

مروری بر الگوریتم‌های خوشه‌بندی فازی

مسعود یقینی^۱؛ مریم رنجپور^۲؛ فرید یوسفی^۳

چکیده

هدف از خوشه‌بندی فازی استخراج مدل‌های فازی از داده‌هاست. کاربردهای متعدد خوشه‌بندی فازی در تحلیل داده‌ها و تشخیص الگو و نیز زمینه‌های پژوهشی موجود در این زمینه از جمله استفاده از آن در حل مسائل مسیریابی، تخصیص و زمان‌بندی نیاز به مطالعه الگوریتم‌های موجود و بهبود و اصلاح آن‌ها را آشکارتر می‌نماید.

در این مقاله، پس از مرور اجمالی مفاهیم پایه‌ای خوشه‌بندی فازی، شامل مبحث خوشه‌بندی، نظریه مجموعه‌های فازی و نهایتاً تعاریف پایه‌ای خوشه‌بندی فازی، به معرفی الگوریتم‌های اصلی و سپس الگوریتم‌های رایج که صورت‌های تغییر یافته‌ی الگوریتم‌های اصلی هستند پرداخته شده و در بخش نهایی مروری اجمالی بر مقالات ۱۰ سال اخیر (۲۰۰۰ الی ۲۰۰۹) ارائه شده است. هدف از این تحقیق ارائه یک دسته‌بندی برای الگوریتم‌های معرفی شده در زمینه خوشه‌بندی فازی است.

کلمات کلیدی

خوشه‌بندی، نظریه مجموعه‌های فازی، خوشه‌بندی فازی.

A Survey of Fuzzy Clustering Algorithms

Masoud Yaghini; Maryam Ranjpour; Farid Yousefi

ABSTRACT

A Fuzzy clustering technique looks for fuzzy model from data points. Because of applications of fuzzy clustering in data analysis, pattern recognition and to solve some problems like VRP, it is necessary to study its techniques.

This paper presents a review on fuzzy clustering algorithms. In the first section we have a glance on basic definitions of clustering, fuzzy set theory and fuzzy clustering. The second section includes basic fuzzy clustering algorithms. The two next sections present the variants of basic algorithms and the last one is a review of papers in fuzzy clustering (۲۰۰۰-۲۰۰۹). The main idea of this paper is classifying new algorithms of fuzzy clustering.

KEYWORDS

Clustering, Fuzzy Set Theory, Fuzzy Clustering.

^۱استادیار دانشکده مهندسی راه آهن دانشگاه علم و صنعت ایران، ۷۷۴۹۱۰۲۷، Yaghini@iust.ac.ir
^۲دانشجوی کارشناسی ارشد مهندسی حمل و نقل ریلی دانشگاه علم و صنعت ایران، Ranjpour@Gmail.com
^۳کارشناس ارشد مهندسی صنایع دانشگاه خواجه نصیرالدین طوسی، Yousefi@engineering.com

۱. مقدمه

خوشه‌بندی از ابزارهای متداول داده‌کاوی بوده که به استخراج دسته‌هایی با حداکثر شباهت بین عناصر داخل دسته و حداقل شباهت با عناصر سایر دسته‌ها می‌پردازد. این تشابه یا عدم تشابه براساس معیارهای اندازه‌گیری فاصله تعریف می‌شود. در واقع خوشه‌بندی یک کلاس‌بندی بدون نظارت است که در آن کلاس‌ها از پیش تعریف نشده‌اند. هدف از انجام خوشه‌بندی ارائه چشم‌انداز مناسبی از اتفاقات در حال وقوع در پایگاه داده‌ها به مصرف‌کننده نهایی اطلاعات می‌باشد. کاربرد دیگر خوشه‌بندی را می‌توان در تعیین داده‌هایی که با سایر داده‌ها تفاوت چشمگیر دارند عنوان نمود.

در خوشه‌بندی کلاسیک، هر نمونه ورودی متعلق به یک و فقط یک خوشه است و نمی‌تواند عضو دو خوشه و یا بیشتر باشد. به عبارتی خوشه‌ها همپوشانی ندارند در حالیکه در خوشه‌بندی فازی یک نمونه می‌تواند متعلق به بیش از یک خوشه باشد. خوشه‌بندی فازی به کشف مدل‌های فازی از داده‌ها می‌پردازد. این مقاله به بررسی الگوریتم‌هایی که هدفشان استخراج خوشه‌های فازی از داده‌های قطعی است، می‌پردازد. یکی از اولین روش‌های خوشه‌بندی فازی که بر مبنای تابع هدف و استفاده از فاصله اقلیدسی بنا شده بود در سال ۱۹۷۴ توسط دان^۲ ارایه و سپس توسط بزدک^۳ تعمیم داده شد. پس از آن یانگ^۴ یک بررسی اجمالی روی روش‌های خوشه‌بندی فازی انجام داد. سپس گوستافسون و کسل^۵ در سال ۱۹۷۹ الگوریتم خوشه‌بندی فازی با استفاده از ماتریس کوواریانس فازی را ارائه نمودند. تلفیق رویکرد امکان در خوشه‌بندی فازی نخستین بار توسط کیم و کریشناپورام^۶ در سال ۱۹۹۳ صورت گرفت. پس از آن نیز اصلاحات و بهبودهای بسیاری روی الگوریتم‌های ارائه شده صورت گرفته‌اند که در این تحقیق به برخی از آن‌ها خواهیم پرداخت. طبق دسته‌بندی صورت گرفته در این مقاله، الگوریتم‌های پایه‌ای در زمینه خوشه‌بندی فازی محدود به Fuzzy C-Means^۸ و Possibilistic C-Means^۷ شد که از Hard C-Means در ادبیات موضوع با عنوان الگوریتم k-Means معرفی شده است استخراج شده‌اند. همه الگوریتم‌های ارائه شده در این بخش مبتنی بر تابع هدف هستند که خوب بودن^۹ خوشه‌بندی را می‌سنجند.

۲. مفاهیم پایه‌ای

برای درک بهتر الگوریتم‌های خوشه‌بندی فازی آشنایی با مفاهیم پایه‌ای این مبحث ضروری به نظر می‌رسد. در این بخش به توضیح مختصر در مورد مباحث خوشه‌بندی، نظریه مجموعه‌های فازی و خوشه‌بندی فازی می‌پردازیم.

۲.۱. خوشه‌بندی

هدف از خوشه‌بندی، تقسیم‌بندی داده‌های موجود به چند گروه به‌قسمی است که داده‌های گروه‌های مختلف حداکثر تفاوت ممکن را از هم داشته باشند و داده‌های موجود در یک گروه بسیار به هم شبیه باشند. خوشه‌بندی با کلاس‌بندی تفاوت دارد و این تفاوت به برچسب دار بودن و نبودن دسته‌ها مربوط می‌شود. در خوشه‌بندی دسته‌ها از پیش معین نیستند.

دسته‌بندی‌های متفاوتی برای روش‌های خوشه‌بندی وجود دارد که رویکرد زیر جامع‌ترین دسته‌بندی این روش‌هاست:

- روش‌های افزایشی
- روش‌های سلسله‌مراتبی
- روش‌های مبتنی بر چگالی
- روش‌های مبتنی بر شبکه‌های مشبک
- روش‌های مبتنی بر مدل

مهم‌ترین و پرکاربردترین الگوریتمی که در بحث خوشه‌بندی وجود دارد الگوریتم K-Means است که جزو روش‌های افزایشی بوده و در آن، هر خوشه با میانگین اشیای آن (مرکز خوشه) نمایش داده می‌شود. این الگوریتم هنگامی که خوشه‌ها به صورت ابرهای فشرده‌ی مجزا از هم هستند به خوبی کار می‌کند. این روش برای پایگاه‌های داده بزرگ نسبتاً کارا و ارتقا پذیر است ولی اغلب به یک بهینه محلی منتهی می‌شود. از معایب روش، تعیین تعداد خوشه‌هاست که باید از قبل معلوم باشد و روش کارایی برای تعیین آن ارائه نشده است. همچنین برای کشف خوشه‌هایی با شکل‌های پیچیده مناسب نیست. یکی از مهم‌ترین نقاط ضعف این روش حساس بودن آن به داده‌های دور از مرکز است، این داده‌ها به راحتی مراکز را تغییر می‌دهند و ممکن است نتایج مطلوبی حاصل نشود.

۲.۲. نظریه مجموعه‌های فازی

نظریه مجموعه‌های فازی برای اولین بار توسط پروفیسور لطفی‌زاده در سال ۱۹۶۴ با هدف مدل‌سازی ابهامات و عدم قطعیت رویدادها مطرح گردید. این نظریه عدم قطعیت غیراحتمالی را پشتیبانی می‌کند و به دنبال ارائه قالبی برای نمایش و مدیریت دانش مبهم و عدم قطعیت است. در نظریه مجموعه‌های فازی مرز قطعی و مشخصی برای مجموعه‌ها وجود ندارد و مانند مجموعه‌های دقیق نمی‌توان تابع مشخصه با برد صفر و یک تعریف نمود.

عناصر می‌توانند با درجات عضویت متفاوت به مجموعه‌های مختلف تعلق داشته باشند. در نظریه مجموعه‌های کلاسیک یک عنصر یا به مجموعه‌ای تعلق دارد یا اینکه در مجموعه مکمل آن عضویت دارد و هیچ عضویتی به مجموعه نامبرده ندارد در حالیکه با رویکرد فازی در نظریه مجموعه‌ها یک عضو هم می‌تواند در مجموعه باشد و هم در مکمل آن.

۳.۲. خوشه‌بندی فازی

خوشه‌بندی فازی را می‌توان بخشی از تحلیل داده فازی دانست که دارای دو بخش است: یکی تحلیل داده‌های فازی و دیگری تحلیل داده‌های قطعی با استفاده از تکنیک‌های فازی. در این مطالعه به تحلیل دوم محدود خواهیم بود. ایده بنیادین در خوشه‌بندی فازی به این ترتیب است که فرض کنیم هر خوشه مجموعه‌ای از عناصر است. سپس با تغییر در تعریف عضویت عناصر در این مجموعه از حالتی که یک عنصر فقط بتواند عضو یک خوشه باشد (حالت افزایی)، به حالتی که هر عنصر می‌تواند با درجه عضویت‌های مختلف داخل چندین خوشه قرار بگیرد، دسته‌بندی‌هایی که انطباق بیشتری با واقعیت دارند ارائه کنیم.

۳. الگوریتم‌های پایه‌ای خوشه‌بندی فازی

خوشه‌بندی فازی به کشف مدل‌های فازی از داده‌ها می‌پردازد. این مقاله به بررسی الگوریتم‌هایی که هدفشان استخراج خوشه‌های فازی از داده‌های قطعی است، پرداخته است. یکی از اولین روش‌های خوشه‌بندی فازی که بر مبنای تابع هدف و استفاده از فاصله اقلیدسی بنا شده بود در سال ۱۹۷۴ توسط دان ارایه و سپس توسط بزدک تعمیم داده شد. پس از آن یانگ یک بررسی اجمالی روی روش‌های خوشه‌بندی فازی انجام داد. سپس گوستافسون و کسل در سال ۱۹۷۹ الگوریتم خوشه‌بندی فازی با استفاده از ماتریس کوواریانس فازی را ارائه نمودند. تلفیق رویکرد امکان در خوشه‌بندی فازی نخستین بار توسط کیم و کریشناپورام در سال ۱۹۹۳ صورت گرفت. پس از آن نیز اصلاحات و بهبودهای بسیاری روی الگوریتم‌های ارائه شده صورت گرفته‌اند که در این تحقیق به برخی از آن‌ها خواهیم پرداخت.

طبق دسته‌بندی صورت گرفته در این مقاله، الگوریتم‌های پایه‌ای در زمینه خوشه‌بندی فازی محدود به Fuzzy C-Means و Possibilistic C-Means شد که از Hard C-Means که در ادبیات موضوع با عنوان الگوریتم k-Means معرفی شده است استخراج شده‌اند و در ادامه توضیح کامل‌تری از آن‌ها ارائه خواهد شد. همه الگوریتم‌های ارائه شده در این بخش مبتنی بر تابع هدف هستند که خوب بودن خوشه‌بندی را می‌سنجند.

۱.۳. الگوریتم Fuzzy C-Means

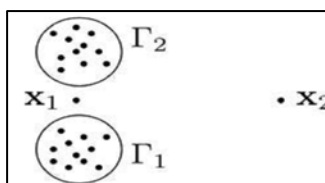
نخستین نسخه این الگوریتم^{۱۰} در سال ۱۹۷۳ توسط دودا و هارت^{۱۱} ارائه شد که یک خوشه‌بندی دقیق انجام می‌داد[۴]. از آنجایی که برخی داده‌ها به خوشه‌های متعدد وابستگی داشتند، امکان قرار دادن آن‌ها در یک خوشه وجود نداشت، بر اساس همین ایده دان نسخه فازی الگوریتم را ارائه نمود[۵]. الگوریتم بارها مورد بازبینی قرار گرفت تا در نهایت ارائه نسخه نهایی الگوریتم و نیز معرفی m به عنوان fuzzifier توسط بزدک صورت گرفت[۶].

الگوریتم حاصل ابرهای کروی از نقاط را در یک فضای p-بعدی شناسایی می‌کند. این خوشه‌ها به طور مفروض تقریباً هم‌اندازه هستند. هر خوشه با مرکز نشانی داده می‌شود. این نحوه نمایش خوشه‌ها، مدل یا نمونه^{۱۲} نیز نامیده می‌شود، زیرا اغلب به عنوان نماینده همه داده‌های تخصیص داده شده به خوشه انگاشته می‌شود. به عنوان یک متر برای فاصله، فاصله اقلیدسی بین یک نقطه و یک نمونه مورد استفاده قرار می‌گیرد. در انتخاب مرکز خوشه، همانگونه که از اسم الگوریتم پیداست مقدار میانگین مورد استفاده قرار می‌گیرد. برای محاسبه مرکز خوشه مجموع درجات عضویت هر عنصر به توان m در خودش به حاصلضرب توان m درجه عضویت‌ها تقسیم می‌شود.

مشکلی که در ارتباط با این الگوریتم مطرح است این است که الگوریتم نمی‌تواند خوشه‌هایی با اشکال و اندازه‌ها و چگالی‌های متفاوت شناسایی کند. برای شناسایی اشکال دیگر به جای ماتریس همانی در تعیین فاصله می‌توان از ماتریس‌های دیگر بهره جست. مانند ماتریس قطری برای تشخیص خوشه‌های بیضوی. از مزایای این الگوریتم، سهولت آن است که منجر به کاهش زمان محاسباتی می‌شود. در عمل با تکرارهای کمی می‌توان به حلی تقریباً نهایی رسید.

۲.۳. الگوریتم Possibilistic C-Means

مشخصه نسبی بودن درجه عضویت احتمالی با وجود مناسب بودن در اکثر مواقع، گاه می‌تواند ایجاد مشکلاتی نماید[۷]. به عنوان نمونه‌ای از این دسته مشکلات می‌توان به مورد ذیل اشاره نمود.



شکل ۱: ایراد وارد بر درجه عضویت احتمالی

نقطه x_1 از هر دو خوشه به یک فاصله است لذا درجه عضویتش به هر دو خوشه ۰.۵ است، این منطقی است، ولی مشکل وقتی پیش می‌آید که همان درجه تعلقات به x_2 هم نسبت داده می‌شود. در حالیکه فاصله این دو نقطه از خوشه‌ها به یک اندازه نیست. دلیل این موضوع نرمال‌سازی و لزوم برابر یک بودن مجموع درجات عضویت یک نقطه در خوشه‌های مختلف است. نرمال کردن درجات عضویت می‌تواند منجر به اثرات نامطلوب در ارائه داده‌های پرت و دور از مرکز شود. اگر شرط نرمال‌سازی را در الگوریتم FCM آزاد کنیم، این اثرات نامطلوب هم کمتر خواهند شد. این رویکرد امکان نام دارد [۸،۹،۱۰]. تابع هدف نیز که قبلاً فقط مربع فواصل را حداقل می‌نمود با رویکرد امکان چندان موافق به نظر نمی‌رسد. با حذف شرط نرمال‌سازی با اخذ درجات عضویت صفر برای همه نقاط در هر خوشه‌ای حداقل تابع هدف حاصل می‌شود و البته این معقول نیست که همه خوشه‌ها خالی باشند. پس بایستی جریمه‌ای در نظر گرفته شود تا درجات عضویت از صفر دور شوند [۱۰].

۴. الگوریتم‌های حاصل از تغییر در تابع فاصله

در هر دو الگوریتم پایه‌ای عنوان شده در بخش قبل، فاصله بین داده‌ها^{۱۲} و مراکز خوشه‌ها با متر اقلیدسی اندازه گرفته می‌شد. با این متر فقط می‌توان به شناسایی خوشه‌های کروی پرداخت. با تغییر متر مورد بحث، صورت‌های مختلفی برای حل مشکل تشخیص اشکال متفاوت ارائه شده است که در این بخش چند نمونه بیان شده که شامل الگوریتم Gustafson-Kessel، Fuzzy shell clustering و Kernel based algorithm می‌باشد. البته الگوریتم کلی‌تری نیز به نام Fuzzy relational clustering معرفی شده است که ورودی‌اش یک ماتریس فاصله است [۱۱].

۱.۴. الگوریتم Gustafson-Kessel

با جایگزین کردن فاصله اقلیدسی با متر دیگری^{۱۴} (که توسط یک ماتریس متقارن، معین و مثبت ایجاد شده) در الگوریتم FCM می‌توان به شناسایی خوشه‌های بیضوی نیز دست یافت. این پیشنهاد توسط گوستافسون و کسل در سال ۱۹۷۹ ارائه شد [۱۲]. در مقایسه با FCM در این روش هر خوشه علاوه بر مرکز خوشه توسط یک ماتریس متقارن، معین و مثبت مشخص می‌شود. این ماتریس برای هر خوشه یک نرم ایجاد می‌کند. باید این موضوع را هم در نظر گرفت که با انتخاب دلخواه و اختیاری ماتریس‌ها، فاصله‌ها می‌توانند به طور دلخواه کوچک شوند. برای اجتناب از حداقل سازی تابع هدف با ماتریس‌های با ورودی‌های تقریباً صفر، نیاز به مقداری ثابت برای خوشه‌ها با ماتریسی با دترمینان یک داریم. بنابراین حالا فقط شکل خوشه‌ها متغیر است نه اندازه‌هایشان. اما گوستافسون و کسل اشکال متفاوت برای خوشه‌ها را نیز مقدور ساختند بدین ترتیب که یک مقدار ثابت e برای هر ماتریس A معرفی کردند و در حالت کلی بایستی $\det(A)=e$ اگرچه انتخاب ثابت‌ها نیز نیاز به یک دانش پیشین در مورد خوشه‌ها دارد. اشکال بیضوی حاصل از GK در مورد مثال‌های یکسان با FCM که اشکال کروی حاصل می‌شد، خوشه‌بندی بهتری حاصل می‌کند، البته موضوع تجمع نقاط در مراکز خوشه‌ها در مورد خوشه‌بندی امکانی با GK نیز به صورت مسئله باقی می‌ماند. اگر داده‌ها با رویکرد امکان خوشه‌بندی شوند، فاکتورهای توسعه برای خوشه‌ها می‌توانند برای تشخیص اشکال در تصاویر مورد استفاده قرار گیرند. موقعیت و جهت را از مرکز خوشه و ماتریس می‌توان بدست آورد. کیفیت نتیجه حاصل از این روش به شدت وابسته به داده‌های در دسترس است. در مقایسه با FCM هزینه‌های محاسباتی GK بسیار بیشتر است. یک روش پیشنهادی مؤثر در کاهش گام‌های تکرار و افزایش سرعت همگرایی، آغاز الگوریتم GK با نتایج حاصل از یک اجرای FCM می‌باشد.

۲.۴. الگوریتم Fuzzy shell clustering

الگوریتم‌هایی که پیش از این به ذکر آن‌ها پرداختیم مناسب برای خوشه‌های فشرده محذب بودند به همین دلیل الگوریتم‌های خوشه‌بندی یکپارچه^{۱۵} نیز نامیده می‌شوند. این الگوریتم‌ها مناسب و مفید در کاربردهای تحلیل داده هستند. اما خوشه‌بندی فازی یک کاربرد دیگر هم در زمینه تشخیص و تحلیل تصویر دارد که نیاز به الگوریتم‌های متفاوتی را آشکار می‌سازد.

صورت‌های مختلفی از FCM و PCM برای تعیین خطوط، دوایر و بیضی‌ها روی مجموعه داده‌های متناظر با زیرساختارهای پیچیده پیشنهاد شده است که الگوریتم‌های shell clustering نامیده می‌شوند [۱۳]. متر مورد استفاده در این الگوریتم‌ها در حالت عمومی به صورت زیر است که در آن p مرکز و r شعاع است:

$$d^2(x_k, (p_i, r_i)) = (\|x_k - p_i\| - r_i)^2$$

از الگوریتم‌های این دسته می‌توان به FCV^{۱۶} اشاره نمود که برای تعیین خطوط، صفحات و ابرصفحات توسعه یافته است. الگوریتم AFCE^{۱۷} که صورت تغییر یافته FCM و PCM است و اجزای متمایز خط را^{۱۸} به خوشه‌های مختلف تخصیص می‌دهد. ترازهای دایره‌ای^{۱۹} را می‌توان با استفاده از Fuzzy C-Shells و Fuzzy C-Spherical Shells یافت. این دسته الگوریتم‌ها نمونه‌های با ذات متفاوت از سایر داده‌ها را بیرون می‌کشند و نیاز به جایگزینی متر اقلیدسی با متر اصلاح شده‌ای برای اندازه‌گیری فاصله بین داده‌ها و نمونه‌ها دارند. الگوریتم FCQS^{۲۰} قادر به تشخیص سهمی‌ها،

هذلولی‌ها و خوشه‌های خطی است. همچنین تکنیک‌های دیگری^{۲۱} از shell clustering برای ساختارهای ناهموار مانند مثلث‌ها و سایر چند ضلعی‌ها توسعه یافته‌اند [۱۴, ۱].

۳,۴. الگوریتم‌های خوشه‌بندی فازی مبتنی بر هسته

این دسته از الگوریتم‌ها، تابع فاصله را برای کاربرد در داده‌های غیر برداری^{۲۲} مانند دنباله‌ها، درخت‌ها یا گراف‌ها بدون نیاز به تغییر کامل در خود الگوریتم‌ها، اصلاح می‌کنند [۱۶, ۱۵]. در حالت کلی روش‌های مبتنی بر هسته مستقل از طبیعت داده و بدون نیاز به تطابق داده بکار بسته می‌شوند. اصول اولیه و تعاریف و مفاهیم پایه‌ای این روش‌ها در [۱۵] ارائه شده است.

بکار بستن قالب هسته برای خوشه‌بندی فازی استفاده از سایر فاصله‌ها را ممکن ساخته است. در بخش قبل دیدیم که در fuzzy shell clustering هم مترهای دیگری مورد استفاده قرار گرفت اما تفاوتی که وجود دارد این است که در آن‌ها هدف استخراج نمونه‌هایی است که از سایر داده‌ها متفاوتند لذا به اصلاح فاصله بین آنها می‌پردازد. این در حالیست که در رویکرد هسته شباهت بین جفت داده‌ها محاسبه می‌شود و مراکز خوشه‌ها را در برنمی‌گیرد. رویکرد هسته مستقل از ذات داده می‌تواند به کار برده شود در حالیکه الگوریتم‌های fuzzy shell clustering باید مختص نوع نمونه باشد. روش‌های مبتنی بر هسته Prototype-Based نیستند [۱۷]. به کار بستن این روش‌ها نیاز به انتخاب هسته و پارامترهای آن دارد که کار دشواری است و مشابه مسأله انتخاب مشخصه^{۲۳} و نماینده داده‌ها در روش‌های دیگر می‌باشد.

۵. الگوریتم‌های حاصل از تغییر در تابع هدف

در این بخش به بررسی الگوریتم‌هایی می‌پردازیم که به تغییرات اساسی‌تری در تابع هدف می‌پردازند و به دنبال بهبود نتایج خوشه‌بندی در موارد خاص هستند مانند حالتی که داده‌های پراکنده^{۲۴} داریم. در ادامه به برخی از اشکال^{۲۵} می‌پردازیم.

دسته اول از الگوریتم‌هایی که در این بخش به آن‌ها اشاره خواهد شد آن‌هایی هستند که هدفشان به طور صریح مواجهه با داده‌های پراکنده است. دسته دیگر به مطالعه و اصلاح fuzzifier می‌پردازد. گروهی به معرفی اصطلاحات جدید در تابع هدف برای بهینه‌سازی تعداد خوشه‌ها (به جای داشتن آن به طور قطعی در ابتدای فرآیند) می‌پردازد و نهایتاً به ارائه بهبودهایی در PCM اشاره خواهد شد. البته حد و مرزی بین این دسته‌ها وجود ندارد و تأثیرات متقابل میان آن‌ها انکارناپذیر است.

۱,۵. Noise Handling Variants

هدف از ارائه این دسته، تعریف الگوریتم‌های خوشه‌بندی فازی قوی^{۲۶} تری است که نتایج آن بستگی به حضور داده‌های پراکنده یا دور از مرکز در مجموعه داده‌ها ندارد. در این دسته سه رویکرد وجود دارد: اول noise clustering که به معرفی یک خوشه معین به نام خوشه noise برای نمایش داده‌های پرت می‌پردازد. دومین روش مبتنی بر استفاده از تخمین زنده‌های توانمند^{۲۷} است و سومی اثر داده‌های پرت را با تعریف وزن کاهش می‌دهد. رویکرد نخست برای اولین بار در [۱۸] پیشنهاد و سپس در [۱۹, ۲۰] توسعه یافت. رویکرد دوم در [۲۱] و رویکرد سوم در [۲۲] ارائه شدند که رویکرد آخر به هر داده وزنی اختصاص می‌دهد تا بتواند اثرات آن‌را بر پارامترهای خوشه کنترل کند.

۲,۵. Fuzzifier Variants

کلاس دیگری از اشکال FCM مبتنی بر مطالعه fuzzifier هستند. این توان برای اینکه درجات عضویت به جای اتخاذ دو مقدار {۰, ۱} در بازه [۰, ۱] مقدار بگیرند معرفی شد [۲۳, ۲۴, ۲۵, ۲۶].

۳,۵. Cluster Number Determination Variants

الگوریتم‌های خوشه‌بندی افزایی به دنبال افزای فازی بهینه داده‌ها به C خوشه هستند که C ورودی الگوریتم است. در اغلب موارد داده‌کاو واقعی، این پارامتر از پیش مشخص نیست و باید تعیین شود. از این دسته مطالعات می‌توان به [۲۷, ۲۸, ۲۹, ۳۰] اشاره نمود.

۴,۵. صورت‌های دیگر Possibilistic C-Means

از آنجاییکه PCM ممکن است منجر به نتایجی شود که رضایت‌بخش نباشد (مانند پدیده خوشه‌های منطبق که دلیل آن تابع هدف بهینه شده است که حداقل سرتاسری آن وقتی به دست می‌آید که همه خوشه‌ها یکسان باشند) می‌توان آن را با اصلاح تابع هدف بهبود بخشید. در ادامه به اختصار دو صورت اصلاح شده از PCM، اولی مبتنی بر الحاق یک اصطلاح جریمه در تابع هدف و دومی ترکیب PCM و FCM، ارائه می‌شود:

۱,۴,۵. دافعه خوشه‌ای

برای جلوگیری از ادغام خوشه‌ای، [۳۱, ۳۲] پیشنهاد شمول عبارتی دادند که بیانگر دفع بین خوشه‌ها باشد و به این ترتیب خوشه‌ها مجبور به تمایز باشند.

۲,۴,۵. صور تغییر یافته PCM، مبتنی بر ترکیب FCM و PCM

در مقالات [۳۳,۳۴]، رویکرد دیگری برای غلبه بر مشکل موجود در PCM ارائه شده است. رویکرد جدید هر دو درجه امکان و عضویت را برای انجام خوشه‌بندی ضروری می‌انگارد. درجه امکان به کاهش اثر داده‌های دور از مرکز می‌انجامد در حالیکه درجه عضویت برای تخصیص نقاط ضروری است. طبق [۳۵] یک خوشه‌بندی خوب هم نیاز به مشخصه افراز بندی FCM و هم مشخصه توانای جستجوی مد^{۲۸} PCM دارد.

۶. مروری بر مقالات سالهای اخیر

همان‌گونه که در بخش‌های قبل عنوان شد، در دو دهه اخیر اصلاحات و بهبودهای متعددی در الگوریتم‌های موجود در زمینه خوشه‌بندی فازی صورت گرفته است و همچنین الگوریتم‌های جدیدی مبتنی بر الگوریتم‌های اصلی و یا از ترکیب این الگوریتم‌ها با سایر الگوریتم‌ها (شامل سایر الگوریتم‌های خوشه‌بندی فازی و یا الگوریتم‌های دیگر مانند الگوریتم‌های تکاملی و...) ارائه شده است. بخشی از ادبیات موضوع در بخش قبل و برای دسته‌بندی کلی مطالب عنوان شد و بخشی دیگر که از مطالعه مقالات سال‌های اخیر حاصل شده است در ادامه ارائه خواهد شد. مقالات را می‌توان در حالت کلی به دو دسته کاربردی و الگوریتمی تقسیم‌بندی نمود. این تقسیم‌بندی از آنجایی ناشی شده است که برخی مقالات صرفاً به استفاده از الگوریتمی در زمینه‌ای خاص پرداخته‌اند. در بررسی اجمالی مقالات کاربردی به یک دسته‌بندی کلی از این مقالات می‌رسیم که به صورت زیر می‌توان به عنوان آن پرداخت:

✓ مقالاتی که هدف اصلی آن‌ها تعیین داده‌های دور از مرکز یا به عبارتی داده‌های پرت است.

✓ مقالاتی که صرفاً با هدف استخراج دسته‌هایی مشابه از الگوریتم‌های خوشه‌بندی استفاده می‌کنند. این حالت می‌تواند پیش‌نیازی برای کلاس‌بندی در نظر گرفته شود.

برای مقالات الگوریتمی نیز سه کلاس در نظر گرفته شد:

کلاس ۱: مقالاتی که به صورت ترکیب الگوریتم‌های خوشه‌بندی فازی با سایر الگوریتم‌ها (شامل سایر الگوریتم‌های خوشه‌بندی فازی و یا الگوریتم‌های دیگر مانند الگوریتم‌های تکاملی و...) هستند،

کلاس ۲: مقالاتی که به بررسی تعداد بهینه خوشه‌ها (اعتبار سنجی) می‌پردازند،

کلاس ۳: مقالاتی که به اصلاح یا بهبود یکی از مقالات موجود می‌پردازند.

مقالات کلاس ۱، ۵۸٪ کل مقالات بررسی شده سالهای ۲۰۰۰ الی ۲۰۰۹، مقالات کلاس ۲ که به بررسی تعداد بهینه خوشه‌ها (اعتبار سنجی) می‌پردازند، ۱۱٪ کل مقالات و مقالاتی که به اصلاح یا بهبود مقالات موجود می‌پردازند، ۲۵٪ از کل مقالات مورد مطالعه را تشکیل می‌دهند. در این بین ۶٪ مقالات، متعلق به هر دو دسته ترکیبی و بهبود یا اصلاح الگوریتم‌های موجود هستند. خلاصه‌ای از مقالات الگوریتمی و کاربردی در پیوست ۱ و ۲ ارائه شده است.

۷. نتیجه گیری

خوشه‌بندی فازی از تلفیق رویکرد فازی در بحث خوشه‌بندی برای کاربردی‌تر نمودن آن و انطباق بیشتر با دنیای واقعی حاصل شده است. در این مقاله به مروری بر الگوریتم‌های متداول خوشه بندی و به نوعی به خوشه بندی این الگوریتم‌ها پرداخته‌ایم. الگوریتم‌های رایج و کاربردی‌تر در جدول شماره ۱ خلاصه شده‌اند. آنچه در این تحقیق آمده است تکنیک‌هایی است که برای کشف مدل‌های فازی از داده‌های قطعی به‌کار می‌رود و در کارهای آتی می‌توان به مطالعه تکنیک‌هایی که برای داده‌های فازی توسعه یافته‌اند، اقدام نمود. استفاده از روش‌های متاهوریستیک برای بهبود نتایج خوشه‌بندی فازی نیز می‌تواند در لیست کارهای آتی قرار گیرد.

جدول ۱: مطرح ترین الگوریتم‌ها در خوشه‌بندی فازی

سال	ارائه کننده (برای نخستین بار)	نام الگوریتم
۱۹۷۳	Duda & Hart	ISODATA
۱۹۷۴	Dunn	Fuzzy C-Means
۱۹۷۹	Gustafson & Kessel	Gustafson-Kessel
۱۹۹۳	Krishnapuram & Keller	PCM
۱۹۹۷	Klawonn & Kruse & Timm	Fuzzy Shell Clustering
۲۰۰۳	Wu & Xie & Yu	Kernel Based Algorithm

پیوست شماره ۱: خلاصه‌ای از مقالات الگوریتمی

ردیف	عنوان	نویسنده/نویسندگان	سال	نوع مقاله	کلاس	الگوریتم
۱	The upper bound of the optimal number of clusters in fuzzy clustering	Jian, et	۲۰۰۱	journal	۲	-
۲	Simultaneous grouping of parts and machines with an integrated fuzzy clustering method	Josien, et.	۲۰۰۲	journal	۱	Fuzzy C-Means, Fuzzy K-Nearest Neighbors
۳	A real time hydrological forecasting system using a fuzzy clustering approach	Luchetta, et.	۲۰۰۳	journal	۳	The Nearest Neighborhood Clustering
۴	Towards a robust fuzzy clustering	Leski	۲۰۰۳	journal	۳	E-Insensitive Fuzzy C-Means
۵	A fuzzy-based customer classification method for demand-responsive logistical distribution operations	Hu, et.	۲۰۰۳	journal	۱	The Hybrid Fuzzy-Hierarchical Clustering Algorithm
۶	On cluster validity index for estimation of the optimal number of fuzzy clusters	Kim, et.	۲۰۰۴	journal	۲	Fuzzy C-Means
۷	Evolutionary Fuzzy Clustering Algorithm with Knowledge-Based Evaluation and Applications for Gene Expression Profiling	Park, et.	۲۰۰۵	journal	۱	Evolutionary Fuzzy Clustering
۸	Improved Validation Index for Fuzzy Clustering	Tang, et.	۲۰۰۵	Conference	۲	-
۹	Participatory Learning in Fuzzy Clustering	Silva, et.	۲۰۰۵	Conference	۱	Participatory Learning Fuzzy Clustering Algorithm
۱۰	Fuzzy Clustering Algorithm Based on Tree for Association Rules	Pi, et.	۲۰۰۶	journal	۱	FCABTAR
۱۱	Maximum entropy fuzzy clustering with application to real-time target tracking	Liamgqun, et.	۲۰۰۶	journal	۱	Maximum Entropy Fuzzy Clustering
۱۲	IFCM: Fuzzy Clustering for Rule Extraction of Interval Type-۲ Fuzzy Logic System	Zhang, et.	۲۰۰۷	Conference	۳	IFCM
۱۳	بررسی و مقایسه روشهای مختلف خوشه‌بندی فازی تفکیکی مبتنی بر روش استاندارد خوشه‌بندی فازی FCM	Sepehr, et.	۲۰۰۷	Conference	۳	FCM, GK
۱۴	A Weighted Fuzzy Clustering Algorithm for Multicast Routing Optimization	Bing, et.	۲۰۰۸	Conference	۱	WF cluster
۱۵	Image segmentation with a fuzzy clustering algorithm based on Ant-Tree	Yang, et.	۲۰۰۸	journal	۱	Ant-Tree Based Fuzzy Clustering
۱۶	Kernelized fuzzy attribute C-means clustering algorithm	Liu, et.	۲۰۰۸	journal	۳	Kernelized Fuzzy Attribute C-Means Clustering
۱۷	Modified Fuzzy Gap Statistic for Estimating Preferable Number of Clusters in Fuzzy k-Means Clustering	Arima, et.	۲۰۰۸	journal	۲	Fuzzy C-Means
۱۸	Robust fuzzy clustering using mixtures of Student's-t distributions	Chatzis, et.	۲۰۰۸	journal	۱	Fuzzy C-Means
۱۹	A fuzzy clustering algorithm based on evolutionary programming	Dong, et.	۲۰۰۹	journal	۱	EPFCM

ردیف	عنوان	نویسنده/نویسندگان	سال	نوع مقاله	کلاس	الگوریتم
۲۰	A fuzzy clustering method of construction of ontology-based user profiles	Han, et.	۲۰۰۹	journal	۱	FCOU
۲۱	Soft decision trees: A genetically optimized cluster oriented approach	Shukla, et.	۲۰۰۹	journal	۱	GA Guided Fuzzy Clustering
۲۲	Applications of axiomatic fuzzy set clustering method on management strategic analysis	Xu, et.	۲۰۰۹	journal	۱,۳	Fuzzy Clustering Method Based On AFS
۲۳	Autocorrelation-based fuzzy clustering of time series	DUrso, et.	۲۰۰۹	journal	۱	The Autocorrelation-Based Fuzzy C-Means
۲۴	Data-driven fuzzy clustering based on maximum entropy principle and PSO	Chen, et.	۲۰۰۹	journal	۱	Fuzzy Clustering Method Based On MEP And PSO
۲۵	Towards improving fuzzy clustering using support vector machine: Application to gene expression data	Mukhopadhyay, et.	۲۰۰۹	journal	۱	Iterated Version Of Fuzzy C-Means-VGA
۲۶	The extension of fuzzy QFD: From product planning to part deployment	Liu	۲۰۰۹	journal	۱,۳	Modified Fuzzy C-Means
۲۷	Simulated annealing based automatic fuzzy clustering combined with ANN classification for analyzing micro array data	Maulik, et.	۲۰۰۹	journal	۱	The VSA Based Fuzzy Clustering Algorithm
۲۸	Robust neural-fuzzy method for function approximation	Shieh, et.	۲۰۰۹	journal	۱	Robust Neural Fuzzy C-Means
۲۹	Robust fuzzy clustering-based image segmentation	Yang, et.	۲۰۰۹	journal	۳	Modified FCM
۳۰	RECM: Relational evidential c-means algorithm	Masson, et.	۲۰۰۹	journal	۳	RECM
۳۱	Fuzzy ensemble clustering based on random projections for DNA micro array data analysis	Avogadri, et.	۲۰۰۹	journal	۱	FCM, Fuzzy Ensemble Clustering
۳۲	KFCE: A dictionary generation algorithm for sparse representation	Xie, et.	۲۰۰۹	journal	۱	KFCE

پیوست شماره ۲: خلاصه‌ای از مقالات کاربردی

ردیف	عنوان	نویسنده/نویسندگان	سال	نوع مقاله	کاربرد	الگوریتم
۱	شناسایی قطبهای همگن از استانهای کشور بر مبنای شاخصهای بهداشتی و جمعیتی	غلامرضا بابایی	۲۰۰۴	Journal	شناسایی قطبهای همگن از استانهای کشور	Fuzzy C-Means
۲	Fuzzy Clustering Algorithms And Their Applications To Chemical Datasets	Jehan zeb shah et.	۲۰۰۵	Seminar	Chemo Informatics	Fuzzy C-Mean, GK, GG
۳	Stabilisation Diagrams: Pole Identification Using Fuzzy Clustering Techniques	M. Scionti et.	۲۰۰۵	Journal	Pole Identification	Fuzzy C-Mean, GK, GG
۴	A Fuzzy Clustering Approach To Real-Time Demand-Responsive Bus Dispatching Control	Jiuh-biing sheu	۲۰۰۵	Journal	Bus Dispatching Control	-
۵	رده بندی استانهای کشور بر مبنای شاخصهای بهداشتی و جمعیتی به کمک تکنیک آماری خوشه بندی فازی	غلامرضا بابایی	۲۰۰۵	Journal	رده بندی استانهای کشور	Fuzzy C-Means

ردیف	عنوان	نویسنده/نویسندگان	سال	نوع مقاله	کاربرد	الگوریتم
۶	A Novel Dynamic Resource Allocation Model For Demand-Responsive City Logistics Distribution Operations	Jiuh-biing sheu	۲۰۰۶	Journal	Demand-Responsive City Logistics Distribution Operations	-
۷	Students' Allocation Using Fuzzy Clustering Algorithms And Fukuyama And Sugeno's Fuzzy Cluster Validity Index	Sani susanto et.	۲۰۰۶	Journal	Students' Allocation	Fuzzy C-Means
۸	An Integrated Artificial Neural Network And Fuzzy Clustering Algorithm For Performance Assessment Of Decision Making Units	A. Azadeh et.	۲۰۰۷	Journal	Performance Assessment Of Decision Making Units	Fuzzy C-Means
۹	A Simplicity-Based Fuzzy Clustering Approach For Detection And Extraction Of Murmurs From The Phonocardiogram	V nigam	۲۰۰۸	Journal	Detection And Extraction Of Murmurs From The Phonocardiogram	Fuzzy C-Means
۱۰	Segmented Software Cost Estimation Models Based On Fuzzy Clustering	Javier aroba	۲۰۰۸	Journal	Segmented Software Cost Estimation Models	Fuzzy C-Means
۱۱	Identification Of Piecewise Affine Systems By Means Of Fuzzy Clustering And Competitive Learning	M.e. gegu' ndez	۲۰۰۸	Journal	Identification Of Piecewise Affine Systems	Fuzzy C-Means
۱۲	Structural Damage Detection By Fuzzy Clustering	Samuel da silva	۲۰۰۸	Journal	Structural Health Monitoring	Fuzzy C-Means ,GK
۱۳	Use Of Fuzzy Clustering Technique And Matrices To Classify Aminoacids And Its Impact To Chou'spsued Oaminoacid Composition	D.n. georgiou	۲۰۰۸	Journal	Classifying Amino Acids	The Fuzzy Equivalence Relation-Based Hierarchical Clustering Method
۱۴	A Knowledge Based Approach For A Fast Image Retrieval System	Alberto amato	۲۰۰۸	Journal	Fast Image Retrieval System	FCM, Femdd
۱۵	A Metamodel Optimization Methodology Based On Multi-Level Fuzzy Clustering Space Reduction Strategy And Its Applications	Wang hu	۲۰۰۸	Journal	The Practical Nonlinear Engineering Problems Optimization	Fuzzy C-Means
۱۶	Fuzzy Clustering Of Locations For Degree Of Accident Proneness Based On Vehicle User Perceptions	Jayanth jacob	۲۰۰۸	Journal	Clustering Of Locations For Degree Of Accident Proneness	Fuzzy C-Means
۱۷	Hybrid Adaptive Predictive Control For The Multi-Vehicle Dynamic Pick-Up And Delivery Problem Based On Genetic Algorithms And Fuzzy Clustering	Doris sáez	۲۰۰۸	Journal	Solving The Dynamic Multivehicle Pick-Up And Delivery Problem	Fuzzy C-Means
۱۸	Product Interface Reengineering Using Fuzzy Clustering	Fei gao	۲۰۰۸	Journal	Product Interface Reengineering	-
۱۹	A Case Study Of Applying Data Mining Techniques In An Outfitter's Customer Value Analysis	Shian-chang huang ,	۲۰۰۹	Journal	Customer Value Analysis	Fuzzy C-Means
۲۰	A Hybrid Intelligent System For Fault Detection And Sensor Fusion	Mohammad abdel kareem jaradat	۲۰۰۹	Journal	Fault Detection And Sensor Fusion	Fuzzy C-Means

ردیف	عنوان	نویسنده/نویسندگان	سال	نوع مقاله	کاربرد	الگوریتم
۲۱	A Spectral Clustering-Based Framework For Detecting Community Structures In Complex Networks	Jeffrey Q. Jiang	۲۰۰۹	Journal	Detecting Community Structures In Complex Networks	Fuzzy C-Means
۲۲	Clustering Analysis Of The Seismic Catalog Of Iran	Anooshiravanansari	۲۰۰۹	Journal	Analysis Of The Seismic Catalog	G&G
۲۳	Customer Pattern Search For After-Sales Service In Manufacturing	Jin sook ahn	۲۰۰۹	Journal	Customer Pattern Search	Fuzzy C-Means
۲۴	Utilization Of Self-Organizing Map And Fuzzy Clustering For Site Characterization Using Piezocone Data	Sarat kumar das et.	۲۰۰۹	Journal	Site Characterization	Fuzzy C-Means
۲۵	Using Conditional FCM To Mine Event-Related Brain Dynamics	Christos n.zigkolis et.	۲۰۰۹	Journal	Neuro Informatics	Conditional FCM
۲۶	Use Of Fuzzy Clustering Technique And Matrices To Classify Aminoacids And Its Impact To Chou'spseudo Aminoacid Composition	D.n. georgiou	۲۰۰۹	Journal	Classifying Aminoacids	The Fuzzy Equivalence Relation-Based Hierarchical Clustering Method
۲۷	T-S Fuzzy Model Identification Based On Anovel Fuzzy C-Regression Model Clustering Algorithm	Chaoshun li	۲۰۰۹	Journal	Identification Of Takagi-Sugeno(T-S)Fuzzy Model	Fuzzy C-Regression Model Clustering
۲۸	Prediction Of Relationships Between Hydrostatic Coefficients And Processing Parameters Of Porous PZT Ceramics By Radial Basis Function	B. Praveenkumar	۲۰۰۹	Journal	Prediction Of Relationships Between Hydrostatic Coefficients	Fuzzy C-Means
۲۹	Identifying Outlying Assessors In Sensory Profiling Using Fuzzy Clustering And Multi-Block Methodology	Tobias dahl	۲۰۰۹	Journal	Identifying Outlying Assessors In Sensory Profiling	Fuzzy C-Means

منابع و مراجع

- [۱] F. Hoppner, F.Klawonn, R.Kruse, T.Runkler; *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition*, John Wiley & Sons, ۲۰۰۰.
- [۲] M.S.Ilic, L.C.Jain; *Innovations In Fuzzy Clustering*, Springer, ۲۰۰۶.
- [۳] J. Valente De Oliveira, W. Pedrycz; *Advances in Fuzzy Clustering and Its Applications*, John Wiley & Sons, ۲۰۰۷.
- [۴] R. Duda, P. Hart; *Pattern Classification and Scene Analysis*, John Wiley & Sons, ۱۹۷۳.
- [۵] J.C. Dunn; "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well Separated Clusters". Journ. Cybern. ۳, ۹۵-۱۰۴, ۱۹۷۴.
- [۶] J.C. Bezdek; *Fuzzy Mathematics in Pattern Classification*, Ph. D. Thesis, Applied Math. Center, Cornell University, Ithaca, ۱۹۷۳.
- [۷] H.Timm, C.Borgelt, C.Do`Ring,R.Andkruse, "An Extension To Possibilistic Fuzzy Cluster Analysis", Fuzzy sets And Systems ۱۴۷, ۳-۱۶, ۲۰۰۴.
- [۸] D.Dubois, H.Prade; *Possibility Theory*, Plenum Press, New York, Usa, ۱۹۸۸.
- [۹] R. Dave', R. Krishnapuram; "Robust Clustering Methods: A Unified View". IEEE Transactions On Fuzzy Systems ۵, ۲۷۰-۲۹۳, ۱۹۹۷.
- [۱۰] R. Krishnapuram, J. Keller; "A Possibilistic Approach to Clustering", IEEE Transactions on Fuzzy Systems ۱, ۹۸-۱۱۰, ۱۹۹۳.
- [۱۱] R. Hathaway, J. Bezdek; "Nerf C-Means: Non-Euclidean Relational Fuzzy Clustering", Pattern Recognition ۲۷(۳), ۴۲۹-۴۳۷, ۱۹۹۴.
- [۱۲] E. E. Gustafson, W. C. Kessel; "Fuzzy Clustering With A Fuzzy Covariance Matrix", Proc. of The IEEE Conference on Decision and Control, San Diego, Californien, Pp. ۷۶۱-۷۶۶, ۱۹۷۹.
- [۱۳] F. Klawonn, R. Kruse, H. Andtim; *Fuzzy Shell Cluster Analysis*, Springer, ۱۹۹۷.
- [۱۴] J. C. Bezdek, J. Keller, R.Krishnapuram; *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing'*, Kluwer Boston, London Chapter ۳. Cluster Analysis for Relational Data, Pp. ۱۳۷-۱۸۲, ۱۹۹۹.
- [۱۵] B. Scho`Lkopf, A. Smola; *Learning With Kernels*. MIT Press, ۲۰۰۲.
- [۱۶] V. Vapnik; *The Nature of Statistical Learning Theory*. Springer, New York, Usa, ۱۹۹۵.
- [۱۷] Z. Wu, W. Xie, J. Yu; "Fuzzy C-Means Clustering Algorithm Based on Kernel Method", proc. of ICCIMA'۰۳, Pp. ۱-۶, ۲۰۰۳.
- [۱۸] R. Dave'; "Characterization and Detection of Noise in Clustering", Pattern Recognition Letters ۱۲, ۶۵۷-۶۶۴, ۱۹۹۱.

- [19] R. Dave, S. Andsen; "On Generalising The Noise Clustering Algorithms", Proc. of The 9th IFSA world congress, IFSA'97, Pp. 205-210, 1997.
- [20] R. Dave, S. Andsen; "Generalized Noise clustering as a Robust Fuzzy C-M-Estimators Model", Proc. of The 19th Int. Conference of the North American Fuzzy Information Processing Society: Nafips'98, Pp. 256-260, 1998.
- [21] H. Frigui, R. Krishnapuram; "A Robust Algorithm for Automatic Extraction of an Unknown Number of Clusters from Noisy Data", Pattern Recognition Letters 17, 1223-1232, 1996.
- [22] A. Keller; "Fuzzy Clustering With Outliers", Proc. of The 19th Int. Conf. of The North American Fuzzy Information Processing Society, Pp. 143-147, 2000.
- [23] F. Klawonn, F. Ho"ppner; "An Alternative Approach to the Fuzzifier in Fuzzy Clustering to Obtain Better Clustering Results", Proceedings 3rd Eusflat, Pp. 73-77, 2003.
- [24] F. Klawonn, F. Ho"ppner; "What Is Fuzzy about Fuzzy clustering? Understanding and Improving the Concept of the Fuzzifier", In Advances In Intelligent Data Analysis, Pp. 256-264, 2003.
- [25] P. Rousseeuw, E. Trauwart, L. Andkaufman; "Fuzzy Clustering With High Contrast", Journal Of Computational And Applied Mathematics 94, 81-90, 1995.
- [26] F. Klawonn; "Fuzzy Clustering: Insights and a New Approach", Math ware And Soft Computing 11, 125-142, 2004.
- [27] H. Sahbi, N. Boujemaa; "Validity of Fuzzy Clustering Using Entropy Regularization", Proc. of the IEEE int. Conf. on Fuzzy Systems, 2005.
- [28] H. Frigui, R. Krishnapuram; "Clustering by Competitive Agglomeration", Pattern Recognition 30 (7), 1109-1119, 1997.
- [29] H. Frigui, R. Krishnapuram; "A Robust Competitive Clustering Algorithm with Applications in Computer Vision", IEEE Transactions on Pattern Analysis and Machine Intelligence 21(5), 450-465, 1999.
- [30] M. Halkidi, Y. Batistakis, M. Vazirgiannis; "Cluster Validity Methods: Part I And Part II", Sigmod Record 31(2), 19-27 and 40-45, 2002.
- [31] H. Timm, R. Kruse; "A Modification to Improve Possibilistic Fuzzy Cluster Analysis", proc. of fuzz-IEEE'02, 2002.
- [32] H. Timm, C. Borgelt, C. Do"Ring, R. kruse; "An Extension to Possibilistic Fuzzy Cluster Analysis", Fuzzy sets And Systems 147, 3-16, 2004.
- [33] N. Pal, K. Pal, J. Bezdek; "A Mixed C-Means Clustering Model", Proc. Of Fuzz'IEEE'97, Pp. 11-21, 1997.
- [34] N. Pal, K. Pal, J. Keller, J. Bezdek; "A New Hybrid C-Means Clustering Model", Proc. of Fuzzy IEEE'04, Pp. 179-184, (2004).
- [35] R. Dave, S. Andsen; "Generalized Noise Clustering as a Robust Fuzzy c-M-Estimators model", proc. of the 19th int. Conference of the North American Fuzzy Information Processing Society: Nafips'98, Pp. 256-260, 1998.
- [36] X. Kong, R. Wang, G. Li; "Fuzzy Clustering Algorithms Based on Resolution and Their Application in Image Compression", Pattern Recognition 35, 2439 - 2444, 2002.
- [37] K. Josien, T. W. Liao; "Simultaneous Grouping of Parts and Machines with an Integrated Fuzzy Clustering Method", Fuzzy Sets and Systems 126, 1-21, 2002
- [38] C.W. Tao; "Unsupervised Fuzzy Clustering With Multi-Center Clusters", Fuzzy Sets and Systems 128, 305-322, 2002
- [39] T. Hu, J. Sheu; "A Fuzzy-Based Customer Classification Method for Demand -Responsive Logistical Distribution Operations", Fuzzy Sets and Systems 139, 431-450, 2003
- [40] A. Luchetta, S. Manetti; "A Real Time Hydrological Forecasting System Using a Fuzzy Clustering Approach", Computers & Geosciences 29, 1111-1117, 2002
- [41] D. Kim, K. H. Lee, D. Lee; "On Cluster Validity Index for Estimation of the Optimal Number of Fuzzy Clusters", Pattern Recognition 37, 2009 - 2025, 2004
- [42] F. Crespo, R. Weber; "A Methodology for Dynamic Data Mining Based On Fuzzy Clustering", Fuzzy Sets and Systems 150, 267-284, 2005
- [43] H. Park, S. Yoo, S. Cho; "Evolutionary Fuzzy Clustering Algorithm with Knowledge-Based Evaluation and Applications for Gene Expression Profiling", Journal of Computational and Theoretical Nano science Vol.2, 1-10, 2005
- [44] M. Ilic, T. Kuwata; "On Fuzzy Clustering Based Self-Organized Methods", The 2005 IEEE International Conference on Fuzzy Systems, 2005
- [45] Y. Tang, F. Sun, Z. Sun; "Improved Validation Index for Fuzzy Clustering", 2005 American Control Conference June 8-10, 2005. Portland, Or, Usa, 2005.
- [46] L. Silva, F. Gomide, R. Yager; "Participatory Learning in Fuzzy Clustering", The 2005 IEEE International Conference On Fuzzy Systems, 2005
- [47] H. Shen, J. Yang, X. Liu, K. Chou; "Using Supervised Fuzzy Clustering to Predict Protein Structural Classes", Biochemical and Biophysical Research Communications 334, 577-581, 2005
- [48] J. Sheu; "A Fuzzy Clustering Approach to Real-Time Demand-Responsive Bus Dispatching Control", Fuzzy Sets and Systems 150, 437-455, 2005
- [49] J. Shah; "Fuzzy Clustering Algorithms and Their Applications to Chemical Datasets", The Postgraduate Annual Research Seminar 2005
- [50] M. Scionti, J.P. Lanslots; "Stabilisation Diagrams: Pole Identification Using Fuzzy Clustering Techniques", Advances in Engineering Software 26, 768-779, 2005
- [51] E. Avineri; "Soft Computing Applications in Traffic and Transport Systems: A Review", Advances in Soft Computing 1, 17-25, 2005
- [52] Z. Bao, B. Han, S. Wu; "A General Weighted Fuzzy Clustering Algorithm", ICIAR 2006, LNCS 4142, Pp. 102 - 109, 2006
- [53] O. Pianyk; "Analytically Tractable Case of Fuzzy C-Means Clustering", Pattern Recognition 39, 35 - 46, 2006
- [54] D. Pi, X. Qin, Q. Wang; "Fuzzy Clustering Algorithm Based on Tree for Association Rules", International Journal of Information Technology, Vol.12, No.3, 2006
- [55] L. Liangqun, J. Hongbing, G. Xinbo; "Maximum Entropy Fuzzy Clustering with Application to Real-Time Target Tracking", Signal Processing 86, 3432-3447, 2006
- [56] J. Sheu; "A Novel Dynamic Resource Allocation Model for Demand-Responsive City Logistics Distribution Operations", Transportation Research Part E 42, 445-472, 2006
- [57] S. Susanto, P. Vasant; "Students' Allocation Using Fuzzy Clustering Algorithms and Fukuyama and Sugeno's Fuzzy Cluster Validity Index", World Transactions on Engineering and Technology Education Vol.5, No.1, 2006
- [58] W. Zhang, W. Liu; "IFCM: Fuzzy Clustering for Rule Extraction of Interval Type-2 Fuzzy Logic System", 49th IEEE Conference on Decision and Control New Orleans, La, Usa, Dec. 12-14, 2007
- [59] A. Azadeh, S.F. Ghaderi, M. Anvari, M. Saberi, H. Izadbakhsh; "An Integrated Artificial Neural Network and Fuzzy Clustering Algorithm for Performance Assessment of Decision Making Units", Applied Mathematics and Computation 187, 584-599, 2007
- [60] W. Pedrycz, H.S. Park, S.K. Oh; "A Granular-Oriented Development of Functional Radial Basis Function Neural Networks", Neuro computing 72, 42-53, 2008
- [61] W. Pedrycz; "A Dynamic Data Granulation through Adjustable Fuzzy Clustering", Pattern Recognition Letters 29, 2059-2066, 2008
- [62] Q. Bing, J. Lu, Y. Long; "A Weighted Fuzzy Clustering Algorithm for Multicast Routing Optimization", Fifth International Conference on Fuzzy Systems and Knowledge Discovery
- [63] N. Cebron, M. R. Berthold; "Adaptive Prototype-Based Fuzzy Classification", Fuzzy Sets and Systems 159, 2806 - 2818, 2008

- [६४] W. Pedrycz, P. Rai; “*Collaborative Clustering With the Use of Fuzzy C-Means and Its Quantification*”, Fuzzy Sets and Systems 159, 2399 – 2427, 2008
- [६5] X. Yang , W. Zhao, Y. Chen, X. Fang; “*Image Segmentation with a Fuzzy Clustering Algorithm Based on Ant-Tree*”, Signal Processing 88, 2453– 2462, 2008
- [६6] J. Liu, M. Xu; “*Kernelized Fuzzy Attribute C-Means Clustering Algorithm*”, Fuzzy Sets and Systems 159, 2428 – 2445, 2008
- [६7] C. Arima, K. Hakamada, M. Okamoto, T. Hanai; “*Modified Fuzzy Gap Statistic for Estimating Preferable Number of Clusters in Fuzzy K-Means Clustering*”, Journal of Bioscience and Bioengineering Vol. 105, No. 3, 273–281, 2008
- [६8] S. Chatzis, T. Varvarigou; “*Robust Fuzzy Clustering Using Mixtures of Student’s-T Distributions*”, Pattern Recognition Letters 29, 1901–1905, 2008
- [६9] A. Amato, V. Di Lecce; “*A Knowledge Based Approach for a Fast Image Retrieval System*”, Image and Vision Computing 26, 1466–1480, 2008
- [70] W. Hu, L. Enying, G.Y. Li, Z.H. Zhong; “*A Meta model Optimization Methodology Based on Multi-Level Fuzzy Clustering Space Reduction Strategy and Its Applications*”, Computers & Industrial Engineering 55, 503–532, 2008
- [71] V. Nigam, R. Priemer; “*A Simplicity-Based Fuzzy Clustering Approach for Detection and Extraction of Murmurs from the Phonocardiogram*”, Physiol. Meas. 29, 33–47, 2008
- [72] D. Sáez, C. E. Cortés, A. Núñez; “*Hybrid Adaptive Predictive Control for the Multi-Vehicle Dynamic Pick-Up and Delivery Problem Based on Genetic Algorithms and Fuzzy Clustering*”, Computers & Operations Research 35, 3412 – 3438, 2008
- [73] M.E. Gegu’ Ndez, J.Aroba, J.M. Bravo; “*Identification of Piecewise Affine Systems by Means of Fuzzy Clustering and Competitive Learning*”, Engineering Applications of Artificial Intelligence 21, 1321– 1329, 2008
- [74] F. Gao, G. Xiao, J. Chen; “*Product Interface Reengineering using Fuzzy Clustering*”, Computer-Aided Design 40, 439–446, 2008
- [75] J. Aroba, J. Gallego, M. Sicilia, I. Ramos, E. Barriocanal; “*Segmented Software Cost Estimation Models Based On Fuzzy Clustering*”, the Journal of Systems and Software 81, 1944–1950, 2008
- [76] S. Silva, M. Ju’ Nior, V. L. Junior, M. J. Brennan; “*Structural Damage Detection by Fuzzy Clustering*”, Mechanical Systems and Signal Processing 22, 1636–1649, 2008
- [77] Y. Wang, C. Li, Y. Zuo; “*A Selection Model for Optimal Fuzzy Clustering Algorithm and Number of Clusters Based on Competitive Comprehensive Fuzzy Evaluation*”, IEEE Transactions On Fuzzy Systems, Vol. 16, No. 3, June 2009
- [78] C. Braak, Y. Kourmpetis; “*Approximating A Similarity Matrix by a Latent Class Model: A Reappraisal of Additive Fuzzy Clustering*”, Computational Statistics and Data Analysis 53, 3182–3193, 2009
- [79] M. Filipponi; “*Dealing with Non-Metric Dissimilarities in Fuzzy Central Clustering Algorithms*”, International Journal of Approximate Reasoning 50, 363–384, 2009
- [80] M. L. Seghier, C. J. Price; “*Dissociating Functional Brain Networks by Decoding the Between-Subject Variability*”, Neuro image 45, 349–359, 2009
- [81] V. Galdi, A. Piccolo, P. Siano; “*Exploiting Maximum Energy from Variable Speed Wind Power Generation Systems by Using an Adaptive Takagi–Sugeno–Kang Fuzzy Model*”, Energy Conversion and Management 50, 413–421, 2009
- [82] A. Celikyilmaz , I. Burhan, R. Aktas, M. Doganay , N. B. Ceylan; “*Increasing Accuracy of Two-Class Pattern Recognition with Enhanced Fuzzy Functions*”, Expert Systems with Applications 36, 1337–1354, 2009
- [83] U. Maulik; “*Modified Differential Evolution Based Fuzzy Clustering for Pixel Classification in Remote Sensing Imagery*”, Pattern Recognition 42, 2135 – 2149, 2009
- [84] W. Pedrycz; “*Statistically Grounded Logic Operators in Fuzzy Sets*”, European Journal of Operational Research 193, 520–529, 2009
- [85] L. Han, G. Chen; “*A Fuzzy Clustering Method of Construction of Ontology-Based User Profiles*”, Advances in Engineering Software 40, 525–540, 2009
- [86] S.K. Shukla, M.K. Tiwari; “*Soft Decision Trees: A Genetically Optimized Cluster Oriented Approach*”, Expert Systems with Applications 36, 551–563, 2009
- [87] X. Xu, X. Liu, Y. Chen; “*Applications Of Axiomatic Fuzzy Set Clustering Method On Management Strategic Analysis*”, European Journal Of Operational Research 198, 297–304, 2009
- [88] D. Chen, C.Zhao; “*Data-Driven Fuzzy Clustering Based on Maximum Entropy Principle and PSO*”, Expert Systems with Applications 36, 625–633, 2009
- [89] C.G. Looney; “*Fuzzy Connectivity Clustering With Radial Basis Kernel Functions*”, Fuzzy Sets and Systems 160, 1868–1885, 2009
- [90] R. Avogadri, G. Valentini; “*Fuzzy Ensemble Clustering Based on Random Projections for DNA Microarray Data Analysis*”, Artificial Intelligence in Medicine 45, 173–183, 2009
- [91] Z. Yang, F. Chung, W. Shitong; “*Robust Fuzzy Clustering-Based Image Segmentation*”, Applied Soft Computing 9, 80–84, 2009
- [92] H. Shieh, Y. Yang, P. Chang, J. Jeng; “*Robust Neural-Fuzzy Method for Function Approximation*”, Expert Systems with Applications 36, 6903–6913, 2009
- [93] H. Liu; “*The Extension of Fuzzy QFD: From Product Planning to Part Deployment*”, Expert Systems with Applications 36, 11131–11144, 2009
- [94] S. Huang , E. Chang, H. Wu; “*A Case Study of Applying Data Mining Techniques in an Outfitter’s Customer Value Analysis*”, Expert Systems With Applications 36, 5909–5915, 2009
- [95] M. Jaradat, R. Langari; “*A Hybrid Intelligent System for Fault Detection and Sensor Fusion*”, Applied Soft Computing 9, 415–422, 2009
- [96] A. Ansari , A. Noorzad , H. Zafarani; “*Clustering Analysis of the Seismic Catalog of Iran*”, Computers & Geosciences 35, 475– 486, 2009
- [97] J. Ahn, S. Sohn; “*Customer Pattern Search for After-Sales Service in Manufacturing*”, Expert Systems with Applications 36, 5371–5375, 2009
- [98] M. Samhouri , M. Abu-Ghoush, E. Yaseen, T. Herald; “*Fuzzy Clustering-Based Modeling of Surface Interactions and Emulsions of Selected Whey Protein Concentrate Combined to I-Carrageenan and Gum Arabic Solutions*”, Journal Of Food Engineering 91, 10–17, 2009
- [99] T. Dahl, T. Naes; “*Identifying Outlying Assessors in Sensory Profiling Using Fuzzy Clustering and Multi-Block Methodology*”, Food Quality and Preference 20, 287–294, 2009
- [100] R. Ceylan, Y. Ozbay, B. Karlik; “*A Novel Approach for Classification of ECG Arrhythmias: Type-2 Fuzzy Clustering Neural Network*”, Expert Systems with Applications 36, 6721–6726, 2009
- [101] D.N. Georgiou, T.E.Karakasidis, J.J.Nieto, A.Torres; “*Use of Fuzzy Clustering Technique and Matrices to Classify Amino Acids and Its Impact to Chou’s Pseudo Amino Acid Composition*”, Journal Of Theoretical Biology 257, 17–26, 2009

- [۱۰۲] C. N. Zigkoulis, Nikolaos A. Laskaris, “Using Conditional FCM to Mine Event-Related Brain Dynamics”, Computers in Biology and Medicine ۳۹, ۳۴۶ – ۳۵۴, ۲۰۰۹
- [۱۰۳] S. K. Das, P. K. Basudhar, “Utilization of Self-Organizing Map and Fuzzy Clustering For Site Characterization Using Piezocone Data”, Computers and Geotechnics ۳۶, ۲۴۱–۲۴۸, ۲۰۰۹

زیر نویس ها

-
- ^۳ dunn
- ^۴ Bezdek
- ^۵ Yang
- ^۶ Kessel & Gustafson
- ^۷ Krishnapuram & Kim
- ^۸ Probabilistic Fuzzy C-Means, Standard Fuzzy C-Means
- ^۹ goodness
- ^{۱۰} Hard C-Means, ISODATA Algorithm
- ^{۱۱} Duda & Hart
- ^{۱۲} prototype
- ^{۱۳} data point
- ^{۱۴} Mahalanobis
- ^{۱۵} Solid clustering algorithms
- ^{۱۶} fuzzy c-varieties
- ^{۱۷} Adaptive Fuzzy C-Elliptotypes Algorithm
- ^{۱۸} disjoint line segments
- ^{۱۹} circle contours
- ^{۲۰} Fuzzy C-Quadric Shells
- ^{۲۱} Fuzzy C-Rectangular Shells, Fuzzy C-۲ Rectangular Shells
- ^{۲۲} non-vectorial
- ^{۲۳} feature
- ^{۲۴} noisy data
- ^{۲۵} variant
- ^{۲۶} robust
- ^{۲۷} robust estimator
- ^{۲۸} mode seeking robust