

الگوریتم مسیریابی نوین برای شبکه‌های حسگر بی سیم زیر آب با استفاده از یادگیری تقویتی و شبکه‌های عصبی متخاصم زایشی

نوید کرمی بنماران، دانشجوی کارشناسی ارشد، نستوه طاهری جوان^{*}، استادیار

گروه مهندسی کامپیوتر- دانشکده فنی و مهندسی - دانشگاه بین المللی امام خمینی (ره)- قزوین - ایران

{navid.karami, nastooh} @eng.ikiu.ac.ir

چکیده

شبکه‌های حسگر بی سیم زیر آب با چالش‌هایی مانند محدودیت انرژی گره‌ها، حرکت حسگرها در اعماق مختلف و نیاز به روش‌های ارتباطی متنوع مواجه هستند. این محدودیت‌ها باعث کاهش کارایی الگوریتم‌های مسیریابی مرسوم در سایر شبکه‌های بی سیم چندگانه در این نوع از شبکه‌ها می‌شوند. در این پژوهش، یک روش مسیریابی نوین مبتنی بر یادگیری تقویتی و شبکه‌های عصبی متخاصم زایشی برای کاربرد در شبکه‌های حسگر بی سیم زیر آب معرفی می‌شود. هدف اصلی در این رویکرد، کشف و ذخیره مسیرهای مناسب برای ارسال داده‌ها است. سپس، با استفاده از الگوی شبکه‌های عصبی متخاصم زایشی (GAN)، این مسیرها برای آموزش یک مدل یادگیری عمیق به کار گرفته می‌شوند و سعی بر تولید مسیرهای جدید در این چرخه است. نتایج شبیه‌سازی نشان می‌دهند که استفاده از این رویکرد به بهبود نرخ تحویل موفق بسته‌ها به مقصد و افزایش طول عمر شبکه نسبت به روش‌های یادگیری تقویتی سنتی در شبکه‌های حسگر زیر آب منجر می‌شود.

واژه‌های کلیدی: شبکه‌های حسگر بی سیم زیر آب، مسیریابی، یادگیری تقویتی، یادگیری عمیق، شبکه عصبی متخاصم زایشی.

^{*} نویسنده مسئول: نستوه طاهری جوان، nastooh@eng.ikiu.ac.ir

A Routing Algorithm for Underwater Wireless Sensor Networks Based on Reinforcement Learning and Generative Adversarial Neural Networks

Navid Karami Benmaran, Master Student, Nastooh Taheri Javan*, Assistant Professor

Computer Engineering Department- Imam Khomeini International University- Qazvin- Iran

{navid.karami, nastooh}@eng.ikiu.ac.ir

Abstract:

Underwater wireless sensor networks (UWSNs) face various constraints such as limited node energy, sensor mobility at different depths, and the need for diverse communication methods. These limitations reduce the efficiency of conventional routing algorithms used in other multi-hop wireless networks. In this study, an innovative routing method based on reinforcement learning and generative adversarial neural networks (GANs) is proposed for UWSNs. The proposed approach aims to discover and store optimal data transmission paths within the network. Subsequently, these paths are used to train a deep learning model following the generative adversarial neural network framework, allowing for the generation of new routes. Simulation results demonstrate that the proposed method improves the successful packet delivery rate and extends network lifetime compared to traditional reinforcement learning techniques in UWSNs.

Keywords:

Underwater Wireless Sensor Network, Routing, Reinforcement Learning, Deep Learning, Generative Adversarial Neural Network.

* Corresponding Author: Nastooh Taheri Javan, nastooh@eng.ikiu.ac.ir

۱. مقدمه

آب استفاده می‌شود. این شبکه‌ها شامل حسگرهای متصل به باتری و وسایل نقلیه خودران هستند که در اعماق مختلف دریا قرار گرفته و داده‌ها را با استفاده از امواج صوتی جمع‌آوری می‌کنند. هدف اصلی این شبکه‌ها، جمع‌آوری پیوسته داده‌ها و ایجاد مسیرهای مؤثر برای تبادل اطلاعات است تا پوشش کاملی از منطقه تحت نظارت فراهم شود [۳].

پیاده‌سازی شبکه‌های حسگر بی‌سیم زیر آب در عمل با چالش‌های متعددی روبروست. از جمله این چالش‌ها، نوع ارتباط متفاوت

کاربردهای اینترنت اشیا در سال‌های اخیر رشد قابل توجهی داشته است [۱]. در این میان، شبکه‌های حسگر بی‌سیم به‌طور گسترده‌ای مورد استفاده قرار می‌گیرند [۲] و یکی از مهم‌ترین شاخه‌های آن‌ها، شبکه‌های حسگر بی‌سیم زیر آب است. اقیانوس‌ها و دریاها بیش از ۹۰ درصد از آب زمین را تشکیل می‌دهند و نقش حیاتی در حفظ زندگی موجودات زنده دارند. برای جمع‌آوری داده‌های مربوط به شرایط محیطی زیر آب، از شبکه‌های حسگر بی‌سیم زیر

گره‌ها (استفاده از امواج صوتی)، محدودیت انرژی حسگرها، و تحرک گره‌ها در اعماق مختلف با فشارهای متفاوت آب است [۴]. این محدودیت‌ها منجر به ارائه راه‌حل‌های متنوعی برای بهبود عملکرد در لایه‌های مختلف معماری شبکه شده‌اند.

یکی از مهم‌ترین موضوعات در این شبکه‌ها که به بهبود کارایی آن‌ها در جنبه‌های مختلف کمک می‌کند، پروتکل‌های مسیریابی است. پروتکل‌های مسیریابی برای شبکه‌های حسگر بی‌سیم زیر آب باید بتوانند مقیاس‌پذیری، بهره‌وری انرژی، طول عمر شبکه و سرعت همگرایی را تضمین کنند. هدف اصلی این پروتکل‌ها، افزایش طول عمر شبکه با ارائه مسیرهای بهینه و کاهش مصرف انرژی است [۵].

کارایی الگوریتم‌های مسیریابی در شبکه‌های حسگر بی‌سیم زیر آب تحت تأثیر عوامل متعددی قرار دارد، از جمله هزینه‌های جستجوی همسایه، هزینه‌های ارتباطی، و هزینه‌های محاسباتی. به همین دلیل طراحی پروتکل‌های مسیریابی مخصوص برای شبکه‌های صوتی زیر آب با چالش‌های خاصی مواجه است [۶]. برای مثال، کانال‌های زیر آب دارای ویژگی‌های منحصربه‌فردی مانند تضعیف بالا، پهنای باند محدود، و ناپایداری هستند. همچنین، حسگرها دارای باتری‌های محدودی هستند و به دلیل قرارگیری در زیر آب، از منابعی مانند انرژی خورشیدی برای شارژ مجدد محروم‌اند. علاوه بر این، ارتباطات زیر آب نسبت به ارتباطات رادیویی در خشکی انرژی بیشتری مصرف می‌کنند. بنابراین، بهبود پروتکل‌های مسیریابی با در نظر گرفتن این چالش‌ها می‌تواند به بهبود کارایی این شبکه‌ها کمک کند [۷].

راه‌حل‌های ارائه‌شده در این زمینه معمولاً محدود به بهبود یک پارامتر خاص و استفاده از رویکردهای سنتی هستند. اما با تعریف دقیق مسائل تصمیم‌گیری در گره‌ها و در نظر گرفتن چالش‌های

مذکور، می‌توان به نتایج بهتری دست یافت [۸]. در این مقاله، ما با استفاده از مدل‌سازی مسئله به کمک مدل تصمیم‌گیری مارکوف و به‌کارگیری رویکرد یادگیری تقویتی، یک راهکار جدید برای مسیریابی در محیط‌های تصادفی شبکه‌های حسگر بی‌سیم زیر آب ارائه کرده‌ایم.

در رویکرد پیشنهادی، عامل یادگیرنده در نقاط تصمیم‌گیری تلاش می‌کند تا پاداش تجمعی خود را به حداکثر برساند. برای این منظور، مسئله به گونه‌ای مدل‌سازی شده که عامل با تصمیم‌گیری‌های مناسب در جهت بهبود کارایی شبکه، پاداش مثبت و در غیر این صورت، پاداش منفی دریافت می‌کند. پس از یافتن سیاست بهینه، عامل فرایند یادگیری را کامل می‌کند.

برای بهبود عملکرد الگوریتم مسیریابی در رویکرد پیشنهادی، از شبکه عمیق متخاصم زایشی (GAN) در طول فرایند یادگیری استفاده می‌شود [۹]. شبکه عصبی عمیق متخاصم زایشی نوعی شبکه عصبی است که برای تولید داده‌های جدید مشابه داده‌های واقعی به کار می‌رود. این الگوریتم شامل دو شبکه عصبی مجزا است: یکی شبکه مولد که وظیفه تولید داده‌های جدید مشابه داده‌های واقعی را بر عهده دارد، و دیگری شبکه متمایزکننده که به تشخیص داده‌های واقعی از داده‌های جعلی تولیدشده توسط مولد می‌پردازد. این دو شبکه در رقابتی خصمانه با یکدیگر عمل می‌کنند؛ شبکه مولد سعی می‌کند داده‌هایی ایجاد کند که بتواند متمایزکننده را فریب دهد، در حالی که متمایزکننده تلاش می‌کند داده‌های واقعی را از جعلی تشخیص دهد [۱۰].

در رویکرد مسیریابی پیشنهادی، شبکه متخاصم با استفاده از اطلاعات گراف همسایگی حسگرها، مسیرهای واقعی موجود را ارائه می‌دهد که این امر موجب بهبود الگوریتم مسیریابی و جلوگیری از پیمایش مسیرهای غیرواقعی می‌شود. در این

✓ **ارائه یک راه حل مبتنی بر الگوریتم یادگیری Q- برای حل مسئله مسیریابی:** ارائه الگوریتمی مبتنی بر یادگیری Q- که به صورت پویا و بر اساس وضعیت جاری گره‌ها در زمان انتقال داده، مسیرهای بهینه را انتخاب می‌کند.

✓ **مقایسه عملکرد با روش‌های یادگیری تقویتی سنتی:** مقایسه عملی و شبیه سازی شده رویکرد پیشنهادی با رویکردهای پیشین که معیار مناسبی برای ارزیابی عملکرد رویکرد پیشنهادی است.

ادامه این مقاله از بخش‌های زیر تشکیل شده است. در فصل دوم پیشینه پژوهش به اختصار مرور می‌شود و در فصل سوم جزئیات رویکرد پیشنهادی به تفصیل بیان می‌گردد. در فصل چهارم نتایج به دست آمده از شبیه سازی ارائه شده و در نهایت فصل پنجم جمع بندی و کارهای آتی را ارائه می‌دهد.

۲. پیشینه پژوهش

شبکه‌های حسگر بی سیم زیر آب به دلیل ویژگی‌های خاص محیطی و محدودیت‌های فنی مانند ساختار پویا و جابجایی مداوم گره‌ها، به پروتکل‌های مسیریابی متفاوتی نسبت به شبکه‌های حسگر زمینی نیاز دارند. این شبکه‌ها با چالش‌های متعددی روبرو هستند؛ از جمله محدودیت‌های توان محاسباتی، ذخیره سازی و مصرف انرژی، که به دلیل شرایط محیطی پیچیده تر زیر آب، طراحی پروتکل‌های مسیریابی خاصی را می‌طلبد [۱۱]. شبکه‌های زیر آب از امواج صوتی برای انتقال داده استفاده می‌کنند و با چالش‌هایی نظیر مصرف انرژی بالا، تأخیر زیاد، و ظرفیت محدود مواجه‌اند. به همین دلیل، روش‌های مسیریابی سنتی برای شبکه‌های زمینی قابل استفاده نیستند و نیاز به رویکردهای نوین دارند [۱۲]. در ادامه این بخش تحقیقات اخیر در زمینه مسیریابی در این شبکه‌ها به اختصار مرور می‌شوند. ما این رویکردها را در دو دسته بررسی کرده ایم: الگوریتم‌های مسیریابی بدون استفاده

الگوریتم دو نوع چرخه وجود دارد: چرخه اول که در آن گره حسگر بدون دانش قبلی از شبکه متخاصم درخواست مسیریابی می‌کند و به طور تصادفی یکی از مسیرهای پیشنهادی را انتخاب می‌کند؛ سپس، ارزش مسیر و پاداش‌ها محاسبه و به GAN بازخورد داده می‌شود. در چرخه دوم، گره با توجه به تجربیات قبلی و دانش کسب شده، یکی از مسیرها را انتخاب می‌کند و پس از ارسال بسته، بازخوردی مشابه به GAN و گره‌ها ارسال می‌گردد. تفاوت اصلی این دو چرخه در کسب تدریجی دانش توسط گره حسگر است.

رئوس مشاورکت و نوآوری مقاله حاضر را به شرح زیر می‌توان خلاصه کرد:

✓ **ارائه یک الگوریتم مسیریابی مبتنی بر یادگیری تقویتی در شبکه‌های حسگر زیر آب:** به کارگیری رویکرد یادگیری تقویتی برای مسیریابی در شبکه‌های حسگر بی سیم زیر آب که به عامل یادگیرنده اجازه می‌دهد با بهینه سازی پاداش جمعی، مسیرهای هوشمندانه تری انتخاب شوند.

✓ **استفاده از شبکه عمیق متخاصم زایشی (GAN) برای بهبود عملکرد رویکرد یادگیری تقویتی در فرآیند انتخاب مسیرها:** استفاده از شبکه GAN برای تولید داده‌های مصنوعی و ترکیب آن با داده‌های واقعی به منظور بهبود فرآیند آموزش در مسیریابی در شبکه‌های حسگر زیر آب.

✓ **مدل سازی مسئله مسیریابی در شبکه‌های حسگر زیر آب با استفاده از فرآیند تصمیم گیری مارکوفی (MDP):** بهره گیری از مدل تصمیم گیری مارکوف برای تعریف دقیق مسئله و بهبود فرآیند تصمیم گیری در مسیریابی در محیط‌های تصادفی و پویای زیر آب.

از یادگیری ماشین و الگوریتم های مسیریابی مبتنی بر یادگیری ماشین.

۲-۱: پروتکل های مسیریابی بدون استفاده از یادگیری

ماشین

استفاده از رویکردهای سنتی در مسیریابی شبکه های حسگر بی-سیم زیرآب کماکان محبوب است. به عنوان نمونه مقاله [۱۳] به بررسی رویکردهای مسیریابی جغرافیایی و فرصت محور برای شبکه های حسگر زیرآبی می پردازد. این پژوهش با در نظر گرفتن چالش های منحصربه فرد محیط های زیرآبی، یک الگوریتم مسیریابی جدید ارائه می دهد که با ترکیب اطلاعات جغرافیایی و مکانیزم های فرصت محور، قابلیت اطمینان و کارایی شبکه را بهبود می بخشد. روش پیشنهادی نه تنها بهره وری مصرف انرژی را افزایش می دهد، بلکه قابلیت تطبیق پذیری در برابر تغییرات محیطی و خرابی های احتمالی را نیز فراهم می کند. نتایج شبیه سازی نشان می دهد که این روش در مقایسه با الگوریتم های پیشین عملکرد بهتری از نظر تأخیر، مصرف انرژی و نرخ تحویل بسته دارد.

یکی از پرکاربردترین رویکردهای سنتی در شبکه های حسگر بی-سیم زیرآب استفاده از مسیریابی مبتنی بر خوشه بندی است [۱۴]. در این دسته از پروتکل ها، گره های حسگر به خوشه های جداگانه تقسیم می شوند تا بهره وری انرژی افزایش یابد. هر خوشه دارای یک سرخوشه است که وظیفه جمع آوری داده ها از اعضای خوشه و ارسال آن ها به ایستگاه پایه یا گره سینک را برعهده دارد. خوشه بندی، با تقسیم منطقه به خوشه های پویا، باعث کاهش مصرف انرژی می شود. با این حال، سرخوشه ها به دلیل مصرف انرژی بالا نیاز به انتخاب هوشمندانه دارند تا تعادل مصرف انرژی در شبکه حفظ شود و طول عمر شبکه افزایش یابد [۱۵].

برای مثال، در [۱۶] با استفاده از قدرت سیگنال برای انتخاب سرخوشه، پروتکلی ارائه شده است که به صورت چرخشی سرخوشه ها را انتخاب می کرد. این روش می تواند طول عمر شبکه را افزایش دهد، اما از تعادل انرژی بین گره ها بهره مند نبود. در پژوهشی دیگر [۱۷]، الگوریتمی ارائه شده است که سرخوشه ها را بر اساس انرژی باقی مانده و ارتباطات گره ها انتخاب می کردند، که نتیجه آن بهبود اطمینان و بهره وری انرژی بود؛ با این حال، این روش با چالش هایی همچون توان عملیاتی کم و تأخیر مواجه است.

یکی دیگر از رویکردها، استفاده از ایده کلونی مورچه ها در رویکرد مسیریابی است [۱۸]. در این روش از رفتار مورچه ها برای مسیریابی تقلید می شود و مسیرهایی که بیشترین بازدید را داشته اند، انتخاب می شوند. در این روش مسیرهایی که بیشتر مورد استفاده قرار می گیرند، تقویت می شوند و به تدریج به مسیرهای بهینه تبدیل می شوند. این رویکرد دارای مزایای زیادی است که آن را برای حل مسائل مسیریابی در شبکه های حسگر بی-سیم جذاب می سازد. یکی از این مزایا، توانایی این روش در انطباق با تغییرات دینامیک محیط است. هنگامی که شرایط شبکه، مانند ناپایداری گره ها یا تغییرات توپولوژی، تغییر می کند، الگوریتم های کلونی مورچه ها می توانند به سرعت مسیرهای جدیدی را شناسایی و بهبود بخشند. این سازگاری موجب می شود که شبکه بتواند به طور مؤثر و بدون نیاز به اطلاعات اولیه دقیق، عملکرد خود را بهینه سازی کند.

از سوی دیگر استفاده از الگوریتم ژنتیک نیز برای این منظور استفاده شده است [۱۹]. در این روش با استفاده از انتخاب طبیعی و تابع شایستگی برای بهینه سازی مسیرها استفاده می شود. این الگوریتم ها با تقلید از فرآیندهای طبیعی انتخاب طبیعی و بقای اصلح، به حل مسائل پیچیده و چندمنظوره مسیریابی پرداخته و توانایی بهینه سازی مسیرها را در شرایط مختلف دارند. در این

روش، هر مسیر به عنوان یک کروموزوم در نظر گرفته می‌شود و به آن‌ها نمره‌ای به نام تابع شایستگی اختصاص داده می‌شود که نشان‌دهنده کیفیت آن مسیر است. این تابع بر اساس معیارهای مختلفی مانند مصرف انرژی، زمان انتقال و تعداد گره‌های واسطه محاسبه می‌شود. سپس با استفاده از تکنیک‌های انتخاب، تقاطع و جهش، نسل‌های جدیدی از کروموزوم‌ها تولید می‌شود که احتمالاً بهبود یافته و به سمت بهینه‌سازی حرکت می‌کنند. این فرآیند تکراری تا زمانی ادامه می‌یابد که یک راه‌حل قابل قبول یا بهینه به دست آید. الگوریتم‌ها می‌توانند زمان انتقال داده و مصرف انرژی را به طور چشمگیری کاهش دهند و در عین حال کارایی شبکه را افزایش دهند. از سوی دیگر، این الگوریتم‌ها به دلیل نیاز به محاسبات پیچیده و زمان‌بر، ممکن است در برخی سناریوهای واقعی دچار چالش‌هایی شوند.

استفاده از منطق فازی در مسیریابی شبکه‌های حسگر زیر آب به دلیل پیچیدگی‌های محیط و متغیرهای متعدد بسیار مفید است [۲۰]. این روش با توانایی خود در مدیریت عدم قطعیت‌ها و پردازش اطلاعات ناقص، به تصمیم‌گیری‌های بهینه در انتخاب مسیر کمک می‌کند. با اعمال قواعد فازی، می‌توان به طور همزمان چندین معیار مانند مصرف انرژی، تأخیر و کیفیت سیگنال را مدنظر قرار داد و مسیرهای مؤثرتری را انتخاب کرد. علاوه بر این، منطق فازی می‌تواند به گره‌ها این امکان را بدهد که با توجه به وضعیت لحظه‌ای شبکه و شرایط محیطی، تصمیمات بهتری اتخاذ کنند. این قابلیت انعطاف‌پذیری به ویژه در شرایطی که داده‌های حسگر ممکن است نادرست یا ناقص باشند، ارزشمند است. با استفاده از منطق فازی، مسیریابی می‌تواند به طور همزمان بهینه‌سازی چندین هدف را دنبال کند و انعطاف بیشتری نسبت به رویکردهای سنتی داشته باشد.

در [۲۱] یک پروتکل مسیریابی جدید مبتنی بر فشار برای شبکه‌های حسگر بی‌سیم زیرآبی ارائه شده است. این پروتکل با

هدف افزایش قابلیت اطمینان و بهبود بهره‌وری انرژی طراحی شده و از اطلاعات فشار حسگرها برای انتخاب مسیرهای بهینه و کاهش مصرف انرژی استفاده می‌کند. نتایج ارزیابی نشان می‌دهد که پروتکل ارائه‌شده در مقایسه با روش‌های موجود، از نظر مصرف انرژی، نرخ تحویل داده و طول عمر شبکه عملکرد بهتری دارد.

۲-۲: پروتکل‌های مسیریابی مبتنی بر یادگیری ماشین

امروزه از رویکردهای مبتنی بر یادگیری ماشین برای مسیریابی در شبکه‌های حسگر زیر آب نیز استفاده می‌شوند [۲۲]. در [۲۳] یک پروتکل مسیریابی تطبیقی برای شبکه‌های حسگر آکوستیک زیرآبی ارائه شده که مبتنی بر شبکه‌های عمیق Q است. این پروتکل با هدف بهینه‌سازی مصرف انرژی و کاهش تأخیر طراحی شده است و از یادگیری عمیق برای اتخاذ تصمیمات هوشمند در انتخاب مسیرها استفاده می‌کند. روش پیشنهادی با تحلیل شرایط محیطی و ویژگی‌های شبکه، مسیرهایی را انتخاب می‌کند که مصرف انرژی را کاهش داده و تأخیر در ارسال داده‌ها را به حداقل برساند. نتایج تجربی نشان می‌دهد که این پروتکل در مقایسه با سایر روش‌ها، عملکرد بهتری از نظر طول عمر شبکه و بهره‌وری انرژی دارد.

از سوی دیگر شبکه‌های عصبی توانایی تشخیص الگوهای پیچیده و غیرخطی را دارند و می‌توانند به‌طور مؤثری برای حل مسائل چندمعیاره در مسیریابی استفاده شوند. به‌ویژه در محیط‌های زیر آب که شرایط متغیر و چالش‌های خاصی نظیر تداخل سیگنال و محدودیت‌های انرژی وجود دارد، این شبکه‌ها می‌توانند به ارائه راهکارهای بهینه کمک کنند. با این حال، استفاده از شبکه‌های عصبی در مسیریابی به منابع محاسباتی بالایی نیاز دارد [۲۴]. به‌علاوه، انتخاب ساختار مناسب شبکه و تنظیم پارامترها برای بهینه‌سازی عملکرد، چالش‌هایی هستند که محققان باید در نظر بگیرند. این الگوریتم‌ها می‌توانند با یادگیری از داده‌های تاریخی،

عملکرد خود را در شرایط مختلف بهبود ببخشند و به عنوان یک ابزار قدرتمند در طراحی پروتکل‌های مسیریابی جدید در شبکه‌های حسگر زیر آب مورد استفاده قرار گیرند.

یکی دیگر از رویکردها برای مسیریابی در این نوع از شبکه‌ها، استفاده از الگوریتم‌های مبتنی بر یادگیری تقویتی است [۲۵]. این الگوریتم‌ها به گره‌ها اجازه می‌دهند تا از طریق آزمون و خطا بهترین مسیرها را کشف کنند و به تدریج عملکرد شبکه را بهینه سازند. در این روش‌ها، گره‌ها با ارسال داده‌ها از مسیرهای تصادفی و دریافت پاداش، مانند تأخیر یا نرخ خطا، بهترین مسیرها را شناسایی می‌کنند. جدول (۱) تناظر مولفه‌های یادگیری تقویتی و مسیریابی در شبکه‌های حسگر بی سیم زیر آب را نشان می‌دهد.

یکی از پرکاربردترین الگوریتم‌های یادگیری تقویتی در این حوزه، الگوریتم یادگیری-Q است. هر گره در این الگوریتم، جدولی به نام جدول Q را نگه می‌دارد که ارزش هر مسیر ممکن به مقصد را ثبت می‌کند. با ارسال داده‌ها و دریافت پاداش، گره جدول خود را به‌روزرسانی می‌کند و در نهایت بهترین مسیر را می‌یابد. این الگوریتم، به‌ویژه در شرایطی که گره‌ها اطلاعات کامل از وضعیت شبکه ندارند، بسیار کارآمد است [۲۵].

جدول (۱): مولفه‌های یادگیری تقویتی در مسیریابی شبکه‌های حسگر

عامل یادگیر	حسگر بی سیم
عمل	انتخاب گره بعدی جهت تحویل بسته داده
وضعیت	وضعیت انرژی و موقعیت و ... گره حسگر بی سیم
پاداش	بازخورد نسبت به حسگر انتخاب شده برای تحویل بسته

در این راستا، در [۲۶] الگوریتمی مبتنی بر یادگیری-Q ارائه دادند که با در نظر گرفتن پارامترهایی مانند انرژی باقی‌مانده و تأخیر، گره بعدی را انتخاب می‌کرد. نتیجه این پژوهش کاهش مصرف

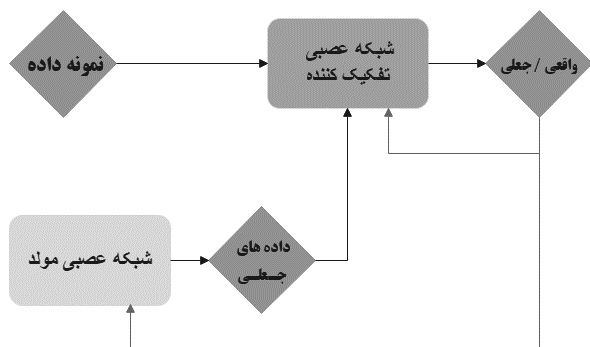
انرژی و تأخیر بود، هرچند گاهی اوقات گره‌های غیر بهینه نیز انتخاب می‌شدند. راهکارهای مبتنی بر یادگیری تقویتی توسعه دادند که به گره‌ها اجازه می‌داد به‌طور فرصت‌طلبانه بسته‌ها را ارسال کنند. این تصمیم‌گیری بر اساس شرایط لحظه‌ای کانال و موقعیت گره‌ها انجام می‌شد، و نتایج نشان داد که مصرف انرژی به‌طور چشمگیری کاهش یافته و تأخیر شبکه بهبود پیدا کرده است.

ترکیب الگوریتم‌های یادگیری تقویتی با نظریه بازی نیز یکی از نوآوری‌های اخیر در حوزه مسیریابی در شبکه‌های حسگر بی سیم زیر آب است [۲۷]. این رویکرد به گره‌ها اجازه می‌دهد تا با استفاده از مدل‌سازی نظریه بازی به عنوان یک بازی چندنفره، مسیریابی توزیع‌شده و کارآمدی را انجام دهند. به‌ویژه در شرایطی که اطلاعات کامل شبکه در دسترس نیست، این ترکیب می‌تواند تصمیم‌گیری‌های هوشمندانه‌ای در مسیریابی ایجاد کند.

۲-۳: جمع‌بندی

روش‌های مختلف مسیریابی در شبکه‌های حسگر بی سیم زیر آب هر کدام دارای مزایا و معایبی هستند. پروتکل‌های مبتنی بر خوشه‌بندی انرژی را بهینه‌تر مدیریت می‌کنند، اما نیازمند انتخاب هوشمندانه سرخوشه‌ها هستند. از سوی دیگر، روش‌های مبتنی بر یادگیری تقویتی و نظریه بازی، با یادگیری از تجربه و آزمون و خطا، مسیریابی توزیع‌شده و پویا را فراهم می‌کنند، هرچند پیاده‌سازی آن‌ها به دلیل پیچیدگی‌های محیطی و محاسباتی چالش‌برانگیز است. در نهایت، انتخاب روش مناسب مسیریابی باید بر اساس نیازها و محدودیت‌های خاص هر شبکه و محیط زیر آب انجام شود.

۳. رویکرد پیشنهادی



شکل (۱): خلاصه عملکرد شبکه های متخاصم زایشی

با توجه به ثابت بودن موقعیت حسگرها و مکانیزم مشخص برای انتخاب مسیرهای بهینه، احتمال وقوع شرایطی وجود دارد که مسیرهای انتخاب شده مشابه یا حتی تکراری باشند. این موضوع به ویژه در شبکه‌هایی که حسگرها از مسیرهای پیشنهادی الگوبرداری می‌کنند، قابل مشاهده است. با این حال، در الگوریتم پیشنهادی ما که مبتنی بر یادگیری Q-طراحی شده، کیفیت مسیرها به صورت دینامیک و بر اساس معیارهای لحظه‌ای ارزیابی می‌شود. بنابراین، حتی در شرایطی که مسیرهای مشابه به عنوان بهینه پیشنهاد شوند، این امر به معنای یک مشکل اساسی در ساختار الگوریتم نیست، زیرا سازوکار یادگیری Q-به گونه‌ای است که قابلیت انطباق با شرایط متغیر شبکه را دارد.

یکی از ویژگی‌های کلیدی الگوریتم یادگیری Q-این است که بر اساس سیستم پاداش‌دهی عمل می‌کند. اگر استفاده مکرر از مسیرهای تکراری منجر به کاهش پاداش شود—برای مثال، به دلیل افزایش مصرف انرژی در گره‌های مشارکت‌کننده، کاهش عمر باتری یا تراکم ترافیکی—الگوریتم این کاهش پاداش را شناسایی کرده و به‌طور خودکار به جستجوی مسیرهای جایگزین می‌پردازد. این فرآیند تطبیقی به الگوریتم اجازه می‌دهد تا نه تنها از مسیرهای کم‌بازده و تکراری اجتناب کند، بلکه بهینه‌ترین مسیرها را با توجه به وضعیت کنونی شبکه شناسایی نماید. به

هدف اصلی رویکرد پیشنهادی در این مقاله، به‌کارگیری شبکه‌های متخاصم زایشی در فرآیند کشف مسیر الگوریتم مسیریابی مبتنی بر یادگیری Q-در شبکه‌های حسگر بی‌سیم زیر آب است. تحقیقات پیشین نشان داده‌اند که شبکه‌های متخاصم زایشی در کاربردهای مختلف، از جمله ارائه مسیرهای بین دو نقطه با استفاده از تصاویر هوایی، بهینه‌سازی تخصیص منابع در شبکه‌های بی‌سیم و تقویت عملکرد سیستم‌های تشخیص نفوذ، بهبودهای چشمگیری ایجاد می‌کنند.

ساختار اصلی GAN شامل دو شبکه عصبی جداگانه به نام‌های مولد و تشخیص‌دهنده است که به صورت رقابتی (متخاصم) در برابر یکدیگر آموزش می‌بینند. هدف اصلی این شبکه‌ها تولید داده‌هایی است که به قدری شبیه به داده‌های واقعی باشند که توسط تشخیص‌دهنده به عنوان داده واقعی شناسایی شوند. نحوه عملکرد GAN به این صورت است که مولد سعی می‌کند نمونه‌هایی شبیه به داده‌های واقعی تولید کند، در حالی که تشخیص‌دهنده وظیفه دارد که بین داده‌های واقعی و داده‌های تولیدشده توسط مولد تفاوت قائل شود. در طول فرآیند آموزش، مولد تلاش می‌کند تا به مرور زمان توانایی خود را در فریب دادن تشخیص‌دهنده افزایش دهد و تشخیص‌دهنده نیز در مقابل بهبود می‌یابد تا تفاوت بین داده‌های واقعی و داده‌های تولیدی را شناسایی کند. این رقابت باعث بهبود کیفیت داده‌های تولیدی می‌شود و در نهایت مولد قادر به تولید نمونه‌های بسیار واقعی از داده‌ها می‌شود. نحوه عملکرد GAN به طور خلاصه در شکل (۱) نشان داده شده است.

عبارت دیگر، سیستم به طور مداوم در حال یادگیری از تجربیات گذشته است و قادر است تصمیمات بهتری برای آینده بگیرد.

این فرآیند تطبیقی به ویژه در شبکه‌های حسگر بی سیم که با محدودیت‌های منابع انرژی، تراکم داده و نوسانات محیطی مواجه هستند، اهمیت بالایی دارد. اگرچه مسیرهای مشابه ممکن است در کوتاه مدت به عنوان بهینه شناسایی شوند، اما الگوریتم پیشنهادی ما به دلیل طراحی پویا و انعطاف پذیر خود، این قابلیت را دارد که در بلندمدت مسیرهای کارآمدتری را پیدا کرده و استفاده از منابع را بهینه کند. این ویژگی همچنین تضمین می کند که شبکه در برابر تغییرات غیرمنتظره مانند خرابی گره ها یا افزایش بار ترافیکی مقاوم باقی می ماند.

در نهایت، این نکته حائز اهمیت است که تکرار مسیرها تنها زمانی به چالشی واقعی تبدیل می شود که بهره وری شبکه به شکل قابل توجهی تحت تأثیر قرار گیرد. در این حالت، الگوریتم یادگیری - Q با اعمال تغییرات در تصمیم گیری های خود و ارائه مسیرهای جدید و کارآمدتر، این مشکل را مدیریت می کند. چنین قابلیت هایی باعث شده که الگوریتم پیشنهادی نه تنها به عنوان راهکاری برای مسیریابی بهینه، بلکه به عنوان سیستمی هوشمند و خودتنظیم برای مدیریت منابع در شبکه های حسگر بی سیم مطرح شود.

در ادامه مراحل رویکرد مسیریابی پیشنهادی برای شبکه های حسگر بی سیم زیر آب تشریح شده اند.

۳-۱: ارزیابی بسته های ارسالی و مسیرهای انتخابی

در روش پیشنهادی، دو معیار کلیدی در تعریف الگوریتم مسیریابی در نظر گرفته شده است، زیرا این دو معیار در شبکه های حسگر بی سیم زیر آب از اهمیت بالایی برخوردارند و به حفظ طول عمر شبکه و ارتباط مؤثر کمک می کنند:

- مقدار و سطح انرژی گره

- فاصله بین دو گره (تراکم) و فاصله نسبت به گره

سینک

فرایند تصمیم مارکوف در رویکرد پیشنهادی به صورت زیر تعریف می شود:

S: وضعیت گره های شبکه از جمله موقعیت و سطح انرژی که بسته از آنجا ارسال می گردد (مبدأ).

A: مسیر و گره ای که برای تحویل بسته به گره پایه انتخاب می شود.

R: پاداش اختصاص یافته به گره مبدأ پس از انتقال بسته از گره به گره دیگر.

تابع پاداش $R(s, a)$ به صورت زیر تعریف می شود:

$$R(s, a) = R(cur, nbr) = \left(-\frac{d(cur, nbr)}{E(cbr)} \right) + E(nbr) \quad (1)$$

که در آن $d(cur, nbr)$ - اختلاف فاصله گره فعلی تا گره بعدی است، $E(nbr)$ انرژی گره بعدی، $E(cbr)$ انرژی گره فعلی، $R(cur, nbr)$ مقدار پاداش جابجایی بسته از وضعیت فعلی به وضعیت بعدی است. شایان ذکر است کسر $\frac{d(cur, nbr)}{E(cbr)}$ در واقع مقدار فاصله حسگر فعلی تا همسایه ای که بسته به آن تحویل داده می شود بر انرژی گره فعلی است و این کسر به صورت منفی در معادله قرار گرفته است و بزرگ تر بودن این عبارت باعث کاهش پاداش خواهد شد و اثر منفی روی مسیریابی خواهد داشت.

انتقال بسته از یک گره به گره دیگر زمانی مؤثر است که در تحویل بسته به گره انتخابی، هم فاصله بین گره ها کم باشد و هم گره مقصد دارای سطح انرژی مناسبی باشد و قبلاً نیز انتخاب نشده باشد. زیرا فاصله زیاد منجر به مصرف انرژی بیشتر توسط گره فعلی می شود. افزایش سطح انرژی گره کاندید (گره بعدی یا همسایه) به عنوان یک پاداش، به طور غیرمستقیم می تواند از

تحويل مکرر بسته به یک گره خاص و در نتیجه انتخاب مکرر آن توسط شبکه متخاصم زایشی جلوگیری کند (بهینه‌سازی محلی). در روش پیشنهادی، مقدار ارزش حالت-عمل به صورت زیر محاسبه می‌شود:

$$\begin{aligned} Q^* &= (1 - \alpha) \text{old}(\text{cur}, \text{nbr}) + \alpha(R(\text{cur}, \text{nbr}) + Q(\text{nbr})) \\ Q &= R(\text{vur}, \text{nbr}) + Q(\text{nbr}) \end{aligned} \quad (2)$$

که در آن α ضریب تاثیر ارزش قبلی گره (یادگیری های قبلی) و $R(\text{cur}, \text{nbr})$ مقدار پاداش جا به جایی بسته از وضعیت فعلی به وضعیت بعدی است. در اینجا Q مقدار ارزش حالت-عمل به ازای یک جابه جایی خواهد بود و Q^* مقدار به روز رسانی آن، به عبارت دیگر Q^* مشخص می‌سازد در قرار دادن ارزش برای یک (حالت - عمل) چقدر به دانش قبلی اهمیت دهیم و چه میزان به دانش و یا همان ارزش محاسبه شده جدید.

در بخشی از فرمول محاسبه ارزش حالت - عمل، مقدار ارزش محض گره $Q(s)$ مشخص می‌شود که از رابطه زیر به دست می‌آید:

$$Q(s) = \frac{E(\text{cur})}{N} \quad (3)$$

که در آن $E(\text{cur})$ انرژی گره جاری و N برابر با تعداد گام فاصله گره جاری تا گره سینک است. این مقدار نشان‌دهنده این است که هر گره‌ای که به گره سینک نزدیک‌تر باشد، می‌تواند ارزش ویژه‌ای داشته باشد. زیرا ارسال بسته به این گره‌ها به صورت دوره‌ای می‌تواند از طولانی شدن مسیر و افزایش تعداد گام‌های نزدیک به هم در مسیرهای ارائه شده توسط شبکه متخاصم زایشی جلوگیری کند. پس از گذر از مسیر و تحويل بسته از گره‌ای به گره دیگر، ارزش کلی مسیر با تقسیم مجموع ارزش‌های گره‌ها به تعداد گره‌های مشارکت‌کننده در مسیر محاسبه می‌شود و به گره‌ها اعلام می‌گردد. این ارزش به شبکه متخاصم زایشی ارائه شده و از طریق رابطه زیر محاسبه می‌شود:

$$Q(\text{Total}) = Q \frac{(\text{Hop } 1 \dots \text{Hop } n - 1)}{N} \quad (4)$$

که در آن $Q(\text{hop } 1, \dots, \text{hop } n)$ برابر است با مقدار مجموع تمام ارزش‌های گره‌های حسگر در طول مسیر و N برابر است با تعداد گام تا گره سینک. در عمل الگوریتم شبکه متخاصم زایشی بازخورد ارزش مسیری را که ارائه کرده است، دریافت می‌کند تا در مراحل بعدی مسیرهای مشابه را تولید کند. همچنین هر گره می‌آموزد که کدام گره‌ها در مسیر می‌توانند مفید باشند تا از بین مسیرهای ارائه شده، انتخاب هوشمندانه‌تری انجام دهد. شبکه متخاصم زایشی تلاش می‌کند با توجه به تجربیات خود، مسیرهای بهتری تولید کند.

در ابتدای کار، پس از استقرار گره‌ها، گره سینک یک بسته حاوی مشخصات خود، از جمله مختصاتش را توزیع می‌کند. هر گره با دریافت این پیام، اطلاعات مربوط به خود را درون بسته قرار داده و آن را مجدداً برادکست می‌کند. در این مرحله، تمام گره‌ها مختصات گره سینک و همسایگان در دسترس خود را دریافت می‌کنند. مسیر به دست آمده از شبکه متخاصم زایشی در سرآیند بسته ارسالی از سوی گره درخواست‌کننده مسیر قرار می‌گیرد تا با جابجایی بسته، گره بعدی بتواند سرآیند را استخراج کرده و دریابد که بسته را به کدام گره تحويل دهد. هر گره حسگر موظف است تا تابع ارزش و تابع پاداش خود را تا گره بعدی محاسبه کند. پس از طی شدن این فرآیند، پیامی مبنی بر صحت انتقال و ارتباط به مبدا از مسیر طی شده بازمی‌گردد. این بسته می‌تواند حاوی اطلاعات لازم برای تخمین ارزش مسیر تولیدشده توسط هر گره در چرخه‌های بعدی باشد. برای تکرار چرخه مسیریابی و استفاده از شبکه متخاصم زایشی و مسیرهای تولیدشده، همواره مسیرها بر اساس دانش فعلی گره و تجربیات پیشین انتخاب می‌شوند، به گونه‌ای که هر مسیری که دارای گره‌های با ارزش‌تری باشد، انتخاب می‌شود.

۳-۲: بررسی ارسال و انتقال بسته ها

در رویکرد پیشنهادی ما در محیط شبکه حسگر بی سیم زیر آب، قبل از شروع فرآیند و در مرحله شناسایی، تمامی گره های حسگر همسایگان خود را شناسایی کرده و موقعیت خود را در شبکه تثبیت می کنند. پس از این مرحله، گره های متخاصم مستقر در محیط اطلاعات مربوط به گراف همسایگی حسگرها را دریافت می کنند. لازم به ذکر است که تمامی حسگرها در موقعیت خود ثابت هستند و حرکات قابل توجهی که بر ارسال و دریافت پیام و تغییر مسیر تأثیرگذار باشد، ندارند.

شبکه متخاصم با در اختیار داشتن اطلاعات گراف همسایگی تمامی حسگرها، قادر است مسیرهای واقعی موجود را ارائه دهد. این ویژگی به الگوریتم کمک می کند تا زودتر به مسیرهای بهینه دست یابد و از پیمایش مسیرهای غیر واقعی اجتناب کند.

در فرآیند مسیریابی گره های حسگر، دو نوع عملکرد وجود دارد که به آن ها چرخه اول و چرخه دوم می گوئیم.

مسیریابی (چرخه اول):

۱. گره حسگر در چرخه اول از شبکه متخاصم درخواست ارسال بسته می کند.

۲. شبکه متخاصم مجموعه ای از مسیرها را در اختیار گره حسگر قرار می دهد.

۳. گره حسگر به طور تصادفی یکی از مسیرها را انتخاب می کند.

۴. مسیر طی شده و ارزش انتقال بسته برای هر پرش حسگر و پاداش محاسبه می شود.

۵. بازخورد به GAN و گره های شرکت کننده در مسیر ارسال می گردد.

مسیریابی (چرخه دوم):

۱. گره حسگر از شبکه متخاصم درخواست ارسال بسته می کند.

۲. شبکه متخاصم مجموعه ای از مسیرها را در اختیار گره حسگر قرار می دهد.

۳. اگر گره حسگر قبلاً سابقه ارسال بسته را داشته باشد، با توجه به دانش قبلی خود یکی از مسیرها را انتخاب می کند (محاسبه ارزش مسیر بر اساس اطلاعات قبلی).

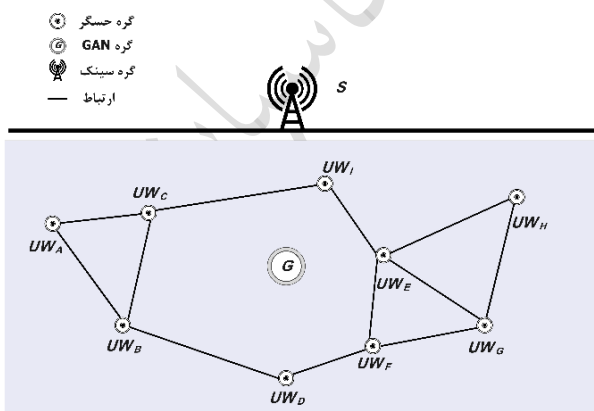
۴. مسیر طی شده و ارزش انتقال بسته برای هر پرش حسگر و پاداش محاسبه می شود.

۵. بازخورد به شبکه متخاصم و گره های شرکت کننده در مسیر ارسال می گردد.

تفاوت میان این دو چرخه در این است که در چرخه اول، گره هیچ دانشی از مسیرها ندارد، اما در چرخه های بعدی، دانش به تدریج کسب می شود.

یک مثال از اجرای رویکرد پیشنهادی در شکل (۲) نشان داده شده است که در آن گره ها به انواع زیر تقسیم می شوند:

- گره های حسگر شبکه
- گره ایستگاه پایه
- گره متخاصم



شکل (۲): نمونه شبکه حسگر بی سیم زیر آب

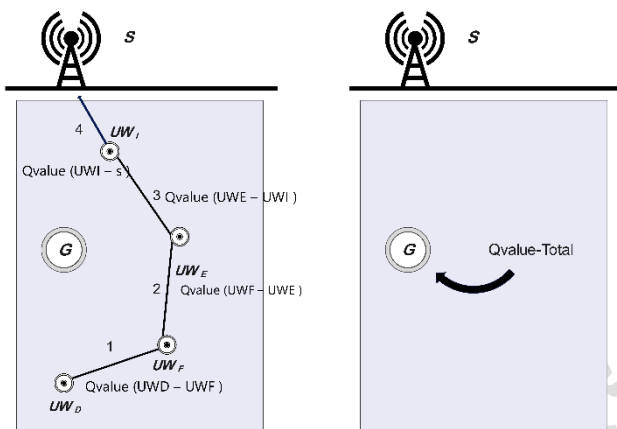
جهت روشن شدن رویکرد پیشنهادی، مثال زیر را در نظر بگیرید.

گره UW_D درخواست ارسال یک بسته دارد. بنابراین گره G

مسیرهای پیشنهادی را به UW_D ارسال می کند. حال گره UW_D

و بسته به سمت مقصد خود ارسال می شود. همچنین همانند چرخه اول مقادیر مربوطه محاسبه شده و در اختیار حسگرها و گره GAN قرار می گیرد.

هریک از گره های حسگر پس از دریافت نتیجه و ارزش انتقال بسته از گره به گره دیگر، جدول مقدار Q را مانند شکل (۵) تکمیل می کنند تا در چرخه ارسال بعدی بتوانند از مسیرهایی که شبکه متخصصم در اختیار آنها قرار می دهد مسیرهایی با ارزش بیشتر را انتخاب کنند.



شکل (۴): ارائه بازخورد به گره ها

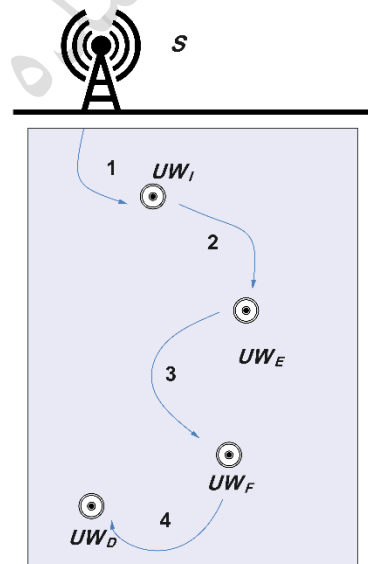
شایان یادآوری است در الگوریتم هایی که بر پایه یادگیری تقویتی بنا شده اند، دو اصل بهره وری و اکتشاف مطرح است. در روش های مبتنی بر یادگیری تقویتی، عامل با حضور در محیط و تجربه حالات جدید می تواند یادگیری را در طی فرآیند پیمایش محیط انجام دهد. از سوی دیگر، عامل می تواند به اطلاعات فعلی خود اتکا کند یا فضای جدیدی از حالات را تجربه کرده و دانش خود را با اطلاعات تازه به روزرسانی کند.

یک مسیر را برای اولین بار به صورت تصادفی انتخاب می کند (مثلاً مسیر تصادفی $UW_D - UW_F - UW_E - UW_I - S$).

در ادامه مسیری که در شکل (۳) نشان داده شده است، پیمایش شده و مقادیر زیر محاسبه می شود:

$$Qvalue-Total = Qvalue(UW_D - UW_F) + Qvalue(UW_F - UW_E) + Qvalue(UW_E - UW_I) + Qvalue(UW_I - S) / HopCount \quad (5)$$

در نهایت نتایج به دست آمده باید برای همه گره ها ارسال شده تا در حافظه خود ذخیره کنند و در نهایت مقدار ارزش تجمیعی مسیر را در اختیار گره GAN قرار دهند.



شکل (۳): نحوه پیمایش مسیر

در ادامه و در چرخه های بعدی، اگر یک گره یک بسته برای ارسال داشته باشد، شبکه متخصصم با بازخوردی که از مسیرهای تولیدی در چرخه های قبلی گرفته سعی خواهد کرد تا مسیر بهینه تری را تولید کند. سپس لیست مسیرها در اختیار حسگر قرار می گیرد و حسگر با توجه به یادگیری قبلی خود تخمین می زند که کدام مسیر ارزش بالاتری دارد (روند بازگرداندن نتیجه مسیریابی و ارزش ها در شکل (۴) نمایش داده شده است). یعنی از میان مسیرهای ارائه شده، انتقال به سمت گره ای صورت می گیرد که ارزش بهتری داشته باشد. در ادامه گره مسیر را انتخاب کرده

در اولین پیمایش، وابستگی مسیریابی به یک گره خاص را از بین می‌برد.

- کشف مسیرهای ناشناخته: این روش قادر است مسیرهایی را کشف کند که ممکن است با روش‌های سنتی مسیریابی قابل شناسایی نباشند. برای نمونه، با تغییر سطح انرژی گره‌ها، گاهی ممکن است بسته‌ها از مسیر کوتاه‌تر به یک مسیر طولانی‌تر جدید هدایت شوند، که این تغییر می‌تواند به افزایش طول عمر شبکه کمک کند.

- انطباق با تغییرات شبکه: این روش می‌تواند خود را با تغییرات در توپولوژی شبکه و شرایط محیطی سازگار کند.

۴. نتایج شبیه سازی

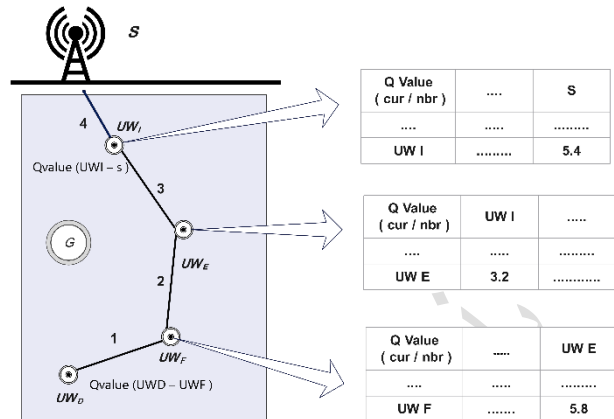
در این بخش جهت ارزیابی رویکرد پیشنهادی، آن را در کنار یک الگوریتم مسیریابی مبتنی بر یادگیری تقویتی [۲۷] پیاده‌سازی کرده‌ایم تا بتوانیم کارایی و عملکرد رویکرد پیشنهادی را به خوبی ارزیابی کنیم. برای شبیه سازی از ابزار متلب نسخه 2024a و کتابخانه یادگیری عمیق استفاده کرده‌ایم. در این شبیه سازی از یک شبکه حسگر بی سیم با تعداد ۵۰ گره در یک محیط با ابعاد ۱۰۰۰ در ۱۰۰۰ متر استفاده شده است و نتایج ارائه شده میانگین ۲۵ بار اجرای هر آزمایش هستند.

۴-۱: جزئیات پیاده سازی رویکرد پیشنهادی

برای پیاده سازی دقیق رویکرد پیشنهادی و ایجاد یک مدل GAN برای تولید مسیر و یکپارچه سازی آن با یادگیری-Q در شبکه حسگر بی سیم مراحل و تنظیمات زیر در نظر گرفته شده اند:

مرحله ۱: ایجاد داده‌های تصادفی مسیر

با توجه به نیاز به داده‌های آموزشی برای آموزش مدل، ابتدا داده‌های آموزشی متناسب با شرایط مسئله تولید می‌شوند. برای



شکل (۵): تکمیل جدول یادگیری برای هر گره حسگر

در راهکار پیشنهادی، از میان مسیرهای ارائه‌شده توسط گره GAN، گره حسگر مسیرهایی را انتخاب می‌کند که ارزش بالاتری را برای او به ارمغان می‌آورد و بدین ترتیب بهره‌وری شبکه را افزایش می‌دهد. در این میان، گره حسگر ممکن است به‌طور تصادفی با حالت‌هایی مواجه شود که تجربه‌ای از آن‌ها ندارد و برای اولین بار با آن‌ها روبه‌رو می‌شود که این حالت به عنوان یک فرآیند اکتشافی در نظر گرفته می‌شود.

الگوریتم GAN که در شبکه حسگر برای کشف مسیر ارائه شده- است به صورت زیر تعریف می‌شود: شبکه تولید کننده به ترتیب از ۱۲۸، ۲۵۶ و ۳۰ لایه و شبکه تشخیص دهنده از لایه‌هایی به تعداد ۱۲۸، ۲۵۶ و ۱ تشکیل می‌شود. ورودی داده‌ها، که شامل اطلاعات گره‌های مسیر است، در تولیدکننده پردازش می‌شود و پس از پردازش به شبکه تفکیک‌کننده تحویل داده می‌شود. سپس این داده‌ها بر اساس بازخوردهای دریافتی توسط الگوریتم یادگیری-Q ارزیابی می‌شوند.

استفاده از شبکه‌های متخاصم زایشی در این روش پیشنهادی می‌تواند نتایج زیر را به دنبال داشته باشد:

- تولید مسیرهای متنوع: این روش می‌تواند مسیرهای بسیار متنوعی ایجاد کند، که شانس یافتن مسیر بهینه را افزایش می‌دهد. به عنوان مثال، انتخاب تصادفی مسیر

مرحله ۴: آموزش مدل شبکه متخاصم زایشی

در این مرحله، پس از تعریف اصطلاحات مربوط به آموزش، به آموزش مدل GAN با استفاده از داده‌های موجود در مجموعه داده‌ها می‌پردازیم. برای این منظور، یک تابع به نام trainGAN ایجاد کردیم که مدل را آموزش می‌دهد.

بهینه‌سازهای مختلفی برای اهداف متفاوت مانند طبقه‌بندی یا رگرسیون وجود دارند، اما برای داده‌های ما، بهینه‌ساز Adam نتایج بهتری ارائه داده است. همچنین، تابع زیان Cross-Entropy انتخاب شده است. این تابع باعث می‌شود که شبکه عصبی نه تنها خروجی‌ای نزدیک به خروجی واقعی تولید کند، بلکه با اطمینان بالا کلاس هدف را تشخیص دهد.

مرحله ۵: تولید مسیر

پس از آموزش شبکه تولیدکننده، تولید مسیرهای جدید آغاز می‌شود. این کار با استفاده از تولیدکننده آموزش‌دیده و تولید ورودی‌های تصادفی برای ایجاد مسیرهای جدید انجام می‌شود.

مرحله ۶: پیاده‌سازی یادگیری-Q برای مسیریابی

یادگیری-Q یک الگوریتم یادگیری تقویتی بدون نیاز به نمونه‌برداری است که ارزش هر عمل را در یک وضعیت خاص یاد می‌گیرد. این الگوریتم از چندتایی (S, A, α, γ) استفاده می‌کند، که در آن S مجموعه حالات، A مجموعه اعمال، α نرخ یادگیری و γ عامل کاهش است. سپس یک جدول Q ایجاد می‌شود و با استفاده از معادله بلمن، هنگام تعامل عامل با محیط (در اینجا مسیرهای تولید شده) به‌روزرسانی می‌گردد.

مرحله ۷: تکرار و ترکیب مسیرها

پس از تولید مجموعه‌ای از مسیرها توسط GAN، از یادگیری-Q استفاده می‌شود تا بهترین مسیر بر اساس معیارهای ارزیابی و پاداش انتخاب شود. در نهایت، عامل یاد می‌گیرد که کدام مسیرها

این منظور، ۱۰۰ ماتریس مسیر با تعداد گره‌های متغیر بین ۲ تا ۳۰ ایجاد می‌شود. هر یک از این مسیرها دارای ویژگی‌هایی هستند که برای حل مسئله ضروری‌اند، از جمله طول مسیر، کیفیت لینک بین ۰ تا ۱، مصرف انرژی در هر گره (با ضربی از ۱۰)، و مقدار پاداش که بین ۰ تا ۱ قرار دارد.

مرحله ۲: بارگذاری مجموعه داده

برای اجرای مدل، مجموعه‌ای از داده‌های مسیرها نیاز است. در این مرحله، داده‌های مربوط به مسیرها در شبکه حسگر بی‌سیم بارگذاری می‌شوند. این مجموعه داده شامل اطلاعاتی مانند مسیر، طول مسیر، کیفیت لینک، میزان مصرف انرژی، و مقدار پاداش هر مسیر است.

مرحله ۳: تعریف معماری شبکه متخاصم زایشی برای تولید مسیر

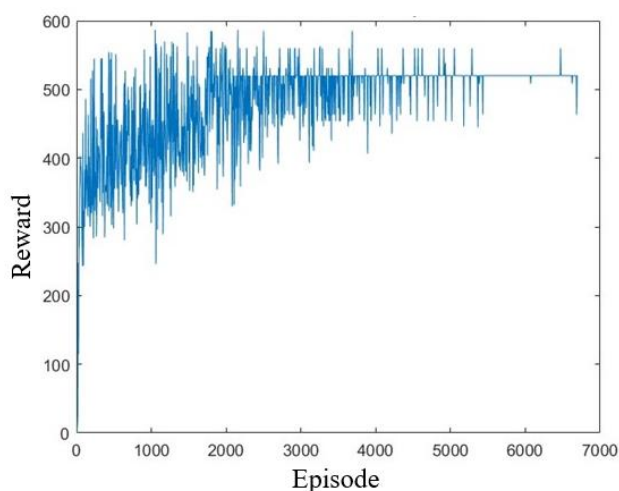
برای ساخت مدل شبکه متخاصم زایشی، باید شبکه‌های تولیدکننده و تشخیص‌دهنده ایجاد شوند. شبکه تولیدکننده دارای سه لایه به ترتیب ۱۲۸، ۲۵۶ و ۳۰ واحد است و هر لایه ۶۴ نرون دارد. شبکه تشخیص‌دهنده نیز شامل سه لایه با ۱۲۸، ۲۵۶ و ۱ واحد است. تعداد نرون‌ها پس از آزمایش‌های مختلف برای دستیابی به بهترین عملکرد تعیین شده‌اند.

در لایه‌های میانی از تابع فعال‌ساز ReLU استفاده شده است که در یادگیری عمیق بسیار رایج است. این تابع مقادیر منفی را صفر و مقادیر مثبت را به همان شکل نگه می‌دارد. در آخرین لایه شبکه تولیدکننده از تابع فعال‌ساز Tanh استفاده شده است؛ زیرا این نوع شبکه‌ها معمولاً خروجی‌های غیرخطی تولید می‌کنند و توابع فعال‌ساز غیرخطی برای این مرحله مناسب‌تر هستند. در شبکه تشخیص‌دهنده، از تابع فعال‌ساز Sigmoid در انتهای مدل استفاده می‌شود تا خروجی به شکل احتمالاتی باشد.

روش پیشنهادی را با یکی از جدیدترین روش‌های موجود (یادگیری تقویتی) مقایسه می‌کنیم.

۴-۲-۱: همگرایی

قبل از هر معیار دیگری، بررسی همگرایی روال یادگیری استفاده شده در رویکرد پیشنهادی در شکل (۶) نشان داده شده است. در این شکل همگرایی بر اساس پاداش تجمعی محاسبه شده است. شایان ذکر است پاداش تجمعی معیاری رایج در یادگیری تقویتی است که مجموع پاداش‌های دریافت‌شده توسط عامل در تمام مراحل را نشان می‌دهد. هر چه مقدار پاداش تجمعی بیشتری حاصل شود، عملکرد الگوریتم بهتر است.



شکل (۶): همگرایی رویکرد پیشنهادی

رویکرد پیشنهادی مبتنی بر GAN، مسیرهای بهینه‌ای را که در الگوریتم یادگیری Q-بیشترین پاداش تجمعی را کسب کرده‌اند، مجدداً به چرخه یادگیری و مسیریابی وارد می‌کند. این فرآیند با حفظ یک هم‌بندی گراف از گره‌های مفید، پایداری مسیرهای بهینه را در مراحل مختلف تضمین می‌کند. با این حال، در صورت کاهش انرژی گره‌های محبوب در مسیرهای پرتدد، الگوریتم به سرعت مسیرهای جایگزین، حتی اگر طولانی‌تر باشند، پیشنهاد می‌دهد. در نهایت، با توجه به اینکه همگرایی این رویکرد بر اساس پاداش دریافتی تنظیم شده، عملکرد مطلوبی از نظر

را برای رسیدن به گره مقصد انتخاب کند و همزمان معیارهای ارزیابی را بهینه‌سازی نماید.

۴-۲: نتایج عددی

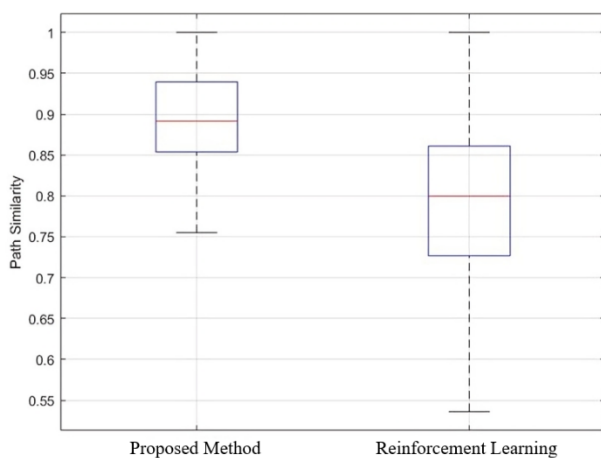
با توجه به اینکه برای آموزش و ارزیابی نتایج به داده‌های آزمایشی نیاز است، از مجموعه داده‌ای تصادفی که شامل ۱۰۰ ماتریس مسیر با تعداد گره‌هایی بین ۲ تا ۳۰ است، استفاده شده است. هر یک از این مسیرها دارای ویژگی‌های مورد نیاز مسئله است که شامل طول مسیر، کیفیت لینک (بین ۰ تا ۱)، مصرف انرژی در هر گره (با ضریب ۱۰)، و مقدار پاداش (بین ۰ تا ۰٫۱) می‌باشد. در طول فرآیند شبیه‌سازی، گره‌های حسگر به صورت تصادفی داده‌هایی برای ارسال به سمت گره سینک تولید می‌کنند. نرخ متوسط تولید داده‌ها در طول فرآیند شبیه‌سازی بین ۱۵۰ کیلو بیت بر ثانیه تا ۵۰۰ کیلو بیت بر ثانیه تنظیم شده است. خلاصه پارامترهای استفاده شده در فرآیند شبیه‌سازی در جدول (۲) آمده است.

جدول (۲): خلاصه پارامترهای شبیه‌سازی

پارامتر	مقدار/توضیحات
تعداد گره‌ها	50
ابعاد محیط شبیه‌سازی	۱۰۰۰ × ۱۰۰۰ متر
تعداد تکرار آزمایش‌ها	۲۵ بار
محدوده کیفیت لینک	بین ۰ تا ۱
مصرف انرژی در هر گره	ضریبی از ۱۰
محدوده پاداش مسیرها	بین ۰ تا ۰٫۱
انرژی اولیه هر گره	۱٫۲۵ ژول
نرخ تولید داده	بین ۱۵۰ کیلو بیت بر ثانیه تا ۵۰۰ کیلو بیت بر ثانیه
عامل کاهش در یادگیری (γ)	0.9
نرخ یادگیری در یادگیری (α)	0.1
ابزار شبیه‌سازی	متلب نسخه ۲۰۲۴ و کتابخانه یادگیری عمیق

برای ارزیابی و بررسی عملکرد روش پیشنهادی، از معیارهای مختلفی استفاده شده است. در ادامه، ضمن معرفی این معیارها،

مختلف است. در مجموع، نتایج به دست آمده نشان می‌دهد که روش پیشنهادی در معیار شباهت مسیر عملکرد بهتری دارد.



شکل (۷): مقایسه دو روش با معیار شباهت مسیر

۴-۲-۳: حفظ توپولوژی گراف

معیار حفظ توپولوژی گراف (Graph Topology Preservation) در عمل میزان حفظ ساختار اولیه گراف را در فرآیند مسیریابی یا شبیه‌سازی ارزیابی می‌کند. هدف از حفظ توپولوژی گراف این است که یک الگوریتم مسیریابی، مسیرهای مناسب را انتخاب کند یا تغییرات مورد نیاز را اعمال کند، بدون اینکه توپولوژی کلی شبکه (مثل تعداد و نوع ارتباطات اصلی بین گره‌ها) تغییرات عمده‌ای پیدا کند. به عبارت دیگر، مسیریابی باید به گونه‌ای انجام شود که شبکه همچنان ساختار اولیه خود را حفظ کند و پایداری و هماهنگی آن به هم نخورد. شکل (۸) دو روش مورد نظر را از نظر حفظ توپولوژی گراف با هم مقایسه می‌کند.

همگرایی نشان می‌دهد. رویکرد پیشنهادی مبتنی بر GAN، مسیرهای بهینه‌ای را که در الگوریتم یادگیری-Q بیشترین پاداش جمعی را کسب کرده‌اند، مجدداً به چرخه یادگیری و مسیریابی وارد می‌کند. این فرآیند با حفظ یک هم‌بندی گراف از گره‌های مفید، پایداری مسیرهای بهینه را در مراحل مختلف تضمین می‌کند. با این حال، در صورت کاهش انرژی گره‌های محبوب در مسیرهای پرتردد، الگوریتم به سرعت مسیرهای جایگزین، حتی اگر طولانی‌تر باشند، پیشنهاد می‌دهد. در نهایت، با توجه به اینکه همگرایی این رویکرد بر اساس پاداش دریافتی تنظیم شده، عملکرد مطلوبی از نظر همگرایی نشان می‌دهد.

۴-۲-۲: شباهت مسیر

معیار ارزیابی شباهت مسیر (Path Similarity) در واقع سطح شباهت بین مسیرهای تولید شده توسط یک الگوریتم و مسیرهای مرجع (یا واقعی) را بررسی می‌کند. از این معیار برای ارزیابی الگوریتم‌های مسیریابی در شبکه‌های حسگر بی‌سیم استفاده می‌شود و نشان می‌دهد که چه میزان مسیرهای پیشنهادی الگوریتم مشابه یکدیگر (و یا مشابه با مسیرهای بهینه) هستند. در اینجا ما مسیرهای استفاده شده توسط دو رویکرد مورد مقایسه را با مسیرهای بهینه موجود در توپولوژی مفروض مقایسه می‌کنیم.

شکل (۷) مقایسه عملکرد روش پیشنهادی و روش یادگیری تقویتی را به کمک نمودار جعبه‌ای بر اساس توزیع داده‌ها، میانه و دامنه مقادیر نمایش می‌دهد. در شکل (۶)، میانگین روش پیشنهادی حدود ۰,۹ با انحراف معیار ۰,۰۵ است که نشان‌دهنده شباهت بالای مسیرها در این روش با مسیرهای بهینه است. در مقابل، روش یادگیری تقویتی با میانگین ۰,۸ و انحراف معیار ۰,۱ عملکرد ضعیف‌تری دارد. از نظر میانگین، روش پیشنهادی شباهت مسیر بیشتری دارد و از نظر پراکندگی، روش یادگیری تقویتی تنوع بیشتری دارد، که نشان از ناپایداری بیشتر آن در شرایط

۴-۲-۴: تعداد بسته‌های تحویل شده به مقصد

این معیار درصد بسته‌های ارسالی که سالم به مقصد تحویل شده‌اند را ارزیابی می‌کند. نتایج حاصل از شبیه‌سازی برای دو رویکرد مورد مقایسه، به ازای نرخ‌های متفاوت ارسال داده در کل شبکه، در جدول (۲) نشان داده شده است.

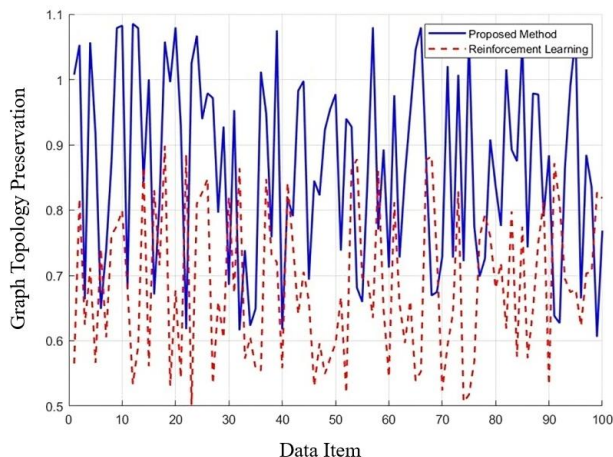
جدول (۲): مقایسه نرخ تحویل بسته‌ها به مقصد

نرخ ترافیک ورودی	نرخ تحویل بسته‌ها به مقصد	
	رویکرد پیشنهادی	یادگیری تقویتی [۲۱]
100 Kbps	98.1%	96.3%
150 Kbps	96.5%	95.0%
200 Kbps	95.1%	93.8%
250 Kbps	93.5%	92.0%
300 Kbps	92.2%	90.3%

همانطور که در جدول (۲) نمایان است، رویکرد پیشنهادی به دلیل استفاده از مسیرهای مناسب‌تر، نرخ تحویل بالاتری نسبت به رویکرد مقابل را به ارمغان می‌آورد. روش پیشنهادی با حفظ مصرف انرژی در سطح مشابه، تعداد بیشتری از بسته‌ها را به مقصد رسانده است. این به دلیل انتخاب مسیرهای طولانی‌تر برای جریان‌های مختلف است که با توزیع مناسب بار ترافیکی میان گره‌ها از تمرکز بیش از حد مصرف انرژی در گره‌های خاص جلوگیری کرده است. این توزیع متوازن انرژی در سراسر شبکه موجب کاهش فشار بر گره‌های خاص و در نتیجه بهبود نرخ تحویل بسته‌ها شده است.

۴-۲-۵: تأخیر انتها به انتها

میانگین تأخیر انتها به انتها یکی از مهمترین معیارهایی است که می‌تواند دو رویکرد مسیریابی مورد نظر را با یکدیگر مقایسه کرد. جدول (۴) نتایج به دست آمده از شبیه‌سازی را نشان می‌دهد. همانطور که در این جدول مشاهده می‌شود، میانگین تأخیر انتها به انتها برای بسته‌های ارسال شده در رویکرد پیشنهادی به دلیل استفاده از مسیرهای مناسب‌تر، از رویکرد مقابل کمتر است.



شکل (۸): مقایسه دو روش با معیار حفظ توپولوژی گراف

همانطور که در شکل (۸) نمایان است، عملکرد روش پیشنهادی با خط آبی نشان می‌دهند که روش پیشنهادی به طور کلی عملکرد بهتری در حفظ توپولوژی گراف دارد. شایان ذکر است با گذر زمان و افزایش مصرف انرژی در گره‌های مختلف، تعدادی از گره‌های شبکه به تدریج با اتمام ذخایر انرژی خود از کار خواهند افتاد و خاموش می‌شوند. این پدیده نه تنها به کاهش تعداد گره‌های فعال در شبکه منجر می‌شود، بلکه تغییرات قابل توجهی را در توپولوژی گراف شبکه ایجاد می‌کند. حذف گره‌ها از ساختار شبکه می‌تواند موجب قطع ارتباط بین بخش‌هایی از شبکه شود و دسترسی به برخی مسیرهای موجود را غیرممکن سازد.

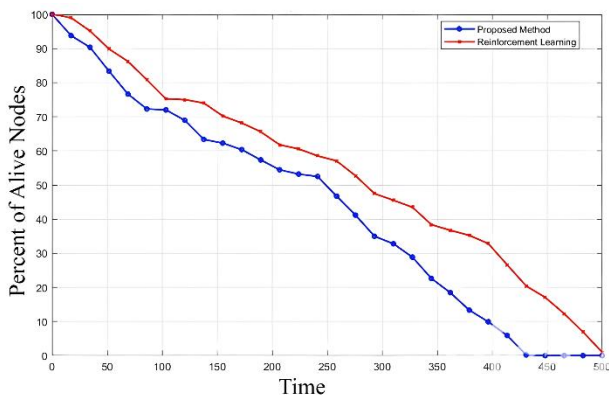
این تغییرات ممکن است به شکل محلی یا گسترده بر عملکرد شبکه تأثیر بگذارند، به‌ویژه اگر گره‌های کلیدی یا گره‌هایی که در مسیرهای پرتردد قرار دارند خاموش شوند. در چنین شرایطی، مسیریابی داده‌ها با چالش بیشتری مواجه می‌شود و الگوریتم‌های شبکه باید با بهره‌گیری از مکانیزم‌های تطبیقی به سرعت خود را با توپولوژی جدید سازگار کنند. علاوه بر این، کاهش تعداد گره‌های فعال می‌تواند تأثیر مستقیمی بر بهره‌وری کلی شبکه، از جمله نرخ تحویل بسته‌ها و افزایش تأخیر انتها به انتها داشته باشد.

جدول (۴): مقایسه میانگین تأخیر انتها به انتها

نرخ ترافیک ورودی	میانگین تأخیر انتها به انتها	
	رویکرد پیشنهادی	یادگیری تقویتی [۲۱]
100 Kbps	1.25 s	1.58 s
150 Kbps	1.37 s	1.71 s
200 Kbps	1.55 s	1.96 s
250 Kbps	1.84 s	2.34 s
300 Kbps	2.21 s	2.79 s

۴-۲-۶: طول عمر شبکه

برای مقایسه طول عمر شبکه و میزان مصرف انرژی در دو رویکرد مدنظر، تعداد گره‌های زنده مانده در برابر تعداد گره‌های خاموش شده در طول شبیه‌سازی به دقت اندازه‌گیری شدند. گره‌های شبکه انرژی اولیه خود را برای انجام محاسبات و ارتباطات خود مصرف می‌کنند. شکل (۹) تعداد گره‌های زنده در طول مدت شبیه‌سازی را برای دو رویکرد نشان می‌دهد.



شکل (۹): مقایسه درصد گره‌های زنده در طول شبیه‌سازی

همانطور که در شکل (۹) نشان داده شده است، در رویکرد یادگیری تقویتی همواره تعداد گره‌های زنده در شبکه قدری بیشتر از رویکرد پیشنهادی است و عملاً شبکه طول عمر بالاتری دارد. این امر نشان از کمتر بودن مصرف انرژی در گره‌های شبکه در رویکرد یادگیری تقویتی نسبت به رویکرد پیشنهادی دارد. در رویکرد پیشنهادی به دلیل محاسبات بالاتر در طول روال یادگیری، گره‌ها قدری انرژی بیشتری مصرف می‌کنند. در کارهای آتی ما قصد داریم تمرکز بیشتری بر روی کاهش مصرف انرژی در رویکرد پیشنهادی قرار دهیم.

۵. نتیجه‌گیری

در این مقاله، با در نظر گرفتن معیارهای مهمی مانند انرژی باقی‌مانده حسگرها و کاهش تأخیر، و با توجه به رفتار عامل‌ها در پژوهش‌های مختلف مبتنی بر یادگیری عمیق، تعداد زیادی مسیر بهینه برای مسیریابی داده در شبکه‌های حسگر بی‌سیم زیرآبی

به طور شهودی می‌توان گفت دلیل کاهش تأخیر انتها به انتها در رویکرد پیشنهادی در مقایسه با رویکرد رقیب، بهره‌گیری از مسیرهای بهینه‌تر و استفاده از یادگیری مبتنی بر GAN است. این رویکرد با ترکیب قدرت یادگیری تقویتی و توانایی شبکه‌های مولد تخصصی در مدل‌سازی پیچیدگی‌های محیطی، امکان انتخاب مسیرهایی را فراهم می‌کند که نه تنها کوتاه‌تر هستند، بلکه با در نظر گرفتن شرایط واقعی شبکه، مانند تراکم گره‌ها، مصرف انرژی و تأخیر انتقال داده، کارآمدتر نیز عمل می‌کنند.

GAN در این فرآیند به شبیه‌سازی و پیش‌بینی بهتر رفتار شبکه کمک می‌کند و با تولید داده‌های مصنوعی نزدیک به واقعیت، الگوریتم یادگیری تقویتی را به سمت انتخاب تصمیمات دقیق‌تر هدایت می‌کند. این ترکیب هوشمندانه از یادگیری تقویتی و GAN، منجر به کاهش قابل توجه تأخیر انتها به انتها می‌شود، زیرا مسیرهای انتخاب‌شده به گونه‌ای طراحی می‌شوند که کمترین مانع و بیشترین بازدهی را در انتقال داده‌ها داشته باشند. علاوه بر این، رویکرد پیشنهادی توانایی انطباق با تغییرات شبکه را نیز داراست؛ به این معنا که در صورت تغییر شرایط مانند خرابی گره‌ها یا افزایش بار ترافیکی، به سرعت می‌تواند مسیرهای جدید و بهینه‌تری را پیشنهاد کند. این ویژگی، عملکرد پایدارتر و کاهش تأخیر بیشتری را نسبت به روش‌های سنتی تضمین می‌کند.

[3] F. Campagnaro, F. Steinmetz, B. C. Renner, "Survey on low-cost underwater sensor networks: From niche applications to everyday use," *Journal of marine science and engineering*, vol. 11, no. 1, art. no. 125, 2023. Doi: 10.3390/jmse11010125.

[4] H. Luo, J. Wang, F. Bu, R. Ruby, K. Wu and Z. Guo, "Recent Progress of Air/Water Cross-Boundary Communications for Underwater Sensor Networks: A Review," *IEEE Sensors Journal*, vol. 22, no. 9, pp. 8360-8382, 2022. Doi: 10.1109/JSEN.2022.3162600.

[5] S. Gupta S, N. P. Singh, "Underwater wireless sensor networks: a review of routing protocols, taxonomy, and future directions," *The Journal of Supercomputing*, vol. 80, no. 4, art. no. 5163-5196, 2024. Doi: 10.1007/s11227-023-05646-w.

[6] S. Khisa and S. Moh, "Survey on Recent Advancements in Energy-Efficient Routing Protocols for Underwater Wireless Sensor Networks," *IEEE Access*, vol. 9, pp. 55045-55062, 2021. Doi: 10.1109/ACCESS.2021.3071490.

[7] Z. Liu, X. Jin, Y. Yang, K. Ma and X. Guan, "Energy-Efficient Guiding-Network-Based Routing for Underwater Wireless Sensor Networks," *IEEE Internet of Things Journal*, vol. 9, no. 21, pp. 21702-21711, 2022. Doi: 10.1109/JIOT.2022.3183128.

[8] H. Khan, S. A. Hassan and H. Jung, "On Underwater Wireless Sensor Networks Routing Protocols: A Review," *IEEE Sensors Journal*, vol. 20, no. 18, pp. 10371-10386, 2020. Doi: 10.1109/JSEN.2020.2994199.

[9] J. Gui, Z. Sun, Y. Wen, D. Tao and J. Ye, "A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 4, pp. 3313-3332, 2023. Doi: 10.1109/TKDE.2021.3130191.

[10] O. L. Striuk, Y. U. Kondratenko, "Generative adversarial neural networks and deep learning: successful cases and advanced approaches," *International Journal of Computing*, vol. 20, no. 3, pp. 339-49, 2021. Doi: 10.47839/ijc.20.3.2278.

[11] K. K. Gola, S. Arya, "Underwater acoustic sensor networks: Taxonomy on applications, architectures, localization methods, deployment techniques, routing techniques, and threats: A systematic review," *Concurrency and Computation: Practice and Experience*, vol. 35, no. 23, art. no. e7815, 2023. Doi: 10.1002/cpe.7815.

[12] J. Luo, Y. Chen, M. Wu and Y. Yang, "A Survey of Routing Protocols for Underwater Wireless Sensor Networks," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 137-160, 2021. Doi: 10.1109/COMST.2020.3048190.

[13] R. W. L. Coutinho, A. Boukerche, L. F. M. Vieira and A. A. F. Loureiro, "Geographic and Opportunistic Routing for Underwater Sensor Networks," *IEEE Transactions on Computers*, vol. 65, no. 2, pp. 548-561, 1 Feb. 2016. Doi: 10.1109/TC.2015.2423677.

[14] G. Tuna, "Clustering-based energy-efficient routing approach for underwater wireless sensor networks,"

شناسایی و ذخیره شد. سپس، این مسیرها به منظور آموزش شبکه عمیق متخاصم زایشی به کار گرفته شدند. فرآیند آموزش به صورت آفلاین و در دستگاه‌هایی با منابع پردازشی کافی انجام گرفت و مدل به دست آمده برای مسیریابی داده‌ها در حسگرهای بی سیم زیرآبی مورد استفاده قرار گرفت. در مرحله انتقال داده، مسیر بهینه بر اساس وضعیت پویای گره حسگر فعلی تولید شد، که این امر منجر به ایجاد الگوریتمی پویا گردید که قادر به تصمیم‌گیری بر اساس شرایط جاری هر گره است. هدف اصلی این الگوریتم، یافتن مسیرهای بهینه با مصرف انرژی بهینه برای حسگرها بود.

علاوه بر این، از شبکه GAN برای تولید داده‌های مصنوعی و ترکیب آن‌ها با داده‌های واقعی استفاده شد. این رویکرد موجب افزایش حجم داده‌های آموزشی و بهبود عملکرد الگوریتم‌های یادگیری ماشین و پروتکل‌های مسیریابی در شبکه‌های حسگر بی سیم زیرآبی شد، که به دلیل محدودیت تعداد گره‌ها و داده‌ها، نیازمند بهینه‌سازی بودند. در نهایت، روش پیشنهادی با روش مبتنی بر یادگیری تقویتی مقایسه گردید. نتایج شبیه‌سازی نشان داد که روش پیشنهادی از نظر معیارهایی همچون شباهت مسیرها، حفظ توپولوژی شبکه، پاداش تجمعی، نرخ بسته‌های تحویلی، و میزان تاخیر انتها به انتها عملکرد بهتری نسبت به روش یادگیری تقویتی داشته است و در مقابل قدری مصرف انرژی در گره‌های شبکه را افزایش می دهد.

مراجع

[1] ا. توکلی و ع. کشاورزحداد، "مروری بر ویژگی‌های فناوری‌های ارتباطی در خانه‌های هوشمند و چالش‌های پیش رو،" *محاسبات نرم*، ۱۴۰۲. Doi: scj.2023.242846.0/10.22052

[2] س. اشکذری طوسی، م. ن. تیموری، و. سبزواری، "مکان‌یابی اهداف متحرک در شبکه حسگر بیسیم مبتنی بر الگوریتم انتشار حداقل میانگین مربعات بر پایه تابع زیان هوپر،" *محاسبات نرم*، ۱۴۰۲. Doi: 10.22052/scj.2023.252719.1141.

- [26] Y. Yuan, M. Liu, X. Zhuo, Y. Wei, X. Tu and F. Qu, "A Q-Learning-Based Hierarchical Routing Protocol with Unequal Clustering for Underwater Acoustic Sensor Networks," *IEEE Sensors Journal*, vol. 23, no. 6, pp. 6312-6325, 2023. Doi: 10.1109/JSEN.2022.3232614.
- [27] C. Wang, X. Shen, H. Wang, W. Xie, H. Zhang and H. Mei, "Multi-Agent Reinforcement Learning-Based Routing Protocol for Underwater Wireless Sensor Networks with Value of Information," *IEEE Sensors Journal*, vol. 24, no. 5, pp. 7042-7054, 2024. Doi: 10.1109/ACCESS.2021.3128516.
- [15] M. U. Khan, P. Otero and M. Aamir, "An Energy Efficient Clustering Routing Protocol Based on Arithmetic Progression for Underwater Acoustic Sensor Networks," *IEEE Sensors Journal*, vol. 24, no. 5, pp. 6964-6975, 2024. Doi: 10.1109/JSEN.2024.3354252.
- [16] Y. Sun, M. Zheng, X. Han, S. Li, J. Yin, "Adaptive clustering routing protocol for underwater sensor networks," *Ad Hoc Networks*, vol. 136, art. no. 102953, 2022. Doi: 10.1016/j.adhoc.2022.102953.
- [17] R. Zhu, A. Boukerche, Y. Chen, Q. Yang, "A reliable cluster-based opportunistic routing protocol for underwater wireless sensor networks," *Computer Networks*, vol. 251, art. no. 110622, 2024. Doi: 10.1016/j.comnet.2024.110622.
- [18] W. Zhu, X. Yang, T. Wu and Y. Qiu, "A Routing Algorithm for Underwater Acoustic-Optical Hybrid Wireless Sensor Networks Based on Intelligent Ant Colony Optimization and Energy-Flexible Global Optimal Path Selection," *IEEE Sensors Journal*, vol. 24, no. 10, pp. 17116-17126, 2024. Doi: 10.1109/JSEN.2024.3386892.
- [19] X. Xiao, H. Huang, W. Wang, "Underwater wireless sensor networks: An energy-efficient clustering routing protocol based on data fusion and genetic algorithms," *Applied Sciences*, vol. 11, no. 1, art. no. 312, 2020. Doi: 10.3390/app11010312.
- [20] D. Han, et. al. "Trust-Aware and Fuzzy Logic-Based Reliable Layering Routing Protocol for Underwater Acoustic Networks," *Sensors*, vol. 23, no. 23, art. no. 9323, 2023. Doi: 10.3390/s23239323.
- [21] A. Khasawneh, M. S. B. A. Latiff, O. Kaiwartya and H. Chizari, "A reliable energy-efficient pressure-based routing protocol for underwater wireless sensor network," *Wireless Networks*, vol. 24, pp. 2061-2075, 2018. Doi: 10.1007/s11276-017-1461-x.
- [22] L. Als Salman and E. Alotaibi, "A Balanced Routing Protocol Based on Machine Learning for Underwater Sensor Networks," *IEEE Access*, vol. 9, pp. 152082-152097, 2021. Doi: 10.1109/ACCESS.2021.3126107.
- [23] Y. Su, R. Fan, X. Fu and Z. Jin, "DQELR: An Adaptive Deep Q-Network-Based Energy- and Latency-Aware Routing Protocol Design for Underwater Acoustic Sensor Networks," *IEEE Access*, vol. 7, pp. 9091-9104, 2019. Doi: 10.1109/ACCESS.2019.2891590.
- [24] X. Zhu, et. al, "Dynamic layered routing protocols based on BP-NN for underwater acoustic sensor networks," *Applied Acoustics*, vol. 211, art. no. 109454, 2023. Doi: 10.1016/j.apacoust.2023.109454.
- [25] C. Wang, X. Shen, H. Wang, H. Zhang and H. Mei, "Reinforcement Learning-Based Opportunistic Routing Protocol Using Depth Information for Energy-Efficient Underwater Wireless Sensor Networks," *IEEE Sensors Journal*, vol. 23, no. 15, pp. 17771-17783, 1 Aug.1, 2023. Doi: 10.1109/JSEN.2023.3285751.