2024 IEEE 40th International Conference on Data Engineering Workshops (ICDEW) | 979-8-3503-8403-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICDEW61823.2024.00020

# Beyond the Dimensions: A Structured Evaluation of Multivariate Time Series Distance Measures

Jens E. d'Hondt
*Eindhoven University of Technology*
Eindhoven, the Netherlands
j.e.d.hondt@tue.nl

Odysseas Papapetrou
*Eindhoven University of Technology*
Eindhoven, the Netherlands
o.papapetrou@tue.nl

John Paparrizos
*The Ohio State University*
Columbus, OH, USA
paparrizos.1@osu.edu

***Abstract*—A variety of distance measures for multivariate time series has been proposed in recent literature. However, evaluations of such measures have been incomplete; comparisons are limited to subsets of similar measures, lacking a holistic view of the field with an appropriate taxonomy of measures. This paper presents a structured evaluation of multivariate time series distance measures. Through a novel taxonomy, measures are categorized based on how they handle the multiple variates; in an atomic or a holistic manner. Experimental evaluation of 12 measures shows that no single measure or approach is superior; the optimal choice depends on the data and the task at hand.**

***Index Terms*—Multivariate Time Series, Distance Measures**

## I. Introduction

With the rapid development of sensor technology, the analysis of time series, and, particularly, of *multivariate time series* **(MTS)**, are becoming increasingly common [13]–[16], [21], [22], [24], [26], [33], [36], [37], [39]. Multivariate time series are collections of time series, measuring different signals at the same resolution, usually sourcing from the same physical object or process [47]. These different signals are referred to as the *dimensions* or *variates* of the MTS. Examples of MTS include data from climate sensor arrays (e.g., an array measuring the temperature, humidity, and air pressure at a certain location) [20] and motion capture data (e.g., the position and acceleration of different body parts) [1], [7], [18]. An open problem in the field of MTS processing is the selection of appropriate distance measures for efficient and effective similarity search [28], [35], [38],[1] a core subroutine in many downstream tasks such as classification [8], [29], [40], [42], [43], [47], clustering [3], [30], [31], and anomaly detection [4]–[6], [9], [10], [27], [32], [44], [45]. Existing MTS distance measures can be classified into two categories; *element-wise* and *cross-wise* distance measures. Element-wise measures are extensions of measures on univariate time series **(UTS)**, and compare MTS by aggregating the per-variate distances (i.e., comparing *identical* variates of *different* MTS). In contrast, cross-wise measures are specialized for MTS, and compare MTS by considering the internal distances between the variates within each MTS (i.e., comparing the MTS *as a whole*). A simple example of an element-wise distance is the sum of per-variate Euclidean distances (i.e., $L_2$) between two MTS [42]. An example of a cross-wise measure is the sum of squared differences of the correlation matrices of two MTS. Figure 1 illustrates the difference between both approaches; the MTS in the first column of plots are highly similar in terms of their $L_2$ distances, whereas the MTS in the second column are similar in terms of their correlation structure, captured by the squared differences, but not in terms of their $L_2$ distances. The MTS in the third column are not similar in terms of either measure. While both approaches were studied in isolation, it is unclear how they relate to one another, both theoretically and empirically. Also, evaluation of cross-wise measures has been limited to performance comparisons with a few other measures over limited datasets. With the wide range of measures serving different purposes, it is important to establish general guidelines for selecting and using MTS distance measures, such that practitioners can make informed decisions.

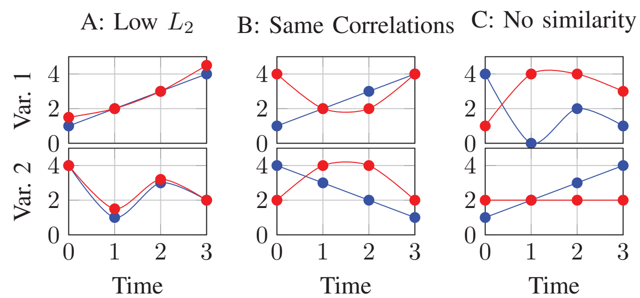


Fig. 1: Examples of pairs of MTS (the red and the blue) with different types of similarity, left to right columns: (a) the $L_2$ distances between the two variates of the two MTS is small, (b) for both MTS, the correlation between their two variates is -1, (c) no clear similarity exists between the MTS.

**Scope.** This paper focuses on the *quality* and theoretical properties of MTS distance measures. The *efficiency* of measures is not considered due to the lack of optimized similarity search solutions for many measures. Our future work will include a thorough evaluation of efficiency to provide a complete review of MTS distance measures.

**Contributions:** (a) We propose a novel taxonomy for MTS

This work has received funding from the European Union's Horizon Europe research and innovation programme STELAR under grant agreement No. 101070122.

[1]Note that similarity is the inverse of distance. These concepts can be used interchangeably in the context of similarity search

979-8-3503-8403-1/24/$31.00 ©2024 IEEE
DOI 10.1109/ICDEW61823.2024.00020

distance measures, which is based on the distinction between *element-wise* and *cross-wise* measures, the constraints on time alignment over the variates, and the support of varying variates between MTS. (b) We provide a comprehensive overview of the state-of-the-art in MTS distance measures, and use them to populate the taxonomy. Additionally, two novel measures are proposed that address some limitations of current approaches. (c) We evaluate the distance measures on a variety of datasets, and conclude that no single measure is universally the best. While the latter insight is not new, we show that it remains true for a larger and more diverse set of measures than previously considered. Besides that, we provide conceptual guidelines on the most suitable measure types for different data characteristics.

**Outline** The remainder of this paper is structured as follows. In Section II we discuss the preliminaries and related work. In Section III we motivate and introduce the taxonomy for MTS distance measures, and demonstrate its relevance through real-world examples. In Section IV we populate the taxonomy with the state-of-the-art and novel measures, and in Section V we present a thorough experimental comparison of the measures. We summarize the work in Section VI.

## II. Preliminaries

### A. Notation

A multivariate time series $\boldsymbol{T}$ of length $m$ with $v$ variates is denoted as a matrix $\boldsymbol{T} = [\boldsymbol{T}_1, ..., \boldsymbol{T}_v]$, where $\boldsymbol{T}_i$ is the UTS with $m$ observations of variate $i$. An MTS distance measure $d$ is a function $d : \mathbb{R}^{m\times v} \times \mathbb{R}^{m\times v} \rightarrow \mathbb{R}$, where $d(\boldsymbol{A}, \boldsymbol{B})$ is the distance between time series $\boldsymbol{A}$ and $\boldsymbol{B}$.

### B. Related work

We will now discuss recent experimental studies of time series distance measures. Relevant works on individual distance measures are discussed separately in Section IV, when the measures are introduced.

Recently, Paparrizos et al. [34], [35] performed a detailed experimental study of UTS distance measures, focusing on the quality of measures rather than their efficiency. In their work, the authors debunk 4 misconceptions about univariate measures. Namely, they show that (a) lesser-known normalization methods outperform z-normalization on multiple distance measures, (b) alternative lock-step measures such as Manhattan outperform the widely used Euclidean distance, (c) sliding measures, such as cross-correlation, are not dominated by elastic measures such as DTW, and (d) DTW is not always the best elastic measure; measures such as Merge-Split-Move outperform DTW on multiple datasets. While their work only considers univariate measures, it provides a solid framework for evaluating time series distances, and many univariate measures can be extended to MTS. The first comparison of MTS distance measures was by Shokoohi-Yekta et al. in [43]. The authors defined two approaches for extending DTW to the multivariate case; through *dependent* (*DTW-D*) or *independent* (*DTW-I*) time warping. With *DTW-D*, the warping path is shared between all variates, while with *DTW-I*, each variate has

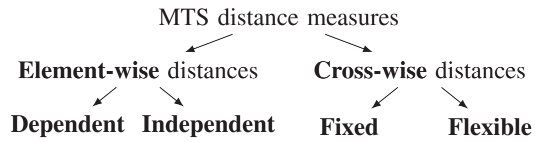


Fig. 2: Taxonomy of MTS distance measures

its own warping path. Comparing to Euclidean distance, the authors showed both qualitatively and quantitatively that the choice for *DTW-D* or *DTW-I* is relevant and domain-specific. Shifaz et al. [42] recently extended 6 more univariate measures to MTS, accounting both dependent and independent time warping, and compared them on the task of classification on the UEA archive [2]. Among the measures are Euclidean distance, LCSS, Edit distance, and Move-Split-Merge. Similar to [43], the authors concluded that no measure nor alignment strategy is superior; all but one measures were the best on at least one dataset.

## III. Taxonomy

In this section we will present a novel taxonomy of MTS distance measures, which separates the measures based on three properties; (a) the way variates are handled during distance computation, (b) the way point alignments are constrained over the variates, and, (c) the support of variate mismatches between MTS. The first property divides MTS measures into *element-wise* and *cross-wise* distances. The second property divides element-wise distances into *dependent* and *independent* point alignments, and the third property divides cross-wise distances into distances on *fixed* and *flexible* representations. The taxonomy is depicted in Fig. 2.

**Handling variates (element-wise/cross-wise)** The first proposed MTS distances involved aggregation of per-variate distances computed with univariate measures [42], [43], [46]. As such, they ignored possible interdependencies between the variates within an MTS. Formally, these distances are functions that only include terms on *matching variates* of *different MTS*. We coin such measures *element-wise distances*. To illustrate, given two MTS $\boldsymbol{X}, \boldsymbol{Y}$, an element-wise distance function $d(\boldsymbol{X}, \boldsymbol{Y})$ only includes terms of the form $f(\boldsymbol{X}_i, \boldsymbol{Y}_i)$. The DTW extensions of Shokoohi-Yekta et al. [43] (*DTW-I* and *DTW-D*) are examples of element-wise distances. Other works propose distance measures that treat the MTS variates holistically, allowing preservation of correlations among the variates. Formally, these measures are functions that, in addition to terms with matching variates, also include terms on *different variates*, i.e., cross-terms of the form $f(\boldsymbol{X}_i, \boldsymbol{Y}_j)$ or $f(\boldsymbol{X}_i, \boldsymbol{X}_j)$, with $i \neq j$. We call such measures *cross-wise distances*. In several cases, these measures capture the internal dependencies of an MTS by *transforming it to a new representation* based on its internal distances, i.e., $d(X_i, X_j)$, and then computing distances on the transformed MTS using a univariate measure [12], [47]. Examples of such distances over transformations are the PCA Similarity Factor ($S_{PCA}$) and Eros [47], which capture similarities in the internal covariance matrices of MTS by aggregating the cosine similarities

| Term type | *Element-wise* | *Cross-wise* |
|---|---|---|
| Two MTS, same variate; $f(\boldsymbol{X}_i, \boldsymbol{Y}_i)$ | ✓ | ✓ |
| Two MTS, different variates; $f(\boldsymbol{X}_i, \boldsymbol{Y}_j)$ | | ✓ |
| One MTS, different variates; $f(\boldsymbol{X}_i, \boldsymbol{X}_j)$ | | ✓ |

TABLE I: Terms used in element- and cross-wise distances.

between the principal components of the MTS [17]. Table I summarizes the terms used in the functions of element-wise and cross-wise distances.

**Handling point-alignments (dependent/independent)** As raw processing of the MTS variates allows for elastic alignment between the time points of two MTS, the question arises whether this alignment should be shared among variates or not. Adopting the terminology of Shifaz et al. [42], we split element-wise distances into subcategories based on the way point alignments are constrained over the MTS variates; *dependent* (when the alignment is shared among variates, e.g., *DTW-D*) or *independent* (e.g., *DTW-I*).

**Handling variate mismatches (fixed/flexible)** Frequently, not all variates are relevant when comparing two MTS, or the MTS do not contain measurements on exactly the same variates.[2] In these cases, the MTS are compared on a subset of the variates. We refer to this phenomenon as *variate mismatches*. Handling variate mismatches can be non-trivial for cross-wise distances when it involves transforming the MTS to a new representation without the query in mind, and the original variates cannot be clearly separated after the transformation. For example, when designing an efficient similarity search algorithm for $S_{PCA}$ [47], one likely stores and indexes the principal components instead of the original data. As these principal components are linear combinations of *all original variates* they cannot be reduced at query time to the query variates, leading to an error on the distance to the query when the original variates did not match. We call such transformations *fixed*. Conversely, representations such as covariance matrices of MTS are *flexible*; one can consider only the sub-matrices corresponding to the query variates when computing distances, thus allowing for matching of variates after the transformation. We therefore differentiate cross-wise measures on this property, to either *fixed* or *flexible* representations.

## IV. Populating the Taxonomy

In this section, we discuss the key distance measures for MTS, and group them by the subcategories of the taxonomy presented in Section III. Table II summarizes all measures.

### A. Element-wise distances

We start by looking into the different element-wise distance measures that will be evaluated. For each measure we will consider both the dependent and independent variant.

**Multivariate $L_p$ distance ($L_p$):** $L_p$ is defined as the $p$-norm of the differences over all variates of two MTS [2]: $L_p(\boldsymbol{X}, \boldsymbol{Y}) = \sqrt[p]{\sum_{i=1}^{m} \sum_{j=1}^{v} (\boldsymbol{X}_{i,j} - \boldsymbol{Y}_{i,j})^p}$. $L_2$ was previously used in [41], [42], [46], [47]. Since $L_p$ lacks any form of time elasticity, it does not have a dependent and independent variant.

[2]This occurs in many real-world applications where the variates are not standardized across measurements (e.g., different measuring standards or medical equipment between hospitals in patient health monitoring).

| Name | Ref | Type | 2nd level class |
|---|---|---|---|
| $L_p$ | [2], [41]–[43], [46], [47] | EW | Dep./Indep. |
| *DTW* | [2], [41]–[43], [46] | EW | Dep./Indep. |
| *LCSS* | [41], [42], [46] | EW | Dep./Indep. |
| $D_{PCA}$ | [47] | CW | Fixed |
| $D_{Eros}$ | [47] | CW | Fixed |
| $D_{KL}$ | | CW | Flexible |
| *Frob* | | CW | Flexible |

TABLE II: Overview of multivariate distance measures.

**Multivariate DTW (*DTW-D*, *DTW-I*):** We start by introducing DTW for UTS, before presenting its extension for MTS. DTW distance is defined as the minimum cost of all possible alignments (i.e., mapping of data points) between two UTS. The cost of an alignment (or *warping path*) is the Euclidean distance taken over the aligned values. The independent extension of DTW (*DTW-I*) is defined as the sum of the univariate DTW distances over the variates of two MTS [43]. The dependent extension of DTW (*DTW-D*) is the same as univariate DTW, with the exception that it uses the multivariate Euclidean distance (i.e., $L_2$) over the aligned points as cost function [43]. Consequently, where *DTW-I* results in $v$ unique alignments, *DTW-D* results in a single alignment shared between all variates.

**Multivariate LCSS (*LCSS-D*, *LCSS-I*):** Univariate Longest Common Subsequence (LCSS) distance is defined as the length of the longest subsequence of two time series, potentially starting at different time points [46]. Real-valued subsequences are considered matching if all data points are within a certain threshold $\epsilon$ of each other and the temporal offset between the subsequences is lower than $\delta \in \mathbb{N}^+$. The independent variant of LCSS (*LCSS-I*) is defined as the sum of the univariate LCSS distances over the variates of two MTS [42]. The dependent variant (*LCSS-D*) introduced in [46], requires subsequences to match on all variates.

### B. Cross-wise distances

We now present the state-of-the-art cross-wise distance measures, including two novel measures: Cross-wise-Frobenius, and the KL-divergence between multivariate gaussians.

**PCA distance ($D_{PCA}$):** The PCA Similarity Factor is defined as the sum of the cosine similarities of all pairs of principal components of two MTS [19], [47]: $S_{PCA,k}(\boldsymbol{X}, \boldsymbol{Y}) = \sum_{i=1}^{k} \sum_{j=1}^{k} cos^2(\omega_{i,j})$ where $\omega_{i,j}$ is the angle between the $i$-th and $j$-th principal component of $\boldsymbol{X}$ and $\boldsymbol{Y}$, and $k \leq v$ is the number of considered principal components. As discussed in Section III, $S_{PCA}$ involves *fixed* representations. As $S_{PCA}$ is similarity measure, we convert it to a distance measure by taking the reciprocal [23], i.e., $D_{PCA,k} = \frac{1}{1+S_{PCA,k}}$.

**Eros distance ($D_{Eros}$):** Eros similarity is defined as the weighted sum of the cosine similarities of the ordered principal components of two MTS [47]: $S_{Eros}(\boldsymbol{X}, \boldsymbol{Y}) = \sum_{i=1}^{v} w_i |\cos(\omega_{i,i})|$ where the weight vector $w$ is based on the (normalized) aggregated eigenvalues of a training set of MTS, as a measure of average importance. As such, Eros is

| | | Element-wise | | | | | Cross-wise | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $L_2$ | DTW-D | DTW-I | LCSS-D | LCSS-I | $D_{Eros}$ | $D_{KL}$ | $D_{PCA}$ | *Frob*-$L_1$ | *Frob*-$L_2$ | *Frob-Cov* | *Frob-DTW* |
| $r = 1$ | Unnormed | 1 | **7** | 3 | 0 | **7** | 1 | 3 | 0 | 1 | 0 | 2 | 1 |
| | Normed | 2 | 4 | 3 | 0 | 4 | 0 | 2 | 1 | 0 | 0 | 0 | 2 |
| $r = 5$ | Unnomed | 10 | **17** | **17** | 2 | 11 | 3 | 9 | 2 | 5 | 6 | 7 | 6 |
| | Normed | 9 | 13 | 14 | 1 | 9 | 4 | 6 | 3 | 2 | 1 | 3 | 4 |
| Median | Unnomed | 0.550 | 0.675 | **0.754** | 0.510 | 0.650 | 0.534 | 0.511 | 0.259 | 0.545 | 0.555 | 0.538 | 0.511 |
| Accuracy | Normed | 0.596 | 0.676 | 0.743 | 0.500 | 0.648 | 0.505 | 0.503 | 0.259 | 0.481 | 0.500 | 0.503 | 0.514 |

TABLE III: Number of datasets where each measure was in the $r$ highest classification accuracies.

a weighted version of $S_{PCA}$ considering only corresponding axes instead of all pairs. Eros similarity can be transformed to a distance measure as $D_{Eros} = \sqrt{2 - 2 * S_{Eros}}$ [47] and has the same *fixed* representations as $D_{PCA}$.

**KL-divergence of Multivariate Gaussians ($D_{KL}$):** As a more theoretically sound alternative to $D_{PCA}$ and $D_{Eros}$, we propose to model MTS as single multivariate gaussians and use the (closed-form) KL-divergence as a distance measure. $D_{KL}$ supports ad-hoc matching of variates by marginalizing the multivariate gaussians to the matched subset of variates. By modelling MTS as a gaussian, we do not address the temporal aspect of MTS at all. Instead, the data points are handled as *i.i.d.* samples from a distribution.

**Cross-wise-Frobenius (*Frob*):** *Frob* takes the Frobenius norm of the element-wise differences of the distance matrices of two MTS. As such, when applied to the covariance matrices of two MTS, *Frob* effectively measures the Euclidean distance between the vectors containing the variances and covariances of the two MTS. Formally, *Frob* is defined as: $\textit{Frob}(\boldsymbol{X}, \boldsymbol{Y}, f) = ||\Sigma_{X,f} - \Sigma_{Y,f}||_F$ where $\Sigma_{X,f}$ indicates the distance matrix of $\boldsymbol{X}$ using some distance measure $f$. Furthermore, as one can select subsets of variates through the submatrices of $\Sigma_{X,f}$, the representations are *flexible*. We refer to instances of *Frob* by including the respective distance measure in the name, e.g., *Frob-DTW*. In our evaluation we consider *Frob*-$L_1$, *Frob*-$L_2$, *Frob-Cov*, and *Frob-DTW*.

## V. Evaluation

The experiments aimed to evaluate the discriminative power of the measures, using classification as the downstream task (Section V-A). This goal is motivated by the fact that classification is a common proxy for such comparisons when direct quantitative evaluation of measures is non-trivial [34], [41]–[43], [47]. We stress again that this work focuses on evaluating the quality of distance measures; computational efficiency is not the focus of this work.

**Implementation and datasets.** All measures are implemented in Python. The code is available on GitHub. Our evaluation uses the popular *UEA archive*, which contains 30 labeled real-world datasets from diverse domains, with varying number of variates and lengths [2].

### *A. Classification performance*

Classification is performed using 5-NN classifiers. Performances are measured through accuracies on leave-one-out cross-validation. Additionally, we investigate the use of z-normalization. Analyzing the performance of the all measures in Table III, we find that element-wise measures on average perform better than cross-wise distances, but they do not dominate them. In fact, there are datasets where cross-wise measures perform significantly better than element-wise measures. This is the case for *BasicMotions*, *Cricket*, and *EigenWorms*, among others. Particularly for the latter, the average accuracies of cross-wise measures are significantly higher than those of element-wise measures. Overall, the results show that there is no 'one-size-fits-all' measure that exclusively outperforms all others. This is also the case for normalization of the data, which does not consistently improve or degrade performance. This is in line with the findings of [2], [41]–[43], [47]. Therefore, the choice of measure and normalization should be made based on the dataset and application at hand, and the measure properties that fit that context. Analyzing the differences among element-wise measures, we find that *DTW-D* ranks highest on average, while *DTW-I* has the highest average accuracy. This tells us that *DTW-I* is more robust compared to *DTW-D*, and that there is also no one best element-wise measure, confirming [41]–[43]. Similar insights are obtained for cross-wise measures, with $D_{KL}$ ranking highest on average, and *Frob*-$L_2$ having the highest average accuracy among cross-wise measures, but no exclusive superiority of any cross-wise measure.

### *B. Future work*

Although this work evaluates 12 diverse measures, it excludes element-wise measures based on Lorentzian distance, Cross-Correlation, and Merge-Split-Move, and normalizations such as UnitLength, which all showed promising performance in [34]. Future work should include these measures and normalizations, and compare them more robustly through statatical tests such as the Friedman test [11] and the post-hoc Nemenyi test [25]. Such work should go in tandem with efficiency evaluations such as the studies of Echihabi et al. [9], [10] to offer a complete handbook for measure selection in MTS applications.

## VI. Conclusions

In this paper, we made the first steps towards a structured evaluation of MTS distance measures. Through a novel taxonomy, distance measures were categorized into two groups based on the way they handle the variates of the MTS. Both groups were split into subcategories based on properties relevant to each group. Experimental results show that no MTS measure is superior, and the optimal choice depends on the data. While this work proposes effective alternatives to existing cross-wise measures, we argue that future work should focus on the development of new measures that capture both aspects of similarity in MTS explicitly and dynamically.

## REFERENCES


[1] Alon, J., Sclaroff, S., Kollios, G., Pavlovic, V.: Discovering clusters in motion time-series data. In: 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings., vol. 1, pp. I–I (2003)

[2] Bagnall, A.J., Dau, H.A., Lines, J., Flynn, M., Large, J., Bostrom, A., Southam, P., Keogh, E.J.: The UEA multivariate time series classification archive, 2018. CoRR **abs/1811.00075** (2018)

[3] Bariya, M., von Meier, A., Paparrizos, J., Franklin, M.J.: k-shapestream: Probabilistic streaming clustering for electric grid events. In: 2021 IEEE Madrid PowerTech, pp. 1–6. IEEE (2021)

[4] Boniol, P., Paparrizos, J., Kang, Y., Palpanas, T., Tsay, R.S., Elmore, A.J., Franklin, M.J.: Theseus: navigating the labyrinth of time-series anomaly detection. Proceedings of the VLDB Endowment **15**(12), 3702–3705 (2022)

[5] Boniol, P., Paparrizos, J., Palpanas, T., Franklin, M.J.: Sand in action: subsequence anomaly detection for streams. Proceedings of the VLDB Endowment **14**(12), 2867–2870 (2021)

[6] Boniol, P., Paparrizos, J., Palpanas, T., Franklin, M.J.: Sand: streaming subsequence anomaly detection. Proceedings of the VLDB Endowment **14**(10), 1717–1729 (2021)

[7] Corradini, A.: Dynamic time warping for off-line recognition of a small gesture vocabulary. In: Proceedings IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, pp. 82–89 (2001)

[8] Dziedzic, A., Paparrizos, J., Krishnan, S., Elmore, A., Franklin, M.: Band-limited training and inference for convolutional neural networks. In: International Conference on Machine Learning, pp. 1745–1754. PMLR (2019)

[9] Echihabi, K., Zoumpatianos, K., Palpanas, T., Benbrahim, H.: The lernaean hydra of data series similarity search: An experimental evaluation of the state of the art. Proc. VLDB Endow. **12**(2), 112–127 (2018)

[10] Echihabi, K., Zoumpatianos, K., Palpanas, T., Benbrahim, H.: Return of the lernaean hydra: experimental evaluation of data series approximate similarity search. Proc. VLDB Endow. **13**(3), 403–420 (2019)

[11] Friedman, M.: The use of ranks to avoid the assumption of normality implicit in the analysis of variance. Journal of the American Statistical Association **32**(200), 675–701 (1937)

[12] Ghassempour, S., Girosi, F., Maeder, A.: Clustering multivariate time series using hidden markov models. Int J Environ Res Public Health **11**(3), 2741–2763 (2014)

[13] Goel, R., Soni, S., Goyal, N., Paparrizos, J., Wallach, H., Diaz, F., Eisenstein, J.: The social dynamics of language change in online networks. In: Social Informatics: 8th International Conference, SocInfo 2016, Bellevue, WA, USA, November 11-14, 2016, Proceedings, Part I 8, pp. 41–57. Springer (2016)

[14] Jeung, H., Sarni, S., Paparrizos, I., Sathe, S., Aberer, K., Dawes, N., Papaioannou, T.G., Lehning, M.: Effective metadata management in federated sensor networks. In: 2010 IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing, pp. 107–114. IEEE (2010)

[15] Jiang, H., Liu, C., Jin, Q., Paparrizos, J., Elmore, A.J.: Pids: attribute decomposition for improved compression and query performance in columnar storage. Proceedings of the VLDB Endowment **13**(6), 925–938 (2020)

[16] Jiang, H., Liu, C., Paparrizos, J., Chien, A.A., Ma, J., Elmore, A.J.: Good to the last bit: Data-driven encoding with codecdb. In: Proceedings of the 2021 International Conference on Management of Data, pp. 843–856 (2021)

[17] Jolliffe, I.T.: Principal component analysis for special types of data. Springer (2002)

[18] Kadous, M.W., et al.: Temporal classification: Extending the classification paradigm to multivariate time series. Citeseer (2002)

[19] Krzanowski, W.J.: Between-groups comparison of principal components. Journal of the American Statistical Association **74**(367), 703–707 (1979)

[20] Liess, S., Agrawal, S., Chatterjee, S., Kumar, V.: A teleconnection between the west siberian plain and the ENSO region. Journal of Climate **30**(1), 301 – 315 (2017)

[21] Liu, C., Jiang, H., Paparrizos, J., Elmore, A.J.: Decomposed bounded floats for fast compression and queries. Proceedings of the VLDB Endowment **14**(11), 2586–2598 (2021)

[22] Liu, S., Mangla, T., Shaowang, T., Zhao, J., Paparrizos, J., Krishnan, S., Feamster, N.: Amir: Active multimodal interaction recognition from video and network traffic in connected environments. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **7**(1), 1–26 (2023)

[23] Mair, P., Groenen, P.J.F., de Leeuw, J.: More on multidimensional scaling and unfolding in r: smacof version 2. Journal of Statistical Software **102**(10), 1–47 (2022)

[24] McKeown, K., Daume III, H., Chaturvedi, S., Paparrizos, J., Thadani, K., Barrio, P., Biran, O., Bothe, S., Collins, M., Fleischmann, K.R., et al.: Predicting the impact of scientific concepts using full-text features. Journal of the Association for Information Science and Technology **67**(11), 2684–2696 (2016)

[25] Nemenyi, P.B.: Distribution-free multiple comparisons. Ph.D. thesis, Princeton University (1963)

[26] Paparrizos, I.: Fast, scalable, and accurate algorithms for time-series analysis. Ph.D. thesis, Columbia University (2018)

[27] Paparrizos, J., Boniol, P., Palpanas, T., Tsay, R.S., Elmore, A., Franklin, M.J.: Volume under the surface: a new accuracy evaluation measure for time-series anomaly detection. Proceedings of the VLDB Endowment **15**(11), 2774–2787 (2022)

[28] Paparrizos, J., Edian, I., Liu, C., Elmore, A.J., Franklin, M.J.: Fast adaptive similarity search through variance-aware quantization. In: 2022 IEEE 38th International Conference on Data Engineering (ICDE), pp. 2969–2983. IEEE (2022)

[29] Paparrizos, J., Franklin, M.J.: Grail: efficient time-series representation learning. Proceedings of the VLDB Endowment **12**(11), 1762–1777 (2019)

[30] Paparrizos, J., Gravano, L.: k-shape: Efficient and accurate clustering of time series. In: Proceedings of the 2015 ACM SIGMOD international conference on management of data, pp. 1855–1870 (2015)

[31] Paparrizos, J., Gravano, L.: Fast and accurate time-series clustering. ACM Transactions on Database Systems (TODS) **42**(2), 1–49 (2017)

[32] Paparrizos, J., Kang, Y., Boniol, P., Tsay, R.S., Palpanas, T., Franklin, M.J.: Tsb-uad: an end-to-end benchmark suite for univariate time-series anomaly detection. Proceedings of the VLDB Endowment **15**(8), 1697–1711 (2022)

[33] Paparrizos, J., Liu, C., Barbarioli, B., Hwang, J., Edian, I., Elmore, A.J., Franklin, M.J., Krishnan, S.: Vergedb: A database for iot analytics on edge devices. In: CIDR (2021)

[34] Paparrizos, J., Liu, C., Elmore, A., Franklin, M.J.: Querying time-series data: A comprehensive comparison of distance measures. IEEE Data Engineering Bulletin (DEB 2023) **47**, 69–88 (2023)

[35] Paparrizos, J., Liu, C., Elmore, A.J., Franklin, M.J.: Debunking four long-standing misconceptions of time-series distance measures. In: Proceedings of the 2020 ACM SIGMOD international conference on management of data, pp. 1887–1905 (2020)

[36] Paparrizos, J., White, R.W., Horvitz, E.: Detecting devastating diseases in search logs. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 559–568 (2016)

[37] Paparrizos, J., White, R.W., Horvitz, E.: Screening for pancreatic adenocarcinoma using signals from web search logs: Feasibility study and results. Journal of oncology practice **12**(8), 737–744 (2016)

[38] Paparrizos, J., Wu, K., Elmore, A., Faloutsos, C., Franklin, M.J.: Accelerating similarity search for elastic measures: A study and new generalization of lower bounding distances. Proceedings of the VLDB Endowment **16**(8), 2019–2032 (2023)

[39] Roddick, J.F., Hornsby, K.S.: Temporal, spatial, and spatio-temporal data mining. In: Lecture Notes in Computer Science (2001)

[40] Ruiz, A.P., Flynn, M., Large, J., Middlehurst, M., Bagnall, A.: The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances. Data Mining and Knowledge Discovery **35**(2), 401–449 (2021)

[41] Salarpour, A., Khotanlou, H.: An empirical comparison of distance measures for multivariate time series clustering. International Journal of Engineering **31**(2), 250–262 (2018)

[42] Shifaz, A., Pelletier, C., Petitjean, F., Webb, G.I.: Elastic similarity and distance measures for multivariate time series. Knowledge and Information Systems **65**(6), 2665–2698 (2023)

[43] Shokoohi-Yekta, M., Hu, B., Jin, H., Wang, J., Keogh, E.J.: Generalizing DTW to the multi-dimensional case requires an adaptive approach. Data Min. Knowl. Discov. **31**(1), 1–31 (2017)

[44] Sylligardos, E., Boniol, P., Paparrizos, J., Trahanias, P., Palpanas, T.: Choose wisely: An extensive evaluation of model selection for anomaly detection in time series. Proceedings of the VLDB Endowment **16**(11), 3418–3432 (2023)
[45] Tuli, S., Casale, G., Jennings, N.R.: Tranad: Deep transformer networks for anomaly detection in multivariate time series data. Proc. VLDB Endow. **15**(6), 1201–1214 (2022)
[46] Vlachos, M., Hadjieleftheriou, M., Gunopulos, D., Keogh, E.: Indexing multi-dimensional time-series with support for multiple distance measures. In: Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '03, p. 216–225. Association for Computing Machinery, New York, NY, USA (2003)
[47] Yang, K., Shahabi, C.: A pca-based similarity measure for multivariate time series. In: Proceedings of the 2nd ACM International Workshop on Multimedia Databases, MMDB '04, p. 65–74. Association for Computing Machinery, New York, NY, USA (2004)