Identification of connectivity changes in multivariate time series using feature selection

Alkiviadis Tsimpiris and Christos Koutlis

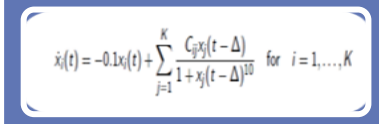
Department of Electrical and Computer Engineering,
Aristotle University of Thessaloniki,
Thessaloniki 54124, Greece

The work is supported by the Greek General Secretariat for Research and Technology (Aristeia II, No 4822)

The Objective

- The problem of this study is the identification of changes in the connectivity structure of a high dimensional dynamical system observed through a multivariate time series.

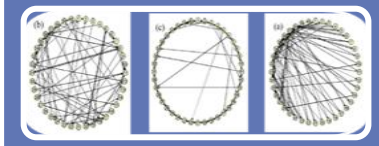
Flow chart



$$\dot{x}_i(t) = -0.1x_i(t) + \sum_{j=1}^{K} \frac{C_{ji}x_j(t-\Delta)}{1+x_j(t-\Delta)^{10}} \quad \text{for} \quad i = 1,\ldots,K$$

Simulation Study on multivariate time series

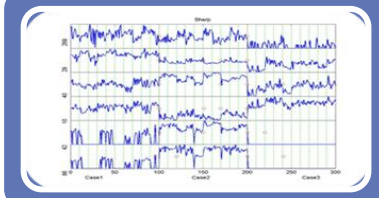
(Mackey-Glass 25 variables)



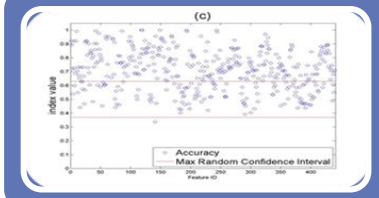
Connectivity structures (Random, Small world, Scale free)

PMIME

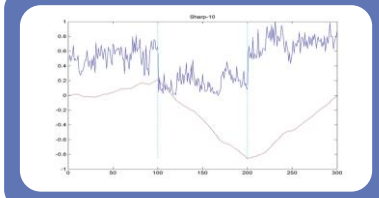
(A). Sharp changes, (B). Smooth changes



469 Network features



Feature selection and classification of different connectivity structures (Bayes classifier, FSS, SVM RFE, CMINN)



On-line detection of time series changes (CUSUM, sliding window of 30 points, alarm)

Simulation Study

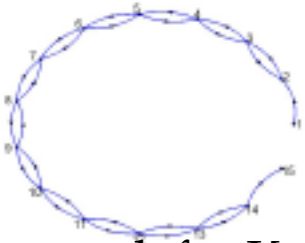
- In the simulation study, the two classes are combinations of the connectivity structures formed by a random network, generated by the Erdös-Renyi model, a small world network and a scale free network
- These structures determine the coupling of 25 variables in the coupled Mackey-Glass system of high complexity generating a long multivariate time series record starting with the random connectivity structure in the middle of the record changing to the small world structure and in the end the record is changing to the scale free structure
- **Scenario A.**
  - Sharp changes between different network structures (also 10 internal sharp changes of each network structure)
- **Scenario B.**
  - Smooth changes between different network structures (also 10 internal smooth changes of each network structure)

Multivariate time series

- Coupled Mackey-Glass differential equations

$$\dot{x}_i(t) = -0.1x_i(t) + \sum_{j=1}^{K} \frac{C_{ij}x_j(t-\Delta)}{1+x_j(t-\Delta)^{10}} \quad \text{for} \quad i = 1, \ldots, K$$

- Δ=100, C=0.2, n=2048, t=10min, K=25



e.g True network for K=15

The connectivity structure

- The connectivity structure is estimated from the multivariate time series using an information measure of Granger causality designed to deal with many observed variables, called partial mutual information from mixed embedding (PMIME).

The causality measure PMIME

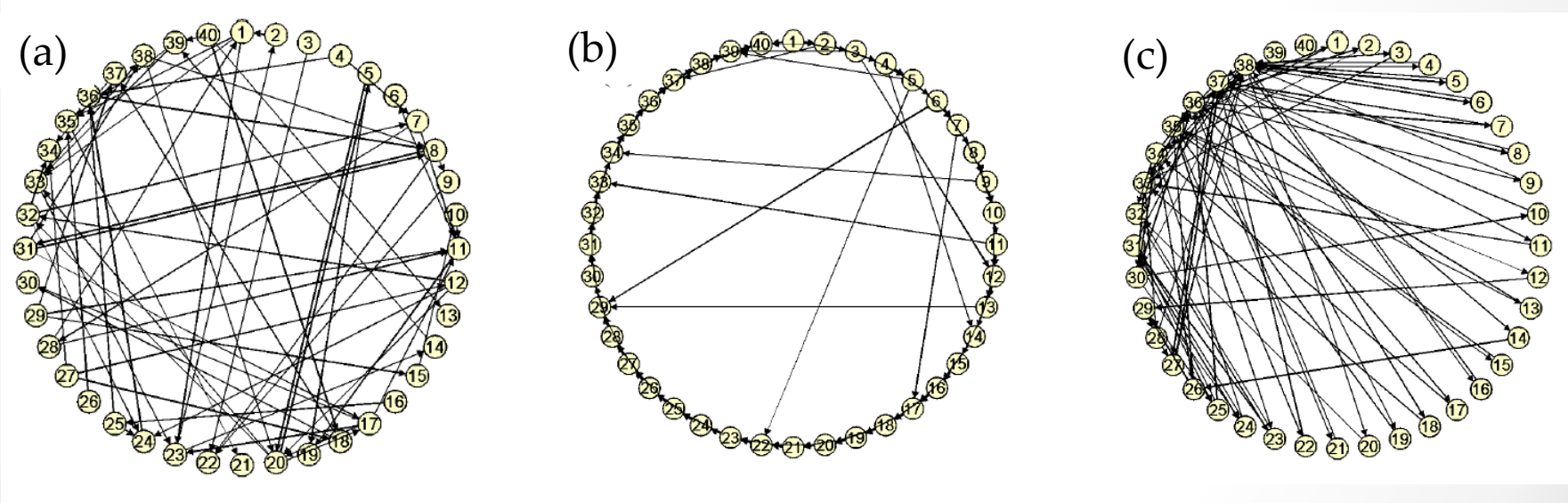
$$R_{X\to Y|Z} = \frac{I(\mathbf{y}_t^T; \mathbf{w}_t^X \mid \mathbf{w}_t^Y, \mathbf{w}_t^Z)}{I(\mathbf{y}_t^T; [\mathbf{w}_t^X, \mathbf{w}_t^Y] \mid \mathbf{w}_t^Z)}$$

$R_{X\to Y|Z}$: information of Y explained only by X-components of the embedding vectors, normalized against the mutual information accounting for the presence of $\mathbf{w}_t^Z$.

$$\mathbf{w}_t = (\underbrace{x_{t-\tau_{x1}}, \ldots, x_{t-\tau_{xm_x}}}_{\mathbf{w}_t^X}, \underbrace{y_{t-\tau_{y1}}, \ldots, y_{t-\tau_{ym_y}}}_{\mathbf{w}_t^Y}, \underbrace{z_{t-\tau_{z1}}, \ldots, z_{t-\tau_{zm_z}}}_{\mathbf{w}_t^Z})$$

The network

- A network is formed having nodes the observed variables and directed connections given by PMIME



Network structures: (a) Random, (b) Small world, (c) Scale-free

The features

- 469 of network indices are computed on the network, which characterize the network topology and are termed features
- Feature values length (300, 150 points)

Classification Problem

- The problem is posed as a classification task, where the classes are the different connectivity structures (seen as networks) to be discriminating on the basis of the generated multivariate times series from each connectivity structure and the features computed on them.
- Bayes Classifier
- Classification with single features

Feature selection

-Feature selection is applied using network features (accuracy >=0.9)

-The best feature subset is found using a majority vote on feature subsets of 3 different feature selection methods

- FSS (forward sequential selection)
- SVM-RFE (Support Vector Machines - Recursive Feature Elimination)
- CMINN (Conditional Mutual Information with Nearest Neighbors estimate)

- 10 realizations
- Random Training set 70%, Test set 30%

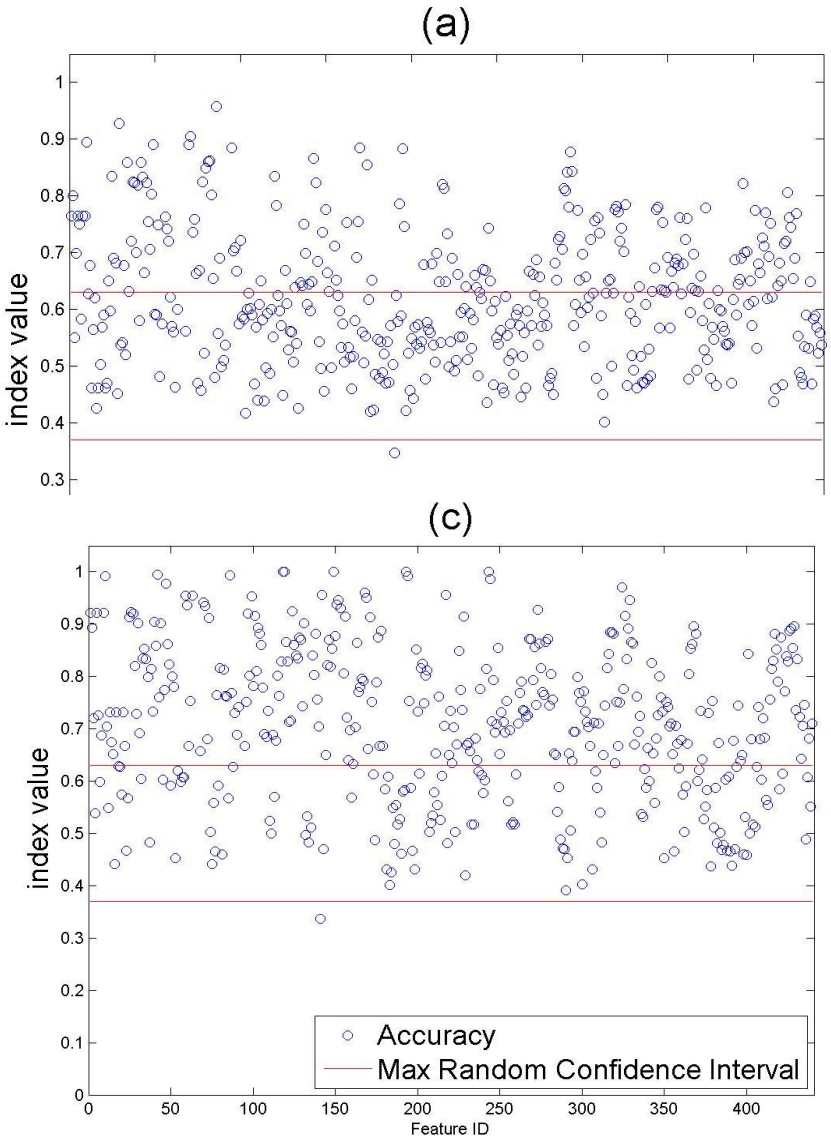
On-line detection of time series changes.

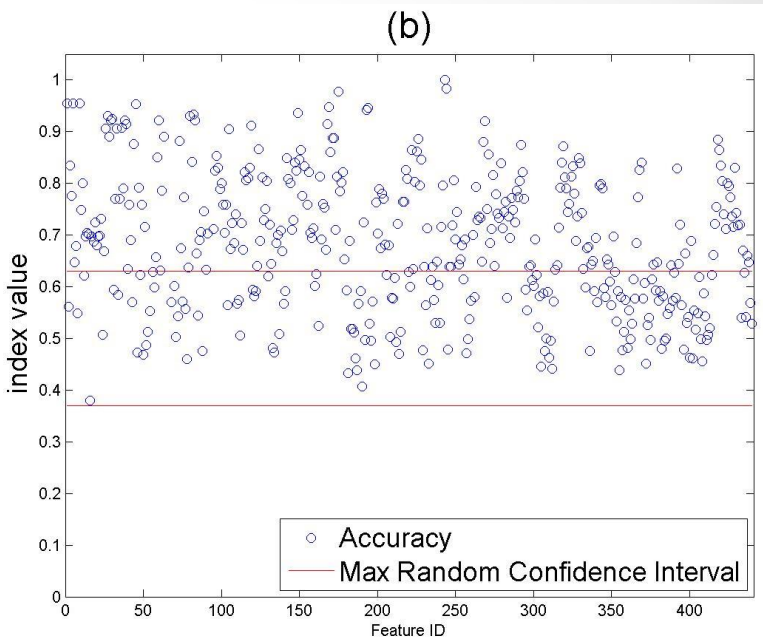
- CUSUM algorithm (cumulative sum control chart) is a sequential analysis technique developed by E. S. Page used for monitoring change detection

$$C_k = \sum_{i=1}^{k} (y_i - \bar{y}), \qquad k = 1, 2, \ldots, N-1$$

- CUSUM application on features with accuracy >=0.9
- Sliding window of 30 feature values (point of min(C_k))
- Alarm (Scenario A >0.8 of window length)
  (Scenario B >0.6 of window length)

Results A (sharp changes)



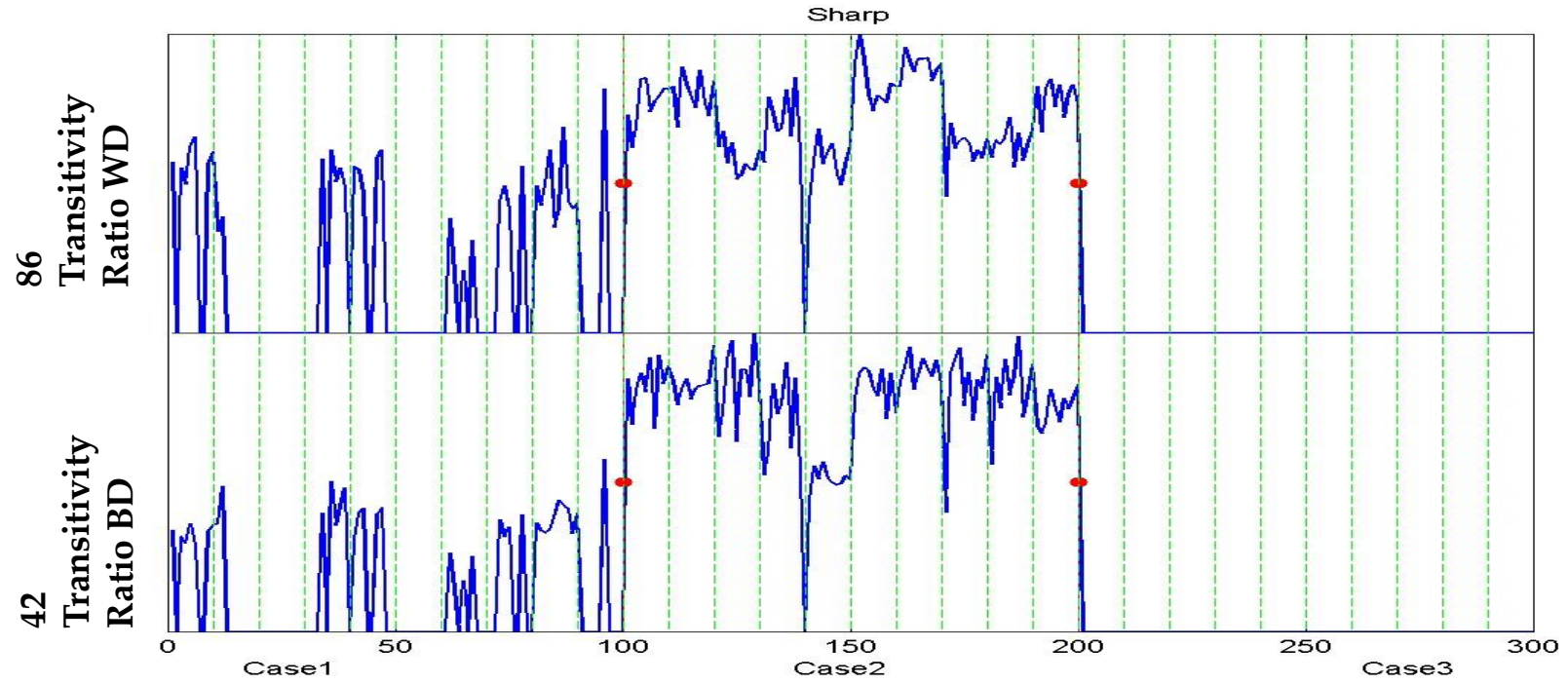


- Accuracy values of single features
- Bayes classifier

a) Random Network vs Small World
b) Random Network vs Scale Free
c) Small World Network vs Scale Free

Results A (sharp changes)

Window length=30
CUSUM threshold=24 (0.8*window)



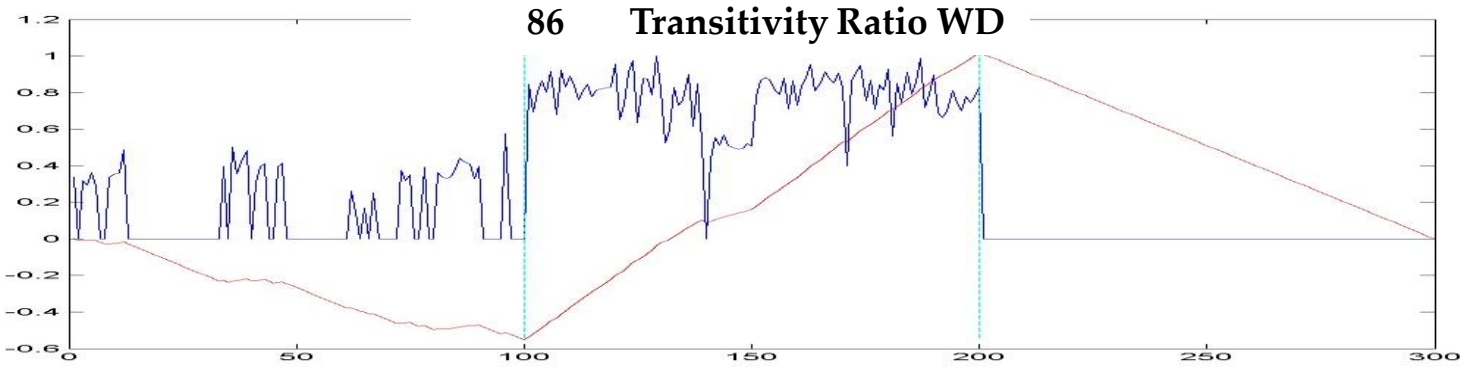
et=[1,5,9,10,25,26,27,29,30,33,36,38,39,40,42,44,45,47,59,60,63,70,71,73,80,82,83,86,97,99,101,105,106,118,119,120,124,131,142,
149,151,152,153,155,156,168,169,171,175,193,194,217,228,243,244,263,269,273,294,324,326,329,394]; 63 features

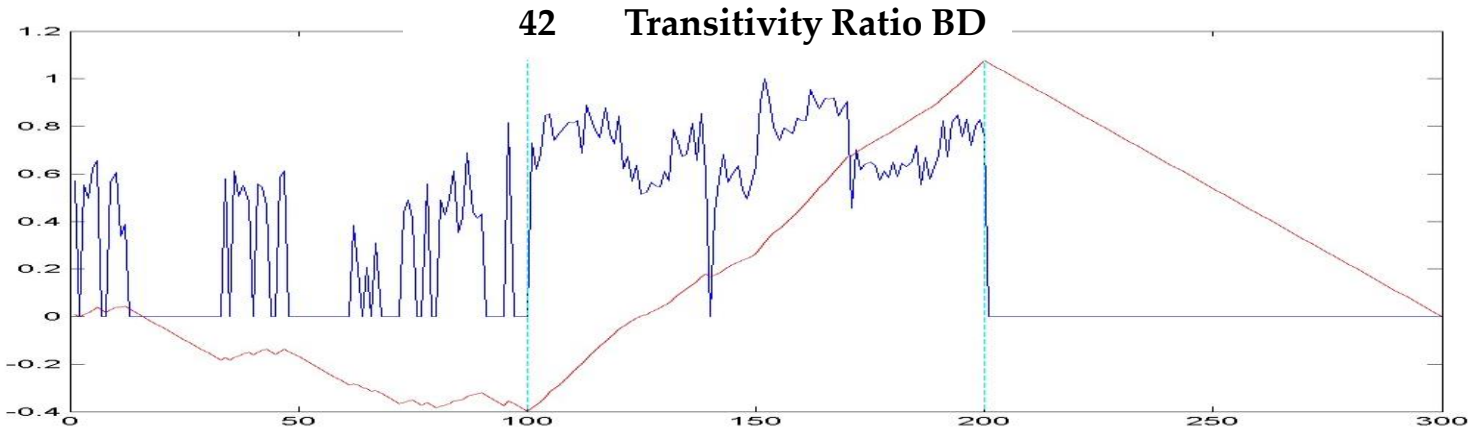
Normalized best feature values selected by CUSUM

Case1 (0-100) Random Network
Case2 (101-200) Small World Network
Case3 (201-300) Scale Free

42 Transitivity Ratio BD
86 Transitivity Ratio WD

Results A (Sharp changes)





Normalized best feature values and CUSUM values

Case1 (0-100) Random Network vs Small World
Case2 (101-200) Random Network vs Scale Free
Case3 (201-300) Small World Network vs Scale Free

Results A (sharp changes)

FSS		
	Average cardinality	Average Accuracy
a)	2.2	1.00
b)	1	1.00
c)	1	1.00

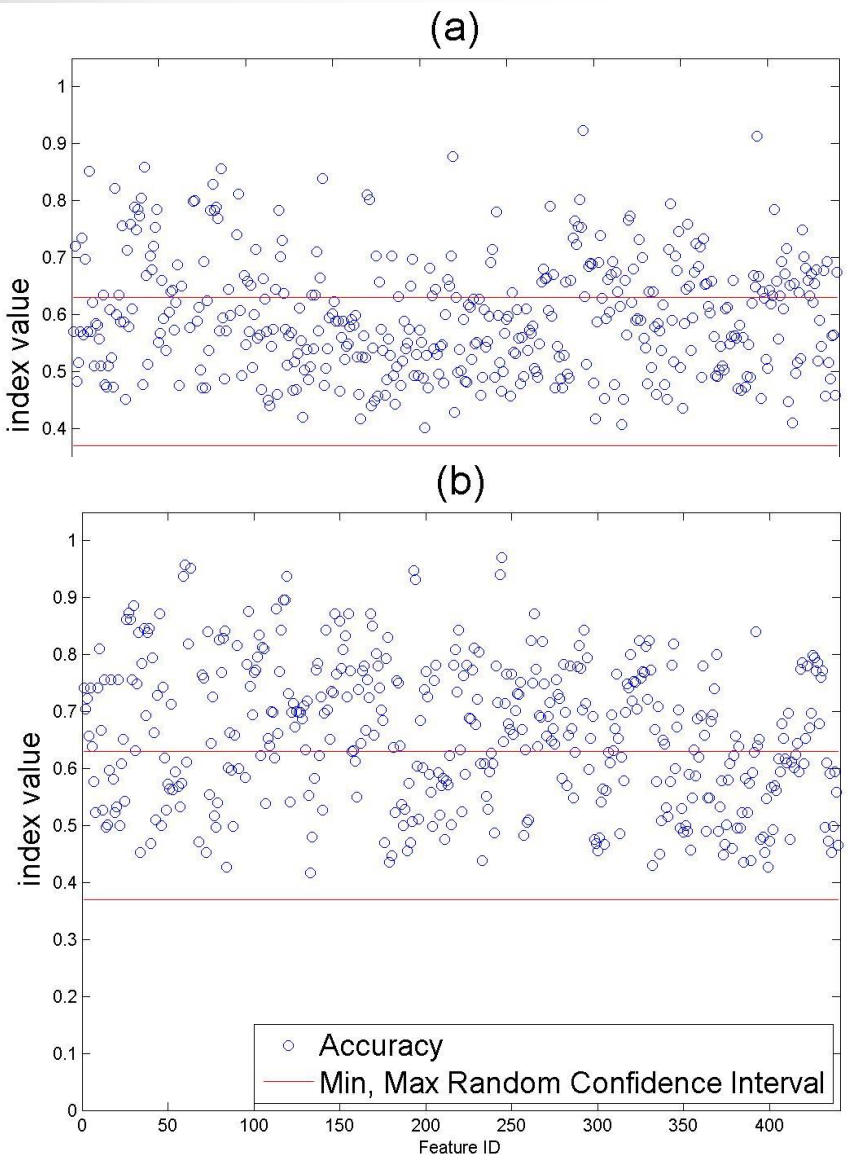
CMINN 5 features	
	Average Accuracy
a)	0.97
b)	1.00
c)	0.98

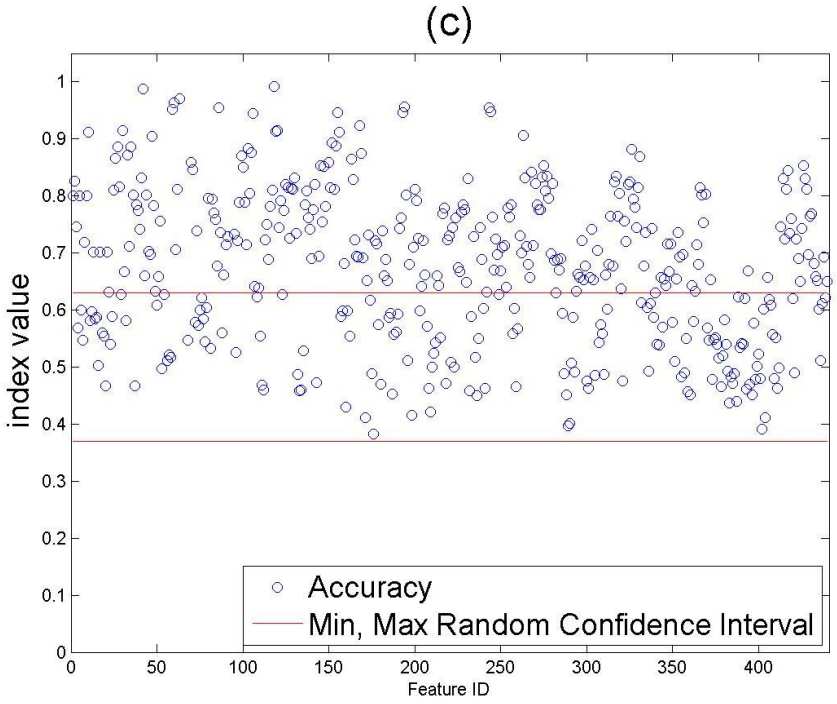
SVM-RFE 5 features	
	Average Accuracy
a)	0.99
b)	1.00
c)	0.99

- Accuracy values for features subsets of FSS
- Bayes classifier

Case a) Random Network vs Small World
Case b) Random Network vs Scale Free
Case c) Small World Network vs Scale Free

Results B (smooth changes)



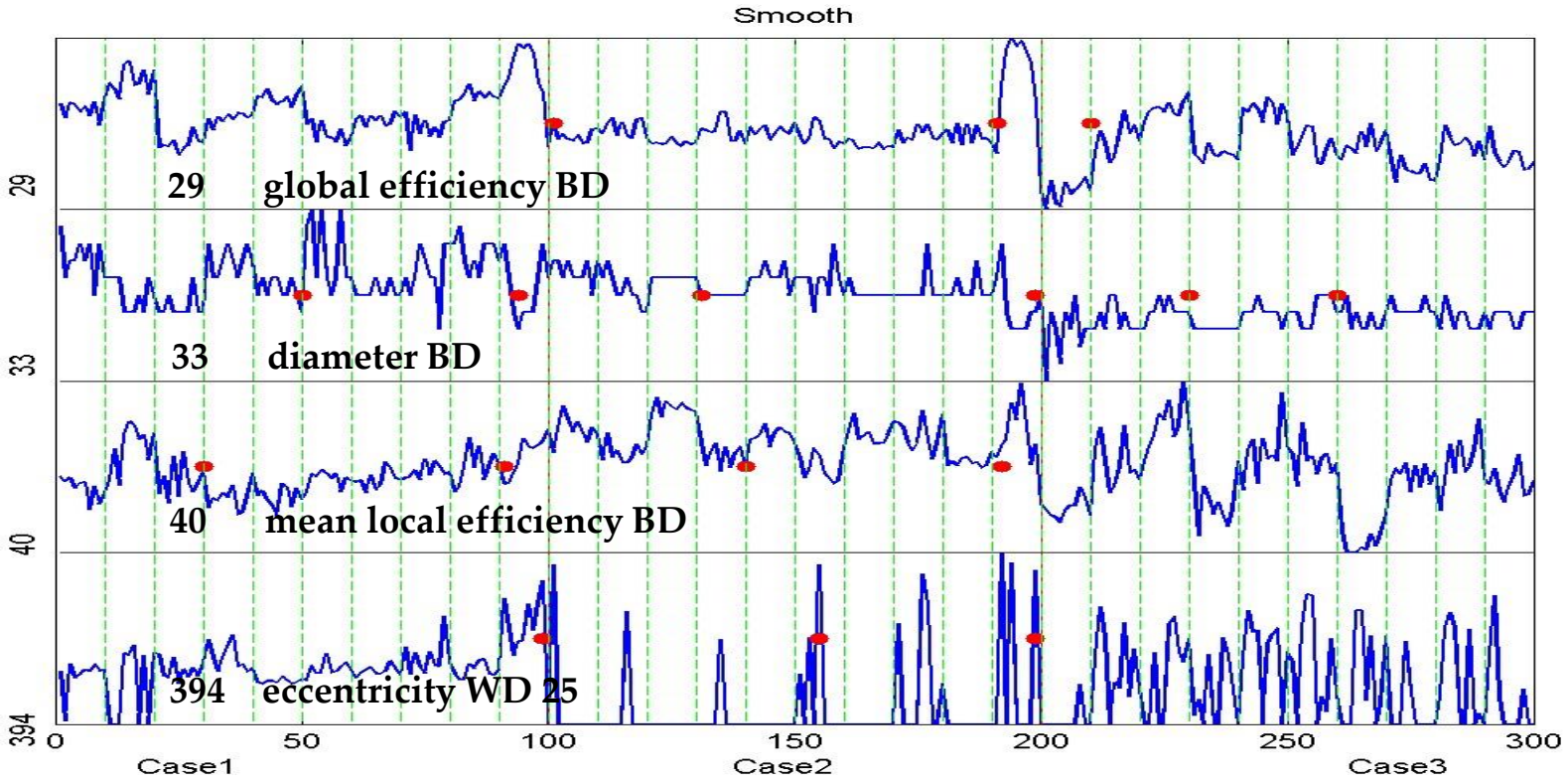


- Accuracy values of single features
- Bayes classifier

a) Random Network vs Small World
b) Random Network vs Scale Free
c) Small World Network vs Scale Free

Results B (smooth changes)

Window length=30
CUSUM threshold>=18 (0.6*window)

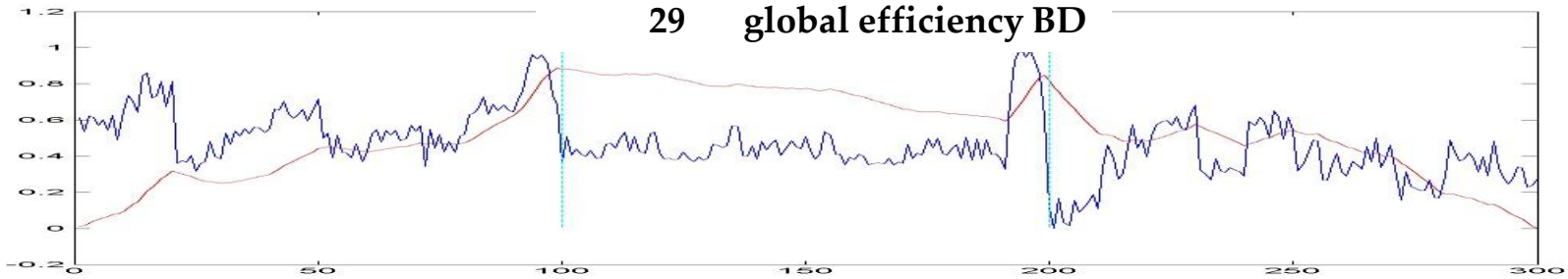


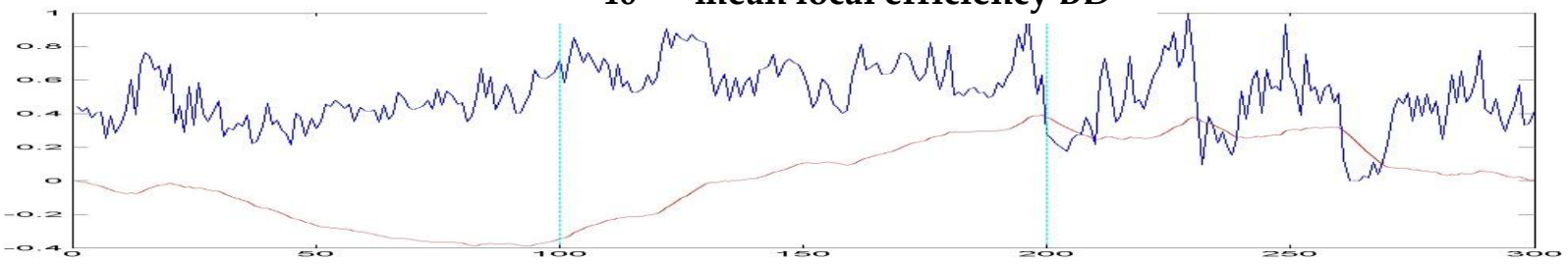
Normalized best feature values

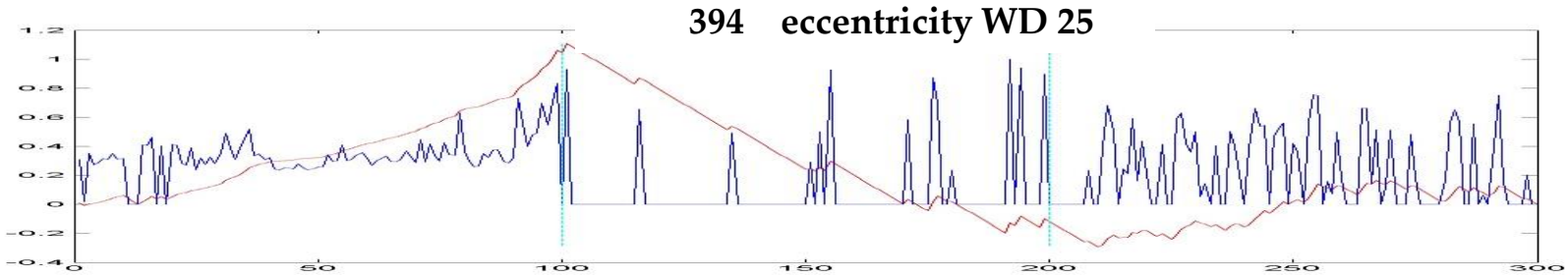
Case1) Random Network vs Small World
Case2) Random Network vs Scale Free
Case3) Small World Network vs Scale Free

29 global efficiency BD
33 diameter BD
40 mean local efficiency BD
394 eccentricity WD 25

Results B (smooth changes)







Normalized best feature values and CUSUM values

Case1 (0-100) Random Network vs Small World
Case2 (101-200) Random Network vs Scale Free
Case3 (201-300) Small World Network vs Scale Free

Results B (smooth changes)

FSS		
	Average cardinality	Average Accuracy
a)	2.5	0.98
b)	1.2	1.00
c)	1.3	1.00

CMINN 5 features	
	Average Accuracy
a)	0.94
b)	0.99
c)	0.97

SVM-RFE 5 features	
	Average Accuracy
a)	0.94
b)	1.00
c)	0.98

- Accuracy values for features subsets of **FSS**
- Bayes classifier

a) Random Network vs Small World
b) Random Network vs Scale Free
c) Small World Network vs Scale Free

Conclusions

- PMIME reduces the dimension of the connectivity problem
- Single Features can classify different network structures
- Feature selection improves classification efficiency in cases where single features are not enough
- CUSUM detect successfully time series changes
- Need for more realizations
- Large computation time for simulations of multivariate time series

The work is supported by the Greek General Secretariat for Research and Technology (Aristeia II, No 4822)

Bibliography

- A. Tsimpiris, and D. Kugiumtzis, Expert Systems, "Feature Selection for Classification of Oscillating Time Series", 29, 456 – 477 (2012).
- A. Tsimpiris, I. Vlachos, and D. Kugiumtzis, "Nearest neighbor estimate of conditional mutual information in feature selection", Expert Systems with Applications, 39, 12697 – 12708 (2012).
- I. Vlachos & D. Kugiumtzis "Non-uniform state space reconstruction and coupling detection", Physical Review E, 82: 016207, (2010)
- A. Papana, D. Kugiumtzis & P.G. Larsson "Detection of direct causal efects and application in the analysis of electroencephalograms from patients with epilepsy", International Journal of Bifurcation and Chaos, 22, 1250222 (2012)
- D. Kugiumtzis, "Direct-coupling information measure from nonuniform embedding". Physical Review E, 87, (2013)
- Read More: http://www.worldscientific.com/doi/abs/10.1142/S0218127412502227
- E. S. Page. " Continuous Inspection Schemes ", Biometrika, 41(1), 100- 115, (1954)
- C. Koen and F.Lombard, "The analysis of indexed astronomical time series – IX. A period change test", MNRAS, Vol. 353 98-104, (2004)