DOR: 20.

DOI: 10.22034/csj.2025.227839

تابستان ۱۴۰۴، دوره ۱۰، شماره ۲، پیاپی ۳۷

دریافت مقاله: ۱۴۰۳/۱۱/۰۱

پذیرش مقاله: ۱۴۰۳/۱۱/۱۷

نوع مقاله: پژوهشی

یک رویکرد مبتنی بر یادگیری تقویتی جهت بهبود مصرف انرژی در شبکه‌های حسگر بی‌سیم به کمک عامل متحرک

علی حیدری پور

دانشجوی کارشناسی گروه مهندسی کامپیوتر، دانشکدهٔ فنی و مهندسی، دانشگاه بین‌المللی امام‌خمینی، قزوین، ایران

پست الکترونیکی: aliheydaripouruce@gmail.com

نستوه طاهری جوان*

استادیار، گروه مهندسی کامپیوتر، دانشکدهٔ فنی و مهندسی، دانشگاه بین‌المللی امام‌خمینی، قزوین، ایران

پست الکترونیکی: nastooh@eng.ikiu.ac.ir

چکیده

محدودیت انرژی و مدیریت مصرف آن یکی از چالش‌های اساسی در شبکه‌های حسگر بی‌سیم محسوب می‌شود. در این شبکه‌ها، گره‌های نزدیک به ایستگاه پایه معمولاً انرژی بیشتری مصرف کرده و زودتر از کار می‌افتند؛ در نتیجه، کارآیی و طول عمر شبکه به شدت کاهش می‌یابد. یکی از راه‌های غلبه بر این مشکل، استفاده از عامل متحرک است که می‌تواند داده‌ها را از گره‌های شبکه جمع‌آوری کرده و حتی آنها را شارژ کند. این پژوهش، یک رویکرد مبتنی بر یادگیری تقویتی برای بهبود عملکرد عامل متحرک ارائه می‌دهد. در رویکرد پیشنهادی، عامل متحرک به صورت هوشمندانه تصمیم می‌گیرد که هنگام رسیدن به یک گره جدید، آیا آن گره را شارژ کند یا داده‌های آن را جمع‌آوری نماید. علاوه بر این، عامل متحرک پس از اتمام عملیات، مقصد بعدی خود را نیز مبتنی بر یادگیری انتخاب می‌کند. برای این منظور، عملکرد عامل متحرک با استفاده از فرآیند تصمیم‌گیری مارکف مدل‌سازی شده، سپس یک راه‌حل مبتنی بر یادگیری تقویتی برای آن ارائه شده است. نتایج شبیه‌سازی نشان می‌دهد که رویکرد پیشنهادی مصرف انرژیِ حسگرها را در کل شبکه به صورت متوازن‌تری توزیع کرده و طول عمر شبکه را در مقایسه با روش‌های پیشین به طور متوسط ۱۶ درصد افزایش می‌دهد.

کلید واژه‌ها: شبکه‌های حسگر بی‌سیم، عامل متحرک، یادگیری تقویتی، فرآیند تصمیم‌گیری مارکفی، بهبود مصرف انرژی، افزایش طول عمر شبکه.

۱- مقدمه

در سال‌های اخیر، شبکه‌های حسگر بی‌سیم و اینترنت اشیاء به علت کارایی بالا استفادهٔ گسترده‌ای پیدا کرده‌اند و در نتیجه تحقیقات زیادی پیرامون آنها صورت می‌پذیرد. شبکه‌های حسگر بی‌سیم شامل حسگرهای کوچکی هستند که می‌توانند اطلاعاتی مانند دما، رطوبت، حرکت و سایر پارامترهای محیطی را جمع‌آوری کرده و به یک ایستگاه پایه (در اینجا گره سینک) ارسال کنند [۱]. این شبکه‌ها در حوزه‌های متفاوتی کاربرد پیدا کرده‌اند، از کاربردهایی

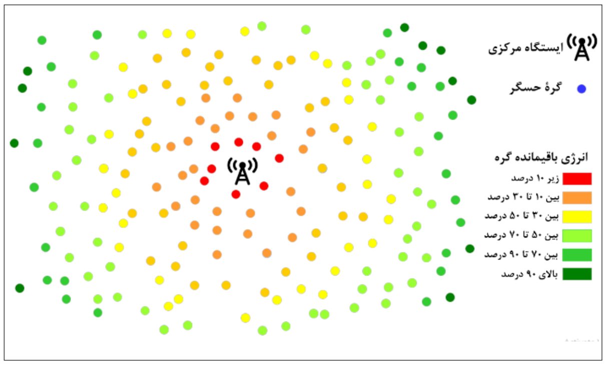
* نویسندهٔ مسئول

مانند کشاورزی هوشمند گرفته تا مسائلی پیرامون نظارت محیطی.

یکی از چالش‌های اصلی در این شبکه‌ها، محدودیت منبع انرژی (عمدتاً باتری) است که باعث می‌شود تلاش برای افزایش طول عمر شبکه در تحقیقات این حوزه اهمیت زیادی پیدا کند [۲]. گره‌های حسگر در عمل برای انجام وظایف مختلف خود، انرژی زیادی مصرف می‌کنند که فرآیند ارسال اطلاعات عموماً بیشتری انرژی را در این شبکه‌ها مصرف می‌کند [۳].

با نگاهی دقیق‌تر به ساختار این شبکه‌ها روشن می‌شود که گره‌ها معمولاً داده‌ها را از طریق مسیرهایی و به صورت گام به گام به سمت ایستگاه پایه ارسال می‌کنند [٤]. به این ترتیب گره‌هایی که به ایستگاه پایه نزدیک‌تر هستند، علاوه بر ارسال داده‌های خود، باید بسته‌های دریافتی از سایر گره‌ها را نیز به سمت ایستگاه پایه جلورانی کنند. این موضوع باعث می‌شود این گره‌ها مصرف انرژی بالاتری داشته باشند و در ادامه زودتر خاموش شده و به گلوگاه تبدیل شوند [٥]. هنگامی که این گره‌ها از کار بیفتند، ارتباط با ایستگاه پایه قطع شده و طول عمر شبکه به شدت کاهش می‌یابد. شکل ۱ میانگین انرژی باقیماندهٔ گره‌های حسگر را در همبندی در نظر گرفته شده نشان می‌دهد، هنگامی که هر گره با نرخ ثابتی ترافیک تولید می‌کند. در این شکل هرچه گره‌ها به ایستگاه پایه (که در این همبندی سعی شده است در مرکز شبکه قرار داده شود) نزدیک‌تر باشند، بار ترافیکی بیشتری را باید جلورانی کنند و در نتیجه انرژی باقیماندهٔ کمتری خواهند داشت. این شکل به خوبی نشان می‌دهد که ایجاد توازن در مصرف انرژی بین حسگرها، کمک شایانی به افزایش طول عمر شبکه خواهد کرد.

در پژوهش‌های پیشین، تلاش‌های زیادی برای بهبود مصرف انرژی در شبکه‌های حسگر بی‌سیم انجام شده است [٦]. یکی از ایده‌های مؤثر، استفاده از ایستگاه پایه متحرک[1] است [٧]. در این رویکردها، ایستگاه پایه به سمت گره‌های حسگر حرکت کرده و خود به عمل جمع‌آوری داده می‌پردازد. این روش کمک می‌کند بار ترافیکی شبکه نسبتاً به‌طور یکنواخت توزیع شده و گره‌ها مصرف انرژی متوازن‌تری داشته باشند، اما در مقابل ثابت نبودن ایستگاه پایه موجب افزایش تاخیر انتها به انتها می‌شود، زیرا در این راه‌کار و به دلیل مشخص نبودن مکان ایستگاه پایه، حسگرها نمی‌توانند داده‌های خود را ارسال کنند و باید حتماً منتظر رسیدن ایستگاه پایه جهت جمع‌آوری داده‌ها بمانند. رویکرد دیگر، استفاده از تکنیک‌های برداشت انرژی از محیط برای گره‌های حسگر است [٨]. در این روش، گره‌های شبکه انرژی خود را از منابع محیطی مانند نور خورشید یا فرکانس‌های رادیویی تامین می‌کنند. این تکنیک باعث کاهش وابستگی گره‌ها به باتری و افزایش دوام شبکه می‌شود، اما در مقابل مجهز کردنِ همهٔ گره‌های شبکه به تکنیک جمع‌آوری انرژیِ محیطی هزینه‌ها را به شدت افزایش می‌دهد و کماکان گره‌های نزدیک به ایستگاه پایه بار ترافیکی بالاتر و در نتیجه مصرف انرژی بیشتری دارند [٩]. در مقابل برخی روش‌های دیگری نیز برای بهبود مصرف انرژی در این شبکه‌ها ارائه شده است، مانند تکنیک‌های فشرده‌سازی، تجمیع اطلاعات و پردازش‌های محلی، که صرفاً سعی دارند حجم داده‌های ارسالی در این شبکه‌ها را کاهش دهند [١٠].



شکل ۱: نحوهٔ مصرف انرژی در شبکه‌های حسگر چندگامه

یکی از رویکردهایی که در سال‌های اخیر در شبکه‌های حسگر بی‌سیم جهت مدیریت مصرف انرژی مورد توجه قرار گرفته است، استفاده از یک عامل متحرک[2] است [١١].

1-Mobile Sink

2-Mobile Agent

عامل متحرک، یک دستگاه یا حسگر متحرک و مستقل است که می‌تواند در محیط شبکه حرکت کرده و داده‌ها را از گره‌ها جمع‌آوری کند یا در حالت‌های پیشرفته‌تر، آن‌ها را شارژ نماید. این روش کمک محسوسی به کاهش بار ترافیکی بر روی گره‌های نزدیک به ایستگاه پایه می‌کند. عامل متحرک با حرکت در شبکه، می‌تواند حسگرها را به‌صورت دوره‌ای شارژ کرده یا داده‌های آن‌ها را جهت حمل به ایستگاه پایه، دریافت کند [۱۲].

پژوهشگران در سال‌های اخیر رویکردهای متنوعی را در استفاده از عامل متحرک در شبکه‌های حسگر بی‌سیم ارائه داده‌اند [۱۳]. به عنوان مثال در تعداد عمده‌ای از تحقیقات، عامل متحرک فقط با حرکت در شبکه به جمع‌آوری اطلاعات از گره‌ها می‌پردازد. از سوی دیگر در برخی از این تحقیقات، عامل متحرک فقط وظیفهٔ شارژ گره‌ها را بر عهده دارد.

نتایج تحقیقات اخیر در حوزهٔ استفاده از عامل متحرک در شبکه‌های بی‌سیم حسگر چندگامه نشان می‌دهند که حسگرهای نزدیک به ایستگاه پایه در مقایسه با حسگرهای دورتر، انرژی بسیار بیشتری مصرف می‌کنند [۱٤]. دلیل این امر مشارکت اجباری حسگرهای نزدیک به ایستگاه پایه در جلورانی بسته‌ها به سمت ایستگاه مرکزی است. این مشاهدات انگیزهٔ اصلیِ این پژوهش را پدید آورده‌اند که به طور خلاصه به صورت زیر بیان می‌گردد: در عمل حسگرهای مختلف خدمات متفاوتی از جانب عامل متحرک نیاز دارند، بعضی بیشتر به شارژ باتری نیاز دارند و بعضی به جمع‌آوری داده.

از سوی دیگر، موضوع شیوهٔ حرکت عامل متحرک در شبکه یکی از چالش‌هایی است که بر کارآیی استفاده از عامل متحرک تاثیرگذار است [۱۵]. در حالت کلی می‌توان سه راه‌کار برای نوع حرکت عامل متحرک در شبکه در نظر گرفت: ۱) عامل متحرک از مسیرهای از پیش تعیین شده استفاده می‌کند، ۲) عامل متحرک به صورت تصادفی در شبکه حرکت می‌کند و ۳) از الگوریتم‌های ابتکاری برای برنامه‌ریزی مسیر استفاده می‌شود. پیرامون تصمیم‌گیری در مورد الگوی حرکت عامل متحرک، رویکرد پیشنهادی، از الگوریتم یادگیری تقویتی استفاده می‌کند تا در هر گام عامل متحرک بتواند از مسیرهای بهتری در راستای بیشینه کردن کارآییِ خود استفاده کند.

به طور خلاصه می‌توان گفت در رویکرد پیشنهادی عامل متحرک هر دو وظیفهٔ جمع‌آوری اطلاعات و شارژ کردن گره‌ها را بر عهده دارد. برای این منظور، عامل متحرک در کل شبکه حرکت می‌کند و به فراخور خدمات مناسب را به گره‌های مختلف ارائه می‌دهد، به این ترتیب که هرگاه عامل متحرک در یک نقطه (و برای ارائهٔ خدمات) توقف کرد، بر اساس شرایط گره‌های آن محدوده سعی می‌کند بهترین گزینه را بین دو گزینهٔ جمع‌آوری اطلاعات یا شارژ گره‌ها انتخاب کند. در این راستا و برای بهبود نحوهٔ تصمیم‌گیریِ عامل متحرک از رویکردی مبتنی بر یادگیری تقویتی [١٦] استفاده می‌شود. در ادامه و مبتنی بر همین رویکرد یادگیری، مقصد بعدی برای عامل متحرک انتخاب می‌شود تا به سوی آن حرکت کند.

در این پژوهش، ابتدا عملکرد عامل متحرک به کمک فرآیند تصمیم‌گیری مارکف۳ [۱۷] مدل‌سازی شده و سپس به کمک الگوریتم یادگیری-Q [۱۸] یک راه‌حل مبتنی بر یادگیری تقویتی برای عملکرد گره پیشنهاد می‌شود. به این ترتیب عامل متحرک پس از مدتی عملکرد خود را بهبود بخشیده و به مرور فرا می‌گیرد که هر کجا باید چه خدمتی را به گره‌های حسگر ارائه دهد و چگونه در شبکه حرکت کند.

یکی از چالش‌های مهم در استفاده از عامل متحرک در شبکه‌های حسگر بی‌سیم، مسئلهٔ اتمام انرژیِ خودِ عامل متحرک است. با توجه به وظایفی که بر عهدهٔ این گره گذاشته می‌شود، مصرف انرژی آن به شدت بالاست و باید فکری به حال آن کرد. در رویکرد پیشنهادی عامل متحرک به ماژول جمع‌آوری انرژی از محیط تجهیز می‌شود، همچنین در صورت نیاز می‌تواند به سمت ایستگاه پایه

3-Markov Decision Process (MDP)

برگشته تا مجدداً شارژ شود.

رئوس نوآوری در این مقاله به قرار زیر هستند:

- تجمیعِ مؤثرِ دو فرآیند جمع‌آوری داده‌ها و شارژ حسگرها توسط عامل متحرک در شبکه‌های حسگر بی‌سیم.
- مدل‌سازی مسئلهٔ تصمیم گیری در عامل متحرک با استفاده از فرآیند تصمیم‌گیری مارکف.
- ارایهٔ یک راه‌حل مبتنی بر یادگیری تقویتی با استفاده از الگوریتم یادگیری-Q.
- انتخاب مسیر بهینه برای حرکت عامل متحرک در شبکه، با استفاده از رویکرد یادگیری تقویتیِ پیشنهادی.
- انجام شبیه‌سازی‌های گسترده و مقایسهٔ کارآیی رویکرد پیشنهادی با کارهای پیشین و مشابه.

ادامهٔ این مقاله به شرح زیر سازماندهی شده است: در بخش ۲، مروری بر کارهای مرتبط ارائه می‌شود. بخش ۳ به معرفی مدل سیستم اختصاص دارد. بخش ٤ جزئیات رویکرد پیشنهادی را تشریح می‌کند. سپس در بخش ٥، عملکرد رویکرد پیشنهادی از طریق آزمایش‌های شبیه‌سازی متعدد ارزیابی و با روش‌های مشابه مقایسه می‌شود. در پایان، بخش ٦ به نتیجه‌گیری و پیشنهادهایی برای پژوهش‌های آتی اختصاص دارد.

۲- پژوهش‌های مرتبط

استفاده از عامل متحرک در شبکه‌های حسگر بی‌سیم به یکی از موضوعات مهم و جذاب تحقیقاتی در دهه اخیر تبدیل شده است. این عامل‌ها وظایفی مانند جمع‌آوری داده و شارژ حسگرها را با بهره‌وری بالا انجام می‌دهند و مزایایی مانند کاهش مصرف انرژی و افزایش طول عمر شبکه را به همراه دارند. در سال‌های اخیر پژوهش‌های مختلفی پیرامون استفاده از عامل‌های متحرک در شبکه‌های حسگر بی‌سیم انجام شده است که در این بخش تعدادی از آنها مرور می‌شوند.

در [۱۹] یک رویکرد برای جمع‌آوری داده‌ها در شبکه‌های حسگر بی‌سیم با استفاده از عامل‌های متحرک پیشنهاد شده است. در این پژوهش، حرکت عامل‌های متحرک با استفاده از الگوریتم‌های ابتکاری[٤] به‌گونه‌ای طراحی شده که مسیرهای بهینه برای کاهش مصرف انرژی و افزایش عمر شبکه انتخاب شوند. نتایج شبیه‌سازی نشان می‌دهند که این روش نه تنها تأخیر را کاهش می‌دهد، بلکه مصرف انرژی را نیز بهبود می‌بخشد. یکی از مزایای این رویکرد کارایی بالای آن در مدیریت انرژی است، اما ایراد آن این است که به دلیل نیاز به هماهنگی دقیق بین عامل‌ها و حسگرها، پیاده‌سازی آن در شبکه های پویا دشوار است.

در [۲۰] راهبردهای شارژ سیار برای شبکه‌های حسگر بی‌سیم بررسی شده است. نویسندگان در این پژوهش روشی را پیشنهاد می‌کنند که با استفاده از دستگاه‌های شارژ سیار، امکان تأمین انرژی گره‌های حسگر را بهینه می‌کند. این روش با در نظر گرفتن عوامل مختلفی مانند مکان گره‌ها، میزان انرژی باقی‌مانده، و نیازمندی‌های زمانی، به کاهش مصرف انرژی و بهبود کارایی کل شبکه کمک می‌کند. نتایج شبیه‌سازی نشان می‌دهد که راهبرد پیشنهادی نسبت به روش‌های موجود، انرژی بیشتری را ذخیره کرده و طول عمر شبکه حسگر را افزایش می‌دهد. با این حال، این روش نیازمند زمان‌بندی دقیق و منابع محاسباتی بالاتری است که در شبکه‌های بزرگ مقیاس غیرعملی است.

در [۲۱] تمرکز بر استفاده از شارژکننده‌های متحرک برای شارژ شبکه‌های حسگر بی‌سیم است. شارژکننده‌های متحرک به‌طور دوره‌ای حرکت کرده و نقاط مشخصی را برای شارژ گره‌ها دنبال می‌کنند. مدل‌سازی صورت گرفته در این پژوهش شامل تحلیل مسیرهای بهینه برای شارژ گره‌ها و زمان‌بندی مناسب آن‌ها است. شبیه‌سازی‌ها نشان داده که این روش می‌تواند عمر شبکه را به طور قابل توجهی افزایش دهد. مزیت این روش، پایداری بالای شبکه است. با این حال اگرچه این روش پایداری شبکه را افزایش می‌دهد، اما در سناریوهای بیدرنگ کارایی مطلوبی ندارد.

در [۲۲] نویسندگان به بررسی روش‌های تجمیع داده‌ها

4-Heuristic

با استفاده از عامل‌های متحرک در شبکه‌های حسگر بی‌سیم می‌پردازنـد. در این رویکرد، حرکـت مأمورهای متحرک هوشـمندانه بوده و بر اساس شرایط شـبکه، مسیرهای آن‌ها تغییر می‌کند. هدف این روش، کاهش انتقال داده‌های تکراری و در نتیجه بهینه‌سـازی مصرف انرژی و کاهش تأخیر در انتقال داده‌ها اسـت. نتایج شبیه‌سـازی نشـان می‌دهد که این روش کارایی کلی شـبکه را افزایش می‌دهد، اما در مقابل نیاز به منابع پردازشی بیشـتری برای اجرای آن دارد. همچنیـن در این روش انرژی گره‌های نزدیک به ایسـتگاه پایه را نادیده گرفته شده و این امر موجب ایجاد گلوگاه در این ناحیه می‌شود.

در [۲۳] یـک رویکرد مبتنی بـر برنامه‌ریزی پویا برای مسیر حرکت عامل‌های متحرک در شبکه‌های حسگر بی‌سیم می‌پـردازد. حرکت عامل‌های متحرک به‌صورت هوشـمند و بر اسـاس تغییرات همبندی شـبکه تنظیم می‌شود و در این راه الگوریتم‌هـای برنامه‌ریزی پویا برای بهبود کارایی شبکه استفاده شـده‌اند. نتایج شبیه‌سازی نشان می‌دهد که ایـن روش می‌تواند مصرف انرژی را کاهش دهد و کارایی شبکه را افزایش دهد. مزیت این روش انعطاف‌پذیری بالای آن است، اما در مقابل عیب بزرگ آن نیاز به توان محاسباتی بالا است که برای گره‌های با منابع محدود مناسب نیست.

در [۲۴] نویسـندگان به بررسی چالش‌های امنیتی در شـبکه‌های حسـگر بی‌سیم مبتنی بـر عامل‌های متحرک پرداخته‌انـد. در این پژوهش حرکـت مأمورهای متحرک به‌صورت دوره‌ای طراحی شـده است، به‌طوری که داده‌ها از مسیرهای مشـخص به مرکز انتقال داده می‌شوند. این پژوهـش همچنین تهدیدات امنیتی موجـود را تحلیل کرده و راهکارهایی بـرای کاهش خطرات ارائـه می‌دهد. نتایج شبیه‌سـازی نشـان می‌دهد که این روش می‌تواند امنیت شـبکه را بهبود بخشـد، با این حال ممکن اسـت به منابع پردازشی بیشـتری نیاز داشـته باشـد. بزرگ‌ترین نقطه ضعف این ایده آن اسـت که تأثیر آن بر افزایش مصرف انرژی نادیده گرفته شده است.

در [۲۵] یک روش برای جمع‌آوری داده‌ها با اسـتفاده از عامل‌های متحرک در شـبکه‌های حسـگر بی‌سیم ارائه شـده اسـت که به طور خـاص مصرف انـرژی را بهینه می‌کند. مدل‌سـازی انجام شـده با در نظر گرفتن حرکت بهینه عامل‌های متحرک به‌گونه‌ای طراحی شـده که نتایج شبیه‌سازی نشان‌دهنده افزایش طول عمر شبکه است. این روش مزیـتی بزرگ در کاهش مصـرف انرژی و افزایش کارایی کلی شـبکه دارد، اما به دلیل اسـتفاده از مدل‌های ایستا، قادر به تطبیق با تغییرات همبندی شبکه نیست.

در [۲۶] روش جدیدی برای بهبود کارایی در جمع‌آوری داده‌ها به کمک عامل متحرک در شبکه‌های حسگر بی‌سیم پیشنهاد شـده است. در این رویکرد از فناوری ZigBee به دلیل ویژگی‌های خود از جمله مصرف کم انرژی و توانایی پشـتیبانی از ارتباطات چندپرشی، به‌عنوان پایه این روش استفاده شده است. نتایج شبیه‌سازی نشان می‌دهد که این رویکرد در مقایسـه با روش‌های سنتی، مصرف انرژی را کاهـش داده و زمان جمع‌آوری داده را بهبود می‌بخشـد. ایراد اصلیِ این روش این اسـت که به دلیل محدودیت‌های پهنای باند در ZigBee در شـبکه‌های بـا ترافیک دادهٔ بالا کارایی کمتری دارد.

مقاله [۲۷] به بررسی طراحی سیسـتمی ترکیبی برای برداشـت انرژی از محیط اطـراف، به‌منظور تأمین انرژی حسـگرهای متحرک و ثابت در شبکه‌های حسگر بی‌سیم با اسـتفاده از عامل های متحرک می‌پردازد. این سیستم با اسـتفاده از انرژی خورشیدی و فرکانس رادیویی به‌طور هم زمان، انـرژی مورد نیاز دسـتگاه‌ها را تأمین می‌کند. نتایج شبیه‌سازی نشان می‌دهد که این روش می‌تواند عمر شبکه را افزایش دهد. نقطه ضعف این رویکرد در پیچیدگی طـراحی و نیاز به هماهنگی بین منابع انرژی مختلف نهفته است.

در [۲۸] روشی برای بهینه‌سازی جمع‌آوری داده‌ها در شبکه‌های حسـگر بی‌سیم به کمک عوامل متحرک معرفی شده است. در این روش، از الگوریتم ژنتیک چندگانه برای

کاهش مصرف انرژی و هزینه‌های انتقال داده در گره‌های حسگر استفاده می‌شود. نتایج ارائه شده نشان می‌دهد که این روش می‌تواند به‌طور مؤثری مصرف انرژی را کاهش دهد. مزیت اصلی این روش، بهینه‌سازی مسیرهای جمع‌آوری داده و کاهش بار روی گره‌های حسگر است. با این حال، پیچیدگی محاسباتی الگوریتم و نیاز به تنظیم دقیق پارامترها می‌تواند به‌عنوان چالشی در مقایسه با روش‌های پیشین مطرح شده باشد.

در [۲۹] روشی برای بهبود کارایی انرژی در شبکه‌های حسگر بی‌سیم ناهمگن با استفاده از عامل‌های متحرک ارائه شده است. در این روش، شبکه شامل سه نوع گره با سطوح انرژی مختلف است: گره‌های برتر، پیشرفته و معمولی. برای گره‌های برتر با انرژی بالا، از یک روش متمرکز برای انتخاب سرخوشه استفاده می‌شود که در آن تابع هزینه‌ای بر اساس فاصله گره‌ها از مرکز ثقل انرژی باقیمانده و نسبت سیگنال به نویز و تداخل به کار می‌رود. سایر گره‌ها با استفاده از عامل‌های متحرک داده‌های خود را به ایستگاه پایه منتقل می‌کنند که این امر منجر به کاهش مصرف انرژی می‌شود. مزیت اصلی این رویکرد بهبود توان عملیاتی شبکه و کاهش تعداد گره‌های از کار افتاده در مقایسه با پروتکل‌های موجود است و در مقابل نقطه ضعف اصلی آن پیچیدگی بیشتر در مدیریت شبکه به دلیل استفاده از روش‌های متمرکز و غیرمتمرکز برای گره‌های مختلف است.

در رویکردهای پیشین در زمینهٔ استفاده از عامل متحرک در شبکه‌های حسگر بی‌سیم، عموماً فقط یکی از وظایف جمع‌آوری داده یا شارژ حسگرها بر عهدهٔ عامل متحرک گذاشته شده است. در رویکرد پیشنهادی در پژوهش، سعی شده است عامل متحرک هر دو وظیفه را به فراخور در قبالِ حسگرها انجام دهد.

۳. مدل سیستم

در این بخش مفروضات اصلیِ در نظر گرفته شده و مدل سیستم در این پژوهش به اختصار آمده‌است.

۳-۱. مدل شبکه

در این پژوهش، شبکه به مربع‌های مختلف تقسیم شده است، به‌طوری که هر مربع نمایانگر یک زیرمجموعه از شبکه است. پهنای باند فرکانس رادیویی عامل متحرک به اندازه‌ای در نظر گرفته شده که تمامی حسگرهای داخل هر مربع را پوشش دهد. در فرآیند توزیع حسگرها، N حسگر به‌صورت یکنواخت در کل شبکه قرار داده شده‌اند. همچنین، تلاش شده است تعداد حسگرها در هر مربع تقریباً یکسان باشد تا توزیع متعادلی حاصل شود.

مختصات قرارگیری هر سنسور از رابطه زیر به دست می‌آید:

$$y = Uniform(y_{min}, y_{max}) \quad x = Uniform(x_{min}, x_{max}) \qquad (۱)$$

هر حسگر همسایه‌های خود را با استفاده از پهنای ارتباطی R_{comm} پیدا می‌کند و ارتباط چندپرشی را به ایستگاه پایه برقرار می‌کند. حسگرها به‌منظور کاهش مصرف انرژی، بسته‌های داده خود را به همسایه‌ای ارسال می‌کنند که کمترین فاصله را با ایستگاه پایه دارد.

حسگرها پس از پخش شدن در شبکه به‌صورت بی‌تحرک باقی می‌مانند. هر حسگر همسایه‌های خود را در حافظه ذخیره کرده است، تا در صورت از کار افتادن نزدیک‌ترین همسایه به ایستگاه پایه، بتواند همسایه دیگری را جایگزین کند همچنین هر حسگر در هر گام زمانی، با توجه به فضای خالی میانگیر خود، بین یک تا سه بسته داده تولید می‌کند. تولید هر بسته به اندازهٔ برابر انرژی مصرف می‌کند. همچنین هر حسگر می‌تواند تا اندازهٔ میانگیر خود، بسته‌ها را ذخیره کند.

در این پژوهش فرض شده است که احتمال انتقال موفقیت‌آمیز داده‌ها توسط گره‌های حسگر وابسته به چندین عامل است. اگر گرهی زنده نباشد یا انرژی کافی نداشته باشد، داده‌ای ارسال نخواهد شد. گره‌ها فقط

می‌توانند داده‌های خود را به گره‌های دیگر یا ایستگاه پایه‌ای که در محدوده ارتباطی آن‌ها قرار دارد، ارسال کنند. اگر گره مقصد ظرفیت ذخیره‌سازی کافی نداشته باشد، انتقال متوقف می‌شود. اگر انرژی گره فرستنده کافی نباشد، نمی‌تواند تمام داده‌های خود را ارسال کند. اگر عامل متحرک در محدوده گره باشد، داده‌ها می‌توانند مستقیماً به او ارسال شوند.

۳- ۲- مدل انرژی

در این پژوهش هر گره حسگر با انرژی اولیهٔ E_{init} شروع به کارکرده و پس از آن انرژی مصرفی برای ارسال بسته به صورت زیر محاسبه می‌شود [۳۰]:

$$E_{tx}(k,d) = E_{elec} \times k + E_{amp} \times k \times d^n \qquad (۲)$$

به طوری که در آن $E_{tx}(k,d)$ انرژی مورد نیاز برای ارسال k بیت داده در فاصله d است و E_{elec} انرژی مصرف شده به ازای هر بیت است که توسط مدار فرستنده یا گیرنده استفاده می‌شود و E_{amp} انرژی مصرف شده توسط تقویت‌کننده فرستنده است و n عددی است که به آن نمای افت مسیر گفته می‌شود. این پارامتر کلیدی در مدل‌های انتشار امواج رادیویی نشان‌دهنده شدت افت قدرت سیگنال در طول مسیر انتقال است، اگر n برابر با ۲ باشد، قدرت سیگنال با افزایش فاصله به صورت ملایم‌تری کاهش پیدا می‌کند. (مانند شرایط فضای باز) و اگر n>2 باشد، این حالت نشان‌دهنده محیطهایی با موانع بیشتر یا جذب انرژی بالاتر است که باعث افت قدرت سیگنال با شدت بیشتری می‌شود (مانند محیطهای شهری یا داخل ساختمان).

به طور مشابه، انرژی مصرف‌شده برای دریافت هر بسته داده به صورت زیر محاسبه می‌شود:

$$E_{ex}(k) = E_{elec} \times k \qquad (۳)$$

۳-۳- مفروضات عامل متحرک

در این پژوهش فرض شده است که عامل متحرک در ایستگاه پایه با شارژ کامل و به اندازهٔ $E_{MA,init}$ شروع به کار می‌کند و در هر وضعیت می‌تواند سه عمل را انجام دهد: ۱) شارژ کردن حسگرهای اطراف خود با کمک امواج رادیویی، ۲) جمع‌آوری داده از حسگرهای اطراف خود با فرستادن پیام درخواست جمع‌آوری داده، و ۳) حرکت به سمت جایگاه مجاور در شبکه.

مدلِ شارژ کردن حسگرها با فاصله d توسط موبایل کارگزار به شکل زیر است [۳۱]:

$$E_{charge} = \varphi \times E_{MA} \times \left(\frac{d.}{d}\right)^n \qquad (٤)$$

که در آن φ مقدار فاکتور بهره‌وری انتقال انرژی بی‌سیم است، به این معنی که نیمی از انرژی به طور موثر به حسگرها می‌رسد و E_{ma} نشان‌دهنده مقدار انرژی موجود برای شارژ کردن است. $d.$ مقدار فاصله مرجع است (این فاصله مرجع برای مدل‌سازی افت مسیر در شبکه‌های بی‌سیم استفاده می‌شود) و n نشان دهندهٔ نمای افت مسیر است و مشخص می‌کند که قدرت سیگنال با چه شدتی در طول مسیر کاهش می‌یابد.

۳- ۴. مدل تصمیم گیری

شیوهٔ تصمیم گیریِ عامل متحرک در این پژوهش با استفاده از فرآیند تصمیم‌گیری مارکف [۱۷] مدلسازی شده و به کمک یک الگوریتم مبتنی بر یادگیری-Q [۱۸] مسئله حل می‌شود تا عامل متحرک بتواند در هر بزنگاه یک تصمیم از بین دو گزینهٔ انجام عملیات شارژ یا جمع آوری داده.

انتخاب بین عملیات شارژ یا جمع‌آوری داده بر اساس مفروضات زیر انجام می‌شود. در هر جایگاه (از شبکه)، عامل متحرک باید ابتدا لیستی از حسگرهای اطراف خود را دریافت کند و میانگین انرژیِ همسایگان را بهمراه میانگین حجم دادهٔ موجود در میانگیر همسایگان را محاسبه کند. سپس بر اساس رویکرد یادگیریِ مبتنی بر یادگیری-Q، سعی می‌کند در هر نقطه تصمیم مناسب را اتخاذ کند.

پـس از ایـن مرحله، نوبـت به انتخـاب مسیر حرکت می‌رسـد. در رویکرد یادگیریِ طراحی شـده، هنگامی که عامل متحرک تصمیم به شارژ یا جمع‌آوری داده می‌گیرد، بر اسـاس پاداشی که دریافت کرده است، به وضعیتی که در آن قرار داشته، ارزشی اختصاص می‌دهد. به این ترتیب هر جایگاه از شـبکه از پیش دارای یک مقدار است و عامل متحرک بر اسـاس ارزش این مقادیر مقصد بعدی خود را از بین چهار گزینهٔ بالا، پایین، راست و چپ انتخاب می‌کند.

۴- رویکرد پیشنهادی

یکی از چالش‌های اصلی در شـبکه‌های حسگر بی‌سیم چند گامی این است که حسگرهایی که در نزدیکی ایستگاه پایه قـرار دارند، به دلیـل حجم زیاد اطلاعـاتی که برای جلورانی به سمت ایستگاه پایه دریافت می‌کنند، به سرعت انرژی خود را مصرف می‌کنند. در این پژوهش، سعی شده است با به‌کارگیری عامل متحرک حسگرهایی که به شارژ نیاز دارند را شارژ کرده و از سوی دیگر دادهٔ حسگرهایی که میانگیر آن‌ها پر شده است را دریافت کرد. با این هدف عامل متحرک در شبکه حرکت کرده و سعی می‌کند در هر بزنـگاه، بهترین تصمیم را برای کمک به افزایش طول عمر شبکه اتخاذ کند.

در شـبکه‌های حسـگر بی‌سیم چندگامی، حسـگرها می‌توانند به‌جای ارسالِ بیدرنگِ داده‌های دریافت شده، آنها را مـدتی در میانگیر خود ذخیره کرده و منتظر سـرکشیِ عامل متحرک بمانند. با این رویکرد صرفه جوییِ زیادی در مصرف انرژی در فرآیند ارسال–دریافت بسته‌ها می‌شود، زیرا در این رویکرد بسته‌ها تنها یک بار و به‌صورت مستقیم به عامل متحرک ارسـال شوند و در عمل انتقال داده‌ها در مسیرهـای چندگامی حذف می‌شـود. این رویکردِ کلی هم موجب کاهش مصرف انرژی در کل شـبکه می‌گردد و هم موجب ایجاد توازنِ مصرف انرژی بین حسگرهای مختلف شـده و با کاهش بار انتقال داده در حسگرهای نزدیک به ایستگاه پایه، طول عمر شبکه را افزایش می‌دهد.

در این رویکرد، عامل متحرک در هر توقف خود با دو مسـئلهٔ تصمیم‌گیری مواجه می‌شود، اولی انتخابِ عملیات مناسـب برای گره‌های اطراف (از بیـن جمع‌آوری داده یا شـارژ) و دومی انتخاب جهت حرکـت برای مرحلهٔ بعدی. در واقع انجام همزمانِ وظایف شـارژ و جمع‌آوری داده‌ها توسـط عامل متحرک در هر توقف بهینه نیست، زیرا این رویکرد باعث هدررفت منابع و انرژی می‌شود. فعال‌سازی مکرر تجهیزات مانند تقویت‌کننده‌ها و ماژول‌های ارتباطی نـه تنها مصرف انرژی عامل متحـرک را افزایش می‌دهد، بلکه کارایی سیسـتم را نیز کاهـش می‌دهد. علاوه بر این، انجـام هر دو وظیفه در هر توقف زمان بیشـتری می‌طلبد و باعث تأخیر در ارائه خدمات به سایر حسگرها می‌شود. در بسیاری از موارد، حسـگرهای شـبکه تنهـا به یکی از این خدمات (شـارژ یا جمع‌آوری داده) نیاز دارند و انجام وظیفه غیرضروری در این شرایط موجب مصرف بیهوده منابـع عامل متحرک در نقاطی می‌شـود که نیاز واقعی به آن وجود ندارد. این مسـئله نه‌تنها زمان خدمات‌رسانی را افزایش می‌دهد، بلکه بهره‌وری انرژی در کل شبکه را نیز با چالش مواجه می‌کند.

در ایـن پژوهش مـا ابتدا عملکرد عامـل متحرک را با اسـتفاده از فرآیند تصمیم‌گیری مارکف مدل‌سازی کرده سـپس با اسـتفاده از یک رویکرد مبتنی بـر یادگیری-Q سـعی در حل آن داریم. در رویکرد یادگیریِ پیشـنهادی به صورت شـهوی می‌توان گفت عامل متحرک (به عنوان عامـل یادگیر)، با محیط اطراف خـود تعامل برقرار کرده و حالت‌های مختلـف را بررسی می‌کند و تا بتواند بهترین عمل را انتخاب می‌کند. پس از انتخاب عمل، عامل بر اساس نتیجه حاصل شده، یک پاداش دریافت می‌کند. هدف اصلی این فرآیند، به حداکثر رساندن میزان پاداش دریافتی است

۴-۱. فرموله کردن مسـئله مبتنی بـر فرآیند تصمیم مارکفی

با توجه بـه خصوصیات فرآینـد تصمیم‌گیریِ عامل

متحرک در مسئلهٔ ما، می‌توان این فرآیند را با استفاده از فرآیند تصمیم‌گیری مارکف فرمول‌بندی کرد. فرآیند تصمیم مارکوفی یک چارچوب تئوری مناسب برای مدل‌سازی و حل مسائل تصمیم‌گیریِ متناوبی است که در آنها یک عاملِ یادگیر باید به صورت مداوم تصمیماتی را اتخاذ کند تا بتواند یک هدف را بهینه‌سازی کند.

ساختار فرآیند تصمیم‌گیری مارکف شامل چهار مولفه اصلی است: ۱) مجموعه‌ای از وضعیت‌ها که نشان‌دهنده وضعیت‌های ممکن سیستم هستند، ۲) مجموعه‌ای از اعمال که عاملِ یادگیر می‌تواند در هر وضعیت انتخاب کند ۳) تابع احتمال انتقال که تغییر از یک وضعیت به وضعیت دیگر را با توجه به یک عمل خاص مشخص می‌کند، و ٤) تابع پاداش که پاداش یا هزینه مرتبط با هر عمل در یک وضعیت خاص را تعیین می‌کند.

در این‌گونه از مسائل، فرآیند تصمیم‌گیری مارکف می‌تواند برای مدل‌سازی رفتار دینامیکی سیستم و یافتن یک سیاست بهینه برای تصمیم‌گیری استفاده شود. این سیاست مشخص می‌کند که عاملِ یادگیر در هر وضعیت چه عملی را انتخاب کند تا پاداش تجمعی در طول زمان بیشینه شود.

وضعیت (S): در مسئلهٔ جاری، وضعیت های عامل متحرک به صورت زیر تعریف می‌شود:

$$s = (Grid\ Index, E_{avg}, B_{avg}) \quad (۵)$$

که در آن Grid Index در واقع مربع (جایگاه) فعلی است که عامل متحرک در بزنگاه تصمیم‌گیری در آن وجود دارد، E_{avg} میانگین انرژی حسگرهای موجود در آن مربع است و B_{avg} میانگین حجم اطلاعات موجود در میانگیر حسگرهای داخل آن مربع است.

اعمال (A): در هر بزنگاه تصمیم گیری، عامل متحرک با گزینه‌های زیر برای انتخاب مواجه است:

- شارژ: شارژ کردن حسگرهای جایگاه فعلی.
- جمع‌آوری داده: دریافت اطلاعات از حسگرهای جایگاه فعلی.

پس از این مرحله، عامل باید تصمیم بگیرد که به کجا حرکت کند. برای این منظور چهار انتخاب دارد: حرکت به سمت بالا، پایین، چپ و راست.

ارزش هر جایگاه بر اساس میزان شارژ و جمع‌آوری اطلاعات، از طریق فرمول زیر محاسبه می‌شود:

$$V_{grid} = w_{collect} \times E_{charged}^{neighbor} + w_{collect} \times D_{collected}^{neighbor} \quad (٦)$$

که در آن $E_{charged}^{neighbor}$ در واقع تخمینی از انرژی است که عامل می‌تواند در جایگاه همسایه شارژ کند و $D_{collected}^{neighbor}$ تخمینی از میزان داده‌ای است که عامل می‌تواند در جایگاه همسایه از حسگرها جمع‌آوری کند و $w_{collect}$ و $w_{collect}$ وزن‌هایی برای متعادل‌سازی اهمیت شارژ و جمع‌آوری داده هستند.

در مرحلهٔ دوم و با توجه به ارزش جایگاه‌های همسایه، عامل متحرک تصمیم می‌گیرد که به بالا، پایین، چپ یا راست حرکت کند. در شکل ۲ شبه کد ارزیابی و انتخاب جایگاه بعدی برای حرکت آمده است.

```
function DecideNextMove (current_grid)
   Initialize V_max ← -∞
   Initialize a_best ← None
   for each action a in A_move do
      neighbor_grid ← Grid resulting from action a
      Estimate E_charged^neighbor and D_collected^neighbor
      Compute V_grid using Equation (15)
      if V_grid > V_max then
         V_max ← V_grid
         a_best ← a
   return a_best
```

شکل ۲: شبکه کد انتخاب مقصد بعدی در رویکرد پیشنهادی

۴- ۲- راه‌کار یادگیری

برای حل مسئلهٔ پیشنهادی از راه‌کار یادگیری تقویتی

استفاده شده است. برای این منظور یک رویکرد مبتنی بر الگوریتم یادگیری-Q ارائه شده است. در حالت کلی می‌توان گفت یادگیری تقویتی یک رویکرد برای حل مسائل تصمیم‌گیری متوالی است که عامل یادگیر با تعامل با محیط، از طریق آزمون و خطا و دریافت پاداش، سیاستی بهینه برای بیشینه‌سازی پاداش بلندمدت یاد می‌گیرد. الگوریتم یادگیری-Q) یکی از روش‌های پایه در یادگیری تقویتی است که بدون نیاز به مدل محیط، یک جدول مقادیر Q(s,a) را برای ارزیابی کیفیت انجام هر عمل a در هر وضعیت s به‌روز می‌کند. هدف، یادگیری سیاستی است که بهترین اعمال را در هر حالت انتخاب کند.

در راهکار پیشنهادی، پاداش‌ها برای تشویق رفتار مطلوب و مجازات‌ها برای تنبیه رفتار غیرمطلوب برای عامل متحرک در نظر گرفته شده اند. مجموع پاداش کل به فاکتورهای مختلفی وابسته است. برای محاسبۀ پاداش عامل متحرک در رابطۀ زیر استفاده می‌شود:

$$r = r_{action} + r_{efficiency} + r_{penalt} \qquad (۷)$$

که در آن r_{action} یا همان پاداش اعمال از رابطه زیر به دست می‌آید:

(۸)

$$r_{action} = \begin{cases} \gamma_{charge} \times E_{charge}, & if\ a = charge \\ \gamma_{collect} \times D_{collected}, & if\ a = collect \\ -\gamma_{move}, & if\ a = move \end{cases}$$

و $r_{efficiency}$ از رابطۀ زیر محاسبه می‌شود:

(۹)

$$r_{efficiency} = \begin{cases} \gamma_{near} \times D_{collected}, & if\ charging\ near\ base \\ \gamma_{far} \times D_{collected}, & if\ collecting\ data\ far\ from\ base \end{cases}$$

و r_{penalt} از رابطۀ زیر محاسبه می‌شود:

(۱۰)

$$r_{penalty} = \{-\gamma_{penalty} \times D_{colleted}, \quad if\ collecting\ near\ base$$

در این رابطه‌ها، a نشان‌دهنده عمل است، E_{charge} میزان انرژی‌ای است که حسگرها شارژ شده‌اند، $D_{collected}$ مقدار داده‌ای است که از میانگیر حسگرها جمع‌آوری شده است و γ_{charge}، γ_{far}، γ_{near}، γ_{move}، و $\gamma_{collect}$ پارامترهای ثابت مثبتی هستند که هر جزء را وزن‌دهی می‌کنند. از سوی دیگر مفاهیم نزدیک و دور از ایستگاه پایه با پارامتری به نام D_{thresh} تعیین می‌شود. اگر فاصله یک جایگاه از این مقدار کمتر باشد نزدیک به ایستگاه پایه محسوب می‌شود، در غیر این صورت دور از ایستگاه پایه به حساب می‌آید.

در رویکرد پیشنهادی، مقادیر-Q، بر اساس رابطۀ زیر به روز می‌شوند [۱۸]:

(۱۱)

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \tau \max_{\acute{a}} Q(\acute{s},\acute{a}) - Q(s,a)]$$

که در آن $\alpha = ۰/۱$ و $\tau = ۰/۹$ است و $\acute{s}$ وضعیت بعدی بعد از انجام دادن عمل a است.

در راهکار یادگیری پیشنهادی، احتمال انتقال به صورت ضمنی پیاده سازی شده است و نه به صورت صریح، به جای آن عامل یادگیر باید از طریق تعامل مکرر با محیط و به روزرسانی مقادیر-Q و مبتنی بر الگوریتم یادگیری-Q این احتمالات را به‌دست آورد.

به طور شهودی می‌توان گفت با اجرای رویکرد یادگیری تقویتی پیشنهادی در یک شبکۀ حسگر بی‌سیم چندگامه، گره‌هایی که به ایستگاه پایه نزدیکترند بیشتر به خدمات شارژ از جانب عامل متحرک نیازمند هستند، زیرا هم به دلیل موقعیت مکانیِ خود بار ترافیکی بیشتری را دریافت می‌کنند، و هم به دلیل نزدیکی به ایستگاه پایه ترجیح می‌دهند بسته‌های خود را مستقیم برای ایستگاه پایه ارسال کنند. از سوی دیگر، گره‌های دور از ایستگاه پایه عموما نیاز به خدمات جمع‌آوری داده دارند، تا از ارسال بسته‌ها به صورت چندگانه در شبکه جلوگیری شود.

۵- نتایج شبیه‌سازی

جهت ارزیابی عملکرد رویکرد پیشنهادی، از شبیه‌سازی استفاده شده است. در این بخش رویکرد پیشنهادی در

کنار سه رویکرد مشابه پیاده‌سازی شده و نتایج حاصل از شبیه‌سازی ارائه شده است.

۵- ۱. محیط شبیه سازی

برای انجام فرآیند شبیه‌سازی، از زبان برنامه‌نویسی پایتون استفاده شده است. جهت پیاده‌سازی فرآیند یادگیری تقویتی از کتابخانهٔ numpy بهره گرفته شده است. هر سناریوی شبیه‌سازی ۳۰ بار اجرا شده است و میانگین نتایج به دست آمده گزارش شده است. خلاصه مقادیر پارامترهای استفاده شده در فرآیند شبیه‌سازی در جدول ۱ آمده است. شایان ذکر است مقادیر پارامترهای مربوط به انرژی در فرآیند شبیه‌سازی عمدتاً بر اساس مرجع [۳۰] انتخاب شده‌اند.

جدول ۱: خلاصه پارامترهای شبیه‌سازی

مقدار	پارامتر
۱۰۰ × ۱۰۰ units	Network area
۱۰۰	Number of nodes (N)
۲۰ units	Initial node energy (E_{init})
۱۵ units	Node communication range (R_{comm})
٤۰۰۰ bits	Data packet size (k)
۵۰ nJ/bit	.E_{elec}
۱۰۰ $pJ/bit/m^۲$	.E_{amp}
۲	Path loss exponent (n)
۵۰ units	Mobile agent initial energy($E_{MA,init}$)
۰,۱ units per move	Movement energy cost(E_{move})
۰,۵	Wireless energy transfer efficiency (μ)
۰,۱	Learning rate (α)
۰,۹	Discount factor (γ)
۰,۱	Exploration rate (ε)

۵- ۱- ۱. رویکردهای مورد مقایسه

جهت ارزیابیِ عملکرد رویکرد پیشنهادی، دو رویکرد دیگر در کنار رویکرد پیشنهادی در فرآیند شبیه‌سازی پیاده‌سازی شدند. رویکردهای مورد مقایسه در این بخش عبارتند از:

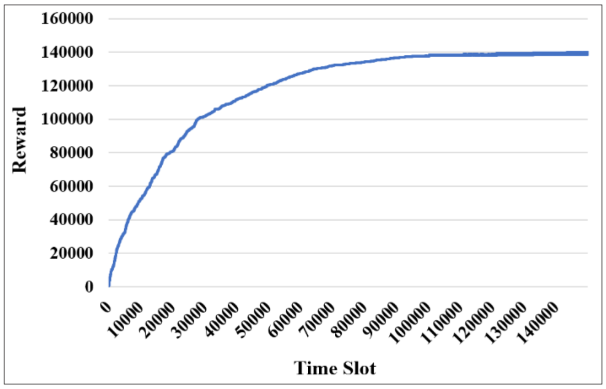
- رویکرد پیشنهادی.
- رویکرد MCh که در آن عامل متحرک فقط خدمت شارژ به حسگرها ارائه می‌دهد [۲۰].
- رویکرد MDC که در آن عامل متحرک فقط خدمت جمع‌آوری داده به حسگرها ارائه می‌دهد [۲٦].

۵- ۲. نتایج عددی

در این بخش نتایج حاصل از شبیه‌سازی ارائه شده است.

۲-۱-۵. همگرایی الگوریتم یادگیری پیشنهادی

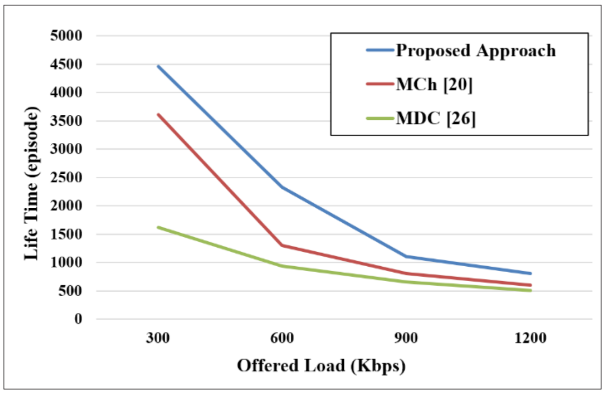
در شکل ۳ همگرایی یادگیری در عامل متحرک نشان داده شده است. همان‌طور که در شکل مشاهده می‌شود تقریبا بعد از هفتاد هزار تکرار الگوریتم یادگیری همگرا می‌شود.



شکل ۳: همگرایی الگوریتم یادگیری پیشنهادی

۵- ۲- ۲- طول عمر شبکه در برابر بار ترافیکی

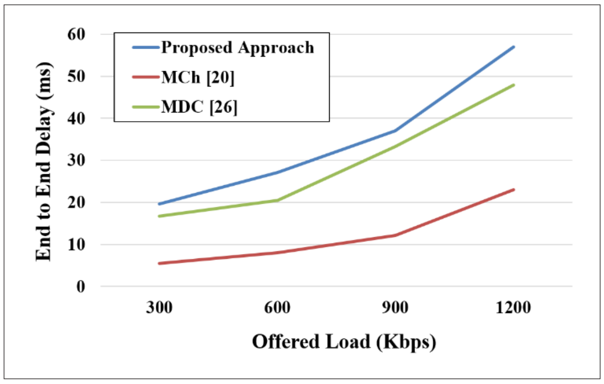
برای بررسی میزان انرژی مصرفی در رویکردهای مورد مقایسه، طول عمر شبکه را در فرآیند شبیه‌سازی محاسبه کرده‌ایم. در اینجا منظور است عمر شبکه عبارت است از مدت زمان از ابتدای فرآیند شبیه‌سازی تا لحظه‌ای که ۵ درصد از گره‌ها به دلیل اتمام ذخیرهٔ انرژی خاموش می‌شوند. شکل ٤ طول عمر شبکه در سه رویکرد مختلف در برابر بار ترافیکی که به شبکه تزریق می‌شود مقایسه شده است. همان‌طور که در این شکل نمایان است، رویکرد پیشنهادی عملکرد بهتری به دلیل توزیع مصرف انرژی در کل شبکه دارد. در واقع در رویکرد پیشنهادی عامل متحرک سعی دارد در هر بزنگاه بهترین تصمیم را برای شبکه اتخاذ کند تا بار ترافیکیِ تحمیل شده به گره‌های نزدیک به ایستگاه مرکزی کاهش یابد.



شکل ۴: مقایسۀ طول عمر شبکه در برابر بار ترافیکی

۵- ۲- ۳. تاخیر انتها-به-انتها

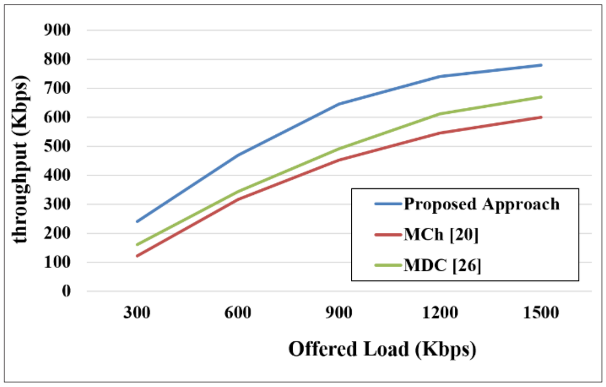
یکی از پارامترهای مهم در ارزیابی کارآیی شبکه‌های حسگر بی‌سیم، میانگین تاخیر انتها به انتهاست. در اینجا منظور از تاخیر انتها برابر است با مدت زمان بین لحظه‌ای که یک داده در یک حسگر تولید می‌شود، تا لحظه‌ای که آن داده به ایستگاه مرکزی تحویل داده می‌شود. شکل ٥ میانگین تاخیر انتها به انتها را به ازای حجم ترافیک اعمال شده به شبکه نشان می‌دهد. همان‌طور که در این شکل نمایان است رویکرد MDC تاخیر زیادی به بسته‌ها اعمال می‌کند. در ادامه رویکرد پیشنهادی نیز قدری تاخیر بالاتری نسبت به رویکرد MCh اعمال می‌کند. به این دلیل که در رویکرد پیشنهادی بخشی از داده‌ها به صورت چندگامه به ایستگاه پایه می‌رسد و بخش دیگر از طریق عامل متحرک منتقل می‌شود.



شکل ۵: مقایسۀ میانگین تاخیر انتها به انتها در برابر بار ترافیکی

۴-۲-۵. نرخ گذردهی

در اینجا نرخ گذردهی را به صورت زیر تعریف کرده‌ایم: حجم داده‌ای که در واحد زمان به صورت صحیح به ایستگاه پایه تحویل می گردد. در شکل ٦ نرخ گذردهی سه رویکرد مورد مقایسه در برابر تغییرات حجم ترافیک تحمیلی به شبکه نشان داده شده است.



شکل ۶: مقایسۀ نرخ گذردهی

۶- نتیجه‌گیری

در این پژوهش، رویکردی جدید برای بهبود مصرف انرژی در شبکه‌های حسگر بی‌سیم با استفاده از عامل متحرک مبتنی بر یادگیری تقویتی ارائه شد. این رویکرد با ترکیب فرآیندهای جمع‌آوری داده‌ها و شارژ گره‌ها توسط عامل متحرک، تلاش می‌کند مصرف انرژی گره‌های نزدیک به ایستگاه پایه را کاهش داده و توزیعِ مصرف انرژی را در شبکه بهبود بخشد. در این راهکار با استفاده از فرآیند تصمیم‌گیری مارکف و الگوریتم یادگیری-Q نحوۀ حرکت و تصمیم‌گیری عامل متحرک بهبود یافته و این امکان را برای عامل متحرک فراهم می‌سازد تا خدمات مناسب را به گره‌های مختلف ارائه دهد. همچنین تجهیز عامل متحرک به ماژول جمع‌آوری انرژی از محیط و امکان شارژ مجدد از ایستگاه پایه، چالش مربوط به اتمام انرژی این عامل را به میزان قابل‌توجهی کاهش داده است. نتایج شبیه‌سازی‌ها نشان دادند که رویکرد پیشنهادی نسبت به روش‌های

پیشین، بهبـــود قابل‌توجهی در افزایش طول عمر شـــبکه، کاهش تأخیر انتها به انتها و توزیع یکنواخت مصرف انرژی دارد.

۷- مراجع

1. Kandris, D., Nakas, C., Vomvas, D. & Koulouras, G. 2020. "Applications of wireless sensor networks: an up-to-date survey," Applied system innovation, vol. 3, no. 1, Art. no. 14.
2. Singh, J., Kaur, R. & Singh, D. 2021. "Energy harvesting in wireless sensor networks: A taxonomic survey," International Journal of Energy Research, vol. 45, no. 1, pp.118-140.
3. Khalaf, O. I., & Abdulsahib, G. M. 2020. "Energy efficient routing and reliable data transmission protocol in WSN," Int. J. Advance Soft Compu. Appl, vol. 12, no. 3, pp.45-53.
4. Karimi-Bidhendi, S., Guo, J., & Jafarkhani, H. 2022. "Energy-Efficient Deployment in Static and Mobile Heterogeneous Multi-Hop Wireless Sensor Networks," IEEE Transactions on Wireless Communications, vol. 21, no. 7, pp. 4973-4988.
5. Altowaijri, S. M. 2022. "Efficient next-hop selection in multi-hop routing for IoT enabled wireless sensor networks," Future Internet, vol. 14, no. 2, Art. no. 35.
6. Sah D. K., & Amgoth, T. 2020. "Renewable energy harvesting schemes in wireless sensor networks: A survey," Information Fusion, vol. 63, pp.223-247.
7. Li, H., Dai, Y., Chen, Q., Liao, D. & Jin, H. 2024."Energy efficient mobile sink driven data collection in wireless sensor network with nonuniform data," Scientific Reports, vol. 14, no. 1, Art. no. 28190.
8. Singh, J., Kaur, R., & Singh, D. 2021. "Energy harvesting in wireless sensor networks: A taxonomic survey," International Journal of Energy Research, vol. 45, no. 1, pp.118-140.
9. Nguyen, M. T., Nguyen, C. V., Tran, H. T. & Viola, F. 2023. "Energy harvesting for mobile agents supporting wireless sensor networks," Energy Harvesting and Systems, vol. 10, no. 2, pp.267-275.
10. Sonai, V., & Bharathi, I. 2024. "A New Statistical Compression-Based Method for Wireless Sensor Networks Energy Efficient Data Transmission," IEEE Sensors Letters, vol. 8, no. 3, pp. 1-4.
11. Saravanan R., & Sathya, V. 2024. "Wireless sensor network data gathering using a multi-fold gravitational search algorithm with a mobile agent," International Journal of System of Systems Engineering, vol. 14, no. 1, pp.23-41.
12. Fissaoui, M. E., Beni-hssane, A., Ouhmad, S. & Makkaoui, K. E. 2021. "A survey on mobile agent itinerary planning for information fusion in wireless sensor networks," Archives of computational methods in engineering, vol. 28, pp.1323-1334.
13. Alsboui, T., Hill, R., Al-Aqrabi, H., Farid, H. M. A., Riaz, M., Iram, S., Shakeel, H. M., & Hussain, M. 2022. "A dynamic multi-mobile agent itinerary planning approach in wireless sensor networks via intuitionistic fuzzy set," Sensors, vol. 22, no. 20, Art. no. 8037.
14. Gao, Z., Liu, C., & Chen, Y. 2024."Scheduling of ERD-assisted charging of a WRSN using a directional mobile charger," IEEE Transactions on Mobile Computing, vol. 23, no. 6, pp. 6681-6696.
15. Kamble, A. A. & Patil, B. M. 2021."Systematic analysis and review of path optimization techniques in WSN with mobile sink," Computer Science Review, vol. 41, Art. no. 100412.
16. Matsuo, Y., LeCun, Y., Sahani, M., Precup, D., Silver, D., Sugiyama, M., Uchibe, E. & Morimoto, J. 2022. "Deep learning, reinforcement learning, and world models, Neural Networks, vol. 152, pp.267-275.
17. Steimle, L. N., Kaufman, D. L. & Denton, B. T. 2021. "Multi-model Markov decision processes," IISE Transactions, vol. 53, no. 10, pp.1124-1139.
18. Clifton, J. & Laber, F. 2020. "Q-learning: Theory and applications," Annual Review of Statistics and Its Application, vol. 7, no. 1, pp.279-301.
19. Xie, G., Ota, K., Dong, M., Pan, F. & Liu, A. 2017. "Energy-efficient routing for mobile data collectors in wireless sensor networks with obstacles," Peer-to-Peer Networking and Applications, vol. 10, no. 3, pp. 472–483.
20. Chen, S T.-S., Chen, J.-J., Gao, X.-Y., & Chen, T.-C. 2022. "Mobile Charging Strategy for Wireless Rechargeable Sensor Networks," Sensors, vol. 22, no. 1, pp. 359.
21. Yang, Y., Li, C., & Pan, W. 2022. "Mobile Charging Strategy for Wireless Rechargeable Sensor Networks," Sensors, vol. 22, no. 1, pp. 359.
22. Lohani, D., & Varma, S. 2016. "Energy Efficient Data Aggregation in Mobile Agent Based Wireless Sensor Network," Wireless Personal Communications, vol. 89, pp. 1165–1176.
23. Alsboui, T., Hill, R., Al-Aqrabi, H., et al., 2022. "A Dynamic Multi-Mobile Agent Itinerary Planning Approach in Wireless Sensor Networks via Intuitionistic Fuzzy Set," Sensors, vol. 22, no. 20, pp. 8037.
24. Teymourzadeh, M., Vahed, R., Alibeygi, S., & Dastanpour, N. 2020. "Security in Wireless Sensor Networks: Issues and Challenges," arXiv preprint arXiv:2007.05111.
25. Lohani, D. & Varma, S. 2016. "Energy Efficient Data Aggregation in Mobile Agent Based Wireless Sensor Network," Wireless Personal Communications, vol. 89, pp. 1165–1176.
26. Tran, H. T., Nguyen, C. V., Phung, N. T., & Nguyen, M. T. 2023. "Mobile Agents Assisted Data Collection in Wireless Sensor Networks Utilizing ZigBee Technology," Bulletin of Electrical Engineering and Informatics, vol. 12, no. 2, pp. 4541–4548.

27. Nguyen, M. T., Nguyen, C. V., Tran, H. T., & Viola, F. 2023. "Energy harvesting for mobile agents supporting wireless sensor networks," Energy Harvesting and Systems, vol. 10, no. 2, pp. 267-275.
28. Saravanan, R. & Sathya, V. 2024. "Wireless sensor network data gathering using a multi-fold gravitational search algorithm with a mobile agent," International Journal of System of Systems Engineering, vol. 14, no. 1, pp. 23-41.
29. Singh, H., Bala, M. & Bamber, S. S. 2020. "Augmenting network lifetime for heterogenous WSN assisted IoT using mobile agent," Wireless Networks, vol. 26, pp. 5965–597.
30. Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. 2000. "Energy-efficient communication protocol for wireless microsensor networks," Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, Maui, HI, USA.
31. Camli, B., Sarioglu, B., & Yalcinkaya, A. D. 2014. "Photodiodes for Monolithic CMOS Circuit Applications," IEEE Journal of Selected Topics in Quantum Electronics, vol. 20, no. 6, pp. 336-343, Art. no. 3803908.