Node localization via analyzing multi-path signals in ultrasonic sensor networks

W.J. Tomlinson Jr., B. Dong, S. Lorenz. and S. Biswas
Michigan State University, Department of Electrical and Computer Engineering, 428 S. Shaw Lane, 2120 Engineering
Building, East Lansing, MI, USA 48824

ABSTRACT

This paper proposes a novel signal analysis based node localization strategy for sensor networks used in structural health monitoring (SHM) applications. The key idea is to analyze location-dependent multipath signal patterns in inter-node ultrasonic signals, and use machine-learning mechanisms to detect such patterns for accurate node localization on metal substrates on target structures. Majority of the traditional mechanisms rely on radio based Time Delay of Arrival (TDOA), coupled with multilateration, and multiple reference nodes. The proposed mechanism attempts to solve the localization problem in an ultrasonic sensor network (USN), avoiding the use of multiple reference beacon nodes. Instead, it relies on signal analysis and multipath signature classification from a single reference node that periodically transmits ultrasonic localization beacons. The approach relies on a key observation that the ultrasonic signal received at any point on the structure from the reference node, is a superposition of the signals received on the direct path and through all possible multi-paths. It is hypothesized that if the location of the reference node and the substrate properties are known a-priori, it should be possible to train a receiver (source node), to identify its own location by observing the exact signature of the received signal. To validate this hypothesis, steps were taken to develop a TI MSP-430 based module for implementing a run-time system from a proposed architecture. Through extensive experimentation within an USN on the 2024 Aluminum substrate, it was demonstrated that localization accuracies up to 92% were achieved in the presence of varying spatial resolutions.

Keywords: Source Localization, Ultrasound, Ultrasonic Sensor Networks, Pattern Recognition, System Design, Structural Health Monitoring

1. INTRODUCTION

This paper introduces an alternative form of localization in an ultrasonic sensor network (USN), through the use of an intelligent ultrasound pulse transmitter and receiver system, while exploiting the known geometry of a metal substrate and the occurrence of multi-path signal propagation. Traditional forms of localization invoke the principles of received signal strength intensity (RSSI), time of arrival (TOA), time difference of arrival (TDOA), and angle of arrival (AOA). In these implementations, heavy demand for reliable clock synchronization and the assistance of multiple reference nodes are required for reasonable localization accuracy. Localization through the use of ultrasound can be achieved in many different environments, through many different mediums in which a signal must propagate, and can have advantageous methods implementation over traditional localization. Primary applications of this form of localization reside in the area of Structural Health Monitoring (SHM), an area of Non-Destructive Evaluation (NDE), where sensor networks that perform advanced data acquisition techniques are used to diagnose damage over long periods of time. This need to externally monitor the physical condition of a structure is presented in order to make repairs under situations where the structure's integrity is questionable. An example of such an application can be seen in [1], where ultrasonic guided waves are to perform SHM on various structures such as the skin of an aircraft wing, with manually inserted defects. Another application [2], involves the use of ultrasound signal classification to characterize signal behavior of defects existing on composite surfaces. This leads to the conclusion that the use of ultrasonic communication has many characteristics that make it suitable for study for the application presented in this paper.

By adopting a system seen in the work done by [3], an energy efficient ultrasonic pulse based WSN is used for binary information exchange. In particular, the work presented by Lorenz et al. focusses on developing an ultrasonic pulse transceiver for the purpose of creating a pulse based sensor network that is better suited for event detection applications in comparison to packet switched forms of networking. In doing so, the establishments of a reliable pulse based

Wireless Sensing, Localization, and Processing IX, edited by Sohail A. Dianat, Michael David Zoltowski,
Proc. of SPIE Vol. 9103, 910306 · © 2014 SPIE · CCC code: 0277-786X/14/$18 · doi: 10.1117/12.2052984

communication link and novel distance estimation technique were developed to support the implementation of a frame structure for multi-hop communication on a metal substrate. However, their research used only the portions of the ultrasound signal in which accurate arrival times of the signal could be determined. And while the multipath components of the signal were used to help determine the appropriate frame structure, only the duration of this portion of the signal became of any importance. In the approach presented in this work, it is possible to develop a system that combines multipath signal analysis, to aid in source node localization on a 2024 Aluminum substrate, whose topology is distributed into a cellular manner.

The contributions of this paper are the following: (1) an offline model is first developed in which a collected signal is pre-processed using MATLAB for accuracy evaluation and analysis of location dependent features in the ultrasound multipath delay spread. (2) A TI MSP-430 based module is developed for implementing a run-time version of the proposed system architecture. For both implementations, machine training is performed offline in which an array of different classification mechanisms such as Logistic Regression Models, Feedforward Artificial Neural Networks, and Naïve Bayes, are experimented with.

2. HYPOTHESIS

The approach taken in this paper relies on a key observation that the ultrasonic signal received at any point on the structure from the reference node, is a superposition of the signals received on the direct path and through all possible multi-paths. Thus, it is hypothesized, that if the location of the reference node and the substrate properties are known a-priori, it should be possible to train a receiver (i.e., source node) to identify its own location by observing the exact signature of the received signal.

Based on this observation, a machine learning based signal analysis model has been developed and applied for a metal substrate (2024 Aluminum Plate, used for aircraft wing skins) structure, to validate and analyze the performance of the proposed node localization concepts in an ultrasonic sensor network.

3. EXPERIMENTAL SETUP

3.1 Substrate geometry

The problem formulation first begins with defining the environment for which localization will take place. The environmental factors play a heavy role in the way ultrasonic signals behave and propagate in the form of lamb waves [4]. For example, the plate thickness is one parameter that dictates how the many modes of the ultrasound signal are generated. In addition to this, the physical boundaries and properties of any plate structure will also cause multi-path reflections to occur. In this paper, 2024 aluminum (material commonly used in aircraft skin construction) is used with a thickness of 1mm, and a length and width of 3.6m and 1.2m respectively.

3.2 Ultrasound communication link

A robust ultrasound communication link is formed on the surface of the aluminum medium, consisting of a single and separate transmitter and receiver. The receiver portion of this link shall act as the source node, while the transmitter, fixed in one position (far bottom right corner) on the plate and shall as a reference for ultrasound localization.

To evaluate the lamb waves that propagate through the aluminum medium, a formal and consistent method of receiving the acoustic waves must be present. For the experimental portion of this paper, a standalone piezo-electric transducer (PZT) disc from APC International (D-9.55mm-1.00mm-850 WFB) is used as a receiver. The general functionality of the PZT for its operation as a receiver is as follows: (1) The lamb waves received create a mechanical vibration at the transducer; (2) The mechanical vibrations facilitate a process that coverts the mechanical waves into a readable voltage from the wires soldered to the PZT. It is also important to note that this PZT is the same model used in [3]. Results from this work indicate that the resonance frequency is approximately 245 kHz; therefore at this frequency, the PZT is most sensitive and has the ability to yield the highest received signal strength.

In order initiate the communication process using ultrasonic pulses, an ultrasonic transmitter is designed and implemented, and is used in conjunction with the Mica2 sensor platform. The transmitter designed is based strongly off

of the work done in [3], with a few slight modifications. The Mica2 controls the transmitter's ability to inject ultrasonic pulses every 30 milliseconds into the medium using a PZT. From the perspective of the transmitter, the PZT functions in the following manner: (1) An induced voltage generates small mechanical vibrations at the transducer disc and (2) The mechanical vibrations produce the lamb waves that propagate through the medium for the signal to be detected by a receiving PZT. The interval in which pulses are sent plays an important role in the number of multipath reflections that may occur. If pulses are sent more frequently, more pulses occur and overlap with pulses from previous transmissions also occurs. In contrast to this, if pulses are sent too infrequently, fewer reflections exist and results may yield a signal lower with amplitude, resulting in a signal with fewer peak distinctions. Additionally, the entire transmitter is designed to function at 245 kHz, since the PZT has a resonance at roughly this frequency. The transmission voltage is kept constant at 6V, as opposed to the option to select 3V or 6V from the work presented in [3]. Transmission voltage is kept at a level of 6V in order to successfully reach the entire span of the aluminum plate (i.e., receive a signal strong enough to be detected).

3.3 Cell distribution

Preliminary experiments were conducted in an attempt to characterize the entire behavior of the multi-path signals present within the substrate, in order to yield precise coordinate based localization. This approach was deemed infeasible for the context and timeframe of this work. However, these preliminary studies gave great indication that the appearance of the raw signal is indeed location dependent. To exploit this discovery, and to achieve control over the granularity of localization, the resolution of localization is mapped to a cell based distribution. Resolution of localization is varied from low to high, increasing the number of cells, while also decreasing the area on the plate that each cell governs. Thus, the overall goal is to determine the correct cell location of the unknown node, whether the position lies in the cell center or near its boundary. The cells are of equal size and the number of cells varies from 3, to 12 to 24. Figure 1 gives an example for three (3) cell distribution in the case of center of cell locations and non-centered locations. For the image on the right, representing non-centered locations, the lettered nomenclature represent the center of each cell, while numbered representations exist for labeling non-centered locations.

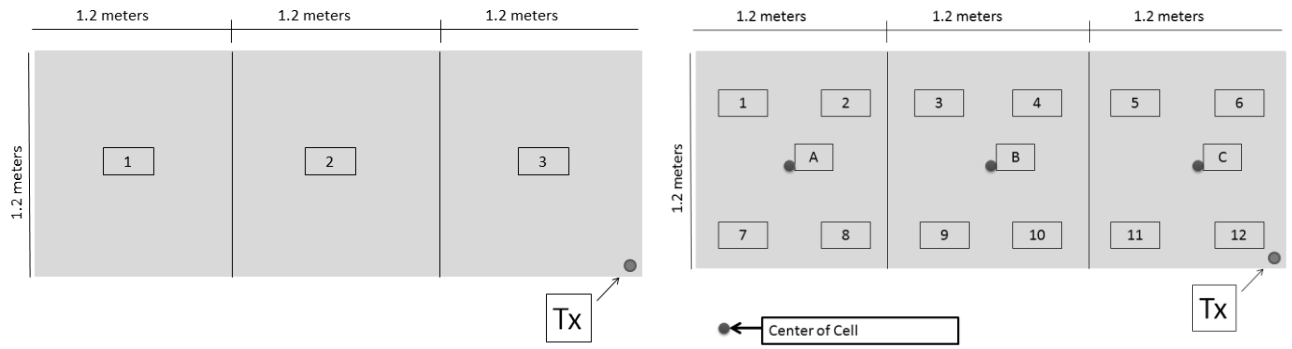

Figure 1. Diagram of 2024 Aluminum Plate, mapped into 3 Cells for center-of-cell and non-centered locations.

3.4 Coupling

Coupling inconsistencies can have a huge impact on received signal strength. Many coupling methods exist in the form of thermo bonding tape, epoxy, coupling gel and more [5]. A coupling method was devised and tested in this paper to ensure that the most consistent form of data was obtained from the aluminum plate. Figure 2 depicts the design of the coupling method used. An earth disc magnet is bonded to the PZT using epoxy glue. The layer of glue forms an even layer that is attached to the magnet. This ensures that the soldering points from the wires attached to the PZT do not provide any unevenness when the magnet is placed on top of the PZT. A magnet of reverse polarity is attached to the opposite side of the plate, to allow the PZT to be fixed to the surface.

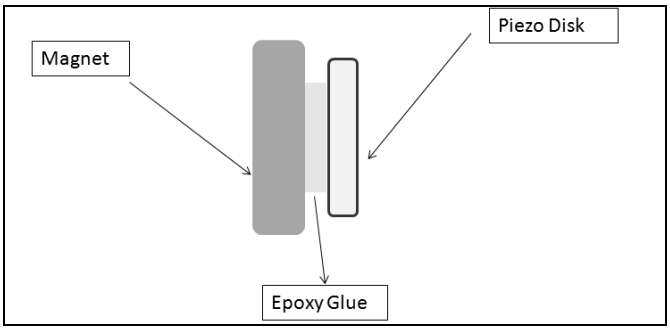

Figure 2. Coupling configuration using a magnet bonding approach.

3.5 Data collection

In order to utilize pattern classification techniques, where training and testing data must be generated, a large amount of data must be collected as an initial stage. Therefore, data was collected from the aluminum plate using the ultrasound transmitter to send the pulses and the stand alone PZT to observe the signal. To sample, view and collect data, the Picoscope 2204 version is used, which has a varying maximum sampling rate in the range of 50 to 100 Mega samples per second. In all experiments, data is sampled at a rate of 3.965 MHz. Proceeding the sampling and storing of a raw received signal, the PZT is re-coupled to the aluminum plate. This instance will be known as one coupling trial. This data collection method of the raw received signals takes place at the center of each cell, in addition to at least two positions that lie outside the center of each cell and lie close to the boundaries of another cell. Each cell centered and non-centered location measurement is a product of data collected from twenty coupling trials, in order to account for variations in the signal due to coupling. The purpose of data collection outside the center is to be able to assess the amount of confusion that may exist when a position is close to a neighboring cell.

For each cell localization resolution, non-centered locations are placed under the same steps for data collection. For 3-Cell Localization Resolution (3CLR), there exist a total of 4 non-centered locations in each cell. For 12-Cell Localization Resolution (12CLR), there exist a total of 2 non-centered locations in each cell, and the same applies to 24-Cell Localization Resolution (24CLR).

Results of data collection support the results from preliminary experiments. By examining the characteristics of the raw signal obtained through data collection, certain observations were discovered that are used to create a localization framework using the appearance of multipath reflections that exist in every signal measured. In the work done by [3], multipath reflections were not of interest. The dominant lamb wave modes, Symmetric (S) and Asymmetric (A) modes are the only parts of the signal that carry useful information in the form of such parameters as wave velocity and wave dispersion. Based on the dispersion curve diagrams for the aluminum plate [6], it can be determined that the frequency at which the ultrasound is transmitted (245 kHz), and the thickness of the plate, allow for only the first incident of the S (S_o) and A (A_o) waves to propagate [6]. Further examination, has led to the conclusion that the patterns of multipath reflections at random positions on the aluminum plate, yielded an almost unique multipath spread from a visual prospective Figure 3 shows an example of such, where 4 signals are depicted randomly from chosen cells within a distribution of 12 cells.

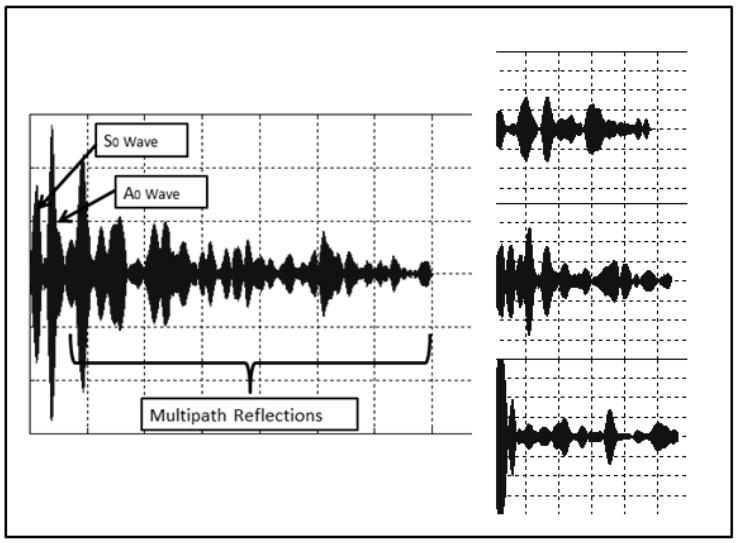

Figure 3. Examples of received signals at various locations on the aluminum substrate.

4. OFFLINE LOCALIZATION MODEL

The development of an offline model is the first step taken to implement node localization by exploiting the properties of ultrasound signal propagation on an aluminum medium. Discriminatory features of the raw received signals will be used to determine which grant the highest degree of differentiation when examining received waveforms on a cell-to-cell basis.

4.1 Pre-processing system design

Pre-processing is a commonly used step in signal processing before extracting useful information from a signal. The benefits of pre-processing include data size reduction, and the elimination of noise and variability in signals that originate from the same conditions in which data was collected. In this paper, a system is implemented in MATLAB in which a number of pre-processing techniques are implemented to produce the envelope of the raw signals obtained at various locations throughout the aluminum plate. From the envelopes, features will be extracted to obtain and improve classification results for the purpose of differentiating locations on the plate in a cell-by-cell case. Figure 4 depicts the block diagram for the pre-processing system implemented in software.

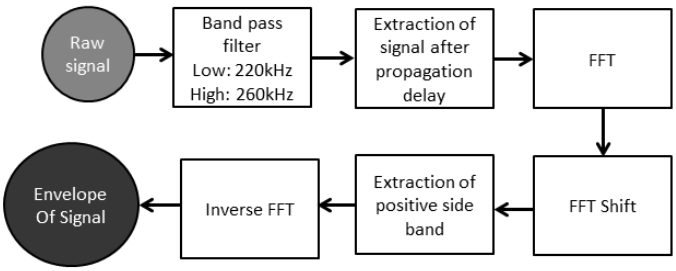

Figure 4. Offline localization model system block diagram.

The first step, after collection of the data, uses a filter to remove any unwanted frequency components. The MATLAB Signal Processing Toolbox was used to implement a 10th order Butterworth band pass filter, with a pass band of 40 kHz, passing frequencies in the range of 220 kHz to 260 kHz. These frequencies represent the high pass filter cutoff and low pass filter cutoff respectively

Based on the work studied in [7], the frequency component of ultrasound signals can provide useful information when utilizing applications that involve some form of machine learning and pattern recognition. Ideally, eventual steps will lead to analyzing the Fourier transform of the band pass filtered signal. However, data collected in the form of the raw signal does not only include the signal itself. It also includes the portion of the collected data that depicts the propagation delay the signal underwent. This delay causes the actual signal to occur later in the set of data and not at the start of when the data is collected. Thus, the samples obtained during this segment of the signal, are not the actual signal and can alter the appearance of the signal in the frequency domain. To remove the effect of such uncertainties, a manual extraction of the signal is done based on its position on the aluminum plate, which corresponds to a certain propagation delay. The index in the data in which the signal began and ended was taken as the new data set for the band pass filtered raw signal. The start of the signal was chosen from the observance of the first dominant wave after the propagation delay. The end of the signal was determined by observing the point at which the last dominant reflection occurred before an interval of inactivity (no pulse transmission) became visible. This indexing process was done for 5 random coupling instances. After averaging the results among the randomly chosen 5 instances, this indexing was applied to all 20 coupling instances belonging to each particular cell. Figure 5 depicts the raw band pass filtered signal, with the manual extraction of the signal minus samples from the effects of propagation delay and period of signal inactivity.

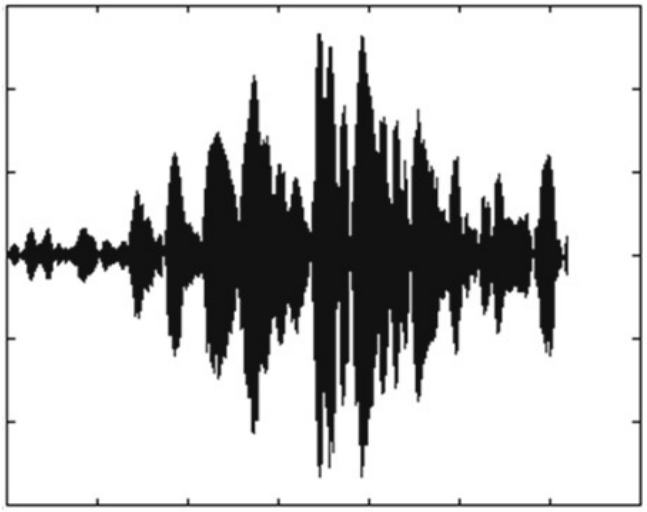

Figure 5. Band pass filtered signal after manual extraction.

To begin the process of generating an envelope of the received ultrasound signal, a discrete Fourier transform of the signal data up until this point in the system chain, is computed using the Fast Fourier Transform (FFT) MATLAB algorithm. The output of the FFT is shifted such that it rearranges the FFT by moving the zero-frequency component to the center of the array. This step is useful for visualization purposes. The symmetric nature of the FFT MATLAB algorithm allows for the signal to appear to have a mirror representation of itself in the negative portion of the frequency axis. Thus, both the positive and negative representations of the signal in the frequency domain hold the same information.

Therefore, while extracting the signal for the purpose of performing the Inverse Fourier Transform, it is presumed that only the positive side band of the signal is needed. Any additional information would just add more redundancy to the data set. To extract the positive side band, manual determination of the bands start and endpoint are determined in the same process applied in the extraction of the band pass filtered raw signal. Once this step is completed, the Inverse Fast Fourier Transform (IFFT) MATLAB Algorithm is computed to produce the desired signal envelope.

4.2 Feature Extraction

To confidently asses what distinguishable features may be present in the envelope, extensive study was conducted on the physical appearance and attributes of the ultrasound signals. This study, combined with the knowledge gained in relation to lamb (surface) wave propagation and the signal's shape dependence on the proximity to the ultrasound transmitter was used to determine a suitable set of features to extract.

The first feature to be chosen and examined was the auto and cross correlation coefficient of each signal, where the statistical measure of similarly is computed against the signal itself and against a different signal, respectively. A set of computations were done for the 3CLR case for proof of concept, and then extended to the higher cell localization resolutions. MATLAB is used to provide the algorithm for the cross correlation (XCORR) coefficients computed in this work. The correlation coefficient was computed and a distribution was plotted for all permutations (with no repetitions) of coupling instances within each cell and all permutations against neighboring cells. Figure 6 shows an example of the inter-cell and intra-cell correlation with respect to cell 1, for 3CLR. From the figure it can clearly be shown that the expected results match the results obtained. The intra-cell correlations yielded a distribution containing higher values of the XCORR coefficient, while the inter-cell correlation, yielded lower values within the plotted distributions. To explain how a tangible feature was produced from the XCORR coefficient, let $L_i^m \; for \; i = 1 \ldots 12 \; and \; m = 1 \ldots 20,$ represent the waveforms for the non-centered locations (3CLR) and let $C_j^k \; for \; j = A, B, C \; while \; k = 1 \ldots 20,$ represent the waveforms for centered locations (3CLR). Thus, the correlation coefficient feature is represented by Equation 1. The values *j* and *i* are used to represent cell indices, while *j* and *k* represent the various coupling trials.

$$Xcorrf_j = \frac{\sum_{all\ k} Xcorr(L_i^m, C_j^k)}{20} \tag{1}$$

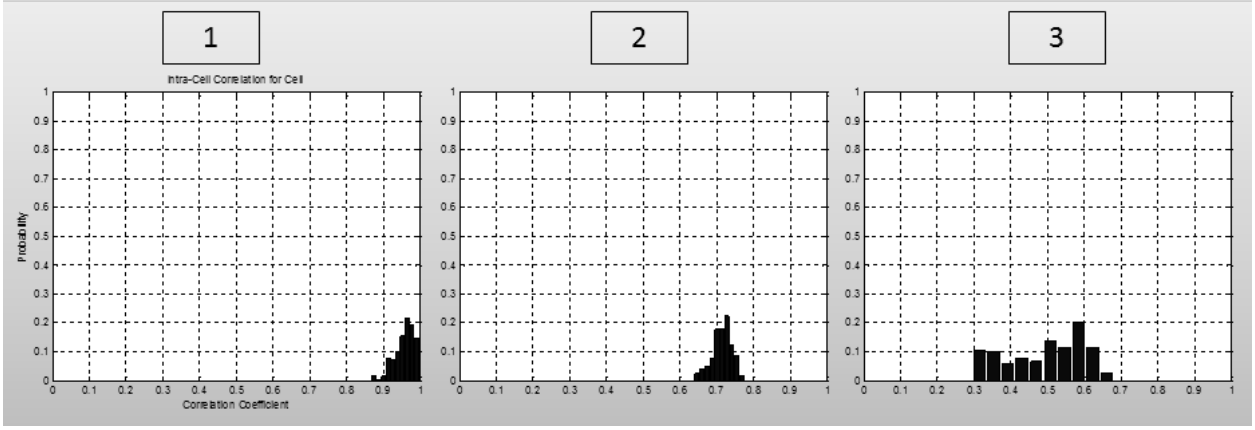

Figure 6. Inter-cell and intra-cell correlation with respect to cell 1 for 3-cell localization resolution.

When expanding this experiment to consider the data for non-centered locations, results exhibit a different trend. Figure 7 depicts the intra-cell correlation between Cell A and all of its respective non-centered locations. It can be shown that there exist correlation coefficients that overlap within the same range of values for the correlation between the center of Cell A and the center of Cell B, as well as other non-centered locations within Cell A. Therefore, the ability to distinguish a non-centered location within Cell A from neighboring locations (in and out of the cell), has been reduced substantially. It is here that the necessity of other discriminating features becomes evident.

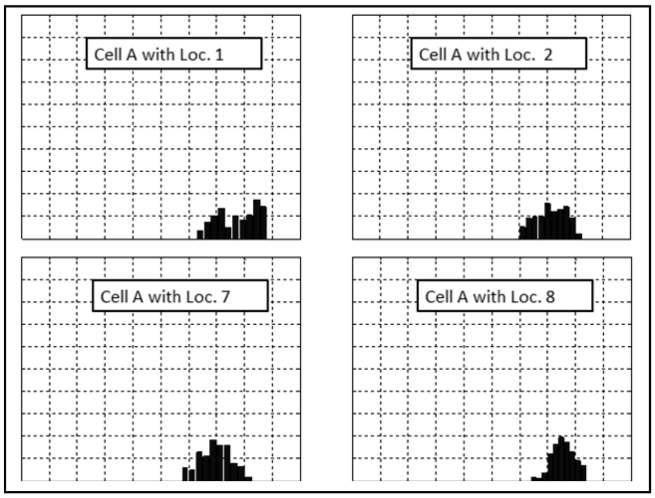

Figure 7. Intra-cell correlation between Cell A and all of its respective non-centered locations.

The notion of signal length (i.e., the number of samples representing the signal), is used for the next set of features extracted from the ultrasound envelope. It was discovered that each non-centered location has a value of signal length that is closest to the cell that it belongs in. For example, cell A's signal length is closest to the lengths of non-centered locations 1, 2, 7 and 8. To be able to make use of such a feature, the manual method of extracting the signal was replaced with script file in MATLAB that does the detection. Following the same notation for the center and non-centered locations in the previous sub section, the feature for the signal length is represented in Equation 2.

$$Lengthf_j = \left| length\left(L_i^m - \frac{\sum_{all\ k} C_j^k}{20} \right) \right| \tag{2}$$

Examining the characteristics of each envelope, a small relationship was discovered between the number of peaks present in the signal and the ultrasound receiver's location in relation to its proximity to the edges of the aluminum plate. For some positions close to the boundary of the plate, the presence of peaks becomes more dominant. This is also slightly intuitive, for the reason that one would expect more reflections to occur close to the edges of the plate, thus creating more noticeable peaks. However, the definition of what actually classifies as a peak needed to be formally set. To mitigate this issue, a threshold is used as a baseline for peak detection. Data points that lie above the threshold are evaluated for their relative maximum peaks. As a result, the feature of *"Number of Peaks above Threshold"* is extracted to take advantage of such observations. For some waveform, $w(n)$, a peak exists at a point n', if the following is satisfied:

$$\begin{gathered} If\ there\ exisits\ some\ value(a) > THR(threshold\ value) \\ if\ w(n') > w(n), when\ |n - n'| < a \\ Where\ THR = \beta\ x\ Global\ Maximum\ Peak \\ And\ \beta = \{.10,\quad .20,\quad .30,\quad .40,\quad .50,\quad .60,\quad .70,\quad .80,\quad .90\} \end{gathered} \tag{3}$$

Using MATLAB, a script is created that varies the value of β in order to find the optimal threshold value that will be evaluated later for the purposes of finding which threshold yields higher inter-class distinguishability. It was determined that the sensitivity to the threshold decreased substantially after 50% of the Global Maximum Peak value.

Upon further examination, another location dependent observation was made. It was discovered that on average, locations close to the ultrasound transmitter will have a lower amount of samples from the start of the envelope to the location of its maximum peak value. In contrast, on average, locations that are further away from the transmitter have a

larger amount of samples present between the start of the envelope and the index of its maximum peak value. From this observation, the feature of *"Number of Samples from the Start of the envelope waveform to the location of its max peak"* was extracted. To define this feature more formally, let μ be the number of samples from the index of start of the envelope (ε), to the index of the first max peak. Using the additions of these two notations, the following is defined:

$$[maxpoint, location] = GlobalMAX\left(L_i^m\right)$$
$$where\ \mu = location - \varepsilon \tag{4}$$

Using the standard deviation as a feature is common practice in applications using pattern classification [2]. The standard deviation shows how much variation is present from the average envelope value. The standard deviation was added as feature to increase the feature space in which classification will be performed. The intent is to increase the overall classification accuracy.

The Kurtosis value, also used in [2] as a feature, seemed plausible due to its relationship with probability distributions. By using the correlation coefficient distributions generated, the kurtosis value can be used to characterize its behavior in accordance to the shape of the distribution. High kurtosis valued distributions have sharper peaks and longer and fatter tails. Low valued kurtosis distributions have more rounded peaks and shorter and thinner tails. By observing the XCORR coefficient distributions, it can be seen that the distributions that yield higher values of correlation, for the center of cell cases, have higher kurtosis values.

The value of skewness, also based on the shape of a probability distribution, is also evaluated for the use of providing similar benefits to that of using the kurtosis value. The skewness can take on positive or negative values. Negative skew indicates that the tail on the left side of the distribution is longer or fatter than the right side. Positive skew indicates that the tail on the right side is longer or fatter than the left side. Distributions that have higher values of correlation with a particular cell or non-centered location tend to skew towards the right, while the contrast case skews to the left. Though the cases presented in the standard deviation, kurtosis and skewness may not be entirely consistent, these features were chosen with the notion of being able to increase the amount of differentiability among cells.

4.3 Localization Method

The learning and evaluating procedures done during pattern classification prove very useful when attempting to generalize and represent data sets. Generalization of the data would allow the system to perform well under data instances that have not yet been seen or processed by the system. Achieving generalization will be discussed in a later section of this paper. Representing the data has been achieved through the data collected from the aluminum plate. Therefore, in this chapter, classification algorithms used to evaluate the data represented in the localization system are introduced. One approach taken at classification is a rather simplistic method, while the remainder of the algorithms are utilized via the popular machine learning software, Weka (Waikato Environment for Knowledge Analysis), developed at the University of Waikato, New Zealand [8].

To properly evaluate the performance of the features extracted for the case of strictly center of cell classification, an algorithm was developed that utilizes a rather simple decision rule. The Maximum Average Correlation Coefficient (MACC) Algorithm is used to classify the center of cell positions on the aluminum plate and it is implemented via MATLAB. Step (1) is used as a training phase, in order to simulate the functionality of the localization system if it were implemented in real time. The selection of a random waveform is chosen from the pool of data collected from each cell and all 20 instances, totaling 60 envelopes for 3CLR, 240 envelopes for 12CLR and 480 envelopes for 24CLR. Step (2), removes the random waveform from the pool of collected data as well as one random waveform in the form of one instance of coupling from every other cell. This allows for the averaging of the result that takes place in Step (3), (where the cross correlation of the random waveform is computed against all cells) via Equation 5, to not be affected by one value of the variable *Xcorr* that will have a correlation coefficient of 1 (due to the comparing of the randomly selected waveform with itself at some point in the summation). The other cells not containing the randomly selected waveform also have one instance removed to account for the removal of the waveform from the cell of origin. Finally, in Step (4), Equation 6 is used to classify the cell of origin of each randomly selected waveform. Once an array of averaged

correlation coefficients is computed, index belonging to the cell with the maximum XCORR coefficient is chosen to be the cell of origin of the random waveform previously selected. The algorithm continues to select random waveforms until the entire pool of collected envelope data has been selected, thus depleting the entire set of data.

$$C_i = \frac{\sum_{j=1}^{j=19} Xcorr(w, w_i^j)}{19} \tag{5}$$

and

$$C_{est} = \max(C_i), for\ i = 1 \ldots (3\ or\ 12\ or\ 24) \tag{6}$$

It is important to note that the overall goal of paper was not to develop novel algorithms for the purpose of creating a pattern classification approach to ultrasound localization, but rather to assess whether such a system was possible with a high degree of accuracy. Thus, Weka was sought out for the purpose of obtaining diversity in classification accuracy by taking advantage of the many classifiers available through Weka's infrastructure. Weka is used to classify the non-centered locations, due to the fact that the XCORR coefficient does not provide similar results as the center of cell cases. The training set for this system is composed of all the features mentioned in Chapter 5, with each center of cell location as the reference for generating the features for the non-centered case. To test each class, 10-Fold Cross Validation is used. Using Weka, the performance of eight (8) common classifiers was evaluated for the same set of training and testing environments.

5. RUN-TIME LOCALIZATION

Thus far, any attempts presented in this paper for successful ultrasound localization have all been implemented via MATLAB, with previously collected data, and have had very little consideration for real world constraints that are always present in the creation of a real system. As a result, a run-time system will be implemented to represent a new version of the ultrasound receiver.

5.1 System constraints

The localization system implemented in MATLAB relies on a few key factors that contribute to its performance. For ultrasound pulse transmission, a constant 6V supply is applied to the transmitter. For data collection via the stand-alone piezo, an oscilloscope with an extremely high sampling rate is used to reconstruct the signal. For processing and storing, a desktop machine is used in combination with powerful software such as MATLAB and Weka to perform the classification directly on the host machine. Unfortunately, some of these elaborate approaches prove infeasible for a run-time system. As a result, the areas mentioned will now become constraints for the development of the system. The transmission level of 6V will now be used as an upper bound for all voltage levels delivered to the run-time system. Data collection will be no longer done at high sampling rates. Instead, an approach at uncovering a lower bound on the sampling rate needed to maintain high localization accuracy was determined. Experiments were done in MATLAB to determine such a value. Results indicate that a sampling frequency of approximately 25 kHz is needed to maintain previous results of localization accuracy for 3CLR and for 12CLR and 24CLR, approximately 50 kHz. Processing of the signal to produce an envelope will need to be accomplished in hardware, to alleviate the level of computational complexity needed in the use of a commercial microcontroller, which will also be used to sample the signal and send the processed data to a base station like machine. Utilization of such a method, allows for the host machine to be alleviated of most of the tasks involved in localization as seen in the offline model.

5.2 System design

Figure 8 depicts the end-to-end block diagram for the entire system, beginning with the input of the raw signal. The run-time system operates on 3.57 Volts (Vcc of the microcontroller), and the performance of each stage of the system was heavily evaluated and compared (when applicable) with the offline model implemented via MATLAB.

The amplifiers are used to increase the signal amplitude to a voltage level suitable enough to be detected by the envelope circuit, and to be able to compensate for the attenuation that it will present. In order to design an amplifier with a gain

capable of providing sufficient amplification of the signal throughout the entire plate, experiments were conducted to measure the received voltage level of the ultrasound signal at maximum plate distances (i.e., the 3 remaining corners of the plate). The gain measurements of the amplification are set in such a way to avoid the loss of signal data in the form of peaks (through the effects of saturation or attenuation).

The creation of the envelope signal is produced in hardware by utilizing a germanium diode in series with a resistor and capacitor in parallel. The performance overall is visually good in terms of the overall shape similarity between the two waveforms. There also exist high frequency components that could alter the performance of the overall system if they are not properly dealt with. Under a closer inspection, it can be seen that the frequency components that are unwanted lie in the 245 kHz range and the harmonics associated with it. The dominant frequency components of the envelope do not exceed 25 kHz.

As a result, a traditional 3rd RC low pass filter (cutoff of 50 kHz) was implemented to remove the frequency components that are unwanted and that may cause aliasing to occur before the sampling done by the Analog-to-Digital Conversion (ADC) Channel used by the microcontroller (MSP430).

Lastly the comparator and IRIS mote (The IRIS Wireless Sensor Network Module), provided by Memsic, Powerful Sensor Solutions, work in conjunction with one another to act as a timer triggered switch that is used to inform the microcontroller to sample the received envelope on its rising edge. This sampled data is then sent to the PC to calculate and display the classification of cell positions in run-time. It is important to note that the experimental localization results of the run-time system are only presented for localizing positions within the center of the cells for 3CLR.

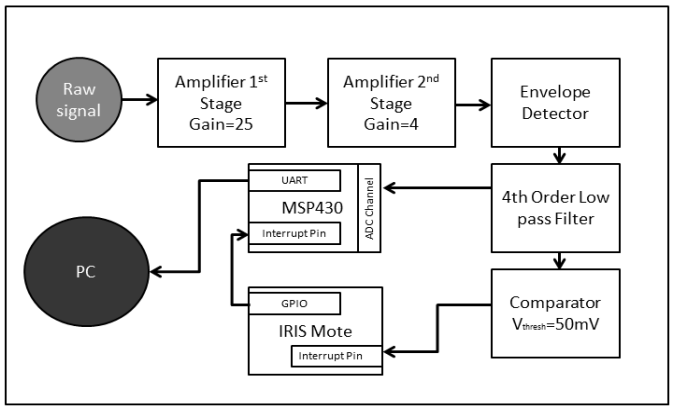

Figure 8.Run –time system block diagram.

5.3 Localization method

Once the data is received on the serial port of the PC, a terminal program equipped with a plot function is used to verify the correctness of the data sent, but MATLAB is again used for the final steps in classification. Figure 9 illustrates the system implemented in MATLAB that handles the post-processing of the sampled data leading up to cell classification. MATLAB serial port functions are used to bring in the data from the MSP430. Once acquired, any additional noise that may be present is also removed through the use of a 2nd order Butterworth low pass filter with a cutoff of 45 kHz, implemented in software. The filtered envelope signal is then down sampled to the appropriate length of sample points needed to give the signal equal length to the software produced envelopes previously generated in the offline model. This length requirement is essential for the computation of the cross correlation coefficient between envelopes. The software produced envelopes will now serve as a basis for comparison between each of the hardware produced envelopes, in order to run the MACC algorithm for the center-of-cell case. Figure 10 shows the comparative performance of the envelope reconstruction via the MSP430 and MATLAB.

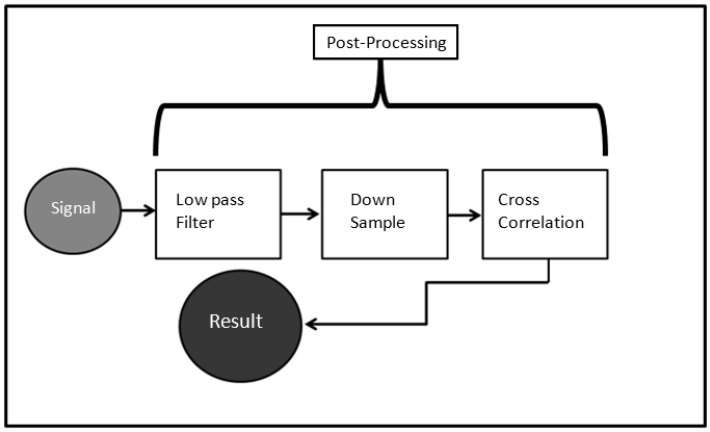

Figure 9.Block diagram of localization method for run-time system.

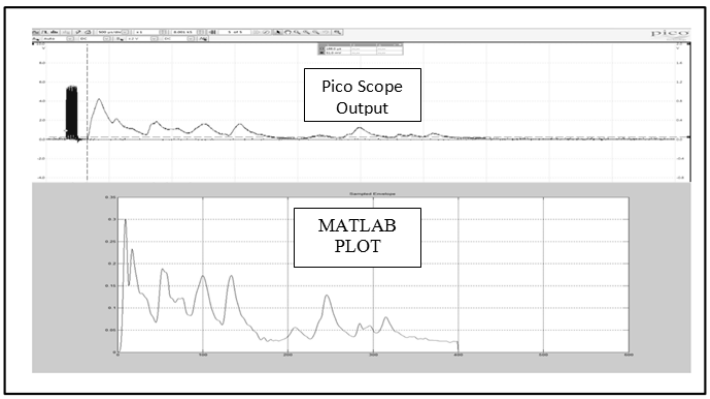

Figure 10. Performance of envelope reconstruction.

6. SYSTEM PERFORMANCE

The performance of the classification algorithms implemented in MATLAB, utilized within Weka and experimentally via a run-time system is presented. Results of the localization accuracy for the offline model are categorized by cell resolution, containing the performance results for center of cell classification and non-centered location classification under the MACC algorithm and the Weka in-house classifiers, respectively. Results of the localization accuracy under the run-time system are presented only for the center of cell case, and for 3CLR.

6.1 Offline localization model

The performance of the MACC algorithm under 3CLR, 12CLR and 24CLR yield promising results for future analysis of the non-centered locations. The results for 3CLR show that there was no confusion between neighboring cells and that the classification accuracy was 100%. For the 12CLR case, the accuracy obtained was 98.75%, with minor confusion existing among adjacent cells. The case for 24CLR also shows a similar trend. Accuracy for this case lies slightly below the 12CLR case, at roughly 98.54%. From these results it can be seen that the MACC algorithm is a sufficient as a stand-alone classifier, needed to correctly assign randomly chosen waveforms to their proper cell center.

Applying the MACC algorithm for the following experiments: (1) the 12 non-centered locations compared with each cell center for the case of 3CLR, (2) the 24 non-centered locations compared with each cell center for the case of 12CLR, and (3) the 48 non-centered locations compared with each cell center for the case of 24CLR, the results differ greatly from the center of cell case. The overall accuracy of the 3CLR case was roughly 58.08%. For 12CLR, the accuracy obtained was 27.92% and for the last case (24CLR), 18.13%. In an attempt to study or determine a pattern of confusion, the data

collected during classification was used to generate more informative means of presenting the confusion that is present within this set of classification results. Table 1 illustrates a comprehensive view of the localization performance under the MACC Algorithm for both the center of cell and non-centered locations. For a given non-centered location, results indicated large amounts of overlap with the correlation coefficient distributions between cells A, B and C. Based on these observations, and the observance that each cell resolution exhibits more confusion as the number of cells increase, it has been concluded that the other resolutions will be equally as bad in terms of overlap. Therefore, the features that were previously extracted were evaluated, in an attempt to improve the accuracy that has been achieved by only using the MACC Algorithm. Table 2 lists the localization accuracy under all 8 classifiers used in Weka, under each cell resolution

Table 1. Localization performance under the MACC Algorithm.

Localization performance under the MACC Algorithm – Offline Localization Model		
Localization Resolution	**Center-of-cell Location Accuracy**	**Non-centered Location Accuracy**
3CLR	100%	58.08%
12CLR	98.75%	27.92%
24CLR	98.54%	18.13%

Using Weka in conjunction with the features extracted, provided an increase in the poor classification rate given by using only the MACC algorithm. For all the cell resolutions, the overall accuracy yielded an increase. The best classifier for 3CLR was the Multilayer Perceptron, with an accuracy of 92.08%. For the 12CLR case, the best classifier in terms of performance was Simple Logistic Regression, with 97.29%. Lastly, the 24CLR case shows a maximum classification accuracy of 96.25% under Logistic Regression. It is important to note that note all classifiers resulted in a similar trend of increasing accuracy. It is also important to note that in Weka, each resolution, existing with a different number of features, counts as a different experiment and does not follow the decreasing trend in accuracy as the number of cells increase. However, the statement can be made that with the addition of each feature, an improvement in the overall accuracy of classifying the non-centered locations into their respective cell of origin is present.

Table 2. Localization accuracy among all 8 classifiers used in Weka.

Localization accuracy for Non-Centered Locations – Offline Localization Model			
Classifier	**3CLR Accuracy**	**12CLR Accuracy**	**24 CLR Accuracy**
Multilayer Perceptron	92.08%	96.46%	93.44%
Meta Bagging	91.25%	89.58%	90.00%
RBF Network	84.17%	95.83%	92.34%
J48 Decision Tree	89.17%	84.17%	78.65%
Logistic Regression	82.92%	94.79%	96.25%
Simple Logistic Regression	84.58%	97.29%	95.52%
SMO	86.67%	95.42%	90.63%
Naïve Bayes	78.75%	87.08%	75.21%

6.2 Run-time localization

The performance for 3-Cell Localization is depicted in the form of the confusion matrix for the experiments done to compare the performance between the offline model and the run-time system. It can be shown that most of the confusion lies at cell 2.

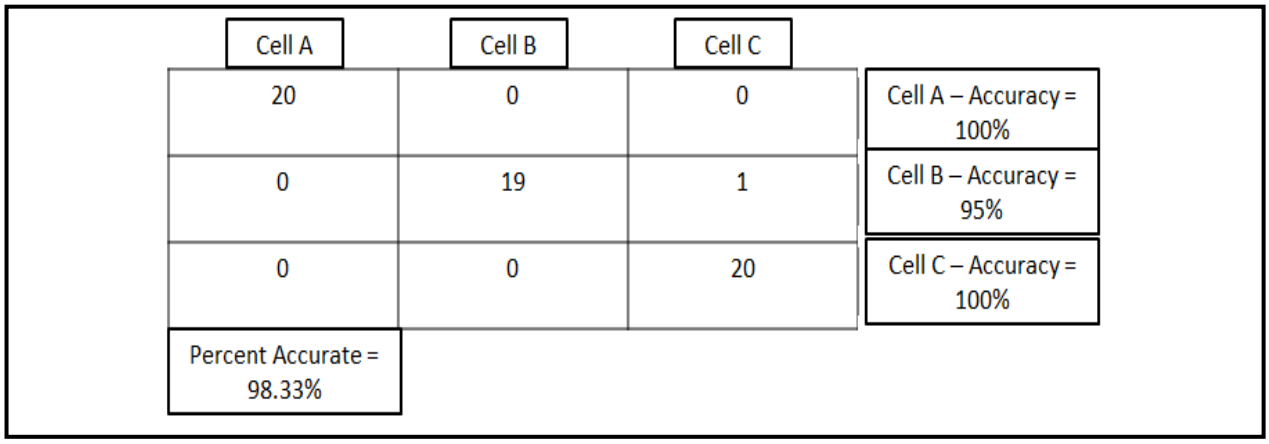

Figure 10.Confusion matrix for 3CLR, run-time system localization.

7. CONCLUSIONS AND FUTURE WORK

In this paper, a localization system for the discovery of the cell-based position of a single unknown node using ultrasound communication on a metal substrate (2024 Aluminum Plate) is introduced. The traditional functionality of localization in the context of wireless sensor networking involves using time of arrival (TOA) family based techniques to accurately estimate the position of an unknown network node in a particular environment, with high accuracy and precision. The use of these hardware and software systems opens up the avenue to many real life constraints during its implementation phase. In the work presented in this paper, the basic properties that govern mechanical carrier waves (lamb waves) are used to aid in the problem of solving localization in an ultrasonic sensor network, without the use of traditional techniques. Using the boundary dependent multipath reflections of the lamb waves, the position of a single node is achieved in a cell-based manner on a metal substrate. Using MATLAB, the raw received waveforms of the ultrasound signal are pre-processed to produce an envelope of that signal that which was discovered to be both location and multipath reflection dependent. Through extensive amounts of data collection and signal processing, the study of the waveform produced on the receiving end of ultrasound communication was conducted. As a result, features were determined, extracted and evaluated using pattern classification techniques to provide distinguishability among received envelopes. Utilizing these features, and varying complexity levels of classification algorithms and techniques, center-of-cell and non-center-of-cell localization is performed on the data collected from all possible cells mapped out on the aluminum plate. This set of experiments deemed it possible to achieve localization using a less understood concept in relation to widely known and used time of flight based approaches. As a result, it yielded the implementation of a run-time system capable of achieving similar localization results. The run-time system generates the envelope of the received waveform purely through hardware and utilizes a microcontroller in conjunction with a popular sensor networking platform to intelligently sample the envelope signal at its correct start time. This sampled data is sent to a base station like device where the localization results are computed in run-time, using the same set of sampled data collected for use in the offline model as a training reference to implement the localization. Final results prove that it is possible to achieve high localization accuracy while generalizing the system to a point where training data of higher resolutions can be utilized in conjunction with removing some of the limitations placed upon traditional approaches of localization in WSNs, both in theory and in a practical implementation.

Future work in regards to this research problem will include the evaluation of the localization accuracy in run-time of higher cell resolutions, increasing the number of cells while still using the same pool of data collected in initial experiments as a basis for cell localization. In addition to this, the use of the machine learning software Weka will be extended to localize the non-center-of-cell locations in the presence of the run-time localization system. The future results stemming from this extension will also be compared to the results generated in the theoretical system. Lastly, an important area of future work lies in attempting to optimize the operation of the run-time system, in the hopes of achieving high cell based localization accuracy with the most minimal of hardware and software requirements for the different stages of the system.

REFERENCES

[1] Zhao, X. et al., "Active health monitoring of an aircraft wing with embedded piezoelectric sensor/actuator network: I. Defect detection, localization and growth monitoring." Smart materials and structures 16(4), 1208 (2007).

[2] Anastassopoulos, A.A., Nikolaidis, V. N. and Philippidis T. P., "A comparative study of pattern recognition algorithms for classification of ultrasonic signals," Neural Computing & Applications 8(1), 53-66 (1999).

[3] Lorenz, S., Dong, B., Huo, Q., Tomlinson, W. J. and Biswas, S., "Pulse based sensor networking using mechanical waves through metal substrates," Proceedings of the SPIE 8753, (2013).

[4] Giurgiutiu, V. and Soutis, C, "Enhanced composites integrity through structural health monitoring," Applied Composite Materials 813-829 (2012).

[5] Ayrulu, B. and Barshan, B., "Comparative analysis of different approaches to target classification and localization with sonar," Multisensor Fusion and Integration for Intelligent Systems International Conference, 25-30 (2001).

[6] Bin, L., Giurgiutiu, V. and Ayman, M.K., "The use of exact Lamb waves modes for modeling the power and energy transduction of structurally bonded piezoelectric wafer active sensors," Proceedings of the SPIE**,** Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 8345, (2012).

[7] Hajrya, R., Verge, M. and Mechbal, N., "Active damage detection and localization applied to a composite structure using piezoceramic patches," Control and Fault-Tolerant Systems (SysTol), 2010 Conference 849-854 (2010).

[8] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and I. Witten, H. "The WEKA Data Mining Software: An Update," SIGKDD Explorations 11, (2009).