2011

Detecting Temporal Patterns using Reconstructed Phase Space and Support Vector Machine in the Dynamic Data System

Wenjing Zhang
Marquette University, wenjing.zhang@marquette.edu

Xin Feng
Marquette University

2011 IEEE International Conference on Information and Automation, (June 6-8, 2011): 209-214. DOI.

Detecting Temporal Patterns Using Reconstructed Phase Space and Support Vector Machine in The Dynamic Data System

Wenjing Zhang
Department of Electrical and Computer Engineering, Marquette University, Milwaukee, WI
Xin Feng
Department of Electrical and Computer Engineering, Marquette University, Milwaukee, WI
Naveen Bansal
Department of Mathematics, Statistics, and Computer Science, Marquette University, Milwaukee, WI

Abstract:

In this paper we present a method for detecting dynamic temporal patterns that are characteristic and predictive of significant events in a dynamic data system. We employ the Gaussian Mixture Model (GMM) to cluster the data sequence into three categories of signals, e.g. normal, patterns and events. The data sequence is then embedded into a Reconstructed Phase Space (RPS) which is topologically equivalent to the dynamics of

the original system. We apply a hybrid method using Support Vector Machines (SVM) and Maximum a Posterior (MAP) to classify temporal pattern signals based on the event function. We performed two experimental applications using chaotic time series and Sludge Volume Index (SVI) series related to the Sludge Bulking problem. The proposed hybrid GMM-SVM phase space approach effectively detects temporal patterns and achieves higher predictive accuracy compared with the original RPS framework.

SECTION I. Introduction

Consider a data sequence of observations representing a measurement of the dynamic data system

$$X = \{x_t, t = 1,2, \dots N\}$$

(1)

where t is the time index and N is the total number of observations. The time sequence of observed data usually contains events of interest of the underlying dynamic system that are often complex and chaotic. These events are often closely related to time ordered structures, called temporal patterns in the sequence. Discovering temporal patterns that are closely related to the events are extremely important in many applications. For example, financial market investors want to retrieve temporal patterns in the historical prices of a specific security to decide a best timing for a long or short position [2]. In scientific research, discovering temporal patterns in large datasets benefit researchers to gain information of the underlying system dynamics [9].

The new method proposed here aims to detect temporal patterns that are characteristic and predictive of significant event of interest. Existing frequency domain approaches using Discrete Fourier Transform (DFT) [4] and Discrete Wavelet Transform (DWT) [5] [10] to classifying or matching time sequence data is based on spectral patterns to reduce dimension of the feature space. The frequency domain transformation chooses fewer but better coefficients to characterize the dynamic data system. Since these approaches focus on overall dynamic characteristics of the system, data sequences with different nonlinear dynamic patterns but similar power spectrum may not be distinguished. The method using piecewise linear representation in [6] is based on representing patterns as a set of simple templates and requires a priori knowledge of the internal structures of the dynamic data system.

Alternative approaches to temporal pattern identification include nonlinear classification using neural networks, decision trees and clustering algorithms [10]. Studies in dynamic systems and chaos theory provide a new pattern identification approach based on the reconstructed phase space [12]–[13][14][15]. It is capable of representing temporal patterns of nonlinear dynamic sequence data that is typically chaotic and irregular. Time Series Data Mining (TSDM) [1] [3] [7] is an effective approach to detect the temporal pattern. The basic concept is to embed the data sequence into a reconstructed phase space (RPS) with an optimized time delay τ and embedding dimension Q. The underlying theory discussed in [12] [13] guarantees that the such an embedding in RPS can describe the dynamic of a system given that the dimension of the phase space is greater than twice of the box counting dimension of the underlying system.

Since the dynamic data sequence discussed here are not pure signals and are significantly contaminated by correlated noise. Directly modeling the original data sequence in a phase space would result in poor performance as the patterns correlated to events are covered by the trajectory of unrelated signals and noise. The method proposed here applies a Gaussian Mixture Model (GMM) to screen potential temporal patterns that are statistically related to the events. From a Bayesian learning perspective, this essentially filter the data points that can be considered statistically irrelevant to the target event and reduce the noise in the resulting embedding in RPS. We then apply a second-stage classification using an SVM in the reconstructed phase space extending some previous work that employed GMM modeling [8].

SECTION II. Related Work

A reconstructed phase space is a Q-dimensional metric space into which a data sequence is unfolded. In [12], Takens showed that if Q is large enough, the phase space is homeomorphic to the state space that generated the data sequence. Takens' Theorem guarantees that the reconstructed dynamics are topologically identical to the true dynamics of the system. This provides the theoretical justification for reconstructing state spaces using time-delay embedding. Given a data sequence X, a state space topologically equivalent to the original state space can be reconstructed by a process called time-delay embedding [11]. The vectors, e.g. data points, $X_t, t = 1,2, \ldots, n$ in $\mathbf{R}^Q$ are represented by:

$$\begin{bmatrix} \mathbf{x}_{1+(Q-1)\tau} \\ \mathbf{x}_{2+(Q-1)\tau} \\ \vdots \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} x_{1+(Q-1)\tau} & \cdots & x_{1+\tau} & x_1 \\ x_{2+(Q-1)\tau} & \cdots & x_{2+\tau} & x_{1+\tau} \\ \vdots & & \ddots & \\ x_N & \cdots & x_{N-(Q-2)\tau} & x_{N-(Q-1)\tau} \end{bmatrix}_{(N-(Q-1)\tau\times 1)}$$

(2)

where τ is the time-delay and Q is the dimension of embedding vector. Povinelli and Feng [3], applied an optimized genetic algorithm to search the phase space for optimal heterogeneous clusters of temporal patterns that are predictive of specific events. Since genetic search algorithm is probabilistic and computational complex, the genetic search optimization approach may experience a high level of computing load and inconsistent results. In [1], A Gaussian shaped fuzzy membership function was applied to temporal pattern clusters and objective functions. The center and radius of the pattern clusters are then calculated by the fuzzy membership function instead of forming a rigid region with fixed center and radius. In [7], the predictability measure was introduced and applied a logistic regression for multiple temporal pattern clusters to evaluate the validity of these patterns.

However, in practice, temporal patterns can be hidden in the immense noise or unrelated signals. In such cases, directly an embedding into phase space may produce unsatisfactory results since the trajectories of patterns of interests may not be separated with noisy trajectories. Therefore, the original data sequence should be filtered based on the statistical modeling and characteristics related to the event of interests. Most of previous work treat temporal patterns by individual clusters and may become insufficient if the patterns in phase space are sparse and forming irregular non-spherical shapes. Due to the nature of event based pattern classification, it is desirable to perform a two class classification that separate the event patterns with nonevent patterns or noises.

SECTION III. GMM-SVM Classification in Phase Space

In this section, we present our approach to temporal pattern recognition by applying an SVM in the phase space and differentiate patterns that statistically significantly related to events. This is done in three stages. The first stage involves preprocessing the dataset and estimating the time delay and the dimension of RPs. The second stage is to apply GMM to the dataset estimating the statistical distribution of three classes which are normal, pattern and events. An MAP (Maximum a posteriori) classifier is used to find the Bayesian optimal decision rule determining the class each point belongs to. In the final stage, the temporal pattern classification is done with support vector machine classifier in the reconstructed phase space.

A. Dimension and Time-Delay Estimation

The time delay τ can be calculated by using the first minimum of mutual information function which provides the quantitative characteristics of spatial patterns in phase space [11]. The minimum of the mutual information function was found to be effective in estimating the time delay. Given a data sequence and time delay τ, the mutual information is computed by:

$$M(X_t, X_{t-tau}) = \sum_{i,j} p_{ij}(\tau) ln \frac{p_{ij}(\tau)}{p_i p_j}$$

(3)

The dimension Q of the RPS is determined using a false nearest-neighbors technique [13]. For each data point $\mathbf{x}_i^Q$ in R^Q, the changing rate of distance is defined by:

$$r_i = \sqrt{\frac{\|\mathbf{x}_i^{Q+1} - \mathbf{x}_j^{Q+1}\|^2 - \|\mathbf{x}_i^Q - \mathbf{x}_j^Q\|^2}{\|\mathbf{x}_i^Q - \mathbf{x}_j^Q\|^2}}$$

(4)

where $\|\mathbf{x}_i^Q - \mathbf{x}_j^Q\|$ is the Euclidean distance between its nearest neighbor $\mathbf{x}_j^Q = \underset{\mathbf{x}_j^Q, i \neq j}{argmin} \|\mathbf{x}_i^Q - \mathbf{x}_j^Q\|$

$\mathbf{x}_i^Q$ is marked as having a false nearest neighbor, if r_i exceeds a given threshold ρ. The criterion for adequate embedding dimension Q is that the number of data points for which $r_i > \rho$ is zero in R^Q.

B. Gaussian Mixture Model Classification

We formulate the temporal pattern detection problem by separating the dataset by three classes of data points: pattern, normal and event. Statistical properties and temporal structures are two important features of patterns in dynamic data sequence. Previous work under the reconstructed phase space framework has been successful in characterizing the temporal structure of the patterns. However, a discriminative approach has not been applied to characterize pattern and normal points based on the statistical properties. The approach proposed here employs a Gaussian mixture model to exploit the discriminative information in the data sequences. A data sequence can then be considered as a mixture of three different distributions which represent three states: normal, pattern and event. A MAP classifier is applied to determine Bayesian optimal thresholds that separate these three states. The classification applied here provides a filtering of data points to be embedded into phase space. For example, if data point x_t is classified as a pattern point, the sequence $\mathbf{x}_{t-1}^p = (x_{t-(Q-1)\tau-1}, x_{t-(Q-2)\tau-1}, \dots x_{t-1}^p)$ will be embedded into phase space. The filtering of noisy signals by setting a MAP classification to the dataset would significantly reduce the amount of data that embedded in phase space and typically improve the performance of classification in phase space. Consider the a Gaussian mixture model defined as

$$p(x) = \sum_{i=1}^{M} p(\omega_i) \mathcal{N}(x | \mu_i, \Sigma_i)$$

(5)

where M is the number of mixture components with each component distributed as $\mathcal{N}(x|\mu_i, \Sigma_i)$ with mean μ_i and covariance matrix Σ_i, and $p(\omega_i)$ is the marginal distribution for ith component of the mixtures, with constraint $\sum_{i=1}^{M} p(\omega_i) = 1$.

The Advantages of this type of model is that it can represent multi-modal distributions and for many cases, in the limit as $M \to \infty$, it can also represent any possible distribution. For each point in the pattern state, we formulate target events by defining the event function:

$$g(\mathbf{x}_t) = max\{x_{t+1}, \dots, x_{t+k}\} - c > 0$$

(6)

where k is the time-step ahead, c is the threshold that defines an event of interests, and ΔT is a preselected maximum time frame for an event prediction. The event function also defines two classes of nonevent points. A data point x_t can be considered as a pattern point if it satisfies (6), otherwise considered as non-pattern point. Under the Phase Space framework, ΔT will be a function of embedding dimension Q and time delay τ.

Denote three states, $\omega_1, \omega_2, \omega_3$ indicating the classes of normal, pattern and event respectively. The Expectation Maximization algorithm [16] is used to estimate the parameters of GMM:

$$\hat{p}(\omega_i) = \frac{1}{n}\sum_{k=1}^{n} \hat{p}(\omega_i|x_k, \hat{\theta})$$

$$\hat{\mu}_i = \frac{\sum_{k=1}^{n} p(\omega_i|x_k, \hat{\theta})x_k)}{\sum_{k=1}^{n} p(\omega_i|x_k, \hat{\theta})}$$

$$\hat{\Sigma}_i = \frac{\sum_{k=1}^{n} P(\omega_i|x_k)(x_k - \hat{\mu}_i)(x_k - \hat{\mu}_i)^T}{\sum_{k=1}^{n} P(\omega_i|x_k)}$$

(7)

where

$$\hat{p}(\omega_i|x_k, \hat{\theta}) = \frac{\mathcal{N}(x_k|\hat{\mu}_i, \hat{\Sigma}_i)\hat{p}(\omega_i)}{\sum_{i=1}^{c} \mathcal{N}(x_k|\hat{\mu}_i, \hat{\Sigma}_i)\hat{p}(\omega_j)}$$

(8)

By using the model given by learning a GMM, we can construct maximum a posteriori (MAP) classifier to separate normal, pattern points. A data point x_k is classified by selecting the class that gives higher posterior likelihood:

$$\hat{\omega} = \arg\max_{i} \; \mathcal{N}(x_k | \hat{\mu}_i, \hat{\Sigma}_i)\hat{p}(\omega_i)$$

(9)

C. Support Vector Machine Classification in RPS

In the first stage of classification, a two class Gaussian mixture model is used to make a first step classification of pattern and normal points. In this step, dataset is classified by based on GMM clustering which does not include the information of whether the event occurs within ΔT. We formulate the problem by applying labels to each embedding $\mathbf{x}_t^p = (x_{t-(Q-1)\tau}, x_{t-(Q-2)\tau}, \dots x_t^p)$ with false and positive indicator $g_i \in \{+1, -1\}$ based on the true occurrence of events defined by the event function (6).

Traditional RPS method typically embeds all the data point into phase space and uses a clustering method in the phase space with an event function to separate the event and nonevent patterns. The approach would perform well for datasets with relatively low level of noise with respect to signals, but may not work well for many applications when the dynamics of the system is significantly affected by noise. In such cases, a well-defined cluster may not exist in a phase space embedding. In our approach, an MAP classifier based on Gaussian mixture modeling effectively reduces the number of data point embedded in phase space and typically result in a sparse, separable embedding which well suited for a SVM classification task.

A Support Vector Machines (SVM) [17] [18] looks for the best solution by finding the optimal separating hyperplane. Consider a linear model form $g(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\phi(\mathbf{x})$, where $\phi(x)$ is a nonlinear transformation that maps the original X into a higher dimensional space. Given vectors $\mathbf{x}_i, i = 1, \dots l$ with an indicator vector $y_i \in \{+1, -1\}$ and slack variables ξ_i, C-SVC solves the following optimization problem:

$$\begin{aligned}
&\min_{\mathrm{w},b,\xi} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{l}\xi_i \\
&\text{subject to } y_i(\mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}_i) + b) \geq 1 - \xi_i \\
&\xi_i \geq 0, i = 1, \dots l
\end{aligned}$$

(10)

By using Karush-Kuhn- Tucker (KKT) condition, we can reformulate the problem as a dual problem

$$\begin{aligned}
&\max_{\alpha} \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j y_i y_j K(x_i, x_j) \\
&\text{subject to} \sum_{i=1}^{l}\alpha_i y_i = 0 \\
&0 \leq \alpha_i \leq C, i = 1, \dots l,
\end{aligned}$$

(11)

where $K(\mathbf{x}_i, \mathbf{x}_j) \equiv \phi(\boldsymbol{x}_i)^T\phi(\boldsymbol{x}_j)$ is the kernel function.

The new vector $\mathbf{x}_t^p$ in phase space will be classified as a member of the pattern class if

$$\sum_{i=1}^{l} \alpha_i g_i K(\mathbf{x}_i^p, \mathbf{x}^p) + b > 0$$

(12)

where α_i is the solution of the dual problem.

SECTION IV. Overview of the Method

The proposed method can be summarized as follows: first, determine the dimension Q of the phase spaces and the lengths of the temporal patterns τ, Then GMM is learned from training dataset by EM algorithm to estimate three mixtures that are normal, pattern and event. A MAP classifier is applied to determine the decision threshold for classifying three mixtures. Then for a data point x_t classified as pattern point, the time sequence $(x_{t-(Q-1)\tau-1}, x_{t-(Q-2)\tau-1}, \dots x_{t-1})$ is embedded into phase space as vector $\mathbf{x}_{t-1}^p$. The next step is to apply a SVM to classify the temporal pattern structure in the phase space based on the true event occurrence defined by the event function. This second stage classification finds a decision function that separate the "false" and "true" patterns that are predictive of events with high confidence. Fig. 2 shows procedure of the method.

SECTION IV. Experiments

In this section, we describe two experiments applying our algorithm on a simulated chaotic time series and a real task. The first dataset is generated by Lorenz equations, a benchmark problem, with 500-second duration and a sampling frequency of 125Hz. The second dataset is Sludge Volume Index (SVI) series which is an effective indicator for controlling the rate of de-sludging in water treatment process. We compare the results by our approach to the original phase space framework.

A. Chaotic Time Series Analysis

A Lorenz time series is generated by setting the initial values: $x_0 = 0, y_0 = -0.01, z_0 = 0.01$, and parameters: $\sigma = 9, r = 25$, and $b = 3.3$. The time delay was estimated as $\tau = 0.2s$ and the embedding dimension $Q = 3$.

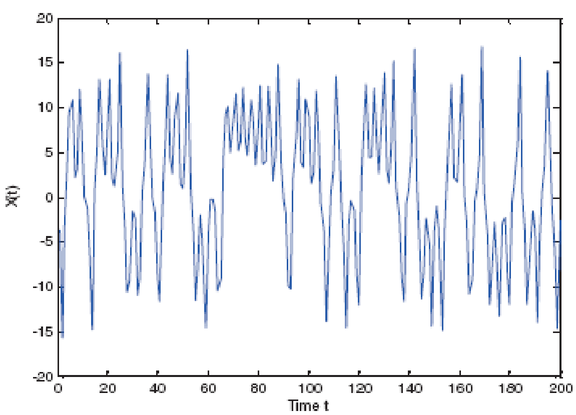

Fig.1. Sample Lorenz series x(t)

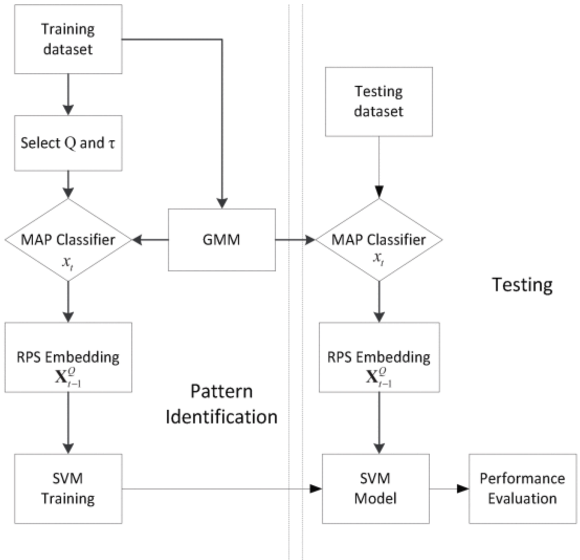

Fig. 2. Block Diagram of the Method

The events are defined by the event function

$$g(x_t) = x_{t+1} - 11.0 > 0$$

(13)

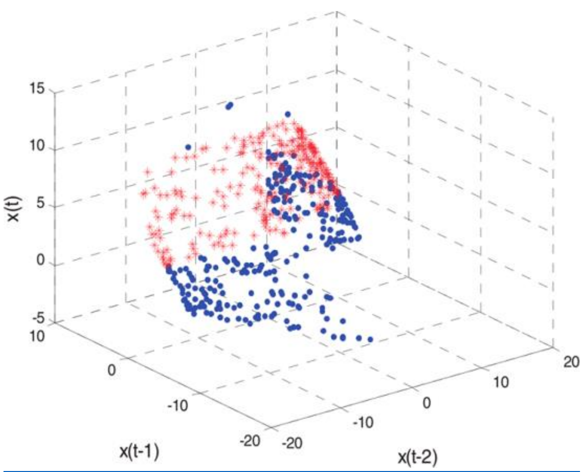

Fig. 3. Temporal patterns in 3D phase space

which means that in the following time point, the magnitude of the series is greater than a threshold, 11.0. The first 1500 data points are used for our training process and the other 700 data points are used for testing. We estimate the distribution of the normal, pattern and event data points using GMM. MAP classifier is then applied to label the data points that can be considered temporal pattern points or irrelevant to defined events. The pattern points are embedded into phase space and classified using SVM based on the labeling defined by (13).

The results of proposed method and previous framework are presented in Table II. As a comparison, this example shows that in the training phase both methods achieves 100%

TABLE I The Test Results of the Gmm-Svm Method of Lorenz series

	Predicted as events	Predicted as nonevents
Actual events	True positive=49	False negative=9
Actual nonevents	False positive =6	True negative=436

TABLE II The Comparison of Prediction Results of Methods

	Training Set		**Test Set**	
	GMM-SVM	**TSDM**	**GMM-SVM**	**TSDM**
Prediction Accuracy	99.56%	99.86%	89.09%	98.52%
Positive Accuracy	87.45%	62.35%	84.48%	48.52%

prediction accuracy and new method has higher positive accuracy. In the testing phase, the results given by the new approach is consistent with the training results, whereas the TSDM method poor recognition rate of events with a low positive accuracy.

B. Sludge Volume Index (SVI) Analysis

Sludge bulking is a phenomenon that occurs in water treatment plant. It is one of the primary causes of water treatment plant failure as the bulking conditions result in exceeding discharge limitations. Efforts have been made to study this problem from a biological point of view, but failed to formulate a deterministic cause-effect relationship. Since the current knowledge is limited, unconventional approaches have been applied including stochastic models and artificial neural systems [9].

The sludge volume index (SVI) used here is the primary indicator representing the bulking conditions. The SVI dataset is authorized by a water treatment company located in Chicago, the measurements of the data sequence are obtained by a daily observation from 2002–2008. According to U.S. government water treatment policy, the events are defined by the event function $g(\mathbf{x}_t) = max\{x_{t+1}, \dots, x_{t+k}\} - 150.0 > 0, k = 3$. Our approach is applied to this dataset and the identified temporal patterns are presented in Fig. 4

The temporal patterns are marked by squares lines with embedding dimension $Q = 4$. The last point of the pattern is the predicting point which indicates the sludge bulking is likely to occur with high probability. The results of proposed method and previous framework are presented in Table III.

In comparing the two methods, we can see that for this SVI prediction task, the GMM-SVM approach outperforms the TSDM approach by 21.28% in the testing phase. The results can be better understood by examining the temporal pattern plotted in Fig. 4. The patterns that related to the events are not simply consistent of their structures and shapes. Instead, the temporal patterns are time-evolving and not obvious to capture by traditional approaches. The results demonstrated that the new approach could be applied in a monitoring system for the sludge bulking conditions of water treatment plants providing early alerts for the potential bulking problems.

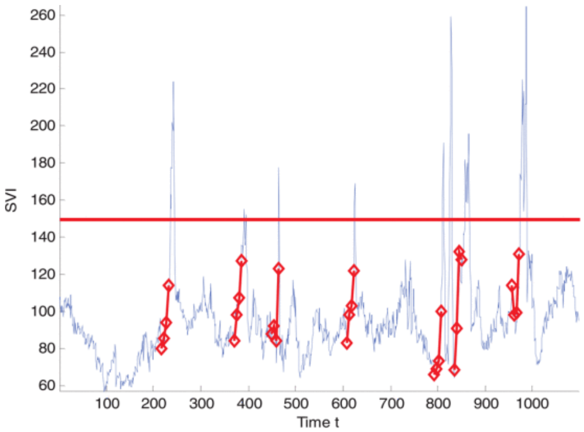

Fig. 4. Temporal Patterns and Prediction of SVI

TABLE II The Comparison of Prediction Results of Methods

	Training Set	**Test Set**		
	GMM-SVM	**TSDM**	**GMM-SVM**	**TSDM**
Prediction Accuracy	90.25%	75.32%	81.25%	65.73%
Detection Accuracy	75.35%	63.58%	70.56%	51.28%

SECTION V. Conclusion

In this paper, we presented a new GMM-SVM algorithm based on the reconstructed phase space. This algorithm provides a discriminative approach that utilizes both by Gaussian mixture model and SVM technique to classify temporal patterns that are statistically correlated with events in dynamic data system.

We assessed the performance of the method by applying this approach to two dataset. Experiments show that this method yields significant improvement in the event detection rate compared with baseline RPS method.

There are several issues need to be further studied. Although we have employed Gaussian mixture model, more distributions can be applied for clustering. Also, we have focus on the single data sequence and another extension can be achieved by combining multiple data sequences.

References

1. X. Feng and H. Huang, "A Fuzzy-Set-Based Reconstruction Phase Space Method for Identification of Temporal Patterns in Complex Time Series", *IEEE Trans. on Knowledge and Data Engineering*, vol. 17, no. 5, pp. 601-613.

2. Leon Lee Chiung-Hon, Alan Liu and Wen-Sung Chen, "Pattern discovery of fuzzy time series for financial prediction", *IEEE Trans. on Knowledge and Data Engineering*, vol. 18, no. 5, may 2006.

3. R.J. Povinelli and X. Feng, "A New Temporal Pattern Identification Method for Characterization and Prediction of Complex Time Series Events", *IEEE Trans. on Knowledge and Data Engineering*, vol. 15, no. 2, pp. 339-352, March/Apri1 2003.

4. C. Faloutsos, M. Ranganathan and Y. Manolopoulos, "Fast subsequence matching in time-series databases", *SIGMOD*, pp. 419-429, 1994.

5. K. Chan and A. Fu, "Efficient time series matching by wavelets", *In: ICDE*, pp. 126-133, 1999.

6. E. Keogh, "A fast and robust method for pattern matching in time series databases", *Proc. 9th Int. Con! on Tools with Artificial Intelligence ICTAL*, pp. 578-584, 1997.

7. X. Feng and O. Senyana, "Mining Multiple Temporal Patterns of complex dynamic data systems", *CIDM 2009*, pp. 411-417.

8. R. J. Povinelli, M. T. Johnson, A. C. Lindgren and J. Ye, "Time series classification using Gaussian mixture models of reconstructed phase spaces", *IEEE Trans. on Knowledge and Data Engineering*, vol. 16, pp. 779-783, 2004.

9. Capodaglio, H. Jones, V. Novotny and X. Feng, "Sludge bulking analysis and forecasting: application of system identification and artificial neural computing technologies", *Water Research*, vol. 25, no. 10, pp. 1217-24.

10. D. B. Percival and A. T. Walden, Wavelet Methods for Time Series Analysis, Cambridge, England:Cambridge University Press, 2000.

11. H. Kantz and T. Schreiber, Nonlinear Time Series Analysis, Cambridge University Press, 1997.

12. F. Takens, "Detecting Strange Attractors in Turbulence" in Lecture Notes in Math, New York:Springer, vol. 898, 1981.

13. M. B. Kennel, R. Brown and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction", *Phys. Rev*, vol. 45, pp. 3403, 1992.

14. T. Sauer, J.A. Yorke and M. Casdagli, "Embedology", *J Statistical Physics*, vol. 65, pp. 579-616, 1991.

15. J. Iwanski and E. Bradley, "Recurrence Plot Analysis: To Embed or not to Embed?", *Chaos*, vol. 8, pp. 861-871, 1998.

16. T.K. Moon, "The expectation-Maximization Algorithm", *IEEE Signal Processing Mag*, vol. 13, pp. 47-59, 1996.

17. J. C. Burges, "A tutorial on support vector machines for pattern recognition", *Data Mining Knowl. Disc*, vol. 2, no. 2, pp. 1-47, 1998.

18. Cortes and V. Vapnik, "Support vector networks", *Machine Learning*, vol. 20, pp. 273-297, 1995.

19. J. V. Hansen and R.D. Nelson, "Neural Networks and Traditional Time Series Methods: A Synergistic Combination in State Economic Forecasts", *IEEE Trans. on Neural Networks*, vol. 8, no. 4, July 1997.

20. S.M. Weiss and N. Indurkhya, Predictive Data Mining: A Practical Guide, San Francisco:Morgan Kaufmann, 1998.

21. S.M. Pandit and S.-M. Wu, Time Series and System Analysis with Applications, New York:Wiley, 1983.