Pattern Recognition 44 (2011) 2231–2240

Contents lists available at ScienceDirect

Pattern Recognition

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

Weighted dynamic time warping for time series classification

Young-Seon Jeong [a], Myong K. Jeong [a,b,c,*], Olufemi A. Omitaomu [d]

[a] Department of Industrial and Systems Engineering, Rutgers University, Piscataway, NJ, USA
[b] Rutgers Center for Operations Research, Rutgers University, Piscataway, NJ, USA
[c] Department of Industrial and Systems Engineering, KAIST, Daejon, Korea
[d] Geographic Information Science & Technology Group, Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

ARTICLE INFO

Article history:
Received 20 November 2009
Received in revised form
29 August 2010
Accepted 30 September 2010
Available online 14 October 2010

Keywords:
Dynamic time warping
Adaptive weights
Weighted dynamic time warping
Modified logistic weight function
Time series classification
Time series clustering

ABSTRACT

Dynamic time warping (DTW), which finds the minimum path by providing non-linear alignments between two time series, has been widely used as a distance measure for time series classification and clustering. However, DTW does not account for the relative importance regarding the phase difference between a reference point and a testing point. This may lead to misclassification especially in applications where the shape similarity between two sequences is a major consideration for an accurate recognition. Therefore, we propose a novel distance measure, called a weighted DTW (WDTW), which is a penalty-based DTW. Our approach penalizes points with higher phase difference between a reference point and a testing point in order to prevent minimum distance distortion caused by outliers. The rationale underlying the proposed distance measure is demonstrated with some illustrative examples. A new weight function, called the modified logistic weight function (MLWF), is also proposed to systematically assign weights as a function of the phase difference between a reference point and a testing point. By applying different weights to adjacent points, the proposed algorithm can enhance the detection of similarity between two time series. We show that some popular distance measures such as DTW and Euclidean distance are special cases of our proposed WDTW measure. We extend the proposed idea to other variants of DTW such as derivative dynamic time warping (DDTW) and propose the weighted version of DDTW. We have compared the performances of our proposed procedures with other popular approaches using public data sets available through the UCR Time Series Data Mining Archive for both time series classification and clustering problems. The experimental results indicate that the proposed approaches can achieve improved accuracy for time series classification and clustering problems.

1. Introduction

There has been a long-standing interest for time series classification and clustering in diverse applications such as pattern recognition, signal processing, biology, aerospace, finance, medicine, and meteorology [1,2,8,12,14,18,23,25,26], and thus some notable techniques have been developed including nearest neighbor classifier with a given distance measure, support vector machines, and neural networks [2,4,20]. The nearest neighbor classifiers with dynamic time warping (DTW) has shown to be effective for time series classification and clustering because of its non-linear mappings capability [7,18,25]. The DTW technique finds an optimal match between two sequences by allowing a non-linear mapping of one sequence to another, and minimizing the distance between two sequences [8,7,12,22]. The sequences are "warped" non-linearly to determine their similarity independent of any non-linear variations in the time dimension. The technique was originally developed for speech recognition, but several researchers have evaluated its application in other domains and have developed several variants such as derivative DTW (DDTW) [11,21,22]. Fig. 1 shows the example of process of aligning two out of phase sequences by DTW.

The methodology for DTW is as follows. Assume a sequence A of length m, $A=a_1, a_2, \ldots, a_i, \ldots, a_m$ and a sequence B of length n, $B=b_1, b_2, \ldots, b_j, \ldots, b_n$. We create an m-by-n path matrix where the (ith, jth) element of matrix contains the distance between the two points a_i and b_j such that $d(a_i,b_j)=||(a_i-b_j)||_p$, where $||\cdot||_p$ represents the l_p norm. The warping path is typically subject to several constraints such as [22]

Endpoint constraint: the starting and ending points of warping path have to be the first and the last points of the path matrix, that is, $u_1=(a_1, b_1)$ and $u_k=(a_m, b_n)$.
Continuity constraint: the path can advance one step at a time. That is, when $u_k=(a_i, b_j)$, $u_{k+1}=(a_{i+1}, b_{j+1})$ where $a_i-a_{i+1} \le 1$ and $b_i-b_{i+1} \le 1$.
Monotonicity: the path does not decrease, i.e., $u_k=(a_i, b_j)$, $u_{k+1}=(a_{i+1}, b_{j+1})$ where $a_i \ge a_{i+1}$ and $b_i \ge b_{i+1}$.

* Corresponding author at: Department of Industrial and Systems Engineering, Rutgers University, 640 Bartholomew Road-Room 115, Piscataway, NJ 08854, USA. Tel.: +1 732 445 4858; fax: +1 732 445 5472.
E-mail address: mjeong@rci.rutgers.edu (M.K. Jeong).

0031-3203/$ - see front matter © 2011 Published by Elsevier Ltd.
doi:10.1016/j.patcog.2010.09.022

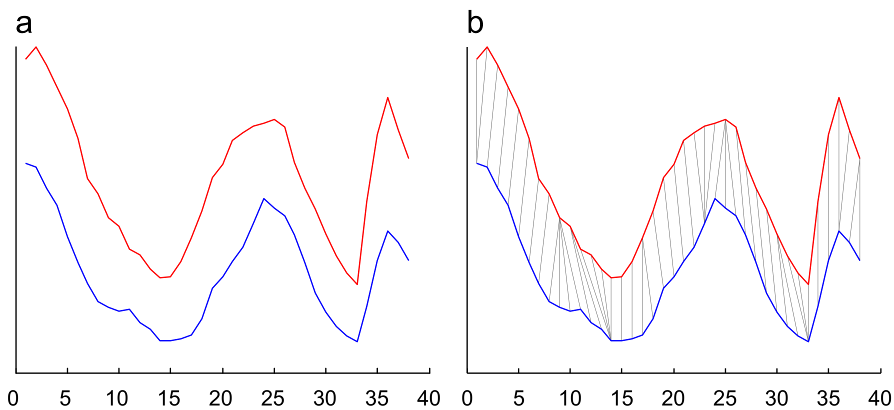

Fig. 1. Alignment of sequences based on DTW: (a) two similar sequences, but out of phase and (b) alignment by DTW.

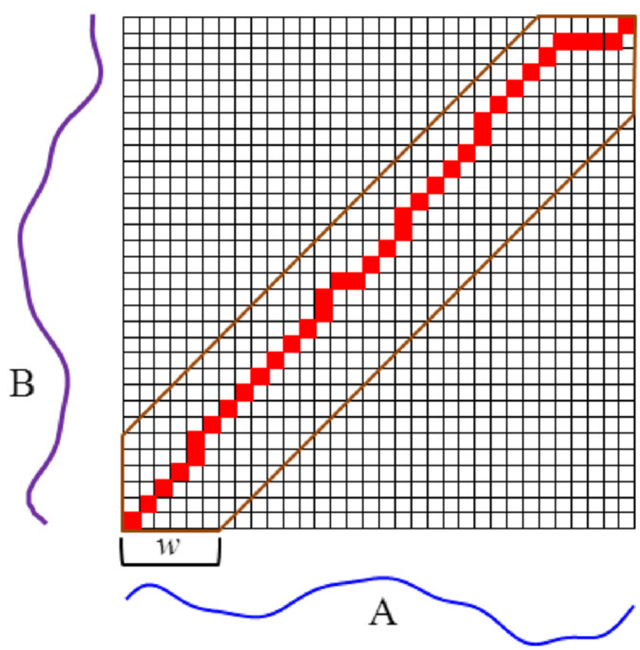

Fig. 2. Warping matrix and optimal warping path by DTW.

The best match between two sequences is the one with the lowest distance path after aligning one sequence to the other. Therefore, the optimal warping path can be found by using recursive formula given by

$$\mathrm{DTW}_p(A,B) = \sqrt[p]{\gamma(i,j)}$$

where $\gamma(i,j)$ is the cumulative distance described by

$$\gamma(i,j) = |a_i - b_j|^p + \min\{\gamma(i-1,j-1), \gamma(i-1,j), \gamma(i,j-1)\} \tag{1}$$

As seen from Eq. (1), given a search space defined by two time series DTW_p guarantees to find the warping path with the minimum cumulative distance among all possible warping paths that are valid in the search space. Thus, DTW_p can be seen as the minimization of warped l_p distance with time complexity of $O(mn)$. By restraining a search space using constraint techniques such as Sakoe–Chuba Band [22] and Itakura Parallelogram [7], the time complexity of DTW can be reduced. Fig. 2 shows the warping matrix and optimal warping path between two sequences by DTW. In Fig. 2, a band with width w is used to constrain the warping.

However, the conventional DTW calculates the distance of all points between two series with equal weight of each point regardless of the phase difference between a reference point and a testing point. This may lead to misclassification especially in applications such as image retrieval where the shape similarity between two sequences is a major consideration for an accurate recognition, thus neighboring points between two sequences are more important than others. In other words, relative significance depending on the phase difference between points should be considered.

Therefore, this paper proposes a novel distance measure, called the weighted dynamic time warping (WDTW), which weights nearer neighbors more heavily depending on the phase difference between a reference point and a testing point. Because WDTW takes into consideration the relative importance of the phase difference between two points, this approach can prevent a point in a sequence from mapping the further points in another one and reduce unexpected singularities, which are alignments between a point of a series with multiple points of the other series. Some practical examples will be presented to graphically illustrate possible situations where WDTW clearly is a better approach.

In addition, a new weight function, called the modified logistic weight function (MLWF), is proposed to assign weights as a function of the phase difference between a reference point and a testing point. The proposed weight function extends the properties of logistic function to enhance the flexibility of setting bounds on weights. By applying different weights to adjacent points, the proposed algorithm can enhance the detection of similarity between series.

Finally, we extend the proposed idea to other variants of DTW such as derivative dynamic time warping (DDTW) and propose the weighted version of DDTW (WDDTW). We compare the performances of our proposed procedures with other popular approaches using public data sets available through UCR Time Series Data Mining Archive [13] for both time series classification and clustering problems. The experimental results show that the proposed procedures achieve improved accuracy for time series classification and clustering problems.

This remainder of the paper is organized as follows. In Section 2, we review some related literatures on times series classification and its methodologies. Section 3 explains the rationale of the advantage of the proposed idea. In Section 4, we describe the proposed WDTW and the modified logistic weight function for automatic time series classifications. The experimental results are presented and discussed in Section 5. The paper ends with concluding remarks and future works in Section 6.