Scientific Journal of Electronical & Cyber Defence
Vol. 10, No. 4, Winter 2023, Serial No. 40

Reducing the Destructive Effect of Misbehaving Users in Cooperative Spectrum Sensing using Reinforcement Learning

S. Z. Majidian*

*PhD student, Aras International Branch, Islamic Azad University, Tabriz, Iran
(Received: 08/09/2021, Accepted: 18/12/2021)

ABSTRACT

The presence of misbehaving users in Cognitive Radio Networks (CRN) can disrupt the process of spectrum sensing and detecting the status of the Primary User (PU). In order to reduce the destructive effect of this group of users in CRNs, in this paper, a new mechanism based on reinforcement learning for cooperative spectrum sensing is presented. The proposed method is a cooperative spectrum sensing mechanism based on user weighting, according to which users receive a weight commensurate with how they behave in spectrum sensing. The reinforcement learning model used in the proposed method is a learning automata which, using reward and penalty processes, allocates more weight to users with normal behavior in sensing the spectrum and less to misbehaving users. In this way, the learning automata updates the users' weight vector based on the response received from the environment, after performing a sensing operation in each repetition. After repeating the sensing operation several times, the learner will be able to optimize the user's weight vector. In order to evaluate the proposed method, its performance in the simulation environment has been tested and the results have been compared with the existing method for cooperative spectrum sensing. The results show that using the proposed method in the presence of misbehaving users will significantly improve network performance. The results show that the use of the proposed method in the presence of misbehaving users significantly improves the network performance and the proposed method can increase the throughput in the network by 9.88% compared to the standard method and compared to the sensing method based on Improve credit by 4.98%. Also, using the proposed method improves energy efficiency by 22.5% compared to the standard method and increases it by 88.8% compared to the credit-based sensing method .

Keywords: Cognitive Radio, Cooperative Spectrum Sensing, Learning Automata, Identifying Misbehaving Users.

*Corresponding Author Email: Mysun7196@gmail.com

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Publisher: Imam Hussein University © Authors

نشریه علمی "پدافند الکترونیکی و سایبری"
سال دهم، شماره ۴، زمستان ۱۴۰۱، ص ۹- ۱

علمی – پژوهشی

کاهش اثر مخرب کاربران بدرفتار در حسگری همکارانه طیف مبتنی بر یادگیری تقویتی

سیده زهره مجیدیان

دانشجوی دکترا ، واحد بین‌الملل ارس، دانشگاه آزاد اسلامی، تبریز، ایران

(دریافت: ۱۴۰۰/۰۶/۱۷، پذیرش: ۱۴۰۰/۰۹/۲۷)

چکیده

وجود کاربران بدرفتار در شبکه‌های رادیوشناختی می‌تواند موجب اخلال در فرآیند حسگری طیف و تشخیص وضعیت کاربر اولیه گردد؛ در این مقاله، به‌منظور کاهش اثر مخرب این دسته از کاربران در شبکه‌های رادیوشناختی، یک سازوکار نوین مبتنی بر راهبرد یادگیری تقویتی به‌منظور حسگری همکارانه طیف ارائه شده است. روش پیشنهادی، یک سازوکار حسگری همکارانه مبتنی بر وزن‌دهی کاربران بوده که بر اساس آن کاربران وزنی متناسب با نحوه رفتار خود در حسگری طیف را دریافت می‌کنند. مدل یادگیری تقویتی به‌کاررفته در روش پیشنهادی یک آتاماتای یادگیر بوده که با استفاده از فرآیندهای پاداش و جریمه، به کاربران دارای رفتار بهنجار در حسگری طیف وزن بیشتر و به کاربران بدرفتار مقادیر وزن کمتری اختصاص می‌دهد. بدین‌صورت که آتاماتای یادگیر پس از انجام عمل حسگری در هر بار تکرار، بردار وزن کاربران را بر اساس پاسخ دریافتی از محیط به‌روزرسانی می‌کند. پس از چند بار تکرار حسگری، آتاماتای یادگیر قادر خواهد بود بردار وزن کاربران را به صورتی بهینه تنظیم کند. به‌منظور ارزیابی روش پیشنهادی، عملکرد آن در محیط شبیه‌سازی مورد آزمایش قرار گرفته و نتایج حاصله با روش موجود برای حسگری همکارانه طیف مقایسه شده است. نتایج حاصله نشان می‌دهد که استفاده از روش پیشنهادی در شرایط حضور کاربران بدرفتار موجب بهبود چشمگیر عملکرد شبکه می‌شود و روش پیشنهادی می‌تواند گذردهی در شبکه را نسبت به روش استاندارد به میزان ۹/۸۸ درصد و در مقایسه با روش حسگری مبتنی بر اعتبار به میزان ۴/۹۸ درصد بهبود بخشد. همچنین، به‌کارگیری روش پیشنهادی موجب بهبود ۲۲/۵ درصدی کارایی انرژی نسبت به روش استاندارد و افزایش ۸/۸۸ درصدی آن نسبت به روش حسگری مبتنی بر اعتبار می‌گردد..

کلیدواژه‌ها: رادیوشناختی، حسگری همکارانه طیف، آتاماتای یادگیر، شناسایی کاربران بدرفتار

۱- مقدمه

رادیوشناختی[1] یک تکنولوژی با هدف استفاده بهینه از طیف رادیویی شبکه‌های بی‌سیم می‌باشد. در شبکه‌های مبتنی بر این تکنولوژی، کاربران به دو دسته کاربران اولیه[2] و کاربران ثانویه[3] تقسیم می‌شوند. کاربران اولیه در هر لحظه توانایی استفاده از طیف را دارند؛ درصورتی‌که کاربران ثانویه تنها زمانی مجاز به استفاده از طیف رادیویی هستند که کاربر اولیه حضور نداشته باشد [۱].

حسگری همکارانه طیف یک راهکار کارآمد برای بهبود دقت تشخیص حضور کاربران اولیه در شبکه‌های رادیوشناختی می‌باشد. بااین‌وجود، عواملی مانند حضور کاربران بدرفتار ممکن است اثرات مخربی بر عملکرد شبکه و حسگری همکارانه طیف در پی داشته باشد. کاربران بدرفتار، کاربران ثانویه‌ای هستند که با هشدار اشتباهِ حضور کاربر اولیه، کارایی شبکه رادیوشناختی را کاهش می‌دهند [۲]. حذف یا کاهش اثر این دسته از کاربران می‌تواند موجب افزایش دقت آشکارسازی در حسگری طیف شود [۳]. یکی از راهکارهای شناسایی این دسته از کاربران، استفاده از روش وزن‌دهی و تعیین اعتبار برای هر کاربر است. بدین‌صورت که با اختصاص وزن مناسب به اعتبار قبلی کاربران و با استفاده از رفتار فعلی آنان، یک آستانه بهینه برای جداسازی کاربران بهنجار از کاربران بدرفتار مشخص کرد. در این مقاله، از سازوکار وزن‌دهی کاربران برای ایجاد تمایز بین کاربران بهنجار و کاربران بدرفتار استفاده می‌شود. وجه تمایز روش پیشنهادی در این تحقیق با روش‌های ارائه شده پیشین، بهره‌گیری از راهبرد یادگیری تقویتی آتاماتای یادگیر به‌منظور وزن‌دهی بهینه کاربران شبکه می‌باشد که در کارهای پیشین به آن پرداخته نشده است. عدم پیچیدگی محاسباتی و همگرایی سریع به راهبرد بهینه برای مسئله وزن‌دهی کاربران؛ از جمله مزایایی است که استفاده از مدل آتاماتای یادگیر در روش پیشنهادی در پی دارد.

استفاده از آتاماتای یادگیر موجب می‌شود که بتوان با گذر زمان به شناخت بالاتری از وضعیت کاربران دست یافته و کاربران بدرفتار را از سایرین متمایز نمود. در یادگیری تقویتی یک عامل

* رایانامه نویسنده مسئول: Mysun7196@gmail.com

[1] Cognitive Radio
[2] Primary User (PU)
[3] Secondary User (SU)

* این مقاله یک مقاله با دسترسی آزاد است که تحت شرایط و ضوابط مجوز Creative Commons Attribution (CC BY) توزیع شده است.
© نویسندگان
ناشر: دانشگاه جامع امام حسین (ع)

یادگیرنده در طی یادگیری با فعل‌وانفعالات مکرر با محیط، به یک سیاست کنترل بهینه می‌رسد. کارایی این فعل‌وانفعالات با محیط به‌وسیله بیشینه (کمینه) بودن پاداش (جریمه) عددی که از محیط گرفته می‌شود، ارزیابی می‌گردد [۴]. مزیت اصلی یادگیری تقویتی نسبت به سایر روش‌های یادگیری، عدم نیاز به هیچ‌گونه اطلاعاتی از محیط است [۵]. در روش پیشنهادی، با تجهیز مرکز ادغام یک مدل آتاماتای یادگیر، عمل به‌روزرسانی وزن کاربران انجام می‌شود. در این روش، آتاماتای تصادفی موجود در مرکز ادغام، بر اساس الگوریتم یادگیر خود وزن کاربران بدرفتار را کاهش داده و از طرفی موجب افزایش وزن مربوط به هر کاربر بهنجار خواهد شد.

ادامه این مقاله به‌صورت زیر سازمان‌دهی شده است: در بخش دوم به‌مرور برخی از تحقیقات مرتبط پرداخته و در بخش سوم روش پیشنهادی ارائه شده است. سپس در بخش چهارم نتایج شبیه‌سازی ارائه شده و در بخش پنجم، پس از نتیجه‌گیری و جمع‌بندی پیشنهادهایی برای ادامه تحقیقات ارائه شده است.

۲- سوابق تحقیق

در [۶] یک روش حسگری همکارانه طیف و آگاه از اثر کاربران بدرفتار ارائه شده است. در این روش احتمال وجود سیگنال جعلی در هنگام حضور یا عدم حضور کاربر اولیه تخمین زده می‌شود. سپس از پارامترهای به‌دست‌آمده برای تعیین آستانه بهینه‌ای که بتواند احتمال خطا را به حداقل برساند استفاده می‌شود. روش پیشنهادی در این مقاله شامل دو الگوریتم متوالی بوده که در الگوریتم اول پارامترهای وجود حمله مقداردهی شده و در مرحله دوم، با استفاده از پارامترهای پیداشده یک آستانه بهینه برای حسگری همکارانه طیف تعیین می‌شود.

در [۷] یک روش حسگری همکارانه طیف در سیستم‌های شناختی صنعتی (CIS[1]) مبتنی بر یادگیری تقویتی و آستانه دوگانه انرژی ارائه شده است. در CIS، مجموعه‌ای از کانال‌ها برای انتخاب توسط کاربر ثانویه در دسترس بوده و در این مقاله تلاش شده از تا با استفاده از تکنیک یادگیری تقویتی، روشی برای پیش‌بینی کانال بهینه ارائه شود. استفاده از آستانه دوگانه انرژی موجب می‌شود که بتوان یک کران پایین برای احتمال آشکارسازی تعیین نمود. این روش همچنین درصورتی‌که انرژی برداشت شده بین دو آستانه قرار گیرد؛ از یک مدل بیزی برای ادغام نتایج حسگری بهره می‌گیرد. این روش در شرایط افزایش نویز یا وجود کاربران بدرفتار عملکرد نامناسبی خواهد داشت.

در [۸] از تکنیک یادگیری عمیق تقویتی به‌منظور کاهش احتمال تداخل در حسگری همکارانه طیف استفاده شده است. در تکنیک یادگیری عمیق تقویتی، یک مدل یادگیری تقویتی با مدل‌های یادگیری عمیق ترکیب شده است. در این مقاله، کاربران ثانویه به این مدل یادگیر تجهیز شده تا به‌واسطه آن راهبرد بهینه حسگری برای اجتناب از تداخل مشخص گردد. پیچیدگی بالای مدل یادگیری عمیق بکار رفته در این روش موجب شده است تا نتوان از آن برای به‌کارگیری در انواع تجهیزات بهره برد.

در [۹] یک راهکار حسگری همکارانه مبتنی بر اعتبار برای شناسایی کاربران بدرفتار در شبکه‌های رادیوشناختی ارائه شده است. در این روش، M گزارش قبلی هر کاربر ثانویه با نتیجه ادغام مقایسه شده و اعتبار کاربر بر اساس این مقایسات تعیین می‌شود. با توجه به اینکه این تکنیک از پاسخ محیط (رفتار کاربر اولیه) برای تعیین اعتبار کاربران استفاده نمی‌کند؛ در نتیجه در شرایط افزایش تعداد کاربران بدرفتار عملکرد نامناسبی دارد.

در [۱۰] یک راهکار مبتنی بر آستانه دوگانه پویا برای حسگری همکارانه طیف فرکانسی در شبکه‌های رادیوشناختی و تحت توان نویز ارائه شده است. در این مقاله، ابتدا یک تابع سودمندی با هدف کمینه‌سازی خطای حسگری طیف تعریف شده و آستانه بهینه انرژی بر اساس آن تعیین شده است. سپس به‌منظور کاهش اثر مخرب ناشی از نویز، یک مکانیسم تنظیم آستانه دوگانه پویا ارائه شده است، و قانون ادغام بهینه برای این آستانه‌ها با هدف حداقل نمودن احتمال خطای سراسری موردبحث واقع شده است.

در [۱۱]، یک طرح حسگری همکارانه طیف با بهره‌گیری از قابلیت یادگیری تقویتی برای کاربران ثانویه پیشنهاد شده است. این مدل یادگیری تقویتی وظیفه تعیین ترتیب اسکن کانال‌ها و همچنین انتخاب کاربران همکار برای حسگری همکارانه طیف فرکانسی را بر عهده دارد. در این روش از مدل یادگیری تقویتی Q-Learning استفاده شده است که در آن، هر کاربر ثانویه الگوی اشغال کانال‌های اصلی را یاد گرفته و بنابراین یک لیست اولویت حسگری پویا را تشکیل می‌دهد تا سربار اسکن و تأخیر دسترسی را کاهش دهد. از طرفی، روش ارائه شده در [۱۲] از تئوری بازی‌های ائتلافی و الگوریتم ژنتیک به‌منظور طراحی یک روش حسگری همکارانه طیف فرکانسی در شبکه‌های رادیوشناختی استفاده کرده است. در این مدل، کاربران ثانویه ائتلاف‌هایی را برای همکاری در حسگری همکارانه شکل می‌دهند و سازمان‌دهی بهینه این ائتلاف‌ها با استفاده از الگوریتم ژنتیک صورت می‌پذیرد.

[1] Cognitive Industrial Systems

در [۱۳] یک روش جدید مبتنی بر ماشین بردار پشتیبان (SVM[1]) برای حسگری همکارانه طیف فرکانسی در شبکه‌های رادیوشناختی ارائه شده است. در این روش، بردارهای حاوی سطوح انرژی نمونه‌برداری شده از کاربر اولیه، به‌عنوان بردارهای ویژگی ورودی ماشین بردار پشتیبان در نظر گرفته می‌شوند و در مرحله آموزش این مدل یادگیر مورداستفاده قرار می‌گیرند. ماشین بردار پشتیبان پیشنهاد شده در این تحقیق، در دو فاز آموزش و آزمون عمل می‌کند. بدین ترتیب؛ در فاز اول، مدل طبقه‌بندی با استفاده از بردارهای ویژگی دارای برچسب آموزش داده می‌شود. پس از آن و در فاز دوم، سطوح انرژی نمونه‌برداری شده از کانال اولیه به‌صورت بردارهای ویژگی آزمون به مدل یادگیر آموزش دیده تحویل داده خواهد شد تا بر اساس الگوهای مشاهده‌شده قبلی، حالت حضور یا عدم حضور کاربر اولیه مشخص گردد.

۳- روش پیشنهادی

در این بخش، یک مدل جدید برای حسگری همکارانه طیف ارائه خواهد شد که با استفاده از راهکار وزن‌دهی کاربران، اثر ناشی از کاربران بدرفتار در شبکه رادیوشناختی را کاهش می‌دهد. بدین منظور، در ادامه پس از بیان فرضیات و تشریح مدل سیستم مفروض، جزئیات روش پیشنهادی تشریح خواهد شد.

۳-۱- فرضیات و مدل سیستم

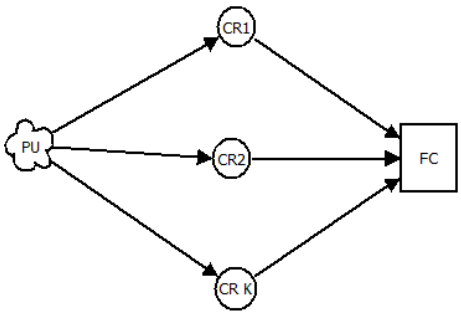
در این تحقیق یک شبکه بی‌سیم با سیستم دستیابی فرصت‌طلبانه به طیف در نظر گرفته می‌شود که دارای یک کانال کنترلی به‌منظور ردوبدل پیام‌های مربوط به سیگنال توسط کاربران ثانویه می‌باشد. کاربران ثانویه نمی‌توانند از وضعیت کانال‌های اولیه آگاه باشند و نتیجه سنجش هر کاربر یک بیت نتیجه به‌صورت $O \in \{0,1\}$ خواهد بود. درصورتی‌که $O=0$ باشد، کاربر ثانویه عدم حضور کاربر اولیه را تشخیص داده است و در غیر این صورت کاربر ثانویه تشخیص داده است که کاربر اولیه حضور دارد. با توجه به امکان سنجش غیرکامل[2]، ممکن است مشاهدات نادرست باشند. در این شرایط حالات زیر رخ می‌دهد:

- درصورتی‌که مشاهدات، حالت اشغال کانال را نمایش دهد و در واقع کانال آزاد باشد، در این حالت یک هشدار اشتباه رخ داده که احتمال آن به‌صورت p_{fa} نمایش داده می‌شود.
- درصورتی‌که مشاهدات حالت اشغال کانال را نمایش دهد و در واقع کانال نیز اشغال باشد، در این حالت یک آشکارسازی صحیح رخ داده که احتمال آن به‌صورت p_d نمایش داده می‌شود.

با توجه به شرایط فوق، سنجش کامل حالتی است که در آن شرایط $p_d=1$ و $p_{fa}=0$ برقرار باشد. در مدل سیستم مفروض تعدادی کاربر بدرفتار نیز در نظر گرفته می‌شود. کاربران بدرفتار، کاربرانی هستند که نتیجه تشخیص آن‌ها موجب بروز خطا در فرآیند حسگری طیف می‌شود. در مدل سیستم مورداستفاده در این تحقیق فرض می‌شود که رفتار این کاربران مبنی بر تعیین حضور یا عدم حضور کاربر اولیه به‌صورت تصادفی بوده و ارتباطی با سیگنال دریافتی نخواهد داشت.

۳-۲- کاهش اثر کاربران بدرفتار مبتنی بر آتاماتای یادگیر

یک راهکار همکارانه برای حسگری طیف مفروض است که در آن K کاربر ثانویه اطلاعات حسگری محلی خود را از طریق یک کانال کنترلی به مرکز ادغام (FC[3]) ارسال می‌کنند. مدل سنتی حسگری همکارانه طیف در شکل (۱) نمایش داده شده است.



شکل (۱). مدل سنتی حسگری همکارانه طیف

با توجه به شکل (۱)، هر کاربر ثانویه شرکت‌کننده در حسگری همکارانه طیف عمل سنجش وضعیت طیف را به‌صورت محلی انجام می‌دهد. سپس با ادغام اطلاعات حسگری مربوط به هر کاربر، خروجی نهایی تولید خواهد شد. وجود کاربران بدرفتار در این فرآیند موجب بروز خطا در حسگری مشارکتی می‌شود. زیرا نتیجه حسگری این کاربران به مرکز ادغام ارسال شده و همانند سایر کاربران ثانویه در تعیین وضعیت حضور کاربر اولیه در نظر گرفته خواهد شد. در روش پیشنهادی، با افزودن یک راهکار مبتنی بر وزن و یادگیری تقویتی آتاماتای یادگیر برای تعیین اهمیت نتیجه ارسالی کاربران مختلف به مرکز ادغام، اثر کاربران بدرفتار کاهش داده خواهد شد. بر این مبنا، ابتدا به تشریح نحوه حسگری طیف در روش‌های موجود پرداخته

[1] Support Vector Machine
[2] Imperfect Sensing
[3] Fusion Center

می‌شود. فرض کنید که سیگنال دریافتی توسط کاربر ثانویه iام به‌صورت زیر باشد:

$$y_i(n) = \begin{cases} w_i(n), & for\ H_0 \\ h_i s(n) + w_i(n), & for\ H_1 \end{cases} \quad (۱)$$

در رابطه فوق، H_0 و H_1 به ترتیب فرض‌های مربوط به عدم حضور و حضور کاربر اولیه در حسگری طیف هستند. $s(n)$ معرف سیگنال ارسال شده توسط کاربر اولیه، h_i ضریب کانال[1] بین کاربر اولیه و کاربر ثانویه iام و $w_i(n)$ یک نمونه از نویز گاوسی AWGN[2] با میانگین صفر و واریانس σ_w^2 می‌باشد. همچنین n نشان‌دهنده اندیس نمونه‌هاست.

برای تصمیم‌گیری محلی در سنجش توسط هر کاربر، تمامی کاربران ثانویه از آشکارساز انرژی با آستانه λ و تعداد نمونه N یکسان استفاده می‌کنند. در این حالت، تصمیم محلی کاربر پس از حسگری برابر خواهد بود با:

$$T_{y,i} = \frac{1}{N}\sum_{n=0}^{N-1}|y_i(n)|^2 \quad (۲)$$

$$D_i = \begin{cases} 0, & if\ T_{y,i} < \lambda \\ 1, & if\ T_{y,i} \geq \lambda \end{cases} \quad (۳)$$

درصورتی‌که هر یک از سه حالات ذکرشده رخ دهد، مرکز ادغام نتیجه حسگری همکارانه طیف بر اساس تصمیم محلی کاربران ثانویه را به‌صورت زیر تعیین خواهد کرد:

$$Y_h = \sum_{i=1}^{K} D_i \underset{H_0}{\overset{H_1}{\gtrless}} k \quad (۴)$$

در رابطه فوق، K تعداد کاربران مشارکت‌کننده در حسگری طیف است. درصورتی‌که $k=K$ باشد، آنگاه روش ادغام نتایج AND خواهد بود. همچنین درصورتی‌که $k=1$ باشد، ادغام به روش OR صورت خواهد پذیرفت. همان‌طور که اشاره شده، وجود کاربران بدرفتار در فرآیند ادغام موجب بروز خطا در تعیین نتیجه حضور یا عدم حضور کاربر اولیه خواهد شد. به‌عنوان مثال، حالتی را در نظر بگیرید که کاربر اولیه حضور ندارد. درصورتی‌که تمام کاربران بهنجار تصمیم‌گیری محلی را به‌صورت درست انجام دهند و تنها یک کاربر بدرفتار حضور کاربر اولیه را مشخص کند، در این حالت قانون OR منجر به تشخیص حضور کاربر اولیه شده و در نتیجه یک هشدار اشتباه رخ خواهد داد. با استدلال مشابه، درصورتی‌که کاربر اولیه حضور داشته باشد و تمامی کاربران بهنجار حضور آن را به‌درستی تشخیص دهند، وجود یک گزارش نادرست توسط کاربران بدرفتار موجب خطا در تشخیص حضور کاربر اولیه برای قانون AND خواهد شد. برای کاهش اثرات این خطاها که از وجود کاربران بدرفتار ناشی می‌شود، در روش پیشنهادی یک راهکار حسگری مشارکتی مبتنی بر معیار وزن و آتاماتای یادگیر ارائه می‌شود. سازوکار روش پیشنهادی به‌صورت فلوچارت در شکل (۲) نمایش داده شده است.

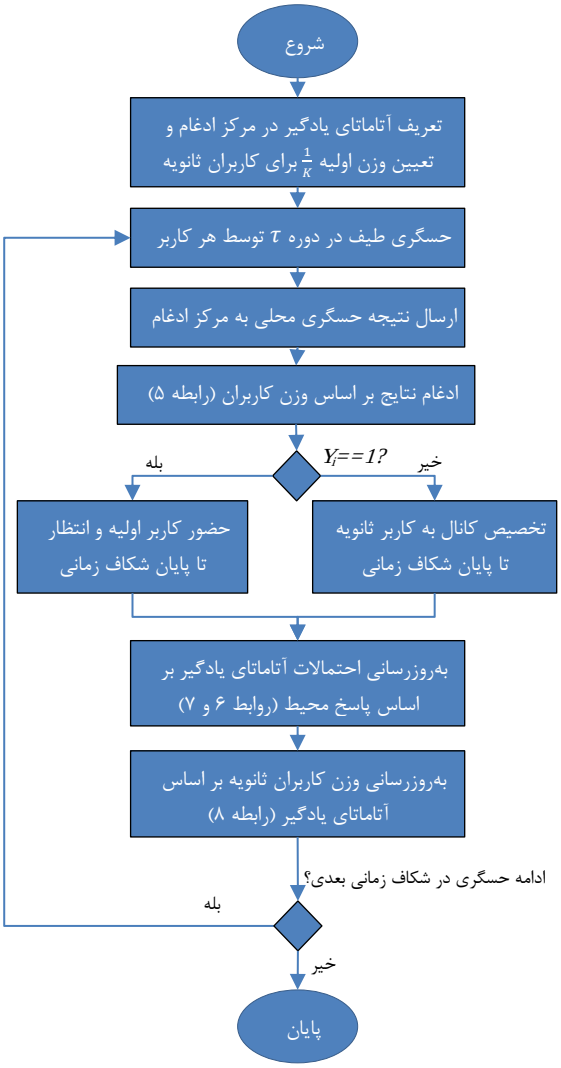
مطابق فلوچارت نمایش داده شده در شکل (۲)، در روش پیشنهادی برای هر کاربر یک پارامتر وزن در نظر گرفته شده است که نشان‌دهنده میزان بهینگی کاربر برای مشارکت در طیف می‌باشد. مقدار بیشتر وزن در روش پیشنهادی موجب می‌شود که اثر مربوط به تصمیم محلی کاربر در حسگری مشارکتی افزایش یابد. از طرفی مقدار وزن پایین‌تر به معنای کاهش اثر مربوط به تصمیم کاربر در حسگری خواهد بود. با فرض وجود K کاربر ثانویه در شبکه می‌توان بردار وزن کاربران در تکرار iام حسگری را به‌صورت $W_i = \{w_1, w_2, \dots, w_K\}$ نمایش داد.

در روش پیشنهادی ابتدا تمامی کاربران عمل حسگری را به‌صورت محلی انجام داده و تصمیم محلی خود مبنی بر حضور یا عدم حضور کاربر اولیه را با استفاده از رابطه (۳) به مرکز ادغام ارسال می‌کنند. سپس با داشتن بردار وزن کاربران در هر دوره، می‌توان با استفاده از رابطه زیر حضور یا عدم حضور کاربر اولیه را تشخیص داد:

$$Y_i = \left[\frac{1}{K}\sum_{k=1}^{K}(w_k \times D_k)\right] \quad (۵)$$

رابطه فوق برای ادغام نتایج تصمیم محلی کاربران در روش پیشنهادی بکار می‌رود. در رابطه فوق، K تعداد کاربران، w_k وزن مربوط به کاربر k-ام و D_k نتیجه تصمیم محلی این کاربر است که به‌صورت یک بیت مشخص می‌شود $(D_k \in \{0,1\})$. همچنین به ازای تمامی کاربران ثانویه همواره رابطه $\sum_{k=1}^{K}(w_k) = 1$ برقرار است. در نتیجه همواره رابطه $0 \leq \frac{1}{K}\sum_{k=1}^{K}(w_k \times D_k) \leq 1$ برقرار خواهد بود. در رابطه (۵)، عملگر [ ] مشخص‌کننده عملگر گرد کردن است. پس از هر بار حسگری طیف و ادغام نتایج، بردار وزن کاربران بر اساس پاسخ دریافت شده از محیط به‌روزرسانی می‌شود. به‌منظور به‌روزرسانی مقادیر وزن مربوط به هر کاربر در روش پیشنهادی از ویژگی یادگیری تقویتی در آتاماتای یادگیر استفاده می‌شود. استفاده از آتاماتای یادگیر موجب می‌شود که بتوان با گذر زمان به شناخت بالاتری از وضعیت کاربران دست یافته و کاربران بدرفتار را از سایرین متمایز نمود.

[1] Channel Gain
[2] Additive White Gaussian Noise



شکل (۲). فلوچارت روش پیشنهادی

در تحقیق جاری، با تجهیز مرکز ادغام به یک آتاماتای یادگیر، عمل به‌روزرسانی وزن کاربران انجام می‌شود. در این روش، آتاماتای تصادفی موجود در مرکز ادغام، بر اساس الگوریتم یادگیر خود وزن کاربران بدرفتار را کاهش داده و از طرفی موجب افزایش وزن مربوط به هر کاربر بهنجار خواهد شد. ساختار آتاماتای مورداستفاده در روش پیشنهادی در شکل (۳) نمایش داده شده است.

شکل (۳). ساختار آتاماتای یادگیر و تعامل آن با محیط

آتاماتای یادگیر نمایش داده شده در شکل (۳)، دارای یک مجموعه از اقدامات برای انتخاب می‌باشد. این مجموعه به‌صورت $A = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ نام‌گذاری می‌شود. هر اقدام در مجموعه A دارای یک احتمال برای انتخاب می‌باشد و انتخاب هر اقدام بر اساس مجموعه احتمالات انجام می‌شود. هر آتاماتا با انتخاب یک اقدام از مجموعه A و اِعمال آن بر محیط، عمل می‌کند. عمل مذکور توسط یک محیط ارزیابی می‌شود و آتاماتا از پاسخ محیط برای انتخاب عمل بعدی خود استفاده می‌کند. در طی این فرایند آتاماتا با تنظیم احتمال اقدامات بر اساس پارامترهای پاداش و جریمه یاد می‌گیرد که کدامین اقدام بهینه بوده و انتخاب آن باید با احتمال بیشتری انجام شود.

در روش پیشنهادی، مجموعه اقدامات آتاماتای یادگیر برابر با تعداد کاربران ثانویه در شبکه می‌باشد. در آغاز کار شبکه، احتمال تمامی اقدامات در آتاماتای یادگیر با هم برابر بوده و برابر با $\frac{1}{K}$ می‌باشد. مجموعه احتمالات آتاماتا به‌صورت $P = \{p_i, p_2, \dots, p_n\}$ نمایش داده می‌شود. هدف آتاماتای یادگیر تعیین این شرایط است که تأثیر کدام کاربر برای مشارکت در حسگری طیف افزایش یافته و اثر کدام کاربر باید کاهش یابد. به‌بیان‌دیگر، آتاماتای یادگیر با تغییر در بردار احتمالات خود، وزن هر کاربر برای مشارکت در فرآیند حسگری را تعیین می‌کند.

بر اساس شرایط ذکر شده، پس از تعیین خروجی سیستم حسگری با استفاده از رابطه (۵)، آتاماتای یادگیر منتظر می‌ماند تا پاسخ محیط را دریافت کند. پس از دریافت پاسخ از محیط ممکن است برای هر اقدام آتاماتا (یا به بیانی دیگر، هر کاربر ثانویه) شرایط زیر رخ دهد:

- اگر نتیجه تصمیم محلی کاربر ثانویه‌ی متناظر با اقدام جاری در آتاماتا با پاسخ محیط یکسان باشد، بنابراین کاربر جاری (اقدام جاری) یک تصمیم بهینه برای آتاماتا محسوب می‌شود. در این حالت، آتاماتای یادگیر، احتمال مربوط به کاربر (اقدام) جاری را با استفاده از رابطه زیر افزایش می‌دهد (اقدام جاری به‌صورت i نمایش داده شده است):

$$p_j(k+1) = \begin{cases} p_j(k) + a[1 - p_j(k)] & j = i, \\ (1-a)p_j(k) & \forall\, j \neq i. \end{cases} \qquad (۶)$$

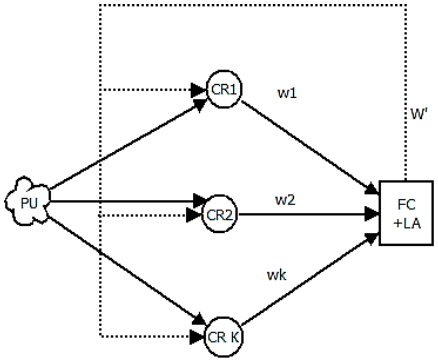
- اگر نتیجه تصمیم محلی کاربر ثانویه‌ی متناظر با اقدام جاری در آتاماتا با پاسخ محیط یکسان نباشد، بنابراین کاربر جاری (اقدام جاری) یک تصمیم اشتباه برای آتاماتا محسوب می‌شود. در این حالت، آتاماتای یادگیر، احتمال مربوط به کاربر (اقدام) جاری را با استفاده از رابطه زیر کاهش می‌دهد:

$$p_j(k+1) = \begin{cases} (1-b)p_j(k) & j = i, \\ \left(\frac{b}{K-1}\right) + (1-b)p_j(k) & \forall\, j \neq i. \end{cases} \qquad (۷)$$

در روابط (۶) و (۷)، پارامترهای a و b، به ترتیب ضرایب پاداش و جریمه می‌باشند و مقدار آن‌ها ۰/۵ در نظر گرفته شده است. همچنین k اندیس زمانی گسسته و K تعداد کاربران ثانویه می‌باشد. پس از اعمال شرایط فوق بر هریک از اقدامات آتاماتای یادگیر (به‌صورت جداگانه برای هر اقدام)، بردار احتمالات آتاماتا به‌روزرسانی می‌شود. در گام بعدی، باید بردار احتمالات به‌روزرسانی شده در آتاماتا را به بردار وزن کاربران ثانویه تغییر داد. با توجه به اینکه باید همواره رابطه $\sum_{k=1}^{K}(w_k)=1$ برای بردار وزن کاربران برقرار باشد، عمل به‌روزرسانی وزن کاربران با استفاده از بردار احتمالات آتاماتا توسط رابطه زیر صورت می‌گیرد:

$$w_i = \frac{p_i}{\sum_{j=1}^{K} p_j} \qquad (۸)$$

در رابطه فوق، w_i وزن مربوط به کاربر iام، p_i احتمال مربوط به اقدام متناظر با کاربر iام و K تعداد کاربران ثانویه می‌باشد. پس از به‌روزرسانی بردار وزن کاربران، از مقادیر به‌دست‌آمده برای حسگری مشارکتی در دور جدید استفاده خواهد شد. ساختار حسگری مشارکتی در روش پیشنهادی به‌صورت شکل (۴) می‌باشد.



شکل (۴). حسگری همکارانه طیف در روش پیشنهادی

مطابق روند نمایش داده شده در شکل (۴)، هر کاربر دارای یک مقدار وزن بوده که تأثیر کاربران ثانویه برای مشارکت در حسگری را مشخص می‌کند. پس از ادغام نتایج بر اساس بردار وزن و تصمیم محلی کاربران، از آتاماتای یادگیر موجود در مرکز ادغام برای به‌روزرسانی وزن کاربران استفاده شده و این روند در طی فرآیند حسگری تکرار خواهد شد.

۴- نتایج و بحث

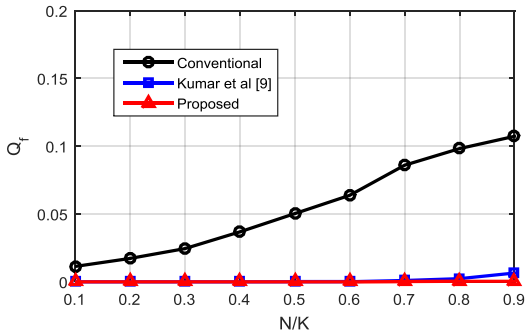
شبیه‌سازی روش پیشنهادی با استفاده از نرم‌افزار MATLAB انجام شده است. در محیط شبیه‌سازی، یک کاربر اولیه با رفتار پویا را در نظر گرفته که وضعیت آن بین حالات فعال و غیرفعال به‌صورت تصادفی تغییر می‌کند. کاربران ثانویه مشارکت‌کننده در سنجش طیف برابر با ۱۰ در نظر گرفته شده است. مقدار پارامتر SNR برابر با dB 10- تعیین شده است. واریانس نویز برابر با $\sigma_w^2 = 1$ بوده و تعداد تکرار شبیه‌سازی ۱۰۰۰۰ دور می‌باشد.

نتایج حاصل از شبیه‌سازی روش پیشنهادی با الگوریتم استاندارد حسگری همکارانه در شبکه‌های رادیوشناختی و همچنین روش حسگری همکارانه مبتنی بر اعتبار در منبع [۹] مقایسه شده است. برای شبیه‌سازی روش پیشنهادی یک سناریو مبتنی بر تعداد کاربران بدرفتار در شبکه در نظر گرفته شده است.

در این آزمایش‌ها فرض شده است که شبکه رادیوشناختی دارای K کاربر ثانویه بوده که N کاربر آن بدرفتار می‌باشند. به‌منظور آزمایش عملکرد روش پیشنهادی، تعداد کاربران بدرفتار را در بازه $1 \le N \le K-1$ تغییر داده و عملکرد شبکه از نظر معیارهای مختلف ارزیابی می‌شود.

در این سناریو، معیارهای احتمال تشخیص صحیح، احتمال هشدار غلط، انرژی مصرفی، گذردهی و کارایی انرژی شبکه محاسبه شده و با روش استاندارد مقایسه می‌شود. مدل محاسباتی این معیارها در [۱] ارائه شده است. در این سناریوی شبیه‌سازی که هدف آن بررسی عملکرد روش پیشنهادی در شرایط تغییر تعداد کاربران بدرفتار است، نسبت کاربران بدرفتار به کل کاربران شبکه در بازه ۰/۱ تا ۰/۹ تغییر می‌یابد. در این بازه، مقدار ۰/۱ بدین معناست که ۱۰ درصد کاربران ثانویه بدرفتار می‌باشند. همچنین نرخ ۰/۹ در این بازه به معنای این است که کاربران بدرفتار، ۹۰ درصد کل کاربران ثانویه را تشکیل می‌دهند.

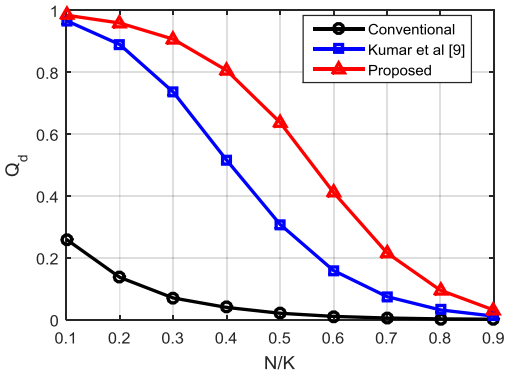
در شکل (۵) نمودار تغییرات احتمال هشدار اشتباه به ازای مقادیر مختلف نسبت کاربران بدرفتار نمایش داده شده است. در این نمودار محور افقی نمایش‌دهنده نسبت بین کاربران بدرفتار به کل کاربران ثانویه در شبکه است. همچنین محور عمودی نشان می‌دهد که به ازای نسبت جاری برای کاربران بدرفتار در شبکه، احتمال هشدار اشتباه به چه مقدار خواهد بود. در این نمودار، نتایج حاصل از پیاده‌سازی روش پیشنهادی با شبکه استاندارد و همچنین روش ارائه شده در منبع [۹] مقایسه شده است.



شکل (۵). احتمال هشدار اشتباه به ازای تعداد کاربران بدرفتار

افزایش تعداد کاربران بدرفتار به معنای افزایش احتمال خطا در حسگری می‌باشد. یعنی با افزایش این دسته از کاربران در شبکه، تشخیص اشتباه حضور کاربر اولیه با احتمال بیشتری رخ می‌دهد. این واقعیت برای روش استاندارد در شکل (۵) کاملاً مشهود است. بر اساس نتایج به‌دست‌آمده، نرخ هشدار اشتباه در روش پیشنهادی به ازای تمامی حالات نزدیک به صفر است. درصورتی‌که در روش استاندارد با افزایش تعداد کاربران بدرفتار، احتمال هشدار اشتباه افزایش می‌یابد. همچنین در روش مبتنی بر اعتبار [۹] احتمال هشدار اشتباه با افزایش نرخ کاربران بدرفتار افزایش می‌یابد. این نمودار نشان می‌دهد که روش پیشنهادی می‌تواند به‌خوبی اثر کاربران بدرفتار را ازنظر نرخ هشدار اشتباه کنترل کند. بدین معنا که در صورت افزایش تعداد کاربران بدرفتار، روش پیشنهادی قادر خواهد بود احتمال هشدار اشتباه را در سطح پایینی نگاه دارد.

در شکل (۶)، نمودار تغییرات احتمال تشخیص صحیح حضور کاربر اولیه به ازای مقادیر مختلف نسبت کاربران بدرفتار نمایش داده شده است.

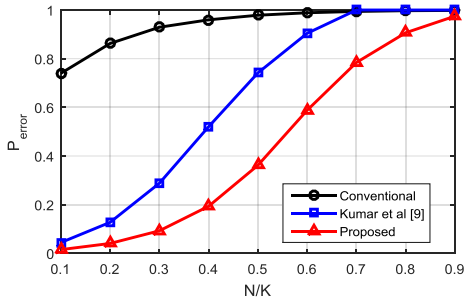


شکل (۶). احتمال تشخیص صحیح به ازای تعداد کاربران بدرفتار

افزایش تعداد کاربران بدرفتار موجب می‌شود که نتوان حضور کاربر اولیه را به‌درستی تشخیص داد. این امر موجب می‌شود که با افزایش تعداد کاربران بدرفتار، احتمال تشخیص صحیح کاربر اولیه کاهش یابد. بر اساس نتایج به‌دست‌آمده، روش استاندار دارای کمترین احتمال تشخیص صحیح حضور کاربر اولیه می‌باشد. زیرا روش استاندارد بین کاربران بهنجار و کاربران بدرفتار تمایزی قائل نمی‌شود. این امر موجب می‌شود که اثر منفی این دسته از کاربران در عملیات حسگری نمایان شود. درصورتی‌که در روش پیشنهادی با به‌کارگیری معیار وزن برای هر کاربر این اثر منفی به حداقل خواهد رسید. این واقعیت در شکل (۶) قابل‌مشاهده است. همان‌طور که در این شکل نمایش داده شده است، روش پیشنهادی می‌تواند همواره حضور کاربر اولیه را با احتمال بیشتری نسبت به روش‌های مقایسه شده تشخیص دهد. این نمودار نشان می‌دهد که روش پیشنهادی می‌تواند به‌خوبی اثر کاربران بدرفتار را ازنظر نرخ تشخیص صحیح کنترل کند. بدین معنا که در صورت افزایش تعداد کاربران بدرفتار، روش پیشنهادی قادر خواهد بود احتمال تشخیص صحیح را در سطح بالایی نگاه دارد.

همچنین در شکل (۷) نمودار تغییرات احتمال خطا در حسگری نمایش داده شده است. احتمال خطا در حسگری به‌صورت زیر قابل‌محاسبه می‌باشد:

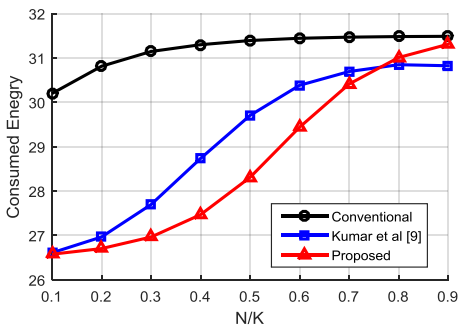
$$P_e = Q_f + Q_m \xrightarrow{Q_m=1-Q_d} P_e = 1 + Q_f - Q_d \qquad (۹)$$



شکل (۷). احتمال خطا در حسگری به ازای تعداد کاربران بدرفتار

بر اساس نتایج حاصل از محاسبه احتمال خطا در حسگری، روش پیشنهادی می‌تواند میزان خطا را به‌صورت قابل‌توجهی کاهش دهد. در شکل (۸) نمودار تغییرات انرژی مصرفی به ازای مقادیر مختلف نسبت کاربران بدرفتار داده شده است. همان‌طور که در این نمودار نمایش داده شده است. مقدار انرژی مصرفی در روش پیشنهادی به ازای مقادیر مختلف نسبت کاربران بدرفتار کمتر از روش استاندارد و روش حسگری مبتنی بر اعتبار [۹] می‌باشد. از طرفی با افزایش نسبت کاربران بدرفتار، میزان انرژی مصرفی در عملیات حسگری افزایش می‌یابد. اما باوجوداین افزایش، روش پیشنهادی می‌تواند میزان انرژی مصرفی را در سطح پایینی نگاه دارد.

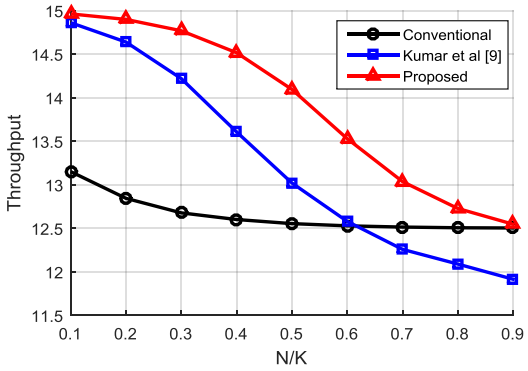
در این آزمایش، میانگین انرژی مصرفی در روش پیشنهادی برابر با ۲۸/۶۹ و در روش‌های استاندارد و مبتنی بر اعتبار به ترتیب برابر با ۳۱/۱۹ و ۲۹,۱۶ می‌باشد. این نتایج نشان می‌دهد که روش پیشنهادی موجب کاهش ۸ درصدی مصرف انرژی نسبت به روش استاندارد و کاهش ۱/۶ درصدی مصرف انرژی نسبت به روش حسگری مبتنی بر اعتبار در منبع [۹] خواهد شد.



شکل (۸). تغییرات انرژی مصرفی به ازای تعداد کاربران بدرفتار

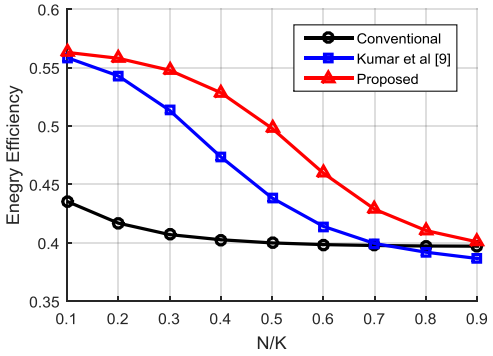
یکی از معیارهای اصلی برای ارزیابی کارایی الگوریتم‌های حسگری طیف، محاسبه گذردهی شبکه می‌باشد. در شکل (۹)، نمودار تغییرات گذردهی به ازای مقادیر مختلف نسبت کاربران بدرفتار نمایش داده شده است.

طبیعی است که با افزایش تعداد کاربران بدرفتار در شبکه، عملکرد شبکه رادیوشناختی در حسگری طیف با مشکل مواجه شده و در نتیجه گذردهی کاهش خواهد یافت. اما استفاده از روش پیشنهادی موجب می‌شود که بتوان به ازای تمامی حالات، میزان گذردهی شبکه را در سطحی بالاتری از روش‌های مقایسه شده نگاه داشت.



شکل (۹). تغییرات گذردهی به ازای تغییرات تعداد کاربران بدرفتار

همان‌طور که در شکل (۹) نمایش داده شده است، میزان گذردهی در روش پیشنهادی به ازای تمامی مقادیر نرخ کاربران بدرفتار، بالاتر از روش‌های مقایسه شده می‌باشد. میانگین گذردهی در روش پیشنهادی برابر با ۱۳/۹ و در روش‌های استاندارد و حسگری مبتنی بر اعتبار [۹] به ترتیب برابر با ۱۲/۶۵ و ۱۳/۲۴ می‌باشد. بنابراین روش پیشنهادی می‌تواند گذردهی در شبکه را نسبت به روش استاندارد به میزان ۹/۸۸ درصد و در مقایسه با روش حسگری مبتنی بر اعتبار به میزان ۴/۹۸ درصد بهبود بخشد. شکل (۱۰) نمودار تغییرات کارایی انرژی به ازای مقادیر مختلف آستانه انرژی را نمایش می‌دهد. کارایی انرژی شبکه رادیوشناختی با محاسبه تقسیم گذردهی بر انرژی مصرفی آن محاسبه می‌شود.



شکل (۱۰). تغییرات کارایی انرژی به ازای تعداد کاربران بدرفتار

مطابق شکل (۱۰)، میزان کارایی انرژی در روش پیشنهادی در تمامی طول تغییرات مقدار آستانه انرژی بیشتر از دو روش مقایسه شده می‌باشد. در این آزمایش، میانگین کارایی انرژی در روش پیشنهادی برابر با ۰/۴۹، در روش استاندارد برابر با ۰/۴ و در حسگری مبتنی بر اعتبار برابر با ۰/۴۵ می‌باشد. بنابراین روش پیشنهادی می‌تواند موجب بهبود ۲۲/۵ درصدی کارایی انرژی نسبت به روش استاندارد و همچنین افزایش ۸/۸۸ درصدی آن نسبت به روش حسگری مبتنی بر اعتبار گردد.

۵- نتیجه‌گیری

در این مقاله یک روش جدید به‌منظور کاهش اثر مخرب کاربران بدرفتار در حسگری همکارانه طیف برای شبکه‌های رادیوشناختی ارائه شد. در روش پیشنهادی از قابلیت یادگیری تقویتی آتاماتای یادگیر استفاده شده است. بدین‌صورت که آتاماتای یادگیر پس از انجام عمل حسگری در هر بار تکرار، بردار وزن کاربران را بر اساس پاسخ دریافتی از محیط (استفاده موفق از طیف یا برخورد) به‌روزرسانی می‌کند. پس از چند بار تکرار حسگری، آتاماتای یادگیر قادر خواهد بود بردار وزن کاربران را به‌صورت بهینه سازمان‌دهی کند. نتایج حاصل از ارزیابی عملکرد روش پیشنهادی در یک محیط شبیه‌سازی‌شده نشان داد که با استفاده از روش پیشنهادی می‌توان انرژی مصرفی شبکه را به میزان ۸ درصد کاهش داده و میانگین گذردهی را نسبت به روش استاندارد به میزان ۹/۸ درصد افزایش داد. همچنین با استفاده از روش پیشنهادی می‌توان میزان کارایی انرژی را به میزان ۲۲/۵ درصد نسبت به راهکار حسگری همکارانه استاندارد افزایش داد.

پیشنهاد می‌شود در کارهای آینده استفاده از مدل‌های احتمالی مانند مدل مخفی مارکوف به‌منظور به‌روزرسانی وزن کاربران شبکه موردبررسی قرار گیرد.

۶- مراجع

[1] S. Khamayseh, & A. Halawani, "Cooperative Spectrum Sensing in Cognitive Radio Networks: A Survey on Machine Learning-based Methods," Journal of Telecommunications and Information Technology, Vol. 82, pp. 1-18, 2020.

[2] P.T. Sivagurunathan, P. Ramakrishnan, & N. Sathishkumar, "Recent Paradigms for Efficient Spectrum Sensing in Cognitive Radio Networks: Issues and Challenges," Journal of Physics: Conference Series, Vol. 1717, pp. 12-57, 2021.

[3] F. Benedetto, & G. Giunta, "A theoretical analysis of asymptotical performance of cooperative spectrum sensing in the presence of malicious users," IEEE Wireless Communications Letters, Vol. 7, pp. 380-383, 2017.

[4] M. Botvinick, S. Ritter, J. Wang, Z. Kurth-Nelson, C. Blundell, & D. Hassabis, "Reinforcement learning, fast and slow," Trends in cognitive sciences, Vol. 23, pp. 408-422, 2019.

[5] S. Levine, A. Kumar, G. Tucker, & J. Fu, "Offline reinforcement learning: Tutorial, review, and perspectives on open problems," arXiv preprint arXiv:2005. 01643, 2020.

[6] A. Sharifi, M. Sharifi, and J. Niya, "Secure cooperative spectrum sensing under primary user emulation attack in cognitive radio networks: Attack-aware threshold selection approach," International Journal of Electronics and Communications (AEÜ), Vol. 10, pp. 1-10, 2015.

[7] X. Liu, C. Sun, M. Zhou, C. Wu, B. Peng, & P. Li, "Reinforcement learning-based multislot double-threshold spectrum sensing with Bayesian fusion for industrial big spectrum data," IEEE Transactions on Industrial Informatics, Vol. 17, pp. 3391-3400, 2020.

[8] Y. Zhang, P. Cai, C. Pan, & S. Zhang, "Multi-agent deep reinforcement learning-based cooperative spectrum sensing with upper confidence bound exploration," IEEE Access, Vol. 7, pp. 118898-118906, 2019.

[9] A. Kumar, N. Gupta, R. Tapwal, & J. Singh, "Trust Aware Scheme based Malicious Nodes Detection under Cooperative Spectrum Sensing for Cognitive Radio Networks," Adjunct Proceedings of the International Conference on Distributed Computing and Networking, Vol. 21, pp. 56-61, 2021.

[10] R. Wan, L. Ding, N. Xiong, W. Shu, & L. Yang, "Dynamic dual threshold cooperative spectrum sensing for cognitive radio under noise power uncertainty," Human-centric Computing and Information Sciences, Vol. 9, pp. 1-21, 2019.

[11] W. Ning, X. Huang, K. Yang, F. Wu, & S. Leng, "Reinforcement learning enabled cooperative spectrum sensing in cognitive radio networks," Journal of Communications and Networks, Vol. 22, pp. 12-22, 2020.

[12] M. Rajendran, & M. Duraisamy, "Distributed coalition formation game for enhancing cooperative spectrum sensing in cognitive radio ad hoc networks," IET Networks, Vol. 9, pp. 12-22, 2020.

[13] E. Ghazizadeh, D. Abbasi-moghadam, & H. Nezamabadi-pour, "An enhanced two-phase SVM algorithm for cooperative spectrum sensing in cognitive radio networks," International Journal of Communication Systems, Vol. 32, pp. 38-56, 2019.