

تخصیص منبع آگاه از معنا با سازوکار هم‌زیستی چرخه کاری انعطاف‌پذیر برای شبکه‌های ناهمگون در شبکه‌های خودرویی متصل با استفاده از شبکه‌های بی‌سیم
نسل پنجم

نویسندگان: ژیهو شائو، چیونگ وو عضو ارشد IEEE، پینگ‌ی فان عضو ارشد IEEE، نان چنگ عضو ارشد IEEE،
چیانگ فان، و جیانگ‌ژو وانگ عضو برجسته IEEE.

چکیده

این پژوهش، چارچوبی برای تخصیص منابع آگاه به معنا¹ (SARA) به نام² (SARADC) را پیشنهاد می‌دهد که مجهز به سازوکار همزیستی چرخه کار³ (DC) انعطاف‌پذیر است. این چارچوب برای شبکه‌های ناهمگن در ارتباطات خودرویی نسل پنجم⁴ (5G-V2X) طراحی شده و بر پایه الگوریتم بهینه‌سازی سیاست مجاورتی⁵ (PPO) از خانواده یادگیری تقویتی عمیق⁶ (DRL) عمل می‌کند. به‌طور مشخص، ما شبکه‌های خودرویی متصل را در ساختاری ناهمگن را در سه‌لایه مورد مطالعه قرار می‌دهیم. برای برآورده ساختن الزامات شبکه‌های طراحی کرده و دو میار ارزیابی جدید شامل: نرخ انتقال معنایی با سرعت بالا⁷ (HSR) و بازدهی طیف معنایی (HSSE)⁸ را برای تخصیص منابع معرفی خواهیم کرد. علاوه بر این، مسئله همزیستی کاربران خودرویی و کاربران وای‌فای در طیف‌های بدون مجوز مربوط به 5G- (NR-U)⁹ مورد توجه قرار می‌گیرد. رویکرد ما به‌طور مشترک، سازوکار همزیستی DC و همچنین تخصیص منابع و ایستگاه‌های پایه¹⁰ (BSs) را بهینه می‌سازد. [۲] برخلاف روش‌های سنتی انتقال مبتنی بر بیت، رویکرد ما مخابرات معناگرا را به‌طور یکپارچه در سیستم ارتباطی ادغام می‌نماید. همچنین نتایج آزمایشگاهی نشان می‌دهند که چارچوب پیشنهادی ما، در مقایسه با روش‌های مرسوم، به‌طور قابل‌توجهی HSSE و توان عملیاتی معنایی¹¹ (ST) را برای هر دو گروه کاربران خودرویی و کاربران وای‌فای بهبود می‌بخشد.

واژگان کلیدی

ارتباط معناگرا، شبکه‌های خودرویی متصل، تخصیص منبع، باندهای بدون مجوز، همزیستی شبکه‌های نسل پنجم جدید بدون مجوز، یادگیری تقویتی عمیق.

مقدمه

¹ semantic-aware resource allocation

² semantic-aware resource Allocation duty cycle

³ duty cycle

⁴ fifth-generation vehicle to everything

⁵ proximal policy optimization

⁶ deep reinforcement learning

⁷ high-speed semantic transmission rate

⁸ high-speed semantic spectrum efficiency

⁹ 5G-new radio unlicensed

¹⁰ base stations

¹¹ semantic throughput

شبکه‌های نسل پنجم، با هدف ارائه خدمات ارتباطی با سرعت بالا، تأخیر پایین و قابلیت اطمینان بالا برای کاربردهایی نظیر ارتباطات خودرو با همه چیز^۱ (V2X) طراحی شده‌اند [۱]. با این حال، افزایش تعداد دستگاه‌ها، چالش‌هایی را در زمینه‌های ظرفیت و بازدهی طیف ایجاد کرده است. در مناطق شهری، سازوکارهای شبکه‌بندی سنتی ممکن است در پاسخگویی به این نیازها ناتوان باشند. استقرار سلول‌های کوچک^۲ و شبکه‌های ناهمگن با افزایش تعداد آنتن‌ها بر روی ایستگاه‌های پایه کوچک‌تر، ظرفیت شبکه را بهبود می‌بخشند [۲] پیشرفت‌های اخیر در حوزه مخابرات معناگرا، به جای داده‌های خام، بر معنا و مفهوم اطلاعات تمرکز دارند و بدین ترتیب، بازدهی طیف را بهبود می‌بخشند [۳]. مطالعات متعددی به بررسی کاربرد این رویکرد در حوزه‌های مختلف از جمله تصاویر، متن، صوت، خدمات واقعیت ترکیبی، و سنجش بی‌سیم پرداخته‌اند. [۴]–[۸]

فناوری 5G-NRU، ظرفیت شبکه‌های ناهمگن نسل پنجم را با استفاده از باندهای فرکانسی بدون مجوز افزایش می‌دهد [۹]؛ ولی با این حال، به دلیل تداخل، باید سازوکارهای همزیستی با سایر شبکه‌های بی‌سیم مانند وای‌فای، از قبیل گوش دادن پیش از مکالمه [۱۰]^۳ (LBT) و انتقال تطبیقی با حسگر حامل (CSAT)^۴ [۱۱]، مورد توجه قرار گیرند. علی‌رغم این تدابیر، عدم قطعیت دسترسی به کانال در NR-U در مقایسه با باندهای دارای مجوز، منجر به عملکرد ضعیف‌تر می‌شود. از این رو، سازوکارهای سنتی همزیستی NR-U و وای‌فای چالش جدیدی را برای شبکه‌های نسل پنجم به وجود آورده‌اند. علاوه بر این، تخصیص منابع سنتی مبتنی بر بیت در شبکه‌های خودرویی، اطلاعات معنایی را نادیده می‌گیرد که موجب بهره‌وری ضعیف طیف و تراکم می‌شود [۱۲]. اگرچه یادگیری تقویتی عمیق امکان بهینه‌سازی پویا را فراهم می‌سازد. [۱۳]، اما همچنان بر چارچوب‌های مبتنی بر بیت متکی است. بهبود بازدهی طیف نیازمند اتخاذ یک رویکرد معناگرا و رسیدگی به مسائل همزیستی-NR-U و وای‌فای است. در حالی که [۱۴] به تخصیص منابع در مخابرات معناگرا پرداخته است، اما فاقد یک چارچوب جامع آگاه به معنا برای شبکه‌های ناهمگن شبکه‌های خودرویی نسل پنجم است و همچنین با ماهیت پویای شبکه‌های خودرویی سازگاری ندارد یا به بهینه‌سازی مشترک چرخه کار و تخصیص منابع نمی‌پردازد. برای حل مشکل فوق‌الذکر، این نامه پژوهشی برای نخستین بار مخابرات معناگرا را در یک سیستم ارتباطی خودرویی شبکه ناهمگن سه‌لایه در محیط‌های شهری معرفی می‌کند. ما یک رویکرد چرخه کار (DC) انعطاف‌پذیر را برای رسیدگی به مسائل همزیستی میان کاربران خودرویی و کاربران وای‌فای در شبکه‌های NR-U پیشنهاد می‌کنیم و سنجه‌های نرخ انتقال معنایی با سرعت بالا (HSR) و بازدهی طیف معنایی با سرعت بالا (HSSE) را در بهینه‌سازی

1 vehicle to everything

2 small Cells

3 listen before talk

4 carrier sensing adaptive transmission

تخصیص منابع وارد می‌نماییم. ما الگوریتم چارچوب تخصیص منابع آگاه به معنا با سازوکار همزیستی DC انعطاف‌پذیر (SARADC) را طراحی می‌کنیم که از الگوریتم بهینه‌سازی سیاست مجاورتی (PPO) از خانواده DRL برای بهینه‌سازی چرخه کار انعطاف‌پذیر و تخصیص منابع، بر مبنای آگاهی معنایی جهت حداکثرسازی توان عملیاتی معنایی (ST) استفاده می‌کند. نتایج آزمایشگاهی نشان می‌دهند که الگوریتم پیشنهادی ما از نظر سنجه‌های HSSE و ST برای هر دو گروه کاربران خودرویی و کاربران وای‌فای، عملکردی بهتر از سایر روش‌های مرجع دارد.

مدل سازی سیستم

همان‌طور که در تصویر ۱ نشان داده شده است، تمرکز ما بر روی ارتباطات خودرو با همه چیز در یک محیط شهری پرسرعت است. در این مدل، N تعداد وسایل نقلیه با سرعت ثابت V و در مسیرهای متغیر در حال حرکت هستند که منجر به ایجاد N لینک خودرو با زیرساخت $^{1}(V2I)$ می‌شود. وضعیت کانال هر t_u میلی‌ثانیه به‌روزرسانی می‌شود.

ما یک ساختار شبکه ناهمگن سه‌لایه 2 را در نظر می‌گیریم:

۱. لایه اول B_1 : شامل ایستگاه‌های پایه ماکرو $^{3}(MaBs)$ است که پوشش گسترده و ارتباطات با ظرفیت بالا را در باندهای فرکانسی دارای مجوز 4 فراهم می‌کنند.

۲. لایه دوم B_2 : شامل ایستگاه‌های پایه میکرو $^{5}(MiBs)$ است که در باندهای فرکانسی بدون مجوز عمل کرده و خدمات با سرعت بالا و تأخیر کم ارائه می‌دهند.

۳. لایه سوم P : شامل نقاط دسترسی وای‌فای است که این‌ها نیز از باندهای بدون مجوز استفاده می‌کنند.

$B = \{B_1, B_2\}$ مجموعه ایستگاه‌های پایه (BSs) را نشان می‌دهد که $MaBs$ و $MiBs$ به ترتیب دارای R_1 و R_2 بلوک منابع هستند.

1 Vehicle to instruments

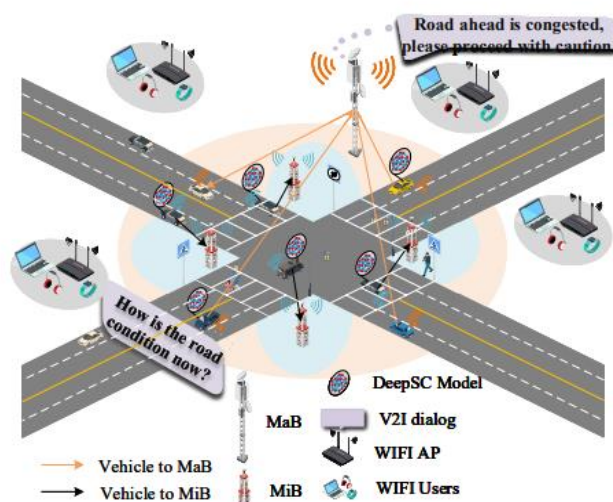
2 Three-tier HetNets

3 macro base station

4 Licensed Frequency Bands

5 micro base station

N وسیله نقلیه، B ایستگاه پایه (BSs) و P نقطه دسترسی وای فای، همگی مجهز به مدل های¹ (DeepSC) هستند. مدل DeepSC شامل یک فرستنده، کانال، و گیرنده است. فرستنده شامل کدگذاری معنایی² و کدگذاری کانال³ است، در حالی که گیرنده شامل کدگشایی معنایی⁴ و کدگشایی کانال⁵ با استفاده از مدل ترانسفورمر⁶ همراه با دانش پس زمینه مشترک⁷ می باشد. برای مثال، یک خودروی مجهز به DeepSC، عبارت "وضعیت جاده چگونه است؟" را به اطلاعات معنایی کدگذاری می کند، که ایستگاه پایه ماکرو (MaB) آن را به "جاده پیش رو پرترافیک است، لطفاً با احتیاط ادامه دهید" کدگشایی می کند. این مدل بر روی MaB از پیش آموزش داده شده است و سپس مدل های فرستنده معنایی DeepSC آموزش دیده، به دلیل پیچیدگی استخراج اطلاعات معنایی در طول فرآیند آموزش، پخش شده و توسط تمامی وسایل نقلیه به کار گرفته می شوند.



شکل (۱) سیستم تخصیص منابع آگاه از معنا

1 Deep Learning Enabled Semantic Communication Systems

2 Semantic Encoding

3 Channel Encoding

4 Semantic Decoding

5 Channel Decoding

6 Transformer Model

7 Shared Background Knowledge

الف (فرستنده-گیرنده‌های DeepSC و معیارهای نوین

خودروی فرستنده‌ی n اُم جمله‌ای S_n با l_q کلمه تولید می‌کند، به طوری که:

$$S_n = [s_{n,1}, s_{n,2}, \dots, s_{n,l}, \dots, s_{n,l_q}] \quad 1-1$$

این جمله توسط بخش کدگذاری DeepSC پردازش می‌شود تا اطلاعات معنایی X_n از S_n استخراج گردد:

$$X_n = ch_\beta(se_\alpha(S_n)) \quad 1-2$$

که در آن $se_\alpha(0)$ و $ch_\beta(0)$ شبکه‌های کدگذار معنایی و کانال، به ترتیب با پارامترها α و β هستند. بردار نماد معنایی به صورت ۱-۳ تعریف می‌شود،

$$X_n = [x_{n,1}, x_{n,2}, \dots, x_{n,l}, \dots, x_{n,l_q}] \quad 1-3$$

که در آن u نشان‌دهنده‌ی میانگین تعداد نمادهای معنایی استفاده شده برای هر کلمه است. سپس، اطلاعات معنایی کدگذاری شده از طریق کانال بی‌سیم ارسال می‌شود تا سیگنال معنایی دریافت شده Y_n حاصل شود:

$$Y_n = H_{n,i} X_n + \zeta \quad 1-4$$

که در آن $H_{n,i}$ بهره‌ی کانال برای i اُمین لینک $V2I$ و ζ نویز است. بنابراین، نسبت سیگنال به تداخل به نویز^۱ (SINR) برای n اُمین خودرو برابر است با:

$$SINR_{n,b,r} = \frac{\eta_{n,b,r} P_{n,b,r} H_{n,b,r}}{D_{n,b,r} + \delta^2} \quad 1-5$$

که در آن $\eta_{n,b,r}$ یک شاخص باینری است؛ $\eta_{n,b,r}=1$ به این معنی است که n اُمین خودرو داده‌ی معنایی را به b اُمین ایستگاه پایه (BS) با استفاده از r اُمین بلوک منابع (RB) ارسال می‌کند.

¹ Signal to noise ratio

$P_{n,b,r}$ و $H_{n,b,r}$ به ترتیب توان ارسالی و بهره کانال از n آمین خودرو به آمین BS در آمین RB هستند و β^2 است .
 $D_{n,b,r}$ **تداخل** روی آمین RB در آمین BS است که توسط سایر BS ها (شامل هر دو BS های ماکرو و BS های میکرو اعمال شده و به صورت زیر تعریف می گردد):

$$D_{n,b,r} = \begin{cases} \sum_{\hat{n}=1}^N \sum_{\hat{b}=1}^{B_2} \eta_{\hat{n},\hat{b},r} p_{\hat{n},\hat{b},r} H_{\hat{n},b,r} & \hat{n} \neq n \\ \sum_{\hat{n}=1}^N \sum_{\hat{b}=1}^{B_1} \eta_{\hat{n},\hat{b},r} p_{\hat{n},\hat{b},r} H_{\hat{n},b,r} & \hat{n} = n \end{cases} \quad ۱-۶$$

پس سیگنال رمزگشایی شده بصورت زیر نمایش داده می شود :

$$\hat{S}_n = se_{\mu}^{-1}(ch_v^{-1}(Y_n)) \quad ۱-۷$$

se_{μ}^{-1} و ch_v^{-1} به ترتیب شبکه های رمزگشای معنایی و رمزگشای کانال می باشند .

برای ارزیابی عملکرد ارتباطات معناگرا، از تشابه معنایی ζ به عنوان یک معیار عملکرد استفاده می شود. این تشابه با استفاده از بازنمایی های رمزگذار دو طرفه ترنسفورمر (BERT) محاسبه می شود و به صورت زیر تعریف می گردد:

$$\xi = \frac{B(s_n)B(\hat{s}_n)}{\|B(s_n)B(\hat{s}_n)\|} \quad ۱-۸$$

همچنین مقادیر HSE و HSSE نیز بصورت زیر تعریف می‌شود :

$$HSR = W \frac{I}{uL} \xi \quad 1-9$$

$$HSSE = \frac{HSR}{W} = \frac{I}{u_q L} \xi \quad 1-10$$

مقدار $\frac{I}{L}$ که متأثر از نوع منبع است، ثابت می‌ماند و در طول فرآیند بهینه‌سازی قابل صرف‌نظر است. پارامتر تشابه معنایی ξ به صورت تابعی از u (تعداد واحدهای معنایی برای هر کلمه) و نسبت سیگنال به تداخل به نوبت نمایش داده می‌شود.

ب) مکانیزم چرخه وظیفه انعطاف‌پذیر

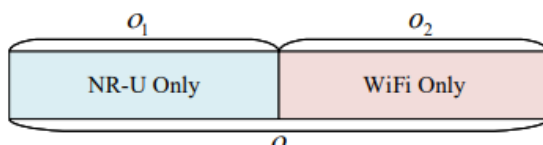
از آنجایی که خودروها از فناوری NR-U برای ارسال معنایی استفاده می‌کنند و WiFi از طیف بدون مجوز بهره می‌برد، احتمال وقوع تداخل (Collisions) به دلیل همپوشانی باندهای فرکانسی وجود دارد.

شکل ۲ نشان می‌دهد که یک بازه زمانی ثابت O به دو بخش تقسیم می‌شود:

- در بخش O_1 خودروها به طور انحصاری بسته‌های داده معنایی را از طریق NR-U ارسال می‌کنند.

- در بخش O_2 ، کاربران WiFi به طیف بدون مجوز دسترسی پیدا می‌کنند.

مدت زمان بخش‌های نام برده شده بر اساس نیازمندی‌های شبکه، قابل انطباق و تغییر است.



شکل ۲) مکانیزم SADC

مسئله‌ی بهینه‌سازی

هدف این مسئله، بهینه‌سازی راندمان طیفی معنایی با سرعت بالا (HSSE) به صورت میانگین است.

این بهینه‌سازی باید به صورت مشترک بر روی چهار متغیر کلیدی انجام شود:

پارامتر تخصیص کانال تصمیم می‌گیرد که کدام خودرو از کدام بلوک منابع (RB) استفاده کند.

پارامتر تخصیص توان تعیین سطح توان ارسالی هر خودرو.

پارامتر دوره زمانی انحصاری تنظیم مدت زمانی که NR-U بدون تداخل WIFI داده‌های معنایی ارسال می‌کند.

$$P_0 : \max_{\eta, b, r, p, \rho_1, \mu} \frac{HSSE_n}{N} \quad 1-10-1$$

$$s.t. \sum_{n=1}^N ST_n \geq ST_n \quad 1-10-2$$

$$\sum_{n=1}^N ST_w \geq ST_w \quad 1-10-3$$

$$\sum_{n=1}^N \eta_{n,b,r} \leq 1 \quad 1-10-4$$

$$\forall b \in B, \forall r \in R, \quad 1-10-5$$

$$\xi \geq \xi_{th}, \quad 1-10-6$$

$$u \in \{0, 1, \dots, u_{\max}\} \quad 1-10-7$$

هدف اصلی این تحقیق، بهینه‌سازی راندمان طیفی معنایی با سرعت بالا (HSSE) به صورت میانگین در محیط پویای شبکه‌های خودرویی است. این راندمان، یک معیار حیاتی است که کیفیت انتقال اطلاعات (سنجیده شده با تشابه معنایی) را با کارایی مصرف منابع طیفی ترکیب می‌کند. برای دستیابی به این هدف چندوجهی، لازم است که سیستم به صورت همزمان و هوشمندانه چهار متغیر اصلی را تنظیم و بهینه کند. این متغیرها شامل تخصیص کانال (تصمیم‌گیری گسسته در مورد اینکه کدام خودرو از کدام بلوک منابع استفاده کند)، تخصیص توان (تنظیم توان ارسالی هر خودرو)، دوره زمانی انحصاری (تنظیم مدت زمان اختصاصی برای انتقال معنایی از طریق طیف بدون مجوز) و تعداد نمادهای معنایی در هر کلمه هستند که عامل اصلی در تعیین میزان فشردگی معنایی و کاهش سرشار انتقال است. به دلیل ماهیت پیچیده‌ی این مسئله که شامل متغیرهای

گسسته و پیوسته است و همچنین وابستگی شدید هدف (HSSE) به ساختار شبکه عصبی (تشابه معنایی) می‌باشد این مسئله به یک مسئله بهینه‌سازی غیرخطی و ترکیبی تبدیل می‌شود که حل آن با روش‌های سنتی دشوار است؛ از این رو، استفاده از رویکردهای هوش مصنوعی پیشرفته مانند یادگیری تقویتی عمیق (DRL) برای یافتن سیاست‌های بهینه‌ی بلندمدت و مؤثر ضروری است.

رای حل مسئله‌ی بهینه‌سازی (حداکثرسازی راندمان طیفی معنایی)، ما الگوریتم SARADC را پیشنهاد می‌کنیم که از یادگیری تقویتی عمیق (DRL)، به‌ویژه الگوریتم PPO، برای بهینه‌سازی تخصیص منابع در شبکه‌های پرسرعت خودرویی استفاده می‌کند.

وضعیت، اقدام و پاداش عامل (Agent)

۱. وضعیت (State): در هر لحظه، هر خودرو (عامل) وضعیت محیط را مشاهده می‌کند. این وضعیت شامل پارامترهای بلندمدت مانند بهره‌ی کانال لحظه‌ای، نسبت سیگنال به تداخل به نویز (SINR)، راندمان طیفی معنایی (HSSE) کنونی خودِ خودرو، HSSE شبکه‌ی WIFI در لحظه قبل، و میزان تداخلی است که در لحظه قبل از سایر خودروها دریافت کرده است.

۲. اقدام (Action): بر اساس وضعیت مشاهده شده، هر عامل یک اقدام را انتخاب می‌کند. این اقدام شامل چهار تصمیم کلیدی است: انتخاب ایستگاه پایه و بلوک منابع، تعیین توان ارسالی، تعیین نسبت چرخه وظیفه O_1 برای انتقال انحصاری NR-U می‌باشد.

۳. پاداش (Reward): پس از اجرای اقدام، عامل یک پاداش دریافت می‌کند تا کیفیت تصمیم‌های ارزیابی شود. پاداش به‌گونه‌ای طراحی شده است که هدف اصلی سیستم (افزایش عملکرد) را تأمین کند، اما در عین حال، جریمه‌هایی نیز برای اقداماتی که منجر به نقض محدودیت‌ها می‌شوند (مانند تخصیص همزمان یک منبع به چندین کاربر) در نظر گرفته می‌شود.

فرآیند یادگیری PPO

الگوریتم SARADC از ساختار Actor-Critic در PPO استفاده می‌کند. در این ساختار، شبکه‌ی Actor مسئول انتخاب بهترین اقدام (Policy) و شبکه‌ی Critic مسئول ارزیابی ارزش وضعیت (Value Function) است.

آموزش شامل دو حلقه‌ی اصلی است: حلقه‌ی بیرونی برای تعداد کل اپیزودها (Episodes)، و حلقه‌ی داخلی برای گام‌های زمانی درون هر اپیزود. در هر گام زمانی، عامل وضعیت را مشاهده، اقدام را بر اساس سیاست قبلی

انتخاب، و پاداش و وضعیت بعدی را دریافت می‌کند. این تجربیات در یک بافر بازپخش (Replay Buffer) ذخیره می‌شوند. سپس از این داده‌ها برای محاسبه‌ی تابع برتری (Advantage Function) استفاده می‌شود تا شبکه‌های Actor و Critic آموزش ببینند و سیاست بهینه به مرور زمان بهبود یابد.

Algorithm 1: SARADC Algorithm solving P_0

Input: $\pi, \theta, \theta^v, \theta_{old}, \theta_{old}^v$
Output: the optimal policy parameter π^*
Initialize parameters $\theta_{old} = \theta, \theta_{old}^v = \theta^v$;
Initialize replay experience buffer R ;
for episode from 1 to E_{max} **do**
 for t from 1 to t_{max} **do**
 Receive observation state s_n^t ;
 Select action $a_n^t = \pi(a_n^t | s_n^t; \theta_{old})$ using the old policy;
 Observe reward r_n^t and next state s_n^{t+1} ;
 Store tuple $(s_n^t, a_n^t, r_n^t, s_n^{t+1})$ in R ;
 Compute advantages A_t according to (13)
 Normalize A_t
 if t can be exactly divided by t_u **then**
 for epochs from 1 to K_{max} **do**
 Randomly sample a batch of bs experience from the replay buffer R ;
 Compute actor loss according to (14)
 Compute critic loss according to (15)
 Compute total loss according to (16)
 Update actor-critic parameters θ and θ^v according to (17) and (18);
 Update old actor-critic parameters θ_{old} and θ_{old}^v with θ and θ^v .

شکل ۳) الگوریتم SARADC برای حل P

$$S_n^t = [h_{n,b,r}^t, SINR_{n,b,r}^t, HSSE_{nn,b,r}^t, HSSE_{wn,b,r}^{t-1}, D_{n,b,r}^{t-1}] \quad 1-11$$

$$a_n^t = [\eta_{n,b,r}^t, P_n^t, o_{1n}^t, \mu_n^t] \quad 1-12$$

$$r_n^t = \frac{\sum_{n=1}^N ST_n}{N \underline{ST_n}} \omega(ST_w, \underline{ST_w}) - C \Psi(\eta_{n,b,r}^t) \quad 1-13$$

$$\Psi(\eta_{n,b,r}^t) = \{1, \sum_{r=1}^R \eta_{n,b,r}^t \succ 1, c = penalty\} \quad 1-14$$

$$1-15$$

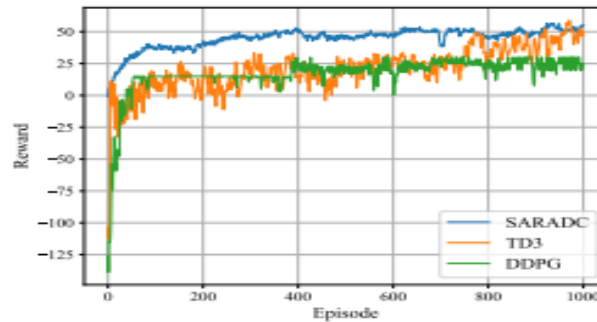
$$L_{actor}(\theta) = - \min (rt(\theta)At, \text{clip} (rt(\theta), 1 - \epsilon, 1 + \epsilon) At)$$

اگر گام زمانی فعلی مضربی از بازه‌ی تعیین‌شده‌ی به‌روزرسانی باشد، الگوریتم برای بهینه‌سازی شبکه‌های Actor (مسئول سیاست‌گذاری) و Critic (مسئول ارزیابی ارزش) وارد یک حلقه‌ی داخلی تکرار (Epoch) می‌شود که تا بار اجرا می‌گردد. در طول هر دوره‌ی آموزش، یک دسته‌ی از تجربیات کامل (شامل وضعیت‌ها، اقدامات، پاداش‌ها و وضعیت‌های بعدی) از بافر بازپخش نمونه‌برداری می‌شود. سپس، با استفاده از این داده‌ها، تابع زیان (Loss Function) شبکه‌ی Actor محاسبه و به‌روزرسانی می‌شود. این تابع زیان با هدف بهینه‌سازی پاداش‌های مورد انتظار طراحی شده است، در حالی که از طریق مکانیزم برش (Clipping)، از تغییرات ناگهانی و شدید در سیاست جدید جلوگیری می‌کند. این محدودیت تضمین می‌کند که سیاست جدید (Actor) بیش از حد از سیاست قدیمی فاصله نگیرد، و این امر پایداری فرآیند یادگیری در شبکه‌های پویا را تضمین می‌کند. همزمان، تابع زیان شبکه‌ی Critic نیز با هدف به حداقل رساندن خطا در تخمین مقدار حالت (Value Function) به‌روزرسانی می‌شود تا ارزیابی دقیقی از پاداش‌های آتی ارائه دهد و به شبکه‌ی Actor در جهت‌گیری صحیح کمک کند. پس از تکمیل این حلقه‌ی داخلی، پارامترهای جدید شبکه‌های Actor و Critic به عنوان پارامترهای جاری ثبت می‌شوند.

$$L_{actor}(\theta) = - \min (rt(\theta)At, \text{clip} (rt(\theta), 1 - \epsilon, 1 + \epsilon) At)$$

جدول ۱) پارامترهای سیستم مدل و SARADC

Parameters of System Model			
Parameter	Value	Parameter	Value
W	15 KHz	RB number	12
u_{\max}	5-25 bits/word	Vehicle height	1.5 m
BS height	25 m	ξ_{th}	0.9
Parameters of SARADC			
Parameter	Value	Parameter	Value
lr	0.0001	∂	(0.9, 0.99)
E_{\max}	1000	t_{\max}	100
t_u	5	K_{\max}	5
γ	0.99	ϵ	0.2



شکل ۴) پاداش در طول یک دوره آموزش

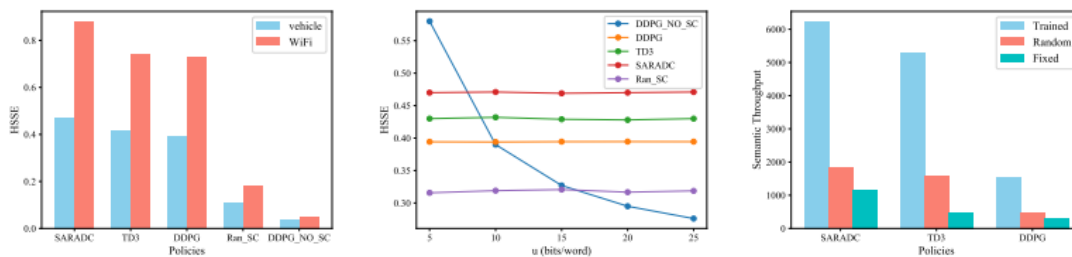
نتایج شبیه سازی

ما الگوریتم SARADC را با استفاده از پایتون ۳.۷ ارزیابی می‌کنیم. سناریوی شبیه‌سازی شامل ۵ خودرو در منطقه‌ای به ابعاد ۱۰۰۰ متر مربع با سرعت ثابت ۳۶ کیلومتر بر ساعت است. وضعیت کانال هر ۵ میلی‌ثانیه به‌روزرسانی می‌شود، که ضمن تعادل با پیچیدگی محاسباتی، تغییرات واقع‌بینانه اطلاعات حالت کانال (CSI) را منعکس کرده و یک محیط آموزشی پایدار برای مدل DRL فراهم می‌کند. تلفات مسیر برابر است با $\log(d)$ که d فاصله بین خودروها و ایستگاه‌های پایه (BSs) است. نرخ یادگیری LR نرخ کاهش و سایر پارامترهای مناسب مطابق جدول ۱ تنظیم شده‌اند.

برای ارزیابی عملکرد الگوریتم پیشنهادی SARADC با قابلیت تخصیص دوز مصرفی منعطف (Flexible DC)، آن را با پنج روش پایه زیر مقایسه می‌کنیم:

- SARA-DDPG تخصیص منابع آگاه به معنا (SARA) مبتنی بر DDPG با DC منعطف.

- SARA-TD3: SARA مبتنی بر TD3 با DC منعطف.
- Ran SC: SARA با انتخاب تصادفی و DC منعطف.
- Fixed و Random: SARA مبتنی بر PPO با DC ثابت (Fixed) و DC تصادفی (Random).
- DDPG NO SC الگوریتم تخصیص منابع با بیت واحد، مبتنی بر DDPG با DC منعطف، که نمادهای معنایی را در نظر نمی‌گیرد.



شکل ۵) نتایج شبیه سازی

نتیجه‌گیری

این مقاله یک چارچوب SARADC را پیشنهاد می‌کند که برای شبکه‌های ناهمگن (ارتباطات خودرویی) با استفاده از الگوریتم یادگیری تقویتی عمیق PPO طراحی شده و ارتباطات معنایی را در شبکه‌های پرسرعت خودرویی ادغام می‌کند. نتایج اصلی به این صورت خلاصه می‌شوند: اولاً، SARADC با بهره‌گیری از داده‌های معنایی، به کارایی معنایی بسیار بالاتر (HSSE) دست یافته و استفاده کارآمد از منابع را تضمین می‌کند. ثانیاً، تخصیص دوز مصرفی ثابت (Fixed DC) یا تصادفی (Random DC) به دلیل عدم سازگاری، منجر به کاهش توان عملیاتی کلی سیستم می‌شود؛ در حالی که DC منعطف پیشنهادی ما با تقاضا سازگار شده و عملکرد را افزایش می‌دهد. نهایتاً، مزایای انتقال اطلاعات معنایی به وضوح زمانی آشکار می‌شود که داده‌ها با کدگذاری سنتی به بیش از ۸ بیت نگاشت می‌شوند.

منابع

- [1] Q. Wu, S. Shi, Z. Wan, Q. Fan, P. Fan and C. Zhang. "Towards V2I Age-aware Fairness Access: A DQN Based Intelligent Vehicular Node Training and Test Method", Chinese J. of Electronics, vol. 32, no. 6, 2023, pp. 1230-1244.
- [2] Q. Wu, W. Wang, P. Fan, Q. Fan, J. Wang and K. B. Letaief. "URLLC-Aware Resource Allocation for Heterogeneous Vehicular Edge Computing," IEEE Trans. on Veh. Technol., doi: 10.1109/TVT.2024.3370196.
- [3] X. Luo, H. -H. Chen and Q. Guo. "Semantic Communications: Overview, Open Issues, and Future Research Directions," IEEE Wireless Commun., vol. 29, no. 1, pp. 210-219, Feb. 2022.
- [4] E. Bourtsoulatz, D. Burth Kurka and D. Gunduz. "Deep Joint Source-Channel Coding for Wireless Image Transmission," IEEE Trans. Cogn. Commun. and Netw., vol. 5, no. 3, pp. 567-579, Sept. 2019.
- [5] Y. Wang, M. Chen, W. Saad, T. Luo, S. Cui and H. V. Poor. "Performance Optimization for Semantic Communications: An Attention-based Learning Approach," in IEEE Global Communications Conference, Madrid, Spain, pp. 1-6, 2021.
- [6] T. Han, Q. Yang, Z. Shi, S. He and Z. Zhang. "Semantic-Preserved Communication System for Highly Efficient Speech Transmission," IEEE J. Sel. Areas in Commun., vol. 41, no. 1, pp. 245-259, 2023.
- [7] H. Du, J. Wang, D. Niyato, J. Kang, Z. Xiong and D. I. Kim. "AI-Generated Incentive Mechanism and Full-Duplex Semantic Communications for Information Sharing," IEEE J. on Sel. Areas in Commun., vol. 41, no. 9, pp. 2981-2997, Sept. 2023.
- [8] H. Du, J. Wang, D. Niyato, J. Kang, Z. Xiong, J. Zhang and X. Shen. "Semantic Communications for Wireless Sensing: RIS-Aided Encoding

and Self-Supervised Decoding,” IEEE J. on Sel. Areas in Commun., vol. 41, no. 8, pp. 2547-2562, Aug. 2023.

[9] N. Patriciello, S. Lagen, B. Bojović and L. Giupponi. “NR-U and IEEE 802.11 Technologies Coexistence in Unlicensed mmWave Spectrum: Models and Evaluation,” IEEE Access, vol. 8, pp. 71254-71271, 2020.

[10] M. Haghshenas and M. Magarini. “NR-U and Wi-Fi Coexistence Enhancement Exploiting Multiple Bandwidth Parts Assignment,” in 2022

IEEE 19th Annual Consumer Communications and Networking Conference, Las Vegas, NV, USA, pp. 260-263, 2022.

[11] B. Yin, H. Hu, B. Xi, Q. Liu, Y. Zheng and Z. Zhang. “Joint Radio Resources Allocation in the Coexisting NR-U and Wi-Fi Networks,” 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications, Helsinki, Finland, pp. 1532-1538, 2021.

[12] M. Noor-A-Rahim, Z. Liu, H. Lee, G. G. M. N. Ali, D. Pesch and P. Xia. “A Survey on Resource Allocation in Vehicular Networks,” IEEE Trans. on Intell. Transp. Syst., vol. 23, no. 2, pp. 701-721, Feb. 2022.

[13] W. Chen, X. Qiu, T. Cai, H. -N. Dai, Z. Zheng and Y. Zhang. “Deep Reinforcement Learning for Internet of Things: A Comprehensive Survey,” IEEE Commun. Surv. Tut., vol. 23, no. 3, pp. 1659-1692, thirdquarter 2021.

[14] L. Yan, Z. Qin, R. Zhang, Y. Li and G. Y. Li. “Resource Allocation for Text Semantic Communications,” IEEE Wireless Commun. Lett., vol. 11, no. 7, pp. 1394-1398, July 2022.

[15] H. Xie, Z. Qin, G. Y. Li and B. -H. Juang. “Deep Learning Enabled Semantic Communication Systems,” IEEE Trans. on Signal Processing.,

vol. 69, pp. 2663-2675, 2021.