Measurement 130 (2018) 290–297

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Scale-varying dynamic time warping based on hesitant fuzzy sets for multivariate time series classification

Shuai Liu[a,b,*], Changliang Liu[a,b]

[a] School of Control and Computer Engineering, North China Electric Power University, Changping District, Beijing, China
[b] State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Changping District, Beijing, China

ARTICLE INFO

Keywords:
Multivariate time series (MTS)
Dynamic time warping (DTW)
Hesitant fuzzy sets
Scale-varying classification

ABSTRACT

Multivariate time series (MTS) data widely exists in daily life. How to classify MTS remains a major problem in data mining, computer science, financial area and other relative industry. MTS data is always treated as a whole object or time instance one by one, while in this paper, time instance segments were paid more attention. A new distance measure named dynamic time warping based on hesitant fuzzy sets (HFS-DTW) is proposed for MTS classification. HFS-DTW is a generalized dynamic time warping algorithm, and due to the characteristic of HFS, it is easy to find optimal alignment between time instance segments. Also, the proposed method could be reduced to original DTW by setting scale parameters. In order to apply the proposed algorithm correctly and efficiently, the parameter constraints were discussed. Furthermore, using 10-fold cross-validation, five MTS data sets selected from the University of California, Irvine machine learning repository, were tested by the proposed algorithm. By comparing with state-of-the-art algorithms, the results demonstrate the proposed method could balance the higher accuracy and lower time-consuming in classification.

1. Introduction

The time series contains time-varying variates. Since the 1940s, the technologies of computer chips, internet of things and sensors make significant development. Nowadays, Multivariate time series (MTS) is widely spread among the process of production, monitoring, and transaction. It is various, like thermal power station [1], movement tracking [2], handwritten signature verification [3], stock trading [4], etc. MTS has many variates and there are inner relations between the variates. The traditional algorithm for data mining should be upgraded to meet the challenge from a new type of data.

Researches in multivariate time series classification become more and more popular during the past decades. It is concluded that there are two basic ways to classify the MTS. They are classification based on time instance distance and classification based on time series similarity. Two-dimensional singular value decomposition (2dSVD) took the shape and the inner relation of MTS into consideration [5]. Using echo state network and adaptive differential evolution algorithm, 18 MTS data sets were tested to prove the proposed approach effective and stable [6]. A new MTS classification named temporal discrete SVM (TDVM) was put forward, based on a temporal extension of discrete support vector machines. It could serve a better accuracy among other against algorithm [7]. Stacking which is an ensemble method was studied to process the MTS. The MTS data was decomposed into univariates. And after level 0 classifier, all univariates were collected to get a final classification. Then benchmark test showed that it worked well [8]. Glowacz investigated several feature extraction methods and different classifiers, keeping focus on the protection of DC motor. The method of selection of amplitudes of frequencies (MSAF-1, MSAF-5) [9,10], SMOFS-25-EXPANDED [11], Fast Fourier Transform (FFT) [9,10], and Linear Discriminant Analysis (LDA) [10,11], K-Nearest Neighbour (KNN) [12], SOM (Self-organizing Map) [11] and other classifiers were tested to show the efficiency and accuracy of the proposed framework. Furthermore, there are methods based on similarity attempt to solve the problem of MTS classification. A new DTW-based Bayesian approach was proposed to deal with sleep and wake classification, and a case study was presented to show the advantage of the proposed method [13]. Luczak had put much effort to classify MTS. Parametric derivative dynamic time warping (DD_{DTW}) [14], parametric integral dynamic time warping (ID_{DTW}) [15], and parametric combined dynamic time warping distance measure (combDTW) [16] proved better performance against the original dynamic time warping (DTW).

Dynamic time warping is a type of elastic algorithm which overcomes the weakness of Euclidean distance. Finding an optimal alignment between the two series is the main function of DTW. Euclidean distance measures the distance of raw data with the same time position,

* Corresponding author at: School of Control and Computer Engineering, North China Electric Power University, Changping District, Beijing, China.
E-mail address: liushuai@ncepu.edu.cn (S. Liu).

https://doi.org/10.1016/j.measurement.2018.07.094
Received 9 May 2018; Received in revised form 30 July 2018; Accepted 31 July 2018
Available online 02 August 2018

sequentially. With DTW, one data point could match more than one data points to minimize the cost function and find the optimal warping path. With this advantage, different length data series could be measured by DTW. DTW has a huge number of application scenes, such as speech recognition [17], hand-drawn symbol recognition [2], etc.

In Ref. [18], a modified algorithm named multi-dimensional DTW based on DTW was proposed, by extending standard DTW of univariate time series into multivariate time series area. And Euclidean distance was applied to compute the local distance between two multivariate time series. Combining the metric learning algorithm large margin nearest neighbor (LMNN) and DTW, LMNN–DTW was introduced [19]. The Ref. [20] proposed a form of DTW called Derivative DTW (DDTW). The proposed algorithm measured the local derivatives of the data to find the correct warping path. Mahalanobis distance based DTW (MDDTW) was introduced to enhance the accuracy of MTS classification [21]. And a LogDet divergence based metric learning with triplet constraint (LDMLT) model is established. We call it LDMLT-MDDTW method in here. There are some practical methods to improve the efficiency of DTW, like SparseDTW [22] or FTW [23].

What's more, the core part of data mining for MTS is distance measure (metric learning). Different distance measure has different features which may be suitable for some type of data. Usually, Euclidean distance is selected as the default distance measure in most popular algorithms. But for each MTS data set, the structure and characteristic are unique, so distance metric should be various. For example, the distance function used by original dynamic time warping (DTW) [24] algorithm is Euclidean distance, but it could be replaced by Mahalanobis distance to get a lower classification error rate [21]. The hesitant fuzzy set is a new kind of distance metric, which has rarely research on MTS classification.

As a new extension of fuzzy sets, hesitant fuzzy sets (HFSs) [25] have been introduced to deal with the uncertainty. HFS has attracted a lot of attention all over the world, because of efficiency in representing uncertainty. The definition, distance, similarity, and correlation for HFSs were shown in Refs. [26,27]. In Ref. [28], the notions of hesitant fuzzy entropy and cross-entropy to tackle with the multi-attribute decision making were defined. The researchers investigated the multiple criteria group decision making (MCGDM) problems with hesitant fuzzy information. And they studied a model based on traditional grey relational analysis (GRA) method, with the help of HFSs [29]. In Ref. [30], a new formula of correlation coefficients between HFSs was established, extending the interval of correlation coefficient from $[0, 1]$ to $[-1, 1]$. The idea of generalized hesitant fuzzy synergetic weighted distance measures was proposed in Ref. [31]. And in Ref. [32], the new formula of similarity and entropy of hesitant fuzzy sets were defined, by summarising the relationship of distance, similarity, and entropy based on hesitant fuzzy sets. The clustering algorithms for clustering the hesitant fuzzy data and two methods of application in the data mining task were discussed in Ref. [33].

From the above literature, the improved DTW algorithm focuses on changing the structure of DTW to achieve better classification accuracy. Also, the combinations of DTW with other algorithm have to learn new properties to process the different type of data. However, the distance metric still lacks in-depth study. On the other hand, the HFSs are influential in system evaluation, but there is no effective application of HFSs in classification.

In this paper, we propose an improved DTW algorithm based on hesitant fuzzy sets (HFSs) to tackle with the classification of multivariate time series. In our new method, the membership degree concept from HFS is brought into MTS classification as a new distance metric. Time instance segments alignment was studied, which is rarely mentioned before. And for reaching lower error rate in MTS classification, scale constraints were built to find the most effective scale parameters. At last, comparisons against state-of-the-art algorithm were discussed, a higher accuracy in most cases and lower time-consuming with suitable scale parameters were achieved.

In the experiments, the nearest neighbor classification (1NN) was deployed to test the proposed method. 1NN classifier and DTW had been shown one of the best combinations for univariate time series classification [34], and state-of-the-art algorithms aiming at MTS classification selected the same combination [14–16,19–21]. So we tested HFS-DTW under equally experiment environment. Also, 10-fold cross-validation technique is required to eliminate the impact of potential unbalance data.

The remained of the paper is organized as follows. Section 2 introduces the basic notion and background information. In Section 3, an improved DTW algorithm based on hesitant fuzzy sets (HFS-DTW) is presented. Then experiments and experimental results are given to demonstrate the effectiveness of the proposed algorithm in Section 4. Also, suggestions for the proposed algorithm application are made. At last, the paper is concluded in Section 5.

2. Related work

We will discuss the DTW and the notion of hesitant fuzzy sets in this section.

2.1. Original dynamic time warping

The DTW is built to solve the univariate time series optimal alignment problem, and Fig. 1 shows the ability and performance of DTW algorithm. The warping path between two time series is completed when the minimum normalized cumulative distance has been found, and there are three basic procedures of DTW algorithm.

Suppose that distance between two univariate time series A and B ($A = a_1, a_2, ...a_i, ...a_I$, $B = b_1, b_2, ...b_j, ...b_J$) is calculated by DTW algorithm. $c(k) = (i(k), j(k))$ represents a warping path through a_i and b_j, then the normalized cumulative distance describes as:

$$D(A, B) = \mathrm{Min}\left[\frac{\sum_{k=1}^{K} d(c(k))\cdot w(k)}{\sum_{k=1}^{K} w(k)}\right] \tag{1}$$

Note: $d(c(k))$ is an operator to generate the distance between the element $a_{i(k)}$ and $b_{j(k)}$, and there are usually several operator options. $w(k)$ describes the weight of the warping path, and K has a constraint: $\max(I, J) \leqslant K < I + J$.

DTW must satisfy three basic constraints:

(1) Monotonicity: any two adjacent elements of the warping path must

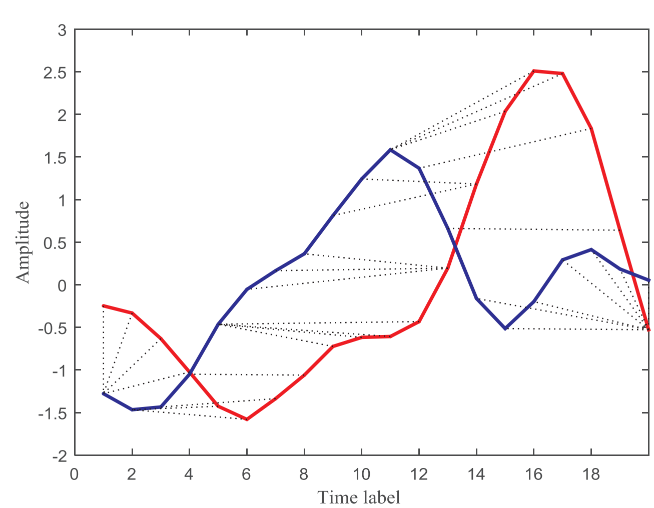

Fig. 1. The utility of DTW for two univariate time series (Red line and blue line denote time series, the dotted line connects the data points which have minimum distance). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)