Knowledge-Based Systems 53 (2013) 20–26

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

A spectral clustering algorithm based on intuitionistic fuzzy information

Dawei Xu, Zeshui Xu *, Shousheng Liu, Hua Zhao

College of Sciences, PLA University of Science and Technology, Nanjing 211101, China

ARTICLE INFO

Article history:
Received 9 December 2012
Received in revised form 3 June 2013
Accepted 26 July 2013
Available online 25 August 2013

Keywords:
Intuitionistic fuzzy set
Intuitionistic fuzzy similarity degree
Graph laplacians
Spectral clustering
Algorithm

ABSTRACT

In this paper, we propose a technique for clustering data with intuitionistic fuzzy information. We first define two new intuitionistic fuzzy similarity measures, and then use it to construct an intuitionistic fuzzy similarity measure matrix, by which we present a spectral algorithm to cluster intuitionistic fuzzy information. At last, two numerical examples are given to illustrate and verify our algorithm.

1. Introduction

In 1983, Atanassov [1] introduced the concept of intuitionistic fuzzy set (IFS), a generalization of fuzzy set, which was originally introduced by Zadeh [2]. Since its appearance, IFS has received more and more attentions and has been applied to a variety of fields, such as decision making [3–6], pattern recognition [7–9], market prediction [10], etc. Since the late 1980s, data mining has gradually become a hot research field. As a sub-field of data mining, clustering analysis plays a very important role, which can gathers the similar samples into the same class, while gathers the dissimilar ones into different classes. Up to now, clustering analysis has been widely applied to various fields [11,12]. Recently, some authors have investigated the clustering techniques for intuitionistic fuzzy information [13–16]. Zhang et al. [13] firstly constructed an intuitionistic fuzzy similarity matrix, then computed the intuitionistic fuzzy equivalence matrix by the transitive closure of the intuitionistic fuzzy similarity matrix, and then they clustered the samples by the λ-cutting matrix of the interval-valued matrix. Xu et al. [14] gave a new clustering method that can transform an association matrix into an equivalent association matrix, and then they utilized the λ-cutting matrix to cluster the given IFSs. Xu et al. [15] proposed the intuitionistic fuzzy C-means (IFCM) algorithms to cluster intuitionistic fuzzy information, and then extended this method to interval-valued intuitionistic fuzzy environment. Wang et al. [16] extended the netting clustering method to intuitionistic fuzzy environment, and achieved some good results. Although the intuitionistic fuzzy clustering techniques described above have their own distinctive features, they are based on either the intuitionistic fuzzy equivalence matrices or the transitive closure technique so that they need a large amount of computation efforts, or brings the data distortion, or even may fall into the local optimal solutions.

Recently, scholars have proposed a new technique called spectral clustering algorithm. It has obvious characteristics that it is relatively simple to implement and uneasy to fall into local optimal solutions by eigen-analysis. Yet, it is worthy of pointing out that this kind of clustering technique can only cluster the real-number data. To overcome the shortcomings of the above existing intuitionistic fuzzy clustering techniques, in this paper we shall extend the spectral clustering technique to intuitionistic fuzzy environments. To do that, the rest of the paper is organized as follows. Section 2 reviews the concept of IFSs and introduces the intuitionistic fuzzy similarity measure matric. Section 3 introduces two kinds of intuitionistic fuzzy similarity measures. Section 4 reviews the normal Spectral Clustering method and Section 5 extends it to intuitionistic fuzzy environment. In Section 6 we use a numerical example to show the effectiveness and the relatively better results of the developed algorithm, comparing it with some other existing intuitionistic fuzzy clustering techniques. Section 7 concludes the paper.

2. Preliminaries

Firstly, we shall give some basic concepts that may lead the readers to have a better understanding of intuitionistic fuzzy sets [1].

Definition 1 [1]. Let a set X be fixed, then an intuitionistic fuzzy set (IFS) A on X is defined as $A = \{\langle x, \mu_A(x), \upsilon_A(x)\rangle | x \in X\}$, where the functions

* Corresponding author. Tel.: +86 25 84483382.
E-mail address: xuzeshui@263.net (Z. Xu).

0950-7051/$ - see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.knosys.2013.07.020

$$\mu_A : X \to [0,1], \quad x \in X \to \mu_A(x) \in [0,1] \tag{1}$$

and

$$v_A : X \to [0,1], \quad x \in X \to v_A(x) \in [0,1] \tag{2}$$

denote the membership and non-membership degrees of the element $x \in X$ to A respectively, with

$$\mu_A(x) + v_A(x) \leqslant 1, \quad \text{for any } x \in X \tag{3}$$

Furthermore, the function $\pi_A(x) = 1 - \mu_A(x) - v_A(x)$ is called the uncertainty (or hesitation) degree of x to A. Especially, if $\pi_A(x) = 0$, then A reduces to a fuzzy set [2]. For convenience, Xu and Yager [17,18] called $\alpha = (\mu_\alpha, v_\alpha)$ an intuitionistic fuzzy number (IFN), where

$$\mu_\alpha \in [0,1], \quad v_\alpha \in [0,1], \quad \mu_\alpha + v_\alpha \leqslant 1 \tag{4}$$

Definition 2 [13]. Let $Z = (z_{ij})_{m\times n}$ be a $m \times n$ matrix, if all z_{ij} ($i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$) are IFNs, then Z is called an intuitionistic fuzzy matrix.

Definition 3 [13]. Let X be a non-empty set, $\Phi(X)$ be the set of all the intuitionistic fuzzy sets on X. Assume that $\vartheta: (\Phi(X))^2 \to [0, 1]$ is a mapping, and $A_j \in \Phi(X)$ ($j = 1, 2$), then $\vartheta(A_1, A_2)$ is called the intuitionistic fuzzy similarity measure of A_1 and A_2, if it satisfies the following conditions:

(1) $0 \leqslant \vartheta(A_1, A_2) \leqslant 1$;
(2) $\vartheta(A_1, A_2) = 1$ if and only if $A_1 = A_2$;
(3) $\vartheta(A_1, A_2) = \vartheta(A_2, A_1)$;
(4) If $A_1 \subseteq A_2 \subseteq A_3$, then $\vartheta(A_1, A_3) \leqslant \vartheta(A_1, A_2)$ and $\vartheta(A_1, A_3) \leqslant \vartheta(A_2, A_3)$.

Based on the intuitionistic fuzzy similarity measure defined above, we shall give the concept of intuitionistic fuzzy similarity measure matrix as below:

Definition 4. Let A_i ($i = 1, 2, \ldots, n$) be a collection of IFSs, $s_{ij} = \vartheta(A_i, A_j)$, $i,j = 1, 2, \ldots, n$ be the intuitionistic fuzzy similarity measure of A_i and A_j, then $S = (s_{ij})_{n\times n}$ is called an intuitionistic fuzzy similarity measure matrix, which has the following properties:

(1) (Reflexivity). $s_{ii} = 1$, $i = 1, 2, \ldots, n$;
(2) (Symmetry). $s_{ij} = s_{ji}$.

3. Two new similarity measures of intuitionistic fuzzy sets

As we know, some of the pre-presented clustering methods are based on the intuitionistic fuzzy similarity measure, whose values are IFNs and cannot easily be dealt with in the spectral clustering method. Besides, the existing intuitionistic fuzzy similarity measures are somewhat complex and may need a lot of computations [19–23]. For the preparation of extending the normal spectral clustering method to intuitionistic fuzzy environment, in the following, we shall present two new ways of constructing intuitionistic fuzzy similarity measures whose values are real numbers.

Fuzzy Hamming closeness degree is a nice way to define the similarity measure, based on which, we try to construct two intuitionistic fuzzy similarity measures:

Definition 5. let A and B be two intuitionistic fuzzy sets, then we define the following function:

$$\vartheta(A,B) = \min\left\{\left(1 - \frac{1}{n}\sum_{i=1}^{n}|\mu_A(x_i) - \mu_B(x_i)|\right),\ \left(1 - \frac{1}{n}\sum_{i=1}^{n}|v_A(x_i) - v_B(x_i)|\right)\right\} \tag{5}$$

Theorem 1. *The function $\vartheta(A, B)$ given in Definition 5 is an intuitionistic fuzzy similarity measure.*

Proof.

(1) Obviously, $0 \leqslant \vartheta(A, B) \leqslant 1$.
(2) If $\vartheta(A, B) = 1$, that is,

$$1 - \frac{1}{n}\sum_{i=1}^{n}|\mu_A(x_i) - \mu_B(x_i)| = 1 - \frac{1}{n}\sum_{i=1}^{n}|v_A(x_i) - v_B(x_i)| = 1 \tag{6}$$

Then

$$\mu_A(x_i) = \mu_B(x_i), \quad v_A(x_i) = v_B(x_i), \quad i = 1, 2, \ldots, n \tag{7}$$

Thus $A = B$.
If $A = B$, then we can obtain:

$$1 - \frac{1}{n}\sum_{i=1}^{n}|\mu_A(x_i) - \mu_B(x_i)| = 1 - \frac{1}{n}\sum_{i=1}^{n}|v_A(x_i) - v_B(x_i)| = 1 \tag{8}$$

So $\vartheta(A, B) = 1$.
From the above analysis we know that $\vartheta(A_1, A_2) = 1$ if and only if $A_1 = A_2$.
(3)

$$\begin{aligned}\vartheta(A,B) &= \min\left\{\left(1 - \frac{1}{n}\sum_{i=1}^{n}|\mu_A(x_i) - \mu_B(x_i)|\right), \left(1 - \frac{1}{n}\sum_{i=1}^{n}|v_A(x_i) - v_B(x_i)|\right)\right\} \\ &= \min\left\{\left(1 - \frac{1}{n}\sum_{i=1}^{n}|\mu_B(x_i) - \mu_A(x_i)|\right), \left(1 - \frac{1}{n}\sum_{i=1}^{n}|v_B(x_i) - v_A(x_i)|\right)\right\} \\ &= \vartheta(B,A)\end{aligned} \tag{9}$$

(4) Suppose that there exists an intuitionistic fuzzy set $C \in IFS(U)$ satisfying $A \subseteq C \subseteq B$, if $1 - \frac{1}{n}\sum_{i=1}^{n}|\mu_A(x_i) - \mu_B(x_i)| \leqslant 1 - \frac{1}{n}\sum_{i=1}^{n}|v_A(x_i) - v_B(x_i)|$, then we have

$$\begin{aligned}\vartheta(A,B) &= \min\left\{\left(1 - \frac{1}{n}\sum_{i=1}^{n}|\mu_A(x_i) - \mu_B(x_i)|\right), \left(1 - \frac{1}{n}\sum_{i=1}^{n}|v_A(x_i) - v_B(x_i)|\right)\right\} \\ &= 1 - \frac{1}{n}\sum_{i=1}^{n}|\mu_A(x_i) - \mu_B(x_i)| \leqslant 1 - \frac{1}{n}\sum_{i=1}^{n}|\mu_A(x_i) - \mu_C(x_i)|\end{aligned} \tag{10}$$

While if $1 - \frac{1}{n}\sum_{i=1}^{n}|\mu_A(x_i) - \mu_B(x_i)| \geqslant 1 - \frac{1}{n}\sum_{i=1}^{n}|v_A(x_i) - v_B(x_i)|$, then we can get

$$\vartheta(A,B) \leqslant 1 - \frac{1}{n}\sum_{i=1}^{n}|v_A(x_i) - v_C(x_i)| \tag{11}$$

Thus,

$$\begin{aligned}\vartheta(A,B) &\leqslant \min\left\{\left(1 - \frac{1}{n}\sum_{i=1}^{n}|\mu_A(x_i) - \mu_C(x_i)|\right), \left(1 - \frac{1}{n}\sum_{i=1}^{n}|v_A(x_i) - v_C(x_i)|\right)\right\} \\ &= \vartheta(A,C)\end{aligned} \tag{12}$$

Similarly, we can prove that $\vartheta(A, B) \leqslant \vartheta(B, C)$. □