Journal of Computational and Applied Mathematics 319 (2017) 514–520

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Adaptive cost dynamic time warping distance in time series analysis for classification

Yuan Wan, Xiao-Li Chen *, Ying Shi

School of Science, Wuhan University of Technology, 430070, Wuhan, China

HIGHLIGHTS

- Adaptive cost is introduced to dynamic time warping for better classification.
- Cost function is proposed for dynamic time warping.
- Experiments on UCR datasets prove that AC-DTW perform better than other methods.

ARTICLE INFO

Article history:
Received 4 September 2016
Received in revised form 20 October 2016

Keywords:
Time series classification
Dynamic time warping
Adaptive cost
Nearest neighbor classifier

ABSTRACT

Dynamic time warping (DTW) distance is commonly used in measuring similarity between time series for classification. In order to obtain the minimum cumulative distance, however, DTW distance may map multiple points on one time series to one point on another, and this makes time series over stretched and compressed, resulting in missing important feature information thus influence the classification accuracy. In this paper, we propose a method called adaptive cost dynamic time warping distance (AC-DTW), which adjusts the number of points on one time series mapped to the points on another. AC-DTW records the trajectories of all points and then adaptively allocates the cost rate to each point by calculating cost function at the next step. The results of the experiments implemented on 17 UCR datasets by using nearest neighbor classifier demonstrate that AC-DTW prevails in criterion of higher accuracy rate in comparison with some existing methods.

1. Introduction

The time series classification is an important topic in time series analysis and it has been applied in many areas, such as finance, biometrics, networking, artificial intelligence, etc. [1–4]. Given a time series sequence p, the goal of classification is to classify p in a dataset D into a prepared class. In this process, one of the important step is to calculate the distance between p and other time series in dataset D. To measure the distance between two time series, Euclidean distance and its variants are used most often. Agrawal et al. [5] proposed using Discrete Fourier Transform and Struzik et al. [6] proposed using Discrete Wavelet Transform. Although the improved Euclidean distance methods can translate the amplitude of time series, they failed to extend and warp the time axis. Dynamic time warping (DTW) is nowadays widely-used for measuring the distance, which was originally used in the field of text data matching and pattern recognition. Due to the elastic metric, it is introduced into the time series mining and classification by Berndt and Clifford [7], they use the dynamic programming [8] to align sequences with different lengths. While applying the DTW distance, the situation may occur that most points in a

* Corresponding author.
E-mail addresses: wanyuan@whut.edu.cn (Y. Wan), 785535947@qq.com (X.-L. Chen), 1329990417@qq.com (Y. Shi).

http://dx.doi.org/10.1016/j.cam.2017.01.004

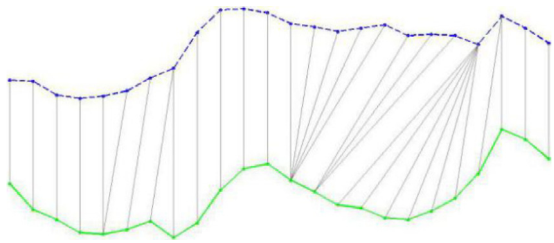

Fig. 1. The map of two time series under DTW distance.

time series are mapped to only a few points on another, and this makes the time series over stretched and compressed, leading to a singularity [9].

To avoid such extreme situation, the optimal warping path is limited to a search window, by the restriction method "Search window". Ratanamahatana and Keogh [10] proposed the adaptive window restriction method, knowing that the walking range of the optimum warping path is limited. It is most likely to miss the correct dynamic time warping distance in this case. Sakoe and Chiba [11] proposed slope constraint DTW, this method corrected the direction and step size of the optimal bending path, but it might destroy the continuity of the optimal warping path. Also, the selection of parameters is a difficult problem as we know already. Keogh and Pazzani [12] introduced the Derivative Dynamic Time Warping (DDTW), this method achieved a new sequence from the derivative of the original time series, then by applying the DTW distance on the new sequence, which can capture information about shape, but the step to find the new sequence will increase the computational complexity. Jeong et al. [13] proposed a weighted DTW (WDTW). Because of the DTW, it does not take relative significance depending on the phase difference between points into consideration. The proposed technique penalizes points with higher phase difference, between a reference point and a testing point to prevent minimum distance distortion resulted through outliers, but how to determine the weight depending on the different cases is difficult. Due to the high computational cost of DTW as the lengths of the signals increase, Barbon et al. [14] introduced an optimized version of the DTW that is based on the Discrete Wavelet Transform. And Salvador and Chan [15] proposed FastDTW, an approximation of DTW that has a linear time and space complexity.

In this paper, we propose an adaptive cost dynamic time warping distance (AC-DTW), which takes into consideration the excessive stretching or compression of the dynamic time warping distance. When choosing the optimal path, we limit the excessive tension or compression by giving the current step base distance an appropriate weight (greater than one), AC-DTW can avoid such an extreme situation that too many points in one sequence are mapped to only a few points in the other sequence.

The remainder of this paper is organized as follows. In Section 2, we introduce the background of the DTW method. In the following section, we describe the proposed method, which is called the AC-DTW, including the cost function, adaptive cost dynamic time warping distance, and together with the algorithm. Experiments and the performance are given in Section 4 and finally, we present our conclusion and the perspective of this work.

2. Original DTW

The dynamic time warping distance is now widely used in time series similarity measurement, which searches for the flexible corresponding relation between two time series. Fig. 1 gives a map of two time series under the DTW distance.

Consider two time series sequences $P = \{p_1, p_2, \ldots, p_n\}$ and $Q = \{q_1, q_2, \ldots, q_m\}$ of length n and m, we construct an $n \times m$ matrix, called distance matrix $D_{base} = (d_{base}(p_i, q_j))_{n\times m}$, where the entry $d_{base}(p_i, q_j)$ represents the base distance between points p_i and q_j. Then, we search an optimal warping path $R = \{r_1, \ldots, r_k, \ldots, r_K\}$, which minimizes the accumulation of corresponding distance, where $\max(m, n) \le K < m + n - 1$, Fig. 2 shows the optimal warping path between P and Q in the distance matrix.

The warping path is typically subjected to the following constraints.

(1) Boundary conditions: the first and the last points in one time series should map to the other series, i.e. $r_1 = D(1, 1)$, $r_K = D(n, m)$;
(2) Monotonicity: for $r_k = D(i_k, j_k)$ and $r_{k+1} = D(i_{k+1}, j_{k+1})$, $i_{k+1} \ge i_k$ and $j_{k+1} \ge j_k$;
(3) Continuity: $i_{k+1} \le i_k + 1$ and $j_{k+1} \le j_k + 1$.

So, the dynamic time warping distance is obtained by the minimum warping path distance. Therefore, the optimal warping path formula can be described as follows

$$DTW(P, Q) = \min_{W} \left[\sum_{k=1}^{K} d_{base}(r_k) \right] \tag{1}$$