Generalizations of Fuzzy C-Means Algorithm to Granular Feature Spaces, based on Underlying Metrics: Issues and Related Works

Vasile Georgescu

Faculty of Economics and Business Administration, University of Craiova
Craiova, Romania
Email: vgeo@central.ucv.ro, v_geo@yahoo.com

Abstract—*This paper considers dissimilarity measures and clustering techniques for two special cases of set-defined objects: fuzzy granules and subsequence time series. To deal with clustering of such kind of objects, we propose two implementations that generalize the Fuzzy C-Means algorithm to granular feature spaces. Granular computing is a paradigm oriented towards capturing and processing meaningful pieces of information, the so-called information granules. In a granular feature space, such as a space populated with p-dimensional fuzzy granules, we are concerning with both granular data samples and granular centroids (center of clusters). In order to accommodate clustering algorithms to work in a granular environment we have to choose and/or define appropriate metrics and descriptors. Either a crisp distance between granules or the defuzzified value of a fuzzy distance has to be chosen. On the other hand, subsequence time series clustering requires a generalization of Fuzzy C-Means algorithm in a similar way. It involves a set-defined centroid and appropriate dissimilarity measures to determine the degree to which time sequences are different from their centroid. Furthermore, we discuss related work in granular clustering and subsequence time series clustering.*

Keywords—Granular clustering, Metrics in granular feature spaces, Fuzzy C-Means clustering, Agglomerative granular clustering algorithms, Subsequence time series clustering.

1 *P*-dimensional fuzzy granules: representation and cardinality measure

In what follows, the attention will be restricted to the class Φ^p of normal fuzzy convex granules on $\Re^p$, whose α-level sets are nonempty compact convex sets for all $\alpha > 0$. Each fuzzy granule $M \in \Phi^p$ is uniquely characterized by its support function $s_M(u,\alpha) = \sup\{ <u,x> \mid x \in M^\alpha \}$, $u \in S^{p-1}, \alpha \in (0,1]$, where S^{p-1} is the (p-1)-dimensional unit sphere of $\Re^p$ (i.e. $\|u\| = 1$) and $<\cdot,\cdot>$ is the inner product in $\Re^p$. The support function is a mapping from the class of fuzzy sets Φ^p into the space of functions $L\left(S^{p-1} \times [0,1]\right)$, which preserves addition and multiplication with non-negative scalars.

A p-dimensional fuzzy granule $A = A^1 \times \cdots \times A^p$ is defined on the product space $X = X_1 \times \cdots \times X_p$. Representation and cardinality measure for A can be given either in terms of membership functions:

$$\mu_A : \Re^p \to [0,1],$$
$$\mu_A(x) = \min\left(\mu_{A^1}(x_1), \ldots, \mu_{A^p}(x_p) \right), \quad \forall\, x \in X$$

$$|A| = \int_{\xi \in Supp(A)} \mu_A(x)\, dx =$$
$$= \int_{x_1 \in Supp(A^1)} \cdots \int_{x_p \in Supp(A^p)} \min\left(\mu_{A^1}(x_1), \ldots, \mu_{A^p}(x_p)\right) dx_1 \ldots dx_p$$

or in terms of $\alpha -$ level intervals:

$$A(\alpha) = A^1(\alpha) \times \cdots \times A^p(\alpha); \qquad \alpha \in [0,1]$$

$$|A| = \int_0^1 |A(\alpha)|\, d\alpha = \int_0^1 \prod_{j=1}^{p} \left|A^j(\alpha)\right|\, d\alpha.$$

In particular, let A^j be an LR-fuzzy set. We have:

$$A^j(\alpha) = \left[A^{jL}(\alpha), A^{jR}(\alpha)\right]_{LR}$$
$$|A| = \int_0^1 \prod_{j=1}^{p} \left|A^{jR}(\alpha) - A^{jL}(\alpha)\right| d\alpha.$$

The latest integral can be evaluated numerically by means of a quadrature formula (e.g. the adaptive Simpson quadrature).

2 Metrics in granular feature spaces

2.1 *A few examples*

We focus on distances in granular feature spaces populated with fuzzy granules. There is an enormous literature that covers this topic (at least in case of one-dimensional fuzzy spaces). There is also a large diversity of approaches, among which we can distinguish:

a) Membership focused distances (vertical)
b) Spatially focused distances (horizontal)
c) Mix of spatial and membership distances (tolerance)
d) Feature distances (low or high dimensional representations)
e) Morphological (mixed focus)

Most of these metrics are defined as crisp distances. As a typical example, we can consider an L_2-metric on the space of normal compact convex fuzzy sets Φ^p using the L_2-metric on the Hilbert space of square-integrable functions

ISBN: 978-989-95079-6-8

$L_2\left(S^{p-1} \times [0,1]\right)$. This space is equipped with the inner product

$$< M, N >_\lambda = p \cdot \int_{[0,1]} \int_{S^{p-1}} s_M(u,\alpha)\ s_N(u,\alpha)\ \mu(du)\ \lambda(d\alpha).$$

where μ is normalized Lebesgue measure on S^{p-1} (i.e. $\mu(S^{p-1}) = 1$) and λ is normalized Lebesgue measure on $[0,1]$. Now, the corresponding L_2-metric results of the following form:

$$\delta_2^{(\lambda)}(M, N) = \left\| s_M - s_N \right\|_2 =$$
$$= \left(p \cdot \int_{[0,1]} \int_{S^{p-1}} | s_M(u,\alpha) - s_N(u,\alpha) |^2 \mu(du)\ \lambda(d\alpha) \right)^{1/2}.$$

Set-defined distances (such as fuzzy distances) have been also proposed ([1]).

For example, a *fuzzy distance* of two non-empty one-dimensional fuzzy sets A and B is defined as the fuzzy set $D_f(A, B) = \left(\Re^+, \mu_{D_f}(A,B)\right)$ with the membership function

$$\mu_{D_f(A,B)}(y) = \begin{cases} \sup\{\alpha \in [0,1] \mid A_\alpha, B_\alpha \neq \emptyset;\ \rho(A_\alpha, B_\alpha) \leq y\} \\ \quad for\ y \leq \lim\limits_{\beta \to \min\{h(M), h(N)\}_-} \rho(A_\beta, B_\beta) \\ 0 \qquad\qquad otherwise \end{cases}$$

A generalization of a fuzzy distance to the *p*-dimensional feature space is easiest if we define fuzzy granules as a Cartesian product of fuzzy sets, where the expression for the fuzzy distance can be simplified using distances in each dimension. A simplification can be obtained by choosing some particular kinds of fuzzy sets.

2.2 *A fuzzy distance between p-dimensional fuzzy granules*

In [5] we introduced a fuzzy distance between *p*-dimensional fuzzy granules defined as a Cartesian product of LR-fuzzy sets. Let us consider a p-dimensional feature space $X = X_1 \times \cdots \times X_p$ and a configuration of n objects, each *one described by a p-dimensional fuzzy granule. Two arbitrary objects u and v are imprecisely located in the feature space by means of two fuzzy granules:* $A_u = A_{u1} \times \cdots \times A_{up}$ *and* $A_v = A_{v1} \times \cdots \times A_{vp}$. *Assume the components* A_{uk} *and* A_{vk} *(*$k = 1, \ldots, p$*) belonging to the* k^{th} *dimension of* X *are LR-fuzzy sets, i.e.* $A_{uk}(\alpha) = (A_{uk}^L(\alpha), A_{uk}^R(\alpha))$ *and* $A_{vk}(\alpha) = (A_{vk}^L(\alpha),\ A_{vk}^R(\alpha))$. *Therefore, the following inequalities hold:*

$$A_{uk}^L(\alpha) \leq x_{uk}(\alpha) \leq A_{uk}^R(\alpha)\ ; \quad A_{vk}^L(\alpha) \leq x_{vk}(\alpha) \leq A_{vk}^R(\alpha)\ .$$

By subtracting $x_{vk}(\alpha)$ from $x_{uk}(\alpha)$, we obtain

$$A_{uk}^L(\alpha) - A_{vk}^R(\alpha) \leq x_{uk}(\alpha) - x_{vk}(\alpha) \leq A_{uk}^R(\alpha) - A_{vk}^L(\alpha)$$

and then, using well-known properties such as $\min\limits_{w \leq x \leq z} | x | = \max(0, w, -z)$ and $\max\limits_{w \leq x \leq z} | x | = \max(w, -z)$, respectively, the range of $\left| x_{uk}(\alpha) - x_{vk}(\alpha) \right|$, for any α in $[0,1]$, can be written as follows:

$$\left| x_{uk}(\alpha) - x_{vk}(\alpha) \right| \in$$
$$\Big[\max\left(0,\ A_{uk}^L(\alpha) - A_{vk}^R(\alpha),\ A_{vk}^L(\alpha) - A_{uk}^R(\alpha)\right),$$
$$\max\left(A_{uk}^R(\alpha) - A_{vk}^L(\alpha),\ A_{vk}^R(\alpha) - A_{uk}^L(\alpha)\right) \Big].$$

Now, the identity

$$\max(w, z) = \frac{1}{2}\left(w + z + | w - z |\right),\ \forall w, z \in \Re,$$

allows us to derive the left and right α – bounds of a granular distance, namely the fuzzy distance. We have:

$$d_{uv}^{\min}(\alpha) = \frac{1}{4}\Bigg(\sum_{k=1}^{p} \Big[A_{uk}^L(\alpha) - A_{vk}^R(\alpha) + A_{vk}^L(\alpha) - A_{uk}^R(\alpha) +$$
$$+ | A_{uk}^L(\alpha) - A_{vk}^R(\alpha) - A_{vk}^L(\alpha) + A_{uk}^R(\alpha) | +$$
$$+ \Big| A_{uk}^L(\alpha) - A_{vk}^R(\alpha) + A_{vk}^L(\alpha) - A_{uk}^R(\alpha) +$$
$$+ | A_{uk}^L(\alpha) - A_{vk}^R(\alpha) - A_{vk}^L(\alpha) + A_{uk}^R(\alpha) | \Big| \Big]^2 \Bigg)^{\frac{1}{2}}$$

$$d_{uv}^{\max}(\alpha) = \frac{1}{2}\Bigg(\sum_{k=1}^{p} \Big[A_{uk}^R(\alpha) - A_{vk}^L(\alpha) + A_{vk}^R(\alpha) - A_{uk}^L(\alpha) +$$
$$+ | A_{uk}^R(\alpha) - A_{vk}^L(\alpha) - A_{vk}^R(\alpha) + A_{uk}^L(\alpha) | \Big]^2 \Bigg)^{\frac{1}{2}}.$$

Finally, the fuzzy distance between the two p-dimensional fuzzy granules A_u and A_v results in the form of a LR-fuzzy set:

$$d_{uv}(\alpha) = \left(d_{uv}^{\min}(\alpha),\ d_{uv}^{\max}(\alpha)\right) = \left(d_{uv}^L(\alpha),\ d_{uv}^R(\alpha)\right).$$

A point-wise distance can then be obtained by defuzzifying the fuzzy distance d_{uv}, e.g. by computing its centroid. An alternative way may be that of averaging the α-level intervals $[d_{uv}^L(\alpha),\ d_{uv}^R(\alpha)]$ over $[0,1]$:

$$Defuzz(d_{uv}) = E\left[d_{uv}(\alpha)\right] = \frac{1}{2} \int_{[0,1]} \left| d_{uv}^R(\alpha) + d_{uv}^L(\alpha) \right| \lambda(d\alpha).$$

where λ is a normalized Lebesgue measure on $[0,1]$, e.g.

$$\lambda([0,1]) = \int_0^1 w(\alpha)\ d\alpha = 1, \text{ where } w(\alpha) = 2\alpha\ .$$

Remark: The fuzzy distance between two p-dimensional fuzzy granules A_u and A_v with empty intersection, whose components A_{uk} and A_{vk} are trapezoidal fuzzy sets for all $k = 1, ..., p$, is still a trapezoidal fuzzy set. This is because, in

ISBN: 978-989-95079-6-8

such a case, $d_{uv}^{L}(\alpha)$ is an increasing function that interpolates linearly between $d_{uv}^{L}(0)$ and $d_{uv}^{L}(1)$ whereas $d_{uv}^{R}(\alpha)$ is a decreasing function that interpolates linearly between $d_{uv}^{R}(1)$ and $d_{uv}^{R}(0)$ (see Fig. 1).

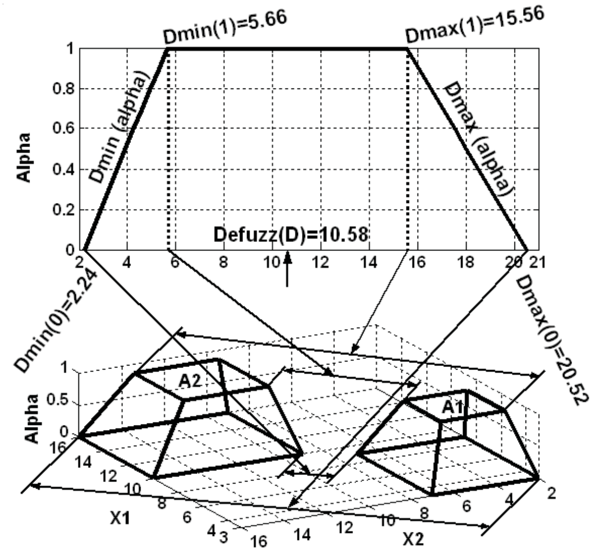


Figure 1: Two-dimensional fuzzy granules, the fuzzy distance between them and its defuzzified value.

The fuzzy distance between pairs of fuzzy granules with non-empty intersection, is linear-shaped, but not necessarily trapezoidal-shaped (see Fig. 2).

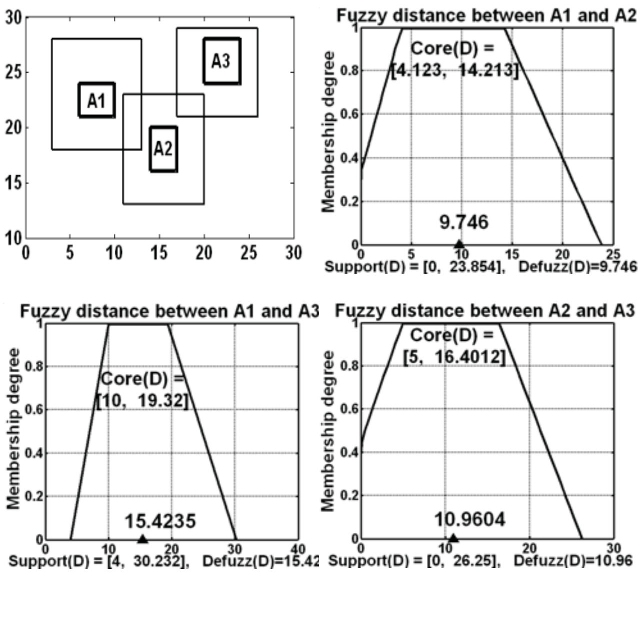


Figure 2: Two-dimensional fuzzy granules with non-empty intersection and the pair-wise fuzzy distances.

Obviously, our formal definition of fuzzy distance introduced above applies similarly to fuzzy granules of any finite dimension p. Fig. 3 illustrates the case of three-dimensional fuzzy granules.

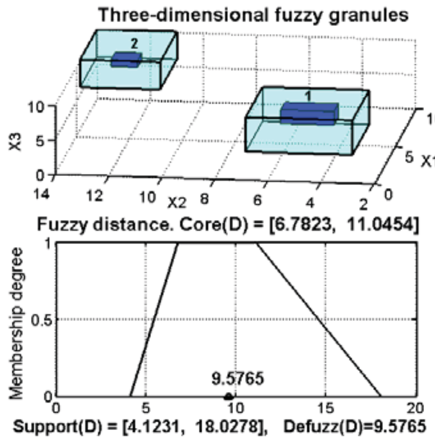


Figure 3: Three-dimensional fuzzy granules, the fuzzy distance and its defuzzified value.

3 Granular Clustering

3.1 A generalization of fuzzy c-means algorithm to granular feature spaces

We propose a generalization of Fuzzy C-Means Algorithm to fuzzy feature spaces, which leads to adopting a well defined distance between p-dimensional fuzzy granules (either a crisp distance, or the defuzzified value of the fuzzy distance defined above).

The Fuzzy C-means clustering algorithm is based on the minimization of an objective function called *C-means functional*:

$$J(X;U,V) = \sum_{i=1}^{c}\sum_{k=1}^{N}(u_{ik})^{m}\|x_k - v_i\|_A^2$$

where

$$V = (v_1, \ldots, v_c), \quad v_i \in \Re^n$$

is a vector of *cluster prototypes* (centers), which have to be determined, and

$$D_{ikA}^2 = \|x_k - v_i\|_A^2 = (x_k - v_i)' A (x_k - v_i)$$

is a squared inner-product distance norm.

In a granular feature space, such as a space populated with p-dimensional fuzzy granules, x_k are granular samples and v_i are granular centroids.

Remark. The granular centroid of the set $\{A_u\}_{u=1,\ldots,n}$ of LR-type p-dimensional fuzzy granules is still an LR-type p-dimensional fuzzy granule.

Fig. 4 shows a set $\{A_1, A_2, A_3\}$ of 3-dimensional fuzzy granules, their granular centroid G, as well as the three pair-wise granular centroids G_{12}, G_{13} and G_{23}, respectively.

Furthermore, the distance D_{ikA}^2 must be replaced with a well defined square distance between p-dimensional fuzzy granules. Since the fuzzy distance introduced above have been defined as a square distance, we can use the square value of the fuzzy distance in order to replace D_{ikA}^2.

ISBN: 978-989-95079-6-8

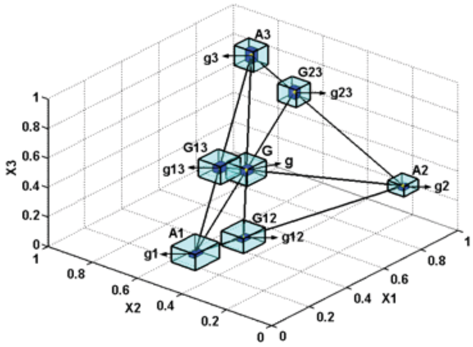


Figure 4. Granular and point-wise weighted centroids of 3D fuzzy granules (the granular version of median theorem).

For testing the granular clustering algorithm we started from a set of crisp data (Fig. 5).

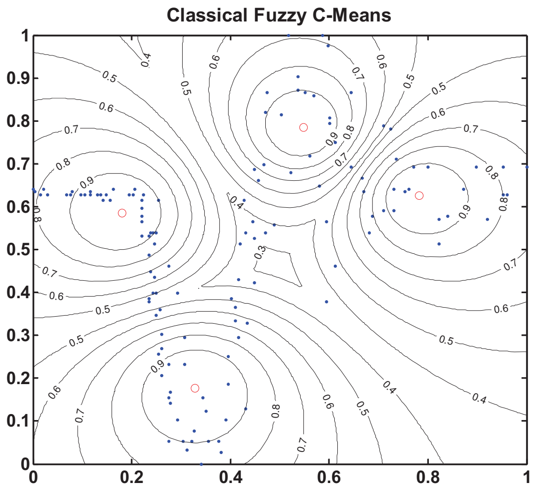


Figure 5: Fuzzy C-means clustering for crisp data.

Then, we transformed these initial data into two-dimensional fuzzy granules, i.e. as a Cartesian product of trapezoidal shaped fuzzy sets on each dimension. Fig. 6 shows 4 fuzzy clusters populated with two-dimensional fuzzy granules

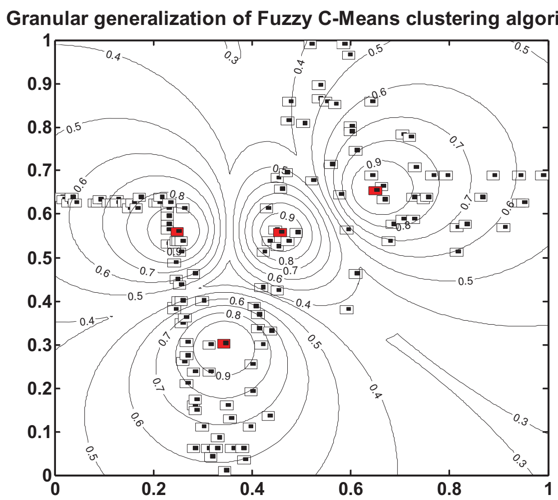


Figure 6: Granular fuzzy C-means clustering.

3.2 Related work

In ([4], [5]) we proposed an agglomerative granular clustering method that generalizes the method in [12], originally proposed by Pedrycz and Bargiela. Since resorting to a different granulation framework (fuzzy set theory instead of interval mathematics) involves a refinement of the metric formalism, the fuzzy distance has been used for allowing fuzzy granules (rather than hyperboxes) to be processed during the cluster growing process. The main step of the iterative process is finding the two closest information granules in order to aggregate them into a more comprehensive one. Let $C = agg(A, B)$ be the resulting granule. In terms of p-dimensional LR-fuzzy granules, the aggregation process can be carried out as follows:

$$C = C_1 \times \cdots \times C_p$$

with $C_i^{\alpha} = \left(x_{C_i}^L - \ell_{C_i} \cdot L^{-1}(\alpha),\ x_{C_i}^R + r_{C_i} \cdot R^{-1}(\alpha) \right).$

The core and the support of C_i is obtained as

$$C_i^1 = \left(x_{C_i}^L, x_{C_i}^R \right) = \left(\min\,(x_{A_i}^L, x_{B_i}^L),\quad \max\,(x_{A_i}^R, x_{B_i}^R) \right)$$

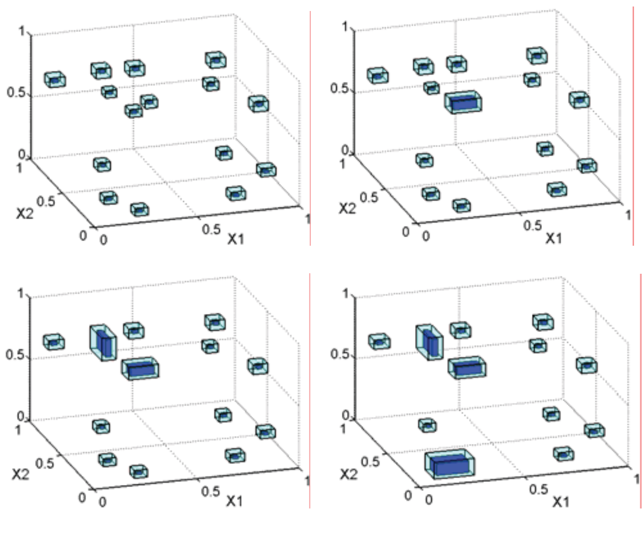
$$C_i^0 = \left(x_{C_i}^L - \ell_{C_i}, x_{C_i}^R + r_{C_i} \right) = \\ = \left(\min(x_{A_i}^L - \ell_{A_i}, x_{B_i}^L - \ell_{B_i}, \max(x_{A_i}^R + r_{A_i}, x_{B_i}^R + r_{B_i}) \right).$$

The compatibility measure guiding the search for the two closest fuzzy granules can now be defined as

$$compat\,(A, B) = 1 - Defuzz\big(d(A, B)\big) \cdot e^{-\beta |C|}$$

where $|C| = \int_0^1 \left| C^{\alpha} \right| d\alpha = \int_0^1 \prod_{i=1}^{p} \left| C_i^{\alpha} \right| \, d\alpha$ is the cardinality of C and β is a tuning coefficient.

Maximizing the compatibility measure means that the pair of candidate fuzzy granules to be clustered should not only be close enough (i.e., the distance between them should be small), but the resulting granule should be compact (i.e., its expansion along every direction must be well-balanced). Fig. 7 shows the sequence of cluster growing over the granular clustering process.



ISBN: 978-989-95079-6-8

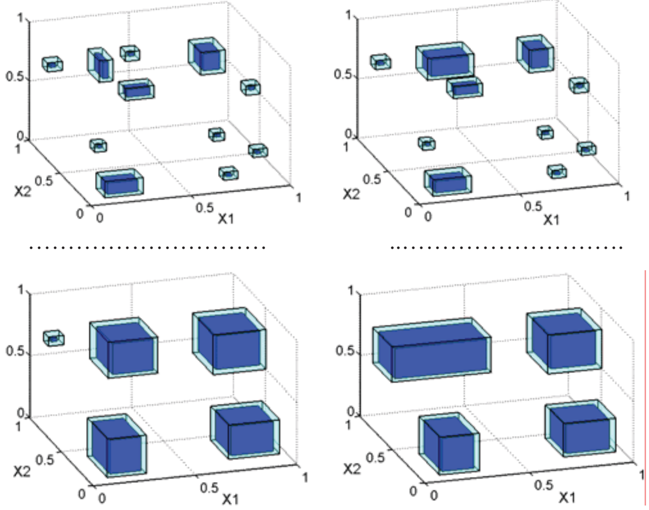


Figure 7: The sequence of cluster growing over the granular clustering process.

4 Time Series Clustering

4.1 Time series (dis)similarity

Roughly speaking, a granule is a set-defined object. In these terms, a subsequence time series can be viewed as a special kind of granule.

Let $S = y_m, \ldots, y_{m+w-1}$ be a subsequence with length w of time series $Y = y_1, \ldots, y_n$, where $1 \leq m \leq n - w + 1$. Subsequences will be represented as vectors in a w-dimensional vector space.

Clustering of subsequence time series requires the definition of a similarity or of a distance measure. Examples of dissimilarity measures are based on Euclidean norms, piecewise linear approximations, dynamic time warping, longest common subsequences, and probabilistic similarity models.

The usage of the Euclidean distance is subject to the constraint that both time subsequences are of the same length w. Thus we can define the dissimilarity between sequences S_1 and S_2 as $L_p(S_1, S_2)$, that is, the distance between the two $w-$dimensional vectors measured by the L_p norm (when $p = 2$, this reduces to the familiar Euclidean distance).

Some major disadvantages of the Euclidean distance are as follows: it does not allow for different *baselines* in the time sequences; it is very sensitive to *phase shifts* in time; it does not allow for *acceleration and deceleration* along the time axis.

Other specialized distance measures have been described for time series clustering, such as Dynamic Time Warping, DTW, and Longest Common Subsequence Similarity, LCSS.

Dynamic Time Warping (DTW) is an extensively used technique in speech recognition and allows acceleration–deceleration of signals along the time dimension.

Following [11], let us consider two sequences (of possibly different lengths) $Q = \{q_1, \ldots, q_n\}$ and $C = \{c_1, \ldots, c_m\}$.

To align two sequences using DTW, we construct an n-by-m matrix where the (i^{th}, j^{th}) element of the matrix contains the distance $d(q_i, c_j)$ between the two points q_i and c_j (i.e. $d(q_i, c_j) = (q_i - c_j)^2$). Each matrix element (i, j) corresponds to the alignment between the points q_i and c_j. A warping path W is a contiguous set of matrix elements that defines a mapping between Q and C. The k^{th} element of W is defined as $w_k = (i, j)_k$, so we have:

$$W = w_1, w_2, \ldots, w_k, \ldots, w_K, \qquad \max(m, n) \leq K < m + n - 1.$$

The warping path is typically subject to several constraints.

- Boundary conditions: $w_1 = (1, 1)$ and $w_K = (m, n)$.
- Continuity: Given $w_k = (a, b)$, then $w_{k-1} = (a', b')$, where $a - a' \leq 1$ and $b - b' \leq 1$. This restricts the allowable steps in the warping path to adjacent cells.
- Monotonicity: Given $w_k = (a, b)$, then $w_{k-1} = (a', b')$, where $a - a' \geq 0$ and $b - b' \geq 0$. This forces the points in W to be monotonically spaced in time.

There are exponentially many warping paths that satisfy the above conditions, however we are interested only in the path which minimizes the warping cost:

$$DTW(Q, C) = \min\left(\sqrt{\sum_{i=1}^{K} w_k}\right) \Big/ K.$$

The K in the denominator is used to compensate for the fact that warping paths may have different lengths.

This path can be found efficiently using dynamic programming to evaluate the following recurrence, which defines the cumulative distance $\gamma(i, j)$ as the distance $d(i, j)$ found in the current cell and the minimum of the cumulative distances of the adjacent elements:

$$\gamma(i, j) = d(q_i, c_j) + \min\{\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)\}.$$

The Euclidean distance between two sequences can be seen as a special case of DTW where the k^{th} element of W is constrained such that $w_k = (i, j)_k$, $i = j = k$.

The warping path is also constrained in a global sense by limiting how far it may stray from the diagonal. The subset of the matrix that the warping path is allowed to visit is called the warping window. The two most common constraints in the literature are the Sakoe-Chiba band and the Itakura parallelogram. We can view a global or local constraint as constraining the indices of the warping path $w_k = (i, j)_k$, such that $j - r \leq i \leq j + r$, where r is a term defining the allowed range of warping, for a given point in a sequence. In the case of the Sakoe-Chiba band, r is independent of i; for the Itakura parallelogram, r is a function of i.

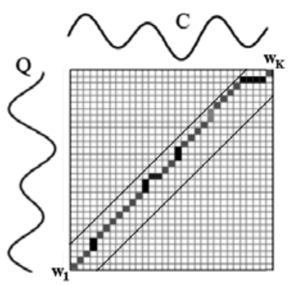


Figure 8: Optimal warping path with the Sakoe-Chiba band as global constraints.

ISBN: 978-989-95079-6-8

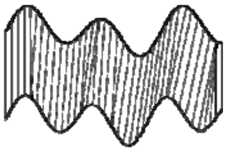


Figure 9: Aligning two time sequences using DTW.

4.2 *Non-overlapping subsequence time series clustering*

Our implementation of Fuzzy C-Means algorithm for non-overlapping subsequence time series clustering is based essentially on using the DTW distance, which is proved to be largely superior to the Euclidian distance.

We also report a real-world application of this method for clustering non-overlapping subsequences of length 30 from Bucharest Stock Exchange Bet Index. The dataset contains 2210 daily values. Fig. 10 shows ten clusters, which consist of several subsequence time series, grouped around a cluster centroid.

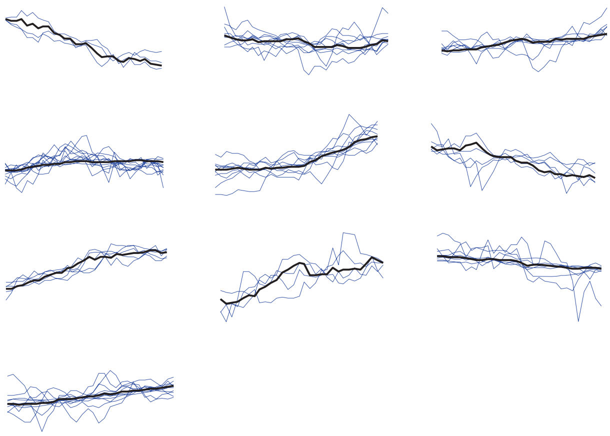


Figure 10: Non-overlapping subsequence time series of length 30: clusters and centroids obtained by fuzzy c-means.

4.3 *Overlapping subsequence time series clustering*

Subsequence Time Series (STS) clustering typically employs a clustering technique to the subsequences of a time series generated using a sliding window technique (Fig. 11). Although it is very popular, Keogh et al. [9] reported, for the first time, a surprising anomaly: cluster centers obtained using STS clustering closely resemble "sine waves", irrespective of the nature of original time series itself, and therefore, the results of STS clustering are meaningless.

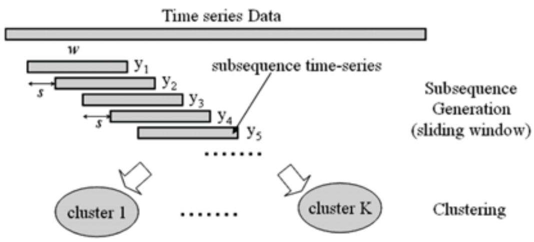


Figure 11: STS clustering using a sliding window technique.

According to the Keogh's criticism, superposition of slightly shifted subsequences causes the generation of sine waves. In order to explicitly exclude superposition, a motif-based clustering algorithm has been proposed in [10]. Another solution to avoid sine wave has been reported in [2], which consists of aligning the phase at the frequency with maximum spectrum power with respect to each time series subsequence, and then, to apply a clustering algorithm to them. Although these approaches are promising, the need of identifying solid mathematical foundations behind STS clustering steel remains.

References

[1] I. Bloch. On fuzzy distances and their use in image processing under imprecision. *Pattern Recognition*, 32:1873–1895, 1999.

[2] R. Fujimaki, S. Hirose and T. Nakata. Theoretical analysis of subsequence time-series clustering from a frequency-analysis viewpoint. *SIAM International Conference on Data Mining*, 506-517, 2008.

[3] V. Georgescu. Granular vs. point-wise metrics and statistics to accommodate machine learning algorithms in granular feature spaces, *Fuzzy Economic Review,* vol.XII, No.2, 45-74, 2007.

[4] V. Georgescu. A Generalization of Symbolic Data Analysis Allowing the Processing of Fuzzy Granules, *Lecture Notes in Artificial Intelligence,* vol. 3131, 215-226, 2004.

[5] V. Georgescu. A Formal Framework Allowing to Deal with Granular Clustering not only in Terms of Hyperboxes, but also in terms of Fuzzy Granules, *Proceedings of MS'2004-France*, Lyon, July 5–7, 2004.

[6] V. Georgescu. On the Foundations of Granular Computing Paradigm, *Fuzzy Economic Review,* vol.VIII, No.2, 73-105, 2003.

[7] V. Georgescu. Multivariate Fuzzy-termed Data Analysis: Issues and Methods, *Fuzzy Economic Review,* vol.VI, No.1, 19-48, 2001.

[8] D. Gunopulos and G. Das. Time series similarity measures and time series indexing. *ACM SIGMOD Conference*, Santa Barbara, CA., 2001.

[9] E. Keogh, J. Lin and W. Truppel. Clustering of time series subsequences is meaningless: implications for previous and future research. *Proceedings of the 3rd IEEE International Conference on Data* Mining, 115–122, 2003.

[10] E. Keogh, B. Chiu and S. Lonardi. Probabilistic discovery of time series motifs. *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003.

[11] E. Keogh and C. A. Ratanamahatana. Exact indexing of dynamic time warping. *Knowledge and Information Systems*, 7: 358-386, 2005.

[12] W. Pedrycz and A. Bargiela Granular Clustering: a Granular Signature of Data, *IEEE Transactions on Systems, Man and Cybernetic.,* vol. 32, No.2, 212-224, 2002.

[13] W. Pedrycz. *Knowledge-Based Clustering. From Data to Information Granules*, John Wiley & Sons, 2005.

[14] W. Pedrycz. (ed.). *Granular Computing: An Emerging Paradigm*, Physica-Verlag, 2001.

[15] L. A. Zadeh. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, *Fuzzy Sets and Systems*, 90: 111-117, 1997.

ISBN: 978-989-95079-6-8