



Computational Intelligence in Electrical Engineering
Vol. 12, No. 2, 2021
Research Paper

Reliability and Battery Lifetime Improvement for IoT Networks: Challenges and AI-powered solutions

Amin Azari ^{1*}, Mohsen NikNejad ², Mahmoud Abbasi ³

¹ Stockholm University, Stockholm, Sweden

² Department of Computer Engineering, Allameh Dehkoda University, Isfahan, Iran

³ Department of Software Engineering, Mashhad Branch, IAU, Mashhad, Iran

Abstract:

Towards realizing an intelligent networked society, enabling low-cost low-energy connectivity for things, also known as the Internet of Things (IoT), is of crucial importance. While the existing wireless access networks require centralized signaling for managing network resources, this approach is of less interest for future generations of wireless networks due to the energy consumption in such signaling and the expected increase in the number of IoT devices. Then, in this work, we investigate leveraging machine learning for distributed control of IoT communications. Towards this end, first, we investigate low-complex learning schemes that are applicable to resource-constrained IoT communications. Then, we propose a lightweight learning scheme which enables the IoT devices to adapt their communication parameters to the environment. Further, we investigate analytical expressions presenting the performance of a centralized control scheme for adapting communication parameters of IoT devices and compare the results with the results from the proposed distributed learning approach. The simulation results confirm that the reliability and energy efficiency of IoT communications could be significantly improved by leveraging the proposed learning approach.

Keywords: IoT, 5G, battery lifetime, reliability, machine learning, Multi-arm bandit.



2252-083X/ © 2021 The Authors. Published by University of Isfahan

This is an open access article under the CC BY-NC-ND/4.0/ License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).



<http://dx.doi.org/10.22108/isee.2020.118921.1271>

بهبود قابلیت اطمینان و عمر باتری در شبکه‌های ارتباطی اینترنت اشیا:

چالش‌ها و راهبردهای مبتنی بر یادگیری ماشینی

امین آذری^{۱*}، محسن نیک نژاد^۲، محمود عباسی^۳

۱- دانشگاه استکهلم، استکهلم، سوئد

aazari@kth.se

۲- کارشناسی ارشد نرم افزار کامپیوتر، دانشگاه علامه دهخدا، اصفهان، ایران

moh3n.niknezhad@gmail.com

۳- کارشناسی ارشد، واحد مشهد، دانشگاه آزاد اسلامی، مشهد، ایران

mahmoud_abbasi_student@yahoo.com

چکیده: در راستای تحقق یک جامعه هوشمند، برقراری ارتباط برای تمام اشیای هوشمند با هزینه و انرژی مصرفی کم یک نیاز اساسی است. شبکه‌های بی‌سیم کنونی برای برقراری بهینه ارتباطات، به مدیریت متمرکز شبکه و منابع نیاز دارند و عواملی مانند انرژی مصرفی در ارسال - دریافت سیگنال‌های کنترلی و تعداد زیاد دستگاه‌های اینترنت اشیا، امکان استفاده از چنین رویکردهای متمرکزی را در آینده غیرممکن خواهند کرد. برای حل این مشکل، در این مقاله امکان استفاده از راه‌حل‌های مبتنی بر یادگیری ماشینی برای شبکه‌های اینترنت اشیا بررسی شده است. در گام نخست برای دستیابی به این هدف، روش‌های یادگیری با پیچیدگی کم بررسی شده‌اند که مناسب پیاده‌سازی در اشیاء‌اند. در ادامه، یک روش یادگیری برای تطبیق پارامترهای مخابراتی در اشیاء با محیط پیرامون آنها ارائه شده است. در این روش پیشنهادی، تابع ارزش هر تصمیم براساس سابقه انرژی مصرفی و میزان موفقیت در اتخاذ آن تصمیم طراحی می‌شود. این طراحی، دستگاه را قادر می‌کند بهترین مصالحه را بین انرژی مصرفی و قابلیت اطمینان ارتباطات به دست آورد. در گام بعدی، مقایسه عملکرد روش پیشنهادی مقاله با رویکرد سیستم‌های متمرکز، با بهره‌گیری از ابزار هندسه تصادفی ارائه شده است. سپس، ارتباط بین پارامترهای الگوریتم یادگیری ماشینی و عملکرد سیستم مخابراتی، مانند انرژی مصرفی و قابلیت اطمینان، تجزیه و تحلیل شده است. نتایج شبیه‌سازی نشان می‌دهند در مقایسه با جدیدترین روش‌های پیشنهاد شده در ادبیات این موضوع، هر دو معیار بهره‌وری انرژی و سطح اطمینان با استفاده از روش یادگیری مندرج در این مقاله به صورت چشمگیری بهبود می‌یابند.

واژه‌های کلیدی: اینترنت اشیا، عمر باتری، یادگیری، نسل پنجم، یادگیری ماشینی، ماشین شانس چند اهرمه.

۱- مقدمه

سیستم‌های ارتباطی در مقابل نویز و تداخل قرار داده‌اند که این مسئله عمدتاً ناشی از محدودیت‌های کانال فیزیکی مانند نویز و تداخل است. به لطف پهنای باند بالا و استفاده از سخت‌افزارها و نرم‌افزارهای پیشرفته در نسل چهارم شبکه تلفن همراه (هم در ایستگاه‌های پایه و هم در دستگاه‌های ارتباطی کاربران)، این شبکه قادر به ارائه ارتباطات با سرعت بالا به صورت یکپارچه و مطمئن به کاربران است [۱]. در مقایسه با شبکه‌های نسل چهارم، شبکه‌های تلفن همراه نسل پنجم، طیف وسیع‌تری از

در طول پیدایش شبکه‌های مخابراتی، بخش چشمگیری از تحقیقات در این حوزه، تمرکز خود را روی مقاوم‌سازی

^۱ تاریخ ارسال مقاله: ۱۳۹۸/۰۶/۰۶

تاریخ پذیرش مقاله: ۱۳۹۹/۰۵/۱۴

نام نویسنده مسئول: امین آذری

نشانی نویسنده مسئول: سوئد، استکهلم، دانشگاه استکهلم

در مناطق روستایی و حومه شهرها نیز یاد کرد. این فناوری‌ها با بهره‌گیری از روش استفاده آزاد از منابع رادیویی (grant-free access) که نیازمند جفت‌سازی، هماهنگ‌سازی و رزرو منابع رادیویی نیست، حجم سیگنالینگ لازم بین دستگاه (به‌طور مثال، دستگاه اینترنت اشیا) و سرورها را بسیار کاهش داده‌اند. این کاهش سیگنالینگ، به افزایش طول عمر باتری دستگاه متصل در یک شبکه لورا یا سیگفاکس منجر می‌شود [۷]. در زمان حاضر، دسترسی آزاد به منابع رادیویی، موضوع بحث در استانداردهای مخابراتی بوده است و انتظار می‌رود در نسخه‌های آینده استاندارد 3GPP LTE نیز پیاده‌سازی شود [۸].

درحالی‌که میزان مصرف انرژی در حالت دسترسی آزاد بسیار کمتر از حالت با رزرو قبلی است، قابلیت اطمینان ارتباطات در دسترسی آزاد به علت تداخل احتمالی بسته‌های ارسال، مشکلی اساسی است [۹ و ۱۰ و ۱۱]. اندازه‌گیری تداخل در فرکانس ۸۶۸ مگاهرتز (باند فرکانسی ISM) در [۹] ارائه شده است. این نتایج نشان می‌دهند باند فرکانسی ISM، با توجه به رشد روزافزون فناوری‌های ارتباطی که به‌صورت هم‌زمان در باند فرکانسی آزاد ISM فعالیت می‌کنند، در برخی زیرباندها و زمان‌ها تداخل زیادی دارد که به معنای احتمال بالای تصادم بسته‌ها در این زیرباندها است.

در سال‌های اخیر علاقه فراوانی به استفاده از ابزارهای یادگیری ماشینی برای توصیف رفتار شبکه‌های مقیاس بزرگ به وجود آمده است [۱۲]. در [۱۳]، یک راه‌حل مبتنی بر یادگیری تقویتی^۲، در سمت شبکه، برای شبکه‌های LTE ارائه شده است که به ترافیک اینترنت اشیا سرویس می‌دهند. در [۱۴]، روش خوشه‌بندی خودسازمان‌یافته و دسترسی به منابع مخابراتی براساس خوشه‌بندی، برای شبکه‌های اینترنت اشیا بررسی شده است. در [۱۵]، استفاده از الگوریتم یادگیری MAB^۳ برای شبکه‌های اینترنت اشیا پیشنهاد شده است که در آن، دستگاهها یاد می‌گیرند از زیرکانال‌های با سطح بالایی از تداخل استاتیک استفاده نکنند. در [۱۶]، برای کنترل زمانبندی و انتقال اطلاعات^۴ دستگاههای اینترنت اشیا در سیستم‌های با پردازش لبه^۵ از

ابزارهای ارتباطی را هدف قرار داده است؛ برای مثال، در نسل پنجم، ایجاد ارتباط برای دستگاههای هوشمند با محدودیت در زمینه انرژی، هزینه و پیچیدگی محاسباتی، یکی از اهداف اصلی است [۲].

براساس یک چشم‌انداز بلندمدت، انتظار می‌رود شبکه‌های نسل پنجم و پس از آن، امکان اتصال در مقیاس گسترده، با هزینه کم و قابلیت اطمینان بالا را برای تمام دستگاههای با قابلیت اتصال به شبکه فراهم کنند. تا به امروز، طراحی و بهینه‌سازی شبکه‌های ارتباطی و به دنبال آن، دستیابی به درک درست از محدودیت‌های فیزیکی مانند نویز و تداخل، مبتنی بر مدل‌های آماری بوده‌اند. همچنین، دستگاههای ارتباطی کاربران که غالباً تلفن همراه هوشمندند، به‌طور معمول هر روز شارژ می‌شوند و برای برقراری ارتباطات نیاز دارند به‌طور مستمر به کانال ارتباطی با ایستگاه پایه مربوطه گوش دهند؛ درحالی‌که ایستگاه‌های پایه وظیفه مدیریت ارتباطات، ارسال دستورالعمل‌های ارتباطی و زمانبندی منابع رادیویی را بر عهده دارند. با در نظر گرفتن پیچیدگی، گستردگی مقیاس و ناهمگونی شبکه‌های بی‌سیم نسل آینده، به‌خصوص با در نظر گرفتن ترافیک اینترنت اشیا، امکان کنترل متمرکز میلیون‌ها ارتباط، در سمت شبکه بیش از پیش مشکل است [۳]. در ادبیات این حوزه برای حل این مشکل دو نوع راه‌حل کلی ارائه شده است؛ دسته اول، راه‌حل‌های تکاملی است که تلاش می‌کنند شبکه‌های سلولی را به نحوی توسعه دهند که ترافیک مربوط به دستگاههای اینترنت اشیا در بستر شبکه‌های سلولی موجود مانند LTE سرویس‌دهی شود [۴، ۵]. دسته دوم شامل راه‌حل‌هایی است که بر بازبینی اساسی ارتباطات اشیا با سرورها، کاهش اساسی حجم سیگنالینگ بین دستگاههای ارتباطی کاربر و شبکه و کاهش کنترل شبکه بر منابع مخابراتی متمرکز شده‌اند [۶]. دسته سوم که دسترسی آزاد به منابع رادیویی نیز نامیده می‌شود، در سال‌های اخیر با استقبال زیاد گروه‌های تحقیقاتی، استانداردهای مخابراتی و حتی صنعت مواجه شده است؛ برای مثال، به پروتکل‌های ارتباطی مانند سیگفاکس (SigFox) و لورا (LoRa) می‌توان اشاره کرد.

فناوری‌های لورا و سیگفاکس دو پروتکل غالب در زمینه اینترنت اشیا، روی باندهای فرکانسی بدون مجوزند. از این دو فناوری می‌توان به‌عنوان دو کاندید اصلی پیاده‌سازی نسل پنجم

² Reinforcement Learning

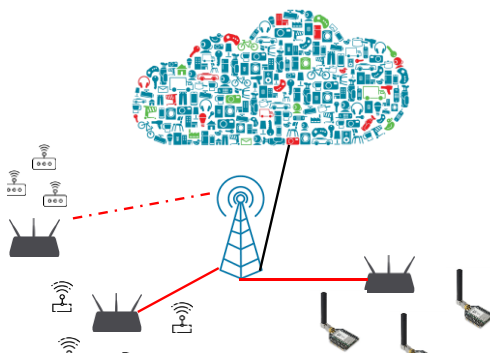
³ Multi-Arm Bandit

⁴ Offloading

⁵ Edge processing

زیر است:

- مدل‌سازی تنظیم پارامترهای مخابراتی اینترنت اشیا به‌عنوان مسئله بهینه‌سازی؛
 - بررسی الگوریتم‌های یادگیری ماشینی کاربردی در بهبود عملکرد شبکه‌های اینترنت اشیا؛
 - ارائه راه‌حل مبتنی بر یادگیری ماشینی با هدف افزایش بهره‌وری انرژی اشیا و قابلیت اطمینان ارتباطات در شبکه‌های اینترنت اشیا؛
 - توسعه یک مدل تحلیلی برای ارزیابی عملکرد رویکردهای یادگیری توزیع‌شده با بهره‌گیری از ابزارهای آماری هندسه تصادفی؛
 - ارزیابی قابلیت اطمینان، بهره‌وری انرژی و مصالحه بین آنها در راه‌حل‌های پیشنهادی و سایر راه‌حل‌های موجود در ادبیات حوزه.
- ادامه این مقاله به این ترتیب سازماندهی شده است: در بخش ۲ سیستم مدنظر در مقاله تشریح می‌شود. بخش ۳ مربوط به بررسی مبتنی بر یادگیری ماشینی کاربردی در اینترنت اشیا است. در بخش ۴، یک الگوریتم مبتنی بر یادگیری ماشینی توزیع‌شده ارائه شده و عملکرد آن با روش بهینه‌سازی متمرکز مقایسه شده است. نتایج مربوط به شبیه‌سازی، در بخش ۵ و نتیجه‌گیری در بخش ۶ ارائه شده است.



شکل ۱: مدل سیستم، اینترنت اشیا با دسترسی آزاد

۲- مدل سیستم و فرموله‌سازی مسئله

سیستم مدنظر در این کار، حالت کلی از ارتباطات اینترنت اشیا بدون نیاز به دسترسی به کانال، مانند سیگفاکس و لورا را

یادگیری تقویتی عمیق استفاده شده است. راه‌حل پیشنهادی در این مقاله قادر است به‌صورت هم‌زمان مسئله زمانبندی و انتقال اطلاعات را حل کند تا بدین ترتیب میزان مصرف انرژی و میانگین تأخیر را به‌صورت چشمگیری کاهش دهد. در [۱۷]، جنبه‌های مختلف استفاده از یادگیری تقویتی عمیق برای اینترنت اشیا خودمختار، شامل مدل‌ها، کاربردها و چالش‌ها بررسی شده‌اند. در [۱۸] تخصیص کانال به‌صورت توزیع‌شده و تصادفی بین کاربران شبکه لورا بررسی شده است. در [۱۹ و ۲۰]، تخصیص کانال در شبکه‌های اینترنت اشیا با استفاده از یادگیری ماشینی بررسی شده، اما به مصرف انرژی در انتخاب پارامترهای مخابراتی توجهی نشده است.

به‌منظور تحقق شبکه‌های اینترنت اشیا خودسازمان‌یافته که قادر به انطباق با محیط پیرامون خود هستند، در این پژوهش سناریوهایی بررسی شده‌اند که در آن اینترنت اشیا روی باند با دسترسی آزاد و با حضور شبکه‌های تداخلی دیگر پیاده‌سازی می‌شود. در این نوع ارتباطات، انتخاب پارامترهایی مانند نرخ داده، زیرکانال استفاده‌شده، توان ارسالی و تعداد تکرار ارسال، مشخص‌کننده ظرفیت شبکه و طول عمر باتری اشیا-اند. هدف اصلی از این پژوهش افزایش سطح اطمینان ارتباطات در دستگاه‌های اینترنت اشیا از زمان قیمت در محیط‌های تداخلی و کاهش میزان مصرف انرژی در این دستگاه‌ها است. در این راستا، یک راه‌حل مبتنی بر یادگیری ماشینی توزیع‌شده، برای کنترل پارامترهای مخابراتی پیشنهاد شده است و نتایج به‌دست آمده با نتایج تحلیلی مربوط به مسئله بهینه‌سازی متمرکز مقایسه شده‌اند.

نتایج ارزیابی نشان‌دهنده کاهش چشمگیر در مصرف انرژی و افزایش احتمال موفقیت در ارسال داده است. نتایج اولیه این پژوهش در [۲۱] ارائه شده که شامل استفاده از یادگیری ماشینی برای تنظیم پارامترهای مخابراتی تکنولوژی لورا در هنگام پیاده‌سازی است. در این نسخه، ما [۲۱] را توسعه دادیم و نتایج جدیدی در استفاده از یادگیری ماشینی برای سازگاری شبکه با تداخل مقطعی، یادگیری کمک‌شده توسط شبکه و تکنیک انتقال یادگیری ارائه می‌کنیم. نوآوری‌های اصلی این مقاله شامل موارد

⁶ Autonomous IoT (AIoT)

مدل می‌کند. مجموعه‌ای از دستگاه‌های اینترنت اشیا (مجموعه Φ) را در یک محدوده جغرافیایی A در نظر بگیرید. انتقال داده از اشیا به سرور از طریق نقاط دسترسی انجام می‌گیرد که در این محدوده قرار داده شده‌اند (شکل ۱). هرگاه یکی از اشیا پاکتی برای ارسال دارد، بدون نیاز به رزرو منبع و زمانبندی، اقدام به ارسال پکت می‌کند. هر نقطه دسترسی این دستگاه‌های مختلف در شبکه اینترنت اشیا، الگوهای متفاوتی در استفاده از منابع رادیویی دارند؛ به طور دقیق‌تر، مدت زمان بین دو ارسال داده، پهنای باند سیگنال استفاده‌شده، توان ارسال، نرخ ارسال داده و زمان انتقال بسته‌ها از دستگاهی به دستگاه دیگر متفاوت است. در این مقاله، پهنای باند فرکانسی به اشتراک گذاشته برای ارتباطات را با W و چگالی طیفی توان نویز را با N نمایش می‌دهیم.

فرض کنید از مجموعه Φ ، زیرمجموعه Φ_s مدنظر ما برای جمع‌آوری اطلاعات‌اند و ترافیک سایر اشیا تداخل محسوب می‌شوند. مسئله بررسی‌شده در این مقاله، کنترل پارامترهای مخابراتی برای مجموعه‌ای از دستگاه‌های اینترنت اشیا، Φ_s ، براساس مشاهده و تعامل با محیط اطراف است. فرض کنید در زمان t ، i امین دستگاه از مجموعه Φ_s نیاز به ارسال داده داشته باشد، در این حالت مسئله مدنظر به صورت زیر فرموله می‌شود:

فرض کنید از مجموعه Φ ، زیرمجموعه Φ_s مدنظر ما برای جمع‌آوری اطلاعات‌اند و ترافیک سایر اشیا تداخل محسوب می‌شوند. مسئله بررسی‌شده در این مقاله، کنترل پارامترهای مخابراتی برای مجموعه‌ای از دستگاه‌های اینترنت اشیا، Φ_s ، براساس مشاهده و تعامل با محیط اطراف است. فرض کنید در زمان t ، i امین دستگاه از مجموعه Φ_s نیاز به ارسال داده داشته باشد، در این حالت مسئله مدنظر به صورت زیر فرموله می‌شود:

$$\begin{aligned} & \text{maximize } F(\text{REL}_i, \text{EC}_i) \\ & p_i, r_i, h_i, m_i, \forall i \in \Phi_s \\ & \text{s.t. } p_i \in \mathbb{P}, r_i \in \mathbb{R}, h_i \in \mathbb{H}, m_i \in \mathbb{M}, \end{aligned} \quad (1)$$

که در آن $F(\cdot)$ بیان‌کننده تابع هدف است و تعادلی بین مصرف انرژی (EC) و قابلیت اطمینان (REL) ارتباطات ایجاد می‌کند. با توجه به تفاوت کیفیت سرویس (QoS) در کاربردهای مختلف اینترنت اشیا، تعریف تابع $F(\cdot)$ در هر کاربرد می‌تواند متفاوت از کاربردهای دیگر باشد. در این پژوهش، مجموع مقادیر وزن‌دار و استاندارد شده به بازه $(0,1)$ قابلیت اطمینان و مصرف انرژی، تابع هدف در نظر گرفته شده‌اند؛ یعنی

$$F(\text{Reli}, \text{EEi}) = (1 - \beta)\text{Rel}_i + \beta\text{EE}_i, \quad (2)$$

که در آن $0 \leq \beta \leq 1$ پارامتر تنظیم‌کننده تعادل بین سطح اطمینان و مصرف انرژی است. همچنین، p_i ، r_i ، h_i و m_i به ترتیب اشاره به توان ارسال، نرخ داده ارسالی، زیرکانال

۳- یادگیری ماشینی در اینترنت اشیا

۳-۱- معرفی و دسته‌بندی

یادگیری ماشینی که در آن، یک واحد هوشمند، اطلاعات موجود را پردازش کرده است و برای خوشه‌بندی، پیش‌بینی یا تصمیم‌گیری به کار می‌رود، در حالت کلی به سه دسته یادگیری با داده‌های نشاندار، با داده‌های بی‌نشان و یادگیری تقویتی^۷ تقسیم می‌شود. در سال‌های اخیر به لطف توسعه قدرت ذخیره و پردازش داده، یادگیری ماشینی در قسمت‌های مختلف شبکه و حتی در سرویس‌های اینترنت اشیا جای خود را باز کرده است [۲۲]؛ برای مثال، در [۲۳] روش‌های ارائه‌شده برای امنیت بیشتر ارتباطات اینترنت اشیا بررسی شده‌اند. یک روش پیشنهادی، یادگیری مشخصه‌های فیزیکی هر شیء، مانند تابع توزیع احتمال آفست فرکانسی آن، از طریق بررسی عمیق ارتباطات گذشته و ارزیابی هویت هر پکت بعدی دریافتی با این مشخصه‌ها است. بزرگ‌ترین مشکل در استفاده از یادگیری با داده‌های نشاندار و با داده‌های بی‌نشان، نیاز به حجم داده نسبتاً بالا برای پردازش و همگراشدن راه‌حل است که کاربرد آن را در راه‌حل‌های توزیع‌شده در سمت اشیا، ناممکن می‌کند و به‌طور معمول، مدنظر در سمت شبکه و سرورها است. همچنین،

⁷ Supervised/unsupervised, reinforcement learning

پاداش در طول زمان و کمینه‌سازی میزان ضررهای^{۱۱} متحمل شده در تصمیم‌گیری‌هایش است. رویکرد مبتنی بر MAB باعث ایجاد تعادل بین فرایند اکتشاف و استخراج^{۱۲} می‌شود. «اکتشاف» به بازه‌های تصمیم‌گیری اشاره دارد که در آن عامل تلاش می‌کند مجموعه گزینه‌های گوناگون را تست کند؛ حتی در صورتی که پاداش‌های قبلی این اقدامات کمتر از دیگر مجموعه اقدامات باشد. «استخراج» بیان‌کننده زمان‌های تصمیم‌گیری است که عامل براساس مشاهدات قبلی، به صورت حریمانه‌ای برای افزایش پاداش تلاش می‌کند. با توجه به کاربردهای گسترده الگوریتم یادگیری MAB در زمینه‌های مختلفی مانند رباتیک، این الگوریتم به خوبی در ادبیات این پژوهش بررسی شده و راه‌حل‌های کارآمدی به منظور کمینه‌سازی میزان ضرر در تصمیم‌گیری‌ها پیشنهاد شده است. در ادامه مقاله تلاش شده است رویکردهایی برای حل مسئله کنترل پارامترهای مخابراتی در دستگاه‌های اینترنت اشیا در دو محیط متفاوت ارائه شوند: (۱) محیط با تداخل روی کانال داده و عدم تداخل روی کانال پس‌خورده (۲) محیط با تداخل روی کانال داده و پس‌خورده.

۳-۳- محیط با تداخل روی کانال داده

در محیط با تداخل روی کانال داده، مسئله یادگیری، MAB تصادفی^{۱۳} نامیده می‌شود. برای MAB تصادفی که در آن، پاداش هر عمل از یک تابع چگالی احتمال به دست می‌آید، خانواده الگوریتم UCB^{۱۴}، کمینه ضرر ممکن را به دست می‌آورد [۲۲]. به عبارت دیگر، این الگوریتم به انتخاب تصمیم با بالاترین کران اطمینان در برآورده کردن پاداش بیشینه همگرا می‌شود. در میان الگوریتم‌های خانواده UCB، ما بر UCB₁، ارائه شده در [۲۴] تمرکز می‌کنیم که در آن ضرر ناشی از تصمیم‌گیری با نرخ $O(\log n)$ رشد می‌کند [۲۵].

۳-۴- محیط با تداخل روی کانال داده و پس‌خورده

در محیط با تداخل روی کانال داده و پس‌خورده با شرایطی

یادگیری تقویتی که در آن هر شیء از طریق ارتباط با محیط و تجربه، به تدریج پاداش درازمدت هر عمل را درک می‌کند، به طور گسترده در سمت ابزارهای هوشمند شایان توجه قرار گرفته است [۲۲]. از مهم‌ترین مدل‌های یادگیری در این حوزه می‌توان به یادگیری چندانتخابی^۸ (MAB) و یادگیری Q، معروف به Q-learning اشاره کرد [۲۲]. در هر دو روش، کاربر تصمیم‌گیرنده برای هر عمل یک نمایه در نظر می‌گیرد و با توجه به تجربه حاصل از انتخاب آن، برای اولویت دادن یا ندادن به آن عمل تصمیم می‌گیرد. تفاوت این دو الگوریتم، به تفاوت نگاه آنها به محیط برمی‌گردد که در اولی، محیط با یک حالت (state) مدل می‌شود؛ درحالی‌که در دومی، محیط پیرامون با تعدادی حالت مدل می‌شود و ارزش هر عمل در هر حالت جداگانه بررسی می‌شود.

در این پژوهش که اینترنت اشیا با نیاز مبرم به بهینه‌گری انرژی بررسی می‌شود، روی مدل یادگیری MAB تمرکز می‌شود تا سربار انرژی و پیچیدگی پردازش داده و تصمیم‌گیری در اشیا به حداقل برسد.

۳-۲- خودسازماندهی با MAB

در این بخش، راه‌حلی مبتنی بر مدل MAB برای خودسازماندهی اشیا متصل به اینترنت ارائه شده است. در روش یادگیری توزیع شده MAB، هر دستگاه برای بیشینه‌سازی مقدار تابع هدف خود، $F(REL_i, EC_i)$ تلاش می‌کند. این کار با استفاده از بهترین انتخاب (A_i) از مجموعه تصمیمات (\mathbb{A}) انجام می‌گیرد: $A_i = \{p_i, c_i, h_i, m_i\} \in \mathbb{A}$. برای انتخاب بهترین تصمیم، پاداش^۹ (نتیجه) اقدامات قبلی باید به نحوی ذخیره و پردازش شوند؛ برای مثال، بعد از انتخاب یک تصمیم در زمان t ، یعنی $A_i(t)$ ، دستگاه می‌تواند گزارش دریافت از سوی نقطه دسترسی یا سرور را یک پاداش در نظر بگیرد. این گزارش دریافت با $\xi(t) \in \{1, 0\}$ نمایش داده می‌شود و در آن، ۱ و ۰ به ترتیب بیان‌کننده پیام تصدیق دریافت و عدم تصدیق^{۱۰} هستند. هدف عامل بیشینه‌کردن میزان مجموع

¹¹ Regret

¹² Exploration and exploitation

¹³ Stochastic MAB

¹⁴ Upper Confidence Bound

⁸ Multi-arm bandit

⁹ Rewards

¹⁰ ACK and NACK

ارسال کرده‌اند، $T_k(t)$ تعداد دفعاتی که تا زمان t تصمیم k انتخاب شده است، $A_k(t)$ تصمیم گرفته‌شده در زمان تصمیم‌گیری t ، و

$$\bar{z}(t) = z(t) \left[(1 - \beta) + \frac{\beta E_k}{E_{min}} \right] \quad (۳)$$

نشان‌دهندهٔ پاداش تغییر یافتهٔ تصمیم انتخابی در زمان t است که تابعی از دریافت شدن یا نشدن ACK، یعنی $z(t) \in \{0,1\}$ است و در ادامه بررسی می‌شود. همچنین، $Z_k(t)$ پاداش تجمعی تصمیم k تا زمان t ، پارامتر ایجاد تعادل بین سطح اطمینان و بهره‌وری انرژی و در نهایت، $I_k(t)$ ارزش انتخاب تصمیم k در زمان t است. همچنین، با استفاده از رویکرد مشابهی، در الگوریتم ۲ به‌جای ذخیره و به‌روزرسانی ارزش یک تصمیم در پارامتر $I_k(t)$ ، احتمال رسیدن به بیشترین پاداش از طریق تصمیم k ، یعنی $p_k(t)$ ذخیره و به‌روزرسانی می‌شود. در این الگوریتم‌ها، α و ρ پارامترهای طراحی‌اند که میان فازهای اکتشاف و استخراج در الگوریتم ۱ و ۲ (به ترتیب) تعادل ایجاد می‌کنند. با مقایسهٔ الگوریتم ۱ و ۲ با مسئله (۱) مشاهده می‌شود $F(\text{Rel}_i, \text{EE}_i)$ در این الگوریتم‌ها با استفاده از تابع پاداش تغییر یافته، یعنی $\bar{z}(t)$ در (۳)، مدل‌سازی شده است. در رابطه (۳)، $z(t)(1 - \beta)$ نشان‌دهندهٔ پاداش خارجی (از طرف گیرنده با توجه به دریافت کردن یا نکردن پکت) و $\frac{\beta E_k}{E_{min}}$ نشان‌دهندهٔ پاداش داخلی با توجه به میزان مصرف انرژی است.

مواجه هستیم که هم یک پکت ممکن است به مقصد نرسد و هم ACK مربوط به یک پکت رسیده به مقصد نیز ممکن است در اثر تداخل به گیرنده نرسد. در این شرایط مسئلهٔ یادگیری به نام MAB غیر تصادفی شناخته می‌شود که در آن، پاداش هر تصمیم لزوماً از یک تابع چگالی احتمال خاص مشتق نمی‌شود (چون پیام حاوی پاداش نیز ممکن است در اثر تداخل از بین برود). نمونه‌هایی از MAB غیر تصادفی، محیط با تداخل خصمانه است که در آن یک عنصر متخاصم می‌تواند با تداخل روی کانال پسخورد، کارایی ارتباطات را دچار مشکل کند. علاوه بر این، در کاربردهای اینترنت اشیا در باند ISM، تداخل از سوی شبکه‌های دیگر می‌تواند در کانال پسخورد تداخل ایجاد کند.

الگوریتم‌های متعدد برای یادگیری MAB با تداخل غیر تصادفی در ادبیات این حوزه حضور دارند که در میان آنها عملکرد الگوریتم EXP3 بهینه است [۲۶]. الگوریتم EXP3 هر زمان تصمیم‌گیری t ، یک تصمیم از مجموعه \mathbb{A} را براساس توزیع‌های احتمال پاداش آنها انتخاب می‌کند؛ یعنی به بیان آماری: $A_t(t) \sim \{p_1(t), \dots, p_{|\mathbb{A}|}(t)\}$. میزان ضرر از تصمیم‌گیری براساس الگوریتم EXP3 با مرتبهٔ زمانی $O(\sqrt{n})$ رشد می‌کند که این میزان، کمینهٔ رشد بین سایر الگوریتم‌ها است [۲۳].

۴- الگوریتم‌های پیشنهادی

به یاد بیاوریم هدف مسئلهٔ بهینه‌سازی (۱)، بیشینه‌کردن سطح اطمینان و بهره‌وری انرژی دستگاهها است. در این بخش و با هدف ارائه یک راه‌حل توزیع‌شده برای این مسئله، دو الگوریتم یادگیری مبتنی بر MAB ارائه شده که در آنها دو معیار میزان موفقیت و مصرف انرژی در فرایند یادگیری در محیط با تداخل روی کانال داده و پسخورد لحاظ شده است. این دو الگوریتم در الگوریتم ۱ و ۲ ارائه شده‌اند. در الگوریتم ۱، در پایان هر انتقال موفقیت‌آمیز داده، فرستنده پیام تصدیق (ACK) دریافت می‌کند و مادامی که این تصدیق دریافت می‌شود، مقدار پاداش تجمعی برای تصمیم مربوطه یک واحد افزایش می‌یابد [۲۲]. در نشانه‌گذاری استفاده‌شده، E_k بیان‌کنندهٔ مقدار انرژی مصرف‌شده برای ارسال بسته با استفاده از اقدام k ، E_{min} حداقل انرژی مصرف‌شده در میان اقداماتی که با موفقیت یک بسته را

الگوریتم ۱: یادگیری در محیط با تداخل روی کانال داده

```

1 Initialization:  $Z_k(1)=0, T_k(1)=1, \forall k \in \mathbb{A}$ ;
2 for  $t = 1, 2, \dots$  do
    - Update actions' values:
       $I_k(t) = Z_k(t) + \sqrt{\alpha \log(t)/T_k(t)}$ ;
    - Take the best action:  $\arg \max_{k \in \mathbb{A}} b_k(t) \rightarrow A(t)$ ;
    - Update reward counter:
       $Z_{A(t)}(t+1) = Z_{A(t)}(t) + \hat{z}(t)$ ;
    - Update visit counter:  $T_{A(t)}(t+1) = T_{A(t)}(t) + 1$  -
    return  $A(t)$ ;
```

الگوریتم ۲: یادگیری در محیط با تداخل روی کانال داده و پسخورد

```

1 Initialization:  $W_k(1) = 1, \forall k \in \mathbb{A}$ ;
2 for  $t = 1, 2, \dots$  do
    - Define distributions:
       $p_k(t) = (1-\rho) \frac{W_k(t)}{\sum_{j=1}^{|\mathbb{A}|} W_j(t)} + \frac{\rho}{|\mathbb{A}|}, \forall k \in \mathbb{A}$ ;
    - Take the best action:  $A(t) \sim \{p_1(t), \dots, p_{|\mathbb{A}|}(t)\}$ ;
    - Update the weight function:
       $W_{A(t)}(t+1) = W_{A(t)}(t) \exp(\frac{\rho \hat{z}(t)}{|\mathbb{A}| p_{A(t)}(t)})$ ;
    - return  $A(t)$ ;
```

۴-۱- انتقال یادگیری

راه حل ارائه شده در الگوریتم ۱ و ۲ یک راه حل توزیع شده است که براساس آن، هر دستگاه پارامترهای مخابراتی خود را با محیط تطبیق می دهد. در این قسمت، تکنیک انتقال یادگیری (transfer learning) برای افزایش سرعت همگرایی و در نتیجه، بهبود عملکرد سیستم بررسی می شود. الگوریتم ۱ در هنگام آماده سازی، ارزش هر تصمیم را برابر صفر و الگوریتم ۲ نیز احتمال گرفتن بیشترین پاداش از انتخاب هر تصمیم را برابر با سایر تصمیم ها در نظر می گیرد. حال اگر یک دستگاه جدید که شروع به فعالیت می کند از دستگاه های اطراف، نقطه دسترسی یا از سابقه ارتباطات قبلی، ارزش تصمیم ها را از یک مقدار اولیه آغاز کند، به این فرایند، انتقال یادگیری می گویند. انتقال یادگیری، چه از سمت دستگاه های قوی تر همسایه و چه از سمت شبکه، سرعت همگرایی الگوریتم یادگیری را به میزان زیادی افزایش می دهد که این افزایش با افزایش تعداد تصمیم ها چشمگیر خواهد بود.

در بخش بعدی، رویکرد یادگیری پیشنهاد شده برای کنترل اقدامات در دستگاه های اینترنت اشیا بررسی شده است که با فناوری لورا به یکدیگر متصل شده اند و نتایج به دست آمده با رویکرد بهینه متمرکز مقایسه شده است.

۵- مطالعه موردی: کاربرد در شبکه لورا

لورا، لایه فیزیکی پروتکل لورا^{۱۵}، تلاش می کند ارتباطات بی سیم در فواصل طولانی، با نرخ داده و مصرف انرژی پایین را میسر کند. ارتباطات در لورا در سه زیرکانال با پهنای باند هر کدام ۱۲۵ کیلوهرتز و در باند فرکانسی ISM صورت می گیرد. مقاومت بالا در برابر نویز و تداخل از ملزومات اساسی برای کار در باند فرکانسی ISM است؛ براساس این، روش مدولاسیون CSS^{۱۶} در فناوری لورا استفاده شده است. مدولاسیون CCC امکان دریافت و تشخیص سیگنال هایی با فاکتورهای پخش (SF) متفاوت، به صورت هم زمان را ممکن

می کند. فاکتور پخش در لورا بین ۷ تا ۱۲ تنظیم می شود که این اعداد به تعداد chirp های استفاده شده برای کد کردن یک بیت اشاره دارند؛ از این رو، نرخ داده برای کد c با استفاده از فرمول زیر محاسبه می شود:

$$R(c) = \frac{c \times BW \times \mu}{2c}, \forall c \in \mathbb{C} = \{7, \dots, 12\} \quad (4)$$

که در آن μ نرخ کد بین صفر و یک است. براساس [۲۷ و ۲۸]، سیگنال به نویز مورد نیاز برای تشخیص درست سیگنال هایی با فاکتور پخش $\{7, \dots, 12\}$ ، به ترتیب برابر با $\gamma_{th}(c) = \{-6, -9, -12, -15, -17.5, -20\}$ است. علاوه بر این، با افزایش c، میزان نرخ داده کاهش می یابد و تداخل نسبت به نویز از کاهش نرخ ارسال داده حاصل می شود. شایان توجه است در لورا^{۱۵} برای ارتباطات از توان های انتقال ۲، ۵، ۸، ۱۱ و ۱۴ (dBm) پشتیبانی می شود [۲۷ و ۲۸].

۵-۱- کنترل پارامترهای مخابراتی در لورا

یک شبکه لورا متشکل از یک نقطه دسترسی و تعدادی دستگاه متصل را فرض کنید که در آن دستگاهها براساس فرایند نقطه ای پواسن^{۱۷} (PPP) و با تراکم λ در یک محدوده پراکنده شده اند. هر فرستنده لورا به طور میانگین در هر T_{rep} ثانیه یک پاکت به گیرنده ارسال می کند. با در نظر گرفتن رابطه (۱)، برای تنظیم پارامترهای ارتباطی هر فرستنده، مسئله زیر باید حل شود:

$$\begin{aligned} & \max_{\{c_i, h_i, m_i, p_i\}} F(Rel_i, EE_i) \\ \text{s. t. } & p_i \in \{2, 5, 8, 11, 14\} \text{ dBm} \\ & c_i \in \{7, 8, 9, 10, 11, 12\}, h_i \in \{1, 2, 3\} \end{aligned} \quad (5)$$

۵-۱-۱- رویکرد توزیع شده

با به کارگیری الگوریتم های (۱)، ارائه شده در بخش ۴، مسئله بهینه سازی ارائه شده در رابطه (۵) حل می شود؛ براساس این، مجموعه \mathbb{A} شامل ۹۰ جفت تصمیم خواهد بود که هر جفت شامل یک توان ارسال، یک زیرکانال و یک کد است. نتایج استفاده از رویکرد توزیع شده در شکل ۲ نشان داده شده اند که در ادامه تحلیل خواهند شد.

¹⁷ Poisson point process (PPP)

¹⁵ LoRa Wide Area Network (LoRaWAN)

¹⁶ Chirp Spread Spectrum

۵-۱-۲- رویکرد متمرکز

این قسمت مربوط به تخصیص بهینه کدهای ارسال $c \in \{1, \dots, 6\}$ بین فرستنده‌ها است و تعداد سطح‌های توان ارسالی، تعداد بازارسال و تعداد زیرکانال برابر ۱ در نظر گرفته می‌شود. علاوه بر این، تخصیص کدهای ارسال شش‌گانه تنها نسبت به تداخل از سوی فرستنده‌های لورا بهینه می‌شوند و سایر منابع تداخل در باند ISM در نظر گرفته نمی‌شوند. با در نظر گرفتن این امر که با افزایش کد ارسال از ۶ به ۱۲، نرخ ارسال داده کاهش می‌یابد و احتمال تصادم و مقاومت در برابر نویز افزایش می‌یابد، انتظار می‌رود فرستنده‌های نزدیک به ایستگاه پایه مقادیر فاکتور پخش کوچک‌تری نسبت به فرستنده‌های دورتر انتخاب کنند و برعکس [۱۱]. براساس این، می‌توان ادعا کرد مسئله تخصیص کد ارسال معادل مسئله پیدا کردن چگالی بهینه فرستنده‌های با یک کد ارسال مشابه در هر نقطه از شبکه است. به منظور حل این مسئله در یک فضای دوبعدی، ناحیه سرویس دهی را مجموعه‌ای از ۶ حلقه تقسیم می‌کنیم که هر حلقه دارای شعاع داخلی $r_{j,1}$ و شعاع خارجی $r_{j,2}$ است و اختصاص به یک کد ارسال دارد. با در نظر گرفتن نتایج به دست آمده در [۲۹] و توسعه آنها، می‌توان تابع لاپلاس تداخل مربوط به دستگاه‌های توزیع شده در زمین حلقه که با $\Phi_{j,c} \subset \Phi$ نشان داده شده را به این صورت به دست آورد:

$$L_{\Phi_{j,c}}(s) = \exp\left(-2\pi \int_{r_{j,1}}^{r_{j,2}} \frac{\lambda_{j,c} T_c / T_{rep} r dr}{1 + s P_t G r^{-\delta}}\right), \quad (6)$$

که در آن $T(c) = \frac{D}{R(c)}$ ، C, D, T_{rep}, P_t به ترتیب اشاره به توان ارسالی، زمان بین دو ارسال پکت، طول بسته (بیت)، کد انتخابی برای ارسال و زمان انتقال دارند و $G r^{-\delta}$ نشان‌دهنده افت مسیر است. اکنون می‌توان تابع لاپلاس تداخل دریافتی از همه دستگاه‌هایی که از فاکتور پخش c استفاده می‌کنند را به صورت $\mathcal{L}_{\Phi_c}(s) = \prod_{j \in \mathbb{J}} \mathcal{L}_{\Phi_{j,c}}(s)$ نوشت. اگر نویز سیستم را با N و تداخل دریافتی از سوی فرستنده‌های گروه Φ را با I_{Φ} نشان دهیم، احتمال موفقیت در ارسال بسته برای یک فرستنده که از کد c استفاده می‌کند و در فاصله Z از ایستگاه پایه

قرار گرفته است، به صورت زیر به دست می‌آید:

$$\begin{aligned} p_s(c, z) &= \Pr(P_t h G z^{-\delta} \geq \gamma_c N) \Pr(P_t h G z^{-\delta} \geq \gamma_{th} I_{\Phi_c}), \\ &= L_{\Phi_c}(s) \Big|_{s=\frac{\gamma_c}{P_t G z^{-\delta}}} L_N(s) \Big|_{s=\frac{\gamma_{th}}{P_t G z^{-\delta}}}, \\ &= \prod_{j \in \mathbb{J}} \exp\left(\int_{r_{j,1}}^{r_{j,2}} \frac{-2\pi \lambda_{j,c} T_c / T_{rep} r dr}{1 + \left(\frac{r}{z}\right)^{\delta} \frac{1}{\gamma_{th}}}\right) \exp\left(-\frac{N \gamma_c(c) z^{\delta}}{P_t G}\right), \quad (7) \\ &= \prod_{j \in \mathbb{J}} \exp\left(-\lambda_{j,c} \frac{T_c}{T_{rep}} [Q(r_{j,2}) - Q(r_{j,1})]\right) \exp\left(-\frac{N \gamma_c z^{\delta}}{P_t G}\right) \end{aligned}$$

که $\gamma_c = \gamma_{th_N}(c)$ برابر سطح آستانه نویز تحمل‌شدنی برای کد c و γ_{th} برابر سطح آستانه تداخل تحمل‌شدنی است. همچنین، برای $\delta = 4$ داریم:

$$Q(x) = \pi \operatorname{atan}\left(\gamma_{th}^{-0.05} \frac{x^2}{z^2}\right) \sqrt{\gamma_{th}} z^2. \quad (8)$$

حال، مسئله (۳) به صورت زیر بازنویسی می‌شود:

$$\begin{aligned} \max_{\lambda_{j,c}} \sum_{j \in \mathbb{J}} \sum_{c \in \mathbb{C}} \frac{\lambda_{j,c}}{\lambda} \left[(1 - \beta) \int_{r_{j,1}}^{r_{j,2}} p_s(c, z) dz + \beta \frac{T_c}{T_c} \right] \\ \text{s.t: } \sum_{c \in \mathbb{C}} \lambda_{j,c} = \lambda. \quad (9) \end{aligned}$$

با حل این مسئله بهینه‌سازی، چگالی استفاده از هر کد در محیط به صورت تابعی از فاصله تا گیرنده به دست می‌آید. مشاهده می‌شود حل مسئله بهینه‌سازی به صورت متمرکز بسیار پیچیده است؛ حتی زمانی که برای ساده‌سازی آن فرض می‌شود دستگاه‌ها در یک فضای دوبعدی و با فرایند نقطه‌ای پواسن توزیع شده‌اند و تنها یک کانال لورا با یک سطح توان ارسال بدون نویز خارجی وجود دارد. در ادامه، کارایی راه‌حل بهینه متمرکز با رویکرد یادگیری توزیع شده مقایسه می‌شود.

۵-۱-۳- مقایسه رویکرد متمرکز و توزیع شده

شکل (۲) احتمال موفقیت در ارسال داده با رویکرد توزیع شده حاصل از الگوریتم ۱ و نتایج مربوط به استراتژی بهینه متمرکز را برای شرایط زیر نشان می‌دهد:

$$C = \{7, 10\}, \gamma_{th_N} = \{-6, -15\}, P_t = 14 \text{ dBm}, \\ N_d = 1000, T_{rep} = 200, D = 100 \text{ Bytes}.$$

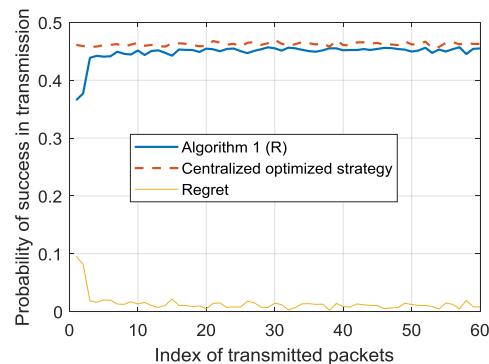
مقادیر مربوط به دیگر پارامترها در جدول ۱ آورده شده‌اند. در شکل (۲)، محور x نشان‌دهنده نمایه پکت‌های ارسال شده است. همان‌طور که در بالا نیز اشاره شد، هر فرستنده به‌طور مستقل

انتخاب می‌شود و توان ارسالی برابر ۱۴ dBm است؛ درحالی‌که نمایه (C,P) نشان می‌دهد توان و کد ارسالی هر دو با الگوریتم انتخاب می‌شوند. در پایان، الگوریتم ۵ نشان‌دهنده روش یادگیری به‌کاررفته در [۲۰ و ۱۹] است که برخلاف روش پیشنهادی ما، در تابع پاداش آن صرفه‌جویی در مصرف انرژی در نظر گرفته نشده است. شکل ۳، احتمال موفقیت در ارسال داده را برای ۲ الگوریتم ارائه‌شده در این کار (الگوریتم ۱ و ۲) و دو الگوریتم پایه برای سنجش (الگوریتم ۳ و ۴) نشان می‌دهد. محور x نشان‌دهنده نمایه پاکت‌های ارسالی (درواقع نشان‌دهنده سیر زمان) است. در بازه میانی از زمان، یک تداخل شدید در ۲ زیرکانال از مجموع ۶ شش زیرکانال اتفاق می‌افتد.

تصمیم می‌گیرد پاکت‌هایش را با استفاده از فاکتور پخش ۷ یا ۱۰ ارسال کند. بعد از تعداد کمی ارسال، مشهود است نتایج به‌دست‌آمده با استفاده از یادگیری توزیع‌شده بسیار نزدیک به رویکرد متمرکز شده است و ضرر ناشی از تصمیم‌گیری (regret) به صفر نزدیک می‌شود. باید توجه کرد یادگیری توزیع‌شده با وجود داشتن تفاوت کمی با رویکرد بهینه متمرکز در احتمال موفقیت، در عمل باعث بهبود چشمگیر در افزایش طول عمر باتری خواهد شد. به این علت که با این روش، فرستنده بدون نیاز به گوش دادن به سیگنال‌های کنترلی قادر است پارامترهای ارتباطی دستگاهها را به‌صورت توزیع‌شده تنظیم کند.

جدول (۱): پارامترهای ارزیابی عملکرد [۱۶]

مقدار	پارامتر
دایره‌ای با شعاع ۲ کیلومتر	ناحیه سرویس
۲/۵ بر ثانیه	نرخ تجمعی ورود بسته
۲۰ بایت	طول بسته
۱ زیرکانال	تعداد زیرکانال‌ها
۱۲۵ کیلوهرتز	پهنای باند: \mathcal{W}
۰/۸	نرخ کد: μ
{6,9,12,15,17.5,20}dB	آستانه سیگنال به نویز
6 dB	آستانه سیگنال به تداخل
{8,14}dBm, 10dBm, 2	توان مصرفی: P_1, P_0, η
۰/۱, ۰/۴, ۰/۴	پارامترهای یادگیری (پیش‌فرض) α, β, ρ



شکل ۲: مقایسه الگوریتم ۱ و رویکرد بهینه متمرکز در تخصیص کد به فرستنده‌های لورا

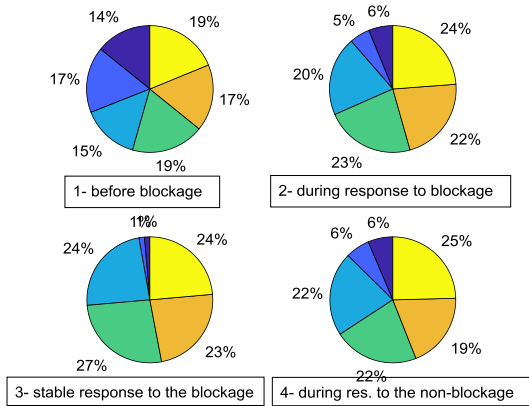
۶- ارزیابی عملکرد

در این بخش نتایج شبیه‌سازی عملکرد الگوریتم مبتنی بر یادگیری ماشینی در فناوری لورا ارائه شده‌اند. در این شبیه‌سازی، ۵۰۰ فرستنده در یک محیط دایروی با شعاع ۲ کیلومتر و به‌طور تصادفی پراکنده شده‌اند. هدف، توزیع دو سطح متفاوت توان ارسالی ۸ و ۱۴ (dBm) و انتخاب یک زیرکانال از ۶ زیرکانال در میان این فرستنده‌ها است. پارامترهای شبیه‌سازی در جدول ۱ آورده شده‌اند. در نتایج مربوط به شبیه‌سازی، الگوریتم ۱ و ۲ به الگوریتم‌های ارائه‌شده در بخش ۴ و الگوریتم ۳ به الگوریتم متمرکز ارائه‌شده در [۱۱] اشاره دارد که در آن، بار به‌طور مساوی بین زیرکانال‌ها توزیع می‌شود. الگوریتم ۴ نشان‌دهنده الگوریتم استفاده‌شده در [۱۸] است که در آن، زیرکانال به‌صورت تصادفی انتخاب می‌شود. در پایان، نمایه C در جلوی نام یک الگوریتم نشان می‌دهد که فقط کد ارسالی با آن الگوریتم

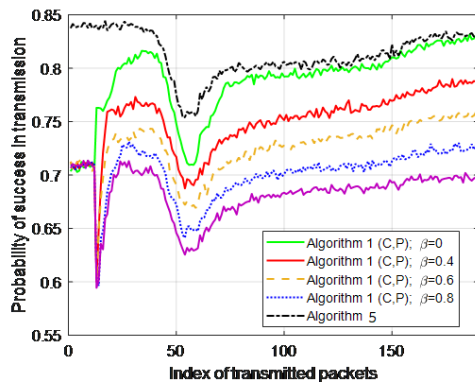
جدول (۲): مقایسه کلی روش‌های بررسی‌شده در مقاله

تابع هدف	محل انتخاب	انتخاب پارامتر	سال	نام روش
احتمال موفقیت	مرکزی	تقسیم مساوی	2019	مرجع [۱۱]
احتمال موفقیت	توزیع‌شده	تصادفی	2019	مرجع [۱۸]
احتمال موفقیت	توزیع‌شده	یادگیری	2019	مرجع [۱۹]
احتمال موفقیت	توزیع‌شده	یادگیری	2020	مرجع [۲۰]
احتمال موفقیت و مصرف انرژی	توزیع‌شده	یادگیری	2020	پیشنهادی

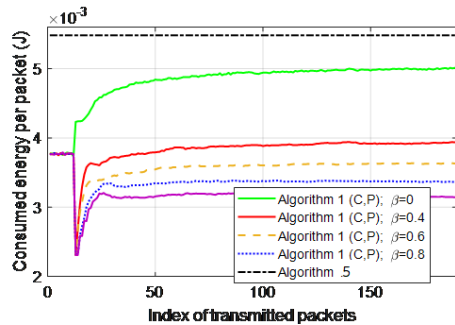
در شکل ۳، مدت زمان پاسخ به تداخل، طول تداخل و زمان پاسخ به رفع تداخل مشخص شده‌اند. مشاهده می‌شود الگوریتم مبتنی بر یادگیری ماشینی قادر است با سرعت خوب خود را با



شکل ۶: نحوه پاسخ الگوریتم ۱ به یک تداخل



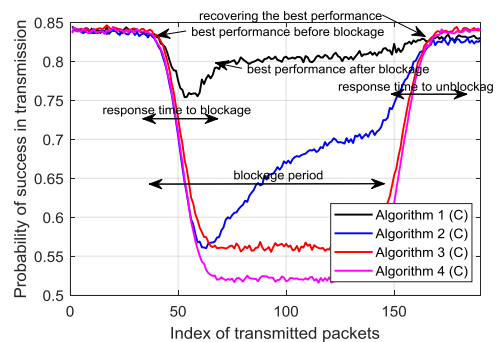
شکل ۷: احتمال موفقیت در ارسال پاکت برحسب زمان برای الگوریتم پیشنهادی با مقادیر مختلف β



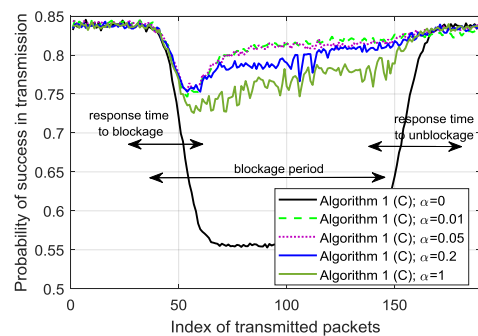
شکل ۸: مصرف انرژی در ارسال یک پاکت برحسب زمان برای الگوریتم پیشنهادی با مقادیر مختلف β

شکل ۴ نشان‌دهنده عملکرد الگوریتم ۱ برای مقادیر مختلف پارامتر α ، پارامتر ایجاد تعادل بین تجربه تصمیم‌های دیگر و استفاده از تجربه گذشته است. مشاهده می‌شود انتخاب یک مقدار بهینه برای این پارامتر که در اینجا ۰/۰۵ است، به سیستم کمک می‌کند تا در مقابل تداخل‌های احتمال، واکنش به موقع

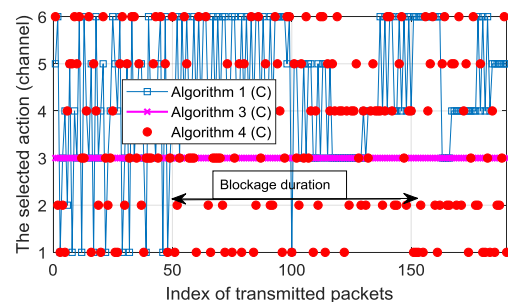
شرایط محیط تطبیق دهد و احتمال موفقیت را در ارسال داده، با وجود و بدون وجود تداخل، به مقدار بیشینه نزدیک کند. همچنین، مشاهده می‌شود الگوریتم ۲، در این شرایط که روی کانال پس‌خورده تداخل نداریم، با سرعت پایین همگرا می‌شود؛ چون به نتایج ACK دریافتی به اندازه الگوریتم ۱ اعتماد ندارد. همین نبودن اعتماد است که در شرایط جمینگ و تداخل روی کانال پس‌خورده، به این الگوریتم کمک می‌کند عملکرد بهینه‌تری نسبت به الگوریتم ۱ داشته باشد [۲۱].



شکل ۳: احتمال موفقیت در ارسال پاکت برحسب زمان برای الگوریتم‌های پیشنهادی و پایه



شکل ۴: احتمال موفقیت در ارسال پاکت برحسب زمان برای الگوریتم پیشنهادی با مقادیر مختلف α



شکل ۵: نمایه زیرکانال انتخاب‌شده

سرویس درخواستی، دقت لازم به کار برده شود.

۷- نتیجه گیری

در این پژوهش، راهکارهای ممکن برای بهبود قابلیت اطمینان و طول عمر باتری اینترنت اشیا بررسی شده است. کاهش سیگنالینگ بین فرستنده‌های اینترنت اشیا و شبکه دسترسی در شبکه‌های ارتباطی مبتنی بر دسترسی آزاد به منابع رادیویی، مانند لورا، به کاهش مصرف انرژی در ارسال داده‌ها و افزایش احتمال تصادم داده‌ها در انتقال منجر می‌شود. در این پژوهش، الگوریتم‌های یادگیری ماشینی با قابلیت پیاده‌سازی در اینترنت اشیا بررسی شده و یک راه‌حل مبتنی بر یادگیری توزیع شده با پیچیدگی کم برای پیاده‌سازی در فرستنده‌های اینترنت اشیا پیشنهاد شده است.

در این راه‌حل، از پاداش داخلی و خارجی به ترتیب برای کمینه‌سازی مصرف انرژی و احتمال تصادم در زمان انتقال داده روی کانال‌های اشتراکی استفاده شده است. سپس عملکرد این روش توزیع شده با روش مبتنی بر راه‌حل بهینه متمرکز توسعه داده شده با استفاده از هندسه تحلیلی، مقایسه شده و همگرایی الگوریتم توزیع شده تأیید شده است. نتایج حاصل از شبیه‌سازی شبکه ارتباطی لورا با کمک الگوریتم توزیع شده نشان‌دهنده بهبود چشمگیر در احتمال موفقیت در ارسال داده و افزایش طول عمر باتری دستگاهها است. این نتایج نشان می‌دهند در صورت استفاده از این روش یادگیری در دستگاههای اینترنت اشیا، این دستگاهها قادر به سازگاری خود با محیط پیرامون و تداخل احتمالی خواهند بود که نتیجه آن، افزایش قابلیت اطمینان در ارتباطات به صورت خودکار است.

مراجع

- [1] Cox, Christopher. "An introduction to LTE: LTE, LTE-advanced, SAE and 4G mobile communications," John Wiley & Sons, 2012.
- [2] C. Mavromoustakis, G. Matorakis, and J. M. Batalla, Internet of Things (IoT) in 5G mobile technologies. Springer, 2016.
- [3] M. Kulin, C. Fortuna, E. De Poorter, D. Deschrijver, and I. Moerman, "Data-driven design of intelligent wireless networks: An overview and tutorial," Sensors, vol. 16, no. 6, p. 790, June 2016.
- [4] A. Azari and G. Miao, "Network lifetime maximization for cellularbased M2M networks," IEEE Access, vol. 5, pp. 18 927–18 940, 2017.

نشان دهد و نیز در هنگام رفع تداخل به سرعت به حالت بهینه سابق برگردد.

شکل ۵ نشان‌دهنده نمایه تصمیم انتخاب شده با الگوریتم ۱، ۳ و ۴ برحسب زمان برای یک دستگاه خاص است. مشاهده می‌شود دستگاه با استفاده از الگوریتم ۱ پس از تجربه موفق بودن روی زیرکانال ۱ و ۲ دیگر به ندرت به آنها مراجعه می‌کند.

شکل ۶ نشان‌دهنده نحوه توزیع زیرکانال‌ها بین کاربران با استفاده از الگوریتم ۱ در قبل، حین و بلافاصله بعد از رفع تداخل است. مشاهده می‌شود کاربران قبل از وقوع تداخل روی کانال‌های مختلف به خوبی پخش شده‌اند. پس از وقوع تداخل در دو زیرکانال، استفاده از این دو زیرکانال به تدریج کم می‌شود. بلافاصله پس از رفع تداخل نیز مشاهده می‌شود کاربران دوباره به استفاده از این زیرکانال‌ها متمایل می‌شوند. دقت شود حتی در میانه تداخل شدید روی این دو زیرکانال، الگوریتم یادگیری ماشینی آنها را گاهی با ارسال داده بررسی می‌کند که این تکرار بررسی با پارامتر α تنظیم می‌شود.

شکل ۷ و ۸ نشان‌دهنده میزان موفقیت و مصرف انرژی با کمک الگوریتم ۱ و ۵ در حالتی است که علاوه بر زیرکانال ارسال داده، توان ارسال نیز با یادگیری ماشینی تنظیم شود. بر طبق رابطه (۳)، β تعادل بین بهره‌وری انرژی و احتمال موفقیت را در تابع هدف برقرار می‌کند. شکل ۷ نشان می‌دهد با افزایش تعداد تصمیم‌ها (در نظر گرفتن توان ارسال نیز به عنوان یک تصمیم)، زمان لازم برای همگرایی نیز در مقایسه با الگوریتم ۵ افزایش می‌یابد که تصمیم‌های کمتری دربردارد. این تأخیر در همگرایی، در پاسخ الگوریتم‌ها به تداخل ایجاد شده نیز درخور توجه است که نمودار سبز نسبت به سیاه عملکرد پایین‌تری دارد. شکل ۸ مصرف انرژی را برای الگوریتم ۱ و هریک از مقادیر β در مقایسه با الگوریتم ۵ را نشان می‌دهد. مشاهده می‌شود استفاده از الگوریتم یادگیری برای انتخاب سطح توان ارسال همواره به کاهش انرژی مصرفی منجر می‌شود و این کاهش انرژی با افزایش β ، پارامتر ایجاد تعادل در تابع هدف (۳)، افزایش می‌یابد. مقایسه تطبیقی شکل ۷ و ۸ نشان می‌دهد کاهش مصرف انرژی ممکن است با کاهش احتمال موفقیت همراه باشد؛ در نتیجه، باید در انتخاب مقدار مناسب برای پارامتر β با توجه به نوع کاربرد اینترنت اشیا مدنظر و کیفیت

- Technology for Vehicle and Asset Tracking in Smart Harbors." IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society. Vol. 1. IEEE, 2019.
- [19] Ta, Duc-Tuyen, et al. "LoRa-MAB: A Flexible Simulator for Decentralized Learning Resource Allocation in IoT Networks." 2019 12th IFIP Wireless and Mobile Networking Conference (WMNC). IEEE, 2019.
- [20] Hasegawa, So, et al. "Performance Evaluation of Machine Learning Based Channel Selection Algorithm Implemented on IoT Sensor Devices and Its Application to Wireless Sensor Network for Building Monitoring System." 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIC). IEEE, 2020.
- [21] Wang, Wenbo, et al. "Decentralized Learning for Channel Allocation in IoT Networks over Unlicensed Bandwidth as a Contextual Multi-player Multi-armed Bandit Game." arXiv preprint arXiv:2003.13314 (2020).
- [22] Azari, Amin, and Cicek Cavdar. "Self-organized Low-power IoT Networks: A distributed learning approach," 2018 IEEE Global Communications Conference (GLOBECOM), 2018.
- [23] M. S. Talebi Mazraeh Shahi, "Minimizing regret in combinatorial bandits and reinforcement learning," Ph.D. dissertation, KTH Royal Institute of Technology, 2017.
- [24] Xiao, Liang, et al. "IoT security techniques based on machine learning," arXiv preprint arXiv:1801.06275, 2018.
- [25] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine learning*, vol. 47, no. 2-3, pp. 235–256, 2002.
- [26] S. Bubeck et al., "The best of both worlds: stochastic and adversarial bandits," in *Conference on Learning Theory*, 2012, pp. 1–23.
- [27] R. Kl'ima et al., "Combining online learning and equilibrium computation in security games," in *International Conference on Decision and Game Theory for Security*. Springer, 2015, pp. 130–149.
- [28] B. Reynders, W. Meert, and S. Pollin, "Power and spreading factor control in low power wide area networks," in *IEEE ICC*, 2017.
- [29] O. Georgiou and U. Raza, "Low power wide area network analysis: Can LoRa scale?" *IEEE Wireless Communications Letters*, vol. 6, no. 2, pp.162–165, 2017.
- [30] A. Azari and C. Cavdar, "Performance Evaluation and Optimization of LPWA IoT Networks: A Stochastic Geometry Approach," 2018 IEEE GLOBECOM, 2018.
- [5] Nokia Networks, "LTE-M – optimizing LTE for the Internet of things," Tech. Rep., 2015.
- [6] W. Yang, M. Wang, J. Zhang, J. Zou, M. Hua, T. Xia, and X. You, "Narrowband wireless access for low-power massive internet of things: A bandwidth perspective," *IEEE Wireless Commun.*, vol. 24, no. 3, pp. 138–145, 2017.
- [7] E. Morin, M. Maman, R. Guizzetti, and A. Duda, "Comparison of the 'device lifetime in wireless networks for the internet of things," *IEEE Access*, vol. 5, pp. 7097–7114, 2017.
- [8] R1-163510, "Candidate NR Multiple Access Schemes," Tech. Rep., April 2016, 3GPP TSG RAN WG1 Meeting 84, Busan, Korea.
- [9] M. Lauridsen et al., "Interference measurements in the european 868 MHz ISM band with focus on LoRa and SigFox," in *IEEE WCNC*, 2017, pp. 1–6.
- [10] M. Masoudi, A. Azari, E. A. Yavuz, and C. Cavdar, "Grant-free Radio Access IoT Networks: Scalability Analysis in Coexistence Scenarios," in *IEEE ICC*, 2018.
- [11] F. Cuomo et al., "EXPLoRa: EXtending the Performance of LoRa by suitable spreading factor allocations," in *IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2017, pp. 1–5.
- [12] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, April 2017.
- [13] Y. J. Liu, S. M. Cheng, and Y. L. Hsueh, "eNB selection for machine type communications using reinforcement learning based markov decision process," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 12, pp. 11 330–11 338, Dec. 2017.
- [14] A. Azari, "Energy-efficient scheduling and grouping for machine-type communications over cellular networks," *Ad Hoc Networks*, vol. 43, pp. 16 – 29, 2016.
- [15] R. Bonnefoi, L. Besson, C. Moy, E. Kaufmann, and J. Palicot, "Multiarmed bandit learning in IoT networks: Learning helps even in nonstationary settings," in *CROWNCOM*, 2017.
- [16] Lei, Lei, et al. "Multiuser Resource Control With Deep Reinforcement Learning in IoT Edge Computing." *IEEE Internet of Things Journal* 6.6 (2019): 10119-10133.
- [17] Lei, Lei, et al. "Deep reinforcement learning for autonomous internet of things: Model, applications and challenges." arXiv preprint arXiv:1907.09059 (2019).
- [18] Priyanta, Irfan Fachrudin, et al. "Evaluation of LoRa