



Time Alignment Measurement for Time Series

Duarte Folgado^{a,*}, Marília Barandas^a, Ricardo Matias^{b,c}, Rodrigo Martins^b, Miguel Carvalho^d, Hugo Gamboa^e

^a Associação Fraunhofer Portugal Research, Rua Alfredo Allen 455/461, Porto, Portugal

^b Physiotherapy Department, School of Health, Polytechnic Institute of Setúbal, Estefanilha, Edifício da ESCE, 2914-503, Setúbal, Portugal

^c Champalimaud Research, Champalimaud Centre for the Unknown, Lisbon, Portugal

^d Minho University, Campus de Azurém, 4800-058, Guimarães, Portugal

^e Laboratório de Instrumentação, Engenharia Biomédica e Física da Radiação (LIBPhys-UNL), Departamento de Física, Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

ARTICLE INFO

Article history:

Received 26 July 2017

Revised 3 January 2018

Accepted 2 April 2018

Available online 3 April 2018

Keywords:

Time series

Time warping

Similarity

Distance

Signal alignment

ABSTRACT

When a comparison between time series is required, measurement functions provide meaningful scores to characterize similarity between sequences. Quite often, time series appear warped in time, i.e., although they may exhibit amplitude and shape similarity, they appear dephased in time. The most common algorithm to overcome this challenge is the Dynamic Time Warping, which aligns each sequence prior establishing distance measurements. However, Dynamic Time Warping takes only into account amplitude similarity. A distance which characterizes the degree of time warping between two sequences can deliver new insights for applications where the timing factor is essential, such well-defined movements during sports or rehabilitation exercises. We propose a novel measurement called Time Alignment Measurement, which delivers similarity information on the temporal domain. We demonstrate the potential of our approach in measuring performance of time series alignment methodologies and in the characterization of synthetic and real time series data acquired during human movement.

© 2018 Published by Elsevier Ltd.

1. Introduction

The comparison of time series have existed in the scenario of sequence matching, subsequence searching, and motif detection. Those challenges are intrinsically related to time series classification applied in several contexts such as pattern recognition [1–4], signal processing [5], shape detection [6], bioinformatics [7,8], human activity recognition [9] and on-line handwritten signature validation [10].

When a comparison of two streams of data with implicit or explicit time information associated is executed, there is the need for a measurement function that provides information on the similarity of the two data streams. Time series comparison may be established using a wide range of available distance measurement functions. Some of the traditional metrics, such the Euclidean distance or some modification thereof, assume that the discrete signals are equidistant points in time and also aligned in the time axis. In some domains, although time series may present amplitude and shape similarity, they can be considered to be out-of-phase. Therefore, similar regions may appear in different instants in time, lead-

ing to different degrees of time distortion, or time warping, among several sequences, since they are not aligned in the temporal domain. In those circumstances, traditional distances fail to measure this distortion since they are very sensitive to small distortions in time and typically unable to directly handle unequal length time series without some sort of preprocessing [11].

In order to overcome these limitations, elastic distances which contemplate temporal elastic shifting have been proposed. Dynamic Time Warping (DTW) and Longest Common Subsequence (LCSS) compensate non-linear temporal distortions by aligning the discrete sequences before establishing amplitude measurements in the discrete domain [12]. Since those algorithms do not take into account the information between inter-sampling points, [13] proposed the Continuous Dynamic Time Warping (CDTW), which extends the classic methodology by allowing mapping between instants that may eventually not belong to the original time vector for each series. The work from [14] uses an optimization approach to calculate a parametric polynomial warping path reflecting the alignment between both series. Therefore, the last two alternatives produce an optimal warping path which translates the alignment between two signals in the continuous domain.

Motivated by the fact that off-the-shelf applications of semi-supervised learning algorithms do not typically work well

* Corresponding author.

E-mail address: duarte.folgado@fraunhofer.pt (D. Folgado).

when applied to time series, the authors from [15] proposed a new distance which tries to minimize this behaviour. The proposed distance is called Dynamic Time Warping Delta (DTW-D) and is the ratio between DTW and Euclidean distances.

Inspired by the well-known *edit distance* for string comparison, which calculates the minimum number of insertions, deletions, and substitution operations to transform a string in another, some authors translated the core idea to time series [16–18]. In order to generalize the concept from strings to time series, two elements of each sequence are matched if the absolute difference between them is bellow a given tolerance value. The common goal of the approaches consists in identifying the smallest number of operations (additions, deletions and substitutions) to transform a sequence in another.

Prior to establish a similarity measurement between time series, most of the aforementioned examples perform a previous alignment between the two sequences. The optimal alignment may also be used for summarizing a set of time series, since it allows to compute a more meaningful average between sequences which may exhibit time warping. The work developed by [19], and more recently by [20], proposes time series averaging methods based on preceding alignments, which demonstrated favourable impacts on clustering performance.

However, whilst we observed a multitude of proposed novel elastic distances over the last years, they are mostly centered in measuring similarity accounting for amplitude differences [21,22]. Those facts motivated our work in the development of a novel time distance able to measure similarity between time series in the temporal domain, namely Time Alignment Measurement (TAM). The proposed methodology is able to describe the behaviour in time between two signals by measuring the fraction of time distortion between them. The distortion may comprise periods of temporal advance or periods of delay. When signals are similar-alike in time they can be considered to be in phase between each other. This approach can deliver useful information to domains where information between the temporal misalignment of time series is needed. Examples of such domains include well-defined human movements executed in sports or rehabilitation exercises. The authors from [23] investigated the feasibility of biofeedback training applied to therapeutic exercises, where repetitive movements should follow well-defined timings to be considered successfully executed. The authors calculated the mean error of the distance between anatomic segments executed by the subject to a previously recorded reference. A distance able to truly characterize temporal misalignment between movements should bring new perspectives for the evaluation of the correctness of the exercises through the complete movement execution.

The literature review allowed to identify that most of the work developed over the last years in the development of new distance functions mostly takes into account amplitude similarity. The major contribution presented on this work is propose a novel distance which measures similarity in time domain.

The remaining content of this paper is organized as follows: in section 2, a brief overview of DTW algorithm is presented, since we use DTW to align two time series prior calculating TAM. Section 3 introduces the TAM distance and presents examples based on synthetic time series to support its potential. In section 4 we present two use cases for the proposed distance based on real time series data. Finally, section 5 contains the conclusions and future work directions.

2. Time series alignment

In this section, we motivate for the utility of DTW algorithm to establish an alignment between two time series in order to calculate TAM. We start with a brief explanation of DTW algorithm and explore some of the challenges arising while aligning signals that present amplitude fluctuation.

2.1. Dynamic Time Warping

The DTW algorithm allows two time-dependent sequences that are similar, but locally out of phase, to align in time. Its main objective consist of identifying an optimal alignment between sequences by warping the time axis iteratively.

In order to align two time series $X := (x_1, x_2, \dots, x_N)$ and $Y := (y_1, y_2, \dots, y_M)$ of length N and M respectively, a N -by- M cost matrix is computed. Each (n^{th}, m^{th}) element of the cost matrix, $C \in \mathbb{R}^{N \times M}$, corresponds to the distance between each pair of elements of the sequences X and Y . The Euclidean distance is usually employed as a distance function to define the cost matrix element as:

$$c(x_n, y_m) = (x_n - y_m)^2 \quad (1)$$

The goal of DTW is to find the optimal warping alignment path between X and Y having minimum overall cost. A warping path, W , is a set of matrix elements that define the relationship between X and Y . The k^{th} element of W is defined as $w_k = (i, j)_k$, $w_k \in \mathbb{R}^2$:

$$W = (w_1, w_2, \dots, w_k, \dots, w_K) \quad \max(N, M) \leq K \leq N + M - 1 \quad (2)$$

The resulted path should be composed by a set of matrix elements satisfying the following conditions:

- **Boundary condition:** Enforces that the first and the last elements of X and Y are aligned to each other $\therefore w_1 = (1, 1)$ and $w_K = (N, M)$.
- **Monotonicity condition:** Forces the points in the warping path to be monotonically spaced in time $\therefore i_1 \leq i_2 \leq \dots \leq i_N$ and $j_1 \leq j_2 \leq \dots \leq j_M$.
- **Step size condition:** Avoids omissions in elements and replications in the alignment of X and Y $\therefore (w_{k+1} - w_k) \in (1, 0), (0, 1), (1, 1)$ for $k \in [1 : K - 1]$.

The optimal warping path is the path that has the minimum total cost among all possible warping paths. One could test every incumbent warping path and determine the minimum cost candidate, but such method will lead to a exponential computational complexity in the lengths of N and M . Using dynamic programming, an accumulated cost matrix, D , is computed in order to find the path that minimizes the warping cost in an $O(N, M)$ complexity [12]. Each accumulated cost matrix element is defined as the local cost measure in the current cell plus the minimum of the local cost measures in the adjacent cells:

$$D(n, m) = \min\{D(n-1, m-1), D(n-1, m), D(n, m-1)\} + c(x_n, y_m) \quad (3)$$

where $n \in [1 : N]$, $m \in [1 : M]$, D is the accumulated cost matrix, and $c(x_n, y_n)$ is the local cost measure found in the current cell.

Using this accumulated matrix, the optimal warping path, $W^* = (w_1, w_2, \dots, w_K)$, is computed in reverse order of indices, starting with $w_K = (N, M)$, by the following algorithm:

$$w_{k-1} = \begin{cases} (1, m-1), & \text{if } n = 1 \\ (n-1, 1), & \text{if } m = 1 \\ \operatorname{argmin}\{D(n-1, m-1), D(n-1, m), D(n, m-1)\}, & \text{otherwise} \end{cases} \quad (4)$$

Download English Version:

<https://daneshyari.com/en/article/6938856>

Download Persian Version:

<https://daneshyari.com/article/6938856>

[Daneshyari.com](https://daneshyari.com)