

# شناسایی حمله تقلید از سیگنال کاربر اولیه در شبکه‌های رادیوشناختی مبتنی بر طبقه‌بندی غیر پارامتریک بیزین

خاطره اکبری و جمشید ابویی

به صورت فرصت‌طلبانه استفاده می‌کنند [۱]. در واقع کاربران ثانویه در زمان عدم حضور کاربران اولیه از باندهای فرکانسی خالی که به آنها حفره‌های طیف گفته می‌شود، استفاده کرده و باعث استفاده بهینه از این منبع نادر می‌شوند. هدف اصلی CR شناسایی طیف خالی برای استفاده کاربر ثانویه است. تمایز بین سیگنال‌های کاربران اولیه و ثانویه در مرحله طیف‌سنجی، به خصوص در محیط مخابراتی ناامن، نیاز به روش‌های آشکارسازی مورد اطمینان دارد [۲]. گسترش موفقیت‌آمیز شبکه‌های رادیوشناختی وابسته به عملکرد الگوریتم‌های امنیتی در برابر سوء استفاده از این شبکه‌ها است. مقاله‌های مربوط به رادیوشناختی، روش‌های مختلفی برای سنجش طیف و مدیریت طیف را بررسی کردند ولی برخی از آنها، مشکلات امنیتی را در نظر نگرفته‌اند [۳].

در حمله تقلید از سیگنال کاربر اولیه<sup>۴</sup> (PUEA) کاربران مخرب سیگنال‌های کاربران اولیه را تقلید می‌کنند و با اختلال در عملکرد و بی‌عدالتی در میان گره‌های رادیوشناختی، باعث کاهش کارایی شبکه رادیوشناختی می‌شوند. روش‌های طیف‌سنجی در تمایز سیگنال‌های کاربران اولیه در حضور تهدید امنیتی PUEA محدودیت‌هایی به همراه دارند. در میان روش‌های پیشنهادی در مقابله با PUEA، برخی از مطالعات بر "روش‌های دفاعی" تمرکز دارند [۴] در حالی که برخی دیگر به "روش‌های شناسایی" PUEA پرداخته‌اند [۵] تا [۷]. اغلب روش‌های دفاعی در برابر PUEA با فرض وجود مهاجم PUE در محیط، از روشی مشابه روش به نام پرش فرکانسی تصادفی<sup>۵</sup> در مقابله با جیمینگ<sup>۶</sup> بهره می‌برند [۸]. مرجع [۹] با کمک نظریه بازی به شناسایی کاربر مخرب می‌پردازد. برخی از روش‌ها به تأیید هویت فرستنده در لایه‌های بالاتر مدل OSI تمرکز دارند [۱۰] (لایه IP با بهره‌گیری از پروتکل IPsec و لایه انتقال با استفاده از SSL). در [۱۱] هر کاربر ثانویه، در مرحله راه‌اندازی شبکه رادیوشناختی، طیف‌سنجی انجام می‌دهد و دوره‌های ON و OFF هر فرستنده سیگنال کاربر اولیه را جمع‌آوری می‌کند. با مقایسه الگوی فعالیت سیگنال ارسالی مجهول با الگوی فعالیت کاربر اولیه، کاربر مخرب شناسایی می‌شود. در روش‌های شناسایی PUEA، اغلب تکنیک‌ها از اطلاعات مکانی کاربران اولیه بهره می‌برند. در روش‌های مبتنی بر مکان‌یابی، فرض می‌شود کاربران اولیه دارای مکانی ثابت هستند و تغییرات کانال نیز قابل توجه نیست. همچنین کاربران مخرب نمی‌توانند توان ارسالی خود را تغییر دهند و این در حالی است که با تغییر مکان کاربر اولیه و یا تغییر پیکربندی، مکان متغیر خواهد بود. افزون بر این اگر کاربر مخرب توانایی تطبیق توان ارسالی سیگنال خود را با کاربران اولیه واقعی داشته باشد یا تغییر مکان دهد، می‌تواند این دو پارامتر را تقلید کند. بنابراین پیاده‌سازی چنین روش‌هایی را با چالش روبه‌رو می‌کند. از سوی

چکیده: رادیوشناختی به عنوان یک فناوری کلیدی برای مقابله با کمبود طیف فرکانسی در شبکه‌های بی‌سیم به طور گسترده مورد توجه قرار گرفته است. یکی از چالش‌های مهم در تحقق شبکه‌های رادیوشناختی، امنیت این نوع شبکه‌ها است. از مهم‌ترین این تهدیدها، حمله تقلید از سیگنال کاربر اولیه است، بدین معنی که کاربر مخرب سعی دارد سیگنالی مشابه با سیگنال کاربر اولیه ارسال کند تا کاربران ثانویه را فریب داده و از ارسال سیگنال‌های این کاربران در حفره‌های طیفی جلوگیری کند و ضمن ایجاد ترافیک در شبکه، با به دست آوردن باند فرکانسی خالی، اطلاعات خود را ارسال کند. در این مقاله، روشی برای شناسایی حمله تقلید از سیگنال کاربر اولیه پیشنهاد می‌گردد که با خوشه‌بندی سیگنال‌های ارسالی کاربران اولیه و کاربران مخرب، این سیگنال‌ها را متمایز می‌کند. در این روش، تعداد سیگنال‌های ارسالی در محدوده شبکه رادیوشناختی در طول خوشه‌بندی سیگنال‌ها به دست می‌آید. با به کارگیری روش طبقه‌بندی مدل مخلوطی فرایند دیرپیشه که بر اساس روش غیر پارامتریک بیزین می‌باشد، سیگنال‌های اولیه فعال در محیط طبقه‌بندی می‌شوند. همچنین برای دستیابی به سرعت همگرایی بالاتر در الگوریتم، روش فرایند رستوران چینی برای مقداردهی اولیه و نمونه‌برداری غیر یکنواخت جهت انتخاب پارامتر خوشه‌ها به الگوریتم اعمال می‌گردد.

کلیدواژه: امنیت، حمله تقلید از سیگنال کاربر اولیه، خوشه‌بندی غیر پارامتریک بیزین، شبکه رادیوشناختی.

## ۱- مقدمه

پیشرفت سریع ارتباطات بی‌سیم و نیاز به نرخ داده بالا، تقاضای طیف را افزایش داده و از سوی دیگر باندهای دارای مجوز دارای باند مجزا هستند. این در حالی است که برخی از باندهای فرکانسی در همه زمان‌ها به طور کامل استفاده نمی‌شود. به همین دلیل ایده فرصت استفاده از این باندها در شبکه‌های رادیوشناختی مطرح شد. رادیوشناختی<sup>۱</sup> (CR) فرصت ارسال در باندهای دارای مجوز را بدون ایجاد تداخل با کاربران مجوزدار فراهم می‌کند. در شبکه رادیوشناختی کاربران به دو دسته تقسیم می‌شوند: (۱) کاربران اولیه (PUs) که مجوز برای استفاده از بخش خاصی از طیف را دارند، اما در طول یک بازه زمانی خاص از تمام باند فرکانسی خود استفاده نمی‌کنند و (۲) کاربران ثانویه<sup>۲</sup> (SUs) که از بخش‌هایی از طیف

این مقاله در تاریخ ۲۴ مهر ماه ۱۳۹۶ دریافت و در تاریخ ۲۸ خرداد ماه ۱۳۹۷ بازنگری شد.

خاطره اکبری، دانشکده مهندسی برق، دانشگاه یزد، یزد، ایران، (email: khatereh.akbari@gmail.com).

جمشید ابویی (نویسنده مسئول)، دانشکده مهندسی برق، دانشگاه یزد، یزد، ایران، (email: abouei@yazd.ac.ir).

4. Primary User Emulation Attack
5. Random Frequency Hopping
6. Jamming

1. Cognitive Radio
2. Primary Users
3. Secondary Users

درصد شناسایی درست تعداد سیگنال‌های فعال در محیط مورد بررسی قرار خواهد گرفت و طرح ارائه‌شده با روش‌های خوشه‌بندی X-means و Mean-Shift مقایسه می‌گردد. نشان داده می‌شود که روش ارائه‌شده، دقت خوشه‌بندی را نسبت به دو روش Mean-Shift و X-Means بهبود می‌دهد. مدل مخلوطی فرایند دیریشله<sup>۲</sup> (DPMM) نیاز به تکرار برای رسیدن به مقدار پایدار دارد. برای افزایش سرعت همگرایی، روش فرایند رستوران چینی<sup>۳</sup> (CRP) در مقداردهی اولیه ارائه خواهد شد و نمونه‌برداری غیر یکنواخت جهت انتخاب پارامتر نمونه‌برداری اعمال می‌شود. در بخش چهارم با جمع‌بندی مطالب، مقاله را به پایان می‌رسانیم.

## ۲- طرح پیشنهادی طبقه‌بندی سیگنال‌های اولیه در شبکه‌های رادیوشناختی

### ۲-۱ مدل سیستم شبکه رادیوشناختی

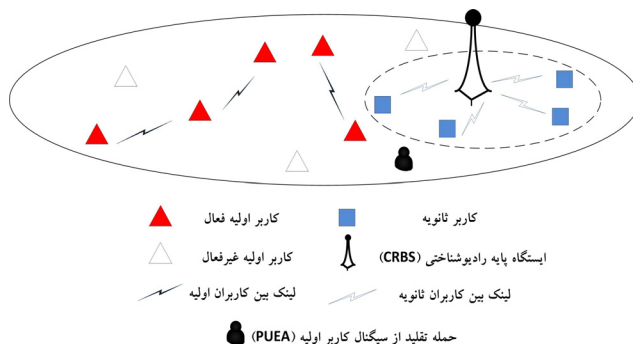
یک شبکه رادیوشناختی شامل  $R$  کاربر ثانویه  $SU_1, \dots, SU_R$  و ایستگاه پایه رادیوشناختی<sup>۴</sup> (CRBS) را در نظر بگیرید (شکل ۱). در محدوده شبکه رادیوشناختی،  $P$  کاربر اولیه واقعی  $PU_1, \dots, PU_p$  وجود دارند. تعداد کاربران اولیه فعال با  $K \leq P$  مشخص می‌شود که در طول زمان تغییر می‌کند. فرض می‌کنیم که پارامتر  $K$  برای شبکه رادیوشناختی نامعلوم است و باند فرکانسی توسط CRBS طیف‌سنجی می‌شود. کاربران اولیه واقعی می‌توانند در  $L$  کانال غیر هم‌پوشان، سیگنال ارسال کنند. فرض می‌شود کانال‌ها دارای پهنای باند یکسان  $W$  هستند. مدولاسیون‌های M-QAM، QPSK و BPSK برای سیگنال‌های کاربر اولیه قابل استفاده است.

کاربران ثانویه اجازه دارند که از حفره‌های طیف برای ارسال اطلاعات خود استفاده کنند. کاربران ثانویه باید کانال مورد استفاده خود را زمانی که کاربر اولیه شروع به ارسال مجدد می‌کند، ترک کرده و کانال‌های خالی دیگر را برای ارسال خود استفاده کنند. هر کاربر ثانویه برای بررسی حضور یا عدم حضور کاربر اولیه در کانال مورد استفاده خود، به صورت فردی طیف‌سنجی انجام می‌دهد و اطلاعات خود را به CRBS ارسال می‌کند. در شبکه رادیوشناختی در نظر گرفته شده، روش طیف‌سنجی مورد استفاده آشکارسازی انرژی است که روشی سریع و با پیاده‌سازی ساده است و نیازمند اطلاعات پیشینی از کاربران اولیه نمی‌باشد. طیف‌سنجی به صورت دوره‌ای در طول زمان  $\tau$  اجرا می‌شود و  $T - \tau$  مدت زمان ارسال سیگنال است. فرض می‌شود که در هر کانال یک فرستنده کاربر اولیه می‌تواند ارسال داشته باشد. یک مهاجم PUE در شبکه را در نظر بگیرید که خصوصیات کاربر اولیه را در هر بازه زمانی، فقط در یک کانال تقلید می‌کند.

### ۲-۲ استخراج ویژگی

سیگنال  $z(t)$  در طول طیف‌سنجی، در باند فرکانسی در دسترس شبکه رادیوشناختی شامل  $L$  کانال دریافت می‌شود که به صورت رابطه زیر نمایش داده می‌شود

$$z[n] = s[n] + \omega[n] = \sum_{i=1}^K s_i[n] e^{j\pi f_i n} + \omega[n] \quad (1)$$



شکل ۱: شبکه رادیوشناختی با حضور یک مهاجم PUE.

دیگر، اغلب روش‌های شناسایی حمله PUE، تعداد و نوع سیگنال‌های اولیه فعال را ثابت و معلوم فرض می‌کنند، این در حالی است که در شبکه رادیوشناختی، اطلاعات مربوط به تعداد و نوع سیگنال‌های اولیه فعال در محیط مخابراتی معلوم نمی‌باشد و در طول زمان تغییر می‌کند و ممکن است سیگنال‌های جدیدی از جمله سیگنال‌های کاربران اولیه واقعی یا کاربران مخرب در محیط ارسال شود. در واقع، ایستگاه پایه رادیوشناختی که فرایند شناسایی PUE را انجام می‌دهد، اطلاعات پیشینی از تعداد کاربران اولیه یا مخرب فعال در محیط ندارد. بنابراین محیط رادیویی ماهیت غیر پارامتریک دارد. در ماهیت غیر پارامتریک، اطلاعاتی از تعداد کاربران اولیه نداریم. در این راستا، [۱۲] از شناسه‌های رادیویی از جمله اختلاف فرکانس حامل و اختلاف شیفیت فازی که مختص فرستنده مخابراتی می‌باشند، استفاده و تعداد فرستنده‌های کاربر اولیه را مشخص می‌کند. با کمک آدرس MAC فرستنده، کاربر مخرب از کاربران اولیه واقعی متمایز می‌شود.

در این مقاله، یک طرح برای تشخیص PUE ارائه می‌شود که بدون دانش پیشینی از سیستم‌های فعال اولیه، خوشه‌بندی سیگنال‌های کاربران اولیه انجام می‌گیرد. ویژگی‌های فرکانس حامل و نرخ سمبل از سیگنال‌های ارسالی استخراج و سیگنال‌های ارسالی با روش خوشه‌بندی متمایز می‌شوند. برای تشخیص کاربر مخرب، شبکه رادیوشناختی الگوریتم شناسایی را به صورت پیوسته در بازه زمانی مشخص تکرار می‌کند تا رفتار کاربران اولیه به دست آید. در واقع، کاربر مخرب ویژگی‌های طیفی سیگنال ارسالی خود را تغییر می‌دهد تا در کانال‌های دیگر نیز امکان ارسال داشته باشد. هدف کاربر مخرب در ارسال به کانال‌های مختلف، میان کاربر مخرب و کاربر اولیه اصلی تمایز ایجاد می‌کند. فرایند شناسایی کاربر مخرب به صورت آفلاین اجرا می‌شود و کاربر مخرب بر اساس رفتار آن در ارسال به کانال‌های دیگر شناسایی می‌شود. در بخش دوم مقاله، طرح شناسایی حمله PUE مبتنی بر طبقه‌بندی سیگنال‌های کاربران اولیه با روش غیر پارامتریک بیزین و تئوری آن و مدل سیستم مورد نظر را شرح می‌دهیم و به روش نمونه‌برداری گیبس<sup>۱</sup> (GS) اصلاح‌شده که از قسمت‌های اصلی مدل مخلوطی فرایند دیریشله است می‌پردازیم. یک سناریو از طبقه‌بندی سیگنال‌های کاربران اولیه در شبیه‌ساز MATLAB را مورد ارزیابی قرار می‌دهیم. در گام نخست، با تولید سیگنال‌ها در فرکانس‌های متفاوت و استخراج ویژگی‌های آن، ماتریس ویژگی را برای طبقه‌بندی‌کننده تولید می‌کنیم. در گام دوم با مشخص کردن تعداد سیگنال‌های ارسالی و نوع آن، سیگنال‌های رادیویی محیط بی‌سیم را شناسایی می‌کنیم. در بخش سوم مقاله، دقت خوشه‌بندی با معیار کارایی

2. Dirichlet Process Mixture Model  
3. Chinese Restaurant Process  
4. Cognitive Radio Base Station

1. Gibbs Sampling

$$\theta^* = \arg \max_{\theta} P(\theta | \{y_i\}_{i=1}^N) = \arg \max_{\theta} \frac{P(\{y_i\}_{i=1}^N | \theta) G(\theta)}{P(\{y_i\}_{i=1}^N)} \quad (5)$$

اطلاعات مربوط به تعداد سیگنال‌های اولیه فعال در محیط مخابراتی معلوم نیست و بنابراین برای توزیع پیشین  $G(\theta)$  در تخمین‌زننده MAP، توزیع ثابتی روی تعداد خوشه‌ها نمی‌توان در نظر گرفت و توزیع خوشه‌ها بر اساس  $\{y_i\}_{i=1}^N$  تعریف می‌شود. به ازای مجموعه مشاهدات  $\{y_i\}_{i=1}^N$ ، مجموعه پارامترهای خوشه  $\{\theta_k^*\}_{k=1}^K$  تعریف می‌شود. در روش غیر پارامتریک بی‌زین، فرض می‌شود توزیع درست‌نمایی معلوم است. پارامترهای  $\theta_k^*$  و  $K$  مجهول هستند و مسأله طبقه‌بندی در چهارچوب بی‌زین<sup>۲</sup> خواهد بود. طبقه‌بندی‌کننده غیر پارامتریک بی‌زین بر اساس مدل مخلوطی فرایند دیریشله (DPMM) مدل می‌شود و فرض می‌کند تعداد خوشه‌ها یک متغیر تصادفی است و اطلاعات پیشینی از تعداد خوشه‌ها ندارد.

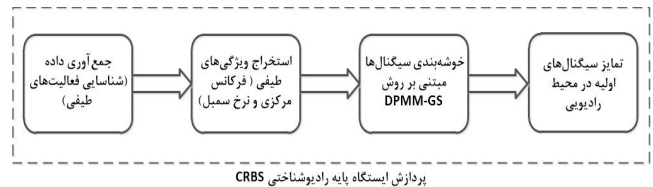
فرایند دیریشله: فرض می‌کنیم که مشاهدات  $\{y_i\}_{i=1}^N$  از یک مدل مخلوطی به دست می‌آید که مؤلفه‌های مدل مخلوطی  $\{\theta_i\}_{i=1}^N$  از توزیع  $G(\theta)$  نتیجه گرفته می‌شود. توزیع احتمال گسسته  $G$  تحقق از فرایند دیریشله  $DP(\alpha, G)$  با پارامتر تمرکز  $\alpha$  و توزیع پایه  $G$  است [۱۳]. از توزیع دیریشله به عنوان توزیع پیشین در مدل‌سازی بی‌زین استفاده می‌شود، چرا که توزیع دیریشله، مزدوج پیشین برای توزیع چندجمله‌ای و توزیع دسته‌ای است و تابع چگالی احتمال توزیع دیریشله، وزن‌های  $K$  مؤلفه مدل مخلوطی یا خوشه‌ها می‌باشد  $(\omega_1, \dots, \omega_K)$ . مدل مخلوطی فرایند دیریشله را می‌توان به صورت زیر بیان نمود [۱۴]

$$\begin{aligned} G &\sim DP(\alpha, G) \\ \theta_i^* &\sim G \\ y_i | \theta_k^* &\sim f(y_i | \theta_k^*) \end{aligned} \quad (6)$$

مسأله خوشه‌بندی، تخمین پارامترهای خوشه  $\theta_i$  برای هر مشاهده  $y_i$  را می‌توان از طریق تخمین MAP در (۵) به دست آورد ولی چون تابع توزیع توأم  $P(\theta | \{y_i\}_{i=1}^N)$  معلوم نیست، باید تخمین  $\theta_i$  را به صورت جداگانه محاسبه کنیم که برای محاسبه آن از روش‌های مونت‌کارلو استفاده می‌کنیم. از روش‌های ساده مونت‌کارلو برای به دست آوردن تخمین توزیع توأم شرطی، نمونه‌برداری گیبس است که از توزیع شرطی حاشیه‌ای  $P(\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_N, \{y_i\}_{i=1}^N)$  نمونه می‌گیرد. توزیع پسین مورد نظر برای تخمین پارامترها به صورت رابطه زیر می‌باشد [۱۵]

$$\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_N, \{y_i\}_{i=1}^N = \begin{cases} \theta_k^* & , p_k = \frac{n_k f(y_i | \theta_k^*)}{\alpha f(y_i) + \sum_{k=1}^K n_k f(y_i | \theta_k^*)} \\ \sim f(\theta_i | y_i) & , p = \frac{\alpha f(y_i)}{\alpha f(y_i) + \sum_{k=1}^K n_k f(y_i | \theta_k^*)} \end{cases} \quad (7)$$

پارامتر  $\theta_i$  مربوط به مشاهده  $y_i$  با احتمال  $p_k$  به خوشه  $k$  ام یا با احتمال  $p$  به خوشه جدید تخصیص می‌یابد. تابع درست‌نمایی  $f(y_i | \theta_k^*)$  معلوم است.  $f(y_i)$  توزیع حاشیه‌ای  $y_i$  می‌باشد که برای



شکل ۲: بلوک دیاگرام روش ارائه‌شده شناسایی سیگنال‌های اولیه جهت تمایز کاربران واقعی و مخرب در شبکه‌های رادیوشناختی.

که  $s_i[n]$  سیگنال باند پایه کاربر اولیه  $i$ ام در فرکانس  $f_{c_i}$  است و  $\omega[n]$  نویز گوسی جمع‌شونده AWGN است.  $z[n](n=0, \dots, N_s-1)$  سیگنال دریافتی نمونه‌برداری شده است به گونه‌ای که  $N_s = \tau/T_s$  تعداد نمونه‌های سیگنال در طول طیف‌سنجی  $\tau$  با زمان نمونه‌برداری  $T_s$  می‌باشد. سیگنال  $z[n]$  مجموع سیگنال‌های دریافتی از کاربران اولیه و  $\omega[n]$  سیگنال نویز است. در اینجا منظور از سیگنال‌های دریافتی از کاربران اولیه، سیگنال‌های شناسایی‌شده هستند که در دسته سیگنال اولیه قرار گرفته‌اند که می‌تواند سیگنال اولیه واقعی یا مخرب باشد

$$Z = \sum_{k=1}^K S_k + W \quad (2)$$

ویژگی‌های مناسب برای استخراج با هدف خوشه‌بندی سیگنال‌های کاربران اولیه به صورت زیر است:  
**فرکانس مرکزی:** برای تعیین فرکانس‌های مرکزی از تخمین‌زننده طیف توان استفاده شده است. برای چگالی طیف توان، پنجره مستطیلی با طول پنجره  $L_\omega$  قرار داده می‌شود

$$S_{zz}[k] = \frac{1}{L_\omega N_s} \sum_{l=-\frac{L_\omega-1}{2}}^{\frac{L_\omega-1}{2}} |Z[k+l]|^2 \quad (3)$$

**فرکانس دوری:** در بسیاری موارد، تابع خودهمبستگی سیگنال‌های مخابراتی متناوب با زمان هستند و بنابراین سیگنال دریافتی  $z(t)$  دارای خاصیت ایستادگی خواهد بود. خاصیت ایستادگی سیگنال می‌تواند ناشی از نرخ سمبل سیگنال باشد. فرکانس‌های دوری  $\alpha$  برای هر نوع سیگنال یکتا است و به عنوان شناسه طیفی برای شناسایی نوع سیگنال استفاده می‌شود. فرکانس‌های دوری منطبق بر نرخ سمبل سیگنال را از تابع همبستگی طیفی به دست می‌آوریم

$$S_{zz}^\alpha[k] = \frac{1}{L_\omega N_s} \sum_{l=-\frac{L_\omega-1}{2}}^{\frac{L_\omega-1}{2}} Z[k-l] Z^*[k+\alpha-l] \quad (4)$$

### ۳- طبقه‌بندی سیگنال‌های رادیویی

با توجه به شکل ۲ با استخراج ویژگی‌های فرکانس حامل و نرخ سمبل از سیگنال‌ها، مشاهدات  $\{y_i\}_{i=1}^N$  تولید می‌شود. هر مشاهده  $y_i$  منطبق بر یک خوشه با پارامتر  $\theta_i$  است. مسأله اصلی طبقه‌بندی، مشخص کردن پارامترهای  $\{\theta_i\}_{i=1}^N$  با در دسترس بودن مشاهدات  $\{y_i\}_{i=1}^N$  با در نظر گرفتن توزیع پیشین برای پارامتر  $\theta$  بر اساس بیشینه‌گر احتمال پسین (MAP) مسأله خوشه‌بندی تعریف می‌شود. در (۵)،  $P(\theta | \{y_i\}_{i=1}^N)$  احتمال پسین خوشه‌ها با وجود مشاهدات  $\{y_i\}_{i=1}^N$  است.  $G(\theta)$  توزیع پیشین برچسب‌های خوشه‌ها و  $P(\{y_i\}_{i=1}^N | \theta)$  احتمال شرطی خوشه است

احتمال حاشیه‌ای  $f(y_i)$  از  $(\lambda)$  با در نظر گرفتن توزیع پیشین  $G(\theta)$  به دست می‌آید

$$f(y_i) = \int_{\theta} f(y_i|\theta)G(\theta)d\theta$$

$$G(\theta) = NIW(\mu_k, \Sigma_k) = N(\mu, \frac{\sum_k}{R})IW(\Psi, \nu) \quad (12)$$

### ۴-۲ افزایش سرعت همگرایی طبقه‌بندی کننده DPMM

DPMM به خاطر ساختار غیر پارامتریک خود، ابزاری مناسب برای طبقه‌بندی سیگنال است، اگرچه باعث پیچیدگی محاسباتی بالایی می‌شود چون نیاز به تعداد تکرار زیاد در نمونه‌برداری گیبس دارد. از این رو برای کاهش تعداد تکرار مورد نیاز برای نمونه‌برداری روش نمونه‌برداری غیر یکنواخت و در مرحله مقداردهی اولیه روش CRP ارائه شده است.

### ۴-۵ مقداردهی اولیه الگوریتم

برای نمونه‌برداری گیبس نیاز به مقداردهی اولیه و تخصیص اولیه خوشه‌ها به مشاهدات است که چند حالت مورد بررسی قرار می‌گیرد. نمونه‌بردار گیبس با هر مقداردهی اولیه دلخواه، به توزیع ایستا همگرا خواهد شد و به ازای تعداد numofiterGS برای هر پارامتر تکرار می‌شود. با این حال مقداردهی اولیه بر تعداد تکرار لازم برای همگرایی تأثیر خواهد داشت. از این رو چند حالت مقداردهی اولیه برای این الگوریتم در نظر می‌گیریم:

۱) N-Initial:  $(K^{(i)} = N)$  در این حالت، هر کدام از مشاهدات به خوشه‌های مجزا تخصیص پیدا می‌کنند، یعنی هر مشاهده یک خوشه به حساب می‌آید  $(\theta_1^{(i)}, \dots, \theta_N^{(i)}) = (y_1, \dots, y_N)$ .

۲) CRP-Initial:  $(K^{(i)} = K)$  بر اساس روش رستوران چینی (CRP)، مشاهدات را به  $K$  خوشه تخصیص می‌دهیم، یعنی در ابتدا مشاهده اول  $y_1$  را به اولین خوشه اختصاص می‌دهیم. مشاهده دوم بر اساس  $(\gamma)$  به خوشه‌ها اختصاص می‌یابد، یعنی ممکن است به خوشه‌های تشکیل‌شده تخصیص یابد یا خوشه جدید تشکیل دهد و به همین ترتیب به ازای همه مشاهدات بر اساس روش CRP عمل می‌شود. با این روش مشاهدات مشابه با احتمال بیشتری در همان مرحله اول، به خوشه یکسان تخصیص می‌یابند و روند همگرایی سریع‌تر خواهد شد.

### ۴-۶ انتخاب پارامتر جهت نمونه‌برداری

اگر یک مشاهده  $y_i$  به خوشه  $k$  با تعداد المان‌های  $n_k^{(j)}$  تخصیص یابد در حالی که  $n_k^{(j)}$  بزرگ باشد، این احتمال زیاد است که در مرحله نمونه‌برداری بعدی دوباره به همان خوشه تخصیص پیدا کند. دلیل این روند، را می‌توان در  $(\gamma)$  یافت. در این رابطه می‌بینیم که احتمال انتساب مشاهده  $i$ ام به خوشه  $k$ ام با  $n_k^{(j)}$  متناسب است. بنابراین می‌توان نتیجه گرفت که نمونه‌برداری از این نوع مشاهدات در هر تکرار الگوریتم با احتمال زیادی، تغییر در خوشه منتسب ایجاد نمی‌کند.

۱) Sequential-select: در این حالت، همه پارامترهای  $\{\theta_i^{(j)}\}_{i=1}^N$  به ازای  $i = 1, \dots, N$  به صورت ترتیبی انتخاب می‌شوند.

۲) Uniform-select: در حالت دوم، پارامترهای  $\{\theta_i^{(j)}\}_{i=1}^N$  به ازای  $i = 1, \dots, N$  به صورت تصادفی با توزیع یکنواخت انتخاب می‌شوند و در این حالت همه پارامترها با احتمال یکسان انتخاب می‌شوند.

ارزیابی رابطه، باید  $f(\theta_i|y_i)$  به فرم بسته به دست بیاید. خروجی الگوریتم گیبس، خوشه‌هایی است که هر خوشه یک نوع سیگنال ارسالی اولیه را در فضای ویژگی نگاشت می‌کند (الگوریتم ۱):

ورودی: مجموعه مشاهدات  $\{y_i\}_{i=1}^N$ ، پارامتر تمرکز  $\alpha$ ، تعداد تکرار مرحله نمونه‌برداری numofiterGS.

مرحله مقداردهی اولیه: به یکی از حالت‌های زیر می‌توان برای این الگوریتم مقداردهی اولیه کرد  $(\theta_1^{(i)}, \dots, \theta_N^{(i)}) = (y_1, \dots, y_N)$ .

$$K^{(i)} = N \quad (1)$$

$$K^{(i)} = K \quad (2)$$

مرحله نمونه‌برداری از توزیع پسین: به‌ازای تعداد تکرار (numofiterGS:  $j = 1$ )

این مرحله اجرا می‌شود تا توزیع پسین  $\{\theta_i\}_{i=1}^N$  به توزیع ایستا همگرا شود. برای این مرحله، طبق حالت‌های عنوان‌شده در زیر، می‌توان پارامترهای خوشه‌ها را انتخاب کرد، سپس  $(\gamma)$  محاسبه شده و نمونه‌برداری می‌شود. برچسب‌های هر مشاهده و تعداد خوشه‌ها در هر تکرار الگوریتم به روز می‌شود.

Sequential-select (۱)

Uniform-select (۲)

Nonuniform-select (۳)

خروجی: برچسب‌های هر بردار ویژگی و تعداد سیگنال‌های ارسالی در محیط  $\{\theta_i^{(j)}\}_{i=1}^N$  و  $K$ .

الگوریتم ۱: NonParametric Bayesian Clustering Algorithm using Gibbs Sampling.

پارامترهای خوشه‌ها بر اساس مدل مخلوطی گوسی بردار میانگین و کواریانس  $\mu_k$  و  $\Sigma_k$  خواهند بود که هر دو پارامتر خوشه‌ها را نامعلوم بر اساس مدل غیر پارامتریک در نظر می‌گیریم

$$\theta_k^* = (\mu_k, \Sigma_k) \quad (8)$$

$$y_i | \theta_k^* \sim N(\mu_k, \Sigma_k)$$

برای توزیع پارامترهای  $\mu_k$  و  $\Sigma_k$  از توزیع نرمال و وارون ویشارت (IW) استفاده می‌شود [۱۶]

$$\mu_k \sim N(\mu, \frac{\Sigma_k}{R}) \quad (9)$$

$$\Sigma_k \sim InverseWishart(\Psi, \nu)$$

پارامتر  $\Psi$  تخمین اولیه از ماتریس کواریانس خوشه‌ها و پارامتر  $\nu$  درجه آزادی توزیع است. برای محاسبه احتمال پسین پارامترهای خوشه‌ها جهت نمونه‌برداری، با توجه به  $(\gamma)$  می‌توان رابطه زیر را برای توزیع مشاهدات نرمال به دست آورد

$$\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_N, \{y_i\}_{i=1}^N = \begin{cases} \theta_k^* & , p_k \propto n_k f(y_i | \theta_k^*) \\ \sim f(\theta_i | y_i) & , p_i \propto \alpha f(y_i) \end{cases} \quad (10)$$

به طوری که مجموع احتمالات باید برابر با یک باشد  $(p_i + \sum_{k=1}^K p_k = 1, k \in \{1, \dots, K\})$ . در واقع  $n_k$  برابر با تعداد مشاهدات خوشه  $k$ ام است و احتمال شرطی خوشه با توجه به توزیع نرمال به فرم زیر به دست می‌آید

$$f(y_i | \theta_k^*) = (\pi)^{\frac{D}{2}} \det(\Sigma_k)^{-\frac{1}{2}} \times \exp(-\frac{1}{2}(y_i - \mu_k)^T inv(\Sigma_k)(y_i - \mu_k)) \quad (11)$$

ویژگی‌های سیگنال‌ها را به خوشه صحیح مشخص می‌نماید که به صورت نسبت تعداد بردارهای ویژگی تخصیص داده شده به خوشه صحیح بر تعداد کل بردارهای ویژگی سیگنال‌ها ( $N$ ) محاسبه می‌شود.

(۳) تخصیص به خوشه غیر صحیح (FCA): تخصیص هر کدام از بردار ویژگی‌های سیگنال‌ها به خوشه معتبر غیر از خوشه خودش را نشان می‌دهد که از نسبت تخصیص تعداد بردارهای ویژگی تخصیص داده شده به خوشه غیر صحیح (FA) بر تعداد کل بردارهای ویژگی ( $N$ ) به دست می‌آید. دو معیار FCA و ACC در صورتی که خوشه غیر معتبری شناسایی نشده باشد معادل هستند.

(۴) واگرایی کولیک-لیبلر<sup>۱</sup> (KLD): شاخص اندازه‌گیری تفاوت بین دو توزیع احتمال می‌باشد [۱۷]. اگر توزیع دو خوشه به صورت  $N(\mu_1, \Sigma_1)$  و  $N(\mu_2, \Sigma_2)$  باشد، شاخص KLD با (۱۳) به دست می‌آید

$$KLD = \frac{1}{2} [\text{trace}(\Sigma_2^{-1} \Sigma_1) + (\mu_2 - \mu_1)^T \Sigma_2^{-1} (\mu_2 - \mu_1) + \ln \left( \frac{|\Sigma_1|}{|\Sigma_2|} - D \right)] \quad (13)$$

که  $\mu_k$  و  $\Sigma_k$  بردار میانگین و ماتریس کواریانس هر کدام از توزیع‌های احتمال خوشه است و  $D$  بعد بردار را نشان می‌دهد که برابر با ۲ می‌باشد. شاخص KLD، فاصله بین خوشه‌ها را محاسبه می‌کند و عملکرد طبقه‌بندی‌کننده را به ازای تغییر فاصله خوشه‌ها از یکدیگر به دست می‌آورد. هرچه این شاخص بزرگ‌تر باشد، دو خوشه از یکدیگر فاصله دارند و بهتر خوشه‌بندی خواهند شد. ما معیار KLD را برای سیگنال‌های آشکار شده جدید و سیگنال‌های شناسایی شده، محاسبه می‌کنیم. با این معیار، سیستم شناسایی حمله PUE می‌تواند برای سیگنال‌های جدید فعال در محیط تصمیم‌گیری کند و آنها را در خوشه‌های شناسایی شده یا خوشه جدید قرار دهد. بردار میانگین و ماتریس کواریانس هر خوشه برای تشخیص نوع سیگنال ارسالی استفاده می‌شود. در شبکه رادیوشناختی مورد نظر برای بررسی روش DPMM-GS در خوشه‌بندی سیگنال‌های ارسالی از کاربران اولیه، فرض می‌کنیم نمونه‌های سیگنال‌های دریافتی از کاربران اولیه فعال در محیط بی‌سیم اعم از کاربر اولیه واقعی یا مهاجم PUE دارای توزیع نرمال است و بردارهای ویژگی استخراج شده از این سیگنال‌ها، از مدل مخلوطی نرمال گرفته شده است  $N(\mu_k, \Sigma_k)$ .

باید خاطر نشان کرد در روش خوشه‌بندی توزیع مشاهدات با توزیع مخلوط گوسی مدل شده و این مدل به توزیع مشاهدات حساس نیست. روش ارائه شده برای توزیع مشاهدات غیر گوسی نسبت به گوسی عملکرد ضعیف‌تری دارد. با این وجود از عملکرد پایداری برخوردار است و با توزیع مشاهدات غیر گوسی نیز نسبت به روش‌های X-Means و MeanShift عملکرد بهتری دارد. روش‌های X-Means و MeanShift نیز برای فضای ویژگی گوسی مدل شده‌اند.

در شبیه‌سازی، روش خوشه‌بندی DPMM-GS را با فرض تعداد سیگنال‌های ارسالی برابر با  $K=2$  و ابعاد فضای ویژگی  $D=2$  و تعداد مشاهدات حس شده در طول زمان طیف‌سنجی برابر با  $N=100$ ، به ازای سه معیار احتمال آشکارسازی صحیح، دقت خوشه‌بندی و تخصیص به خوشه غیر صحیح مقایسه می‌کنیم. شکل ۳ احتمال

(۳) Nonuniform-select: برای کاهش پیچیدگی محاسباتی یک روش انتخاب پارامتر ارائه می‌کنیم که پارامترهای خوشه را به طور مؤثر در هر تکرار نمونه‌برداری گیبس انتخاب می‌کند و باعث می‌شود سریع‌تر به مقدار نهایی همگرا شود و کاهش قابل توجه محاسبات طبقه‌بندی‌کننده DPMM را نتیجه دهد. این حالت با انتخاب هوشمندانه‌تر پارامترها در هر تکرار الگوریتم، سعی بر کاهش تکرار مورد نیاز برای همگرایی الگوریتم دارد و پارامترهایی را برای نمونه‌برداری انتخاب می‌کند که احتمال بیشتری دارد که خوشه متناسب به آن تغییر کند. برخلاف حالت دوم، پارامترهای  $\{\theta_i^{(j)}\}_{i=1}^N$  به ازای  $i=1, \dots, N$  با احتمال غیر یکنواخت  $\pi_i^{(j)}$  انتخاب می‌شوند. احتمال  $\pi_i^{(j)}$  با تعداد مشاهدات در خوشه متناسب به  $y_i$ ،  $n_k^{(j)}$  رابطه عکس دارد  $(\pi_i^{(j)} = (\sqrt{n_{k,i}^{(j)}}) / \sum_{i=1}^N \sqrt{n_{k,i}^{(j)}})$ . در نتیجه، احتمال انتخاب پارامترها با تعداد المان کمتر در خوشه متناسب به آن بیشتر می‌شود.

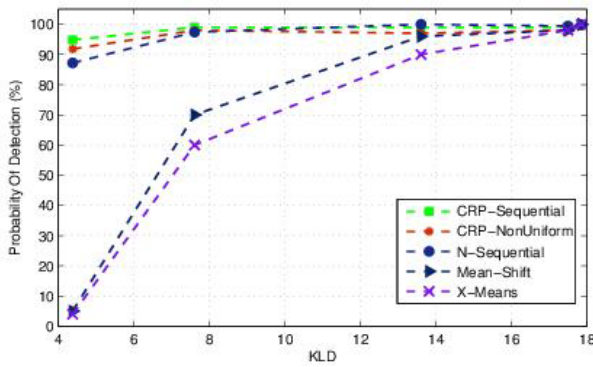
### ۳- نتایج شبیه‌سازی

با شبیه‌سازی یک سناریوی نمونه از شبکه رادیوشناختی، روش ارائه شده خوشه‌بندی سیگنال‌های کاربران اولیه مبتنی بر DPMM-GS مورد ارزیابی قرار می‌گیرد و با روش‌های خوشه‌بندی مانند X-Means و Mean-Shift از لحاظ دقت شناسایی تعداد سیگنال‌های اولیه ارسالی در محیط رادیویی و تخصیص صحیح ویژگی‌های سیگنال به هر خوشه مقایسه می‌شود. در شبیه‌سازی‌ها، تأثیر تغییر تعداد سیگنال‌های ارسالی و فاصله دو خوشه در فضای ویژگی بر روی شاخص‌های عملکرد مذکور (دقت شناسایی تعداد سیگنال‌های اولیه ارسالی و تخصیص صحیح ویژگی‌های سیگنال به هر خوشه)، ارزیابی می‌گردد. تنها سیگنال‌های ارسالی کاربران اولیه در نظر گرفته می‌شوند و ویژگی‌های سیگنال‌های اولیه خوشه‌بندی می‌گردد. این فرض به این دلیل مورد استفاده قرار گرفته که سیگنال‌های کاربران ثانویه توسط شبکه رادیوشناختی قابل شناسایی است و از خوشه‌بندی و متمایز کردن سیگنال‌های آنها از سیگنال‌های اولیه صرف نظر می‌شود. تعداد کانال‌های در دسترس شبکه رادیوشناختی برابر با  $L=10$  و طول دوره طیف‌سنجی برابر با  $30 \mu s$  در نظر گرفته شده است. در این مدت زمان، ویژگی‌های سیگنال دریافتی  $z(t)$  استخراج می‌شود. این مرحله ۱۰۰ بار تکرار می‌گردد و فرض می‌شود در طول این مدت، سیگنال  $z(t)$  تغییر نمی‌کند. در این مدت زمان سنخش طیف، ویژگی‌های سیگنال‌های تولید شده استخراج می‌شود. سیگنال به نویز همه سیگنال‌ها یکسان و ثابت برابر با مقدار  $SNR=10 \text{ dB}$  در نظر گرفته شده است. کانال مخابراتی کانال AWGN می‌باشد که نویز سفید جمع‌شونده با سیگنال‌های ارسالی همراه است. برای شروع در محیط رادیویی، سیگنال‌های مخابراتی با فرکانس‌های مرکزی ۱۰، ۵۰، ۹۰ و ۱۴۰ مگاهرتز در باند پایه در حال ارسال می‌باشند که به ترتیب دارای مدولاسیون‌های BPSK، QPSK، QAM-۱۶ و QAM-۶۴ هستند و نرخ سمبل سیگنال‌ها به ترتیب برابر با ۲، ۲، ۵ و ۵ بر حسب MBaud است. به منظور ارزیابی طرح خوشه‌بندی سیگنال‌های اولیه، معیارهای کارایی زیر تعریف می‌گردند:

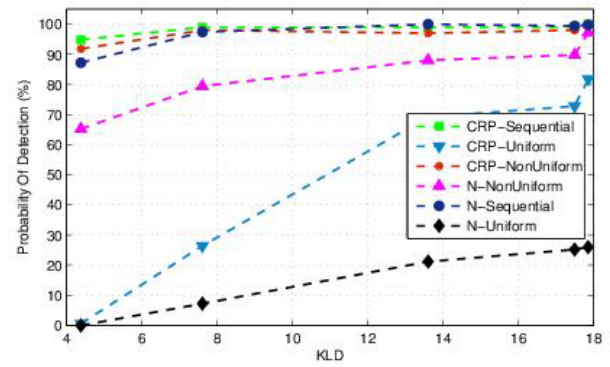
(۱) احتمال آشکارسازی صحیح (POD): دقت الگوریتم خوشه‌بندی را در شناسایی تعداد سیگنال‌های کاربران اولیه، مشخص می‌کند که از نسبت شمارش صحیح تعداد سیگنال‌های اولیه  $n(\text{Correct})$  بر کل تکرارهای اجرا شده  $n(\text{Total})$  به دست می‌آید.

(۲) دقت خوشه‌بندی (ACC): درستی تخصیص هر کدام از بردار

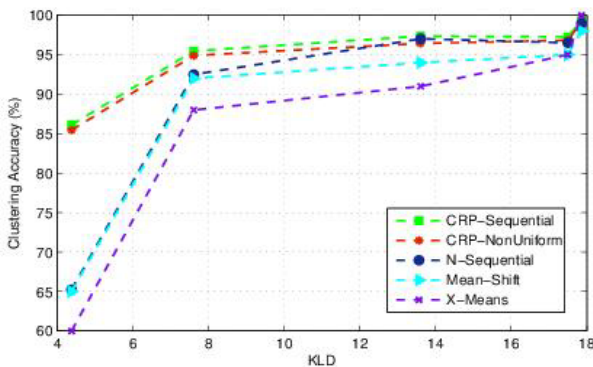
1. Kullback Leibler-Divergence



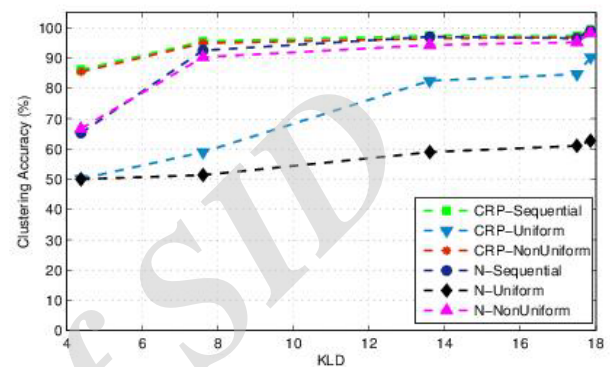
شکل ۲: احتمال آشکارسازی صحیح روش‌های مورد مطالعه به ازای  $K = D = 2$  و  $N = 100$  بر حسب KLD.



شکل ۳: احتمال آشکارسازی صحیح روش DPMM به ازای  $K = D = 2$  و  $N = 100$  بر حسب KLD.



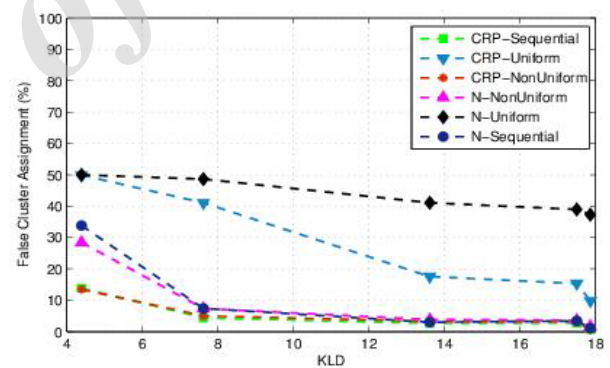
شکل ۴: مقایسه دقت خوشه‌بندی در روش‌های DPMM و X-Means و Mean-Shift به ازای  $K = D = 2$  و  $N = 100$  بر حسب KLD.



شکل ۵: دقت خوشه‌بندی به ازای  $K = D = 2$  و  $N = 100$  بر حسب KLD.

دارند. باید خاطر نشان کرد همه حالت‌های شبیه‌سازی شده به مقدار پایدار همگرا خواهند شد، با این تفاوت که برخی از حالت‌ها به ازای مقدار KLD برابر، به تعداد تکرار بیشتری برای همگرایی نیاز خواهند داشت. شکل ۴ دقت خوشه‌بندی (ACC) الگوریتم DPMM را بر حسب KLD نشان می‌دهد. حالت CRP-Nonuniform و CRP-Sequential با دقت بالاتر به ازای همه مقادیر KLD، بردارهای ویژگی را به خوشه‌ها به درستی تخصیص می‌دهد. در شکل ۵ نیز تخصیص به خوشه غیر صحیح (FCA) را می‌توان مشاهده کرد. عملکرد حالت‌های CRP-Nonuniform و CRP-Sequential نسبت به حالت‌های دیگر بهتر می‌باشد و تخصیص به خوشه غیر صحیح در محدوده ۱۵٪ تا ۲۰٪ به ازای افزایش KLD است.

در شکل ۶ عملکرد روش‌های خوشه‌بندی DPMM، X-Means و Mean-Shift را در شناسایی درست سیگنال‌های ارسالی در محیط به ازای مقادیر متفاوت KLD می‌بینیم. به ازای KLDهای کوچک، هنگامی که مرکز خوشه‌ها به یکدیگر نزدیک باشند، عملکرد روش DPMM بهتر است. هرچه مقدار KLD بیشتر شود و شباهت خوشه‌ها کمتر باشد، عملکرد هر سه روش بهبود می‌یابد. به ازای  $KLD = 17.86$ ، احتمال آشکارسازی صحیح هر سه روش، ۱۰۰ درصد می‌باشد. در واقع در  $KLD = 17.86$  سیگنال‌های ارسالی در دو خوشه کاملاً مجزا می‌باشند و هیچ همپوشانی در فضای ویژگی ندارند. بنابراین روش‌های X-Means و Mean-Shift نیز دارای عملکرد ۱۰۰ درصد می‌باشند. هرچه مقدار KLD کمتر شود عملکرد این دو روش نسبت به DPMM افت می‌کند. در  $KLD = 4.39$  دو خوشه سیگنال در فضای ویژگی در فاصله نزدیکی قرار دارند و میانگین یک خوشه در محدوده خوشه دیگر قرار دارد. در این مقدار KLD، دو روش X-Means و Mean-Shift عملکرد پایین‌تری نسبت به روش DPMM دارند. شکل ۷ عملکرد سه روش مذکور را در



شکل ۶: تخصیص به خوشه غیر صحیح به ازای  $K = D = 2$  و  $N = 100$  بر حسب KLD.

KLD آشکارسازی صحیح را بر حسب پارامتر KLD نشان می‌دهد. توجه به شکل ۳، درصد شناسایی تعداد سیگنال‌های آشکار شده با افزایش شاخص KLD بهبود می‌یابد. در حقیقت، مقدار کوچک شاخص نشان می‌دهد بردارهای ویژگی دو سیگنال آشکار شده و در واقع فاصله خوشه‌ها به یکدیگر نزدیک هستند و انتظار می‌رود خوشه‌بندی با خطای بیشتری همراه باشد. همان‌طور که بیان شد کاربر مخرب نیز سعی دارد سیگنال ارسالی خود را به سیگنال ارسالی کاربر اولیه واقعی نزدیک کند و بنابراین برای بررسی عملکرد تمایز سیگنال‌های ارسالی در محیط، از معیار KLD بهره بردیم.

در شرایط یکسان DPMM-GS، حالت CRP-Uniform و N-Uniform، عملکرد ضعیف‌تری نسبت به حالت‌های دیگر به ازای مقادیر متفاوت KLD دارند. در این حالت، انتخاب نمونه‌های احتمال تخصیص هر بردار ویژگی به خوشه‌ها به صورت تصادفی با توزیع یکنخواه خواهد بود. حالت CRP-Sequential، CRP-Nonuniform و N-Sequential به ترتیب عملکرد قابل قبول و متوسط احتمال آشکارسازی صحیح ۹۰٪

- access networks," in *Proc IEEE Symp. on New Frontiers in Dynamic Spectrum*, 11 pp., Singapore, Singapore, 6-9 Apr. 2010.
- [5] C. Zhao, L. Xie, X. Jiang, L. Huang, and Y. Yao, "A PHY-layer authentication approach for transmitter identification in cognitive radio networks," in *Proc. Int. Conf. on Communications and Mobile Computing, CMC'10*, vol. 2, pp. 154-158, Shenzhen, China, 12-14 Apr. 2010.
- [6] Z. Caidan, W. Wumei, H. Lianfen, and Y. Yan, "Anti-PUE attack base on the transmitter fingerprint identification in cognitive radio," in *Proc. Int. Conf. on Wireless Communications, Networking and Mobile Computing*, 5 pp., Beijing, China, 24-26 Sept. 2009.
- [7] W. L. Chin, et al., "Channel-based detection of primary user emulation attacks in cognitive radios," in *Proc. IEEE Vehicular Technology Conf. (VTC Spring)*, 5 pp., Yokohama, Japan, 6-9 May 2012.
- [8] H. Li and Z. Han, "Dogfight in spectrum: combating primary user emulation attacks in cognitive radio systems, part i: known channel statistics," *IEEE Trans. on Wireless Communications*, vol. 9, no. 11, pp. 3566-3577, Nov. 2010.
- [9] N. Nguyen-Thanh, P. Ciblat, A. Pham, and V. Nguyen, "Surveillance strategies against primary user emulation attack in cognitive radio networks," *IEEE Trans. on Wireless Communications*, vol. 14, no. 9, pp. 4981-4993, Sep. 2015.
- [10] C. Biao, K. M. Borle, and D. Wenliang, "A physical layer authentication scheme for countering primary user emulation attack," in *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP'13*, pp. 2935-2939, Vancouver, BC, Canada, 26-31 May 2013.
- [11] X. Chunsheng and M. Song, "Detection of PUE attacks in cognitive radio networks based on signal activity pattern," *IEEE Trans. on Mobile Computing*, vol. 13, no. 5, pp. 1022-1034, May 2014.
- [12] N. T. Nguyen, R. Zheng, and Z. Han, "On identifying primary user emulation attacks in cognitive radio systems using nonparametric Bayesian classification," *IEEE Trans. on Signal Processing*, vol. 60, no. 3, pp. 1432-1445, Mar. 2012.
- [13] Y. W. Teh, *Dirichlet Processes*, Encyclopedia of Machine Learning, Springer, 2010.
- [14] M. Bkassiny, S. K. Jayaweera, and L. Yang, "Multidimensional Dirichlet process-based non-parametric signal classification for autonomous self-learning cognitive radios," *IEEE Trans. on Wireless Communications*, vol. 12, no. 11, pp. 5413-5423, Nov. 2013.
- [15] M. D. Escobar and M. West, "Bayesian density estimation and inference using mixtures," *J. of the American Statistical Association*, vol. 90, no. 430, pp. 577-588, Jun. 1995.
- [16] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, 2012.
- [17] J. R. Hershey and P. A. Olsen, "Approximating the Kullback Leibler divergence between Gaussian mixture models," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, ICASSP'07*, vol. 4, pp. 317-320, Honolulu, HI, USA, 15-20, Apr. 2007.

**خاطره اکبری** تحصیلات خود را در مقطع کارشناسی در رشته مهندسی برق الکترونیک از دانشگاه اصفهان و در مقطع کارشناسی ارشد در رشته مهندسی برق مخابرات از دانشگاه یزد به ترتیب در سال‌های ۱۳۹۰ و ۱۳۹۴ به پایان رسانده است. زمینه‌های تحقیقاتی مورد علاقه نامبرده عبارتند از: رادیوشناختی، نسل پنجم شبکه‌های بی‌سیم، شبکه‌های حسگر بی‌سیم و شناسایی آماری الگو.

**جمشید ابویی** مدرک کارشناسی و کارشناسی ارشد (رتبه نخست) مهندسی برق خود را از دانشگاه صنعتی اصفهان به ترتیب در سال‌های ۱۳۷۲ و ۱۳۷۵ دریافت نمود. از سال ۱۳۷۷ الی ۱۳۸۲ نامبرده به عنوان عضو هیأت علمی و سرپرست آزمایشگاه‌های مهندسی برق در دانشکده مهندسی برق و کامپیوتر دانشگاه یزد مشغول بود. پس از آن به دوره دکترای مهندسی برق و کامپیوتر در دانشگاه واترلو در کانادا وارد گردید و در سال ۱۳۸۸ موفق به اخذ درجه دکترا در مهندسی برق از دانشگاه مذکور گردید. دکتر ابویی دوره پس‌دکتررا را نیز در دانشگاه تورنتو در کانادا در سال ۱۳۸۹ به پایان رساند و هم‌اینک دانشیار دانشکده مهندسی برق دانشگاه یزد می‌باشد. زمینه‌های علمی مورد علاقه نامبرده متنوع بوده و شامل موضوعاتی مانند نسل پنجم شبکه‌های سیار، شبکه‌های فمتوسل، رادیوشناختی، تخصیص منابع در شبکه‌های بی‌سیم، شبکه‌های حسگر بی‌سیم و سیستم‌های نوری فضای باز می‌باشد.

دقت خوشه‌بندی (ACC) نشان می‌دهد. همان طور که مشخص است برای تمایز سیگنال‌ها، عملکرد روش DPMM به ازای KLDهای کوچک نیز بهتر از دو روش دیگر می‌باشد.

آشکارسازی کاربر حمله تقلید از سیگنال کاربر اولیه: تاکنون روشی برای خوشه‌بندی سیگنال‌های کاربران اولیه و تعیین تعداد کاربران اولیه فعال در محیط مخابراتی ارائه شد. در صورت در دسترس بودن شناسه‌های مختص کاربران اولیه (ID) در شبکه رادیوشناختی، کاربر مخرب تنها با مقایسه شناسه کاربران شناسایی شده و شناسه‌های معتبر قابل شناسایی است (مانند آنچه در [۱۲] اشاره شده است). در حالت دوم با فرض این که کاربر مخرب، توانایی تقلید از شناسه‌های کاربران اولیه را داشته باشد، کاربران متمایز با شناسه‌های مشابه شناسایی می‌شوند. وجود شناسه‌های مشابه نشان‌دهنده وجود کاربر مخرب در شبکه رادیوشناختی است. از سوی دیگر، هدف کاربر مخرب در اشغال کانال‌های مخابراتی با هدف کاربر اولیه واقعی متفاوت است. بنابراین الگوی رفتاری کاربر مخرب با کاربر اولیه واقعی متفاوت خواهد بود. این الگوی رفتاری در افزایش دوره‌های فعال بودن در کانال‌ها آشکار می‌شود. به طور دقیق‌تر، مقدار انرژی کاربر اولیه واقعی در طول زمان تغییر می‌کند. کاربر مخرب با هدف حمله رد سرویس، مقدار انرژی نسبتاً یکنواختی خواهد داشت. دوره‌های فعال کاربر مخرب از دوره‌های غیر فعال بیشتر و تابع خودهمبستگی انرژی PUE تقریباً ثابت است.

#### ۴- جمع‌بندی

در این مقاله با هدف شناسایی حمله PUE طرحی مبتنی بر طبقه‌بندی سیگنال‌های کاربران اولیه با روش غیر پارامتریک بی‌زین در شبکه رادیوشناختی ارائه گردید. با اعمال الگوریتم خوشه‌بندی، تعداد و نوع سیگنال‌های ارسالی به دست می‌آید. تعداد و نوع کاربران اولیه فعال برای شبکه رادیوشناختی، مجهول در نظر گرفته شد. برای دستیابی به سرعت همگرایی بالاتر به روش نمونه‌برداری گیبس اصلاح‌شده که از قسمت‌های اصلی مدل مخلوطی فرایند دیربشله است پرداختیم. با به کارگیری روش نمونه‌برداری غیر یکنواخت و مقداردهی اولیه از طریق روش CRP، سرعت نمونه‌برداری گیبس بهبود یافت. دقت خوشه‌بندی با معیار کارایی درصد شناسایی درست تعداد سیگنال‌های فعال در محیط و تخصیص مشاهدات درست به خوشه متناظر مورد بررسی قرار گرفت و طرح ارائه‌شده با روش‌های خوشه‌بندی X-means و Mean-Shift مقایسه گردید و نشان داده شد که روش ارائه‌شده، دقت خوشه‌بندی را نسبت به دو روش Mean-Shift و X-Means بهبود داده است.

#### مراجع

- [1] B. Wang and K. J. R. Liu, "Advances in cognitive radio networks: a survey," *IEEE J. of Selected Topics in Signal Processing*, vol. 5, no. 1, pp. 5-23, Feb. 2011.
- [2] X. Shengli, L. Yi, Z. Yan, and Y. Rong, "A parallel cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. on Vehicular Technology*, vol. 59, no. 8, pp. 4079-4092, Oct. 2010.
- [3] R. Saifan, A. E. Kamal, and Y. Guan, "Efficient spectrum searching and monitoring in cognitive radio network," in *Proc. IEEE 8th Int. Conf. on Mobile Adhoc and Sensor Systems, MASS'11*, pp. 520-529, Valencia, Spain, 17-22 Oct. 2011.
- [4] R. W. Thomas, R. S. Komali, B. J. Borghetti, and P. Mahonen, "A Bayesian game analysis of emulation attacks in dynamic spectrum