

بهینه‌سازی شبکه‌های G5 مبتنی بر هوش مصنوعی: مروری جامع بر تخصیص منابع، مدیریت ترافیک و برش پویا شبکه

Dileesh Chandra Bikkasani^{1, *}, Malleswar Reddy Yerabolu²

¹Department of Technology Management, University of Bridgeport, Bridgeport, USA

²Independent Researcher, North Carolina, USA

چکیده

پیشرفت سریع شبکه‌های G5 همراه با پیچیدگی روزافزون مدیریت منابع، کنترل ترافیک و نیازهای پویا در سرویس‌ها، ضرورت استفاده از روش‌های هوشمند برای بهینه‌سازی شبکه را برجسته می‌کند. این مقاله به‌طور جامع روش‌های مبتنی بر هوش مصنوعی برای بهینه‌سازی شبکه‌های G5 را مرور می‌کند، با تمرکز بر تخصیص منابع، مدیریت ترافیک و برش پویا شبکه. مدل‌های سنتی در سازگاری با ماهیت پویا و پیچیده مخابرات مدرن محدودیت دارند، در حالی‌که تکنیک‌های هوش مصنوعی — به‌ویژه یادگیری ماشین (ML) و یادگیری تقویتی عمیق (DRL) — راهکارهای مقیاس‌پذیر و تطبیقی ارائه می‌دهند. این رویکردها امکان بهینه‌سازی بلادرنگ را با یادگیری از شرایط شبکه، پیش‌بینی الگوهای ترافیک و مدیریت هوشمند منابع در اسلایس‌های مجازی فراهم می‌کنند. ادغام هوش مصنوعی در شبکه‌های G5 عملکرد را ارتقا می‌دهد، تأخیر را کاهش می‌دهد و بهره‌وری پهنای‌بند را تضمین می‌کند؛ که برای پشتیبانی از کاربردهای نوظهور مانند اینترنت اشیا، سیستم‌های خودران و واقعیت افزوده ضروری است. این مقاله همچنین تکنیک‌های کلیدی هوش مصنوعی و کاربردهای آن‌ها در چالش‌های G5 را برجسته کرده و نشان می‌دهد چگونه می‌توانند نوآوری‌های آینده در مدیریت شبکه را هدایت کنند. با ایجاد زمینه برای عملیات خودکار شبکه در G6 و فراتر از آن، این پژوهش بر تأثیر تحول‌آفرین هوش مصنوعی بر زیرساخت‌های مخابراتی و نقش آن در شکل‌دهی آینده اتصال تأکید می‌کند.

کلیدواژه‌ها

G5، مخابرات، ارتباطات بی‌سیم، هوش مصنوعی، عملکرد شبکه

1. مقدمه

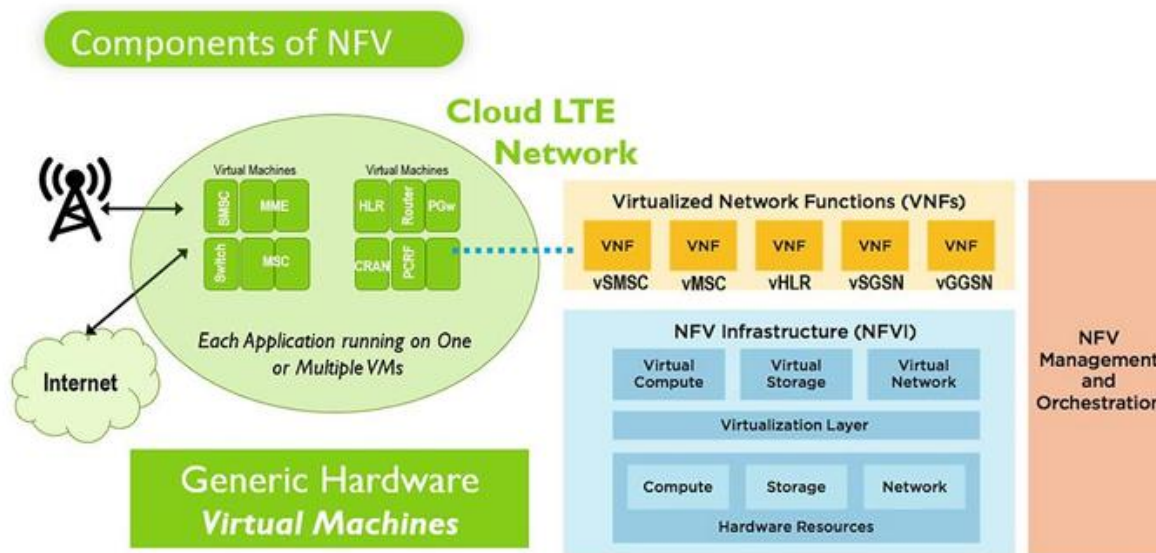
اتصال بسیار قابل‌اعتماد و با تأخیر بسیار پایین است. ارتباطات ماشینی انبوه (mMTC): پشتیبانی از تعداد بسیار زیادی دستگاه متصل، به‌ویژه در حوزه اینترنت اشیا. پهنای‌بند موبایل ارتقا یافته (eMBB): تمرکز بر افزایش پهنای‌بند برای دسترسی اینترنت پرسرعت.

ارتباطات URLLC به دنبال ارائه اتصال بسیار قابل‌اعتماد و با تأخیر پایین هستند؛ eMBB بر افزایش پهنای‌بند برای دسترسی اینترنت پرسرعت تمرکز دارد؛ و mMTC از تعداد زیادی دستگاه متصل پشتیبانی می‌کند و امکان پیاده‌سازی اینترنت اشیا در مقیاس وسیع را فراهم می‌سازد [2].

با تکامل ارتباطات بی‌سیم، پیشرفت‌های چشمگیری در حوزه مخابرات رخ داده است و تقاضای داده از نسل چهارم (G4) تا نسل پنجم (G5) هزار برابر افزایش یافته است [1]. هر نسل جدید کاستی‌های نسل قبلی را برطرف کرده و بهبودهای چشمگیری ارائه داده است. به‌ویژه، ظهور فناوری نسل پنجم شبکه‌های بی‌سیم (G5) نوید سرعت‌های بی‌سابقه، تأخیر فوق‌العاده پایین و اتصال همزمان چندین دستگاه را می‌دهد. استاندارد جدید 5G-NR (New Radio) به سه دسته خدمات متمایز تقسیم می‌شود: ارتباطات فوق‌قابل‌اعتماد با تأخیر کم (URLLC): هدف آن ارائه

شبکه G5 بر پایه فناوری‌هایی مانند طیف امواج میلی‌متری (mmWave)، سیستم‌های MIMO عظیم (Massive MIMO) و مجازی‌سازی عملکرد شبکه (NFV) [3] بنا شده است و وعده‌ی تحول در بسیاری از صنایع را می‌دهد.

بهینه‌سازی عملکرد G5 برای کاربردهای نوظهور مانند خودروهای خودران، چندرسانه‌ای، واقعیت افزوده و واقعیت مجازی (AR/VR)، اینترنت اشیا، ارتباطات ماشین به ماشین (M2M) و شهرهای هوشمند حیاتی است.



شکل ۱. اجزای NFV (مجازی‌سازی عملکرد شبکه)

برای نیازهای خاص برنامه‌هایی مانند خودروهای خودران یا پزشکی از راه دور — که به قابلیت اعتماد بالا و تأخیر کم نیاز دارند — بهینه کنند. شکل ۱ عناصر معماری NFV را نشان می‌دهد، از جمله لایه مجازی‌سازی، منابع سخت‌افزاری و بخش مدیریت و ارکستراسیون که کنترل تخصیص و مقیاس‌دهی منابع را بر عهده دارند. NFV نقشی محوری در فعال‌سازی Network Slicing دارد؛ قابلیت‌های حیاتی در G5 که به اپراتورها اجازه می‌دهد شبکه‌های مجازی متناسب با نیازهای خاص هر کاربرد ایجاد کنند.

با این حال، پیچیدگی و ناهمگونی شبکه‌های G5 چالش‌های متعددی ایجاد می‌کند، از جمله تضمین کیفیت سرویس (QoS)، مدیریت منابع و بهینه‌سازی شبکه. با گسترش شبکه‌های G5، رویکردهای سنتی مبتنی بر قوانین برای مدیریت شبکه ناکافی می‌شوند. تخصیص کارآمد منابع، مدیریت ترافیک و برش پویا شبکه برای پاسخگویی به موارد استفاده سنگین بدون کاهش سرعت یا قابلیت اعتماد ضروری است [4]. علاوه بر این، با افزایش جریان ترافیک

شکل ۱ اجزای مجازی‌سازی عملکرد شبکه (NFV) را نشان می‌دهد که یکی از قابلیت‌های کلیدی در G5 است. NFV عملکردهای شبکه را از سخت‌افزار اختصاصی جدا می‌کند و اجازه می‌دهد این عملکردها به صورت نرم‌افزار روی سخت‌افزار استاندارد اجرا شوند. با مجازی‌سازی عملکردهایی مانند فایروال‌ها، بالانس‌کننده‌های بار و دروازه‌ها، NFV مدیریت شبکه را پویا و مقیاس‌پذیر می‌سازد و تخصیص منابع را در اسلایس‌های مختلف شبکه و کاربردهای گوناگون آسان‌تر می‌کند. این انعطاف‌پذیری برای مدیریت تقاضاهای روبه‌رشد برنامه‌های G5 حیاتی است، جایی که سازگاری بلندرنج و بهینه‌سازی منابع اهمیت بالایی دارد.

اهمیت NFV در توانایی آن برای جداسازی نرم‌افزار از سخت‌افزار نهفته است، به گونه‌ای که اپراتورهای شبکه می‌توانند خدمات جدید را سریع‌تر ارائه دهند و خدمات موجود را کارآمدتر مقیاس‌دهی کنند. برای مثال، در یک شبکه G5، اپراتورها می‌توانند منابع را به طور پویا به عملکردهای مجازی شبکه (VNFs) تخصیص دهند و آن‌ها را

موبایل، برآورده سازی به موقع نیازهای مشتری مستلزم رسیدگی به تخصیص پهنای باند برای خدمات ابری ناهمگون است [5].

تخصیص منابع شبکه (RA) در G5 نقشی حیاتی در بهینه سازی استفاده کارآمد از طیف، توان محاسباتی و انرژی دارد تا نیازهای ارتباطات بی سیم مدرن را برآورده کند. تخصیص منابع برای برنامه های داده محور، دستگاه های IoT و فناوری های نوظهوری مانند خودروهای خودران (AV) و واقعیت افزوده (AR) ضروری است. این فرآیند تضمین می کند که این فناوری ها منابع کافی شبکه را دریافت کنند و عملکرد کلی و QoS در محیط های پویا و ناهمگون ارتقا یابد. روش های سنتی تخصیص منابع به اطلاعات وضعیت کانال (CSI) متکی هستند که هزینه های سربار قابل توجهی ایجاد می کند و هزینه کلی فرآیند را افزایش می دهد [6].

بخش ۲ مقاله بر تکنیک های مختلف هوش مصنوعی در مدیریت منابع تمرکز دارد و تأثیر آن ها بر برش شبکه، بهره وری انرژی و کیفیت کلی سرویس (QoS) را برجسته می کند. با بهره گیری از یادگیری تقویتی (RL)، روش های بهینه سازی و مدل های یادگیری ماشین (ML)، این راهبردهای پیشرفته نیازهای پویا و پیچیده شبکه های G5 را برطرف کرده و راهکارهای تطبیقی و هوشمند برای اتصال بهتر و پایداری ارائه می دهند.

برش شبکه امکان ایجاد چندین شبکه مجازی بر روی یک زیرساخت فیزیکی مشترک را فراهم می کند، که هرکدام برای نیازهای خاص سرویس بهینه سازی می شوند.

شبکه های مجازی (Network Slicing) می توانند به طور مستقل برای پشتیبانی از کاربردهای متنوع با نیازهای عملکردی متفاوت پیکربندی شوند؛ مانند ارتباطات با تأخیر کم، اتصال انبوه دستگاه ها یا خدمات داده با توان عملیاتی بالا. با استفاده از برش شبکه، G5 می تواند تجربه های متناسب با انواع مختلف کاربران ارائه دهد و در عین حال استفاده از منابع شبکه را به حداکثر برساند. این قابلیت برای موارد نوظهوری مانند کارخانه های هوشمند، پزشکی از راه دور و سیستم های خودران حیاتی است، جایی که نیازهای

عملکردی می تواند به طور قابل توجهی بین کاربردها متفاوت باشد.

ساختار مقاله به این صورت است: بخش اول درک کلی موضوع و ضرورت هوش مصنوعی را توضیح می دهد؛ بخش دوم روش های سنتی هوش مصنوعی و یادگیری ماشین را معرفی می کند؛ و بخش سوم شامل تکنیک های مبتنی بر یادگیری عمیق است.

2. داده کاوی و تکنیک های هوش مصنوعی برای تخصیص منابع

هوش مصنوعی در تخصیص منابع نتایج امیدوارکننده ای نشان داده است، زیرا از طریق یادگیری مداوم و سازگاری با تغییرات شبکه می تواند کارایی را افزایش دهد. برخلاف مدل های سنتی مبتنی بر معادلات ریاضی، یادگیری تقویتی (RL) با رویکرد داده محور منابع را بهینه تخصیص می دهد و باعث بهبود نرخ داده و کاهش تأخیر می شود [7]. با این حال، دستیابی به تخصیص کاملاً توزیع شده منابع به دلیل محدودیت های جداسازی مجازی در هر اسلایس شبکه امکان پذیر نیست [8].

در شبکه های G LTE4 و LTE Advanced (LTE/A)، سوئیچینگ بسته های IP عمدتاً بر مدیریت تعداد متغیر کاربران در یک ناحیه تمرکز دارد. با استفاده از MIMO مبتنی بر ML برای پیش بینی کانال و تخصیص منابع، دقت CSI افزایش یافته و تأخیر پردازش کاهش می یابد، هرچند این روش در برابر پیچیدگی و بار ترافیکی G5 ناکارآمد است [9]. تخصیص منابع سنتی مبتنی بر CSI با سربار سیستم مواجه است که می تواند تا ۲۵٪ ظرفیت کل سیستم را مصرف کند و برای کاربردهای 5G Cloud RAN (CRAN) بهینه نیست [10].

روش های سنتی بهینه سازی شامل الگوریتم های تقریبی برای اتصال کاربران نهایی به Remote Radio Heads (RRH) هستند. این الگوریتم تعداد کاربران متصل به هر RRH را تخمین زده و ارتباط بین کاربران، RRH ها و واحدهای پردازش پایه (BBU) را برقرار می کند [11]. چالش هایی مانند Beamforming در امواج میلی متری (mmWave) و بهینه سازی تأخیر با استفاده از Beamforming ترکیبی برای

سناریوهای تک‌کاربره و چندکاربره نیز بررسی شده‌اند [12]. در نمونه‌ای دیگر، الگوریتم Random Forest برای تخصیص منابع پیشنهاد شد که خروجی‌های یک طبقه‌بند دودویی را اعتبارسنجی می‌کند؛ هرچند نیازمند تحقیقات بیشتر است [13].

بسیاری از تکنیک‌های یادگیری تقویتی عمیق (DRL) برای برش شبکه در G5 به کار گرفته شده‌اند و امکان تخصیص پویا منابع را فراهم کرده‌اند که باعث افزایش نرخ داده و کاهش تأخیر می‌شود [14]. برای مثال، Balevi و Gitlin (2018) الگوریتم خوشه‌بندی‌ای پیشنهاد کردند که با استفاده از تکنیک‌های ML کارایی شبکه را افزایش داده و تخصیص منابع را بر اساس شرایط بلادرنگ شبکه تنظیم می‌کند [15]. همچنین طرحی مبتنی بر شبکه‌های عصبی کانولوشنی گرافی (GCN) برای تخصیص منابع بررسی شد.

روش‌های بهینه‌سازی مانند الگوریتم‌های ابتکاری (Heuristic) و الگوریتم‌های ژنتیک (GA) نیز برای حل مسائل تخصیص منابع به کار گرفته می‌شوند تا تداخل را کاهش داده و بهره‌وری طیفی را افزایش دهند. الگوریتم‌های ژنتیک با اصول انتخاب، ترکیب و جهش راه‌حل‌ها را به سمت پیکربندی‌های بهینه تخصیص منابع تکامل می‌دهند. روش‌های ابتکاری مانند Simulated Annealing و Particle Swarm Optimization (PSO) نیز برای بهبود مدیریت منابع استفاده می‌شوند.

ادغام الگوریتم‌های مبتنی بر هوش مصنوعی مانند RL و DRL در شبکه‌های G5 امکان تخصیص منابع تطبیقی و بلادرنگ را بر اساس شرایط شبکه و نیازهای کاربران فراهم می‌کند و عملکرد شبکه را به طور قابل توجهی بهبود می‌بخشد [16].

به طور کلی، تکنیک‌های هوش مصنوعی و یادگیری ماشین در حال متحول کردن تخصیص منابع در شبکه‌های G5 هستند. با حرکت از مدل‌های سنتی به رویکردهای داده‌محور و تطبیقی مانند RL و DRL، این فناوری‌ها می‌توانند نرخ داده را افزایش دهند، تأخیر را کاهش دهند و سربار سیستم را به طور کارآمد مدیریت کنند. در حالی که روش‌های سنتی با پیچیدگی و تقاضای روزافزون کاربران دچار مشکل می‌شوند،

بهینه‌سازی مبتنی بر AI راهکارهای پویا ارائه می‌دهد که با شرایط بلادرنگ شبکه سازگار می‌شوند و مدیریت منابع را در عصر G5 و فراتر از آن کارآمدتر می‌سازند.

3. داده‌کاوی و هوش مصنوعی در مدیریت ترافیک

یادگیری ماشین (ML) کاربردهای متنوعی در مدیریت ترافیک شبکه دارد؛ از پیش‌بینی حجم ترافیک و شناسایی تهدیدات امنیتی گرفته تا بهینه‌سازی مهندسی ترافیک (TE). این قابلیت‌ها امکان پایش فعال شبکه، تقویت اقدامات امنیتی و بهبود مدیریت جریان ترافیک را فراهم می‌کنند و در نتیجه عملیات شبکه کارآمدتر و مقاوم‌تر می‌شود [18]. با ترکیب اطلاعات زمانی، مکانی و فرکانسی، پژوهشگران پنج الگوی پایه در حوزه زمان برای ترافیک موبایل شناسایی کرده‌اند که با انواع مختلف مناطق شهری مانند مسکونی، تجاری و مراکز حمل‌ونقل مطابقت دارند [19].

مدل‌های سنتی مانند ARIMA به طور گسترده برای مدل‌سازی فصلی استفاده شده‌اند، زیرا می‌توانند وابستگی‌های زمانی در داده‌های سری زمانی را مدل کنند. مدل ARIMA ترکیبی از اجزای خودرگرسیون (AR)، انتگرال‌گیری (I) و میانگین متحرک (MA) است که مقادیر آینده را بر اساس مشاهدات گذشته پیش‌بینی می‌کند [20]. نسخه‌های توسعه‌یافته مانند SARIMA از تحلیل طیفی برای توصیف الگوهای ترافیک و برآورد پارامترها با روش بیشینه درست‌نمایی استفاده می‌کنند [21]. با وجود اثربخشی، ARIMA و SARIMA اغلب در برابر الگوهای غیرخطی و پیچیده ترافیک که ویژگی شبکه‌های G5 هستند ناکارآمدند.

مدل‌های یادگیری ماشین مانند SVM و Random Forest برای رفع این محدودیت‌ها به کار گرفته شده‌اند. SVM روابط غیرخطی را به ویژه در حالت رگرسیون (SVR) مدل می‌کند [22]. الگوریتم Random Forest مجموعه‌ای از درخت‌های تصمیم را با داده‌ها و ویژگی‌های انتخاب شده تصادفی آموزش می‌دهد و سپس آن‌ها را ترکیب می‌کند تا دقت و مقاومت پیش‌بینی افزایش یابد [23]. این روش ناهمگونی ترافیک شبکه G5 را به خوبی مدیریت می‌کند.

یادگیری عمیق (DL) بسیاری از جنبه‌های مدیریت ترافیک شبکه را متحول کرده است؛ از پیش‌بینی و برآورد ترافیک

گرفته تا مسیریابی هوشمند. مدل‌هایی مانند LSTM به‌طور ویژه مؤثر هستند، زیرا می‌توانند وابستگی‌های بلندمدت در داده‌های ترتیبی را یاد بگیرند و برای پیش‌بینی ترافیک شبکه بسیار مناسب‌اند. DL همچنین جایگزین‌های نویدبخشی برای مدیریت تداخل، مدیریت طیف، استفاده چندمسیره، سازگاری لینک، دسترسی چندکاناله و کنترل ازدحام ارائه می‌دهد [24]. برای مثال، یک زمان‌بند مبتنی بر شبکه عصبی با دو لایه پنهان کاملاً متصل توانست برخوردها را در یک شبکه حسگر بی‌سیم پنج‌گره‌ای تا ۵۰٪ کاهش دهد [25].

روش‌های پیشرفته مانند معماری‌های مبتنی بر Transformer با استفاده از مکانیزم توجه (Self-Attention) داده‌های حجیم را به‌طور کارآمد پردازش می‌کنند [26]. تکنیک‌های DRL نیز برای زمان‌بندی ترافیک‌های حجیم و انعطاف‌پذیر (HVFT) در شبکه‌های موبایل پیشنهاد شده‌اند. این مدل با استفاده از الگوریتم DDPG سیاست کنترلی برای زمان‌بندی ترافیک IoT و سایر ترافیک‌های با تأخیر قابل‌تحمل یاد می‌گیرد و هدف آن بهینه‌سازی سرویس‌دهی به HVFT در عین کمینه‌سازی افت کیفیت برای ترافیک‌های حساس به تأخیر است [27].

مطالعات متعددی کاربردهای نوآورانه DRL و چارچوب‌های AI را در این زمینه نشان داده‌اند. یک مطالعه رویکرد DRL برای زمان‌بندی مکان‌یابی مشارکتی غیرمتمرکز در شبکه‌های خودرویی معرفی کرد [28]. چارچوبی مبتنی بر CNN و RNN توانست توان عملیاتی را حدود ۳۶٪ افزایش دهد، هرچند هزینه زمان آموزش و مصرف حافظه بالایی داشت [29]. مدل دیگری مبتنی بر LSTM به ایستگاه‌های پایه کوچک امکان داد به‌طور پویا به طیف بدون مجوز دسترسی پیدا کنند و انتخاب کانال بی‌سیم را بهینه کنند [30]. همچنین یک رویکرد DRL برای بهینه‌سازی مسیریابی در SDN به نتایجی مشابه روش‌های سنتی با تأخیر کم دست یافت [31].

کار روی مسیریابی و مدیریت تداخل که معمولاً به الگوریتم‌های پرهزینه‌ای مانند WMMSE وابسته است، با تقریب این الگوریتم‌ها توسط شبکه‌های عصبی کوچک پیشرفت کرده و پتانسیل بالایی برای بهبود سیستم‌های Massive MIMO نشان داده است.

خلاصه: ادغام فناوری‌های هوش مصنوعی در مدیریت ترافیک شبکه‌های G5 پیشرفت‌های چشمگیری در حوزه‌هایی مانند پیش‌بینی ترافیک، تخصیص منابع و مدیریت شبکه ایجاد کرده است. تکنیک‌هایی مانند ML و DL با استفاده از مدل‌هایی مانند LSTM و چارچوب‌های پیشرفته مبتنی بر CNN، RNN و DRL ماهیت پیچیده و پویا شبکه‌های G5 را پوشش می‌دهند. راهکارهای مبتنی بر AI کارایی و قابلیت اعتماد شبکه را با بهبود مدیریت تداخل، دسترسی طیف و قابلیت‌های مسیریابی ارتقا می‌دهند و خود را با الگوهای متغیر ترافیک سازگار می‌کنند. این نوآوری‌ها پتانسیل تحول‌آفرین AI را در دستیابی به عملیات شبکه مقاوم، تطبیقی و کارآمد در G5 برجسته می‌کنند و مسیر تحقیقات آینده را هموار می‌سازند.

خودکارسازی برش پویا شبکه با استفاده از هوش مصنوعی امکان‌پذیر است. رویکرد سنتی و دستی مدیریت شبکه برای محیط‌های بزرگ و بسیار ناهمگون که توسط G5 ایجاد می‌شوند کافی نیست. هوش مصنوعی، به‌ویژه یادگیری ماشین (ML)، قابلیت‌های پیشرفته‌ای در تصمیم‌گیری بلادرنگ، تحلیل پیش‌بینانه و کنترل تطبیقی ارائه می‌دهد که برای استقرار و مدیریت کارآمد اسلایس‌های شبکه حیاتی هستند.

مدل‌های هوش مصنوعی می‌توانند الگوهای ترافیک را پیش‌بینی کنند، شرایط شبکه را تحلیل نمایند و تخصیص منابع را به‌طور پویا تنظیم کنند تا نیازهای خاص هر اسلایس برآورده شود. این امر تضمین می‌کند که اسلایس‌ها حتی در شرایط ترافیک متغیر و تقاضاهای متفاوت سرویس، عملکرد بهینه خود را حفظ کنند. الگوریتم‌های یادگیری تقویتی (RL) و یادگیری عمیق (DL) اغلب برای مدیریت فرآیندهای پیچیده تصمیم‌گیری مورد نیاز در ارکستراسیون اسلایس‌ها استفاده می‌شوند. این الگوریتم‌ها می‌توانند به‌طور خودکار از داده‌های شبکه یاد بگیرند، منابع را بهینه کنند و بار را بین اسلایس‌ها بدون دخالت انسانی متعادل سازند.

تخصیص منابع مبتنی بر هوش مصنوعی نقش حیاتی در موفقیت برش شبکه دارد. هر اسلایس ممکن است نیازهای متفاوتی از نظر پهنای باند، تأخیر و قابلیت اعتماد داشته باشد، بنابراین تخصیص پویا منابع ضروری است. هوش

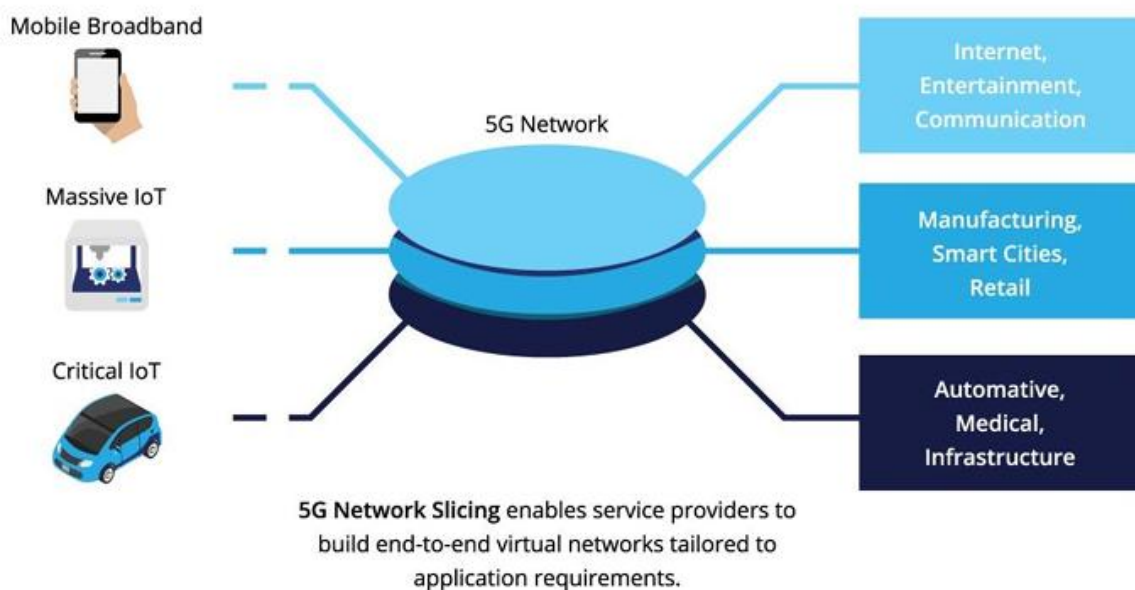
دریافت منابع مناسب توسط هر اسلایس برای حفظ توافق‌نامه‌های سطح سرویس (SLAs) است.

مدیریت ترافیک در برش شبکه یکی دیگر از حوزه‌هایی است که هوش مصنوعی در آن برتری دارد. تنوع خدمات در یک شبکه 5G—مانند eMBB، URLLC و IoT انبوه—نیازمند اولویت‌بندی هوشمند ترافیک است. الگوریتم‌های هوش مصنوعی الگوهای ترافیک را در زمان واقعی تحلیل می‌کنند و به سیستم امکان می‌دهند اسلایس‌هایی را که به تأخیر کمتر یا قابلیت اعتماد بالاتر نیاز دارند به‌طور خودکار در اولویت قرار دهد. این مدیریت پویا تضمین می‌کند که خدمات حیاتی مانند خودروهای خودران یا جراحی‌های از راه دور نسبت به کاربردهای کم‌اهمیت‌تر مانند پخش ویدئو در اولویت قرار گیرند.

مصنوعی می‌تواند بر اساس داده‌های تاریخی و الگوهای ترافیک بلادرنگ منابع را پیش‌بینی و پیش‌تخصیص دهد. برای مثال، الگوریتم‌های ML مانند شبکه‌های عصبی می‌توانند زمان‌های اوج ترافیک برای سرویس‌های خاص را پیش‌بینی کنند و تخصیص منابع را به‌صورت پیشگیرانه انجام دهند تا از ازدحام جلوگیری شود.

یادگیری تقویتی، به‌ویژه در محیط‌های چندعاملی (Multi-Agent Reinforcement Learning – MARL)، نیز برای تخصیص منابع در برش شبکه محبوب شده است. MARL به موجودیت‌های مختلف شبکه مانند ایستگاه‌های پایه و تجهیزات کاربر اجازه می‌دهد به‌عنوان عامل‌های مستقل همکاری کنند تا عملکرد کلی شبکه به حداکثر برسد. نتیجه این کار استفاده کارآمدتر از منابع، کاهش اتلاف و تضمین

5G Network Slicing



شکل ۲. برش شبکه در 5G

گلوگاه‌های احتمالی را پیش‌بینی کنند و به شبکه اجازه دهند اقدامات پیشگیرانه انجام دهد تا حتی در دوره‌های اوج مصرف، عملیات روان و بدون اختلال ادامه یابد.

یکی از مزایای کلیدی ادغام هوش مصنوعی در برش شبکه، قابلیت خودبهینه‌سازی (Self-Optimization) است.

مدیریت ترافیک مبتنی بر هوش مصنوعی همچنین می‌تواند ازدحام را کاهش دهد و کیفیت کلی سرویس (QoS) را با مسیریابی مجدد ترافیک از مسیرهای کمتر شلوغ یا تنظیم تخصیص پهنای باند بهبود بخشد. مدل‌های پیش‌بینانه که بر اساس داده‌های تاریخی ترافیک آموزش دیده‌اند، می‌توانند

هوش مصنوعی می‌تواند به‌طور مداوم شاخص‌های عملکرد شبکه مانند تأخیر، توان عملیاتی و نرخ خطا را در اسلایس‌های مختلف پایش کند. هنگامی که انحرافی از عملکرد مورد انتظار شناسایی شود، سیستم‌های هوش مصنوعی می‌توانند به‌طور خودکار پیکربندی‌ها را تنظیم کنند، منابع را دوباره تخصیص دهند یا حتی معماری اسلایس را تغییر دهند تا عملکرد بهینه‌تر بازیابی شود.

برای مثال، در شرایطی که یک اسلایس مخصوص برنامه‌های اینترنت اشیا (IoT) با افزایش ناگهانی تعداد دستگاه‌های متصل مواجه شود، هوش مصنوعی می‌تواند ظرفیت یک اسلایس را با تخصیص مجدد منابع از اسلایس‌های کم‌اهمیت‌تر افزایش دهد. به همین ترتیب، اسلایس‌هایی که نیازمند تأخیر فوق‌العاده پایین هستند می‌توانند به‌طور پویا بازپیکربندی شوند تا مسیرهای با تأخیر کمتر در اولویت قرار گیرند.

رویکردهای مبتنی بر هوش مصنوعی برای غلبه بر پیچیدگی برش شبکه در 5G بنیادی هستند. با بهره‌گیری از فناوری‌هایی مانند یادگیری تقویتی، شبکه‌های عصبی و سیستم‌های چندعاملی، شبکه‌های 5G می‌توانند کارایی، سازگاری و مقیاس‌پذیری بیشتری به دست آورند. هوش مصنوعی تضمین می‌کند که اسلایس‌های شبکه به‌طور پویا ایجاد، نگهداری و بهینه‌شوند و خدمات متناسب با نیازهای متنوع اکوسیستم دیجیتال مدرن ارائه دهند.

5. چالش‌ها و مسیرهای آینده

ادغام هوش مصنوعی در شبکه‌های 5G برای تخصیص منابع، مدیریت ترافیک و برش شبکه پتانسیل بالایی دارد اما با چالش‌های متعددی نیز همراه است. یکی از چالش‌های اصلی، پیچیدگی مدیریت شبکه‌های متراکم و ناهمگون است. از آنجا که 5G از کاربردهای مختلفی مانند eMBB، URLLC و IoT انبوه پشتیبانی می‌کند، نیاز به بهینه‌سازی بلادرنگ منابع حیاتی می‌شود. سیستم‌های سنتی مبتنی بر قوانین قادر به مدیریت کارآمد ترافیک پویا و تقاضای کاربران نیستند و راهکارهای تطبیقی مبتنی بر AI ضرورت پیدا می‌کنند.

با این حال، اجرای مدل‌های هوش مصنوعی برای تصمیم‌گیری بلادرنگ در مقیاس وسیع نیازمند توان محاسباتی بالا و الگوریتم‌های یادگیری کارآمد است تا از تأخیر و گلوگاه‌های سیستمی جلوگیری شود. یکی از مشکلات کلیدی، سربار و تأخیر در تخصیص منابع مبتنی بر AI است، به‌ویژه هنگام استفاده از مدل‌های یادگیری تقویتی عمیق (DRL). این مدل‌ها اگرچه قادرند از محیط شبکه یاد بگیرند و تنظیمات پویا انجام دهند، اما هزینه‌های آموزشی و مصرف حافظه بالایی دارند که می‌تواند در عملیات بلادرنگ ناکارآمدی ایجاد کند، به‌خصوص در شبکه‌های بزرگ مانند شهرهای هوشمند یا خودروهای خودران.

روش‌های MARL (یادگیری تقویتی چندعاملی) نیز به هماهنگی گسترده بین موجودیت‌های شبکه نیاز دارند که اگر به‌درستی مدیریت نشود، منجر به سربار و ائتلاف منابع خواهد شد. چالش دیگر، وابستگی به اطلاعات وضعیت کانال (CSI) برای تخصیص منابع است که سربار زیادی ایجاد می‌کند و در کاربردهای CRAN و mmWave ناکارآمد است.

راهکارهای موجود مانند الگوریتم‌های ابتکاری، الگوریتم‌های ژنتیک یا روش‌های خوشه‌بندی تنها بهبودهای جزئی ارائه می‌دهند و اغلب با افزایش تقاضای کاربران مقیاس‌پذیر نیستند. مسیرهای آینده شامل بهبود کارایی و مقیاس‌پذیری راهکارهای مبتنی بر AI در 5G است. تحقیقات باید بر بهینه‌سازی الگوریتم‌های یادگیری برای کاهش هزینه‌های آموزشی و مصرف حافظه تمرکز کنند؛ احتمالاً از طریق یادگیری فدرال (Federated Learning) یا رایانش لبه‌ای (Edge Computing) که پردازش را نزدیک‌تر به لبه شبکه توزیع می‌کند. همچنین مدل‌های ترکیبی هوش مصنوعی که چندین تکنیک ML را ترکیب می‌کنند — مانند CNN برای پیش‌بینی ترافیک و RL برای تخصیص منابع — می‌توانند راهکارهای سازگارتر برای محیط‌های ناهمگون 5G ارائه دهند.

برش شبکه در 5G نیز به مکانیزم‌های پیشرفته‌تر ارکستراسیون مبتنی بر AI نیاز دارد. پیش‌بینی و سازگاری بلادرنگ اسلایس‌ها بر اساس الگوریتم‌های AI حیاتی خواهد بود، به‌ویژه در مدیریت نیازهای متفاوت سرویس‌ها

از نظر تأخیر، قابلیت اعتماد و پهنای باند. ادغام مدل‌های AI با شبکه‌های نرم‌افزاری (SDN) و مجازی‌سازی عملکرد شبکه (NFV) می‌تواند مدیریت اسلایس‌ها را به‌طور پویا بهینه کند.

6. نتیجه‌گیری

ادغام تکنیک‌های مبتنی بر هوش مصنوعی در شبکه‌های G5 رویکردی تحول‌آفرین برای غلبه بر چالش‌های ذاتی تخصیص منابع، مدیریت ترافیک و برش شبکه است. با افزایش پیچیدگی شبکه‌های G5، روش‌های سنتی قادر به ارائه سازگاری بلادرنگ مورد نیاز برای محیط‌های پویا و با عملکرد بالا نیستند. مدل‌های AI، به‌ویژه مبتنی بر ML و DRL، راهکارهای تطبیقی و داده‌محور ارائه می‌دهند که می‌توانند به‌طور مداوم از شرایط شبکه یاد بگیرند و عملکرد را بهینه کنند، تأخیر را کاهش دهند و سربار سیستم را مدیریت نمایند.

تخصیص منابع در G5 به‌ویژه با رشد برنامه‌های داده‌محور مانند خودروهای خودران، واقعیت افزوده و استقرارهای گسترده IoT حیاتی است. روش‌های مبتنی بر AI مانند DRL و الگوریتم‌های ژنتیک رویکردهای مقیاس‌پذیر برای مدیریت کارآمد طیف، توان محاسباتی و انرژی ارائه می‌دهند. این روش‌های هوشمند کاستی‌های مدل‌های سنتی مانند تخصیص مبتنی بر CSI را برطرف کرده و سربار کمتر و سازگاری بهتر با شرایط متغیر فراهم می‌کنند.

مدیریت ترافیک نیز یکی دیگر از حوزه‌هایی است که AI به‌طور چشمگیری عملکرد شبکه‌های G5 را ارتقا می‌دهد. از طریق پیش‌بینی پیشرفته ترافیک و تحلیل بلادرنگ، مدل‌های AI مانند LSTM و معماری‌های مبتنی بر Transformer ابزارهای قدرتمندی برای پیش‌بینی الگوهای ترافیک و بهینه‌سازی توزیع بار شبکه ارائه می‌دهند. این قابلیت‌ها برای مدیریت افزایش نمایی ترافیک موبایل حیاتی هستند و بهره‌وری پهنای باند و مقاومت شبکه را حتی در شرایط تقاضای بالا تضمین می‌کنند.

علاوه بر این، برش شبکه — به‌عنوان ستون فقرات معماری G5 — به‌شدت از توانایی AI در ارکستراسیون و بهینه‌سازی اسلایس‌های مجازی در زمان واقعی بهره‌مند می‌شود.

تکنیک‌هایی مانند MARL کنترل دقیق‌تر تخصیص منابع بین اسلایس‌ها را ممکن می‌سازند و تضمین می‌کنند که هر اسلایس توافق‌نامه سطح سرویس (SLA) خود را برآورده کند و در عین حال کارایی کلی شبکه بهینه شود.

ادغام هوش مصنوعی در G5 نه یک فناوری مکمل، بلکه یک ضرورت است تا پتانسیل کامل شبکه‌های نسل بعدی محقق شود. گذار از سیستم‌های سنتی به الگوریتم‌های هوشمند و تطبیقی یک تغییر پارادایم است که آینده‌ی مخابرات را تعریف خواهد کرد. این تحول امکان عملیات شبکه‌ای مقاوم‌تر، کارآمدتر و مقیاس‌پذیرتر را فراهم می‌کند که طیف وسیعی از فناوری‌های نوظهور را پشتیبانی می‌کنند. همگرایی هوش مصنوعی و G5 پایه‌ای برای شبکه‌های خودکار ایجاد کرده و مسیرهای پژوهشی جدیدی را برای ارتقای عملکرد، کارایی و مقیاس‌پذیری در عصر G6 و فراتر از آن می‌گشاید.

اختصارات (Abbreviations)

- ML: یادگیری ماشین (Machine Learning)
- DL: یادگیری عمیق (Deep Learning)
- DRL: یادگیری تقویتی عمیق (Deep Reinforcement Learning)
- NR: رادیوی جدید (New Radio)
- URLLC: ارتباطات فوق‌قابل اعتماد با تأخیر کم (Ultra-Reliable Low-Latency Communications)
- MARL: یادگیری تقویتی چندعاملی (Multi-agent Reinforcement Learning)
- HVFT: ترافیک حجیم و انعطاف‌پذیر (High-volume Flexible Traffic)
- DDPG: الگوریتم گرادینت سیاست قطعی عمیق (Deep Deterministic Policy Gradient)
- mMTC: ارتباطات ماشینی انبوه (Massive Machine-Type Communications)

[5] Abadi, A., T. Rajabioun, and P. A. Ioannou, Traffic flow prediction for road transportation networks with limited traffic data. IEEE transactions on intelligent transportation systems, 2014. 16(2): p. 653-662.

[6] Imtiaz, S., et al. Random forests resource allocation for 5G systems: Performance and robustness study. in 2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW). 2018. IEEE.

[7] Wang, T., S. Wang, and Z.-H. Zhou, Machine learning for 5G and beyond: From model-based to data-driven mobile wireless networks. China Communications, 2019. 16(1): p. 165-175.

[8] Baghani, M., S. Parsaeefard, and T. Le-Ngoc, Multi-objective resource allocation in density-aware design of C-RAN in 5G. IEEE Access, 2018. 6: p. 45177-45190.

[9] Shehzad, M. K., et al., ML-based massive MIMO channel prediction: Does it work on real-world data? IEEE Wireless Communications Letters, 2022. 11(4): p. 811-815.

[10] Chughtai, N. A., et al., Energy efficient resource allocation for energy harvesting aided H-CRAN. IEEE Access, 2018. 6: p. 43990-44001.

[11] Zarin, N. and A. Agarwal, Hybrid radio resource management for time-varying 5G heterogeneous wireless access network. IEEE Transactions on Cognitive Communications and Networking, 2021. 7(2): p. 594-608.

[12] Huang, H., et al., Optical true time delay pool based hybrid beamformer enabling centralized beamforming control in millimeter-

• MIMO: سیستم‌های ورودی-خروجی چندگانه عظیم (Massive Multiple-input Multiple-output) (Systems

• CSI: اطلاعات وضعیت کانال (Channel State) (Information

مشارکت نویسندگان (Author Contributions)

• Dileesh Chandra Bikkasani نویسنده اصلی است.

• Malleswar Reddy Yerabolu نویسنده همکار است.

• نویسندگان نسخه نهایی مقاله را خوانده و تأیید کرده‌اند.

تضاد منافع (Conflicts of Interest)

نویسنده اعلام می‌کند که هیچ تضاد منافی وجود ندارد.

مراجع

[1] An, J., et al., Achieving sustainable ultra-dense heterogeneous networks for 5G. IEEE Communications Magazine, 2017. 55(12): p. 84-90.

[2] ITU. Setting the Scene for 5G: Opportunities & Challenges. 2020 [cited 2024 07/13]; Available from: <https://www.itu.int/hub/2020/03/setting-the-scene-for-5g-opportunities-challenges/>

[3] Sakaguchi, K., et al., Where, when, and how mmWave is used in 5G and beyond. IEICE Transactions on Electronics, 2017. 100(10): p. 790-808.

[4] Foukas, X., et al., Network slicing in 5G: Survey and challenges. IEEE communications magazine, 2017. 55(5): p. 94-100.

- [20] Box, G. E., et al., Time series analysis: forecasting and control. 2015: John Wiley & Sons.
- [21] Shu, Y., et al., Wireless traffic modeling and prediction using seasonal ARIMA models. *IEICE transactions on communications*, 2005. 88(10): p. 3992-3999.
- [22] Kumari, A., J. Chandra, and A. S. Sairam. Predictive flow modeling in software defined network. in *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)*. 2019. IEEE
- [23] Moore, J. S., A fast majority vote algorithm. *Automated Reasoning: Essays in Honor of Woody Bledsoe*, 1981: p. 105-108.
- [24] Arjoune, Y. and S. Faruque. Artificial intelligence for 5g wireless systems: Opportunities, challenges, and future research direction. in *2020 10th annual computing and communication workshop and conference (CCWC)*. 2020. IEEE.
- [25] Mennes, R., et al. A neural-network-based MF-TDMA MAC scheduler for collaborative wireless networks. in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. 2018. IEEE.
- [26] Vaswani, A., et al., Attention is all you need. *Advances in neural information processing systems*, 2017. 30.
- [27] Chinchali, S., et al. Cellular network traffic scheduling with deep reinforcement learning. in *Proceedings of the AAAI Conference on Artificial Intelligence*. 2018.
- [28] Peng, B., et al., Decentralized scheduling for cooperative localization with deep reinforcement wave C-RAN systems. *Science China Information Sciences*, 2021. 64(9): p. 192304.
- [13] Lin, X. and S. Wang. Efficient remote radio head switching scheme in cloud radio access network: A load balancing perspective. in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*. 2017. IEEE.
- [14] Gowri, S. and S. Vimalanand, QoS-Aware Resource Allocation Scheme for Improved Transmission in 5G Networks with IOT. *SN Computer Science*, 2024. 5(2): p. 234.
- [15] Bouras, C. J., E. Michos, and I. Prokopiou. *Applying Machine Learning and Dynamic Resource Allocation Techniques in Fifth Generation Networks*. 2022. Cham: Springer International Publishing.
- [16] Li, R., et al., Intelligent 5G: When cellular networks meet artificial intelligence. *IEEE Wireless communications*, 2017. 24(5): p. 175-183.
- [17] Ericsson. 5G to account for around 75 percent of mobile data traffic in 2029. [cited 2024 07/13]; Available from: <https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/mobile-traffic-forecast>
- [18] Amaral, P., et al. Machine learning in software defined networks: Data collection and traffic classification. in *2016 IEEE 24th International conference on network protocols (ICNP)*. 2016. IEEE.
- [19] Wang, H., et al. Understanding mobile traffic patterns of large scale cellular towers in urban environment. in *Proceedings of the 2015 Internet Measurement Conference*. 2015.

learning. IEEE Transactions on Vehicular Technology, 2019. 68(5): p. 4295-4305.

[29] Cao, G., et al., AIF: An artificial intelligence framework for smart wireless network management. IEEE Communications Letters, 2017. 22(2): p. 400-403.

[30] Challita, U., L. Dong, and W. Saad, Proactive resource management for LTE in unlicensed spectrum: A deep learning perspective. IEEE transactions on wireless communications, 2018. 17(7): p. 4674-4689.

[31] Stampa, G., et al., A deep-reinforcement learning approach for software-defined networking routing optimization. arXiv preprint arXiv:1709.07080, 2017.