 DOI: 10.1784/insi.2007.49.8.451

Damage diagnosis using a kernel-based method

A Chattopadhyay, S Das and C K Coelho

Paper presented at WCEAM-CM2007, the Second World Congress on Engineering Asset Management and Fourth International Conference on Condition Monitoring, Harrogate, UK, June 2007.

This paper presents the use of a kernel-based machine learning technique, popular in the field of pattern recognition, to detect and classify various forms of damage states in both isotropic and anisotropic structures. A classification algorithm based on one-class Support Vector Machines (SVMs) is used for damage detection. The SVMs use a Gaussian kernel to map the input attributes to the high dimensional feature space and the transformed features are linearly separated by a decision plane. A procedure for obtaining the optimal value of the hyperparameter that controls the smoothness of the kernel is described. The type of damage addressed in this paper includes a combination of loose bolt and fatigue crack damage in a single lap, Al 6061-T651, bolted joints. Graphite/epoxy composite laminates with different types of damage are also studied, taking into account uncertainties in the measurement and material properties. The results show that the algorithm is able to accurately distinguish between different torque states and changes in crack length in the bolted joint sample. In anisotropic media, the algorithm was able to detect and classify various damage signatures with significant accuracy, using mutual information of two sensors. The algorithm was able to produce similar levels of accuracy when variability due to material properties was introduced.

Keywords: Structural health monitoring, time embedding technique, kernel, Support Vector Machines, fatigue crack, bolted joint, classification, anomaly detection.

Aditi Chattopadhyay is a Professor of Mechanical and Aerospace Engineering and the Director of the Adaptive, Intelligent, Materials and Systems (AIMS) Center, Ira A Fulton School of Engineering at Arizona State University. She received her MS and doctorate degrees in Aerospace Engineering from the Georgia Institute of Technology. She has held a Research Scientist position at the NASA Langley Research Center and summer faculty position at the NASA Ames Research Center. Her research focuses on smart materials and adaptive structures, structural health monitoring, mechanics of composites and structural dynamics. She is a Fellow of AIAA and ASME.

Santanu Das PhD is currently a postdoctoral associate at the Department of Mechanical and Aerospace Engineering at Arizona State University. He received his BE degree in Electrical Engineering from JGEC, India and MS degree in Microsystems Engineering from University of Applied Sciences, Furtwangen, Germany. After his masters degree, he worked for the European Aeronautics Defense and Space Company (EADS), Munich, Germany. His research focuses on structural health monitoring for heterogeneous structures.

Clyde Coelho is a PhD student under the supervision of Dr Aditi Chattopadhyay at the Department of Mechanical and Aerospace Engineering at Arizona State University. He received his B.S.E degree in Aerospace Engineering from Arizona State. His current research focuses on damage detection in complex metallic structures, in particular, different types of aerospace joints using machine-based learning techniques.

Aditi Chattopadhyay, Santanu Das and Clyde K Coelho are with the Department of Mechanical and Aerospace Engineering, Arizona State University, PO Box 876106, Tempe, AZ 85287, USA. Tel: +1 (480) 965-9342/+1 (650) 604-0964/+1 (480) 727-0007; E-mail: Aditi@asu.edu/Santanu.Das@asu.edu/Clyde.Coelho@asu.edu

Nomenclature

α_i	=	Lagrange multiplier
Θ	=	Feature map
w	=	Vector from origin to hyperplane
ρ	=	Offset of hyperplane from the origin
ξ_i	=	Slack parameter
K	=	Radial Basis Function kernel
ν	=	Classification error limit
F	=	Decision function
x	=	Input signals
z	=	Input signals mapped to high dimensional space

Introduction

The principal objective of structural health monitoring (SHM) is to be able to detect the presence of defects close to the nucleation stage so that steps can be taken to avoid system or sub-system level failure. Capabilities for *in situ* interrogation, therefore, must be advanced by coordinated research in materials physics, fundamental sensor design, sensor electronic systems and their implementation, sensor management and data processing, and information-level analysis. Although data management occurs last in the process of analysing physical material properties and determining required actions, it is first and most important at defining systems-level requirements on sensor performance, the correlation between sensor information and materials degradation and damage, and the associated levels of confidence in that relationship. This paper presents a kernel-based method to diagnose structural defects and presents some examples on the implementation of this technique for isotropic media with complex geometries and anisotropic laminates with different types of damage. Thus, the performance of the classifier is tested taking into account changes in the type of material, geometry and damage type. An anomaly detection tool is developed using one-class Support Vector Machines (SVMs) which utilise only the nominal state of the system while looking for different fault modes that might occur during its operation.

The fundamental idea behind the kernel-based method is to use a function Θ that maps a given dataset, $x \in X$, into the Hilbert space and the transformed features are then manipulated to perform a specific task. Here the objective is to separate different defect signatures (patterns) and, specifically, Radial Basis Function (RBF) has been used to do the mapping. Throughout this research, Support Vector Machines have been used as a data driven tool to separate the transformed images of the original dataset. There are many different applications of SVM methods currently being used for anomaly detection and pattern recognition. The advantage of using these methods for classification lies in the fact that they produce reasonably accurate results while using only a fraction of the computational time of other commonly used algorithms[15-17]. Data driven methods can be supervised or unsupervised in nature, based on the way they are trained on the available historical data. In an unsupervised learning technique, the model experiences the nominal behaviour of the system only and would identify unseen abnormalities, if that occurs. Here, the system is trained only with the

dataset or attributes extracted from those datasets that characterise the normal behaviour of the system and considered as 'observed' features. However, in supervised learning it is assumed that the given system features have already been classified by a human expert into *m* of *n* categories, based on some prior knowledge, and the training is performed on all possible categories. The drawback of this method is that every possible behaviour for a system must be known for accurate classification, which is not practical for real world problems where an infinite number of off-normal behaviours can exist. The one-class SVM, an unsupervised technique used in this study, identifies outliers in the test dataset and characterises anomalous behaviours in vibration signatures.

In the current research, single lap bolted joints have been chosen as the test specimens because fatigue damage is often induced in them due to the stress concentrations caused by fretting at the interfaces and variations in part geometry. One particular problem of fatigue in bolted joints occurs when a bolt comes loose and the life of the joint is dramatically reduced[2-6]. The role of torque or preload in a bolted joint is to provide a clamping load at the interface resulting in a friction type joint where the load is transferred directly from one lap to another through friction. A loose bolt converts the joint into a bearing-type joint where load is transferred through the bolthole, resulting in faster nucleation and propagation of cracks. A wave-based technique has been adopted to detect and characterise defects in the single lap bolted joint using piezoceramic transducers, mounted on the surface of complex geometry. These transducers can be used both as actuator or sensors depending on the requirement of application and have a wide variety of applications in the field of active diagnosis. The paper uses an active wave-based technique with narrow band burst signal and the choice of the input excitations makes the detection and quantification task much simpler because of low dispersion effects and distinct 'wave packet' characteristics when compared to broad band excitations like chirp signals. Experimental data has been collected from a single lap bolted joint set-up subjected to fatigue loading. An optical telescope was used to measure the length of the crack every time the fatigue loading was paused for measurements. Various cases of torque level and crack length were studied using data collected from the multiple, surface-mounted sensors. The goal is to classify different defect modes due to fatigue under the influence of torque loss in the bolt.

High-performance lightweight composites are increasingly popular in the field of aerospace, automobile, weapons and mechanical systems. Though composites offer a high strength-to-weight ratio and can be tailored based on application, they also exhibit specific types of multiscale and complex damages such as delaminations, matrix cracks, fibre breakage, etc. The second phase of this research demonstrates the application of the proposed kernel-based classifier to detect and classify the signature characteristics due to the presence of various types of damage such as delaminations, drilled holes, notches and saw-cut in composite laminates so that the status of the structure can be ascertained. A sufficient number of test samples was used to account for uncertainties in material properties.

The paper is organised as follows. Section 2 discusses the theory behind the SVMs classifier and provides a brief description on the time-embedding technique used to generate the features of the input space. This is followed by some details on the conducted experiments. Section 3 presents some conclusions and discussion on the outcomes of the detection algorithm.

Support Vector Machines: a kernel classifier

The Support Vector Machines (SVMs) is a machine learning technique that maps the extracted feature vector of input space to high-dimensional domain known as feature space and thereafter constructs an optimal hyperplane to separate the features by solving a quadratic optimisation problem. Often, in real world, patterns are nonlinearly separable in input space. The idea is to map the n-dimensional vectors of the input space into a high-dimensional (possibly infinite dimensional) feature space (Figure 1) where the transformed image of the input patterns are linearly separable. This can be achieved using Cover's theorem, which states that a multi-dimensional input space can be transformed to a feature space where the transformed image of the input patterns are linearly separable provided the transformation is nonlinear and the dimensionality of the feature space is high enough.

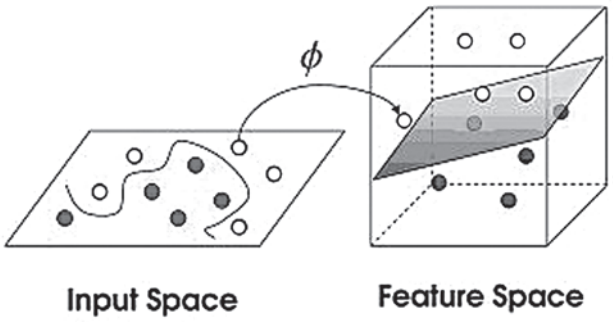

Figure 1. Features mapped and separated in high-dimensional space

The high dimensionality of the feature space enables the construction of a linear separating hyperplane in the space. However, numerical optimisation schemes in high dimension would suffer from the problems associated with dimensionality. Such computational complexities can be avoided by taking the advantage of the inner-product kernel where the dot product in the feature map is implicitly computed by evaluating the simple kernel, thus avoiding the explicit calculation of the feature map. In the present study, the input data is mapped into an infinite-dimensional feature space using a Radial Basis Function (RBF) kernel and can be expressed as:

$$K\left(\vec{x}, \vec{x}_i\right) = \exp\left(-\frac{1}{2\sigma^2}\left\|\vec{x} - \vec{x}_i\right\|^2\right) \quad \text{.....(1)}$$

where the two vectors correspond to the input vector $\vec{x}$ and the i^{th} input pattern $\vec{x}_i$. RBFs[13] are popular for interpolating scattered data as the associated system of linear equations is guaranteed to be invertible under very mild conditions on the locations of the data points. For example, the thin-plate spline only requires that the points are not co-linear while the Gaussian and multiquadric place no restrictions on the locations of the points. In particular, RBFs do not require that the data lie on any sort of regular grid. Once the data is mapped to the N dimensional space, an N-1 dimensional hyperplane is constructed that maximises the separation of the data from the origin taking into account the nonlinear relationship of the data, treating the origin as belonging to the second class. Traditional methods utilise a structure risk minimisation method to minimise the empirical training error[13]. The method employed here maximises the separation of the data from the origin, thus maximising the observable difference between the training set and data that belongs to a different class. The classifier uses the outliers as representatives of data that has not been observed in the training set. This optimisation problem is aimed at finding the optimal set of hyperplane parameters for which the margin of separation between the origin and the support vectors are maximised. This is the same as minimising the Euclidean norm of the weight vector ($\vec{w}$). For non-separable patterns, the primal problem can be formulated as follows:

$$\min_{w,\rho,\xi,b} \frac{1}{lv}\sum_{i=1}^{l}\xi_i - \rho + \frac{1}{2}\langle w, w\rangle \quad \text{.....(2)}$$

subject to $\langle w,\Theta(x)\rangle \geq \rho - \xi_i, \xi_i \geq 0$, for $\nu \in [0,1]$ where Θ is the feature map, ρ is the separation of the hyperplane from the origin, ν is the training error limit and ξ is the non-zero slack variable. Setting up the dual problem (Eq. 6) is achieved by first constructing the Lagrangian function, which is expressed as:

$$J\left(\vec{x},\rho,\alpha\right) = \frac{1}{2}\vec{w}^{T}\vec{w} - \sum_{i=1}^{N}\alpha_i\left[y_i\left(\vec{w}^{T}\vec{x}_i + \rho\right) - 1\right] \quad \text{.....(3)}$$

The two optimality conditions are:

$$\frac{\partial J\left(\vec{w},\rho,\alpha\right)}{\partial w} = 0 \quad \text{.....(4)}$$

$$\frac{\partial J\left(\vec{w},\rho,\alpha\right)}{\partial \rho} = 0 \quad \text{.....(5)}$$

Using the two conditions of optimality and Kuhn-Tucker condition on the Lagrangian function, with some simple manipulations it is possible to construct the dual problem which is expressed as:

$$\min_{\alpha} \frac{1}{2}\sum_{i,j}\alpha_i\alpha_j K(x_i,x_j) \quad \text{.....(6)}$$

subject to $0 \leq \alpha_i \leq \frac{1}{lv}, \sum\alpha_i = 1$, where α_i is the Lagrange multiplier. The offset parameter (ρ) can be recovered for all values of α_i. Once this optimisation problem is solved, all the parameters necessary to construct the optimal hyperplane are known. Mathematically, features with non-zero Lagrangian multipliers ($\alpha_i \geq 0$) are termed as support vectors. Once $\Theta(x)$ and α_i are available, the offset can be calculated from the following relation:

$$\rho = \sum_i \alpha_i K(x_i,x_j) \quad \text{.....(7)}$$

Although all the points in Θ are separable, it may not be computationally effective to compute the canonical hyperplanes for data sets that are not linearly separable; therefore an allowable training error ν is introduced[10]. Even though this results in a hyperplane that is not canonical, it yields acceptable solutions quickly. In this paper, an allowable error of 10% was the maximum allowed which means that a 90% classification rate must be achieved when constructing the hyperplane using the training data. This parameter allows the user to make a trade-off between model complexity and training error. The main advantage of the one-class SVM becomes apparent while training the algorithm because only one class of data belonging to any arbitrary reference can be used[11]. Since there is often no data pertaining to the healthy state of an existing structure, an algorithm must not be greatly affected by a change in the training data. To achieve the optimum classification rate, the algorithm minimises the upper bound of the generalisation error by maximising the separations of the most similar patters or support vectors in hyperspace. As a consequence, the margin between the data and the separating hyperplane is also maximised[9]. During classification, if the test data points are sufficiently different from the training data they are placed near the origin, on the opposite side of the decision plane. This represents a change in the state of the system that is most likely due to structural or sensor damage. To determine on which side of the hyperplane a particular point belongs, a decision function is used to test each new data point. The decision function for a given test vector $\Theta(z)$ is expressed in terms of the RBF kernel by the following function:

$$F(z) = sign\left(\sum_{i=1}^{l}\alpha_i K(x_i,x_j) - \rho\right) \quad \text{.....(8)}$$

where $F(z)$ is the decision function that decides whether a training point should lie on the same side of the hyperplane as the training data or on near the origin if it is from a different class. A data point that is from the same class is assigned a value of +1 and data that is not from the same class is assigned a value of -1.

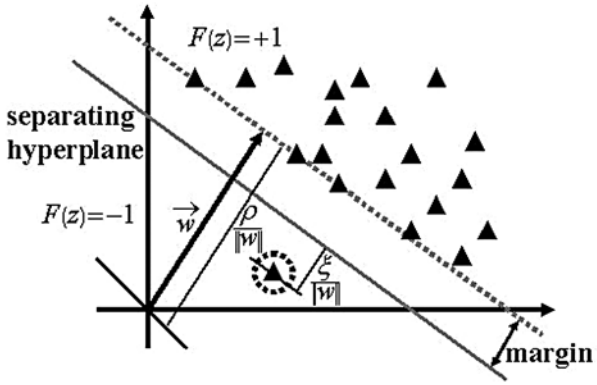

Figure 2. Graphical representation of hyperplane construction

In order to design the SVMs classifier, it is necessary to select the appropriate kernel parameter for each class of data. The parameter controls the smoothness of the kernel function and is tuned based on the model parameter, such that the upper bound on the classification error is satisfied. The parameter σ acts as an in-built tuning parameter that controls the quality of the training. The kernel width has a regularisation effect on the cost function that is minimised during training in order to achieve the upper bound of the allowable error ν[14]. Since choice of this parameter is application specific, it is important to calculate an optimum value for a given training set. A simple brute force method is adopted to do this[12]. First, α_i and ρ are calculated for a range of kernel parameters and the data is classified using these parameters. The optimal σ is selected as the smallest kernel parameter that reduces the misclassified data to v percent of the training data. A graphical representation of this selection is shown in Figure 3.

Time-embedded features

In machine learning based characterisation and classification schemes, the outcome is typically dependent on the input features that have been extracted while preprocessing[18] the raw data. In a wave-based approach, the response of the distributed sensors is basically a time series which can be defined as a sequence of measurements $x(t)$, at different instants of time, of an observable x acquired at regular time intervals. In time series applications, the dynamical information of the system can be extracted for a given dataset of scalar observations where each of this observation corresponds to the projection of the systems' state vector in one dimension. Taken's theorem[19] states that it is possible to reconstruct the attractor in the phase space, given $x(t)$. This can be achieved using the time-embedding approach where one can construct a one-to-one differential mapping between a finite windowed time series. Given a time series $x(t)$ with N number of data points, the state space vectors can be represented as follows:

$$y(t) = \begin{bmatrix} x(t) \\ x(t+\tau) \\ x(t+2\tau) \\ . \\ . \\ . \\ x\left(t+\left(D_E-1\right)\tau\right) \end{bmatrix} \quad \text{.....(9)}$$

where the time instant $t = nT_s$, T_s being the sampling time. Here D_E is the embedding dimension, τ is the time delay and n is the reassigned data points of the state vector where $n = N - (D_E - 1)\,\tau$. The delay reconstruction makes it possible to view the dynamics in-terms of the scalar field and hence the best surface fitting these points would represent the approximate dynamics of the system[20].

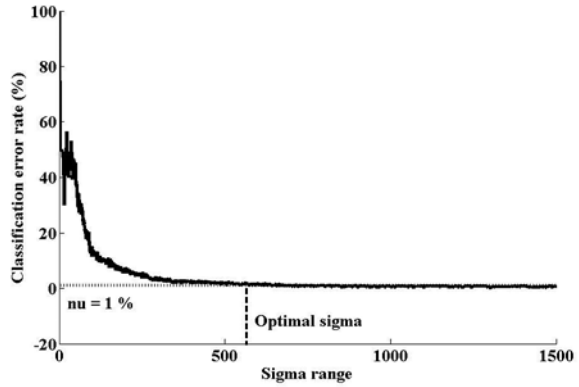

Figure 3. Demonstration of optimal σ selection

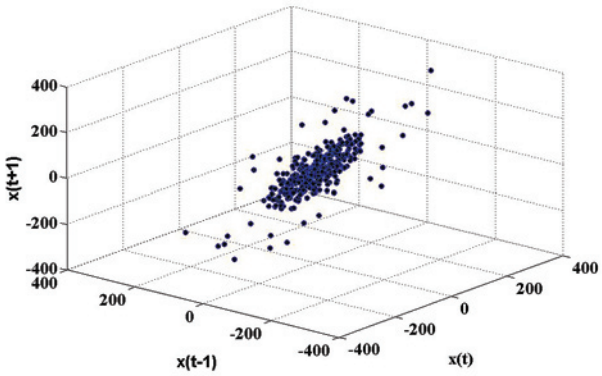

Figure 4. Two-dimensional delay construction (sensor 7 response: under 100% torque 0 Kcycs)

Figure 4 represents a similar interpretation of two-dimensional embedding of the sensor 7 response under 100% torque condition and before even the loading is applied.

The time-embedding approach is a very popular technique in the field of nonlinear dynamics and is commonly used as a tool in order to predict the future dynamics of the system. To ensure proper reconstruction the reconstructing parameters that is the embedding dimension and the time delay has to be assigned properly. In data driven approaches, while introducing delay in experimental data sets, the choice of the time delay is considered to be a very important step. This is because when τ is chosen to be very small delay compared to the internal time scale of the system, the successive components of the delay vectors $x(t)$ and $x(t + \tau)$ are almost linearly dependent on each other, that is they are highly correlated. Very large delay τ, if chosen, can result in 'over-folding' of the attractor[19]. A practical way to choose the proper τ is to assign an interval $0 < \tau < 0.5T_{fm}$, where T_{fm} is the reciprocal of the dominant frequency present in the spectrum. The dimension D_E of the reconstruction provides a proper embedding if

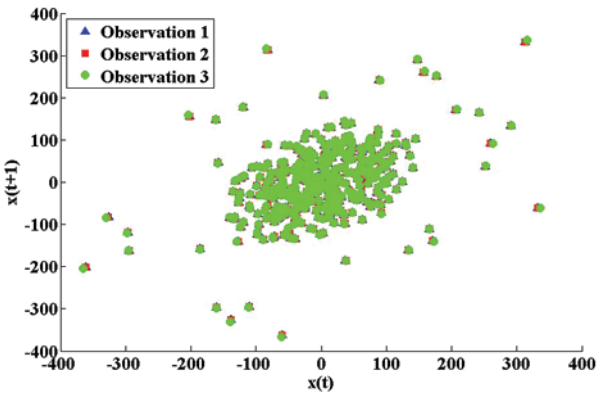

Figure 5. Phase portrait x(t) *vs* x(t+1) for three separate measurements under same torque condition

$D_E \geq 2d_s + 1$, where d_s is the dimension of the flow of the original dimension of the original flow. A significantly high value of D_E would be sufficient to provide a proper embedding except it can lead to high computational expenses depending on the nature of the application. In the present work, D_E has been taken as 11 and τ as the interval between two consecutive sample points. This means that each measurements *x(t)* would lead to an 11 dimensional input vector. Figure 5 represents the closeness of the phase portrait for three different measurements from the sensor 7 response under 100% torque condition with 0 kcycles. Figure 6 shows how the phase portrait differs when the dynamics of the system changes under different torque conditions (100%/60%/30% cases).

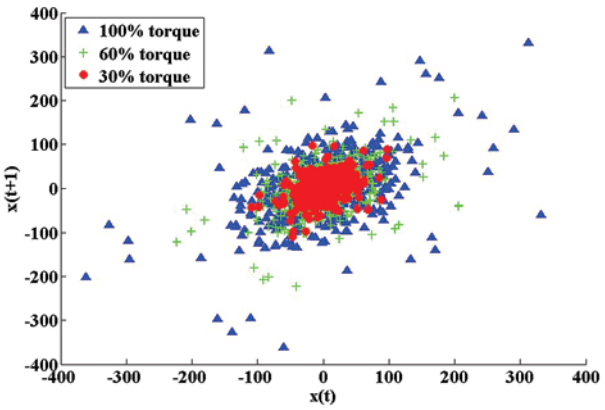

Figure 6. Phase portrait x(t) *vs* x(t+1) for three separate measurements under different torque condition

Experimental procedure for the bolted joint specimen

The test specimens that were used for testing were machined out of Al 6061-T651 plate, 1/8" thick. The dimensions of the laps were chosen so that they were representative of commonly used joints in aerospace structures. Since the location of the notch was known before testing, only the centre lap was mounted with sensors. The single lap bolted joint was instrumented with 0.25 in diameter, 0.01 in thick, APC 850 PZTs in an optimised arrangement. The fatigue loading of the bolted joint was carried out on an Instron 1331, 22 kip capacity servo hydraulic load frame. The joint was subjected to a 2 kip max load (R=0.1) at 20 Hz. To speed up the testing, one bolt was left completely loose while the other bolts were tightened to 100 in-lb of torque. Figure 7 shows the set-up that was used.

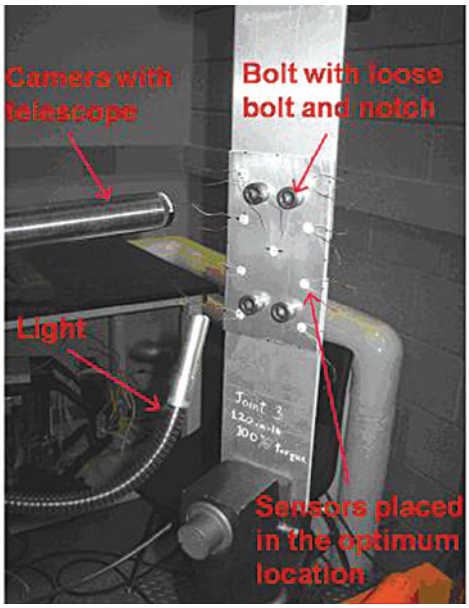

Figure 7. Experimental set-up used to fatigue bolted joints

In order to know the exact initiation site for the crack, the bolthole with the loosened bolt was notched using an electrical discharge machining (EDM) wire. The fatigue test was paused frequently during the test to take PZT measurements and pictures of the damage. Pictures were taken using a camera attached to an optical camera focused on the notch. In order to improve the picture quality and make it easier to distinguish surface scratches from a growing crack, only the viewing surface was polished using 1200 grit paper. The rest of the contact surfaces were untouched so that the fatigue behaviour of the structure was not significantly altered. The bolted joint was left on the frame when data was collected to ensure uniform boundary conditions for all the readings. The level of torque in the loosened bolt was also adjusted to 30, 60 and 100% to show the effect of torque on the observed signal at a given crack length. A 130 kHz, Gaussian windowed sine wave was used as the excitation signal. The acquisition was carried out at a rate of 2 MSa/s with 100 observations taken for each sensor.

Before the test specimen could be instrumented with detection hardware, it is necessary to determine the wave attenuation in the aluminium media. By understanding the extent of the attenuation in the material, distance between sensors placed on the structure can be determined so that there is sufficient overlap of the sensing regions in the structure. An optimal sensor placement technique[7] is used to determine the placement of the sensors on the structure. The optimisation was constrained so that sensors were not placed too close to the edges or the bolts. The resulting sensor placements are shown in Figure 8. Once the placement of the sensors is determined, it is necessary to decide the minimum size of crack that can be detected so that a threshold voltage corresponding to that minimum crack size can be determined and any signal with a lower voltage is ignored as noise.

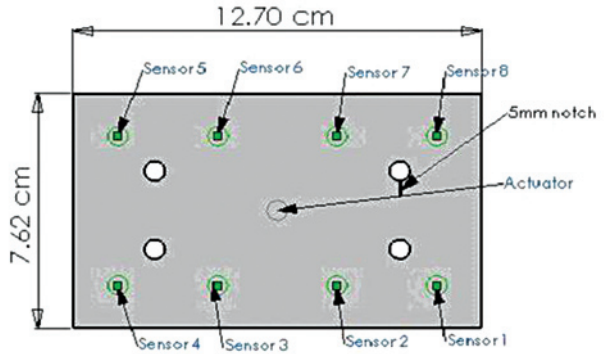

Figure 8. Optimised sensor placement used for data collection

Experimental procedure for the composite specimens

The composite samples were manufactured using a composite prepreg laid up in a $[0,90]_{4s}$ configuration and cured in a hot press using a standard curing heat and pressure cycle. The sample was clamped at one end and mounted with Thunder PZTs as shown in Figure 9, with sensors mounted on either side of the damaged region. Four types of damage, namely delamination, drilled hole, notch, saw cut were tested.

A similar input signal to that used for the bolted joint was applied to the different composite. The excitation signal used was an 8 kHz tone burst signal with a sampling frequency of 100 kHz. To account for the material variability, data was collected from four coupons from each damage category. Ten observations were made for each coupon subjected to identical boundary conditions to account for experimental uncertainties associated with data acquisition.

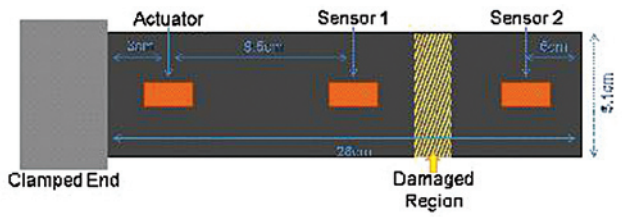

Figure 9. Experimental set-up for composite

Signal conditioning

Since the data for the bolted joint specimen was collected while it was still mounted on the servo hydraulic frame, the acquired signal had some parasitic parameters that had to be removed before classification. These parameters included broadband noise, a low frequency wiggle since the sample was still mounted on the frame, non-uniform signal energy and DC clamping. In order to remove these parameters, the data was first processed using band pass filter, allowing frequencies between 10-300 kHz to remove the low frequency wiggle, noise and DC clamping. The signals were then normalised and down sampled to make the signal processing faster. A matching pursuit decomposition using 60 iterations was used to decompose the signal and then reconstruct it to make the input signal smoother for classification. This step was not necessary for the data collected for the composite sample since it was collected while the sample was in a completely static condition and data was collected using a lower excitation and sampling frequency.

Results and discussion

Table 1 shows the result of the SVMs classification of damage on the metallic bolted joint specimen at a constant torque. All the data in this Table refers to data collected after the sample exhibited signs of fatigue damage and the loosened bolt is at 0% torque. The matrix R_{ij} represents the ability of the classifier to correctly distinguish between the i^{th} and j^{th} class of damage. To obtain this matrix, a matrix Q_{ij} which represents percentage of correct classification when training with each row (i^{th} class) and testing with each column (j^{th} class) is calculated. To understand the selection criteria used to create R from Q, consider the following.

Let A and B represent 2 classes of signals that may or may not be distinct. If we train the algorithm with A and test with B it will yield a classification rate X. Similarly, training with B and testing with A yields a classification rate Y. If the absolute value of the difference between X and Y is less than a defined threshold, then the two signals A and B belong to the same class and this has been used as a selection criteria. In the present analysis, this difference is taken to be less than five percent (Eq. 10) of the required classification rate (1-ν). Geometrically, the selection criteria means that the hyperplane constructed for the first case is very similar to the hyperplane constructed in the second case in that most of the data lies on the same side for both cases. In this study, R_{ij} is the number of observations classified as being in the same class in Q_{ij} divided by the total number of observation combinations.

Table 1 represents the outcome of the kernel-based classifier when trained with each class of data and trained across all

Table 1. Outcome of the classifier using data from sensor 7 for constant torque

R_{ij}		kcycles						
ν = 0.01		130	135	140	145	155	200	285
kcycles	130	1	0.0124	0	0	0	0	0
	135	0.0124	1	0.0136	0	0	0	0
	140	0	0.0136	1	0	0	0	0
	145	0	0	0	0.9848	0.0196	0	0
	155	0	0	0	0.0196	1	0	0
	200	0	0	0	0	0	0.8392	0
	285	0	0	0	0	0	0	0.8816

possible datasets from different classes. For example, if trained with 130 kcycles, the algorithm is able to correctly classify 100% of the test sets of 130 kcycles but shows 1.2% misclassification on the 135 kcycles data. One important observation is that the misclassification, though small, mainly occurs when the training and testing sets belong to the classes when the effects due to the change in crack sizes are very small (for example, between 145/155 kcycles test cases). The classifier can accurately categorise the signatures from different crack sizes as long as the change in crack length is sufficiently large. When the difference between the crack size for the training and testing class is large enough, the classifier is able to classify 100% of the signals as belonging to a different class.

Table 2 shows a similar classification outcome matrix, but this time all the data is selected from the healthy state (0 kcycles) and the amount of correct classification as a function of torque level is shown. Every damage state consists of 100 observations, of length 400. The algorithm used 50 waveforms for training and 50 waveforms for testing. The Table shows that the algorithm is able to classify all torques with accuracy greater than 94%. From these results, it can be seen that the algorithm can very accurately classify fatigue damage and torque loss states for a constant torque level and constant crack size respectively.

Table 2. Outcome of the classifier using data from sensor 7 for fatigue damage state

R_{ij}		% Torque		
ν = 0.01		30	60	100
% Torque	30	0.94	0	0
	60	0	1	0
	100	0	0	1

For the data analysed for the composite sample, Table 3 shows the different damage cases used for training and testing. Ten observations from each class were first used as training data and used to classify the remaining ten signals from each class. Then the training set is used as the testing data and the training data is used as the testing data. Tables 4 and 5 show the results of the classification for each combination of training and testing data. From these Tables it can be seen that there are some cases of misclassification, when the training and testing cases are reversed for certain cases. The selection criterion applied can be expressed as follows:

$$\begin{aligned} &If,\ \left\| R_{ij} - R_{ji} \right\| \le 0.05\left(1-\nu\right) \\ &Else\ Q_{ij}^{k} = Q_{ji}^{k} = 1 \\ &\qquad Q_{ij}^{k} = Q_{ji}^{k} = 0 \end{aligned} \qquad \text{.....(10)}$$

In order to improve the classification result, the R_{ij} matrices of sensor 1 and sensor 2 were combined to construct the matrix M as shown below:

$$M = Q^{sensor\ 1} \cap Q^{sensor\ 2} \text{ .....(11)}$$

The matrix M represents the final outcome of the classifier based on the mutual information of both sensors. The difference between the classification rate X and Y had to be greater than five percent

Table 3. Training and testing sets for different damage types in composite material

Damage class	Training class	No. of observations	Testing class	No. of observations
Healthy	TRC1	10	TEC1	10
Delaminated	TRC2	10	TEC2	10
Drilled hole	TRC3	10	TEC3	10
Notch	TRC4	10	TEC4	10
Saw cut	TRC5	10	TEC5	10

Table 4. Output of the classifier for case 1, sensor 1

R_{ij}		kcycles (testing data)									
ν = 0.05		TRC1	TRC2	TRC3	TRC4	TRC5	TEC1	TEC2	TEC3	TEC4	TEC5
kcycles (training data)	TRC1	0.95	0.84	0.81	0.8	0.75	0.93	0.83	0.8	0.79	0.75
	TRC2	0.97	0.95	0.86	0.94	0.83	0.97	0.93	0.86	0.94	0.85
	TRC3	0.98	0.93	0.95	0.97	0.91	0.98	0.95	0.94	0.96	0.92
	TRC4	0.97	0.89	0.86	0.95	0.83	0.97	0.89	0.87	0.93	0.84
	TRC5	1	0.95	0.94	0.99	0.95	1	0.95	0.94	0.99	0.92
	TEC1	0.95	0.85	0.83	0.87	0.79	0.95	0.84	0.83	0.84	0.8
	TEC2	0.97	0.92	0.88	0.95	0.88	0.96	0.95	0.88	0.94	0.88
	TEC3	1	0.95	0.93	0.99	0.92	1	0.95	0.95	0.99	0.93
	TEC4	0.97	0.88	0.87	0.96	0.82	0.96	0.88	0.88	0.95	0.83
	TEC5	1	0.96	0.94	0.99	0.93	1	0.96	0.95	0.99	0.95

Table 5. Output of the classifier for case 1, sensor 2

R_{ij}		kcycles (testing data)									
ν = 0.05		TRC1	TRC2	TRC3	TRC4	TRC5	TEC1	TEC2	TEC3	TEC4	TEC5
kcycles (training data)	TRC1	0.95	0.95	0.89	0.92	0.91	0.94	0.96	0.89	0.91	0.91
	TRC2	0.9	0.95	0.84	0.91	0.9	0.9	0.94	0.85	0.89	0.9
	TRC3	0.97	0.95	0.95	0.76	0.91	0.97	0.96	0.95	0.72	0.9
	TRC4	0.96	0.99	0.9	0.95	0.92	0.96	0.99	0.91	0.93	0.91
	TRC5	0.94	0.96	0.96	0.92	0.95	0.95	0.96	0.95	0.9	0.95
	TEC1	0.94	0.96	0.89	0.93	0.91	0.95	0.96	0.89	0.92	0.9
	TEC2	0.81	0.92	0.84	0.81	0.87	0.83	0.95	0.83	0.8	0.88
	TEC3	0.95	0.94	0.92	0.77	0.9	0.94	0.94	0.95	0.72	0.9
	TEC4	0.96	0.99	0.91	0.95	0.92	0.96	1	0.91	0.96	0.92
	TEC5	0.95	0.96	0.95	0.92	0.94	0.95	0.96	0.96	0.91	0.95

of the maximum classification rate for the 2 classes to belong to different data sets. Table 6 shows the final outcome of the classifier and it can be seen that the classifier was able to accurately classify all the damage states that were tested with no misclassification. Tables 7 and 8 show the classification results for a second case where the effect of material variability was taken into consideration by testing two coupons with the same damage. These Tables show very slight deviations from the values shown in Tables 4 and 5 and there were several non-unique cases. As expected, since the individual sensor results did not vary much with a change in material, a similar result was obtained when the information from both sensors was combined (Table 9).

Table 6. Outcome of the classifier for composite data using two sensors

M_{ij}		kcycles (testing data)									
ν = 0.05		TRC1	TRC2	TRC3	TRC4	TRC5	TEC1	TEC2	TEC3	TEC4	TEC5
kcycles (training data)	TRC1	1	0	0	0	0	1	0	0	0	0
	TRC2	0	1	0	0	0	0	1	0	0	0
	TRC3	0	0	1	0	0	0	0	1	0	0
	TRC4	0	0	0	1	0	0	0	0	1	0
	TRC5	0	0	0	0	1	0	0	0	0	1
	TEC1	1	0	0	0	0	1	0	0	0	0
	TEC2	0	1	0	0	0	0	1	0	0	0
	TEC3	0	0	1	0	0	0	0	1	0	0
	TEC4	0	0	0	1	0	0	0	0	1	0
	TEC5	0	0	0	0	1	0	0	0	0	1

Conclusion

A kernel-based classification technique has been used along with active wave-based approach to detect and classify damage in both metallic and composite structures. The procedure has been used to classify fatigue damage caused by the loss of torque in a bolt and associated crack growth. Composite laminates with different

Table 7. Output of the classifier for case 2, sensor 1

R_{ij}		kcycles (testing data)									
ν = 0.05		TRC1	TRC2	TRC3	TRC4	TRC5	TEC1	TEC2	TEC3	TEC4	TEC5
kcycles (training data)	TRC1	0.95	0.69	0.87	0.62	0.81	0.76	0.63	0.81	0.49	0.72
	TRC2	0.96	0.95	0.85	0.79	0.82	0.92	0.77	0.87	0.79	0.82
	TRC3	0.92	0.72	0.95	0.65	0.83	0.88	0.7	0.86	0.64	0.82
	TRC4	0.85	0.81	0.86	0.95	0.82	0.9	0.7	0.85	0.67	0.81
	TRC5	0.88	0.7	0.91	0.66	0.96	0.88	0.7	0.91	0.6	0.9
	TEC1	0.8	0.65	0.85	0.54	0.76	0.95	0.69	0.87	0.65	0.83
	TEC2	0.91	0.76	0.87	0.79	0.82	0.96	0.95	0.85	0.79	0.82
	TEC3	0.9	0.72	0.88	0.68	0.84	0.94	0.73	0.95	0.67	0.83
	TEC4	0.88	0.7	0.84	0.66	0.8	0.85	0.81	0.85	0.95	0.81
	TEC5	0.83	0.69	0.9	0.58	0.9	0.87	0.68	0.89	0.62	0.95

Table 8. Output of the classifier for case 2, sensor 2

R_{ij}		kcycles (testing data)									
ν = 0.05		TRC1	TRC2	TRC3	TRC4	TRC5	TEC1	TEC2	TEC3	TEC4	TEC5
kcycles (training data)	TRC1	0.95	0.93	0.89	0.79	0.92	0.92	0.91	0.89	0.8	0.92
	TRC2	0.67	0.95	0.62	0.72	0.67	0.64	0.86	0.69	0.66	0.61
	TRC3	0.84	0.86	0.95	0.72	0.86	0.83	0.89	0.84	0.67	0.85
	TRC4	0.88	0.91	0.81	0.95	0.88	0.8	0.91	0.79	0.74	0.86
	TRC5	0.85	0.89	0.87	0.74	0.95	0.89	0.89	0.87	0.72	0.89
	TEC1	0.92	0.91	0.9	0.81	0.92	0.95	0.92	0.89	0.8	0.92
	TEC2	0.64	0.86	0.69	0.66	0.61	0.66	0.96	0.61	0.71	0.66
	TEC3	0.82	0.87	0.83	0.66	0.85	0.85	0.87	0.95	0.71	0.85
	TEC4	0.81	0.91	0.8	0.76	0.87	0.9	0.92	0.82	0.96	0.88
	TEC5	0.84	0.89	0.87	0.73	0.89	0.87	0.89	0.86	0.73	0.95

Table 9. Outcome of the classifier for composite data including material variability

M_{ij}		kcycles (testing data)									
ν = 0.05		TRC1	TRC2	TRC3	TRC4	TRC5	TEC1	TEC2	TEC3	TEC4	TEC5
kcycles (training data)	TRC1	1	0	0	0	0	1	0	0	0	0
	TRC2	0	1	0	0	0	0	1	0	0	0
	TRC3	0	0	1	0	0	0	0	1	0	0
	TRC4	0	0	0	1	0	0	0	0	1	0
	TRC5	0	0	0	0	1	0	0	0	0	1
	TEC1	1	0	0	0	0	1	0	0	0	0
	TEC2	0	1	0	0	0	0	1	0	0	0
	TEC3	0	0	1	0	0	0	0	1	0	0
	TEC4	0	0	0	1	0	0	0	0	1	0
	TEC5	0	0	0	0	1	0	0	0	0	1

damage types have been used and the effects of material and measurement uncertainty were taken into consideration. In both cases, a time-embedding preprocessing tool has been used to view the dynamics of the system and the state of the system has been classified using one-class Support Vector Machine algorithm. The developed technique uses the Gaussian kernel to map the input patterns to high-dimensional space, while taking advantage of the kernel where the dot product is computed implicitly by evaluating the simple kernel. Moreover, the application of one-class algorithm does not require the labelling of all states different from the nominal training state. The results for the bolted joint samples showed good classification results for damage due to a change in torque at constant crack length and for a change in crack length for a constant torque. In the composite samples, the technique was able to accurately distinguish between the five cases of states considered: healthy, delaminated, drill hole, notch, and saw cut.

The results showed that the individual sensors were able to classify a large number of cases distinctly but a few non-distinct cases were still present. To improve the detection method, information from both sensors was combined yielding very accurate results. The effect of material variability was also included and the classifier was still able to correctly identify 100% of the classes using mutual sharing of information from multiple sensors. This was expected since the effect of material variability did not significantly alter the classification rate of the individual sensors. In both cases, combining information from both sensors produced 100% classification results.

Acknowledgements

The research was supported in parts by: Air Force Office of Scientific Research, grant FA95550-06-1-0309, technical monitor, Dr Victor Giurgiutiu, and NASA Ames Research Center, grant NNA07CN64A, technical monitor Dr Rodney Martin.

References

1. A Chattopadhyay, A Papandreou-Suppappola, D Cochran, J B Spicer, P D Peralta, R O Claus and R G Ghanem, 'A multidisciplinary approach to structural health monitoring of metallic aerospace systems', Proc., Integrated Vehicles Systems Health Management Conference, March 2006.
2. P Lazzarin, V Milani and M Quaresimin, 'Scatter bands summarising the fatigue strength of aluminum alloy bolted joints', International Journal of Fatigue, Vol 19, No 5, pp 401-407, 1997.
3. J María Mínguez and J Vogwell, 'Effect of torque tightening on the fatigue strength of bolted joints', Engineering Failure Analysis, Vol 13, pp 1410-1421, December 2006.
4. A Aragón, J M Alegre and F Gutiérrez-Solana, 'Effect of clamping force on the fatigue behavior of punched plates subjected to axial loading', Engineering Failure Analysis, Vol 13, pp 271-281, March 2006.
5. M P Szolwinski, G Harish and T N Farris, 'Experimental observation of the effect of contact parameters on fretting fatigue crack nucleation', in Proceedings of USAF Structural Integrity Program Conference, pp 237-265, 1995.
6. M P Szolwinski, G Harish, P A Mc Veigh and T N Farris, 'The role of fretting crack nucleation in the onset of widespread fatigue damage: analysis and experiments', FAA-NASA Symposium on the Continued Airworthiness of Aircraft Structures, pp 585-596, 1996.
7. A Chattopadhyay, S Das and D Miller, 'Acoustic based structural health monitoring for composites using optimal sensor placement: analysis and experiments', ICCES, 2004.
8. P Tongpadungrod, T D L Rhys and P N Brett, 'An approach to optimize the critical sensor locations in one-dimensional novel distributive tactile surface to maximize performance', Sensors and Actuators A: Physical, Vol 105, pp 47-54, 2003.
9. S L Padula and R K Kincaid, 'Optimization Strategies for Sensor and Actuator Placement', NASA/TM-1999-209126, 1999.
10. S Kwang-Kyu, 'An application of one class support vector machines in content based image retrieval', Expert Systems with Applications: An International Journal, Vol 33, pp 491-498, August 2007.
11. S Das, A N Srivastava and A Chattopadhyay, 'Classification of damage signatures in composite plates using one-class SVMs', Proceedings of IEEE Aerospace Conference, 2007.
12. R Unnthorsson, T P Runarsson and M T Jonsson, 'Model selection in one class nu-SVMs using RBF kernels', 16th Conference on Condition Monitoring and Diagnostic Engineering Management, pp 27-29, August 2003.
13. A J Smola, B Schölkopf and K R Müller, 'The connection between regularization operators and support vector kernels', Neural Networks Vol 11, pp 637-649, 1998.

14. N E Ayat, M Cheriet and C Y Suen, 'Automatic model selection for the optimization of SVM kernels', Pattern Recognition, Vol 38, pp 1733-1745, 2005.
15. S Taeshik and M Jongsub, 'A hybrid machine learning approach to network anomaly detection', Information Sciences, March 2007.
16. H W Park, H Sohn, K H Law and C R Farrar, 'Time reversal active sensing for health monitoring of a composite plate', Journal of Sound and Vibration, Vol 302, pp 50-66, 2007.
17. Z Liang and A K Nandi, 'Fault classification using genetic programming', Mechanical Systems and Signal Processing', Vol 21, pp 1273-1284, April 2007.
18. N Tanaka, H Okamoto and M Naito, 'Estimating the active dimension of the dynamics in a time series based on an information criterion', Physica D: Nonlinear Phenomena, Vol 158, pp 19-31, October 2001.
19. T Buzug and G Pfister, 'Optimal delay time and embedding dimension for delay-time coordinates by analysis of the global static and local dynamical behavior of strange attractors', Phys. Rev. A, Vol 45, pp 7073-7084, May 1992.
20. P E McSharry, 'Innovations in consistent nonlinear deterministic prediction', D.Phil. Thesis, University of Oxford, 1999.