Preprint, accepted for MFPT59, Virgina Beach, US, April 2005

EXTRACTING INFORMATION FROM CONVENTIONAL AE FEATURES FOR ONSET DAMAGE DETECTION IN CARBON FIBER COMPOSITES

Runar Unnthorsson[a] and Niels Henrik Pontoppidan[b] and Magnus Thor Jonsson[a]

[a] Department of Mechanical Engineering, University of Iceland
Hjarðarhaga 2-6, IS-107 Reykjavik, Iceland
[b] Informatics and Mathematical Modeling, Technical University of Denmark
Richard Petersens Plads 321, 2800 Lyngby, Denmark
Corresponding author Runar Unnthorsson, runson@hi.is

Abstract: We have analyzed simple data fusion and preprocessing methods on Acoustic Emission measurements of prosthetic feets made of carbon fiber reinforced composites. This paper presents the initial research steps; aiming at reducing the time spent on the fatigue test. With a simple single feature probabilistic scheme we have showed that these methods can lead to increased classification performance. We conclude that: the derived features of the TTL count leads to increased classification under supervised conditions. The probabilistic classification scheme was founded on the histogram, however different approaches can readily be investigated using the improved features, possibly improving the performance using multiple feature classifiers, e.g., Voting systems; Support Vector Machines and Gaussian Mixtures.

Key Words: Acoustic Emission; Carbon fibres; Data fusion; Fatigue testing; Probabilistic classification; Supervised learning

Introduction: During the design phase of a prosthetic carbon fiber foot at Össur hf., prototype models are built, evaluated, and improved to meet the design criteria. For evaluation several testing methods are used; visual tests, stiffness measurements and fatigue testing. Feet are mainly subjected to dynamic loading, which sometimes requires low loads to initiate and propagate faults [16]. Because of this fatigue testing is extremely important part of the design process. Fatigue tests of composites can take very long time, up to several weeks. It is therefore valuable for an engineer, that wants to test a prototype, to be able to see if it will fail early during the test. Considerable time can be saved by this. By shortening the test time, feet can be; designed and developed in less time and at a lower cost than previously possible. During the fatigue testing of an prosthetic feet, the stiffness is monitored and visual tests are also performed regularly. However, visual tests can only detect faults that extend to the surface of a foot or affect it in some way. In order to be able to predict the fatigue strength of a foot undergoing fatigue testing, other NDT methods can be used for obtaining data.

When microstructural changes occur in composites, energy is released and transient stress waves are generated. These stress waves are called called Acoustic Emissions (AE) [16]. The stress waves travel through the composite and when they reach the surface it will vibrate. The small surface displacements generated by the vibration of the composite can be detected by using appropriate sensors. According to Duesling [3] piezoelectric sensors are most popular. AE can be generated by several types of damage that occurs in the material, e.g. fibre breakage, matrix cracking, delamination etc. Because microstructural damage generates AE the method has the potential of detecting damage in its early stages, much earlier than possible by monitoring deflection or performing visual inspection. Therefore a microstructural damage that starts to formate in the last cycles of a fatigue test can be detected, a damage that would pass undetected by the other two previously mentioned methods.

AE signals are not only generated by damage, they can also come from other sources, such as; the testing machine [16], electrical disturbances [16, 5], friction due to rubbing of parts [12] and the friction generated by opening and closing of matrix cracks [13]. According to Hamstad [7] AE is also generated under loading because of the different material properties of the fibers and the matrix.

Friction generated AE can provide important information about the condition of an composite [1]. Some researchers have attempted to filter the AE signal generated by friction from the total signal in order to better detect AE from damage, but this can be difficult [15]. The fact that a damage only emits AE once but friction many times suggests that approaches based on friction will be more robust.

An AE parameter analysis is the conventional way of performing an AE monitoring. The results from using standard AE parameters, especially amplitude, are contradicting in the literature [5, 2, 8, 3]. Amplitude suffers from attenuation. The main reasons for the attenuation are; geometric spreading, dispersion, internal friction and scattering [11]. Also, Prosser et al. [10] reported that the same type of damage doesn't always produce AE with the same amplitude. This indicates that amplitude is not a very good feature. Godin et al. [6] claimed that conventional AE analysis cannot distinguish between different AE sources and suggested that more advanced methods, such as multivariate analysis and classifiers, should be used. Similar comments were made by Tsamtsakis et al. [15]. They suggested that new parameters, like force or displacement, should be added.

This paper reports the results of using several methods to extract information from conventional AE features. The features were obtained using a commercial AE acquisition system (AESmart 2000 from DECI Inc) and consist of a fixed set of standard AE features. The emphasis was put on generating features which could be used for fast and reliable detection of damage onset. It was also considered important for the detection method to be robust. Robustness was believed to be obtained by basing the features on friction based AE.

The outline of the remaining of the paper is as follows: First the setup of the acquisition system is explained then the methodology is explained. It starts with explaining how new features are derived from the available feature set by both applying sensor fusion and using moving window second order moment. The section ends with a description a simple classification system based on the assumption that faulty specimens have inhibit greater AE activity than normal ones, thus threshold on the features can be used for classification. We compare classification properties of the different features by using using receiver-operator characteristics (ROC) curves. The features are compared and discussed – identifying the useful features. Finally the concluding remarks are made and future work is outlined.

System Setup: The fatigue test specimens were two types of prosthetic feet differing slightly in stiffness and size. Foot no. 1 is stiffer and smaller than foot no 2. The construction of both feet is the same, they are made from unidirectional carbon fiber reinforced epoxy and woven mats are used for the top and bottom layers for nicer look. The test was performed according to ISO 10328 specifications, with a “Foot/Limb” test system at Össur’s testing facilities. During the fatigue test, two actuators were used to flex the foot for 2×10^6 cycles at 1.5 Hz. The maximum load for each actuator was kept constant and the deflection was monitored. No change in maximum deflection was measured, which means that no stiffness reduction was observed during the tests. For data collection SE9125-MI data transducers and the AESmart 2000 AE acquisition system, from from DECI Inc., was used.

According to Dunegan [4] the system splits the AE signal into low frequency (LF) and high frequency (HF) signals. The LF signal contains frequencies between 20 to 60 kHz and the HF signal contains frequencies between 100 to 500 kHz. Data is only collected from one sensor at a time, each time the system switches between the sensors it stores the data in an Excel file. The data that is stored is not the actual AE signal but instead the following features that the system extracts from this signal; time of the recorded event, HF counts above threshold (TTL counts), HF peak amplitude, LF peak amplitude, ratio of the HF/LF peak amplitudes, event count, time difference between HF and LF signals. Other features are also stored by the system but, they are irrelevant to the test.

The system was set up to switch between the two sensors, i.e. one on each foot, using 6 second dwell time. The monitoring was performed twice a day for 4 to 6 hours each time. The gain for both the LF and HF gain was set to 60 dB. The threshold setting for the LH and HF was set to 200 mV. The time interval for counting the TTL counts was set to 1 ms.

Data preprocessing: During data acquisition each foot was monitored for periods of six seconds, before the system switched started monitoring the other foot. Each entry in the data file is the result of the following events

1. AE signal crosses the trigger threshold
2. Acquisition board computes the set of features from a 1 ms analysis window
3. Acquisition board goes back to detecting threshold crossing in AE signal. Further the system might also change foot if the 6 s period has ended

<u>Moving window second order moment</u>: We compute the second order moment over a fixed number of consecutive data entries, move the window and repeat. This gives the moving window second order moment that peaks when the local mean value change, and also when

the local variance increase. So this is just a simple "amplification" of changes in a time series. Besides from giving a different measure (variance instead of values) this method also performs filtering since a number of consecutive samples are used for each value.

$$\check{\mu}(n) = \frac{1}{N} \sum_{n'=0}^{N-1} x(n - n') \tag{1}$$

$$\check{\sigma}^2(n) = \frac{1}{N-1} \sum_{n'=0}^{N-1} \left(x(n - n') - \hat{\mu}(n)\right)^2 \tag{2}$$

Time normalization: Each data entry comes from a 1 ms analysis window triggered by a threshold crossing. So we can imagine several regimes (explaining the data entries in the 6 s periods where the system monitors each foot).

1. few data entries with a few high readings – occasional low activity
2. many data entries with a few high readings – steady low activity
3. few data entries with many high readings – occasional high activity
4. many data entries with many high readings – steady high activity

This means that if we omit the time information and look at the feature values, we cannot differ between the occasional and steady activity. Instead we fuse the features in each 6 s period. We deliberately use the term fuse, even though most features are just summed up during the period. But for the time entry we take the earliest value, for the event count we take the maximum, and for the ratio we recalculate; since the original ratio computation was constrained. Another benefit of time normalization is reduction of data size as several events are combined into one.

We can also apply the moving window second order moment to the time normalized data in order to further enhance changes.

Energy normalization: In the LF/HF peak amplitude we have the peak amplitudes, we already have the number of TTL counts in a 6 s period. Multiplying those two quantities and sum over the events in the period gives an upper bound estimate of the "energy".

$$E = \sum_i TTL_i P_i \tag{3}$$

Filtering: One way of decreasing feature variance is low-pass filtering, essentially this forces the classification system to evaluate the feature value for consecutive examples. This smears out occasional high values, but also the sudden steps in the feature, thus this delays the significant changes. Further notice that this is related to the post classification processing described last in the following section.

Feature processing: We adopt a very simple setup for supervised single feature classification of the specimen condition. Assuming that a damaged specimen generate a "loud" AE signal, it will also generate many TTL counts with high peak amplitudes and many event counts. Thus as a rule of thumb: The feature value in a faulty specimen should be larger than the normal. In order to compute the feature thresholds, we compute the histograms of each feature on two sets of data, a normal and a faulty. Histograms are not directly well suited for classification tasks, so instead we use the normalized cumulated sum of the histogram, corresponding to a sample of the true cumulated density function (CDF). The CDF gives the probability that the feature is less or equal to τ.

$$P(x \leq \tau) = \int_{x=-\infty}^{\tau} p(x) \tag{4}$$

A good feature for classification will behave such that $P_{\text{normal}}(x \leq \tau) \gg P_{\text{faulty}}(x \leq \tau)$ is fulfilled, not everywhere but in an interval. Actually $P_{\text{normal}}(x \leq \tau) = X, X \in [0;1]$ says that X of the normal examples have a feature value x less or equal to τ. Obviously this means that thresholding with this value of τ corresponds to $1 - X$ false alarms, and that $1 - P_{\text{faulty}}(x \leq \tau)$ of the faulty specimens is correctly detected (see further [14]). Varying τ from the smallest to the largest observed value whilst tabulating $1 - P_{\text{normal}}(x \leq \tau)$ and $1 - P_{\text{faulty}}(x \leq \tau)$ gives the Reciever Operator Characteristics (ROC) shown in Figure 1.

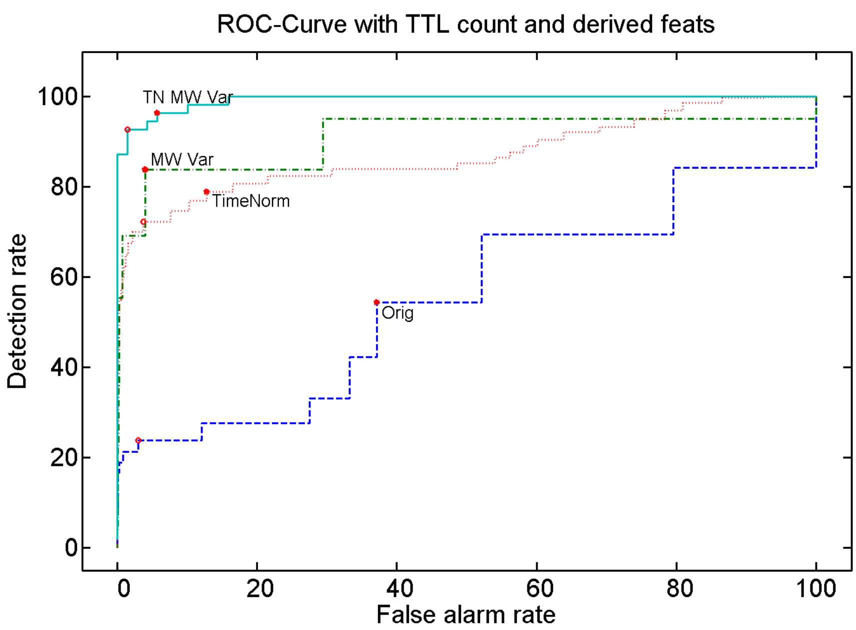

Figure 1: ROC-curves for the TTL count and the three derived features. The o's and stars on the curves indicate the performance of the two optimal thresholds. Notice the consensus between the different optimality criterions wrt. ranking of the features – also the Neyman-Pearson criterion at 5% agrees. The figure also show our conclusion, that the time normalization (data fusion) and moving window variance (processing) increase the performance of the system.

$$\text{Detection rate } D(\tau) = 1 - P_{\text{faulty}}(x \leq \tau)s \tag{5}$$
$$\text{False alarm rate } F(\tau) = 1 - P_{\text{normal}}(x \leq \tau) \tag{6}$$

This approach with a single dimension classifier is very simple, with multiple dimensions, like using several of the obtained features, either pattern recognition or voting systems should be explored.

We adopt two rules for obtaining the optimal threshold:

1. Maximize the difference between the detection rate and false alarm rate
2. Minimize the absolute distance to the optimal point (100% detection, 0% false alarm)

Also applicable is the Neyman-Pearson Criterion[14], where the best classifier/threshold is the one with the highest detection rate given a set false alarm rate, say 5%.

Having decided on a threshold we also have knowledge of the false alarm rate of the classifier under normal conditions. In many cases occasional false alarms will occur, an we could filter out with a moving average low pass filter. The length of the filter and the significant number of alarms for the filtered signal can be obtained through the Binomial hypothesis testing (this is further described for a unsupervised setting in [9]).

In short we can vary the threshold in order to balance the rate of false alarms versus detection, however it is a trade-off. Direct improvement is only available through a better classifier. Another available trade-off is between classification accuracy and delay. If we know that the false alarm rate is 5 out of 100, we could wait until 12 out of 100 was classified as faulty; however this will introduce a delay since we need to gather enough fault-classifications.

Results and Discussion: Table 1 shows the statistical parameters for the ROC curves corresponding to the original features, i.e. those provided by the acquisition system.

Feature	false 1	detect 1	false 2	detect 2	Area
TTL count	3	24	37	54	0.55
HF amp	3	10	18	15	0.15
LF amp	0	2	25	8	0.07

Table 1: Original data

None of these features are able to discriminate well between normal and faulty signals under the assumption that faulty signals are louder. The LF amp with the assumption that normal signals are loud works quite well (area under ROC-curve 0.93), this could fit the "steady low amplitude" regime descried under *time normalization*.

Applying a *moving window variance* on the original features the classification performances is considerably improved. Now take a look at the features generated by taking a moving window variance of the original data, shown in Table 2, then we have features with considerably better detection performance.

This increase in performance is in some cases followed by increased false detection rate. However, these features are generally better suited for classifying as indicated by the much larger area under the ROC curves. The performance of the time normalized data is shown in Table 3.

Feature	false 1	detect 1	false 2	detect 2	Area
TTL count	4	84	4	84	0.91
HF amp	16	87	16	87	0.84
LF amp	36	79	36	79	0.67

Table 2: Moving Window Variance of Original data

Feature	false 1	detect 1	false 2	detect 2	Area
TTL count	4	72	13	79	0.87
HF amp	11	49	22	56	0.60
LF amp	3	8	48	45	0.31
Event count	43	91	26	73	0.78

Table 3: Time Normalized data

These features result in better classification than with the original features (Table 1), however not as good as with the moving window variance. However the combination, by applying moving window variance on those features we hoped for similar improvement as was observed for the original data. Table 4 lists the ROC performance of the new set of features.

Feature	false 1	detect 1	false 2	detect 2	Area
TTL count	1	93	6	96	0.99
HF amp	13	93	13	93	0.92
LF amp	23	84	23	84	0.81
Event count	83	100	54	55	0.52

Table 4: Moving Window Variance of Time Normalized data

According to the Moving window variance of the normalized TTL counts is definitely a good feature.

Feature	false 1	detect 1	false 2	detect 2	Area
TTLxHF	13	56	13	56	0.54
TTLxLF	5	19	40	35	0.29
TTLx(HF+LF)	5	29	42	51	0.43

Table 5: Energy Normalized data

In order to see if the classification performance of the feature created by applying moving window variance on the normalized TTL counts could be improved further, a filtering was applied. As Table 6 shows that the filtering improved the false alarm rate compared to Table 4, but the area below the curve is slightly decreased as the detection rate is also decreased. Figure 2 shows the time series of both features derived from the TTL count as well as their respective classification. The various preprocessing steps reduce the local variance of the classification.

Feature	false 1	detect 1	false 2	detect 2	Area
NTTL variance	0	95	0	95	0.98

Table 6: Filtered moving window variance of the normalized TTL counts

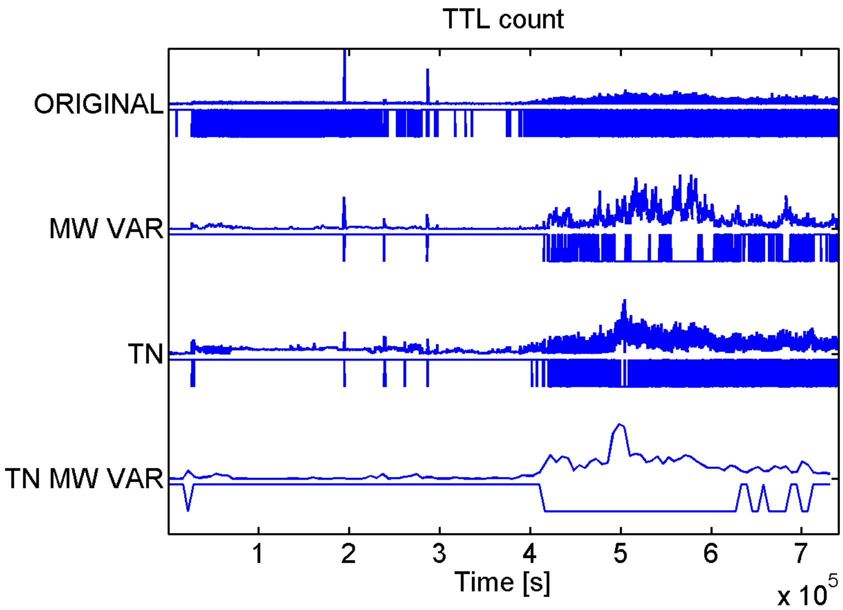

Figure 2: The time series of the TTL count and three derived features. Each entry in the features has been classified; this binary output is plotted just below the feature (the lower value is faulty, upper is normal). It is believed that the specimen turns faulty after $4 \cdot 10^5 s$. The uncertainty using the original TTL count is clearly visible during under both normal and faulty conditions, as the classification switches all the time.

Conclusion: We have analyzed simple data fusion and preprocessing methods on Acoustic Emission measurements of prosthetic feets made of carbon fiber reinforced composites. This paper presents the initial research steps; aiming at reducing the time spent on the fatigue test. With a simple single feature probabilistic scheme we have showed that these methods can lead to increased classification performance. We conclude that: the derived features of the TTL count leads to improved classification and damage detection in a supervised setup.

The probabilistic classification scheme was founded on the histogram, however different approaches can readily be investigated using the improved features, possibly improving the performance using multiple features classifiers, e.g., Support Vector Machines and Gaussian Mixtures.

Acknowledgements: The authors gratefully acknowledge Össur hf. (`www.ossur.com`) for providing samples and access to their testing facility. The research of Runar Unnthorson was supported by a research grant from RANNIS, the Icelandic Research Council (`www.rannis.is`). His visit at DTU was supported by grants from the Helga Kristjánsdóttir and Sigurliði Kristjánsson Memorial Fund and Landsvirkjun's (`www.lv.is`). Niels H. Pontoppidan was supported by the European Commission through Growth Project AE-WATT (`mole.imm.dtu.dk/aewatt/`).

References:

[1] J. Awerbuch and S. Ghaffari. Monitoring progression of matrix splitting during fatigue loading through acoustic emission in notched unidirectional graphite/epoxy composite. *Journal of Reinforced Plastics and Composites*, 7(3):245–64, 1988.

[2] S. Barre and M.L. Benzeggagh. On the use of acoustic emission to investigate damage mechanism in glass-fibre-reinforced polypropylene. *Composites Science and Technology*, 52(3):369–376, 1994.

[3] L.A. Duesing. Acoustic emission testing of composite materials. *Reliability and Maintainability Symposium, 1989. Proceedings., Annual*, pages 128–134, 1989.

[4] H.L. Dunegan. An alternative to pencil lead breaks for simulation of acoustic emission signal sources, August 2000.

[5] M. Giordano, A. Calabro, C. Esposito, A. D'Amore, and L. Nicolais. An acoustic-emission characterization of the failure modes in polymer-composite materials. *Composites Science and Technology*, 58(12):1923–1928, 1998.

[6] N. Godin, S. Huguet, R. Gaertner, and L. Salmon. Clustering of acoustic emission signals collected during tensile tests on unidirectional glass/polyester composite using supervised and unsupervised classifiers. *NDT and E International*, 37(4):253–264, 2004.

[7] Marvin A. Hamstad. Acoustic Emission Primer, Lecture delivered by Dr. Hamstad at the acoustic emission working group meeting on august 4, 2003 (aewg-46). internet, 2004. http://www.engr.du.edu/profile/frame_bio.html.

[8] K. Ono. Acoustic emission behavior of flawed unidirectional carbon fiber-epoxy composites. *Journal of Reinforced Plastics and Composites*, 7(1):90–105, 1988.

[9] Niels Henrik Pontoppidan, Jan Larsen, and Sigurdur Sigurdsson. Non-statrionary Condition Monitoring with the AEWATT Toolbox. Submitted to MFPT 59, 2005.

[10] W.H. Prosser, K.E. Jackson, S. Kellas, B.T. Smith, J. McKeon, and A. Friedman. Advanced waveform-based acoustic emission detection of matrix cracking in composites. *NDT and E International*, 30(2):108, 1997.

[11] William H. Prosser. Applications of advanced waveform-based ae techniques for testing composite materials. *Proceedings of the SPIE conference on Nondestructive Evaluation Techniques for Aging Infrastructure and Manufacturing: Materials and Composites*, pages 146–153, 1996.

[12] S.R. Ravishankar and C.R.L. Murthy. Characteristics of ae signals obtained during drilling composite laminates. *NDT and E International*, 33(5):341–348, 2000.

[13] N. Sato and T. Kurauchi. Interpretation of acoustic emission signal from composite materials and its application to design of automotive composite components. *Research in Nondestructive Evaluation*, 9(3):119–136, 1997.

[14] Clayton Scott and Jeffrey Silverman. The Newman-Pearson Criterion. internet, 2004. `http://cnx.rice.edu/content/m11548/latest`.

[15] D. Tsamtsakis, M. Wevers, and P. De Meester. Acoustic emission from cfrp laminates during fatigue loading. *Journal of Reinforced Plastics and Composites*, 17(13):1185–1201, 1998.

[16] M. Wevers. Listening to the sound of materials: acoustic emission for the analysis of material behaviour. *NDT and E International*, 30(2):99–106, 1997.