Accepted Manuscript

E-Embed: A Time Series Visualization Framework based on Earth Mover's Distance

Bingkun Chen, Hong Zhou, Xiaojun Chen

PII: S1045-926X(18)30121-6
DOI: https://doi.org/10.1016/j.jvlc.2018.08.002
Reference: YJVLC 850

To appear in: *Journal of Visual Languages and Computing*

Received date: 20 July 2018
Accepted date: 6 August 2018

Please cite this article as: Bingkun Chen, Hong Zhou, Xiaojun Chen, E-Embed: A Time Series Visualization Framework based on Earth Mover's Distance, *Journal of Visual Languages and Computing* (2018), doi: https://doi.org/10.1016/j.jvlc.2018.08.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

E-Embed: A Time Series Visualization Framework based on Earth Mover's Distance

Bingkun Chen, Hong Zhou*, Xiaojun Chen

College of Computer Science and Software Engineering, Shenzhen University, China

Abstract

Time series analysis is an important topic in machine learning and a suitable visualization method can be used to facilitate the work of data mining. In this paper, we propose E-Embed: a novel framework to visualize time series data by projecting them into a low-dimensional space while capturing the underlying data structure. In the E-Embed framework, we use discrete distributions to model time series and measure the distances between time series by earth mover's distance (EMD). After calculating the distances between time series, we can visualize the data by dimensionality reduction algorithms. In order to better combine with different dimensionality reduction methods (such as Isomap) that depend on K-nearest neighbor (KNN) graph, we propose an algorithm for constructing a KNN graph based on the earth mover's distance. We evaluate our visualization framework on both univariate time series data and multivariate time series data. Experimental results demonstrate that the E-Embed can provide high quality visualization with low computational cost.

Keywords: Time series, visualization, EMD, dimensionality reduction, similarity measures

1. Introduction

A time series data consists of a sequence of data points listed in time order and can reflect the temporal changes of various properties of complex systems. Time series analysis plays an important role in many fields, such as finance, motion capture, business, meteorology, biology, etc. However, Due to the characteristics of high-dimensional and dynamic changes in time series data, the analysis of time series data faces great challenges. To solve these challenges, scholars have done a lot of research on machine learning, data mining and data visualization. Data visualization is a tool to reveal the potential patterns and trends of data in a visual way, we can use data visualization to facilitate data analysis.

In this article, we propose E-Embed: A framework for visualizing time-series data in low-dimensional spaces that has the ability to expose the relationships and structures between time-series data. E-Embed can be used to visualize univariate time series data (UTS) and multivariate time series data (MTS). Firstly, we propose a time series similarity measure based on earth mover's distance (EMD) [1]. Experiments show that using EMD to measure distance between time series has good performance, and the speed is superior to the traditional methods such as DTW and LCSS. After measuring the pairwise distance between time series data with EMD, we use dimensionality reduction algorithm to project the time series into a low-dimensional space. In the E-Embed framework, in order to better integrate with difference dimensionality reduction algorithms that rely on KNN graphs, we propose a KNN graph construction algorithm based on EMD. To the best of our knowledge, E-Embed is the first method to combine EMD with time series visualization and various dimensionality reduction algorithms.

The outline of this paper is as follows. In Section 2, we mainly introduce the work related to time series visualization. In Section 3, we discuss the concept of earth mover's distance, which can be used to measure the distance

*Corresponding author

Email addresses: chenbingkun03@163.com (Bingkun Chen), hzhou@szu.edu.cn (Hong Zhou), xjchen@szu.edu.cn (Xiaojun Chen)

Download English Version:

https://daneshyari.com/en/article/9952371

Download Persian Version:

https://daneshyari.com/article/9952371

Daneshyari.com