Weighted Multi-view Partitioning of Time Series

Yves Lechevallier[1]
Francisco de A. T. de Carvalho[2]

(1)
Yves.Lechevallier@laposte.net

(2) Centro de Informatica-CIn/UFPE
Recife-PE, Brasil,
fatc@cin.ufpe.br

International Workshop Advances in Data Science
Dauphine University, Paris

Outline

1. Introduction

2. Choice of the D'urso and Vichi distance between time series

3. Dynamic Clustering Algorithm with Adaptive Weights

4. Applications : Solar Irradiance data sets and Satellite time trajectory
    - Application I : Solar Irradiance
    - Application II : Satellite time trajectory

5. References

Introduction

Context

- Multi-View Clustering models can be viewed as a way to extract information from different data table ;
- These tables have been generated using different sets of variables associated to different views ;
- Provide a partition and a prototype for each cluster and learn a relevance weight vectors for each table and each variable.
- Optimize an **adequacy criterion** that measures the fitting between clusters and their prototypes using adaptive weights.
- In our context, each view represents a time series or curve.
- A specific time series distance (D'urso and Vichi distance) is assigned to each time series

Definitions

INPUT : Data

- $E=\{e_1,\ldots,e_n\}$ be a set of n objects ;
- V is the number of tables (time series) and d_v is the number of variables (discretization point) in the table v ;
- Φ is the global representation space, Φ^v is the representation space of the table v and Φ^v_j is the representation space of the variable j belonging to the table v ;
- Each object $e_i \in E$ is described by a *vector* $(X_i^1,\ldots,X_i^v,\ldots,X_i^V)$ where $X_i^v \in \Phi^v$ is the description of the object e_i in the table v ;
- $x_i^{(j,v)} \in \Phi^v_j$ is the value of the object e_i for the variable j belonging to the table v.

OUTPUT : Partition and prototype set

A prototype is a Fréchet mean and a prototype set is a set of Fréchet means.

- A partition $P=(C_1,\ldots,C_K)$ of E into K clusters ;
- The prototype set is $(G_1,\ldots,G_K)$ where G_k is the prototype of the cluster k ;
- $G_k=(G_k^1,\ldots,G_k^v,\ldots,G_k^V)$ where G_k^v is an element of Φ^v ;

Multiple Time Series and Cluster Prototypes

Multiple time series and Cluster Prototype (also time series) are in the same space Φ

Multiple Time Series

time series_1	...	time series_v	...	time series_V
X_i^1	...	X_i^v	...	X_i^V
$x_1^{(1,1)} \dots x_1^{(d_1,1)}$	...	$x_1^{(1,v)} \dots x_1^{(d_v,v)}$	...	$x_1^{(1,V)} \dots x_1^{(d_V,V)}$
...	...	...	...	...
$x_n^{(1,1)} \dots x_n^{(d_1,1)}$	...	$x_n^{(1,v)} \dots x_n^{(d_v,v)}$	...	$x_n^{(1,V)} \dots x_n^{(d_V,V)}$

Cluster Prototype defined in each time series

time series_1	...	time series_v	...	time series_V
G_i^1	...	G_i^v	...	G_i^V
$g_1^{(1,1)} \dots g_1^{(d_1,1)}$	...	$g_1^{(1,v)} \dots g_1^{(d_v,v)}$	...	$g_1^{(1,V)} \dots g_1^{(d_V,V)}$
...	...	...	...	...
$g_K^{(1,1)} \dots g_K^{(d_1,1)}$	...	$g_K^{(1,v)} \dots g_K^{(d_v,v)}$	...	$g_K^{(1,V)} \dots g_K^{(d_V,V)}$

Clustering process reduces the error between Multiple Time series and Cluster Prototype of time series by minimization of the criterion W

Type of Variables

Single-valued variables (Classical variables)

- quantitative or numerical variables.
- qualitative or categorical variables.
- binary variables.

Multi-valued variables (Symbolic variables)

- interval variables or quantitative multi-valued variables.
- categorical multi-valued variables.
- modal or histogram variables.

Time series variables

We propose to use the D'urso and Vichi distance that is a compromise between these 3 relational data tables (position, velocity and acceleration) where the weights are learned by our method. With time series or curves, the relationship between variables is an order on the variables.

Choice of the distance $d_{(j,v)}$ in the space Φ_j^v

This choice should be realized by user.

$d_{(j,v)}$ is the distance of the variable j belonging to the table v then $(\Phi_j^v, d_{(j,v)})$ is a metric space

Time series data

- From the original time series, we propose to construct three data sets corresponding to the **position**, **velocity** and **acceleration**
- The distance $d_{(j,v)}$ is the D'urso and Vichi distance based on the Euclidean distance.

Interval time series

- The representation space Φ_j^v is a finite and closed interval set of $\Re$.
- we propose to construct three interval data sets corresponding to the **position**, **velocity** and **acceleration** of the interval time series.
- The distance $d_{(j,v)}$ is an extension of the D'urso and Vichi distance based on the Hausdorff distance.

Global distance

The proposed global distance d is determined by a positive weighted linear combination of distances corresponding to the different time series by the following formula :

Weighted linear combination of distance

$$d^2(e_i, e_l) = \sum_{v=1}^{V} \omega_v \sum_{j=1}^{d^v} \tau_j^v d^2_{(j,v)}(e_i, e_l)$$

where :

- τ_j^v is the weight of the variable j of the time series v and $\tau_j^v \geq 0$;
- ω_v is the weight of the time series v and $\omega_v \geq 0$;
- $d_{(j,v)}$ is the distance associated to the variable j of the time series v ;

$\Lambda = (\Lambda_1, \ldots, \Lambda_v, \ldots, \Lambda_V)$ with $\Lambda_v = (\tau_1^v, \ldots, \tau_j^v, \ldots, \tau_{d_v}^v)$ where τ_j^v is the weight value of the variable j of the table v.
$\Omega = (\omega_1, \ldots, \omega_V)$ is the weight vector of the time series.
As $(\Phi_j^v, d_{(j,v)})$ is a metric space then (Φ, d) is, also, a metric space.

Clustering times series strategies

Two strategies :

- Clustering on Multiple Dissimilarity Matrices (F. A. T. De Carvalho, Y. Lechevallier, and F. M. De Melo, 2012)
- Clustering on times series with D'urso and Vichi distance. It is our approach

Advantages or not of Multiple Dissimilarity Matrices

Multiple Dissimilarity Matrices

- gives a collaborative role between the dissimilarity matrices (Pedrycz, 2002) and allows to obtain a final consensus partition (Leclerc and Cucumel, 1987).
- solves the problem of multiple representations of the objects (Cleziou and al 2009) (called also Multi-View Data).

In rare cases the similarity or dissimilarity matrix are native. Generally similarity or dissimilarity measures are computed on the data table.
The choice of the distance depends on the variable type. See (F. A. T. De Carvalho, Y. Lechevallier, and F. M. De Melo,2012) where D'urso and Vichi distance is proposed to aggregate time series.
The cluster prototype is an observation (not a aggregative data), it is a time series of the set E. The prediction of the cluster on a new time series is based on the distance between this new time series and this observation.
Our approach on the relational data tables uses D'urso and Vichi distance and allows to associate an aggregative data at each cluster.

Distances on time series

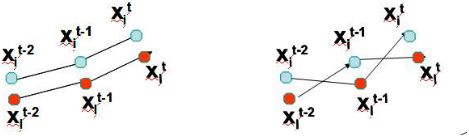


FIGURE – Time series (a) time series (b)

If the Euclidean distance is used the distances between two times series (a) and two times series (b) are equal. The permutation of the order of the variables does not change the result.

It is important that the distances between time series (a) and (b) are different and the distance on (a) computed is smaller that the distance on (b).

The D'urso and Vichi distance verifies these properties.

D'urso and Vichi distance

In order to compare time series, we propose a cross sectional-longitudinal dissimilarity function proposed by D'urso and Vichi (1998).

From the i-th time series, three time series are computed :
The position of the i-th time series is a vector $\mathbf{x}_i{=}(x_i^t)_{t\in T}$ where the value $x_i^{t_h}$ is the h-th discretization on the position t_h of the time series i .
The velocity of the i-th time series vector is defined as $\mathbf{v}_i{=}(v_i^{t_2},\ldots,v_i^{t_q})$, where $v_i^{t_h}{=}\frac{x_i^{t_h}-x_i^{t_{h-1}}}{t_h-t_{h-1}}$.
The acceleration of the i-th time series vector is defined as $\mathbf{a}_i{=}(a_i^{t_3},\ldots,a_i^{t_q})$, where $a_i^{t_h}{=}\frac{v_i^{t_h}-v_i^{t_{h-1}}}{t_h-t_{h-2}}$.

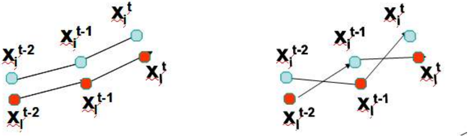
The compromise dissimilarity between the i-th and l-th time series is defined as $d^2(i,l){=}\alpha_1\|\mathbf{x}_i-\mathbf{x}_l\|^2{+}\alpha_2\|\mathbf{v}_i-\mathbf{v}_l\|^2{+}\alpha_3\|\mathbf{a}_i-\mathbf{a}_l\|^2$. It is weighted sum of three Euclidean distances. In D'urso and Vichi, the weights α of each component is determined by the global objective criterion.

Distances on Velocity and Acceleration vectors

The velocity distance is $d^2_{Ve}(i,l){=}\omega_{Ve}\sum_{j=2}^{T}\tau_j^{Ve}(v_i^t-v_l^t)^2$ where

$$(v_i^t-v_l^t)^2{=}((x_i^t-x_i^{t-1})-(x_l^t-x_l^{t-1}))^2{=}((x_i^t-x_l^t)-(x_i^{t-1}-x_l^{t-1}))^2$$

$$(v_i^t-v_l^t)^2{=}(x_i^t-x_l^t)^2{+}(x_i^{t-1}-x_l^{t-1})^2-2(x_i^t-x_l^t).(x_i^{t-1}-x_l^{t-1})$$



If $(x_i^t-x_l^t)$ and $(x_i^{t-1}-x_l^{t-1})$ are same sign (positive or negative) the velocity distance is smaller than the sum of the squared Euclidean distances between the points x_i^t and x_l^t and between the points x_i^{t-1} and x_l^{t-1}.
If $(x_i^t-x_l^t)$ and $(x_i^{t-1}-x_l^{t-1})$ are different signs the velocity distance is greater.
With acceleration distance these relationships apply to velocities.

Criterion W measures the homogeneity of the partition P on the set of tables

The goal of the clustering process is to optimize the following criterion W.

Criterion W on the partition P

$$\begin{aligned} W(P) &= \sum_{k=1}^{K} \sum_{e_i \in C_k} d^2(e_i, G_K) \\ &= \sum_{k=1}^{K} W_k = \sum_{k=1}^{K} \sum_{e_i \in C_k} \sum_{v=1}^{V} \omega_v \sum_{j=1}^{d^v} \tau_j^v d^2_{(j,v)}(e_i, g_k^{(j,v)}) \\ &= \sum_{v=1}^{V} \omega_v J_v(P) = \sum_{v=1}^{V} \omega_v \sum_{j=1}^{d^v} \tau_j^v \sum_{k=1}^{K} \sum_{e_i \in C_k} d^2_{(j,v)}(e_i, g_k^{(j,v)}) \end{aligned}$$

where

- W_k measures the homogeneity of the cluster C_k ;
- $J_v(P)$ measures the homogeneity of the table v ;

Optimization of the adequacy Criterion W

Criterion $W(P)$

$$W(P) = \sum_{v=1}^{V} \omega_v \sum_{j=1}^{d^v} \tau_j^v \sum_{k=1}^{K} \sum_{e_i \in C_k} d^2_{(j,v)}(e_i, g_k^{(j,v)})$$

Trivial solutions should be obtained :

- $\omega_v = 0, \forall v = 1$ to V
- $\tau_j^v = 0, \forall v = 1$ to V and $\forall j = 1$ to d^v

also it is necessary to introduce some constraints.

We propose the following constraint set :

Constraints proposed

$$\prod_{v\ =\ 1}^{V} \omega_v = 1 \ , \ \omega_v \geq 0$$

and

$$\prod_{j\ =\ 1}^{d^v} \tau_j^v = 1 \ , \ \tau_j^v \geq 0 \ , \ \forall v = 1 \text{ to } V$$

Dynamic Clustering Algorithm with Adaptive Weights on Times series

The objective is to **build** a partition P and a prototype G_k for each cluster and to **learn** two weight vectors Ω and Λ associated to the variable set and time series set.

Four steps are used and, during these steps, partition P, prototype set G and weight vectors Ω and Λ are estimated and change at each iteration ;

clustering algorithm

It starts with an initialization step and alternates four steps.

- ***Initialization***.

and repeat these following steps

- ***Step 1 : Build the Partition***.
- ***Step 2 : Compute the Fréchet mean set***.
- ***Step 3 : Compute the Weight vector on the Variable set***.
- ***Step 4 : Compute the Weight vector of the time series***.

until the convergence. The adequacy criterion W reaches a stationary value representing a local minimum.

Initialization Step

Parameters

- Fix K(number of clusters) ;
- Select the distance $d_{(j,v)}$ for each variable j and each time series v ;
- Randomly select K prototypes $G_k \in E$;
- Fix Λ where τ_j^v is the weight of the variable j from the table v ;
- Fix Ω where ω_v is the weight of the time series v.

$\Lambda = (\Lambda_1, \ldots, \Lambda_v, \ldots, \Lambda_V)$ where $\Lambda_v = (\tau_1^v, \ldots, \tau_j^v, \ldots, \tau_{d_v}^v)$ where τ_j^v is the weight of the variable j in the time series v.
$\Omega = (\omega_1, \ldots, \omega_V)$ is the weight vector of the time series.

Trivial initial solution

A trivial initial solution is to fix

- $\tau_j^v = 1$, $\forall v = 1$ to V and $\forall j = 1$ to d_v ;
- $\omega_v = 1$, $\forall v = 1$ to V.

Step 1 : Build the Best Partition

Fixed elements

The prototype set $(G_1, \ldots, G_K)$ and the weight vectors Ω and Λ.

This step is the classical **k-means** affectation step applied with the distance d

Update the partition $P = (C_1, \ldots, C_K)$

$$\begin{aligned} C_k &= \{e_i \in E : d^2(e_i, G_k) = \sum_{v=1}^{V} \omega_v \sum_{j=1}^{d^v} \tau_j^v d^2_{(j,v)}(e_i, g_k^{(j,v)}) \\ &\leq d^2(e_i, G_h) = \sum_{v=1}^{V} \omega_v \sum_{j=1}^{d^v} \tau_j^v d^2_{(j,v)}(e_i, g_h^{(j,v)}) \end{aligned}$$

Step 2 : Find the Best Fréchet mean set

Fixed elements

The partition $P{=}(C_1,\dots,C_K)$ and the weight vectors Ω and Λ.

This step is the classical **k-means** representation step using the Fréchet function with the distance d.

Find the $G_k \in \Phi$ that minimizes the Fréchet function

$$
\begin{aligned}
G_k &= \arg\min_{G\in\Phi} \sum_{e_i\in C_k} d^2(e_i,G) \\
g_k^{(j,v)} &= \arg\min_{g\in\Phi_j^v} \sum_{e_i\in C_k} d^2_{(j,v)}(e_i,g)
\end{aligned}
$$

G_k is a Fréchet mean or Karcher mean of C_k.
$g_k^{(j,v)}$ is a Fréchet mean of C_k on the variable j belonging to the table v.

Step 3 : Compute of the Best Weight τ_j^v of the variable set

Fixed elements

The partition $P{=}(C_1,\ldots,C_K)$ of E into K clusters, the prototype sets $(G_1,\ldots,G_K)$ and the weight vector Ω of the table set.

For each view v we use the Lagrange formula :

$$L(\Lambda^v,\lambda){=}J(\omega,\tau,P){+}\lambda(\prod_{l\,=\,1}^{d^v}\tau_j^v-1)$$

The weight τ_j^v that minimizes the criterion W under the constraints $\tau_j^v{>}0$ and $\prod_l^{d^v}{=}_1\,\tau_j^v\;=1$ is

$$\tau_j^v{=}\frac{\prod_{u=1}^{d^v}J_u^v(P)^{1/d^v}}{J_j^v(P)}$$

where $J_j^v(P){=}\sum_{k=1}^K\sum_{e_i\in C_k}d^2_{(j,v)}(e_i,g_k^{(j,v)})$ measures the homogeneity of the variable j in the table v.

Step 4 : Compute of the Best Weight Ω of the time series

Fixed elements

The partition $P{=}(C_1,\ldots,C_K)$ of E into K clusters, the prototype sets $(G_1,\ldots,G_K)$ and the weight vector Λ of the variable set.

We use the Lagrange formula :

$$L(\omega^v,\lambda){=}J(\omega,\tau,P){+}\lambda(\prod_{l\;=\;1}^{V}\omega_v-1)$$

The solution is the weight ω^v that minimizes the criterion W under the constraints $\omega^v{>}0$ and $\prod_l^V=_1\omega^v=1$:

$$\omega^v{=}\frac{\prod_{u=1}^{V}J_u(P)^{1/V}}{J_v(P)}$$

where $J_v(P){=}\sum_{j=1}^{d^v}\tau_j^v J_j^v(P){=}\sum_{j=1}^{d^v}\tau_j^v\sum_{k=1}^{K}\sum_{e_i\in C_k}d^2_{(j,v)}(e_i,g_k^{(j,v)})$ measures the homogeneity of the time series v

1. Introduction

2. Choice of the D'urso and Vichi distance between time series

3. Dynamic Clustering Algorithm with Adaptive Weights

4. Applications : Solar Irradiance data sets and Satellite time trajectory
   - Application I : Solar Irradiance
   - Application II : Satellite time trajectory

5. References

The RCI-GS project

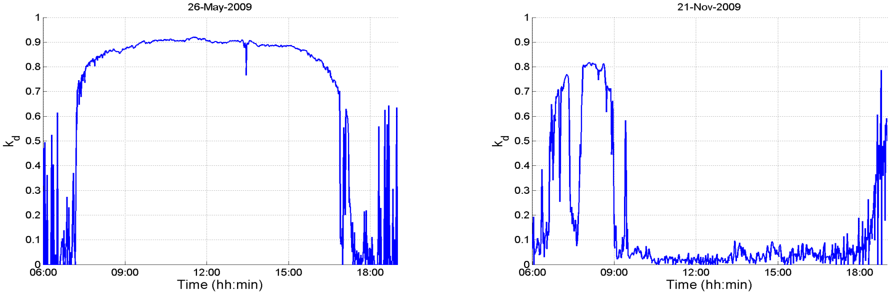
The development of grid-connected photovoltaic power systems leads to new challenges. The short or medium term prediction of the solar irradiance is definitively a solution to reduce the storage capacities.

The aim of the RCI-GS project (P.O.E. 2007-2013 / FEDER) is to predict the solar irradiance at a short and medium term for the Reunion Island.

Following the studies on clustering recently performed by (Muselli et al. (2000)) and (Soubdhan et al. (2009)) it is necessary to have a clustering approach before prediction modeling.

Solar Irradiance data set

(Jeanty et al. (2013)) propose the solar radiance measure k_b given by $k_b = \frac{Direct_{Radiation}}{Global_{Radiation}}$ which integrates all environmental influencing factors. When k_b is close to 1, the direct radiation level is close to the global radiation level, indicating we are in presence of a sunny day. A value close to 0 is the signature of a very cloudy day.



Clear Sky Day

Intermittent Cloudy Day

FIGURE – Two Day Examples Described by k_b Measure

Solar irradiance data : clustering on dissimilarity matrices

Two methods are applied to three dissimilarity matrices (position, velocity, acceleration) presented in De Carvalho *et al.*(2012).

- *MRDCA – RWG* Dynamic clustering algorithm with weight for each dissimilarity matrix estimated globally.
- *MRDCA – RWL* Dynamic clustering algorithm with weight for each dissimilarity matrix estimated for each cluster (locally estimated) ;

Table	MRDCA-SWG	MRDCA-SWL				
		C1	C2	C3	C4	C5
Position	2.14	2.10	2.19	1.87	1.97	2.55
Velocity	0.78	0.80	0.77	0.81	0.78	0.71
Acceleration	0.59	0.58	0.58	0.65	0.64	0.54

TABLE – Vector or Matrix of Relevance Weight

The position dissimilarity matrix has the highest relevant weight for the five clusters of MRDCA-SWL method also we selected MRDCA-SWG solution where the weight of the position dissimilarity matrix has also the highest weight.

Solar irradiance data : clustering results

Five clusters are obtained by $MRDCA-RWL$ method.

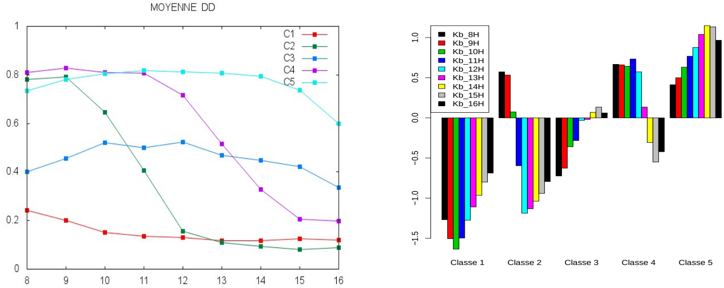


FIGURE – K_b average and t_k for each class of D'urso and Vichi partition

- C1 : *Cloudy days* (15 %)
- C2 : *Intermittent bad days* (20 %)
- C3 : *Disturbed days* (14 %)
- C4 : *Intermittent good days* (24 %)
- C5 : *Clear sky days* (27 %)

Solar irradiance data : clustering on time series

Three methods are applied to three time series views (position, velocity, acceleration).

- $MND - V$ Dynamic clustering algorithm with weight for each time series view ;
- $MND - W$ Dynamic clustering algorithm with weight for each variable ;
- $MND - WV$ Dynamic clustering algorithm with weight for each variable and for each time series view.

σ=1 = Data are normalized σ#1 = Data are not normalized

Time series	MND-V		MND-WV	
	σ=1	σ#1	σ=1	σ#1
Position	1.50	1.43	1.37	0.87
Velocity	0.80	0.95	0.81	1.45
Acceleration	0.83	0.72	0.89	0.78

TABLE – Weights of the time series

The position has the highest weight when the data are normalized and shows that this view is important in this classification. When the data are not normalized the interpretation of weights is impossible because it depends on the variances of these views.

Solar irradiance data set : Validation

- σ=1 Yes= Data are not normalized, No = Data are normalized ;
- Λ=1 Yes= The weights of the variables are fixed, No = The weights of the variables are adaptive ;
- Ω=1 Yes= The weights of the tables are fixed, No = The weights of the tables are adaptive ;
- *Criterion* : minimum value of the criterion observed in the run set ;

Methods	Parameters			Results
	σ=1	Λ=1	Ω=1	*Criterion*
K-means(1)	No	Yes		14925.73
(2)	Yes	Yes		873.27
$MND-W$(1)	No	No	Yes	14189.48
(2)	Yes	No	Yes	844.65
$MND-V$ (1)	No	Yes	No	14418.76
(2)	Yes	Yes	No	840.96
$MND-WV$ (1)	No	No	No	13899.29
(2)	Yes	No	No	816.28

The efficient strategy is to select an variable and table weighting approach ($MND-WV$ (1) and (2)). The role of normalization is not determinant. The optimal partitions are the same for $MND-WV$ (1) and (2) approaches.

Solar irradiance data set : Cluster averages

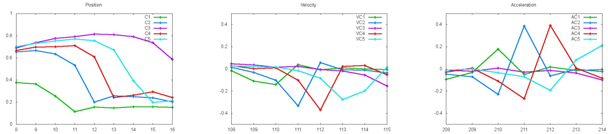


FIGURE – Average clusters : Position, Velocity, Acceleration

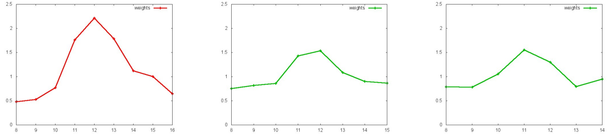


FIGURE – Weight vectors : Position, Velocity, Acceleration

The weights between 11h and 13h are high.

Solar irradiance data set : comparison between $MND - WV$ and $MRDCA - SWL$ approaches

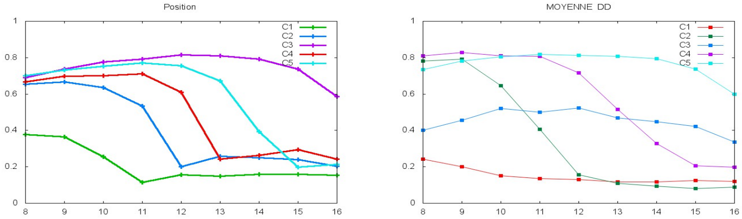


FIGURE – Clusters : $MND - WV$ and $MRDCA - SWL$

- C1 : *Cloudy days* (15 %) Clusters are similar
- C2 : *Intermittent bad days* (20 %) Clusters are similar
- C4 : *Intermittent average days* (24 %) *Disturbed days* (24 %)
- C5 : *Intermittent good days* (27 %) $MRDCA - SWL$ (24 %)
- C3 : *Clear sky days* (14 %) $MRDCA - SWL$ (27 %)

The average curves between *Intermittent average days* and *Disturbed day* are different.

1. Introduction

2. Choice of the D'urso and Vichi distance between time series

3. Dynamic Clustering Algorithm with Adaptive Weights

4. Applications : Solar Irradiance data sets and Satellite time trajectory
   - Application I : Solar Irradiance
   - Application II : Satellite time trajectory

5. References

Satellite data set

This data set concerns 472 radar waveforms. The data were registered by the Topex/Poseidon satellite upon the Amazon River.
Each object (time trajectory) is represented by 70 discretization positions.
We propose to use the D'urso and Vichi distance that is a compromise dissimilarity between these 3 relational data tables (position, velocity and acceleration) where the weights are learned by our method.

- Two methods are applied to three dissimilarity matrices (position, velocity, acceleration) presented in De Carvalho *et al.*(2012).

  $MRDCA-RWG$ Dynamic clustering algorithm with weight for each dissimilarity matrix estimated globally.
  $MRDCA-RWL$ Dynamic clustering algorithm with weight for each dissimilarity matrix estimated for each cluster (locally estimated) ;
- Two methods are applied to three time series views (position, velocity, acceleration).It is Our approach.

  $MND-V$ Dynamic clustering algorithm with weight for each time series view ;
  $MND-WV$ Dynamic clustering algorithm with weight for each variable and for each time series view.

For each K, the clustering algorithm is run 100 times and the best result according to the adequacy criterion is selected.

Satellite data set : vector or matrix of weights of the dissimilarity matrices

Data Matrix	$MRDCA - RWG$	$MRDCA - RWL$						
		C1	C2	C3	C4	C5	C6	C7
Position	1.4309	3.43	0.78	1.67	2.30	2.60	2.16	0.75
Velocity	0.8447	0.60	1.13	0.76	0.67	0.64	0.70	1.06
Acceleration	0.8272	0.48	1.11	0.78	0.64	0.59	0.65	1.25

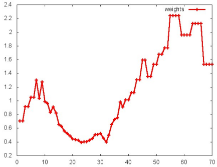
The partition given by $MRDCA - RWG$, position dissimilarity matrix has the highest relevant weight.

For clusters 1, 3, 4, 5 and 6 of the partition given by $MRDCA - RWL$, position dissimilarity matrix has the highest relevant weight.
For cluster 2 and 7 velocity and acceleration dissimilarity matrices have the highest relevant weights.

Satellite data set : vector or matrix of weights by our approach

Time series	$MND-V$	$MND-WV$
Position	1.49	1.57
Velocity	0.94	0.90
Acceleration	0.72	0.71

The weight associated to position view has the highest and allowing weight positions to be adjusted does not change the result. The weights of the positions are very different.



After the 45th position the weights are high and between 5th and 15th position the weights are important.

Satellite data set : Cluster means

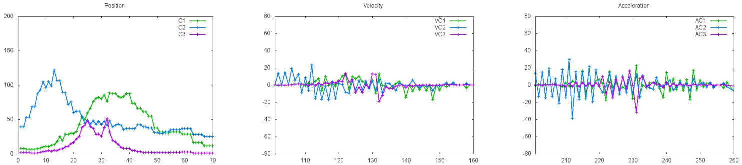


FIGURE – C1, C2 and C3 clusters : Position, Velocity, Acceleration

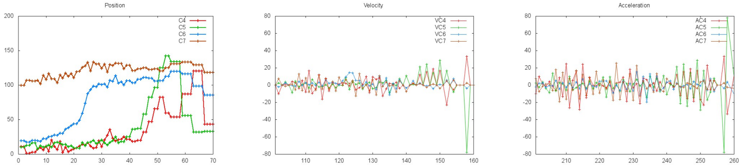


FIGURE – C4 , C5, C6 and C7 clusters : Position, Velocity, Acceleration

Conclusion

- we have proposed a **novel multi-table clustering methods** which simultaneously perform adaptive weights for variables and tables and use batch K-means like algorithm ;
- for each variables it is possible to choose a specific distance ;
- A global criterion is proposed and this criterion is optimized by our clustering process ;
- Experiments have been conducted on two real datasets.
- In the future we propose to add a new step where the adaptive weights are computed locally for each class.

References I (others approaches)

- H. Frigui, C. Hwang, F. C.-H. Rhee, *Clustering and aggregation of relational data with applications to image database categorization*. Pattern Recognition 40, 3053-3068, (2007).
- G. Cleuziou, M. Exbrayat, L. Martin, and J.-H. Sublemontier, *Cofkm : A centralized method for multiple-view clustering*. In ICDM 2009 Ninth IEEE International Conference on Data Mining, Miami, USA, pp. 752-757, (2009).
- G. Tzortzis,A. Likas : *Kernel-based weighted multi-view clustering* In :Proceedings of the 12th International Conference on Data Mining, 675Ű 684, (2012).
- A. Kumar,P. Rai,H. Daumé : *Co-regularized multi-view spectral clustering*. In : Neural Information Processing Systems(NIPS), 1413-1421, (2011).
- H. Zeng,Y.-m. Cheung : *Feature selection and kernel learning for local learning- based clustering*. IEEE Trans, Pattern Anal.Mach.Intell. 33, 1532-1547 (2011).

References II

(Thank You !)

- E. Diday, and G. Govaert, *Classification automatique avec distances adaptatives*. R.A.I.R.O. Informatique Computer Science 11(4), 329-349, (1977)
- M. Chavent, F. A. T. De Carvalho, Y. Lechevallier, and R. Verde, *New clustering methods for interval data*. Computational Statistics 21(2), 211-229, (2006)
- F. A. T. De Carvalho, Y. Lechevallier, and F. M. De Melo, *Partitioning hard clustering algorithms based on multiple dissimilarity matrices*. Pattern Recognition 45(1), 447-464, (2012)
- F. A. T. De Carvalho, Y. Lechevallier, F.M. de Melo, *Relational partitioning fuzzy clustering based on multiple dissimilarity matrices*, Fuzzy Sets Syst. 215 , 1-28, (2013)
- F. A. T. De Carvalho, F.M. de Melo, Y. Lechevallier, *A multi-view relational fuzzy c-medoid vectors clustering algorithm*, Neurocomputing. 163, 115-123, (2015)