2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

Trend Detection and Data Mining via Wavelet and Hilbert-Huang Transforms[‡]

Murat Yasar
Techno-Sciences, Inc.
Beltsville, MD 20705 USA
Email: yasarm@technosci.com

Asok Ray
Pennsylvania State University
University Park, PA 16802 USA
Email: axr2@psu.edu

***Abstract*— This paper presents the formulation and evaluation of effective algorithms of reliable data analysis for real-time monitoring of incipient faults and anomalies, data fusion and event classification. The objective is to alleviate the shortcomings of the existing techniques for data mining by taking advantage of nonlinear filtering to handle non-Gaussian and non-stationary multiplicative noise and uncertainties. New concepts have been developed toward characterization of the data features and behavior interpretation of the underlying processes to evaluate their performance. In particular, the techniques of wavelet transform, Hilbert-Huang transform, and symbolic encoding are investigated to explore their effectiveness and relative simplicity to interpret and implement data mining tasks.**

***Index Terms*—Wavelet transform, Hilbert transform, Data compression, Signal analysis, Fault detection**

I. Introduction

DATA analysis is undeniably the most pertinent part of science and technology related to health monitoring, since data is generated and processed in all systems one way or another. Regardless of whether the source of data is physical measurements from a sensor signal or numerical modeling, the goal of data analysis is to reveal the underlying processes. However, a vast majority of the raw data sets are not well suited for analysis and may encounter several inherent difficulties such as (i) short data span, (ii) non-stationary behavior of data, (iii) underlying nonlinear process dynamics. Some of these problems can be simplified to a certain extent or approximated by simpler solutions. For example, the linearization around an equilibrium point usually works well for mildly nonlinear data sets.

For a general solution to be tractable over a wide range, data analysis may require innovative techniques and computationally efficient algorithms. Moreover, data analysis should address the problems of detection of trends, anomalies and faults to infer mathematically and computationally tractable models from the data sets for information fusion from multiple sources and simultaneous classification of instances and features. This makes the algorithms challenging and significant since the outcome of such analysis methods can be applicable to many areas from simple time-frequency decomposition of signals to advanced health monitoring and threat identification.

Data analysis has a long history starting from Fourier transform in late 1700s and early 1800s [1]. However, traditional data analysis methods are largely based on linear and stationary assumptions. For example, Fourier spectral analysis provides a general solution for examining the global energy-frequency distributions. Mainly because of its simplicity, Fourier analysis dominated the data analysis efforts and has been applied to various kinds of data. Although the Fourier analysis is valid under extremely general conditions [1], the data sets still must be periodic or statistically stationary. The stationary requirement is not particular to the Fourier spectral analysis and is required for most of the available data analysis techniques.

In recent years, new methods and algorithms have emerged for analysis of non-stationary and nonlinear data. For example, the Wigner-Ville distribution [2] was designed for linear but non-stationary data. Furthermore, various nonlinear time-series-analysis methods [3][4] were designed for nonlinear but stationary and deterministic systems. Nonetheless, most natural and human engineered complex systems produce data, which are most likely to be both nonlinear and non-stationary.

A necessary condition to represent nonlinear and non-stationary data is to have an adaptive basis. For non-stationary and nonlinear data, where adaptation is absolutely necessary, no available methods could be found until recently [5]. In this paper, we address the problem of analyzing nonlinear and non-stationary data to provide the information about incipient faults, anomalies, trends, and data classifications in complex systems.

The primary goal of the work reported in this paper is to develop effective and robust algorithms for real-time monitoring of incipient faults and anomalies, fusion of data and classification of events. To achieve these goals, it is necessary to characterize the data features and interpret the characteristic of underlying processes to evaluate the system performance. Thus, it is possible to meet the increasing demands for data mining of sensor signals and analytical observables for health monitoring applications. Outputs from these analytical tools will be useful for scheduling both routine and preventive maintenance, threat and risk

[‡]This work was supported in part by NASA Glenn Research Center under Contract No. NNC07QA08P and by U.S. Army Research Laboratory and U.S. Army Research Office under Grant No. W911NF-07-1-0376.

assessment and reconfiguring the resilient and fault tolerant control.

This paper is organized in five sections including the present one. Section 2 introduces the key data analysis methods for algorithm development, and the trend detection problem for incipient faults and anomalies is formulated in Section 3. Section 4 presents an example application of the developed techniques. The paper is summarized and concluded in Section 5.

II. Data Analysis Methods

A key element in data analysis for anomaly and fault detection is to transform the sensor data sets into formats that enable efficient, effective detection and classification of anomaly events and trends of interest. The sensor data are characterized by signals that have features that can be intermittent in both time and space. To maximize the information content derived from the data, the raw signals can be described in multiple resolutions in time, frequency, and space.

The proposed method aims to alleviate the shortcomings of the existing techniques for data mining. Emphasis is laid on nonlinear filtering to deal with (possibly) non-Gaussian and non-stationary multiplicative noise and uncertainties. To this end, several advanced methods that have been developed in recent years are utilized to form the basic functional components of the analysis. Among the few possibilities of nonlinear non-stationary data analysis, wavelet transform (WT) [6][7], Hilbert-Huang transform (HHT) [5][8] and symbolic encoding (SE) [9][10] techniques are chosen due to their effectiveness, their relative simplicity for interpretation and implementation.

A. Wavelet Analysis of Time Series Data for Time-Frequency Localization

Time series data of signals are often preprocessed to characterize the signal properties and extract pertinent information. It is well known that for the stationary signals with known time periods, Fourier analysis is sufficient to represent the frequency domain characteristics of the signal. However, Fourier analysis may not be appropriate if the signal has non-stationary characteristics such as drifts, abrupt asynchronous changes, and frequency trends. Wavelet analysis, on the other hand, uses adaptive windows to overcome these difficulties. Long windows retrieve low-frequency information, whereas short windows are used for high-frequency information [7]. The ability to perform flexible localized analysis is one of the striking features of wavelet transform.

The wavelet approach is essentially an adjustable windowed Fourier spectral analysis. Defining the wavelet function as:

$$\psi_s(\tau) = |s|^{-p}\,\psi_1\left(\frac{\tau}{s}\right) \quad \forall p \in [0, \infty)$$

where ψ_1 is called the 'Mother wavelet' and p is usually chosen as 1/2, the wavelet transform (WT) can be defined as

$$W(s,t) = |s|^{-p} \int_{-\infty}^{\infty} f(\tau)\psi^*\left(\frac{\tau - t}{s}\right) d\tau .$$

Therefore, wavelets provide scale-independence for analysis and synthesis of signals, which means continuous wavelet transform (WT) provides resolutions for all the scales. This very appealing feature of wavelets and other relevant reasons such as WT analysis being linear and having an analytic form for the result make wavelets widely accepted among researchers and engineers. Indeed, WT is very useful in analyzing data with gradual frequency changes. Most of its applications have been in signal denoising and image compression, as well as heath monitoring [11].

Among other signal processing choices, *wavelet transforms* offer an attractive means for developing multi-resolution representations of signals suitable for hierarchical classification algorithms. They have been used in this way in speech and image processing and in radar data processing, among many other applications. The success of wavelet transforms in these and other areas can be attributed to the inherent ability of wavelet representations to reveal the superposition of different structures occurring in these signals on *different time scales at different times* (or on *different spatial scales at different locations*).

Wavelet representations efficiently separate and sort the constituent structures of a complex signal. Software tools for construction of wavelet transforms have recently become available, offering the user a collection of standard libraries of waveforms that can be chosen to fit specific classes of signals. The rapid development of wavelet theory and computational methods provides a versatile and powerful set of tools for the analysis and manipulation of signals. However, the leakage due to the compact support of the mother wavelet function makes the quantitative definition of the energy-frequency-time distribution difficult. In spite of these problems, wavelet analysis is still among the best available non-stationary data analysis method.

B. Hilbert-Huang Transform of Time Series Data for Time-Frequency-Energy Analysis

While many other techniques may fail in analyzing non-stationary and nonlinear systems, Hilbert-Huang transform (HHT) meet the challenges of time-frequency-energy representation of the data. Consisting of empirical mode decomposition and Hilbert spectral analysis, HHT is suitable for computing instantaneous frequency and energy of the signal in time domain. Therefore, the ability to perform flexible localized analysis in the three independent variables, namely, time, frequency and energy, is a striking feature of HHT [5]. HHT has good computation efficiency and does not involve the concepts of time and frequency resolution. These features alleviate noise and signal distortion problems and make HHT as an excellent candidate for data mining in health management applications [12].

The Hilbert transform is performed on real-valued time-series data and produces an analytical signal for which the instantaneous frequency can be defined. The Hilbert transform of a signal $x(t)$ is defined as its convolution with $1/t$ as

$$H(t) = \frac{P}{\pi}\int_{-\infty}^{\infty}\frac{x(\tau)}{t-\tau}d\tau$$

where P is the Cauchy principal value. Combining $x(t)$ and $H(t)$ as $Z(t) = x(t) + iH(t) = A(t)e^{i\phi(t)}$ yields the analytical signal $Z(t)$. One important property of the Hilbert transform is that if the signal $x(t)$ is monocomponent, then the time derivative of instantaneous phase $\phi(t)$ will be the physical meaning of instantaneous frequency

$$\omega(t) = \frac{d\phi(t)}{dt}.$$

The aim of HHT is to decompose the signal into monocomponent pieces by using an empirical mode decomposition algorithm. An intrinsic mode function (IMF) is a function that satisfies almost monocomponent requirements and it represents one and only one oscillation mode embedded in the data. Therefore, by applying the Hilbert transform to each IMF component, which is obtained from mode decomposition, it is possible to estimate instantaneous frequencies in the data. HHT also enables us to represent the amplitude and the instantaneous frequency as functions of time in a three-dimensional plot, in which the energy density of the signal can be contoured on the frequency-time plane to produce the Hilbert energy spectrum [8].

Among other signal processing choices, *HHT* offers an attractive means for developing time-frequency-energy representations of signals suitable for anomaly, fault and trend detection algorithms. They have been used in this way in signal processing and fault identification, among other applications. The success of HHT in these and other areas can be attributed to the inherent ability of localized representations to reveal the superposition of different structures occurring in these signals on *different frequencies at different time instances* (or on *different energy levels at different locations*).

C. Symbolic Dynamics and Encoding

The concept of symbolic dynamics and its merits for encoding nonlinear system dynamics from observed data sequence has utmost importance from data mining point of view. It not only helps us to describe the process dynamics in a lower dimensional space but also overcome the difficulties due to uncertainties and noise. Let a continuously varying physical process be modeled as

$$\frac{dx(t)}{dt} = f\left(x(t), \theta(t_s)\right) \quad x(0) = x_0$$

where $t \in [0, \infty)$ denotes the time; $x \in \mathbb{R}^n$ is the state vector in the phase space; and $\theta \in \mathbb{R}^m$ is the parameter vector varying in time scale t_s. Sole usage of the model may not always be feasible due to unknown parametric and non-parametric uncertainties and noise. A convenient way of learning the dynamical behavior is to rely on the additional information provided by sensory data [13]. A tool for behavior description of nonlinear dynamical systems is based on the concept of symbolic encoding for transitions from smooth dynamics to a discrete symbolic description [9].

The stationary motion of the process dynamics are confined in compact (i.e., closed and bounded) region $\Omega \in \mathbb{R}^n$ if the stability conditions are satisfied. The phase space of the dynamical system is partitioned into a finite number of cells, so as to obtain a coordinate grid of the space [14]. Encoding of Ω is accomplished by introducing a mutually exclusive and exhaustive partitioning $B = \{b_1, \ldots, b_m\}$ consisting of m cells. The process dynamics describes an orbit in Ω as $\{x_0, \ldots, x_k\}$ $x_i(t_i) \in \Omega$, which passes through or touches the cells of the partition B. Let us denote the cell visited by the trajectory $\{x_0, \ldots, x_k\}$ at a time instant t_i as a random variable S that takes a symbol value $\sigma \in \Sigma$. The set Σ of symbols that label the partition elements is called the symbol alphabet. Each state $x_i \in \Omega$ generates a symbol defined by a mapping from the phase space into the symbol space as $x_i \to \sigma_j$. This mapping is called Symbolic Encoding (SE) as it attributes a legal (i.e., physically admissible) symbol sequence $\{\sigma_0, \ldots, \sigma_k\}$ to the system dynamics starting $\{x_0, \ldots, x_k\}$ from an initial state x_0. (Note: A symbol alphabet Σ is called a generating partition of the phase space Ω if every legal symbol sequence uniquely determines a specific initial condition x_0, i.e., every symbolic orbit uniquely identifies one continuous space orbit [10].)

This creates a spatial and temporal discretization of the system dynamics defined by the trajectories. Symbolic encoding can be viewed as coarse graining of the phase space, which is subjected to (possible) loss of information resulting from granular imprecision of partitioning boxes. However, the essential robust features (e.g., periodicity and chaotic behavior of an orbit) are preserved in the symbol sequences through an appropriate partitioning of the phase space [9] and usual difficulties that arise from measurement noise and errors, and sensitivity to initial conditions to analyze the process data is mostly evaded due to granular partitioning.

III. PROBLEM FORMULATION FOR TREND DETECTION

Physical faults that gradually evolve over a prolonged period of operation influence the performance of components in electrical and mechanical system. Therefore from health monitoring point of view, it is imperative to detect the trends of incipient faults and anomalies at an early stage and predict the future effects and consequences. This provides a prognostic capability as well as diagnostics in advanced health management systems. Similarly, component degradation occurs on a slow time scale with respect to the system dynamics, which constitutes the fast time scale, and is observed in very small magnitude changes in system behavior before failure. The precursors to this type of slowly approaching failure have to be monitored in order to predict the remaining life of a component with the best accuracy.

Often a small part of the energy of the signals to be processed can be attributed to fault conditions and there is a substantial level of noise and unmodeled dynamics associated with the observed process variables [15]. Our technique for trend detection of anomalies has two features (i) isolation of signal features that have a high degree of correlation with anomalies; and (ii) application of a decision-algorithm to classify the set of signal features to identify nominal and anomalous behavior. The basic technique include preprocessing of time series data using the wavelet analysis [7], which is well suited for time-frequency analysis of non-stationary signals, for attenuation of noise and spurious disturbances in the raw time series data without any significant loss of pertinent information. The wavelet-transformed data is partitioned using the maximum entropy principle [6] to generate the symbol sequences. Then, the technique builds probabilistic automata [16] from temporal and spatial data series generated by a simple nonlinear spatial system [10]. The goal is to let the process describe itself, without appealing to "a priori" assumptions about the process's structure. Computational mechanics shows, from either empirical data or from a probabilistic description of behavior, how to infer a model of the hidden process that generated the observed behavior. This representation captures the patterns and regularities in the observations in a way that reflects the causal structure of the process [15].

The procedure for trend detection of incipient faults and anomalies using the concept of wavelet transform-based symbolic encoding (WT-SE) requires the following steps:

1. Time series data acquisition on the fast scale from sensors and/or analytical measurements (i.e., outputs of physics-based or empirical model). Data sets are collected either at different slow time epochs or as single set over the whole span of the data.
2. Transformation of the data from time domain to the wavelet domain for attenuation of noise and spurious disturbances in the raw time series.
3. Transformation of wavelet data from the continuous domain to the symbolic domain by partitioning into finitely many discrete blocks to generate symbol sequences at different slow time epochs.
4. Construction of finite state automaton from the generated symbol sequence and computation of the instantaneous statistical pattern vectors, whose elements are obtained online from frequency counting of the automaton states.
5. Identification of changes in statistical behavior based on the derived information as the evolution of instantaneous probability vectors relative to the nominal condition.

This procedure specifically accommodates localization of the time-frequency characteristics of the data in a lower dimension than the process space, leading to data compression without losing pertinent information.

For trend detection and classification, Hilbert-Huang transform (HHT) of the data can also be utilized since HHT is best suited to get a broad picture of the data in time-frequency-energy space. This improves the solution to the classification problem and detection of the trends in the data as the information relevant to the data is enhanced by HHT. A natural extension of trend detection will be to infer the model and determine the model parameters from the data.

Based on above discussion, the second algorithm makes use of the Hilbert-Huang transform of the signal which represents nonlinear, non-stationary process dynamics. The changes in the instantaneous frequency and the localized energy of the data may be associated with parametric and non-parametric anomalies in the system. The HHT procedure differs from the previous one as follows:

1. Time series data acquisition on the fast scale from sensors and/or analytical measurements (i.e., outputs of physics-based or empirical model). Data sets are collected either at different slow time epochs or as single set over the whole span of the data.
2. Transformation of the data from time domain to the Hilbert domain for information enhancement and localization of the raw time series.
3. Identification of changes in signal behavior based on the derived information as the evolution of instantaneous frequency and amplitude relative to the nominal condition.

The major advantages of the developed techniques for small anomaly detection and trend identification for faults are listed below:

- Robustness to measurement noise and spurious signals
- Adaptability to low-resolution sensing due to the coarse graining in space partitions
- Capability for early detection of anomalies because of sensitivity to signal distortion and real-time execution on commercially available inexpensive platforms.

IV. ALGORITHM VALIDATION

As an example, we illustrate an application of WT-SE based anomaly discovery and HHT based trend detection algorithms for validation purposes. The example system implements a second-order non-autonomous, forced Duffing equation with a cubic nonlinearity in one of the state variables [10][5], given by

$$\frac{d^2x}{dt^2}+\alpha\frac{dx}{dt}+x\left(1+\varepsilon x^2\right)=A\cos(\omega t)$$

where ε is not necessarily small. The numeric values for the system parameter and input excitation are chosen as follows: $\varepsilon=1$, $A=22$ and $\omega=5$. The challenge is detection of the slowly varying dissipation parameter α, which is associated with parametric anomaly during a fixed time frame. Furthermore, the non-stationary data of the system is corrupted with Gaussian noise to examine the robustness of the algorithms. The phase plot of the data is shown in Figure 1, where the dissipation parameter α is increased gradually from 0.10 to 0.35 in a span of 400 seconds. The sampling frequency of the data is 100 Hz. As seen in the figure, the signal content practically overlaps in the domain of α.

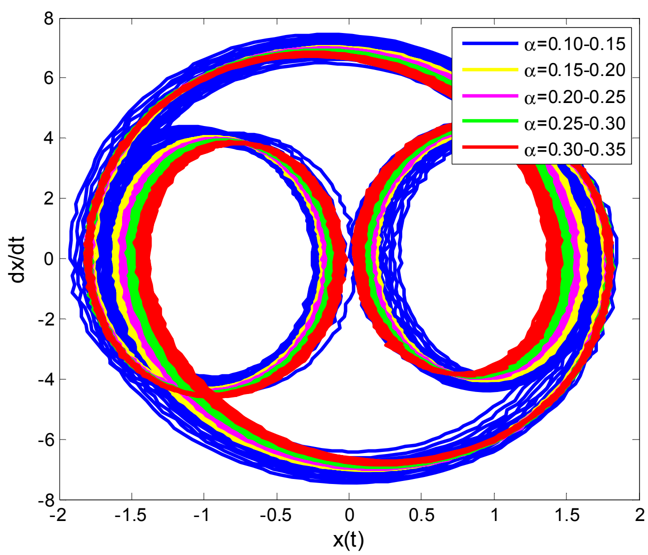

Figure 1. Phase plot of Duffing system with variable dissipation parameter α

It should be noted that both the wavelet-based and HHT-based algorithms utilize only time series data of the available measurements. Also, the noise content of the signal is not filtered so that we are able to evaluate the robustness of the algorithms. Undoubtedly, full observability and noise filtering would enhance the analysis results.

First, the WT-SE algorithm is employed for detection of the trend of anomaly in the time series data of the Duffing system. As described earlier, WT is used for attenuation of noise and spurious disturbances in the raw time series as well as time localized analysis of the data. For this example, we used 1st order Daubechies (Haar) wavelet in the analysis. Although the usual practice is to select a wavelet basis that has similar characteristics with the signal, our approach for symbolization of the wavelet coefficients via maximum entropy principle relaxes this criteria. The more important factor in the analysis is selection of the pseudo-frequency of the wavelet, since the wavelet coefficients of the signal are usually larger when this pseudo-frequency corresponds to the locally dominant frequencies in the underlying signal.

In the context of this paper, Symbolic Encoding helps estimating instantaneous probability vectors based on wavelet transformed. In a recent study [17], Hilbert transformed data is also used for symbolic encoding, albeit a two dimensional symbolization is performed in angular and radial directions. For this example, the number of states for finite state machine is chosen as 8 and the initial probability vector, P^0, is computed from first ten seconds of data. Subsequently, a sliding window of length 10 seconds is used for computation of instantaneous probabilities in 0.5 seconds intervals. Figure 2 presents the time evolution of angle between instantaneous probability vectors and the unit vector in the direction of the first component, which is chosen as a reference.

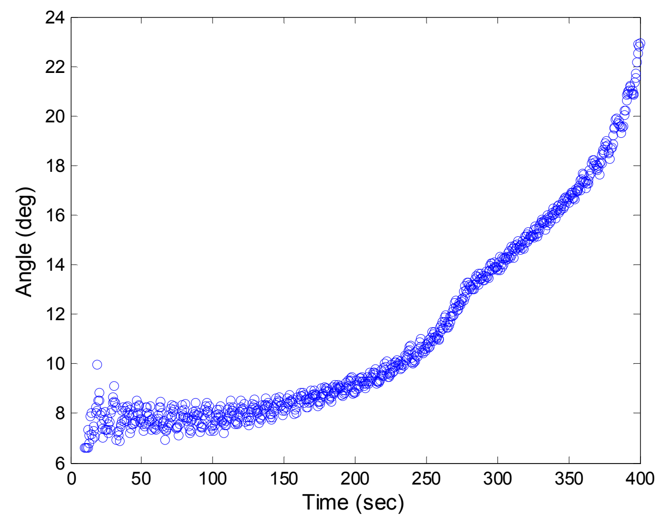

Figure 2. Time evolution of instantaneous probability vector using WT-SE

The time-frequency-energy localizations of the data using HHT are presented in Figure 3. The frequency spectrum of the data is unaffected as the parameter changes. This makes the frequency based analysis such as FFT ineffective for this kind of anomalies and detection of trend in the data becomes conspicuously difficult. Yet, in Figure 3 we can see the decrease in the energy of the lower fundamental frequency of the signal as parameter α increases (i.e. more the dissipation, less the energy remaining in the system). Just observing the energy content of this frequency makes a straightforward trend detection of the parametric anomaly. Another important feature of the analysis is rejection of noise and distortions in the data. This is achieved by checking the cross-correlations between the signal and each IMF component during the HTT process. Only the instantaneous frequencies and amplitudes of highly correlated IMF components are evaluated since the rest constitutes the noise and residuals in the signal. In this example, only two IMF components are analyzed since mean of their correlation sequences are at least three times that of remaining three components.

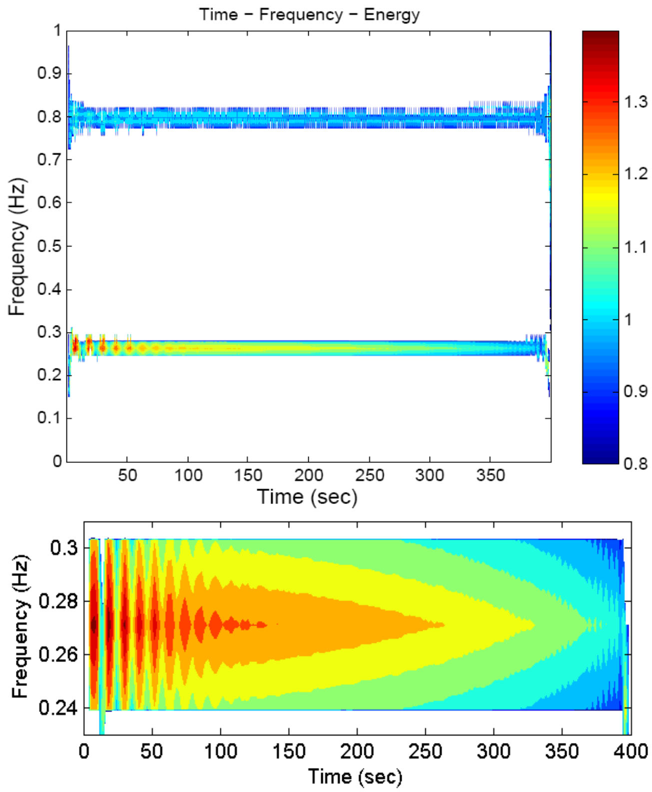

Figure 3. Time/Frequency/Energy evolutions of the data using Hilbert-Huang transform

The computational complexity is obviously a big concern for data mining applications when real time analysis or online implementation of the algorithms is required. It should be noted that we don't claim the algorithms to be optimal with respect to execution time or memory requirement. Yet, it is observed that both algorithms meet real time execution requirement as they finish analysis of 400 seconds of data (i.e. 40000 data points) on the average of 4.756 seconds and 3.492 seconds for WT-SE and HHT algorithms, respectively, on a 3.40 GHz Pentium 4 processor in MATLAB 7.1 environment. From memory point of view, the requirement of (random access) memory is in WT-SE case ~30 MB whereas in HHT case ~50 MB; both are reasonable for a commercially available laptop computer.

V. Summary and Conclusions

This paper presents two novel data mining methods and their comparison based on real-time implementation for trend detection of anomalies and faults. The first method uses the concept of wavelet transform-based symbolic encoding (WT-SE), since adaptive window characteristic of wavelet is suitable for analysis of non-stationary data and symbolic encoding allows the process to describe itself, without resorting to assumptions about its structure. The second approach utilized Hilbert-Huang Transform (HHT) to analyze non-stationary and nonlinear data. HHT offers the ability to perform flexible localized analysis in three independent variables (i.e. time-frequency-energy).

The Duffing system is used to validate the algorithms on a noise-corrupted nonlinear and non-stationary data set. WT-SE and HHT algorithms have been executed in real time with low memory requirements and both of them effectively detected the trend of a parametric anomaly in the time series. Both algorithms are able to capture the non-linear nature of the Duffing system as the dissipation parameter changes. The future work includes incorporating symbolic encoding for the HHT data and performing the signal analysis under distortions and non-Gaussian noise to investigate the robustness and adaptability of the algorithms. It is also aimed to optimize the algorithms to obtain better memory requirement and execution times.

References

[1] D. C. Champeney, *A Handbook of Fourier Theorems*, Cambridge University Press, 1987.

[2] P. Flandrin, *Time-frequency/Time-scale analysis*, Academic Press, 1999.

[3] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, 2004.

[4] J. Franke, W. Härdle and D. Martin, *Robust and Nonlinear Time Series Analysis*, Springer-Verlag, 1984.

[5] N. E. Huang and S. S. P. Shen, *The Hilbert-Huang Transform and Its Applications*, World Scientific, 2005.

[6] V. Rajagopalan and A. Ray, "Symbolic time series analysis via wavelet-based partitioning," *Signal Processing,* vol. 86, no.11, pp. 3309-3320, 2006.

[7] G. Kaiser, *A Friendly Guide to Wavelets*, Birkhauser, 1994.

[8] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", *Proc. Royal Society, A: Math. Phys. Eng. Sci.*, vol. 454, pp. 903-995, 1998.

[9] R. Badii and A. Politi, *Complexity: Hierarchical Structures and Scaling in Physics*, Cambridge University Press, 1999.

[10] A. Ray, "Symbolic dynamic analysis of complex systems for anomaly detection," *Signal Processing*, vol. 84, no. 7, pp. 1115-1130, 2004.

[11] M. M. Reda Taha, A. Noureldin, J. L. Lucero and T. J. Baca, "Wavelet transform for structural health monitoring: A compendium of uses and features," *Structural Health Monitoring*, vol. 5, no. 3, pp. 267-295, 2006.

[12] Z. K. Peng, P. W. Tse and F. L. Chub, "A comparison study of improved Hilbert–Huang transform and wavelet transform: application to fault diagnosis for rolling bearing," *Mechanical Systems and Signal Processing*, no. 19, pp. 974-988, 2005.

[13] H. D. I. Abarbanel, *Analysis of Observed Chaotic Data*, Springer, 1996.

[14] D. Lind and M. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, 1995.

[15] R. Samsi, *A Probabilistic Framework for Fault Detection in Induction Motors*, PhD. Thesis in Electrical Engineering, Pennsylvania State University, 2005.

[16] J. E. Hopcroft, R. Motwani and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley, 2003.

[17] A. Subbu and A. Ray, "Space partitioning via Hilbert transform for symbolic time series analysis," *Applied Physics Letters***,** vol. 92, no. 8, pp. 084107-1 to 084107-3, 2008.