Anomaly Prediction in Non-Stationary Signals using Neural Network Based Multi-Perspective Analysis

Abdullah Alshehri[1], Aaron Waibel[2], Soundararajan Ezekiel[2]

[1] Electrical Engineering Department, King Abdulaziz University-Rabigh, Saudi Arabia
[2]Department of Computer Science Indiana University of Pennsylvania, Indiana PA, USA
*E-mail: ashehri@kau.edu.sa

Abstract: A new technique for predicting anomalies in the near future of an observed signal is being presented. Before any data analysis can be performed on an observed signal, the signal's underlying pattern must be cleared. A wavelet de-noising scheme is used because it provides a better result compared to other de-noising algorithms and it is simple from a computational standpoint. Robust peak-finding algorithm is used to find smaller anomalies that appear frequently throughout the signal pattern. In addition to or in place of wavelet de-noising, other views of the signal may be generated for analysis. The generated perspectives are used as input to a feed-forward neural network that will predict the likelihood of an anomalous event occurring later in the signal. The neural network is trained using the Resilient Backpropagation of Errors (Rprop) supervised learning algorithm with data sets consisting of a mix of signals known to precede anomalous events as well as signals known to be free of significant anomalies. This paper provides a means of predicting large or abnormal events in signals such as seismograms, EKGs, EEGs, and other non-stationary signals. Our algorithm has been tested on a large collection of seismic and EKG (electrocardiogram) signals. The obtained accuracy as high as 70% with EKG signals and as high as 83% with seismic signals, when the test data is taken from within the same time frame as the training set. Though there was greater consistency found at a lower degree of accuracy for seismic signals.

[Abdullah Alshehri, Aaron Waibel, Soundararajan Ezekiel. **Anomaly Prediction in Non-Stationary Signals using Neural Network Based Multi-Perspective Analysis**. *Life Sci J* 2014; 11(6):685-693]. (ISSN: 1097-8135). http://www.lifesciencesite.com. 104

Keywords: Wavelet, Neural Network, Pattern Recognition, Seismic Signal, Anomaly Prediction, Signal Processing

1. INTRODUCTION

Accurately detecting and predicting anomalies in signals is an issue of importance in many fields. Techniques for finding and predicting future anomalies in signals are an asset in the fields of seismology (Andreas et al. 2010), medicine (Mikheld, Daqrouq 2008), meteorology (Zhang et al. 2007), and speech recognition (Mitsuru et al. 1998). EKG, EEG, and seismic signals are part of a class of time dependent signals known as non-stationary signals. Because of their dependence on time, non-stationary signals can be particularly hard to analyze for anomalies.

The electrocardiogram (EKG) measures the electrical activity of the heart over time. Physicians use the EKG to diagnose heart conditions with symptoms that manifest themselves as anomalies in the cardiac rhythm such as heart attacks and arrhythmia (Fensli et al. 2005). Detecting anomalies in EKGs then is a vital part of accurately diagnosing patients and providing proper treatment. Many current methods for finding anomalies in the heart beat focus on statistical time series analysis (Chuah, Fu 2007; Boucheham 2011). Time series analysis techniques, however, struggle to detect anomalies in noisy signals as well as those that have varying amplitudes between data sets (Cheboli 2010). The technique we present for anomaly detection is centered on the use of neural networks to recognize patterns indicative of anomalies in EKGs.

The electroencephalogram (EEG), similar to the EKG, is a measure of electrical activity over time. The EEG, however, records electrical activity in the scalp to characterize brain activity. The EEG varies much more than the EKG as the brain may be in a variety of states (Teplan 2002). Because of this the EEG has many applications including: monitoring anesthesia depth, find areas of brain injury, investigate epilepsy, and many more (Teplan 2002). Though we do not test our method with EEGs, similar techniques to ours have shown promising results in classifying EEG signals (Neep et al. 1997).

Accurately predicting anomalies in seismic signals is an important subject of ongoing research in the field of seismology. A seismogram is usually composed of three time series collected from two horizontal directions and one vertical direction (Ramirez 2012). Analysis of seismic signals is useful for finding natural resources underground and for prediction of earthquakes (Ramirez 2012). The technique that we propose in this paper provides a means of analyzing seismic signals for anomalies indicative of future anomalous events such as earthquakes.

An important primary step in signal processing is de-noising of the signal. Noise obscures the characterizing properties of a signal, making analysis difficult. A popular method for de-noising of a signal is wavelet based de-noising methods. Wavelet based methods are particularly useful for non-stationary signals such as EKGs, EEGs, and seismic signals. Other methods for de-noising of non-stationary signals include the Short Time Fourier Analysis (Gabor transform) and Spline and Kernel estimators among others. These alternative methods are not as well-suited to de-noising non-stationary signals as the wavelet transform. The Gabor transform is very inflexible since its precision is dependent on the window size of the transform. Spline and Kernel estimators often suppress the signal along with any noise. In contrast, Donoho and Johnstone's work on wavelet de-noising (1994; 1995; 1995; 1995) demonstrated the exceptional utility wavelet based de-noising methods have when applied to non-stationary signals.

Another common pre-processing step in signal analysis is identifying prominent peaks within the signal. Peak-finding or filtering, as this method is called, is very useful in anomaly detection. This technique has already proven itself to be valuable in identifying the major events of seismic signals (Mitchell et al. 1998). Peak-finding methods may be applied in various ways. More commonly, they are applied to a signal's frequency domain in order to find the limits for a thresholding algorithm (Gao et al. 2006). There are many general peak-finding algorithms such as nonlinear filtering, Gaussian second derivative filtering, and Kalman filtering, that are designed to be used in a broad range of applications. However, applying these general algorithms requires the selection of many free parameters (Scholkman et al. 2012). Because of this downside to general algorithms, we develop a simple yet robust statistically based peak-finding algorithm to suit our methodology.

The Fourier transform provides an overall view of a signal's component frequencies. While not as suitable as the wavelet transform for non-stationary signals, the Fourier transform's comprehensive picture of the frequencies can be useful for anomaly prediction. Applications of the Fourier transform include noise reduction in audio signals as well as less obvious uses such as efficient polynomial multiplication (Shatkay 1995). In our methodology, we use the discrete Fourier transform to give our neural network an alternative perspective of the input signal.

After a signal is de-noised its anomalies can then be found by means of a pattern recognition method. Artificial neural networks (ANN) provide just such a method making them a valuable tool in signal processing as well as an alternative to statistical methods of pattern classification. ANNs excel at recognizing patterns because of their ability to learn and make inferences about new inputs based on previous input patterns (Kasthurirangan 2010). Because of their ability to generalize from training sets, ANNs are already utilized for pattern detection in many fields including speech (Mitsuru et al. 1998), EEG signals (Subasi 2005), and financial data (Kim 2004). Feed-forward ANNs have a basic but flexible structure that allows them to be easily adapted for various applications. We employ the feed-forward network model in our technique to analyze perspectives of a signal and determine the existence of anomalies.

The rest of this paper is organized as follows. Section II provides background on wavelets, wavelet transform, wavelet thresholding of coefficients, Fourier transform, artificial neural networks, as well as the training algorithms for neural networks. Section III introduces the methodology used to obtain our results. The differences between the EKG anomaly detection experiment and both the single and multi-perspective cases of the seismic signal anomaly prediction experiments are also detailed in section III. Then Sections IV and V are the experimental results and conclusion respectively.

The pattern and relationships between species diversity and ecosystem functioning are the current areas of great ecological interest throughout the world. Species diversity incorporates two components (Stirling and Wilsey, 2001); evenness (how evenly abundance or biomass is distributed among species) and richness (number of species per unit area). High evenness can increase invasion resistance, below-ground productivity and reduce total extinction rates (Smith et al., 2004). The spatial variations in biodiversity generally include species diversity in relation to size of the area, relationship between local and regional species diversity and diversity along gradients across space, and environmental factors such as latitude, altitude, depth, isolation, moisture and productivity (Gaston, 2000). In addition, species richness of a taxon is not only sufficient to express diversity but the equitability is also a important factor because communities however vary in properties of the total importance of the species and share their functional contribution (Tilman, 2000).

2. BACKGROUND AND RELATED WORKS

2.1 WAVELETS

Wavelets have proven themselves to be an effective tool in signal analysis. Alfred Haar discovered the first wavelet in 1909, though it was not called this at the time (1910). Since Haar's time, wavelet theory has been advanced by many scientists including John Morlet (1984), Stephane Mallet (1987), Y. Meyer (1999), Ingrid Daubechies (1992), and others. The field of wavelet analysis continues to be heavily researched and has many applications to problems in signal processing.

A wavelet is similar to a waveform. However, unlike waveforms, wavelets have a finite duration. A wavelet is defined as a basis function, $W_{jk}(t)$, that produces the group of functions f(t) where:

$$f(t) = \sum_{j,k} b_{jk} w_{jk}(t) \quad (1)$$

Every wavelet, $W_{jk}(t)$, is generated by compressing a mother wavelet, *w(t)*, *j* times and translating it *k* times as shown in the formula:

$$w_{jk}(t) = w(2^j t - k) \quad (2)$$

2.2 Wavelet Transform

The wavelet transform is a powerful tool for analyzing a broad range of signals. Wavelet analysis has very useful properties such as unconditional basis. This means that for many signals the expansion coefficients will decrease very quickly. Additionally, the wavelet transform yields a more accurate local description of the signal than alternatives such as the Gabor transform and is more flexible as well. Beyond these useful properties, the wavelet transform is also easy to compute as it requires only the basic operations multiplication and addition (Burrus et al. 1998).

Now that we have defined the concept of a wavelet, we can define the discrete wavelet transform (DWT), which is used to decompose a signal into high and low frequency components also called the approximate (low frequency) and detail (high frequency) components. The DWT allows any square integrable function to be defined by the series:

$$c_j(k) = \sum_m h(m-2k) c_{j+1}(m) \quad (3a)$$

$$d_j(k) = \sum_m h_1(m-2k) c_{j+1}(m) \quad (3b)$$

where $c_j(k)$ denotes the approximate coefficient and $d_j(k)$ denotes the detail coefficient of f(t).

Wavelet Thresholding of Coefficients (De-noising)

After a signal has been decomposed by applying the DWT a finite number of times, it can be de-noised. A popular method for de-noising is thresholding of wavelet coefficients. This method also known as wavelet shrinkage, involves comparing the coefficients to an arbitrary threshold value and reducing the coefficients' value if it is within the absolute value of the threshold. There are two common ways of reducing those coefficients that fall within the threshold value: hard thresholding and soft thresholding. In hard thresholding the coefficients within the threshold are reduced to zero, effectively cutting them from the signal. The alternative, soft thresholding subtracts the threshold value from the smaller coefficients rather than setting them to zero. Soft thresholding reduces the impact that values just beyond the threshold range have on the signal.

Once wavelet thresholding is used to reduce noise to a satisfactory level the signal may be reconstructed via the inverse discrete wavelet transform (IDWT). The IDWT is defined as:

$$f(t) = \sum_k c_j(k) 2^{j/2} \phi(2^j t - k) + \sum_k d_j(k) 2^{j/2} \psi(2^j t - k) \quad (4)$$

2.3 Fourier Transform

The Fourier transform, named for John Baptiste Joseph Fourier, transforms a signal from the time domain to the frequency domain. The discrete Fourier transform (DFT) is a variation of the Fourier transform that is used with sampled waveforms. While the Fourier transform is not as suited to non-stationary signals as the wavelet transform because it does not preserve temporal information, it provides an excellent view of all of the frequencies that compose a signal. Essentially, the Fourier transform approximates a signal with a series of sine and cosine waveforms. This yields a set of coefficients that is called the Fourier transform. These may then be graphed in the frequency domain to show all the component frequency bands that compose the original signal.

From Euler's formula we have:

$$e^{ix} = \cos(x) + i\sin(x) \quad (5)$$

Now substituting e^{ix} we may define the discrete Fourier transform (DFT). The DFT transforms a set of N complex numbers into an N-periodic sequence of complex numbers by the formula:

$$X_k = \sum_{n=0}^{N-1} X_n * e^{\frac{-i2k\pi n}{N}} \quad (6)$$

where X_k is the kth complex number in the N-periodic sequence.

2.4 ARTIFICIAL NEURAL NETWORK

Artificial neural networks are based on the work of McCullogh-Pitts and their artificial neuron model (1943). Early ANNs, also called perceptrons, had only a single layer and were therefore incapable of accurately classifying many input patterns or modeling functions such as XOR (Minsky, Papert 1972). Multi-layered networks such as feed-forward networks were able to overcome the shortcomings of the earlier perceptrons and are widely used for pattern recognition tasks.

An artificial neural network is defined by the sorted triple (N, V, w). Where N is defined as the set of neurons, V is defined as the set of connections (i, j) such that i,j ∈ N. And w is a function that defines the weights for each connection (i, j). That is w: V →N. The weight of connection (i, j) is denoted as W_{ij}.

FEED-FORWARD NEURAL NETWORK

A feed-forward neural network is an artificial neural network with an acyclic structure. That is it has no loops, all paths through the network end at the output layer and only visit one neuron per layer. Layered feed-forward neural networks are suitable for many applications. The general structure of a feed-forward network has three types of layers. The initial layer is called the input layer and receives the input pattern. The next layer or layers are called the hidden layer(s). The terminal layer is called the output layer. Feed-forward networks are commonly trained using the backwards propagation of errors learning algorithm.

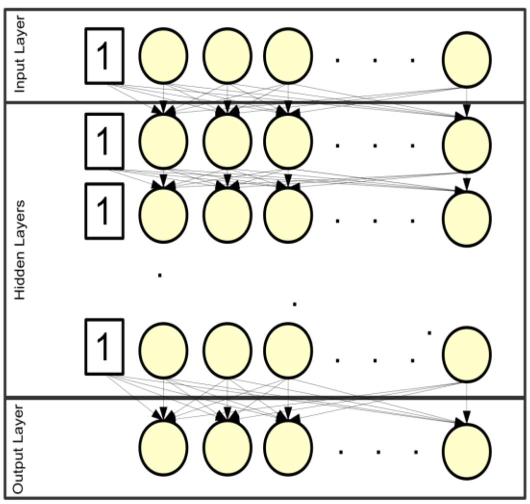

Figure 1: General Structure of a Feed-Forward Neural Network

2.5 RESILIENT BACKWARDS PROPAGATION OF ERRORS

Backwards propagation of errors (backprop) is a supervised learning algorithm for training neural networks. However, the backprop algorithm is prone to converge on local minima rather than the absolute minimum. To solve this issue the resilient backwards propagation of errors algorithm was proposed. The resilient backwards propagtion of errors (Rprop) algorithm is another supervised learning algorithm. Rprop solves the problem that the backprop algorithm faces by observing the sign rather than the magnitude of the partial derivative of the error function. The general algorithm for Rprop is described in the next few sentences and detailed in figure 2. The current connection weight is updated by evaluating its partial derivative at each iteration. The result of evaluating the partial derivative with respect to the current weight is then multiplied by the result of the previous iteration. The sign of the product may be negative, positive, or zero, and each case is handled separately. If the product is negative the change value is multiplied by the learning coefficient. The new change value is then added to the direction that provides the maximum decrease in error. If the product is positive the change value is again updated by multiplication with the learning coefficient. If the product is zero, we add the product to the weight with its sign adjusted to the direction that yields the greatest decrease in error. When using Rprop to train a neural network, the network must first be evaluated for error. If the error is greater than the target error, all weights in the network are updated by the technique detailed above. This process is repeated until the network error is less than the target error.

3. TOOLS METHODS AND MATERIALS

3.1 THE DATA SET AND TOOLS USED

Data for seismic signals was collected from an Incorporated Research Institutions for Seismology (IRIS) database. IRIS provided the largest selection of data among seismic signal databases. We accessed the seismic signal data using the JWEED utility. The data was provided as plain text, but converted to CSV before processing.

The EKG samples were collected from the PTB Diagnostic EKG Database that can be found at Physionet.org. Each signal file contained 15 leads recorded from various locations on the human subject. All 15 leads were analyzed by the neural network after de-noising.

To implement our feed-forward neural network we used the ENCOG library for Java. ENCOG greatly eases the programming involved to build and train a neural network. Wavelet de-noising was performed using the Matlab utilities wthcoef(), wavedec() and waverec(). The wavedec() function is used to decompose the input signal. Then the wthcoef() function thresholds the coefficients of a particular wavelet decomposition level.

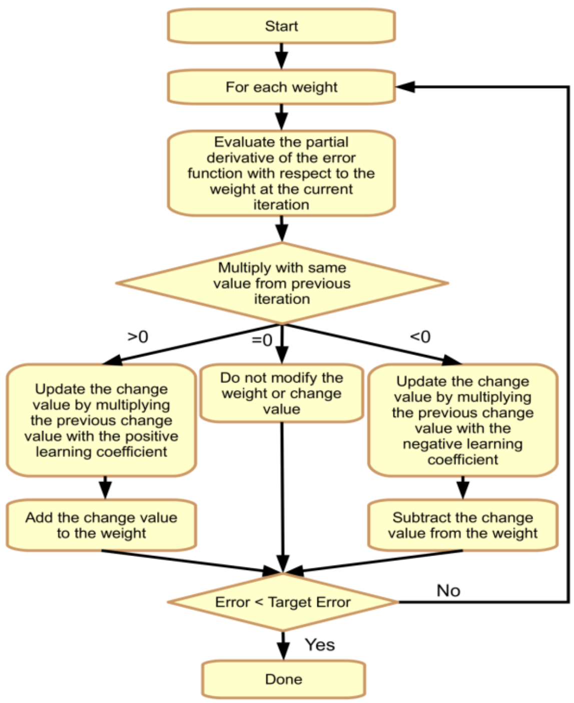

Figure 2: Flowchart of Rprop Algorithm

After thresholding the signal is reconstructed with waverec(). Fourier transforms were carried out in Java using the JWave signal processing library.

3.2 METHODS

The general outline of our method is as follows. First we collect test signal data and produce one or more perspectives of each signal in the data set. A perspective is defined here as a view representing some aspect of the original signal. Each perspective must differ from the other perspectives in its abstract meaning. For example, a wavelet de-noised perspective represents the time-domain of the signal and a Fourier transform perspective represents the frequency-domain. Once generated, we input these perspectives into a feed-forward neural network. The network is trained on data that is pre-processed into perspectives. The network is then tested using the same perspectives applied to the corresponding test signals.

Below we present three implementations of our methodology. The first is application of our technique to anomaly detection in EKG signals using one perspective. The other two are applications of our technique to anomaly prediction in seismic signals. The first of these uses a single perspective for analysis, the other expands on the methodology of the first and uses two perspectives. Because the third application is an expansion of the second we will describe only the third and note those aspects that apply to both and those that apply only to the multi-perspective method to save from being redundant.

Before we discuss the specific techniques we used to generate the perspectives of the input signals, it is important to note that there are many possible alternative perspectives that could also be used in our method. For example, using multiple sensors recording at the same time to generate a multi-dimensional seismogram would allow us to present each dimension of the seismogram to the network as a different perspective. Other possible perspectives could be generated by using other transforms or even just feeding the raw time-magnitude domain of the signal as a perspective. Our choice in perspectives differed, though only slightly, between applications to EKG and seismic signals. For EKGs we used wavelet de-noising to generate the only perspective of the signal. We felt that it was not necessary to generate more perspectives or to use a peak-filter on the de-noised signal because the EKG signal has a clear pattern with distinct peaks (refer to figure 3). We chose two perspectives for anomaly prediction in seismic signals. The first perspective was generated by first using wavelet de-noising followed by application of a peak-finding algorithm. The second perspective was generated by application of the discrete Fourier transform to the original signal.

In the next two sub-sections we expound the details of perspective generation for use with EKGs and seismic signals. The first sub-section describes how the single perspective of each EKG sample is created and the second sub-section details both of the perspectives used for seismic signal anomaly prediction. After we detail the different perspective generation techniques we describe the training and testing of the feed-forward neural network used to analyze the perspectives of these signals. Since the differences in the design and training of the network for EKG and seismic signals differ only subtly we will not create separate sub-sections for each.

3.2.1 EKG PERSPECTIVE GENERATION METHODOLOGY

As mentioned above the EKG data was collected from the PTB Diagnostic EKG Database. Each signal sample contained signals recorded from 15 different leads. A lead in this context is a sensor recording electrical activity from some location on the human body. The signals are then de-noised via the discrete wavelet transform and wavelet thresholding techniques. We chose the Coiflet 5 wavelet as the mother wavelet for the discrete wavelet transform. The Coiflet 5 wavelet was chosen because it does a good job of characterizing the EKG signal. To de-noise the signal we apply the discrete wavelet transform using Coiflet 5 as the mother wavelet. Once decomposed the rigsure thresholding algorithm is used for coefficient thresholding. Rigsure thresholding was chosen because it has been shown to provide good results when paired with the Coiflet 5 wavelet (Rami 2012). Then the signal is reconstructed with the inverse discrete wavelet transform.

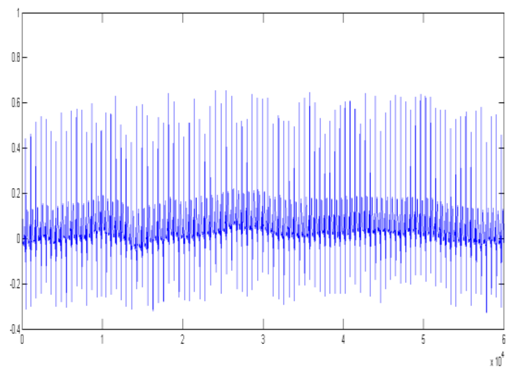

Figure 3: Raw EKG Signal

3.2.2 SEISMIC SIGNAL PERSPECTIVE GENERATION METHODOLOGY

The first perspective used in our multi-perspective seismic signal analysis was created by wavelet de-noising followed by peak-filtering. This perspective is common to both seismic signal analysis experiments. The wavelet de-noising method used here is similar to that used on EKG signals. The DWT was applied for 10 levels of decomposition followed by wavelet thresholding of coefficients for 8 levels. The de-noised signal was then reconstructed with the IDWT. The mother wavelet used for decomposition was the Haar wavelet. We chose the Haar wavelet because it provides a more discrete set of values for the input layer of the neural network in comparison to smoother wavelets such as the Daubechies-4 wavelet.

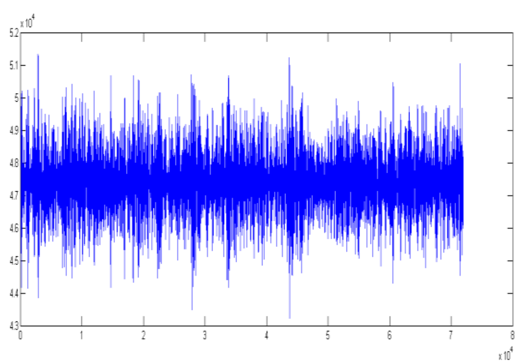

Figure 4: Raw Seismic Signal

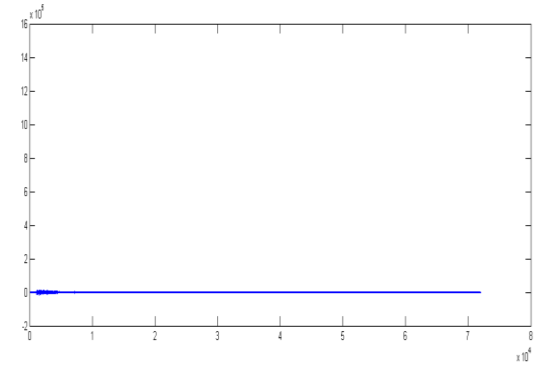

Figure 5: Haar Wavelet Decomposed Seismic Signal

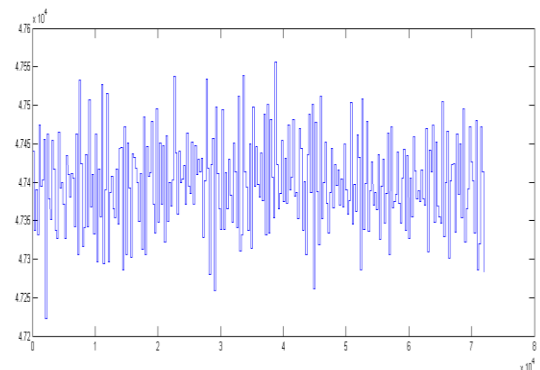

Figure 6: Reconstructed Seismic Signal after De-noising

Unlike EKG signals, seismic signals do not have a clear underlying pattern. This renders anomaly detection difficult. To alleviate this to some extent we apply a peak-filter. There are many peak-filtering algorithms available; however, we chose to implement our own statistically based peak-finding algorithm. Our peak-filtering method uses a threshold interval to separate potentially anomalous peaks from normal low frequency events. The threshold is the interval one local standard deviation above and below a local average for each point in the signal sample. The peak-filtering algorithm is given by the function pk(i):

$$pk(i) = \begin{cases} s(i) : |s(i)| > avg(s,w,i) + lstd(i) \\ 0 : otherwise \end{cases} \quad (7)$$

where

$$lstd(i) = \sqrt{\frac{1}{w} \sum_{k=1-\frac{w}{2}}^{i+\frac{w}{2}} ((s(k) - avg(s,w,i))^2)}$$

and

$$\text{avg}(s,w,i) = \frac{1}{w} \sum_{l=i-\frac{w}{2}}^{i+\frac{w}{2}} s(l)$$

The result of applying the peak-filter is a set of potentially anomalous peaks. The computational complexity of our peak-filter is O(n) because the local average needs to be computed just once per i.

The second perspective was generated by a Fourier transform of the original signal. The Fourier transform transforms a signal from the time-domain to the frequency-domain. Because our signal is sampled we used the DFT. The result of applying the DFT is a holistic view of the frequencies in the signal. This perspective provides the neural network with more data, potentially allowing for more accurate prediction of anomalies.

.

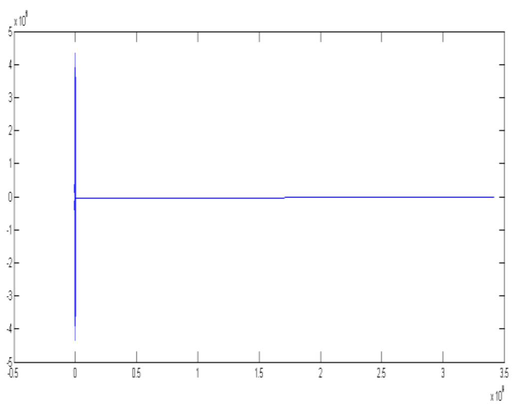

Figure 7: Result of DFT on Seismic Signal

3.2.3 NEURAL NETWORK STRUCTURE, TRAINING, AND TESTING

For all three of our experiments we used the feed-forward neural network design. The one perspective EKG and seismic signal experiments used identical neural network structures. The input layer for each contains 1024 nodes, the hidden layer contains 512 nodes, and the output layer contains only a single node. Because there was more data to input and evaluate for the multi-perspective seismic signal experiment, 2048 nodes were needed in the input layer. The first 1024 were used to hold the wavelet de-noised perspective of the signal. The second 1024 nodes contain the Fourier transform perspective. The hidden layer is also enlarged to handle the greater volume of data, it contains 1024 nodes. The output layer, however, is still a single node. The number of nodes needed for each layer was determined experimentally.

After determining the structure of the neural network, we trained the network. In general the training signals are serialized into the input layer in the same order that a test signal is to be processed. For example, in the multi-perspective case of the seismic signal experiment the training signal perspectives will be presented to the network in the same order. That is, the first 1024 nodes must always contain a wavelet de-noised signal perspective and the second 1024 nodes must always contain a Fourier transform of the original signal. The network is then trained until it converges. Ideally, the error of the network converges to zero but this is not always the case. The network trained on EKG signals converged to .01% while both seismic signal networks converged to roughly 24%. Down sampling may be used to fit longer signals into the network. We used down-sampling for the EKG analysis.

After training, the neural network is then fed new inputs known as test signals. The neural network then processes the test signals and outputs a value between 0 and 1. If the value is greater than or equal to 0.5 we interpret it as a prediction of an anomaly in the future of the signal (seismic signal experiments) or detection of one or more significant anomalies in the signal (EKG analysis). An output value of less than 0.5 indicates that there were no significant anomalies in the signal and that it is unlikely there will be anomalies in the future of the signal.

4. EXPERIMENTS AND RESULTS

4.1 EKG ANOMALY DETECTION RESULTS

The feed-forward neural network used in this experiment was trained on two sets of 60 signals. One set contained healthy signals and the other set had anomalous signals. The network was then tested on two sets of 30 signals. Again, one set contained anomalous signals, and the other contained healthy signals. Each signal was composed of 15 leads. The table below shows the network accuracy for each of the 15 leads.

Table: EKG Anomaly Detection Results by Lead

Network Accuracy (%)	
Lead i	65
Lead ii	45
Lead iii	65
Lead avr	52.5
Lead avl	70
Lead avf	57.5
Lead v1	50
Lead v2	40
Lead v3	55
Lead v4	55
Lead v5	67.5
Lead v6	60
Lead vx	52.5
Lead vy	55
Lead vz	55

As one can see from the table above, the 'avl' lead provided the best results at 70%. With more data to train the network with, the results are likely to improve.

4.2 Single Perspective Seismic Signal Anomaly Prediction

The results from the single perspective version of the seismic signal anomaly prediction ranged from 60% to 83%. However, these results varied greatly. Due to the lack of consistency it is believed by the authors that these results may not hold for larger test sets or test sets taken outside of the time range of the training set.

4.3 Multi-perspective Seismic Signal Anomaly Prediction

In the multi-perspective case, the results were within a much smaller range: 54.7% to 56%. These results were much more consistent than the single perspective results. While not as accurate as the single perspective case, their consistency shows that the network was able to identify some patterns preceding anomalous events. With improvements or variations in the perspectives used to represent the signal, it is possible to increase the accuracy. Using this multi-perspective technique these percentages hold for data collected years after the training data.

5. CONCLUSION

In this paper, we presented a method for predicting possible anomalies in the future of a given signal. Before we processed the signal we made one or more perspectives of the signal that emphasized the signal's base properties. The methods we used to generate perspectives included wavelet de-noising with peak-filtering and Fourier transform. After we created our perspectives of the original signals we used them as input to a feed-forward neural network. The network was trained with the Rprop algorithm on two sets of signals: one with anomalies and one without. The network then outputs the likelihood of an anomaly occurring in the future of the signal. Our results showed that predicting and detecting anomalies in signals such as EKGs, EEGs, and seismic signals by application of artificial neural networks is possible. We achieved results as high as 70% accuracy with EKG signals and as high as 83% with seismic signals, though with inconsistent results.

Acknowledgements:

This research was funded by the Deanship of Scientific Research DSR, King Abdulaziz University in Jeddah under grant no. (18/829/1433). The authors, therefore, acknowledge with thanks DSR technical and financial support.

Corresponding Author:
Dr. Abdullah Ali Alshehri
Department of Electrical Engineering
Kingabdulaziz University-Rabigh, Saudi Arabia
E-mail: ashehri@kau.edu.sa

References

1. Andreas, K., Ohrnberger, M., Scherbaum. F. (2010) Unsupervised Pattern Recognition in Continuous Seismic Wavefield Records Using Self-Organizing Maps. Geophysical Journal International 182(3): 619-630. Print.
2. Boucheham, B. (2011) Abnormality Detection in Electrocardiograms by Time Series Alignment. Communications in Information Science and Management Engineering.1(3):6-10.
3. Burrus, C. S., Gopinath, R. A., Guo, H. (1998) Introduction to Wavelets and Wavelet Transforms A Primer. Prentice Hall. New Jersey. Print.
4. Cheboli, D. (2010) "Anomaly Detection of Time Series." May 2010.
5. Chuah, M. C., Fu, F. (2007) EKG Anomaly Detection via Time Series Analysis.
6. Daubechies I. (1992) Ten Lectures on Wavelets. Society for Industrial and Applied Mathematics.
7. Donoho, D. L., Johnstone, I. M. (1995) Adapting to unknown smoothness via wavelet shrinkage Journal of the American Statistical Assoc., 90(432):1200–224.
8. Donoho, D. L. (1995) De-noising by soft-thresholding IEEE Trans. Inform. Theory, vol. 41, pp. 613–627.
9. Donoho, D. L., Johnstone, I. M. (1994) Ideal spatial adaptation via wavelet shrinkage. Biometrika, vol. 81, pp. 425–455.
10. Donoho, D. L. , Johnstone, I. M. (1995) Wavelet shrinkage: Asymptopia? J. R. Stat. Soc. B, ser. B, 57(2): 301–369.
11. Fensli, R., Gunnarson, E., Gunderson, T. (2005) A Wearable ECG-recording System for Continuous Arrhythmia Monitoring in a Wireless Tele-Home-Care Situation. IEEE International Symposium on Computer-Based Medical Systems.
12. Gao, Y., Guo, Y., Chi, Y. L., Qin, S. R. (2006) Order Tracking Based on Robust Peak Search Instantaneous Frequency Estimation. J.Phys. : Conf. Ser. 48 479.
13. Goupillaud, P. Grossman, A. Morlet, J. (1984) Cyclo-octave and related transforms in seismic signal analysis. SIAM J. Math. Anal. 15:723-736.

14. Haar, A. (1910), "Zur Theorie der orthogonalen Funktionensysteme." Mathematische Annalen 69 (3): 331–371.
15. Kasthurirangan, G. (2010) Effect of Training Algorithms on Neural Networks Aided Pavement Diagnosis. International Journal of Engineering, Science and Technology 2(2):83-92. Print.
16. Kim, T. (2004) Artificial Neural Networks for Non-stationary Time Series. Neurocomputing 61: 439-47. Print.
17. McCullough, W., Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, 7:115 - 133.
18. Meyer, Y. (1987) Orthonormal wavelets. In J.M. Combes A. Grossmann Tchamitchian Ph editors. Wavelets, Time-Frequency Methods and Phase Space.21-37. Springer-Verlag. Berlin. 1989. Proceedings of International Colloquium on Wavelets and Applications. Marseille, France. December 1987.
19. Mikheld, A., Daqrouq, K. (2008) EKG Signal Denoising By Wavelet Transform Thresholding." American Journal of Applied Sciences 5(3): n. pag. Print.
20. Minsky, M., Papert. S. (1972) Perceptrons: An Introduction to Computational Geometry, The MIT Press, Cambridge MA.
21. Mitchell, W., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., Trujillo, J. (1998) A Comparison of Select Trigger Algorithms for Automated Global Seismic Phase and Event Detection. Bulletin of the Seismological Society of America 88(1): 95-106. Print.
22. Mitsuru, K., Matsuoka, K., Ohnishi, N. (1998) A Method of Blind Separation for Convolved Non-stationary Signals." Neurocomputing 22(1-3):157-71. Print.
23. Neep, H., Zhu, C. J., Chung, T. A. Alex. S. (1997) Classification of EEG signals using the wavelet transform. Signal Processing. 59(1):61-72.
24. Rami, C. (2012) Signal Denoising Using Wavelets.

6/2/2014