نشریه علمی پدافند غیرعامل

سال دهم، شماره ٤، زمستان ۱۳۹۸، (پیاپی ٤٠): صص ۳۱-۲۳

طراحی الگوریتم بهینه مبتنی‌بر خوشه‌بندی سلسله مراتبی جهت تفکیک رشته پالس‌های راداری در پردازنده ESM

محمدجواد قلندری[۱]، سید محمد علوی[۲*]

تاریخ دریافت: ۱۳۹۷/۱۲/۲۴
تاریخ پذیرش: ۱۳۹۸/۰۵/۱۶

چکیده

استخراج اطلاعات سیگنال‌هایی که توسط یک سیستم ESM در محیط شلوغ پالسی دریافت می‌شود، اولین گام شنود راداری است. قدم بعدی تفکیک پالس‌های راداری بر اساس داده‌های مرحله قبل، شناسایی منابع انتشاری و عوامل تهدید است. از مشکلات اساسی در بیشتر روش‌های تفکیک، انتخاب مناسب مراکز اولیه جهت خوشه‌بندی است. در الگوریتم پیشنهادی، ابتدا ماتریس تشابه برای داده‌های ورودی بر اساس شعاع همسایگی تعریف‌شده، تولید می‌گردد. سپس با تحلیل کدهای ماتریس تشابه و انتخاب سطرهای با بیشترین کد تشابه، دسته‌های متراکم‌تر به ترتیب جدا می‌شوند. ساختار این الگوریتم روش بهینه‌ای از الگوریتم‌های سلسله مراتبی است. در این الگوریتم معیار شباهت نقش به‌سزایی دارد. با روش به‌کار رفته در این الگوریتم مقدار محاسبات و تکرار در خوشه‌بندی سلسله مراتبی به شدت کاهش می‌یابد. الگوریتم حاصل نسبت به توابع موجود، داده را بر اساس شعاع همسایگی تعریف‌شده به تعداد خوشه‌های از پیش تعیین نشده با اولویت انتخاب دسته‌های متراکم تفکیک می‌نماید. از محاسن این روش نسبت به الگوریتم‌های مبتنی‌بر روش k-mean، دقت در انتخاب اولیه مراکز خوشه‌ها است. نتایج خوشه‌بندی روش پیشنهادی برای نمونه داده‌های راداری شامل ۲۰۰ پالس با نتایج روش خوشه‌بندی حول رهبر که یکی از الگوریتم‌های مطرح درزمینه خوشه‌بندی پالس‌های راداری است، مقایسه و نتایج مطلوب به‌دست آمده است..

کلیدواژه‌ها: خرابی خوشه‌بندی پالس‌های راداری، الگوریتم سلسله مراتبی، Radar Pulse Sorting

۱- دانشجوی دکتری مهندسی برق، دانشگاه جامع امام حسین(ع)، (Jghalandari@ihu.ac.ir)
۲- دانشیار، دانشگاه جامع امام حسین(ع)، (malavi@ihu.ac.ir)- نویسنده مسئول

۱. مقدمه

تحلیل خوشه‌بندی[1] یا به‌طور خلاصه خوشه‌بندی، فرآیندی است که به کمک آن می‌توان مجموعه‌ای از اشیاء را به گروه‌های مجزا افراز کرد. هر افراز یک خوشه نامیده می‌شود. اعضاء هر خوشه با توجه به ویژگی‌هایی که دارند به یکدیگر بسیار شبیه هستند و در عوض میزان شباهت بین خوشه‌ها کمترین مقدار است. در چنین حالتی هدف از خوشه‌بندی، نسبت دادن برچسب‌هایی به اشیاء است که نشان‌دهنده عضویت هر شیء به خوشه است. معمولا ۴ گروه اصلی برای الگوریتم‌های خوشه‌بندی وجود دارد. الگوریتم‌های خوشه‌بندی تفکیکی، الگوریتم‌های خوشه‌بندی سلسله مراتبی، الگوریتم‌های خوشه‌بندی برمبنای چگالی و الگوریتم‌های خوشه‌بندی برمبنای مدل. در ادامه هر یک از این گروه‌ها توضیح داده می‌شود.

۱-۱. خوشه‌بندی تفکیکی[2]

در این روش، براساس n مشاهده و k گروه، عملیات خوشه‌بندی انجام می‌شود. به این ترتیب تعداد خوشه‌ها یا گروه‌ها از قبل در این الگوریتم مشخص است. با طی مراحل خوشه‌بندی تفکیکی، هر شیء فقط و فقط به یک خوشه تعلق خواهد داشت و هیچ خوشه‌ای بدون عضو باقی نمی‌ماند.

معمولا الگوریتم‌های تفکیکی بر مبنای بهینه‌سازی یک تابع هدف عمل می‌کنند. این کار براساس تکرار مراحلی از الگوریتم‌های بهینه‌سازی انجام می‌شود. الگوریتم‌های مختلفی بر این مبنا ایجاد شده‌اند.

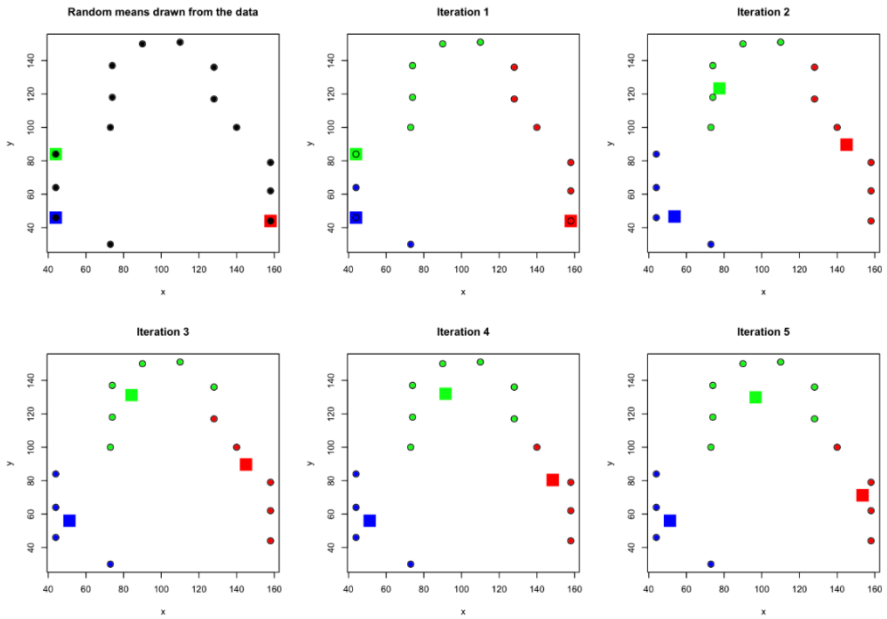
برای مثال الگوریتم k میانگین (K-means) با تعیین تابع هدف براساس میانگین فاصله اعضای هر خوشه نسبت به میانگین، عمل می‌کند. این الگوریتم به شکلی اشیاء را در خوشه‌ها قرار می‌دهد تا میانگین مجموع مربعات فاصله‌ها در خوشه‌ها، کمترین مقدار را داشته باشد.

به این ترتیب با استفاده از روش‌های مختلف بهینه‌سازی می‌توان به جواب مناسب برای خوشه‌بندی تفکیکی رسید. ولی از آنجایی که تعداد خوشه‌ها یا مراکز اولیه باید به الگوریتم داده شود، ممکن است با تغییر نقاط اولیه نتایج متفاوتی در خوشه‌بندی به‌دست آید. در میان الگوریتم‌های خوشه‌بندی تفکیکی، الگوریتمk میانگین که توسط مک‌کوئین[3] جامعه شناس و ریاضیدان در سال ۱۹۶۵ ابداع شد، از محبوبیت خاصی برخوردار است.

الگوریتم پایه برای این روش به شرح زیراست:

۱- در ابتدا k نقطه به‌صورت تصادفی به‌عنوان مراکز خوشه‌ها انتخاب می شوند.
۲- هر نمونه داده به خوشه‌ای که مرکز آن خوشه کمترین فاصله تا آن داده را داراست، نسبت داده می‌شود.
۳- بعد از تعلق تمام داده‌ها به یکی از خوشه‌ها، برای هر خوشه یک نقطه جدید به‌عنوان مرکز خوشه محاسبه می‌شود (میانگین نقاط متعلق به هر خوشه).
۴- مراحل ۲ و۳ به اندازه تعداد تکرار مطلوب انجام می‌شود.

در شکل (۱) انتخاب اولیه و تصادفی مراکز خوشه‌ها و تصحیح مراکز در تکرارهای بعدی الگوریتم نمایش داده شده است.



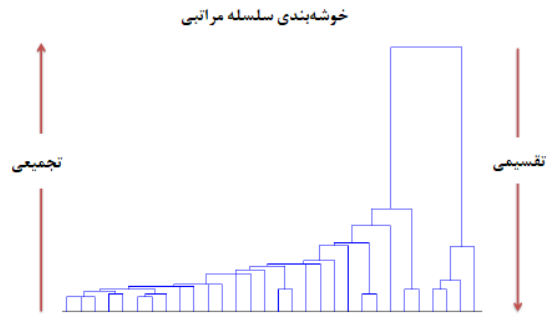
شکل (۱): نمایش روال خوشه‌بندی با الگوریتم kmeans

1- Cluster Analysis
2- Partitioning Clustering
3- McQueen

مشکلات روش خوشه‌بندی K-Means: علی‌رغم این که خاتمه‌پذیری الگوریتم بالا تضمین شده است ولی جواب نهایی آن واحد نبوده و همواره جوابی بهینه نمی‌باشد. به‌طور کلی روش ساده بالا دارای مشکلات زیر است. جواب نهایی به انتخاب خوشه‌های اولیه وابستگی دارد. روالی مشخص برای محاسبهٔ اولیهٔ مراکز خوشه‌ها وجود ندارد. اگر در تکراری از الگوریتم تعداد داده‌های متعلق به خوشه‌ای صفر شد راهی برای تغییر و بهبود ادامه روش وجود ندارد. در این روش فرض شده است که تعداد خوشه‌ها از ابتدا مشخص است. اما معمولا در کاربردهای زیادی تعداد خوشه‌ها مشخص نمی‌باشد.

۱-۲. خوشه‌بندی سلسله مراتبی[1]

برعکس خوشه‌بندی تفکیکی که اشیاء را در گروه‌های مجزا تقسیم می‌کند، خوشه‌بندی سلسله مراتبی در هر سطح از فاصله، نتیجه خوشه‌بندی را نشان می‌دهد. این سطوح به‌صورت سلسله مراتبی هستند. برای نمایش نتایج خوشه‌بندی به‌صورت سلسله مراتبی از درختواره[2] مطابق شکل (۲) استفاده می‌شود. این شیوه، روشی مؤثر برای نمایش نتایج خوشه‌بندی سلسله مراتبی است. در الگوریتم خوشه‌بندی سلسله مراتبی معمولا با کوچکترین خوشه‌ها شروع می‌شود. به این معنی که در ابتدا هر مقدار بیانگر یک خوشه است. سپس دو مقداری که بیشترین شباهت (کمترین فاصله) را دارند با یکدیگر ادغام شده و یک خوشه جدید می‌سازند. بعد از این مرحله، ممکن است دو مقدار یا یک مقدار با یک خوشه یا حتی دو خوشه با یکدیگر ادغام شده و خوشه جدیدی ایجاد کنند. این عملیات با رسیدن به بزرگترین خوشه که از همه مقدارها تشکیل یافته پایان می‌یابد.



شکل (۲): نمایش خوشه‌بندی سلسله مراتبی به روش تجمیعی و تقسیمی

به این شیوه خوشه‌بندی سلسله مراتبی، روش تجمیعی[3] می‌گویند. برعکس اگر عمل خوشه‌بندی سلسله مراتبی به شکلی باشد که با بزرگترین خوشه، عملیات خوشه‌بندی شروع شده و با تقسیم هر خوشه به خوشه‌های کوچکتر ادامه پیدا کند، به آن خوشه‌بندی سلسله مراتبی تقسیمی[4] می‌گویند. در آخرین مرحله این روش، هر مقدار تشکیل یک خوشه را داده و n خوشه خواهیم داشت. هر چند شیوه نمایش خوشه‌بندی سلسله مراتبی تقسیمی نیز توسط درختواره انجام می‌شود ولی از آنجایی که محدودیت‌هایی برای تقسیم یک خوشه به زیر خوشه‌های دیگر وجود دارد، روش تقسیمی کمتر به کار می‌رود.

برای اندازه‌گیری فاصله بین دو خوشه یا یک مقدار با یک خوشه از روش‌های پیوند استفاده‌می‌شود. از انواع روش‌های پیوند می‌توان به نزدیکترین همسایه یا پیوند تکی دورترین همسایه یا پیوند کامل و همچنین پیوند میانگین اشاره کرد. تفاوت اصلی در بین تمام این روش‌ها به نحوهٔ محاسبهٔ شباهت بین خوشه‌ها مربوط می‌شود.

۱-۳. خوشه‌بندی برمبنای چگالی[5]

در الگوریتم‌ها خوشه‌بندی برمبنای چگالی، نقاط با تراکم زیاد شناسایی شده و در یک خوشه قرار می‌گیرند. از الگوریتم‌های معروف در این زمینه می‌توان به DBSCAN اشاره کرد که در سال ۱۹۹۶ توسط استر[6] معرفی شد. این الگوریتم توانایی شناسایی نقاط دورافتاده را داشته و می‌تواند تاثیر آن‌ها را در نتایج خوشه‌بندی از بین ببرد.

در این روش ابتدا یک نقطه را به صورت اختیاری انتخاب کرده و نقاط قابل دسترس برای آن پیدا می‌شود. اگر X یک نقطه مرکزی باشد (نقطه‌ای که درون خوشه قرار داشته باشد)، خوشه را تشکیل می‌دهد ولی اگر X نقطه‌ای مرزی باشد (نقطه‌ای که در لبه‌های خوشه قرار داشته باشد) هیچ نقطه قابل دسترسی برای X وجود نخواهد داشت و الگوریتم از نقطه دیگری ادامه پیدا می‌کند. در مراحل خوشه‌بندی ممکن است دو خوشه که به یکدیگر نزدیک باشند، ادغام شوند. فاصله بین دو خوشه برحسب حداقل فاصله بین اعضای دو خوشه (پیوند تکی) سنجیده می‌شود.

۱-۴. خوشه‌بندی برمبنای مدل[7]

در روش‌های پیشین، فرضیاتی مبنی بر وجود یک توزیع آماری برای داده‌ها وجود نداشت. در روش خوشه‌بندی برمبنای مدل یک توزیع آماری برای داده‌ها فرض می‌شود. هدف در اجرای خوشه‌بندی برمبنای مدل، برآورد پارامترهای توزیع آماری به همراه متغیر پنهانی است که به‌عنوان برچسب خوشه‌ها در مدل

1- Hierarchical Clustering
2- Dendrogram
3-Agglomerative
4- Divisive Hierarchical Clustering
5- Density-Bases Clustering
6- Ester
7- Model-Based Clustering

معرفی شده است.

۱-۵. اعتبارسنجی الگوریتم‌های خوشه‌بندی

هدف از اعتبارسنجی خوشه‌ها، یافتن خوشه‌هایی است که بهترین تناسب را با داده‌های موردنظر داشته باشند. دو معیار پایه اندازه‌گیری پیشنهادشده برای ارزیابی و انتخاب خوشه‌های بهینه عبارت‌اند از:

تراکم[1]: داده‌های متعلق به یک خوشه بایستی تا حد ممکن به یکدیگر نزدیک باشند. معیار رایج برای تعیین میزان تراکم داده‌ها واریانس داده‌ها است. جدایی[2]: خوشه‌ها خود بایستی به‌اندازه کافی از یکدیگر جدا باشند. بعضی از شاخص‌های استفاده‌شده در جدول (۱) در این بخش ارائه شده است:

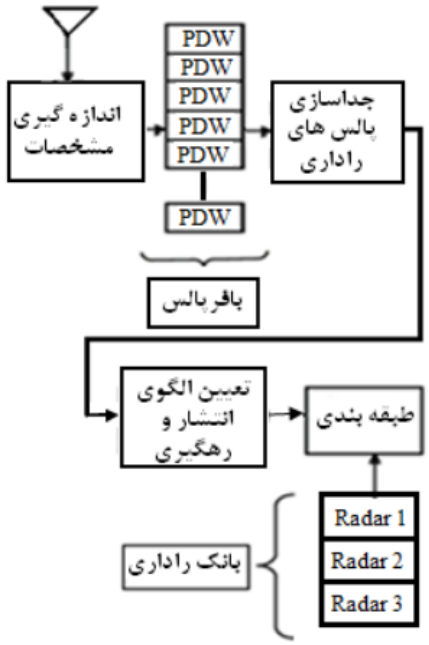
روش silhouette میانگین: این معیار مشخص می‌کند که پراکندگی داده‌ها در خوشه‌ها به چه صورت است. هر چه مقدار سیلوئت بالاتر باشد، کیفیت خوشه‌بندی نیز بالاتر است. در روش سیلوئت میانگین، الگوریتم خوشه‌بندی به‌ازای مقادیر مختلف k اجرا شده و به‌ازای هر اجرا، معیار سیلوئت برای هر یک از اعضای خوشه‌ها محاسبه می‌شود. سپس از سیلوئت‌های به‌دست‌آمده معدل گرفته می‌شود. مقدار بهینه k مقداری است که به‌ازای آن، سیلوئت میانگین حداکثر شود.

شاخص دون (Dunn Index)**:** در این معیار اگر مجموعه داده، دارای خوشه‌های جداپذیر باشد، فاصله بین خوشه‌ها زیاد و قطر خوشه‌های آن کوچک خواهد بود. درنتیجه مقداری بزرگ‌تر برای این معیار مطلوب‌تر است.

در مرجع [۱] انواع روش‌های اعتبارسنجی بیان شده است. برای بررسی و مقایسه تعداد بهینه خوشه‌های مجزا می‌توان از برنامه Cvap[3] متلب استفاده کرد. متلب در این جعبه‌ابزار، چند روش خوشه‌بندی متداول و شاخص‌های معتبر برای ارزیابی خوشه‌های به‌دست‌آمده را فراهم نموده است.

۲. عملکرد کلی یک سیستم ESM

عملکرد کلی یک سیستم ESM در شکل (۳) نشان داده شده است. آنچه در ابتدا به ورودی گیرنده می‌رسد، قطاری از پالس‌های متداخل است. گیرنده ESM پس از آشکارسازی پالس‌های راداری، آنها را مرتب نموده و اندازه‌گیری‌های عددی شامل زمان دریافت پالس، عرض پالس، دامنه پالس، نوع مدولاسیون پالس، فرکانس رادیویی و زاویه دریافت پالس از واحد پردازشگر گیرنده به‌دست می‌آید.



شکل (۳): شرح عملکرد سیستم ESM مرجع [۲]

این مقادیر در قالب فایل‌های توصیف‌کننده پالس یا PDW[4] وارد پردازشگر اطلاعات می‌شود. در واحد پردازشگر اطلاعات، پالس‌های مربوط به رادارهای مختلف از هم تفکیک می‌شوند. به این عمل جداسازی رشته پالس‌های راداری[5] می‌گویند. بعد از یافتن رشته پالس‌ها، واحد شناسایی و نمایش با بررسی پارامترهای رشته پالس و استخراج اطلاعات لازم دیگر (مانند نوع اسکن و نرخ آن) به جستجو در بانک اطلاعات راداری جهت شناسایی اهداف می‌پردازد. درنهایت، اطلاعات رادارهای فعال، نوع تهدید و اولویت آنها به سامانه‌های هدایت و کنترل، اخلالگرها، سامانه‌های دفاعی و کاربر داده می‌شود.

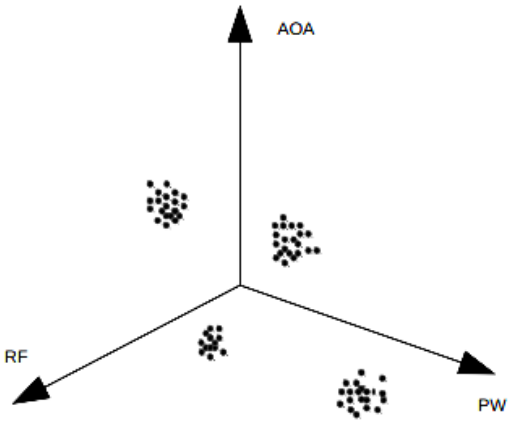
پردازشگر اطلاعات ESM ارتباط پالس‌های متفاوت را با گروه‌بندی آنها به‌دست آورده و پالس‌ها را تفکیک می‌کند. هدف از این کار جدا کردن پالس‌های متداخل و به‌دست آوردن پالس‌های هم‌قطاری است که مربوط به یک منبع انتشاری هستند. در حالت بهینه این واحد از بلوک خوشه‌بندی پالس‌های راداری و بلوک تجزیه وتحلیل زمانی تشکیل شده است. خوشه‌بندی، طبقه‌بندی نمونه داده‌های مشابه و تشکیل خوشه‌هایی است که پارامترهای آنها تا حد امکان به هم شبیه بوده و نسبت به خوشه‌های دیگر متفاوت باشند.

1- Compactness
2- Separation
3- Cluster Validity Analysis Platform

4- Pulse Description Word
5- Deinterleaving

روش‌های مختلف تفکیک پالس‌های راداری در مراجع مختلف مانند [۱۱-۳] توضیح داده شده است. الگوریتم‌های متعدد طراحی‌شده، سعی در افزایش دقت، کاهش زمان محاسبه و پاسخ سیستم دارند.

در خوشه‌بندی پالس‌های هم زاویه، هم فرکانس و با پهنای پالس تقریباً مساوی در یک گروه خوشه قرار می‌گیرند. شکل (۴) خوشه‌بندی در فضای سه‌بعدی زاویه، فرکانس و عرض پالس را نشان می‌دهد. در بلوک بعدی پردازشگر، اطلاعات پالس‌های مربوط به هر لانه یا رشته پالس ازنظر پریود زمانی (زمان دریافت) مورد تجزیه‌وتحلیل قرار می‌گیرد [۱۲]. آنالیز زمان ورود پالس به دنبال یافتن ارتباط زمانی بین پالس‌های یک مجموعه و تشخیص نوع PRI و مقدار آن است [۳].



شکل (۴): خوشه‌بندی در ابعاد زاویه، فرکانس و عرض پالس مرجع [۴]

۳. انتخاب روش بهینه تفکیک و دسته بندی پالس راداری

انتخاب روش بهینه تفکیک و دسته‌بندی پالس راداری به سناریوی ورودی نیز بستگی دارد. در پایان‌نامه‌های مراجع [۱۷] و [۱۹] روش‌های مختلف و مزایا و معایب هریک بررسی شده است. داده‌ها ماتریس‌هایی با تعداد سطرهای متغیر هستند که هر سطر حاصل اندازه‌گیری‌های به‌عمل‌آمده بر روی یک پالس (PDW) است. در هر فریم تعداد داده‌ها یا سطرهای ورودی متغیر است.

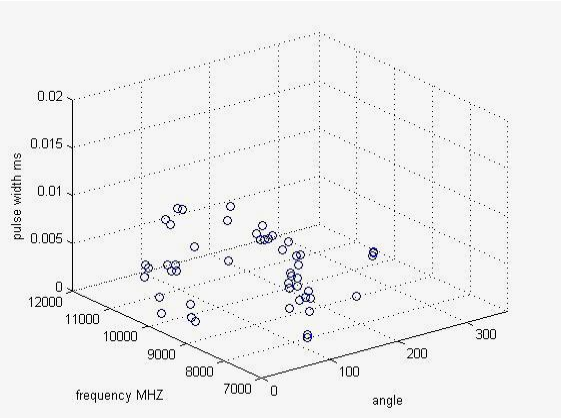
روش خوشه‌بندی سلسله مراتبی با پیوند نزدیک‌ترین همسایه بهینه شده با تحلیل کدهای تشابه برای تفکیک فریم‌های ورودی با روش سلولی یا خوشه‌بندی به‌صورت انتخاب رهبر (ترتیبی) مقایسه می شود. در روش رهبر هر پالس ورودی به عنوان رهبر یک خوشه معرفی شده و اگر فاصله متریک پالس بعدی از حد آستانه کمتر بود، به خوشه‌های موجود پیوسته و در غیر این صورت خوشه جدیدی تشکیل می‌دهد.

۳-۱. روش پیشنهادی برای تفکیک پالس

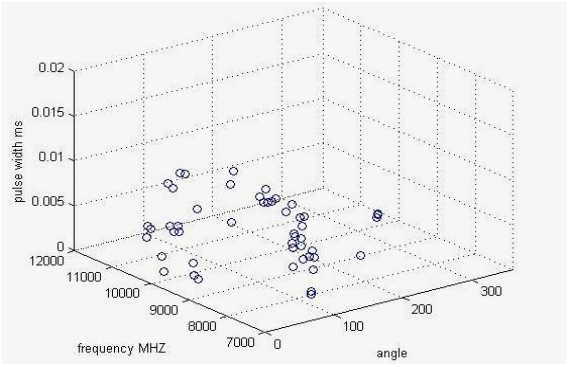
در روش پیشنهادی برای تفکیک پالس‌های ورودی که به نوعی خوشه بندی سلسله مراتبی است، از پیوند نزدیک‌ترین همسایه مانند مرجع [۱۵] با اولویت انتخاب مراکز خوشه‌های متراکم‌تر در تحلیل ماتریس تشابه استفاده شده است [۱]. در این روش بعد از بررسی اولیه و تعیین نقاط متراکم، پالس رهبر به همراه گروه خود از میانگین خوشه متراکم انتخاب می‌شود. از محاسن این روش تفکیک نسبت به الگوریتم‌های مبتنی بر روش تفکیکی k-mean انتخاب اولیه مراکز خوشه‌ها است. برای رسیدن به این هدف از حداکثر کدهای ماتریس تشابه و انتخاب نزدیک‌ترین همسایه در شعاع آستانه موردقبول استفاده شده است. با مقایسه شعاع همسایگی و تعیین کد تشابه، ردیف‌هایی از ماتریس داده PDW که بیشترین تشابه را با بقیه سطرها دارند، تعیین و به‌عنوان مراکز خوشه‌ها معرفی خواهند شد.

هرچند این الگوریتم از روش خوشه‌بندی برمبنای چگالی نیز بهره برداری می کند، علت نامگذاری آن به روش سلسله مراتبی در این نکته است که با انتخاب فاصله شباهت، تعداد خوشه ها می تواند از سطح کل داده‌ها تا سطح داده‌های منفرد تغییر کند.

در برنامه مجزا جهت نمایش الگوریتم همه پارامترهای سطرهای مشابه جمع شده و میانگین آنها به‌عنوان نماینده خوشه معرفی می‌گردد. نتایج حاصل از الگوریتم برای یک نمونه داده ورودی و خروجی‌های تفکیک‌شده در شکل‌های (۶-۵) نمایش داده شده است. داده‌های ورودی مطابق اطلاعات ماتریس PDW شبیه‌سازی و برای آزمایش الگوریتم به‌کار می‌روند. هر سطر ماتریس، پارامترهای PDW مربوط به یک نمونه داده یا یک پالس راداری است.



شکل (۵): منحنی پراکنگی ۹۲ پالس راداری در یک محیط پرتراکم، برای یک بسته از اطلاعات قبل از خوشه‌بندی



شکل (۶): منحنی پراکنگی تعداد ۴۰ هدف مجزا (با توجه به حد آستانه تعریف‌شده) حاصل از خوشه‌بندی.

۳-۲. شبیه‌سازی ورودی

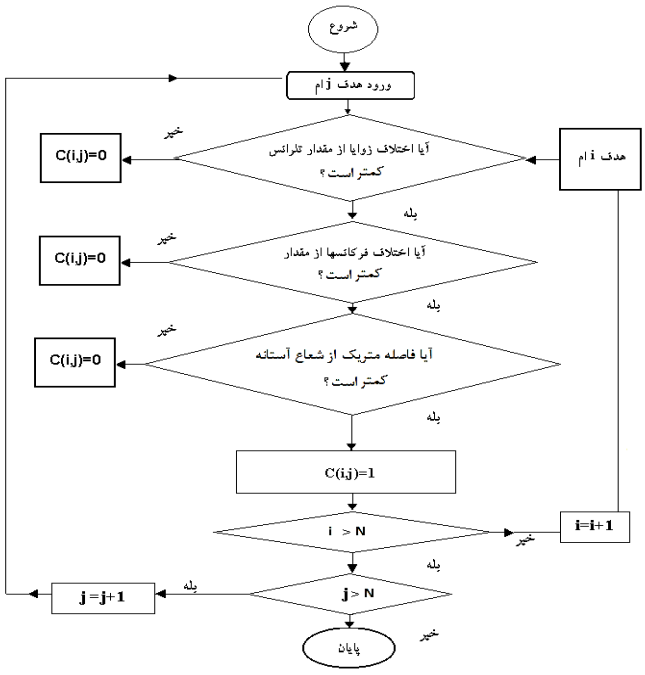
ستون‌های اصلی ماتریس ورودی به‌عنوان خصیصه‌هایی که تفکیک بر اساس آن‌ها انجام می‌شود، به ترتیب از چپ: زاویه برحسب درجه، فرکانس برحسب مگاهرتز، عرض پالس برحسب میکروثانیه و باقی ستون‌ها شامل زمان ورود، دامنه، نوع مدولاسیون و ... می‌باشند.

۳-۳. محاسبه کد تشابه

برای محاسبه کدهای تشابه از فلوچارت شکل (۷) استفاده می‌شود. با گروه‌بندی اولیه چنانچه اختلاف زوایا در سطرهای ماتریس کمتر از محدوده قابل‌قبول، اختلاف فرکانس‌ها و درنهایت اختلاف عرض پالس نیز قابل‌قبول باشد، کد تشابه مربوطه یک وگرنه صفر خواهد شد. در نسخه نهایی به‌جای مقایسه عرض پالس، فاصله متریک مناسب جایگزین و با شعاع همسایگی مقایسه می‌گردد. اگر این فاصله کمتر باشد، به‌عنوان همسایه پذیرفته می‌شود.

به‌عنوان‌مثال، به‌ازای ورودی p پیوست -که یک ماتریس داده با i=۶ نمونه ورودی است- کدهای ماتریس تشابه به نام c با ابعاد ۶×۶ به‌صورت زیر نتیجه می‌شود. بدون توجه به قطر اصلی عدد هر آرایه این ماتریس متقارن، وضعیت تشابه نمونه‌های ردیف یک تا ۶ را با بقیه ورودی‌ها با توجه به شعاع همسایگی نشان می‌دهد. در این مثال سطرهای اول و دوم بیشترین کد ۱ را دارند. آرایه‌های ۱، تشابه نمونه‌های (۲ و ۱) و (۴ و ۱) در سطر اول و (۱ و ۲) و (۳ و ۲) را در سطر دوم نشان می‌دهند. محدوده همسایگی تعریف‌شده در برنامه برای زاویه ۵ درجه، فرکانس ۰/۵٪ و برای عرض پالس ۲۰٪ است [۶].

$$c_{(i\times i)}=\begin{array}{c} \\ ۱ \\ ۲ \\ ۳ \\ ۴ \\ ۵ \\ ۶ \end{array}\begin{array}{c} \begin{array}{cccccc} ۱ & ۲ & ۳ & ۴ & ۵ & ۶ \end{array} \\ \begin{bmatrix} ۱ & ۱ & ۰ & ۱ & ۰ & ۰ \\ ۱ & ۱ & ۱ & ۰ & ۰ & ۰ \\ ۰ & ۱ & ۱ & ۰ & ۰ & ۰ \\ ۱ & ۰ & ۰ & ۱ & ۰ & ۰ \\ ۰ & ۰ & ۰ & ۰ & ۱ & ۰ \\ ۰ & ۰ & ۰ & ۰ & ۰ & ۱ \end{bmatrix} \end{array}$$



شکل (۷): الگوریتم محاسبه ماتریس شباهت مرجع [۲]

۴. توضیح الگوریتم برنامه تفکیک

بعد از محاسبه کد تشابه، ردیف‌هایی که بیشترین تشابه را با بقیه سطرها دارند، با جمع تعداد یک سطرها به دست می‌آید. به‌این‌ترتیب سطری که بیشترین کد تشابه را با همسایه‌های خود دارد، تعیین می‌شود. این سطر تقریباً مرکز خوشه خواهد شد. اگر چند خوشه متراکم پیدا شود، ابتدا خوشه‌ای که کمترین انحراف معیار را در شعاع همسایگی دارد، جدا می‌گردد. برای این کار از فاصله متریک در ماتریس اختلاف‌ها و محاسبه مجموع آن‌ها استفاده شده است.

در برنامه مجزا جهت نمایش این الگوریتم همه پارامترهای سطرهایی که طبق شعاع آستانه، با این سطر (سطری تقریباً میانی و در وسط خوشه متراکم) مشابه هستند، جمع نموده و میانگین پارامترهای آن‌ها به‌عنوان نماینده خوشه معرفی می‌گردد.

بعد از تعیین اولین خوشه متراکم و حذف کدهای خوشه متراکم مربوطه در ماتریس تشابه، جمع کدهای هر سطر ماتریس تشابه محاسبه و همین روال دنبال می‌شود. برای این کار از همان کدهای تولیدشده در مقایسه اولیه استفاده شده تا مقدار پردازش کم شود. درواقع هدف این است که خوشه‌های متراکم‌تر در ابتدا به‌عنوان خوشه‌های مورد تائید و بااهمیت جدا شود.

از محاسن این روش تفکیک نسبت به الگوریتم‌های مبتنی بر روش k-mean، انتخاب بهینه‌مراکز خوشه‌های اولیه است. روش پیشنهادی می‌تواند برای تفکیک انواع دیگر داده نیز به کار رود. مراحل اجرای الگوریتم به‌صورت زیر است:

۱- محاسبه کد تشابه،

۲- محاسبه ماتریس اختلاف‌ها برای سطرهای مشابه هم‌زمان با محاسبه کد تشابه،

۳- تعیین ردیف‌هایی که بیشترین تشابه را با بقیه سطرها دارند (با محاسبه جمع کدهای هر سطر).

۴- از بین خوشه‌های متراکم، انتخاب خوشه‌ای که درمجموع کمترین انحراف معیار را در مقدار شعاع همسایگی دارد.

۵- محاسبه میانگین پارامترها به‌عنوان نماینده خوشه و جدا کردن سطرهای مربوطه در لانه مجزا

۶- صفر نمودن کد مربوطه در ماتریس تشابه بعد از اختصاص به یک خوشه،

۷- تعیین لانه متراکم بعدی و شروع از مرحله سوم،

۸- ادامه تا پایان داده یا رسیدن به حداقل تراکم تعریف‌شده

در این الگوریتم تعداد خوشه‌ها و تکرار برنامه به شعاع همسایگی بستگی دارد. به همین دلیل روش سلسله مراتبی بهینه برای آن انتخاب شده است. جدول مقایسه (۱) به کمک برنامه Cvap متلب به‌دست آمده که در آن از شاخص‌های Dunn و silhouette برای تعیین تعداد بهینه خوشه‌ها و بررسی تعداد خوشه‌های جدول استفاده شده است. نتایج مقایسه بالاتر بودن شاخص Dunn و مقدار بهینه تعداد خوشه‌های روش پیشنهادی را نشان می‌دهد. همچنین شاخص silhouette برای ۳۸ خوشه روش پیشنهادی از روش رهبر (ترتیبی) که به ۴۳ خوشه تفکیک شده، بالاتر است.

جدول (۱): مقایسه الگوریتم پیشنهادی با یک الگوریتم کلاسیک مانند روش سلولی یا حول رهبر

پارامتر	روش رهبر	روش پیشنهادی	حالت بهینه
ni	۲۰۰	۲۰۰	-------
clusters	۴۳	۳۸	-------
DB	۰/۴۷۳۰۷	۰/۳۳۴۵۳	max
Dunn	۰/۰۲۶۵۸۲	۰/۱۰۰۸۲	max
Silhouette	۰/۲۸۷۳۸	۰/۰۳۳۵۸	max
RMSSD	۴/۸۱۸۶	۴/۱۸۴۶	min

۵. نتیجه‌گیری

درروش پیشنهادی برای تفکیک پالس‌های راداری از خوشه‌بندی سلسله مراتبی با پیوند نزدیک‌ترین همسایه با اولویت انتخاب خوشه‌های متراکم‌تر در تحلیل ماتریس تشابه به‌عنوان یک ایده جدید استفاده شده است. پس از محاسبه و تحلیل کدهای ماتریس تشابه، پردازش‌های بعدی بر اساس آن به حداقل رسیده است. با اختصاص بیشترین تعداد همسایه‌ها به یک خوشه بر اساس ماتریس تشابه، نیاز به تکرار خوشه‌بندی در مقایسه با روش‌های معمول که ماتریس فاصله در آنها محاسبه می‌گردد، تا حد زیادی کاهش می‌یابد. در این روش، نقاط متراکم‌تر به‌عنوان مراکز خوشه‌های با اولویت بالاتر انتخاب و خوشه‌های متراکم به‌عنوان رشته پالس‌های راداری جدا می‌شوند. از محاسن این روش تفکیک نسبت به الگوریتم‌های مبتنی بر روش k-mean آن است که انتخاب اولیه مراکز خوشه‌ها در آن با دقت انجام می‌گیرد. مقایسه نتایج خوشه‌بندی یک نمونه داده راداری

الگوریتم پیشنهادی با روش رهبر که یکی از الگوریتم‌های مطرح در خوشه‌بندی پالس‌های راداری است، نشان می‌دهد که این الگوریتم دارای شاخص Dunn و Silhouette بالاتر و انحراف میانگین کمتر است؛ یعنی خوشه‌ها متراکم‌تر و دقت خوشه‌بندی بالاتر است. استفاده از این الگوریتم در سیستم‌های پردازنده ESM می‌تواند به‌طور مؤثری سبب افزایش دقت و سرعت تفکیک پالس‌های راداری گردد.

۶. منابع

۱. قلنـدری، محمـدجواد، آقابابـایی، مجیـد، مـرادی احسـان، خوشه‌بندی به روش نزدیک‌تـرین همسـایه بـا اولویـت انتخـاب خوشه‌های متراکم تر در تحلیل ماتریس تشـابه جهـت تفکیـک پالس‌های راداری، فصلنامه علمی پژوهشـی دریـافنون، دوره ۴، شماره ۴، صفحات ۲۴-۱۳، ۱۳۹۶.

۲. قلندری، محمدجواد، خوگر، احمدرضا، طراحـی و شـبیه‌سـازی الگـوریتم ارتبـاط و دسـته‌بنـدی اهـداف راداری در پردازشـگر ESM، پایان‌نامه کارشناسی ارشد دانشگاه صنعتی مالـک اشـتر، ۱۳۸۱.

۳. احمدی، معین، محامدپور،کمال، تشخیص مدولاسـیون فاصـله تکـرار پـالس‌هـای راداری، مجلـه انجمـن مهندسـین بـرق و الکترونیک ایران، سال ششم، شماره دوم، ۱۳۸۸.

۴. نادری، محمد، موسوی، احمد، جداسازی و نمایش پـالس‌هـای راداری در سیستم‌های پشتیبانی جنگ الکترونیـک، پایـان‌نامـه کارشناسی ارشد، دانشگاه علم و صنعت، ۱۳۷۹.

۵. حسنی، حمید، شیخی، عباس، بیغش، مهرزاد، بررسی روشـهای جداسازی پالس‌های راداری براساسPDW، پایان‌نامه برای اخـذ درجه کارشناسی‌ارشد رشته مهندسی برق– مخابرات سیسـتم، دانشگاه شیراز، ۱۳۹۶.

۶. ناصـری، علـی، سـعادتی مقـدم، گـودرز، ارائـه یـک الگـوریتم هوشـمند بـرای جداسـازی سـیگنال‌هـای متـداخل راداری، هجدهمین کنفرانس مهندسی بـرق ایـران، دانشـگاه صـنعتی اصفهان، صفحات ۱۲۱-۱۱، ۱۳۸۹.

۷. نصرالهی، پویان، خداپرست، فرزاد، حسام‌پـور، کـریم، دادگرنیـا، ابوالفضل، خوشه‌بندی پالس‌های متداخل راداری با اسـتفاده از مدل مخلوط گوسی در محیط‌های متراکم سـیگنال، ششـمین کنفرانس ملی جنگ الکترونیک ایران، مهر، ۱۳۹۲.

۸. ناصـری، علـی، ارائـه الگـوریتم مناسـب بـرای بخـش پـردازش سیسـتم‌هـای شناسـایی رادار ELINT و ESM، نشـریه علمـی پژوهشی انجمن کامپیوتر ایران، مجلد ۱۰، ۱۳۹۱.

9. A. M. Kilincarslan, "A Study on Identification of Radar Emiters," Atılım University, 2011.

10. R. Wiley, "ELINT: The Interception and Analysis of Radar Signals," Artech House, 2006.

11. A. Naseri, H. S. Shahhoseini, and M. Naderi, "Data Clustering by Minimum Difference Tree and PRI Transform. IEEE proceedings, MELECON," pp. 183-187, 2002.

12. Ai-Ling He, De-Guo Zeng, Jun Wang, and Bin Tang, "Multi-Parameter Signal Sorting Algorithm Based on Dynamic Distance Clustering," Journal of Electronic Science and Tecnology of China, vol. 7, no. 3, 2009.

13. Dong-Weon Lee, Jin-Woo Hun, and Won-Don lee, "Adaptive Radar Pulses Clustering Based on Density Cluster Window," 23rd ITS-CSCC,1377-1380, 2008.

14. H. S. Shahhoseini, A. Naseri, and M. Naderi, "A New Matrix Method for Pulse Train Identification," IEEE proceedings, MELECON, pp. 183-187, 2002.

15. M. W. Maier, "Processing Throughput Estimation for Radar Intercept Receivers," IEEE Transactions on Aerospace and Electronic System, vol. 34, no. 1, 1998.

16. C. Malika, N. Ghazzali, V. Boiteau, and A. Niknafs, "NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set," Journal of Statistical Software, vol. 61, pp. 1–36, 2014.

17. P. Hansson, "analysis of some methods for deinterleaving of pulse trains," Stockholm, Sweden, 2007.

18. S. Bailie, "An FPGA Implemention of Incremental Clustering for Radar Pulse Deinterleaving," Master of Science at Northeastern University, Boston, 2010.

19. M. G. S. A, B. Tang, "Sorting radar signal from symmetry clustering perspective," Journal of Systems Engineering and Electronics, vol. 28, no. 4, pp. 690 – 696, 2017.

20. S. Cao, S. Wang, and Y. Zhang, "Density-Based Fuzzy C-Means Multi-Center Re-Clustering Radar Signal Sorting Algorithm," 17th IEEE International Conference on Machine Learning and Applications, 2018.

۷- پیوست

الف- داده‌های فایل نمونه ورودی نامرتب p:

p = AOA F PW .....

20	9320	0.001	5500	400	1	0.002	45
23	9310	0.0011	5010	20	1	0.002	25
25	9300	0.0012	5010	20	1	2.001	25
28	9310	0.0011	2010	20	1	0.001	25
29	9310	0.0011	2010	20	1	0.001	25
33	9300	0.0012	4500	100	1	0.001	20

ب- خروجی تفکیک شده p1:

p1 =

20	9320	0.001	5500	400	0.002
23	9310	0.0011	5010	20	0.002
28	9310	0.0011	2010	20	0.001
29	9310	0.0011	2010	20	0.001
25	9300	0.0012	5010	20	2.001
0	0	0	0	0	0
0	0	0	0	0	0
33	9300	0.0012	4500	100	0.001

Design an Optimum Hierarchical Clustering Algorithm for Radar Pulse Clustering in ESM Processor

M. J. Ghalandari, S. M. Alavi*

Abstract

The first step of an ESM system is extracting features of the received signal in a noisy environment. The next step is clustering and deinterleaving of radar pulses on the basis of data achieved from the previous procedure, and the final step is identifying the emitters and threat factors. Proper selection of initial centers for clusters is the fundamental problem in clustering methods. In the proposed algorithm, first a similarity matrix for the input data based on defined neighborhood radius is formed. Then, with analysis of the similarity matrix and selection of lines with maximum similar codes, denser clusters are separated sequentially. The structure of this algorithm is an optimum hierarchical clustering method. In this algorithm similarity criterion has a very important role. This technique reduces the iterations and computations considerably and separates data into a not predetermined number of clusters using neighborhood radius. The selection priority is with dense sets. One of the advantages of this algorithm compared to the algorithms based on k-mean is careful selection of the initial center of clusters. Results of the proposed method for a data sample consisting of 200 radar pulses is compared with the results of clustering around the leader which is one of the main clustering algorithms in the field of radar pulses and the desired advantage is achieved. In this way, regarding the high stream radar pulse and without iteration requirement, we have optimum pulse strings separation.

Key Words: *radar pulse clustering, Hierarchical clustering algorithm, Radar Pulse Sorting*

* Imam Hossein University - (malavi@ihu.ac.ir) - Writer-in-Charge