Partition problems for fuzzy graphs

Daniel Gómez
Complutense University of Madrid
dagomez@estad.ucm.es

Javier Montero and Javier Yáñez
Complutense University of Madrid
$\{monty, jayage\}$@mat.ucm.es

Abstract

The main contribution of this paper is to obtain a polynomial algorithm that allows us to classify step by step a set of items that are related by means of a fuzzy graph extending the algorithm published in [8] to any class of graph.

Keywords: Partition problems, fuzzy graph.

1 Introduction

Classification could be defined in a very simple way as a procedure in which individual items are placed into groups based on its quantitative or qualitative inherent information. Classification problems are one of the most statistical studied problems in literature and there exists a huge amount of papers dedicated to its applications in the field of decision theory, statistics, economics, biology, geology, industry, etc. In contrast with the huge amount of classification papers few algorithms can found that deals with the problem of classification when the objects are related in some way.

In the case in which the relation between targets can be modelled as a graph, the previous problem could be addressed as a *partition problem* in graphs. In partition problems, a partition of the graph is founded attending to some good properties. During the last three decades, many researchers have developed models, built algorithms, and implemented solutions for districting-partition problems. Such problems can be viewed as a grouping process of elementary units or atoms of a given territory into larger pieces of land or zones, thus giving rise to a partition, also called a district map. There are many practical questions/applications related to districting problems (see [13] for more details).

The classification-coloring algorithm proposed in this paper is an extension of the coloring algorithm proposed in [7] but for general graphs. In the field of image classification and with the main aim to obtain a segmentation of a digital image or a remote sensing image, in [7], a coloring algorithm for valued graphs was developed. This algorithm was developed to classify pixels in an image and to find homogeneous regions for a further supervised classification. Later, in [8], the authors extended this coloring algorithm to the case in which the graph was fuzzy and planar. This new algorithm permitted the modelization of the digital image in a more realistic way since the fuzzyness appears naturally when we work with real problems and images. It is important to emphasize that the two algorithms previously developed for coloring and classifying images are only valid for a specific class of graphs (see [7, 8]) and had an specific application in the field of classification images.

The main contribution of this paper, is to develop a classification algorithm in which it is supposed that the target or objets that have to be classified are not independent and the relation among this items are given by a graph. In order to do that we extend the ideas of the coloring algorithms presented in [7, 8] to any class of graph.

2 The model: The relations among items and its associated fuzzy graph

We will denote by P the set of items and by k the number of characteristic or variables associated to each $p \in P$. We will identify by p_i for $i \in \{1, \ldots, k\}$, the i-characteristic or variable of item p. Let $G = (V = P, E)$ be the graph that shows the relations among items. Now, mathematically, the problem that we want to solve is to classified the set P when the a priori information that we have is the characteristic associated to the objects and the graph that show the relations among them.

One of the first steps in any classification system is to define for any pair of items a distance or dissimilitude measure. Once the relations among targets-items have been modeled, we shall consider a fuzzy distance to express the dissimilitude between the measured properties of items, $d : P \times P \longrightarrow \widetilde{[0, \infty)}$, where $\widetilde{[0, \infty)}$ will be here the set of fuzzy numbers with domain in R^+ ([2] see also [11] and [14]). We will denote by $\widetilde{d_{pp'}} = d(p, p')$ the fuzzy distance between the items p and p'. We will denote its membership function by $\mu_{pp'} : R^+ \to [0, 1]$ and by $\widetilde{D} = \{\widetilde{d_{pp'}} \ / \ (p, p') \in P \times P\}$ we will denote its associated fuzzy distance matrix.

Definition. *Given the classification problem previously defined, and given a fuzzy distance function d, the* item fuzzy graph *can be defined as the pair $\widetilde{G(P)} = (P, \tilde{E})$, when the fuzzy value for the edge associated to two adjacent items p and p' is the fuzzy distance between them, i.e. $\widetilde{E_{p,p'}} = \widetilde{d_{pp'}}$.*

3 The coloring-segmentation partition problem for graphs

The classification that we propose in this work can be viewed as a dendrogram where initially all items are in the same group. In the first step this group (that is modeled as a fuzzy graph) is colored in two color classes (0-1). This coloring process will establish a first segmentation of the graph taking the connected components of the subgraph defined by the item that have been colored as 0 and by 1 respectively. This process is iteratively repeated until (as happen with the classical dendrogram) all items are grouped individually. So, we propose to successively apply a basic binary coloring process, leading to a hierarchical coloring of the image.

As can be viewed in this coloring process the key point is how we establish in each iteration the binary coloring for the fuzzy graph and the election of the α value.

Once the binary coloring process is defined and how to determine the value of alpha in each iteration, the binary coloring procedure can be successively applied to each family of items belonging to the same color class, meanwhile there are adjacent items being different. It is important to note that the color classes are not necessary connected. We will define the term of homogeneous region or group in the graph as a set of connected items with the same color class. This homogeneous regions will be the final partition or classification of the items in the graph.

Definition. *Given a set P with associate fuzzy graph $\widetilde{G(P)}$ and let C a coloring of the fuzzy graph, a connected component of each color class induced by C will be said a* region.

3.1 The basic binary coloring procedure

The basic binary coloring we propose will assign values "0" and "1" to each pair of adjacent items depending on the fuzzy distance between their measured descriptions, when compared to a prescribed threshold α. So, we can obtain the first binary coloring assigning an arbitrary color ("0" or "1") to an arbitrary item, and fixing the order in which items will be colored. Once a initial item has been colored, the remainder of the items that are in the same connected component of the fuzzy graph $\widetilde{G(P)} = (P, \widetilde{(E)})$ are colored following this rule. If two items p and q are adjacent then

$$col(q) = \begin{cases} col(p) & if\ \widetilde{E_{p,q}} \tilde{<} \alpha \\ 1 - col(p) & if\ \widetilde{E(p,q)} \tilde{\geq} \alpha \end{cases}$$

for all (p, q) *adjacent in* P.

It is important to note that this binary coloring process is not always consistent. Inconsistent situations are therefore present when an item can be colored by two different ways, *i.e.* when there is a cycle in the graph that allows us to color depending on the path that is chosen.

If the graph is acyclic, we can obtain a consistent coloring by choosing, randomly, any initial item from every connected component, and assigning to each one of these items either color 0 or color 1, arbitrarily. In this case, once a value α has been fixed, coloring of each adjacent item is unique. Otherwise, if the connected graph is not acyclic (this is the case for most of the real situations) this coloring rule could produce some problems.

In general, given an item already colored and a fixed valued of α, it could be exist a cycle that produces an *inconsistent* coloring, i.e. for a given item on that cycle, you can used two different colors depending on the path that you use.

In order to deal with these inconsistencies, in [7] the authors defined an arbitrary spanning tree for valued graphs. Once a spanning tree $T(G)$ of the graph has been defined, we can produce a binary coloring of the valued-fuzzy graph without inconsistencies. Now the question is to decide for a given valued of α what is the best spanning tree of the graph for our proposes (a related problem with this can be seen in [9]). Taking into account that our idea is to classify the items of P in homogeneous groups we are going to choose the spanning tree of fuzzy graph $\widetilde{G(P)}$ as follows:

- **1 Step**. We build the fuzzy graph $\widetilde{G(P)}^* = (P, \tilde{E}^*)$, when the fuzzy value for the edge associated to two adjacent items p and q is $\widetilde{E^*_{p,q}} = \widetilde{E_{p,q}}$ if $\widetilde{E_{p,q}} \tilde{\leq} \alpha$ and $-\widetilde{E_{p,q}}$ otherwise.
- **2 Step**. We obtain the minimum spanning tree of the fuzzy graph $\widetilde{G(P)}^*$. Let us denote by $\widetilde{MST}$ the minimum spanning tree maintaining the original values of the edges, i.e. $(\widetilde{MST}, E)$ instead of $(\widetilde{MST}, E^*)$.

3.2 Alpha value election

It is clear that this process is strongly dependent on the α value selection (α_i) in each iteration. Of course, this selection is not a trivial task. If all selected values are close to the minimum dissimilitude measure, we shall be able to identify only regions being very different. If we consider all values close to the maximum of this measures, we will be sensitive to small variations and a non informative segmentation will be given with too many regions. In order to discriminate in terms of similar size sets of pixels at each step, a possibility is to choose those α values taking into account quartiles of the distance distribution. In this way we can impose certain equilibrium in coloration, of course subject to the final goal of segmentation.

3.3 Ranking fuzzy numbers

From the previous algorithm we can see that we tackle sometimes with the problem of ranking fuzzy sets or fuzzy numbers. The problem of ordering fuzzy numbers has been studied by many authors (see, e.g., [3, 4]) or [14]). An interesting approach is to transform fuzzy numbers into real numbers by means of a ranking function (see [3]).

In this work we will say that a fuzzy set $\widetilde{A}$ with membership function μ_A and domain in R^+ is considered as a fuzzy number in the sense of Kerre (see [14]).

Definition. *Let* $\aleph$ *be the set of fuzzy numbers and let* $a, b \in \aleph$. *Then* $a \tilde{\geq} b \longleftrightarrow F(a) \geq F(b)$.

4 Final comments

The main contribution of this paper is to obtain a polynomial algorithm that allows us to classify step by step (by means of a dendogram) a set of items that are related by means of a graph. This problem was partially addressed in [8] for a special class of graphs and in this paper it has been extended to any class of fuzzy graph. In this sense, we can also extend the applications of this classification algorithm to any classification problem in which the relation among items are modeled by means of a graph. In [7, 8], it can be see the application to the field of image processing and remote sensing classification.

Acknowledgements

This Research has been partially supported by grant TIN2006-06190, from the Government of Spain.

References

[1] J.C. Bezdek, J.D. Harris, Fuzzy partitions and relations: an axiomatic basis for clustering, Fuzzy Sets and Systems 1 (1978) 111-127.

[2] D. Dubois, H. Prade, Fuzzy sets and Systems, Theory and Applications. Academic Press, Orlando, Florida (1997).

[3] D. Dubois, H. Prade, Ranking fuzzy numbers in the setting of possibility theory, Information Sciences 30 (1983) 183-224.

[4] G. Facchinetti, R.G. Ricci, A characterization of a general class of ranking functions on triangular fuzzy numbers, Fuzzy Sets and Systems 146 (2004) 297-312.

[5] D. Feng, S. Wenkang, C. Liangzhou, D. Yong, Z. Zhenfu, Infrared image segmentation with 2-D maximum entropy method based on particle swarm optimization, Pattern Recognition Letters 26 (2005) 597-603.

[6] G.M. Foody, The continuum of classification fuzziness in thematics mapping, Photogrammetric Engineering and Remote Sensing 65 (1999) 443-451.

[7] D. Gómez, J. Montero, J. Yáñez, C. Poidomani (2007): A graph coloring algorithm approach for image segmentation. *Omega* 35:173–183.

[8] D. Gómez, J. Montero and J. Yánez (2006): A coloring algorithm for image classification. *Information Sciences 176, 3645-3657.*

[9] J. Yez and J. Montero (2008) : Analyzing the inconsistencies for the coloring problem in fuzzy graphs. *Proocedings of the FLINS conference 2008.*

[10] O. Kabva, S. Seikkala, On fuzzy metric spaces, Fuzzy Sets and Systems 12 (1984) 215-229.

[11] A. Kaufman, M. Gupta, Introduction to Fuzzy Arithmetic. Van Nostrand Reinhold, Mew York (1985).

[12] S. Muñoz, M.T. Ortuño, J. Ramirez, J. Yáñez, Coloring fuzzy graphs, Omega 33 (2005) 211-221.

[13] F. Tavares-Pereira, JR Figueira, V. Mousseau, B. Roy, Multiple criteria districting problems - The public transportation network pricing system of the Paris region, Annals of Operations Research 154 (2007) 69-92.

[14] X. Wang, E.E. Kerre, Reasonable properties for the ordering of fuzzy quantities (I), Fuzzy Sets and Systems 118(3) (2001) 375-385.

[15] L.A. Zadeh, Fuzzy sets, Information and Control 8 (1965) 338-353.