Journal of Soft Computing and Information Technology (JSCIT)
Babol Noshirvani University of Technology, Babol, Iran
Journal Homepage: jscit.nit.ac.ir
DOI: 10.22034/jscit.2025.510640.2111
Volume 14, Number 2, Summer 2025, pp. 60-82
Received:03/04/2025, Revised:05/18/2025, Accepted: 07/14/2025

A Trust-Aware Reinforcement Learning Approach for Secure IoT Routing

Fatemeh Vafaei Nejad[1], Mohammad Mehdi Gilanian Sadeghi [2*] and Mohammad Hossein Rezvani [3]
1- Department of Computer and Information Technology Engineering, Qa.C., Islamic Azad University, Qazvin, Iran.
2*- Department of Computer and Information Technology Engineering, Qa.C., Islamic Azad University, Qazvin, Iran.
3- Department of Computer and Information Technology Engineering, Qa.C., Islamic Azad University, Qazvin, Iran.
[1] f.vafaeinejad@iau.ac.ir, [2*] gilaniansadeghi@iau.ac.ir, and [3] mh.rezvani@iau.ac.ir

Corresponding author's address: Mohammad Mehdi Gilanian Sadeghi, Department of Computer and Information Technology Engineering, Qa.C., Islamic Azad University, Qazvin, Iran.

Abstract- The Internet of Things (IoT) comprises devices with limited processing power and energy, operating within low-power and lossy networks (LLNs). These networks are susceptible to various vulnerabilities due to constrained resources, weak physical security, dynamic network topologies, and unstable communication links. Developing security solutions for such networks is challenging because of these limitations. Moreover, the Routing Protocol for Low-Power and Lossy Networks (RPL), recognized as the standard routing protocol in IoT, offers minimal defense against various routing attacks. Attackers can exploit vulnerabilities in this protocol to launch malicious attacks against IoT networks. Therefore, considering the limited resources of the nodes, in this paper, a Trust-Aware Reinforcement Learning-based method is proposed for secure routing in the Internet of Things. This approach combines the reliability of trust-based methods with the high accuracy of machine learning techniques to assess trust levels at both node and link levels. The primary goal of this approach is to effectively identify and isolate malicious nodes, providing secure and trust-based routing in IoT networks. Simulation results against decreased rank attack demonstrate that the proposed approach outperforms existing methods. Specifically, it improves the packet delivery ratio (PDR) by 17.46% over MRHOF, 3.2% over SecTrust-RPL, and 1.69% over SMTrust. Additionally, it enhances throughput and achieves a higher attack detection rate, while introducing only a negligible increase in control overhead and energy consumption. Therefore, the proposed approach is well-suited for real-time applications, as it minimizes packet loss by routing through trusted nodes and effectively isolating malicious ones.

Keywords- Internet of Things - IoT, Secure Routing, RPL Protocol, Rank Attack, Reinforcement Learning, routing, Trust System, Low-power and Lossy Networks - LLNs.

مجله علمی رایانش نرم و فناوری اطلاعات
دانشگاه صنعتی نوشیروانی بابل
صفحه مجله: jscit.nit.ac.ir
شناسه دیجیتال: 10.22034/jscit.2025.510640.2111
جلد ۱۴، شماره ۲، تابستان ۱۴۰۴، صفحه ۶۰-۸۲
دریافت: ۱۴۰۳/۱۲/۱۴، بازنگری: ۱۴۰۴/۰۲/۲۸، پذیرش: ۱۴۰۴/۰۴/۲۳

رویکرد یادگیری تقویتی آگاه از اعتماد برای مسیریابی امن اینترنت اشیا

فاطمه وفائی نژاد[۱]، محمدمهدی گیلانیان صادقی*[۲]، محمدحسین رضوانی[۳]
۱- گروه مهندسی کامپیوتر و فناوری اطلاعات، واحد قزوین، دانشگاه آزاد اسلامی، قزوین، ایران.
*۲- گروه مهندسی کامپیوتر و فناوری اطلاعات، واحد قزوین، دانشگاه آزاد اسلامی، قزوین، ایران.
۳- گروه مهندسی کامپیوتر و فناوری اطلاعات، واحد قزوین، دانشگاه آزاد اسلامی، قزوین، ایران.
[1]f.vafaeinejad@iau.ac.ir, [2*] gilaniansadeghi@iau.ac.ir, [3] mh.rezvani@iau.ac.ir

* نشانی نویسنده مسئول: محمدمهدی گیلانیان صادقی، قزوین، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه آزاد اسلامی، واحد قزوین.

چکیده- اینترنت اشیا شامل دستگاه‌هایی با توان پردازشی و انرژی محدود است که در شبکه‌های کم‌توان و پر اتلاف فعالیت می‌کنند. محدودیت منابع، امنیت فیزیکی ضعیف، پویایی توپولوژی شبکه و ناپایداری پیوندهای ارتباطی، این شبکه‌ها را در برابر انواع آسیب‌پذیری‌ها قرار می‌دهد. توسعه راهکارهای امنیتی برای این شبکه‌ها به دلیل این محدودیت‌ها چالش‌برانگیز است. علاوه بر این، پروتکل مسیریابی شبکه‌های کم‌توان و پر اتلاف (RPL) که به‌عنوان استاندارد مسیریابی در اینترنت اشیا شناخته‌شده است، در برابر حملات مسیریابی مختلف، حداقل سطح دفاعی را ارائه می‌دهد. مهاجمان می‌توانند با سوءاستفاده از آسیب‌پذیری‌های این پروتکل، حملات مخربی را علیه شبکه‌های اینترنت اشیا اجرا کنند. در این مقاله، یک روش مبتنی بر یادگیری تقویتی آگاه از اعتماد برای مسیریابی امن در اینترنت اشیا پیشنهاد شده است. این روش با ترکیب قابلیت اطمینان روش‌های مبتنی بر اعتماد و دقت بالای روش‌های مبتنی بر یادگیری ماشین، به ارزیابی سطح اعتماد در سطح گره و پیوند می‌پردازد. هدف اصلی این رویکرد، شناسایی و جداسازی مؤثر گره‌های مخرب و فراهم‌سازی مسیریابی ایمن و مبتنی بر اعتماد در شبکه‌های اینترنت است. نتایج شبیه‌سازی در برابر حمله کاهش رتبه نشان می‌دهد که روش پیشنهادی عملکرد بهتری نسبت به روش‌های موجود دارد. این روش نرخ تحویل بسته (PDR) را نسبت به MRHOF به میزان ۱۷/۴۶ درصد، نسبت به SecTrust-RPL به میزان ۳/۲ درصد و نسبت به SMTrust به میزان ۱/۶۹ درصد بهبود می‌بخشد. علاوه بر این، توان عملیاتی شبکه را افزایش داده و نرخ شناسایی حملات را بالا می‌برد، درحالی‌که افزایش سربار کنترلی و مصرف انرژی آن ناچیز است؛ بنابراین، این روش برای کاربردهای بلادرنگ مناسب بوده و با مسیریابی از طریق گره‌های مورد اعتماد و حذف گره‌های مخرب، نرخ از دست رفتن بسته‌ها را کاهش می‌دهد.

واژه‌های کلیدی: اینترنت اشیا؛ مسیریابی امن؛ پروتکل مسیریابی RPL؛ حمله کاهش رتبه؛ یادگیری تقویتی؛ سیستم اعتماد؛ شبکه‌های کم‌توان و پر اتلاف.

۱- مقدمه

اینترنت اشیا شامل مجموعه‌ای از دستگاه‌های هوشمند متصل به شبکه است که امکان تبادل و پردازش داده‌ها را فراهم می‌کند. با افزایش تعداد این دستگاه‌ها و گسترش کاربردهای آن، اهمیت این فناوری به‌طور چشمگیری افزایش‌یافته است [۱]. بااین‌حال، چالش‌های امنیتی و حریم خصوصی از موانع اصلی پذیرش گسترده آن محسوب می‌شوند در بسیاری از حوزه‌های حیاتی مانند مراقبت‌های بهداشتی، تأمین امنیت داده‌ها و شبکه برای کاربران و ارائه‌دهندگان خدمات ضروری است [۲]. هم‌زمان با رشد اینترنت اشیا، حملات سایبری نیز افزایش‌یافته و امنیت این شبکه‌ها را به چالش کشیده است؛ بنابراین، برای پذیرش گسترده اینترنت اشیا،

باید اصول امنیتی ازجمله محرمانگی، یکپارچگی، در دسترس‌پذیری، احراز هویت و انکارناپذیری رعایت شوند [۳].

یکی از مهم‌ترین چالش‌های امنیتی در بستر اینترنت اشیا، ضعف پروتکل‌های مسیریابی در برابر تهدیدات داخلی و خارجی است. پروتکل مسیریابی RPL که به‌عنوان استاندارد مسیریابی در شبکه‌های کم‌توان و پر اتلاف شناخته می‌شود، باوجود سه حالت امنیتی خود، همچنان در برابر طیف وسیعی از حملات آسیب‌پذیر باقی‌مانده است [۴]. این مسئله به‌ویژه در کاربردهای حیاتی مانند سلامت و انرژی که نیازمند قابلیت اطمینان بالا هستند، یک چالش امنیتی جدی تلقی می‌شود. مطالعات نشان داده‌اند که حتی در حالت امن نیز RPL در برابر حملات مختلف مانند کاهش رتبه، سیبیل، سیاه‌چاله، کرم‌چاله و ارسال انتخابی مقاوم نیست. ضعف‌های ذاتیRPL ، ازجمله حملات مبتنی بر مقدار رتبه و تغییر شماره نسخه، این پروتکل را به هدفی جذاب برای مهاجمان تبدیل کرده است [۵]. همچنین، راهکارهای امنیتی ارائه‌شده توسط گروه کاری[۱] ROLLدر کارگروه مهندسی اینترنت[۲]IETF، بیشتر بر نیازهای کلی امنیتی تمرکز دارند و روش‌های مشخصی برای مقابله با این تهدیدات ارائه نمی‌دهند [۳]. از سوی دیگر، استفاده از روش‌های سنتی مانند رمزنگاری به دلیل مصرف بالای انرژی و سربار پردازشی بالا، برای شبکه‌های اینترنت اشیا که دارای منابع محدود هستند، کارآمد نیست [۶]؛ بنابراین، نیاز به یک راهکار نوین و کارآمد برای پایش رفتار گره‌ها، شناسایی گره‌های مخرب و امن سازی مسیرها بیش‌ازپیش احساس می‌شود.

برای حل این چالش، یک روش مسیریابی امن مبتنی بر یادگیری تقویتی و آگاه از اعتماد پیشنهادشده است که امنیت و کارایی پروتکل RPL را در برابر حمله کاهش رتبه بهبود می‌بخشد. در این رویکرد، گره‌های اینترنت اشیا به‌عنوان عامل‌های یادگیرنده در نظر گرفته می‌شوند که از محیط پیرامون خود اطلاعات را دریافت کرده و بر اساس راهکار یادگیری Q، تصمیمات بهینه‌ای برای مسیریابی اتخاذ می‌کنند. این روش با ترکیب سطح اعتماد گره‌ها، میزان اعتماد پیوندها و انرژی باقی‌مانده آن‌ها، فرآیند مسیریابی را بهبود می‌بخشد و به کاهش نرخ از دست رفتن بسته‌ها، افزایش دقت در شناسایی حملات و کاهش سربار کنترلی کمک می‌کند. علاوه بر این، برخلاف روش‌های پیشین مانند SecTrust-RPL [7] که تنها بر سطح اعتماد گره‌ها تکیه‌دارند، روش پیشنهادی ما معیارهای بیشتری را در فرآیند تصمیم‌گیری لحاظ می‌کند.

نوآوری‌های اصلی این مقاله به شرح ذیل است.

- استفاده از یادگیری تقویتی در مسیریابی :گره‌های اینترنت اشیا به‌عنوان عامل‌های یادگیرنده عمل کرده و از پایش محیط پیرامون خود برای تصمیم‌گیری‌های بهتر در مسیریابی داده‌ها بهره می‌برند.
- ترکیب معیارهای اعتماد در سطوح گره و پیوند: رویکرد پیشنهادی علاوه بر در نظر گرفتن اعتماد سطح گره در هدایت داده‌ها، از اعتماد سطح پیوند (ETX[۳]) و انرژی باقی‌مانده نیز بهره می‌برد.
- بهره‌گیری از راهکار تصمیم‌گیری توزیع‌شده: هر گره به‌صورت مستقل میزان اعتماد همسایگان خود را محاسبه می‌کند که منجر به کاهش سربار و حذف نقطه شکست مرکزی در شبکه می‌شود.

ادامه این مقاله به شرح ذیل سازمان‌دهی شده است.

بخش ۲ به بررسی پیش‌زمینه‌ای از ساختار پروتکل RPL و بررسی حمله کاهش رتبه پرداخته و مفاهیم پایه‌ای یادگیری تقویتی را موردبررسی قرار می‌دهد. بخش ۳ به بررسی پژوهش‌های مرتبط در زمینه روش‌های مبتنی بر اعتماد در شبکه‌های اینترنت اشیا مبتنی بر پروتکل مسیریابی RPL می‌پردازد. بخش ۴ جزئیات روش پیشنهادی را شرح می‌دهد. بخش ۵ به ارزیابی عملکرد روش پیشنهادی اختصاص دارد. در بخش ۶ به آنالیز و تجزیه‌وتحلیل آماری پرداخته می‌شود و درنهایت بخش ۷ مقاله را با یک جمع‌بندی و پیشنهادهایی برای پژوهش‌های آینده به پایان می‌رساند.

۲- پیش‌زمینه

در این بخش معرفی ساختار پروتکل مسیریابی RPL، نحوه اجرای حمله کاهش رتبه و مفاهیم پایه‌ای یادگیری تقویتی، موردبررسی قرار می‌گیرد.

۲-۱- مروری بر ساختار پروتکل مسیریابی RPL

پروتکل مسیریابی RPL از روش مسیریابی بردار فاصله استفاده کرده و با ساختار گراف جهت‌دار غیر مدور مقصد گرا که به‌اختصار[۴]DODAG نامیده می‌شود کار می‌کند. هر DODAG شامل گره‌ها و یک ریشه است که ارتباط با اینترنت را فراهم می‌کند. این ساختار از شناسه‌های منحصربه‌فردی مانند شناسه نمونه RPL، شناسه DODAG و شماره نسخه تشکیل‌شده است. ریشه، شکل‌گیری شبکه را با ارسال دوره‌ای پیام‌های DIO آغاز می‌کند [۸].

RPL مدیریت DODAG را از طریق چهار نوع پیام شامل پیام DIO،

جهت انتشار اطلاعات مسیر، پیام DAO، برای اعلام مقصد و برقراری ارتباط باریشه، پیام DIS برای کشف مسیر توسط گره‌های جدید و پیام DAO-ACK برای تأیید دریافت پیام DAO توسط ریشه انجام می‌دهد. پیام‌های DIO شامل اطلاعاتی مانند مقدار رتبه، شناسه نمونهRPL ، شماره نسخه و شناسه DODAG هستند. مقدار رتبه نشان‌دهنده موقعیت گره در DODAG است؛ هرچه گره به ریشه نزدیک‌تر باشد، مقدار آن کمتر است. گره‌ها با ارسال پیام‌های DAO مقصد خود را مشخص کرده و ریشه نیز باDAO-ACK به آن پاسخ می‌دهد. پیام‌های DIS توسط گره‌های جدید برای اعلام درخواست پیوستن به شبکه ارسال می‌شوند [۹].

انتخاب گره والد در DODAG بر اساس تابع هدف انجام می‌شود . IETF دو تابع استاندارد را برای این منظور معرفی کرده است [10]OF0[۵]:که بر اساس تعداد گام، کوتاه‌ترین مسیر به ریشه را انتخاب می‌کند و[۶] [11]MRHOFکه از معیار ETX برای انتخاب مسیر با کمترین هزینه انتقال استفاده می‌کند. این روش از تغییرات غیرضروری در مسیرهای شبکه جلوگیری می‌کند[۱۱].

۲-۲- حمله کاهش رتبه در RPL

حملات رتبه در شبکه‌های RPL شامل دست‌کاری پیام‌های کنترلی یا تغییر موقعیت جعلی گره‌ها به‌عنوان والد ایده آل به‌منظور تغییر توپولوژی DODAG است. مهاجمان خود را به‌عنوان والد بهینه معرفی کرده و بسته‌های داده را دور انداخته و یا به مسیرهای ناکارآمد هدایت می‌کنند. هدف اصلی این حملات، تغییر ساختار DODAG از طریق حملات افزایش و کاهش رتبه است.

در حمله کاهش رتبه، مهاجمان مقدار رتبه خود را کمتر از مقدار واقعی اعلام می‌کنند تا همسایگانشان آن‌ها را به‌عنوان والد انتخاب کنند [۱۲, ۱۳]. در این شرایط، تعداد زیادی از گره‌های سالم شبکه از طریق مهاجم به DODAG متصل می‌شوند درنتیجه ترافیک بسیار زیادی به گره مهاجم منتقل می‌شود. پس‌ازآن، مهاجم می‌تواند به‌عنوان یک گره فرو چاله عمل کرده و بسته‌های دریافتی از گره‌های زیرمجموعه خود را دست‌کاری یا حذف کند [۱۴].

۲-۳- مفاهیم پایه‌ای یادگیری تقویتی(RL[۷])

یادگیری تقویتی یکی از شاخه‌های مهم یادگیری ماشین است که بر مبنای تعامل مستمر یک عامل با محیط و یادگیری از طریق بازخورد (پاداش یا تنبیه) طراحی‌شده است[۱۵]. در این چارچوب، عامل باهدف بیشینه‌سازی پاداش تجمعی، به‌صورت مکرر اقداماتی را انجام می‌دهد، نتایج آن را مشاهده نموده و بر اساس آن سیاست بهینه‌ای برای تصمیم‌گیری در وضعیت‌های مختلف محیط می‌آموزد.

در مدل استاندارد یادگیری تقویتی، عامل به‌طور هم‌زمان نقش یادگیرنده و تصمیم‌گیرنده را ایفا می‌کند و فرآیند یادگیری از طریق آزمون‌وخطا و بر پایه تعاملات مستقیم با محیط صورت می‌گیرد. این تعامل شامل مشاهده وضعیت فعلی، انتخاب اقدام مناسب، دریافت پاداش و به‌روزرسانی سیاست تصمیم‌گیری عامل است.

در این مدل، پاداش‌ها مشخص می‌کنند که یک اقدام برای عامل مفید یا مضر است و محیط در هر گام زمانی با اختصاص پاداش، بازخورد لازم را ارائه می‌دهد. تابع ارزش میزان پاداش آینده را تخمین می‌زند و تأثیر بلندمدت اقدامات را نشان می‌دهد، درحالی‌که پاداش‌ها نشان‌دهنده نتایج فوری هر اقدام هستند.

چارچوب RL معمولاً به‌صورت یک مجموعه چهارتایی (S, A, P, R) نمایش داده می‌شود که در آن S مجموعه‌ای از وضعیت‌ها، A مجموعه‌ای از اقدامات ممکن، P احتمال تغییر وضعیت‌ها و R تابع پاداش است.

عامل در فواصل زمانی گسسته (مانند گام‌های زمانی یا اپیزودها) با محیط تعامل دارد [۱۶]. در هر گام زمانی، عامل وضعیت فعلی خود را مشاهده کرده، یک اقدام انتخاب می‌کند، پاداش مربوط به اقدام انتخابی را از محیط دریافت می‌کند و سپس به وضعیت جدیدی منتقل می‌شود.

۳- کارهای مرتبط

حملات علیه RPL را می توان به سه دسته اصلی شامل حملات به منابع، توپولوژی و ترافیک شبکه تقسیم بندی کرد. برای نمونه، حمله سیبیل با جعل هویت‌های متعدد توسط یک گره مخرب، باعث اختلال در فرآیند مسیریابی می شود. حمله سیاه چاله با جذب و حذف بسته‌ها توسط یک گره مخرب، مسیر داده را قطع می‌کند. در حمله فروچاله، گره مهاجم با تبلیغ دروغین یک مسیر بهینه، حجم زیادی از ترافیک را به سمت خود جذب می‌کند. همچنین در حمله مرد میانی، گره مهاجم با قرارگیری در مسیر ارتباطی، داده‌ها را شنود یا دستکاری می‌کند [۱۴].

محققان روش‌های امنیتی RPL برای مقابله با حملات را به دودسته اصلی تقسیم می‌کنند: راهکارهای مبتنی بر پروتکل امن و راهکارهای مبتنی بر سیستم تشخیص نفوذ (IDS[۸]). روش‌های مبتنی بر پروتکل امن شامل رمزنگاری، بهره‌گیری از اعتماد و رویکردهای مبتنی بر آستانه هستند. در مقابل، راهکارهای مبتنی بر سیستم تشخیص نفوذ با شناسایی ناهنجاری‌ها، امنیت RPL را افزایش می‌دهند و شامل روش‌های مبتنی بر امضا، ناهنجاری، ویژگی و ترکیبی هستند [۱۷, ۱۸].

با توجه به اینکه تمرکز رهیافت پیشنهادی بر مبنای اعتماد است بنابراین در این بخش، مقالات قبلی امن سازی پروتکل RPL مبتنی بر اعتماد، بررسی می‌شود.

رویکرد SecTrust-RPL [7] یک پروتکل مسیریابی مبتنی بر اعتماد است که گره‌های مخرب را شناسایی و ایزوله می‌کند. این مدل برای مسیریابی امن تنها بر اساس نرخ هدایت بسته‌ها مورد اعتمادترین گره را انتخاب می‌کند و سایر پارامترهای محیطی تأثیرگذار نظیر اعتماد سطح پیوند ارتباطی را در مسیریابی در نظر نمی‌گیرد. در روش ارائه شده ، نویسندگان از مشاهده رفتار گره ها برای هدایت بسته ها رو به جلو جهت ارزیابی اعتماد گره های همسایه بهره برده اند . بدین صورت که در صورتی که گره i بخواهد گره j را در بازه زمان ارزیابی t مورد ارزیابی قرار بدهد و میزان اعتماد گره j را بدست بیاورد ، از طریق فعال سازی مد شنود بررسی می کند که گره j به چه میزان از بسته هایی را که برایش ارسال کرده است ، به جلو هدایت کرده است . همچنین یک پارامتر تنبیه در نظر گرفته است که گره انتخاب شده به عنوان والد به میزان بسته هایی که هدایت نکند مجازات می شود و از میزان اعتماد آن کاسته می شود . بدین تریتب هر گره برای انتخاب والد قابل اعتماد همواره گره های همسایه را ارزیابی نموده و گره ای که بالاترین اعتماد را از بین همسایگان مجاز داشته باشد برای مسیریابی انتخاب می کند. هاشمی و شمس‌علیئی یک مدل چندبعدی مبتنی بر اعتماد به نامDCTM-IoT را در [۱۹] ارائه کردند که کیفیت سرویس(QoS)[۹] ، کیفیت داده‌های زمینه‌ای(QPC)[۱۰] ، و اطلاعات متنی را در نظر می‌گیرد. بااینکه این روش در برابر حملات سیاه‌چاله، کاهش رتبه، و سیبیل مقاوم است، اما به دلیل وابستگی زیاد به داده‌های تاریخی، برای دستگاه‌های اینترنت اشیا که با محدودیت شدید منابع مواجه هستند، مناسب نیست .همچنین FDTM-IoT [20] نیز یک مدل سلسله‌مراتبی پویا مبتنی بر منطق فازی است که در محیط‌های متحرک عملکرد خوبی دارد، اما مانند رویکرد DCTM-IoT به منابع زیاد نیاز دارد.

رویکرد MRTS [21] با استفاده از نظریه بازی‌ها همکاری گره‌ها را ارزیابی و گره‌های مخرب را شناسایی می‌کند، اما در محاسبات مربوط به اعتماد گره‌ها از توصیه‌های گره‌های همسایه بهره می‌برد ولی اعتبارسنجی پیشنهادها را در نظر نمی‌گیرد که این امر خودش ممکن است منجر به حملات مبتنی بر توصیه‌های نادرست شود. رویکردهای SIEWEو T-SIEWE که در [۲۲] ارائه‌شده است، بر شناسایی حملات سیاه‌چاله با محدود کردن تعداد گره‌های نظارتی و فعال کردن حالت شنود تمرکز کردند. گره‌های نظارتی رفتارهای ارسال بسته همسایگان خود را مشاهده می‌کنند. سطح اعتماد گره‌ها در این رویکرد با نسبت بسته‌های ارسال‌شده به بسته‌های دریافتی تعیین‌شده است. اگرچه این رویکرد دارای نقاط قوتی نظیر محدود نمودن حالت شنود به گره‌های نظارتی است، اما با تولید بسته‌های اعتماد جداگانه برای انتقال اطلاعات اعتماد سربار محاسباتی شبکه را افزایش می‌دهند که این اطلاعات می‌توانستند در پیام‌های DIO موجود در پروتکل RPL گنجانده شوند. در رویکرد CTrust-RPL [۲۳] نویسندگان ، تنها اعتماد گره‌ها را بررسی کرده‌اند و اعتماد پیوندهای ارتباطی را نادیده گرفته‌اند که در نظر گرفتن اعتماد پیوندهای ارتباطی می‌توانست کارایی مسیریابی را بهبود بخشد. جانانی و همکارش یک مدل اعتماد به نام RFTrust را در [۲۴] معرفی کردند که از اعتماد مستقیم به همراه اعتماد غیرمستقیم مبنی بر توصیه بهره برده است. این رویکرد از الگوریتم جنگل تصادفی برای محاسبه اعتماد مستقیم استفاده می‌کند، اما فیلتر کردن پیشنهادهای نامعتبر را در نظر نگرفته است.

روش مبتنی بر رگرسیون لجستیک که در [۲۵] ارائه‌شده است، رفتار گره‌ها را برای شناسایی مهاجمان تحلیل نموده است، اما به دلیل وابستگی به کنترل‌کننده مرکزی، دچار نقطه شکست مرکزی است. مطالعه‌ای دیگر در [۲۶] اعتماد گره‌ها را بر اساس مصرف انرژی، تأخیر و نرخ تحویل بسته محاسبه کرده، اما به دلیل اینکه محاسبات مربوطه به اعتماد به گره ریشه واگذار می‌شود بار پردازشی زیادی به گره ریشه تحمیل می‌شود. مدل ارائه‌شده در [۲۷] با ترکیب مکان‌یابی جغرافیایی و انتخاب والدین مبتنی بر اعتماد حملات سیبیل را کاهش داده، اما با اضافه کردن بسته‌های درخواست-پاسخ، جدید در شبکه باعث ازدحام شبکه می‌شود.

رویکرد SecRPL-MS [28] با استفاده از روش مبتنی بر اعتماد جهت انتخاب گره‌ها برای مسیریابی بسته‌ها و الگوریتم بهینه‌سازی سیل فیش و شبکه عصبی QINN[۱۱] جهت شناسایی مهاجمان، حملات مختلف مانند کاهش رتبه، مردمیانی، سیبیل و سیاه‌چاله را شناسایی کرده است، اما به یک‌نهاد امنیتی وابسته است که می‌تواند نقطه شکست شبکه باشد. راهکار ارائه‌شده در [۲۹] با نام THC-RPL، محاسبات اعتماد را به گره ریشه واگذار کرده تا بار پردازشی گره‌ها را کاهش دهد، اما با جمع‌آوری اعتماد غیرمستقیم از همه همسایگان، سربار محاسباتی بالایی دارد.

درنهایت رویکرد SMTrust [30]شامل دو فاز تشکیل اعتماد و تشخیص حمله است و از معیارهایی مانند میزان موفقیت در ارسال بسته‌ها، انرژی باقی‌مانده، موقعیت مبتنی بر شاخص قدرت سیگنال دریافتی RSSI[۱۲] و مشاهدات تاریخی برای ارزیابی اعتماد استفاده می‌کند. در روش ارائه شده از جمع وزنی این معیارها برای محاسبه

اعتماد گره ها در شبکه استفاده شده است. این روش در برابر حملات سیاه‌چاله و کاهش رتبه عملکرد بهتری نسبت به MRHOF، DCTM [19] و SecTrust-RPL [7] دارد اما درباره نحوه ارسال مقادیر اعتماد در پیام‌های DIO توضیح کافی ارائه نکرده است. با این وجود نویسندگان در مقاله دیگر خود [۳۱] پژوهش مشابه ای انجام دادند که روش محاسبه اعتماد غیر مستقیم را بر اساس توصیه های همسایگان تشریح نموده اند.

در پژوهش [۳۲]، محققان اثرات حملات تخلیه انرژی، به‌ویژه حملات سیل پیام Helloو تغییر شماره نسخه را بر پروتکلRPL که به‌طور گسترده در شبکه‌های اینترنت اشیاء استفاده می‌شود، بررسی کرده‌اند. نویسندگان نشان داده‌اند که این حملات با تخلیه انرژی گره‌های محدود از نظر منابع، باعث ناپایداری شبکه می‌شوند. برای مقابله با این تهدیدات، یک مکانیزم مبتنی بر اعتماد ارائه شده است که پیش از پردازش پیام‌های کنترلی، قابلیت اعتماد گره‌ها را ارزیابی می‌کند. اثربخشی این راهکار از طریق شبیه‌سازی‌های گسترده و مقایسه با چهار سناریوی پایه بدون گره‌های مخرب نشان داده شده است. نتایج حاکی از آن است که روش پیشنهادی به‌طور قابل‌توجهی تأثیرات مخرب این حملات را بر عملکرد شبکه کاهش می‌دهد.

جدول ۱: مقایسه تحقیقات در مورد امن سازی پروتکل مسیریابی RPL با استفاده از روش‌های مبتنی بر اعتماد.

مرجع/ سال انتشار	راهکار اصلی پژوهش	مزایا	معایب	حمله موردبررسی
[۱۹] ۲۰۱۹	ارائه مدل اعتماد چندبعدی با ترکیب اعتماد مستقیم و غیرمستقیم	اندازه‌گیری اعتماد گره‌ها به‌صورت پویا با پارامترهای قابل تنظیم اعتبارسنجی توصیه‌های همسایگان برای جلوگیری از حملات مبتنی بر توصیه	نیاز به تاریخچه گسترده‌ای از داده‌ها جهت ارزیابی اعتماد افزایش سربار پیام‌های کنترلی ناشی از توصیه‌ها پیچیدگی بالا در پیاده‌سازی	حمله سیاه‌چاله، حمله رتبه و حمله سیبیل
[۲۰] ۲۰۲۰	ارائه مدل اعتماد سلسله‌مراتبی، پویا و فازی با ابعاد چندگانه	استفاده از منطق فازی برای مقابله با عدم‌قطعیت محاسبه اعتماد دومرحله‌ای (مستقل و تجمیع‌شده) ابعاد اعتماد انعطاف‌پذیر برای انطباق‌پذیری	وابستگی به تاریخچه زیادی از داده‌ها پیچیدگی محاسباتی بالا افزایش سربار ارتباطی به دلیل محاسبات و تبادل مکرر معیارهای اعتماد نیازمندی به سفارشی‌سازی در سناریوهای وابسته به اطلاعات زمینه‌ای	حمله سیاه‌چاله، حمله رتبه و حمله سیبیل
[۲۱] ۲۰۲۰	مسیریابی مبتنی بر ارزیابی اعتماد و تحلیل نظریه بازی برای مقابله با رفتارهای خودخواهانه	جلوگیری از حملات داخلی و رفتارهای خودخواهانه برای افزایش امنیت بهینه‌سازی انتخاب گره‌ها با استفاده از تعادل مصرف انرژی ارزیابی اعتماد با استفاده از معیارهای متنوع (انرژی، کیفیت پیوند ارتباطی، صداقت، خودخواهی)	نیاز به منابع محاسباتی بالا آسیب‌پذیری به توصیه‌های نامعتبر گره‌ها عدم به‌کارگیری راهکار اعتبارسنجی توصیه ها چالش‌های مقیاس‌پذیری در شبکه‌های بزرگ یا در پیاده‌سازی در دنیای واقعی پیچیدگی در پیاده‌سازی	حمله سیاه‌چاله و حمله رتبه
[۲۲] ۲۰۲۱	ترکیب مدل اعتماد و روش‌های آماری برای شناسایی و فیلتر کردن گره‌های مشکوک	کاهش مصرف انرژی با محدود کردن گره‌ها در حالت استراق سمع ترکیب اعتماد و فیلترهای آماری برای شناسایی دقیق (راهکارهای اعتماد و غربالگری)	عدم توجه به اعتماد در سطح پیوندهای ارتباطی افزایش سربار ارتباطی به دلیل تبادل داده‌های مرتبط با اعتماد پیچیدگی بیشتر به دلیل مشارکت چندین گره در محاسبه مقدار اعتماد عدم استحکام کافی در محیط‌های پویا به دلیل وابستگی به فیلترهای آماری	حمله سیاه‌چاله
[۲۳] ۲۰۲۱	مدل اعتماد سلسله مراتبی برای شناسایی حملات با انتقال محاسبات اعتماد به لایه کنترل	انتقال محاسبات به لایه کنترل برای صرفه‌جویی در انرژی مصرفی گره‌ها متمرکز کردن محاسبات اعتماد برای کاهش بار پردازشی گره‌ها	عدم توجه به اعتماد پیوندهای ارتباطی ایجاد گلوگاه به دلیل متمرکز بودن محاسبات اعتماد آسیب‌پذیری‌های امنیتی بالقوه به دلیل دستکاری در اعتماد مشکلات مقیاس‌پذیری در شبکه‌های بزرگ و توزیعی	حمله سیاه‌چاله
[۷] ۲۰۱۹	مدل اعتماد مستقیم بر اساس رفتار گره‌ها برای هدایت بسته‌ها و تنبیه گره‌های مهاجم با اتخاذ پارامتر تنبیه	سادگی رویکرد ارائه‌شده برای محاسبه اعتماد گره‌ها ارزیابی مدل از طریق شبیه‌سازی و پیاده‌سازی سناریوی مربوطه در دنیای واقعی	افزایش نرخ از دست دادن بسته‌ها به دلیل انتخاب گره‌های مورد اعتماد باانرژی کم برای هدایت بسته‌ها عدم در نظر گرفتن اعتماد پیوندهای ارتباطی	حمله رتبه و حمله سیبیل
[۲۴] ۲۰۲۱	مدل مبتنی بر اعتماد با استفاده از منطق موضوعی Subjective)	کاهش سربار ارتباطی با محاسبه کارآمد اعتماد از طریق اعتماد مستقیم یا غیرمستقیم	عدم در نظر گرفتن شرایط محیطی مانند افت سیگنال و تداخل	حمله فرو چاله

جدول ۱: مقایسه تحقیقات در مورد امن سازی پروتکل مسیریابی RPL با استفاده از روش‌های مبتنی بر اعتماد.

مرجع/ سال انتشار	راهکار اصلی پژوهش	مزایا	معایب	حمله موردبررسی
	(Logic و جنگل تصادفی (Random Forest) برای شناسایی و مقابله با حملات	اعتبارسنجی علمی از طریق تحلیل ریاضی برای تضمین اثربخشی مدل بهبود دقت ارزیابی اعتماد با ترکیب روش‌های تخفیف و اجماع در اعتماد غیرمستقیم	بی‌توجهی به کیفیت پیوند ارتباطی در ارزیابی اعتماد مستقیم	
[۲۵] ۲۰۲۱	مدل مبتنی بر اعتماد با استفاده از رگرسیون لجستیک برای شناسایی و جداسازی گره‌های مخرب در شبکه‌های اینترنت اشیا مبتنی بر RPL	ارزیابی چندمعیاره اعتماد با ترکیب اعتماد مستقیم، شهرت و تجربه دقت بالا در شناسایی گره‌های مخرب با استفاده از رگرسیون لجستیک	ایجاد نقطه شکست به دلیل وابستگی به کنترل‌کننده مرکزی پیچیدگی پردازشی بالا به دلیل استفاده از رگرسیون لجستیک و ارزیابی چندمعیاره وابستگی عملکرد مدل به داده‌های آموزشی افزایش مصرف انرژی عدم توجه به کیفیت پیوند ارتباطی	حمله سیاه‌چاله
[۲۶] ۲۰۲۱	مدل مبتنی بر اعتماد با به‌کارگیری گره‌های شاهد برای ارزیابی قابلیت اطمینان گره‌ها و شناسایی گره‌های تکراری	شناسایی گره‌های در معرض خطر از طریق آزمون تکرار تشخیص مؤثر حملات تکرار کاهش نرخ شکست ارتباطات	وابستگی به گره ریشه برای ارزیابی اعتماد افزایش سربار محاسباتی و ارتباطی عدم در نظر گرفتن کیفیت پیوند ارتباطی	حمله تکرار یا باز پخش
[۲۷] ۲۰۲۱	انتخاب والد بر اساس ترکیبی از مدل اعتماد و موقعیت جغرافیایی و بهره‌گیری از گره‌های ناظر برای شناسایی گره‌های مخرب	استفاده از جداول ردیابی برای پایش و شناسایی گره‌های مخرب جداسازی سریع گره‌های مخرب و کاهش تأثیر آن‌ها بر شبکه	افزایش سربار شبکه به دلیل استفاده از بسته‌های درخواست و پاسخ افزایش پیچیدگی استقرار به علت وابستگی به گره‌های نظارتی سربار ذخیره‌سازی و پردازشی بالا در شبکه‌های بزرگ عدم سازگاری اطلاعات جغرافیایی ایستا با شبکه‌های پویا	حمله سیبیل
[۲۸] ۲۰۲۱	ارائه پروتکل امن و احراز هویت شده مبتنی بر اعتماد برای شبکه‌های اینترنت اشیاء مبتنی بر RPL با پشتیبانی از سینک متحرک و بهره‌گیری از الگوریتم Sailfish برای مسیریابی، الگوریتم Prince برای رمزنگاری و QINN برای انتخاب مسیر سینک متحرک	ثبت تمامی گره‌ها در یک‌نهاد امنیتی در مرحله اولیه راه‌اندازی استفاده از احراز هویت قوی برای افزایش امنیت بکارگیری روش‌های احراز هویت امن تأمین امنیت داده‌ها با استفاده از الگوریتم رمزنگاری افزایش طول عمر شبکه با بهره‌گیری از سینک متحرک برای جمع‌آوری داده‌ها بهبود مقیاس‌پذیری با تقسیم شبکه به توری‌های[۱۳] هم‌اندازه با مدیریت سر شبکه	آسیب‌پذیری به دلیل وابستگی به یک‌نهاد امنیتی مرکزی پیچیدگی بالای پیاده‌سازی افزایش سربار محاسباتی الگوریتم‌های پیچیده و رمزنگاری درروش ارائه‌شده افزایش سربار ارتباطی ناشی از فرآیندهای ثبت‌نام و احراز هویت افزایش مصرف انرژی مراحل ثبت‌نام، احراز هویت و رمزنگاری محدودیت در پذیرش در همه سناریوهای اینترنت اشیا به دلیل وابستگی به سینک متحرک	حمله رتبه، حمله سیبیل، حمله سیاه‌چاله و حمله مردمیانی
[۲۹] ۲۰۲۲	مدل ترکیبی اعتماد محور (اعتماد مستقیم و مبتنی بر توصیه) با همکاری گره‌ها برای شناسایی گره‌های مخرب	بهینه‌سازی مصرف منابع با انتقال پردازش‌ها به گره‌های قدرتمند استفاده از لیست گره‌های مخرب برای اطمینان از مسیریابی امن مصرف انرژی کم و افزایش عمر عملیاتی دستگاه‌های اینترنت اشیا	افزایش بار محاسباتی به دلیل پردازش اعتماد غیرمستقیم وابستگی به گره ریشه و ایجاد یک نقطه شکست مرکزی افزایش سربار پردازشی گره ریشه در شبکه‌های بزرگ تأخیر در ارتباطات ناشی از انتشار لیست گره‌های مخرب	حمله سیبیل
[۳۰] ۲۰۲۲	مدل ترکیبی اعتماد مبتنی بر رفتار گره‌ها در هدایت بسته‌ها، کیفیت پیوند ارتباطی، توصیه‌های دریافتی و معیارهای تحرک	پشتیبانی از گره‌های ثابت و متحرک با محاسبه اعتماد مبتنی بر تحرک سازگاری پویا با تغییرات شبکه و حفظ پایداری توپولوژی	عدم توضیح دقیق درباره محاسبه اعتماد مبتنی بر توصیه، پیچیدگی بالای محاسبات اعتماد افزایش تأخیر و سربار پردازشی تمرکز محدود بر شبکه‌های کاملاً پویا با تغییرات مداوم گره‌ها	حمله رتبه و حمله سیاه‌چاله
[۳۲] ۲۰۲۳	آنالیز رفتار گره ها جهت ارزیابی اعتماد آنها با استفاده از تطبیق پیام های کنترلی با مدل مجموعه کاری	تشخیص رفتاری گره‌های مهاجم تحلیل دقیق انرژی گره‌های قربانی	عدم تحلیل دقیق شاخص های مهم شبکه بررسی محدود پارمترهای کارایی (مصرف انرژی)	حمله سیل پیام hello و حمله تغییر نسخه

جدول ۱، مطالعات مربوط به حملات کاهش رتبه و سایر حملات مرتبط با توپولوژی را خلاصه می‌کند و نقاط قوت و محدودیت‌های هرکدام را نشان می‌دهد. از این جدول مشخص است که بسیاری از رویکردهای موجود با چالش‌هایی مانند نیاز بیش‌ازحد به منابع، در نظر نگرفتن پارامترهای کافی یا ایجاد سربار محاسباتی بالا مواجه هستند. رویکرد پیشنهادی ما با اعمال بهبودهای معماری در پروتکلRPL ، این محدودیت‌ها را رفع می‌کند. به‌طور خاص، نوآوری‌های اصلی ما به شرح زیر است:

- افزودن کیفیت پیوند ارتباطی به محاسبه اعتماد:

برخلاف سایر روش ها [۷, ۲۳-۲۶]، مدل پیشنهادی کیفیت پیوند ارتباطی را به‌عنوان عاملی حیاتی در ارزیابی اعتماد وارد می‌کند. این امر باعث می‌شود که انتخاب والد دقیق‌تر انجام‌شده و ارتباطات بین گره‌ها قابل‌اطمینان‌تر باشد.

- در نظر گرفتن انرژی باقیمانده گره:

در مقابل [7] SecTrust-RPL مدل پیشنهادی انرژی باقی‌مانده گره را به‌عنوان یک معیار کلیدی برای سنجش اعتماد در نظر می‌گیرد. این ویژگی امکان تصمیم‌گیری‌های بهینه و صرفه‌جویی در مصرف انرژی را فراهم می‌آورد.

- استفاده بهینه از بسته‌های DIO برای اشتراک‌گذاری انرژی باقیمانده گره‌ها:

برخلاف مدلی که در [۲۷] پیشنهادشده و بسته‌های پرسش و پاسخ جدیدی به شبکه اضافه می‌کند، رویکرد ما از طریق اصلاح پیام DIOموجود، از ارسال بسته‌های جداگانه جلوگیری می‌کند. این بسته‌های تغییریافته به‌طور دوره‌ای بر اساس تایمر Trickle[33] ارسال می‌شوند تا همسایگان بدون افزایش سربار ارتباطی، از وضعیت انرژی گره‌ها مطلع شوند.

- حذف اعتماد مبتنی بر توصیه‌ها:

برای ساده‌سازی و بهینه‌سازی مدل اعتماد، همانند SecTrust-RPL [۷]، ما صرفاً به اعتماد مستقیم تکیه می‌کنیم و از اعتماد مبتنی بر توصیه که در برخی رویکردها مانند [۱۹-۲۱, ۲۹, ۳۰] به‌کاررفته است، اجتناب می‌کنیم. این امر نیاز به ارزیابی و اعتبارسنجی توصیه‌ها را حذف کرده و مدل را سبک و کارآمد نگه می‌دارد.

- محاسبه اعتماد به‌صورت توزیعی مابین گره‌ها:

مدل پیشنهادی به‌صورت کاملاً توزیع‌شده عمل می‌کند؛ به این معنی که هر گره به‌طور مستقل اعتماد همسایگان خود را ارزیابی می‌کند. این ویژگی نیاز به یک کنترل‌کننده مرکزی را برطرف کرده که ممکن است به‌عنوان نقطه شکست واحد عمل کند. برخلاف مدل‌های محاسبه اعتماد متمرکز [۲۳, ۲۵, ۲۶, ۲۸, ۲۹] که وابستگی زیادی به کنترل مرکزی دارند، در مدل پیشنهادی در صورت بروز مشکل در یک گره، سایر بخش‌های شبکه بدون اختلال به فعالیت خود ادامه می‌دهند.

- استفاده کارآمد از معیارهای مرتبط برای محاسبه اعتماد:

برخلاف روش‌های دیگری [۱۹, ۲۰, ۳۰] که به داده‌های تاریخی گسترده متکی هستند، رویکرد ما از معیارهای مهمی مانند نرخ انتقال، کیفیت پیوند ارتباطی و انرژی باقیمانده بهره می‌برد. این معیارها برای محاسبه اعتماد کافی بوده و باعث کاهش سربار محاسباتی شده و مدل را برای دستگاه‌های اینترنت اشیا با منابع محدود مناسب می‌سازد.

با اعمال این بهبودها و بهره‌گیری از یادگیری تقویتی برای ارزیابی تطبیقی اعتماد، روش پیشنهادی ما امنیت مسیریابی را به‌طور چشمگیری افزایش داده درنتیجه منجر به افزایش دقت مسیریابی در انتخاب گره‌های مورد اعتماد شده و پایداری شبکه را در برابر حملات کاهش رتبه بهبود می‌بخشد.

۴- روش پیشنهادی

در این بخش، ساختار کلی روش پیشنهادی برای افزایش امنیت مسیریابی در شبکه‌های مبتنی بر RPL ارائه می‌شود. با توجه به آسیب‌پذیری این پروتکل در برابر حملات داخلی مانند کاهش رتبه و سیبیل، یک مدل اعتماد ترکیبی مبتنی بر رفتار گره‌ها و کیفیت پیوندها طراحی شده است. تصمیم‌گیری عامل‌ها برای انتخاب مسیر امن، با بهره‌گیری از یادگیری تقویتی Q انجام می‌گیرد تا در مواجهه با رفتارهای مشکوک، انتخاب والدین مسیریابی بهینه و پویا باشد.

بخش‌بندی روش پیشنهادی بدین‌صورت است که در بخش ۴-۱ به معرفی ساختار پروتکل مسیریابی امن پیشنهادی پرداخته می‌شود. در بخش ۴-۲ رویکرد محاسبه اعتماد در سطح گره و در سطح پیوند ارتباطی موردبررسی قرار می‌گیرد. در بخش ۴-۳ تابع پاداش برای تصمیم‌گیری بهینه عامل‌ها برای مسیریابی مبتنی بر اعتماد ارائه می‌شود و درنهایت فرموله کردن مسئله مسیریابی امن توسط تابع هدف مبتنی بر یادگیری-Q برای انتخاب والد امن در بخش ۴-۴ شرح داده می‌شود.

۴-۱- ساختار پروتکل مسیریابی امن پیشنهادی

ایده اصلی پشت پروتکل مسیریابی امن طراحی‌شده در این مقاله

بسیار ساده است و یک روش مسیریابی امن برای شبکه‌های اینترنت اشیا فراهم می‌کند که با شناسایی و جداسازی گره‌های مخرب، ارتباطات امن را تضمین می‌کند. این رهیافت با بهره‌گیری از یادگیری تقویتی مبتنی بر اعتماد، امن‌ترین مسیر را برای ارسال داده‌ها انتخاب می‌کند. در این روش، اعتماد در دو سطح گره و پیوند ارتباطی محاسبه می‌شود:

- اعتماد سطح گره بر اساس رفتار گره‌ها در ارسال بسته‌ها و میزان انرژی باقیمانده آن‌ها محاسبه می‌شود.
- اعتماد سطح پیوند ارتباطی با ارزیابی کیفیت پیوند ارتباطی از طریق مقدار ETX تعیین می‌شود.

مراحل کارکرد این پروتکل به‌صورت شکل ۱ است. از آنجایی که در روش پیشنهادی انتخاب گره های والد بر اساس یادگیری Q انجام می شود بنابراین در ابتدا مقداردهی اولیه پارامترهای مربوط به یادگیری Q نظیر α,γ انجام می شود. از آنجایی که گره های عامل دانشی از محیط ندارند بنابراین در ابتدای راه اندازی شبکه مقدار ثابت ε_0 نیز از یک مقدار ثابت مقدار دهی اولیه می شود و به مرور زمان که نرخ یادگیری افزایش می یابد از مقدار ε کاسته می شود. همچنین مقدار اعتماد همه گره ها برابر اعتماد آستانه لحاظ می گردد. بعد از آن فرآیندهایی که در مدل پیشنهادی برای هدایت بسته ها به مقصد که ریشه DoDAG است، انجام می‌شود به شرح ذیل است.

- **جمع‌آوری اطلاعات و محاسبه اعتماد در سطح گره و پیوند ارتباطی**

- گره‌ها اطلاعات مربوط به ارسال و دریافت بسته‌ها را با استفاده از حالت شنود جمع‌آوری می‌کنند.
- انرژی باقی‌مانده گره‌ها از طریق پیام‌های DIO درون شبکه بین همسایگان اعلان می‌شود.
- کیفیت پیوند ارتباطی با مقدار ETX محاسبه‌شده و در جداول همسایگی به‌روز می‌شود.

- **انتخاب والد بر اساس سیاست حریصانه اپسیلون[۱۴]**

در این مرحله گره جاری از لیست همسایگان مجاز برای مسیریابی که رتبه آنها کمتر از رتبه خودش باشد بر اساس سیاست حریصانه اپسیلون با احتمال ε، به صورت تصادفی یکی از گره ها را به عنوان والد انتخاب می کند و با احتمال $\boldsymbol{\varepsilon - 1}$ والدی که ماکزیمم مقدار Q را داشته باشد به عنوان والد انتخاب می شود.

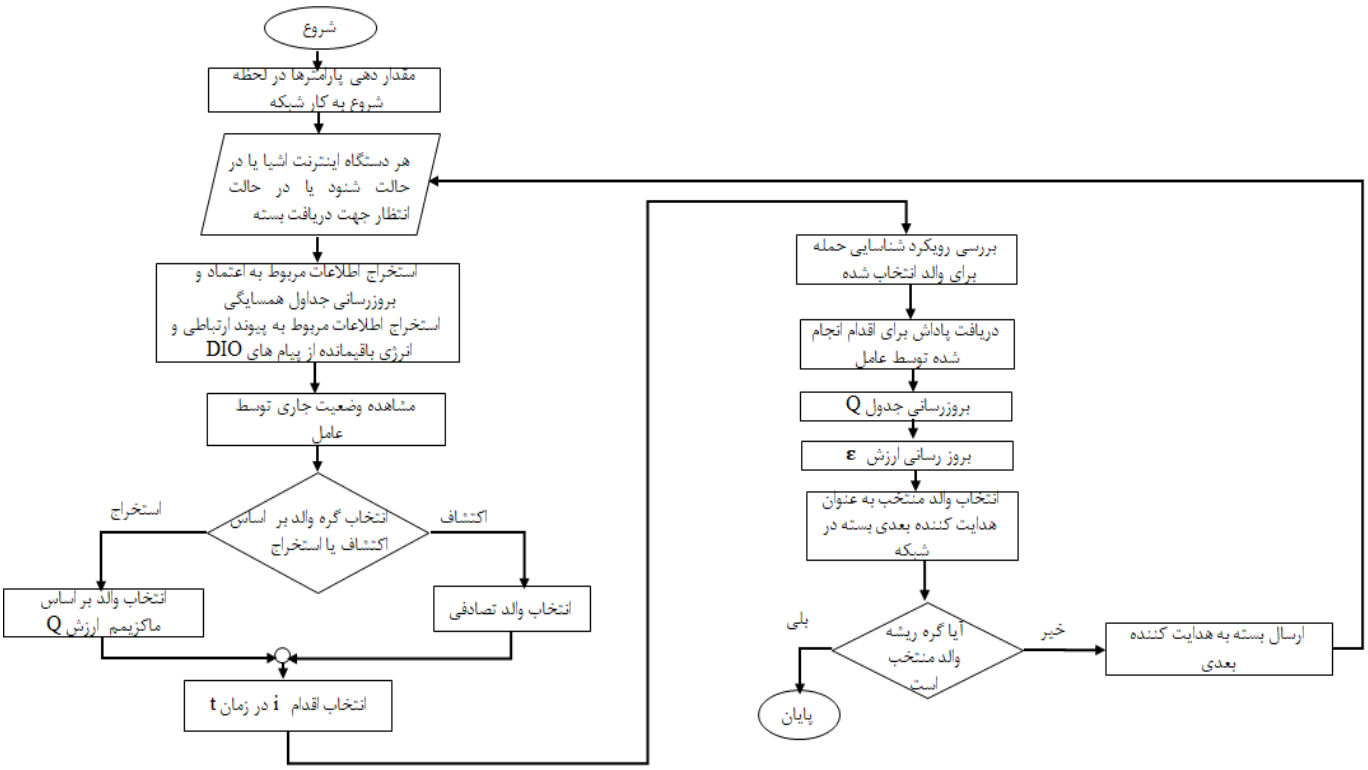
- **شناسایی گره‌های مخرب و محاسبه مقدار پاداش**

- با بررسی رفتار گره‌ها، کیفیت پیوند ارتباطی و میزان انرژی باقیمانده گره‌ها مقادیر اعتماد سطح گره و سطح پیوند ارتباطی محاسبه‌شده و بر اساس آن پاداش عامل از محیط توسط انتخاب والد موردنظر تخصیص داده می‌شود.
- در این پژوهش، تمرکز بر شناسایی گره‌های مخرب داخلی با تمرکز بر حمله کاهش رتبه است؛ برای این منظور، از رویکردی مشابه با روش SecTrust-RPL [7] جهت شناسایی و حذف گره‌های مخرب استفاده‌شده است.
- هنگامی که هر گره، گره ای را به عنوان والد انتخاب کند، تابع پاداش گره انتخاب شده را مورد ارزیابی قرار می دهد در صورتی که گره انتخاب شده گره مخرب باشد پاداش منفی خیلی بزرگ یا جریمه به گره داده می شود در غیر این صورت پاداش مثبتی بر اساس میزان اعتماد سطح گره و سطح لینک به گره اختصاص داده می شود.

- **مسیریابی امن مبتنی بر یادگیری تقویتی و به‌روزرسانی Q-Table**

- گره‌ها با استفاده از الگوریتم یادگیری- Qو استراتژی-ε greedy بهترین والد را انتخاب می‌کنند.
- بر اساس پاداش کسب شده مقدار مربوط به ارزش Q در هر مرحله به‌روزرسانی شده و در جدول Q ذخیره می‌شود.
- ارزش ε، با افزایش دانش عامل، به‌طور خودکار کاهش‌یافته و وابستگی به بهره‌برداری را افزایش می‌دهد.
- گره منتخب برای مسیریابی به عنوان والد مورد ارزیابی قرار می گیرد در صورتی که گره والد ریشه باشد بسته به مقصد منتقل شده و کار به اتمام می رسد در غیر این صورت گره والد به عنوان هدایت کننده بسته رو به جلو می باشد و بسته را ارسال کرده و فرآیند برای گره‌های بعدی تکرار می‌شود تا بسته به مقصد برسد.

بنابراین پروتکل پیشنهادی نه‌تنها امنیت شبکه را افزایش می‌دهد، بلکه با مدیریت انرژی و کیفیت پیوند ارتباطی، عملکرد مسیریابی را نیز بهینه می‌کند.



شکل ۱: نمودار جریانی روش پیشنهادی

۴-۲- محاسبه اعتماد در سطح گره و پیوندهای ارتباطی

در فرآیند مسیریابیRPL ، انتخاب بهترین والد از میان گره‌های همسایه برای حفظ ثبات و طول عمر شبکه بسیار مهم است. در سطح گره، انرژی باقی‌مانده عامل اصلی در تعیین ظرفیت گره‌ها برای انتقال مؤثر داده‌ها است. هر گره حسگر در شبکه چهار معیار توان را نظارت می‌کند :حالت توان پایین (LPM[۱۵])، توان استفاده از پردازنده، توان شنود رادیویی و توان ارسال رادیویی. این معیارها به‌طور جمعی مصرف انرژی را در حالت‌های عملیاتی مختلف ارزیابی می‌کنند. انرژی باقی‌مانده یک گره سطح اعتماد انرژی آن را نشان می‌دهد که با استفاده از معادله (۱) محاسبه می‌شود. این معادله از تفاوت بین سطح باتری حداکثری یا اولیه $E_{initial}$ و انرژی مصرف‌شده $E_{Depleted}$ توسط گره به‌دست می‌آید [۳۴].

$$R_E = E_{initial} - E_{Depleted} \qquad (۱)$$

انرژی مصرف‌شده نیز از طریق معادله (۲) محاسبه می‌شود.

$$E_{Depleted} = V \times \frac{T_{CPU}I_{CPU} + T_{LPM}I_{LPM} + T_{RX}I_{RX} + T_{TX}I_{IX}}{R_{timer}} \qquad (۲)$$

متغیرهای T_{CPU},T_{LPM},T_{RX} ,T_{TX} مدت‌زمان فعالیت میکروکنترلر را در حالت‌های فعال، کم‌مصرف، دریافت و ارسال نشان می‌دهند. به همین ترتیب، متغیرهای I_{CPU},I_{LPM},I_{RX},I_{TX} جریان موردنیاز میکروکنترلر در این حالت‌ها را مشخص می‌کنند. همچنین، دو متغیر R_{timer} و V به ترتیب تعداد تیک‌ها در هر ثانیه و ولتاژ باتری را نشان می‌دهند.

در سطح گره معیار دیگر، اعتماد به ارتباطات با بررسی موفقیت‌آمیز تبادل بسته‌ها بین گره‌ها تعیین می‌شود. اعتبارسنجی ارسال مؤثر بسته‌ها با نظارت مستمر بر رفتار گره‌های ارتباطی در ارسال بسته‌های دریافتی انجام می‌شود. درروش پیشنهادی، معادله (۳) سطح اعتماد هر گره را محاسبه می‌کند، جایی که پارامترβ نشان‌دهنده مجازات بوده و گره‌های مخرب را جریمه می‌کند. مقدار اولیه این پارامتر ۰،۰۱ است و با هر مورد شناسایی رفتار مخرب در عدم ارسال بسته‌ها این پارامتر بر اساس معادله (۴) افزایش می‌یابد. معادله (۵) سطح اعتماد اولیه تمامی گره‌ها را برابر آستانه تعیین‌شده در شروع راه‌اندازی شبکه تعیین می‌کند.

در معادله (۳)، مقدار $T_t(N_i,N_j)$ نشان‌دهنده سطح اعتمادی است که گره i به توانایی گره j در ارسال بسته‌ها در زمان t دارد. در اینجا، گره i به‌عنوان ارزیاب و گره j به‌عنوان گره ارزیابی‌شونده در نظر گرفته می‌شوند. هر گره دو مقدار را برای بررسی سطح اعتماد همسایگان خود ذخیره می‌کند: $PF_{ji}(t)$ حجم کل بسته‌هایی که گره j در بازه ارزیابی t ارسال کرده است و $PS_{ij}(t)$حجم کل

بسته‌هایی که گره i در این بازه به گره j منتقل کرده است.

$$T_t(N_i,N_j) = \frac{PF_{ji}(t)}{PF_{ji}(t) + \beta[PS_{ij}(t) - PF_{ji}(t)]} \quad (۳)$$

$$\beta_i = \beta_i + 0.005 \,, \beta_0 = 0.01 \quad (۴)$$

$$T_0(N_i,N_j) = T_{threshold} \quad \forall i,j \,, t = t_0 \quad (۵)$$

برای تعیین اعتماد پیوند ارتباطی، از معیار ETX استفاده می‌شود. بهره‌برداری از این معیار عملکرد RPL را بهبود می‌بخشد و مقادیر پایین‌تر ETX کیفیت پیوند ارتباطی بهتری را نشان می‌دهند .این معیار تعداد بسته‌های انتقالی، شامل ارسال مجددها را برای تحویل موفقیت‌آمیز در یک مسیر مشخص تخمین می‌زند. محاسبه مقدار اعتماد پیوند ارتباطی از طریق معادله (۶) و (۷) بیان می‌شود. سه پارامتر اصلی اعتماد—اعتماد ارتباطی، اعتماد انرژی باقی‌مانده و اعتماد پیوند ارتباطی—برای ارزیابی تابع پاداش جهت تخصیص پاداش در الگوریتم یادگیری تقویتی از روش پیشنهادی ما نقش کلیدی دارند.

$$ETX(t) = \frac{ETX(t) - ETX_{min}}{ETX_{max} - ETX_{min}} = \frac{ETX(t)}{ETX_{max}} \quad (۶)$$

$$Link_Trust = 1 - ETX(t) \quad (۷)$$

۴-۳- تابع پاداش جهت تصمیم‌گیری بهینه عامل‌ها در مسیریابی

عامل با انتخاب والد خاص برای مسیریابی بسته از محیط پیرامون پاداش یا مجازات دریافت می‌کند. درصورتی‌که گره مهاجم را به‌عنوان والد خود انتخاب کند جریمه بزرگی دریافت می‌کند. در صورت انتخاب والد مورد اعتماد پاداش مربوطه را دریافت می‌کند که این پاداش بر اساس جمع وزنی اعتماد سطح گره و سطح پیوند ارتباطی است. اعتماد سطح گره به‌نوبه خود شامل اعتماد ارتباطی و اعتماد انرژی باقی‌مانده گره منتخب است.

تابع پاداش نقش کلیدی در یادگیری Q ایفا می‌کند، زیرا رفتار و عملکرد عامل را تعیین می‌کند. تابع پاداش در طرح پیشنهادی از معادله (۸) پیروی می‌کند:

$$r = \begin{cases} R^+ & normal\ node \\ R^- & Malicious\ node \end{cases} \quad (۸)$$

هنگامی‌که یک عامل، گره مخرب را به‌عنوان ارسال‌کننده بسته انتخاب کند، مقدار پاداش منفی قابل‌توجهی را به‌عنوان جریمه دریافت می‌کند. در مقابل، اگر گره انتخاب‌شده مخرب نباشد، مقدار پاداش طبق معادله (۹) محاسبه می‌شود.

$$R^+ = w_n * Node_trust + w_l * Link_Trust \quad (۹)$$
$$w_n + w_l = 1$$

مقدار پاداش برای عامل در بازه ۰ تا ۱ قرار دارد که جمع وزنی اعتماد سطح گره $Node_trust$ و اعتماد سطح پیوند ارتباطی $Link_Trust$ است. در این معادله w_n و w_l به ترتیب ارزش وزن‌های مربوط به اعتماد سطح گره و اعتماد سطح پیوند ارتباطی است . اگر یک گره ریشه در میان همسایگان ظاهر شود، به‌عنوان بهترین گزینه برای گره والد انتخاب می‌شود که منجر به دریافت بالاترین مقدار پاداش ($R^+ = 1$) خواهد شد. این انتخاب به این دلیل صورت می‌گیرد که در فرآیند مسیریابی، گره مبدأ، مقصد را در جدول همسایگان خود شناسایی کرده و نشان می‌دهد که مقصد در محدوده انتقال گره مبدأ قرار دارد، که این امر پاداش پیوند ارتباطی را به حداکثر مقدار خود می‌رساند. در غیر این صورت، مقدار پاداش بر اساس کیفیت پیوند ارتباطی، میزان اعتماد گره و سطح انرژی گره محاسبه می‌شود. پاداش گره i درصورتی‌که گره j را به‌عنوان والد خود انتخاب کند، طبق معادله (۱۰) تعیین می‌شود.

$$R^+ = w_1 \times \frac{Trust_j}{MAX_{TRUST}} + w_2 \times \frac{RE_j}{MAX_{ENERGY}} + w_3 \times \frac{ETX_j}{MAX_{ETX}} \quad (۱۰)$$
$$w_1 + w_2 + w_3 = 1$$

در این معادله، میزان پاداش جمع وزنی اعتمادهای مبتنی بر ارتباطات در هدایت بسته‌ها، انرژی باقیمانده و اعتماد سطح پیوند ارتباطی محاسبه می‌شود که w_1 , w_2 , w_3 به ترتیب ارزش وزن‌های مربوط به این پارامترها می‌باشند .سطح سطحرژی گره از نسبت انرژی باقی‌مانده گره (RE) به حداکثر انرژی آن MAX_{ENERGY} یا ظرفیت باتری محاسبه می‌شود. مقدار حداکثر سطح انرژی هر گره برابر با ۱ است که نشان‌دهنده گرهی با شارژ کامل است. MAX_{ETX} و MAX_{Trust} نیز به ترتیب حداکثر مقدار ETX و سطح اعتماد هستند.

۴-۴- فرموله کردن مسئله

پروتکل مسیریابی RPL انتخاب بهترین گره والد را بر اساس یک تابع هدف تعیین می‌کند. این تابع معیارهای مشخصی را برای

انتخاب والد از میان گره‌های همسایه در DODAG تعیین می‌کند. درروش پیشنهادی، از یادگیری-Q برای بهینه‌سازی انتخاب والد استفاده‌شده است. هر عامل پس از دریافت بسته از گره منبع، مناسب‌ترین گره همسایه را برای ارسال بسته انتخاب می‌کند. این فرآیند تا رسیدن بسته به گره ریشه ادامه می‌یابد.

هدف عامل، حداکثر سازی پاداش تجمعی در طول عملیات است، به این صورت که با انتخاب بهترین والد در هر مرحله، بازده کلی تخفیف یافته را بهینه می‌کند. این مقدار با استفاده از معادله زیر محاسبه می‌شود:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}$$

$$\gamma \in [0,1] \qquad (۱۱)$$

در این رابطه، γ ضریب تخفیف بوده که مشخص می‌کند عامل چقدر به پاداش‌های آینده اهمیت می‌دهد. اگر مقدار این ضریب نزدیک به صفر باشد، عامل دید کوتاه‌مدت دارد و تنها پاداش‌های آنی را در نظر می‌گیرد.؛ اما با نزدیک شدن این مقدار به ۱، عامل اهمیت بیشتری را برای پاداش‌های آینده قائل می‌شود.

تابع ارزش حالت(V(s ، مقدار پاداش مورد انتظار برای پیروی از سیاست π در حالت s را نشان می‌دهد که مطابق معادله (۱۲) محاسبه می‌شود.

$$V^{\pi}(s) = E_{\pi}[G_t|s_t = s] = E_{\pi}[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1}|s_t \ = s] \qquad (۱۲)$$

هدف عامل، حداکثر سازی پاداش تجمعی در طول عملیات است، به این صورت که با انتخاب بهترین والد در هر مرحله، بازده کلی تخفیف یافته را بهینه می‌کند. تابع ارزش حالت بهینه V^* مقدار پاداش مورد انتظار برای پیروی از یک سیاست بهینه π^* را نشان می‌دهد. این مقاله از یادگیری-Q، یک روش یادگیری تقویتی بدون مدل و خارج از سیاست، برای محاسبه مقادیر بهینه حالت-عمل Q^* استفاده می‌کند. بنابراین بر اساس معادله ۱۳ و ۱۴ داریم :

$$V^*(s) = \max_{\pi} V^{\pi}(s), \ \forall\, s \in S \qquad (۱۳)$$

$$Q^*(s,a) = \max_{\pi} Q^{\pi}(s,a)\ , \forall s \in S, \forall a \in A(s) \qquad (۱۴)$$

تابع ارزش بهینه را می‌توان با استفاده از معادله بلمن مدل‌سازی کرد [۱۵] و به‌صورت معادله (۱۵) نشان داد که در آن احتمال انتقال بین حالات $p(\acute{s}|s,a)$پس از اجرای یک عمل مشخص به‌صورت معادله (۱۶) نشان داده می‌شود.

$$V^*(s) = \max_{a\in A}[r + \gamma \sum_{\acute{s}\in S} p(\acute{s}\,|s,a)V^*(\acute{s})] \qquad (۱۵)$$

$$p(\acute{s}|s,a) = Pr\{S_{t+1} = \acute{s}\ |\ S_t = s\ , a_t = a\} \qquad (۱۶)$$

بنابراین بر اساس معادلات (۱۴،۱۵ و ۱۶) سیاست مسیریابی بر مبنای استفاده از یادگیری-Q به‌صورت معادله (۱۷) خواهد بود.

$$Q^{\pi}(s,a) = E_{\pi}[G_t|s_t = s\ ,a_t = a]$$

$$= E_{\pi}[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1}|s_t = s\ ,a_t = a]$$

$$= r + \gamma \sum_{\acute{s}\in S} p(\acute{s}|s,a)V^{\pi}(\acute{s}) \qquad (۱۷)$$

درنهایت با ترکیب روابط (۱۵) و (۱۷) مقدار بهینه Q به دست می‌آید که به‌صورت معادله (۱۸) است.

$$V^*(s) = \max_{a\in A} Q^*(s,a) \qquad (۱۸)$$

درنتیجه در یادگیری-Q، فرآیند یادگیری عامل طی چندین مرحله به نام دوره‌های یادگیری انجام می‌شود. در هر دوره t ، عامل در حالت sقرار دارد، عمل a را اجرا می‌کند، پاداش r را دریافت کرده و به حالت جدید 's منتقل می‌شود. مقدار Q در هر مرحله با معادله (۱۹) به‌روزرسانی می‌شود.

$$Q(,a_t) \leftarrow\ (1-\alpha)\ Q(s_t,a_t) + \alpha[r_{t+1} + \gamma \max_{a\in A} Q(s_{t+1},a_{t+1})] \qquad (۱۹)$$

در این رابطه، αنرخ یادگیری را نشان می‌دهد که تعیین می‌کند تا چه حد دانش جدید و قبلی بر تصمیمات عامل تأثیر می‌گذارد. درروش پیشنهادی، عامل‌ها از استراتژی ε-greedy پیروی می‌کنند[۱۵]. در مرحله بهره‌برداری، آن‌ها گره والدی را انتخاب می‌کنند که بیشترین مقدار Q را داشته باشد. در مرحله کاوش، گره ارزیابی‌کننده با احتمال ε به‌طور تصادفی یک همسایه را انتخاب می‌کند تا از گرفتار شدن در بیشینه محلی جلوگیری کند. مقدار ε به‌صورت خودکار در طول زمان کاهش می‌یابد. در ابتدا، مقدار ε بالا است، زیرا عامل دانش محدودی از محیط دارد، اما باگذشت زمان کاهش می‌یابد و عامل می‌تواند بیشتر به بهره‌برداری متکی باشد. مقدار ε بر اساس معادله (۲۰) به‌روزرسانی می‌شود[۳۵].

$$\varepsilon_{t_i} = k \times e^{(-1 \times t_i \times decay_rate)} \quad (۲۰)$$

$$k = cnt$$

در این روش، مقدار اولیه ε ثابت در نظر گرفته می‌شود و باگذشت زمان کاهش می‌یابد؛ یعنی از یک مقدار ثابت k نرخ اکتشاف آغاز می شود و با گذر زمان از نرخ اکتشاف کاسته می‌شود و به نرخ بهره‌برداری افزوده می‌شود. در معادله ۲۰، $decay_rate$ به‌عنوان پارامتر تنزیل است . این روند نشان می‌دهد که سطح کاوش در ابتدا بالا است و به‌تدریج با افزایش بهره‌برداری کاهش می‌یابد.

به‌طور خلاصه، روش پیشنهادی از یادگیری- Qبه‌عنوان یادگیری تقویتی برای مدیریت مسیریابی در شبکه‌های با مقیاس وسیع و سازگار با محدودیت‌های منابع در دستگاه‌های اینترنت اشیا بهره می‌برد. در طول فرآیند یادگیری، عامل در مراحل مختلف اقدام به انتخاب عمل، دریافت پاداش بر اساس اعتماد سطح گره و سطح پیوند ارتباطی و به‌روزرسانی مقدار Q می‌کند. سیاست بهینه درنهایت بر اساس بیشترین مقدار Q انتخاب می‌شود و الگوریتم می‌تواند به‌صورت آنلاین در زمان اجرا برای بهینه‌سازی مسیریابی به‌روز شود. درنتیجه روش پیشنهادی برای حل مسائل مسیریابی در زمان اجرا مناسب است، زیرا سیاست آن به‌صورت پویا و بر اساس تصمیمات اخیر به‌روزرسانی می‌شود.

۵- ارزیابی کارایی

ما کارایی پروتکل پیشنهادی را با ادغام آن در چارچوب مسیریابی RPL از طریق مطالعه شبیه‌سازی دقیق ارزیابی کردیم. این بخش عملکرد پروتکل پیشنهادی را با پروتکل‌های موجود مقایسه می‌کند. شبیه‌سازی‌ها در پلتفرم سبک و متن‌باز Contiki-NG/Cooja انجام‌شده است که برای ارزیابی شبکه‌های اینترنت اشیا مناسب است.

شبیه‌سازی شامل ۳۰ گره است که در قالب ۵ سناریوی مختلف، به‌صورت ثابت در یک محیط مربعی با ابعاد ۱۰۰×۱۰۰ مترمربع قرارگرفته‌اند. در هر سناریو، یک گره ریشه (سرور) در مرکز شبکه مستقرشده و سایر گره‌های کلاینت به‌صورت تصادفی در اطراف آن توزیع‌شده‌اند. برای پیاده‌سازی سخت‌افزاری گره‌ها، از پلتفرم Z1 (Zolertia)مبتنی بر میکروکنترلر کم‌مصرف MSP430 استفاده‌شده است که با فرکانس ۲,۴ گیگاهرتز عمل می‌کند [۳۶].

در هر سناریو، ۱۰ درصد از گره‌ها به‌عنوان گره مهاجم انتخاب‌شده و در موقعیت مشخصی برای اجرای حمله کاهش رتبه قرار داده‌شده‌اند. هدف اصلی از شبیه‌سازی، بهبود قابلیت اطمینان مسیر از طریق انتخاب گره‌های مورد اعتماد به‌عنوان والد بوده است. عملکرد روش پیشنهادی در مقایسه با سه روش مرجع MRHOF-RPL[11]، SecTrust-RPL[7] و SMTrust[30]، بر اساس چندین معیار عملکردی ازجمله نرخ تحویل بسته‌ها، مصرف انرژی، سربار کنترلی، نرخ تشخیص حمله و توان عملیاتی شبکه مورد ارزیابی قرارگرفته است.

به‌منظور بررسی تأثیر تداخل سیگنال و اتلاف بسته‌ها در سمت گیرنده، شعاع ارسال گره‌ها ۳۰ متر و شعاع تداخل ۳۵ متر در نظر گرفته‌شده است. برای مدل‌سازی شرایط محیطی همراه با نویز، چهار نرخ دریافت ۳۰، ۵۰، ۷۰ و ۹۰ درصد لحاظ شده است که معادل با نرخ اتلاف بسته ۷۰، ۵۰، ۳۰ و ۱۰ درصد در سمت گیرنده است. در هر سناریو، برای هر نرخ دریافت، آزمایش‌ها با ۲۰ مقدار متفاوت از دانه تصادفی (Seed Number) انجام‌شده‌اند که درمجموع، ۱۰۰ اجرای مستقل برای هر مدل صورت گرفته است. در تمام اجراها، توپولوژی گره‌ها و موقعیت گره‌های مهاجم ثابت باقی‌مانده‌اند و تنها تفاوت، مربوط به مقدار دانه تصادفی بوده که باعث ایجاد تغییر در زمان‌بندی‌های داخلی سیستم و رفتارهای تصادفی گره‌ها شده است. این رویکرد امکان تحلیل دقیق رفتار مدل پیشنهادی در شرایط متنوع را فراهم کرده و تکرارپذیری نتایج را تضمین می‌کند.

مدت‌زمان اجرای هر آزمایش ۶۰ دقیقه در نظر گرفته‌شده و در این مدت، هر گره در هر دقیقه یک بسته داده ارسال کرده است. برای تحلیل نتایج، میانگین شاخص‌ها در تمام اجراها محاسبه‌شده و با استفاده از فاصله اطمینان ۹۵ درصد، در نمودارها نمایش داده‌شده‌اند. ازآنجاکه محیط Contiki-NG ابزارهای تحلیل مستقیم لاگ‌ها را فراهم نمی‌کند، اسکریپت‌های پردازش لاگ با زبان Python 3.11توسعه داده‌شده‌اند. این اسکریپت‌ها، خروجی لاگ گره‌ها را خط به خط پردازش کرده و مقادیر دقیق شاخص‌های عملکردی را استخراج می‌نمایند.

با توجه به استفاده از چندین سناریو، تنوع نرخ دریافت، تکرار آزمایش‌ها با مقادیر مختلف تصادفی و در نظر گرفتن شرایط واقعی مانند نویز و تداخل، می‌توان نتایج این شبیه‌سازی را با اطمینان نسبی به شبکه‌های واقعی اینترنت اشیا تعمیم داد. همچنین استفاده از سخت‌افزارهای رایج و معتبر مانند Zolertia Z1 که در پروژه‌های اینترنت اشیا واقعی کاربرد دارند، اعتبار بیشتری به نتایج این شبیه‌سازی می‌بخشد. شکل ۱ ، بخش پیوست نشان‌دهنده محیط شبیه‌سازی در حال اجرا از سناریو شماره ۵ است.

در بخش ۵-۱ نحوه محاسبه معیارهای کارایی را شرح می‌دهیم، در بخش ۵-۲ نتایج را به‌طور مفصل بررسی کرده و بهبودهای

عملکردی حاصل از پروتکل پیشنهادی را نشان می‌دهیم. درنهایت در بخش ۵-۳ به بحث و بررسی پروتکل پیشنهادی و مزایای روش پیشنهادی و موازنه بین امنیت و کارایی روش پیشنهادی می‌پردازیم.

۵-۱- معیارهای عملکرد جهت ارزیابی روش پیشنهادی

در این بخش، شش معیار اساسی برای ارزیابی عملکرد RPL با استفاده از چهار تابع هدف مختلفMRHOF ،SecTrust-RPL، SMTrust و رهیافت پیشنهادی ارائه می‌شود. هدف این معیارها شناسایی و جداسازی گره‌های مخرب و ایجاد یک مسیر مطمئن است. ازآنجاکه کارایی RPL تحت تأثیر توپولوژی شبکه و تابع هدف قرار دارد، این معیارها تأثیر مستقیمی بر فرآیند مسیریابی دارند. عملکردRPL با استفاده از معیارهای زیر ارزیابی‌شده است:

- **ترافیک سربار پیام‌های کنترلی**:

پروتکل RPL از سه نوع پیام کنترلی اصلیDIO ، DAO، DIS استفاده می‌کند. میزان کل ترافیک کنترلی سربار با جمع کل پیام‌های کنترلی ارسالی از گره‌های کلاینت به گره ریشه یا سینک اندازه‌گیری می‌شود[۳۷]. این معیار توسط معادلات (۲۱ و ۲۲) محاسبه می‌شود. درواقع این پارامتر نسبت پیام‌های کنترلی به تعداد کل پیام‌های ارسالی در شبکه است.

$$Total\ control\ packets = \sum_{1}^{n} DIS + \sum_{1}^{n} DIO + \sum_{1}^{n} DAO \quad (۲۱)$$

$$Control\ traffic\ overhead = \frac{Total\ control\ packets}{Total\ UDP_sent\ packet + Total\ control\ packets} \quad (۲۲)$$

- **نرخ تحویل بسته(PDR)**:

این معیار مهم، نسبت بسته‌های دریافتی موفق در گره مقصد یا ریشه به‌کل بسته‌های ارسالی از همه گره‌های کلاینت را ارزیابی می‌کند. مقدار PDR بالاتر نشان‌دهنده‌ی عملکرد بهتر پروتکل مسیریابی است. این معیار توسط معادله (۲۳) محاسبه می‌شود.

$$PDR = \frac{Total\ Packet\ Recived}{Total\ Packet\ Sent} \times 100 \quad (۲۳)$$

- **مصرف انرژی:**

این معیار در شبکه‌های با تلفات بالا اهمیت ویژه‌ای دارد، زیرا بر عمر عملیاتی شبکه‌های اینترنت اشیا تأثیر می‌گذارد. میزان مصرف انرژی در حین اجرای پروتکل با استفاده از ماژول Energest در Contiki-NG محاسبه می‌شود. برای تعیین مقدار دقیق مصرف انرژی، از داده برگ یا دیتاشیت سخت‌افزار Zolertia Z1 استفاده کردیم[۳۶]. بنابراین کل انرژی مصرفی به‌صورت معادله (۲۴) محاسبه می‌شود.

$$Total_Energy = E_{CPU} + E_{LPM} + E_{RX} + E_{TX} \quad (۲۴)$$

$$= V \times \frac{T_{CPU} \times 10 + T_{LPM} \times 0.023 + T_{RX} \times 18.8 + T_{TX} \times 17.4}{R_{timer}}$$

درواقع انرژی کل مصرفی، برابر با مجموع انرژی که CPU در حالت فعال، در حالت کم‌مصرف، گره در حال ارسال و دریافت داده مصرف می‌کند؛ بنابراین انرژی مصرفی از حاصل‌ضرب مدت‌زمانی که صرف این امور می‌شود در جریانی که در هر مرحله مصرف می‌شود محاسبه می‌شود. در این معادله جریان مصرفی درزمانی که CPU در حالت فعال و در حالت کم‌مصرف است به ترتیب معادل ۱۰ میلی‌آمپر و ۲۳ میکرو آمپر است. جریان مصرفی رادیو در حالت دریافت و ارسال بسته نیز به ترتیب معادل ۱۸,۸ و ۱۷,۴ میلی‌آمپر است. همچنین ولتاژ گره Z1 برابر با ۳ ولت است و مقدارR_{timer} نیز ۳۲۷۶۸ تیک در هر ثانیه می‌باشد.

- **توان عملیاتی شبکه**(Throughput):

این معیار، نشان‌دهنده‌ی تعداد کل بیت‌هایی است که در هر ثانیه به گره مقصد یا ریشه با موفقیت تحویل داده‌شده‌اند. توان عملیاتی شبکه از طریق معادله‌ی (۲۵) محاسبه می‌شود که مبتنی بر ضرب تعداد کل بسته‌های تحویل داده‌شده، اندازه‌ی بسته (۱۲۸ بایت) و مقدار ۸ (برای تبدیل بایت به بیت) و تقسیم آن بر کل مدت‌زمان شبیه‌سازی (۳۶۰۰ ثانیه) است.

$$Throughput = \frac{Total\ delivered\ packets \times packet_size \times 8}{3600} \quad (۲۵)$$

- **نرخ شناسایی حمله:**

این معیار، توانایی مدل در شناسایی و جداسازی مهاجمان در فرآیند مسیریابی را اندازه‌گیری می‌کند. نرخ شناسایی حمله، تعداد گره‌های مخربی را که در یک بازه‌ی زمانی مشخص شناسایی‌شده‌اند را محاسبه می‌کند.

۵-۲- نتایج شبیه‌سازی

در این بخش، نتایج شبیه‌سازی در چندین سناریو با توپولوژی‌های مختلف شبکه شامل ۲۹ گره کلاینت و یک گره ریشه مورد تحلیل و مقایسه قرارگرفته است. هدف اصلی این ارزیابی، بررسی عملکرد پروتکل پیشنهادی و مقایسه آن با MRHOF-RPL و SecTrust-RPL [7] و SMTrust[30] است تا مزایای روش پیشنهادی نشان داده شود. در ادامه به ارزیابی معیارهای کارایی و مقایسه آن‌ها در سه مدل مختلف می‌پردازیم. جدول ۲ شامل پارامترهای مربوط به محیط شبیه‌سازی است.

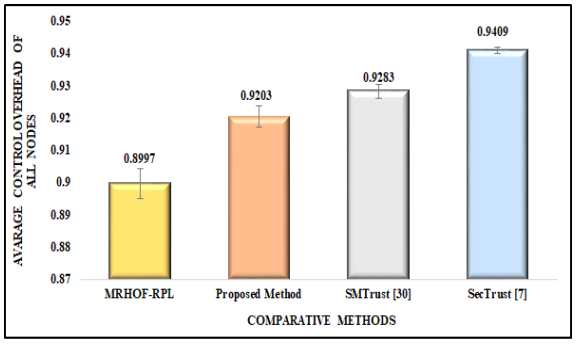
• ترافیک سربار پیام‌های کنترلی

پروتکل RPL از پیام‌های کنترلی ICMPv6 شامل DIO، DIS و DAO برای تبادل اطلاعات میان گره‌ها استفاده می‌کند. افزایش تعداد این پیام‌ها منجر به افزایش سربار شبکه می‌شود. در زمان تشکیلDODAG ، گره ریشه پیام DIO را پخش می‌کند، گره‌های کلاینت با پیام DAO پاسخ داده و گره‌هایی که هنوز عضوی از DODAG نیستند، پیام DIS ارسال می‌کنند تا درخواست عضویت دهند.

همان‌طور که در شکل ۲ قابل‌مشاهده است، روش پیشنهادی نسبت به نسبت به SMTrust و SecTrust-RPLبه‌طورکلی سربار کمتری دارد؛ اما روش پیشنهادی و SecTrust-RPL وSMTrust هر سه نسبت به MRHOF سربار بیشتری را تجربه می‌کنند و این سربار اضافی، ناشی از به‌روزرسانی معیار اعتماد، تغییرات والدین و تغییرات رتبه است که موجب ارسال بیشتر پیام‌های DIO می‌شود.

جدول ۲: پارامترها و مقادیر شبیه‌سازی شبکه

پارامترهای شبیه‌سازی	مقادیر	پارامترهای شبیه‌سازی	مقادیر
برد یا محدوده ارسال سیگنال	۳۰ متر	ابزار شبیه‌سازی	Contiki-NG/cooja
برد یا محدوده تداخل سیگنال	۳۵ متر	نرم‌افزار آنالیز فایل‌های متنی لاگ های خروجی	Python
تأخیر اولیه جهت راه‌اندازی شبکه	۵ ثانیه	مدل تجهیز (موت) مورداستفاده	Z1 Zolertia
مدل خرابی پیوند ارتباطی	UDGM with distance	محدوده شبیه‌سازی	۱۰۰ متر در ۱۰۰ متر
زمان شبیه‌سازی	۶۰ دقیقه	پروتکل لایه انتقال	UDP
تعداد سناریوها	۵ سناریو	پروتکل لایه شبکه	IPv6/RPL
تعداد اجرای هر سناریو	۲۰ بار	لایه سازگاری یا انطباق	6LoWPAN
αنرخ یادگیری	۰/۲	پروتکل لایه پیوند داده	CSMA
γنرخ تنزیل	۰/۹	پروتکل لایه فیزیکی	IEEE 802.15.4
میزان اعتماد آستانه	۶۲	سایز بسته اطلاعات	۱۲۸ بایت
فاصله اطمینان	۰/۹۵	تعداد گره‌های مهاجم	۱۰ درصد تعداد گره‌های شبکه
نرخ ارسال بسته‌ها در هر گره	۱۰۰ درصد	نرخ دریافت بسته‌ها در هر گره	۳۰، ۵۰، ۷۰ و ۹۰ درصد



شکل ۲: سربار پیام‌های کنترلی در برابر حمله کاهش رتبه

درروش پیشنهادی در مرحله یادگیری اولیه، تعداد تغییرات والدین بیشتر است، اما با پیشرفت یادگیری، گره‌ها بیشتر به بهره‌برداری متکی می‌شوند و تغییرات والد کاهش می‌یابد، درنتیجه سربار کنترلی روش پیشنهادی در مقایسه با SecTrust-RPL کمتر شده و کارایی بهتری در شبکه‌هایی با مقیاس وسیع ارائه می‌دهد. از سوی دیگر روش SMTrust نسبت به روش پیشنهادی سربار بالاتری را تجربه می‌کند؛ زیرا برای به‌روزرسانی مداوم مقادیر اعتماد نیازمند تبادل مکرر پیام‌های توصیه بین گره‌های همسایه است. این فرآیند باعث افزایش تعداد پیام‌های کنترلی در شبکه می‌شود، به‌ویژه در شبکه‌هایی با تغییرات زیاد یا حضور گره‌های مهاجم. ازآنجایی‌که روش پیشنهادی تنها بر اعتماد مستقیم متکی است درنتیجه بدون نیاز به انتشار توصیه‌های اضافی و با استفاده از داده‌های محلی، فرآیند ارزیابی اعتماد را انجام می‌دهد. به همین دلیل، حجم پیام‌های کنترلی کاهش‌یافته و سربار کنترلی شبکه در مقایسه با SMTrustبه‌طور چشم‌گیری کمتر است.

• نرخ تحویل بسته

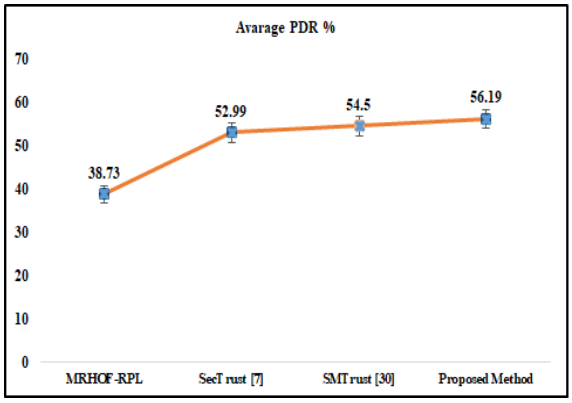
نرخ تحویل بسته (PDR) یکی از معیارهای اساسی برای ارزیابی عملکرد شبکه است و در سامانه‌های مقابله با شرایط اضطراری،

دستیابی به PDR بالا نقش حیاتی در تضمین ارتباطات قابل‌اعتماد دارد. این معیار نشان‌دهنده‌ی درصد بسته‌هایی است که با موفقیت در شبکه منتقل‌شده‌اند و عملکرد انتها به انتهای پروتکل‌های مسیریابی را مشخص می‌کند. برای نشان دادن عملکرد مسیریابی برتر، PDR روش پیشنهادی باید بالاتر از روش‌های موجود باشد.

در این تحلیل، عملکرد PDR درروش پیشنهادی با روش‌های دیگر مقایسه شده است. همان‌طور که در شکل ۳ نشان داده‌شده است، روش پیشنهادی نسبت به و [7]SecTrust-RPL ، MRHOF و SMTrust عملکرد بهتری دارد. این برتری به دلیل استفاده از چندین معیار اعتماد مانند ارزیابی اعتماد سطح ارتباطات گره‌ها برای هدایت بسته‌ها، اعتماد پیوندهای ارتباطی و انرژی باقی‌مانده است. علاوه بر این، روش پیشنهادی گره‌های مهاجم را شناسایی و حذف می‌کند و باعث می‌شود فقط گره‌های قابل‌اعتماد در فرآیند مسیریابی شرکت کنند که درنهایت منجر به افزایش PDR درروش پیشنهادی می‌شود.

در مقابل، [7] SecTrust-RPL تنها بر اساس یک معیار یعنی نسبت هدایت بسته‌ها توسط گره انتخاب‌شده به‌عنوان والد، میزان اعتماد را ارزیابی می‌کند که در مقایسه با مدل پیشنهادی، دقت پایین‌تری دارد؛ همچنین روش پیشنهادی با تمرکز بر شناسایی دقیق گره‌های غیرقابل‌اعتماد و جلوگیری از مشارکت آن‌ها در فرآیند مسیریابی، موجب کاهش ارسال بسته‌ها از مسیرهای مخرب یا ناپایدار شده و از گم‌شدن یا تخریب بسته‌ها جلوگیری می‌کند. برخلاف SMTrust که بخشی از اعتماد خود را از توصیه گره‌های همسایه دریافت می‌کند (که ممکن است تحت تأثیر گره‌های مهاجم قرارگرفته باشند)، مدل پیشنهادی تنها بر معیارهای قابل‌اعتماد محلی تکیه دارد. درنتیجه، مسیرهای انتخاب‌شده درروش پیشنهادی پایدارتر و امن‌تر هستند که این موضوع منجر به افزایش نرخ تحویل بسته‌ها نسبت به SMTrust می‌شود.

بنابراین، می‌توان نتیجه گرفت که روش پیشنهادی بالاترین میزان تحویل بسته را در میان این چهار مدل دارد، پس‌ازآن به ترتیب SMTrust،SecTrust-RPL و درنهایت MRHOF قرار می‌گیرد. در MRHOF میزان PDRبه‌طور قابل‌توجهی پایین‌تر از سه مدل دیگر است که عمدتاً به دلیل ازدحام شبکه و برخورد بسته‌ها رخ می‌دهد و منجر به افزایش نرخ از دست رفتن بسته‌ها می‌شود. علاوه بر این، مهاجمان موجود در شبکه نیز عملکرد آن را بیشتر کاهش می‌دهند، زیرا استانداردMRHOF فاقد راهکار امنیتی برای شناسایی و جداسازی مهاجمان است.



شکل ۳: نرخ تحویل بسته در برابر حمله کاهش رتبه

• مصرف انرژی

در این بخش، مصرف انرژی پروتکل پیشنهادی با سایر پروتکل‌ها مقایسه شده است. مدیریت بهینه‌ی مصرف انرژی در شبکه‌های اینترنت اشیا اهمیت بالایی دارد، زیرا این دستگاه‌ها معمولاً دارای منابع باتری محدودی هستند. درروش پیشنهادی، به دلیل پردازش‌های اضافی و تبادل پیام‌هایDIO ، مصرف انرژی در ابتدا افزایش می‌یابد، اما پس از شناسایی و حذف گره‌های مخرب، شبکه به ثبات رسیده و مصرف انرژی کاهش پیدا می‌کند. علاوه بر این، اگر دو گره والد دارای مقدار اعتماد برابر باشند، گره‌ای که انرژی باقی‌مانده‌ی بیشتری داشته باشد یا پیوند ارتباطی قابل‌اعتمادتری ارائه دهد، در اولویت انتخاب قرار می‌گیرد.

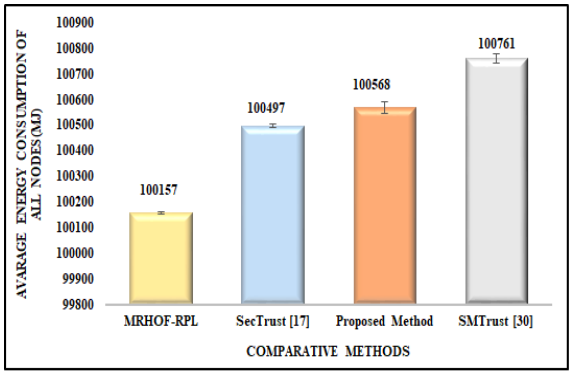
Contiki-NG شامل یک ماژول Energest برای برآورد مصرف انرژی است که امکان توسعه‌ی یک روش سبک و نرم‌افزاری را برای دستگاه‌های اینترنت اشیا با محدودیت انرژی فراهم می‌کند. با فعال‌سازی این ماژول در فایلMakefile از مسیریابی RPL می‌توان اطلاعات مصرف انرژی هر گره را در بازه‌های زمانی ۶۰ ثانیه‌ای ثبت کرد.

در این مقاله، از ماژول Energest در Contiki-NG برای محاسبه‌ی مصرف انرژی استفاده‌شده و مقدار آن برای همه سناریوها مطابق معادله (۲۴) محاسبه‌شده است. سپس، میانگین مصرف انرژی برای کل سناریوها تعیین‌شده است. همان‌طور که در شکل ۴ مشاهده می‌شود، بخشی از مصرف انرژی درروش پیشنهادی و-SecTrust RPLبه دلیل نیاز به فعال‌سازی حالت شنود جهت نظارت بر ارسال بسته‌ها توسط گره‌های همسایه است. در مقابل، پروتکل MRHOF این قابلیت را فعال نمی‌کند. بااین‌حال، به دلیل عدم وجود راهکارهای کاهش حمله، این پروتکل با افزایش ازدحام شبکه و از دست رفتن بسته‌ها و ارسال مجدد مواجه می‌شود که منجر به

مصرف بالای انرژی می‌شود.

در حالت کلی روش پیشنهادی به‌طور میانگین مصرف انرژی بیشتری نسبت به SecTrust-RPL دارد که این مقدار بسیار ناچیز است؛ زیرا در هر مرحله، عامل تصمیم‌گیرنده گره‌هایی را با بالاترین سطح اعتماد برای مسیریابی انتخاب کرده و از ارسال بسته‌ها به گره‌های غیرقابل‌اعتماد جلوگیری می‌کند. این امر به بهینه‌سازی مصرف انرژی و کاهش مصرف کلی شبکه کمک می‌کند؛ اما به‌هرحال محاسباتی که برای اعتماد درروش پیشنهادی در نظر گرفته می‌شود نسبت به SecTrust-RPL بیشتر است؛ زیرا درروش پیشنهادی عامل پاداشی را که دریافت می‌کند بر اساس ارزیابی اعتماد سطح گره (که بر اساس اعتماد ارتباطات و انرژی باقیمانده) و اعتماد سطح پیوند ارتباطی است؛ اما درروش SecTrust-RPL تنها از یک معیار برای محاسبه اعتماد بهره می‌برد. نتایج حاکی از این است که مدل SMTrust بالاترین نرخ میزان انرژی مصرفی را دارد زیرا محاسبه اعتماد نیازمند بررسی اعتماد گره‌های همسایه و محاسبه اعتماد غیرمستقیم و انتقال این مقدار محاسبه‌شده به گره در حال ارزیابی است؛ بنابراین مدل SMTrust به دلیل استفاده از چندین مؤلفه برای محاسبه اعتماد ازجمله انرژی باقی‌مانده، مقدار RSSI، توصیه‌های گره‌های همسایه، اعتماد قبلی و اعتماد جاری، نیازمند انجام محاسبات سنگین‌تر و تبادل پیام‌های بیشتری بین گره‌ها است همین موضوع منجر به پیچیدگی پردازشی بیشتر و درنتیجه مصرف انرژی بالاتر شده است. این مسئله به‌ویژه در شبکه‌های با تراکم بالا یا با تعداد زیاد گره‌های مهاجم منجر به افزایش چشمگیر مصرف انرژی می‌شود. در مقابل، روش پیشنهادی با بهره‌گیری از معیارهای محلی و بدون وابستگی به توصیه‌های دیگر گره‌ها، پیچیدگی پردازش را کاهش داده و تعداد تبادل پیام‌ها را محدود کرده است. این ویژگی باعث می‌شود مصرف انرژی درروش پیشنهادی به‌طور محسوسی کمتر از SMTrust باشد و درنتیجه، طول عمر شبکه نیز افزایش یابد.

همان‌طور که در شکل ۴ نشان داده‌شده است مجموع مصرف انرژی درروش پیشنهادی برابر با ۱۰۰۵۶۸ میلی ژول است که اندکی بیشتر از SecTrust-RPLکه ۱۰۰۴۹۷ میلی ژول و MRHOFکه ۱۰۱۵۷ میلی ژول است. همچنین میزان انرژی مصرفی درروش SMTrust برابر ۱۰۰۷۶۱ میلی ژول است که در حد ۱۹۳ میلی ژول از روش پیشنهادی بیشتر است. افزایش جزئی مصرف انرژی مربوط به روش پیشنهادی، به دلیل افزایش امنیت و قابلیت اطمینان شبکه درروش پیشنهادی قابل توجیه است. بااین‌حال، این مقدار ناچیز بوده و روش پیشنهادی گزینه‌ی مناسبی برای شبکه‌های اینترنت اشیا با اولویت امنیت محسوب می‌شود.



شکل ۴: میزان مصرف انرژی در برابر حمله کاهش رتبه

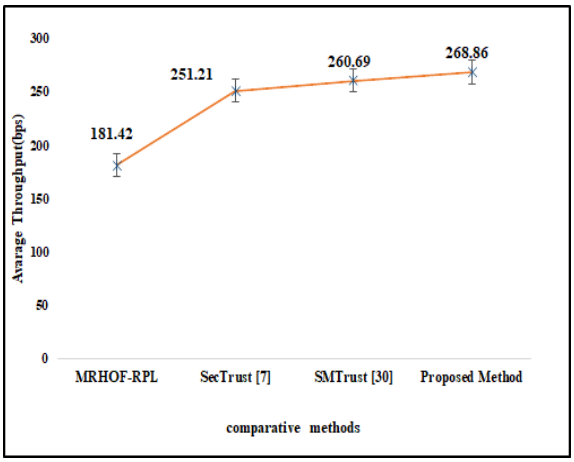
به‌صورت خلاصه، برخلاف MRHOF که فاقد راهکار کاهش حملات است و به دلیل ارسال‌های مکرر، انرژی زیادی هدر می‌دهد، روش پیشنهادی به‌صورت پویا، والدین قابل‌اعتماد را انتخاب می‌کند و از اتلاف انرژی ناشی از دخالت‌های گره‌های مخرب جلوگیری می‌کند. با کاهش ارسال‌های مجدد و جلوگیری از اختلال در شبکه، مصرف انرژی را بهینه می‌کند. در مقابل، SecTrust-RPLتنها از نسبت ارسال موفق بسته‌ها به‌عنوان معیار اعتماد استفاده می‌کند و عواملی مانند کیفیت پیوند ارتباطی و انرژی باقی‌مانده را در نظر نمی‌گیرد. این موضوع باعث کاهش اثربخشی آن در شناسایی حملات و درنتیجه، مصرف انرژی کمی پایین‌تر از روش پیشنهادی می‌شود؛ اما روش پیشنهادی با ارزیابی چندین معیار ازجمله کیفیت پیوند ارتباطی، میزان انرژی و رفتار گره‌ها در ارسال بسته‌ها، یک مدل جامع‌تر از اعتماد را ارائه می‌دهد که امنیت مسیریابی را تقویت می‌کند.

- **توان عملیاتی شبکه**

توان عملیاتی شبکه که نشان‌دهنده تعداد کل بسته‌های داده‌ای است که در واحد زمان با موفقیت ارسال می‌شوند، برحسب بیت بر ثانیه (bps) اندازه‌گیری می‌شود. در این مجموعه آزمایش‌ها، نرخ ارسال داده ثابت است و در هر دقیقه یک بسته داده ارسال می‌شود. شکل ۵ نشان می‌دهد که روش پیشنهادی توان عملیاتی بالاتری نسبت به سایر مدل‌ها دارد. این عملکرد به‌ویژه در حضور گره‌های مهاجم برجسته‌تر است، زیرا مدل پیشنهادی با بهره‌گیری از ارزیابی اعتماد در سطح گره و پیوند ارتباطی، به انتخاب گره‌های والد قابل‌اعتمادتر اولویت می‌دهد. راهکار پاداش‌دهی در مدل یادگیری تقویتی پیشنهادی موجب می‌شود گره‌ها به‌صورت خودکار مسیرهایی را انتخاب کنند که به افزایش قابلیت اطمینان شبکه منجر می‌شود. همچنین، اعمال جریمه شدید برای انتخاب گره‌های مخرب، شناسایی آن‌ها را تسهیل کرده و موجب حذف مؤثر آن‌ها از

فرآیند مسیریابی می‌گردد. درنتیجه، مدل پیشنهادی با اجتناب از گره‌های ناسالم، مسیرهای امن‌تری را ایجاد کرده و میانگین توان عملیاتی شبکه را به‌طور محسوسی افزایش می‌دهد.

در مقابل، مدل MRHOF به دلیل فقدان راهکار تشخیص گره‌های مهاجم، دچار افت توان عملیاتی می‌شود. این مدل فاقد معیارهای اعتماد است. بنابراین، در حضور گره‌های مخرب، بسته‌های بیشتری قبل از رسیدن به گره ریشه از بین می‌روند که منجر به کاهش نرخ تحویل بسته‌ها (PDR) و درنتیجه کاهش توان عملیاتی کل شبکه می‌گردد.



شکل ۵: میزان توان عملیاتی شبکه در برابر حمله کاهش رتبه

همچنین، مدل SMTrust علی‌رغم بهره‌گیری از معیارهای متنوعی مانندRSSI ، انرژی باقی‌مانده و توصیه همسایگان، به دلیل وابستگی بالا به تبادل پیام‌های کنترلی برای محاسبه اعتماد، با افزایش تأخیر در به‌روزرسانی مسیرها مواجه است. این وابستگی باعث ایجاد سربار کنترلی بیشتر، افزایش پیچیدگی پردازش و کاهش واکنش‌پذیری گره‌ها می‌شود که درنهایت توان عملیاتی شبکه را تحت تأثیر قرار می‌دهد. انتخاب مسیرهای ناپایدار به دلیل اتکای بیش‌ازحد به اعتماد توصیه‌شده نیز از دیگر نقاط ضعف SMTrust است.

مدلSecTrust-RPL تنها از یک معیار برای ارزیابی اعتماد استفاده می‌کند که آن نیز نسبت تعداد بسته‌های هدایت‌شده توسط هر گره به‌عنوان والد است. این رویکرد ساده‌سازی شده موجب می‌شود که مدل در محیط‌های پر نویز، متغیر یا با حضور گره‌های مهاجم نتواند ارزیابی دقیقی از رفتار واقعی گره‌ها ارائه دهد. درنتیجه، گره‌های مهاجم ممکن است به‌اشتباه به‌عنوان گره‌های والد انتخاب شوند. این انتخاب‌های نادرست منجر به اتلاف بسته‌های داده، تغییرات متناوب در مسیرها و افزایش ازدحام یا اختلال در شبکه می‌شود که همگی توان عملیاتی شبکه را به‌طور محسوس کاهش می‌دهند.

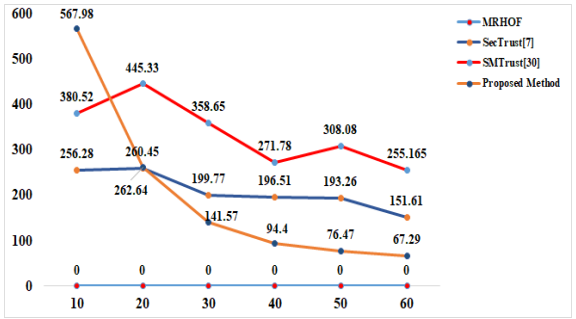
در مقابل، مسیرهای انتخابی در مدل SMTrust نسبت به SecTrust به دلیل ترکیب چندین شاخص برای محاسبه اعتماد، پایداری و قابلیت اطمینان بیشتری دارند. هرچند SMTrust سربار بیشتری دارد، اما در شرایطی که مهاجم‌ها حضور دارند، انتخاب‌های هوشمندتر آن باعث حفظ نسبی توان عملیاتی می‌شود.

درمجموع، روش پیشنهادی با بهره‌گیری از ساختار یادگیری تقویتی، پردازش سریع محلی و حداقل وابستگی به تبادل پیام‌های کنترلی، مسیرهای پایدارتر و سریع‌تری را ایجاد می‌کند و درنتیجه، توان عملیاتی بالاتری را در شرایط مختلف شبکه، ازجمله در سناریوهای پر تداخل یا با حضور گره‌های مهاجم، ارائه می‌دهد.

- **نرخ شناسایی حمله**

در این بخش، توانایی مدل پیشنهادی در شناسایی و ایزوله سازی گره‌های مهاجم مورد ارزیابی قرارگرفته و با روش‌های موجود شامل SecTrust-RPL، SMTrustو MRHOF مقایسه شده است. برای یکسان‌سازی معیار شناسایی مهاجمان، از الگوی تعریف‌شده در شبه کد [7] SecTrust-RPL استفاده‌شده است. شکل ۶ نشان می‌دهد که روش پیشنهادی در تشخیص حملات مبتنی بر کاهش رتبه به‌طور مؤثر عمل می‌کند. این مدل با تخصیص پاداش منفی چشمگیر به رفتارهای مشکوک، روند شناسایی را سرعت می‌بخشد.

در مراحل ابتدایی عملکرد شبکه، روش پیشنهادی با بهره‌گیری از استراتژی اکتشاف در یادگیری تقویتی، نرخ کشف حمله بالایی دارد؛ به‌طوری‌که طی ۱۰ دقیقه نخست، به‌طور میانگین ۵۶۷ حمله را شناسایی کرده است. این نرخ بالا ناشی از پایش فعال رفتار گره‌ها و ارزیابی هم‌زمان اعتماد در دو سطح گره و پیوند ارتباطی است. باگذشت زمان و شناسایی و حذف تدریجی گره‌های مهاجم، از یک‌سو نیاز به شناسایی کاهش می‌یابد و از سوی دیگر، با کاهش



شکل ۶: نرخ شناسایی حملات کاهش رتبه

نرخ ارسال بسته‌های کنترلی DIO طبق الگوریتمTrickle ، تعداد فرصت‌های مشاهده نیز کاهش می‌یابد. به همین دلیل، بین دقیقه

۳۰ تا ۴۰، نرخ تشخیص به‌طور میانگین به ۹۴ مورد می‌رسد. این روند نزولی، برخلاف تصور ضعف عملکرد، نشانه‌ای از پایداری و تثبیت شبکه در حذف گره‌های مخرب است؛ زیرا با ایزوله شدن مهاجمان، دیگر نیازی به شناسایی مکرر آن‌ها وجود ندارد.

در مقایسه، SecTrust-RPLتنها ۲۵۶ حمله را در ده دقیقه اول شناسایی کرده است که تقریباً نیمی از نرخ روش پیشنهادی است. علاوه‌براین، نمودار این مدل نوسانات زیادی در نرخ تشخیص دارد که به استفاده صرف از اعتماد ایستا و وابستگی بالا به مشاهدات مستقیم از گره‌های همسایه برمی‌گردد. این نوسانات باعث کاهش پایداری در تصمیم‌گیری می‌شود.

مدل SMTrust نیز ازنظر دقت شناسایی عملکرد بالاتری نسبت به SecTrustدارد، زیرا از ترکیب چند معیار مانندRSSI ، انرژی باقی‌مانده و توصیه گره‌های همسایه برای ارزیابی اعتماد بهره می‌گیرد. بااین‌حال، به دلیل پردازش سنگین و تأخیر در ارزیابی ترکیبی، در محیط‌های پویا و مهاجم، پایداری شناسایی و زمان واکنش آن کاهش می‌یابد.

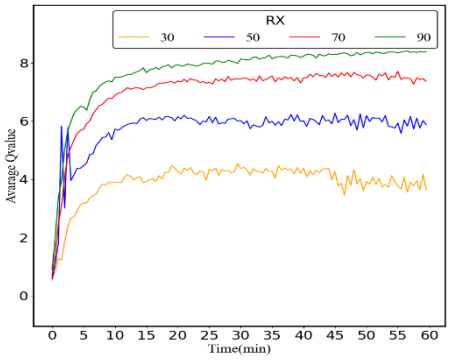
از سوی دیگر، MRHOF-RPL فاقد هرگونه راهکار برای شناسایی و ایزوله سازی حملات مبتنی بر کاهش رتبه بوده و توانایی مقابله با این تهدیدات را ندارد. همان‌طور که در شکل مشاهده می‌شود این استاندارد هیچ حمله‌ای را شناسایی نکرده است.

درنهایت، روش پیشنهادی با بهره‌گیری از ترکیب اکتشاف و بهره‌برداری در چارچوب یادگیری تقویتی، نه‌تنها گره‌های مخرب را در مراحل اولیه به‌خوبی شناسایی می‌کند، بلکه با یادگیری تدریجی، آن‌ها را از فرآیند مسیریابی حذف کرده و پایداری بیشتری نسبت به سایر روش‌ها ایجاد می‌کند؛ بنابراین، کاهش نرخ شناسایی در مدل پیشنهادی نشانه‌ای از ضعف نیست، بلکه حاکی از موفقیت آن در ایزوله سازی کامل مهاجمان و تثبیت وضعیت شبکه در بلندمدت است.

- **میزان همگرایی در تخمین یادگیری روش پیشنهادی**

الگوریتم یادگیری-Q یک پیشرفت مهم در یادگیری تقویتی محسوب می‌شود، زیرا همگرایی آن به سیاست بهینه تضمین‌شده است. همان‌طور که در [۳۸] نشان داده‌شده، این امر ثابت‌شده است که یادگیری- Q همواره راه‌حل بهینه را پیدا می‌کند. همگرایی یادگیری- Q نشان می‌دهد که گره‌های موجود در شبکه به‌طور مؤثر دانش لازم درباره محیط خود را کسب می‌کنند. شکل ۷ پایداری تخمین‌های یادگیری درروش پیشنهادی ما را در شبکه با نسبت نرخ دریافت‌های (RX Ratio)متفاوت در سناریوهای حمله کاهش رتبه را نشان می‌دهد.

در مراحل اولیه یادگیری، عامل‌ها همچنان در حال کاوش محیط هستند که این امر منجر به مقادیر Q نامنظم می‌شود و نشان می‌دهد که تابع ارزش هنوز بهینه نشده است. باگذشت زمان، گره‌ها از اطلاعات آموخته‌شده برای تصمیم‌گیری استفاده می‌کنند و به‌تدریج به سمت وضعیت‌هایی گرایش پیدا می‌کنند که پاداش مثبت بیشتری دارند. به‌طورکلی، شکل ۷ نشان می‌دهد که پس از تقریباً ده دقیقه، شبکه به پایداری رسیده و عملکرد بهینه می‌شود. همگرایی مقادیر Q تأیید می‌کند که یک راه‌حل بهینه برای محیط وجود دارد و درنتیجه شبکه می‌تواند به یک تعادل پایدار دست یابد.



شکل ۷: همگرایی تخمین یادگیری طرح پیشنهادی برای حمله رتبه برای نرخ دریافت‌های متفاوت

مقادیر نمایش داده‌شده در شکل ۷ میانگین مقادیر Q را در تمام گره‌های شبکه نشان می‌دهند. نتایج نشان می‌دهد که در نسبت RXبالاتر (نرخ دریافت بالاتر)، همگرایی نسبتاً سریع‌تر رخ می‌دهد. در مقابل، زمانی که نسبت RX برابر ۳۰ باشد، تخمین‌های یادگیری مقدار Q نوسانات بیشتری را نشان می‌دهند که به دلیل افزایش از دست رفتن بسته‌ها است. این موضوع باعث همگرایی کند تر می‌شود، زیرا گره‌ها به‌دفعات بیشتری بین حالت‌های مختلف جابه‌جا می‌شوند. این یافته‌ها منحنی یادگیری را از فاز اولیه‌ای نشان می‌دهند که الگوریتم بلافاصله پس از راه‌اندازی شبکه آغاز می‌شود. همگرایی سریع مقادیر Q نشان‌دهنده توانایی الگوریتم در سازگاری سریع با تغییرات سیستم پس از مرحله اولیه است.

۵-۳- بحث و بررسی

مدل پیشنهادی با انتخاب گره‌های قابل‌اعتماد برای مسیریابی امن، بهبود چشمگیری در نرخ تحویل بسته(PDR) ، توان عملیاتی شبکه و دقت شناسایی حملات را نشان می‌دهد، درحالی‌که سربار محاسباتی پایینی دارد. این روش تغییرات بیشتری در والدین گره‌ها

ایجاد می‌کند که نشان‌دهنده سازگاری بهتر در انتخاب مسیر است. افزایش مصرف انرژی جزئی و ناچیز بوده و عمدتاً ناشی از فعال‌سازی حالت شنود است. بااین‌حال، ارزیابی اعتماد گره‌ها و پیوندهای ارتباطی و به‌روزرسانی مقادیر Q درروش پیشنهادی، زمان پردازش بیشتری نسبت به MRHOF و SecTrust-RPL نیاز دارد، اما باعث افزایش امنیت و قابلیت اطمینان در مسیریابی می‌شود. درروش پیشنهادی یادگیری -Q به دلیل توانایی آن در یادگیری در محیط‌های بدون مدل و کارایی بالا برای اینترنت اشیا انتخاب‌شده است.

مدل پیشنهادی قابلیت ادغام در کاربردهای مختلف اینترنت اشیا مانند شهرهای هوشمند، اینترنت اشیا صنعتی و اینترنت اشیا پزشکی را دارد و می‌تواند امنیت و قابلیت اطمینان را در شبکه‌های کم‌مصرف اینترنت اشیا افزایش دهد. همچنین، با بهینه‌سازی‌های لازم، می‌توان آن را با سخت‌افزارهای کم‌مصرف سازگار کرده و در شبکه‌های مبتنی بر RPL ادغام کرد.

مزایای کلیدی مدل پیشنهادی عبارت‌اند از:

- تضمین مسیریابی امن در گره‌های اینترنت اشیا با منابع محدود و افزایش قابلیت اطمینان ارتباطات
- شناسایی و مقابله هوشمند با حملات در شبکه‌های مبتنی بر RPLبا استفاده از یادگیری تقویتی
- فراهم‌سازی ارتباطات ایمن و مطمئن در محیط‌های اینترنت اشیا
- امکان ادغام در شبکه‌های همتا به همتا توزیع‌شده متشکل از گره‌های محدود ازنظر منابع
- اثربخشی بالای روش پیشنهادی با توجه به عملکرد برتر نسبت به MRHOF SMTrust وSecTrust-RPL

۶- آنالیز و تجزیه‌وتحلیل آماری

آزمون تحلیل واریانس یک‌طرفه (ANOVA) در نرم‌افزار SPSS برای مقایسه میانگین یک متغیر کمی بین چندین گروه مستقل استفاده می‌شود. جدول ۱، بخش پیوست آمار توصیفی داده‌ها را در آزمون ANOVA که در فاصله اطمینان ۰,۹۵ انجام‌شده است، نشان می‌دهد. جدول ۳ یکی از زیر آزمون‌های ANOVA به نام آزمون لون[۱۶] را نشان می‌دهد که برای بررسی برابری یا نابرابری واریانس‌ها استفاده می‌شود. طبق جدول ۳ سطح معنی‌داری (.Sig) برای پارامترهای نرخ تحویل بسته و توان عملیاتی شبکه بالاتر از ۰,۰۵ است؛ بنابراین، برابری بین واریانس داده‌های حاصل از آزمایش روش‌های MRHOF، SecTrust-RPL، SMTrust و روش پیشنهادی برقرار است.

آزمون لون برای سایر معیارها (Sig < 0.001) نقض فرض همگنی واریانس‌ها را نشان داد که بیانگر تفاوت معنادار در واریانس بین مدل‌ها است.

جدول ۳: آزمون لون برای برابری یا نابرابری واریانس‌ها

		Levene Statistic	df1	df2	Sig.
PDR	Based on Mean	1.673	3	796	.171
	Based on Median	1.642	3	796	.178
	Based on Median and with adjusted df	1.642	3	788.993	.178
	Based on trimmed mean	1.673	3	796	.171
Throughput	Based on Mean	.272	3	796	.846
	Based on Median	.266	3	796	.850
	Based on Median and with adjusted df	.266	3	790.532	.850
	Based on trimmed mean	.272	3	796	.846
Overhead	Based on Mean	97.038	3	796	< .001
	Based on Median	95.942	3	796	< .001
	Based on Median and with adjusted df	95.942	3	466.527	< .001
	Based on trimmed mean	96.804	3	796	< .001
Total Energy	Based on Mean	110.862	3	796	< .001
	Based on Median	68.146	3	796	< .001
	Based on Median and with adjusted df	68.146	3	398.709	< .001
	Based on trimmed mean	97.793	3	796	< .001

جدول ۴ خروجی‌های نهایی اصلی آنالیز واریانس (ANOVA) را ارائه می‌دهد. ازآنجایی‌که سطح معنی‌داری از ۰,۰۵ کمتر است، واریانس داده‌های به‌دست‌آمده از روش پیشنهادی تفاوت معنی‌داری با روش‌های MRHOF، SecTrust-RPL، SMTrust دارد.

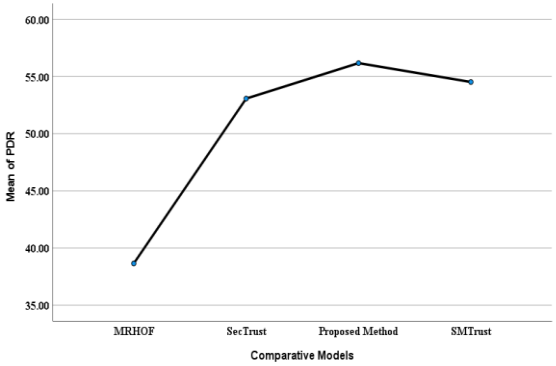
آزمون ANOVA صرفاً تعیین می‌کند که آیا تفاوت معناداری بین سه روش MRHOF، SecTrust-RPL، SMTrust و روش پیشنهادی وجود دارد یا خیر، اما هیچ اطلاعاتی در مورد مقادیر میانگین ارائه نمی‌دهد. برای بررسی این موضوع، باید از تحلیل Hoc Post [۳۹] نیز استفاده کنیم. این روش مقایسه‌های زوجی بین چهار روش را ارائه می‌دهد. آزمون‌های Post Hoc زیادی در آمار وجود دارد. در اینجا، ما از روش حداقل تفاوت معنادار فیشر (LSD) [۴۰] استفاده می‌کنیم.

جدول ۴: آنالیز واریانس یک‌طرفه

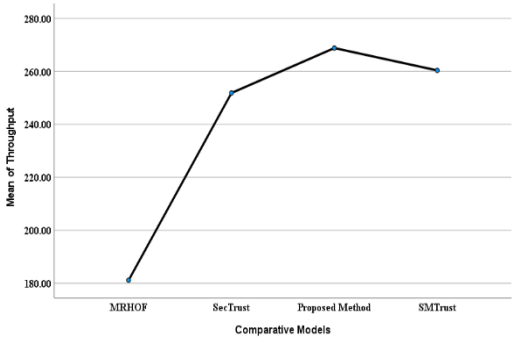
		Sum of Squares	df	Mean Square	F	Sig.
PDR	Between Groups	39040.622	3	13013.541	26.891	< .001
	Within Groups	385220.284	796	483.945		
	Total	424260.906	799			
Throughput	Between Groups	968897.178	3	322965.726	25.862	< .001
	Within Groups	9940445.744	796	12487.997		
	Total	10909342.922	799			
Overhead	Between Groups	.178	3	.059	59.003	< .001
	Within Groups	.802	796	.001		
	Total	.980	799			
Total Energy	Between Groups	37519625.844	3	12506541.948	563.156	< .001
	Within Groups	17677519.623	796	22207.939		
	Total	55197145.467	799			

شکل ۸ میانگین نرخ تحویل بسته از تحلیل LSD مربوط به چهار روش را نشان می‌دهد. همان‌طور که در شکل ۸ مشهود است، میانگین نرخ تحویل بسته روش پیشنهادی از روش‌های دیگر بیشتر است. شکل ۹ میانگین توان عملیاتی شبکه از تحلیل LSD مربوط

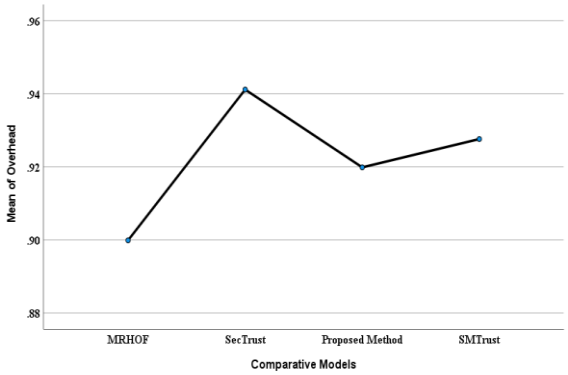
به چهار روش را نشان می‌دهد. بر اساس این شکل، میانگین توان عملیاتی شبکه روش پیشنهادی از روش‌های دیگر بیشتر است. شکل ۱۰ میانگین سربار پیام‌های کنترلی از تحلیل LSD مربوط به چهار روش را نشان می‌دهد. همان‌طور که در شکل ۱۰ نشان داده‌شده است، میانگین سربار کنترلی روش پیشنهادی از روش‌های SecTrust-RPL و SMTrust کمتر است.



شکل ۸: میانگین نرخ تحویل بسته از تحلیل LSD

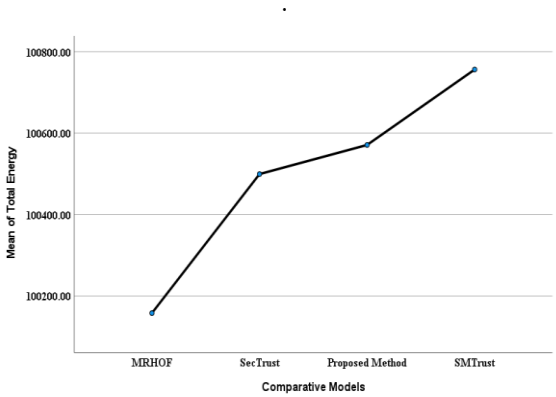


شکل ۹: میانگین توان عملیاتی شبکه از تحلیل LSD



شکل ۱۰: میانگین سربار پیام‌های کنترلی از تحلیل LSD

شکل ۱۱ میانگین مصرف انرژی از تحلیل LSD مربوط به چهار روش را نشان می‌دهد. همان‌طور که در شکل ۱۱ نشان داده‌شده است، میانگین مصرف انرژی روش پیشنهادی از روش SMTrust کمتر است



شکل ۱۱: میانگین مصرف انرژی از تحلیل LSD

۷- نتیجه‌گیری

این مقاله یک راهکار توزیع‌شده برای مسیریابی امن در شبکه‌های اینترنت اشیا ارائه داد که با ترکیب یادگیری تقویتی و راهکار اعتماد، امنیت مسیریابی را افزایش می‌دهد. تصمیم‌گیری در این روش بر اساس ارزیابی اعتماد در سطح پیوند ارتباطی و گره و استفاده از پاداش‌ها و مجازات محیطی انجام می‌شود.

نتایج شبیه‌سازی‌ها در برابر حمله کاهش رتبه نشان داد که روش پیشنهادی عملکرد بهتری نسبت به روش‌های موجود داشته است. این روش نرخ تحویل بسته (PDR) را در مقایسه با MRHOF تا ۱۷,۴۶ درصد، نسبت به SecTrust-RPL تا ۳,۲ درصد و نسبت به SMTrust تا ۱,۶۹ درصد افزایش داده است. همچنین توان عملیاتی شبکه بیشتر ازMRHOF ، SecTrust-RPL و SMTrust بود. علاوه بر این، روش پیشنهادی دقت بالاتری در شناسایی حملات داشته و سربار کنترلی کمتری نسبت بهSMTrust و SecTrust-RPLدارد. باوجود افزایش جزئی مصرف انرژی نسبت به SecTrust-RPL، این هزینه در برابر امنیت و پایداری بالاتر شبکه قابل توجیه است. این یافته‌ها نشان می‌دهد که روش پیشنهادی یک راهکار مؤثر برای بهبود امنیت مسیریابی در شبکه‌های اینترنت اشیا است.

مسیرهای آینده پژوهش شامل بررسی سایر حملات RPL و ارزیابی کارایی روش پیشنهادی در مقابله با آن‌ها است. بهبود دقت محاسبه اعتماد با استفاده از راهکارهای مبتنی بر حداقل توصیه های همسایگان و فیلتر کردن پیشنهادهای نامعتبر نیز می‌تواند دقت تصمیم‌گیری را افزایش دهد. علاوه بر این، آزمایش روش پیشنهادی در مقیاس بزرگ و شبکه‌های پویا به سنجش مقیاس‌پذیری و سازگاری آن کمک خواهد کرد. تطبیق این مدل برای محیط‌های متحرک اینترنت اشیا و بررسی تأثیر تحرک بر تصمیم‌گیری‌های

مسیریابی نیز از دیگر زمینه‌های مهم است. استفاده از یادگیری تقویتی پیشرفته مانند یادگیری Q دوگانه می‌تواند یادگیری آنلاین را بهبود داده و پیچیدگی محاسباتی را کاهش دهد. همچنین، یکپارچه‌سازی روش پیشنهادی با چارچوب‌های امنیتی موجود، امکان استقرار آن را در شبکه‌های متنوع اینترنت اشیا افزایش می‌دهد. درنهایت، بررسی یادگیری عمیق تقویتی می‌تواند به بهبود کارایی و مقیاس‌پذیری در کاربردهای گسترده‌تر اینترنت اشیا کمک کند.

مراجع

[1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surv. Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[2] I. Sadek, S. U. Rehman, J. Codjo, and B. Abdulrazak, "Privacy and security of IoT based healthcare systems: concerns, solutions, and recommendations," in *Proc. 17th Int. Conf. How AI Impacts Urban Living and Public Health (ICOST)*, New York City, NY, USA, Oct. 14–16, pp. 3–17, 2019.

[3] S. Görmüş, H. Aydın, and G. Ulutaş, "Security for the Internet of Things: A Survey of Existing Mechanisms, Protocols and Open Research Issues," *J. Fac. Eng. Archit. Gazi Univ.*, vol. 33, no. 4, pp. 1247–1272, 2018.

[4] T. Winter, P. Thubert, A. Brandt, J. Hui, R. Kelsey, P. Levis, K. Pister, R. Struik, J. P. Vasseur, and R. Alexander, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks," *RFC 6550*, Mar. 2012.

[5] I. S. Alsukayti and A. Singh, "A Lightweight Scheme for Mitigating RPL Version Number Attacks in IoT Networks," *IEEE Access*, vol. 10, pp. 111115–111133, 2022.

[6] M. Hassan, N. Tariq, A. Alsirhani, A. Alomari, F. A. Khan, and M. M. Alshahrani, "GITM: A GINI Index-Based Trust Mechanism to Mitigate and Isolate Sybil Attack in RPL-Enabled Smart Grid Advanced Metering Infrastructures," *IEEE Access*, vol. 11, pp. 62697–62720, 2023.

[7] D. Airehrour, J. A. Gutierrez, and S. K. Ray, "SecTrust-RPL: A secure trust-aware RPL routing protocol for Internet of Things," *Future Generation Computer Systems*, vol. 93, pp. 860–876, 2019.

[8] S. Murali and A. Jamalipour, "Mobility-aware energy-efficient parent selection algorithm for low power and lossy networks," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2593–2601, 2018.

[9] H. Kharrufa, H. A. Al-Kashoash, and A. H. Kemp, "RPL-based routing protocols in IoT applications: A review," *IEEE Sensors Journal*, vol. 19, no. 15, pp. 5952–5967, 2019.

[10] P. Thubert, "Objective function zero for the routing protocol for low-power and lossy networks (RPL)," RFC 6550, 2012.

[11] O. Gnawali and P. Levis, "The minimum rank with hysteresis objective function," *RFC 6719*, Sep. 2012.

[12] P. Nandhini *et al.*, "Enhanced Rank Attack Detection Algorithm (E-RAD) for securing RPL-based IoT networks by early detection and isolation of rank attackers," *The Journal of Supercomputing*, 2022, pp. 1–24.

[13] F. V. Nejad, M. M. G. Sadeghi, and M. H. Rezvani, "Analysis of Decreased Rank Attack on RPL-Based IoT Networks," in *2024 10th International Conference on Web Research (ICWR)*, 2024, IEEE.

[14] A. Mayzaud, R. Badonnel, and I. Chrisment, "A Taxonomy of Attacks in RPL-based Internet of Things," *International Journal of Network Security*, vol. 18, no. 3, pp. 459–473, 2016.

[15] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press, 2018.

[16] Wang, X., Yang, Z., Chen, G., and Liu, Y., "A Reinforcement Learning Method of Solving Markov Decision Processes: An Adaptive Exploration Model Based on Temporal Difference Error," *Electronics*, vol. 12, no. 19, p. 4176, 2023.

[17] B. B. Zarpelão, R. S. Miani, C. T. Kawakani, and S. C. de Alvarenga, "A survey of intrusion detection in Internet of Things," Journal of Network and Computer Applications, vol. 84, pp. 25–37, 2017.

[18] S. Raza, L. Wallgren, and T. Voigt, "SVELTE: Real-time intrusion detection in the Internet of Things," *Ad Hoc Networks*, vol. 11, no. 8, pp. 2661–2674, 2013.

[19] S. Y. Hashemi and F. S. Aliee, "Dynamic and comprehensive trust model for IoT and its integration into RPL," *The Journal of Supercomputing*, vol. 75, no. 7, pp. 3555–3584, 2019.

[20] S. Y. Hashemi and F. Shams Aliee, "Fuzzy, dynamic and trust based routing protocol for IoT," *Journal of Network and Systems Management*, vol. 28, no. 4, pp. 1248–1278, 2020.

[21] Nabil Djedjig, Djamal Tandjaoui, Faouzi Medjek, and Issa Romdhani, "Trust-aware and cooperative routing protocol for IoT security," *Journal of Information Security and Applications*, vol. 52, p. 102467, 2020.

[22] Himanshu B. Patel and Devesh C. Jinwala, "Trust and Strainer Based Approach for Mitigating Blackhole Attack in 6LoWPAN: A Hybrid Approach," *IAENG International Journal of Computer Science*, vol. 48, no. 4, pp. 1062–1074, 2021.

[23] Temur ul Hassan, Muhammad Asim, Thar Baker, Jawad Hassan, and Noshina Tariq, "CTrust-RPL: A control layer-based trust mechanism for supporting secure routing in routing protocol for low power and lossy networks-based Internet of Things applications," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 3, p. e4224, 2021.

[24] K. Prathapchandran and T. Janani, "A trust aware security mechanism to detect sinkhole attack in RPL-based IoT environment using random forest–RFTRUST," *Computer Networks*, vol. 198, p. 108413, 2021.

[25] K. Prathapchandran and T. Janani, "A trust-based security model to detect misbehaving nodes in Internet of Things (IoT) environment using logistic regression," in *Journal of Physics: Conference Series*, IOP Publishing, 2021.

[26] Bacem Mbarek, Mouzhi Ge, and Tomáš Pitner, "Proactive trust classification for detection of replication attacks in 6LoWPAN-based IoT," *Internet of Things*, vol. 16, p. 100442, 2021.

[27] Anil O. Bang and U. Prasad Rao, "A novel decentralized security architecture against Sybil attack in RPL-based IoT networks: a focus on smart home use case," *The Journal of Supercomputing*, vol. 77, no. 12, pp. 13703–13738, 2021.

[28] B. Rakesh, "Novel Authentication and Secure Trust based RPL Routing in Mobile sink supported Internet of Things," *Cyber-Physical Systems*, 2021, pp. 1–34.

[29] Danish Arshad, Muhammad Usman, Abdul Rauf, and Syed Ali Raza Zaidi, "THC-RPL: A lightweight Trust-enabled routing in RPL-based IoT networks against Sybil attack," *PLOS ONE*, vol. 17, no. 7, p. e0271277, 2022.

[30] S. M. Muzammal, S. Khan, S. A. Khan, and M. S. Khan, "A trust-based model for secure routing against RPL attacks in Internet of Things," *Sensors*, vol. 22, no. 18, p. 7052, 2022.

[31] S. M. Muzammal, S. Khan, S. A. Khan, and M. S. Khan, "Trust and Mobility-Based Protocol for Secure Routing in Internet of Things," *Sensors*, vol. 22, no. 16, p. 6215, 2022.

[32] F. Azzedin, "Mitigating denial of service attacks in RPL-based IoT environments: trust-based approach," *IEEE Access*, vol. 11, pp. 129077–129089, 2023.

[33] P. Levis, N. Patel, D. Culler, and S. Shenker, "The Trickle Algorithm," 2011.

[34] L. Lassouaoui, H. Ouabida, and M. Ouzzif, "Evaluation of energy aware routing metrics for RPL," in *2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2016, pp. 1–8.

[35] A. Fotouhi, M. Ding, and M. Hassan, "Deep Q-Learning for two-hop communications of drone base stations," *Sensors*, vol. 21, no. 6, p. 1960, 2021.

[36] Z1 Datasheet. [Online]. Available: https://zolertia.sourceforge.net/wiki/images/e/e8/Z1_RevC_Datasheet.pdf (accessed Aug. 10, 2024).

[37] G. Violettas, S. Petridou, and L. Mamatas, "Evolutionary software defined networking-inspired routing control strategies for the Internet of Things," *IEEE Access*, vol. 7, pp. 132173–132192, 2019.

[38] C. J. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, pp. 279–292, 1992.

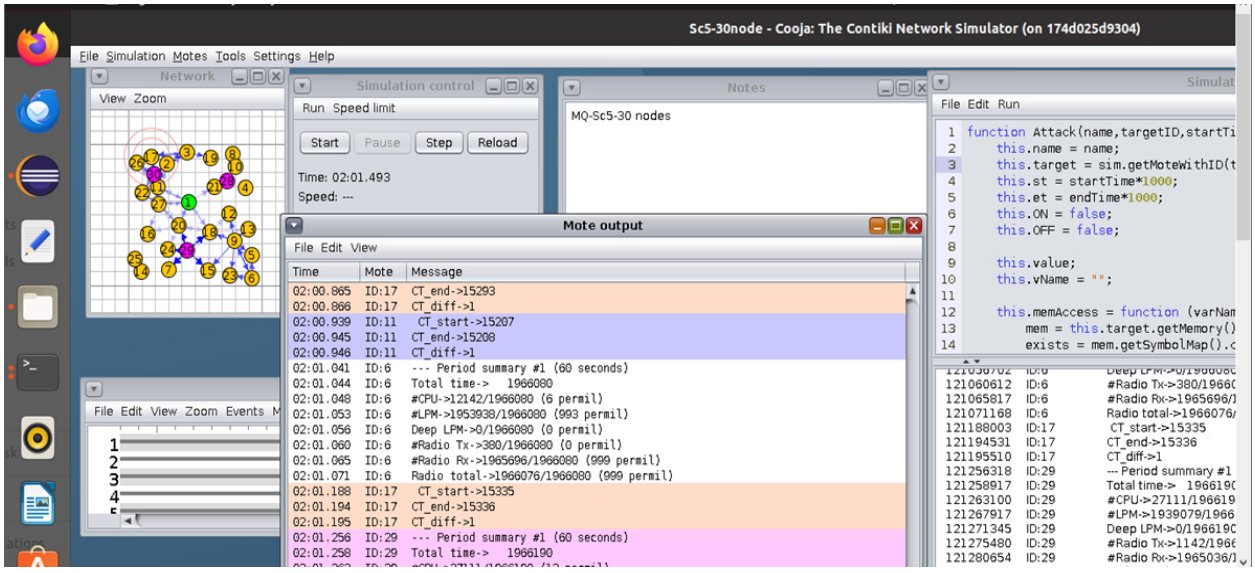
[39] D. G. Pereira, A. Afonso, and F. M. Medeiros, "Overview of Friedman's test and post-hoc analysis," *Communications in Statistics-Simulation and Computation*, vol. 44, no. 10, pp. 2636–2653, 2015.

[40] L. J. Williams and H. Abdi, "Fisher's least significant difference (LSD) test," in *Encyclopedia of Research Design*, 2010, vol. 218, no. 4, pp. 840–853.

پیوست

جدول۱: آمار توصیفی داده‌ها در آزمون ANOVA

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
PDR	MRHOF	200	38.6550464887	20.3653806410	1.44004987527	35.8153307131	41.4947622643	8.32769127	73.97101449
	SecTrust	200	53.0680221604	22.4416923782	1.58686728619	49.9387888305	56.1972554903	20.72322670	85.88374852
	Proposed Method	200	56.1740364772	22.4975636886	1.59081798444	53.0370125421	59.3110604124	23.73199750	88.38559814
	SMTrust	200	54.5192151632	22.6110960341	1.59884593357	51.3663604613	57.6720698652	19.53551913	87.74912075
	Total	800	50.6040800724	23.0432174420	.81470076568	49.0048734198	52.2032867249	8.32769127	88.38559814
Throughput	MRHOF	200	181.161244444	108.340129734	7.66080404096	166.054471782	196.268017107	21.90222222	362.95111111
	SecTrust	200	251.837155556	111.886138785	7.91154474555	236.235933119	267.438377992	62.00888889	411.87555556
	Proposed Method	200	268.791466667	114.740171555	8.11335533811	252.792282512	284.790650821	90.16888889	432.92444444
	SMTrust	200	260.361155556	111.940117613	7.91536162510	244.752406399	275.969904712	84.78222222	425.81333333
	Total	800	240.537755556	116.849244049	4.13124464220	232.428380686	248.647130425	21.90222222	432.92444444
Overhead	MRHOF	200	.8999092833	.04671354490	.00330314644	.8933956220	.9064229445	.81702586	.98917429
	SecTrust	200	.9411743416	.01073190618	.00075886036	.9396779019	.9426707812	.90857174	.98071407
	Proposed Method	200	.9198841521	.03534878903	.00249953684	.9149551741	.9248131302	.84286641	.98078013
	SMTrust	200	.9276277557	.02198365513	.00155447916	.9245623903	.9306931210	.87423496	.97258313
	Total	800	.9221488832	.03502801044	.00123842719	.9197179281	.9245798383	.81702586	.98917429
Total Energy	MRHOF	200	100158.196758	42.0689329472	2.97472277642	100152.330734	100164.062782	100076.949225	100297.901325
	SecTrust	200	100499.619504	75.2700682986	5.32239757143	100489.123968	100510.115041	100320.812335	100785.344690
	Proposed Method	200	100571.142649	230.311418095	16.2854765520	100539.028397	100603.256901	100294.165588	101514.342507
	SMTrust	200	100756.438224	168.383577856	11.9065169742	100732.959090	100779.917358	100512.354568	101410.996753
	Total	800	100496.349284	262.836042649	9.29265740487	100478.108379	100514.590189	100076.949225	101514.342507



شکل ۱: محیط شبیه سازی Contiki-NG در حال اجرا

پاورقی‌ها:

[1] Routing Over Low power and Lossy networks

[2] Internet Engineering Task Force

[3] Expected Transmission Count

[4] Destination Oriented Directed Acyclic Graph
[5] Objective Function Zero
[6] Minimum Rank with Hysteresis Objective Function
[7] Reinforcement Learning
[8] Intrusion Detection System
[9] Quality of Service
[10] Quality of p2p communication
[11] Quantum Inspired Neural Network
[12] Received Signal Strength Indicator
[13] Grid
[14] ε-greedy
[15] Low Power Mode
[16] Levene's test