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

Jens E. d'Hondt[1] Odysseas Papapetrou[1] John Paparrizos[2]

[1] Eindhoven University of Technology, The Netherlands

[2] The Ohio State University, USA

May 13, 2024

Introduction - Multivariate Time Series

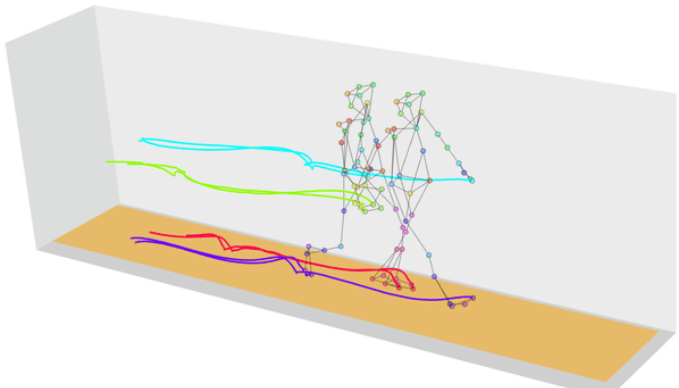
Multivariate time series are becoming increasingly common.

Definition

Multivariate time series (MTS): A time-ordered sequence of data with multiple signals, meaning that each time point has a vector of values, rather than a single scalar. These signals are called the *variates*, *dimensions*, or *channels* of the MTS.

Example

Motion capture data; the position, speed and acceleration of different body parts over time.



Introduction - Distance Measures

- **Open problem:** Unveiling the SOTA in MTS distance measures for Similarity Search.
- **KPI:** Ability to differentiate MTS based on features relevant to the downstream task at hand.

Definition

Similarity Search: Given a query MTS $\vec{Q}$ and a database of other MTS $\mathcal{D}$, find all $\vec{T} \in \mathcal{D}$ similar to $\vec{Q}$, based on some distance $d(\vec{Q}, \vec{T})$.

Introduction - Contributions

- **Novel taxonomy** of MTS distance measures.
- **Population of the taxonomy** with 7 common distance measures (+ 2 new measures).
- **Evaluation of distance measures** through kNN classification on UEA-archive.

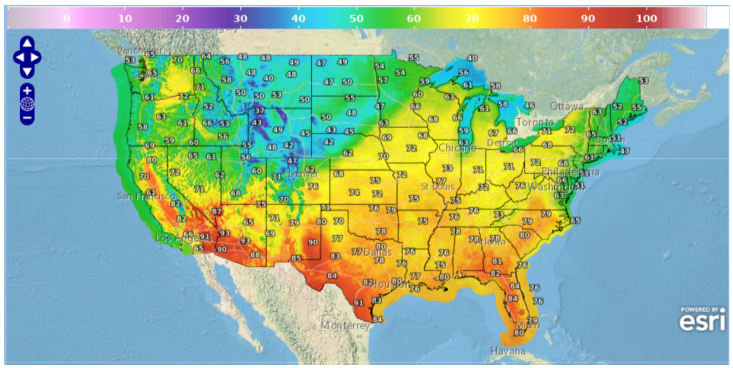
Taxonomy - Level 1: Handling the channels

- How does a distance measure handle the channels of the MTS? Holistically or individually?
- This is reflected by the *type of terms* in the distance formula.
- **Element-wise** distance measures treat the channels individually.
  - $d(\vec{X}, \vec{Y}) = f(x_1, y_1) + f(x_2, y_2) + \ldots + f(x_n, y_n)$
  - *Example*: Multivariate L_2 distance, which boils down to the sum of channel-wise L_2 distances.
- **Cross-wise** distance measures treat the channels holistically.
  - $d(\vec{X}, \vec{Y}) = f(x_1, y_1) + \underbrace{|f(x_1, x_2) - f(y_1, y_2)|}_{\text{internal distance}} + \cdots$
  - Often transform the MTS to a new representation.
  - *Example*: PCA similarity factor, which is the sum of cosine similarities between the principal components.

Taxonomy - Level 1: Handling the channels

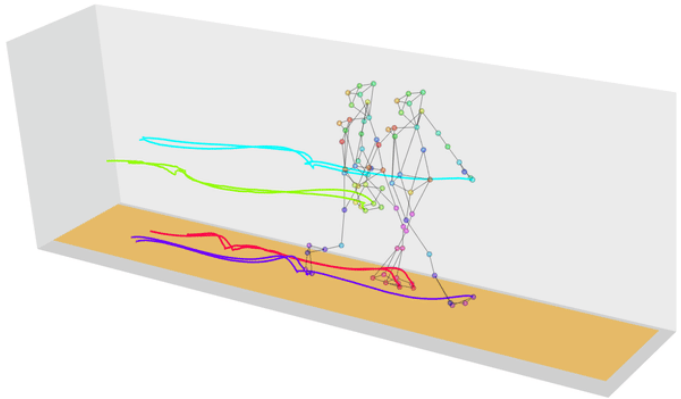
Element-wise distance:

- Captures *direct* similarities.
- Good when little noise, and no/fixed inter-channel correlations.



Cross-wise distance

- Captures *indirect* similarities.
- Good when inter-channel correlations hold predictive power.



Taxonomy - Element-wise Level 2: Time-alignments

- **Independent** measures allow for different alignments for each channel (e.g., DTW with v warping paths).
- **Dependent** measures constrain the alignments to be the same for all channels (e.g., DTW with 1 warping path).

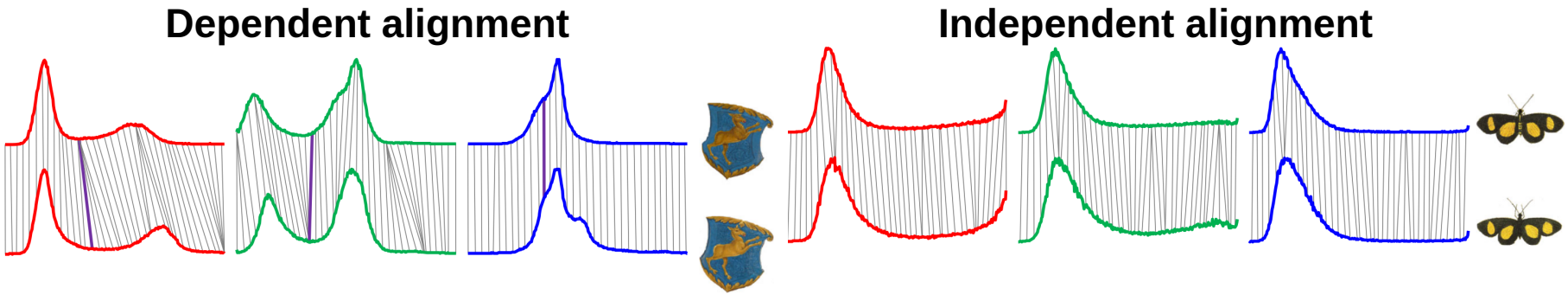


Figure: Dependent vs. independent time-alignments, by Yekta et. al [1]

[1] M. Shokoohi-Yekta, B. Hu, H. Jin, J. Wang, and E. J. Keogh (2017). "Generalizing DTW to the multi-dimensional case requires an adaptive approach". In: *Data Min. Knowl. Discov.* 31.1, pp. 1–31

Taxonomy - Cross-wise Level 2: Representing MTS

- Cross-wise measures often transform MTS to a new representation that captures *internal dependencies*.
- **Fixed** representations **do not** allow separation after transformation (e.g., PCA).
- **Flexible** representations **do** allow separation of the original channels after transformation (e.g., Multivariate Gaussian).
- This is important when query channels are selected adhoc (at query time).

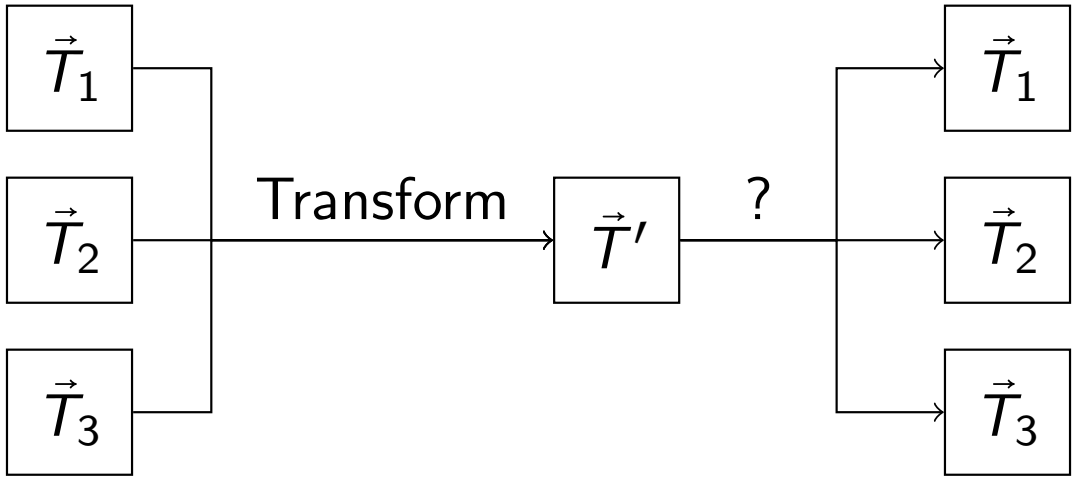


Figure: Are the channels separable after transformation?

Populating the taxonomy

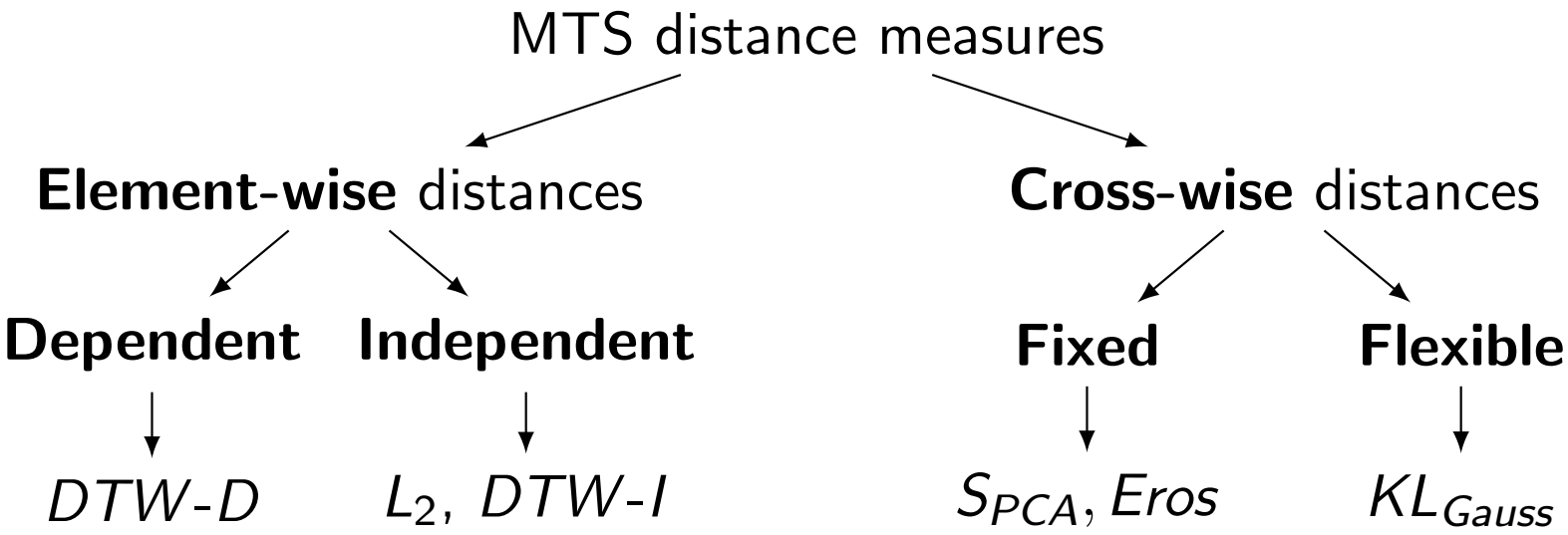


Figure: Taxonomy of MTS distance measures with examples

Evaluation - kNN Classification

- **Dataset:** UEA MTS Classification Archive (30 labeled real-world datasets).
- **KPI:** Classification accuracy rank.

Evaluation - kNN Classification

- **Dataset:** UEA MTS Classification Archive (30 labeled real-world datasets).
- **KPI:** Classification accuracy rank.
- **Results:**
  - No indication of a single best measure.
    - *DTW-D* ranks highest (17/30 in top-5)
    - *DTW-I* has highest median accuracy (0.75).
    - Cross-wise not dominated (D_{KL} 9/30 in top-5).
  - Also, no clear dominance of measure types / normalization.
- **Future work:** More datasets, more measures, more normalization techniques.
  - Goal: Provide guidelines to practitioners and researchers on **measure design** and **measure selection**.

Questions?

Evaluation - kNN Classification

		Element-wise			Cross-wise		
		L_2	DTW-D	DTW-I	D_{KL}	D_{PCA}	$Frob\text{-}L_2$
#Top-1	Unnormed	1	**7**	3	3	0	0
	Normed	2	4	3	2	1	0
#Top-5	Unnormed	10	**17**	**17**	9	2	6
	Normed	9	13	14	6	3	1
Median	Unnomed	0.55	0.68	**0.75**	0.51	0.26	0.56
Accuracy	Normed	0.60	0.68	0.74	0.50	0.26	0.50

Table: Number of datasets where each measure was in the top-x highest accuracies.