

# Thuật toán phân cụm mờ khả thi C-means

## A Possibilistic Fuzzy c-Means Clustering Algorithm

Nikhil R. Pal, Kuhu Pal, James M. Keller, and James C. Bezdek

**Tóm tắt** - Năm 1997, chúng tôi đã đề xuất mô hình C-means mờ khả dĩ (FPCM) và thuật toán tạo ra cả giá trị thành viên và giá trị điển hình khi phân cụm dữ liệu không có nhãn. PFCM giới hạn các giá trị phụ thuộc sao cho tổng giá trị phụ thuộc trên tất cả các điểm dữ liệu bằng 1. Tuy nhiên, ràng buộc tổng hàng trở lên không thực tế đối với các tập dữ liệu lớn. Trong bài báo này, chúng tôi đề xuất một mô hình mới có tên là mô hình phân cụm mờ khả dĩ C-means. PFCM tạo ra đồng thời các thành viên và khả năng, cùng với các điểm nguyên mẫu thông thường hoặc tâm cụm cho mỗi cụm. PFCM là sự kết hợp của C-means khả dĩ (PCM) và C-means mờ (FCM) thường tránh được nhiều vấn đề khác nhau của PCM, FCM, và FPCM. PFCM xử lý được các hạn chế về độ nhiễu của FCM, khắc phục được vấn đề cụm trùng nhau của PCM và loại bỏ các ràng buộc tổng hàng của FPCM. Chúng tôi đưa ra các điều kiện cần thiết nhất cho cực trị của hàm mục tiêu PFCM. Một số ví dụ số được đưa ra để so sánh FCM và PCM với PFCM. Các ví dụ của chúng tôi cho thấy PFCM được so sánh là thuận lợi hơn so với cả 2 mô hình trước đó. Vì các mô hình nguyên mẫu ít nhạy cảm hơn so với các giá trị ngoại lai và có thể tránh các cụm trùng nhau. PFCM là ứng cử viên mạnh cho việc nhận dạng hệ thống dựa trên quy tắc mờ.

**Thuật ngữ chỉ mục**- *c-means, cụm mờ, cụm lai, cụm khả thi.*

**Abstract**—In 1997, we proposed the fuzzy-possibilistic c-means (FPCM) model and algorithm that generated both membership and typicality values when clustering unlabeled data. FPCM constrains the typicality values so that the sum over all data points of typicalities to a cluster is one. The row sum constraint produces unrealistic typicality values for large data sets. In this paper, we propose a new model called possibilistic-fuzzy c-means (PFCM) model. PFCM produces memberships and possibilities simultaneously, along with the usual point prototypes or cluster centers for each cluster. PFCM is a hybridization of possibilistic c-means (PCM) and fuzzy c-means (FCM) that often avoids various problems of PCM, FCM and FPCM. PFCM solves the noise sensitivity defect of FCM, overcomes the coincident clusters problem of PCM and eliminates the row sum constraints of FPCM. We derive the first-order necessary conditions for extrema of the PFCM objective function, and use them as the basis for a standard alternating optimization approach to finding local minima of the PFCM objective functional. Several numerical examples are given that compare FCM and PCM to PFCM. Our examples show that PFCM compares favorably to both of the previous models. Since PFCM prototypes are less sensitive

to outliers and can avoid coincident clusters, PFCM is a strong candidate for fuzzy rule-based system identification.

**Index Terms**—c-means models, fuzzy clustering, hybrid clustering, possibilistic clustering.

### I. GIỚI THIỆU

#### I. INTRODUCTION

Phân cụm một tập hợp dữ liệu không có nhãn  $X = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^p$  là sự phân tán của  $X$  thành  $1 < c < n$  nhóm con sao cho mỗi nhóm con biểu diễn cấu trúc con "tự nhiên" trong  $X$ . Điều này được thực hiện bằng cách gán nhãn cho các vector trong  $X$  và do đó, cho các đối tượng tạo ra  $X$ . Phân vùng của  $X$  là một tập hợp các giá trị của  $\{k\}$  có thể được sắp xếp thuận tiện thành ma trận cỡ  $(c \times n)$   $U = [u_{ki}]$ . Có 3 tập hợp ma trận phân vùng.

$$M_{pcn} = \{U \in \mathbb{R}^{cn} : 0 \leq u_{ki} \leq 1 \forall i, k; \forall k \exists i \ni u_{ki} > 0\} \quad (1a)$$

$$M_{fcm} = \{U \in M_{pcn} : \sum_{i=1}^c \forall k; \sum_{k=1}^n > 0 \forall i\} \quad (1b)$$

**CLUSTERING** an unlabeled data set  $X = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^p$  is the partitioning of  $X$  into  $1 < c < n$  subgroups such that each subgroup represents "natural" substructure in  $X$ . This is done by assigning labels to the vectors in  $X$ , and hence, to the objects generating  $X$ . A *c-partition* of  $X$  is a set of  $(cn)$  values  $\{u_{ik}\}$  that can be conveniently arrayed as a  $(c \times n)$  matrix  $U = [u_{ki}]$ . There are three sets of partition matrices.

$$M_{pcn} = \{U \in \mathbb{R}^{cn} : 0 \leq u_{ki} \leq 1 \forall i, k; \forall k \exists i \ni u_{ki} > 0\} \quad (1a)$$

$$M_{fcm} = \{U \in M_{pcn} : \sum_{i=1}^c \forall k; \sum_{k=1}^n > 0 \forall i\} \quad (1b)$$

Bản thảo nhận ngày 30 tháng 1 năm 2004; sửa đổi ngày 8 tháng 7 năm 2004.. Công trình của K. Pal đã được hỗ trợ một phần bởi Bộ Khoa học và Công nghệ, chính phủ Ấn Độ theo lệnh SR/FTP/ET-248/2001.

Manuscript received January 30, 2004; revised July 8, 2004. The work of K. Pal was supported in part by the Department of Science and Technology, Government of India, under Sanction Order SR/FTP/ET-248/2001.

N. R. Pal làm việc tại Đơn vị Khoa học Điện tử và Truyền thông, Ấn Độ Viện Thống kê, Calcutta 7000108, Ấn Độ (e-mail: nikhil@isical).

ac.in).

N. R. Pal is with the Electronics and Communications Sciences Unit, Indian Statistical Institute, Calcutta 7000108, India (e-mail: [nikhil@isical.ac.in](mailto:nikhil@isical.ac.in)).

. Pal làm việc tại Viện Kỹ thuật và Quản lý, Điện tử Salt Lake Khu phức hợp điện tử, Calcutta 700091, Ấn Độ (e-mail: [kuhu305@hotmail.com](mailto:kuhu305@hotmail.com)).

K. Pal is with the Institute of Engineering and Management, Salt Lake Elec- tronics Complex, Calcutta 700091, India (e-mail: [kuhu305@hotmail.com](mailto:kuhu305@hotmail.com)).

J. M. Keller làm việc tại Phòng Kỹ thuật Điện và Máy tính, Đại học Missouri-Columbia, Columbia, MO 65211-2060 Hoa Kỳ (e-mail: [kellerj@missouri.edu](mailto:kellerj@missouri.edu)).

J. M. Keller is with the Electrical and Computer Engineering Department, University of Missouri-Columbia, Columbia, MO 65211-2060 USA (e-mail: [kellerj@missouri.edu](mailto:kellerj@missouri.edu)).

. C. Bezdek làm việc tại Khoa Khoa học Máy tính, Đại học West Florida, Pensacola, FL 32514 Hoa Kỳ (e-mail: [jbezdek@uwf.edu](mailto:jbezdek@uwf.edu)).

J. C. Bezdek is with the Computer Science Department, University of West Florida, Pensacola, FL 32514 USA (e-mail: [jbezdek@uwf.edu](mailto:jbezdek@uwf.edu)).

Bộ nhận dạng đối tượng kỹ thuật số 10.1109/TFUZZ.2004.840099

Digital Object Identifier 10.1109/TFUZZ.2004.840099

$$M_{hcn} = \{U \in M_{fcm} : u_{ik} = 0 \text{ or } 1 \forall i \text{ and } k\} \quad (1c)$$

Phương trình (1) lần lượt định nghĩa các tập hợp *phân vùng* khả thi, mờ hoặc xác suất và rõ ràng của  $X$ . Vì vậy có 4 loại vector nhãn, nhưng vector nhãn mờ và xác suất lại giống hệt nhau về mặt toán học, có các mục từ 0 đến 1 có tổng bằng 1 trên mỗi cột. Lý do các ma trận này được gọi là *phân vùng* xuất phát từ cách giải thích các mục của chúng. Nếu  $U$  rõ ràng hoặc mờ,  $u_{ik}$  được coi là *thành viên* của  $x_k$  trong tập hợp phân vùng mờ (cụm) thứ  $i$  của  $X$ . Nếu  $U$  trong  $M_{fcm}$  là xác suất,  $u_{ik}$  thường là xác suất sau  $p(i|x_k)$  rằng, với  $x_k$  cho trước, nó xuất phát từ lớp  $i$ . Nếu  $U$  trong  $M_{pcn}$  là xác suất khả thi, có giá trị từ 0 đến 1 và không nhất thiết có tổng bằng 1 trên bất kỳ cột nào. Trong trường hợp cuối cùng này,  $u_{ik}$  được coi rằng  $x_k$  có khả năng thuộc về lớp  $i$ . Một cách giải thích khác của khả năng  $u_{ik}$  là do tính *tính điển hình* của  $x_k$  đối với cụm  $i$ . Lưu ý rằng  $M_{hcn} \subset M_{fcm} \subset M_{pcn}$ .

Equation (1) defines, respectively, the sets of possibilistic, fuzzy or probabilistic, and crisp *c-partitions* of  $X$ . So, there *four* kinds of labels vectors, but fuzzy and probabilistic label vectors are mathematically identical, having entries between 0 and 1 that sum to 1 over each column. The reason these matrices are called *partitions* follows from the interpretation of their entries. If  $U$  is crisp or fuzzy,  $u_{ik}$  is taken as the membership of  $x_k$  in the  $i$ -th partitioning fuzzy subset (cluster) of  $X$ . If  $U$  in  $M_{pcn}$  is probabilistic,  $u_{ik}$  is usually the (posterior) probability  $p(i|x_k)$  that, given  $x_k$ , it came from class  $i$ . And if  $U$  in  $M_{pcn}$  is pos-sibilistic, it has entries between 0 and 1 that do not necessarily sum to 1 over any column. In the last case,  $u_{ik}$  is take as the possibility that  $u_{ik}$  belong to class  $i$ . An alternate interpretation of possibility  $u_{ik}$  is the that it measures the *typicality* of  $x_k$  to cluster  $i$ . Observe that  $M_{hcn} \subset M_{fcm} \subset M_{pcn}$ .

Thuật toán phân cụm C tìm thấy một  $U \in M_{hcn}(M_{fcm}, M_{pcn})$  (hi vọng) "tốt nhất" giải thích và biểu diễn cấu trúc(chưa biết)

trong  $X$  liên quan đến mô hình xác định  $C$ . Đối với  $U$  in  $M_{fcm}$ ,  $c = 1$ , được biểu diễn duy nhất bởi 1 phân vùng cứng  $1_n = \underbrace{(1 \ 1 \dots 1)}_{n \text{ lần}}$  phân vùng này gán các đối tượng vào 1 cụm

duy nhất, và  $c = n$  được biểu diễn duy nhất bởi  $U = I_n$ , ma trận đơn vị  $n \times n$ , lên đến một hoán vị của các cột. Trong trường hợp này, mỗi đối tượng nằm trong cụm đơn lẻ của riêng nó. Việc chọn  $c = 1$  hoặc  $c = n$  bác bỏ giả thuyết rằng  $X$  chứa các cụm.

A clustering algorithm C finds a  $U \in M_{hcn}(M_{fcm}, M_{pcn})$  which (hopefully) "best" explains and represents (unknown) structure in  $X$  with respect to the model that defines  $C$ . For  $U$  in  $M_{fcm}$ ,  $c = 1$  is represented uniquely by the hard  $1_{partition}$

$1_n = \underbrace{(1 \ 1 \dots 1)}_{n \text{ times}}$ , which unequivocally assigns all  $n$  objects to a

single cluster, and  $c = n$  is represented uniquely by  $U = I_n$ , the  $n \times n$  identity matrix up to a permutation of columns. In the case, each object is in its own singleton cluster. Choosing  $c = 1$  or  $c = n$  rejects the hypothesis that  $X$  contains cluster.

Một trong những mô hình phân cụm mờ được sử dụng rộng rãi nhất là phân cụm *c-means mờ* (FCM) [1]. Thuật toán FCM gán các thành viên cho  $x_k$  có mối quan hệ nghịch đảo với khoảng cách tương đối của  $x_k$  đến các nguyên mẫu điểm  $c\{v_i\}$  là tâm cụm trong mô hình FCM. Giả sử  $c = 2$ . Nếu  $x_k$  cách đều hai nguyên mẫu, thì thành viên của  $x$  trong mỗi cụm sẽ giống nhau ( $= 0.5$ ), bất kể giá trị tuyệt đối của khoảng cách của  $x_k$  từ hai tâm cụm (cũng như từ các điểm khác trong dữ liệu). Vấn đề mà điều này tạo ra là các điểm nhiễu, cách xa nhưng cách đều cấu trúc trung tâm của hai cụm, tuy nhiên vẫn có thể được gán thành viên bằng nhau trong cả hai, khi có vẻ tự nhiên hơn nhiều khi các điểm như vậy được gán thành viên rất thấp (hoặc thậm chí không có) trong bất kỳ cụm nào.

One of the most widely used fuzzy clustering models is *fuzzy c-means* (FCM) [1]. The FCM algorithm assigns memberships to  $x_k$  which are inversely related to the relative distance of  $x_k$  to the  $c$  point prototypes  $\{v_i\}$  that are cluster centers in the FCM model. Suppose  $c = 2$ . If  $x_k$  is equidistant from two prototypes, the membership of  $x_k$  in each cluster will be the same ( $= 0.5$ ), regardless of the absolute value of the distance of  $x_k$  from the two centroids (as well as from the other points in the data). The problem this creates is that noise points, far but equidistant from the central structure of the two clusters, can nonetheless be given equal membership in both, when it seems far more natural that such points be given very low (or even no) membership in either cluster.

Để khắc phục vấn đề này, Krishnapuram và Keller [2] đã đề xuất một mô hình cụm mới có tên là *c-means khả thi* (PCM), mô hình này nói rằng buộc tổng cột trong (1b) sao cho tổng của mỗi cột thỏa mãn ràng buộc lỏng lẻo hơn  $0 < \sum_{i=1}^c u_{ik} \leq c$ . Nói cách khác, mỗi phần tử của cột thứ  $k$  có thể là bất kỳ số nào giữa 0 và 1, miễn là ít nhất một trong số chúng là số dương. Họ đề xuất rằng trong trường hợp này, giá trị  $u_{ik}$  nên được hiểu là *tính điển hình* của  $x$  so với cụm  $i$  (thay vì tư cách thành viên của nó trong cụm). Họ diễn giải mỗi hàng của  $U$  như một sự phân phối khả năng trên  $X$ . Thuật toán PCM mà họ đề xuất để tối ưu hóa hàm mục tiêu PCM đôi khi giúp xác định các giá trị ngoại lai (điểm nhiễu). Tuy nhiên như Barni và cộng sự đã chỉ ra [3], cái giá mà PCM phải trả cho sự tự bỏ qua các điểm nhiễu là PCM rất nhạy cảm với các sự khởi tạo và đôi khi nó tạo ra các cụm trùng nhau. Hơn nữa, các đặc trưng có thể nhạy cảm với sự lựa chọn các tham số bổ sung cần thiết cho mô hình PCM.

To overcome this problem, Krishnapuram and Keller [2] proposed a new clustering model named *possibilistic c-means* (PCM), which relaxes the column sum constraint in (1b) so that the sum of each

column satisfies the looser constraint  $0 < \sum_{i=1}^c u_{ik} \leq c$ . In other words, each element of the  $k$ -th  $C$  column can be any number between 0 and 1, as long as at least one of them is positive. They suggested that in this case the value  $u_{ik}$  should be interpreted as the *typicality* of  $x_k$  relative to cluster  $i$  (rather than its membership in the cluster). They interpreted each row of  $U$  as a possibility distribution over  $X$ . The PCM algorithm they suggested for optimization of the PCM objective function sometimes helps to identify outliers (noise points). However, as pointed out by Barni *et al.* [3], the price PCM pays for its freedom to ignore noise points is that PCM is very sensitive to initializations, and it sometimes generates coincident clusters. Moreover, typicalities can be very sensitive to the choice of the additional parameters needed by the PCM model.

Timm *et al.* [13]-[15] đã đề xuất hai thuật toán phân cụm mờ khả thi có thể tránh được vấn đề phân cụm trùng nhau của PCM. Trong [13] và [14], các tác giả đã sửa đổi hàm mục tiêu PCM bằng cách thêm một hàm nghịch đảo của khoảng cách giữa các tâm cụm. Thuật ngữ bổ sung này hoạt động như một lực đẩy và giữ cho các cụm tách biệt (tránh các cụm trùng nhau). Trong [14] và [15], Timm *et al.* sử dụng cùng một khái niệm để sửa đổi hàm mục tiêu như được sử dụng trong thuật toán phân cụm Gustafson và Kessel [16]. Các thuật toán này, mặc dù chỉ sử dụng các đặc điểm điển hình (khả năng), nhưng cố gắng khai thác lợi ích của cả phân cụm mờ và phân cụm khả thi.

Timm *et al.* [13]-[15] proposed two possibilistic fuzzy clustering algorithms that can avoid the coincident cluster problem of PCM. In [13] and [14], the authors modified the PCM objective function adding an inverse function of the distances between cluster centers. This extra term acts as a repulsive force and keeps the clusters separate (avoids coincident clusters). In [14] and [15], Timm *et al.* use the same concept to modify the objective function as used in Gustafson and Kessel [16] clustering algorithm. These algorithms, although use only the typicalities (possibilities), attempt to exploit the benefits of both fuzzy and possibilistic clustering.

Trong [12], chúng tôi đã chứng minh được nhu cầu về cả giá trị khả năng (tức là tính điển hình) và giá trị thành viên, đồng thời đề xuất một mô hình và thuật toán đi kèm để tối ưu hóa nó. Bài báo năm 1997 của chúng tôi gọi thuật toán này là FPCM. FPCM chuẩn hóa các giá trị khả năng, do đó tổng các khả năng của tất cả các điểm dữ liệu trong một cụm là 1. Mặc dù FPCM ít gặp phải các vấn đề của cả FCM và PCM vừa mô tả, nhưng các giá trị khả năng rất nhỏ khi kích thước của tập dữ liệu tăng lên. Trong bài báo này, chúng tôi đề xuất một mô hình mới kết hợp FCM và PCM, tận dụng lợi ích của cả hai mô hình và loại bỏ vấn đề của FPCM. Để tránh nhầm lẫn, chúng tôi gọi mô hình mới này là *c-means mờ khả năng* (PFCM). Phần còn lại của bài báo được tổ chức như sau. Phần II thảo luận về FCM. Phần III thực hiện tương tự đối với PCM. Trong Phần IV, chúng tôi thảo luận về mô hình cụm FPCM, cùng với các điều kiện cần thiết bậc nhất cho hàm FPCM. Trong Phần V, chúng tôi trình bày mô hình PFCM mới, Phần VI bao gồm một số ví dụ số so sánh FCM và PCM với PFCM. Phần VII có phần thảo luận và kết luận của chúng tôi.

In [12], we justified the need for *both* possibility (i.e., typicality) and membership values, and proposed a model and companion algorithm to optimize it. Our 1997 paper called this algorithm FPCM. FPCM normalizes the possibility values, so that the sum of possibilities of all data points in a cluster is 1. Although FPCM is much less prone to the problems of both FCM and PCM just described, the possibility values are very small when the size of the data set increases. In this paper we propose a new model that hybridizes FCM and PCM, enjoys the benefits of both models, and eliminates the problem of FPCM. To avoid confusion, we call this new model *possibilistic fuzzy c-means* (PFCM). The rest of the paper

is organized as follows. Section II discusses FCM. Section III does the same for PCM. In Section IV we discuss the FPCM clustering model, along with the first order necessary conditions for the FPCM functional. In Section V we present the new PFCM model, Section VI includes some numerical examples that compare FCM and PCM to PFCM. Section VII has our discussion and conclusions.

## II. CÓ VẤN ĐỀ GÌ VỚI PHÂN CỤM MỜ VỚI PHÂN CỤM KHẢ THI

### II. WHAT'S WRONG WITH FUZZY AND PROBABILISTIC PARTITIONS?

Mô hình FCM là bài toán tối ưu bị ràng buộc

$$\min_{(U,V)} \{J_m(U, V, X) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m \|x_k - v_i\|^2\} \quad (2)$$

trong đó  $U \in M_{fcn}$ ,  $V = (v_1, v_2, \dots, v_c)$  là một véc tơ của các tâm cụm (trọng số hoặc nguyên mẫu) (không xác định) vì  $v_i \in \mathbb{R}^p$  đối với  $1 \leq i \leq c$ , and  $\|x\|_A = \sqrt{x^T A x}$  là bất kì chuẩn tích vô hướng nào. Phân vùng tối ưu  $U^*$  của  $X$  được lấy từ cặp  $(U^*, V^*)$  là bộ tối thiểu cục bộ của  $J_m$ . Tối ưu hóa gần đúng của  $J_m$  bằng thuật toán FCM-AO dựa trên phép lặp thông qua các điều kiện cần thiết cho vị trí cục bộ của nó.

The FCM model is the constrained optimization problem

$$\min_{(U,V)} \{J_m(U, V, X) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m \|x_k - v_i\|^2\} \quad (2)$$

where  $U \in M_{fcn}$ ,  $V = (v_1, v_2, \dots, v_c)$  is a vector of (unknown) cluster centers (weights or prototypes),  $v_i \in \mathbb{R}$  for  $1 \leq i \leq c$ , and  $\|x\|_A = \sqrt{x^T A x}$  is any inner product norm. Optimal partitions  $U^*$  of  $X$  are taken from pairs  $(U^*, V^*)$  that are local minimizers of  $J_m$ . Approximate optimization of  $J_m$  by the FCM-AO algorithm is based on iteration through the following necessary conditions for its local extrema.

**Định lý FCM [1]** Nếu  $D_{ikA} = \|x_k - v_i\|_A > 0$  với mọi  $i$  và  $k$ ,  $m > 1$ , và  $X$  chứa ít nhất  $c$  điểm phân biệt, thì  $(U, V) \in M_{fcn} \times \mathbb{R}^{cp}$  chỉ có thể tối thiểu  $J_m$  nếu

$$u_{ik} = \left( \sum_{j=1}^c \left( \frac{D_{ikA}}{D_{jkA}} \right)^2 / (m-1) \right)^{-1} \quad 1 \leq i \leq c; 1 \leq k \leq n \text{ và} \quad (3a)$$

$$v_i = \frac{\sum_{k=1}^n u_{ik}^m x_k}{\sum_{k=1}^n u_{ik}^m} \quad 1 \leq i \leq c \quad (3b)$$

**Theorem FCM [1]:** If  $D_{ikA} = \|x_k - v_i\|_A > 0$  for all  $i$  and  $k$ ,  $m > 1$ , and  $X$  contains at least  $c$  distinct points, then  $(U, V) \in M_{fcn} \times \mathbb{R}^{cp}$  may minimize  $J_m$  only if

$$u_{ik} = \left( \sum_{j=1}^c \left( \frac{D_{ikA}}{D_{jkA}} \right)^2 / (m-1) \right)^{-1} \quad 1 \leq i \leq c; 1 \leq k \leq n \text{ and} \quad (3a)$$

$$v_i = \frac{\sum_{k=1}^n u_{ik}^m x_k}{\sum_{k=1}^n u_{ik}^m} \quad 1 \leq i \leq c \quad (3b)$$

Kỳ dị trong FCM xảy ra khi một hoặc nhiều khoảng cách  $D_{ikA} = 0$  tại bất kỳ lần lặp nào. Trong trường hợp này (hiếm khi xảy ra trong thực tế), hãy gán 0 cho mỗi  $u_{ik}$  mà  $D_{ikA} > 0$  và phân bố các thành viên tùy ý trên các  $v_i$  mà  $D_{ikA} = 0$ , tuân theo các ràng buộc trong (1b). Thuật toán phổ biến nhất để tìm ra các nghiệm xấp xỉ (2) là lặp Picard thông qua (3a) và (3b). Kiểu lặp này thường được gọi là *tối ưu hóa xen kẽ* vì nó chỉ lặp qua một chu kỳ ước tính

cho  $V_{t-1} \Rightarrow U_t \Rightarrow V_t$  và sau đó kiểm tra  $\|V_t - V_{t-1}\|_{err} < \epsilon$ . Tương đương, toàn bộ quy trình có thể được dịch chuyển một nửa chu kỳ, do đó việc khởi tạo và kết thúc được thực hiện trên  $U$  và các lần lặp trở thành  $U_{t-1} \Rightarrow V \Rightarrow U_t$ , với tiêu chí kết thúc xen kẽ  $\|U_t - U_{t-1}\|_{err} < \epsilon$ . Tài liệu chứa cả hai thông số kỹ thuật; lý thuyết hội tụ là giống nhau trong cả hai trường hợp. Có một số lợi thế rõ ràng khi khởi tạo và kết thúc trên  $V$  về mặt tiện lợi, tốc độ và lưu trữ. Dạng thay thế khởi tạo và kết thúc trên  $U$  nghiêm ngặt hơn, vì nhiều tham số hơn phải trở nên gần nhau trước khi kết thúc đạt được. Có thể xảy ra trường hợp các kết quả khác nhau xảy ra khi sử dụng cùng  $\epsilon$  với cả hai dạng. Phụ lục trình bày dạng FCM-AO được sử dụng trong các ví dụ của chúng tôi. Một tính chất giới hạn của (3) quan trọng đối với nghiên cứu này là [1]:

$$\lim_{m \rightarrow 1^+} \{u_{ik}\} = \begin{cases} 1, & D_{ikA} < D_{jkA} \quad \forall j \neq i \\ 0, & \text{otherwise} \end{cases}$$

$$1 \leq i \leq c; 1 \leq k \leq n. \quad (4a)$$

Sử dụng kết quả này, chúng ta lấy cùng giới hạn trong (3b), thu được

$$\lim_{m \rightarrow 1^+} \left\{ \left( \mathbf{v}_i = \frac{\sum_{k=1}^n (u_{ik})^m \mathbf{x}_k}{\sum_{k=1}^n (u_{ik})^m} \right) = \frac{\sum_{\mathbf{x}_k \in X_i} \mathbf{x}_k}{n_i} = \bar{\mathbf{v}}_i \right\}, \quad 1 \leq i \leq c$$

trong đó  $X = X_1 \cup \dots \cup X_i \cup \dots \cup X_c$  là phân cụm cứng của  $X$  được xác định bởi vế phải của (4a) với  $\sum_{k=1}^n u_{ik} = n_i = |X_i|$  và  $\mathbf{v}_i$  là vectơ trung bình của  $X_i$ . Nếu chúng ta sử dụng các kết quả này trong (2), chúng ta sẽ có  $J_1(U, V; X)$  tại (5), như được hiển thị ở cuối trang tiếp theo.  $J(U, V; X)$  là hàm mục tiêu tổng bình phương lỗi trong nhóm cổ điển. Phương trình (5) là bài toán tối ưu hóa xác định mô hình *c-means cứng* (HCM). Hơn nữa, vế phải của (4a) và (4b) là các điều kiện cần thiết cho cực trị cục bộ của  $J_1$ . Lấy giới hạn của (3) ở cực trị còn lại, ta có

$$\lim_{m \rightarrow \infty} \{u_{ik}\} = \frac{1}{c} \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (6a)$$

$$\lim_{m \rightarrow \infty} \left\{ \left( \mathbf{v}_i = \frac{\sum_{k=1}^n u_{ik} \mathbf{x}_k}{\sum_{k=1}^n u_{ik}} \right) \right\} = \frac{\sum_{k=1}^n \mathbf{x}_k}{n} = \bar{\mathbf{v}} \quad 1 \leq i \leq c \quad (6b)$$

Trong (6b),  $\bar{\mathbf{v}}$  là trung bình trên của  $X$ . Nếu  $U \in M_{fcm}$  ràng buộc  $\sum_{i=1}^c u_{ik} = 1 \forall k$  khiến việc diễn giải  $u_{ik}$  là tính điển hình của  $\mathbf{x}_k$  đối với cụm thứ  $i$  trở nên khó khăn. Để minh họa cho vấn đề phân vùng mờ này do FCM tạo ra, Hình 1 và 2 hiển thị các tập dữ liệu 2 chiều  $X_{10}$  and  $X_{12}$  bởi 10 và 12 điểm có tọa độ được ra trong cột 2 và 3 của bảng I.

Chúng tôi biểu thị  $\{x_1, x_2, \dots, x_{10}\}$  bởi  $X_{10}$  và  $X_{10} \cup \{X_{11}, X_{12}\}$  as  $X_{12}$ . Một số tác giả gọi  $x_{11}$  là cầu nối và  $x_{12}$  là nhiễu.  $X_{10}$  có 2 cụm hình thoi với năm điểm ở bên trái và bên phải của trục  $y$ .  $X_{11}$  và  $X_{12}$  cách đều tất cả các cặp điểm tương ứng trong 2 cụm.

Singularity in FCM occurs when one or more of the distances  $D_{ikA} = 0$  at any iterate. In this case (rare in practice), assign 0's to each  $u_{ik}$  for which  $D_{ikA} > 0$ , and distribute memberships arbitrarily across the  $\mathbf{v}_i$ 's for which  $D_{ikA} = 0$ , subject to the constraints in (1b). The most popular algorithm for approximating solutions of (2) is Picard iteration through (3a) and (3b). This type of iteration is often called *alternating optimization* (AO) as it simply loops through one cycle of estimates for  $V_{t-1} \Rightarrow U_t \Rightarrow V_t$  and then checks  $\|V_t - V_{t-1}\|_{err} < \epsilon$ . Equivalently, the entire procedure can be shifted one half cycle, so that initialization and termination is done on  $U$ , and the iterates become  $U_{t-1} \Rightarrow V_t \Rightarrow U_t$ , with the alternate

termination criterion  $\|U_t - U_{t-1}\|_{err} < \epsilon$ . The literature contains both specifications; the convergence theory is the same in either case. There are some obvious advantages to initializing and terminating on  $V$  in terms of convenience, speed and storage. The alternate form that initializes and terminates on  $U$ 's is more stringent, since many more parameters must become close before termination is achieved. It can happen that different results ensue by using the same  $\epsilon$  with both forms. The Appendix exhibits the form of FCM-AO used in our examples. A limit property of (3) that is important for this study is [1]:

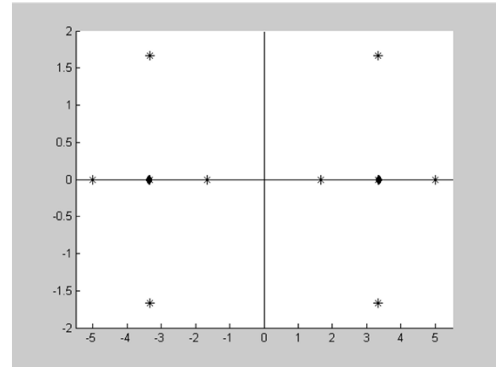
$$\lim_{m \rightarrow 1^+} \{u_{ik}\} = \begin{cases} 1, & D_{ikA} < D_{jkA} \quad \forall j \neq i \\ 0, & \text{otherwise} \end{cases}$$

$$1 \leq i \leq c; 1 \leq k \leq n. \quad (4a)$$

Using this result, we take the same limit in (3b), obtaining

$$\lim_{m \rightarrow 1^+} \left\{ \left( \mathbf{v}_i = \frac{\sum_{k=1}^n (u_{ik})^m \mathbf{x}_k}{\sum_{k=1}^n (u_{ik})^m} \right) \right\} = \frac{\sum_{\mathbf{x}_k \in X_i} \mathbf{x}_k}{n_i} = \bar{\mathbf{v}}_i, \quad 1 \leq i \leq c \quad (4b)$$

where  $X = X_1 \cup \dots \cup X_i \cup \dots \cup X_c$  is the hard  $c$ -partition of  $X$  defined by the right side of (4a) with  $\sum_{k=1}^n u_{ik} = n_i = |X_i|$  and  $\mathbf{v}_i$  the mean vector of  $X_i$ . If we use these results in (2), we get  $J_1(U, V; X)$  at (5), as shown at the bottom of the next page.  $J_1(U, V; X)$  is the classical within-groups sum of squared errors objective function. Equation (5) is the optimization problem that defines the *hard c-means* (HCM) model. Moreover, the right sides of (4a) and (4b) are the necessary conditions for local extrema of  $J_1$ . Taking the limit of (3) at the other extreme, we get



Hình 1. Tập dữ liệu  $X_{10}$  và  $V_{FCM}^{10} = [v_{1,FCM}^{10}, v_{2,FCM}^{10}]$  hiển thị như hình thoi

$$\lim_{m \rightarrow \infty} \{u_{ik}\} = \frac{1}{c} \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (6a)$$

$$\lim_{m \rightarrow \infty} \left\{ \left( \mathbf{v}_i = \frac{\sum_{k=1}^n u_{ik} \mathbf{x}_k}{\sum_{k=1}^n u_{ik}} \right) \right\} = \frac{\sum_{k=1}^n \mathbf{x}_k}{n} = \bar{\mathbf{v}} \quad 1 \leq i \leq c \quad (6b)$$

In (6b),  $\bar{\mathbf{v}}$  is the grand mean of  $X$ . If  $U \in M_{fcm}$ , the constraint  $\sum_{i=1}^c u_{ik} = 1 \forall k$  makes it difficult to interpret  $u_{ik}$  as the typicality of  $\mathbf{x}_k$  to the  $i^{th}$  cluster. To illustrate this problem of fuzzy partitions generated by FCM, Figs. 1 and 2 show two-dimensional data sets  $X_{10}$  and  $X_{12}$  with 10 and 12 points whose coordinates are given in columns 2 and 3 of Table I.

We denote  $\{x_1, x_2, \dots, x_{10}\}$  by  $X_{10}$  and  $X_{10} \cup \{X_{11}, X_{12}\}$  as  $X_{12}$ . Some authors call  $X_{11}$  an "inlier" (bridge) and  $X_{12}$  an "outlier" (noise).  $X_{10}$  has two diamond shaped clusters with five points each on the left and right sides of the  $y$  axis.  $X_{11}$  and  $X_{12}$  are equidistant



from all corresponding pairs of points in the two clusters.

Ví dụ 1:

Tập dữ liệu:  $X_{10}, X_{12}$ .

Mô hình: FCM

Khởi tạo:

$$V_0 = \begin{bmatrix} 0.07 & 0.36 \\ 0.40 & 0.99 \end{bmatrix} \quad (7a)$$

Các tham số:  $c = m = 2$ .

Để hiểu cách các giá trị nội tại và ngoại tại có thể ảnh hưởng đến các phân vùng được FCM tìm thấy, chúng tôi đã áp dụng FCM cho  $X_{10}$  và  $X_{12}$  với  $c = m = 2$  và các giao thức FCM khác như được liệt kê trong Phần V. Chúng tôi đã khởi tạo FCM trên  $V$  bằng cách sử dụng các giá trị ngẫu nhiên được lấy từ  $[0, 1]$  như thể hiện trong (7a) - các giá trị thực được làm tròn thành hai chữ số có nghĩa để hiển thị. Bảng I hiển thị các giá trị thành viên (các hàng  $U_i$  của  $U$  được hiển thị chuyển thành các cột trong Bảng, được làm tròn đến độ chính xác hai chữ số) thu được cho mỗi điểm tại điểm kết thúc sau 12 lần lặp của FCM-AO. Tâm cụm cuối cùng  $V_{FCM}^{10} = [V_{1,FCM}^{10}, V_{2,FCM}^{10}]$  được thể hiện trong (7b) và cũng được thể hiện trong Hình 1 bằng ký hiệu hình thoi.  $V_{FCM}^{10} = \begin{bmatrix} -3.36 & 3.36 \\ 0.00 & 0.00 \end{bmatrix}$

$$V_{FCM}^{10} = \begin{bmatrix} -2.99 & 2.99 \\ 0.54 & 0.54 \end{bmatrix} \quad (7b)$$

Bây giờ, giả sử chúng ta thêm các điểm  $x_{11}$  và  $x_{12} = (0.0, 10.0)^T$  vào  $X_{10}$ . Điểm  $X_{10}$  cách  $X_{11}$  mười đơn vị trực tiếp như thể hiện trong Hình 2. Áp dụng FCM cho  $X_{12}$  với cùng các tham số và khởi tạo như trước, chúng ta có phân vùng FCM được hiển thị trong các cột 6 và 7 của Bảng I và cụm cuối cùng

$$\begin{aligned} \lim_{m \rightarrow \infty} \{ \min_{(U,V)} \{ J_m(U, V, X) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik}^m) \|x_k - v_i\|^2 \} \} \\ = \min_{(U,V)} \{ J_1(U, V, X) \\ = \sum_{k=1}^n \sum_{i=1}^c u_{ik} \|x_k - v_i\|_A^2 \} \end{aligned} \quad (5)$$

Example 1:

Data set:  $X_{10}, X_{12}$ .

Algorithm: FCM.

Initialization:

$$V_0 = \begin{bmatrix} 0.07 & 0.36 \\ 0.40 & 0.99 \end{bmatrix} \quad (7a)$$

To understand how inliers and outliers can affect partitions found by FCM, we applied FCM to  $X_{10}$  and  $X_{12}$  with  $c = m = 2$  and other FCM protocols as listed in Section V. We initialized FCM on  $V$  using random values drawn from  $[0, 1]$  as shown in (7a)-the actual values are truncated to two significant digits for display purpose. Table I shows the membership values (rows  $U$  of  $U$  are shown transposed as columns in the Table, rounded to two-digit accuracy) obtained for each point at termination after 12 iterations of FCM-AO. The terminal centroids

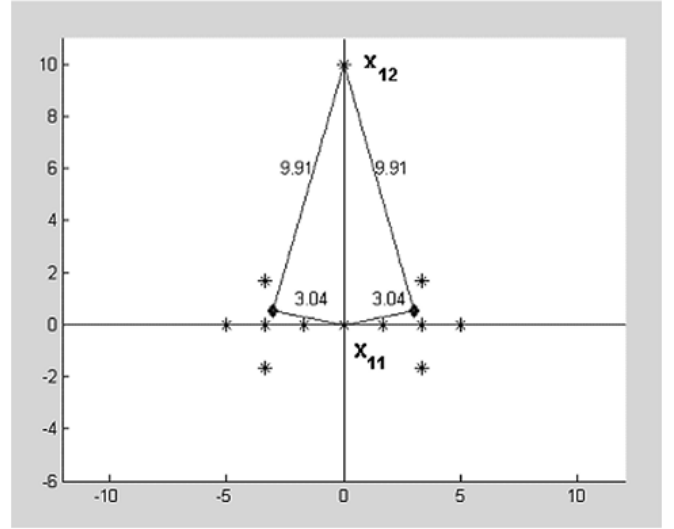
$V_{FCM}^{10} = [V_{1,FCM}^{10}, V_{2,FCM}^{10}]$  are shown in (7b) and also shown in Fig. 1 by the diamond symbol.

$$V_{FCM}^{10} = \begin{bmatrix} -3.36 & 3.36 \\ 0.00 & 0.00 \end{bmatrix}$$

$$V_{FCM}^{10} = \begin{bmatrix} -2.99 & 2.99 \\ 0.54 & 0.54 \end{bmatrix} \quad (7b)$$

Now, suppose we add the points  $x_{12} = (0.0, 10.0)^T$  and  $X_{12} = (0.0, 10.0)$  to  $X_{10}$ . The point  $X_{12}$  is ten units directly above  $X_{11}$  as shown in Fig. 2. Applying FCM to  $X_{12}$  with the same parameters and initialization as before, we get the FCM partition shown in columns 6 and 7 of Table I, and the terminal cluster

$$\begin{aligned} \lim_{m \rightarrow \infty} \{ \min_{(U,V)} \{ J_m(U, V, X) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik}^m) \|x_k - v_i\|^2 \} \} \\ = \min_{(U,V)} \{ J_1(U, V, X) = \sum_{k=1}^n \sum_{i=1}^c u_{ik} \|x_k - v_i\|_A^2 \} \end{aligned} \quad (5)$$



Hình 2 Tập dữ liệu  $X_{12}$  (cột 1 và 2 của Bảng I) và  $v_{12}^{FCM}$

Fig. 2. Data set  $X_{12}$  (columns 1 and 2 of Table I) and  $v_{12}^{FCM}$

Pt.	Data		FCM on $X_{10}$		FCM on $X_{12}$	
	x	y	$U_1^T$	$U_2^T$	$U_1^T$	$U_2^T$
1	-5.00	0.00	0.96	0.04	0.94	0.06
2	-3.34	1.67	0.95	0.05	0.97	0.03
3	-3.34	0.00	1.0	0.00	0.99	0.01
4	-3.34	-1.67	0.95	0.05	0.90	0.10
5	-1.67	0.00	0.90	0.10	0.92	0.08
6	1.67	0.00	0.10	0.90	0.08	0.92
7	3.34	1.67	0.06	0.94	0.03	0.97
8	3.34	0.00	0.00	1.00	0.01	0.99
9	3.34	-1.67	0.06	0.94	0.10	0.90
10	5.00	0.00	0.04	0.96	0.06	0.94
11	0.00	0.00			0.50	0.50
12	0.00	10.00			0.50	0.50

Bảng I. Tập dữ liệu  $X_{10}, X_{12}$  và  $U_{FCM}$  khi kết thúc tại  $X_{10}$  và  $X_{12}$

TABLE I. DATA SETS  $X_{10}, X_{12}$  AND  $U_{FCM}$  AT TERMINATION ON  $X_{10}$  AND  $X_{12}$

Tâm  $V_{FCM}^{12}$  cũng được thể hiện trong (7b) và trong Hình 2 bằng ký hiệu kim cương. Các điểm  $x_{11}$  và  $x_{12}$  có giá trị thành viên là 0.50 trong mỗi cụm vì cả hai đều cách đều tâm  $V_{FCM}^{12}$ , mặc dù  $x_{12}$  cách xa  $V_{FCM}^{12}$  hơn  $x_{11}$  hơn ba lần. Điều này được minh họa trong Hình 2.

Centers  $V_{FCM}^{12}$  are also shown in (7b) and in Fig. 2 by the diamond symbol. Points  $x_{11}$  and  $x_{12}$  have membership values of 0.50 in each cluster because both are equidistant from the centroids  $V_{FCM}^{12}$ , even though  $x_{12}$  is more than three times further away from  $V_{FCM}^{12}$  than is  $x_{11}$ . This is illustrated in Fig. 2.

Việc thêm  $x_{11}$  và  $x_{12}$  vào  $X_{10}$  không làm thay đổi nhiều thành viên đầu cuối của các điểm  $x_1, \dots, x_{10}$  (thay đổi tối đa là 0,05). Tuy nhiên,  $x_{11}$  có vẻ điển hình hơn nhiều so với cấu trúc tổng thể của dữ liệu so với  $x_{12}$ . Điều này minh họa cách một điểm nhiễu như  $x_{12}$  có thể ảnh hưởng xấu đến việc giải thích dữ liệu bằng cách sử dụng thành viên mờ. Vấn đề nằm ở khái niệm cơ bản về phân vùng mờ (hoặc xác suất!) của các tập dữ liệu. Cụ thể, ràng buộc tổng hợp trong (1b) buộc tình huống không mong muốn này xảy ra.

The addition of  $x_{11}$  and  $x_{12}$  to  $X_{10}$  does *not* change the terminal memberships of points  $x_1, \dots, x_{10}$  very much at all (the maximum change is 0.05). However,  $x_{11}$  seems far more typical of the overall structure of the data than  $x_{12}$ . This illustrates how a noise point such as  $x_{12}$  can adversely affect data interpretation using fuzzy memberships. The problem lies with the basic notion of fuzzy (or probabilistic!) partitioning of data sets. Specifically, the summation constraint in (1b) forces this undesirable situation to occur.

### III. CÓ VẤN ĐỀ GÌ VỚI THÀNH PHẦN KHẢ THI III. WHAT'S WRONG WITH POSSIBILISTIC PARTITIONS?

Để tránh các kết quả phản trực giác vừa hiển thị, Krishnapuram và Keller [2] đề xuất nối lỏng ràng buộc cột  $\sum_{i=1}^c u_{ik} = 1 \quad \forall k$  cho các phân vùng mờ trong (1b) để  $u_{ik}$  phản ánh tốt hơn những gì chúng ta mong đợi đối với tính điển hình của  $x_k$  với cụm thứ  $i$ . Chúng tôi biểu diễn tính điển hình bằng  $t_{ik}$  và ma trận tính điển hình là  $T = [t_{ik}]_{c \times n}$ . Krishnapuram và Keller [2] đã đề xuất mô hình PCM

$$\min_{(T,V)} \{P_m(T, V; X, \gamma) = \sum_{k=1}^n \sum_{i=1}^c (t_{ik})^m \|x_k - v_i\|_A^2 + \sum_{i=1}^c \gamma_i \sum_{k=1}^n (1 - t_{ik})^m\} \quad (8)$$

trong đó  $\|x\|_A = \sqrt{x^T A x}$  là bất kỳ chuẩn tích vô hướng nào,  $T \in M_{pcn}$ ,  $V = (v_1, v_2, \dots, v_c)$  là một vvector của các tâm cụm,  $v_i \in \mathbb{R}$  và  $\gamma_i > 0$  là hằng số do người dùng định nghĩa,  $1 < i < c$ . Vì các hàng và cột của  $T$  độc lập, nên việc tối thiểu hóa  $P_m(T, V) = \sum_{i=1}^c \sum_{k=1}^n P_m^{ij}(T, V)$  có thể được thực hiện bằng cách tối thiểu hóa số hạng thứ  $ij$  của  $P_m$  đối với  $T$

$$P_{ij}(T, V) = t_{ij}^m D_{ijA}^2 + \gamma_i (1 - t_{ij})^m \quad (9)$$

Tối ưu hóa gần đúng  $P_m$  bằng thuật toán PCM-AO dựa trên phép lặp qua các điều kiện cần thiết sau đây cho cực trị cục bộ của nó.

To circumvent the counterintuitive results just displayed, Krishnapuram and Keller [2] suggest relaxing the column constraint  $\sum_{i=1}^c u_{ik} = 1 \quad \forall k$  for fuzzy partitions in (1b) so that  $u_{ik}$  better reflects what we expect for the typicality of  $x_k$  to the  $i$ th cluster. We represent typicality by  $t_{ik}$  and the typicality matrix as  $T = [t_{ik}]_{c \times n}$ . Krishnapuram and Keller [2] proposed the PCM model

$$\min_{(T,V)} \{P_m(T, V; X, \gamma) = \sum_{k=1}^n \sum_{i=1}^c (t_{ik})^m \|x_k - v_i\|_A^2 + \sum_{i=1}^c \gamma_i \sum_{k=1}^n (1 - t_{ik})^m\} \quad (8)$$

where  $\|x\|_A = \sqrt{x^T A x}$  is any inner product norm,  $T \in M_{pcn}$ ,  $V = (v_1, v_2, \dots, v_c)$  is a vector of cluster centers,  $v_i \in \mathbb{R}$  and  $\gamma_i > 0$  is a user-defined constant,  $1 < i < c$ . Since the rows and columns of  $T$  are independent, minimization of  $P_m(T, V) =$

$\sum_{i=1}^c \sum_{k=1}^n P_m^{ij}(T, V)$  can be done by minimizing the  $ij$ th term of  $P_m$  with respect to  $T$

$$P_{ij}(T, V) = t_{ij}^m D_{ijA}^2 + \gamma_i (1 - t_{ij})^m \quad (9)$$

Approximate optimization of  $P_m$  by the PCM-AO algorithm is based on iteration through the following necessary conditions for its local extrema.

Định lý PCM [2]: Nếu  $\gamma_i > 0, 1 \leq i \leq c, m > 1$  và  $X$  chứa ít nhất  $c$  điểm phân biệt, thì  $(U, V) \in M_{pcn} \times \mathbb{R}^{cp}$  chỉ có thể tối thiểu  $P_m$  nếu

$$t_{ik} = \frac{1}{1 + \left(\frac{D_{ikA}^2}{\gamma_i}\right)^{\frac{1}{m-1}}}, 1 \leq i \leq c; 1 \leq k \leq n \quad (10a)$$

$$v_i = \frac{\sum_{k=1}^n t_{ik}^m x_k}{\sum_{k=1}^n t_{ik}^m}, 1 \leq i \leq c; 1 \leq k \leq n. \quad (10b)$$

Ở đây, i) không giống như phần tương ứng của nó trong (3a),  $D_{ikA}$  có thể bằng 0 trong (10a), do đó PCM không gặp phải cùng một vấn đề kỳ dị như FCM, và ii) dạng hàm trong (10b) giống hệt với dạng hàm trong (3b). Thuật toán PCM-AO để giải (8) do Krishnapuram và Keller đề xuất chỉ cần thay thế (3) bằng (10) trong FCM-AO và thêm thông số kỹ thuật của  $\{\gamma_i\}$  vào "danh sách chọn" trong Phụ lục. PCM-AO cũng có thể được khởi tạo bằng  $T_0$  hoặc  $V_0$ . Krishnapuram và Keller khuyên bạn nên sử dụng đầu ra đầu cuối của FCM-AO như một cách tốt để khởi tạo PCM-AO. Họ cũng đề xuất chọn  $\{\gamma_i\}$  bằng cách tính toán

$$\gamma_i = K \frac{\sum_{k=1}^n u_{ik}^m D_{ikA}^2}{\sum_{k=1}^n u_{ik}^m} \quad 1 \quad K > 0 \text{ (thường chọn } K = 1) \quad (11)$$

trong đó  $u_{ik}$  là phân vùng FCM đầu cuối của  $X$ . Một số khía cạnh của PCM được thảo luận trong [4] và [5]. PCM đôi khi hữu ích khi dữ liệu bị nhiễu. Tuy nhiên, người dùng cần phải rất cẩn thận về việc lựa chọn  $V_0, \gamma_i$  và  $K$  khi sử dụng PCM-AO [3], [5].

**Theorem PCM [2]:** If  $\gamma_i > 0, 1 \leq i \leq c, m > 1$ , and  $X$  contains at least  $c$  distinct points, then  $(U, V) \in M_{pcn} \times \mathbb{R}^{cp}$  may minimize  $P_m$  only if

$$t_{ik} = \frac{1}{1 + \left(\frac{D_{ikA}^2}{\gamma_i}\right)^{\frac{1}{m-1}}}, 1 \leq i \leq c; 1 \leq k \leq n \quad (10a)$$

$$v_i = \frac{\sum_{k=1}^n t_{ik}^m x_k}{\sum_{k=1}^n t_{ik}^m}, 1 \leq i \leq c; 1 \leq k \leq n. \quad (10b)$$

Here, i) unlike its counterpart in (3a),  $D_{ikA}$  could be zero in (10a), so PCM does not suffer from the same singularity problem that FCM does, and ii) the functional form in (10b) is identical to that in (3b). The PCM-AO algorithm for solving (8) suggested by Krishnapuram and Keller simply replaced (3) with (10) in FCM-AO and added the specification of the  $\{\gamma_i\}$  to the "pick list" in Appendix. PCM-AO can also be initialized with either  $T_0$  or  $V_0$ . Krishnapuram and Keller recommend using terminal outputs of FCM-AO as a good way to initialize PCM-AO. They also suggest choosing the  $\{\gamma_i\}$  by computing

$$\gamma_i = K \frac{\sum_{k=1}^n u_{ik}^m D_{ikA}^2}{\sum_{k=1}^n u_{ik}^m} \quad K > 0 \text{ (the most common choice is } K = 1) \quad (11)$$

where the  $u_{ik}$  are a terminal FCM partition of  $X$ . Several aspects of PCM are discussed in [4] and [5]. PCM sometimes helps when the data are noisy. However, the user needs to be very careful about the choice of  $V_0$ , the  $\{\gamma_i\}$  and  $K$  when using PCM-AO [3], [5].

Hạng tử đầu tiên của  $P_m$  chỉ là  $J_m$ , và khi không có hạng tử thứ hai, tối ưu hóa không bị ràng buộc sẽ dẫn đến giải pháp tầm thường  $t_{ik} = 0 \forall i, k$ . Hạng tử thứ hai của  $P_m$  hoạt động như một hình phạt cố gắng đưa  $t_{ik}$  về 1. Do đó, như Krishnapuram và Keller

[2], [5] đã chỉ ra, ma trận đặc trưng phụ thuộc mạnh mẽ vào sự lựa chọn  $\{\gamma_i\}$ . Nếu  $\gamma_i$  thấp,  $t_{ik}$  sẽ nhỏ và nếu  $\gamma_i$  cao,  $t_{ik}$  sẽ cao. Một số tính chất giới hạn của  $t_{ik}$  trong (10a) sẽ giúp giải thích độ nhạy này của PCM-AO đối với  $\{\gamma_i\}$ ; xem (12)-(14), như được thể hiện ở cuối trang. Vì các cột và hàng của ma trận đặc trưng độc lập với nhau,  $P_m(T, V)$  có thể được chia thành (cn) các hàm mục tiêu phụ như thể hiện trong (9). Nếu khối tạo của mỗi hàng không đủ khác biệt, có thể tạo ra các cụm trùng nhau. Đây là một lý do để sử dụng FCM-AO để tạo các tham số khởi tạo cho PCM-AO. Đôi khi các cụm trùng nhau có thể có lợi. Ví dụ, khi chúng ta bắt đầu với một giá trị lớn của  $c$  và việc giải mờ phân vùng PCM-AO cuối cùng dẫn đến  $c' < c$  cụm riêng biệt, điều này có thể chỉ ra rằng giá trị chính xác của  $c$  là  $c'$ . Tuy nhiên, điều này chỉ hữu ích khi dữ liệu thực sự có  $c'$  cụm. PCM-AO có thể kết thúc với  $c' < c$  cụm ngay cả khi thực sự có  $c > c'$  cụm riêng biệt trong dữ liệu. Nói cách khác, không có gì đảm bảo rằng  $c = c'$  là số cụm phù hợp ngay cả khi PCM-AO dường như gợi ý điều này.

The first term of  $P_m$  is just  $J_m$ , and in the absence of the second term, unconstrained optimization will lead to the trivial solution  $t_{ik} = 0 \forall i, k$ . The second term of  $P_m$  acts as a penalty which tries to bring  $t_{ik}$  toward 1. Thus, as pointed out by Krishnapuram and Keller [2], [5], the terminal typicality matrix is strongly dependent on the choice of the  $\{\gamma_i\}$ . If  $\gamma_i$  is low,  $t_{ik}$  will be small, and if  $\gamma_i$  is high,  $t_{ik}$  will be high. Some limiting properties of  $t_{ik}$  in (10a) will help explain this sensitivity of PCM-AO to the  $\{\gamma_i\}$ ; see (12)-(14), as shown at the bottom of the page. Since the columns and rows of the typicality matrix are independent of each other,  $P_m(T, V)$  can be split into (cn) sub-objective functions as shown in (9). If the initialization of each row is not sufficiently distinct, coincident clusters may result. This is one reason for using FCM-AO to generate the initializing parameters for PCM-AO. Sometimes coincident clusters can be advantageous. For example, when we start with a large value of  $c$ , and defuzzification of the terminal PCM-AO partition results in  $c' < c$  distinct clusters, this may indicate that the correct value of  $c$  is  $c'$ . However, this is useful only when the data really has  $c'$  clusters. PCM-AO can terminate with  $c' < c$  clusters even when there are really  $c' > c$  distinct clusters in the data. In other words, there is no guarantee that  $c = c'$  is the right number of clusters even when PCM-AO seems to suggest this.

Vấn đề tạo cụm trùng nhau không chỉ riêng cho  $P_m$ . Vấn đề này sẽ phát sinh đối với bất kỳ hàm mục tiêu tách biệt nào, tức là một hàm có thể được biểu thị dưới dạng tổng của các hàm mục tiêu phụ độc lập. Vấn đề này không phải do lựa chọn sai các điều kiện; mà là do thiếu các ràng buộc được đặt trên ma trận đặc trưng.

The problem of coincident cluster generation is not specific to  $P_m$ . This problem will arise for any *separable* objective function, that is, a function that can be expressed as a sum of independent subobjective functions. This problem is not caused by a poor choice of penalty terms; rather, it is due to the *lack* of constraints placed on the typicality matrix.

Tóm tắt Phần II và III:  $\sum_{i=1}^c u_{ik} = 1$  đối với mờ (hoặc xác suất)  $U's \in M_{fcm}$  quá mạnh - nó buộc các giá trị ngoại lai phải thuộc về một hoặc nhiều cụm và do đó, ảnh hưởng không đáng có đến cấu trúc chính của  $X$ . Mặt khác, ràng buộc  $0 < \sum_{i=1}^c t_{ik} \leq c$  là quá yếu - nó cho phép điểm dữ liệu hoạt động gần như độc lập với dữ liệu khác trong  $X$ , dẫn đến  $T \in (M_{pcn} - M_{fcn})$  rất mong manh trước các lựa chọn tham số của nó. Mô hình FPCM được đề xuất để khai thác lợi ích của mô hình hóa mờ và khả thi trong khi khắc phục điểm yếu của chúng.

To summarize Sections II and III:  $\sum_{i=1}^c u_{ik} = 1$  for fuzzy(or probabilistic)  $U's \in M_{fcm}$  is *too strong-it forces outliers to belong to one or more clusters and, therefore, unduly influence the main*

structure of  $X$ . On the other hand, the constraint  $0 < \sum_{i=1}^c t_{ik} \leq c$  is *too weak-it allows data point to behave almost independently of the other data in  $X$ , resulting in  $T \in (M_{pcn} - M_{fcn})$  that is very brittle to the choices of its parameters. The FPCM model was proposed to exploit the benefits of fuzzy and possibilistic modeling while circumventing their weaknesses.*

#### IV. CÓ VẤN ĐỀ GÌ VỚI FPCM? IV. WHAT'S WRONG WITH FPCM?

Chúng tôi tin rằng tư cách thành viên (hoặc tính điển hình tương đối) và khả năng (hoặc tính điển hình tuyệt đối) đều quan trọng đối với việc diễn giải chính xác cấu trúc dữ liệu phụ. Khi chúng ta muốn gán nhãn rõ ràng cho một điểm dữ liệu, tư cách thành viên là một lựa chọn hợp lý vì việc gán một điểm cho cụm có nguyên mẫu gần nhất với điểm đó là điều tự nhiên. Mặt khác, trong khi dự đoán tâm cụm, tính điển hình là một phương tiện quan trọng để giảm bớt các tác động không mong muốn của các điểm ngoại lai.

We believe that memberships (or relative typicalities) and possibilities (or absolute typicalities) are *both* important for correct interpretation of data substructure. When we want to crisply label a data point, membership is a plausible choice as it is natural to assign a point to the cluster whose prototype is closest to the point. On the other hand, while estimating the centroids, typicality is an important means for alleviating the undesirable effects of outliers.

Điều kiện cần thiết FCM trong (3a) đối với  $u_{ik}$  là một hàm của  $x_k$  và tất cả các tâm  $c$ . Mặt khác, điều kiện cần thiết cho tính điển hình  $t_{ik}$  trong (10a) là một hàm của riêng  $x_k$  và  $v_i$ . Nghĩa là,  $u_{ik}$  bị ảnh hưởng bởi vị trí của tất cả các tâm cụm  $c$ , trong khi  $t_{ik}$  chỉ bị ảnh hưởng bởi một. Vì  $t_{ik}$  chỉ phụ thuộc vào (và không phụ thuộc vào các giá trị khác) khoảng cách từ  $x_k$  đến  $v_i$  và vào hằng số  $\gamma_i$ , nên chúng ta coi  $u_{ik}$  là tính điển hình tương đối và  $t_{ik}$  là tính điển hình tuyệt đối (hoặc chỉ là tính điển hình) của  $x_k$  đối với cụm  $i$ .

$$\lim_{\gamma_i \rightarrow \infty} \{t_{ik}\} = 1, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (12)$$

$$\lim_{m \rightarrow 1+} \{t_{ik}\} = \begin{cases} 1, & \text{nếu } D_{ikA}^2 < \gamma_i \\ \frac{1}{2}, & \text{nếu } D_{ikA}^2 = \gamma_i \\ 0, & \text{nếu } D_{ikA}^2 > \gamma_i \end{cases} \quad (13)$$

$$t_{ik} = \frac{1}{2} \text{ nếu } D_{ikA}^2 = \gamma_i, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (14)$$

The FCM necessary condition in (3a) for  $u_{ik}$  is a function of  $x_k$  and all  $c$  centroids. On the other hand, the necessary condition for typicality  $t_{ik}$  in (10a) is a function of  $X$  and  $v_i$  alone. That is,  $U_{ik}$  is influenced by the positions of *all*  $c$  cluster centers, whereas  $t_{ik}$  is affected by only one. Since  $t_{ik}$  depends on only (and not relative to others) the distance from  $x_k$  to  $v_i$  and on the constant  $\gamma_i$ , we regard  $u_{ik}$  as the relative typicality and  $t_{ik}$  the *absolute typicality* (or just typicality) of  $x_k$  with respect to cluster  $i$ .

$$\lim_{\gamma_i \rightarrow \infty} \{t_{ik}\} = 1, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (12)$$

$$\lim_{m \rightarrow 1+} \{t_{ik}\} = \begin{cases} 1, & \text{if } D_{ikA}^2 < \gamma_i \\ \frac{1}{2}, & \text{if } D_{ikA}^2 = \gamma_i \\ 0, & \text{if } D_{ikA}^2 > \gamma_i \end{cases} \quad (13)$$

$$t_{ik} = \frac{1}{2} \text{ if } D_{ikA}^2 = \gamma_i, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (14)$$

Ngay cả khi chúng ta thay đổi các điều kiện của hàm PCM trong (8), chúng ta vẫn sẽ gặp vấn đề về các cụm trùng nhau miễn là hàm mục tiêu có thể tách biệt được. Trong [12], đối với mô hình FPCM,

chúng tôi đã đề xuất bài toán tối ưu hóa sau:

$$\min_{(U,T,V)} \{J_{m,\eta}(U,T,V;X) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik}^m + t_{ik}^\eta) D_{ikA}^2\} \quad (15)$$

tuân theo các ràng buộc  $m > 1, \eta > 1, 0 \leq u_{ik}, t_{ik} \leq 1, D_{ikA} = \|x_k - v_i\|_A$  và

$$\sum_{i=1}^c u_{ik} = 1 \quad \forall k, i.e., U \in M_{fcn} \quad \text{và} \quad (16)$$

$$\sum_{k=1}^n t_{ik} = 1 \quad \forall i, i.e., T^t \in M_{nc} \quad (17)$$

Even if we change the penalty terms of the PCM functional in (8), we will still have the problem of coincident clusters as long as the objective function is separable. In [12], for the FPCM model, we proposed the following optimization problem:

$$\min_{(U,T,V)} \{J_{m,\eta}(U,T,V;X) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik}^m + t_{ik}^\eta) D_{ikA}^2\} \quad (15)$$

subject to the constraints  $m > 1, \eta > 1, 0 \leq u_{ik}, t_{ik} \leq 1, D_{ikA} = \|x_k - v_i\|_A$  và

$$\sum_{i=1}^c u_{ik} = 1 \quad \forall k, i.e., U \in M_{fcn} \quad \text{và} \quad (16)$$

$$\sum_{k=1}^n t_{ik} = 1 \quad \forall i, i.e., T^t \in M_{nc} \quad (17)$$

Chuyển vị của  $T$  có thể được chấp nhận là thành viên của tập  $M_{fnc}$ . Chúng tôi xem  $T$  như một phép gán đặc trưng của  $n$  đối tượng cho  $c$  cụm. Thuật ngữ khả năng  $\sum_{i=1}^c \sum_{k=1}^n t_{ik}^n D_{ikA}^2$  sẽ phân bố  $\{t_{ik}\}$  đối với tất cả  $n$  điểm dữ liệu, nhưng không phải đối với tất cả  $c$  cụm. Trong các điều kiện thông thường được đặt trên các bài toán tối ưu hóa c-means, chúng tôi đã thu được các điều kiện cần thiết bậc nhất cho cực trị của  $J_{m,n}$  mà chúng tôi nêu ở đây là một định lý.

The transpose of admissible  $T$ 's are members of the set  $M_{fnc}$ . We view  $T$  as a typicality assignment of the  $n$  objects to the  $c$  clusters. The possibilistic term  $\sum_{i=1}^c \sum_{k=1}^n t_{ik}^n D_{ikA}^2$  will distribute the  $\{t_{ik}\}$  with respect to all  $n$  data points, but not with respect to all  $c$  clusters. Under the usual conditions placed on c-means optimization problems, we obtained the first order necessary conditions for extrema of  $J_{m,n}$ , which we state here as a theorem.

**Định lý FPCM [12]:** Nếu  $D_{ikA} = \|x_k - v_i\|_A > 0$  đối với mọi  $i$  và  $k, m, \eta > 1$  và  $X$  chứa ít nhất  $c$  điểm dữ liệu riêng biệt, thì  $(U, T^t, V) \in M_{fcm} \times M_{fcn} \times \mathbb{R}^p$  chỉ có thể tối thiểu  $J_{m,n}$  nếu

$$u_{ik} = \left( \sum_{j=1}^c \left( \frac{D_{ikA}}{D_{jkA}} \right)^{\frac{2}{(m-1)}} \right)^{-1} \quad 1 \leq i \leq c, 1 \leq k \leq n \quad (18a)$$

$$t_{ik} = \left( \sum_{j=1}^n \left( \frac{D_{ikA}}{D_{ijA}} \right)^{\frac{2}{(\eta-1)}} \right)^{-1} \quad 1 \leq i \leq c; 1 \leq k \leq n \quad \text{và} \quad (18b)$$

$$v_i = \frac{\sum_{k=1}^n (u_{ik}^m + t_{ik}^\eta) x_k}{\sum_{k=1}^n (u_{ik}^m + t_{ik}^\eta)} \quad 1 \leq i \leq c \quad (18c)$$

**Theorem FPCM [12]:** If  $D_{ikA} = \|x_k - v_i\|_A > 0$  for all  $i$  and  $k, m, \eta > 1$ , and  $X$  contains at least  $c$  distinct data points, then  $(U, T^t, V) \in M_{fcm} \times M_{fcn} \times \mathbb{R}^p$  may minimize  $J_{m,\eta}$  only if

$$u_{ik} = \left( \sum_{j=1}^c \left( \frac{D_{ikA}}{D_{jkA}} \right)^{\frac{2}{(m-1)}} \right)^{-1} \quad 1 \leq i \leq c, 1 \leq k \leq n \quad (18a)$$

$$t_{ik} = \left( \sum_{j=1}^n \left( \frac{D_{ikA}}{D_{ijA}} \right)^{\frac{2}{(\eta-1)}} \right)^{-1} \quad 1 \leq i \leq c; 1 \leq k \leq n \quad \text{và} \quad (18b)$$

$$v_i = \frac{\sum_{k=1}^n (u_{ik}^m + t_{ik}^\eta) x_k}{\sum_{k=1}^n (u_{ik}^m + t_{ik}^\eta)} \quad 1 \leq i \leq c \quad (18c)$$

**Chứng minh:** Các biểu thức trong (18a) và (18b) đều có dạng hàm trong (3a) và thu được chính xác theo cùng một cách. Hai phương trình này theo ngay định lý nhân Lagrange. Trên thực tế, (18a) thu được bằng cách giải bài toán rút gọn  $\min_{U_k \in N_{fcn}} \{J_{m,\eta}^k(U_k) = \sum_{i=1}^c (u_{ik}^m + t_{ik}^\eta) D_{ikA}^2\}$  với  $T$  và  $V$  cố định

cho cột thứ  $k$  của  $U$ . Tập hợp mà  $J_{m,\eta}^k$  được tối thiểu hóa là  $N_{fcn}$ , tập hợp tất cả các vectơ nhãn mờ trong  $\mathbb{R}^c$ . Công thức (18b) thu được theo cùng một cách chính xác bằng cách giải nửa bài toán còn lại cho hàng thứ  $i$  của  $T^t$  của  $T, i.e.$ , tức là

$\min_{T^i \in N_{fnc}} \{J_{m,\eta}^i(T^i) = \sum_{k=1}^n (u_{ik}^m + t_{ik}^\eta) D_{ikA}^2\}$ . Tập trên đó  $J_{m,\eta}^i$  được tối thiểu hóa là  $N_{fnc}$ , tập hợp tất cả các vectơ nhãn mờ trong  $\mathbb{R}^n$ . Cả hai phân tích của  $J_{m,\eta}$  đều có thể thực hiện được vì  $J_{m,\eta}$  là tổng các số hạng không âm, do đó tổng các giá trị tối thiểu là giá trị tối thiểu của các tổng.

**Proof:** The expressions in (18a) and (18b) both have the functional form in (3a) and are obtained exactly in the same manner. These two equations follow immediately with the Lagrange multiplier theorem. Actually, (18a) is obtained by solving the reduced problem  $\min_{U_k \in N_{fcn}} \{J_{m,\eta}^k(U_k) = \sum_{i=1}^c (u_{ik}^m + t_{ik}^\eta) D_{ikA}^2\}$  with  $T$  and

$V$  fixed for the  $k$ -th column  $U_k$  of  $U$ . The set over which  $J_{m,\eta}^k$  is minimized is  $N_{fcn}$ , the set of all fuzzy label vectors in  $\mathbb{R}^c$ . Formula (18b) is obtained in exactly the same way by solving the other half problem for the  $i$ -th row  $T^i$  of  $T, i.e.$ ,

$\min_{T^i \in N_{fnc}} \{J_{m,\eta}^i(T^i) = \sum_{k=1}^n (u_{ik}^m + t_{ik}^\eta) D_{ikA}^2\}$ . The set over which  $J_{m,\eta}^i$  is minimized is  $N_{fnc}$ , the set of all fuzzy label vectors in  $\mathbb{R}^n$ .

Both decompositions of  $J_{m,\eta}$  are possible because  $J_{m,\eta}$  is a sum of nonnegative terms, so the sum of the minimums is the minimum of the sums.

Nếu một hoặc nhiều  $D_{ikA} = \|x_k - v_i\|_A = 0$ , chúng ta có thể sử dụng các điều kiện kỳ dị thông thường để tiếp tục (18a) và (18b). Nghĩa là, gán trọng số bằng 0 cho tất cả  $u_{ik}$  và  $t_{ik}$  sao cho  $D_{ikA} = \|x_k - v_i\|_A > 0$ , và phân phối các thành viên và khả năng khác 0 tùy ý theo  $\sum_{i=1}^c u_{ik} = 1$  và  $\sum_{k=1}^n t_{ik} = 1$ . Tính đúng đắn của các phép gán này tuân theo chính xác như trong chứng minh của Định lý FCM.

If one or more  $D_{ikA} = \|x_k - v_i\|_A = 0$ , we can use the usual singularity conditions to continue (18a) and (18b). That is, assign zero weights to all  $u_{ik}$  and  $t_{ik}$  such that  $D_{ikA} = \|x_k - v_i\|_A > 0$ , and distribute nonzero memberships and possibilities arbitrarily subject to  $\sum_{i=1}^c u_{ik} = 1$ . The correctness of these assignments follows exactly as it does in the proof of Theorem FCM.

Đối với (18c), bài toán rút gọn  $\min_{V \in \mathbb{R}^{cp}} \{\hat{J}_{m,\eta}(V) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik}^m + t_{ik}^\eta) D_{ikA}^2\}$  có thể được giải bằng cách cố định  $U$  và  $T$ , và đưa gradient của  $\hat{J}_{m,\eta}$  về 0 đối với mỗi  $v_i$ . Điều này dẫn đến (18c), miễn là  $D_{ikA} = \|x_k - v_i\|_A$  là một chuẩn tích vô hướng, và do đó, có thể phân biệt được trong mỗi  $v_i$ .

For (18c), the reduced problem  $\min_{V \in \mathbb{R}^{cp}} \{\hat{J}_{m,\eta}(V) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik}^m + t_{ik}^\eta) D_{ikA}^2\}$  can be VERCP solved by fixing  $U$  and  $T$ , and zeroing the gradient of  $\hat{J}_{m,\eta}$  with respect to each  $v_i$ . This results in (18c), as long as  $D_{ikA} = \|x_k - v_i\|_A$  is an inner product norm, and hence, differentiable in each  $v_i$ .



Phương trình (18a) có dạng giống hệt với (3a), công thức thành viên FCM. Điều này không có nghĩa là FCM và FPCM sẽ tạo ra cùng một giá trị thành viên, ngay cả khi cả hai thuật toán đều được bắt đầu với cùng một khởi tạo. Tại sao? Bởi vì (18c) khác với (3b); trong FPCM,  $U$  và  $T$  đều được sử dụng để cập nhật  $V$ , do đó ước tính tiếp theo của  $U$  sẽ khác với ước tính do FCM tạo ra. Tiếp theo, chúng tôi nêu mà không cần chứng minh một số kết quả giới hạn cho FFPCM

$$\lim_{m \rightarrow 1+} \{u_{ik}\} = \begin{cases} 1, & D_{ikA} < D_{jkA} \quad \forall j \neq i, \\ 0, & \text{ngược lại} \end{cases},$$

$$1 \leq i \leq c; 1 \leq k \leq n \quad (19a)$$

$$\lim_{m \rightarrow \infty} \{u_{ik}\} = \frac{1}{c}, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (19b)$$

$$\lim_{\eta \rightarrow 1+} \{t_{ik}\} = \begin{cases} 1, & D_{ikA} < D_{jkA} \quad \forall j \neq k \\ 0, & \text{ngược lại} \end{cases}$$

$$1 \leq i \leq c; 1 \leq k \leq n \quad (19c)$$

$$\lim_{\eta \rightarrow \infty} \{t_{ik}\} = \frac{1}{n}, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (19d)$$

$$\lim_{m \rightarrow \infty, \eta \rightarrow \infty} \{v_i\} = \frac{\sum_{k=1}^n x_k}{n} = \bar{v}, \quad 1 \leq i \leq c. \quad (19e)$$

Equation (18a) is identical in form to (3a), the FCM membership formula. This does not mean that FCM and FPCM will generate the same membership values, even if both algorithms are started with the same initialization. Why? Because (18c) is different from (3b); in FPCM,  $U$  and  $T$  are both used for the update to  $V$ , so the succeeding estimate of  $U$  will differ from the one produced by FCM. We next state without proof some limiting results for FPCM

$$\lim_{m \rightarrow 1+} \{u_{ik}\} = \begin{cases} 1, & D_{ikA} < D_{jkA} \quad \forall j \neq i, \\ 0, & \text{otherwise} \end{cases},$$

$$1 \leq i \leq c; 1 \leq k \leq n \quad (19a)$$

$$\lim_{m \rightarrow \infty} \{u_{ik}\} = \frac{1}{c}, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (19b)$$

$$\lim_{\eta \rightarrow 1+} \{t_{ik}\} = \begin{cases} 1, & D_{ikA} < D_{jkA} \quad \forall j \neq k \\ 0, & \text{otherwise} \end{cases}$$

$$1 \leq i \leq c; 1 \leq k \leq n \quad (19c)$$

$$\lim_{\eta \rightarrow \infty} \{t_{ik}\} = \frac{1}{n}, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (19d)$$

$$\lim_{m \rightarrow \infty, \eta \rightarrow \infty} \{v_i\} = \frac{\sum_{k=1}^n x_k}{n} = \bar{v}, \quad 1 \leq i \leq c. \quad (19e)$$

FPCM có cùng loại kỳ dị như FCM. FPCM không gặp phải vấn đề nhạy cảm mà PCM dường như thể hiện [do kết quả giới hạn trong (12)]. Thật không may, khi số  $n$  điểm dữ liệu lớn, các giá trị điển hình được tính theo (18b) sẽ rất nhỏ. Do đó, sau khi thuật toán FPCM-AO để xấp xỉ các giải pháp cho (15) dựa trên phép lặp qua (18) kết thúc, các giá trị điển hình có thể cần được mở rộng quy mô. Về mặt khái niệm, điều này không khác gì việc mở rộng quy mô các điển hình đối với  $\gamma_i$  như được thực hiện trong PCM. Trong PCM,  $\gamma_i$  được sử dụng để mở rộng quy mô  $t_{ik}$  sao cho tại  $D_{ikA} = \sqrt{\gamma_i}$ , quy mô là 0.5. Trong khi việc mở rộng quy mô dường như "giải quyết" vấn đề giá trị nhỏ (do ràng buộc tổng hàng trên  $T$  gây ra), các giá trị được mở rộng quy mô không sở hữu bất kỳ thông tin bổ sung nào về các điểm trong dữ liệu. Do đó, việc mở rộng quy mô  $\{t_{ik}\}$  là một bản sửa lỗi nhân tạo cho một khiếm khuyết toán học của FPCM. Chúng ta có thể tránh bước điều chỉnh tỷ lệ và làm cho FPCM hữu ích hơn cho các tập dữ liệu lớn với mô hình PFCM mới.

FPCM has the same type of singularity as FCM. FPCM does not suffer from the sensitivity problem that PCM seems to exhibit

[because of the limit result in (12)]. Unfortunately, when the number  $n$  of data points is large, the typicality values computed by (18b) will be very small. Thus, after the FPCM-AO algorithm for approximating solutions to (15) based on iteration through (18) terminates, the typicality values may need to be scaled up. Conceptually, this is no different than scaling typicalities with respect to  $\gamma_i$  as is done in PCM. In PCM,  $\gamma_i$  is used to scale  $t_{ik}$  such that at  $D_{ikA} = \sqrt{\gamma_i}$ , the typicality is 0.5. While scaling seems to "solve" the small value problem (which is caused by the row sum constraint on  $T$ ), the scaled values do not possess any additional information about points in the data. Thus scaling the  $\{t_{ik}\}$  is an artificial fix for a mathematical defect of FPCM. We can avoid the scaling step and make FPCM more useful for large data sets, with the new PFCM model.

## V. MỘT PFCM MỚI

### V. A NEW PFCM

Vấn đề rõ ràng của FPCM là nó áp đặt một ràng buộc lên các giá trị điển hình (tổng các giá trị điển hình trên tất cả các điểm dữ liệu cho một cụm cụ thể là 1). Chúng tôi nối lỏng ràng buộc (tổng hàng = 1) trên các giá trị điển hình nhưng vẫn giữ nguyên ràng buộc cột trên các giá trị thành viên. Điều này dẫn đến vấn đề tối ưu hóa sau:

$$\min_{(U,T,V)} \{J_{m,\eta}(U,T,V;X) = \sum_{k=1}^n \sum_{i=1}^c (au_{ik}^m + bt_{ik}^\eta) \|x_k - v_i\|_A^2 + \sum_{i=1}^c \gamma_i \sum_{k=1}^n (1 - t_{ik})^\eta\} \quad (20)$$

tùy thuộc vào các ràng buộc  $\sum_{i=1}^c u_{ik} = 1 \quad \forall k$ , và  $0 \leq u_{ik}, t_{ik} \leq 1$ . Tại đây,  $a > 0, b > 0, m > 1$  và  $\eta > 1$ . Trong (20),  $\gamma_i > 0$  là các hằng số do người dùng định nghĩa. Các hằng số  $a$  và  $b$  định nghĩa tầm quan trọng tương đối của các giá trị thành viên mờ và tính điển hình trong hàm mục tiêu. Lưu ý rằng, trong (20)  $u_{ik}$  có cùng ý nghĩa về thành viên như trong FCM. Tương tự,  $t_{ik}$  có cùng cách diễn giải về tính điển hình như trong PCM. Tại thời điểm này, một câu hỏi tự nhiên nảy sinh: chúng ta có nên giới hạn mình ở  $a + b = 1$  không? Khi làm như vậy, chúng ta có thể loại bỏ một tham số, nhưng chúng ta có thể gán tầm quan trọng tương đối khác nhau cho  $u_{ik}$  và  $t_{ik}$ . Tuy nhiên, điều này có tác dụng không mong muốn. Nếu chúng ta tăng tầm quan trọng (trọng số) của thành viên thì điều đó buộc chúng ta phải giảm trọng số của tính điển hình theo cùng một lượng. Điều này quá hạn chế. Ngoài ra, sau này chúng ta sẽ thấy rằng các giá trị tính điển hình tối ưu phụ thuộc vào độ lớn của  $b$ . Vì vậy, bằng cách ràng buộc  $a + b = 1$ , chúng ta mất đi tính linh hoạt của mô hình.

The apparent problem of FPCM is that it imposes a constraint on the typicality values (sum of the typicalities over all data points to a particular cluster is 1). We relax the constraint (row sum = 1) on the typicality values but retain the column constraint on the membership values. This leads to the following optimization problem:

$$\min_{(U,T,V)} \{J_{m,\eta}(U,T,V;X) = \sum_{k=1}^n \sum_{i=1}^c (au_{ik}^m + bt_{ik}^\eta) \|x_k - v_i\|_A^2 + \sum_{i=1}^c \gamma_i \sum_{k=1}^n (1 - t_{ik})^\eta\} \quad (20)$$

subject to the constraints  $\sum_{i=1}^c u_{ik} = 1 \quad \forall k$ , and  $0 \leq u_{ik}, t_{ik} \leq 1$ . Here,  $a > 0, b > 0, m > 1$  and  $\eta > 1$ . In (20), the  $\gamma_i > 0$  are user defined constants. The constants  $a$  and  $b$  define the relative importance of fuzzy membership and typicality values in the objective function. Note that, in (20)  $u_{ik}$  has the same meaning of membership as that in FCM. Similarly,  $t_{ik}$  has the same interpretation of typicality as in PCM. At this point, a natural question comes: should we constrain ourselves to  $a + b = 1$ ? Doing so, we can eliminate one parameter, yet we can assign different relative importance to  $u_{ik}$  and  $t_{ik}$ . However, this has an undesirable effect. If we increase the importance (weight) of membership then that necessarily forces us to reduce the importance of typicality by

the same amount. This is too restrictive. Also, we will see later that the optimal typicality values depend on the magnitude of  $b$ . So by constraining  $a + b = 1$ , we lose modeling flexibility.

Nếu  $b = 0$ , và  $\gamma_i = 0$  cho mọi  $i$ , thì (20) được rút gọn thành bài toán tối ưu hóa FCM trong (2); trong khi  $a = 0$  chuyển đổi nó thành mô hình PCM thông thường trong (8). Sau đó, chúng ta sẽ thấy rằng khi  $b = 0$ , ngay cả khi chúng ta không đặt  $\gamma_i = 0$  cho mọi  $i$ , (20) ngầm trở nên tương đương với mô hình FCM. Giống như FPCM, trong các điều kiện thông thường đặt vào các bài toán tối ưu hóa c-means, chúng ta có được các điều kiện cần thiết bậc nhất cho cực trị của  $J_{m,\gamma}$ .

If  $b = 0$ , and  $\gamma_i = 0$  for all  $i$ , then (20) reduces to the FCM optimization problem in (2); while  $a = 0$  converts it to the usual PCM model in (8). Later, we will see that when  $b = 0$ , even if we do not set  $\gamma_i = 0$  for all  $i$ , (20) implicitly becomes equivalent to the FCM model. Like FPCM, under the usual conditions placed on c-means optimization problems, we get the first-order necessary conditions for extrema of  $J_{m,\gamma}$ .

**Định lý PFCM:** Nếu  $D_{ikA} = \|x_k - v_i\| > 0$  đối với mọi  $i$  và  $k, m, \eta > 1$  và  $X$  chứa ít nhất  $c$  điểm dữ liệu riêng biệt, thì  $(U, T, V) \in M_{fcn} \times M_{pcn} \times \mathbb{R}^P$  chỉ có thể tối thiểu  $J_{m,\eta}$  nếu

$$u_{ik} = \left( \sum_{j=1}^c \left( \frac{D_{ikA}}{D_{jkA}} \right)^{\frac{2}{m-1}} \right)^{-1} \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (21)$$

$$t_{ik} = \frac{1}{1 + \left( \frac{b}{\gamma_i} D_{ikA}^2 \right)^{\frac{1}{n-1}}} \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (22)$$

$$v_i = \frac{\sum_{k=1}^n (a u_{ik}^m + b t_{ik}^\eta) x_k}{\sum_{k=1}^n (a u_{ik}^m + b t_{ik}^\eta)} \quad 1 \leq i \leq c \quad (23)$$

**Theorem PFCM:** If  $D_{ikA} = \|x_k - v_i\| > 0$  for all  $i$  and  $k, m, \eta > 1$  and  $X$  contains at least  $c$  distinct data points, then  $(U, T, V) \in M_{fcn} \times M_{pcn} \times \mathbb{R}^P$  may minimize  $J_{m,\eta}$  only if

$$u_{ik} = \left( \sum_{j=1}^c \left( \frac{D_{ikA}}{D_{jkA}} \right)^{\frac{2}{m-1}} \right)^{-1} \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (21)$$

$$t_{ik} = \frac{1}{1 + \left( \frac{b}{\gamma_i} D_{ikA}^2 \right)^{\frac{1}{n-1}}} \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (22)$$

$$v_i = \frac{\sum_{k=1}^n (a u_{ik}^m + b t_{ik}^\eta) x_k}{\sum_{k=1}^n (a u_{ik}^m + b t_{ik}^\eta)} \quad 1 \leq i \leq c \quad (23)$$

**Chứng minh:** Các phương trình (21) và (22) tuân theo ngay định lý nhân Lagrange. Phương trình (21) thu được bằng cách giải bài toán rút gọn

$$\min_{U_k \in N_{fcn}} \{J_{m,\eta}^k(U_k) = \sum_{i=1}^c (a u_{ik}^m + b t_{ik}^\eta) D_{ikA}^2\} \text{ với } T \text{ và } V \text{ cố định}$$

cho cột thứ  $k$  của  $U$ . Hàm  $J_{m,\eta}^k$  được tối thiểu hóa trên  $N_{fcn}$ . Biểu thức thành viên không phụ thuộc vào hằng số  $a$ . Ảnh hưởng của  $a$  đến các thành viên đến từ các tâm cụm (23). Phương trình (22) thu được khi giải bài toán  $\min \{J_{m,\eta}^{ik} = (a u_{ik}^m + b t_{ik}^\eta) D_{ikA}^2 + \gamma_i (1 - t_{ik})\}$ . Hằng số  $b$  có ảnh hưởng trực tiếp đến các giá trị điển hình. Các phân tích này của  $J_{m,\eta}$  là có thể vì  $J_{m,\eta}$  là tổng các số hạng không âm, do đó tổng các giá trị cực tiểu là giá trị cực tiểu của các tổng.

**Proof:** Equations (21) and (22) follow immediately with the Lagrange multiplier theorem. Equation (21) is obtained by solving the reduced problem

$$\min_{U_k \in N_{fcn}} \{J_{m,\eta}^k(U_k) = \sum_{i=1}^c (a u_{ik}^m + b t_{ik}^\eta) D_{ikA}^2\} \text{ with } T \text{ and } V$$

fixed for the  $k$ th column  $U_k$  of  $U$ . The function  $J_{m,\eta}^k$  is minimized over  $N_{fcn}$ . The membership expression is independent of the constant  $a$ . The influence of  $a$  on the memberships comes via the centroids (23). Equation (22) is obtained solving the problem  $\min \{J_{m,\eta}^{ik} = (a u_{ik}^m + b t_{ik}^\eta) D_{ikA}^2 + \gamma_i (1 - t_{ik})\}$ . The constant  $b$  has

a direct influence on the typicality values. These decompositions of  $J_{m,\eta}$  are possible because  $J_{m,\eta}$  is a sum of nonnegative terms, so the sum of the minimums is the minimum of the sums.

Nếu một hoặc nhiều  $D_{ikA} = \|x_k - v_i\|_A = 0$ , gán các giá trị bằng không cho tất cả  $u_{ik}$  và  $t_{ik}$  sao cho  $D_{ikA} = \|x_k - v_i\|_A > 0$ , và phân phối các thành viên khác không tùy ý theo  $\sum_{i=1}^c u_{ik} = 1$  và gán  $t_{ik} = 1$  cho những  $v_i$  mà  $D_{ikA} = 0$ . Tính đúng đắn của các phép gán này tuân theo chính xác như trong chứng minh của Định lý FCM.

If one or more  $D_{ikA} = \|x_k - v_i\|_A = 0$ , assign zero values to all  $u_{ik}$  and  $t_{ik}$  such that  $D_{ikA} = \|x_k - v_i\|_A > 0$ , and distribute nonzero memberships arbitrarily subject to  $\sum_{i=1}^c u_{ik} = 1$  and C assign  $t_{ik} = 1$  to those  $v_i$  for which  $D_{ikA} = 0$ . The correctness of these assignments follows exactly as it does in the proof of Theorem FCM.

Cuối cùng, vấn đề được giảm tải

$$\min_{V \in \mathbb{R}^{cP}} \{J_{m,\eta}(V) = \sum_{i=1}^c \sum_{k=1}^n (a u_{ik}^m + b t_{ik}^\eta) D_{ikA}^2\}$$

được giải quyết bằng cách cố định  $U$  và  $T$ , và đưa độ giảm của  $J_{m,\eta}(V)$  về 0 đối với mỗi  $v_i$ . Điều này dẫn đến (23), miễn là  $D_{ikA} = \|x_k - v_i\|_A$  là một chuẩn cảm ứng tích trong.

Finally, the reduced problem

$$\min_{V \in \mathbb{R}^{cP}} \{J_{m,\eta}(V) = \sum_{i=1}^c \sum_{k=1}^n (a u_{ik}^m + b t_{ik}^\eta) D_{ikA}^2\}$$

is solved by fixing  $U$  and  $T$ , and zeroing the gradient of  $J_{m,\eta}(V)$  with respect to each  $v_i$ . This results in (23), as long as  $D_{ikA} = \|x_k - v_i\|_A$  is an inner product induced norm.

Chúng tôi nêu một số tính chất thú vị của PFCM; xem (24a)-(g), như được hiển thị ở cuối trang tiếp theo. Tính chất P6 cho thấy PFCM hoạt động giống như FCM khi các số mũ tăng không giới hạn. Nghĩa là, bất kể giá trị của các hằng số  $a$  và  $b$ , tất cả các tâm cụm  $c$  đều tiếp cận giá trị trung bình tổng thể (lớn) khi  $m \rightarrow \infty$  và  $\eta \rightarrow \infty$ . Phương trình (23) cho thấy nếu chúng ta sử dụng giá trị  $b$  cao so với  $a$ , thì các tâm sẽ bị ảnh hưởng nhiều hơn bởi các giá trị điển hình so với các giá trị thành viên. Mặt khác, nếu chúng ta sử dụng giá trị  $a$  cao hơn thì các tâm cụm sẽ bị ảnh hưởng nhiều hơn bởi các giá trị thành viên. Do đó, để giảm tác động của các giá trị ngoại lai, chúng ta nên sử dụng giá trị  $b$  lớn hơn  $a$ . Các hiệu ứng tương tự cũng có thể thu được bằng cách kiểm soát sự lựa chọn của  $m$  và  $\eta$ . Ví dụ, nếu chúng ta sử dụng giá trị  $m$  lớn và giá trị  $n$  nhỏ hơn, thì tác động của các giá trị ngoại lai lên các tâm cụm sẽ giảm đi. Tuy nhiên, giá trị  $m$  rất lớn sẽ làm giảm hiệu ứng của các thành viên trên các nguyên mẫu và mô hình sẽ hoạt động giống mô hình PCM hơn, kết quả từ (24b). Thuật toán PFCM cũng được đưa vào Phụ lục.

We state some interesting properties of PFCM; see (24a)-(g), as shown at the bottom of the next page. Property P6 shows that PFCM behaves like FCM as the exponents grow without bound. That is, irrespective of the values of the constants  $a$  and  $b$ , all  $c$  centroids approach the overall (grand) mean as  $m \rightarrow \infty$  and  $\eta \rightarrow \infty$ . Equation (23) shows that if we use a high value of  $b$  compared to  $a$ , then the centroids will be more influenced by the typicality values than the membership values. On the other hand, if we use a higher value of  $a$  then the centroids will be more influenced by the membership values. Thus, to reduce the effect of outliers, we should use a bigger value for  $b$  than  $a$ . Similar effects can also be obtained by controlling the choice of  $m$  and  $\eta$ . For example, if we use a large value of  $m$  and a smaller value for  $\eta$ , then the effect of outliers on the centroids will be reduced. However, a very large value of  $m$  will

reduce the effect of memberships on the prototypes and the model will behave more like the PCM model, resulting from (24b). The PFCM algorithm is also included in Appendix.

## VI. VÍ DỤ SỐ LIỆU VI. NUMERICAL EXAMPLES

Chúng tôi so sánh nhiều khía cạnh khác nhau của FCM, PCM và PFCM với năm tập dữ liệu:  $X_{10}$ ,  $X_{12}$ ,  $X_{400}$ ,  $X_{550}$  và IRIS

We compare various aspects of FCM, PCM, and PFCM with five data sets:  $X_{10}$ ,  $X_{12}$ ,  $X_{400}$ ,  $X_{550}$  and IRIS.

Đối với tất cả các tập dữ liệu, chúng tôi sử dụng các giao thức tính toán sau:  $\epsilon = 0.000001$ ,  $T_{max}$  = số lần lặp tối đa = 100 và  $\| \cdot \|_A$  là chuẩn Euclid. Số cụm  $c$  là 3 đối với IRIS và đối với tất cả các tập dữ liệu khác, nó được cố định ở  $c = 2$ . Đối với cả PCM và PFCM, trước tiên chúng tôi chạy FCM để hoàn tất và sử dụng (11) với  $K = 1$  để tìm giá trị của  $\gamma_i$ . Tất cả các thử nghiệm đều kết thúc với tiêu chí hội tụ sau một vài lần lặp.

For all data sets we use the following Computational protocols:  $\epsilon = 0.000001$ ,  $T_{max}$  = maximum number of iterations = 100, and  $\| \cdot \|_A$  is the Euclidean norm. The number of clusters  $c$  is 3 for IRIS and for all other data sets it is fixed at  $c = 2$ . For both PCM and PFCM we first run FCM to termination and use (11) with  $K = 1$  to find the values of  $\gamma_i$ . All trials terminated with the convergence criteria after a few iterations.

	FCM : m=2		PCM : $\eta=2$		PFCM : a=1, b=1, m=2, $\eta=2$	
$X_{10}$	-3.36	3.36	-3.24	3.24	-3.34	3.34
	0.00	0.00	0.00	0.00	0.00	0.00
$X_{12}$	-2.99	2.99	-2.15	2.15	-2.84	2.84
	0.54	0.54	0.02	0.02	0.36	0.36

BẢNG II MẪU THIẾT BỊ ĐẦU CUỐI ĐƯỢC SẢN XUẤT BỞI FCM, PCM VÀ PFCM TRÊN  $X_{10}$  VÀ  $X_{12}$   
TABLE II TERMINAL PROTOTYPES PRODUCED BY FCM, PCM, AND PFCM ON  $X_{10}$  AND  $X_{12}$

Ví dụ 2:

Bộ dữ liệu:  $X_{10}$ ,  $X_{12}$ .

Thuật toán: FCM, PCM, PFCM.

$$\text{Khởi tạo: } V_0 = \begin{bmatrix} -3.34 & 1.67 \\ 1.67 & 0.00 \end{bmatrix} \quad (25)$$

Khởi tạo trong (25) thu được bằng cách chọn ngẫu nhiên hai điểm dữ liệu từ  $X_{10}$  (mỗi cột của  $V_0$  là một điểm dữ liệu). Bảng II hiển thị các trọng tâm do FCM, PCM và PFCM tạo ra. Đối với  $X_{10}$ , cả ba thuật toán đều tạo ra các nguyên mẫu tốt. Tuy nhiên, khi xem xét kỹ hơn, ta thấy rằng PFCM tạo ra các nguyên mẫu khớp chính xác (lên đến làm tròn) với các trọng tâm lý tưởng (thực)

$$V_{ideal} = \begin{bmatrix} -3.34 & 3.34 \\ 0 & 0 \end{bmatrix} \text{ sẽ không thu được giá trị ngoại lai nào.}$$

Example 2: Data sets:  $X_{10}$ ,  $X_{12}$ . Algorithm: FCM, PCM, PFCM.

$$\text{Initialization: } V_0 = \begin{bmatrix} -3.34 & 1.67 \\ 1.67 & 0.00 \end{bmatrix} \quad (25)$$

The initialization in (25) is obtained by randomly selecting two data points from  $X_{10}$  (each column of  $V_0$  is a data point). Table II shows the centroids produced by FCM, PCM and PFCM. For  $V_0$  all three algorithms produce good prototypes. However, a closer look reveals that PFCM produces prototypes that match exactly (up to

rounding) with the ideal (true) centroids  $V_{ideal} = \begin{bmatrix} -3.34 & 3.34 \\ 0 & 0 \end{bmatrix}$  that would be obtained in absence of outliers.

Nếu  $X_{11}$  và  $X_{12}$  bị "bỏ qua", chúng ta hy vọng tìm thấy tâm cụm thực sự của  $X_{10}$  khi  $X_{12}$  được xử lý. Từ Bảng II, chúng ta thấy rằng  $V_{FCM}^{12} = \begin{bmatrix} -2.99 & 2.99 \\ 0.54 & 0.54 \end{bmatrix}$ ,  $V_{PCM}^{12} = \begin{bmatrix} -2.15 & 2.15 \\ 0.02 & 0.02 \end{bmatrix}$  và  $V_{PFCM}^{12} = \begin{bmatrix} -2.84 & 2.84 \\ 0.36 & 0.36 \end{bmatrix}$ . Để đánh giá sơ bộ cách mỗi phương pháp tính đến giá trị nội tại và giá trị ngoại lai, chúng ta tính  $E_X = \|V_{ideal} - V_X^{12}\|^2$ , trong đó  $X$  là FCM/PCM/PFCM:  $E_{FCM} = 0.4141$ ,  $= 1.3161$  và  $E_{PFCM} = 0.3796$ . Mặc dù PCM gán các giá trị đặc trưng khác nhau cho  $X_{11}$  và  $X_{12}$ , nhưng các tâm do PCM tạo ra không tốt lắm so với các tâm do FCM và PFCM tạo ra. Với một khởi tạo và lựa chọn khác nhau, PCM sẽ tạo ra kết quả tốt hơn.

If  $X_{11}$  and  $X_{12}$  are to be "ignored," we hope to find the true centroids of  $X_{10}$  when  $X_{12}$  is processed. From Table II, we see that  $V_{FCM}^{12} = \begin{bmatrix} -2.99 & 2.99 \\ 0.54 & 0.54 \end{bmatrix}$ ,  $V_{PCM}^{12} = \begin{bmatrix} -2.15 & 2.15 \\ 0.02 & 0.02 \end{bmatrix}$  and  $V_{PFCM}^{12} = \begin{bmatrix} -2.84 & 2.84 \\ 0.36 & 0.36 \end{bmatrix}$ . To make a crude assessment how each method has accounted for the inlier and outlier, we compute  $E_X = \|V_{ideal} - V_X^{12}\|^2$ , where  $X$  is FCM/PCM/PFCM:  $E_{FCM} = 0.4141$ ,  $= 1.3161$  and  $E_{PFCM} = 0.3796$ . Although PCM does assign different typicality values to  $X_{11}$  and  $X_{12}$ , the centroids produced by PCM are not very good compared to those generated by FCM and PFCM. With a different initialization and different choice of , PCM should produce better results.

Bảng III thể hiện  $U$  và  $T$  đầu cuối từ FCM, PCM và PFCM với  $a = 1.0$ ,  $b = 1.0$ ,  $m = 2.0$  và  $\eta = 2.0$  (tương ứng với Bảng II). Bảng III cho thấy cả UPFCM và UFCM đều dẫn đến cùng một phân vùng thành viên tối đa. Thứ tự tương đối của các điểm (theo giá trị thành viên) cũng vẫn giữ nguyên.

Table III exhibits the terminal U and T from FCM, PCM and PFCM with  $a = 1.0$ ,  $b = 1.0$ ,  $m = 2.0$  and  $\eta = 2.0$  (corresponding to Table II). Table III shows that both UPFCM and UFCM result in the same max-membership hard partition. The relative ordering of points (in terms of membership values) also remains the same.

PFCM cung cấp mô tả dữ liệu nhiều thông tin hơn FCM, vì nó cung cấp thông tin thành viên gần như giống nhau nhưng cũng cho thấy, thông qua các đặc điểm, ví dụ, rằng  $x_{12}$  không điển hình hơn nhiều so với  $x_{11}$  đối với cả hai cụm. PFCM gán đặc điểm điển hình cao nhất trong cụm bên trái cho  $x_3$ . Điều này rất hợp lý vì  $x_3$  nằm ở trung tâm của cụm bên trái. Tương tự,  $x_8$  là đặc điểm điển hình nhất của cụm bên phải. Ở đây,  $x_{12}$  ít điển hình nhất đối với cả hai cụm, cho thấy rằng nó là một giá trị ngoại lệ. Đặc điểm điển hình cho  $x_{12}$  trong cả hai cụm nhỏ hơn một bậc độ lớn so với giá trị đặc điểm điển hình cho  $x_{11}$ . Điều này cho phép chúng tôi giảm bớt các giá trị ngoại lai khỏi dữ liệu để giảm tác động của nhiễu (như PCM đã làm trong [2]). Đối với khởi tạo này, PCM gán các giá trị điển hình cao nhất cho  $x_5$  và  $x_6$  vì các trọng tâm PCM nằm tại  $V_{PCM}^{12} = \begin{bmatrix} -2.15 & 2.15 \\ 0.02 & 0.02 \end{bmatrix}$ . Cả  $T_{PCM}$  và  $T_{PFCM}$  đều dẫn đến cùng một phân vùng (được làm cứng). Mặc dù đối với  $x_{12}$ , cả PFCM và PCM đều gán gần như cùng các giá trị điển hình, nhưng đối với  $x_{11}$ , các giá trị điển hình do PFCM gán nhỏ hơn đáng kể so với các giá trị do PCM gán. Do đó, các nguyên mẫu PFCM ít bị ảnh hưởng bởi các điểm nhiễu hơn.

PFCM provides a more informative description of the data than FCM, since it provides roughly the same membership in- formation

but also shows, via the typicalities, for example, that  $x_{12}$  is much more atypical than  $x_{11}$  for either cluster. PFCM assigns the highest typicality in the left cluster to  $x_3$ . This is very reasonable because  $x_3$  is located at the center of the left cluster. Similarly,  $x_8$  is most typical of the right cluster. Here,  $x_{12}$  is least typical to either cluster indicating that it is an outlier. The typicality for  $x_{12}$  in both clusters is an order of magnitude smaller than the typicality value for  $x_{11}$ . This enables us to prune outliers from the data to reduce the effects of noise (like PCM did in [2]). For this initialization, PCM assigns the highest typicality values to  $x_5$  and  $x_6$  because the PCM centroids are located at  $V_{PCM}^{12} = \begin{bmatrix} -2.15 & 2.15 \\ 0.02 & 0.02 \end{bmatrix}$ . Both TPCM and TPFCM result in the same (hardened) partition. Although for  $x_{12}$  both PFCM and PCM assign almost the same typicality values, for  $x_{11}$  the typicality values assigned by PFCM are significantly smaller than those assigned by PCM. Hence, the PFCM prototypes are less influenced by the noise points.

Ví dụ 3:

Bộ dữ liệu: X10.

Thuật toán: PFCM.

Khởi tạo:  $V_0 = \begin{bmatrix} 0.07 & 0.36 \\ 0.40 & 0.99 \end{bmatrix}$

Bây giờ chúng tôi nghiên cứu tác động của nhiều tham số khác nhau lên hiệu suất của thuật toán PFCM. Khởi tạo từ (7a) không phải là khởi tạo tốt, vì hai vectơ trong  $V_0$  rất gần nhau đối với siêu hộp nhỏ nhất chứa dữ liệu. Tuy nhiên, chúng tôi sử dụng nó để thể hiện "sự mạnh mẽ" của các thuật toán khác nhau trên các khởi tạo. Với khởi tạo (7a), PCM tạo ra các cụm trùng nhau. Krishnapuram và Keller [2], [5] đề xuất sử dụng các trọng tâm FCM cuối cùng để khởi tạo PCM, do đó chúng tôi cũng báo cáo một số kết quả do PCM tạo ra (trong Ví dụ 5) và PFCM khi việc khởi tạo được thực hiện bởi FCM nguyên mẫu.

Example 3:

Data set: X10.

Algorithm: PFCM.

Initialization:  $V_0 = \begin{bmatrix} 0.07 & 0.36 \\ 0.40 & 0.99 \end{bmatrix}$

We now investigate the impact of various parameters on the performance of the PFCM algorithm. The initialization from (7a) is not a good one, because the two vectors in  $V_0$  are very close to each other with respect to the smallest hyperbox containing the data. However, we use it to show the "robustness" of different algorithms on the initializations. With initialization (7a) PCM generates coincident clusters. Krishnapuram and Keller [2], [5] suggest using the terminal FCM centroids to initialize PCM, so we also report a few results generated by PCM (in Example 5) and PFCM when the initialization is done by the FCM prototypes.

$$P1: \lim_{m \rightarrow 1+} \{u_{ik}\} = \begin{cases} 1, & D_{ikA} < D_{jkA} \forall j \neq i \\ 0, & \text{ngược lại} \end{cases}$$

$$, 1 \leq i \leq c; 1 \leq k \leq n \quad (24)$$

$$P2: \lim_{m \rightarrow \infty} \{u_{ik}\} = \frac{1}{c}, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (24b)$$

$$P3: \lim_{\eta \rightarrow \infty} \{t_{ik}\} = 0.5, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (24c)$$

$$P4: \lim_{\eta \rightarrow 1+} \{t_{ik}\} = \begin{cases} 11, & \text{nếu } bD_{ikA}^2 < \gamma_i \\ \frac{1}{2}, & \text{nếu } bD_{ikA}^2 = \gamma_i, 1 \leq i \leq c; 1 \leq k \leq n \\ 0, & \text{nếu } bD_{ikA}^2 > \gamma_i \end{cases}$$

$$P5: t_{ik} = \frac{1}{2} \text{ nếu } bD_{ikA}^2 = \gamma_i, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (24e)$$

$$P6: \lim_{m, n \rightarrow \infty} \{v_i\} = \sum_{k=1}^n x_k = \bar{v}, \quad 1 \leq i \leq c \quad (24f)$$

$$P7: \text{ nếu } b = 0, \text{ thì } t_{ik} = 1, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (24g)$$

$$P1: \lim_{m \rightarrow 1+} \{u_{ik}\} = \begin{cases} 1, & D_{ikA} < D_{jkA} \forall j \neq i \\ 0, & \text{otherwise} \end{cases}$$

$$, 1 \leq i \leq c; 1 \leq k \leq n \quad (24)$$

$$P2: \lim_{m \rightarrow \infty} \{u_{ik}\} = \frac{1}{c}, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (24b)$$

$$P3: \lim_{\eta \rightarrow \infty} \{t_{ik}\} = 0.5, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (24c)$$

$$P4: \lim_{\eta \rightarrow 1+} \{t_{ik}\} = \begin{cases} 11, & \text{if } bD_{ikA}^2 < \gamma_i \\ \frac{1}{2}, & \text{if } bD_{ikA}^2 = \gamma_i, 1 \leq i \leq c; 1 \leq k \leq n \\ 0, & \text{if } bD_{ikA}^2 > \gamma_i \end{cases} \quad (24d)$$

$$P5: t_{ik} = \frac{1}{2} \text{ nếu } bD_{ikA}^2 = \gamma_i, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (24e)$$

$$P6: \lim_{m, n \rightarrow \infty} \{v_i\} = \sum_{k=1}^n x_k = \bar{v}, \quad 1 \leq i \leq c \quad (24f)$$

$$P7: \text{ if } b = 0, \text{ then } t_{ik} = 1, \quad 1 \leq i \leq c; 1 \leq k \leq n \quad (24g)$$

Point	$U_{FCM} \in M_{fc}$ (from FCM)		$U_{PFCM} \in M_{fcn}$ (from PFCM)		$T_{PFCM} \in M_{pcn}$ (from PFCM)		$T_{PCM} \in M_{pcn}$ (from PCM)	
1	0.936	0.064	0.928	0.072	0.621	0.113	0.492	0.134
2	0.967	0.033	0.953	0.047	0.801	0.165	0.655	0.194
3	0.99	0.01	0.990	0.01	0.953	0.171	0.847	0.208
4	0.899	0.101	0.906	0.094	0.642	0.157	0.648	0.193
5	0.916	0.084	0.932	0.068	0.840	0.278	0.972	0.351
6	0.084	0.916	0.068	0.932	0.278	0.840	0.351	0.972
7	0.033	0.967	0.047	0.953	0.165	0.801	0.194	0.655
8	0.01	0.99	0.01	0.990	0.171	0.953	0.208	0.847
9	0.101	0.899	0.094	0.906	0.157	0.642	0.193	0.648
10	0.064	0.936	0.072	0.928	0.113	0.621	0.134	0.492
11	0.500	0.500	0.500	0.500	0.490	0.490	0.631	0.631
12	0.500	0.500	0.500	0.500	0.072	0.072	0.070	0.070

BẢNG III U VÀ T THU ĐƯỢC TỪ FCM ( $m = 2$ ), PCM ( $\eta = 2$ ), VÀ PFCM ( $a = 1, b = 1, m = 2, \eta = 2$ ) CHO  $X_{12}$  VỚI KHỞI TẠO ĐƯỢC THỰC HIỆN NHƯ TRONG (25)

TABLE III TERMINAL U AND T FROM FCM ( $m = 2$ ), PCM ( $\eta = 2$ ), AND PFCM ( $a = 1, b = 1, m = 2, \eta = 2$ ) FOR  $X_{12}$  WITH INITIALIZATION PERFORMED AS IN (25)

	a	b	m	$\eta$	$v_1$	$v_2$
<u>1</u>	1	1	2	2	-3.34 0.00	3.34 0.00
<u>2</u>	1	3	2	2	-3.35 0.00	3.35 0.00
<u>3</u>	1	6	2	2	-3.35 0.00	3.35 0.00
<u>4</u>	1	7	2	2	-3.35 0.00	3.35 0.00
<u>5</u>	1	1	5	1.5	-3.33 0.00	3.33 0.00
<u>6</u>	1	1	7	1.5	-2.01 0.00	2.01 0.00
<u>7</u>	1	5	5	1.5	-1.76 0.00	1.76 0.00
<u>8</u>	1	5	5	10	-3.32 0.00	3.32 0.00
<u>9</u>	1	1	2	7	-3.35 0.00	3.35 0.00
<u>10</u>	1	4	3	2	-3.34 0.00	3.34 0.00
<u>11</u>	1	0 (=FCM)	5	-----	-3.32 0.00	3.32 0.00
<u>12</u>	0.5	0.5	2	2	-3.297 0.00	3.297 0.00
<u>13</u>	0.25	0.75	2	2	-3.289 0.00	3.289 0.00
<u>14</u>	0.1667	0.8333	5	1.5	2.71 0.00	2.71 0.00

BẢNG IV KẾT QUẢ DO PFCM ĐẠT ĐƯỢC CHO CÁC GIÁ TRỊ KHÁC NHAU CỦA THAM SỐ VỚI  $X_{10}$

TABLE IV RESULTS PRODUCED BY PFCM FOR DIFFERENT VALUES OF THE PARAMETERS WITH  $X_{10}$

Bảng IV cho thấy kết quả do PFCM tạo ra trên  $X_{10}$  đối với các giá trị khác nhau của  $a, b, m$  và  $\eta$ . Các tâm cụm tương ứng với lần chạy 1 trong Bảng IV giống hệt (làm tròn) với các tâm cụm PFCM



trong Bảng II, được PFCM tạo ra bằng cách sử dụng một khối tạo khác. Bốn kết quả đầu tiên trong Bảng IV chứng minh rằng nếu chúng ta thay đổi từ 1 đến 7, giữ nguyên tất cả các tham số khác, thì các nguyên mẫu PFCM đầu cuối về cơ bản không thay đổi. Tại sao? Việc tăng  $b$  sẽ gán tầm quan trọng hơn cho tính điển hình, nhưng kết quả không thay đổi nhiều vì  $X_{10}$  có các cụm được phân tách tốt. Tuy nhiên, nếu dữ liệu có các giá trị ngoại lai thì chúng ta sẽ thấy rằng việc gán tầm quan trọng hơn cho tính điển hình bằng cách tăng  $b$  sẽ cải thiện các nguyên mẫu. (Có thể sử dụng thuộc tính này để hiểu rõ hơn về dữ liệu).

Table IV shows the results produced by PFCM on  $X_{10}$  for different values of  $a, b, m$  and  $\eta$ . The centroids corresponding to run 1 in Table IV are identical (up to rounding) to the PFCM centroids in Table II, which are generated by PFCM using a different initialization. The first four results in Table IV demonstrate that if we vary  $b$  from 1 to 7, keeping all other parameters fixed, the terminal PFCM prototypes are essentially unchanged. Why? Increasing  $b$  assigns more importance to typicality, but the results do not change much since  $X_{10}$  has nice well separated clusters. However, if the data have outliers then, we shall see that giving more importance to typicality by increasing  $b$  improves the prototypes. (This property can be used to get a better understanding of the data.)

So sánh 5 với 6, ta thấy rằng việc tăng giá trị của  $m$  từ 5.0 lên 7.0, khi giá trị của  $\eta$  thấp (1.5), làm giảm các nguyên mẫu vì các giá trị thành viên trở nên đồng đều hơn. Nghĩa là, thành phần thành viên của (23) cố gắng đẩy các nguyên mẫu về phía trung bình lớn và điều này làm giảm các nguyên mẫu, khiến tính điển hình trở nên quan trọng hơn đối với việc tính toán các tâm. Chạy 7 cho thấy các tâm rất kém. So sánh lựa chọn tham số trong 5 với các tham số trong 7, ta thấy rằng chỉ có  $b$  thay đổi từ 1.0 đến 5.0. Giá trị cao của  $m$  loại bỏ độ lệch trong các giá trị thành viên (các giá trị thành viên trở nên gần 0.5 cho tất cả các điểm dữ liệu). Do đó, như vừa giải thích, phần FCM của hàm mục tiêu cố gắng di chuyển các tâm về phía vectơ trung bình lớn và giá trị  $\eta$  thấp làm cho các phép gán tính điển hình gần như rõ ràng. Vì vậy, các tâm cụm sẽ trôi về phía trung bình lớn và tính điển hình sẽ cao chỉ đối với một vài điểm gần với tâm. Kiểm tra  $T$  cho thấy chỉ đối với một điểm trong mỗi cụm, tính điển hình gần bằng 1.0, trong khi đối với tất cả các điểm khác trong cùng cụm, các giá trị tính điển hình rất nhỏ. Do đó, kết quả kém. Nếu có một vài điểm nhiễu, kết quả thậm chí có thể tệ hơn tùy thuộc vào vị trí của các nguyên mẫu. Tuy nhiên, nếu chúng ta tăng giá trị của  $\eta$ , chúng ta có thể phục hồi từ tình huống này.

Comparing 5 with 6, we see that increasing the value of  $m$  from 5.0 to 7.0, when the value of  $\eta$  is low (1.5), degrades the prototypes because the membership values become more uniform. That is, the membership component of (23) tries to push the prototypes toward the grand mean, and this degrades the prototypes, making typicality more important for computation of centroids. Run 7 shows very poor centroids. Comparing the choice of parameters in 5 with those in 7, we find that only  $b$  changes from 1.0 to 5.0. The high value of  $m$  removes the skewness in the membership values (membership values become close to 0.5 for all data points). Hence, as just explained, the FCM part of the objective function tries to move the centroids toward the grand mean vector and a low value of  $\eta$  makes typicality assignments almost crisp. So the cluster centers will drift toward the grand mean and typicality will be high for only a few points close to the centroids. Inspection of  $T$  reveals that for just one point in each cluster the typicality is almost 1.0, while for all other points in the same cluster the typicality values are very small. Hence, the poor results. If there are a few noise points, the results could even be worse depending on the locations the prototypes. However, if we increase the value of  $\eta$ , we can recover from this situation.

Các tâm được tạo ra bởi run 8 tốt hơn nhiều so với các tâm trong run 7. Giữa các lần chạy 7 và 8 chỉ có  $\eta$  thay đổi từ 1.5 đến 10.

Đối với lựa chọn này, các giá trị điển hình gần bằng 0.5 cho tất cả các điểm dữ liệu. Do đó, các giá trị điển hình không có nhiều ảnh hưởng đến các nguyên mẫu. Vì vậy, đối với lựa chọn tham số này, PFCM hoạt động giống FCM hơn. Chạy 11 trong Bảng IV tương ứng với FCM với  $m = 5.0$ . Thật thú vị khi thấy rằng chúng ta có cùng (lên đến làm tròn) các tâm cụm như trong 8. Đối với tập dữ liệu này ngay cả với  $m = 2$  và  $\eta = 7$ , chúng ta vẫn có được kết quả khá tốt, hãy xem chạy 9. Đây thực tế là kết quả FCM. Vì  $\eta = 7.0$ , hiệu ứng của tính điển hình giảm mạnh và chúng ta có được kết quả tương tự như FCM. Bảng IV bao gồm các nguyên mẫu PFCM cho một số lựa chọn tham số khác, cũng khá tốt.

The centroids generated by run 8 are much better than those in run 7. Between runs 7 and 8 only  $\eta$  changes from 1.5 to 10. For this choice, the typicality values are close to 0.5 for all data points. Consequently, typicality values do not have much influence on the prototypes. So, for this choice of parameters, PFCM behaves more like FCM. Run 11 in Table IV corresponds to FCM with  $m = 5.0$ . It is interesting to see that we get the same (up to rounding) centroids as in 8. For this data set even with  $m = 2$  and  $\eta = 7$ , we get reasonably good results, see run 9. This is practically the FCM result. Since  $\eta = 7.0$ , the effect of typicality is drastically reduced and we get results similar to the FCM. Table IV includes the PFCM prototypes for a few other choices of parameters, which are also quite good.

Chúng tôi đã thảo luận trước đó rằng việc ràng buộc  $a + b = 1$  là không được như mong đợi sẽ tạo ra các nguyên mẫu tốt. Để minh họa cho điều này, chúng tôi báo cáo thêm ba kết quả nữa. Tỷ lệ  $a/b$  cho lần chạy 12 giống với tỷ lệ của lần chạy 1, tuy nhiên các tâm do lần chạy 12 tạo ra lại khác với các tâm do lần chạy 1 tạo ra, mặc dù tất cả các giao thức khác bao gồm khởi tạo vẫn giữ nguyên như trước. Tương tự, lần chạy 13 có cùng tỷ lệ  $a/b$  như lần chạy 2. Trong trường hợp này, cũng có sự khác biệt đáng chú ý giữa hai tập trọng tâm. Lần chạy 14 với  $a/b = 1/6$  tương ứng với lần chạy 3. Tuy nhiên, lần chạy 14 dẫn đến các cụm trùng nhau. Vì trọng số tuyệt đối ( $a = 0.1667$ ) được gán cho thành phần thành viên của hàm mục tiêu trong (20) khá nhỏ khiến thuật toán hoạt động gần giống như PCM.

We discussed earlier that constraining  $a + b = 1$  is not expected to result in good prototypes. To illustrate this we report three more results. The ratio  $a/b$  for run 12 is the same as that of Run 1, yet the centroids produced by run 12 are different from those produced by run 1, though all other protocols including initialization remain the same as before. Similarly, run 13 has the same ratio of  $a/b$  as that of run 2. In this case too, there is a noticeable difference between the two sets of centroids. Run 14 with  $a/b = 1/6$  corresponds to run 3. However, run 14 results in coincident clusters. Because the absolute weight ( $a = 0.1667$ ) assigned to the membership component of the objective function in (20) is quite small making the algorithm behave almost like PCM.

*Bài học từ Ví dụ 3:* Chúng ta không được sử dụng giá trị  $\eta$  quá lớn, vì như (24) gợi ý,  $\eta$  lớn hơn sẽ làm giảm tác động của tính điển hình. Ngoài ra, có vẻ như  $b$  phải lớn hơn  $a$  để các giá trị điển hình ảnh hưởng đến phép tính của các nguyên mẫu, do đó làm giảm tác động của các giá trị ngoại lai. Nhưng  $b$  không nên "quá lớn", vì khi đó tác động của các giá trị thành viên sẽ bị loại bỏ thực tế.

*Lessons From Example.3:* We must not use a very large value of  $\eta$ , since as (24) suggests, larger  $\eta$ 's reduce the effect of typicality. Also, it seems that  $b$  should be larger than  $a$  so that typicality values influence the computation of prototypes, thereby reducing the effects of outliers. But  $b$  should not be "too large," for then the effect of membership values will be practically eliminated.

	a	b	m	$\eta$	$v_1$	$v_2$
<b>1</b>	1	1	2	2	-2.84 0.36	2.84 0.36
<b>2</b>	1	3	2	2	-3.00 0.33	3.00 0.33
<b>3</b>	1	6	2	2	-3.11 0.54	3.11 0.54
<b>4</b>	1	7	2	2	-3.13 0.26	3.13 0.26
<b>5</b>	1	1	5	1.5	-3.03 0.08	3.03 0.08
<b>6</b>	1	1	7	1.5	-3.03 0.03	3.03 0.03
<b>7</b>	1	5	5	1.5	-0.00 0.05	0.00 0.05
<b>8</b>	1	5	5	10	-2.93 0.47	2.93 0.47
<b>9</b>	1	1	2	7	-2.97 0.54	2.97 0.54
<b>10</b>	1	4	3	2	-3.11 0.21	3.11 0.21
<b>11</b>	1	0 (=FCM)	5	----	-2.96 0.47	2.96 0.47
<b>12</b>	0.5	0.5	2	2	-2.53 0.32	2.53 0.32
<b>13</b>	0.25	0.75	2	2	-2.48 0.20	2.48 0.20
<b>14</b>	0.1667	0.8333	5	1.5	2.82 0.02	2.82 -0.02

BẢNG V KẾT QUẢ DO PFCM ĐẠT ĐƯỢC CHO CÁC GIÁ TRỊ KHÁC NHAU CỦA THAM SỐ VỚI  $X_{10}$   
TABLE V RESULTS PRODUCED BY PFCM FOR DIFFERENT VALUES OF THE PARAMETERS WITH  $X_{10}$

Ví dụ 4:

Bộ dữ liệu:  $X_{12}$ .

Thuật toán: PFCM.

$$\text{Khởi tạo: } V_0 = \begin{bmatrix} 0007 & 0.36 \\ 0.40 & 0.99 \end{bmatrix}.$$

Bảng V mô tả các trọng tâm được tạo ra bởi PFCM trên  $X_{12}$  bằng cách sử dụng các tham số giống như trong Bảng IV. Trong khi chúng ta không thấy bất kỳ thay đổi nào trong các nguyên mẫu khi  $b$  tăng lên trong khi giữ nguyên tất cả các tham số khác cho  $X_{10}$ , các trọng tâm thực sự được cải thiện đối với  $X_{12}$  (xem Chạy **1-4** trong Bảng V). Điều này xảy ra vì các giá trị  $b$  lớn hơn làm tăng tầm quan trọng tương đối của các đặc trưng trong việc xác định các nguyên mẫu. Ngoại trừ lần chạy **7**, PFCM đã tạo ra các nguyên mẫu tốt. Chạy **7** tạo ra các cụm trùng nhau. Tại sao? Chúng tôi đã đề cập trước đó rằng chúng ta có thể chọn các tham số của PFCM để nó hoạt động giống PCM hơn. Nếu  $a$  rất nhỏ và các tâm cụm ban đầu không được đặt đúng vị trí (tách biệt), thì PFCM, giống như PCM, có thể tạo ra các tâm rất gần nhau (tức là các cụm thực tế trùng nhau). Chạy **7** trong Bảng V tương ứng với tình huống này. Để minh họa thêm cho điểm này, chúng tôi đã chạy PFCM bằng cùng một bộ tham số như trong **7**, nhưng được khởi tạo với các trung tâm FCM đầu cuối. Trong trường hợp này, PFCM tạo ra  $V_{PFCM}^{12} = \begin{bmatrix} -3.27 & 3.27 \\ 0.09 & 0.09 \end{bmatrix}$ . Điều này chỉ ra rằng trong lần chạy **7**, chúng tôi đã thu được các cụm trùng nhau vì các nguyên mẫu ban đầu rất gần nhau. Lần chạy **11** cho thấy các nguyên mẫu do FCM tạo ra với  $m = 5, 0(a = 1, b = 0)$ . Trong bối cảnh của Bảng IV, chúng tôi đã lập luận và chứng minh bằng số liệu rằng việc lựa chọn các tham số trong lần chạy **8** sẽ khiến PFCM hoạt động giống như FCM. Điều này cũng đúng đối với  $X_{12}$ , lần chạy **8** và lần chạy **11** tạo ra hầu như cùng một nguyên mẫu. Đối với  $X_{12}$ , giống như  $X_{10}$ , chúng tôi đã thực hiện ba lần chạy(**12-14**) ràng buộc  $a + b = 1$ , nhưng vẫn duy trì cùng tỷ lệ  $a/b$  tương ứng với các lần chạy **1-3**. Các kết luận rút ra từ lần chạy **12-14** trong Bảng IV cũng có thể áp dụng ở đây.

Example 4:

Data set:  $X_{12}$ .

Algorithm: PFCM.

$$\text{Initialization: } V_0 = \begin{bmatrix} 0007 & 0.36 \\ 0.40 & 0.99 \end{bmatrix}.$$

Table V depicts the centroids generated by PFCM on  $X_{12}$  using the same parameters as in Table IV. While we did not see any change in the prototypes when  $b$  was increased keeping all other parameters fixed for  $X_{10}$ , the centroids do improve for  $X_{12}$  (see Runs **1-4** in Table V). This happens because larger values of  $b$  increase the relative importance of the typicalities in determining the prototypes. Except for run 7, PFCM generated good prototypes. Run 7 generates coincident clusters. Why? We mentioned earlier that we can choose the parameters of PFCM so that it behaves more like PCM. If  $a$  is very small, and the initial centroids are not well placed (separated), then PFCM, like PCM, can generate centers that are very close to each other (i.e., practically coincident clusters). Run 7 in Table V corresponds to this situation. To further illustrate this point, we ran PFCM using the same set of parameters as in 7, but initialized with the terminal FCM centers. In this case PFCM generates  $V_{PFCM}^{12} = \begin{bmatrix} -3.27 & 3.27 \\ 0.09 & 0.09 \end{bmatrix}$ . This indicates that in run 7 we obtained coincident clusters because the initial prototypes were very close to each other. Run **11** shows the prototypes produced by FCM with  $m = 5, 0(a = 1, b = 0)$ . In the context of Table IV, we argued and numerically demonstrated that the choice of parameters in run 8 will make PFCM behave like FCM. This is also true for  $X_{12}$ , run **8** and run **11** generate almost the same prototypes. For  $X_{12}$ , like  $X_{10}$ , we made three runs (**12-14**) constraining  $a + b = 1$ , but maintaining the same ratios  $a/b$  corresponding to runs **1-3**. The conclusions drawn from runs **12-14** in Table IV are equally applicable here.

Ví dụ 5:

Bộ dữ liệu:  $X_{12}$

Thuật toán: PCM.

Khởi tạo: Tâm cụm FCM.

Các lần chạy từ **1** đến **5** của Bảng VI cho thấy các tâm cụm PCM cho các giá trị  $\eta$  khác nhau khi thuật toán PCM được khởi tạo với các nguyên mẫu FCM đầu cuối. Với sự gia tăng  $\eta$  vượt quá 2, PCM tạo ra các cụm gần như trùng khớp. Ngoại trừ một điểm, các giá trị điển hình trở nên gần 0.5. Nói cách khác, với giá trị  $\eta$  khá lớn, tính chất giới hạn của tính điển hình trong (14) đã đạt được. Tuy nhiên, khi sử dụng khuôn mẫu PFCM của chúng tôi để hiện thực hóa PCM, chúng tôi có thể khắc phục tình trạng này bằng cách tăng giá trị của  $b$ . Chạy **6** và chạy **7** trong Bảng VI biểu diễn hai trường hợp như vậy mà chúng tôi sử dụng lần lượt  $b = 3.0$  và  $b = 10.0$  với  $a = 0.0$  (PFCM tương đương với PCM). Với  $b = 10.0$ , PCM (thực hiện bởi PFCM) có thể tạo ra các nguyên mẫu tuyệt vời ngay cả với  $\eta = 3.0$ . Sử dụng PCM, người ta cũng có thể nhận được kết quả như chạy **7** với các giá trị được chia tỷ lệ phù hợp của  $\gamma_i$ .

Example 5:

Data set:  $X_{12}$

Algorithm: PCM.

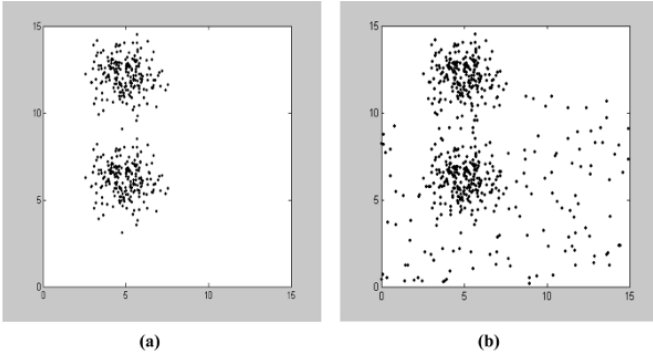
Initialization: FCM centroids.

Runs **1** through **5** of Table VI show the PCM centers for different values of  $\eta$  when the PCM algorithm is initialized with the terminal FCM prototypes. With an increase in  $\eta$  beyond 2, PCM generates almost coincident clusters. Except for one point the typicality values

become close to 0.5. In other words, with a moderately large value of  $\eta$  the limit property of typicality in (14) is attained. However, using our PFCM framework to realize PCM, we can overcome this situation by increasing the value of  $b$ . Run 6 and run 7 in Table VI represent two such cases for which we use respectively  $b = 3.0$  and  $b = 10.0$  with  $a = 0.0$  (PFCM is equivalent to PCM). With  $b = 10.0$ , PCM (realized by PFCM) can produce excellent prototypes even with  $\eta = 3.0$ . Using PCM also one can get results like run 7 with appropriately scaled values of  $\gamma_i$ .

Tiếp theo, chúng tôi thảo luận về kết quả trên  $X_{400}$  và  $X_{550}$ .  $X_{400}$  là hỗn hợp của hai phân phối chuẩn 2 biến với các vectơ trung bình  $\begin{pmatrix} 5.0 \\ 6.0 \end{pmatrix}$  và  $\begin{pmatrix} 5.0 \\ 12.0 \end{pmatrix}$ . Mỗi cụm có 200 điểm, trong khi  $X_{550}$  là phiên bản mở rộng của  $X_{400}$  với 150 điểm bổ sung được phân phối đều trên  $[0, 15][0, 11]$ . Hình 3 cho thấy biểu đồ phân tán của  $X_{400}$  và  $X_{550}$ .

Next, we discuss results on  $X_{400}$  and  $X_{550}$ .  $X_{400}$  is a mixture of two 2-variate normal distributions with mean vectors  $\begin{pmatrix} 5.0 \\ 6.0 \end{pmatrix}$  and  $\begin{pmatrix} 5.0 \\ 12.0 \end{pmatrix}$ . Each cluster has 200 points, while  $X_{550}$  is an augmented version of  $X_{400}$  with an additional 150 points uniformly distributed over  $[0, 15][0, 11]$ . Fig. 3 shows the scatterplot of  $X_{400}$  and  $X_{550}$ .



Hình 3. (a) Đồ thị phân tán của  $X_{400}$ . (b) Đồ thị phân tán của  $X_{550}$

	a	b	m	$\eta$	$\mathbf{v}_1$	$\mathbf{v}_2$
<u>1</u>	0	1	-----	2	-2.146 0.019	2.146 0.019
<u>2</u>	0	1	-----	1.5	-2.96 0.54	2.96 0.54
<u>3</u>	0	1	-----	2.5	-0.00 0.04	0.00 0.04
<u>4</u>	0	1	-----	3	-0.00 0.06	0.00 0.06
<u>5</u>	0	1	-----	5	-0.00 0.07	0.00 0.07
<u>6</u>	0	5	-----	3	-1.67 0.01	1.67 0.01
<u>7</u>	0	10	-----	3	-3.20 0.01	3.20 0.01

BẢNG VI KẾT QUẢ DO PCM TẠO RA CHO CÁC GIÁ TRỊ KHÁC NHAU CỦA CÁC THAM SỐ VỚI  $X_{12}$  KHI KHỞI TẠO VỚI FCM CENTROIDS.

TABLE VI RESULTS PRODUCED BY PCM FOR DIFFERENT VALUES OF THE PARAMETERS WITH  $X$  WHEN INITIALIZED WITH THE FCM CENTROIDS.

Ví dụ 6:

Bộ dữ liệu:  $X_{400}$  và  $X_{550}$ .

Thuật toán: FCM, PFCM và PCM.

$$\text{Khởi tạo: } V_0 = \begin{pmatrix} 5.97 & 4.75 \\ 6.12 & 5.14 \end{pmatrix}. \quad (26)$$

Khởi tạo trong (26) (chỉ hiển thị với hai chữ số có nghĩa) thu được bằng hai điểm dữ liệu được chọn ngẫu nhiên từ  $X_{400}$ . Bảng VII mô tả các nguyên mẫu được tạo ra bởi ba thuật toán khi khởi tạo trong (26) được sử dụng. Giá trị trung bình thực của  $X_{400}$  là  $V_{true} = \begin{bmatrix} 4497 & 4.97 \\ 6.15 & 12.15 \end{bmatrix}$ . Thuật toán PCM không tạo ra các nguyên mẫu tốt mặc dù các trọng tâm ban đầu được tách biệt hợp lý, nhưng dần từ cùng một cụm. Cả FCM và PFCM đều tạo ra các nguyên mẫu tốt, rất gần với  $V_{true}$ . Đối với  $X_{550}$ , giống như  $X_{12}$ , tác động của các điểm nhiễu nhỏ hơn đối với các nguyên mẫu PFCM so với các tâm FCM. Cột 4 của Bảng VII hiển thị các tâm PFCM cho cả hai tập dữ liệu với  $a = 0.5, b = 4.0, m = 2.0$  và  $\eta = 1, 5$ . Vì  $b$  lớn hơn  $a$  tương đối, chúng tôi mong đợi phép tính tâm cụm trong (23) sẽ bị ảnh hưởng nhiều hơn bởi các giá trị điển hình và do đó, các tâm không được mong đợi bị ảnh hưởng nhiều bởi các điểm nhiễu. Cột 4 cho thấy rằng thực tế là như vậy. Trọng tâm PFCM thu được cho  $X_{550}$  ở cột 4 gần với  $V_{true}$  hơn trọng tâm ở cột 3. Mặc dù PCM tạo ra các cụm trùng với (26), nhưng nếu chúng ta sử dụng các nguyên mẫu FCM đầu cuối để khởi tạo PCM, nó sẽ tạo ra các nguyên mẫu tốt như thể hiện ở cột 5 của Bảng VII.

Example 6:

Data Sets:  $X_{400}$  and  $X_{550}$ .

Algorithms: FCM, PFCM, and PCM.

$$\text{Initialization: } V_0 = \begin{pmatrix} 5.97 & 4.75 \\ 6.12 & 5.14 \end{pmatrix}. \quad (26)$$

The initialization in (26) (which is displayed with only two significant digits) is obtained by two randomly selected data points from  $X_{400}$ . Table VII depicts the prototypes generated by the three algorithms when the initialization in (26) is used. The true mean of  $X_{400}$  is  $V_{true} = \begin{bmatrix} 4497 & 4.97 \\ 6.15 & 12.15 \end{bmatrix}$ . The PCM algorithm did not produce good prototypes although the initial centroids are reasonably separated, but coming from the same cluster. Both FCM and PFCM result in good prototypes, which are very close to  $V_{true}$ . For  $X_{550}$ , like  $X_{12}$ , the effect of the noise points is smaller on the PFCM prototypes than it is on the FCM centroids. Column 4 of Table VII shows the PFCM centroids for both data sets with  $a = 0.5, b = 4.0, m = 2.0$  and  $\eta = 1, 5$ . Since  $b$  is comparatively larger than  $a$ , we expect the centroid computation in (23) to be more influenced by the typicality values and, hence, the centroids are not expected to be affected much by the noise points. Column 4 reveals that it is indeed the case. The PFCM centroids obtained for  $X_{550}$  in column 4 are closer to  $V_{true}$  than the centroids in column 3. Although, PCM produces coincident clusters with (26), if we use the terminal FCM prototypes to initialize PCM, it results in good prototypes as shown in column 5 of Table VII.

Results on $X_{400}$									
FCM (m=2.0)		PCM ( $\eta=2.0$ )		PFCM (a=1.0, b=1.0, m=2.0, $\eta=2.0$ )		PFCM (a=0.5 b=4.0, m=2.0, $\eta=1.5$ )		PCM ( $\eta=2.0$ , initialized with FCM centroids)	
4.97	4.97	5.01	5.01	4.98	4.97	5.02	5.01	4.98	4.99
6.14	12.18	6.28	6.28	6.16	12.18	6.18	12.22	6.30	12.09
Results on $X_{550}$ (with 150 noise points)									
FCM (m=2.0)		PCM ( $\eta=2.0$ )		PFCM (a=1.0, b=1.0, m=2.0, $\eta=2.0$ )		PFCM (a=0.5 b=4.0, m=2.0, $\eta=1.5$ )		PCM ( $\eta=2.0$ , initialized with FCM centroids)	
5.68	5.25	5.03	5.03	5.49	5.21	5.21	5.12	5.07	5.02
5.45	11.65	6.41	6.33	5.70	11.71	5.97	11.99	6.74	11.51

BẢNG VII CÁC LOẠI MẪU ĐƯỢC SẢN XUẤT BỞI FCM, PCM VÀ PFCM CÓ KHỞI TẠO NHƯ TRONG (26)

Ví dụ 7:

Bộ dữ liệu:  $X_{400}$  và  $X_{550}$ .

Thuật toán: FCM, PFCM và PCM.

$$\text{Khởi tạo: } V_0 = \begin{pmatrix} 7.19 & 5.02 \\ 12.37 & 5.50 \end{pmatrix}. \quad (27)$$

Bây giờ chúng ta so sánh FCM và PCM với PFCM với khởi tạo trong (27). Giống như Ví dụ 6, khởi tạo trong (27) thu được bằng cách chọn ngẫu nhiên hai điểm dữ liệu, nhưng hai tâm cụm ban đầu được tách biệt rõ ràng và chúng đến từ hai cụm khác nhau. Cột 1 hiển thị tâm điểm FCM trong khi cột 2 và 3 mô tả tâm điểm PCM và PFCM. Trong trường hợp này, cả thuật toán PCM và PFCM đều tạo ra các nguyên mẫu tốt cho  $X_{400}$ . Tâm điểm PCM và PFCM trên  $X_{400}$  rất gần nhau. Đối với  $X_{550}$ , cả PCM và PFCM đều tạo ra các nguyên mẫu tốt hơn FCM.

Example 7:

Data Sets:  $X_{400}$  and  $X_{550}$ .

Algorithms: FCM, PFCM and PCM.

$$\text{Initialization: } V_0 = \begin{pmatrix} 7.19 & 5.02 \\ 12.37 & 5.50 \end{pmatrix}. \quad (27)$$

We now compare FCM and PCM with PFCM with initialization in (27). Like Example 6, the initialization in (27) is obtained by randomly selecting two data points, but the two initial centroids are well separated and they come from two different clusters. Column 1 shows the FCM centroids while columns 2 and 3 depict the PCM and PFCM centroids. In this case, both PCM and PFCM algorithms result in good prototypes for  $X_{400}$ . The PCM and PFCM centroids on  $X_{400}$  are very close to each other. For  $X_{550}$ , both PCM and PFCM generate better prototypes than FCM.

Results on $X_{400}$							
FCM (m=2.0)			PCM ( $\eta=2.0$ )		PFCM (a=1.0 b=1.0, m=2.0, $\eta=2.0$ )		PCM ( $\eta=2.0$ )
4.97	4.97	4.98	4.99	4.98	4.97	4.98	4.99
6.14	12.18	6.30	12.09	6.17	12.17	6.30	12.09
Results on $X_{550}$ (with 150 noise points)							
FCM (m=2.0)			PCM ( $\eta=2.0$ )		PFCM (a=1.0 b=1.0, m=2.0, $\eta=2.0$ )		PCM ( $\eta=2.0$ )
5.25	5.68	5.07	5.02	5.44	5.19	5.07	5.02
11.65	5.45	6.74	11.51	5.87	11.63	6.74	11.51

BẢNG VIII CÁC LOẠI MẪU ĐƯỢC SẢN XUẤT BỞI FCM, PCM VÀ PFCM KHI ĐƯỢC KHỞI TẠO BỞI (27)

TABLE VIII PROTOTYPES PRODUCED BY FCM, PCM AND PFCM WHEN THEY ARE INITIALIZED BY (27)

Initial centroids			FCM (m=2, $M_{FCM}=16$ )			PCM ( $\eta=2$ , $M_{PCM}=50$ )			PFCM (a=1.0, b=1.0, m=2.0 $\eta=2.0$ , $M_{PFCM}=13$ $M_{PFCM}=14$ )		
5.80	5.00	5.10	6.78	5.89	5.00	6.17	6.17	5.00	6.62	5.92	5.00
2.70	2.00	3.50	3.05	2.76	3.41	2.88	2.88	3.40	3.01	2.79	3.41
5.10	3.50	1.40	5.65	4.36	1.48	4.76	4.76	1.48	5.46	4.40	1.48
1.90	1.00	0.30	2.05	1.40	0.25	1.61	1.61	0.25	1.99	1.41	0.25
Initial centroids			FCM (m=2, $M_{FCM}=16$ )			PCM ( $\eta=2$ , $M_{PCM}=50$ )			PFCM (a=1.0, b=1.0, m=2.0 $\eta=2.0$ , $M_{PFCM}=13$ $M_{PFCM}=14$ )		
7.20	5.50	5.10	6.77	5.89	5.00	6.17	6.17	5.00	6.62	5.92	5.00
3.00	2.40	3.30	3.05	2.76	3.41	2.88	2.88	3.40	3.01	2.79	3.41
5.80	3.70	1.70	5.65	4.36	1.48	4.76	4.76	1.48	5.46	4.40	1.48
1.60	1.00	0.50	2.05	1.40	0.25	1.61	1.61	0.25	1.99	1.41	0.25

BẢNG IX KẾT QUẢ VỀ DỮ LIỆU IRIS KHI KHỞI TẠO Thuật toán BẢNG BA ĐIỂM DỮ LIỆU ĐƯỢC CHỌN NGẪU NHIÊN

TABLE IX RESULTS ON IRIS DATA WHEN ALGORITHMS ARE INITIALIZED WITH THREE RANDOMLY SELECTED DATA POINTS

Ví dụ 8:

Tập dữ liệu: IRIS.

Thuật toán:

FCM, PFCM và PCM.

Khởi tạo: Được nêu ở những nơi tương ứng.

Bây giờ chúng tôi trình bày PFCM trên một tập dữ liệu thực, IRIS, với ba cụm. IRIS [7], [8] là một tập dữ liệu bốn chiều chứa 50 mẫu của mỗi loại hoa Iris. Một trong ba cụm (lớp 1) tách biệt rõ ràng với hai cụm còn lại, trong khi lớp 2 và 3 có một số chồng chéo. Kết quả điển hình của việc so sánh các phân vùng FCM hoặc HCM đã được làm cứng với các nhân vật lý chính xác của Iris là 14-17 lỗi. Chúng tôi đã thực hiện một số lần chạy PFCM trên IRIS với các khởi tạo khác nhau và các lựa chọn tham số khác nhau. Đầu tiên, chúng tôi báo cáo kết quả của một vài lần chạy khi FCM, PCM và PFCM được khởi tạo với ba điểm dữ liệu được chọn ngẫu nhiên (Bảng IX). Chúng tôi đã thực hiện một số lần chạy như vậy và trong mỗi trường hợp, FCM và PFCM đều tạo ra các nguyên mẫu tốt, nhưng PCM, ngay cả khi ba tâm cụm ban đầu đến từ ba lớp khác nhau, vẫn tạo ra các cụm trùng nhau (tức là thu được hai cụm riêng biệt). Bảng IX hiển thị một số kết quả với các khởi tạo được hiển thị trong cột đầu tiên của cùng một bảng. Các lỗi thay thế lại ( $M$ ) với  $U_{PFCM}$  và  $T_{PFCM}$  tốt hơn so với  $U_{FCM}$ . PCM nhận được hai cụm trùng nhau tạo ra 50 lỗi thay thế lại.

Example 8:

Data Set: IRIS.

Algorithms:

FCM, PFCM, and PCM.

We now demonstrate PFCM on a real data set, IRIS, with three clusters. IRIS [7], [8] is a four-dimensional data set containing 50 samples each of three types of Iris flowers. One of the three clusters (class 1) is well separated from the other two, while classes 2 and 3 have some overlap. The typical result of comparing hardened FCM or HCM partitions to the physically correct labels of Iris is 14-17 errors. We made several runs of PFCM on IRIS with different initializations and different choices of parameters. First, we report results of a few runs when FCM, PCM, and PFCM are initialized with three randomly selected data points (Table IX). We have made several such runs and in each case FCM and PFCM generated good prototypes, but PCM, even when the three initial centroids come from three different classes, resulted in coincident clusters (i.e., obtained two distinct clusters). Table IX displays some typical results with initializations shown in the first column of the same table. The resubstitution errors ( $M$ ) with  $U_{PFCM}$  and  $T_{PFCM}$  are better than that with  $U_{FCM}$ . PCM gets two coincident clusters making 50 resubstitution errors.

Chúng tôi cũng đã chạy một vài lần PCM và PFCM khi thuật toán được khởi tạo bằng các nguyên mẫu đầu cuối FCM. Trong Bảng X, cột 1, hàng 2 hiển thị các nguyên mẫu đầu cuối FCM được sử dụng để khởi tạo các thuật toán PCM và PFCM (kết quả hiển thị ở cột 2-3 của hàng 2). Trong trường hợp này, như mong đợi, PFCM tạo ra các nguyên mẫu tốt, nhưng PCM lại tạo ra hai cụm



trùng khớp (Bảng X). Trong hàng 3, cột 1, chúng tôi hiển thị một tập hợp khác các nguyên mẫu đầu cuối FCM thu được với  $m = 3$ . Khi các nguyên mẫu FCM này được sử dụng để khởi tạo, PCM lại tạo ra hai cụm trùng khớp mà chúng tôi không hiển thị trong Bảng X. Thuật toán PFCM trong trường hợp này cũng vậy, tạo ra lỗi thay thế nhỏ về cả tư cách thành viên và tính điển hình. Trong IRIS vì lớp 2 và lớp 3 chồng lên nhau, người ta có thể lập luận rằng IRIS có hai cụm. Vì vậy, về mặt này, PCM thực hiện công việc tuyệt vời khi chỉ tìm thấy hai cụm ngay cả khi thuật toán được yêu cầu tìm kiếm ba cụm. Tuy nhiên, có những ứng dụng mà ngay cả khi không có cấu trúc cụm rõ ràng, chúng ta vẫn muốn tìm cụm. Ví dụ, mặc dù IRIS có thể được coi là có hai cụm, nhưng nếu chúng ta muốn trích xuất các quy tắc hoặc nguyên mẫu để thiết kế bộ phân loại, chúng ta cần tìm ít nhất ba cụm. Với dữ liệu IRIS, PCM không thể thực hiện điều này với tất cả các khởi tạo mà chúng ta đã thử nhưng PFCM có thể. Do đó, đối với các ứng dụng như vậy, các nguyên mẫu PFCM sẽ hữu ích hơn các nguyên mẫu FCM và PCM vì các nguyên mẫu PFCM không nhạy cảm với các giá trị ngoại lai và PFCM có thể tránh các cụm trùng nhau.

We also made a few runs of PCM and PFCM when the algorithm is initialized with FCM terminal prototypes. In Table X, column 1, row 2 shows the FCM terminal prototypes that are used to initialize the PCM and PFCM algorithms (results shown in columns 2-3 of row 2). In this case, as expected PFCM produced good prototypes, but PCM again resulted in two coincident clusters (Table X). In row 3, column 1, we show another set of FCM terminal prototypes that are obtained with  $m = 3$ . When these FCM prototypes are used to initialize, PCM again produced two coincident clusters that we do not show in Table X. The PFCM algorithm in this case too, produced low resubstitution error both with respect to membership and typicality. In IRIS since class 2 and class 3 overlap, one can argue that IRIS has two clusters. So, in this regard, PCM does an excellent job in finding only two clusters even when the algorithm is asked to look for three clusters. However, there are applications where even when there is no clear cluster substructure, we want to find clusters. For example, although IRIS may be thought of having two clusters, if we want to extract rules or prototypes for classifier design, we need to find at least *three* clusters. With IRIS data PCM could not do this with all initializations that we tried but PFCM could. Thus for such applications, PFCM prototypes will be more useful than the FCM and PCM prototypes because PFCM prototypes are not sensitive to outliers and PFCM can avoid coincident clusters.

FCM terminal prototypes ( $m=2, M_{FCMU}=16$ )			PCM ( $\eta=2, M_{FCMU}=50$ )			PFCM ( $a=1.0, b=1.0, m=2.0, \eta=2.0$ $M_{PFCMU}=13, M_{PFCMT}=14$ )		
6.77	5.00	5.89	6.17	5.00	6.17	6.62	5.92	5.00
3.05	3.41	2.76	2.88	3.40	2.88	3.01	2.79	3.41
5.65	1.48	4.36	4.76	1.48	4.76	5.46	4.40	1.48
2.05	0.25	1.40	1.61	0.25	1.61	1.99	1.41	0.25
FCM terminal prototype ( $m=3, M_{FCMU}=16$ )			FPCM ( $a=1.0, b=1.0, m=3.0, \eta=3.0, M_{FPCM}=11, M_{FPCMT}=10$ )			FPCM ( $a=1.0, b=3.0, m=2.0, \eta=1.5, M_{FPCM}=13, M_{FPCMT}=11$ )		
6.69	5.00	5.91	6.41	5.01	6.05	6.67	5.02	5.83
3.04	3.40	2.79	2.95	3.40	2.85	3.04	3.42	2.79
5.55	1.49	4.38	5.14	1.51	4.54	5.55	1.47	4.29
2.04	0.25	1.40	1.83	0.26	1.48	2.07	0.25	1.34

BẢNG X KẾT QUẢ VỀ DỮ LIỆU IRIS KHI KHỞI TẠO Thuật toán BẢNG FCM NGUYÊN TẮC ĐẦU CUỐI

TABLE X RESULTS ON IRIS DATA WHEN ALGORITHMS ARE INITIALIZED WITH FCM TERMINAL PROTOTYPES

## VII. PHẦN KẾT LUẬN VII. CONCLUSION

Chúng tôi đã lập luận về yêu cầu của cả giá trị thành viên và giá trị điển hình trong phân cụm, và đã đề xuất một mô hình phân cụm khả năng-mờ có tên là PFCM. Không giống như hầu hết các thuật toán phân cụm mờ và khả năng, PFCM tạo ra **ba đầu ra: một phân vùng mờ**  $c \times n$  hoặc **ma trận thành viên**  $U$  của  $X$ ; **một ma trận khả năng**  $T(c \times n)$  của các đặc trưng trong  $X$ ; và một tập hợp  $c$  tâm cụm  $V$  biểu diễn chặt chẽ các cụm trong  $X$ . PFCM đã được thử nghiệm trên năm tập dữ liệu với nhiều lần chạy (không phải tất cả đều được thảo luận ở đây) và hiệu suất ban đầu của nó cho thấy rằng nó cải thiện các vấn đề mà FCM, PCM và FPCM gặp phải. Các cặp  $(U, T)$  từ PFCM không giống như  $U$  từ FCM và  $T$  từ PCM, nhưng chúng có vẻ chia sẻ cùng các thuộc tính định tính như các ma trận ước tính riêng lẻ. PFCM có hai tham số bổ sung,  $a$  và  $b$  xác định **tâm quan trọng tương đối của thành viên** và **đặc trưng** trong tính toán tâm cụm. Bằng cách kết hợp phù hợp các tham số này, chúng ta có thể khiến PFCM hoạt động giống FCM hoặc PCM hơn. Cần phải nghiên cứu thêm trước khi có thể khẳng định nhiều về phạm vi lựa chọn tốt cho các tham số  $(a, b, m, \eta)$ .

We have argued the need for both membership and typicality values in clustering, and have proposed a possibilistic-fuzzy clustering model named PFCM. Unlike most fuzzy and possibilistic clustering algorithms, PFCM produces three outputs: a  $c \times n$  fuzzy partition or membership matrix  $U$  of  $X$ ; a  $c \times n$  possibility matrix  $T$  of typicalities in  $X$ ; and a set of  $c$  point prototypes  $V$  that compactly represents the clusters in  $X$ . PFCM has been tested on five data sets with many runs (not all discussed here) and its initial performance indicates that it does ameliorate problems suffered by FCM, PCM and FPCM.  $(U, T)$  pairs from PFCM are not the same as  $U$  from FCM and  $T$  from PCM, but they seem to share the same qualitative properties as the individually estimated matrices. PFCM has two additional parameters,  $a$  and  $b$  that define the relative importance of membership and typicality in the computation of centroids. By suitable combination of these parameters we can make PFCM behave more like FCM or PCM. Further investigation is required before much can be asserted about a good range of choices for the parameters  $(a, b, m, \eta)$ .

Các điều kiện cần thiết trong (21)-(23) đối với mô hình PFCM giữ nguyên đối với bất kỳ chuẩn tích vô hướng nào, ví dụ, đối với chuẩn Mahalanobis được chia tỷ lệ [6], do đó công thức này khá tổng quát. Hai nhánh chính của khái quát hóa đối với các mô hình c-means là cục bộ "tương ứng" này như của Gustafson và Kessel [9] hoặc Dave và Bhaswan [10]; và các phần mở rộng của các nguyên mẫu thành các bề mặt giống như vỏ, hãy xem ví dụ Krishnapuram et al. [11]. Kiến trúc cơ bản của thuật toán PFCM-AO rõ ràng sẽ vẫn như vậy, nhưng các phương trình cập nhật cho các phần mở rộng theo cả hai hướng sẽ cần được sửa đổi theo các điều kiện cần thiết thích hợp. Nhiều phần mở rộng này sẽ dễ hiểu và chúng tôi hy vọng sẽ sớm viết về một số phần mở rộng trong số đó.

The necessary conditions in (21)-(23) for the PFCM model hold for any inner product norm, e.g., for the scaled Mahalanobis norm [6], so the formulation is quite general. The two main branches of generalization for c-means models are locally "adaptive" schemes such as those of Gustafson and Kessel [9] or Dave and Bhaswan [10]; and extensions of the prototypes to shelllike surfaces, see for example Krishnapuram et al. [11]. The basic architecture of the PFCM-AO algorithm will clearly remain the same, but the update equations for extensions in either direction will need to be modified by the appropriate necessary conditions. Many of these extensions will be straightforward, and we hope to write about some of them soon.

Store	Unlabeled Object Data $X = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^p$
	<ul style="list-style-type: none"> <li>* <math>1 &lt; c &lt; n</math></li> <li>* <math>m &gt; 1</math> : [for PCM-AO, FPCM-AO and PFCM-AO : <math>\eta &gt; 1</math>]</li> <li>* a, b [for PFCM-AO]</li> <li>* <math>T_{\max}</math> = iteration limit</li> </ul> <p>Pick * Norm for <math>J_m</math> : <math>\ x\ _A = \sqrt{x^T A x}</math></p> <ul style="list-style-type: none"> <li>* <math>E_t = \ V_t - V_{t-1}\ _1 / pc</math></li> <li>* <math>0 &lt; \varepsilon</math> = termination criterion</li> <li>* For PFCM-AO and PCM-AO : <math>(\gamma_1, \dots, \gamma_c), \gamma_i &gt; 0</math></li> </ul>
	Guess $V_0 = (v_{1,0}, v_{2,0}, \dots, v_{c,0}) \in \mathbb{R}^{cp}$
Do	<p>While (<math>T \leq T_{\max}</math> and <math>E_t &gt; \varepsilon</math>)</p> <p>Calculate <math>U_t</math> with <math>V_{t-1}</math> and (3a) or (10a) or (18a) or (21)</p> <p>[For FPCM-AO : Calculate <math>T_t</math> with <math>V_{t-1}</math> and (18b)]</p> <p>[For PFCM-AO : Calculate <math>T_t</math> with <math>V_{t-1}</math> and (22)]</p> <p>Calculate <math>V_t</math> with (3b) or (10b) or (18c) or (23)</p> <p>End While</p> <p><math>(U, T, V) \leftarrow (U_t, T_t, V_t)</math></p>

BẢNG XI Thuật toán FCM-AO, PCM-AO, FPCM-AO và PFCM-AO CHO CHỈ TIÊU SẢN PHẨM BÊN TRONG  
TABLE XI FCM-AO, PCM-AO, FPCM-AO, AND PFCM-AO ALGORITHMS FOR INNER PRODUCT NORMS

## TÀI LIỆU THAM KHẢO

## REFERENCES

[1] J. C. Bezdek, Nhận dạng mẫu với thuật toán hàm mục tiêu mờ. New York: Plenum, 1981.

[1] J. C. Bezdek, Pattern Recognition With Fuzzy Objective Function Algorithms. New York: Plenum, 1981.

[2] R. Krishnapuram và J. Keller, "Một cách tiếp cận khả thi đối với việc phân cụm", IEEE Trans. Fuzzy Syst., tập 1, số 2, trang 98-110, tháng 4 năm 1993.

[2] R. Krishnapuram and J. Keller, "A possibilistic approach to clustering," IEEE Trans. Fuzzy Syst., vol. 1, no. 2, pp. 98-110, Apr. 1993.

[3] M. Barni, V. Cappellini, và A. Mecocci, "Một cách tiếp cận khả thi đối với việc phân cụm", IEEE Trans. Fuzzy Syst., tập 4, số 3, trang 393-396, tháng 6 năm 1996.

[3] M. Barni, V. Cappellini, and A. Mecocci, "A possibilistic approach to clustering," IEEE Trans. Fuzzy Syst., vol. 4, no. 3, pp. 393-396, Jun. 1996.

[4] O. Nasraoui và R. Krishnapuram, "Diễn giải rõ ràng về thuật toán phân cụm mờ và khả thi," trong Proc. EUFIT, Aachen, Đức, 1995, trang 1312-1318.

[4] O. Nasraoui and R. Krishnapuram, "Crisp interpretations of fuzzy and possibilistic clustering algorithm," in Proc. EUFIT, Aachen, Germany, 1995, pp. 1312-1318.

[5] R. Krishnapuram và J. Keller, "Thuật toán c-Means khả thi: Thông tin chi tiết và khuyến nghị", IEEE Trans. Fuzzy Syst., tập 4, số 3, trang 385-393, tháng 6 năm 1996.

[5] R. Krishnapuram and J. Keller, "The possibilistic c-Means algorithm: Insights and recommendations," IEEE Trans. Fuzzy Syst.,

vol. 4, no. 3, pp. 385-393, Jun. 1996.

[6] J. C. Bezdek và S. K. Pal, Mô hình mờ để nhận dạng mẫu. New York: IEEE Press, 1992.

[6] J. C. Bezdek and S. K. Pal, Fuzzy Models for Pattern Recognition. New York: IEEE Press, 1992.

[7] E. Anderson, "Mống mắt của bán đảo GASPE," trong Bull. Amer. Iris Soc., tập 59, 1935, trang 2-5.

[7] E. Anderson, "The irises of the GASPE peninsula," in Bull. Amer. Iris Soc., vol. 59, 1935, pp. 2-5.

[8] R. A. Johnson và D. W. Wichern, Phân tích thống kê đa biến ứng dụng, ấn bản lần thứ 3. Upper Saddle River, NJ: Prentice-Hall, 1992.

[8] R. A. Johnson and D. W. Wichern, Applied Multivariate Statistical Analysis, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1992.

[9] E. E. Gustafson và W. Kessel, "Phân cụm mờ với ma trận hiệp phương sai mờ," trong Proc. 1978 IEEE Conf. Quyết định và Kiểm soát, Piscataway, NJ, 1979, trang 761-766.

[9] E. E. Gustafson and W. Kessel, "Fuzzy clustering with a fuzzy covariance matrix," in Proc. 1978 IEEE Conf. Decision and Control, Piscataway, NJ, 1979, pp. 761-766.

[10] R. N. Davé và K. Bhaswan, "Phân cụm c-shell mờ thích ứng và phát hiện hình elip", IEEE Trans. Neural Networks, tập 3, số 5, trang 643-662, tháng 9 năm 1992.

[10] R. N. Davé and K. Bhaswan, "Adaptive fuzzy c-shells clustering and detection of ellipses," IEEE Trans. Neural Networks, vol. 3, no. 5, pp. 643-662, Sep. 1992.

[11] R. Krishnapuram, H. Frigui và O. Nasroui, "Các thuật toán phân cụm vỏ mờ và khả thi và ứng dụng của chúng vào phát hiện ranh giới và xấp xỉ bề mặt", IEEE Trans. Fuzzy Syst., phần 1, 2, tập 3, số 1, trang 29-43-44-60, tháng 2 năm 1995.

[11] R. Krishnapuram, H. Frigui, and O. Nasroui, "Fuzzy and possibilistic shell clustering algorithms and their application to boundary detection and surface approximation," IEEE Trans. Fuzzy Syst., pt. 1, 2, vol. 3, no. 1, pp. 29-43-44-60, Feb. 1995.

[12] N. R. Pal, K. Pal và J. C. Bezdek, "Mô hình cụm c-means hỗn hợp", trong IEEE Int. Conf. Fuzzy Systems, Tây Ban Nha, 1997, trang 11-21.

[12] N. R. Pal, K. Pal, and J. C. Bezdek, "A mixed c-means clustering model," in IEEE Int. Conf. Fuzzy Systems, Spain, 1997, pp. 11-21.

[13] H. Timm, C. Borgelt, C. Doring và R. Kruse, "Phân tích cụm mờ với lực đẩy cụm", trình bày tại Hội nghị Công nghệ thông minh Euro (EUNITE), Tenerife, Tây Ban Nha, 2001.

[13] H. Timm, C. Borgelt, C. Doring, and R. Kruse, "Fuzzy cluster analysis with cluster repulsion," presented at the Euro. Symp. Intelligent Technologies (EUNITE), Tenerife, Spain, 2001.

[14] H. Timm và R. Kruse, "Một sửa đổi để cải thiện phân tích cụm mờ khả thi", được trình bày tại Hội nghị quốc tế IEEE về Hệ

thống mờ, FUZZ-IEEE' 2002, Honolulu, HI, 2002.

[14] H. Timm and R. Kruse, "A modification to improve possibilistic fuzzy cluster analysis," presented at the IEEE Int. Conf. Fuzzy Systems, FUZZ-IEEE' 2002, Honolulu, HI, 2002.

[15] H. Timm, C. Borgelt, C. Doring và R. Kruse, "Một phần mở rộng cho phân tích cụm mờ khả thi", Fuzzy Sets Syst., tập 2004, trang 3-16.

[15] H. Timm, C. Borgelt, C. Doring, and R. Kruse, "An extension to possibilistic fuzzy cluster analysis," Fuzzy Sets Syst., vol. 2004, pp. 3-16.

[16] E. E. Gustafson và W. C. Kessel, "Phân cụm mờ với ma trận hiệp phương sai mờ," trong Proc. IEEE Conf. Quyết định và Kiểm soát, San Diego, CA, 1979, trang 761-766.

[16] E. E. Gustafson and W. C. Kessel, "Fuzzy clustering with a fuzzy covariance matrix," in Proc. IEEE Conf. Decision and Control, San Diego, CA, 1979, pp. 761-766.

Các tác giả

Authors



Nikhil R. Pal

Nikhil R. Pal nhận bằng Thạc sĩ Quản trị Kinh doanh từ Đại học Calcutta, Calcutta, IN, năm 1982, và bằng Thạc sĩ Công nghệ và Tiến sĩ, cả hai đều chuyên ngành khoa học máy tính, từ Viện Thống kê Ấn Độ, Calcutta, lần lượt vào các năm 1984 và 1991.

Nikhil R. Pal received the Master of Business Management degree from the University of Calcutta, Calcutta, IN, in 1982, and the M.Tech. and Ph.D. degrees, both in computer science, from the Indian Statistical Institute, Calcutta, in 1984 and 1991, respectively.

Hiện tại, ông là Giáo sư tại Đơn vị Khoa học Điện tử và Truyền thông của Viện Thống kê Ấn Độ, Calcutta. Các lĩnh vực nghiên cứu của ông bao gồm xử lý hình ảnh, nhận dạng mẫu, lý thuyết tập mờ, mạng nơ-ron, tính toán tiến hóa và tin sinh học. Ông đồng sáng tác một cuốn sách có tựa đề Fuzzy Models and Algorithms for Pattern Recognition and Image Processing (Norwell, MA: Kluwer, 1999), đồng biên tập hai tập có tựa đề Advances in Pattern Recognition and Digital Techniques (ICAPRDT'99, Narosa) và Advances in Soft Computing (AFSS 2002, New York: Springer-Verlag) và biên tập một cuốn sách có tựa đề Pattern Recognition in Soft Computing Paradigm (Singapore: World Scientific, 2001).

Currently, he is a Professor in the Electronics and Communication Sciences Unit of the Indian Statistical Institute, Calcutta. His research interest includes image processing, pattern recognition, fuzzy sets theory, neural networks, evolutionary computation, and bioinformatics. He coauthored a book titled Fuzzy Models and Algorithms for Pattern Recognition and Image Processing (Norwell, MA: Kluwer, 1999), coedited two volumes titled Advances in Pattern Recognition and Digital Techniques (ICAPRDT'99, Narosa) and Advances in Soft Computing (AFSS 2002, New York: Springer-Verlag), and edited a book titled Pattern Recognition in Soft Computing Paradigm (Singapore: World Scientific, 2001).

Giáo sư Pal phục vụ trong Ban biên tập/Cố vấn của Tạp chí quốc tế về hệ thống thần kinh, Tạp chí quốc tế về lý luận xấp xỉ, Tạp chí quốc tế về hệ thống thông minh lai, Xử lý thông tin thần kinh - Thư và Đánh giá, Tạp chí quốc tế về hệ thống kỹ thuật thông minh dựa trên tri thức, Tạp chí Iran về hệ thống mờ, tập mờ và hệ thống, và IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS B. Ông là Tổng biên tập của IEEE TRANSACTIONS ON FUZZY SYSTEMS và là Thành viên Ủy ban chỉ đạo của tạp chí Applied Soft Computing, Elsevier Science. Ông là Chủ tịch của Hội đồng mạng nơ-ron châu Á - Thái Bình Dương. Ông là Chủ tịch Chương trình của Hội nghị quốc tế lần thứ 4 về Tiến bộ trong Nhận dạng

mẫu và Kỹ thuật số, tháng 12 năm 1999, Calcutta, Ấn Độ, và là Chủ tịch chung của Hội nghị quốc tế AFSS về Hệ thống mờ năm 2002, Calcutta, Ấn Độ, năm 2002. Ông là Chủ tịch chung của Hội nghị quốc tế lần thứ 11 về Xử lý thông tin thần kinh, ICONIP 2004, và là Đồng chủ tịch Chương trình của Hội nghị quốc tế IEEE về Hệ thống mờ năm 2005, Reno, NV, năm 2005.

Prof. Pal serves on the Editorial/Advisory Board of the International Journal of Neural Systems, International Journal of Approximate Reasoning, International Journal of Hybrid Intelligent Systems, Neural Information Processing-Letters and Reviews, International Journal of Knowledge-Based Intelligent Engineering Systems, Iranian Journal of Fuzzy Systems, Fuzzy Sets and Systems, and the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS B. He is the Editor-in-Chief of the IEEE TRANSACTIONS ON FUZZY SYSTEMS and a Steering Committee Member of the journal Applied Soft Computing, Elsevier Science. He was the President of the Asia Pacific Neural Net Assembly. He was the Program Chair of the 4th International Conference on Advances in Pattern Recognition and Digital Techniques, Dec. 1999, Calcutta, India, and the General Chair of 2002 AFSS International Conference on Fuzzy Systems, Calcutta, India, 2002. He is the General Chair of the 11th International Conference on Neural Information Processing, ICONIP 2004, and a Co-Program Chair of the 2005 IEEE International Conference on Fuzzy Systems, Reno, NV, 2005.



Kuhu Pal

Kuhu Pal đã nhận bằng Cử nhân (danh dự) chuyên ngành Vật lý từ Đại học Burdwan, Ấn Độ vào năm 1984, và bằng Thạc sĩ và Tiến sĩ chuyên ngành Vật lý từ Đại học Banaras Hindu, Ấn Độ vào các năm 1987 và 1993.

Kuhu Pal obtained the B.Sc. degree (with honors) in physics from the University of Burdwan, India, in 1984, and the M.Sc. and Ph.D. degrees in physics from Banaras Hindu University, India, in 1987 and 1993, respectively.

Sau đó, bà làm việc với tư cách là Trợ lý nghiên cứu đầu tiên tại Khoa Vật lý của Đại học Banaras Hindu, và sau đó từ tháng 9 năm 1995, làm việc tại Đơn vị Trí tuệ máy móc của Viện Thống kê Ấn Độ, Calcutta. Vào tháng 9 năm 1999, bà gia nhập Viện Kỹ thuật MCKV với tư cách là Giảng viên và sau đó là vị trí thỉnh giảng tại Khoa Khoa học máy tính, Đại học West Florida, Pensacola, từ tháng 1 năm 2000 trong thời gian sáu tháng. Mỗi quan tâm nghiên cứu của bà bao gồm nhận dạng mẫu, lý thuyết tập mờ, bộ điều khiển logic mờ, mạng nơ-ron và khoa học vật liệu tính toán. Bà là Nhà nghiên cứu tại Viện Kỹ thuật và Quản lý, Calcutta, Ấn Độ.

After that, she worked as a Research Associate first in the Physics Department of Banaras Hindu University, and then from September 1995, in the Machine Intelligence Unit of the Indian Statistical Institute, Calcutta. In September 1999, she joined the MCKV Institute of Engineering as a Lecturer and later that for a visiting position with the Computer Science Department, the University of West Florida, Pensacola, from January 2000 for a period of six months. Her research interest includes pattern recognition, fuzzy sets theory, fuzzy logic controllers, neural networks, and computational material science. She was a Researcher at the Institute of Engineering and Management, Calcutta, India.



James M. Keller

James M. Keller nhận bằng Tiến sĩ toán học vào năm 1978.

James M. Keller received the Ph.D. in mathematics in 1978.

Ông đã được bổ nhiệm làm giảng viên tại Chương trình Kỹ thuật sinh học/Tự động hóa nâng cao, Khoa Kỹ thuật máy tính và Khoa học máy tính, và Khoa Kỹ thuật điện và máy tính tại Đại học Missouri-Columbia, nơi ông hiện đang giữ chức Giáo sư. Ông cũng là Giáo sư nghiên cứu R. L. Tatum tại Khoa Kỹ thuật. Các mối quan tâm nghiên cứu của ông bao gồm thị giác máy tính, nhận dạng mẫu, lý thuyết tập mờ và logic mờ, hình học fractal và mạng nơ-ron. Ông đã được tài trợ bởi một số tổ chức công nghiệp và chính phủ, bao gồm Electronics and Space Corporation, Union Electric, Geo-Centers, National Science Foundation, NASA/JSC, Văn

phòng nghiên cứu khoa học của Không quân, Văn phòng nghiên cứu của Lục quân, Văn phòng nghiên cứu của Hải quân và Ban giám đốc cảm biến điện tử và tầm nhìn ban đêm của Lục quân. Ông đã đồng sáng tác hơn 225 ấn phẩm kỹ thuật.

He has had faculty appointments in the Bio- engineering/Advanced Automation Program, the Computer Engineering and Computer Science Department, and the Electrical and Computer Engineering Department at the University of Missouri-Columbia, where he currently holds the rank of Professor. He is also the R. L. Tatum Research Professor in the College of Engineering. His research interests include computer vision, pattern recognition, fuzzy set theory and fuzzy logic, fractal geometry, and neural networks. He has been funded by several industrial and government institutions, including the Electronics and Space Corporation, Union Electric, Geo-Centers, the National Science Foundation, NASA/JSC, the Air Force Office of Scientific Research, the Army Research Office, the Office of Naval Research, and the Army Night Vision and Electronic Sensors Directorate. He has coauthored over 225 technical publications.

Giáo sư Keller là Giảng viên quốc gia của Hiệp hội máy tính (ACM), Giảng viên danh dự của Hội đồng mạng nơ-ron IEEE và là cựu Chủ tịch của Hiệp hội xử lý thông tin mờ Bắc Mỹ (NAFIPS). Ông là cựu Tổng biên tập của IEEE TRANSACTIONS ON FUZZY SYSTEMS, Biên tập viên liên kết của Tạp chí quốc tế về lý luận xấp xỉ và là thành viên Ban biên tập của Pattern Analysis and Application, Fuzzy Sets and Systems, International Journal of Fuzzy Systems và Journal of Intelligent and Fuzzy Systems. Hiện ông đang phục vụ nhiệm kỳ ba năm với tư cách là Thành viên được bầu của Ủy ban hành chính của Hiệp hội IEEE SMC. Ông là Chủ tịch Hội nghị của Hội thảo NAFIPS năm 1991, Đồng chủ tịch Chương trình của Cuộc họp NAFIPS năm 1996, Đồng chủ tịch Chương trình của Hội nghị quốc tế IEEE về Mạng nơ-ron năm 1997 và Chủ tịch Chương trình của Hội nghị quốc tế IEEE về Hệ thống mờ năm 1998. Ông là Chủ tịch chung của Hội nghị quốc tế IEEE về Hệ thống mờ năm 2003.

Prof. Keller is a National Lecturer for the Association for Computing Machinery (ACM), an IEEE Neural Networks Council Distinguished Lecturer, and is a Past President of the North American Fuzzy Information Processing Society (NAFIPS). He is the former Editor-in-Chief of the IEEE TRANSACTIONS ON FUZZY SYSTEMS, an Associate Editor of the International Journal of Approximate Reasoning, and is on the Editorial Board of Pattern Analysis and Applications, Fuzzy Sets and Systems, International Journal of Fuzzy Systems, and the Journal of Intelligent and Fuzzy Systems. He is currently serving a three-year term as an Elected Member of the IEEE SMC Society Administrative Committee. He was the Conference Chair of the 1991 NAFIPS Workshop, Program Co-Chair of the 1996 NAFIPS Meeting, Program Co-Chair of the 1997 IEEE International Conference on Neural Networks, and the Program Chair of the 1998 IEEE International Conference on Fuzzy Systems. He was the General Chair for the 2003 IEEE International Conference on Fuzzy Systems.



James C.Bezdek

James C. Bezdek nhận bằng Tiến sĩ từ Đại học Cornell, Ithaca, NY, vào năm 1973.

James C. Bezdek received the Ph.D. degree from Cornell University, Ithaca, NY, in 1973.

Ông quan tâm đến chế biến gỗ, tối ưu hóa, xe máy, nhận dạng mẫu, làm vườn, câu cá, xử lý hình ảnh, mạng nơ-ron tính toán, nhạc blues và y học tính toán.

His interests include woodworking, optimization, motorcycles, pattern recognition, gardening, fishing, image processing, computational neural networks, blues music, and computational medicine.

Tiến sĩ Bezdek là Biên tập viên sáng lập của Tạp chí quốc tế về lý luận xấp xỉ và IEEE TRANSACTIONS ON FUZZY SYSTEMS, là thành viên của IFSA và là người nhận huy chương IEEE 3rd Millenium và Fuzzy Systems Pioneer.

Dr. Bezdek is the Founding Editor of the Inter- national Journal of Approximate Reasoning and the IEEE TRANSACTIONS ON FUZZY SYSTEMS, a Fellow of the IFSA, and recipient of the IEEE 3rd Millennium and Fuzzy Systems Pioneer medals