Bayesian network approach to classify damages and f-folds feature subset selection method in laminated composite materials

A. O. Addin[b], S. M. Sapuan[a], E. Mahdi[c], M. Osman[d], N. Ismail[a]

[a]*Department of Mechanical and Manufacturing Engineering*

[b]*Department of Intelligent Systems and Robotics, Institute of Advanced Technology*

[c]*Department of Aerospace Engineering*

[d]*Department of Communication Technology and Networks*

Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

Abstract

Structural health monitoring is defined as a computational modeling system used to quantify the health of a structure and its life time. A variety of techniques have been introduced to quantify damage detection and identification in structures. Nevertheless, the Bayesian networks have rarely been used for damage detection in laminated composite materials. This paper is intended to introduce the Bayesian network in general and the Naive-bayes in specific as one of the most successful classification systems to simulate damage detection in laminated composite materials. A method for feature subset selection based on intervals between the amplitudes of waves used for damage detection is also introduced. The method utilizes clustering in the process of feature subset selection. The Bayesian classification and the feature selection method are analyzed based on theoretical point of view and only preliminary tests were conducted based on artificial damages created in quasi-isotopic laminates of the AS4/3501-6 graphite/epoxy system.

1 Introduction

Recently, there has been a tremendous growth in the usage of laminated composite materials (*LCMs*) in all types of engineering structures (e.g. aerospace, automotive, and sports). *LCMs* are fabricated by stacking plates or plies of composite materials together to acquire unique properties (e.g high strength and stiffness, and light weight) that can not be guaranteed by individual constituents of the laminate. One of the drawbacks of such materials is their vulnerability to different kind of damages. Some of the damages are manufacturing related, like foreign object inclusion, porosity, resin rich area, etc. Others are service related, which may result from a bird strike to an aircraft and rain hail during flight, depress hitting in runway during take off or landing, and tools dropped during maintenance procedures [1].

The damages have the potential of growing and leading to catastrophic loss of human life, and decrease in economy. Examples of real-life damages can be shown as airline crashes, space shuttle explosions, and building and bridge collapses. It is very important to prevent the materials from catastrophic failures and to prolong their service life by early detection of the damages. There are a variety of nondestructive evaluation (*NDE*) techniques developed to detect the damages (e.g. ultrasonic, eddy-current, *C*-scan, and etc.). Sometimes, the *NDE* is called nondestructive testing (*NDT*) [2].

One of the smart potential solutions used for damage detection is the structural health monitoring system (*SHM*). The literature defines the *SHM* as the acquisition, validation, and analysis of technical data to facilitate the life-cycle management decisions [3]. *Kessler et al.* [4] say, *SHM* denotes a reliable system with the ability to detect and interpret adverse changes in a structure due to damage or normal operation.

One of the most important and exciting areas of the *SHM* systems is the development of quantitative modeling techniques to predict the presence of damages in *LCMs*. Researchers in the *SHM* field have borrowed and implemented a variety of artificial intelligent and machine learning techniques to quantitatively identify and detect damages in structures. Neural network (*NN*) is one of these techniques which has been widely adopted by many researchers in this area [5–8]. Chakraborty [9] introduces an approach that predicts the presence of embedded delamination (in terms of location, shape, and size) in fiber reinforced plastic composite laminates by using back propagation (*BP*) *NN* with *3* layers (input, hidden, and output). The network has been tested to predict the presence of delamination along with its size, shape, and location. *Su* and *Ye* [10] have demonstrated a lamb wave (*LW*) propagation-based quantitative identification scheme for delamination in carbon-fiber reinforced polymer (*CFRP*) composite structures by using a multi-layer *BP NN*. Another methods like rule-based, fuzzy logic, and genetic algorithms have been used for damage detection and identification.

Recently Bayesian networks (*BNs*) have emerged as a machine learning technique and a generalizing graph-based framework for creating statistical models of domains with inherent uncertainty. *BNs* have attracted a great deal of attention in research institutions as well as in industry as modeling tools for medical systems, risk prediction, forecasting, robotics, computer games, and etc. [11, 12].

The objective of this paper is to introduce *BN* in general and Naive-bayes in specifics as classifiers to simulate damage detection in *LCMs* for the *NDT* and *SHM* communities. The paper also aims to present a novel method for feature subset selection of wave amplitudes for damage detection.

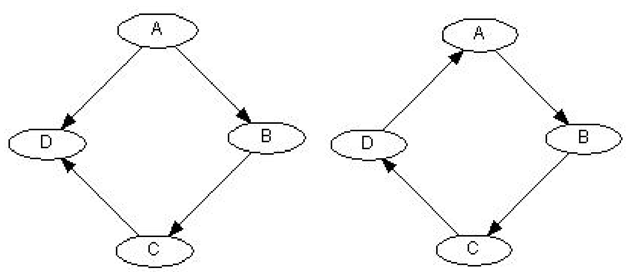
This paper is organized as follows, section two gives a preliminary overview to Bayesian networks based on laminated composite materials and how the Bayeisan network can be used as a classification technique. Section three, shows the Naive bayes in *SHM* systems and how it can be utilized for damage detection employing waves of some *NDT* techniques. Section four introduces the novel f-fold feature subset selections. Section five shows the experiments used in analyzing the f-folds feature subset selection method. Section six concludes the paper and shows some future works.

2 Bayesian Networks

BNs are defined as graphical models that allow us to encode relationships between variables of interest and reason about uncertain domains. They consist of a qualitative part, where features from graph theory are used, and a quantitative part consisting of potentials, which are real-valued functions over a set of variables from the graph. They consist of the following:

- A network structure $G = \{V, E\}$, where $V = \{V_1, V_2, \ldots, V_n\}$ represents a set of variables and E represents a set of directed arcs between the variables (see Figure 6(a)).
- Each variable has a finite set of mutually exclusive states.
- A set of conditional probability tables (*CPTs*) associated with each variable.

The directions of the arcs in *BNs* often represent causal dependency between variables. *BNs* model the quantitative strength of the connections between them, allowing their probabilistic beliefs to be updated automatically as new information arrive. The arcs in any *BNs* are not permitted to be directed cycles, we cannot start from a variable and simply come back to it by following the direction of the arcs in the network (see Figure 1). For this reason the networks are known as directed acyclic graphs (*DAGs*) [11, 12].



(a) Acyclic graph. (b) Cyclic graph.

Figure 1: Figure *a* represents a *BN*, where as Figure *b* does not represent a *BN*. In the later, if we start from *A* following the arcs, we go to *B*, *C*, *D*, then back to *A* forming a cycle (this is not allowed in *BN*).

In *BNs*, a variable is a parent of a child, if there is an arc from the former to the later. In Figure 6(a), variable *A* is a parent to the variable *B*, variable *B* is a child to the variable *A* and a parent to the variable *C*, and variable *D* is a child to the variables *A* and *C*.

BNs can be build by an expert on the domain of study, a structure learning algorithm that automatically extract the structure from a data set, or a combination of both.

The values of each variable should be mutually exclusive and exhaustive, that means the variable must take on exactly one of these values at a time. For example, if we considered building a model to predict the presence of a damage in a composite material, many factors might be taken into account, e.g. the age of the material (*Age*), whether a tool dropped on the material or not (*ToolDrop*), and etc. These factors can be represented as variables in the model connected by directed links according to the direction of impacts (see Figure 4). In Figure 4, the variables *ToolDrop* and *Age* have an impact on the variable *Damage*. That means the presence of the damage can be determined by the states of *ToolDrop* and *Age*. No one can argue that the damage on the material has caused the dropping of the tool on the material or has an impact on the age of the material. Every variable can take one of a different type of discrete values (the states of the variable). The variables *Damage* and *ToolDrop* might be represented by states, which take boolean values *yes* and *no*. And the variable *Age* might be represented by states that take ordered values, *new*, *medium*, and *old*.

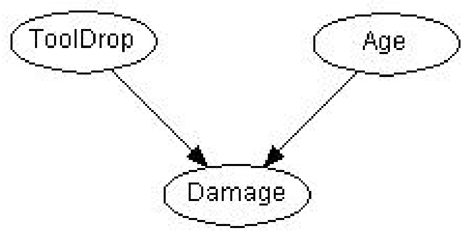


Figure 2: The figure shows a small *BN* structure for damage detection in *LCM*.

If we assume *A* is a variable with *n* states a_1, a_2, $\ldots$, a_n, then *P*(*A*) denotes a probability distribution over these states:

$$P(A) = (x_1, x_2, \ldots, x_n); \qquad x_i \leq 0; \qquad \sum_{i=1} x_i = 1 \tag{1}$$

where x_i is the probability of *A* being in state a_i. This can be written as *P* (*A* = a_i) = x_i or *P* (a_i) = x_i, e.g. P(*Age* = *new*) = 0.8.

The basic concept in the *BN* treatment of certainties in causal networks is conditional probabilities. If the variable *B* has *m* states b_1, b_2, $\ldots$, b_n, the conditional probability statement can be shown as follows"

"The probability of the event *a* given the event *b* is *x*."

which can be written as *P*(*a* | *b*) = *x*. The probability *P*(*A* | *B*) implies an *n* $\times$ *m* table including the probabilities *P*(a_i | b_j) (Table 1).

Table 1: An example of *P*(a_i | b_j), where $1 \leq i \leq 2$ and $1 \leq j \leq 3$. The columns sum to 1.

	b_1	b_2	b_3
a_1	0.4	0.3	0.6
a_2	0.6	0.7	0.4

The fundamental rule for probability calculus is:

$$P(a|b)P(b) = P(a, b), \tag{2}$$

where *P*(*a*, *b*) is the probability of the joint event *a* and *b*. From this we can say *P*(*a* | *b*) *P*(*b*) = *P*(*a* | *b*) *P*(*b*), and this yields the well known *Bayes*' rule:

$$P(b|a) = \frac{P(a|b)P(b)}{P(a)} \tag{3}$$

In Figure 4, the variable *Damage* has two parents and the variables *ToolDrop* and *Age* have no any parents. The joint probability distributions for the variables are shown as *P*(*Damage* | *Age*, *ToolDrop*) *P*(*ToolDrop*), and *P*(*Age*). These probabilities are determined by an expert or automatically extracted from a data set. Since the variables *ToolDorp* and *Age* have no parents, their prior probabilities can be specified as follows:

- *P*(*ToolDrop* = *yes*) = *0.8* and *P*(*ToolDrop* = *no*) = *0.2*
- *P*(*Age* = *new*) = *0.2*, *P*(*Age* = *medium*) = *0.7*, and *P*(*Age* = *old*) = *0.1*.

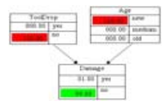
The variable *Damage* has *3* states and two parents each with *2* states. The conditional probability distribution of this variable can be shown as on Table 2. The table has *12* probability values (*3* × *2* × *2*).

ToolDrop	*yes*			*no*		
Age	*new*	*medium*	*old*	*new*	*medium*	*old*
yes	*0.2*	*0.4*	*0.9*	*0.01*	*0.5*	*0.4*
no	*0.8*	*0.6*	*0.1*	*0.99*	*0.5*	*0.6*

Table 2: CPT for *P*(*Damage* | *Age*, *ToolDrop*). The *yes* and *no* in the first column are the states of *Damage*. The first value (*0.2*) in raw *3* and column *2*.

BNs give full representation of probability distributions over their variables. They can be conditioned on any subset of their variables, supporting any direction of reasoning. That means any variables may be query variables and any may be evidence variables. Whenever new information have arrived new beliefs can be calculated. We have shown that *P*(*ToolDrop* = *yes*) = *0.8* and *P*(*Age* = *old*) = *0.1*. Suppose that we have discovered that a tool is dropped on the material and the material is very old, then *P*(*ToolDrop* = *yes*) = *1.0* and *P*(*Age* = *old*) = *1.0*. They are shown in Figure 3(a) as percentages (*100.00* and *00.00*) with red colors. This kind of probabilities is sometimes referred as evidence or instantiation. In *BNs*, when new evidence arrive to some variables, the beliefs on other variables may be changed. This can be shown by carefully studying Figure 3. This process of conditioning on some variables, when observing the value of other variables is known as probability propagation, inference, or belief updating.

(a) The evidence that the tool is dropped and the material is old (*Age* is *old*), increased our belief on the damage to 90%.



(b) The evidence that the tool is not dropped and the material is new (*Age* is *new*), decreased our belief on the damage to 1%.

Figure 3: Changing of believes on *BNs*, when new evidences arrive.

In *SHM* systems many classifiers have been used to detect damages in laminated composite materials, e.g. neural network. *BNs* are powerful tools for knowledge representation and inference under uncertainties. Nevertheless, they are not considered very well as classifiers in *SHM* systems. Naive-bayes is one of the *BNs* classifiers that surprisingly can outperform many sophisticated classifiers on data sets where the features are not strongly correlated.

3 Naive-bayes Classifier in *SHM*

Recently, the state of the art in supervised learning has shown that simple Naive-bayes is surprisingly a competitive classifier and can outperform many *BNs* classifiers (e.g. *C4.5*), when working on data sets where the features are not strongly correlated. It has a strong assumption that all variables in the network are independent of the classification variable (Figure 4). It is very easy to build a Naive-bayes network structure, it does not require a structure learning algorithm.

The amplitudes shown in Figure 5 represent voltage amplitudes of Lamb-waves produced and collected by *PZT* sensors and actuators mounted on the surface of quasi-isotropic graphit/epoxy laminates. The first specimen is a control unit (laminate without damage), the rest of the specimen contain artificial damages. These damages are delamination, crack, and hole. The figure shows that sound waves behave differently when passing through the laminate without and with damage,

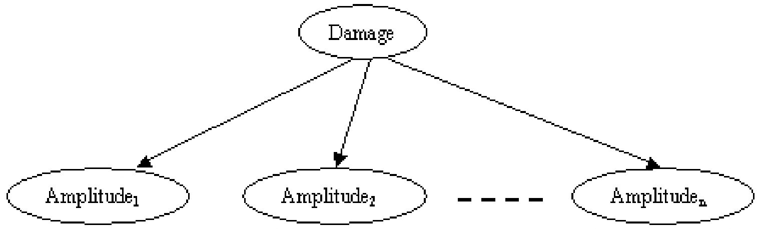
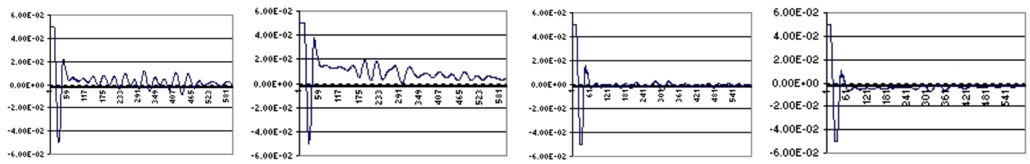


Figure 4: The graph shows a Naive-bayes for damage detection by using amplitudes of waves.

and every damage produces differing amplitudes. Amplitudes like these ones with many cases and different kind of damages can be used to learn the conditional probability tables of variables (*P*($Amplitude_i$ || *Damage*) in the network. Ultimately, the model can be used to predict the damages in laminated composite materials with the highest posterior probability. The probabilities of the damages are determined by entering the new evidence obtained from the amplitudes of the new case to the network.



(a) Without damage (cronrol). (b) Delmination. (c) Crack. (d) Hole.

Figure 5: Time (microseconds) trace of voltage amplitudes collected from quasi-isotropic graphit/epoxy laminates, one without damage and the rest with damages.

The amplitudes shown in Figure 5 were generated by using a constant interval of time (microseconds). For every laminate a set of *600* amplitudes collected. If all of these amplitudes were used as variables on the damage detection model, the model would be overwhelmed, complicated, and its accuracy might slightly be decreased. Different techniques have been adopted for feature subset selections to decrease the size of the data and increase the accuracy. Some of these techniques extract the peaks of the amplitudes as feature subsets, but it is very difficult to be sure whether these peaks can be representative to the whole wave. The rest of the techniques have different kind of limitations and disadvantages. So as to overcome some of these limitations and tackle some of these disadvantages, the f-folds feature subset selection algorithm has been developed.

4 f-folds Feature Subset Selection Algorithm

The amplitudes in Figure 5 have been collected by using a constant interval of time (microseconds). A different data set might be acquired, if the interval value had been changed. If it had been assumed that the interval was increased *10* times more than the original one, then the original amplitudes would be divided into *60* folds (*10* amplitudes in each fold). In this case *10* different data sets would be formed each with *60* amplitudes. The amplitudes included in each set depend on the first amplitude selected from the first fold, if the first amplitude was the first to be included, then the first amplitudes in other folds would be included to the data set, if the second one was the first one to be included, then the seconds in all other folds would be included in the data set, and etc.

Different wave forms can be extracted from these data sets, which some of them produce waves with similar shapes (Figure 6). The data sets with similar wave shapes can be grouped or clustered together and the mean values of every cluster can be considered as representatives of the cluster (the maximum and minimum values can also be considered). The feature extracted from these data depends on the number of folds (f), the number of amplitudes (n), and wether the maximum and minimum values included or not. This notion has been utilized and encoded as f-folds feature subset selection algorithm (Algorithm 1) to extract features from waves for damage detection in laminated composite materials. In this paper, it has been assumed that all of these data sets represent one data set where each data set is included as a separate record.

Algorithm 1 (k-folds feature subset selection algorithm)
Input:
Amps = amp_1, amp_2, . . ., amp_n *(Amplitudes to be clustered).*
k *(number of clusters).*
f *(number of folds).*
Outputs:
Means = $\{$*m(c1), m(c2), . . ., m(ck)*$\}$
Maxs = $\{$*max(c1), max(c2), . . ., max(ck)*$\}$
Mins = $\{$*min(c1), min(c2), . . ., min(ck)*$\}$
procedure Clustering
1. Divide Amps into f folds (fold(1), fold(2), . . ., fold(f)),
where $|fold(1)|$ = $|fold(2)|$ = . . . = $|fold(f)|$
2. Create a new data set $NewAmp$ = $nAmp(1)$, $nAmp(2)$, . . ., $nAmp(m)$,
where $\forall A = fold(k)_i, 1 \leq i \leq m, and 1 \leq k \leq f, A \in nAmp(i)$
(the number of elements in each fold is m = n / f).
3. Implement a clustering algorithm (e.g. k-means) on NewAmp, to return k clusters.
4. Return the mean, maximum, and minumum values of the clusters.

The input to the f-folds feature subset selection algorithm is a set of n amplitudes (*Amps* = amp_1, amp_2, . . ., amp_n). In step *1* the algorithm divides the data set into f folds (Algorithm 1). All folds contain the same number of m amplitudes, where m = n / f. In step *2* the algorithm forms a new set of data containing *m* records by assigning the amplitudes with the same index in all folds to the data set as one record (e.g. the first amplitudes in all folds form the first record and so on). This creates the data set $NewAmp$ ($nAmp(1)$, $nAmp(2)$, $\ldots$, $nAmp(m)$). The number of variables in each record is *f* (the number of folds). In step *3* the algorithm implements a clustering algorithm (e.g. k-means algorithm or *EM* algorithm) on $NewAmp$ to divide their instances into *k* clusters. Since each record has *f* variables, the algorithm returns *f* mean values, *f* maximum values, and *f* minimum values of each cluster. These values will be considered as representatives to the clusters and when combined together they can replace the original data set. For example, if we have *100* instances in the cluster we use only *3* instances (means, maximums, and minimums). The total number of the variables (*t*) in each damage type will be reduced to $3 \times f \times k$, when the means, maximums, and minimums of the clusters are considered. And it will be reduced to $f \times k$, if only the means are considered. The values of *f* and *k* must be determined by the user such that $t \ll n$, which believed to decrease the variables to a minimum value that highly increase the accuracy of the model and simplify it.

5 Experiments and Results

The experiments were *25 cm* $\times$ *5 cm* rectangular $[90/\pm 45/0]_s$ quasi-isotropic laminates of the AS4/3501-6 graphite/epoxy system. Three *PZT* piezoceramic patches were mounted on the surface of each specimen. The *PZT* was cut into *2 cm* $\times$ *0.5 cm* patches so that the longitudinal wave would be favored over the transverse one, and three patches were used on each specimen to actuate and accurately measure the transmitted and reflected waves. The first channel, which was served as the trigger for all of the channels, was connected to the output channel and actuating *PZT*, two others were connected to the sensing piezoceramic patches to the specimen to serve as a control channel in order to zero out drift. A few shapes fo piezoceramic patches were used to produce Lamb waves, and as expected waves propagated parallel to each edge, i.e. longitudinally and transversely for a rectangular patch and circumferentially from a circular piezo.

Various types of damages were introduced to the specimens. The groups of the specimens were formed by:

1. Drilling *6.4 mm* diameter holes into the center of each specimen as a stress concentration.
2. Compressively loading in a four-point bending fixture until audible fiber fracture damage was heard.
3. Cyclically loading in the previous fixture for 2000 cycles as *80*$\%$ of this load to create matrix cracks.
4. Using a thin utility blade to cut a *5 cm* $\times$ *2.5 cm* slot in side to form delamination.
5. Using a Telfon strip cured into the center mid-plane of the laminate to form a delmaination.

An x-ray radiograph was taken for every specimen after introducing the damages. The radiographs were taken using a die-penetrant to help document the type, degree, and location of the damage. Lamb waves were propagated to the specimens by using *15* and *50 KHz* frequencies.

Everyone of these data sets were divided into different number of f folds ($3 \leq f \leq 10$) and a subsets of data were created from these folds for every data set. When the graphs of the subsets of every data set were plotted, there were many subsets showed similar shape of graphs as shown in Figure 6. This let us to believe that the subsets of the data set can be divided into clusters, where the means of these clusters can be used as representatives to these clusters for damage detection.

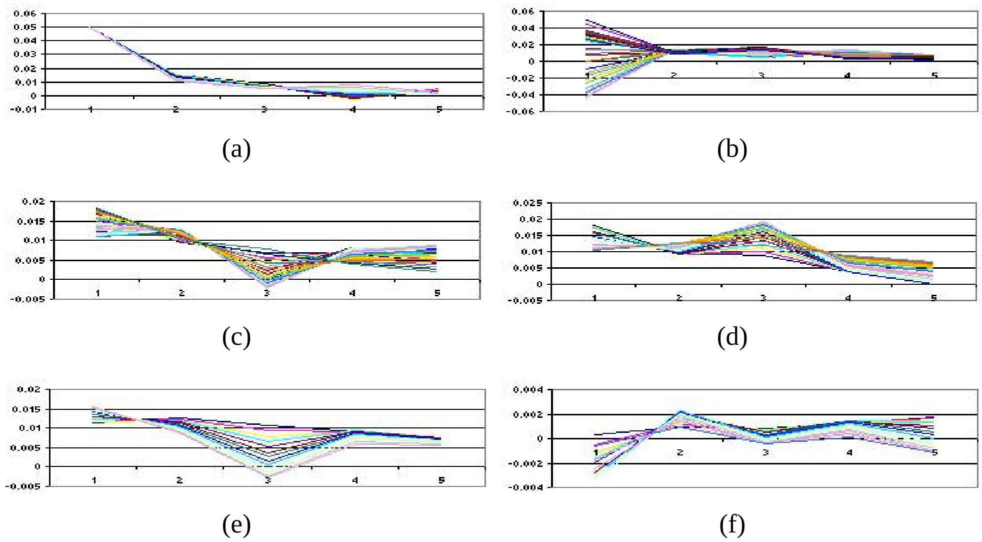


Figure 6: The figures show the similarity of wave shapes of the data sets created by dividing wave of *600* amplitudes into *10* folds.

6 Conclusions and Future Work

Bayesian networks in general and Naive-bayes in specific are a powerful formalism for reasoning under uncertainty that can be employed as classification techniques for damage detection in laminated composite materials. When using waves for damage detection in laminated composite materials, subset of features can be selected by dividing the amplitudes of the wave into folds, form new data sets, applying a clustering algorithm in these data sets, and the mean values of these clusters with or without the maximum and minimum values of the amplitudes of the clusters can be used as feature subsets for the damage detection.

In this paper, the implementation of Bayesian network and Naive-bayes as classification techniques and the f-folds feature subset selection algorithm for damage detection in laminated composite material have been shown theoretically without enough experiments, only a limited experiments for the f-folds algorithm. In the future, we are planning to conduct experimental test to determine the effectiveness of these techniques, and to determine their accuracy. We plan also to develop a technique to automatically determine the number of folds and the number of clusters from the data set.

Acknowledgements

The authors would like to acknowledge the support received from the Ministry of Science, Technology, and Innovation, Malaysia through *IRPA* by offering the grant number *09-02-04-0824-EA-0011* and the assistance provided by Rosnah Nawang of the Advanced Material Laboratory. They would like also to thank Dr. Seth Kessler for offering us the data used in this paper.

References

[1] J. Haywood, P. Coverley, W. Staszewski, and K. Worden, An automatic impact monitor for a composite panel employing smart sensor technology, Smart Materials and Structures, 14:265271, 2005.

[2] R. Miller and P. McIntire, Nondestructive Testing Handbook, Vol. 5, Second Edition, Acoustic Emission Testing, American Society for Nondestructive Testing, 1987.

[3] S. Hall, The effective management and use of structural health data, Proceedings of the 2nd International Workshop on Structural Health Monitoring, Technomic, Lancaster PA, 265-275, 1999.

[4] S. Kessler, S. Spearing, M. Atalla, E. Cesnika, and C. Soutisb, Structural health monitoring in composite materials using frequency response methods, Composites Part B, 33:87-95, 2002.

[5] Y. Shenfang, W. Lei, and P. Ge, Neural network method based on a new damage signature for structural health monitoring, Thin-Walled Structures, 43:553-563, 2005.

[6] D. Grahama, P. Maas, G. B. Donaldsona, and C. Carr, Impact damage detection in carbon fibre composites using HTS SQUIDs and neural networks Carrb, NDT&E International, 37:565-570, 2004.

[7] Z. Su and L. Ye, Lamb wave-based quantitative identification of delmaination in CF/EP composite structures using artificial neural algorithm, Composite Structures, 66:627-637, 2004.

[8] P. C. Kaminsk, The approximate location of damage through the analysis of natural frequencies with artificial networks, Process Mechanical Engineering, 209:117-123, 1995.

[9] D. Chakraborty, Artificial neural network based delamination prediction in laminated composites, Material and Design, 26:1-7, 2004.

[10] Z. Su and L. Ye, Lamb wave-based quantitative identification of delamination in CF/EP composite structures using artificial neural algorithm, Composite Structures 66:627637, 2004.

[11] F. V. Jensen, Bayesian Networks and Decision Graphs , Springer-Verlag New York, Inc., 2001.

[12] K. B. Korb and A. E. Nickolson, Bayesian Artificial Intelligence , Chaman & HallCRC, 2004.

[13] D. Heckerman, A Tutorial on Learning With Bayesian Networks, Microsoft Research Technical Report, 1995.