detection system that has low false positive rates and can work with higher accuracies in real life applications where the SNR of AE are low.

Fig. 8. Crack trajectories of Test 1 (left) and Test 2 (right) projected on the X-Y plane.

Fig. 9. Crack trajectories and locations of clusters in Test 2 (no notch beam).

Acknowledgements

Partial support was provided by the National Science Foundation, Grant CMMI-0825454.

References

Fig. 10. The distribution of event indices of each cluster for both datasets.