Neurocomputing 99 (2013) 496–508

Contents lists available at SciVerse ScienceDirect

Neurocomputing

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

Self-organizing time map: An abstraction of temporal multivariate patterns

Peter Sarlin *

Department of Information Technologies, Turku Centre for Computer Science, Åbo Akademi University, Joukahaisenkatu 3-5, 20520 Turku, Finland

ARTICLE INFO

Article history:
Received 8 November 2011
Received in revised form
3 July 2012
Accepted 13 July 2012
Communicated by B. Hammer
Available online 10 August 2012

Keywords:
Self-organizing time map
Self-organizing map
Exploratory temporal structure analysis
Dynamic visual clustering
Exploratory data analysis

ABSTRACT

This paper adopts and adapts Kohonen's standard self-organizing map (SOM) for exploratory temporal structure analysis. The self-organizing time map (SOTM) implements SOM-type learning to one-dimensional arrays for individual time units, preserves the orientation with short-term memory and arranges the arrays in an ascending order of time. The two-dimensional representation of the SOTM attempts thus twofold topology preservation, where the horizontal direction preserves time topology and the vertical direction data topology. This enables discovering the occurrence and exploring the properties of temporal structural changes in data. For representing qualities and properties of SOTMs, we adapt measures and visualizations from the standard SOM paradigm, as well as introduce a measure of temporal structural changes. The functioning of the SOTM, and its visualizations and quality and property measures, are illustrated on artificial toy data. The usefulness of the SOTM in a real-world setting is shown on poverty, welfare and development indicators.

1. Introduction

During an era of increasing access to complex datasets, abstraction of multivariate temporal patterns is a central issue. However, exploring and extracting patterns in high-dimensional panel data, i.e., along multivariate, temporal and cross-sectional dimensions, is a demanding task. While exploratory data analysis commonly concerns either individual univariate and multivariate time-series or static cross-sectional analysis, a question of central importance is how to combine these tasks. That is, how to identify the occurrence and explore the properties of temporal structural changes in data, as well as their specific locations in the cross section. This type of exploratory data analysis will in the sequel be referred to as exploratory temporal structure analysis.

Kohonen's [1,2] self-organizing map (SOM) is an effective general-purpose tool for abstraction of multivariate mean profiles through projection into a lower dimension. The SOM differs from standard methods for exploratory data analysis by at the same time performing a clustering via vector quantization and projection via neighborhood preservation, as well as by possessing the advantages of a regular grid shape for linking visualizations and a simple and fast learning algorithm. While exploratory analysis with the SOM mainly concerns cross-sectional applications, it is a common tool for classification, clustering and prediction of time-dependent data in a wide range of domains, such as engineering, geographical and environmental sciences, economics and finance (see e.g., [2–4]). The main rationale for using the SOM over more traditional methods for time-series prediction is the inherent local modeling property and topology preservation of units that enhances interpretability of dynamics as well as the availability of growing architectures that facilitate the choice of parsimony (for a thorough review see [5]).

For exploratory analysis on multivariate panel data, however, it is critical to visualize, or present an abstraction across, all dimensions (i.e., multivariate, temporal and cross-sectional spaces). Using a standard two-dimensional SOM for exploratory temporal structure analysis, processing of the time dimension has thus far been proposed along two suboptimal directions: computing separate maps per time unit (e.g., [6–8]) or one map on pooled panel data (e.g., [9–11]). Owing to a possibly high number of time units and temporal differences in correlations and distributions, comparing separate maps per time unit is a laborious task while their structure may not in the least even be comparable. However, SOMs trained with pooled data, for which time can be inferred as a type of latent dimension that is definable but unordered, fail in describing the structure in each cross section. The literature has provided several improvements to the SOM paradigm for temporal processing. We reduce these into four groups: (1) those implicitly introducing time in pre- or post-processing (e.g., trajectories [12]), (2) adaptations of the standard activation and learning rule (e.g., the Hypermap [13]), (3) adaptations of the standard network topology through feedback connections and hierarchical layers (e.g., Temporal SOM [14]) and (4) combinations with other visualization techniques (e.g., interactive spatiotemporal visualization systems [15,16]). Yet, the problem of visualizing changes in inherent data structures over time has not been

* Tel.: +358 2 215 4670.
E-mail address: psarlin@abo.fi

0925-2312/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2012.07.011

entirely addressed. The existing SOM literature has thus shortcomings in disentangling the temporal dimensions and cross-sectional structures for exploratory temporal structure analysis, which is the main focus of this paper.

In this paper, we propose a self-organizing time map (SOTM) for abstraction of the structure in temporal multivariate patterns. In general, the processing of the SOTM is depicted by standard SOM-type learning to one-dimensional arrays for individual time units. We attempt to preserve a stable orientation of the SOTM over time with an initialization based upon short-term memory. When arranging the one-dimensional arrays in an ascending order of time, the SOTM enables a two-dimensional representation with multivariate data structures on the vertical direction and the temporal dimension on the horizontal direction. This output can, when combined with visual aids, be used for dynamic visual cluster analysis, where local distances between SOTM units can be treated as cluster structure information across both directions (i.e., identification of changing, emerging and lost clusters). An ordered SOTM can also be used for projecting individual or grouped data onto the map (constrained by the units of its own time unit). The projections, in conjunction with the structure of the SOTM, enable a temporal version of Bertin's [17] three "levels of reading": elementary level (a view of single multivariate time series), intermediate level (a view of groups of multivariate time series) and global level (a view of temporal multivariate data structures).

For measuring qualities and properties of SOTMs, we adapt several measures and visualizations from the standard SOM paradigm. We also propose a measure for indicating the degree of temporal structural changes in data. A limitation of this work is the absence of a quantitative evaluation, such as commonly performed prediction comparisons to alternative methods. This is, however, due to the lack of a comparable evaluation function. Instead, we illustrate the functioning, output and usefulness of the SOTM on an artificial toy dataset with expected patterns. The generated toy data exhibit multivariate clustered patterns along cross-sectional and temporal dimensions. In addition, we also illustrate drawbacks of a naïve SOM model on these data and provide a guide for interpreting patterns on the SOTM. We also illustrate a real-world application of the SOTM on a temporal multivariate dataset of development and welfare indicators with patterns over the past two decades. The indicators illustrate the progress in fulfilling the Millennium Development Goals (MDGs)-eight goals representing commitments to reduce poverty and hunger and to tackle ill-health, gender inequality, environmental degradation as well as lack of education and access to clean water.

The paper is structured as follows. Section 2 gives an overview of related literature concerning temporal processing with the SOM and attempts to reduce it into four groups. In Section 3, the functioning, visualization and quality and property measures of the SOTM are described. Section 4 illustrates the usefulness of the SOTM, its visualizations and its quality and property measures, as well as a guide for interpreting them, on two datasets: artificial toy data and indicators of the progress towards the MDGs. Section 5 concludes by presenting our key findings and directions for future research.

2. Related work

There is a wide range of literature adapting and extending the standard SOM for temporal processing. While the literature on time in SOMs has been thoroughly reviewed in [6,18–21], a unanimous classification dividing it into distinct groups of studies is far from clear-cut. Drawing upon the above reviews, we attempt to reduce the literature related to the SOTM into four groups of works: those with an implicit consideration of time, those adapting the learning or activation rule, those adapting the topology, and those combining SOMs with other visualization techniques.

The first group applies the standard Kohonen SOM algorithm and illustrates the temporal dimension either as a pre- or post-processing step. The pre-processing concerns embedding a time series into one input vector, such as tapped delay (e.g., [22]). A time-related visualization through post-processing is, however, more common. A connected time series of best-matching units (BMUs), i.e., a trajectory, has been used in the literature to illustrate temporal transitions (e.g., [12,23]). By exploiting the topological ordering of the SOM, visualization of the current and past states enables visual tracking of the process dynamics. In [24,25], the trajectory approach has been extended to show membership degrees of each time-series point to each cluster. However, while temporal patterns require large datasets for generalization and significance, trajectories can only be visualized for a limited set of data. Thus, strengths and actual directions of the patterns can be obtained by probabilistic modeling of state transitions on the SOM (see e.g., [26,27]).

The second group of works adapts the standard SOM activation or learning rule. Those decomposing the learning rule of the standard SOM into two parts, past and future, for time-series prediction have their basis in the Hypermap [13]. The past part is used for finding best-matching units (BMUs), while the entire input vector is used within the updates of the reference vectors. For predicting out-of-sample data, the past part is again used for finding BMUs while the future part of that unit is the predicted value. This type of learning has been used for standard time-series prediction (see e.g., [28,29]) and predictions through non-linear regression (see e.g., [10,11]). The latter type of decomposition can still be divided into supervised and semi-supervised SOMs, where the difference depends on whether [10] or not [11] the present part is used for matching in training. Instead of considering the context explicitly in SOM training, it can be treated as the neighborhood of the previous BMU. Kangas [30], for instance, constrains the choice of a BMU to the neighborhood of the previous BMU and thus has a behavior that resembles the functioning of SOMs with feedback in the next group.

The third group deals with adaptations of the standard SOM network topology through feedback connections and hierarchical layers. The feedback SOMs have their basis in the seminal Temporal SOM (TSOM) [14] that performs leaky integration to the outputs of the SOM. The recurrent SOM (RSOM) [31,32] differs by moving the leaky integration from the output units to the input vectors. A recent recurrent model is the Merge SOM (MSOM) [33] whose context combines the current pattern with the past by a merged form of the properties of the BMU. The recursive SOM (RecSOM) [34] keeps information by considering the previous activation of the SOM as part of the input to the next time unit, while the feedback SOM (FSOM) [35] differs by integrating an additional leaky loop onto itself. The SOM for structured data (SOMSD) labels, on the other hand, directed acyclic graphs to regular [36] and arbitrary [37] grid structures. Finally, Hammer et al. [38] define a general formal framework and show that a large number of SOMs with feedback can be recovered as special cases of it. The hierarchical network architectures, on the other hand, use at each layer one or more SOMs operating at different time scales. The next level in the hierarchy can either use the lower level SOM as input vectors without any processing, such as two-level clustering commonly does, or use transformed input vectors by computing distances between units or concatenating a time series to one input vector, for instance. Kangas [22] introduced hierarchical network architectures to SOMs, and shows that a hierarchical SOM without any additional