

FESM: An Analytical Framework for Elastic Similarity Measures Based Time Series Pattern Recognition

Nafas Esmaeili
Young Researchers and Elite Club,
Qazvin Branch, Islamic
Azad University, Qazvin, Iran

Karim Faez
Department of
Electrical Engineering,
Amirkabir University of Technology,
Tehran, Iran

ABSTRACT

Compare the similarity of time series is a key for most tasks and there are various similarity measures which measure the similarity of time series. Similarity measures are the basis of time series research, they are quite important for improving the efficiency and accuracy of the time series pattern recognition tasks. Therefore selection the best similarity measures are very essential. On this issue, in this paper an analytical framework for elastic similarity measures based time series pattern recognition as, FESM for short, is proposed. FESM consists of three main components: 1) Classification of elastic similarity measures of time series, 2) Comparative evaluation of classified similarity measures based on proposed qualitative evaluation criteria, and 3) Application scopes of classified similarity measures. FESM will be proper for the quick understanding and comparing of time series similarity measures, and selection the best of existing similarity measures for respective time series pattern recognition tasks.

General Terms

Pattern Recognition

Keywords

Time Series, Pattern Recognition, Elastic Similarity Measures.

1. INTRODUCTION

Time series pattern recognition have a broad range of real-world applications such as astronomy, stock-market, phenology, video mining, energy and power, and many other application scopes. There are a huge amount of data in these scopes and the effect of similarity measures for pattern recognition in correlated data such as time series in these applications scopes is an important factor which should be considered. In astronomy [1], producing time series from millions of sky objects is done for classify stars and other phenomena into one of the known types of variability, and the huge discovery space that remains to be explored for a vast number of new unknown discoveries in astronomy domain, the similarity measures between different astronomy time series are used for their learning and pattern recognition tasks. In stock-market [2], there is the challenge of increasing continuous stock-market data with unexpected extrema over time, the similarity measure between stock-market time series data, can be used as input for hierarchical clustering algorithms in order to handle the challenge of increasing

Stock-market time series data. In plant phenology study [3], selection the best similarity measure for classification plant species and providing measures to estimate the change in phenological events, such as loss and discoloration of leaves is an important object. The video mining [4] is one of the active research areas which is related to many application domains, such as semantic indexing and retrieval to intelligent video surveillance, and various approaches to detect video events have been proposed in the literature, the objective of video mining is to discover and describe interesting patterns from the huge amount of video data. There are structural patterns in recurrence plots which can be used to determine the similarity between two video sequences, which is necessary for classification, so measuring the similarity between two recurrence plots, needed for video classification in order to event detection in video sequences. In [5] the CK-1 distance used to measure the similarity between unthresholded recurrence plots that were generated from time series and results show the combination of the CK-1 distance measure together with unthresholded recurrence plots results in higher classification accuracy for time series which represent the shape. Related to energy and power [6], check the effect of similarity measures is a necessary step for the optimized design and development of efficient clustering based models, predictors and controllers of time dependent processes such as building energy consumption patterns. Figure 1 shows the similarity measures based time series pattern recognition.

All above mentioned application scopes show that one of the main challenges of time series pattern recognition is selecting a fit measure of similarity or distance between time series [7], so selection the best similarity measure in order to improve the efficiency and accuracy of the time series pattern recognition tasks is a key and very essential.

The lock step similarity measures are simple and very intuitive for time series data, they have a known weakness of sensitivity to distortion in the time axis. These measures do not match any of the types of robustness [8] so in order to overcome this weakness many elastic measures were proposed. They can generally handle the problems of lock step similarity measures [1].

The elastic similarity measures have elastic steps in their measuring, and work with time series with different lengths, because of these flexibility, they are widely used in science,



Fig 1: Similarity measures based time series pattern recognition

medicine, industry, finance and many other areas.

On the issue of elastic similarity measure, in this paper an analytical framework for elastic similarity measure based time series pattern recognition as, FESM for short, is proposed.

The proposed framework consists of three main components: 1) Classification of elastic similarity measures of time series, 2) Comparative evaluation of classified similarity measures based on proposed qualitative evaluation criteria, and 3) Application scopes of classified similarity measures. FESM will be proper for the quick and simple understanding of time series similarity measures, and selection the best of them fit the need.

The rest of the paper is outlined as follows: Section 2 describes the related concepts, Section 3 presents our proposed framework and Section 4 is devoted to the conclusion and future directions.

2. RELATED CONCEPTS

In this section, some of the important related concepts will be covered, it will begin by introducing the key definitions as follow:

Definition 1 Time Series. A time series $T = (t_1, \dots, t_n)$, $t_i \in \mathbb{R}$ is an ordered sequence of n real-valued variables [8], it is a set of measurements arranged in order either by time or spatial location [9].

Definition 2 Subsequence. Given a time series T of length m , a subsequence S of T is a sampling of length $l \leq m$ of contiguous positions from T , that is, $S = t_p, \dots, t_{p+l-1}$, for $1 \leq p \leq m - l + 1$ [10].

Definition 3 Similarity Measure. The similarity measure $\text{Dist}(T_a, T_b)$, the distance between two time series T_a and T_b , which Dist is a function which gets two time series as inputs and returning the distance d between these two series T_a and T_b . The distance between time series needs to be carefully defined in order to reflect the underlying similarity of such data, this is particularly desirable for segmentation, classification, clustering, similarity-based retrieval and other mining procedures of time series [11].

Definition 4 Subsequence similarity measure. The subsequence similarity measure $D_{\text{subseq}}(T, S)$ is defined as $D_{\text{subseq}}(T, S) = \min(D(T, S'))$ for $S' \in S_s^{|T|}$. It represents the

distance between T and its best matching location in S [8].

Definition 5 Metric distance: The distance between two time series T_a and T_b is a metric if $\text{Dist}(T_a, T_b) > 0$ and $\text{Dist}(T_a, T_b) = \text{Dist}(T_b, T_a)$, and there is triangle inequality as $\text{Dist}(T_a, T_c) \leq \text{Dist}(T_a, T_b) + \text{Dist}(T_b, T_c)$ [1]. More formally, a metric is a function that behaves according to a specific set of conditions (non-negativity, identity of indiscernible, symmetry, triangle inequality) [11].

Definition 6 Elastic similarity measures. Elastic similarity measures are the similarity measures which have the data adaptation in conditions such as offset translation and amplitude scaling.

Definition 7 Outliers and Noise. These two definitions are different from together. Noise is random error or variance which should be removed. Outliers violate the mechanism that generates the normal data and should be detection. In order to detection the outlier firstly the noise should be removed.

Definition 8 Global constraints. All elastic similarity measures are based on dynamic programming with quadratic computational complexity, the Sakoe-Chiba band, and the Itakura parallelogram are two global constraints which use for significantly speed up the calculation of similarities and improve the accuracy of classification [20].

3. FESM: PROPOSED FRAMEWORK

In this section of the paper, an analytical framework for elastic similarity measures based pattern recognition as, FESM for short, is proposed. FESM consists of three main components: 1) Classification of elastic similarity measures of time series, 2) Comparative evaluation of classified similarity measures based on proposed qualitative evaluation criteria, and 3) Application scopes of classified similarity measures. Details will be described.

3.1 The First Component of FESM: Classification

The first component of FESM is a classification of elastic similarity measures of time series. Based on this classification there are two main approaches as: 1) Elastic similarity measures based on the L_p norms, and 2) Elastic similarity measures based on the matching threshold. These two main approaches have different types, which are shown in Figure 2. Let T_a and T_b be two time series with m and n data points respectively, by given T_a and T_b , in the following details of this classification will be described.

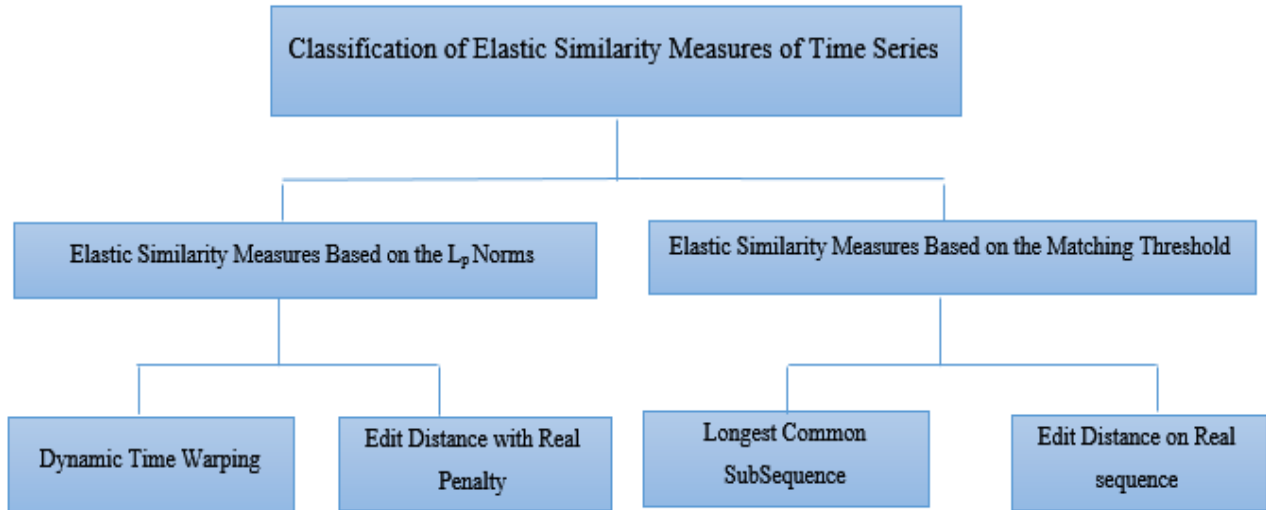


Fig 2: Classification of elastic similarity measures of time series

3.1.1 Elastic Similarity Measures Based on the L_p Norms

The first approach includes elastic similarity measures based on the L_p norms as Manhattan Distance (L_1) [12] and Euclidean Distance (L_2) [13]. This approach represents similarity measures which satisfy the triangle inequality and are metric distances. Two main types of similarity measures are considered in this approach as 1) Dynamic Time Warping (DTW) [14] and 2) Edit Distance with Real Penalty (ERP) [22]. An illustration of an elastic similarity measures based on the L_p norms is shown in Figure 3.

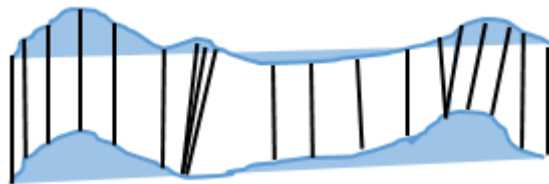


Fig 3: An illustration of an elastic similarity measures based on the L_p norms

• Dynamic Time Warping (DTW)

DTW searches for the best alignment between two time series by attempting to minimize the distance between them. DTW is a technique for effectively achieving the warping [17]. It computes the similarity of time series by finding the optimal warping path in the matrix of distances between points of the two series T_a and T_b with m and n data points respectively, DTW by zero value of the global constraint illustrate L_2 . DTW defined by equation (1) as follows [14];

$$\text{Dist}_{\text{DTW}}(T_a, T_b) = \text{DTW}(m, n) \quad (1)$$

$$\text{DTW}(i, j) = \begin{cases} 0, & \text{if } i = 0, j = 0 \\ \infty, & \text{if } i = 0 \\ \infty, & \text{if } j = 0 \\ \text{dist}(T_a[i], T_b[j]) + \min\{\text{DTW}(i-1, j-1), \text{DTW}(i-1, j), \\ \text{DTW}(i, j-1)\}, & \text{Otherwise} \end{cases}$$

• Edit Distance with Real Penalty (ERP)

ERP introduces a constant value g as the gap of the edit distance and uses L_1 distance between elements as the penalty to handle local time shifting [22]. ERP have their roots in the classic string edit distance. If the distance between two points is too large, ERP simply uses the distance value between one of those points and the reference point [7]. ERP defined by equation (2) as follows;

$$\text{ERP}(T_a, T_b) = \begin{cases} \sum_{i=1}^n \text{dist}(T_b[i], g), & \text{if } m = 0 \\ \sum_{i=1}^m \text{dist}(T_a[i], g), & \text{if } n = 0 \\ \min\{\text{ERP}(\text{Rest}(T_a), \text{Rest}(T_b)) + \text{dist}(T_a[1], T_b[1]), \\ \text{ERP}(\text{Rest}(T_a), T_b) + \text{dist}(T_a[1], g) \\ \text{ERP}(T_a, \text{Rest}(T_b)) + \text{dist}(T_b[1], g), \\ \text{Otherwise} \end{cases} \quad (2)$$

Where $\text{dist}(a, b)$ is the distance between two elements and g is a gap of edit distance. The real distance between elements ($\text{dist}(T_a[1], T_b[1])$) as the penalty used to handle local time shifting.

3.1.2 Elastic Similarity Measures Based on the Matching Threshold

The second approach includes elastic similarity measure based on the matching threshold. This approach represents similarity measures which cannot satisfy the triangle inequality and are not metric distances. Two main types of similarity measures are considered in this approach as: 1) Longest Common SubSequence (LCSS) [19] and 2) and the Edit Distance on Real Sequence (EDR) [23]. An illustration of an elastic similarity measures based on the on the matching threshold is shown in Figure 4.

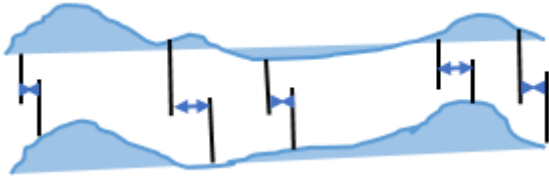


Fig 4: An illustration of an elastic similarity measures based on the matching threshold

- **Longest Common SubSequence (LCSS)**

In LCSS, the similarity between two time series is expressed as the length of the longest common subsequence of both time series [19], to adapt the concept of matching characters in the settings of time series, a threshold parameter ϵ was introduced, stating that two points from two time series are considered to match if their distance is less than ϵ [18]. LCSS uses a threshold parameter ϵ for point matching and a warping threshold δ . It is defined by equation (3) as follows [19, 20];

$$\text{Dist}_{\text{LCSS}}(T_a, T_b) = 1 - \frac{\text{LCSS}(i, j)}{\min(n, m)} \quad (3)$$

$$\text{LCSS}(i, j) = \begin{cases} 0, & \text{if } i = 0 \\ 0, & \text{if } j = 0 \\ 1 + \text{LCSS}(i - 1, j - 1), & \text{if } |T_a[i] - T_b[i]| \leq \epsilon \\ \max\{\text{LCSS}(i - 1, j), \text{LCSS}(i, j - 1)\}, & \text{Otherwise} \end{cases}$$

- **Edit Distance on Real sequence (EDR)**

EDR is another elastic similarity measure based on the matching threshold, it uses a threshold parameter and assigns penalties to gaps between two similar subsequences of time series according to the lengths of the gaps. In EDR, the matching threshold reduces effects of noise by quantizing the distance between a pair of elements to two values, 0 and 1, and reduce the effect of outliers. It's defined by equation (4) as follows [23];

$$\text{EDR}(T_a, T_b) = \begin{cases} n, & \text{if } m = 0 \\ m, & \text{if } n = 0 \\ \text{EDR}(i - 1, j - 1), & \text{if } |T_a[i] - T_b[i]| \leq \epsilon \\ \min\{\text{EDR}(\text{Rest}(T_a), \text{Rest}(T_b)) + \text{subcost}, \text{EDR}(\text{Rest}(T_a), T_b) + 1, \\ \text{EDR}(T_a, \text{Rest}(T_b)) + 1\}, & \text{Otherwise} \end{cases} \quad (4)$$

Where $\text{subcost} = 0$, if $\text{match}(T_a[1], T_b[1]) = \text{true}$ and $\text{subcost} = 1$, otherwise. $\text{Rest}(T)$ stands for the time series obtained from T by eliminating the first element.

3.2 The Second Component of FESM: Comparative Evaluation of Classified Elastic Similarity Measures Based on Proposed Qualitative Evaluation Criteria

This section describes the second component of FESM, which will propose the comparative evaluation of classified similarity measures based on qualitative evaluation criteria. Firstly the qualitative evaluation criteria with their ranking

will be proposed and then analyse details of comparative evaluation of classified similarity measures based on proposed qualitative evaluation criteria will be described. The results are shown in Table 1.

3.2.1 Proposed Qualitative Evaluation Criteria

- **Metric:** If the intended similarity measure satisfies the triangle inequality it is a metric, else it is nonmetric. Ranking is Yes and No.
- **Noise handling:** Refers to robust of the intended similarity measure in handling noise. Ranking is Yes and No.
- **Outliers handling:** Refers to robust of the intended similarity measure in handling outliers. Ranking is Yes and No.
- **Performance:** Refers to compare the classifier average error rates with different elastic similarity measures of our proposed framework. It is based on statistically significant wins and losses counts for the INN classifier with different similarity measures, across many different data sets using the Wilcoxon sign-rank test [20]. Ranking is High, Low and Average.
- **Global constraints effects:** Refers to effects of global constraints on different elastic similarity measures. Some elastic similarity measure are most sensitive to the global constraints so have maximum effect, while others have different behaves sensitively and effects. Ranking is Maximum, Medium and Minimum.
- **Data points mapping:** Refers to allowing the feasibility of mapping of the data points. Ranking is one-to-one, one-to-many and one-to-none.

3.2.2 Comparative Evaluation of Classified Elastic Similarity Measures Based on Proposed Qualitative Evaluation Criteria

- **Dynamic Time Warping**

DTW is an elastic metric similarity measure which can give a satisfactory result on the different length of both input time series with short lengths. In DTW, all elements from both sequences must be used, even the outliers [17], so it is sensitive to the noise and outlier and cannot handle both of them. Experimental results in [20] show that DTW has the most sensitive to the Sakoe-Chiba band global constraints regarding the INN graph, and has the best performance in all elastic similarity measure. This distance measures allow comparison of one-to-one and one-to-many points.

- **Edit Distance with Real Penalty**

ERP is an elastic metric similarity measure. It is sensitive to the noise and cannot handle noises, but it is robust to outliers. ERP measure has intermedia behavior to the Sakoe-Chiba band global constraints regarding the INN graph, and its performance as the generally is worst [20]. This distance measures allow comparison of one-to-one, one-to-many and one-to-none points.

- **Longest Common SubSequence**

LCSS is an elastic nonmetric similarity measure with subsequence matching which solves the problem of the presence of noise by taking into account only sufficiently similar points [19], also some elements such as outliers may

be unmatched or left out [17], so LCSS can handle both of the noise and outliers. Experimental results in [20] show that LCSS measure have intermedia behavior to the application of the Sakoe-Chiba band, and its performance is average. LCSS can compare the one-to-one, one-to-many and one-to-none points.

- Edit Distance on Real sequence

EDR is elastic nonmetric similarity measure, same to LCSS both of the noise and outliers can be handled by the threshold setting. The application of the Sakoe-Chiba band exerts the lowest influence on EDR, it has average error rate across the different data sets so has average performance [20]. EDR can compare the one-to-one, one-to-many and one-to-none points.

Table 1. Comparative evaluation of classified similarity measures based on evaluation criteria

		Metric Distance	Noise Handling	Outliers Handling	Global Constraints Effects	Performance	Data Points Mapping	
FESM	Similarity measures based on the L_p norms	Dynamic Time Warping	Yes	No	No	Maximum	High	one-to-one one-to-many
		Edit Distance with Real Penalty	Yes	No	Yes	Medium	Low	one-to-one one-to-many one-to-none
	Similarity measures based on the matching threshold	Longest Common SubSequence	No	Yes	Yes	Medium	Average	one-to-one one-to-many one-to-none
		Edit Distance on Real sequence	No	Yes	Yes	Minimum	Average	one-to-one one-to-many one-to-none

3.3 The Third Component of FESM: Application Scopes of the Classified Similarity Measure

This section describes the third component of FESM, there are several time series applications in diverse domains which increasingly mining with different similarity measures. Here the proper similarity measures which applied in some of these applications will be shown, these applications come from different domains. A good choice of similarity measure has strong results in these applications, the analysis details are as follows;

3.3.1 Application Scope of the Elastic Similarity Measures Based on the L_p Norms: Trajectory Analysis

The similarity measures in the first approach of proposed classifications have the metric characteristic, as mentioned in section 3.2 the metric distance criteria satisfies the triangle inequality. The triangle inequality is an efficient way to apply pruning strategies [22]. The trajectory data have numerous potential applications in traffic control, flight data, urban planning, astronomy, transportation systems and animal science. For large trajectory databases, it is important to

minimize the computation of the distance between trajectories in the database [27].

- **Application 1: DTW in flight systems**

DTW is one of the most commonly used distance measures for trajectory data. The basic idea behind DTW distance is to find out the warping path between two trajectories that minimizes the warping cost [26].

An automated technique for clustering flight and trajectory data using a Particle Swarm Optimization (PSO) approach has been proposed in [15], which considered the DTW distance as one of the most commonly used metric distance measures for flight trajectory data, the proposed technique is able to find (near) optimal number of clusters as well as (near) optimal cluster centers during the clustering process and reduce the dimensionality of the search space and improve the performance.

- **Application 2: ERP in transportation systems**

Early detection of the outlier is critical for monitoring and managing the condition of railway transportation systems. Railway transportation requires railway systems to be fault tolerant and safe. Failures of railway point systems often lead to service delays or hazardous situations. The condition monitoring system used to detect the early signs of the deteriorated condition of railway point systems for investigation of anomalies and prevent failures.

A methodology for early warning of possible point failure through early detection of changes in the current drawn by the point motor proposed in [28], which use the one class support vector machine classification method with the similarity measure of ERP taking into account specific features of the data in railway transportation systems field. It is able to detect the changes in the measurements of the current of the point operating equipment with greater accuracy compared with the commonly used threshold based technique.

3.3.2 Application Scope of the Elastic Similarity Measures Based on the Matching Threshold: Image and Video Analysis

The image and video analysis include a closely related broad fields such as image pattern recognition, image and video retrieval and authentication, image and video classification, video surveillance, video event recognition, video tracking, computer vision, machine vision, etc.

The similarity measures in the second approach of proposed classifications have threshold setting characteristic which can handle noise and outliers properly. One of a major challenge in these fields is the detection of abnormal events and noise handling, so the second approach similarity measures of our framework classifications can be used on this issue.

- **Application 1: LCSS in video surveillance**

Detection of abnormal events in video surveillance systems is based on the analysis of the trajectories of moving objects in a controlled scene. The comparative experimental results in [25] demonstrate that the LCSS is the most accurate and efficient for the clustering task even in the case of different sampling rates and noise in the four distances widely used as trajectories' similarity measure. Based on their generic adopted process for an abnormal event detection, firstly normal/abnormal clusters from saved trajectories through an unsupervised clustering algorithm has been extracted and then, a new detected trajectory considered and classify as either normal or abnormal.

The similarity between trajectories is a critical step while analyzing trajectories since it affects the quality of further applications such as clustering and classification.

- **Application 2: EDR in computer vision**

There is a need for simulation of computer vision systems applied to crowd monitoring. Simulation the most important aspects of crowds for performance analysis of computer based video surveillance systems is done in [29].

The optimized crowd simulation algorithms can be utilized in computer vision research, for example, provide predictions of pedestrian locations in multi people tracking tasks, in [24] refine crowd simulation algorithms by optimizing their parameters based on EDR has been done and the result demonstrates that this approach significantly reduces the distance between the simulated trajectories of individuals and the trajectories extracted from real video.

4. CONCLUSION AND FUTURE DIRECTIONS

In this paper, an analytical framework for elastic similarity measures based time series pattern recognition as, FESM for short, was developed. FESM consists of three main components: 1) Classification of elastic similarity measures of time series, 2) Comparative evaluation of classified similarity measures based on proposed qualitative evaluation criteria, and 3) Application scopes of classified similarity measures. Two main approaches were presented in the first component as: 1) Elastic similarity measures based on the L_p norms, and 2) Elastic similarity measures based on the matching threshold. These two main approaches had different types, which were described in details. The second components of FESM was done the comparative evaluation of classified similarity measures based on proposed qualitative evaluation criteria and finally in the third component, the application scopes of classified similarity measures were demonstrated.

The outcomes of this research will help to researchers for the quick understanding and compare of time series similarity measures, and selection the best of existing similarity measures fits the need.

Many researchers have proposed hybrid methods which use multiple similarity measures. Hybrid methods have improved the performance of existing similarity measures in various application scopes. In the future work, a framework for novel hybrid similarity measures which have proposed in recent years, will be proposed.

5. REFERENCES

- [1] Lin, J., Williamson, S., Borne, K., DeBarr, D. 2012. Pattern recognition in time series. *Advances in Machine Learning and Data Mining for Astronomy*. Eds. Kamal, A., Srivastava, A., Way, M., and Scargle, J. Chapman & Hall. To Appear.
- [2] Cerioli, A., Laurini, F., and Corbellini, A. 2005. Functional cluster analysis of financial time series. In *New Developments in Classification and Data Analysis*, eds. Vichi, M., Monari, P., Mignani, S., Montanari, A. Springer-Verlag, Berlin, pages 333-341.
- [3] Conti, J.C., Fariol, F.A., Almeida, J., Alberton, B., Morellato, P. C. L., Camolesi, L., Torres, R. 2014. Evaluation of Time Series Distance Functions in the Task of Detecting Remote Phenology Patterns. *ICPR 2014*: 3126-3131

- [4] Vijayakumar, V., Nedunchezian, R. 2012. A study on video data mining. *IJMIR* 1(3): 153-172
- [5] Silva, D. F., Souza, V., De, M., Batista, G. E. 2013. Time series classification using compression distance of recurrence plots. In *Data Mining (ICDM), 2013 IEEE 13th International Conference on*, pages 687–696. IEEE.
- [6] Iglesias, F., Kastner, W. 2013. Analysis of Similarity Measures in Times Series Clustering for the Discovery of Building Energy Patterns, *Energies*, vol. 6, no. 2, pp. 579–597.
- [7] Wang, X., Mueen, A., Ding, H., Trajcevski, G., Scheuermann, P., Keogh, E. 2012. Experimental comparison of representation methods and distance measures for time series data. *Data Min. Knowl. Disc.*
- [8] Esling, P., Agón, C. 2012. Time-series data mining. *ACM Comput. Surv.* 45(1):12.
- [9] Shams, MB., Haji, S., Salman, A., Abdali, H., Alsaffar, A. 2016. Time series analysis of Bahrain's first hybrid renewable energy system. *Energy* 103, 1-15.
- [10] Ye, L., Keogh, E. 2011. Time series shapelets: a novel technique that allows accurate, interpretable and fast classification *Data Min Knowl Disc.* Volume 22, Issue 1, pp 149–182.
- [11] Spiegel, S. 2015. Time series distance measures. Segmentation, classification, and clustering of temporal data. Doctoral Thesis. <http://dx.doi.org/10.14279/depositonce-4619>.
- [12] Yi, B.-K. Faloutsos, C. 2000. Fast time sequence indexing for arbitrary lp norms. *International Conference on Very Large Data Bases*, pp. 385–394.
- [13] Faloutsos, C., Ranganathan, M., Manolopoulos, Y.(1994). Fast subsequence matching in time-series databases. In *proceedings of the ACM SIGMOD Int'l Conference on Management of Data*.pp 419-429.
- [14] Berndt, D. J., Clifford, J. 1994. Using Dynamic Time Warping to Find Patterns in Time Series, in *KDD Workshop*, 1994, pp. 359–370.
- [15] Izakian, Z., Mesgari, MS., Abraham, A. 2016. Clustering of trajectory data using a particle swarm optimization *Computers, Environment and Urban Systems-Elsevier*. Volume 55. Pages 55–65.
- [16] Keogh, E. 2002. Exact Indexing of Dynamic Time Warping. In *Proceedings of the 28th international Conference on Very Large Data Bases*. Hong Kong, China.
- [17] Ratanamahatana, CA., Lin, J., Gunopulos, D., Keogh, E., Vlachos M, Das G. 2005. Mining time series data. *Data Mining and Knowledge Discovery Handbook*, Maimon O, Rokach L (eds.).Springer: Berlin.ISBN: 978-0-387-24435-8.
- [18] Ding, H., Trajcevski, G., Scheuermann, P., Wang, X., & Keogh, E. 2008. Querying and mining of time series data: Experimental comparison of representations and distance measures. *Proceeding of the VLDB Endowment*, 1(2), 1542–1552.
- [19] Vlachos, M., Kollios, G., Gunopulos, D. 2002. Discovering similar multidimensional trajectories. In: *Proceedings 18th International Conference on Data Engineering*. IEEE Comput. Soc, pp. 673–684.
- [20] Kurbalija, V., Radovanović, M., Geler, Z., Ivanović, M. 2014. The influence of global constraints on similarity measures for time-series databases. *Knowledge-Based Systems* 56, 49–67.
- [21] Vlachos, M., Hadjieleftheriou, M., Gunopulos, D., Keogh, E.J. 2006. Indexing Multidimensional Time-Series. *VLDB J.*, 15(1).
- [22] Chen, L., Ng, R. 2004. On the marriage of lp-norms and edit distance. pp. 792–803.
- [23] Chen, L., Oszu, M.T., Oria, V. 2005. Robust and fast similarity search for moving object trajectories. In: *Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data*. ACM, New York, NY, USA, pp. 491–502.
- [24] Jin, Z., Bhanu, B. 2013. Optimizing crowd simulation based on real video data, in: *IEEE International Conference on Image Processing (ICIP)*, Melbourne, VIC, Australia, pp. 3186–3190. doi:10.1109/ICIP.2013.6738656.
- [25] Bouarada Ghrab, N., Fendri, E., Hammami, M. 2016. Clustering Based Abnormal Event Detection: Experimental Comparison for Similarity Measures' Efficiency. *ICIAR* .367-374.
- [26] Zhang, Z., Huang, K., Tan, T. 2006. Comparison of similarity measures for trajectory clustering in outdoor surveillance scenes. *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*.1135-1138
- [27] Abdelbar, M., Buehrer R.M. 2016. Improving Cellular Positioning Indoors Through Trajectory Matching. *Conference:IEEE/ION Position, Location and Navigation Symposium-DOI:10.1109/PLANS.2016.7479705*
- [28] Vileiniskis, M., Remenyte-PreScott, R., Rama, D. 2015. A fault detection method for railway point systems. *Proc. Inst. Mech. Eng. F J. Rail Rapid Transit*.
- [29] Andrade, E.L., Fisher, R.B. 2005. Simulation of crowd problems for computer vision. In *Proceedings First International Workshop on Crowd Simulation*, vol. 3. pp.7