Statistics and Computing (2024) 34:214
https://doi.org/10.1007/s11222-024-10495-5

ORIGINAL PAPER

Fuzzy clustering with Barber modularity regularization

Pierpaolo D'Urso[1] · Livia De Giovanni[2,3] · Lorenzo Federico[2,3] · Vincenzina Vitale[1]

Received: 28 March 2024 / Accepted: 1 September 2024

Abstract
In this paper, we propose a new algorithm for the joint clustering of two sets of statistical units $\mathcal{N}$ and $\mathcal{M}$ which are also equipped with an adjacency structure which is represented by a bipartite network. Our model is based on the fuzzy Partition Around Medoids, and it combines it with techniques for community detection in bipartite complex networks based on Barber modularity maximization. The goal is to produce a partition of $\mathcal{N} \cup \mathcal{M}$ into clusters, each of which is also identified by two medoids, one in $\mathcal{N}$ and one in $\mathcal{M}$, which represent the typical units in the cluster for each set. Such clusters are optimized so that units in the same cluster both have similar values on their attributes and are likely to be adjacent. We test the algorithm on both simulated and real data, to show how it is able to capture a wide range of different interactions between the distribution of the attributes and the network structure.

Keywords Fuzzy C-medoids clustering · Bipartite networks · Community detection · Entropy term

1 Introduction

The aim of this study is to enhance existing research on cluster methodologies within the fuzzy framework by integrating information from a bipartite graph into the clustering process. A novel clustering technique is proposed, the *Fuzzy C-medoids clustering with Barber modularity regularization*, that can provide a fuzzy partition of two disjoint sets of statistical units $\mathcal{N}$ and $\mathcal{M}$ [1] into C different clusters leading to two partition matrices $\mathbf{U} := \{u_{n,c}, n \leq N, c \leq C\}$ and $\mathbf{W} := \{w_{m,c}, m \leq M, c \leq C\}$ respectively.

Units in the sets $\mathcal{N}$ and $\mathcal{M}$ are equipped with attributes in feature spaces $\mathcal{I}$ and $\mathcal{J}$ respectively, and also are the nodes of a bipartite network represented by the biadjacency matrix $\mathbf{A}$. Matrices $\mathbf{U}$ and $\mathbf{W}$ are derived taking into account the information provided by both the attributes on each side and the adjacency structure in the network. Furthermore, as detailed later in this paper, we embed information about the bipartite graph structure as a regularization term within the objective function, aimed at maximizing its effect. The contribution of this regularization term to the clustering criterion is controlled by a tuning parameter.

Typical situations in which the proposed clustering techniques can find practical application are outlined below:

- $\mathcal{N}$: social media users. $\mathcal{I}$: account metrics (creation date, number of followers, verified status...). $\mathcal{M}$: social media posts. $\mathcal{J}$: post metrics (date of posting, engagement metric, language, sentiment...). $a_{n,m}$: indicator that user n replied to tweet m.
- $\mathcal{N}$: nations. $\mathcal{I}$: demographic indicators (average income, life expectancy, most common religion...). $\mathcal{M}$: international organizations or treaties. $\mathcal{J}$: fields of relevance of organizations or treaties (if they are about military, technological cooperation, etc...). $a_{n,m}$: indicator that nation n subscribed to organization or treaty m.

[1] Throughout the entire paper curly letters are sets, normal letters are their cardinality.

✉ Lorenzo Federico
lfederico@luiss.it

Pierpaolo D'Urso
pierpaolo.durso@uniroma1.it

Livia De Giovanni
ldegiovanni@luiss.it

Vincenzina Vitale
vincenzina.vitale@uniroma1.it

[1] Department of Social Science and Economics, Sapienza University, Piazzale Aldo Moro 5, 00185 Rome, Lazio, Italy

[2] Department of Political Science, Luiss University, Viale Romania 32, 00197 Rome, Lazio, Italy

[3] Data Lab, Luiss University, Viale Pola 12, 00198 Rome, Lazio, Italy

Published online: 08 November 2024

- $\mathcal{N}$: tennis players. $\mathcal{I}$: bio-metric measure or match statistics (age, weight, percentage of break points converted...). $\mathcal{M}$: tournaments. $\mathcal{J}$: tournament characteristics (surface, prize money, host nation...). $a_{n,m}$ number of matches played by player n in tournament m.

From a methodological point of view, the proposed *Fuzzy C-medoids clustering with Barber modularity regularization* integrates several advantageous elements. Firstly, it adopts the fuzzy methodology, incorporating an entropy regularization term. Secondly, it incorporates the Partitioning Around Medoids (PAM) approach and, lastly, it introduces another regularization term, represented by the Barber modularity extended to the fuzzy framework.

The rationale behind adopting a fuzzy approach lies in its departure from conventional clustering methods, which necessitate partitions characterized by non-empty and pairwise disjoint subsets. Cluster Analysis using Fuzzy Theory (Zadeh 1965) relaxes this stringent requirement, allowing units to belong to multiple clusters, thus well-suited to grouping complex data or objects with an inherent and unavoidable degree of "imprecision".

The concept of "degree of membership" replaces that of "crisp membership" in such a way that the generic element u_{nc} (or w_{mc}) of partition matrix $\mathbf{U}_{N\times C}$ (or $\mathbf{W}_{M\times C}$), is now a real value in [0, 1] rather than being either 0 or 1.

Because of these advantages, fuzzy clustering techniques have become increasingly popular leading to the development of numerous fuzzy clustering algorithms. Key contributions were made by Bellman et al. (1966) and Ruspini (1969, 1970, 1973); however, the Fuzzy C-Means (FCM), which was first introduced by Dunn (1973) and Bezdek (1974, 1981), is considered the most representative one and is still widely used today. In this formulation, the "fuzziness" is obtained by raising $u_{n,c}$ to an exponent $m > 1$, serving as the fuzziness coefficient regulating the extent of membership shared among the fuzzy clusters.

Another variant of the C-Means clustering method in a fuzzy perspective was introduced by Li and Mukaidono (1995, 1999) and Miyamoto and Mukaidono (1997). In the objective function,[2] an entropy regularization term is included, specifically the Shannon entropy. When applied to membership degrees, this term can be referred to as fuzzy entropy (Coppi and D'Urso 2006). This different formulation of the fuzzy C-Means responds to criticism from some researchers about the role of the fuzziness coefficient, which is often viewed as an artificial, unnatural device lacking physical interpretation. By adopting this approach, the functional is optimized by maximizing both internal cohesion and the assigned measure of entropy, thereby increasing the overall amount of information. Various other entropy-based variants have been proposed in the literature, some of which can be found in Yao et al. (2000), Ichihashi (2000), Zarinbal et al. (2014), Kahali et al. (2019), Gao et al. (2019). Another interesting extension to time-varying data can be found in Coppi and D'Urso (2006). However, within the framework of soft membership models, it is worth mentioning the Partial Membership model (Heller et al. 2008) and the Mixed Membership model (Erosheva et al. 2004; Airoldi et al. 2008), two individual-level mixture models that allow units to belong to more than one cluster.

The rationale behind incorporating a Partitioning Around Medoids (PAM) approach into a fuzzy framework (Krishnapuram et al. 1999, 2001) is rooted in several key properties. Firstly, each cluster is represented by its own *medoid*, distinct from the centroid used in traditional C-means clustering. Unlike centroids, which can be virtual and may not correspond to actual data points, medoids are observed objects from the dataset itself. Specifically, each medoid is selected to minimize the overall distance from all other objects within the same cluster.

Furthermore, the Fuzzy C-Medoids (FCMd) approach offers a notable advantage over Fuzzy C-Means (FCM) when it comes to handling noise and outliers. This advantage stems from the fact that medoids are less influenced by extreme values than the average.

Recent proposals that combine the fuzzy entropy and the medoids-based approach can be found in D'Urso and Vitale (2022), D'Urso et al. (2023a, b, c, d).

As a spatial regularization term, to favor partitions of the units so that units in the same cluster are more likely to be adjacent than units in different clusters, we use a fuzzy extension of the Barber Modularity. In general, modularity is a well-established parameter (Newman 2006) to evaluate how well a (in its original formulation crisp) partition of a graph captures its *community structure*. In network theory, communities are subsets of nodes of a larger networks that have more connections within themselves than with the rest of the network. The modularity of a partition of a network is defined as the difference between the number of links between nodes in the same set in the observed network, and its expected value in a suitable null model, the most popular choices being the Rank-1 inhomogeneous random graph, as defined in Norros and Reittu (2006) or the hard configuration model (see e.g Luczak (1992)). In these models, the expected number of connections between two nodes is proportional to the product of their respective total numbers of connections (usually called degree or strength). A popular way to cluster the nodes of a network according to its community structure is to maximize the modularity over all the possible partitions of its vertices. Exact modularity maximization is NP-Hard (Brandes et al. 2007), so the solution is approximated, usually either by a greedy hierarchical clustering (Clauset et al. 2004) or by more sophisticated algorithms that are allowed

[2] with $m = 1$.

also to split communities or move nodes between different communities (Blondel et al. 2008; Traag et al. 2019). Barber Modularity (Barber 2007) was introduced to adapt the notion of modularity for bipartite networks, using a different null model that would reflect the bipartition. Modularity was later extended to fuzzy partitions of the node set of a network, by defining the modularity of a fuzzy partition as the expected value of the modularity in a crisp partition where nodes are independently randomly assigned to clusters proportionally to their fuzzy memberships (Nepusz et al. 2008).

To the best of our knowledge, there are currently no other proposals in the existing literature that attempt to integrate both the fuzzy clustering approach and the modularity-based approach. We emphasize that our proposed methodology has the potential to reveal deeper insights into the complex structure of data serving as a bridge for further integrations of these two methodologies. This involves leveraging the strengths of each method to enhance the capabilities of the other, thereby contributing to a deeper understanding of complex data structures.

The paper is structured as follows. Section 2 outlines the data structure (2.1) and presents the formal definition of Fuzzy Barber modularity (2.2), concluding with the definition of the proposed clustering method (2.3). Section 3.1 introduces the algorithmic functioning, while Sect. 3.2 describes the newly proposed internal validity index. The simulation plan and results are presented in Sect. 4, followed by the application to real data in Sect. 5. Finally, concluding remarks are provided in the last section.

2 Instruments and methods

2.1 Data structure

We consider a model in which we have two sets of statistical units $\mathcal{N}$ and $\mathcal{M}$, with associated sets of numerical attributes $\mathcal{I}, \mathcal{J}$. These attribute values are represented by the data matrices

$$\mathbf{X} := \{x_{n,i}, n \leq N, i \leq I\}, \ \mathbf{Y} := \{y_{m,j}, m \leq M, j \leq J\}, \tag{1}$$

where $x_{n,i}$ is the value of the attribute i measured on the unit n, and equivalently $y_{m,j}$. The sets $\mathcal{I}$ and $\mathcal{J}$ need not be disjoint, while the sets $\mathcal{N}$ and $\mathcal{M}$ are required to.

We also assign a bipartite adjacency structure between the sets $\mathcal{N}$ and $\mathcal{M}$, represented by the matrix

$$\mathbf{A} := \{a_{n,m}, n \leq N, m \leq M\}, \tag{2}$$

this can be a binary matrix, where $a_{n,m} = 1$ indicates that n and m are contiguous and $a_{n,m} = 0$ means that they are not, or a general non-negative matrix where $a_{n,m}$ represents the intensity of the relation between n and m. We refer to two unites $n \leq N$, $m \leq M$ as adjacent if $a_{n,m} > 0$.

Our goal is to generate a fuzzy partition of the set $\mathcal{V} := \mathcal{N} \cup \mathcal{M}$ in C clusters that exhibit both of the following properties:

- **Cluster Cohesion**: Couples of units $n_1, n_2 \in \mathcal{N}$ or $m_1, m_2 \in \mathcal{M}$ which both have high membership to the same cluster c, also have similar values of their attributes.
- **Community Validity** : Couples of units $n \in \mathcal{N}, m \in \mathcal{M}$ which both have high membership to the same cluster c, are more likely to be adjacent than the average.

The rest of this section will be dedicated to the definition of an adequate objective function to be optimized to find a partition that satisfies both conditions.

2.2 Fuzzy Barber modularity

Here, we give a formal definition of the *Fuzzy Barber modularity*, which we use as a spatial regularization term. Here, like in the baseline modularity introduced in Newman (2006), we compare the number of connections within clusters to its expectation in a suitable null model. Said null model is designed to take into account the bipartite nature of the network, so that, in particular, the expected number of connections between units on the same side is always 0, while between units on opposite sides it retains its product structure. This is equivalent to selecting as a null model the bipartite configuration model (see e.g. van der Hofstad et al. (2022)).

We now make this formal. Let us define:

$$Q = \sum_{n=1}^{N}\sum_{m=1}^{M} a_{n,m}, \quad s_n = \sum_{m=1}^{M} a_{n,m}, \ \forall n \leq N,$$
$$t_m = \sum_{n=1}^{N} a_{n,m}, \ \forall m \leq M. \tag{3}$$

The *Barber modularity matrix* is constructed as:

$$\mathbf{B} = \{b_{n,m}, n \leq N, m \leq M\},$$

where

$$b_{n,m} := a_{n,m} - \frac{s_n t_m}{Q}. \tag{4}$$

In practice, the matrix $\mathbf{B}$ encodes for each couple of units $n \in \mathcal{N}, m \in \mathcal{M}$ the difference between the observed connections

and the expected ones in the null model. The (crisp) *Barber modularity* of a partition of the set $\mathcal{N} \cup \mathcal{M}$ is defined in Barber (2007), up to a normalization constant, as

$$\sum_{n=1}^{N}\sum_{m=1}^{M} b_{n,m}\delta_{n,m}, \tag{5}$$

where $\delta_{n,m}$ is the indicator that n and m are in the same cluster. Note that, by definition,

$$\sum_{n=1}^{N}\sum_{m=1}^{M} b_{n,m} = 0. \tag{6}$$

Consequently, a partition has a positive Barber modularity if the connections within it sets are stronger than their expectation.

We extend the Barber modularity within the fuzzy framework. Consider a fuzzy partition of the set $\mathcal{V} := \mathcal{N} \cup \mathcal{M}$ into C different clusters, identified by the membership matrices

$$\mathbf{U} := \{u_{n,c}, n \le N, c \le C\}, \quad \mathbf{W} := \{w_{m,c}, m \le M, c \le C\}, \tag{7}$$

where $u_{n,c}$ (resp. $w_{m,c}$) represents the degree of membership of unit n (resp. m) to cluster c.

We define the Fuzzy Barber modularity of the partition $(\mathbf{U}, \mathbf{W})$ as

$$\sum_{n=1}^{N}\sum_{c=1}^{C}\sum_{m=1}^{M} u_{n,c}b_{n,m}w_{m,c}. \tag{8}$$

This can also be written in matrix form as $\mathrm{Tr}(\mathbf{U}^T\mathbf{B}\mathbf{W})$. This corresponds, up to a constant, to the expected value of the Barber modularity in a crisp partition where node $n \in \mathcal{N}$ is assigned to cluster c with probability $u_{n,c}$, and node $m \in \mathcal{M}$ to cluster c with probability $w_{m,c}$, all independently of each other. If $(\mathbf{U}, \mathbf{W})$ is a crisp partition, that is, if $u_{n,c}, w_{m,c} \in \{0, 1\}$ for all n, m, c, then this is equal to the usual Barber modularity. This is similar to the extension to a fuzzy context of the baseline modularity in Nepusz et al. (2008).

2.3 The fuzzy C-medoids clustering with Barber modularity regularization

We can now give a formal definition of our clustering model. Given two disjoint sets of statistical units, $\mathcal{N}$ and $\mathcal{M}$, two sets of numerical attributes $\mathcal{I}$ and $\mathcal{J}$—the former only defined on $\mathcal{N}$, and the latter on $\mathcal{M}$—and the adjacency matrix $\mathbf{A}$, the Fuzzy C-medoids clustering with Barber modularity regularization aims to merge the two sources of information on the clustering structure provided by each set of numerical attributes and the network.

Our core idea is to create an objective function to be minimized that depends on the data in the matrices $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{A}$ and on the fuzzy partition $(\mathbf{U}, \mathbf{W})$. To balance the information provided by the numerical attributes and by the network structure, we choose a linear combination of the objective functions of two fuzzy entropic C-medoids algorithm over the two sets $\mathcal{N}$ and $\mathcal{M}$ and the fuzzy Barber modularity of the bipartite network defined by the Barber modularity matrix $\mathbf{B}$ obtained from the adjacency matrix $\mathbf{A}$.

This function is to be minimized over all the valid fuzzy partitions $(\mathbf{U}, \mathbf{W})$ of the set $\mathcal{V} := \mathcal{N} \cup \mathcal{M}$ into C different clusters. Furthermore, associated with each partition matrix, C prototypes, called medoids, are also provided, i.e. the subset $(\mathbf{x}_1, \ldots, \mathbf{x}_c, \ldots, \mathbf{x}_C)$ and $(\mathbf{y}_1, \ldots, \mathbf{y}_c, \ldots, \mathbf{y}_C)$, whose generic $\mathbf{x}_c$, for $c \in C$, is chosen among the N observed units $\mathbf{x}_n = (x_{n,1}, \ldots, x_{n,I}) \in \mathbb{R}^I$, with $n \le N$, and $\mathbf{y}_c$ among the M observed units $\mathbf{y}_n = (y_{m,1}, \ldots, y_{m,J}) \in \mathbb{R}^J$, with $m \le M$ respectively. Both the partition $(\mathbf{U}, \mathbf{W})$ and the medoids $(\mathbf{x}_c, \mathbf{y}_c)_{c \le C}$ are found by solving the following minimization problem:

$$\begin{aligned} &J_{p_1,p_2,C,\gamma}(\mathbf{U}, \mathbf{W}, \mathbf{x}_c, \mathbf{y}_c) \\ &= \sum_{n=1}^{N}\sum_{c=1}^{C} u_{n,c}d^2(\mathbf{x}_n, \mathbf{x}_c) + p_1\sum_{n=1}^{N}\sum_{c=1}^{C} u_{n,c}\log(u_{n,c}) \\ &\quad + \sum_{m=1}^{M}\sum_{c=1}^{C} w_{m,c}d^2(\mathbf{y}_m, \mathbf{y}_c) + p_2\sum_{m=1}^{M}\sum_{c=1}^{C} w_{m,c}\log(w_{m,c}) \\ &\quad - \gamma\sum_{n=1}^{N}\sum_{c=1}^{C}\sum_{m=1}^{M} u_{n,c}b_{n,m}w_{m,c}, \end{aligned} \tag{9}$$

subject to

$$\begin{aligned} &u_{n,c} \ge 0, \quad \sum_{c=1}^{C} u_{n,c} = 1, \qquad w_{m,c} \ge 0, \quad \sum_{c=1}^{C} w_{m,c} = 1, \\ &\qquad \forall\, n \le N,\ m \le M,\ c \le C, \end{aligned} \tag{10}$$

for fixed values of the tuning parameters $\gamma > 0$, $p_1, p_2 \ge 0$.

It should be noted that the function is not properly defined if any of $u_{n,c} = 0$ or $w_{m,c} = 0$ for some n, m, c, since $\log 0 = -\infty$. In such cases, we use the continuous extension of the function $x \log x$ such that $0 \log 0 = 0$.
Specifically,

- $\sum_{n=1}^{N}\sum_{c=1}^{C} u_{n,c}d^2(\mathbf{x}_n, \mathbf{x}_c) + p_1\sum_{n=1}^{N}\sum_{c=1}^{C} u_{n,c}\log(u_{n,c})$

  is the objective function for set $\mathcal{N}$. $u_{n,c}$ denotes the degree of membership of unit n to cluster c, and $d^2(\mathbf{x}_n, \mathbf{x}_c)$ represents the squared Euclidean distance

between the n-th unit and the c-th medoid. The term $\sum_{n=1}^{N}\sum_{c=1}^{C} u_{n,c}\log(u_{n,c})$ signifies the fuzzy entropy, where p_1 acts as the weight factor controlling the degree of fuzziness in the partition $\mathbf{U}$. As p_1 increases, so does the level of fuzziness in the corresponding partition.

- $$\sum_{m=1}^{M}\sum_{c=1}^{C} w_{m,c}d^2(\mathbf{y}_m,\mathbf{y}_c) + p_2\sum_{m=1}^{M}\sum_{c=1}^{C} w_{m,c}\log(w_{m,c})$$

  is the objective function for set $\mathcal{M}$. $w_{m,c}$ denotes the degree of membership of unit m to cluster c, and $d^2(\mathbf{y}_m,\mathbf{y}_c)$ represents the squared Euclidean distance between the m-th unit and the c-th medoid. The term $\sum_{c=1}^{C} w_{m,c}\log(w_{m,c})$ is the corresponding fuzzy entropy, where p_2 is the weight factor controlling the degree of fuzziness in the partition $\mathbf{W}$. As p_2 increases, so does the level of fuzziness in the corresponding partition.

- $$\gamma\sum_{n=1}^{N}\sum_{c=1}^{C}\sum_{m=1}^{M} u_{n,c}b_{n,m}w_{m,c}$$

  is the Fuzzy Barber modularity as defined in (8) where γ is the tuning parameter controlling the network's contribution to the clustering criterion, which would otherwise be based on attributes alone.

In particular, the higher the value of γ, the more important is the network structure compared to the attributes for the choice of the fuzzy partition. As we will see later in the simulation studies in Sect. 4, if the cluster structure is similar for both the attributes and the network, the partitions found for different values of γ will not differ substantially. If instead the clusters identified by the attributes and the network are different, the choice of the parameter γ becomes very important, as it will determine if the algorithm will prioritize one of the two conflicting information or will try to balance both. Thus, it can be concluded that when $\gamma = 0$, the functional is identical to the sum of the objective functions of two separate fuzzy clustering algorithms.

This means that the algorithm would still output a joint fuzzy partition of both sides but identification of the clusters on the set $\mathcal{N}$ and $\mathcal{M}$ would be completely arbitrary as the algorithm would just be equivalent to running two independent entropic fuzzy clustering algorithms one on each set.

It is worth pointing out that the *fuzzy Barber modularity* of the partition $(\mathbf{U},\mathbf{W})$ is subtracted since it requires to be maximized.

An exact solution to the constrained minimization problem described in (9)–(10) is not feasible, consequently we will use an iterative algorithm to approximate at least a local minimum. We will alternatively optimize the objective function with respect to the cluster memberships $(\mathbf{U},\mathbf{W})$ given the medoids $(\mathbf{x}_c,\mathbf{y}_c)_{c\le C}$, and then optimize it with respect to the medoids given the memberships. Note that, in the second step, we can separately optimize the two sets of medoids, as the terms of the objective function that explicitly depend on each of the two are different. The optimization with respect to the medoids can be done by brute force, to optimize with respect to the memberships we use the Lagrangian multiplier method to find that at each step the optimal values of $(\mathbf{U},\mathbf{W})$ given $(\mathbf{x}_c,\mathbf{y}_c)_{c\le C}$ are found as

$$u_{n,c} = \frac{\exp\left\{-\frac{1}{p_1}d^2(\mathbf{x}_n,\mathbf{x}_c) + \frac{\gamma}{p_1}\sum_{m=1}^{M} b_{n,m}w_{m,c}\right\}}{\sum_{c=1}^{C}\exp\left\{-\frac{1}{p_1}d^2(\mathbf{x}_n,\mathbf{x}_c) + \frac{\gamma}{p_1}\sum_{m=1}^{M} b_{n,m}w_{m,c}\right\}}, \tag{11}$$

for all $n \le N$, $c \le C$, and,

$$w_{m,c} = \frac{\exp\left\{-\frac{1}{p_2}d^2(\mathbf{y}_m,\mathbf{y}_c) + \frac{\gamma}{p_2}\sum_{n=1}^{N} b_{n,m}u_{n,c}\right\}}{\sum_{c=1}^{C}\exp\left\{-\frac{1}{p_2}d^2(\mathbf{y}_m,\mathbf{y}_c) + \frac{\gamma}{p_2}\sum_{n=1}^{N} b_{n,m}u_{n,c}\right\}}, \tag{12}$$

for all $m \le M$, $c \le C$. The details on how (11) and (12) are obtained are shown in Appendix A.

3 The algorithm

We now describe the procedure we employ to find the optimal clustering. It consists of two parts: an algorithm for the minimization of the objective function for fixed values of the hyperparameters C and γ, which is presented in Sect. 3.1, and a validity measure that is used *a posteriori* to select the optimal solution among tose found for all the possible values of the hyperparameters, which is presented in Sect. 3.2.

3.1 Minimization algorithm

We now describe the algorithmic optimization of the objective function we described in (9). As we mentioned, the exact solution of the problem is not feasible, so we proceed with an iterative approach, in which we initialize randomly the values of the membership matrices $\mathbf{U}$, $\mathbf{W}$ and the identities of the medoids $(\mathbf{x}_1,\ldots,\mathbf{x}_c,\ldots,\mathbf{x}_C)$ and $(\mathbf{y}_1,\ldots,\mathbf{y}_c,\ldots,\mathbf{y}_C)$, and then alternatively optimize the function with respect to the memberships given the medoids and the medoids given the memberships, as described in Algorithm 1.

Algorithm 1 Fuzzy C-medoids clustering with Barber modularity regularization (FMd-BMR)

Fix C, $max.iter$, $conv = 1 \times 10^{-9}$ and initialize randomly the membership degree matrices $\mathbf{U}$,$\mathbf{W}$;
Set $iter = 0$;
Set *medoids.left*:= $(\mathbf{x}_1, \ldots, \mathbf{x}_C)$, and *medoids.right*:=$(\mathbf{y}_1, \ldots, \mathbf{y}_C)$ arbitrarily;
repeat
  Set $\mathbf{U_{old}} = \mathbf{U}$,$\mathbf{W_{old}} = \mathbf{W}$;
  Update *medoids.left* as follows:
  for $c = 1$ to C **do**
    Define *members*=$\{i \leq N : c = \arg\max_{1 \leq k \leq C} u_{i,k}\}$
    if *members* is not empty **then**
      $q = \arg\min_{q \in members} \sum_{n=1}^{N} u_{n,c} d^2(\mathbf{x}_n, \mathbf{x}_q)$
      Set $\Rightarrow \mathbf{x}_c = \mathbf{x}_q$
    end if
  end for
  Update *medoids.right* as follows:
  for $c = 1$ to C **do**
    Define *members*=$\{i \leq M : c = \arg\max_{1 \leq k \leq C} w_{i,k}\}$
    if *members* is not empty **then**
      $q = \arg\min_{q \in members} \sum_{m=1}^{M} w_{m,c} d^2(\mathbf{y}_n, \mathbf{y}_q)$
      Set $\Rightarrow \mathbf{y}_c = \mathbf{y}_q$
    end if
  end for
  Update $\mathbf{U}$ using (11);
  Update $\mathbf{W}$ using (12);
  $iter \leftarrow iter + 1$;
until ($\|\mathbf{W_{old}} - \mathbf{W}\|_1 < conv$ and $\|\mathbf{U_{old}} - \mathbf{U}\|_1 < conv$) or $iter = max.iter$
return $\mathbf{U}$, $\mathbf{W}$, $(\mathbf{x}_1, \ldots, \mathbf{x}_C)$, $(\mathbf{y}_1, \ldots, \mathbf{y}_C)$

the idea that the medoids are meant to represent prototypes of the units in their cluster and so we forbid the algorithm from choosing so. The time complexity of Algorithm 1 in the worst case scenario is $O(max.iter(N + M)^2 C)$. This algorithm is based on greedy optimization of the objective function and thus it is prone to getting stuck in local minima. To minimize such risk, we run the algorithm multiple times with random starting memberships and medoids and then choose *a posteriori* the optimal solution among all the ones obtained.

3.2 Validity measure

In order to choose the optimal number of groups, we propose a new measure of internal validity that ensures that the model has adequately captured both network and attribute information so that the fuzzy partitions that privilege one of the two, ignoring almost completely the other one will be heavily penalized. This model is designed to objectively consider all information encoded in the three matrices $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{B}$. If the user is willing to forego any of this information, there are numerous algorithms available for fuzzy clustering or community detection. The proposed internal validity index is defined as the ratio of between-cluster separation and fuzzy modularity to within-cluster cohesion, normalized by the number of degrees of freedom to make them comparable across different numbers of clusters, i.e.:

$$F(\mathbf{U}, \mathbf{V}) = \frac{\frac{1}{C} \cdot [min_{c \neq c'} d^2(\tilde{\mathbf{x}}_c, \tilde{\mathbf{x}}_{c'}) + min_{c \neq c'} d^2(\tilde{\mathbf{y}}_c, \tilde{\mathbf{y}}_{c'}) + \sum_{n=1}^{N} \sum_{c=1}^{C} \sum_{m=1}^{M} u_{n,c} b_{n,m} w_{m,c}]}{\frac{\sum_{n=1}^{N} \sum_{c=1}^{C} u_{n,c} d^2(\mathbf{x}_n, \tilde{\mathbf{x}}_c)}{n-C} + \frac{\sum_{m=1}^{M} \sum_{c=1}^{C} w_{m,c} d^2(\tilde{\mathbf{y}}_m, \mathbf{y}_c)}{m-C}}, \tag{13}$$

In the optimization of the medoid of the cluster c we consider as candidates only the units whose maximal membership is to cluster c. This has the goal of avoiding pathological behavior in cases in which two clusters c and c' have similar values for the attributes of their units but are clearly separated by the adjacency structure in $\mathbf{A}$. In such a situation, it is possible for an algorithm that uses unconstrained medoids optimization to choose as the left medoid of cluster c a unit n such that $u_{nc'} > u_{nc}$ for some cluster $c' \neq c$ (or, equivalently, as the right medoid a unit m such that $w_{mc'} > w_{mc}$). This would be highly inconsistent with where $min_{c \neq c'} d^2(\tilde{\mathbf{x}}_c, \tilde{\mathbf{x}}_{c'})$, for $c, c' \in \{1 \ldots C\}$, is the minimum squared distance of cluster medoids (the same holds in terms of y). Thus, higher index values indicate a better clustering result. We specify that the index allows for selecting the optimal combination of C and γ from a range of plausible values while keeping the weight factors p_1 and p_2 fixed, as optimizing all exogenous parameters simultaneously imposes a significant computational burden in practice.

We note that $F(\mathbf{U}, \mathbf{V})$ can take any real value, including negative values, which are observed when the Fuzzy Barber modularity of the optimal solution is negative. These solutions are thus never selected as it means that they behave worse than random with respect to the network structure.

We emphasize that this index is proposed specifically for this context, as existing indices are not directly applicable. Consequently, it represents an original proposal within the class of internal validity indices known in the literature.

4 Simulation study

In order to assess the effectiveness of the proposed clustering technique, three simulation scenarios representing extreme real-world situations were examined. For all scenarios, we consider two disjoint sets of statistical units $\mathcal{N}$ and $\mathcal{M}$ with $N = 50$ and $I = 2$ and $M = 20$ and $J = 3$ respectively.

In the *first scenario*, two groups are formed for each set of units based on their numerical attributes, with the network matching the clustering structure. In detail, the two groups are simulated according to the following scheme.

For the set $\mathcal{N}$, units $n_1, \ldots, n_{29}$ are generated according to $X_1 \sim U(0, 5)$ and $X_2 \sim U(20, 40)$; units $n_{31}, \ldots, n_{49}$ according to $X_1 \sim U(10, 15)$ and $X_2 \sim U(50, 70)$. Units n_{30}, n_{50} are generated as fuzzy units from $X_1 \sim U(6, 9)$ and $X_2 \sim U(41, 49)$.

For the set $\mathcal{M}$, units $m_1, \ldots, m_{10}$ are generated according to $Y_1 \sim U(-20, -10)$, $Y_2 \sim U(0, 10)$, and $Y_3 \sim U(1, 2)$; units $m_{11}, \ldots, m_{19}$ according to $Y_1 \sim U(20, 30)$, $Y_2 \sim U(-15, -5)$, and $Y_3 \sim U(-1, 0)$. Unit m_{20} is a fuzzy unit generated from $Y_1 \sim U(-9, 19)$, $Y_2 \sim U(-4, -1)$, and $Y_3 \sim U(0.1, 0.9)$.

Figure 6 in Appendix B shows the corresponding biadiacency matrix which associates $n_1, \ldots, n_{30}$ with $m_1, \ldots, m_{10}$ and $n_{31}, \ldots, n_{50}$ with $m_{11}, \ldots, m_{20}$. Units belonging to the same cluster based on the attributes set are coloured accordingly. Fuzzy units are coloured cyan for the set $\mathcal{N}$ and black for the set $\mathcal{M}$.

In the *second scenario*, the identical clustering structure is used within the set $\mathcal{N}$, while no clustering structure is defined for $\mathcal{M}$; hence, all 20 units are generated from $Y_1 \sim U(-20, -10)$, $Y_2 \sim U(0, 10)$, and $Y_3 \sim U(1, 2)$. The network still matches the clustering structure on the set $\mathcal{N}$ as shown in Fig. 7.

The *third scenario* differs from the first only in terms of network. In this case, the only set that matches the network is $\mathcal{N}$ in such a way that: $n_1, \ldots, n_{30}$ are connected to $m_1, \ldots, m_5$ and $m_{16}, \ldots, m_{20}$, while $n_{31}, \ldots, n_{50}$ are connected to $m_6, \ldots, m_{15}$, as shown in Fig. 8.

We implemented the clustering technique varying γ from 0 to 3 with a step size of 0.15 and C ranging from 2 to 6 leading to 126 possible combinations. This scheme has been used twice, first with $p_1 = 1$ and $p_2 = 1.5$ and then with $p_1 = 1.5$ and $p_2 = 2$.[3] We ran the algorithm with 100 random restarts. The code used was implemented in R. The execution time, which refers to all 100 random restarts, was a few seconds.

First scenario

In this scenario, it is expected that both sets will be grouped into two clusters, with the cluster membership of each unit increasing as the value of the γ parameter increases.

The proposed internal validity index (13) was used to select the optimal C^* and γ from the 126 partitions resulting from the combination of their values. Based on the results shown in Tables 8 and 9, the index reaches its maximum (172) at $C = 2$ in both cases, associated with $\gamma = 0.75$ and $\gamma = 1.35$ respectively, and flattens out at this maximum value for higher values of the tuning parameter, as shown in Fig. 1. Since the results are the same for the two sets of weighting factors, we focused on the fuzzier solution with $p_1 = 1.5$ and $p_2 = 2$.

To discuss the main features and benefits of the proposed approach, we examine the $\mathbf{U}$ and $\mathbf{W}$ matrices corresponding to the best solution, i.e. $C^* = 2$ and $\gamma = 1.35$.[4] These are compared to the baseline solution, which uses a γ value of 0, and an intermediate solution with $\gamma = 0.45$. The three membership degrees matrices corresponding to the increasing values of the γ parameter are shown in Tables 10 and 11 for the set $\mathcal{N}$ and $\mathcal{M}$ respectively. The same comparisons for $p_1 = 1$, $p_2 = 1.5$ are shown in Tables 12 and 13. As expected, as γ increases, both partitions become crisper. This is especially noticeable for fuzzy units. The logic behind this behavior is that the γ value determines the network's contribution to the clustering criterion. The degree of membership in the original cluster is strengthened as the contamination between the two structures increases as they match in this case.

It is important to note that, in the baseline scenario, there should not be a match between the groups in $\mathbf{U}$ and those in $\mathbf{W}$. However, in the case of $\gamma > 0$, matching among the clusters of the two sets is meaningful as it must respect the relationships encoded in the network. Indeed, as we can note in Tables 10 and 11, in the scenario in which $\gamma = 0$, while the clusters on each side respect, as they should, the attribute structure, the matching between clusters in on the left and the right does not agree with the network structure. In the scenarios where $\gamma > 0$, instead, we observe that the clusters on the left and right are matched according to the connections in the network. In detail, this means that since the first thirty units of set $\mathcal{N}$ form the second cluster, the first ten units of set

[3] The values of p_1 and p_2 are chosen by the user and strictly depend on the scaling of the dissimilarity used and the degree of separation among groups.

[4] We choose the first value of γ that maximizes the validity index.

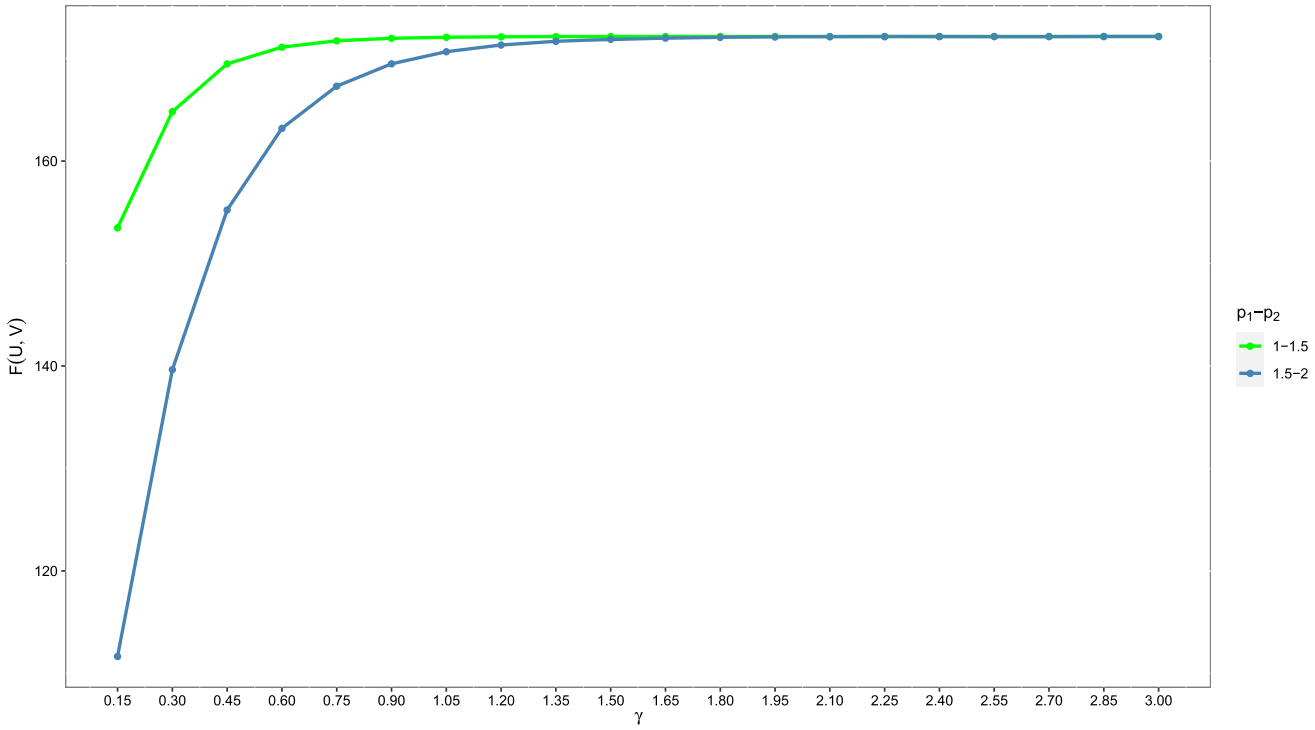

Fig. 1 First scenario: values of the internal validity index when C = 2 setting $p_1 = 1,\ p_2 = 1.5$ (green line) and $p_1 = 1.5,\ p_2 = 2$ (blue line)

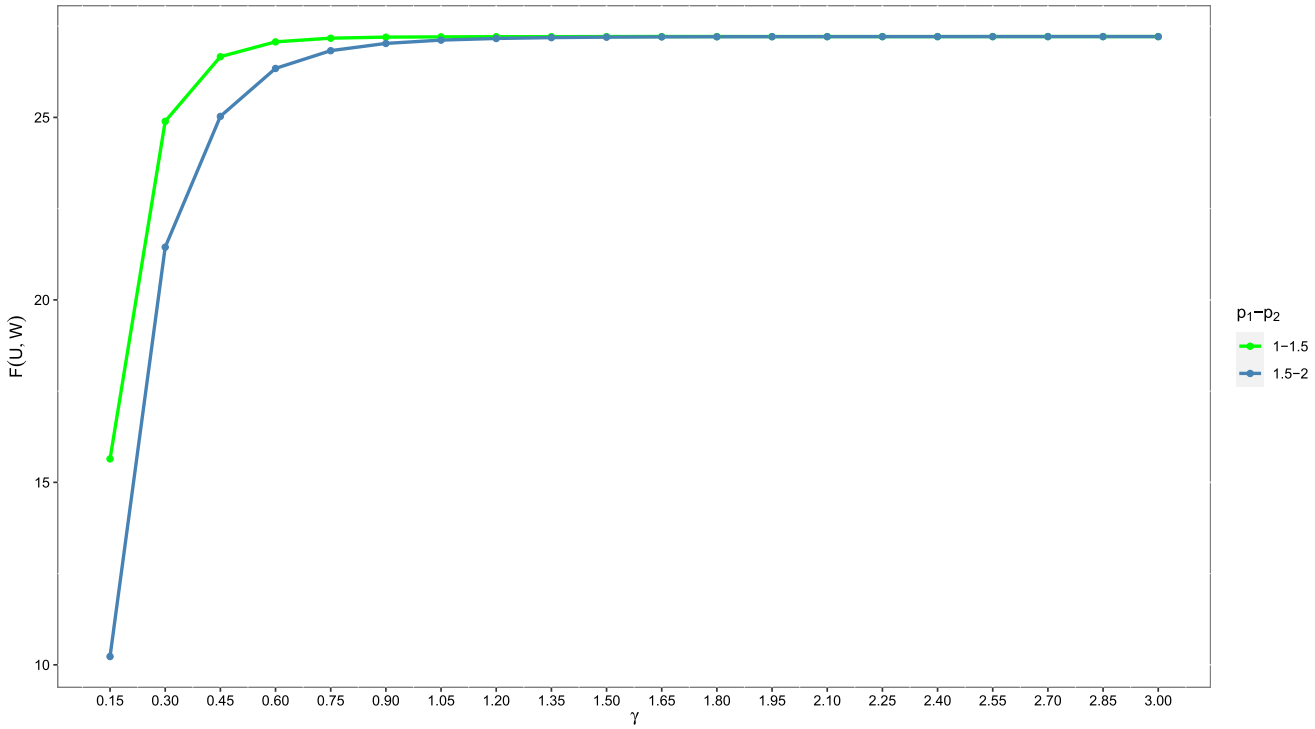

Fig. 2 Second scenario: values of the internal validity index when C = 2 setting $p_1 = 1,\ p_2 = 1.5$ (green line) and $p_1 = 1.5,\ p_2 = 2$ (blue line)

Table 1 Second scenario with $p_1 = 1.5,\ p_2 = 2$: **U** membership degrees matrix for $\gamma \in \{0, 0.30, 0.75\}$

	$\gamma = 0$		$\gamma = 0.30$		$\gamma = 0.75$	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
1	0.85	0.15	0.97	0.03	1.00	0.00
2	0.89	0.11	0.98	0.02	1.00	0.00
3	0.80	0.20	0.96	0.04	0.99	0.01
4	0.89	0.11	0.98	0.02	1.00	0.00
5	0.87	0.13	0.97	0.03	1.00	0.00
6	0.71	0.29	0.92	0.08	0.99	0.01
7	0.88	0.12	0.98	0.02	1.00	0.00
8	0.96	0.04	0.99	0.01	1.00	0.00
9	0.95	0.05	0.99	0.01	1.00	0.00
10	0.96	0.04	0.99	0.01	1.00	0.00
11	0.94	0.06	0.99	0.01	1.00	0.00
12	0.95	0.05	0.99	0.01	1.00	0.00
13	0.95	0.05	0.99	0.01	1.00	0.00
14	0.95	0.05	0.99	0.01	1.00	0.00
15	0.88	0.12	0.98	0.02	1.00	0.00
16	0.84	0.16	0.97	0.03	0.99	0.01
17	0.83	0.17	0.96	0.04	0.99	0.01
18	0.84	0.16	0.97	0.03	0.99	0.01
19	0.80	0.20	0.95	0.05	0.99	0.01
20	0.64	0.36	0.89	0.11	0.98	0.02
21	0.76	0.24	0.94	0.06	0.99	0.01
22	0.93	0.07	0.99	0.01	1.00	0.00
23	0.90	0.10	0.98	0.02	1.00	0.00
24	0.90	0.10	0.98	0.02	1.00	0.00
25	0.88	0.12	0.98	0.02	1.00	0.00
26	0.74	0.26	0.93	0.07	0.99	0.01
27	0.94	0.06	0.99	0.01	1.00	0.00
28	0.95	0.05	0.99	0.01	1.00	0.00
29	**1.00**	**0.00**	**1.00**	**0.00**	**1.00**	**0.00**
30	*0.42*	*0.58*	*0.72*	*0.28*	*0.94*	*0.06*
31	0.07	0.93	0.02	0.98	0.00	1.00
32	0.01	0.99	0.00	1.00	0.00	1.00
33	0.06	0.94	0.02	0.98	0.00	1.00
34	0.05	0.95	0.01	0.99	0.00	1.00
35	0.04	0.96	0.01	0.99	0.00	1.00
36	0.03	0.97	0.01	0.99	0.00	1.00
37	0.02	0.98	0.00	1.00	0.00	1.00
38	0.02	0.98	0.00	1.00	0.00	1.00
39	0.03	0.97	0.01	0.99	0.00	1.00
40	0.04	0.96	0.01	0.99	0.00	1.00
41	0.03	0.97	0.01	0.99	0.00	1.00
42	0.02	0.98	0.00	1.00	0.00	1.00
43	0.06	0.94	**0.00**	**1.00**	**0.00**	**1.00**
44	0.05	0.95	0.01	0.99	0.00	1.00
45	0.14	0.86	0.04	0.96	0.00	1.00

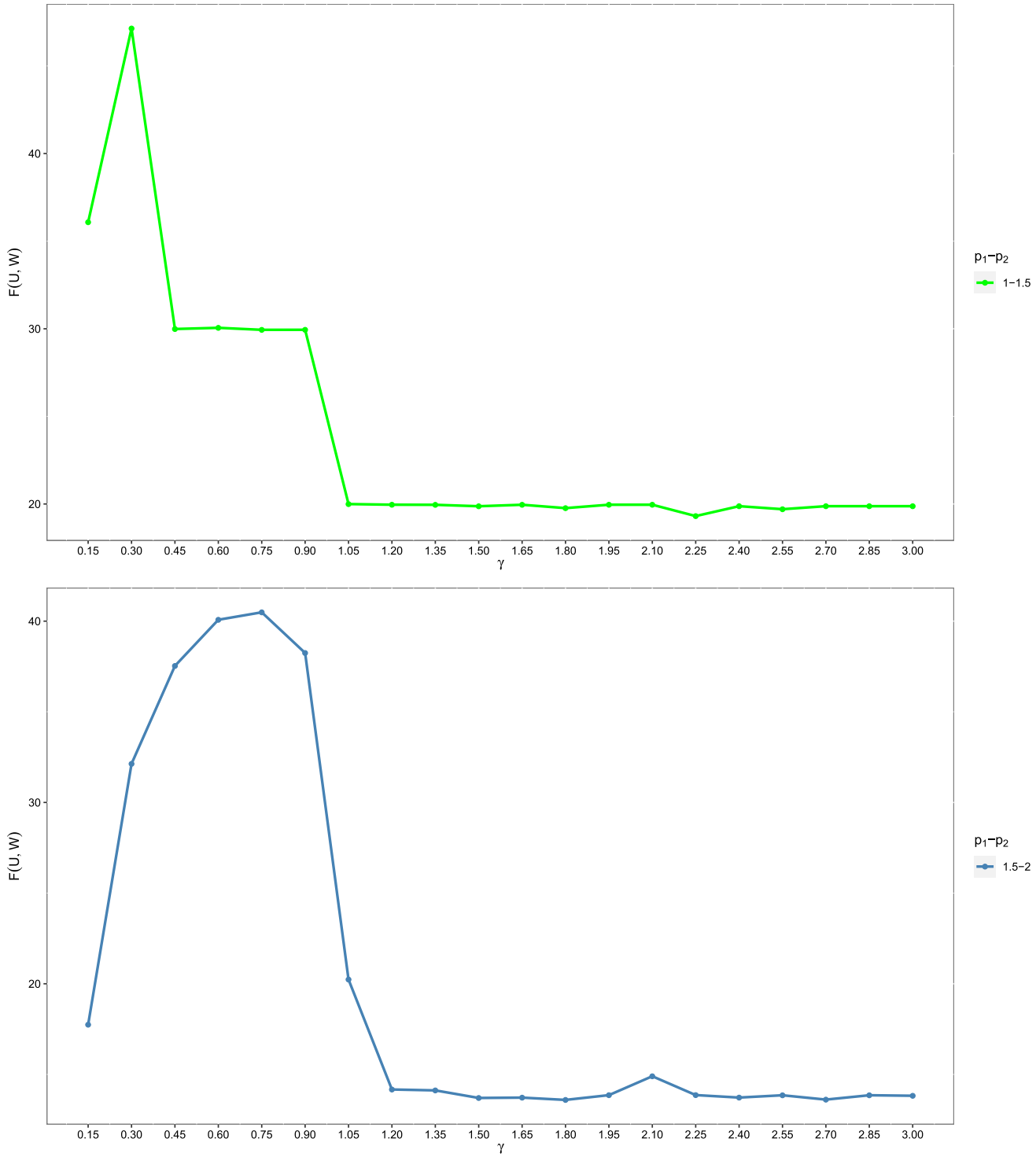

Fig. 3 Third scenario: values of the internal validity index when C = 3 setting $p_1 = 1,\ p_2 = 1.5$ (green line) and C = 4 and $p_1 = 1.5,\ p_2 = 2$ (blue line)

Table 1 continued

	$\gamma = 0$		$\gamma = 0.30$		$\gamma = 0.75$	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
46	0.13	0.87	0.04	0.96	0.00	1.00
47	0.04	0.96	0.01	0.99	0.00	1.00
48	0.10	0.90	0.03	0.97	0.00	1.00
49	**0.00**	**1.00**	0.04	0.96	0.00	1.00
50	*0.35*	*0.65*	*0.16*	*0.84*	*0.01*	*0.99*

The medoids are highlighted in bold, while the fuzzy units are in italics

Table 2 Second scenario with $p_1 = 1.5$, $p_2 = 2$: **W** membership degrees matrix for $\gamma \in \{0, 0.30, 0.75\}$

	$\gamma = 0$		$\gamma = 0.30$		$\gamma = 0.75$	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
1	0.20	0.80	**1.00**	**0.00**	**1.00**	**0.00**
2	0.89	0.11	0.91	0.09	1.00	0.00
3	0.81	0.19	0.91	0.09	1.00	0.00
4	0.12	0.88	0.85	0.15	1.00	0.00
5	0.17	0.83	0.87	0.13	1.00	0.00
6	0.03	0.97	0.98	0.02	1.00	0.00
7	0.94	0.06	0.92	0.08	1.00	0.00
8	0.65	0.35	0.97	0.03	1.00	0.00
9	0.88	0.12	0.89	0.11	1.00	0.00
10	0.13	0.87	0.98	0.02	1.00	0.00
11	0.50	0.50	0.02	0.98	0.00	1.00
12	0.29	0.71	0.37	0.63	0.01	0.99
13	0.65	0.35	0.05	0.95	0.00	1.00
14	0.09	0.91	0.20	0.80	0.01	0.99
15	**1.00**	**0.00**	0.10	0.90	0.00	1.00
16	0.59	0.41	0.03	0.97	0.00	1.00
17	0.07	0.93	0.06	0.94	0.00	1.00
18	**0.00**	**1.00**	0.16	0.84	0.00	1.00
19	0.54	0.46	0.08	0.92	0.00	1.00
20	0.29	0.71	**0.00**	**1.00**	**0.00**	**1.00**

The medoids are highlighted in bold, while the fuzzy units are in italics

$\mathcal{M}$ must necessarily be placed in the second cluster, while the remaining units $(m_{11}, \ldots, m_{20})$ should be in the first cluster, along with units $(n_{31}, \ldots, n_{50})$ in order to respect the network topology.

Second scenario

In the second scenario, there is a clustering structure in terms of attributes only for the set $\mathcal{N}$, while the structure of the network is the same as in the first scenario, distinguishing the same two groups within both the set $\mathcal{N}$ and the set $\mathcal{M}$.

Looking at the values taken by the internal validity index, as reported in Tables 14 and 15, the expected best solution is still based on two groups when $\gamma > 0$. Obviously, in this case, the baseline scenario does not make sense for the set $\mathcal{M}$. This is because there is no clustering structure for the attributes. Figure 2 leads to the same conclusions as in the first scenario. We examine the membership degrees matrices **U** and **W** in Tables 1 and 2, focusing on the fuzzier solution as before, comparing the optimal solution $C^* = 2$ and $\gamma = 0.75$ with the intermediate ($\gamma = 0.30$) and baseline ($\gamma = 0$) solutions based on the same number of groups.

Once again, as γ increases, the **U** matrix becomes crisper: the degree of membership in the original cluster gets stronger with increasing contribution of the network. More importantly, when γ is set to its optimal value, the partitioning in **W** reflects that of the network, assigning units in the two clusters according to the network's connections between sets $\mathcal{N}$ and $\mathcal{M}$.

Similar results were obtained when a crisper solution was considered, as can be seen by looking at the partition in the Tables 16 and 17.

Third scenario

Table 3 Third scenario with $p_1 = 1.5,\ p_2 = 2$: **U** membership degrees matrix for $\gamma \in \{0, 0.30, 0.75\}$

	$\gamma = 0$		$\gamma = 0.30$			$\gamma = 0.75$			
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3	Cluster 4
1	0.85	0.15	0.38	0.61	0.00	0.37	0.63	0.00	0.00
2	0.89	0.11	0.41	0.59	0.00	0.40	0.60	0.00	0.00
3	0.80	0.20	0.37	0.62	0.01	0.35	0.65	0.00	0.00
4	0.89	0.11	0.32	0.68	0.00	0.29	0.71	0.00	0.00
5	0.87	0.13	0.41	0.59	0.00	0.40	0.60	0.00	0.00
6	0.71	0.29	0.28	0.70	0.02	0.25	0.75	0.00	0.00
7	0.88	0.12	0.32	0.67	0.00	0.30	0.70	0.00	0.00
8	0.96	0.04	0.42	0.58	0.00	0.41	0.59	0.00	0.00
9	0.95	0.05	0.41	0.59	0.00	0.40	0.60	0.00	0.00
10	0.96	0.04	0.41	0.59	0.00	0.40	0.60	0.00	0.00
11	0.94	0.06	0.42	0.58	0.00	0.41	0.59	0.00	0.00
12	0.95	0.05	0.42	0.58	0.00	0.41	0.59	0.00	0.00
13	0.95	0.05	0.43	0.57	0.00	0.43	0.57	0.00	0.00
14	0.95	0.05	0.40	0.60	0.00	0.39	0.61	0.00	0.00
15	0.88	0.12	0.31	0.68	0.00	0.28	0.72	0.00	0.00
16	0.84	0.16	0.41	0.59	0.00	0.40	0.60	0.00	0.00
17	0.83	0.17	0.38	0.61	0.01	0.37	0.63	0.00	0.00
18	0.84	0.16	0.39	0.61	0.00	0.38	0.62	0.00	0.00
19	0.80	0.20	0.33	0.66	0.01	0.30	0.70	0.00	0.00
20	0.64	0.36	0.27	0.69	0.04	0.25	0.75	0.00	0.00
21	0.76	0.24	0.28	0.70	0.01	0.25	0.74	0.00	0.00
22	0.93	0.07	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**	**0.00**	**0.00**
23	0.90	0.10	0.41	0.59	0.00	0.40	0.60	0.00	0.00
24	0.90	0.10	0.42	0.57	0.00	0.42	0.58	0.00	0.00
25	0.88	0.12	0.31	0.68	0.00	0.29	0.71	0.00	0.00
26	0.74	0.26	0.28	0.70	0.02	0.26	0.74	0.00	0.00
27	0.94	0.06	0.35	0.65	0.00	0.33	0.67	0.00	0.00
28	0.95	0.05	0.43	0.57	0.00	0.42	0.58	0.00	0.00
29	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**	**0.00**
30	*0.42*	*0.58*	*0.19*	*0.60*	*0.21*	*0.19*	*0.76*	*0.03*	*0.02*
31	0.07	0.93	0.00	0.00	0.99	0.00	0.00	0.49	0.51
32	0.01	0.99	0.00	0.00	1.00	0.00	0.00	0.46	0.54
33	0.06	0.94	0.00	0.00	1.00	0.00	0.00	0.51	0.49
34	0.05	0.95	0.00	0.00	1.00	0.00	0.00	0.46	0.54
35	0.04	0.96	0.00	0.00	1.00	0.00	0.00	0.55	0.45
36	0.03	0.97	0.00	0.00	1.00	0.00	0.00	0.47	0.53
37	0.02	0.98	0.00	0.00	1.00	0.00	0.00	0.46	0.54
38	0.02	0.98	0.00	0.00	1.00	0.00	0.00	0.44	0.56
39	0.03	0.97	0.00	0.00	1.00	0.00	0.00	0.53	0.47
40	0.04	0.96	0.00	0.00	1.00	0.00	0.00	0.47	0.53
41	0.03	0.97	0.00	0.00	1.00	0.00	0.00	0.47	0.53

Table 3 continued

	$\gamma = 0$		$\gamma = 0.30$			$\gamma = 0.75$			
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3	Cluster 4
42	0.02	0.98	0.00	0.00	1.00	0.00	0.00	0.47	0.53
43	0.06	0.94	**0.00**	**0.00**	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**
44	0.05	0.95	0.00	0.00	1.00	0.00	0.00	0.47	0.53
45	0.14	0.86	0.01	0.01	0.98	0.00	0.00	0.56	0.44
46	0.13	0.87	0.01	0.01	0.98	0.00	0.00	0.56	0.44
47	0.04	0.96	0.00	0.00	1.00	**0.00**	**0.00**	**0.00**	**1.00**
48	0.10	0.90	0.00	0.01	0.99	0.00	0.00	0.56	0.44
49	**0.00**	**1.00**	0.01	0.01	0.98	0.00	0.00	0.57	0.43
50	*0.35*	*0.65*	*0.12*	*0.12*	*0.76*	*0.00*	*0.00*	*0.64*	*0.35*

The medoids are highlighted in bold, while the fuzzy units are in italics

Table 4 Third scenario with $p_1 = 1.5,\ p_2 = 2$: **W** membership degrees matrix for $\gamma \in \{0, 0.30, 0.75\}$

	$\gamma = 0$		$\gamma = 0.30$			$\gamma = 0.75$			
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3	Cluster 4
1	0.04	0.96	0.00	0.96	0.04	0.00	1.00	0.00	0.00
2	0.02	0.98	0.00	0.95	0.05	0.00	1.00	0.00	0.00
3	0.08	0.92	0.00	0.97	0.03	0.00	1.00	0.00	0.00
4	**0.00**	**1.00**	0.00	0.95	0.04	0.00	1.00	0.00	0.00
5	0.07	0.93	**0.00**	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**	**0.00**
6	0.01	0.99	0.00	0.04	0.96	0.00	0.00	1.00	0.00
7	0.04	0.96	0.00	0.05	0.95	0.00	0.00	1.00	0.00
8	0.05	0.95	**0.00**	**0.00**	**1.00**	0.00	0.00	1.00	0.00
9	0.03	0.97	0.00	0.04	0.96	**0.00**	**0.00**	**1.00**	**0.00**
10	0.03	0.97	0.00	0.03	0.97	0.00	0.00	1.00	0.00
11	0.98	0.02	0.98	0.00	0.02	0.01	0.00	0.00	0.99
12	0.98	0.02	0.97	0.00	0.03	0.01	0.00	0.00	0.99
13	0.98	0.02	0.98	0.00	0.02	**0.00**	**0.00**	**0.00**	**1.00**
14	0.90	0.10	0.61	0.02	0.37	0.01	0.00	0.01	0.99
15	0.96	0.04	0.91	0.00	0.08	0.01	0.00	0.00	0.99
16	0.97	0.03	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**	**0.00**	**0.00**
17	**1.00**	**0.00**	0.97	0.03	0.00	0.96	0.04	0.00	0.00
18	0.98	0.02	0.99	0.01	0.00	0.98	0.01	0.00	0.00
19	0.99	0.01	1.00	0.00	0.00	0.99	0.00	0.00	0.01
20	*0.48*	*0.52*	*0.19*	*0.78*	*0.03*	*0.05*	*0.95*	*0.00*	*0.00*

The medoids are highlighted in bold, while the fuzzy units are in italics

This scenario is particularly interesting because the algorithm must balance two different grouping structures, leading to non-trivial expected results. If γ is greater than zero, the optimal solution should be based on four groups. This requires a partition of the units in the set $\mathcal{N}$ which may seem counter-intuitive and is discussed further below.

Depending on which set of weighting factors $p_1,\ p_2$ is considered, the internal validity index (13) leads to a different solution for both optimal C^* and γ, as shown in Tables 18 and 19.

When $p_1 = 1,\ p_2 = 1.5$, the index reaches its maximum value at $C^* = 3$ and $\gamma = 0.3$ while for $p_1 = 1.5,\ p_2 = 2$ at $C^* = 4$ and $\gamma = 0.75$. Figure 3 shows that the internal validity index decreases significantly after reaching its maximum. This is because, for higher values of γ parameter, the optimal solution must necessarily become more similar to that of the network, resulting in the selection of $C^* = 2$. Focusing on the fuzzier solution, shown in Tables 3 and 4, the membership of the first thirty units $(n_1, \ldots, n_{30})$ is almost equally distributed between only two clusters. These clusters match

those in the **W** matrix including the set of units $m_1, \ldots, m_5$ and $m_{16}, \ldots, m_{20}$, respectively. Meanwhile, the membership of units $n_{31}, \ldots, n_{50}$ are distributed almost equally between the other two clusters, which correspond in the **W** matrix to the two clusters that contain the set of units $m_6, \ldots, m_{10}$, and $m_{11}, \ldots, m_{15}$, respectively. This solution may not seem intuitive, but it is the most plausible option considering the associated network structure and a certain degree of fuzziness. When the gamma value is less than the optimal one, the best solution relies on three groups. For set $\mathcal{N}$, the membership degrees of units $n_1, \ldots, n_{30}$ are still almost equally distributed between two clusters, while units $n_{31}, \ldots, n_{50}$ belong to a single group.

As far as the **W** matrix is concerned, the algorithm is not able to split the units $m_{11}, \ldots, m_{19}$ according to the network structure, so they are all assigned in the same first cluster. They matched the first 30 units of set $\mathcal{N}$ instead of the last 20 because there were more existing connections.

The fuzzy unit m_{20} is assigned to the second group along with units $m_1, \ldots, m_5$. The results were the same when a crisper solution was considered, as can be seen by looking at the partition in the Tables 20 and 21.

5 Application to real data: bibliometric data

In this section, we test our algorithm on bibliometric data extracted using the Scopus API. We extracted the data on the analysis of the effectiveness of nutritonal value labels on food packages, which is something on which both scientists and policy-makers have been divided over the last decade. As of 2023, 7 European countries have officially adopted the Nutriscore package labeling (Belgium, France, Germany, Luxembourg, The Netherlands, Spain and Switzerland), and there has been a significant push to have it adopted at the European Union level (de Ruiter 2023). We used the following search query to extract all the papers on the topic since year 2014:

```
(TITLE-ABS-KEY(Nutri Score) OR TITLE-ABS-KEY(Nutri-Score) OR TITLE-ABS-KEY (NutriScore) OR TITLE-ABS-KEY(Front-of -pack nutritional label) OR TITLE-ABS -KEY(FOP nutritional label)) AND PUBYEAR > 2013
```

This search returned a total of 588 papers from 2047 different authors. Considering the extreme sparsity of such data, with roughly 80% of the authors having only a single paper among those we collected, we restricted ourselves to the 64 authors with at least 5 papers in the dataset, and the 254 papers they authored.

We defined $\mathcal{N}$ as the set of authors and $\mathcal{M}$ as the set of papers, and for each of them we collected the following sets of numerical attributes:

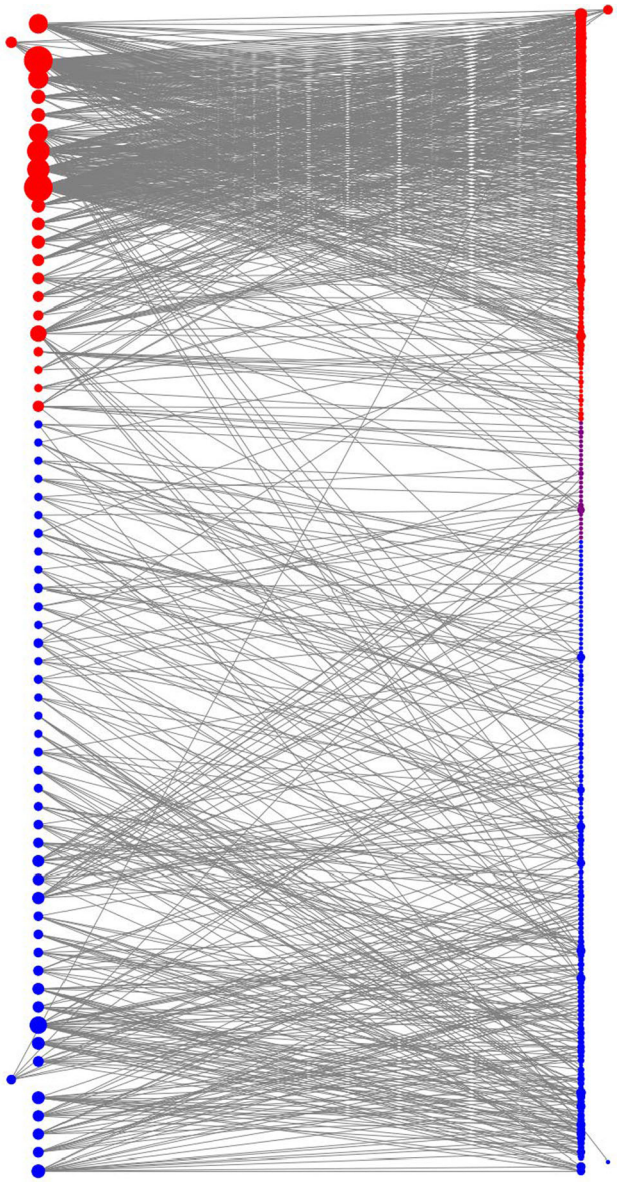

Fig. 4 The associated bipartite graph for the application. Authors are placed on the left and papers on the right, the size of a unit represents its degree in the graph. The vertices representing units belonging to the same cluster, as detected from the algorithm (with $\gamma = 2.25$, $p_1 = 3$, $p_2 = 1.5$, $C = 2$) are coloured red for cluster 1 and blue for cluster 2. Fuzzy units are coloured purple. The left and right medoids of each cluster are highlighted by placing them out of line

- $\mathcal{I}$:= {Years since first paper published, average papers published per year, average citations received per published paper}.
- $\mathcal{J}$:= {Years since publication, number of references in the paper, number of citations received}.

We also defined the adjacency matrix **A** such that $a_{nm} = 1$ if author n wrote paper m and $a_{ij} = 0$ otherwise. The

Table 5 Application with $p_1 = 3,\ p_2 = 1.5$: the proposed internal validity index for the 126 partitions (values are rounded to the nearest unit)

C	γ																			
	0.15	0.30	0.45	0.60	0.75	0.90	1.05	1.20	1.35	1.50	1.65	1.80	1.95	2.10	2.25	2.40	2.55	2.70	2.85	3.00
2	0.8	1.2	1.8	2.9	5.6	8.6	16.2	18.4	20.4	21.9	23.1	24.1	25.5	26.4	27.0	27.6	28.1	28.5	28.9	29.2
3	1.9	3.1	5.0	7.9	11.1	13.9	16.6	18.7	20.0	21.0	21.8	22.5	23.0	23.0	23.3	23.5	23.8	24.1	24.2	24.5
4	0.3	2.6	4.8	6.5	9.0	12.0	15.1	16.6	17.4	18.3	18.4	19.3	20.0	19.5	19.6	19.5	19.6	19.7	19.7	19.7
5	1.2	2.6	4.0	5.3	7.1	10.0	10.6	13.9	15.1	15.7	16.3	16.5	16.8	17.0	17.0	17.4	17.5	17.6	17.5	17.5
6	0.5	1.4	3.1	3.9	6.4	9.0	10.8	11.9	12.4	14.1	15.3	15.6	14.9	14.8	15.2	15.1	15.1	15.2	14.7	14.8

Table 6 Application with $p_1 = 3,\ p_2 = 1.5$: **U** membership degrees matrix for $\gamma \in \{0, 0.90, 2.25\}$

	$\gamma = 0$			$\gamma = 0.90$			$\gamma = 2.25$	
	C1	C2	C3	C1	C2	C3	C1	C2
Cruz-Casarrubias C	0.30	0.34	0.37	0.02	0.11	0.87	0.05	0.95
Barquera S	0.52	0.31	0.17	0.70	0.11	0.19	0.00	1.00
Pettigrew S	0.33	0.35	0.32	0.85	0.12	0.04	1.00	0.00
Jongenelis M.I	0.32	0.35	0.33	0.25	0.27	0.47	0.83	0.17
Ares G	0.38	0.36	0.26	0.00	0.99	0.01	0.00	1.00
Machin L	0.32	0.33	0.35	0.01	0.91	0.09	0.00	1.00
Antunez L	0.33	0.34	0.33	0.02	0.84	0.14	0.00	1.00
Muzzioli L	0.28	0.32	0.40	0.02	0.07	0.91	0.02	0.98
Pinto A	0.30	0.32	0.38	0.04	0.09	0.87	0.01	0.99
Finkelstein E.A	0.38	0.34	0.28	**1.00**	**0.00**	**0.00**	0.09	0.91
Curutchet M.R	0.32	0.34	0.34	0.01	0.90	0.09	0.00	1.00
Gimenez A	0.33	0.33	0.34	0.01	0.97	0.02	0.00	1.00
Sarda B	0.29	0.33	0.38	0.06	0.22	0.72	0.85	0.15
Hercberg S	0.39	0.34	0.27	1.00	0.00	0.00	1.00	0.00
Galan P	0.38	0.34	0.28	1.00	0.00	0.00	1.00	0.00
Touvier M	0.48	0.34	0.18	0.99	0.01	0.00	1.00	0.00
Deschasaux M	0.35	0.35	0.31	0.53	0.36	0.11	1.00	0.00
Ducrot P	0.31	0.33	0.35	0.02	0.89	0.10	1.00	0.00
Kesse-Guyot E	0.37	0.35	0.28	0.66	0.34	0.00	1.00	0.00
Julia C	0.36	0.35	0.29	1.00	0.00	0.00	1.00	0.00
Graham D.J	0.31	0.32	0.37	0.06	0.57	0.38	0.14	0.86
Mazzu M.F	0.30	0.34	0.36	0.00	0.02	0.97	0.00	1.00
Baccelloni A	0.30	0.34	0.36	0.00	0.03	0.96	0.00	1.00
Salas-Salvado J	0.37	0.34	0.30	0.96	0.03	0.01	0.97	0.03
Babio N	0.33	0.34	0.33	0.70	0.17	0.13	**1.00**	**0.00**
Srour B	0.35	0.35	0.30	0.65	0.19	0.16	0.97	0.03
Fialon M	0.30	0.34	0.36	0.45	0.11	0.44	0.99	0.01
Donini L.M	0.34	0.33	0.33	0.15	0.13	0.71	0.01	0.99
Pellegrini N	0.47	0.31	0.22	0.85	0.11	0.04	0.07	0.93
Talati Z	0.31	0.33	0.37	0.19	0.58	0.23	0.95	0.05
Ahmed M	0.29	0.32	0.38	0.04	0.19	0.76	0.07	0.93
Mejean C	0.34	0.34	0.32	0.01	0.95	0.03	0.99	0.01
Andreeva V.A	**0.00**	**0.00**	**1.00**	0.59	0.19	0.22	0.96	0.04
Hall M.G	0.33	0.35	0.32	0.07	0.17	0.76	0.01	0.99
Vanderlee L	0.31	0.34	0.35	0.03	0.13	0.84	0.00	1.00

Table 6 continued

	$\gamma = 0$			$\gamma = 0.90$			$\gamma = 2.25$	
	C1	C2	C3	C1	C2	C3	C1	C2
Hammond D	0.37	0.36	0.27	0.39	0.26	0.35	0.00	1.00
Romani S	0.33	0.33	0.35	0.03	0.11	0.86	0.02	0.98
Taillie L.S	0.35	0.36	0.29	0.15	0.31	0.54	0.01	0.99
Drewnowski A	0.35	0.32	0.33	0.83	0.08	0.09	0.01	0.99
Vandevijvere S	0.35	0.35	0.29	0.77	0.12	0.11	0.81	0.19
Contreras-Manzano A	0.30	0.33	0.37	0.01	0.12	0.87	0.01	0.99
Jauregui A	0.38	0.34	0.29	0.08	0.17	0.75	0.01	0.99
Vargas-Meza J	0.29	0.33	0.37	0.01	0.12	0.86	0.04	0.96
Nieto C	0.29	0.33	0.38	0.02	0.11	0.87	0.04	0.96
Miller C	0.29	0.32	0.38	0.08	0.59	0.33	0.08	0.92
Martinez-Gonzalez M.A	0.38	0.35	0.27	0.93	0.06	0.01	0.13	0.87
Bes-Rastrollo M	0.36	0.35	0.29	0.53	0.28	0.20	0.09	0.91
Anastacio L.R	0.29	0.33	0.38	0.06	0.18	0.76	0.14	0.86
Egnell M	0.33	0.34	0.34	0.89	0.05	0.06	1.00	0.00
Peneau S	**0.00**	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**	1.00	0.00
Fezeu L.K	0.34	0.33	0.33	0.04	0.91	0.05	0.99	0.01
Woods J.L	0.30	0.32	0.38	0.05	0.23	0.72	0.10	0.90
Kelly B	**1.00**	**0.00**	**0.00**	0.10	0.77	0.13	**0.00**	**1.00**
Acton R.B	0.30	0.33	0.37	0.02	0.24	0.73	0.02	0.98
Mhurchu C.N	0.34	0.33	0.33	0.32	0.48	0.20	0.08	0.92
Neal B	0.52	0.32	0.16	0.71	0.28	0.00	0.01	0.99
Gregorio M.J	0.30	0.33	0.37	**0.00**	**0.00**	**1.00**	0.25	0.75
Rempe C	0.36	0.38	0.26	0.91	0.07	0.02	0.88	0.12
L'Abbe M.R	0.29	0.31	0.40	0.19	0.48	0.33	0.06	0.94
Ruffieux B	0.30	0.31	0.39	0.06	0.62	0.33	0.12	0.88
Muller L	0.29	0.32	0.38	0.04	0.54	0.42	0.16	0.84
Hamlin R	0.28	0.31	0.41	0.04	0.56	0.40	0.15	0.85
Aschemann-Witzel J	0.34	0.33	0.32	0.06	0.82	0.12	0.02	0.98
Dixon H	0.30	0.32	0.38	0.05	0.75	0.20	0.04	0.96

The medoids are highlighted in bold

visualization of the associated bipartite network is given in Fig. 4.

We ran the clustering algorithm using as input the attribute matrices $\mathbf{X}$ of author attributes and $\mathbf{Y}$ of paper attributes and the adjacency matrix $\mathbf{A}$, and setting $p_1 = 3$, $p_2 = 1.5$, for different values of γ ranging from 0.15 to 3 and of C ranging from 2 to 6. In Table 5 are reported the values of validity indices. As we can observe, the validity index tends to favor the outputs with high values of γ and $C = 2$. This can be interpreted as a signal that in the data utilised, the network-based community structure is stronger than the attribute-based cluster structure. We show in Tables 6 and 7 the outcomes of the algorithm for $\gamma = 0$, for $\gamma = 0.90$, that is, a value such that the attribute structure makes the algorithm choose 3 clusters, and $\gamma = 2.25$, that is, a value of γ for which the validity index starts plateauing around the optimal value. We understand the fact that the optimal value of the validity index keeps growing with γ as the algorithm considering the network structure of the data to have a stronger "signal" than the attribute structure, and thus being worth prioritizing, similar to the Second Scenario of the simulation. This is unsurprising as the dataset was chosen specifically about a topic which is divisive, so that it would contain a community structure in the adjacency matrix.

We look at how well the clustering algorithm is able to extract insight into the ongoing academic (and political) debate around the Nutriscore adoption in the European Union. We observe that the current division is reflected in the optimal partition found by the clustering. Out of 32 authors in the database with affiliation to a European university, 23 come from one of the 7 countries which have already adopted the Nutriscore at a nationwide level (see Fig. 5). We find that 19 of them are classified into cluster 1 and only 4 into cluster 2, while all the 9 authors from other European countries (of

which 7 are from Italy, the strongest opponent to the EU-wide adoption) are classified into cluster 2.

As expected, given that the validity index favored the outputs with a high value of γ, that is, which give more weight to the network structure than the attributes, we see that the two clusters found for $\gamma = 2.25$ are valid communities in the network (see Fig. 4).

6 Conclusions

In this paper we have defined the Fuzzy C-medoids clustering with Barber modularity (FMd-BMR) regularization as a way to produce a joint classification of two disjoint sets $\mathcal{N}$ and $\mathcal{M}$ of statistical units, for which we are provided data about the values of two different sets of attributes, $\mathcal{I}$ for the units in $\mathcal{N}$ and $\mathcal{J}$ for the units in $\mathcal{M}$, and about a bipartite adjacency structure. We tested our algorithm on both real and simulated data, and verified that indeed it is able to extract many different types of information from the data. In particular, using the validity index we defined in Sect. 3.2 to optimize not only the number of clusters but also the value of the parameter γ we can get different outputs based on the relation between the attributes and the adjacency structure. If, as in Scenario 1, the partition we would identify looking only at the attributes and only at the adjacency are the same, the FMd-BMR correctly outputs that partition. Instead, when one of the two attribute matrices, as in Scenario 2, contains no valuable information, the FMd-BMR outputs a meaningful partition of the corresponding set based on the adjacency structure of the bipartite network. Further, when, as in Scenario 3, the partitions we would find based either only on the attributes or only on the network would be different, the FMd-BMR algorithm is able to combine both information to produce a partition that gives valuable information about both. Finally, in the application to real data, we show how the FMd-BMR algorithm is able to realize that the division in the scientific community makes so that the network community structure in the autorship network is stronger than the associated attribute cluster structure and so it is prioritized in the choice of γ.

We believe this versatility to be a great strength of the model, as it is able to capture a lot of different data structures. Moreover, the way in which the fuzzyness is used in Scenario 3, not to indicate units that are between two distinct clusters but to match two different clusters in the set $\mathcal{M}$ to a single cluster in the set $\mathcal{N}$ is, as far as we know, novel. This shows that there is a lot of untapped potential in the use fuzzy clustering algorithm to capture different patterns in data than the one they were originally designed for.

We also believe that the use of the Barber modularity as a spatial regularization term, instead of the usual spatial penalties has greater potential to incorporate the network structure into the clustering process as equal to the attributes and not only as a correction term. Indeed, spatial penalty terms are limited by the fact that they only penalize partitions in which adjacent units are in different clusters, but not those in which non-adjacent units are in the same cluster. Consequently, if the relative weight of the spatial term compared to the attribute-based term in the objective function is too high, the optimal partition becomes the one in which all elements are in the same cluster. The Barber modularity avoids this, as the matrix $\mathbf{B}$ contains negative entries for non-adjacent units.

In conclusion, this methodological framework can be further extended to incorporate alternative distances, such as the Gower distance for mixed qualitative and quantitative data or the exponential distance for datasets containing outliers. These enhancements would broaden the applicability and robustness of the proposed approach.

Appendix A Derivation of Eqs. (11) and (12).

In this appendix we show the computations required to obtain the optimal values of $u_{n,c}$ and $w_{m,c}$ use to update the matrices $\mathbf{U}$ and $\mathbf{W}$ in Algorithm 1.

We start by writing the Lagrangian function for the units in the set $\mathcal{N}$:

$$\begin{aligned} L(u_{n,c},\lambda) = & \sum_{n=1}^{N}\sum_{c=1}^{C} u_{n,c} d^2(\mathbf{x}_n,\mathbf{x}_c) + \sum_{m=1}^{M}\sum_{c=1}^{C} w_{m,c} d^2(\mathbf{y}_m,\mathbf{y}_c) \\ & + p_1 \sum_{n=1}^{N}\sum_{c=1}^{C} u_{n,c}\log(u_{n,c}) \\ & + p_2 \sum_{m=1}^{M}\sum_{c=1}^{C} w_{m,c}\log(w_{m,c}) \\ & - \gamma \sum_{n=1}^{N}\sum_{c=1}^{C}\sum_{m=1}^{M} u_{n,c} b_{n,m} w_{m,c} \\ & - \lambda\Big(\sum_{c=1}^{C} u_{n,c} - 1\Big). \end{aligned} \tag{A1}$$

We compute next its derivative with respect to λ. We obtain:

$$\frac{\partial L(u_{n,c},\lambda)}{\partial \lambda} = \sum_{c=1}^{C} u_{n,c} - 1, \tag{A2}$$

so that

$$\frac{\partial L(u_{n,c},\lambda)}{\partial \lambda} = 0 \Longleftrightarrow \sum_{c=1}^{C} u_{n,c} = 1. \tag{A3}$$

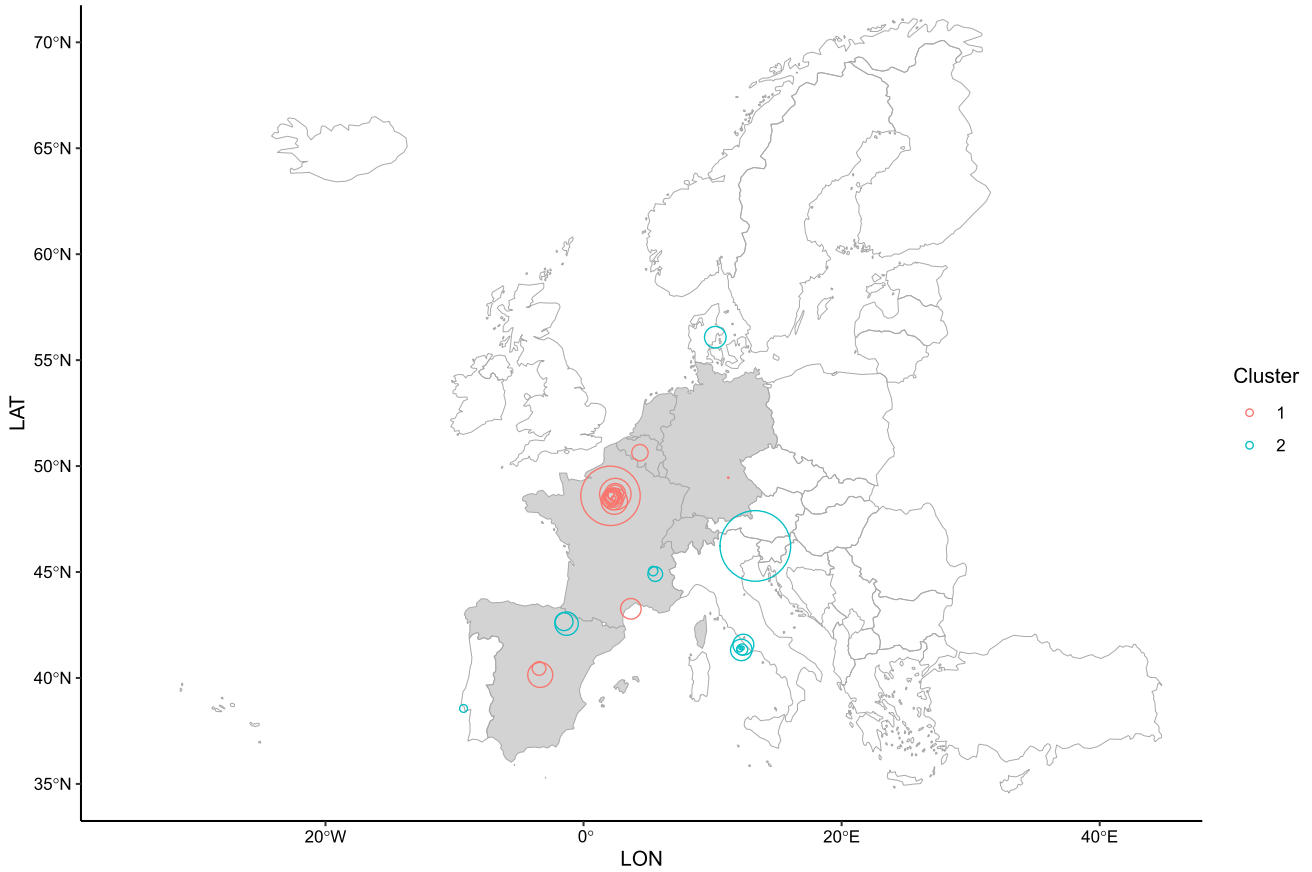

Fig. 5 The map depicts clustering outcomes, with circles colored according to distinct crisp cluster memberships and their sizes reflecting the rescaled ratio between citation numbers and papers. The seven countries that officially adopted the Nutriscore package are shaded in gray

We next compute the partial derivative with respect to $u_{n,c}, n \leq N$:

$$\frac{\partial L(u_{n,c}, \lambda)}{\partial u_{n,c}} = d^2(\mathbf{x}_l, \mathbf{x}_c) + p_1(1 + \log u_{n,c}) - \gamma \sum_{m=1}^{M} b_{n,m} w_{m,c} - \lambda. \tag{A4}$$

We then solve in $u_{n,c}$:

$$\log u_{n,c} = \frac{\lambda}{p_1} - \frac{1}{p_1} d^2(\mathbf{x}_n, \mathbf{x}_c) + \frac{\gamma}{p_1} \sum_{m=1}^{M} b_{n,m} w_{m,c} - 1$$
$$\iff \frac{\partial L(u_{l,c}, \lambda)}{\partial u_{n,c}} = 0 \quad \forall c \leq C \quad n \leq N. \tag{A5}$$

Consequently,

$$u_{n,c} = \exp\left\{ \frac{\lambda}{p_1} - \frac{1}{p_1} d^2(\mathbf{x}_n, \mathbf{x}_c) + \frac{\gamma}{p_1} \sum_{m=1}^{M} b_{n,m} w_{m,c} - 1 \right\}$$
$$c \leq C \quad n \leq N. \tag{A6}$$

Recalling (A3) we find that

$$\exp\left\{ \frac{\lambda}{p_1} - 1 \right\} = \frac{1}{\sum_{c=1}^{C} \exp\left\{ -\frac{1}{p_1} d^2(\mathbf{x}_n, \mathbf{x}_c) + \frac{\gamma}{p_1} \sum_{m=1}^{M} b_{n,m} w_{m,c} \right\}}, \tag{A7}$$

so that, for all $n \leq N$, $c \leq C$,

$$u_{n,c} = \frac{\exp\left\{ -\frac{1}{p_1} d^2(\mathbf{x}_n, \mathbf{x}_c) + \frac{\gamma}{p_1} \sum_{m=1}^{M} b_{n,m} w_{m,c} \right\}}{\sum_{c=1}^{C} \exp\left\{ -\frac{1}{p_1} d^2(\mathbf{x}_n, \mathbf{x}_c) + \frac{\gamma}{p_1} \sum_{m=1}^{M} b_{n,m} w_{m,c} \right\}}. \tag{A8}$$

We have thus obtained (11).

Equivalently, for the units in the set $\mathcal{M}$, the Lagrangian function is

$$L(w_{m,c}, \lambda) = \sum_{n=1}^{N} \sum_{c=1}^{C} u_{n,c} d^2(\mathbf{x}_n, \mathbf{x}_c) + \sum_{m=1}^{M} \sum_{c=1}^{C} w_{m,c} d^2(\mathbf{y}_m, \mathbf{y}_c)$$

$$
\begin{aligned}
&+ p_1 \sum_{n=1}^{N} u_{n,c} \log(u_{n,c}) \\
&+ p_2 \sum_{m=1}^{M} w_{m,c} \log(w_{m,c}) \\
&- \gamma \sum_{n=1}^{N} \sum_{c=1}^{C} \sum_{m=1}^{M} u_{n,c} b_{n,m} w_{m,c} \\
&- \lambda \Big(\sum_{c=1}^{C} w_{m,c} - 1 \Big).
\end{aligned} \tag{A9}
$$

so that

$$
\frac{\partial L(w_{m,c}, \lambda)}{\partial \lambda} = 0 \Longleftrightarrow \sum_{c=1}^{C} w_{m,c} = 1. \tag{A10}
$$

For the elements of $\mathcal{M}$,

$$
\begin{aligned}
\frac{\partial L(w_{m,c}, \lambda)}{\partial w_{m,c}} &= d^2(\mathbf{y}_m, \mathbf{y}_c) + p_2(1 + \log w_{m,c}) \\
&\quad - \gamma \sum_{n=1}^{N} b_{n,m} u_{n,c} - \lambda.
\end{aligned} \tag{A11}
$$

We thus compute,

$$
\begin{aligned}
\log w_{m,c} &= \frac{\lambda}{p_2} - \frac{1}{p_2} d^2(\mathbf{y}_m, \mathbf{y}_c) \\
&\quad + \frac{\gamma}{p_2} \sum_{n=1}^{N} b_{n,m} u_{n,c} - 1 \\
&\Longleftrightarrow \frac{\partial L(w_{m,c}, \lambda)}{\partial w_{m,c}} = 0, \qquad \forall c \le C, \quad m \le M.
\end{aligned} \tag{A12}
$$

Consequently,

$$
\begin{aligned}
w_{m,c} &= \exp\left\{ \frac{\lambda}{p_2} - \frac{1}{p_2} d^2(\mathbf{y}_m, \mathbf{y}_c) + \frac{\gamma}{p_2} \sum_{n=1}^{N} b_{n,m} u_{n,c} - 1 \right\}, \\
&\quad c \le C, \quad m \le M.
\end{aligned} \tag{A13}
$$

Recalling (A10) we find that

$$
\exp\left\{ \frac{\lambda}{p_2} - 1 \right\} = \frac{1}{\sum_{c=1}^{C} \exp\left\{ -\frac{1}{p_2} d^2(\mathbf{y}_m, \mathbf{y}_c) + \frac{\gamma}{p_2} \sum_{n=1}^{N} b_{n,m} u_{n,c} \right\}}, \tag{A14}
$$

so that, for all $m \le M$, $c \le C$,

$$
w_{m,c} = \frac{\exp\left\{ -\frac{1}{p_2} d^2(\mathbf{y}_m, \mathbf{y}_c) + \frac{\gamma}{p_2} \sum_{n=1}^{N} b_{n,m} u_{n,c} \right\}}{\sum_{c=1}^{C} \exp\left\{ -\frac{1}{p_2} d^2(\mathbf{y}_m, \mathbf{y}_c) + \frac{\gamma}{p_2} \sum_{n=1}^{N} b_{n,m} u_{n,c} \right\}}. \tag{A15}
$$

We have thus obtained (12).

Appendix B

In this appendix we show all figures and tables we did not fit into the main body of the paperin order not to make it too long (Figs. 6, 7, 8, Tables 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21).

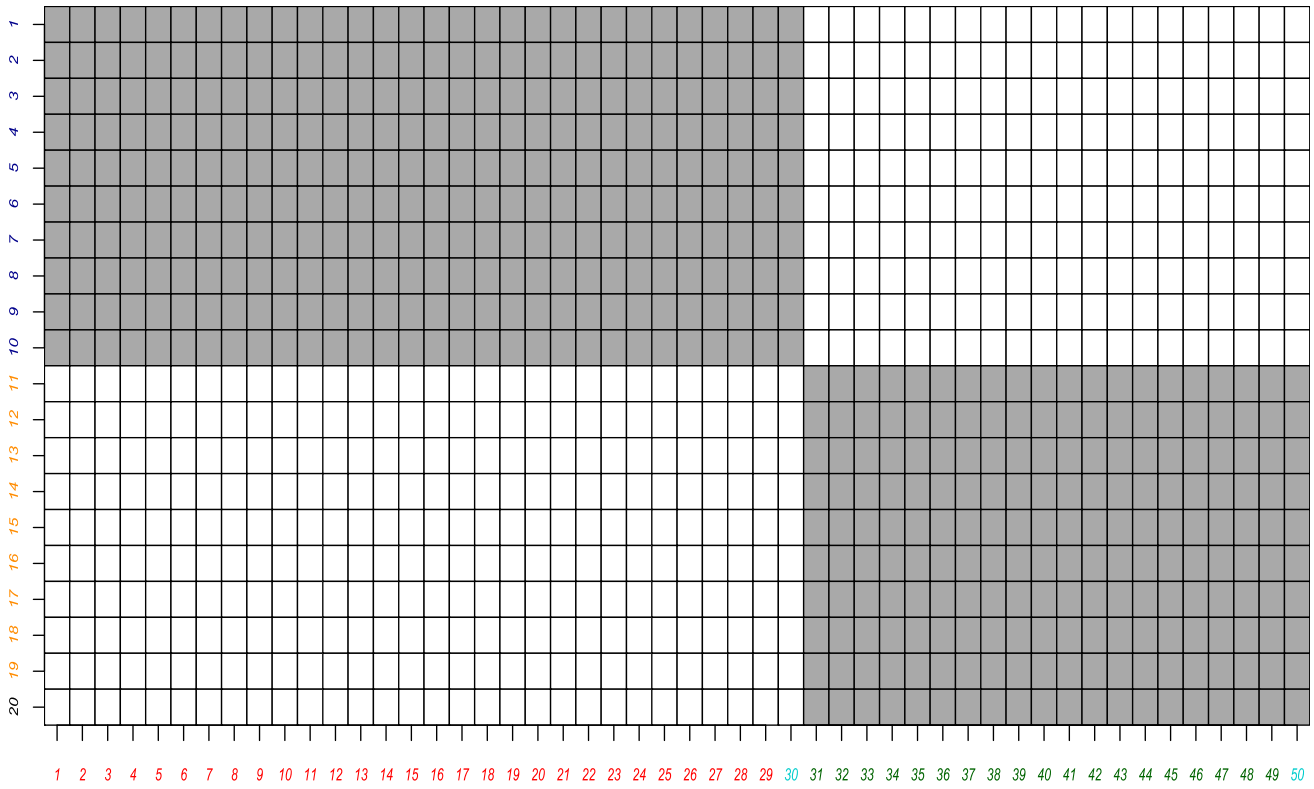

Fig. 6 First scenario: the associated bi-adiacency matrix. Units belonging to the same cluster based on their attributes are coloured accordingly. Fuzzy units are coloured cyan for the set $\mathcal{N}$ and black for the set $\mathcal{M}$

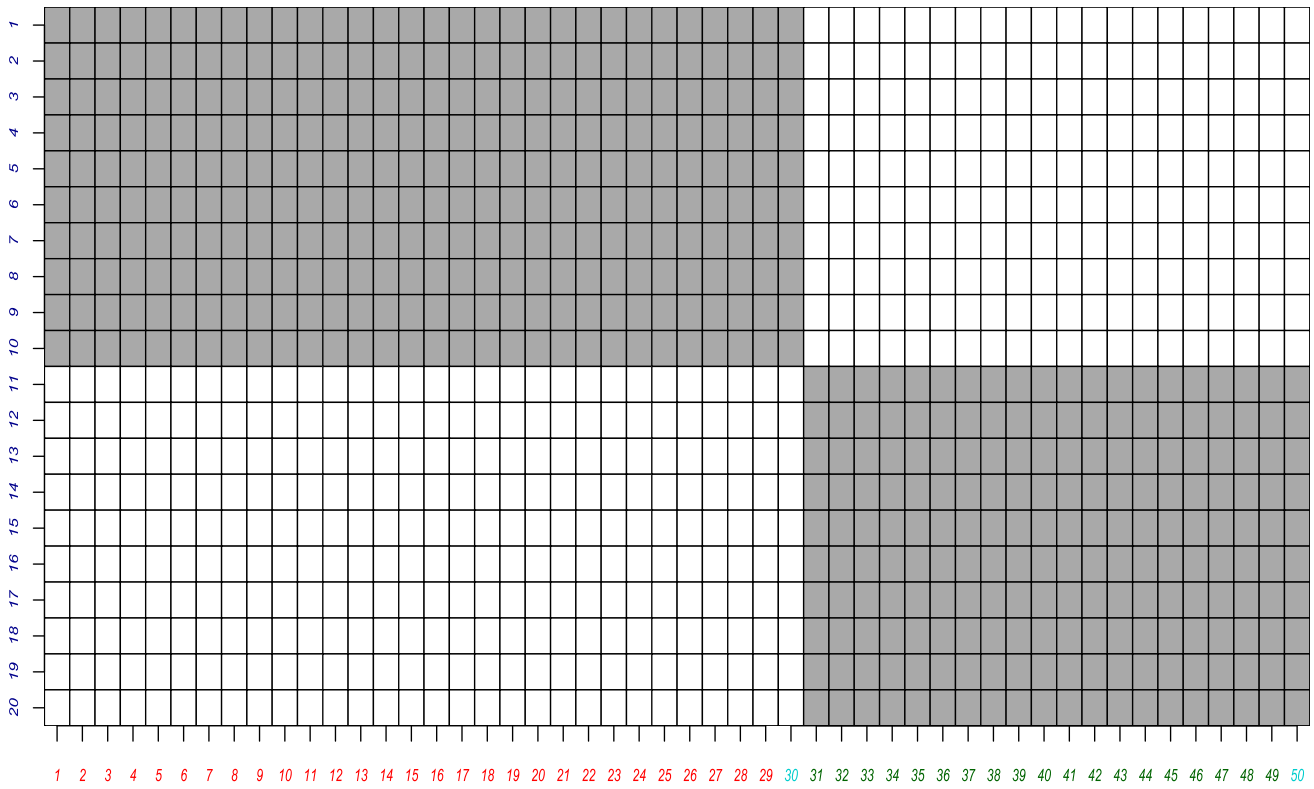

Fig. 7 Second scenario: the associated bi-adiacency matrix. Units belonging to the same cluster based on the attributes set are coloured accordingly. Fuzzy units only for the set $\mathcal{N}$ are coloured cyan

Table 7 Application with $p_1 = 3$, $p_2 = 1.5$: **V** membership degrees matrix for $\gamma \in \{0, 0.90, 2.25\}$

	$\gamma = 0$			$\gamma = 0.90$			$\gamma = 2.25$	
	C1	C2	C3	C1	C2	C3	C1	C2
1	0.53	0.02	0.46	0.25	0.01	0.74	0.05	0.95
2	0.52	0.02	0.46	0.41	0.01	0.58	0.71	0.29
3	0.47	0.02	0.52	0.48	0.09	0.42	0.03	0.97
4	0.42	0.01	0.56	0.55	0.01	0.44	0.16	0.84
5	0.51	0.02	0.47	0.48	0.02	0.50	0.20	0.80
6	0.55	0.02	0.43	0.11	0.12	0.77	0.00	1.00
7	0.48	0.02	0.50	0.84	0.01	0.15	1.00	0.00
8	0.53	0.02	0.45	0.27	0.02	0.71	0.19	0.81
9	0.61	0.02	0.38	0.02	0.00	0.98	0.02	0.98
10	0.62	0.01	0.37	0.05	0.02	0.92	0.02	0.98
11	0.51	0.02	0.47	0.14	0.01	0.86	0.06	0.94
12	0.52	0.02	0.46	0.44	0.01	0.54	0.19	0.81
13	0.48	0.02	0.50	0.73	0.01	0.26	0.82	0.18
14	0.52	0.02	0.47	0.60	0.01	0.39	0.99	0.01
15	0.54	0.02	0.44	0.12	0.14	0.75	0.00	1.00
16	0.44	0.02	0.55	0.34	0.01	0.66	0.04	0.96
17	0.46	0.02	0.52	0.78	0.01	0.21	0.32	0.68
18	0.48	0.02	0.50	0.92	0.00	0.08	1.00	0.00
19	0.48	0.02	0.51	0.18	0.01	0.81	0.03	0.97
20	0.57	0.02	0.41	0.34	0.01	0.65	0.87	0.13
21	0.42	0.02	0.57	0.96	0.00	0.04	0.97	0.03
22	0.49	0.02	0.49	0.72	0.01	0.27	0.97	0.03
23	0.50	0.02	0.49	0.81	0.00	0.18	0.95	0.05
24	0.49	0.02	0.49	0.50	0.03	0.46	0.69	0.31
25	0.51	0.02	0.47	0.84	0.00	0.15	1.00	0.00
26	0.51	0.02	0.48	0.17	0.01	0.82	0.02	0.98
27	0.51	0.02	0.47	0.07	0.00	0.93	0.02	0.98
28	0.51	0.02	0.48	0.71	0.01	0.28	0.94	0.06
29	0.46	0.02	0.53	0.64	0.02	0.34	0.31	0.69
30	0.50	0.02	0.49	0.49	0.01	0.50	0.41	0.59
31	0.61	0.02	0.37	0.11	0.01	0.88	0.06	0.94
32	0.62	0.02	0.37	0.05	0.01	0.95	0.06	0.94
33	0.54	0.02	0.44	0.04	0.00	0.96	0.01	0.99
34	0.43	0.02	0.55	0.92	0.00	0.08	0.97	0.03
35	0.62	0.03	0.34	0.01	0.00	0.99	0.00	1.00
36	0.47	0.03	0.49	0.44	0.03	0.53	0.27	0.73
37	0.47	0.03	0.50	0.67	0.03	0.30	0.64	0.36
38	0.49	0.04	0.48	0.03	0.01	0.97	0.00	1.00
39	0.51	0.03	0.46	0.40	0.04	0.56	0.35	0.65
40	0.47	0.03	0.50	0.68	0.03	0.29	0.64	0.36
41	0.70	0.03	0.27	0.00	0.00	1.00	0.00	1.00
42	0.42	0.03	0.55	0.93	0.01	0.06	0.96	0.04
43	0.44	0.04	0.52	0.77	0.03	0.20	0.31	0.69
44	0.47	0.04	0.49	0.65	0.03	0.32	0.85	0.15
45	0.51	0.03	0.46	0.51	0.03	0.47	0.08	0.92
46	0.41	0.03	0.56	0.77	0.03	0.21	0.49	0.51
47	0.52	0.03	0.45	0.44	0.03	0.53	0.17	0.83
48	0.71	0.03	0.25	0.00	0.00	1.00	0.00	1.00
49	0.56	0.03	0.41	0.11	0.02	0.87	0.12	0.88
50	0.53	0.03	0.44	0.19	0.02	0.78	0.17	0.83
51	0.44	0.03	0.52	0.91	0.01	0.08	1.00	0.00
52	0.51	0.04	0.45	0.41	0.04	0.56	0.18	0.82
53	0.49	0.03	0.48	0.82	0.05	0.14	1.00	0.00
54	0.49	0.03	0.47	0.44	0.03	0.54	0.07	0.93
55	0.48	0.03	0.49	0.41	0.03	0.56	0.28	0.72
56	0.52	0.04	0.45	0.22	0.03	0.75	0.17	0.83
57	0.76	0.03	0.21	0.00	0.00	1.00	0.00	1.00
58	0.49	0.03	0.48	0.23	0.02	0.75	0.02	0.98
59	0.55	0.04	0.41	0.13	0.03	0.84	0.10	0.90
60	0.49	0.04	0.48	0.55	0.05	0.40	0.82	0.18
61	0.42	0.04	0.54	0.84	0.02	0.14	0.37	0.63
62	0.41	0.03	0.55	0.95	0.01	0.04	0.96	0.04
63	0.49	0.03	0.48	0.76	0.02	0.23	0.34	0.66
64	0.47	0.06	0.47	0.73	0.03	0.24	0.89	0.11
65	0.50	0.03	0.46	0.38	0.07	0.55	0.18	0.82
66	0.41	0.03	0.56	0.97	0.00	0.02	1.00	0.00
67	0.54	0.03	0.42	0.08	0.01	0.91	0.01	0.99
68	0.50	0.03	0.47	0.16	0.02	0.82	0.02	0.98
69	0.50	0.03	0.46	0.30	0.03	0.67	0.21	0.79
70	0.45	0.03	0.51	0.74	0.03	0.23	0.28	0.72
71	0.43	0.03	0.54	0.36	0.02	0.63	0.05	0.95
72	0.46	0.03	0.50	0.98	0.00	0.02	1.00	0.00
73	0.45	0.06	0.49	0.62	0.06	0.32	0.25	0.75
74	0.51	0.03	0.46	0.02	0.00	0.98	0.00	1.00
75	0.50	0.03	0.47	0.36	0.04	0.61	0.19	0.81
76	0.53	0.04	0.44	0.36	0.03	0.61	0.13	0.87
77	0.54	0.03	0.43	0.14	0.02	0.84	0.13	0.87
78	0.41	0.03	0.56	0.98	0.00	0.02	0.99	0.01
79	0.46	0.04	0.50	0.36	0.04	0.60	0.09	0.91
80	0.57	0.05	0.38	0.01	0.01	0.98	0.01	0.99
81	0.41	0.03	0.56	0.89	0.02	0.09	0.42	0.58
82	0.47	0.04	0.49	0.63	0.03	0.34	0.23	0.77
83	0.48	0.04	0.48	0.92	0.01	0.08	1.00	0.00
84	0.48	0.03	0.48	0.62	0.02	0.36	0.83	0.17
85	0.61	0.04	0.35	0.01	0.00	0.99	0.00	1.00
86	0.49	0.04	0.47	0.26	0.04	0.70	0.06	0.94
87	0.45	0.08	0.48	0.22	0.05	0.73	0.08	0.92
88	0.46	0.08	0.45	0.82	0.04	0.14	1.00	0.00
89	0.44	0.09	0.47	0.87	0.02	0.11	1.00	0.00
90	0.42	0.08	0.49	0.70	0.07	0.23	0.68	0.32
91	0.46	0.09	0.45	0.09	0.05	0.86	0.01	0.99

Table 7 continued

	$\gamma = 0$			$\gamma = 0.90$			$\gamma = 2.25$	
	C1	C2	C3	C1	C2	C3	C1	C2
92	0.48	0.09	0.43	0.28	0.06	0.66	0.08	0.92
93	0.43	0.10	0.47	0.88	0.03	0.08	1.00	0.00
94	0.44	0.08	0.48	**1.00**	**0.00**	**0.00**	0.93	0.07
95	0.47	0.11	0.43	0.38	0.10	0.51	0.15	0.85
96	0.40	0.08	0.51	0.67	0.06	0.27	0.15	0.85
97	0.49	0.08	0.43	**0.00**	**0.00**	**1.00**	0.15	0.85
98	0.38	0.23	0.39	0.21	0.34	0.45	0.17	0.83
99	0.42	0.08	0.50	0.93	0.01	0.06	0.99	0.01
100	0.42	0.08	0.49	0.66	0.09	0.24	0.87	0.13
101	0.46	0.06	0.47	0.58	0.08	0.34	0.09	0.91
102	0.48	0.07	0.44	0.27	0.06	0.66	0.22	0.78
103	0.48	0.07	0.45	0.60	0.05	0.35	0.77	0.23
104	0.49	0.10	0.41	0.30	0.09	0.61	0.12	0.88
105	0.46	0.11	0.43	0.40	0.11	0.49	0.16	0.84
106	0.45	0.13	0.41	0.15	0.38	0.47	0.05	0.95
107	0.44	0.08	0.47	0.58	0.09	0.33	0.27	0.73
108	0.50	0.12	0.38	0.22	0.10	0.68	0.04	0.96
109	**1.00**	**0.00**	**0.00**	0.80	0.07	0.13	1.00	0.00
110	0.49	0.07	0.44	0.26	0.07	0.67	0.24	0.76
111	0.39	0.07	0.54	0.98	0.01	0.02	1.00	0.00
112	0.47	0.20	0.33	0.02	0.07	0.90	0.00	1.00
113	0.48	0.11	0.41	0.04	0.03	0.94	0.01	0.99
114	0.39	0.06	0.54	0.88	0.03	0.08	0.82	0.18
115	0.43	0.07	0.51	0.80	0.04	0.15	0.17	0.83
116	0.41	0.10	0.49	0.95	0.01	0.04	0.99	0.01
117	0.45	0.08	0.47	0.91	0.01	0.08	1.00	0.00
118	0.39	0.06	0.54	0.88	0.03	0.08	0.82	0.18
119	**0.00**	**0.00**	**1.00**	0.15	0.09	0.76	0.03	0.97
120	0.49	0.10	0.41	0.10	0.07	0.83	0.04	0.96
121	0.47	0.11	0.42	0.09	0.05	0.85	0.05	0.95
122	0.25	0.48	0.27	0.07	0.67	0.26	0.17	0.83
123	0.37	0.13	0.50	0.97	0.01	0.02	1.00	0.00
124	0.31	0.33	0.35	0.80	0.11	0.09	1.00	0.00
125	0.35	0.18	0.48	0.77	0.13	0.11	0.78	0.22
126	0.37	0.21	0.42	0.49	0.27	0.24	0.10	0.90
127	0.44	0.16	0.40	0.24	0.17	0.59	0.16	0.84
128	0.41	0.20	0.40	0.25	0.21	0.54	**0.00**	**1.00**
129	0.34	0.36	0.30	0.56	0.20	0.24	0.99	0.01
130	0.43	0.18	0.40	0.24	0.23	0.53	0.86	0.14
131	0.45	0.16	0.38	0.13	0.23	0.64	0.07	0.93
132	0.44	0.15	0.42	0.30	0.15	0.55	0.28	0.72
133	0.40	0.24	0.36	0.73	0.07	0.20	1.00	0.00
134	0.42	0.16	0.42	0.35	0.29	0.36	0.67	0.33
135	0.45	0.14	0.41	0.26	0.13	0.60	0.21	0.79
136	0.42	0.18	0.40	0.38	0.19	0.43	0.20	0.80
137	0.45	0.16	0.39	0.15	0.13	0.72	0.05	0.95

Table 7 continued

	$\gamma = 0$			$\gamma = 0.90$			$\gamma = 2.25$	
	C1	C2	C3	C1	C2	C3	C1	C2
138	0.37	0.18	0.45	0.94	0.03	0.04	1.00	0.00
139	0.33	0.24	0.43	0.60	0.25	0.15	0.32	0.68
140	0.45	0.15	0.40	0.24	0.19	0.57	0.20	0.80
141	0.33	0.38	0.29	0.03	0.26	0.70	0.01	0.99
142	0.37	0.12	0.51	0.85	0.07	0.08	0.82	0.18
143	0.46	0.18	0.36	0.11	0.14	0.75	0.12	0.88
144	0.27	0.38	0.34	0.84	0.10	0.06	1.00	0.00
145	0.37	0.27	0.36	0.02	0.05	0.93	0.00	1.00
146	0.37	0.20	0.43	0.90	0.04	0.05	**1.00**	**0.00**
147	0.41	0.15	0.44	0.58	0.16	0.27	0.24	0.76
148	0.41	0.19	0.40	0.41	0.19	0.40	0.54	0.46
149	0.37	0.12	0.51	0.85	0.07	0.08	0.82	0.18
150	0.32	0.21	0.47	0.62	0.22	0.16	0.55	0.45
151	0.33	0.31	0.36	0.19	0.55	0.26	0.19	0.81
152	0.46	0.16	0.38	0.06	0.10	0.84	0.02	0.98
153	0.21	0.63	0.16	0.01	0.46	0.52	0.02	0.98
154	0.37	0.15	0.48	0.72	0.15	0.13	0.91	0.09
155	0.42	0.19	0.38	0.39	0.18	0.44	0.58	0.42
156	0.38	0.15	0.47	0.62	0.17	0.21	0.36	0.64
157	0.38	0.17	0.45	0.96	0.01	0.02	1.00	0.00
158	0.37	0.12	0.51	0.85	0.07	0.08	0.82	0.18
159	0.41	0.28	0.31	0.41	0.16	0.43	0.97	0.03
160	0.29	0.40	0.30	0.05	0.15	0.80	0.00	1.00
161	0.32	0.23	0.45	0.79	0.13	0.07	0.82	0.18
162	0.27	0.45	0.28	0.08	0.27	0.65	0.01	0.99
163	0.34	0.42	0.24	0.11	0.35	0.54	0.03	0.97
164	0.35	0.30	0.35	0.09	0.75	0.16	0.02	0.98
165	0.32	0.23	0.45	0.79	0.13	0.07	0.82	0.18
166	0.33	0.28	0.39	0.77	0.13	0.10	0.93	0.07
167	0.23	0.49	0.28	0.30	0.55	0.15	0.30	0.70
168	0.38	0.28	0.34	0.72	0.09	0.19	0.97	0.03
169	0.19	0.59	0.21	0.39	0.49	0.12	0.98	0.02
170	0.27	0.49	0.24	0.47	0.31	0.22	0.98	0.02
171	0.22	0.54	0.24	0.36	0.50	0.15	0.88	0.12
172	0.38	0.28	0.34	0.82	0.07	0.11	0.99	0.01
173	0.20	0.52	0.28	0.78	0.18	0.04	0.98	0.02
174	0.32	0.24	0.44	0.94	0.04	0.02	0.99	0.01
175	0.38	0.28	0.34	0.09	0.58	0.33	0.09	0.91
176	0.27	0.48	0.25	0.05	0.79	0.16	0.04	0.96
177	0.31	0.29	0.40	0.61	0.22	0.17	0.67	0.33
178	0.06	0.84	0.09	0.04	0.92	0.04	0.25	0.75
179	0.32	0.23	0.45	0.92	0.05	0.03	0.97	0.03
180	0.22	0.52	0.25	0.51	0.38	0.11	0.98	0.02
181	0.36	0.25	0.39	0.90	0.05	0.05	0.99	0.01
182	0.33	0.30	0.37	0.84	0.09	0.07	0.99	0.01
183	0.29	0.38	0.33	0.65	0.29	0.06	1.00	0.00

Table 7 continued

	$\gamma = 0$			$\gamma = 0.90$			$\gamma = 2.25$	
	C1	C2	C3	C1	C2	C3	C1	C2
184	0.22	0.56	0.22	0.13	0.57	0.30	0.21	0.79
185	0.26	0.50	0.24	0.06	0.80	0.14	0.03	0.97
186	0.23	0.54	0.23	0.13	0.55	0.33	0.06	0.94
187	0.22	0.58	0.21	0.01	0.93	0.05	0.00	1.00
188	**0.00**	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**	0.02	0.98
189	0.01	0.98	0.01	0.01	0.98	0.01	0.97	0.03
190	0.23	0.54	0.23	0.08	0.44	0.48	0.02	0.98
191	0.13	0.70	0.17	0.09	0.79	0.12	0.12	0.88
192	0.20	0.56	0.24	0.82	0.15	0.03	0.99	0.01
193	0.13	0.73	0.14	0.34	0.58	0.08	0.99	0.01
194	0.11	0.77	0.12	0.18	0.74	0.08	0.90	0.10
195	0.07	0.86	0.07	0.00	1.00	0.00	0.00	1.00
196	0.10	0.79	0.10	0.15	0.78	0.07	0.99	0.01
197	0.26	0.50	0.24	0.12	0.57	0.31	0.19	0.81
198	0.22	0.56	0.22	0.15	0.65	0.20	0.22	0.78
199	0.23	0.54	0.23	0.01	0.96	0.03	0.00	1.00
200	0.24	0.43	0.33	0.77	0.19	0.05	0.93	0.07
201	0.23	0.54	0.23	0.13	0.64	0.22	0.23	0.77
202	0.22	0.59	0.19	0.00	0.97	0.02	0.00	1.00
203	0.08	0.78	0.13	0.31	0.66	0.03	0.92	0.08
204	0.26	0.52	0.23	0.09	0.70	0.21	0.14	0.86
205	0.21	0.50	0.29	0.69	0.26	0.05	0.92	0.08
206	0.26	0.39	0.35	0.87	0.10	0.03	0.97	0.03
207	0.10	0.79	0.12	0.05	0.89	0.06	0.28	0.72
208	0.07	0.87	0.06	0.01	0.96	0.03	0.02	0.98
209	0.06	0.87	0.08	0.00	1.00	0.00	0.00	1.00
210	0.00	1.00	0.00	0.00	1.00	0.00	0.01	0.99
211	0.02	0.97	0.02	0.00	1.00	0.00	0.00	1.00
212	0.02	0.97	0.02	0.00	1.00	0.00	0.01	0.99
213	0.03	0.93	0.04	0.02	0.97	0.01	0.33	0.67
214	0.10	0.79	0.11	0.11	0.80	0.08	0.20	0.80
215	0.14	0.68	0.18	0.23	0.71	0.06	0.14	0.86
216	0.09	0.78	0.13	0.06	0.92	0.02	0.20	0.80
217	0.07	0.86	0.07	0.07	0.89	0.05	0.97	0.03
218	0.12	0.75	0.13	0.03	0.95	0.01	1.00	0.00
219	0.12	0.72	0.15	0.08	0.86	0.06	0.19	0.81
220	0.18	0.61	0.21	0.22	0.64	0.15	0.28	0.72
221	0.11	0.78	0.11	0.02	0.96	0.01	1.00	0.00
222	0.21	0.59	0.20	0.55	0.33	0.12	0.92	0.08
223	0.18	0.68	0.15	0.00	0.97	0.03	0.00	1.00
224	0.13	0.75	0.12	0.00	0.98	0.02	0.00	1.00
225	0.12	0.75	0.13	0.07	0.83	0.10	0.25	0.75
226	0.11	0.79	0.10	0.04	0.75	0.22	0.12	0.88
227	0.09	0.83	0.08	0.01	0.94	0.05	0.07	0.93
228	0.03	0.94	0.03	0.01	0.97	0.03	0.13	0.87
229	0.13	0.76	0.11	0.02	0.91	0.07	0.04	0.96
230	0.04	0.92	0.04	0.01	0.99	0.00	1.00	0.00
231	0.04	0.93	0.04	0.01	0.99	0.00	1.00	0.00
232	0.05	0.88	0.07	0.03	0.94	0.02	0.19	0.81
233	0.09	0.81	0.10	0.15	0.81	0.04	0.30	0.70
234	0.06	0.87	0.06	0.00	0.99	0.01	0.01	0.99
235	0.07	0.84	0.09	0.04	0.92	0.04	0.18	0.82
236	0.09	0.81	0.10	0.05	0.93	0.02	1.00	0.00
237	0.03	0.93	0.04	0.02	0.97	0.02	0.29	0.71
238	0.01	0.99	0.01	0.00	1.00	0.00	1.00	0.00
239	0.03	0.95	0.03	0.01	0.99	0.00	1.00	0.00
240	0.05	0.90	0.05	0.03	0.93	0.04	0.28	0.72
241	0.02	0.95	0.03	0.02	0.97	0.01	0.38	0.62
242	0.03	0.95	0.03	0.01	0.99	0.00	1.00	0.00
243	0.02	0.95	0.03	0.01	0.99	0.00	1.00	0.00
244	0.01	0.98	0.01	0.00	1.00	0.00	1.00	0.00
245	0.03	0.93	0.04	0.03	0.96	0.01	0.39	0.61
246	0.02	0.96	0.02	0.00	0.99	0.01	0.05	0.95
247	0.02	0.96	0.02	0.00	1.00	0.00	0.02	0.98
248	0.04	0.91	0.05	0.05	0.93	0.02	0.32	0.68
249	0.03	0.93	0.04	0.06	0.93	0.01	0.91	0.09
250	0.00	1.00	0.00	0.00	1.00	0.00	0.36	0.64
251	0.02	0.95	0.03	0.02	0.97	0.01	0.36	0.64
252	0.01	0.98	0.01	0.00	1.00	0.00	0.02	0.98
253	0.02	0.96	0.02	0.04	0.95	0.01	0.84	0.16
254	0.00	0.99	0.00	0.00	1.00	0.00	0.04	0.96

The medoids are highlighted in bold

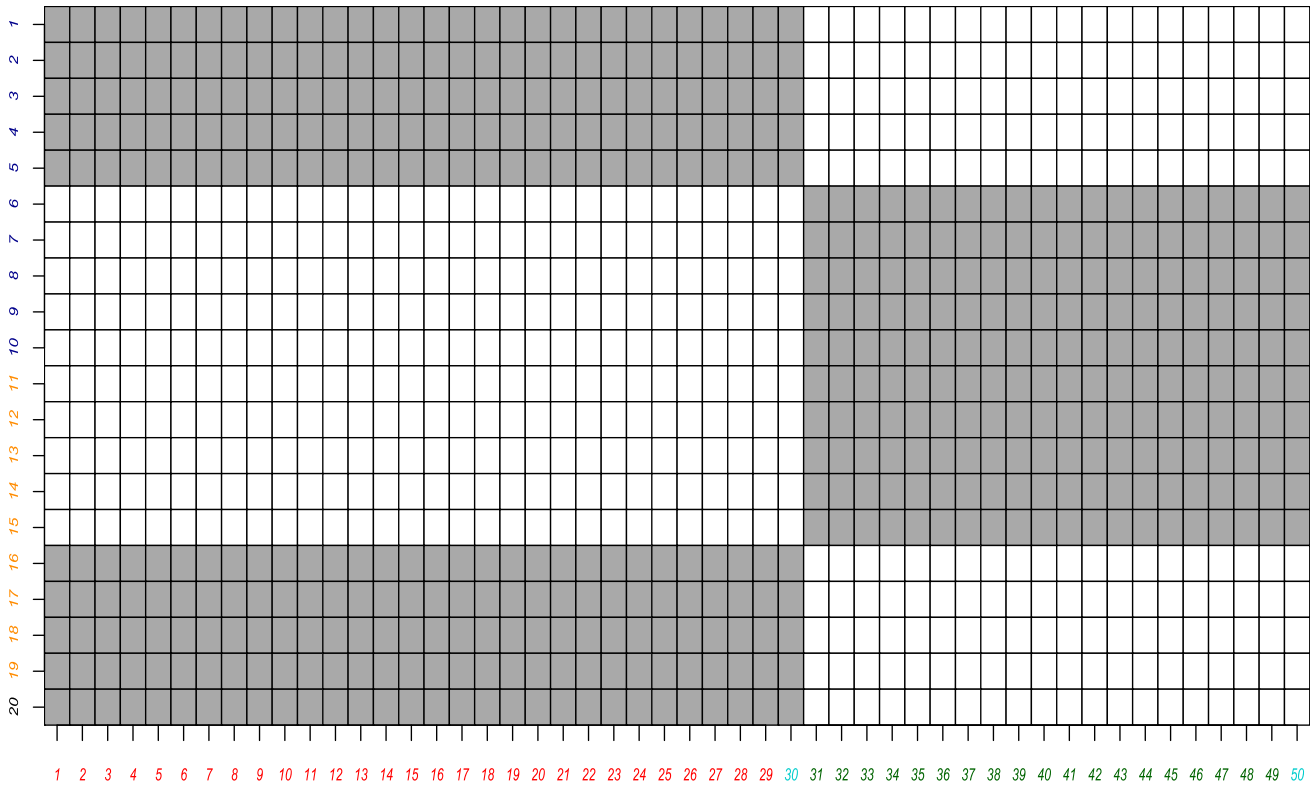

Fig. 8 Third scenario: the associated bi-adiacency matrix. Units belonging to the same cluster based on the attributes set are coloured accordingly. Fuzzy units are coloured cyan for the set $\mathcal{N}$ and black for the set $\mathcal{M}$

Table 8 First scenario with $p_1 = 1,\ p_2 = 1.5$: the proposed internal validity index for the 126 partitions (values are rounded to the nearest unit)

C	γ																			
	0.15	0.30	0.45	0.60	0.75	0.90	1.05	1.20	1.35	1.50	1.65	1.80	1.95	2.10	2.25	2.40	2.55	2.70	2.85	3.00
2	153	165	169	171	172	172	172	172	172	172	172	172	172	172	172	172	172	172	172	172
3	75	83	93	122	124	125	125	125	125	125	105	105	105	105	105	105	105	105	105	105
4	49	56	92	85	85	87	90	87	87	86	87	72	72	72	72	72	72	73	73	73
5	27	58	67	63	62	62	62	63	63	63	63	63	63	63	63	62	64	64	64	64
6	21	42	46	48	50	49	50	40	51	51	51	40	40	40	39	40	40	40	40	40

Table 9 First scenario with $p_1 = 1.5,\ p_2 = 2$: the proposed internal validity index for the 126 partitions (values are rounded to the nearest unit)

C	γ																			
	0.15	0.30	0.45	0.60	0.75	0.90	1.05	1.20	1.35	1.50	1.65	1.80	1.95	2.10	2.25	2.40	2.55	2.70	2.85	3.00
2	112	140	155	163	167	169	171	171	172	172	172	172	172	172	172	172	172	172	172	172
3	67	74	80	89	93	97	100	123	124	104	105	105	105	105	105	105	105	105	105	105
4	35	48	57	65	65	66	68	71	72	86	87	72	72	72	72	72	72	73	73	73
5	23	39	50	56	60	61	62	65	62	62	64	62	62	62	62	62	62	64	64	64
6	16	32	42	46	48	52	53	40	40	51	51	40	40	40	40	40	40	40	40	40

Table 10 First scenario with $p_1 = 1.5,\ p_2 = 2$: **U** membership degrees matrix for $\gamma \in \{0, 0.45, 1.35\}$

	$\gamma = 0$		$\gamma = 0.45$		$\gamma = 1.35$	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
1	0.15	0.85	0.01	0.99	0.00	1.00
2	0.11	0.89	0.01	0.99	0.00	1.00
3	0.20	0.80	0.02	0.98	0.00	1.00
4	0.11	0.89	0.01	0.99	0.00	1.00
5	0.13	0.87	0.01	0.99	0.00	1.00
6	0.29	0.71	0.04	0.96	0.00	1.00
7	0.12	0.88	0.01	0.99	0.00	1.00
8	0.04	0.96	0.00	1.00	0.00	1.00
9	0.05	0.95	0.00	1.00	0.00	1.00
10	0.04	0.96	0.00	1.00	0.00	1.00
11	0.06	0.94	0.00	1.00	0.00	1.00
12	0.05	0.95	0.00	1.00	0.00	1.00
13	0.05	0.95	0.00	1.00	0.00	1.00
14	0.05	0.95	0.00	1.00	0.00	1.00
15	0.12	0.88	0.01	0.99	0.00	1.00
16	0.16	0.84	0.02	0.98	0.00	1.00
17	0.17	0.83	0.02	0.98	0.00	1.00
18	0.16	0.84	0.02	0.98	0.00	1.00
19	0.20	0.80	0.02	0.98	0.00	1.00
20	0.36	0.64	0.06	0.94	0.00	1.00
21	0.24	0.76	0.03	0.97	0.00	1.00
22	0.07	0.93	0.01	0.99	0.00	1.00
23	0.10	0.90	0.01	0.99	0.00	1.00
24	0.10	0.90	0.01	0.99	0.00	1.00
25	0.12	0.88	0.01	0.99	0.00	1.00
26	0.26	0.74	0.03	0.97	0.00	1.00
27	0.06	0.94	0.00	1.00	0.00	1.00
28	0.05	0.95	0.00	1.00	0.00	1.00
29	**0.00**	**1.00**	**0.00**	**1.00**	**0.00**	**1.00**
30	*0.58*	*0.42*	*0.15*	*0.85*	*0.01*	*0.99*
31	0.93	0.07	0.99	0.01	1.00	0.00
32	0.99	0.01	1.00	0.00	1.00	0.00
33	0.94	0.06	0.99	0.01	1.00	0.00
34	0.95	0.05	1.00	0.00	1.00	0.00
35	0.96	0.04	1.00	0.00	1.00	0.00
36	0.97	0.03	1.00	0.00	1.00	0.00
37	0.98	0.02	1.00	0.00	1.00	0.00
38	0.98	0.02	1.00	0.00	1.00	0.00
39	0.97	0.03	1.00	0.00	1.00	0.00
40	0.96	0.04	1.00	0.00	1.00	0.00
41	0.97	0.03	1.00	0.00	1.00	0.00
42	0.98	0.02	1.00	0.00	1.00	0.00
43	0.94	0.06	**1.00**	**0.00**	**1.00**	**0.00**
44	0.95	0.05	1.00	0.00	1.00	0.00

Table 10 continued

	$\gamma = 0$		$\gamma = 0.45$		$\gamma = 1.35$	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
45	0.86	0.14	0.99	0.01	1.00	0.00
46	0.87	0.13	0.99	0.01	1.00	0.00
47	0.96	0.04	1.00	0.00	1.00	0.00
48	0.90	0.10	0.99	0.01	1.00	0.00
49	**1.00**	**0.00**	0.99	0.01	1.00	0.00
50	*0.65*	*0.35*	*0.94*	*0.06*	*1.00*	*0.00*

The medoids are highlighted in bold, while the fuzzy units are in italics

Table 11 First scenario with $p_1 = 1.5,\ p_2 = 2$: $\mathbf{W}$ membership degrees matrix for $\gamma \in \{0, 0.45, 1.35\}$

	$\gamma = 0$		$\gamma = 0.45$		$\gamma = 1.35$	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
1	0.96	0.04	**0.00**	**1.00**	**0.00**	**1.00**
2	0.98	0.02	0.00	1.00	0.00	1.00
3	0.92	0.08	0.00	1.00	0.00	1.00
4	**1.00**	**0.00**	0.00	1.00	0.00	1.00
5	0.93	0.07	0.00	1.00	0.00	1.00
6	0.99	0.01	0.00	1.00	0.00	1.00
7	0.96	0.04	0.00	1.00	0.00	1.00
8	0.95	0.05	0.00	1.00	0.00	1.00
9	0.97	0.03	0.00	1.00	0.00	1.00
10	0.97	0.03	0.00	1.00	0.00	1.00
11	0.02	0.98	1.00	0.00	1.00	0.00
12	0.02	0.98	1.00	0.00	1.00	0.00
13	0.02	0.98	1.00	0.00	1.00	0.00
14	0.10	0.90	1.00	0.00	1.00	0.00
15	0.04	0.96	1.00	0.00	1.00	0.00
16	0.03	0.97	1.00	0.00	1.00	0.00
17	**0.00**	**1.00**	**1.00**	**0.00**	**1.00**	**0.00**
18	0.02	0.98	1.00	0.00	1.00	0.00
19	0.01	0.99	1.00	0.00	1.00	0.00
20	*0.52*	*0.48*	*0.98*	*0.02*	*1.00*	*0.00*

The medoids are highlighted in bold, while the fuzzy units are in italics

Table 12 First scenario with $p_1 = 1,\ p_2 = 1.5$: **U** membership degrees matrix for $\gamma \in \{0, 0.30, 0.75\}$

	$\gamma = 0$		$\gamma = 0.30$		$\gamma = 0.75$	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
1	0.02	0.98	0.00	1.00	0.00	1.00
2	0.01	0.99	0.00	1.00	0.00	1.00
3	0.04	0.96	0.01	0.99	0.00	1.00
4	0.01	0.99	0.00	1.00	0.00	1.00
5	0.02	0.98	0.00	1.00	0.00	1.00
6	0.09	0.91	0.02	0.98	0.00	1.00
7	0.01	0.99	0.00	1.00	0.00	1.00
8	0.00	1.00	0.00	1.00	0.00	1.00
9	0.00	1.00	0.00	1.00	0.00	1.00
10	0.00	1.00	0.00	1.00	0.00	1.00
11	0.00	1.00	0.00	1.00	0.00	1.00
12	0.00	1.00	0.00	1.00	0.00	1.00
13	0.00	1.00	0.00	1.00	0.00	1.00
14	0.00	1.00	0.00	1.00	0.00	1.00
15	0.01	0.99	0.00	1.00	0.00	1.00
16	0.02	0.98	0.00	1.00	0.00	1.00
17	0.03	0.97	0.01	0.99	0.00	1.00
18	0.02	0.98	0.00	1.00	0.00	1.00
19	0.04	0.96	0.01	0.99	0.00	1.00
20	0.15	0.85	0.03	0.97	0.00	1.00
21	0.06	0.94	0.01	0.99	0.00	1.00
22	0.00	1.00	0.00	1.00	0.00	1.00
23	0.01	0.99	0.00	1.00	0.00	1.00
24	0.01	0.99	0.00	1.00	0.00	1.00
25	0.02	0.98	0.00	1.00	0.00	1.00
26	0.07	0.93	0.02	0.98	0.00	1.00
27	0.00	1.00	0.00	1.00	0.00	1.00
28	0.00	1.00	0.00	1.00	0.00	1.00
29	**0.00**	**1.00**	**0.00**	**1.00**	**0.00**	**1.00**
30	*0.48*	*0.52*	*0.15*	*0.85*	*0.01*	*0.99*
31	0.98	0.02	1.00	0.00	1.00	0.00
32	1.00	0.00	1.00	0.00	1.00	0.00
33	0.98	0.02	1.00	0.00	1.00	0.00
34	0.99	0.01	1.00	0.00	1.00	0.00
35	0.99	0.01	1.00	0.00	1.00	0.00
36	1.00	0.00	1.00	0.00	1.00	0.00
37	1.00	0.00	1.00	0.00	1.00	0.00
38	1.00	0.00	1.00	0.00	1.00	0.00
39	1.00	0.00	1.00	0.00	1.00	0.00
40	0.99	0.01	1.00	0.00	1.00	0.00
41	1.00	0.00	1.00	0.00	1.00	0.00
42	1.00	0.00	1.00	0.00	1.00	0.00

Table 12 continued

	$\gamma = 0$		$\gamma = 0.30$		$\gamma = 0.75$	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
43	**1.00**	**0.00**	**1.00**	**0.00**	**1.00**	**0.00**
44	0.99	0.01	1.00	0.00	1.00	0.00
45	0.93	0.07	0.99	0.01	1.00	0.00
46	0.94	0.06	0.99	0.01	1.00	0.00
47	0.99	0.01	1.00	0.00	1.00	0.00
48	0.96	0.04	1.00	0.00	1.00	0.00
49	0.94	0.06	0.99	0.01	1.00	0.00
50	*0.60*	*0.40*	*0.95*	*0.05*	*1.00*	*0.00*

The medoids are highlighted in bold, while the fuzzy units are in italics

Table 13 First scenario with $p_1 = 1,\ p_2 = 1.5$: **W** membership degrees matrix for $\gamma \in \{0, 0.30, 0.75\}$

	$\gamma = 0$		$\gamma = 0.30$		$\gamma = 0.75$	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
1	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**	**1.00**
2	1.00	0.00	0.00	1.00	0.00	1.00
3	0.99	0.01	0.00	1.00	0.00	1.00
4	0.99	0.01	0.00	1.00	0.00	1.00
5	0.99	0.01	0.00	1.00	0.00	1.00
6	1.00	0.00	0.00	1.00	0.00	1.00
7	0.99	0.01	0.00	1.00	0.00	1.00
8	0.99	0.01	0.00	1.00	0.00	1.00
9	1.00	0.00	0.00	1.00	0.00	1.00
10	1.00	0.00	0.00	1.00	0.00	1.00
11	0.00	1.00	1.00	0.00	1.00	0.00
12	0.00	1.00	1.00	0.00	1.00	0.00
13	0.00	1.00	1.00	0.00	1.00	0.00
14	0.04	0.96	1.00	0.00	1.00	0.00
15	0.01	0.99	1.00	0.00	1.00	0.00
16	**0.00**	**1.00**	1.00	0.00	1.00	0.00
17	0.01	0.99	**1.00**	**0.00**	**1.00**	**0.00**
18	0.00	1.00	1.00	0.00	1.00	0.00
19	0.00	1.00	1.00	0.00	1.00	0.00
20	*0.59*	*0.41*	*0.97*	*0.03*	*1.00*	*0.00*

The medoids are highlighted in bold, while the fuzzy units are in italics

Table 14 Second scenario with $p_1 = 1,\ p_2 = 1.5$: the proposed internal validity index for the 126 partitions (values are rounded to the nearest unit)

C	γ																			
	0.15	0.30	0.45	0.60	0.75	0.90	1.05	1.20	1.35	1.50	1.65	1.80	1.95	2.10	2.25	2.40	2.55	2.70	2.85	3.00
2	16	25	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
3	13	16	17	17	17	17	17	18	18	17	18	18	17	17	17	18	18	17	17	18
4	10	12	12	13	13	13	13	13	13	12	12	12	13	13	12	12	12	12	12	12
5	6	10	10	10	10	10	10	10	10	10	10	10	10	9	9	9	9	9	9	9
6	5	9	8	8	8	9	8	8	8	8	8	7	8	8	7	7	8	7	7	7

Table 15 Second scenario with $p_1 = 1.5,\ p_2 = 2$: the proposed internal validity index for the 126 partitions (values are rounded to the nearest unit)

C	γ																			
	0.15	0.30	0.45	0.60	0.75	0.90	1.05	1.20	1.35	1.50	1.65	1.80	1.95	2.10	2.25	2.40	2.55	2.70	2.85	3.00
2	10	21	25	26	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
3	11	14	15	16	17	17	17	17	17	17	17	18	17	17	18	18	17	17	17	17
4	9	11	11	12	12	13	13	13	13	13	12	13	13	12	13	12	13	12	12	13
5	5	7	10	10	10	9	10	10	10	10	10	10	10	10	9	9	9	9	9	9
6	4	8	9	8	8	8	9	8	8	8	8	8	8	8	8	8	8	7	7	8

Table 16 Second scenario with $p_1 = 1,\ p_2 = 1.5$: **U** membership degrees matrix for $\gamma \in \{0, 0.15, 0.45\}$

	$\gamma = 0$		$\gamma = 0.15$		$\gamma = 0.45$	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
1	0.98	0.02	0.99	0.01	1.00	0.00
2	0.99	0.01	0.99	0.01	1.00	0.00
3	0.96	0.04	0.97	0.03	1.00	0.00
4	0.99	0.01	0.99	0.01	1.00	0.00
5	0.98	0.02	0.99	0.01	1.00	0.00
6	0.91	0.09	0.94	0.06	0.99	0.01
7	0.99	0.01	0.99	0.01	1.00	0.00
8	1.00	0.00	1.00	0.00	1.00	0.00
9	1.00	0.00	1.00	0.00	1.00	0.00
10	1.00	0.00	1.00	0.00	1.00	0.00
11	1.00	0.00	1.00	0.00	1.00	0.00
12	1.00	0.00	1.00	0.00	1.00	0.00
13	1.00	0.00	1.00	0.00	1.00	0.00
14	1.00	0.00	1.00	0.00	1.00	0.00
15	0.99	0.01	0.99	0.01	1.00	0.00
16	0.98	0.02	0.98	0.02	1.00	0.00
17	0.97	0.03	0.98	0.02	1.00	0.00
18	0.98	0.02	0.98	0.02	1.00	0.00
19	0.96	0.04	0.97	0.03	1.00	0.00
20	0.85	0.15	0.89	0.11	0.99	0.01
21	0.94	0.06	0.96	0.04	0.99	0.01
22	1.00	0.00	1.00	0.00	1.00	0.00
23	0.99	0.01	0.99	0.01	1.00	0.00
24	0.99	0.01	0.99	0.01	1.00	0.00
25	0.98	0.02	0.99	0.01	1.00	0.00
26	0.93	0.07	0.95	0.05	0.99	0.01
27	1.00	0.00	1.00	0.00	1.00	0.00

Table 16 continued

	$\gamma = 0$		$\gamma = 0.15$		$\gamma = 0.45$	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
28	1.00	0.00	1.00	0.00	1.00	0.00
29	**1.00**	**0.00**	**1.00**	**0.00**	**1.00**	**0.00**
30	*0.52*	*0.48*	*0.61*	*0.39*	*0.93*	*0.07*
31	0.02	0.98	0.01	0.99	0.00	1.00
32	0.00	1.00	0.00	1.00	0.00	1.00
33	0.02	0.98	0.01	0.99	0.00	1.00
34	0.01	0.99	0.01	0.99	0.00	1.00
35	0.01	0.99	0.00	1.00	0.00	1.00
36	0.00	1.00	0.00	1.00	0.00	1.00
37	0.00	1.00	0.00	1.00	0.00	1.00
38	0.00	1.00	0.00	1.00	0.00	1.00
39	0.00	1.00	0.00	1.00	0.00	1.00
40	0.01	0.99	0.00	1.00	0.00	1.00
41	0.00	1.00	0.00	1.00	0.00	1.00
42	0.00	1.00	0.00	1.00	0.00	1.00
43	**0.00**	**1.00**	**0.00**	**1.00**	**0.00**	**1.00**
44	0.01	0.99	0.00	1.00	0.00	1.00
45	0.07	0.93	0.04	0.96	0.00	1.00
46	0.06	0.94	0.04	0.96	0.00	1.00
47	0.01	0.99	0.00	1.00	0.00	1.00
48	0.04	0.96	0.02	0.98	0.00	1.00
49	0.06	0.94	0.03	0.97	0.00	1.00
50	*0.40*	*0.60*	*0.27*	*0.73*	*0.02*	*0.98*

The medoids are highlighted in bold, while the fuzzy units are in italics

Table 17 Second scenario with $p_1 = 1,\ p_2 = 1.5$: **W** membership degrees matrix for $\gamma \in \{0, 0.15, 0.45\}$

	$\gamma = 0$		$\gamma = 0.15$		$\gamma = 0.45$	
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
1	0.14	0.86	0.43	0.57	**1.00**	**0.00**
2	0.94	0.06	0.99	0.01	0.99	0.01
3	0.88	0.12	0.97	0.03	0.99	0.01
4	0.07	0.93	0.26	0.74	0.99	0.01
5	0.11	0.89	0.38	0.62	0.99	0.01
6	0.01	0.99	0.04	0.96	1.00	0.00
7	0.97	0.03	0.99	0.01	0.99	0.01
8	0.70	0.30	0.92	0.08	1.00	0.00
9	0.93	0.07	0.98	0.02	0.99	0.01
10	0.08	0.92	0.29	0.71	1.00	0.00
11	0.50	0.50	0.17	0.83	0.00	1.00
12	0.23	0.77	0.06	0.94	0.07	0.93
13	0.69	0.31	0.32	0.68	0.00	1.00
14	0.04	0.96	0.01	0.99	0.02	0.98
15	**1.00**	**0.00**	**1.00**	**0.00**	0.01	0.99
16	0.62	0.38	0.25	0.75	0.00	1.00
17	0.03	0.97	0.01	0.99	0.00	1.00
18	**0.00**	**1.00**	**0.00**	**1.00**	0.02	0.98
19	0.55	0.45	0.20	0.80	0.01	0.99
20	0.24	0.76	0.06	0.94	**0.00**	**1.00**

The medoids are highlighted in bold, while the fuzzy units are in italics

Table 18 Third scenario with $p_1 = 1,\ p_2 = 1.5$: the proposed internal validity index for the 126 partitions (values are rounded to the nearest unit)

C	γ																			
	0.15	0.30	0.45	0.60	0.75	0.90	1.05	1.20	1.35	1.50	1.65	1.80	1.95	2.10	2.25	2.40	2.55	2.70	2.85	3.00
2	22	26	29	29	29	29	30	30	30	30	30	30	30	30	30	30	30	30	30	30
3	36	47	30	30	30	30	20	20	20	20	20	20	20	20	19	20	20	20	20	20
4	23	39	42	43	43	16	14	14	14	14	14	14	14	14	14	14	14	14	13	15
5	20	25	29	32	32	16	11	10	11	10	10	10	10	10	10	10	10	10	10	10
6	13	23	28	29	22	11	8	9	8	9	8	8	8	8	8	8	8	7	7	7

Table 19 Third scenario with $p_1 = 1.5,\ p_2 = 2$: the proposed internal validity index for the 126 partitions (values are rounded to the nearest unit)

C	γ																			
	0.15	0.30	0.45	0.60	0.75	0.90	1.05	1.20	1.35	1.50	1.65	1.80	1.95	2.10	2.25	2.40	2.55	2.70	2.85	3.00
2	11	22	27	29	29	29	29	29	29	29	30	30	30	30	30	30	30	30	30	30
3	22	33	29	30	29	30	30	20	20	20	20	20	20	20	20	19	19	20	20	20
4	18	32	38	40	40	38	20	14	14	14	14	14	14	15	14	14	14	14	14	14
5	15	21	26	27	27	25	16	11	10	11	11	10	10	10	10	10	10	10	11	10
6	5	18	21	19	24	8	7	8	8	8	8	8	8	8	8	8	8	8	8	8

Table 20 Third scenario with $p_1 = 1,\ p_2 = 1.5$: **U** membership degrees matrix for $\gamma \in \{0, 0.15, 0.30\}$

	$\gamma = 0$		$\gamma = 0.15$			$\gamma = 0.30$		
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3
1	0.98	0.02	0.42	0.58	0.00	0.35	0.65	0.00
2	0.99	0.01	0.46	0.54	0.00	0.38	0.62	0.00
3	0.96	0.04	0.40	0.60	0.00	0.35	0.65	0.00
4	0.99	0.01	0.32	0.68	0.00	0.18	0.82	0.00
5	0.98	0.02	0.46	0.54	0.00	0.40	0.60	0.00
6	0.91	0.09	0.27	0.72	0.01	0.21	0.78	0.00
7	0.99	0.01	0.33	0.67	0.00	0.20	0.80	0.00
8	1.00	0.00	0.47	0.53	0.00	0.28	0.72	0.00
9	1.00	0.00	0.46	0.54	0.00	0.28	0.72	0.00
10	1.00	0.00	0.46	0.54	0.00	0.26	0.74	0.00
11	1.00	0.00	0.48	0.52	0.00	0.32	0.68	0.00
12	1.00	0.00	0.48	0.52	0.00	0.30	0.70	0.00
13	1.00	0.00	0.49	0.51	0.00	0.32	0.68	0.00
14	1.00	0.00	0.45	0.55	0.00	0.28	0.72	0.00
15	0.99	0.01	0.31	0.69	0.00	0.18	0.81	0.00
16	0.98	0.02	0.46	0.54	0.00	0.42	0.58	0.00
17	0.97	0.03	0.42	0.58	0.00	0.36	0.64	0.00
18	0.98	0.02	0.43	0.57	0.00	0.37	0.63	0.00
19	0.96	0.04	0.34	0.66	0.00	0.26	0.74	0.00
20	0.85	0.15	0.26	0.71	0.03	0.23	0.76	0.01
21	0.94	0.06	0.28	0.72	0.00	0.20	0.80	0.00
22	1.00	0.00	**1.00**	**0.00**	**0.00**	0.33	0.67	0.00
23	0.99	0.01	0.46	0.54	0.00	**1.00**	**0.00**	**0.00**
24	0.99	0.01	0.48	0.52	0.00	0.39	0.61	0.00
25	0.98	0.02	0.32	0.68	0.00	0.19	0.81	0.00
26	0.93	0.07	0.28	0.72	0.01	0.21	0.79	0.00
27	1.00	0.00	0.37	0.63	0.00	0.20	0.80	0.00
28	1.00	0.00	0.48	0.52	0.00	0.31	0.69	0.00
29	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**
30	*0.52*	*0.48*	*0.15*	*0.57*	*0.28*	*0.18*	*0.68*	*0.14*
31	0.02	0.98	0.00	0.00	1.00	0.00	0.00	1.00
32	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
33	0.02	0.98	0.00	0.00	1.00	0.00	0.00	1.00
34	0.01	0.99	0.00	0.00	1.00	0.00	0.00	1.00
35	0.01	0.99	0.00	0.00	1.00	0.00	0.00	1.00
36	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
37	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
38	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
39	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
40	0.01	0.99	0.00	0.00	1.00	0.00	0.00	1.00
41	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
42	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
43	**0.00**	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**	**0.00**	**1.00**

Table 20 continued

	$\gamma = 0$		$\gamma = 0.15$			$\gamma = 0.30$		
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3
44	0.01	0.99	0.00	0.00	1.00	0.00	0.00	1.00
45	0.07	0.93	0.00	0.01	0.99	0.00	0.00	1.00
46	0.06	0.94	0.00	0.01	0.99	0.00	0.00	1.00
47	0.01	0.99	0.00	0.00	1.00	0.00	0.00	1.00
48	0.04	0.96	0.00	0.00	1.00	0.00	0.00	1.00
49	0.06	0.94	0.00	0.01	0.99	0.00	0.00	1.00
50	*0.40*	*0.60*	*0.09*	*0.16*	*0.75*	*0.11*	*0.05*	*0.84*

The medoids are highlighted in bold, while the fuzzy units are in italics

Table 21 Third scenario with $p_1 = 1$, $p_2 = 1.5$: $\mathbf{W}$ membership degrees matrix for $\gamma \in \{0, 0.15, 0.30\}$

	$\gamma = 0$		$\gamma = 0.15$			$\gamma = 0.30$		
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3
1	**0.00**	**1.00**	0.00	0.88	0.12	0.00	0.99	0.01
2	0.00	1.00	0.00	0.85	0.15	0.00	0.98	0.02
3	0.01	0.99	0.00	0.91	0.09	0.00	0.99	0.01
4	0.01	0.99	0.00	0.88	0.12	0.00	0.99	0.01
5	0.01	0.99	**0.00**	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**
6	0.00	1.00	0.00	0.08	0.92	0.00	0.01	0.99
7	0.01	0.99	0.00	0.11	0.89	0.00	0.01	0.99
8	0.01	0.99	0.00	0.11	0.89	0.00	0.01	0.99
9	0.00	1.00	**0.00**	**0.00**	**1.00**	**0.00**	**0.00**	**1.00**
10	0.00	1.00	0.00	0.10	0.90	0.00	0.01	0.99
11	1.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
12	1.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
13	1.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
14	0.96	0.04	0.96	0.02	0.02	0.90	0.01	0.10
15	0.99	0.01	1.00	0.00	0.00	0.99	0.00	0.01
16	**1.00**	**0.00**	**1.00**	**0.00**	**0.00**	**1.00**	**0.00**	**0.00**
17	0.99	0.01	0.99	0.01	0.00	0.99	0.01	0.00
18	1.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
19	1.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
20	*0.41*	*0.59*	*0.18*	*0.78*	*0.04*	*0.09*	*0.91*	*0.00*

The medoids are highlighted in bold, while the fuzzy units are in italics

Funding Open access funding provided by Luiss University within the CRUI-CARE Agreement. The research of Livia De Giovanni and Lorenzo Federico is supported by the European Media and Information Fund through the grant Leveraging argument technology for impartial fact-checking (LATIF, Grant Agreement Number 268755).

Availability of data and materials Data sets and code for the implementation of the algorithm are available at https://github.com/LopiJ90/Fuzzy-Clustering-with-Barber-modularity-regularization/tree/main.

Declarations

Conflict of interest The authors have no conflict of interest to declare.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

Airoldi, E.M., Blei, D.M., Fienberg, S.E., et al.: Mixed membership stochastic blockmodels. Journal of Machine Learning Research **9**, 1981–2014 (2008)

Barber, M.J.: Modularity and community detection in bipartite networks. Phys Rev E **76**, 066102 (2007). https://doi.org/10.1103/PhysRevE.76.066102. https://link.aps.org/doi/10.1103/PhysRevE.76.066102

Bellman, R., Kalaba, R., Zadeh, L.: Abstraction and pattern classification. Journal of Math Anal and Appl **13**(1), 1–7 (1966)

Bezdek, J.C.: Numerical taxonomy with fuzzy sets. Journal of Math Biology **1**, 57–71 (1974). https://doi.org/10.1007/BF02339490

Bezdek, J.C.: Objective function clustering. In: Pattern recognition with fuzzy objective function algorithms. Springer, p 43–93 (1981)

Blondel, V.D., Guillaume, J.L., Lambiotte, R., et al.: Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment **2008**(10), P10008 (2008)

Brandes, U., Delling, D., Gaertler, M., et al.: On modularity clustering. IEEE transactions on knowledge and data engineering **20**(2), 172–188 (2007)

Clauset, A., Newman, M.E.J., Moore, C.: Finding community structure in very large networks. Phys Rev E **70**, 066111 (2004)

Coppi, R., D'Urso, P.: Fuzzy unsupervised classification of multivariate time trajectories with the shannon entropy regularization. Comput Stat Data Anal **50**(6), 1452–1477 (2006)

Dunn, J.C.: A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. Journal of Cybernetics **3**(3), 32–57 (1973)

D'Urso, P., Vitale, V.: A kemeny distance-based robust fuzzy clustering for preference data. Journal of Classification **39**(3), 600–647 (2022)

D'Urso, P., De Giovanni, L., Alaimo, L.S., et al.: Fuzzy clustering with entropy regularization for interval-valued data with an application to scientific journal citations. Annals of Operations Research pp 1–24 (2023a)

D'Urso, P., De Giovanni, L., Federico, L., et al.: Fuzzy clustering of spatial interval-valued data. Spatial Statistics **57**, 100764 (2023)

D'Urso, P., De Giovanni, L., Vitale, V.: Robust dtw-based entropy fuzzy clustering of time series. Annals of Operations Research pp 1–35 (2023c)

D'Urso, P., De Luca, G., Vitale, V., et al.: Tail dependence-based fuzzy clustering of financial time series. Annals of Operations Research pp 1–27 (2023d)

Erosheva, E.A., Fienberg, S.E., Lafferty, J.D.: Mixed membership models of scientific publications. Proceedings of the National Academy of Sciences **101**(Suppl 1), 5220–5227 (2004). https://doi.org/10.1073/pnas.0307757101

Gao, Y., Wang, D., Pan, J., et al.: A novel fuzzy c-means clustering algorithm using adaptive norm. International Journal of Fuzzy Systems **21**(8), 2632–2649 (2019)

Heller, K.A., Williamson, S., Ghahramani, Z.: Statistical models for partial membership. In: Proceedings of the 25th International Conference on Machine Learning (ICML '08), ACM. ACM, New York, NY, USA, pp 392–399 (2008)

van der Hofstad, R., Komjáthy, J., Vadon, V.: Phase transition in random intersection graphs with communities. Random Structures & Algorithms **60**(3), 406–461 (2022)

Ichihashi, H.: Gaussian mixture pdf approximation and fuzzy c-means clustering with entropy regularization. In: Proc. 4th Asian Fuzzy Systems Symposium, 2000, pp 217–221 (2000)

Kahali, S., Sing, J.K., Saha, P.K.: A new entropy-based approach for fuzzy c-means clustering and its application to brain mr image segmentation. Soft Computing **23**(20), 10407–10414 (2019)

Krishnapuram, R., Joshi, A., Yi, L.: A fuzzy relative of the k-medoids algorithm with application to web document and snippet clustering. In: International Fuzzy Systems Conference (FUZZIEEE99), Seoul, IEEE, pp 1281–1286 (1999)

Krishnapuram, R., Joshi, A., Nasraoui, O., et al.: Low-complexity fuzzy relational clustering algorithms for web mining. IEEE transactions on Fuzzy Systems **9**(4), 595–607 (2001)

Li, R., Mukaidono, M.: A maximum entropy approach to fuzzy clustering. In: Proceedings of the Fourth IEEE Conference on Fuzzy Systems (FUZZ-IEEE/IFES '95), pp 2227—-2232 (1995)

Li, R.P., Mukaidono, M.: Gaussian clustering method based on maximum-fuzzy-entropy interpretation. Fuzzy Sets and systems **102**(2), 253–258 (1999)

Luczak, T.: Sparse random graphs with a given degree sequence. In: Random graphs, pp 165–182 (1992)

Miyamoto, S., Mukaidono, M.: Fuzzy c-means as a regularization and maximum entropy approach. In: Proc. of 7th International Fuzzy Systems Association World Congress (IFSA'97), II, pp 86–92 (1997)

Nepusz, T., Petróczi, A., Négyessy, L., et al.: Fuzzy communities and the concept of bridgeness in complex networks. Physical Review E **77**(1), 016107 (2008)

Newman, M.E.: Modularity and community structure in networks. Proceedings of the national academy of sciences **103**(23), 8577–8582 (2006)

Norros, I., Reittu, H.: On a conditionally Poissonian graph process. Adv in Appl Probab **38**(1), 59–75 (2006). https://doi.org/10.1239/aap/1143936140. https://0-doi-org.pugwash.lib.warwick.ac.uk/10.1239/aap/1143936140

de Ruiter, S.: Report: Why the european commission must choose the nutri-score nutrition label. https://unescochair-ghe.org/2023/05/12/report-why-the-european-commission-must-choose-the-nutri-score-nutrition-label/, accessed: (11/3/2024) (2023)

Ruspini, E.: A new approach to clustering. Information and control **16**, 22–32 (1969)

Ruspini, E.: Numerical methods of fuzzy clustering. Information Science **2**, 319–350 (1970)

Ruspini, E.: New experimental results in fuzzy clustering. Information Science **6**, 273–284 (1973)

Traag, V.A., Waltman, L., Van Eck, N.J.: From louvain to leiden: guaranteeing well-connected communities. Scientific reports **9**(1), 5233 (2019)

Yao, J., Dash, M., Tan, S., et al.: Entropy-based fuzzy clustering and fuzzy modeling. Fuzzy sets and Systems **113**(3), 381–388 (2000)

Zadeh, L.A.: Information and control. Fuzzy sets **8**(3), 338–353 (1965)

Zarinbal, M., Zarandi, M.F., Turksen, I.: Relative entropy fuzzy c-means clustering. Information Sciences **260**, 74–97 (2014)