

دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده مهندسی برق و کامپیوتر
گروه شبکه



تخصیص منابع و قسمت کردن شبکه در ساختار شبکه‌های دسترسی باز

رساله برای دریافت درجهٔ دکتری در رشتهٔ مهندسی برق
گرایش مخابرات شبکه

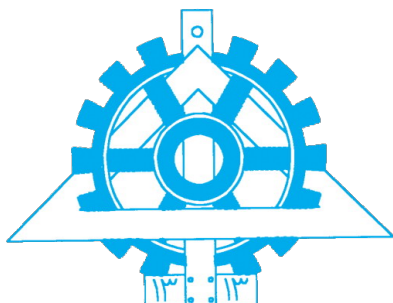
مژده کربلایی مطلب

استاد راهنما

دکتر شاه منصوری

شهریور ۱۳۹۹





دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده مهندسی برق و کامپیوتر
گروه شبکه



تخصیص منابع و قسمت کردن شبکه در ساختار شبکه‌های دسترسی باز

رساله برای دریافت درجهٔ دکتری در رشتهٔ مهندسی برق
گرایش مخابرات شبکه

مژده کربلایی مطلب

استاد راهنما

دکتر شاه منصوری

شهریور ۱۳۹۹



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده مهندسی برق و کامپیوتر



گواهی دفاع از رساله دکتری

هیأت داوران رساله دکتری آقای / خانم مژده کربلایی مطلب به شماره دانشجویی ۸۱۰۱۹۶۰۷۴ در رشته مهندسی برق - گرایش مخابرات شبکه را در تاریخ با عنوان «تخصیص منابع و قسمت کردن شبکه در ساختار شبکه‌های دسترسی باز»

به عدد	به حروف
<input type="text"/>	<input type="text"/>

با نمره نهایی

و درجه

ارزیابی کرد.

ردیف	مشخصات هیأت داوران	نام و نام خانوادگی	مرتبه دانشگاهی	دانشگاه یا مؤسسه	امضا
۱	استاد راهنما	دکتر شاه منصوری	دانشیار	دانشگاه تهران	
۲	استاد داور داخلی	دکتر داور داخلی	دانشیار	دانشگاه تهران	
۳	استاد مدعو	دکتر داور خارجی	دانشیار	دانشگاه داور خارجی	
۴	نماینده تحصیلات تکمیلی دانشکده	دکتر نماینده	دانشیار	دانشگاه تهران	

نام و نام خانوادگی معاون آموزشی و تحصیلات

تکمیلی پردیس دانشکده‌های فنی:

تاریخ و امضا:

نام و نام خانوادگی معاون تحصیلات تکمیلی و

پژوهشی دانشکده / گروه:

تاریخ و امضا:

تعهدنامه اصالت اثر

باسمه تعالی

اینجانب مؤده کربلایی مطلب تأیید می‌کنم که مطالب مندرج در این رساله حاصل کار پژوهشی اینجانب است و به دستاوردهای پژوهشی دیگران که در این نوشته از آنها استفاده شده است مطابق مقررات ارجاع گردیده است. این رساله قبلاً برای احراز هیچ مدرک هم‌سطح یا بالاتری ارائه نشده است.

نام و نام خانوادگی دانشجو: مؤده کربلایی مطلب
تاریخ و امضای دانشجو:

کلیه حقوق مادی و معنوی این اثر
متعلق به دانشگاه تهران است.

تقديم به:

پدر و مادرم

قدردانی

سپاس خداوندگار حکیم را که با لطف بی کران خود، آدمی را به زیور عقل آراست.
در آغاز وظیفه خود می دانم از زحمات بی دریغ اساتید راهنمای خود، جناب آقای دکتر ... و ...، صمیمانه
تشکر و قدردانی کنم که در طول انجام این پایان نامه با نهایت صبوری همواره راهنما و مشوق من بودند و قطعاً
بدون راهنمایی های ارزنده ایشان، این مجموعه به انجام نمی رسید.
از جناب آقای دکتر ... که زحمت مشاوره، بازبینی و تصحیح این پایان نامه را تقبل فرمودند کمال امتنان را
دارم.

با سپاس بی دریغ خدمت دوستان گرانمایه ام، خانم ها ... و آقایان ... در آزمایشگاه ...، که با همفکری مرا
صمیمانه و مشفقانه یاری داده اند.

و در پایان، بوسه می زنم بر دستان خداوندگاران مهر و مهربانی، پدر و مادر عزیزم و بعد از خدا، ستایش می کنم
وجود مقدس شان را و تشکر می کنم از خانواده عزیزم به پاس عاطفه سرشار و گرمای امیدبخش وجودشان، که
بهترین پشتیبان من بودند.

مژده کربلایی مطلب

شهریور ۱۳۹۹

چکیده

این راهنما، نمونه‌ای از قالب پروژه، پایان‌نامه و رساله دانشگاه تهران می‌باشد که با استفاده از کلاس -tehran thesis و بسته‌ی پرشین در \LaTeX تهیه شده است. این قالب به گونه‌ای طراحی شده است که مطابق با دستورالعمل نگارش و تدوین پایان‌نامه کارشناسی ارشد و دکتری، مورخ ۹۳/۰۶/۰۳ پردیس دانشکده‌های فنی دانشگاه تهران باشد و حروف چینی بسیاری از قسمت‌های آن، مطابق با استاندارد قالب‌های فارسی پایان‌نامه در لاتک، به طور خودکار انجام می‌شود.

چکیده بخشی از پایان‌نامه است که خواننده را به مطالعه آن علاقمند می‌کند و یا از آن می‌گریزند. چکیده باید ترجیحاً در یک صفحه باشد. در نگارش چکیده نکات زیر باید رعایت شود. متن چکیده باید مزین به کلمه‌ها و عبارات سلیس، آشنا، بامعنی و روشن باشد. بگونه‌ای که با حدود ۳۰۰ تا ۵۰۰ کلمه بتواند خواننده را به خواندن پایان‌نامه راغب نماید. چکیده، جدای از پایان‌نامه باید به تنهایی گویا و مستقل باشد. در چکیده باید از ذکر منابع، اشاره به جداول و نمودارها اجتناب شود. تمیز بودن مطلب، نداشتن غلط‌های املایی یا دستور زبانی و رعایت دقت و تسلسل روند نگارش چکیده از نکات مهم دیگری است که باید در نظر گرفته شود. در چکیده پایان‌نامه باید از درج مشخصات مربوط به پایان‌نامه خودداری شود. چکیده باید منعکس‌کننده اصل موضوع باشد. در چکیده باید اهداف تحقیق مورد توجه قرار گیرد. تأکید روی اطلاعات تازه (یافته‌ها) و اصطلاحات جدید یا نظریه‌ها، فرضیه‌ها، نتایج و پیشنهادها متمرکز شود. اگر در پایان‌نامه روش نوینی برای اولین بار ارائه می‌شود و تا به حال معمول نبوده است، با جزئیات بیشتری ذکر شود. شایان ذکر است چکیده فارسی و انگلیسی باید حتماً به تأیید استاد راهنما رسیده باشد.

کلمات کلیدی در انتهای چکیده فارسی و انگلیسی آورده می‌شود. محتوای چکیده‌ها بر اساس موضوع و گرایش تحقیق طبقه‌بندی می‌شود و به همین جهت وجود کلمات شاخص و کلیدی، مراکز اطلاعاتی را در طبقه‌بندی دقیق و سریع پایان‌نامه یاری می‌دهد. کلمات کلیدی، راهنمای نکات مهم موجود در پایان‌نامه هستند. بنابراین باید در حد امکان کلمه‌ها یا عباراتی انتخاب شود که ماهیت، محتوا و گرایش کار را به وضوح روشن نماید.

واژگان کلیدی حداکثر ۵ کلمه یا عبارت، متناسب با عنوان، قالب پایان‌نامه، لاتک

فهرست مطالب

ث	فهرست تصاویر
چ	فهرست جداول
ح	فهرست الگوریتم‌ها
ح	فهرست برنامه‌ها
خ	فهرست اختصارات
۱	فصل ۱: مقدمه
۱	۱.۰.۱ مقدمه ای بر 5G
۳	۲.۰.۱ مروری بر نسل‌های اخیر مخابرات
۴	۱.۲.۰.۱ نسل چهارم مخابرات
۴	۲.۲.۰.۱ نسل پنجم مخابرات
۵	۱.۱ مقدمه ای بر ساختار ORAN
۶	۱.۱.۱ مقدمه ای بر ساختار شبکه های دسترسی رادیویی C-RAN
۷	۱.۱.۱.۱ ساختار شبکه های مختلف
۸	۲.۱.۱.۱ ساختار سنتی ایستگاه پایه
۸	۳.۱.۱.۱ ساختار ایستگاه پایه و واحد رادیویی
۹	۴.۱.۱.۱ ساختار C-RAN
۱۴	۲.۱.۱ xRAN

۱۵	۱.۲.۱.۱	مزایای ساختار xRAN
۱۵	۳.۱.۱	ORAN
۱۹	۲.۱	مجازی سازی توابع شبکه
۲۰	۱.۲.۱	ساختار شبکه
۲۲	۳.۱	شبکه تعریف شده نرم افزار (SDN)
۲۳	۴.۱	برش شبکه
۲۵	۵.۱	نتیجه گیری
۲۶	۶.۱	خلاصه ای از فصل های آتی
۲۷		فصل ۲: مروری بر کارهای پیشین
۲۷	۱.۲	مقدمه
۲۷	۲.۲	مسائل پیش رو
۲۷	۱.۲.۲	برش شبکه
۳۱	۲.۲.۲	رفتن به سمت شبکه های دسترسی رادیویی باز
۳۲	۳.۲.۲	قرار دادن VNF ها
۳۳	۳.۲	روش های حل
۳۳	۱.۳.۲	مسائل Np-Hard
۳۳	۱.۱.۳.۲	مسئله ی کوله پشتی
۳۴	۲.۱.۳.۲	مسئله ی بسته بندی جعبه
۳۵	۲.۳.۲	روشهای یادگیری تقویتی
۳۸		فصل ۳: تخصیص منابع در شبکه های Open RAN
۳۸	۱.۳	مقدمه
۴۰	۲.۳	مدل سیستم
۴۰	۱.۲.۳	نرخ قابل دسترس
۴۳	۲.۲.۳	میانگین تاخیر
۴۴	۳.۲.۳	مرکز داده ی فیزیکی

۴۵	۴.۲.۳ صورت مساله
۴۷	۳.۳ روش ابتکاری استفاده شده
۴۷	۱.۳.۳ بخش اول مسئله اول
۴۹	۲.۳.۳ بخش دوم مسئله اول
۵۰	۳.۳.۳ حل دو بخش مسئله اول به صورت تکراری
۵۰	۴.۳.۳ مسئله دوم
۵۲	۴.۳ نتایج عددی
۵۷	۵.۳ نتیجه گیری

۶۰	فصل ۴: تخصیص برش شبکه به صورت دینامیکی
۶۰	۱.۴ مقدمه
۶۰	۲.۴ مدل سیستم و صورت مسئله اول
۶۲	۳.۴ مدل سیستم و صورت مسئله دوم
۶۳	۴.۴ حل مسئله به روش یادگیری تقویتی
۶۵	۱.۴.۴ نتایج عددی مسئله اول
۶۵	۲.۴.۴ نتایج عددی مسئله دوم
۶۶	۵.۴ نتیجه گیری

۶۷	فصل ۵: پیشنهادات و کارهای آتی
۶۷	۱.۵ مقدمه
۶۷	۲.۵ نتیجه گیری
۶۸	۱.۲.۵ مزایای این چالش و حل آن
۶۸	۲.۲.۵ معایب پروژه انجام شده
۶۸	۳.۲.۵ نوآوری های این پروژه

۶۹	کتاب نامه
----	-----------

فهرست تصاویر

۱.۱	مقایسه قابلیت های کلیدی IMT-Advanced (نسل ۴) با IMT-2020 (نسل ۵) با توجه به ITU-R M.2083 [۱]	۳
۲.۱	ساختار سنتی ایستگاه پایه [۲]	۷
۳.۱	ساختار ایستگاه پایه و واحد رادیویی [۲]	۸
۴.۱	ساختار C-RAN [۲]	۹
۵.۱	ساختار شبکه ی C-RAN [۲]	۱۰
۶.۱	ساختار شبکه های دسترسی ابری نامتجانس [۳]	۱۱
۷.۱	مدل سیستم F-RAN [۴]	۱۳
۸.۱	ساختار شبکه ی ORAN [۵]	۱۶
۹.۱	ساختار NFV [۶]	۲۰
۱۰.۱	ساختار SDN [۷]	۲۳
۱۱.۱	سه ساختار برش شبکه [۸]	۲۴
۱.۲	روند برش شبکه [۹]	۲۹
۲.۲	مدل پردازشی شبکه صف [۱۰]	۳۱
۳.۲	یادگیری تقویتی	۳۵
۴.۲	سناریوی ارسال لینک بالا و پایین در برشهای شبکه	۳۶
۱.۳	برش شبکه در ORAN	۳۹
۲.۳	بهره‌وری انرژی براساس تعداد کاربران در هر سرویس	۵۴
۳.۳	بهره‌وری انرژی براساس تعداد بلوکهای منابع فیزیکی	۵۵

۴.۳	بهره‌وری انرژی براساس تعداد سرویسهای مختلف	۵۶
۵.۳	نسبت برش های پذیرفته شده فقط به یک DC در مقابل برش های کل	۵۸
۶.۳	نرمالیزه‌ی مصرف منابع براساس تعداد برشها	۵۸

فهرست جداول

۵۲	۱.۳ پارامترهای شبیه‌سازی
۵۷	۲.۳ پارامترهای شبیه‌سازی

فهرست الگوریتم‌ها

۴۸	۱.۳	اتصال سرویس به برش شبکه
۵۱	۲.۳	برش شبکه و تخصیص منابع
۵۳	۳.۳	قرار گیری منابع مجازی در منابع فیزیکی

فهرست برنامه‌ها

فصل ۱

مقدمه

۱.۰.۱ مقدمه ای بر 5G

5G، مخابرات نسل پنجم سیستم های بیسیم^۱ و شبکه های مخابراتی بعد از نسل چهارم می باشد که تکاملی از لایه ی فیزیکی در تکنولوژی شبکه های مخابراتی سیار همانند LTE است که نسبت به 4G سرعت و پوشش بهتری را فراهم می کند. 5G نوع جدیدی از شبکه را ایجاد می کند که به منظور اتصال تقریباً همه و همه چیز با هم از جمله ماشین ها، اشیاء و دستگاه ها ساخته شده است. 5G فناوری بی سیم برای ارائه سرعت داده های چند گیگابیت بر ثانیه، تأخیر فوق العاده کم، قابلیت اطمینان بیشتر، ظرفیت شبکه گسترده، افزایش در دسترس بودن و تجربه کاربری یکنواخت تر به کاربران بیشتر است. عملکرد بالاتر و بهره وری بهبود یافته باعث افزایش تجربیات کاربر جدید شده و صنایع جدیدی را به هم متصل می کند.

تکنولوژی سیگنال 5G برای پوشش فراگیرتر و بازدهی بهتر سیگنال ایجاد شده است. این پیشرفت ها منجر به تغییراتی از قبیل IOT^۲ و Pervasive Computing در آینده ی نزدیک خواهد شد. همچنین 5G منجر به توسعه و بهبود سرویس های مخابراتی و اینترنتی سیار و در ورای آن، ایجاد تجربه ی بهتری برای مصرف کنندگان خواهد شد.

برای توسعه ی اینترنت سیار و IOT، نیازمند استفاده از شبکه ی نسل پنجم هستیم تا به سادگی منجر به دسترسی شبکه برای ارتباط انسان ها با یکدیگر و ارتباط ماشین با انسان گردد.

به طور کلی، 5G در سه نوع سرویس اصلی متصل از جمله پهن باند تلفن همراه، IoT عظیم و ارتباطات مهم برای مأموریت استفاده می شود.

¹Wireless

²Internet of Things

۱. پهن باند تلفن همراه پیشرفته (eMBB) برای مقابله با نرخ داده‌های بسیار زیاد، تراکم بالای کاربران و ظرفیت ترافیک بسیار بالا برای سناریوهای مختلف و همچنین پوشش یکپارچه و سناریوهای تحرک بالا با نرخ داده‌های استفاده شده بهبود یافته است.

۲. ارتباطات عظیم ماشین (mMTC) برای IoT، برای تعداد بسیار زیاد دستگاه‌های متصل به مصرف کم و نرخ داده کم نیازمند می‌باشد.

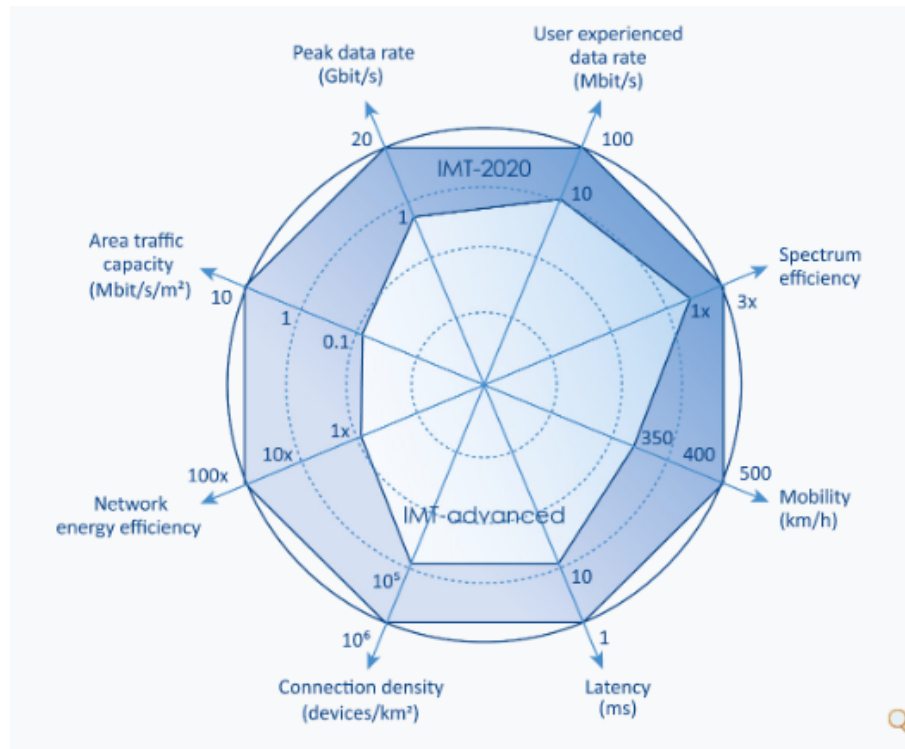
۳. ارتباطات بسیار مطمئن و با تأخیر کم (URLLC) برای برنامه‌های کاربردی مهم برای ایمنی و مأموریت مورد توجه است.

از آنجا که ساختار 5G کمتر به زیرساخت‌های 4G وابسته می‌شود و طیف بیشتری در دسترس قرار می‌دهد، تخمین‌ها سرعت بارگیری را حداکثر ۱۰۰۰ برابر سریعتر از 4G در نظر دارد، که بالقوه از 10Gbps بیشتر است، که به شما امکان می‌دهد تا در کمتر از یک ثانیه فیلم کامل HD را بارگیری کنید. برخی تخمین‌ها محافظه کارانه تر هستند، اما حتی محافظه کارانه‌ترین تخمین نیز این نسل را چندین ده برابر سریعتر از 4G قرار می‌دهد. دلایل نیاز به نسل پنجم اینترنت به طور خلاصه در ادامه بیان شده است [۱].

- ترافیک داده‌های تلفن همراه به دلیل پخش ویدئو به سرعت، رو به افزایش است.
- با در اختیار داشتن چندین دستگاه به طور همزمان، هر کاربر تعداد فزاینده‌ای از اتصالات را در اختیار دارد.
- اینترنت اشیاء به شبکه‌هایی نیاز دارد که میلیاردها دستگاه را اداره کنند.
- با وجود تعداد فزاینده‌ای از دستگاه‌های ارتباطی و افزایش ترافیک داده‌ها، هم دستگاه‌ها و هم شبکه‌ی آن نیازمند افزایش بهره‌وری انرژی هستند.
- به دلیل تحت فشار قرار گرفتن اپراتورهای شبکه برای کاهش هزینه‌های عملیاتی و همچنین به دلیل اینکه کاربران به تعرفه‌های نرخ مسطح عادت می‌کنند و مایل نیستند مبلغ بیشتری بپردازند.
- فناوری ارتباطات سیار می‌تواند موارد استفاده جدیدی را ایجاد کند (به عنوان مثال موارد تاخیر فوق العاده کم یا قابلیت اطمینان بالا) و برنامه‌های جدید برای صنعت که منجر به درآمدزایی بیشتر اپراتورها می‌گردد.

بنابراین عملکرد عملیاتی نسل پنجم می‌بایست به طور قابل توجهی افزایش یابد (به عنوان مثال افزایش راندمان طیفی، سرعت بالاتر داده، تأخیر کم). زیرساخت 5G می‌بایست در حالی که هنوز سطح قابل قبولی از مصرف

انرژی، هزینه تجهیزات و استقرار شبکه و هزینه بهره‌برداری را ارائه می‌دهد، اینترنت اشیاء را به طور گسترده نیز تأمین کند. همچنین از طیف گسترده‌ای از برنامه‌ها و خدمات پشتیبانی کند.



شکل ۱.۱: مقایسه قابلیت‌های کلیدی IMT-Advanced (نسل ۴) با IMT-2020 (نسل ۵) با توجه به ITU-R M.2083 [۱]

۲.۰.۱ مروری بر نسل‌های اخیر مخابرات

در ابتدا می‌خواهیم بدانیم که چه چیزی منجر به رفتن محققان به سوی 5G شده است. یکی از دلایل مهم، سرعت و نرخ انتقال بیشتری است که در ادامه به آن می‌پردازیم. نیاز بشریت به ارتباط تلفنی (انتقال بدون سیم به صورت زمان حقیقی^۳ انسان را به سمت نسل اول ارتباطات 1G سوق داده است. نسل دوم ارتباطات 2G با سرویس‌های انتقال پیام کوتاه ایجاد شد. همچنین با موفقیت تکنولوژی شبکه‌های منطقه‌ای بیسیم، اتصال به داده‌های اینترنتی مورد توجه عموم مردم قرار گرفت که پلی به سوی نسل سوم ارتباطات 3G را فراهم نمود. به طور منطقی پله‌ی بعدی گام برداشتن در راستای کوچک شدن لپ‌تاپ و در آمیختن آن با تلفن که امروزه به صورت تلفن هوشمند^۴ است و دسترسی به اینترنت، پهنای باند بالا و داده‌ها در نقاط مختلف جهان بوده است که 4G یا

^۳Real Time

^۴smart phone

نسل چهارم را به همراه داشته است. با توجه به افزایش تعداد کاربران تلفن‌های هوشمند و تبلت‌ها و افزایش نرخ ارسال اطلاعات و داده‌ها در طی سالهای اخیر طبق پیش بینی‌های سیسکو میزان ترافیک IP طی سالهای اخیر چندین برابر افزایش خواهد یافت. در نتیجه اپراتورها برای حل این مشکل و خدمات‌دهی بهتر ناچار به افزایش ظرفیت شبکه می‌باشند. در ادامه به طور مختصر به نسلهای اخیر مخابراتی می‌پردازیم [۱۱].

۱.۲.۰.۱ نسل چهارم مخابرات

4G یک فناوری بسیار متفاوت در مقایسه با 3G است و هدف از آن، فراهم آوردن سرعت بالا، کیفیت بالا و ظرفیت بالا برای کاربران در عین بهبود امنیت و کاهش هزینه خدمات صوتی و دیتا، چندرسانه‌ای و اینترنت از طریق IP می‌باشد. برنامه‌های کاربردی بالقوه و جاری شامل دسترسی به وب موبایل اصلاح شده، تلفن تلفنی IP، خدمات بازی، تلویزیون همراه با کیفیت بالا، کنفرانس ویدیویی، تلویزیون سه بعدی و محاسبات ابری از قابلیت‌های پشتیبانی آن می‌باشد.

فن آوری‌های کلیدی که این امکان را ایجاد کرده اند MIMO^۵ و OFDM^۶ می‌باشد. دو استاندارد مهم آن LTE^۷ و WiMAX می‌باشد. حداکثر سرعت یک شبکه 4G هنگام حرکت دستگاه ۱۰۰ مگابیت بر ثانیه یا ۱ گیگابیت بر ثانیه برای ارتباطات کم تحرک مانند هنگام ایستادن یا راه رفتن است. تأخیر از حدود 300ms به 100ms با کاهش تراکم دست می‌یابد.

۲.۲.۰.۱ نسل پنجم مخابرات

تکنولوژی 5G یک استاندارد صنعتی است که جایگزین استاندارد رایج کنونی یعنی 4G LTE خواهد شد. این فناوری پنجمین نسل از استاندارد سلولی است. طراحی این استاندارد به گونه‌ای است که سرعت آن از تکنولوژی 4G LTE بسیار سریع تر است. البته هدف این استاندارد صرفاً افزایش سرعت اتصالات اینترنتی تلفن‌های هوشمند نیست. این استاندارد، اینترنت بی سیم بسیار پر سرعتی را در همه جا و برای همه چیزها از جمله خودروهای متصل، خانه‌های هوشمند و ابزارهای اینترنت اشیا (IoT) فراهم خواهد کرد. کاهش مصرف انرژی معیاری است که در این نسل به آن توجه شده است و دستگاه‌های فرستنده و گیرنده اپراتورها باید در ساعت کم مصرف به حالت صرفه‌جویی انرژی وارد شده و به سرعت فعال شوند که این معیار در نسل چهارم قید نشده

⁵Multiple Output Multiple Output

⁶Multiplexing Division Frequency Division

⁷Long Term Evolution

بوده‌است. با توجه به این که نرخ داده و ظرفیت در سیستم‌های نسل چهارم به ظرفیت شانون نزدیک شده است، در نتیجه روش‌هایی که برای افزایش ظرفیت شبکه مورد استفاده می‌گیرند که به شرح زیر است:

- استفاده از تکنیک Massive MIMO
- استفاده از روش‌های پردازش‌های ابری
- شبکه‌ی تعریف شده‌ی نرم افزاری SDN^۸
- موج میلیمتری^۹
- ساختار شبکه‌های دسترسی رادیویی باز ORAN^{۱۰}
- مجازی‌سازی توابع شبکه NFV^{۱۱}
- برش شبکه^{۱۲}

۱.۱ مقدمه‌ای بر ساختار ORAN

مجازی‌سازی سازی RAN توجه زیادی را از طرف اپراتورها به خود جلب می‌کند، زیرا منجر به کاهش هزینه‌های اپراتور و opex می‌شود و همچنین این امکان را برای آنها فراهم کند تا با سرعت بیشتری قابلیت‌های جدیدی به شبکه اضافه کنند.

این احتمال وجود دارد که همه این علاقه‌ها در ایجاد سه گروه مختلف باشد - انجمن xRAN، گروه OpenRAN شرکت Telecom Infra و ابتکار عمل Open VRAN که برای شرکت سیسکو می‌باشد. اگرچه همه این گروه‌ها می‌گویند که در حال کار بر روی یک چیز هستند، که اساساً برای باز کردن RAN با استفاده از رابط‌های استاندارد و عناصر شبکه جعبه سفید است، اما در بررسی دقیق‌تر اختلافاتی نیز وجود دارد. Open RAN (ORAN) تبسیط و ترکیبی از دو ساختار C-RAN^{۱۳} و xRAN می‌باشد که انتظار می‌رود که در فناوری نسل پنجم مخابرات مورد استفاده قرار گرفته و منجر به بهبود عملکرد شبکه‌های دسترسی رادیویی RAN گردد. این ساختار یک شبکه‌ی باز، انعطاف پذیر و هوشمند است.

⁸Software Defined Networking

⁹mm Wave

¹⁰Open Radio Access Network

¹¹Network Function Virtualization

¹²Network Slicing

¹³Cloud Radio Access Network

ORAN توابع شبکه‌ی دسترسی رادیویی را به سه قسمت تقسیم می‌کند، که قسمت اول واحد از راه دور (RU)^{۱۴}، واحد توزیع شده (DU)^{۱۵} و واحد مرکزی (CU)^{۱۶} می‌باشد. در حالی که RU دارای توابع (PHY)^{۱۷} پایین تر است، DU حاوی (PHY) بالاتر، MAC^{۱۸} و RLC^{۱۹} است و (CU) حاوی RRC^{۲۰}، PDCP^{۲۱} و SDAP^{۲۲} است.

DU و CU به عنوان توابع شبکه مجازی (VNFs) پیاده سازی می‌شوند، که در یک محیط ابر اجرا می‌شود. رابط های بین RU، CU و DU رابط های استاندارد باز هستند.

۱.۱.۱ مقدمه ای بر ساختار شبکه های دسترسی رادیویی C-RAN

شبکه های دسترسی ابری منجر به افزایش پوشش ارسالی می‌گردد. با توجه به ساختار شبکه C-RAN که معماری جدیدی را برای شبکه های نسل آینده ارائه می‌دهد، نه تنها ظرفیت شبکه افزایش می‌یابد بلکه مشکلاتی که در روش های دیگر وجود دارد را نیز هموار می‌سازد. مفهوم شبکه دسترسی رادیو ابر C-RAN، به مجازی سازی کارکردهای ایستگاه پایه^{۲۳} با استفاده از تکنولوژی رایانش ابری^{۲۴} اشاره می‌نماید. این مفهوم به ایجاد یک ساختار سلولی جدید منجر می‌شود که در آن، نقاط دسترسی بیسیم کم هزینه که با عنوان واحدهای رادیویی^{۲۵} و یا رادیو هد های راه دور^{۲۶} شناخته می‌شوند- با استفاده از یک ابر متمرکز با قابلیت پیکربندی مجدد و یا واحد مرکزی^{۲۷} مدیریت می‌شوند. شبکه امکان کاهش هزینه های سرمایه گذاری و عملیاتی مورد نیاز برای اپراتورها به منظور توسعه و نگهداری شبکه های ناهمگن متراکم را فراهم می‌آورد. این مزیت مهم در کنار بازده طیفی، تسهیم آماری^{۲۸}، و مزیت های متعادل سازی بار باعث می‌شود تا شبکه C-RAN به عنوان یکی از تکنولوژی های کلیدی در توسعه سیستم های 5G در جایگاه بسیار مناسبی قرار بگیرد. در ادامه، یک بررسی کلی و مختصر از تحقیقات جدید در مورد ساختار C-RAN ارائه می‌شود و موضوعات مورد تاکید عبارتند از فشرده سازی لینک fronthaul

¹⁴remote unit

¹⁵Distributed unit

¹⁶Central unit

¹⁷Physical layer

¹⁸Medium Access Control

¹⁹Radio Link Control

²⁰Radio Resource Control

²¹Packet Data Convergence Protocol

²²Service Data Adaptation Protocol

²³Base Station-BS

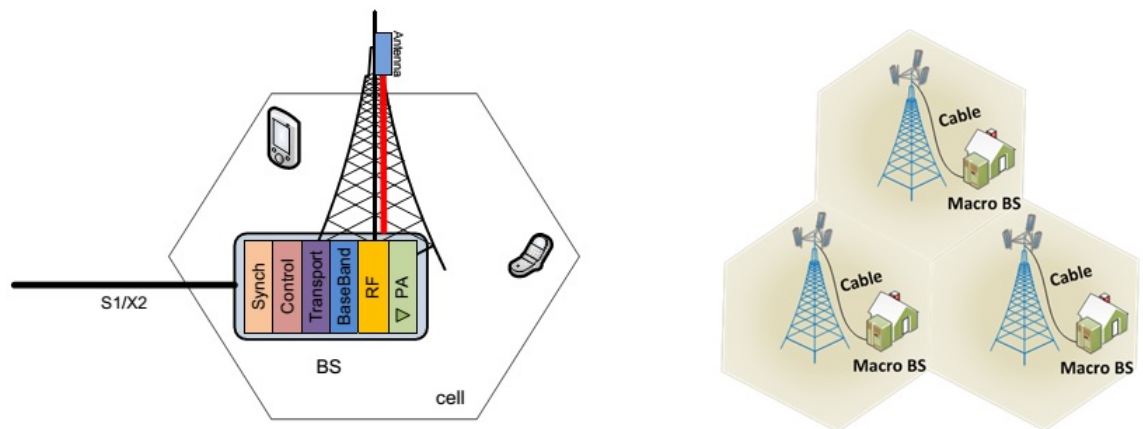
²⁴Cloud Computing

²⁵Radio Units

²⁶Radio Remote Heads

²⁷Control Unit

²⁸Statistical Multiplexing



شکل ۲.۱: ساختار سنتی ایستگاه پایه [۲]

پردازش باند پایه، کنترل دسترسی به محیط واسط، تخصیص منابع، ملاحظات سطح سیستم، و تلاش‌های انجام شده در راستای ارائه استاندارد‌ها.

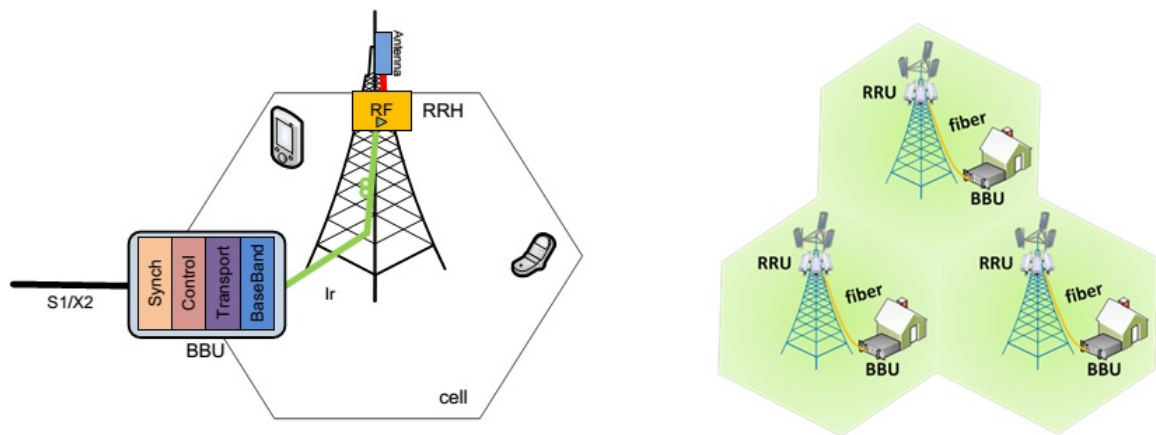
۱.۱.۱.۱ ساختار شبکه‌های مختلف

با توجه به مقاله ی [۲]، هر ایستگاه پایه دو نوع پردازش انجام می‌دهد: پردازش رادیویی که توسط واحد رادیویی^{۲۹} انجام می‌شود و شامل پردازش دیجیتالی، فیلترینگ فرکانسی، تقویت توان و می‌باشد و پردازش باند پایه که توسط واحد باند پایه^{۳۰} که همان واحد کنترل است^{۳۱} انجام شده و از جمله مهمترین وظایف آن می‌توان به کدینگ، مدولاسیون و تبدیل فوریه ی سریع اشاره کرد. در ساختار جدیدی که تحت عنوان C-RAN معرفی خواهیم نمود نحوه ی ارتباط پردازشگرهای رادیویی و باند پایه متحول شده و در نتیجه مزایایی برای شبکه حاصل خواهد شد. در ادامه، انواع ساختارها را بیان خواهد شد.

²⁹RRH

³⁰BBU

³¹CU



شکل ۳.۱: ساختار ایستگاه پایه و واحد رادیویی [۲]

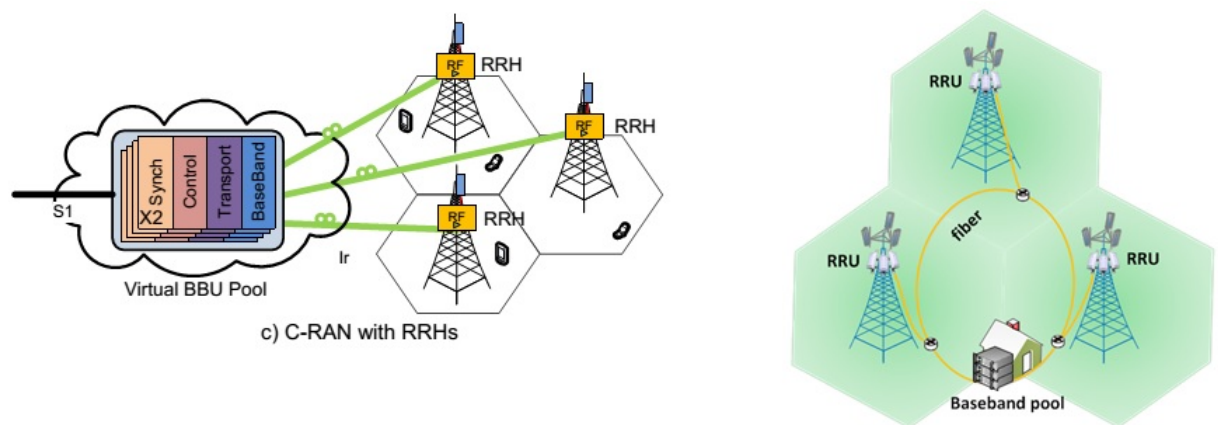
۲.۱.۱.۱ ساختار سنتی ایستگاه پایه

در ساختارهای سنتی ایستگاه پایه، پردازش‌های رادیویی و باند پایه در داخل ایستگاه پایه انجام می‌شد و مدول آنتن نیز در فاصله‌ی چند متری از مدول رادیویی نصب شده و ارتباط آنها توسط کابل کوکسیال برقرار می‌شد که همین امر سبب افزایش تلفات در شبکه می‌باشد. این نوع ساختار در شکل ۱.۳ نشان داده شده است. همان گونه که مشاهده می‌کنید ارتباط بین ایستگاه‌های پایه توسط ارتباط X_2 و ارتباط بین ایستگاه پایه و شبکه‌ی هسته توسط ارتباط S_1 برقرار می‌شود. این نوع ساختار در شبکه‌های 1G و 2G به کار گرفته شده است [۲].

۳.۱.۱.۱ ساختار ایستگاه پایه و واحد رادیویی

در این ساختار واحد رادیویی و واحد پردازشی سیگنال، از هم مجزا شده و واحد رادیویی که تحت عنوان RRH یا RRU نیز شناخته می‌شود، توسط فیبر نوری به واحد باند پایه یا BBU اتصال می‌یابد. همان طور که پیشتر بیان شد واحد رادیویی مسئولیت انجام پردازش‌های دیجیتال از جمله تبدیل آنالوگ به دیجیتال، دیجیتال به آنالوگ، تقویت توان و فیلترینگ را بر عهده دارد، که تفکیک وظایف واحد پردازشی و واحد رادیویی در این ساختار در شکل ۳.۱ قابل مشاهده است. این نوع ساختار برای شبکه‌های نسل سوم معرفی شده و امروزه نیز بیشتر ایستگاه‌های پایه از همین ساختار بهره می‌گیرند. از جمله ویژگی‌های بارز این ساختار امکان ایجاد فاصله بین واحد رادیویی

و پردازی می باشد، که این فاصله به دلیل تاخیر پردازشی و انتشاری نمی تواند از ۴۰ کیلومتر فراتر رود. در این ساختار تجهیزات مرتبط با BBU می توانند به مکانی مناسبتر که قابل دسترس تر بوده و هزینه ی اجاره و نگهداری کمتری را به اپراتورها تحمیل می کنند منتقل شوند و واحد های رادیویی نیز در پشت بام ساختمان ها و مکان های مرتفع نصب می شوند که این خود سبب کاهش هزینه های خنک سازی ادوات موجود می شود. نحوه ی ارتباط بین RRH و BBU مشابه ساختار سنتی بوده و RRH ها نیز توسط معماری زنجیروار با هم در ارتباطند.



شکل ۴.۱: ساختار C-RAN [۲]

۴.۱.۱.۱ ساختار C-RAN

در ادامه ساختار های شبکه دسترسی رادیویی ابری و ساختارهای بهبود یافته ی آن را معرفی می نماییم.

• شبکه های دسترسی رادیویی ابری

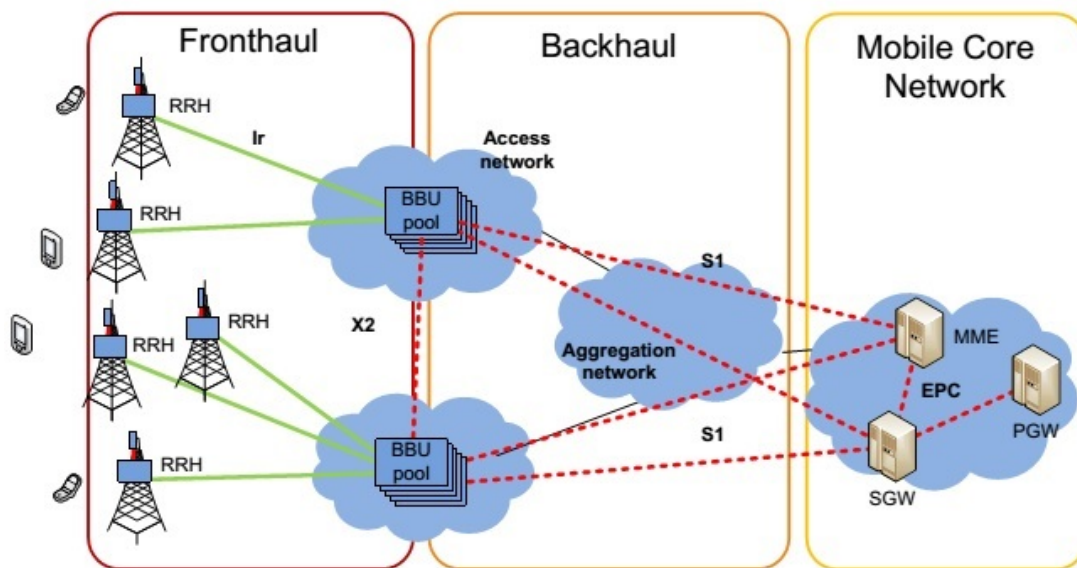
ایده اصلی C-RAN جداسازی بخش رادیویی (RRH)^{۳۲} از واحد پردازشی باند پایه (BBU)^{۳۳} است. از تجمع BBU ها بر روی سرور ابری، BBU-Pool ایجاد می شود. در این ساختار، در راستای بهینه سازی عملکرد BBU ها در مواجهه با ایستگاههای پایه پر ترافیک و کم ترافیک، BBU ها به صورت یک مجموعه ی واحد تحت عنوان BBU Pool در آمده اند که این مجموعه بین چندین سلول به اشتراک

^{۳۲}Radio Remote Head

^{۳۳}Baseband Unit

گزارده شده و مطابق شکل زیر مجازی سازی می شود. در توضیح بیشتر این ساختار می توان این گونه عنوان کرد که BBU Pool به عنوان یک خوشه ی مجازی در نظر گرفته می شود که شامل پردازش گرهایی می باشد که پردازش های باند پایه را انجام می دهند. ارتباط بین BBU ها در ساختار های فعلی به شکل X_2 برقرار می شود که در این ساختار ارتباط بین خوشه ها از فرم جدید X_2 تحت عنوان $X_2 + X_2$ برقرار می شود.

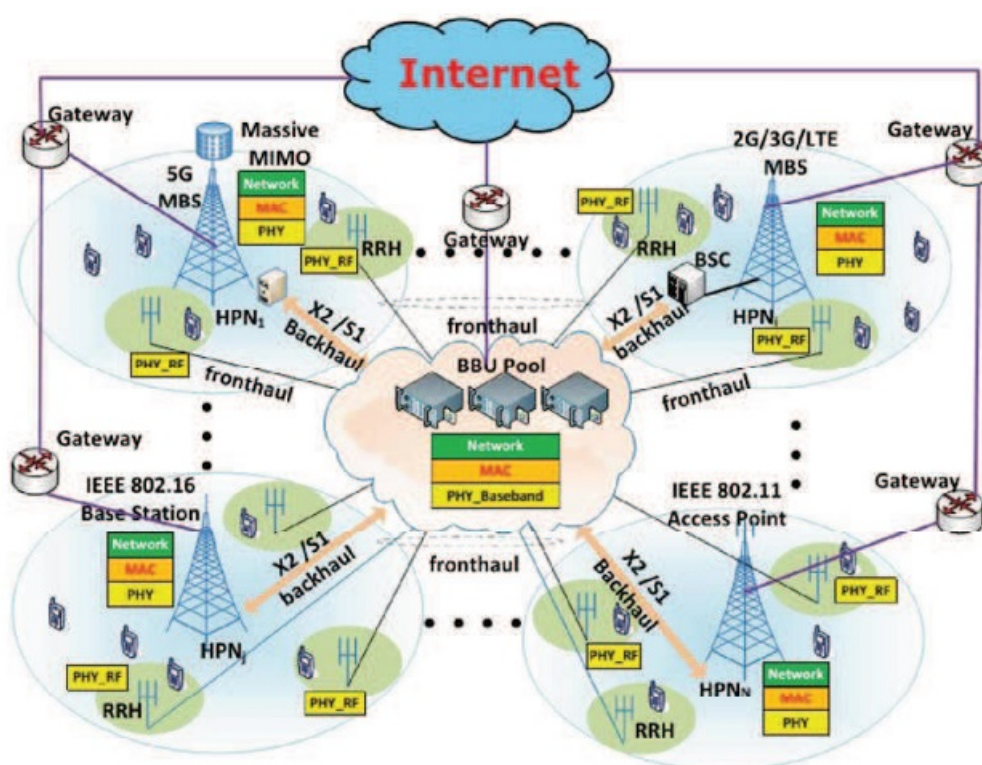
در شکل ۵.۱ ساختار کلی شبکه ی C-RAN در سیستم های LTE نمایش داده شده است. همان طور که در شکل قابل مشاهده می باشد ساختار کلی شبکه C-RAN به دو بخش backhaul و fronthaul تقسیم بندی شده است. بخش fronthaul شبکه به مرحله ی اتصال سایت های RRH به BBU Pool به اتصال backhaul و بخش BBU Pool هسته ی شبکه ی سیار اطلاق می شود. همان گونه که قبلا ذکر شد RRH ها در نزدیکی انتن نصب شده و از طریق لینک های انتقالی نوری با پهنای باند وسیع و تاخیر کم به پردازشگرهای قوی در BBU متصل می شوند. توسط این لینک های انتقالی است که سیگنال های دیجیتال باند پایه از نوع IQ بین RRH و BBU انتقال می یابند [۲].



شکل ۵.۱: ساختار شبکه ی C-RAN [۲]

• شبکه های دسترسی رادیویی ابری نامتجانس (H-CRAN)

برای غلبه بر چالش های شبکه های C-RAN با محدودیت های fronthaul ، شبکه های دسترسی ابری نامتجانس (H-CRAN) معرفی می شود [۴، ۳، ۱۲].



شکل ۶.۱: ساختار شبکه‌های دسترسی ابری نامتجانس [۲]

کاربر و صفحه ی کنترلگر در چنین شبکه هایی از هم مجزا می باشند. که در این شبکه ها ، نودهای توان بالا^{۳۴} HPN ، عمدتاً برای فراهم کردن پوشش بدون درز و اجرای عملکرد صفحه کنترل می باشد. در حالی که RRH ها برای فراهم نمودن سرعت بالای نرخ داده برای انتقال بسته در ترافیک قرار گرفته اند. HPN ها از طریق لینکهای backhaul به BBU Pool متصلند (برای هماهنگ کردن تداخل). ساختار این شبکه شبیه به ساختار C-RAN می باشد . همانطور که در شکل (۶.۱) نشان داده شده است ، تعداد زیادی RRH ، همراه با انرژی مصرفی کم در ساختار H-CRAN ، با یکدیگر در BBU Pool مرکزی ، همکاری می کنند تا گین مشترک بالایی بدست آورند. تنها ، فرکانس رادیویی جلو ، (RF) و عملکردهای پردازشی ساده ، در RRH ، صورت می گیرد ، در حالی که پردازشهای مهم دیگر ، در BBU Pool انجام می گیرد. همچنین تنها بخشی از عملکردها در لایه ی PHY در RRH به مشارکت می انجامد که این مدل در شکل (۶.۱) نشان داده شده است.

اگرچه ، برخلاف C-RAN ، BBU Pool در H-CRAN ، به HPN ها متصلند که این، برای کاهش تداخل متقابل بین RRH ها و HPN ها از طریق محاسبات ابری متمرکز بر اساس تکنیکهای پردازشی مشترک می باشد. همچنین ، داده و واسط کنترل ، بین BBU Pool و HPN های S_1 و X_2 شناخته شده اند که تعریف آنها بر اساس تعریف استاندارد 3G ایجاد شده است.

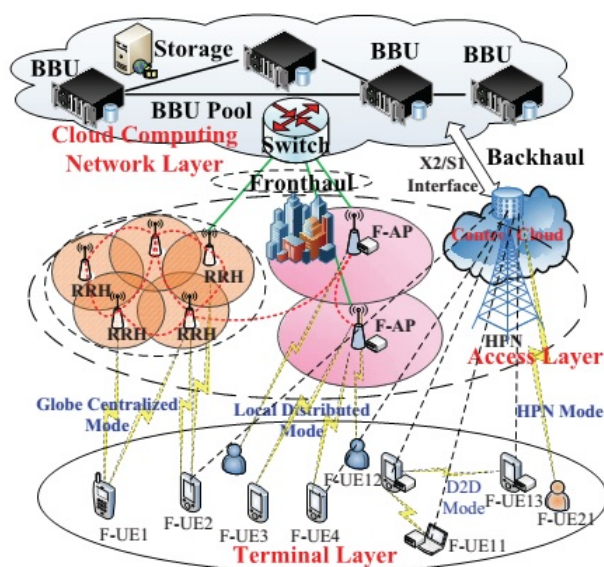
همانطور که سرویسهای صدا ، می توانند به صورت بهینه در طول مد سوئیچ بسته در 4G فراهم گردند ، H-CRAN می تواند به طور همزمان سرویس صدا و داده را پشتیبانی کند. سرویس صدا مرجح به اداره از طریق HPN ها می باشد ، در حالی که ترافیک بسته ی پر داده ، بیشتر توسط RRH اداره می گردد. در مقایسه با ساختار C-RAN ، ساختار H-CRAN ، نیازهای fronthaul را بوسیله ی مشارکت HPN ها برطرف می سازد. با توجه به حضور HPN ها ، سیگنالهای کنترلی و سنبلهای داده در H-CRAN جدا از هم می باشند. تمام کنترل کننده های سیگنال و سیستم هایی که اطلاعات را ارسال می نمایند ، توسط HPN ها به UE ، منتقل می گردد که منجر به سادگی در ظرفیت و در محدودیت تاخیر زمان در لینکهای fronthaul بین RRH ها و BBU Pool می گردد و منجر به صرفه جویی در مصرف انرژی می گردد. همچنین ، برخی از ترافیک های شدید و ناگهانی^{۳۵} و یا سرویس پیام همراه با مقدار داده ی کم ، می تواند به صورت بهینه توسط HPN ها پشتیبانی گردد. مکانیزم کنترل بین ارتباط داشتن و نبود ارتباط ، توسط H-CRAN پشتیبانی می گردد که منجر به حفظ کردن مقدار قابل توجهی Overhead در رادیو بوسیله ی مکانیزم ارتباط جهت دار خالص می گردد. در RRH ، تکنولوژی های مختلف انتقال در لایه ی PHY ، قابل استفاده برای بهبود نرخ انتقال (همانند موج میلیمتری و نور مرئی) می گردد. در HPN ها،

³⁴High Power Node³⁵Burst Traffic

MIMO^{۳۶}، یکی از راه‌های افزایش پوشش در بهبود ظرفیت می‌باشد.

• ساختار دسترسی رادیویی مهی

برای حل کردن مشکلات H-CRAN و C-RAN، نیاز به معرفی ساختار جدید دیگری می‌باشیم که آن را F-RAN می‌نامیم. F-RAN تمام ویژگی‌های مثبت محاسبات ابری و شبکه‌های نامتجانس و محاسبات مهی را همزمان در بر می‌گیرد. محاسبات مهی، اصطلاحی برای جایگزین کردن محاسبات ابری است که مقدار قابل توجهی از ذخیره‌سازی، ارتباطات، کنترل کردن، اندازه‌گیری و مدیریت را در لبه شبکه انجام می‌دهد (نه در کانال و ابر مرکزی) [۴، ۱۲]. سیستم‌های F-RAN تحولی از سیستم‌های C-RAN می‌باشد. برخی از ارتباطات توزیع شده و عملکردهای ذخیره‌سازی در منطق لایه می قرار دارد. همچنین چهار نوع ارتباطات ابری تعریف شده است.



شکل ۷.۱: مدل سیستم F-RAN [۴]

- ابر ذخیره‌گر و ارتباطات مرکزی جامع: که همانند ابر مرکزی C-RAN می‌باشد
- ابر کنترل مرکزی: که برای تکمیل عملکردهای کنترلی می‌باشد و در HPN ها قرار دارد
- ابر ارتباطات منطقی توزیع شده که در برنامه‌های محاسبات مهی و ابزارهای این محاسبات قرار دارد.

³⁶Multiple Input Multiple Output

- ابر ذخیره گر منطق توزیع شده: که همانند قبل در F-RAN قرار دارد.

در این ساختار، برای کاهش تاخیر ناشی از انتقال داده‌ها به ابر مرکزی، ساختارهای RRH را دارای حافظه قرار می‌دهیم که برای ارتباطات محلی، به جای اینکه پردازش‌ها در BBU Pool صورت بگیرد، بدون نیاز به انتقال به ابر مرکزی، درون RRH‌ها انجام پذیرد.

۲.۱.۱ xRAN

xRAN در سال ۲۰۱۶ با هدف استانداردسازی یک جایگزین انعطاف پذیر و باز برای RAN مبتنی بر سخت افزار سنتی بدست آمده است. در این ساختار، سه حوزه‌ی مهم مورد بررسی قرار گرفته است. اولین حوزه‌ی مورد بررسی، جداسازی بخش صفحه‌ی کنترل^{۳۷} از صفحه‌ی کاربر^{۳۸} می‌باشد. حوزه‌ی دوم، ساختن یک پشته نرم افزاری eNodeB مدولار که از سخت افزار COTS استفاده می‌کند، می‌باشد. حوزه‌ی سوم مورد بررسی، انتشار رابط‌های باز شمال و جنوب است [۱۳]. در ادامه این سه حوزه به طور دقیق تر مورد بررسی قرار می‌گیرد [۱۴].

- **جداسازی بخش صفحه‌ی کنترل از صفحه‌ی کاربر:** این انتقال صفحه‌ی کنترل، که قبلاً کاملاً به دستگاههای سخت افزاری RAN متصل بود، به دستگاههای محاسباتی در دسترس امکان می‌دهد RAN بتواند به عنوان یک استخر منطقی از ظرفیت، با کارایی بیشتری کار کند. نرم افزار eNodeB از سخت افزار خاص فروشنده جدا می‌شود و الهام بخش نوآوری در هر دو نرم افزار و سخت افزار به صورت مشارکتی اما به طور مستقل است. برنامه نویسی و کنترل زمان واقعی بی سابقه در زیرساخت های RAN به دست آمده است، که به راحتی از برنامه های کاربردی تلفن همراه و خدمات تجاری پشتیبانی می‌کند.
- **ساختن یک پشته نرم افزاری eNodeB مدولار:** رویکرد xRAN به خوبی با طرح های مجازی سازی عملکرد شبکه حامل (NFV) مطابقت دارد، و همچنین منجر به کنترل عملکرد ترافیک با کارایی بالا، مدیریت تداخل و کنترل منابع رادیویی روی سیستم عامل های استاندارد x86 و می‌شود.
- **انتشار رابط های باز شمال و جنوب:** رابط های استاندارد و باز قابیت پشتیبانی از فروشنده های متعدد همکاری اثبات شده دارند. xRAN.org و اعضای آن به تصویب رساندن این رابط ها از طریق فرآیندهای استاندارد منجر به در دسترس قرار دادن معماری xRAN و پشتیبانی مورد نیاز می‌شوند.

³⁷control plane

³⁸user plane

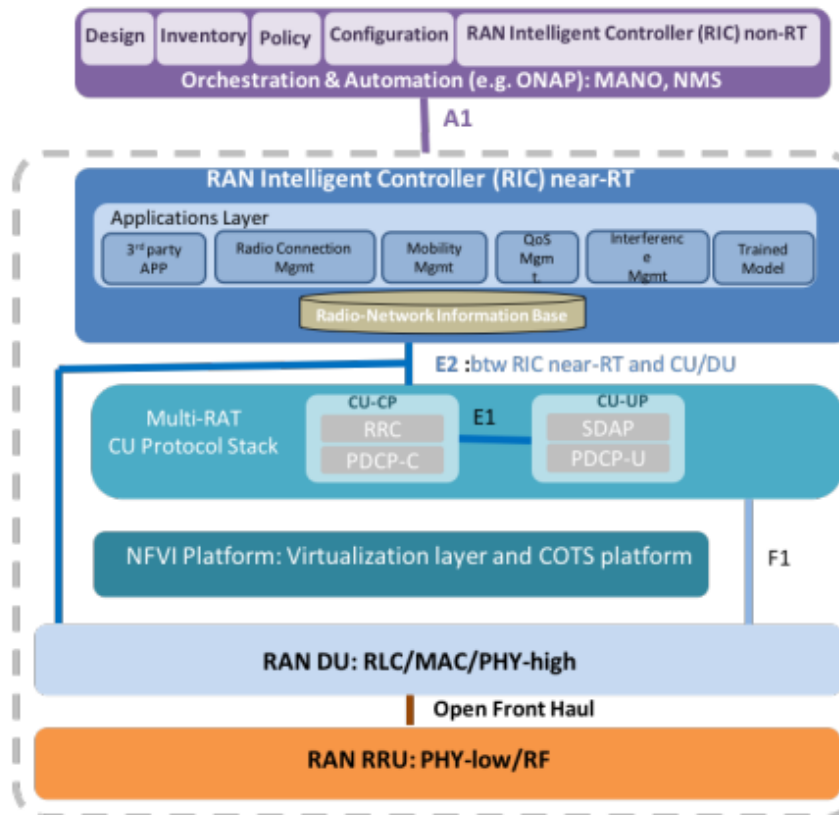
۱.۲.۱.۱ مزایای ساختار xRAN

- جداسازی بخش صفحه ی کنترل از صفحه ی کاربر منجر به برنامه ریزی زمان واقعی بی سابقه و کنترل در زیرساخت RAN می شود که به راحتی برنامه های کاربردی تلفن همراه و خدمات تجاری را پشتیبانی می کند.
- یک پشته eNB مدولار مبتنی بر نرم افزار ، منجر به امکان قرارگیری انعطاف پذیر توابع eNB و کنترل ترکیبی آن با یک برنامه ریز امکان پذیر می شود تا بتواند زمان تاخیر متغیر در fronthaul را کنترل کند.
- رابط های مرزی جنوبی استاندارد، پیاده سازی شبکه با خرید سیستم از چندین شرکت متفاوت را امکان پذیر می سازد و رابط های شمال مرزی، برش کامل شبکه برای بهینه سازی QoE^{۳۹} کاربر را فراهم می کند. رابط های xRAN به خوبی با لبه ابر حامل هماهنگ هستند و اجازه می دهد تا محاسبه و ذخیره سازی منابع در شبکه تلفن همراه به صورت دینامیکی مدیریت شود.
- این ساختار هزینه ی رشد ظرفیت دسترسی رادیویی و هزینه ی بهره برداری را کاهش می دهد.

۳.۱.۱ ORAN

معماری ORAN برای ایجاد زیرساخت های RAN نسل بعدی طراحی شده است. معماری ORAN با تکیه بر اصول هوشمندی و باز بودن، پایه و اساس ساخت RAN مجازی بر روی سخت افزار آزاد ، با کنترل رادیویی ایجاد شده توسط هوش مصنوعی است که توسط اپراتورهای سراسر جهان پیش بینی شده است. این معماری بر روی رابط های استاندارد و تعریف شده ای بنا شده است تا یک زنجیره اکوسیستم با قابلیت باز ایجاد کند که دارای پشتیبانی کامل از استانداردهای تبلیغ شده توسط 3GPP و سایر سازمان های استاندارد صنعت فراهم شود.

³⁹Quality of Experience



شکل ۸.۱: ساختار شبکه‌ی ORAN [۵]

اتحاد ORAN در جستجوی چشم انداز باز بودن و هوشمندی برای شبکه‌های بی سیم نسل بعدی و فراتر از آن است [۵].

- **باز بودن:** ایجاد یک RAN مقرون به صرفه نیاز به باز بودن ارتباط‌ها دارد. رابط‌های باز برای فعال کردن فروشندگان و اپراتورهای کوچکتر به سرعت می‌توانند خدمات خود را معرفی کنند و یا اپراتورها را قادر می‌سازد تا شبکه را متناسب با نیازهای منحصر به فرد خود تنظیم کنند. رابط‌های باز همچنین استقرار چند سازنده‌ای را قادر می‌سازد و اکوسیستم تأمین‌کننده رقابتی‌تر و پر جنب و جوش بیشتری را ایجاد می‌کند. همچنین نرم افزارهای منبع باز و طرحهای مرجع سخت افزار باعث نوآوری سریعتر و دموکراتیک‌تر می‌شود.

- **هوشمندی شبکه‌ها:** با ظهور برنامه 5G پیچیده‌تر و متراکم‌تر شده و خواستار برنامه‌های غنی‌تر می‌شوند. برای کاستن این پیچیدگی نمی‌توان از ابزارهای سنتی انسانی برای استقرار، بهینه‌سازی و بهره‌برداری از شبکه استفاده کرد. در نتیجه، شبکه‌ها باید خود متحرک شوند تا بتوانند از فن‌آوری‌های جدید

مبتنی بر یادگیری برای خودکارسازی عملکرد شبکه‌های عملیاتی و کاهش OPEX استفاده کنند. اتحاد ORAN تلاش خواهد کرد تا از تکنیک‌های یادگیری عمیق در حال ظهور استفاده کند تا بتواند هر لایه از معماری RAN را به طور هوشمند پیاده‌سازی کند. پیاده‌سازی هوشمند هم در مولفه‌ها و خم در سطح شبکه اعمال می‌گردد و منجر به تخصیص دینامیکی منابع رادیویی و بهینه‌سازی بازدهی شبکه می‌گردد. همراه با رابط‌های باز ORAN، اتوماسیون حلقه بسته بهینه‌شده با هوش مصنوعی دست‌یافتنی است و دوره جدیدی را برای عملیات شبکه امکان‌پذیر می‌کند.

- **روش‌های هوش مصنوعی AI^{۴۰} منجر به هوشمندسازی بخش رادیویی با استفاده از نرم افزار تعریف شده^{۴۱} می‌شود:** مفهوم SDN^{۴۲} که مبنی بر جداسازی بخش صفحه‌ی کنترل CP از صفحه‌ی کاربر UP می‌باشد، در ساختار ORAN مورد بررسی قرار می‌گیرد. این جداسازی منجر به بهبود RRM برای استفاده از زمان غیر واقعی و زمان نزدیک به واقعی در کنترلگر هوشمند شبکه‌ی دسترسی رادیویی RIC^{۴۳} با استفاده از رابط‌های A1 و E2 می‌گردد. همچنین منجر به جداسازی CU از CP/UP می‌شود که از طریق رابط E1 در 3GPP توسعه می‌یابد.

- **مجازی‌سازی بخش RAN:** ابری‌سازی RAN یکی از اصول مهم ساختار ORAN می‌باشد. اپراتورها برای پشتیبانی از شکافهای مختلف در شبکه، الزامات NFVI/VIM را برای تقویت سیستم عامل مجازی ارائه می‌دهند. به عنوان مثال: لایه‌ی بالا بین PDCCP و RLC تقسیم می‌شود و لایه‌ی پایین در PHY تقسیم می‌شود.

- **رابط‌های باز:** معماری مرجع ORAN بر روی مجموعه‌ای از رابط‌های کلیدی بین چندین جزء جدا شده‌ی RAN ساخته شده است. اینها شامل رابط‌های 3GPP پیشرفته (Xn, X2, E1, W1, F1) برای قابلیت همکاری بین چندین شرکت مختلف تولیدکننده است. رابط‌های مشخص شده ORAN Alliance شامل یک رابط fronthaul باز بین DU و RRU، رابط E2 و یک رابط A1 بین لایه Or- chestration/NMS است که شامل عملکرد غیر واقعی زمانی^{۴۴} و عملکرد gNB/eNB حاوی عملکرد RIC نزدیک به زمان واقعی^{۴۵} است.

- **سخت افزار جعبه سفید:** برای بهره‌مندی کامل از مقیاسی از اقتصاد ارائه شده توسط یک رویکرد محاسباتی باز، O-RAN Alliance طرح‌های مرجع سخت‌افزاری و ایستگاه پایه به صورت جعبه

⁴⁰Artificial Intelligent

⁴¹Software Defined

⁴²software defined network

⁴³RAN Intelligent Controller

⁴⁴non real time RIC

⁴⁵near-real time RIC

سفید با کارایی بالا را مشخص می‌کند. سیستم عامل‌های مرجع از یک رویکرد جدا شده پشتیبانی می‌کنند و نقشه‌های مفصلی را برای معماری سخت افزار و نرم افزار ارائه می‌دهند تا هم BBU و RRU را فعال کنند.

- **نرم افزار منبع باز:** اتحادیه ORAN ارزش انجمن‌هایی که منابع باز ارائه می‌دهند را درک کرده و از آنها پشتیبانی می‌کند. بسیاری از مؤلفه‌های معماری ORAN به صورت منبع باز از طریق جوامع موجود تحویل داده می‌شود. این مؤلفه‌ها عبارتند از: کنترلر هوشمند RAN، پشته پروتکل، پردازش لایه PHY و بستر مجازی سازی. چارچوب نرم افزار منبع باز ORAN نه تنها رابط‌های (E1, W1, F1)، (Xn, X2, E2) را پیاده سازی می‌کند، بلکه انتظار دارد که طراحی مرجع را برای نسل بعدی RRM با هوش جاسازی شده ارائه دهد تا RIC را امکان پذیر کند.

ORAN، المانهای شبکه‌ی دسترسی رادیویی را مجازی می‌کند، آنها را جدا کرده و رابط‌های باز مناسب را برای اتصال این عناصر تعیین می‌کند. همچنین، ORAN از روشهای یادگیری ماشین برای هوشمندسازی لایه‌های RAN استفاده می‌نماید. در ساختار نوآورانه‌ی ORAN نرم افزار قابل برنامه ریزی RAN از سخت افزار جدا می‌شود. یکی از مهم ترین خصوصیات ORAN رابط کاربری باز است که به اپراتورهای موبایل این قابلیت را می‌دهد تا بتوانند سرویس‌های مورد نیاز خود را تعریف نمایند.

در ساختار ORAN، واحد توزیع شده DU، نود منطقی می‌باشد که شامل لایه‌های RLC، MAC، و High-PHY است. علاوه بر این، واحد مرکزی CU نود منطقی است که شامل لایه‌های RRC، SDAP و PDCCP می‌باشد. نود منطقی واحد رادیویی RU نیز، شامل لایه‌ی Low-PHY و بخش پردازش رادیویی می‌باشد. ORAN، رابط‌هایی از جمله رابط fronthaul باز را شامل می‌شود که بخش DU را به RU متصل می‌نماید (رابط E2). همچنین رابط A1 بین لایه‌ی orchestration/NMS که شامل تابع غیر واقعی زمان است و eNB/qNB که شامل تابع نزدیک به زمان است.

با افزایش ترافیک تلفن همراه، شبکه‌های تلفن همراه و تجهیزات که آنها را اجرا می‌کند باید نرم افزاری تر، مجازی، انعطاف پذیر، هوشمند و کارآمدتر شوند. اتحادیه‌ی ORAN متعهد است در حال تکامل شبکه‌های دسترسی رادیویی باشد که باعث می‌شود آنها نسبت به نسل‌های قبل بازتر و باهوش تر شوند. تجزیه و تحلیل در زمان واقعی که توسط سیستم‌های یادگیری ماشین تعبیه شده است و مازول‌های پایانی هوش مصنوعی را هدایت می‌کند، باعث تقویت هوش شبکه می‌شود. عناصر شبکه مجازی با رابط‌های باز و استاندارد، جنبه‌های اصلی طرح‌های مرجع توسعه یافته توسط اتحادیه‌ی ORAN خواهد بود. فن آوری‌های موجود از عناصر شبکه منبع باز و جعبه سفید، نرم افزار و اجزای سخت افزاری مهم این طرح‌های مرجع خواهد بود.

۲.۱ مجازی سازی توابع شبکه

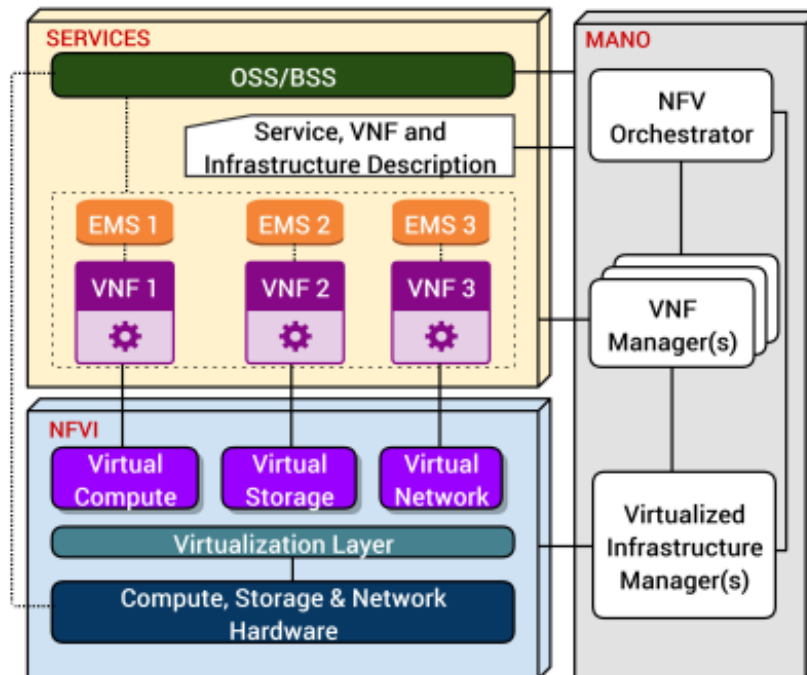
برای بهبود سرویس دهی در نسل پنجم مخابرات، جداسازی المان های نرم افزاری و سخت افزاری شبکه صورت گرفته است و به عنوان مجازی سازی توابع شبکه (NFV)^{۴۶} معرفی شده است. حال توابع شبکه ی مجازی VNF^{۴۷}، بلوکهای توابع سیستم هستند. در نسل پنجم مخابرات انتظار می رود که میزبان چندین سرویس با نیازهای مختلف به طور همزمان باشند. ایده اصلی NFV جداسازی تجهیزات شبکه فیزیکی از توابع اجرا شده بر روی آنها است. این بدان معنی است که یک عملکرد شبکه - مانند فایروال - می تواند به عنوان نمونه ای از نرم افزارهای ساده به فراهم آورندگان سرویس (SP)^{۴۸} ارسال شود. این امر امکان ادغام بسیاری از انواع تجهیزات شبکه بر روی سرورهای با حجم بالا، سوئیچ ها و انبارها را فراهم می کند، که می توانند در مراکز داده، نودهای شبکه توزیع شده و در محل کاربر نهایی قرار بگیرند. به این ترتیب، یک سرویس خاص می تواند به مجموعه ای از توابع شبکه مجازی (VNFs) تجزیه شود، که می تواند در نرم افزارهایی که روی یک یا چند سرور فیزیکی استاندارد در صنعت قرار دارند، اجرا شود. سپس VNF ها ممکن است در مکانهای مختلف شبکه (به عنوان مثال، با هدف معرفی خدمات هدفمند به مشتریان در یک موقعیت جغرافیایی خاص) جابجا شده و خدمات رسانی کنند، بدون اینکه لزوماً به خرید و نصب سخت افزار جدید نیاز داشته باشند. NFV به SPها با انعطاف پذیری بیشتری وعده می دهد تا بتواند بیشتر قابلیت ها و خدمات شبکه خود را به کاربران و سایر خدمات باز کنند و امکان استقرار یا پشتیبانی از سرویس های جدید شبکه را به طور سریعتر و ارزانتر داشته باشند تا بتوانند سرویس بهتری داشته باشند. برای دستیابی به این مزایا، NFV مسیر را برای کاهش اختلافات در نحوه ارائه خدمات شبکه در مقایسه با عملکرد فعلی ایجاد می کند. خلاصه این ویژگی ها به شرح زیر است [۱۵].

- **جدا سازی بخش نرم افزار از سخت افزار:** از آنجا که عنصر شبکه، ترکیبی از سخت افزارها و نرم افزارهای یکپارچه نخواهد بود، تکامل هر دو مستقل از یکدیگر می باشد. که این ویژگی منجر به جداسازی زمان بندی توسعه و نگهداری نرم افزار و سخت افزار می گردد.
- **استقرار عملکرد شبکه انعطاف پذیر:** جدا کردن نرم افزار از سخت افزار به تنظیم مجدد و به اشتراک گذاری منابع زیرساختی کمک می کند، بنابراین، سخت افزار و نرم افزار، باهمدیگر می توانند در زمان های مختلف عملکردهای مختلفی را انجام دهد که به اپراتورهای شبکه کمک می کند تا خدمات جدید شبکه را سریعتر در همان پلت فرم فیزیکی مستقر کنند. بنابراین، مؤلفه ها را می توان در هر دستگاه با قابلیت NFV در شبکه قرار داد و اتصالات آنها به روشی انعطاف پذیر تنظیم کرد.

^{۴۶}network function virtualization

^{۴۷}Virtual network function

^{۴۸}Service Provider



شکل ۹.۱: ساختار NFV [۶]

- **مقیاس گذاری پویا:** جداسدن عملکرد شبکه به اجزای نرم افزاری نعطاف پذیری بیشتری را برای عملکرد واقعی VNF به روشی پویاتر، با توجه به ترافیک واقعی که اپراتور شبکه برای تأمین ظرفیت نیاز دارد، فراهم می کند.

۱.۲.۱ ساختار شبکه

VNF ها برای به اشتراک گذاشتن منابع مختلف فیزیکی و مجازی زیرساخت ها می توانند مستقر و مجدداً تنظیم شوند، تا مقیاس پذیری و کارآمدی سیستم را تضمین کنند که منجر می شود SP ها به سرعت سرویس های جدید را در سیستم وارد کنند. به طور کلی، سه مؤلفه اصلی در NFV وجود دارد: خدمات، NFVI و مدیریت NFV و orchestration^{۴۹} که در شکل (۹.۱) دیده می شود. این مؤلفه ها به شرح زیر بیان می گردد [۶].

۱. خدمات: یک سرویس مجموعه ای از VNF ها است که می توانند در یک یا چند ماشین مجازی پیاده سازی شوند. در بعضی مواقع، VNF ها می توانند در ماشینهای مجازی نصب شده در سیستم عامل یا

⁴⁹NFV-MANO

سخت افزار بطور مستقیم نصب شوند. آنها توسط سرپرستان بومی یا مانیتورهای ماشین مجازی اداره می شوند. معمولاً توسط یک سیستم مدیریت عناصر^{۵۰}، (EMS) که مسئولیت ایجاد، تنظیمات، نظارت، عملکرد و امنیت آن است، اداره می شود. EMS اطلاعات ضروری مورد نیاز سیستم پشتیبانی عملیات^{۵۱} (OSS) را در یک محیط SP فراهم می کند. OSS سیستم مدیریت عمومی است، که همراه با سیستم پشتیبانی از تجارت^{۵۲} (BSS)، به ارائه دهندگان کمک می کند تا چندین سرویس ارتباطی از راه دور را به کار ببندند و مدیریت کنند. (به عنوان مثال سفارش، صورتحساب، تمدید، عیب یابی مشکل و غیره). مشخصات NFV بر ادغام با راه حل های موجود BSS / OSS متمرکز است.

۲. NFVI: زیرساخت های NFV تمام منابع سخت افزاری و نرم افزاری را که شامل محیط NFV است، پوشش می دهد. NFVI شامل اتصال شبکه بین مکان ها، به عنوان مثال، بین مراکز داده و ابرهای ترکیبی عمومی یا خصوصی است. منابع فیزیکی به طور معمول شامل محاسبات، ذخیره سازی و سخت افزار شبکه است که وظیفه ی آن پردازش، ذخیره سازی و اتصال VNF ها از طریق لایه مجازی سازی است و دقیقاً بالای سخت افزار قرار دارد و منابع فیزیکی را چکیده می کند (که به صورت منطقی تقسیم شده و به VNF ها اختصاص می یابد). هیچ راه حل خاصی برای استقرار NFV وجود ندارد. در عوض معماری NFV می تواند از یک لایه مجازی سازی موجود مانند Hypervisor با ویژگی های استاندارد که منابع سخت افزاری را به راحتی استخراج می کند و آنها را به VNF ها اختصاص می دهد، استفاده کند. وقتی این پشتیبانی در دسترس نباشد، اغلب، لایه مجازی سازی از طریق یک سیستم عامل حاصل می شود که نرم افزاری را در بالای سرور غیر مجازی یا با اجرای یک VNF به عنوان یک برنامه اضافه می کند.

۳. NFV-MANO: NFV-MANO از این موارد تشکیل شده است: orchestrator، مدیران VNFs و مدیران زیرساخت مجازی. چنین بلوکی عملکردهای مورد نیاز برای کارهای مدیریتی را که برای VNF ها اعمال می شود، به عنوان مثال تهیه و پیکربندی را ارائه می دهد. NFV-MANO شامل orchestration و مدیریت چرخه منابع فیزیکی یا مجازی است که از مجازی سازی زیرساخت ها و مدیریت چرخه VNF ها پشتیبانی می کند. همچنین شامل بانکهای اطلاعاتی است که برای ذخیره اطلاعات و مدل های داده استفاده می شود که ویژگی های چرخه عمر توابع، خدمات و منابع را تعریف می کند. NFV-MANO روی کلیه وظایف مدیریتی مجازی سازی ویژه لازم در چارچوب NFV تمرکز دارد. علاوه بر این، این چارچوب رابط هایی را تعیین می کند که می توانند برای ارتباطات بین مؤلفه های مختلف NFV، MANO، و همچنین هماهنگی با سیستم های سنتی مدیریت شبکه (یعنی OSS و BSS) مورد استفاده

⁵⁰Element Management System

⁵¹Operations Support System

⁵²Business Support System

قرار گیرند تا امکان عملکرد هر دو VNF و کارکردهای اجرا شده بر روی تجهیزات فراهم شود. به طور خلاصه، اگر برش شبکه با استفاده از فایروال و DPI مستقر شده باشد، آنگاه NFV-MANO وظیفه دارد بگوید این VNF ها در کجای شبکه فیزیکی قرار دارند. همچنین این VNF ها توسط EMS و همان MANO کنترل می شوند.

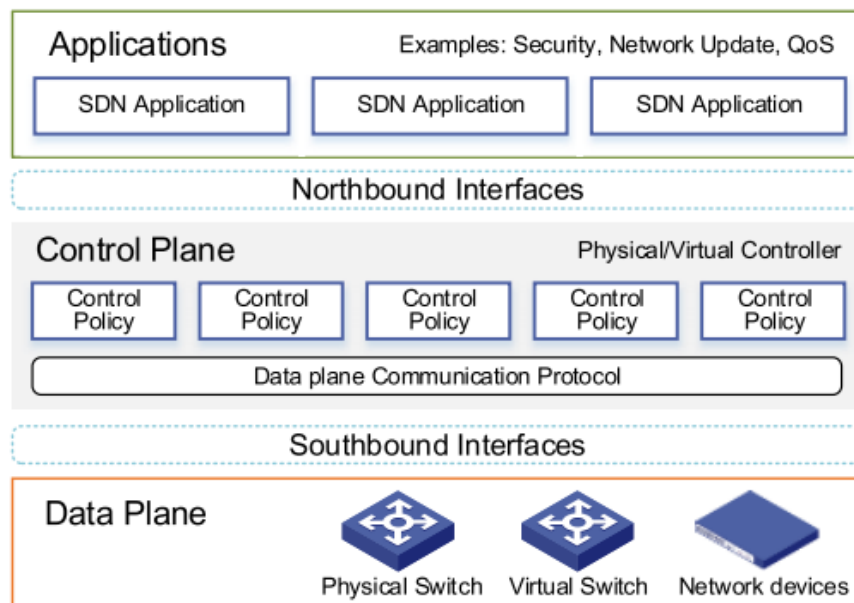
۳.۱ شبکه تعریف شده نرم افزار (SDN)

بنیاد شبکه باز^{۵۳} (ONF) یک مجموعه ای است که به توسعه، استاندارد سازی و تجاری سازی SDN^{۵۴} پرداخت. ONF به طور صریح و دقیق SDN را بدین صورت تعریف کرد: شبکه تعریف شده توسط نرم افزار (SDN) یک معماری شبکه است که کنترل شبکه از ارسال جدا می شود و به طور مستقیم قابل برنامه ریزی است. SDN توسط دو ویژگی تعریف می شود، یعنی جدا شدن صفحه ی کنترل و داده و قابلیت برنامه ریزی در صفحه کنترل. با این وجود، هیچ یک از این دو امضای SDN در معماری شبکه کاملاً جدید نیستند [۱۶]. SDN در اصل یک الگوی شبکه سازی متمرکز است که در آن هوش شبکه (یعنی عملکرد کنترل یا صفحه کنترل) به طور منطقی در یک یا مجموعه ای از موجودیت های کنترل (یعنی کنترل کننده های SDN) متمرکز می شود در حالی که صفحه ی انتقال داده، ساده و چکیده شده برای برنامه های کاربردی می باشد و سرویس های شبکه درخواست خود را از طریق کنترل کننده های SDN بیان می کنند. در حالی که در مورد هسته اصلی شبکه موبایل LTE، EPC صحبت می کنیم، مفهوم SDN برای دستیابی به جدایی واضح بین صفحات کنترل و کاربر در اشخاص SGW و PGW استفاده می شود. با تقسیم دروازه به این روش (یعنی از SGW به SGW-C و SGW-U و از PGW به PGW-C و PGW-U) مقیاس بندی این مؤلفه ها به طور مستقل امکان پذیر است و طیف وسیعی از گزینه های استقرار را نیز ممکن می کند.

پروتکل مورد استفاده بین صفحه ی کنترل و صفحه ی کاربر می تواند یا افزونه پروتکل موجود OpenFlow باشد، که توسط گروه کاری بی سیم و موبایل ONF با (WMWG) رابط های جدید، یعنی Sxa و Sxb ساخته می شود، که توسط 3GPP CUPS تعریف و مشخص می شوند [۱۷]. جداسازی صفحه ی کنترل از کاربر منجر به کنترل بیشتر شبکه بوسیله ی برنامه می گردد که منجر به بهبود تنظیمات و کارآمدی سیستم می گردد. SDN با ساختار برنامه ریزی شده ی قوانین ترافیک، جایگزین امیدوار کننده ای برای فرماندهی ترافیک ارائه می دهد. ساختار SDN در شکل (۱۰.۱) آورده شده است. در این ساختار ۳ لایه ی مختلف وجود دارد که در ادامه بیان می کنیم [۷].

⁵³Open Networking Foundation

⁵⁴Software Defined Network



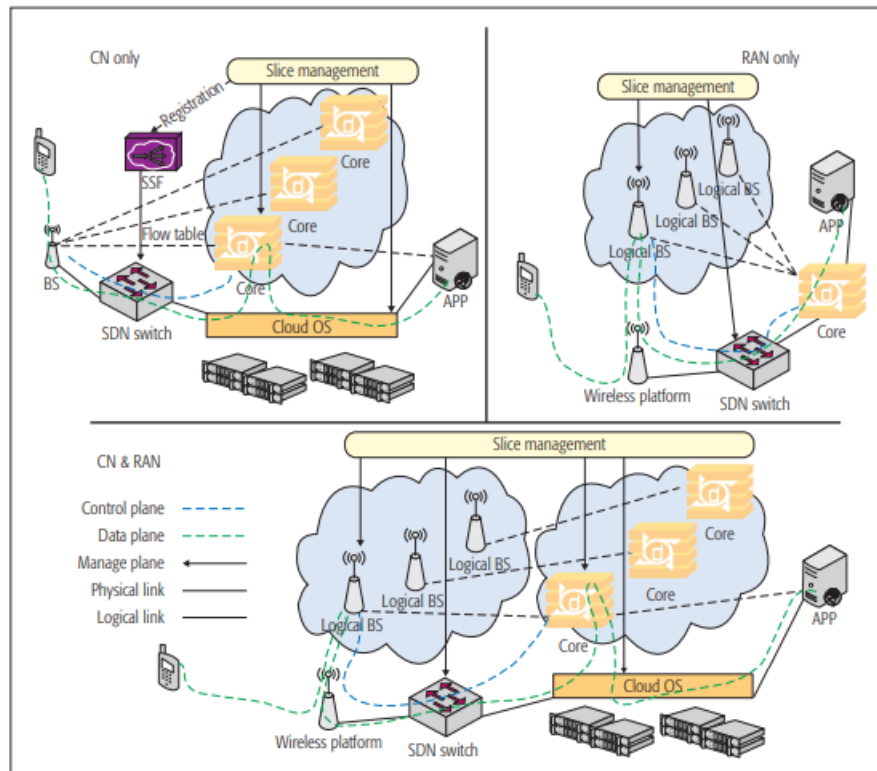
شکل ۱۰.۱: ساختار SDN [۷]

۱. لایه ی برنامه : این لایه مجموعه ای از برنامه های متمرکز بر خدمات شبکه را پوشش می دهد و آنها عمدتاً برنامه های نرم افزاری هستند که با لایه کنترل ارتباط برقرار می کنند.
۲. لایه ی کنترل: به عنوان هسته اصلی SDN ، لایه کنترل از یک کنترلر متمرکز تشکیل شده است که منطقاً نمای شبکه جهانی و پویا را حفظ می کند ، که از لایه برنامه درخواست می کند و دستگاه های شبکه را از طریق پروتکل های استاندارد مدیریت می کند.
۳. لایه ی داده: این لایه، زیرساخت ها شامل سوئیچ ها، روترها و لوازم شبکه می باشد. در زمینه SDN ، این دستگاه ها قابل برنامه ریزی هستند و از رابط های استاندارد پشتیبانی می کنند.

۴.۱ برش شبکه

پیش بینی می شود شبکه های 5G چندین سرویس را با نیازهای مختلف به طور همزمان پشتیبانی کند. برش شبکه^{۵۵} به عنوان راه حلی برای چنین تقاضا در نظر گرفته شده است. یک برش شبکه، یک شبکه منطقی end-to-end است که خدمات با نیازهای خاص را ارائه می دهد. چندین برش شبکه در یک زیرساخت یکسان اجرا و مدیریت

⁵⁵Network Slicing



شکل ۱۱.۱: سه ساختار برش شبکه [۸]

می شوند و به طور مستقل کار می کنند. برش شبکه با هدف تقسیم منطقی مجموعه توابع و منابع شبکه در یک نهاد شبکه در نظر گرفته شده است که مطابق با خواسته های فنی یا تجاری خاص می باشد. با خرد کردن یک شبکه فیزیکی به چندین شبکه منطقی، برش شبکه می تواند از خدمات متناسب با تقاضا برای سناریوهای برنامه مشخص در همان زمان با استفاده از همان شبکه فیزیکی پشتیبانی کند. با استفاده از برش شبکه، منابع شبکه می توانند به صورت پویا و کارآمد به برش های شبکه منطقی با توجه به خواسته های QoS مربوطه اختصاص داده شوند [۱۸].

پیاده سازی های مختلفی از برش شبکه وجود دارد که شامل برش هسته ی شبکه، برش RAN و برش هر دو بخش می باشد [۸].

- **برش هسته:** هسته ی شبکه (CN) ^{۵۶} به عنوان برش های شبکه، مجازی سازی می شوند که با ویژگی هایی مانند ویژگی های قابل برنامه ریزی و قابل اعتماد بودن که شامل مدیریت حرکت و تأیید اعتبار می باشد. برش های شبکه فقط در CN وجود دارد. بنابراین، نه RAN و نه تجهیزات کاربر (UE) برای های CN برش داده شده نیاز به تنظیم ویژه ندارند. در برش هسته ی شبکه، برش تنها در بخش هسته ی

⁵⁶core network

شبکه است و تمام واسط‌ها و فرایندها، بدون تغییر باقی می‌مانند به جز مواردی که در ابتدا UE ها به شبکه ها وصل می‌شوند ، زیرا UE ها باید به برش صحیح CN ها اختصاص داده شوند.

- **برش شبکه ی دسترسی رادیویی:** برخلاف برش ، CN برش های RAN روی سخت افزار رادیویی و استخر منابع باند پایه، به نام یک سطح بی سیم، اجرا می‌شوند که دارای کشش کمتری نسبت به زیرساخت مجازی بالغ شده در ها CN هستند. با چند BS منطقی، برش های RAN پارامترهای مختلفی از رابط های هوا (به عنوان مثال ، طول نماد ، فاصله زیر حامل ، طول پیشوند چرخه و پارامترهای درخواست تکرار خودکار هیبریدی^{۵۷}) را اعمال می‌کند. علاوه بر این، پارامترهای دیگری مانند انتخاب سلول و آستانه انتقال، و همچنین سیاست های انتقال هماهنگ را می‌توان برای هر برش تعریف کرد تا یک تجربه بی سیم برجسته را به کاربران ارائه دهد.

- **برش هسته و شبکه ی دسترسی رادیویی:** در این سناریو، هر برش از RAN به یک برش از هسته متصل می‌شود، بنابراین اپراتورها می‌توانند یک شبکه منطقی انتهای به مشتریان ارائه دهند. روش انتخاب برش همان روش برش RAN است، بنابراین کاربران پس از دسترسی به سیستم، نیازی به انتخاب برش CN ندارند. این مدل از برش مزایای هر دو مدل از برش را باهم دارد. در نتیجه این روش برش، قادر به برنامه ریزی ویژگی های CN و همچنین دارای قابلیت تغییر رابط های هوایی RAN می‌باشد.

۵.۱ نتیجه گیری

در این فصل ابتدا مروری بر تاریخچه ی مخابرات و ۵ نسل مخابراتی شد. سپس ساختار های مختلف دسترسی رادیویی به طور خلاصه بیان شد و در نتیجه ی آن ساختار CRAN که ساختار ابری است تعریف شد. سپس ساختار xRAN مورد توجه قرار گرفت و در نهایت ساختار ORAN که ترکیب و تکاملی از CRAN و xRAN می‌باشد مورد توجه قرار گرفت.

بعد از بیان ساختارهای رادیویی، ساختار هسته ی شبکه را در نسل پنجم بیان کردیم که شامل SDN و NFV می‌باشد که منجر به جداسازی صفحه ی کنترل از کاربر می‌شود و سیستم هوشمندتر همراه با قابلیت برنامه ریزی بیشتر می‌گردد. در ادامه برش شبکه در بخش رادیویی و هسته و هر دو مورد توجه قرار گرفته شد.

⁵⁷HARQ

۶.۱ خلاصه ای از فصل های آتی

در فصل دوم مروری بر ادبیات پیشین و خلاصه ای از مدل سیستم مقالات موجود، بیان می گردد. در فصل سوم مدل سیستم در نظر گرفته بیان می شود و صورت مسئله به نمایش گذاشته می شود و روش های حل آن بیان می گردد. در فصل چهارم نتایج شبیه سازی قرار داده می شود. در فصل پنجم نیز نتیجه گیری و کارهای آتی مورد نظر بیان می شود.

فصل ۲

مروری بر کارهای پیشین

۱.۲ مقدمه

در این فصل، به مرور کارهای گذشته می پردازیم. ابتدا صورت مسئله مقالات مختلف را بررسی می نمایم که به ترتیب شامل مقالاتی هستند که از برش شبکه استفاده کرده اند. برش شبکه در سه بخش رادیویی، بخش هسته و هردو بخش هسته و رادیویی صورت می گیرد. سپس در زمینه ی سیر عبور از شبکه های دسترسی رادیویی ابری به شبکه های دسترسی رادیویی باز مطالعه نموده و بعد از آن درباره ی جاگیری VNF ها صحبت می کنیم. سپس در مورد روش حل مسئله صحبت می کنیم که شامل مسائل کوله پشتی و بسته بندی جعبه می باشد و در نهایت در مورد روشهای یادگیری تقویتی صحبت می کنیم.

۲.۲ مسائل پیش رو

در این بخش به مطالعه ی مسائل پیش رو می پردازیم.

۱.۲.۲ برش شبکه

برش شبکه یک شبکه منطقی انتها به انتهای مستقل است که بر روی یک زیرساخت فیزیکی مشترک کار می کند و قادر به ارائه خدمات می باشد. در این بخش برش شبکه در بخش رادیویی و هسته و هردو را بررسی می کنیم.

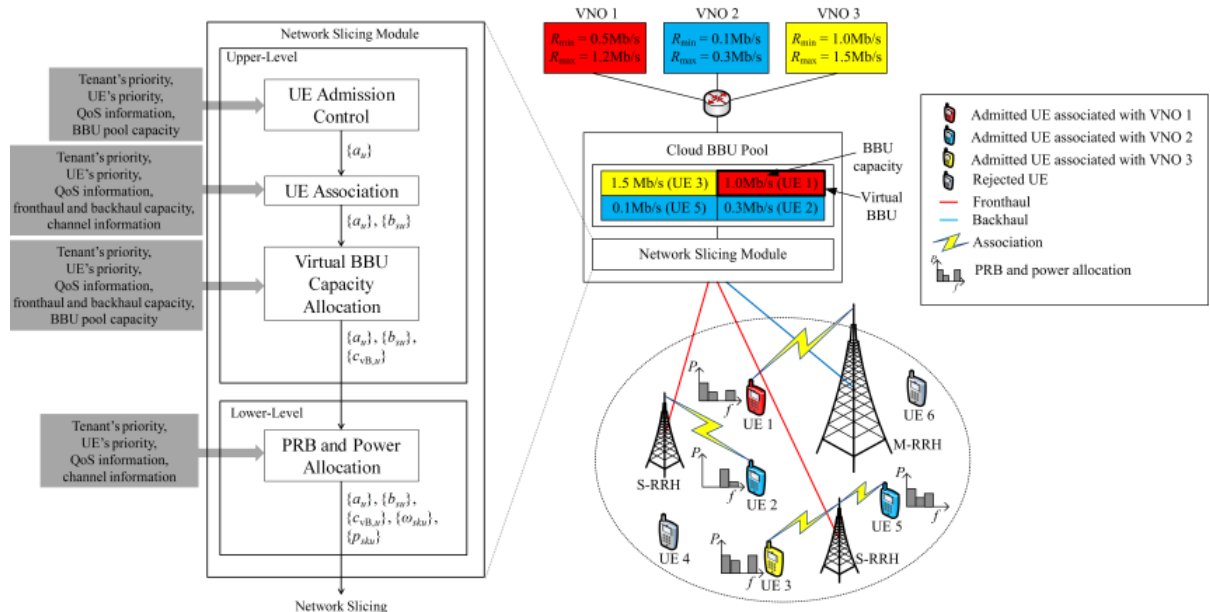
برش RAN یکی از کلیدهای اصلی برای انعطاف پذیری سفارشات و مدیریت مجازی سازی ایستگاه پایه می باشد تا بتواند منابع رادیویی را در میان سرویس های مختلف تقسیم کرده و منجر به سازگاری اپراتورها و برطرف کردن نیاز سرویس ها گردد.

شبکه های دسترسی رادیویی ابر (C-RAN) به عنوان یک چارچوب امیدوار کننده برای سیستم های ارتباط بی سیم نسل پنجم ظاهر شده اند. از آنجا که آنها می توانند پیچیدگی رمزگشایی، مصرف انرژی و دخالت های ناشی از افزایش تراکم تلفن همراه را کاهش دهند [۱۹]. در ادامه در مورد برش شبکه در بخش رادیویی شبکه های دسترسی رادیویی ابری صحبت می کنیم.

در مقاله ی [۹] برش شبکه به صورت دینامیکی در بخش رادیویی مورد بررسی قرار گرفته شده است. برش شبکه در اینجا به عنوان فرآیند تخصیص منابع شبکه به کاربران انجام می شود چارچوب طرح برش شبکه شامل یک سطح بالاتر، که مدیریت کنترل پذیرش کاربران، ارتباط کاربر که شامل تخصیص واحد رادیویی (RRH) برای بیشینه سازی نرخ کاربران و تخصیص ظرفیت منابع باند پایه (BBU) و یک سطح پایین تر، که تخصیص توان و بلوک منابع فیزیکی (PRB) در میان کاربران می باشد. در این مدل فرض می کنیم که هر سرویس دارای شبکه اصلی خود (یا قطعه اصلی شبکه) است که به H-CRAN متصل می شود. سلول بزرگ RRH (M-RRH) و سلولهای کوچک RRHs (S-RRHs) به ترتیب از طریق پیوندهای پشتی و fronthaul به یک استخر ابر BBU متصل می شوند. همچنین، تقسیم C/U در مدل سیستم فرض می شود، که به موجب آن صفحات کنترل و داده از هم جدا می شوند به گونه ای که صفحات کنترل توسط M-RRH در شبکه مدیریت می شود. همانطور که در شکل (۱.۲) مشخص شده است ابتدا پذیرش کاربر مورد توجه قرار می گیرد و سپس کاربر به RRH متصل می شود و پس از آن ظرفیت BBU به آن تخصیص می دهد که تا این بخش از کار در سطح بالا قرار داریم. در سطح بالا، یک مسئله کنترل پذیرش با برنامه نویسی پویا می باشد که در آن پیچیدگی را می توان تنظیم کرد. این مسئله از جنس مسئله ی کوله پشتی^۱ باینری می باشد که با الگوریتم دینامیکی جواب بهینه ی آن بدست می آید. همچنین مسئله ی ارتباط کاربر نیز یک مسئله ی کوله پشتی باینری است که با استفاده از یک الگوریتم حریص با پیچیدگی کم بهینه و حل می شود. مسئله ی تخصیص ظرفیت BBU نیز فرموله شده و با برنامه ریزی خطی حل می شود. حال وارد الگوریتم سطح پایین تر می شویم که تخصیص توان و منبع فیزیکی می باشد. برای مساله ی سطح پایین تر، مشکل تخصیص منابع به عنوان یک مشکل برنامه نویسی mixed-integer غیر محدب است که با استفاده از روش دوگانه لاگرانژ حل می شود.

در مقاله ی [۲۰، ۲۱] برش شبکه در شبکه های دسترسی رادیویی ابری مورد توجه قرار گرفته است. در بخش fronthaul مشکلاتی از قبیل پیچیدگی شبکه و محدودیت نرخ وجود دارد که در برش شبکه، منجر به بهبود آن می شود. علاوه بر این، C-RAN می تواند مجازی سازی مجموعه ای از توابع RAN را امکان پذیر کرده و راه

¹Knapsack



شکل ۱.۲: روند برش شبکه [۹]

را برای اصطلاحاً RAN مجازی باز کند. با این کار می‌توان چندین شبکه مجازی یا برش ایجاد کرد. مقاله ی [۲۰] نشان داده است که استفاده از برش شبکه و برخورداری از سویچ بسته در fronthaul مزایای زیادی را به همراه خواهد داشت که از جمله برخورداری از تقسیمات عملکردی مختلف خواهد بود. همچنین از معایب این کار تاخیر نسبتاً اندکی می‌باشد.

در مقاله ی [۲۲] برش شبکه در بخش رادیویی برای ساختار مه^۲ یا F-RAN در نظر گرفته شده است که در آن دو نمونه برش شبکه برای هات اسپات و سناریوهای وسیله نقلیه با زیرساخت مربوط تنظیم می‌شود. به طور خاص، چارچوب برای برش RAN به عنوان یک مشکل بهینه سازی مشترک برای مقابله با ذخیره کردن و انتخاب حالت است. با توجه به خواسته های کاربران مختلف و منابع محدود، پیچیدگی مسئله بهینه سازی اصلی بسیار زیاد است و همین امر باعث می‌شود که رویکردهای بهینه سازی سنتی به طور مستقیم سخت باشد. برای مقابله با این معضل، یک الگوریتم یادگیری تقویت عمیق ارائه شده است، که ایده اصلی آن این است که سرور ابر تصمیمات صحیحی را در زمینه ذخیره محتوا و انتخاب حالت برای به حداکثر رساندن عملکرد پاداش در وضعیت کانال پویا و وضعیت حافظه نهان ارائه می‌دهد.

در مقاله ی [۲۳، ۲۴] اجرای مفهوم برش در سطح RAN توسط اپراتور شبکه تلفن همراه (MNO) برای پاسخگویی به نیازها می‌باشد. همچنین مساله ی تخصیص منابع (در اینجا پهنای باند) مورد توجه قرار گرفته شد.

^۲Fog Radio Access Network

چالش‌های پیش رو برش RAN نیز مورد بررسی قرار گرفته است که یکی از چالش‌ها شامل طراحی و مدیریت چندین برش در زیرساخت مشترک به روشی کارآمد و در عین حال ضمانت SLA توافق شده برای هر یک از آنها است. این چالش ما را نیازمند مفهوم ایزولاسیون برش می‌کند.

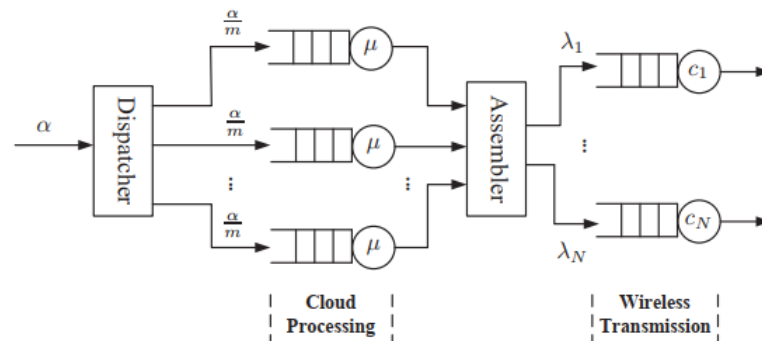
در مقاله ی [۲۵] برش در بخش RAN مورد توجه قرار گرفته است. همچنین در این مقاله یک برنامه تخصیص منابع پویا، با هدف به طور مشترک بهینه سازی مصرف برق و تخصیص پهنای باند در حالی که رضایت از تأخیر مربوطه برای ورود ترافیک پراکنده uRLLC و کیفیت خدمات eMBB را تا حد ممکن ارائه می‌دهد، پیشنهاد می‌کند. طرح پیشنهادی براساس کنترل بهینه توان برای تخصیص منابع آگاه از تأخیر است. در نتیجه در این سیستم هدف مینیمم کردن توان با شروط برآورده شدن شروط پهنای باند و تأخیر که با شرط صف پردازش نشان داده، می‌باشد.

در مقاله ی [۲۶] اجرای عملی برش شبکه پیشنهاد شده است. در مدل پیشنهادی، نویسندگان فرض می‌کنند که در هر شکاف زمانی مشخص، کاربران فقط می‌توانند یک برش شبکه واحد را درخواست کنند. در اینجا تابع هدفی بر اساس نسبت میزان منابع اختصاص داده شده به کاربران در هر زمان t به ظرفیت کل منابع مشخص شده است و هدف بیشینه سازی آن می‌باشد. مدل پیشنهادی بر اساس مسئله multi armed bandit ساخته شده است و نویسندگان سه نوع آن را برای حل جنبه‌های مختلف تخصیص برش شبکه معرفی کرده‌اند. آنها با استفاده از MATLAB مدل بهینه سازی را شبیه سازی کرده و نتایج را با یک الگوریتم حریص مقایسه کردند. آنها همچنین اثبات مفهوم برش شبکه را ارائه دادند.

برش شبکه یکی از فناوری‌های کلیدی است که به شبکه‌های ۵G اجازه می‌دهد منابع اختصاصی به صنایع مختلف (خدمات) ارائه دهند. در مقاله ی [۲۷] نویسندگان یک روش تخصیص منابع (تأخیر بهینه) برای برشهای شبکه حمل و نقل ۵G برای پشتیبانی از خدمات URLLC ارائه داده‌اند. آنها ویژگی‌های منبع شبکه و ویژگی‌های توپولوژی تخصیص منابع در تقسیم شبکه را معرفی کردند.

در [۲۸، ۲۹] ایزوله کردن برش شبکه ی هسته مورد توجه قرار گرفته است. [۲۸] برای کاهش تأثیر حملات DDoS در احراز هویت برش، از ایزوله کردن برش شبکه ی هسته استفاده شده و حل آن با ترکیبی از شبیه سازی و یک آزمایش عملی ارزیابی شده است. نویسندگان [۲۹] دو چالش مهم برش شبکه در بخش هسته مورد توجه قرار داده‌اند که شامل ایزوله کردن برش شبکه و تضمین میزان تأخیر انتها به انتها می‌باشد. در این مقاله، مساله ی بهینه سازی به صورت mixed integer linear programming می‌باشد که تابع هدف درخواست‌های برش ورودی را به سروری که کمترین میزان استفاده از آن شده است، اختصاص داده و مسیری را با حداقل تأخیر پیدا می‌کند. خروجی این مساله VNF ها را به سرور اختصاص می‌دهد.

در این دسته مقالات، سرویس‌ها به دو بخش تقسیم می‌شوند در بخش اول سرویس‌هایی که نسبت به تأخیر حساسند و دسته ی دوم سرویس‌هایی که نسبت به نرخ انتقال حساسند. همچنین در برخی مقالات هر دو



شکل ۲.۲: مدل پردازشی شبکه صف [۱۰].

ویژگی برای یک سرویس مد نظر می باشد. در این مدل های سیستم، تاخیر با استفاده از $M/M/1$ در ساده ترین حالت یا برای نزدیک تر شدن به حالت حقیقی از $M/D/1$ نیز استفاده می شود. می توان در این مدل ها تاخیر را کمینه و نرخ انتقال را بیشینه کرده و یا برای کاربران نرخ را از حد مورد نیاز بیشتر و تاخیر را کمتر از حد مورد نیاز فرض کرد [۳۲، ۳۱، ۳۰، ۱۰]. همانطور که در شکل (۲.۲)، مشخص است، در این شبکه برای هر بخش تعدادی VNF قرار دارند که پردازش ها را انجام می دهند. در مسیر لینک پایین بسته ها با نرخ α به صف های مختلف وارد شده و پس از پردازش با همدیگر ادغام شده و سپس بسته ی هر کاربر از طریق وایرلس منتقل می شوند. در این پردازش ها، از روش $M/M/1$ استفاده شده است. در این مدل مقالات اشاره ی مستقیم به برش شبکه نشده است ولی در آنها ترکیبی از مفهوم برش RAN و Core به چشم می خورد.

۲.۲.۲ رفتن به سمت شبکه های دسترسی رادیویی باز

در فوریه ۲۰۱۸، شبکه دسترسی رادیویی آزاد (O-RAN) با ادغام xRAN و اتحاد C-RAN برای ایجاد سطح جدیدی از باز بودن در شبکه دسترسی رادیویی ایجاد شد که از نسل ۵G و ۶G پشتیبانی می کند. هدف اصلی O-RAN افزایش عملکرد RAN از طریق عناصر شبکه مجازی و واسط های باز است که دارای هوش در RAN است. صراحت و هوش دو ستون اصلی تلاش های انجام شده توسط اتحاد O-RAN است که یک نیروی جهانی متشکل از بیش از ۱۶۰ شرکت کننده از فروشندگان بزرگ، شرکت های کوچک و متوسط، اپراتورهای شبکه، مبتدیان و مؤسسات دانشگاهی است [۳۳].

در مقاله ی [۳۴] مقدمه ای در مورد مفاهیم، اصول و الزامات Open RAN که توسط اتحاد O-RAN مشخص شده، بیان شده است. در این مقاله، به منظور نشان دادن نقش هوش در O-RAN طرح مدیریت منابع رادیویی هوشمندی را برای رسیدگی به ازدحام ترافیک و نشان دادن اثربخشی آن در یک مجموعه داده در دنیای

واقعی پیشنهاد شده است. یک معماری سطح بالا از این سناریوی استقرار که سازگار با الزامات O-RAN است نیز مورد بحث قرار گرفته است. مقاله با چالشهای کلیدی فنی و مشکلات باز برای تحقیقات و توسعه آینده به پایان می‌رسد.

در مقاله ی [۳۵] تعاریف عمومی، ویژگی های اساسی و روند تحقیقاتی فعلی در شبکه های دسترسی رادیویی ابری و مشتقات آن، شبکه های دسترسی رادیویی مجازی و شبکه های دسترسی رادیویی باز ارائه شده است. علاوه بر این، نتایج عملی و آموزه های آموخته شده در مورد محدودیت ها و مسائل پیش بینی نشده مجازی سازی شبکه های دسترسی رادیویی را ارائه داده شده است.

در مقاله ی [۳۶، ۳۷] ساختار و مدیریت منابع رادیویی (RRM) هوشمند و همچنین نقش مدیریت لینک رادیویی (RLM) در بهینه سازی انرژی در RRM در نظر گرفته شده است. ساختار RLM در زیرساخت O-RAN مورد بررسی قرار گرفته است. علاوه بر این، دیدگاه O-RAN و معماری آن مورد توجه قرار گرفته است.

۳.۲.۲ قرار دادن VNF ها

NFV الگویی است که عملکردهای شبکه سنتی را مجازی می‌کند و آنها را در سخت افزارهای عمومی و ابرها در مقابل سخت افزارهای تعیین شده، قرار می‌دهد. در واقع NFV بخش نرم افزار را از سخت افزار جدا می‌نماید. بنابراین یک سرویس داده شده می‌تواند به مجموعه ای از توابع شبکه مجازی (VNF) تجزیه شود، سپس می‌توان آن را در نرم افزارهایی که روی یک یا چند سرور استاندارد فیزیکی صنعت اجرا می‌شوند، پیاده سازی کرد. اپراتورهای شبکه تلفن همراه همراه عهده دار تصمیم گیری مدیریت زیرساخت است. این وظیفه بخشی از تنظیمات شبکه است و شامل تصمیم گیری در مورد قرار دادن VNF های مورد نیاز در سراسر زیرساخت و اختصاص پردازنده، حافظه و منابع ذخیره سازی به VNF ها و مسیریابی داده ها از طریق گره های شبکه می‌باشد. به لطف برش شبکه، شبکه های ۵G از انواع خدمات به روشی انعطاف پذیر و سریع پشتیبانی می‌کنند. در این زمینه، ما به دنبال تصمیم گیری بهینه و با کیفیت بالا در مورد قرار دادن VNF در میان میزبانهای فیزیکی برای تحقق بخشیدن به خدمات هستیم.

در مقالات [۳۸، ۳۹، ۴۰] هدف یافتن تعداد بهینه ی VNF ها در یک زنجیره ی سرویس و قرار گیری VNF های مورد نظر بر روی سرور در هر بازه ی زمانی می باشد تا بتوان میزان هزینه را در سیستم به حداقل رساند. در این مقالات هدف کمینه کردن انرژی های مصرفی در هر بازه ی زمانی می باشد که شامل هزینه ی انرژی مصرفی هر VNF مستقر بر روی سرور در حال کار و هزینه ی استقرار VNF های جدید در هر لحظه ی زمانی می باشد. همچنین مجموع منابع مصرفی VNF های مستقر بر روی هر سرور در هر لحظه می بایست از منابع آن سرور کمتر باشد تا مساله عملی شود. با استفاده از الگوریتم آنلاین این مساله حل شده است. در مقاله ی [۳۹، ۴۰]

نرخ جریان هر VNF و سرور نیز در نظر گرفته شده است و دیتا سنترها و VNF ها به صورت گرافی شبیه سازی شده اند. در مقاله ی [۴۰] الگوریتم روند کردن استفاده کرده که نتیجه ی خوبی را در مقابل مساله ی آفلاین دارد. در مقاله ی [۴۱] مساله ی قرار دادن VNF ها در لبه مورد بررسی قرار می گیرد که در اینجا تخصیص VNF ها در سیستمی با زیرساخت لبه مورد توجه قرار گرفته است و هدف کمینه کردن تاخیر انتها به انتها از هر کاربر به VNF مورد نظر آن می باشد و از روش دینامیکی و پویا برای حل مساله استفاده شده است. در مقاله ی [۴۲] مساله ی قرار دادن VNF در شبکه های فعال SDN/NFV مطالعه شده است، که به طور طبیعی به عنوان یک مساله ی برنامه نویسی باینری (BIP) فرموله شده است. در این مساله قرارگیری VNF زنجیر عملکرد سرویس مورد بررسی قرار گرفته شده است. با استفاده از روش یادگیری تقویت عمیق، الگوریتم قرارگیری VNF مبتنی بر شبکه DDQN^۳ پیشنهاد می کنیم. در مقاله ی [۴۳] مسئله ی بهینه سازی مشترک قرار دادن VNF ها و زمانبندی جریان مطالعه شده است. این مساله از نوع برنامه نویسی عدد صحیح می باشد. برای حالت تک جریان، مساله به سادگی قابل حل است اما برای چندین جریان مساله NP-hard خواهد بود و با استفاده از روش relax کردن لاگرانژ قابل حل می باشد.

۳.۲ روش های حل

در این بخش مروری بر حل مسائل توسط مقالات می نماییم. ابتدا به دو مسئله ی معروف NP-Hard اشاره می کنیم سپس به حل مسائل با استفاده از روش یادگیری تقویتی می پردازیم.

۱.۳.۲ مسائل Np-Hard

در اینجا به دو مسئله ی کوله پشتی و بسته بندی جعبه می پردازیم.

۱.۱.۳.۲ مسئله ی کوله پشتی

یکی از مسائل پیش رو، مسئله ی کوله پشتی^۴ می باشد. این مسئله، از جنس NP-hard می باشد که در این مسئله می خواهیم تعدادی شی با وزنهای مختلف را در تعدادی جایگاه با ظرفیت مشخص قرار دهیم. هدف

^۳Double Deep Q-learning

^۴knapsack

در این مسئله جاگیری بیشترین تعداد اشیاء در این جایگاه‌ها می‌باشد. حل این مسئله با استفاده از روش‌های مختلف صورت می‌گیرد.

در مقاله ی [۹] همانطور که قبل تر اشاره شد، مسئله ی پذیرش کاربر و ارتباط کاربر از جنس کوله پشتی می‌باشد که به ترتیب با استفاده از الگوریتم دینامیکی و الگوریتم حریص تعریف شده در مقاله حل می‌گردد.

در مقاله ی [۴۴] یک راه حل جامع شامل برش شبکه، پیش بینی ترافیک، کنترل پذیرش و برنامه ریزی برای یک سیستم شامل برش شبکه ۵G ارائه شده است. راه حل کنترل پذیرش به یک مسئله کوله پشتی هندسی (دو بعدی) ترسیم شده و دو الگوریتم کم پیچیدگی به ترتیب برای درخواست های برش شبکه منظم و نامنظم طراحی شده اند.

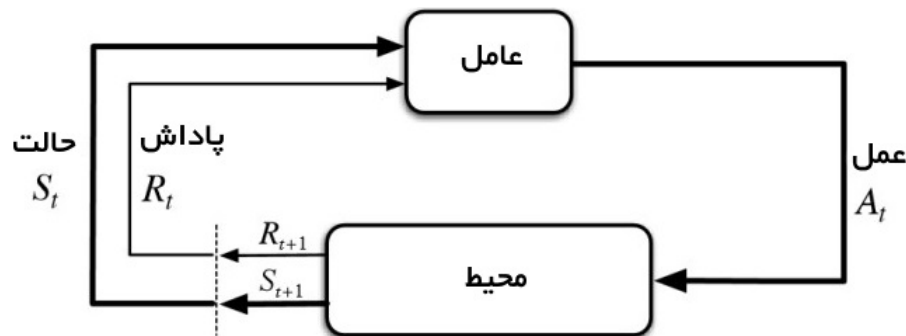
۲.۱.۳.۲ مسئله ی بسته بندی جعبه

در این مسئله هدف قرار دادن تعدادی شیء در تعدادی جعبه با ظرفیت مشخص می‌باشد. در مسئله ی بسته بندی جعبه ^۵ هدف کمینه کردن تعداد جعبه های ورودی با فرض اینکه همه ی اشیا در آن جا شوند.

در مقاله ی [۴۵] مسئله ی تخصیص منابع و بدست آوردن انرژی بهینه در ساختار H-CRAN می‌باشد. در این مقاله هدف تخصیص همزمان منابع ایستگاه رادیویی RRH و باند پایه BBU می‌باشد. این مسئله به دو بخش مجزا برای تخصیص منابع هر بخش شکسته می‌شود. بخش دوم که مربوط به زمان بندی و برنامه ریزی BBU می‌باشد به فرم یک مسئله ی باینری بسته بندی جعبه نوشته می‌شود. برای حل این الگوریتم از روش های مختلفی از جمله First-Fit(NF)، First-Fit(FF)، First-Fit-Decreasing(FFD) و Best-Fit-Decreasing(BFD) می‌باشد که الگوریتم به کار رفته در مسئله از نوع روش BFD است که جواب بهتری در مقایسه با روشهای دیگر دارد.

در مقاله ی [۴۶] به صورت همزمان قرار دادن VNF ها و تخصیص منابع محاسباتی مورد هدف قرار داده شده است. این مسئله به صورت برنامه نویسی خطی mixed-integer می‌شود که بعد از تغییرات به صورت برنامه نویسی خطی عدد صحیح می‌توان نوشت که به فرم مسئله ی بسته بندی جعبه خواهد بود که با استفاده از الگوریتم مشابه BFD حل می‌گردد.

⁵bin packing



شکل ۳.۲: یادگیری تقویتی

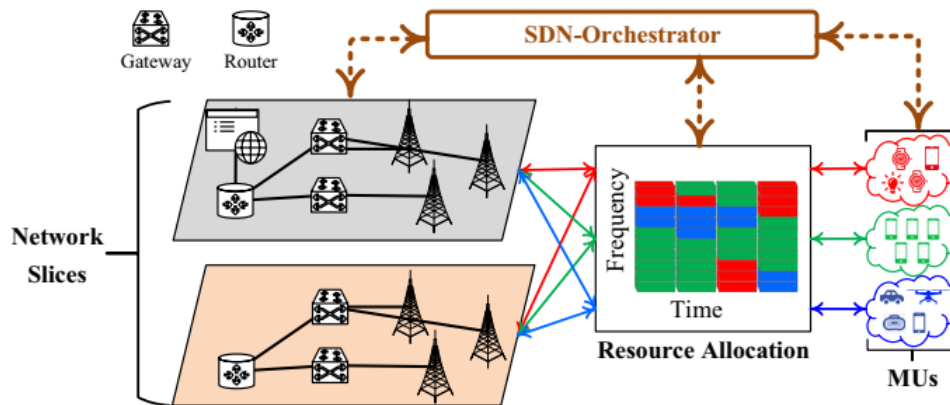
۲.۳.۲ روشهای یادگیری تقویتی

در این بخش تمرکز ما بر روی مقالاتی است که از روش یادگیری عمیق در حل مسئله استفاده می‌کند. یادگیری تقویتی^۶ در حال حاضر یکی از موضوعات داغ پژوهشی محسوب می‌شود و محبوبیت آن روز به روز در حال افزایش است. یادگیری تقویتی گونه‌ای از روش‌های یادگیری ماشین است که یک عامل (agent) را قادر به یادگیری در محیطی تعاملی با استفاده از آزمون و خطاها و استفاده از بازخوردهای اعمال و تجربیات خود می‌سازد. اگرچه هم یادگیری نظارت شده و هم یادگیری تقویتی از نداشت بین ورودی و خروجی استفاده می‌کنند، اما در یادگیری تقویتی که در آن بازخوردهای فراهم شده برای عامل، مجموعه صحیحی از اعمال، جهت انجام دادن یک وظیفه هستند، برخلاف یادگیری نظارت شده از پاداش‌ها و تنبیه‌ها به عنوان سیگنال‌هایی برای رفتار مثبت و منفی بهره برده می‌شود. در یادگیری تقویتی هدف پیدا کردن مدل داده مناسبی است که پاداش کل را برای عامل، بیشینه می‌کند. تصویر زیر ایده اساسی و عناصر درگیر در یک مدل یادگیری تقویتی را نشان می‌دهد. برخی از اصطلاحاتی که عناصر یک مساله یادگیری تقویتی را تشریح می‌کنند در ادامه بیان شده است.

- محیط: (Environment) جهان فیزیکی که عامل در آن عمل می‌کند.
 - حالت: (State) موقعیت کنونی عامل
 - پاداش: (Reward) بازخورد از محیط
 - سیاست: (Policy) روشی برای نگاشت حالت عامل به عمل
 - ارزش: (Value) پاداش آینده که یک عامل با اقدام به یک عمل در یک حالت خاص به آن دست می‌یابد.
- یادگیری تقویتی عمیق DRL^۷، از شبکه‌های عصبی عمیق برای حل مسائل یادگیری تقویتی استفاده می‌کند،

^۶Reinforcement Learning

^۷Deep Reinforcement Learning



شکل ۴.۲: سناریوی ارسال لینک بالا و پایین در برشهای شبکه

از این رو در نام آن از کلمه عمیق استفاده شده است. با در نظر گرفتن Q-Learning که یادگیری تقویتی کلاسیک محسوب می‌شود و Deep-Q-Learning می‌توان تفاوت آن‌ها با یکدیگر را دید. در رویکرد اول، از الگوریتم‌های سنتی برای ساخت جدول Q استفاده می‌شود تا به عامل در یافتن اقدامی که باید در هر حالت انجام شود کمک کند. در دومین رویکرد، از شبکه عصبی (برای تخمین پاداش بر مبنای حالت: مقدار Q) استفاده می‌شود.

در مقاله‌ی [۴۷، ۴۸] سناریویی در نظر گرفته شده است که شامل چندین برش در یک شبکه دسترسی رادیویی با ایستگاههای پایه است که از منابع فیزیکی مشترک (به عنوان مثال، پهنای باند) استفاده می‌کنند. با استفاده از یادگیری تقویت عمیق (DRL) با در نظر گرفتن تقاضای مختلف خدمات به عنوان وضعیت محیط و منابع اختصاص یافته به عنوان عمل محیط، این مشکل حل می‌شود. برای کاهش نویز و رسیدن به سطح انتظار خدمات، از روش GAN در بخش عمیق الگوریتم استفاده شده است که منجر به حداقل رساندن اختلاف بین توزیع مقدار-عمل تخمین زده شده و توزیع ارزش عمل هدف می‌شود. برای یافتن سیاست بهینه ی تخصیص منابع از روش DDQN^۸ استفاده می‌شود.

در مقاله ی [۴۹] الگوریتم تخصیص منابع برش شبکه انتها به انتها مبتنی بر DQN^۹ پیشنهاد شده است که برای سناریوهای چند برش و چند سرویس مناسب است. در این سیستم دو مدل سرویس ارائه شده است که اولی بر مبنای نیاز به رسیدن به نرخ خاص و دومی نیازمند داشتن تاخیر کم می باشد. هدف در این سیستم رسیدن به بیشینه نرخ دسترسی است. برای رسیدن به هدف مورد نظر برشها به دو بخش دسترسی و اصلی تقسیم شده اند. در اینجا الگوریتم به طور مشترک برشهای شبکه دسترسی رادیویی و برشهای شبکه اصلی را در نظر می گیرد تا

^۸Double Deep Q-Network

^۹deep Q-Network

منابع را به صورت دینامیکی طوری اختصاص دهد که حداکثر میزان کاربران به شبکه دسترسی داشته و به بیشینه نرخ برسد. این سیستم به صورت برنامه‌ی mixed-integer نوشته می‌شود و مسئله به صورت دو مسئله‌ی کوله‌پشتی و اتصال لینک‌ها بیان می‌شود و با استفاده از روش DQN حل می‌گردد.

فصل ۳

تخصیص منابع در شبکه های Open RAN

۱.۳ مقدمه

در این فصل، هدف تخصیص منابع در شبکه های Open RAN در لینک فروسو می باشد که شامل تخصیص توان و برش های شبکه است. در این بخش فرض بر این است که در شبکه ی نسل پنجم مخابرات از زیرساخت Open RAN استفاده شده است. این شبکه سرویس هایی از قبیل IoT، سرویس تلفن، پیامک و ... در اختیار کاربران قرار می دهد. در اینجا مفهوم برش شبکه در سیستم بدین صورت به کار رفته که به جای دیدن کاربران به صورت مجزا، کاربرانی که از یک سرویس خاص استفاده می نمایند در دسته ی آن سرویس قرار گرفته و دسته بندی می شوند. همچنین برش هایی از شبکه در اختیار کاربران هر سرویس خاص، قرار می گیرد که هر برش شبکه شامل تعدادی واحد رادیویی^۱ (RU)، بلوک های منبع فیزیکی^۲ (PRB)، یک واحد توزیع شده^۳ (DU)، یک واحد مرکزی^۴ (CU) می باشد که هر واحد توزیع شده و مرکزی شامل تعدادی توابع شبکه ی مجازی شده^۵ (VNF) می باشد.

در این سیستم هدف، بررسی مسئله ی برش شبکه در بخش RAN و بخشی از هسته در شبکه های دسترسی باز ORAN می باشد. ما مسئله زمانبندی لینک بی سیم، اختصاص برشها به سرویسها و اختصاص مراکز داده به برشها را که از مشکلات بسته بندی جعبه ۳-بعدی است را فرموله می کنیم. هدف این است که به طور مشترک

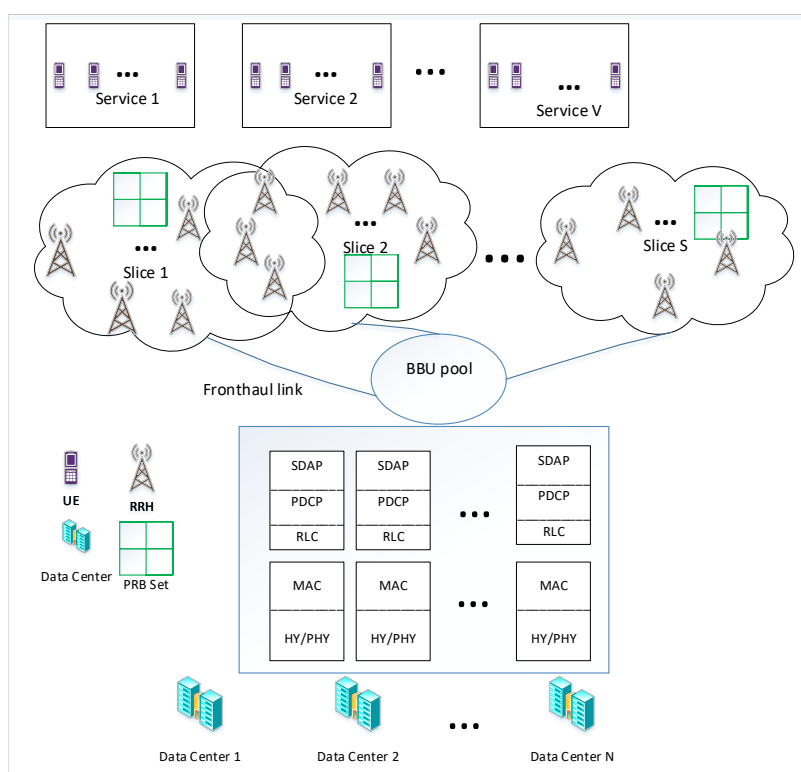
¹Radio Unit

²Physical Resource Block

³Distributed Unit

⁴Centralized Unit

⁵Virtual Network Function



شکل ۱.۳: برش شبکه در ORAN

بهره‌وری انرژی را به حداکثر برسانیم و مصرف توان واحدهای رادیویی را کاهش داده تا هزینه منابع فیزیکی را در یک کانال فروسو به حداقل برسانیم. این مسئله به صورت یک مسئله mixed-integer بیان می‌شود و به دو مسئله‌ی کوچکتر قابل جداسازی است. با استفاده از الگوریتمهای ابتکاری، این مسائل حل شده‌اند. در شکل ۱.۳ لینک فروسو در سیستم ORAN نشان داده شده است که بخش RAN به سه لایه تقسیم می‌گردد که شامل بخش، CU RU و بخش DU می‌باشد. بخش و DU CU بر روی مرکز داده اجرا می‌شوند. کاربران براساس نیاز سرویسهای مختلف تقسیم شده‌اند. نوآوری‌های این کار در ادامه بیان شده است.

- به طور همزمان برش شبکه و تخصیص منابع در سیستم ORAN در نظر گرفته شده است.
- مسئله‌ی پذیرش کاربر و تخصیص خدمات به برش‌ها، و منابع فیزیکی بی‌سیم و مرکز داده به برش‌ها را به عنوان یک مسئله‌ی بهینه‌سازی حل می‌کنیم.
- از الگوریتمهای ابتکاری در جهت حل این مسائل برای رسیدن به پاسخ نسبتاً بهینه استفاده کرده تا در کمترین زمان پاسخ نسبتاً بهینه را بدهد.

۲.۳ مدل سیستم

در این قسمت، سیستم مدل به صورت کامل مورد بررسی قرار می گیرد. فرض می کنیم S برش شبکه داریم که قرار است V سرویس مختلف که شامل کاربرانی است که از سرویس خاص استفاده می نمایند را سرویس دهی نماید. هر سرویس $v \in \{1, 2, \dots, V\}$ شامل U_v کاربر تک آنتنه می باشند که سرویس خاصی را درخواست می نماید. هر برش شبکه $s \in \{1, 2, \dots, S\}$ شامل R_s واحد رادیویی، K_s بلوک های منابع فیزیکی، یک واحد توزیع شده و یک واحد مرکزی که شامل VNF هایی می باشند. همچنین برش های شبکه می توانند منابع مشترک داشته باشند. تمام RU هایی که به یک سرویس خاص خدمت رسانی می کنند به صورت مشارکتی سیگنال به تمام کاربران در آن سرویس ارسال می نمایند. [۵۰، ۵۱] هر واحد رادیویی $r \in \{1, 2, \dots, R\}$ به یک واحد توزیع شده از طریق لینک فیبر نوری با ظرفیت fronthaul محدود متصل می باشد. در سیستم Open RAN دو لایه ی پردازشی که اولی در CU و دومی در DU قرار گرفته است که پردازش ها با VNF ها صورت می گیرد. لایه پایین تر (DU) شامل high-PHY، MAC و RLC می باشد و لایه ی بالاتر (CU) شامل PDCP، SDAP و است. فرض بر این است که M_1 VNF در DU و M_2 VNF در CU قرار دارد. هر VNF به یک برش یا بیشتر تعلق دارد. در s^{th} امین برش شبکه $M_{s,1}$ VNF در DU و $M_{s,2}$ VNF در CU می باشد. VNF های موجود در لایه ی DU و VU به ترتیب دارای ظرفیت محاسباتی μ_1 و μ_2 می باشند.

۱.۲.۳ نرخ قابل دسترس

نرخ قابل دسترس برای i^{th} امین کاربر در v^{th} امین سرویس به صورت زیر نوشته می شود

$$\mathcal{R}_{u(v,i)} = B \log_2(1 + \rho_{u(v,i)}), \quad (1.3)$$

که B پهنای باند سیستم و $\rho_{u(v,i)}$ نسبت سیگنال به نویز i^{th} امین کاربر در v^{th} امین سرویس می باشد که از رابطه ی زیر بدست می آید

$$\rho_{u(v,i)} = \frac{p_{u(v,i)} \sum_{s=1}^S |\mathbf{h}_{R_s,u(v,i)}^H \mathbf{w}_{R_s,u(v,i)}|^2 a_{v,s}}{BN_o + I_{u(v,i)}}, \quad (2.3)$$

که $p_{u(v,i)}$ نشان دهنده ی توان ارسالی توسط RU به i^{th} امین کاربر در v^{th} امین سرویس است و $\mathbf{h}_{R_s,u(v,i)} \in \mathbb{C}^{R_s}$ بردار کانال گین لینک وایرلس از RU ها در s^{th} امین برش شبکه می باشد. همچنین $\mathbf{w}_{R_s,u(v,i)} \in \mathbb{C}^{R_s}$

نشان دهنده ی بردار بیم فرمینگ ارسالی برای RUها در s امین برش شبکه به کاربر i ام در سرویس v ام می باشد. به علاوه، BN_s نشان دهنده ی توان نویز اضافه شونده ی گوسی می باشد و $I_{u(v,i)}$ توان سیگنال تداخلی است. همچنین $a_{v,s} \in \{0, 1\}$ متغیر باینری است که نشان دهنده ی این است که برش شبکه ی s به سرویس v خدمات رسانی می کند یا نه. در صورتی که $a_{v,s} = 1$ برش شبکه ی s به سرویس v خدمات رسانی می کند. در غیر این صورت خدمت رسانی نمی کند.

برای بدست آوردن SNR در فرمول (۲.۳)، فرض می شود که $\mathbf{y}_{U_v} \in \mathbb{C}^{U_v}$ بردار سیگنال دریافتی از همه ی کاربران در سرویس v می باشد

$$\mathbf{y}_{U_v} = \sum_{s=1}^S \sum_{k=1}^{K_s} \mathbf{H}_{\mathcal{R}_s, \mathcal{U}_v}^H \eta_{\mathcal{R}_s} \zeta_{U_v, k, s} a_{v, s} + \mathbf{z}_{U_v}, \quad (3.3)$$

که $\mathbf{x}_{U_v} = [x_{u(v,1)}, \dots, x_{u(v,U_v)}]^T \in \mathbb{C}^{R_s}$ و $\eta_{\mathcal{R}_s} = \mathbf{W}_{\mathcal{R}_s, \mathcal{U}_v} \mathbf{P}_{U_v}^{\frac{1}{2}} \mathbf{x}_{U_v} + \mathbf{q}_{\mathcal{R}_s}$ نشان دهنده ی بردار سمبل ارسالی کاربران سرویس v می باشد. $\mathbf{z}_{U_v} \sim \mathcal{N}(0, N_s \mathbf{I}_{U_v})$ و همچنین N_s توان نویز می باشد. علاوه بر این $\mathbf{q}_{\mathcal{R}_s} \in \mathbb{C}^{R_s}$ نشان دهنده ی نویز کوانتیزاسیون می باشد که از فشرده سازی سیگنال در واحد توزیع شده بدست آمده است. $\mathbf{P}_{U_v} = \text{diag}(p_{u(v,1)}, \dots, p_{u(v,U_v)})$ همچنین در اینجا، $\zeta_{k,s}^{u(v,i)} \in \{0, 1\}$ پارامتر باینری می باشد که نشان دهنده ی این است که i امین کاربر در v امین سرویس امکان ارسال سیگنال خود را از طریق PRB k ام دارد یا نه، در ضمن این PRB متعلق به برش s ام می باشد و یا نه. $\mathbf{h}_{\mathcal{R}_s, \mathcal{U}_v} = [\mathbf{h}_{\mathcal{R}_s, u(v,1)}, \dots, \mathbf{h}_{\mathcal{R}_s, u(v,U_v)}]^T \in \mathbb{C}^{R_s \times U_v}$ نشان دهنده ی ماتریس کانال بین دسته واحد رادیویی \mathcal{R}_s به دسته \mathcal{U}_v کاربران می باشد. بردار کانال بین s امین برش و i امین کاربر در v امین سرویس $\mathbf{h}_{\mathcal{R}_s, u(v,i)} \in \mathbb{C}^{R_s}$ به صورت زیر نشان داده شده است

$$\mathbf{h}_{\mathcal{R}_s, u(v,i)} = \beta_{\mathcal{R}_s, u(v,i)}^{\frac{1}{2}} \mathbf{g}_{\mathcal{R}_s, u(v,i)}, \quad (4.3)$$

که در اینجا $\mathbf{g}_{\mathcal{R}_s, u(v,i)} \sim \mathcal{N}(0, N_s \mathbf{I}_{U_v})$ نشان دهنده ی بردار کانال فیدینگ سریع و تخت می باشد و $\beta_{\mathcal{R}_s, u(v,i)} = \text{diag}(b_{r(s,1), u(v,i)}, \dots, b_{r(s, R_s), u(v,i)})$ نشان دهنده ی ماتریس فیدینگ مقیاس بزرگ می باشد. در اینجا فرض بر این است که اطلاعات حالت کانال CSI، به صورت کامل می باشد.

بردار بیم فرمینگ zero forcing می باشد که برای $\mathbf{W}_{\mathcal{R}_s, \mathcal{U}_v} = [\mathbf{w}_{\mathcal{R}_s, u(v,1)}, \dots, \mathbf{w}_{\mathcal{R}_s, u(v,U_v)}] \in \mathbb{C}^{R_s \times U_v}$ مینیمم کردن تداخل می باشد و بدین صورت است

$$\mathbf{W}_{\mathcal{R}_s, \mathcal{U}_v} = \mathbf{H}_{\mathcal{R}_s, \mathcal{U}_v} (\mathbf{H}_{\mathcal{R}_s, \mathcal{U}_v}^H \mathbf{H}_{\mathcal{R}_s, \mathcal{U}_v})^{-1}. \quad (5.3)$$

توان تداخلی کاربر i ام به سرویس v ام به صورت زیر بیان می شود

$$\begin{aligned}
 I_{u(v,i)} = & \underbrace{\sum_{s=1}^S \sum_{n=1}^S \sum_{\substack{l=1 \\ l \neq i}}^{U_v} \gamma_{\setminus} p_{u(v,l)} a_{v,s} \zeta_{u(v,i),n,s} \zeta_{u(v,l),n,s}}_{\text{interference (intra-service)}} \\
 & + \underbrace{\sum_{\substack{y=1 \\ l \neq v}}^V \sum_{s=1}^S \sum_{n=1}^S \sum_{l=1}^{U_y} \gamma_{\setminus} p_{u(y,l)} a_{y,s} \zeta_{u(v,i),n,s} \zeta_{u(y,l),n,s}}_{\text{interference (inter-service)}} \\
 & + \underbrace{\sum_{s=1}^S \sum_{j=1}^{R_s} \sigma_{q_{r(s,j)}}^2 |h_{r(s,j),u(v,i)}|^2 a_{v,s}}_{\text{interference noise (quantization)}}
 \end{aligned} \tag{۶.۳}$$

که $\gamma_{\setminus} = |h_{\mathcal{R}_s, u(v,i)}^H \mathbf{w}_{\mathcal{R}_s, u(y,l)}|^2$ و $\gamma_{\setminus} = |h_{\mathcal{R}_s, u(v,i)}^H \mathbf{w}_{\mathcal{R}_s, u(v,l)}|^2$ همچنین

$\sigma_{q_{r(s,j)}}$ واریانس نویز کوانتیزاسیون j امین واحد رادیویی در برش s می باشد. سیگنال تداخلی برای هر کاربر از سیگنالهای کاربرانی بدست می آید که از PRB مشترکی استفاده نموده اند. در صورت قرار دادن P_{max} به جای $p_{u(y,l)}$ و $p_{u(v,l)}$ باند بالایی $\bar{I}_{u(v,i)}$ برای $I_{u(v,i)}$ بدست می آید بنابراین، $\forall v, \forall i$ از قرار گرفتن $\bar{I}_{u(v,i)}$ به جای $I_{u(v,i)}$ در رابطه ی (۱.۳) و (۲.۳) بدست می آید.

فرض کنید $\bar{p}_{r(s,j)}$ نشان دهنده ی سیگنال ارسالی از j امین واحد رادیویی در s امین برش می باشد. از رابطه ی (۳.۳) داریم

$$\bar{p}_{r(s,j)} = \sum_{v=1}^V \mathbf{w}_{r(s,j), \mathcal{U}_v} P_{\mathcal{U}_v}^{\setminus} P_{\mathcal{U}_v}^{H \setminus} \mathbf{w}_{r(s,j), \mathcal{U}_v}^H a_{v,s} + \sigma_{q_{r(s,j)}}^2. \tag{۷.۳}$$

در این صورت نرخ کاربران در لینک fronthaul بین j امین واحد رادیویی در برش s ام با واحد توزیع شده ی موجود در این برش، بدین صورت می باشد [۵۲، ۵۳]

$$C_{R(s,j)} = \log \left(1 + \sum_{v=1}^V \frac{w_{r(s,j), \mathcal{D}_s} P_{\mathcal{U}_v}^{\setminus} P_{\mathcal{U}_v}^{H \setminus} w_{r(s,j), \mathcal{U}_v}^H a_{v,s}}{\sigma_{q_{r(s,j)}}^2} \right), \tag{۸.۳}$$

که در اینجا $a_{v,s}$ متغیر باینری است که نشان دهنده ی این است که برش s ام به سرویس v خدمات رسانی می

کند یا نه.

۲.۲.۳ میانگین تاخیر

فرض کنید ورود بسته‌های کاربران، فرآیند پواسون با نرخ ورود $\lambda_{u(v,i)}$ برای i امین کاربر در سرویس v ام می باشد. بنابراین، در لایه ی واحد مرکزی (CU)، میانگین نرخ ورود داده ی کاربری که از خدمات برش s استفاده می نماید $\alpha_{s_1} = \sum_{v=1}^V \sum_{u=1}^{U_v} a_{v,s} \lambda_{u(v,i)}$ می باشد. همچنین، میانگین نرخ داده ی ورودی در لایه ی DU، تقریباً مساوی میانگین نرخ داده ی ورودی لایه ی اول $\alpha_{s_1} \approx \alpha_{s_2}$ می باشد. میانگین نرخ داده ی ورودی در لایه ی DU، تقریباً برابر میانگین نرخ داده ی ورودی در لایه ی CU می باشد $\alpha_s = \alpha_{s_1} \approx \alpha_{s_2}$. در اینجا فرض بر این است که در هر لایه متعادل کننده ی ترافیک موجود است که ترافیک ورودی را به طور مساوی بین VNF ها تقسیم می نماید [۱۰، ۳۰، ۳۱]. فرض کنید پردازش باند پایه هر VNF، بوسیله ی پردازش صف M/M/1 نشان داده می شود. هر بسته بوسیله ی یکی از VNF های برش شبکه، پردازش می شود. بنابراین، میانگین تاخیر برش s در لایه ی اول و دوم با استفاده از مدل M/M.1 به این صورت نشان داده می شود

$$\begin{aligned} d_{s_1} &= \frac{1}{\mu_1 - \alpha_s / M_{s,1}}, \\ d_{s_2} &= \frac{1}{\mu_2 - \alpha_s / M_{s,2}}. \end{aligned} \quad (۹.۳)$$

در اینجا $1/\mu_1$ و $1/\mu_2$ میانگین زمان سرویس دهی در لایه ی اول و دوم به ترتیب می باشند. α_s نرخ ورودی است که توسط متعادل کننده ی ترافیک به طور مساوی بین VNF ها تقسیم می شود. نرخ ورودی هر VNF در هر لایه از برش s ، $\alpha_s / M_{s,i}$ ، $i \in \{1, 2\}$ می باشد. d_{str} تاخیر ارسال برای برش s در لینک وایرلس می باشد. نرخ داده ی ورودی لینک وایرلس برابر نرخ داده ی ورودی متعادل کننده ی ترافیک برای هر برش است [۱۰]. همچنین زمان سرویس برای صف ارسال در هر برش s دارای توزیع نمایی با میانگین $1/(R_{tot_s})$ است و می توان به صورت صف M/M/1 مدل کرد [۱۰، ۳۰، ۳۱، ۳۲]. بنابراین میانگین تاخیر لایه ی ارسال بدین صورت است

$$d_{str} = \frac{1}{R_{tot_s} - \alpha_s}; \quad (۱۰.۳)$$

فصل ۳: تخصیص منابع در شبکه های Open RAN تخصیص منابع و قسمت کردن شبکه در ساختار شبکه های دسترسی باز

$R_{tot_s} = \sum_{v=1}^V \sum_{u=2}^{U_v} a_{v,s} R_{u(v,i)}$ کل نرخ قابل دسترس در هر برش می باشد که به سرویس خاص، تخصیص داده است. میانگین تاخیر در هر برش نیز بدین صورت است

$$D_s = d_{s_1} + d_{s_2} + d_{s_{tr}} \forall s. \quad (11.3)$$

۳.۲.۳ مرکز داده ی فیزیکی

هر VNF نیازمند منابع فیزیکی است که شامل حافظه، نگهدارنده و پردازشگر می باشد. فرض کنید منابع مورد نیاز برای f امین VNF در برش s ام به صورت زیر می باشد

$$\bar{\Omega}_s^f = \{\Omega_{M,s}^f, \Omega_{S,s}^f, \Omega_{C,s}^f\}, \quad (12.3)$$

که در اینجا $\bar{\Omega}_s^f \in \mathbb{C}^3$ و $\Omega_{M,s}^f, \Omega_{S,s}^f, \Omega_{C,s}^f$ به ترتیب نشان دهنده ی مقدار حافظه، نگهدارنده و پردازشگر می باشد. همچنین مقدار کل حافظه، نگهدارنده و پردازشگر برای همه VNF ها در یک برش به این صورت تعریف می شود

$$\bar{\Omega}_{j,s}^{tot} = \sum_{f=1}^{M_{s_1}+M_{s_2}} \bar{\Omega}_{j,s}^f \quad j \in \{M, S, C\}. \quad (13.3)$$

همچنین D_c مرکز داده برای سرویس دهی به VNF ها می باشد. هر مرکز داده شامل تعدادی سرور برای سرویس دهی است. مقدار حافظه، نگهدارنده و پردازشگر به ترتیب با τ_{M_j}, τ_{S_j} و τ_{C_j} برای j^{th} امین مرکز داده به ترتیب زیر می باشد.

$$\tau_j = \{\tau_{M_j}, \tau_{S_j}, \tau_{C_j}\},$$

در این مدل سیستم، تخصیص منابع فیزیکی به VNF ها در نظر گرفته شده است. در اینجا فرض بر این است که $y_{s,d}$ متغیر صفر و یکی است که نشان می دهد مرکز داده ی d ام به s امین برش متصل است یا نه.

۴.۲.۳ صورت مساله

یک معیار مهم برای سنجش بهینگی یک سیستم، بهره‌وری انرژی است که نسبت نرخ کل به توان کل می‌باشد.

$$\eta(\mathbf{P}, \mathbf{A}) := \frac{\sum_{v=1}^V \sum_{k=1}^{U_v} \mathcal{R}_{u(v,k)}}{\sum_{s=1}^S \sum_{i=1}^{R_s} \bar{p}_{r(s,i)}} = \frac{\mathfrak{R}_{tot}(\mathbf{P}, \mathbf{A})}{P_r^{tot}(\mathbf{P}, \mathbf{A})}, \quad (14.3)$$

که در اینجا، $P_r^{tot}(\mathbf{P}, \mathbf{A}) = \sum_{s=1}^S \sum_{i=1}^{R_s} \bar{p}_{r(s,i)}$ توان کل مصرفی تمام واحدهای رادیویی در یک برشی باشد. همچنین $\mathfrak{R}_{tot}(\mathbf{P}, \mathbf{A}) = \sum_{v=1}^V \sum_{k=1}^{U_v} \mathcal{R}_{u(v,k)}$ نرخ تمام کاربران مصرفی تمام سرویس‌ها می‌باشد. فرض کنید توان مصرفی پردازش باند پایه در هر مرکز داده‌ی d که به VNF های یک برش s متصل می‌باشد با $\phi_{s,d}$ نشان داده شده است. بنابراین می‌توان توان کل سیستم را برای کلیه مرکز داده‌های فعال که به برش متصل هستند، بدین صورت نشان داد

$$\phi_{tot} = \sum_{s=1}^S \sum_{d=1}^{D_c} y_{s,d} \phi_{s,d}.$$

همچنین، یک تابع هزینه برای قرار دادن VNF ها در مرکز داده‌ها بدین عنوان تعریف شده است

$$\psi_{tot} = \phi_{tot} - \nu \sum_{d=1}^{D_c} \sum_{v=1}^V y_{s,d} a_{v,s} \quad (15.3)$$

که ν متغیر طراحی برای ارزش بین اولین ترم (۴.۴) است که عبارت است از کل مصرف انرژی منابع فیزیکی و ترم دوم که مقدار برش‌های پذیرفته شده برای داشتن منابع فیزیکی نشان داده شده است. هدف ما این است که مجموع نرخ را به حداکثر رسانده و توان حاصل از کل (حداکثر توان تمام واحدهای رادیویی و کل مصرف انرژی پردازش باند پایه را در کلیه مرکز داده‌ها) به طور همزمان و با وجود محدودیت‌هایی که به شرح زیر نوشته شده است، به حداکثر برساند.

$$\max_{\mathbf{P}, \mathbf{A}, \mathbf{Y}} \quad \eta(\mathbf{P}, \mathbf{A}) + \varphi \frac{1}{\psi_{tot}(\mathbf{Y})} \quad (16.3)$$

$$\text{to subject} \quad \bar{p}_{r(s,i)} \leq P_{max} \quad \forall s, \forall i, \quad (16.3 \text{ ب})$$

$$p_{u(v,k)} \geq 0 \quad \forall v, \forall k, \quad (16.3 \text{ ج})$$

$$\mathcal{R}_{u(v,k)} \geq \mathcal{R}_{u(v,k)}^{min} \quad \forall v, \forall k, \quad (16.3 \text{ د})$$

$$C_{r(s,i)} \leq C_{r(s,i)}^{max} \quad \forall s, \forall i, \quad (۱۶.۳)$$

$$D_s \leq D_s^{max} \quad \forall s, \quad (۱۶.۳)$$

$$\sum_{s=1}^S a_{v,s} \geq 1 \quad \forall s, \quad (۱۶.۳)$$

$$\sum_{d=1}^{D_c} \sum_{v=1}^V y_{s,d} a_{v,s} \geq 1 \times \sum_{v=1}^V a_{v,s} \forall s, \quad (۱۶.۳)$$

$$\sum_{s=1}^S y_{s,d} \bar{\Omega}_{j,s}^{tot} \leq \tau_{jd} \forall d, \forall j \in \mathcal{E}; \quad (۱۶.۳)$$

در اینجا، $P = [p_{u(v,k)}] \forall v, \forall k$ ماتریس توان برای کاربران است، $A = [a_{v,s}] \forall v, \forall s$ نشان دهنده ی متغیر باینری برای اتصال برشها به سرویسها و $Y = [y_{s,d}] \forall s, \forall d$ متغیر باینری برای اتصال مرکز داده ی فیزیکی به ها VNF می باشد. همچنین φ متغیر وزنی برای ارزش گزاری بین ترم اول و دوم تابع هدف است. (۱۶.۳ب) و (۱۶.۳ج) به ترتیب نشان دهنده ی این است که توان واحد رادیویی از حدی بیشتر نشده و توان هر کاربر یک عدد صحیح مثبت است. علاوه براین (۱۶.۳د) منجر به افزایش نرخ هرکاربر از حد مورد نیاز می باشد. (۱۶.۳ه) و (۱۶.۳و) به ترتیب محدودیت ظرفیت لینک fronthaul و محدودیت مورد نیاز برای تاخیر سیگنال ارسالی را نشان می دهد. (۱۶.۳ز) تضمین می کند که هر سرویس به حداقل یک برش شبکه متصل شده است. همچنین، (۱۶.۳ح) تضمین می کند که هر برش که شامل VNF هایی در دو لایه است، به حداقل یک مرکز داده متصل می شود. در (۱۶.۳ط) $\mathcal{E} = \{M, S, C\}$ و شرطها ضمانت می کند که منابع فیزیکی کافی برای ها VNF وجود دارد.

مسئله ی بهینه سازی (۱۶.۳) قابل جداسازی به دو مسئله ی بهینه سازی کوچکتر می باشد زیرا متغیرها می توانند به صورت مجزا بدست آیند. ابتدا نیاز به حل مسئله ی اول داریم و پس از حل آن P و A بدست می آید. سپس با حل مسئله ی دوم و قرار دادن مقدار بهینه ی A در مسئله خروجی Y بدست می آید. مسئله ی اول به صورت زیر می باشد.

$$\max_{P,A} \quad \eta(P, A) \quad (۱۷.۳)$$

$$\text{subject to} \quad \bar{p}_{r(s,i)} \leq P_{max} \quad \forall s, \forall i, \quad (۱۷.۳ب)$$

$$p_{u(v,k)} \geq 0 \quad \forall v, \forall k, \quad (۱۷.۳ج)$$

$$\mathcal{R}_{u(v,k)} \geq \mathcal{R}_{u(v,k)}^{min} \quad \forall v, \forall k, \quad (۱۷.۳د)$$

$$C_{r(s,i)} \leq C_{r(s,i)}^{max} \quad \forall s, \forall i, \quad (۱۷.۳ه)$$

$$D_s \leq D_s^{max} \quad \forall s, \quad (۱۷.۳)$$

$$\sum_{s=1}^S a_{v,s} \geq 1 \quad \forall s. \quad (ز۱۷.۳)$$

مسئله‌ی دوم نیز بدین صورت بیان می‌شود.

$$\min_y \quad \psi_{tot}(\mathbf{Y}) \quad (۱۸.۳)$$

$$\text{subject to} \quad \sum_{d=1}^{D_c} \sum_{v=1}^V y_{s,d} a_{v,s} \geq 1 \times \sum_{v=1}^V a_{v,s} \forall s, \quad (ب۱۸.۳)$$

$$\sum_{s=1}^S y_{s,d} \bar{\Omega}_{\delta,s}^{tot} \leq \tau_{\delta d} \forall d, \forall \delta \in \mathcal{E}; \quad (ج۱۸.۳)$$

۳.۳ روش ابتکاری استفاده شده

در این بخش از روش ابتکاری برای حل مسئله (۱۷.۳) استفاده می‌کنیم که یک مسئله‌ی غیرمحدب همراه با محاسبات طولانی است. برای حل مسئله همانطور که قبلاً بیان شد، مسئله به دو بخش مختلف و مجزا تجزیه می‌شود که به صورت تکراری حل می‌گردد. در بخش اول در ابتدای کار در اولین حلقه با قرار دادن $P = P_{max}$ و $\eta = 0$ در رابطه‌ی (۱۷.۳) مقدار A بدست می‌آید. سپس با A بدست آمده، P را در معادله‌ی (۱۷.۳) بروزرسانی می‌کنیم. این معادله قابل تبدیل به یک مسئله‌ی محدب است و با روشهای محدب حل می‌شود. بعد از حل این مسئله، P و η بروزرسانی شده و این چرخه‌ی دو بخشی تا همگرا شدن ادامه دارد.

۱.۳.۳ بخش اول مسئله‌ی اول

الگوریتم تکرار شده برای بدست آوردن A در اینجا اجرا می‌گردد. بخش اول مسئله‌ی اول از جنس بسته‌بندی جعبه به صورت سه بعدی می‌باشد که یک مسئله‌ی NP-Hard است و برای حل آن نیاز به الگوریتمهای ابتکاری داریم. در [۵۴]، الگوریتم ابتکاری براساس مرتب‌سازی جعبه‌ها براساس اندازه‌گیری بیان شده است. جزئیات الگوریتم موردنظر در ۱.۳ و براساس الگوریتم [۵۴] و روش BFD^۶ بیان شده است.

^۶Best Fit Decreasing

الگوریتم ۱۰.۳ اتصال سرویس به برش شبکه

- 1: Sort services according to the number of UE's requirements in the descending order.
 - 2: Sort slices according to the weighted linear combination of number of PRBs, RUs and VNFs and the capacity of their resources in the descending order.
 - 3: **for** $i \leftarrow 1$ to S **do**
 - 4: **for** $j \leftarrow 1$ to V **do**
 - 5: Set $a_{i,j} = 1$
 - 6: Obtain Parameters of Systems (power and rate , capacity of fronthaul links)
 - 7: **if** conditions (3.16), (3.16), (3.16) and (3.16) are not applied **then**
 - 8: Set $a_{i,j} = 0$;
 - 9: **else**
 - 10: break from inner loop;
 - 11: **end if**
 - 12: **end for**
 - 13: **end for**
-

۲.۳.۳ بخش دوم مسئله‌ی اول

در این بخش با فرض ثابت قرار دادن A توان کاربران در هر سرویس بدست می‌آید.

قضیه ۱.۳.۳. بهره‌وری انرژی زمانی بدست می‌آید که

$$\max_{\mathbf{P}} (\mathfrak{R}_{tot}(\mathbf{P}) - \eta^* P_{rtot}(\mathbf{P})) = \mathfrak{R}_{tot}(\mathbf{P}^*) - \eta^* P_{rtot}(\mathbf{P}^*) = 0. \quad (۱۹.۳)$$

□

برهان. مراجعه به [Appendix A، ۵۵]

در بخش دوم مسئله‌ی اول از الگوریتم تکراری و تابع لاگرانژ برای حل استفاده شده است. از آنجایی که تداخل تابعی از توان کاربران است، برای سادگی فرض باند بالای $\bar{I}_{u(v,i)}$ برای آن می‌کنیم. برای اینکه بتوانیم (۱۷.۳) را به صورت فرم استاندارد بهینه سازی محدب درآوریم، نیازمند این هستیم که متغیرهای روابط (۱۷.۳) و (۱۷.۳) را $\gamma_{r(s,i)}^{C_{r(s,i)}} \times \sigma_{q_{r(s,i)}}^2 = p_{r(s,i)} + \alpha_s$ و $1/(D_s - d_{s_1} + d_{s_2}) + \alpha_s$ تغییر دهیم. فرض کنید μ, λ, ξ و κ ماتریسهای ضرایب لاگرانژ می‌باشد که دارای مقادیر مثبت غیرصفرند. تابع لاگرانژ بدین صورت نوشته می‌شود.

$$\mathcal{L}(\mathbf{P}; \lambda, \chi, \mu, \xi, \kappa) = \sum_{v=1}^V \sum_{k=1}^{U_v} \bar{\mathcal{R}}_{u(v,k)} - \eta \sum_{v=1}^V \sum_{i=1}^{R_s} \bar{p}_{r(s,i)} \quad (۱۲۰.۳)$$

$$+ \sum_{v=1}^V \sum_{k=1}^{U_v} \lambda_{u(v,k)} (\bar{\mathcal{R}}_{u(v,k)} - \mathcal{R}_{u(v,k)}^{max}) \quad (۲۰.۳ ب)$$

$$- \sum_{s=1}^S \sum_{i=1}^{R_s} \mu_{r(s,i)} (\bar{p}_{r(s,i)} - P_{max}) \quad (۲۰.۳ ج)$$

$$- \sum_{s=1}^S \sum_{i=1}^{R_s} \xi_{r(s,i)} (\bar{p}_{r(s,i)} - \sigma_{q_{r(s,i)}}^2 \gamma_{r(s,i)}^{C_{r(s,i)}}). \quad (۲۰.۳ د)$$

$$+ \sum_{v=1}^V \sum_{k=1}^{U_v} \kappa_{u(v,k)} \sum_{s=1}^S (R_{u(v,k)} - \mathfrak{D}_s) a_{v,s}. \quad (۲۰.۳ ه)$$

که در اینجا $\mathfrak{D}_s = \frac{1}{D_s^{max} - d_{s_1} - d_{s_2}} + \alpha_s$ توان بهینه از (۲۰.۳) بدین صورت خواهد بود

$$p_{u(v,i)}^* = \left[\frac{\eta_{u(v,i)} \mathbf{w}_{u(v,i)} - \mathbf{x}_{u(v,i)} \mathbf{z}_{u(v,i)}}{\mathbf{x}_{u(v,i)} \mathbf{w}_{u(v,i)}} \right]^+ \quad (21.3)$$

که در اینجا $\eta_{u(v,i)} = (\lambda_{u(v,i)} + \kappa_{u(v,k)} + 1) \frac{B}{Ln}$ و $\mathbf{w}_{u(v,i)} = \sum_{s=1}^S |\mathbf{h}_{R_s, u(v,i)}^H \mathbf{w}_{R_s, u(v,i)}|^2 a_{v,s}$ همچنین $\mathbf{z}_{u(v,i)} = BN_s + \bar{I}_{u(v,i)}$ و $\mathbf{x}_{u(v,i)} = \sum_{s=1}^S \sum_{i=1}^{R_s} (\mu_{r_{u(s,i)}} + \xi_{r_{(s,i)}} + \eta) \|w_{r_{(s,i)}, u(v,i)}\|^2$ با استفاده از روش sub-gradient توان بهینه P بدست می آید [۵۱].

۳.۳.۳ حل دو بخش مسئله ی اول به صورت تکراری

در (۱.۳.۳) و (۲.۳.۳) جزئیات حل هر بخش از مسئله ی اول آورده شده است. در ابتدا با ثابت نگه داشتن $P = P_{max}$ و $\eta = 0$ مقدار A در مسئله ی (۱۷.۳) و با استفاده از (۱.۳) بدست می آید. سپس A از الگوریتم ۱.۳ بدست می آید. پس از آن در بخش دوم، با استفاده از روش sub-gradient P بدست می آید. بعد از بدست آوردن P و η الگوریتم با مقادیر جدید توان و بهره وری انرژی دوباره تکرار می گردد تا همگرا شود. در اینجا الگوریتم حل مسئله ی اول به صورت الگوریتم (۲.۳) است.

۴.۳.۳ مسئله ی دوم

در این بخش می خواهیم (۲.۴) را حل نماییم که قرارگیری منابع مجازی در منابع فیزیکی می باشد به طوری که تابع هزینه ψ_{tot} کمینه شود. این مسئله از جنس مسئله ی سه بعدی بسته بندی جعبه است که NP-Hard می باشد. برای بدست آوردن مقدار بهینه ی Y الگوریتم ابتکاری برای حل مسئله ی سه بعدی بسته بندی جعبه ارائه شده است. جزئیات الگوریتم در (۳.۳) نوشته شده است. در این الگوریتم، در ابتدا برشهای شبکه و مرکز داده ها را بر اساس جمع وزنی آنها مرتب می کنیم (خط ۱ و ۲ الگوریتم ۳.۳). در اینجا، پارامتر وزندار برای $\Omega_{3,s}^{tot}$ و τ_j^3 بدین صورت تعریف می کنیم.

$$\hat{\Omega}_s^{tot} = w_M \bar{\Omega}_{M,s}^{tot} + w_S \bar{\Omega}_{S,s}^{tot} + w_C \bar{\Omega}_{C,s}^{tot} \quad (22.3)$$

$$\hat{\tau}_j = w_M \tau_j^M + w_S \tau_j^S + w_C \tau_j^C,$$

الگوریتم ۲.۳ برش شبکه و تخصیص منابع

- 1: Set the maximum number of iterations I_{max} , convergence condition ϵ_η and the initial value $\eta^{(1)} = 0$
 - 2: Set $\mathbf{P} = \mathbf{P}_{max}$
 - 3: **for** $counter \leftarrow 1$ to I_{max} **do**
 - 4: Achieve \mathbf{A} by applying Algorithm (3.1)
 - 5: Obtain \mathbf{P} by using sub-gradient method which is mentioned in (3.3.2).
 - 6: **if** $\Re_{tot}(\mathbf{P}^{(i)}, \mathbf{A}^{(i)}) - \eta^{(i)} P_{tot}(\mathbf{P}^{(i)}, \mathbf{A}^{(i)}) < \epsilon_\eta$ **then**
 - 7: Set $\mathbf{P}^* = \mathbf{P}^{(i)}$, $\mathbf{A}^* = \mathbf{A}^{(i)}$ and $\eta^* = \eta^{(i)}$;
 - 8: **break**;
 - 9: **else**
 - 10: $i = i + 1$, Setting $\mathbf{P} = \mathbf{P}^{(i)}$;
 - 11: **end if**
 - 12: **end for**
-

جدول ۱.۳: پارامترهای شبیه سازی

Parameter	Value
Noise power	-174dBm
Bandwidth	120 KHZ
Maximum transmit Power of each RU	40dBm
Minimum delay	300usec
Maximum fronthaul capacity	200 bits/sec/Hz
Minimum data rate	10 bits/sec/Hz

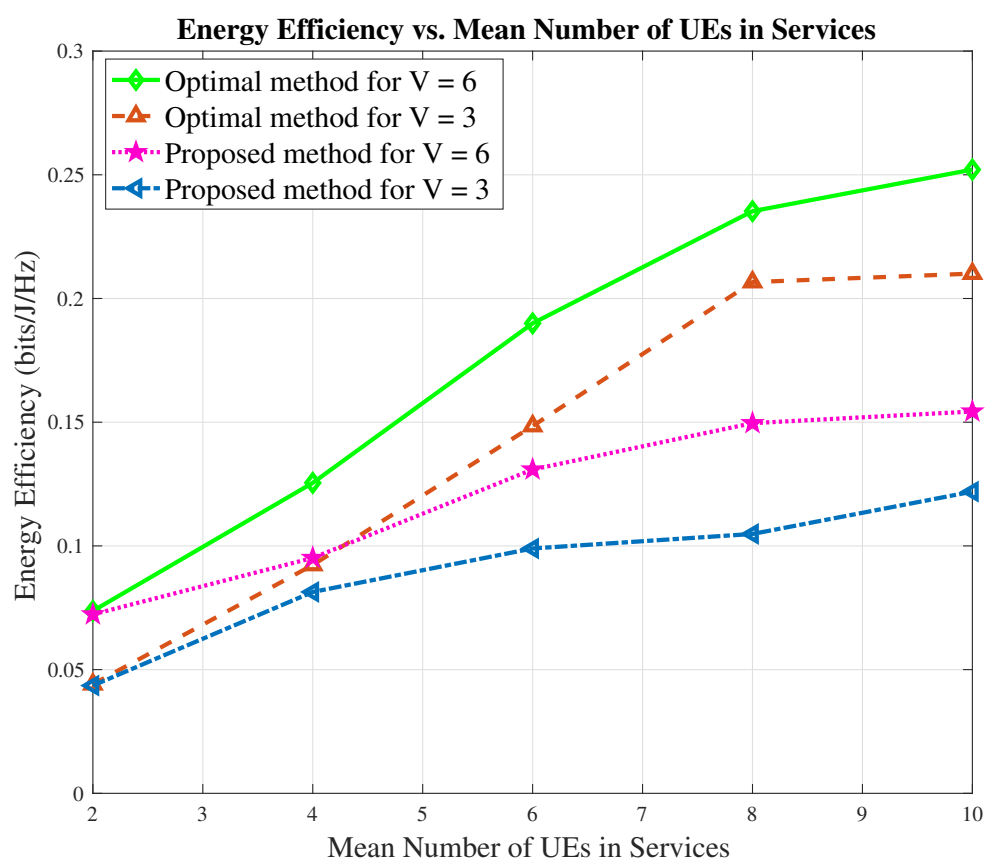
که در اینجا، $w = \{w_M, w_S, w_C\}$ وزن حافظه، نگهداری و CPU می باشد [۵۴]. در مرحله دوم، ما تخصیص را از بیشترین برش های مورد نیاز به DC با بیشترین منابع فیزیکی شروع می کنیم (از خط ۴ تا ۱۱ الگوریتم ۳.۳). بعد از تخصیص برش شبکه به مرکز داده، در صورتی که یکی از برشهای شبکه توسط مرکز داده پذیرش نشود، برای بخش بعدی جاگذاری می رود (از خط ۱۲ تا ۲۵ الگوریتم ۳.۳). در آخر، اگر مرکز داده ای با میزان منابع کمتر آزاد باشد و بتواند برشهای یک مرکز داده ی دیگر با میزان منابع بیشتر را سرویس دهی کند، تخصیص منابع بین این دو مرکز داده جابه جا می شود (خط ۲۶ از الگوریتم ۳.۳).

۴.۳ نتایج عددی

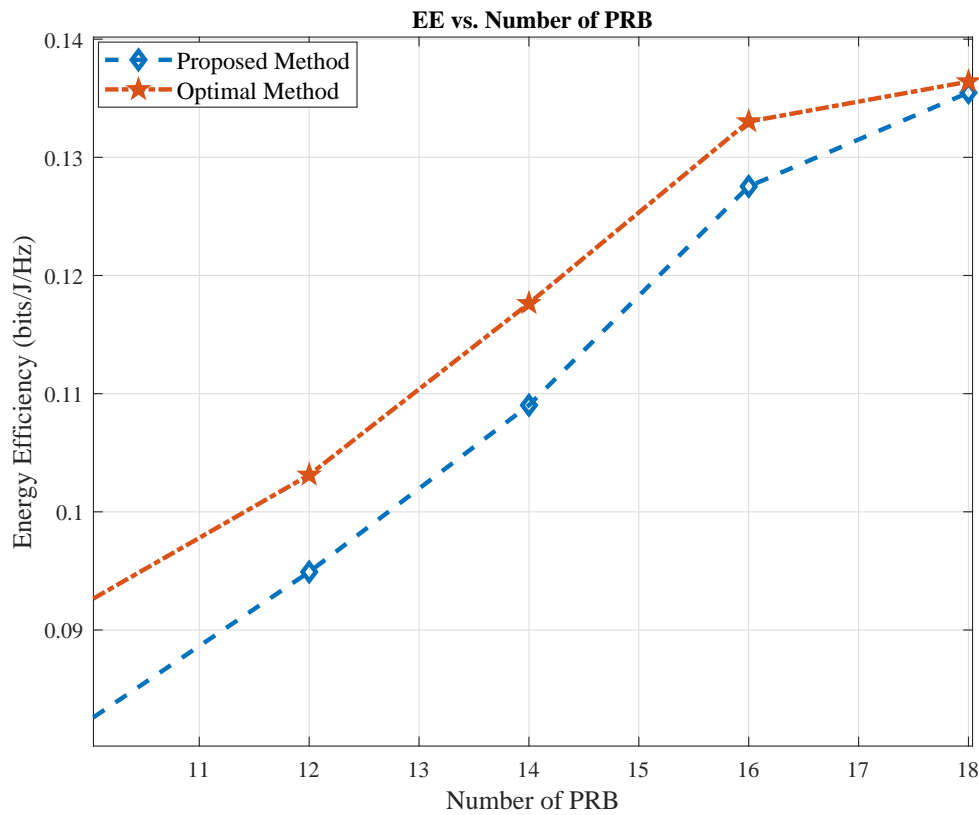
در این بخش نتایج عددی برای مسئله ی اصلی بیان می شود. ابتدا برای مسئله ی اول نمودارهای شبیه سازی را رسم می کنیم. در شکل ۲.۳، بهره وری انرژی برای دو سری تعداد سرویس مختلف با تعداد کاربر متفاوت در هر سرویس بدست آمده است. پارامترهای شبیه سازی در جدول ۱.۳ قرار داده شده است. تعداد منابع فیزیکی ۱۰ تا می باشد. روش بهینه برای بدست آوردن A توسط MOSEK بدست آمده و برای بدست آوردن P از CVX در متلب استفاده شده است و به صورت متناوب این دو ماتریس به روزرسانی می شوند. روش بهینه، $0.9 \text{ bits}/J/Hz$ از روش آورده شده بهتر است (برای $V = 6$ و $E[U_v] = 10$). همچنین $0.9 \text{ bits}/J/Hz$ از روش آورده شده بهتر است (برای $V = 3$ و $E[U_v] = 10$). همانطور که در شکل مشخص است بهره وری انرژی با افزایش تعداد کاربران افزایش می یابد. در شکل ۳.۳ بهره وری انرژی براساس تعداد بلوکهای منابع فیزیکی

الگوریتم ۳.۳ قرار گیری منابع مجازی در منابع فیزیکی

- 1: Sort Slices according to $\hat{\Omega}_s^{tot}, \forall s$ in descending order.
- 2: Sort DCs according to $\hat{\tau}_j, \forall j$ in descending order.
- 3: $\mathbf{Y} = \mathbf{0}$
- 4: **for** $d \leftarrow 1$ to D_c **do**
- 5: **for** $s \leftarrow 1$ to S **do**
- 6: **if** $\sum_{d=1}^{D_c} y_{s,d} == 0$ and $\bar{\Omega}_{\mathfrak{z}(s)}^{tot} \leq \tau_{\mathfrak{z}} \forall \mathfrak{z}, \forall s$ **then**
- 7: Set $y_{s,d} = 1$;
- 8: $\tau_j^{\mathfrak{z}} \leftarrow \tau_j^{\mathfrak{z}} - \bar{\Omega}_{\mathfrak{z},s}^{tot} \mathfrak{z} \in \{M, S, C\}$
- 9: **end if**
- 10: **end for**
- 11: **end for**
- 12: $ind_{rem} = \{s | (\sum_{d=1}^{D_c} y_{s,d} == 0)\}$
- 13: Sort remaining amount of DCs same as before in descending order.
- 14: Sort remaining slices same as before in descending order.
- 15: **for** $r \leftarrow 1$ to S_{rem} **do**
- 16: **for** $n \leftarrow 1$ to D_c **do**
- 17: Set $y_{s,d} = 1$;
- 18: $\bar{\Omega}_{\mathfrak{z},s}^{tot} \leftarrow \bar{\Omega}_{\mathfrak{z},s}^{tot} - \tau_j^{\mathfrak{z}}$
- 19: **if** $\bar{\Omega}_s^{tot} == 0$ **then**
- 20: Set $y_{s,d} = 1$;
- 21: $\tau_j^{\mathfrak{z}} \leftarrow \tau_j^{\mathfrak{z}} - \bar{\Omega}_{\mathfrak{z},s}^{tot} \mathfrak{z} \in \{M, S, C\}$
- 22: break inner loop
- 23: **end if**
- 24: **end for**
- 25: **end for**
- 26: Remapping DCs must be done to prevent wasting Energy



شکل ۲.۳: بهره‌وری انرژی براساس تعداد کاربران در هر سرویس

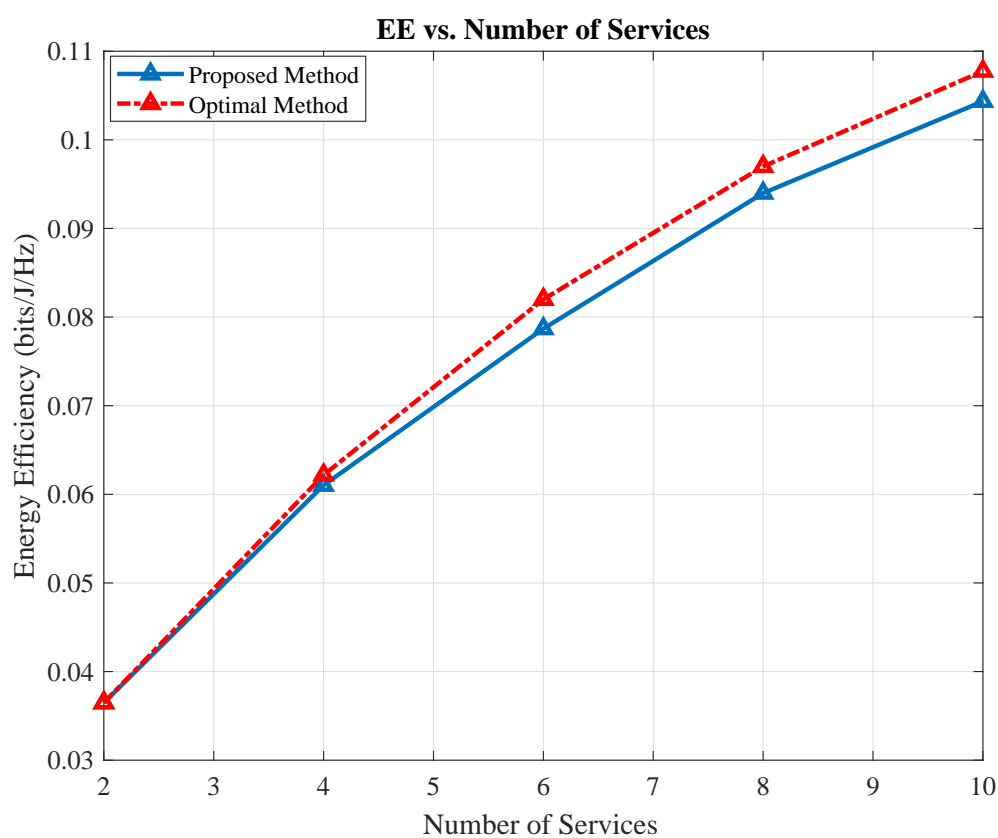


شکل ۳.۳: بهره‌وری انرژی براساس تعداد بلوکهای منابع فیزیکی

برای ۶ سرویس با تعداد کاربر به طور میانگین ۳ تا رسم شده است. همانطور که در نمودارها مشخص است با افزایش تعداد، PRB تداخل کاربران کم شده و بهره‌وری انرژی افزایش می‌یابد و الگوریتم پیشنهادی به روش بهینه نزدیک می‌گردد.

در شکل ۴.۳ بهره‌وری انرژی براساس تعداد سرویسهای مختلف با فرض وجود ۲ کاربر در هر سرویس به طور میانگین، رسم شده است. همانطور که در نمودارها مشخص است با افزایش تعداد سرویسها، بهره‌وری انرژی که مجموع نرخهای کاربران بر روی آن است، افزایش می‌یابد و الگوریتم پیشنهادی به روش بهینه به نسبت نزدیک است.

حال به نتایج عددی مسئله دوم می‌پردازیم. در شکل ۵.۳، نسبت برشهای پذیرفته شده برای دو تعداد مختلف مرکز داده با تعداد برشهای مختلف رسم شده است. پارامترهای شبیه‌سازی در جدول ۲.۳ آورده شده است. همچنین داریم $w_C = 320$ ، $w_S = 100$ و $w_M = 1$. در این شبیه‌سازی ترم دوم رابطه‌ی (۴.۴)، مورد اهمیت قرار گرفته شده است و پارامتر طراحی ν مقدار بالایی دارد. همچنین فرض شده که تنها یک مرکز



شکل ۴.۳: بهره‌وری انرژی براساس تعداد سرویسهای مختلف

جدول ۲.۳: پارامترهای شبیه‌سازی

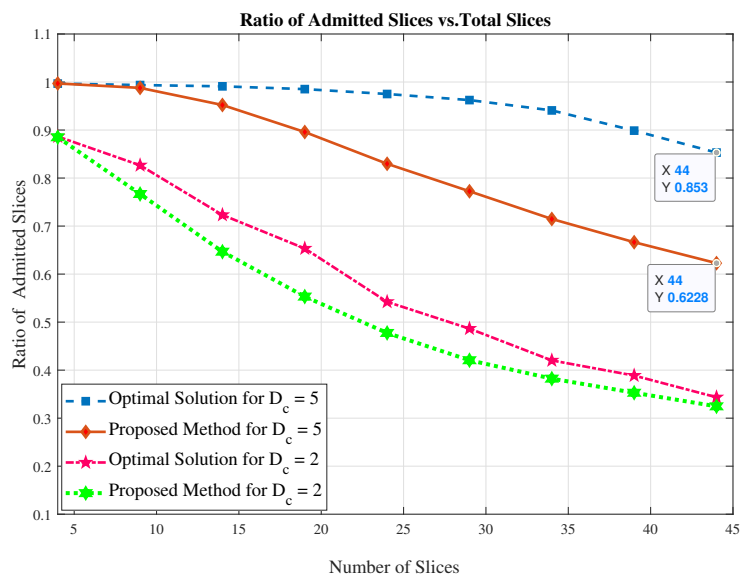
Parameter	Value
Mean of CPU for DCs	320GHz
Mean of Memory for DCs	1T
Mean of Storage for DCs	100T
Mean of CPU for Slices	32GHz
Mean of Memory for Slices	100G
Mean of Storage for Slices	10T

داده می‌تواند هر برش را سرویس دهد. روش ارائه شده در الگوریتم ۳.۳ آورده شده و روش بهینه با استفاده از MOSEK بدست می‌آید. وقتی دو مرکز داده یا DC داریم، روش پیشنهادی و روش بهینه تقریباً همان نسبت برش‌های پذیرفته شده را دارند. اما با افزایش تعداد DC ها به پنج، عملکرد روش پیشنهادی کاهش می‌یابد. با استفاده از پنج DC، تفاوت بین روش پیشنهادی و روش بهینه در بدترین حالت (۴۴ برش) حدود ۲۳ درصد است.

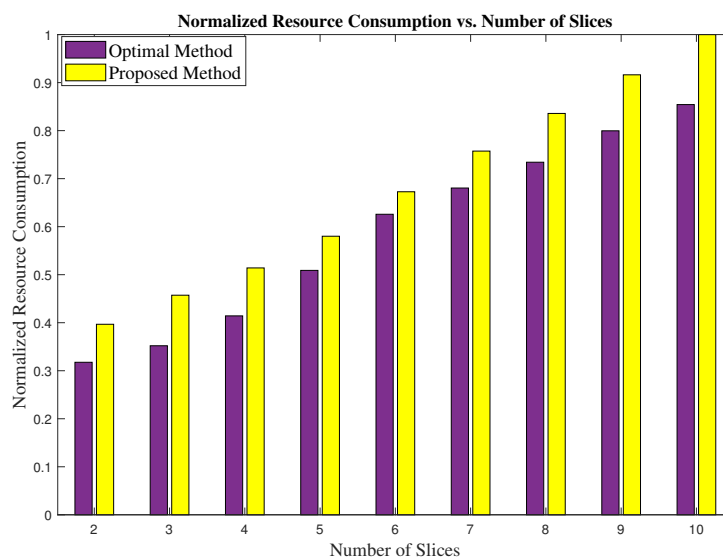
در شکل ۶.۳ نرمالیزه‌ی مصرف منابع براساس تعداد برشهای شبکه آورده شده است. پارامترهای شبیه‌سازی در جدول ۲.۳، قرار داده شده و ($w_M = 1, w_S = 100, w_C = 320$). در این شبیه‌سازی، فرض کنید تعداد DC ها کاملاً کافی باشد تا تمام برش‌ها را پوشش دهد و پارامتر طراحی شده nu کم فرض شود، بنابراین ما روی اولین ترم (۴.۴) تمرکز کردیم. با این حال، اگر هر قطعه باقی مانده باشد، می‌تواند توسط بیش از یک DC سرویس دهی شود. بهینه بودن محل قرارگیری برشها در منابع DC بر اساس مصرف توان DC ها اندازه‌گیری می‌شود. در اینجا مقدار منابع مصرف DC نشده نشان داده شده است. برای ۱۰ برش شبکه اختلاف بین مقدار بهینه و روش اعمال شده ۱۵ درصد می‌باشد.

۵.۳ نتیجه‌گیری

در این فصل، همزمانی برش شبکه و تخصیص توان در سیستم ORAN آمده است. فرض بر این است که کاربران بر اساس نیازهایشان به سرویسهای مختلف طبقه‌بندی می‌شوند. همچنین، تعدادی برش نیز در خدمت این سرویسها قرار می‌گیرد. هر برش شبکه شامل تعدادی، PRB، RU و هایی VNF است که اعمال CU و DU را



شکل ۵.۳: نسبت برش های پذیرفته شده فقط به یک DC در مقابل برش های کل



شکل ۶.۳: نرمالیزه ی مصرف منابع براساس تعداد برشها

انجام می‌دهد. ظرفیت محدود لینک fronthaul برای اتصالات فیبر بین RU و DU در نظر گرفته شده است. هدف به حداکثر رساندن مجموع نرخ و به حداقل رساندن مصرف توان و هزینه انرژی مراکز داده به طور همزمان است. این مسئله به دوزیر مسئله تجزیه می‌شود. هر زیر مسئله به طور جداگانه توسط یک الگوریتم ابتکاری حل می‌شود. با استفاده از نتایج عددی، روش اکتشافی را تأیید کرده و عملکرد الگوریتم‌ها را مطالعه می‌کنیم. برای مسئله‌ی اول بازدهی انرژی در مقابل تعداد UE در هر سرویس به تصویر کشیده شده است. با افزایش میانگین تعداد کاربران در هر سرویس، از بازدهی انرژی بیشتر می‌شود. برای مسئله دو، دو شکل نشان داده شده است. در شکل اول، نسبت برش‌های پذیرفته شده که فقط به یک DC متصل می‌شوند، برای تعداد برش‌های مختلف مشخص شده است. در شکل دوم، مصرف منابع نرمال DC‌ها به تصویر کشیده شده است. در هر شکل، الگوریتم ابتکاری با روش بهینه مقایسه شده و تفاوت بین آنها بحث شده است.

فصل ۴

تخصیص برش شبکه به صورت دینامیکی

۱.۴ مقدمه

در این فصل هدف تخصیص برش شبکه به صورت می باشد. در فصل قبلی مدل سیستم به طور کامل نوشته شده است و در حالت آفلاین حل گردیده است، در این فصل پارامترها مورد نیاز را نسبت به فصل قبلی کمتر کرده و با استفاده از روش دینامیکی در هر لحظه از زمان به حل سیستم می پردازیم. برای حل این سیستم از روش یادگیری تقویتی عمیق استفاده می کنیم. در بخش اول صورت مسئله ی بخش رادیویی نوشته می شود. سپس به صورت مسئله و مدل سیستم بخش هسته می پردازیم و در نهایت روش حل هر دو مسئله و نتایج عددی آن بیان می شود.

۲.۴ مدل سیستم و صورت مسئله ی اول

در این بخش هدف برش شبکه در بخش رادیویی سیستم می باشد. در اینجا، مسئله ی اول فصل قبلی ساده شده و به روش دینامیکی حل می شود. همانند سیستم فصل قبل، فرض می کنیم S برش شبکه داریم که قرار است V سرویس مختلف که شامل کاربرانی است که از سرویس خاص استفاده می نمایند را سرویس دهی نماید. هر سرویس $v \in \{1, 2, \dots, V\}$ شامل تعدادی کاربر تک آنتنه می باشند که سرویس خاصی را درخواست می نماید. هر برش شبکه $s \in \{1, 2, \dots, S\}$ شامل تعدادی RU, PRB, BBU و VNF می باشد. در این بخش سعی بر این است که در ابتدا مسئله را به ساده ترین حالت ممکن حل نماییم. فرض می کنیم سه مدل سرویس مختلف داریم که سرویسهای دسته ی اول نیازمند تاخیر خاص و سرویسهای دسته ی دوم نیازمند داشتن تاخیر کم هستند و سرویس

سوم نیازمند داشتن هر دو حالت تاخیر کم و نرخ زیاد است. در بخش اول این مسئله، هدف بیشینه‌سازی تعدادی سرویسهای پذیرفته شده می‌باشد. در اینجا فرض براین است که تعداد برشهای شبکه محدود می‌باشد. فرض می‌کنیم هر سرویس v دارای اولویت p_v می‌باشد. همچنین فرض براین است که هر سرویس شامل ماکسیمم U_v کاربر است و به طور میانگین کاربران آن نیازمند داشتن نرخ بیشتر از R_v و تاخیر کمتر از D_v هستند. در صورتی که کاربری از دسته‌ی اول باشد $D_v = M$ که برای تاخیر یک عدد بزرگ می‌باشد. و در صورتی که سرویس از دسته‌ی دوم باشد $R_v = N$ که یک عدد کوچک برای نرخ می‌باشد. صورت مسئله به صورت (۱.۴ج) می‌باشد. در اینجا برای سادگی فرض براین است که هر سرویس به ماکسیمم یک برش شبکه متصل می‌گردد. در اینجا هدف حل مسئله در هر اسلات زمانی t می‌باشد. هدف در اینجا بیشینه‌سازی تعداد سرویسهای پذیرفته شده توسط برشهای شبکه می‌باشد به صورتی که شرط تاخیر و نرخ سرویس را ضمانت کنند.

$$\max_{a(t)} \sum_{s=1}^{S(t)} \sum_{v=1}^{V(t)} p_v a_{v,s}(t) \quad (1.4\text{آ})$$

$$\text{subject to} \quad \sum_{s=1}^{S(t)} D_s(t) a_{v,s} \leq D_v(t) \forall v, \quad (1.4\text{ب})$$

$$\sum_{s=1}^{S(t)} R_s(t) a_{v,s}(t) \geq R_v(t) \forall v \quad (1.4\text{ج})$$

برای اینکه معادله‌ی (۱.۴ج) را به فرم مسئله‌ی کوله‌پشتی دریاوریم، از آنجایی که فرض کردیم هر سرویس به ماکسیمم یک برش شبکه متصل می‌شود، می‌توان معادله بدین صورت نوشت:

$$\max_{a(t)} \sum_{s=1}^{S(t)} \sum_{v=1}^{V(t)} p_v a_{v,s}(t) \quad (1.4\text{آ}')$$

$$\text{subject to} \quad \sum_{s=1}^{S(t)} D_s(t) a_{v,s} \leq D_v(t) \forall v, \quad (1.4\text{ب}')$$

$$\sum_{s=1}^{S(t)} \frac{1}{R_s(t)} a_{v,s}(t) \leq \frac{1}{R_v(t)} \forall v \quad (1.4\text{ج}')$$

که در اینجا، معادله‌ی (۱.۴ج) یک معادله‌ی کوله‌پشتی دو بعدی می‌باشد. برای حل این مسئله، از روش یادگیری تقویتی عمیق استفاده می‌شود.

۳.۴ مدل سیستم و صورت مسئله‌ی دوم

در این بخش سعی شده، مسئله‌ی دوم فصل قبل به صورت ساده شده با روش دینامیکی حل شود. عنوان این مسئله، جاگذاری VNFها بر روی مراکز داده می‌باشد. فرض بر این است که S برش شبکه داریم که هر برش شبکه $s \in \{1, 2, \dots, S\}$ می‌باشد. هر برش شبکه شامل تعدادی VNF است که هر VNF نیازمند منابع فیزیکی است که شامل حافظه، نگهدارنده و پردازشگر می‌باشد. فرض کنید برای سادگی مسئله برای هر VNF به مقدار کافی حافظه و نگهدارنده در مراکز داده داریم و تنها منبع مورد نیاز برای f امین VNF در برش s ام مقدار پردازنده است که به صورت $\bar{\Omega}_s^f$ می‌باشد که در اینجا $\bar{\Omega}_s^f \in \mathbb{C}^1$ و

$$\bar{\Omega}_s^{tot} = \sum_{f=1}^{M_s} \bar{\Omega}_s^f \quad (۳.۴)$$

همچنین D_c مرکز داده برای سرویس دهی به VNFها می‌باشد. هر مرکز داده شامل تعدادی سرور برای سرویس دهی است. M_s تعداد کل VNFها در امین s برش شبکه است. مقدار پردازشگر τ_j برای j امین مرکز داده می‌باشد. در این مدل سیستم، تخصیص منابع فیزیکی به VNFها در نظر گرفته شده است. در اینجا فرض بر این است که $y_{s,d}$ متغیر صفر و یکی است که نشان می‌دهد مرکز داده d ام به s امین برش متصل است یا نه. همانند فصل قبل، فرض کنید توان مصرفی پردازش باند پایه در هر مرکز داده d که به VNFهای یک برش s متصل می‌باشد در هر زمان t با $\phi_{s,d}(t)$ نشان داده شده است. بنابراین می‌توان توان کل سیستم را برای کلیه مرکز داده‌های فعال که به برش متصل هستند، بدین صورت نشان داد

$$\phi_{tot}(t) = \sum_{s=1}^S \sum_{d=1}^{D_c} y_{s,d} \phi_{s,d}(t).$$

همچنین، یک تابع هزینه برای قرار همچنین فرض کنید در هر زمان قرار دادن هر مجموعه‌ی جدید VNF برش شبکه s بر روی مرکز داده d مقدار انرژی اضافی را بدین صورت به سیستم اعمال کنند.

$$\phi_{diff}(t) = \sum_{s=1}^S \sum_{d=1}^{D_c} [y_{s,d}(t) - y_{s,d}(t-1)]^+ \phi_{s,d}^{new}(t).$$

تابع هزینه‌ی قرارگیری VNFها بر روی DCها بدین صورت است

$$\psi_{tot}(t) = \phi_{tot}(t) + \phi_{diff}(t) \quad (۴.۴)$$

در اینجا هدف کمینه کردن مقدار انرژی کل در هر زمان، با فرض اینکه مجموع های VNF برشهای تخصیص یافته به هر مرکز داده مقدار کافی پردازنده داشته باشند. در اینجا فرض براین است که برشهای شبکه که قبلاً به سرویسها اختصاص داده شده، می‌بایست در مرکز داده قرار داده شوند و تعداد مراکز داده به اندازه‌ی کافی زیاد هستند، هدف کمینه کردن انرژی است به صورتی که کمترین تعداد مراکز داده‌ها استفاده شوند.

$$\min_{\mathbf{y}} \quad \psi_{tot}(\mathbf{Y})(t) \quad (آ۵.۴)$$

$$\text{t. s.} \quad \sum_{d=1}^{D_c} \sum_{v=1}^V y_{s,d}(t) \geq 1 \forall s, \quad (ب۵.۴)$$

$$\sum_{s=1}^S y_{s,d}(t) \bar{\Omega}_s^{tot} \leq \tau_d \forall d, , \forall; \quad (ج۵.۴)$$

این مسئله به فرم مسئله‌ی بسته‌بندی جعبه قابل بیان است.

۴.۴ حل مسئله به‌روش یادگیری تقویتی

در اینجا از روش یادگیری تقویتی در حل دو مسئله‌ی بالا استفاده می‌نماییم. در روش یادگیری تقویتی، یک عامل سعی می‌کند رفتار بهینه را در یک محیط مشخص پیدا کند. این پروسه‌ی تعاملی، به صورت پروسه‌ی تصمیم‌گیری مارکوف مدل می‌شود که شامل (S, A, R, P, γ) می‌باشد. S نشان‌دهنده‌ی ماتریس فضای حالت و A بردار رفتار می‌باشد. R نیز پاداش عمل است. $P(\cdot|s, a)$ احتمال انتقال و $\gamma \in (0, 1]$ فاکتور تخفیف می‌باشد. سیاست $\Pi(\cdot|s)$ نگرانی از حالت به توزیع اعمال می‌باشد. تابع مقدار-حالت برای حالت s تحت سیاست $\Pi(\cdot|s)$ را با $V^\Pi(s)$ نشان می‌دهند که مقدار بازده مورد انتظار در حالت s تحت سیاست $\Pi(\cdot|s)$ می‌باشد. مقدار ارزش انجام عمل a در حالت s تحت سیاست $\Pi(\cdot|s)$ را با $Q^\Pi(s, a)$ نمایش می‌دهیم. که روابط زیر را براین اساس داریم.

$$V^\Pi(s) = \mathbb{E}_{\Pi, P} \left[\sum_{t=0}^{\infty} \gamma^t R_t | S_0 = s \right] \quad (۶.۴)$$

و تابع ارزش عمل بدین صورت است

$$Q^{\Pi}(s, a) = \mathbb{E}_{\Pi, P} \left[\sum_{t=0}^{\infty} \gamma^t R_t | S_0 = s, A_0 = a \right] \quad (۷.۴)$$

که در اینجا \mathbb{E} میانگین می‌گیرد. باتوجه به رابطه‌ی بلمن، داریم:

$$V^{\Pi}(s) = \mathbb{E}_{\Pi, P} [R + \gamma V^{\Pi}(s')] \quad (۸.۴)$$

و داریم:

$$Q^{\Pi}(s, a) = \mathbb{E}_{\Pi, P} [R + \gamma Q^{\Pi}(s', a')] \quad (۹.۴)$$

که در اینجا s' و a' قابل اتخاذ از $P(\cdot | s, a)$ و سیاست $\Pi(\cdot | s')$ هستند. هدف یادگیری تقویتی، بدست آوردن سیاست بهینه می‌باشد به صورتی که $Q^{\Pi}(s, a)$ بیشینه شود. مقدار بهینه‌ی تابع ارزش عمل در رابطه‌ی بلمن بدین صورت است

$$Q^*(s, a) = \mathbb{E}_{\Pi^*, P} [R + \gamma Q^*(s', a')] \quad (۱۰.۴)$$

همچنین در صورت تعریف اپراتور T^* بلمن داریم

$$T^*Q(s, a) = \mathbb{E}_{\Pi^*, P} [R + \gamma Q(s', a')] \quad (۱۱.۴)$$

که به صورت تکراری اعمال این اپراتور $T^*Q(s, a) \leftarrow Q_{t+1}(s, a)$ منجر به همگرا شدن $Q_t(s, a) \rightarrow Q^*(s, a)$ و $t \rightarrow \infty$ می‌شود [۴۷، ۵۶]. در ابعاد بالاتر، بهتر است از تابع تقریبی استفاده شود. فرض کنید $Q_{\theta}(s, a)$ تابع تقریب با پارامتر θ می‌باشد که تقریبی از جدول مقدار عمل با (s, a) می‌باشد. هدف این بهینه‌سازی، بدست آوردن θ است به طوری که $Q_{\theta}(s, a) \cong Q^*(s, a)$ و مقدار بهینه، به طور تکراری با اعمال اپراتور T^* بدست می‌آید. مقدار بهینه‌ی θ قابل دستیابی با استفاده از مینیمم کردن خطای تفاوت زمانی (TD)

error) بر روی (s, a, r, s') می‌باشد که به طور رندم انتخاب می‌شوند.

$$\zeta^2 = [r + \gamma \max_{a' \in A} Q_\theta(s', a') - Q_\theta(s, a)]^2 \quad (12.4)$$

روشهای مختلفی برای دستیابی به مینیمم خطا هست که ما در ادامه‌ی کار از روش Q-learning استفاده می‌کنیم. در روش Q-learning در هر بروزرسانی تابع Q داریم:

$$Q(s_{t+1}, a_{t+1}) = Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_{a \in A} Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (13.4)$$

که در اینجا، α نرخ یادگیری می‌باشد.

۱.۴.۴ نتایج عددی مسئله اول

در این بخش، برای مسئله‌ی اول حالت، عامل و پاداش را تعیین می‌نماییم و سپس نتایج عملی را نشان می‌دهیم. در این مسئله، عامل، orchestrator است که وظیفه‌ی مدیریت شبکه را برعهده دارد. همچنین، حالت در هر بازه‌ی زمانی برشهایی از شبکه است که به سرویسها متصل شده و اینکه کدام برش به کدام سرویس متصل است. پاداش طوری تعیین شده که بیشترین تعداد سرویسها پذیرفته شود به طوریکه شرط تاخیر و نرخ برآورده گردد.

۲.۴.۴ نتایج عددی مسئله دوم

در این بخش، برای مسئله‌ی دوم حالت، عامل و پاداش را تعیین می‌نماییم و سپس نتایج عملی را نشان می‌دهیم. در این مسئله، عامل، orchestrator است که وظیفه‌ی مدیریت شبکه را برعهده دارد. همچنین، حالت در هر بازه‌ی زمانی برشهایی از شبکه است که به مراکز داده متصل شده و اینکه کدام برش به کدام مرکز داده متصل است. پاداش طوری تعیین شده که کمترین تعداد مراکز داده استفاده گردد و هر لحظه کمترین تعداد مراکز داده‌ی خاموش، روشن شود.

۵.۴ نتیجه‌گیری

در این فصل، دو مسئله‌ی فصل قبلی به صورت ساده شده نوشته شد و مسئله‌ی اول در بخش رادیویی از جنس کوله‌پشتی و مسئله‌ی دوم در بخش هسته از نوع بسته‌بندی جعبه می‌باشد. این دو مسئله به صورت دینامیکی در هر لحظه از زمان حل شده‌اند. برای حل این دو مسئله از روش یادگیری تقویتی عمیق استفاده شده و حالتها و اعمال بیان برای یک عامل در این مسئله بیان کرده است. نتایج عملی آن نیز رسم گردید.

فصل ۵

پیشنهادهای و کارهای آتی

۱.۵ مقدمه

در فصل اول، مقدمه‌ای بر مفاهیم مورد استفاده را بیان کردیم و در مورد نسل پنجم مخابرات و مفاهیم آن صحبت نمودیم. سپس در فصل دوم مروری بر کارهای انجام شده کردیم و مقالات مرتبط با برش شبکه و شبکه‌های دسترسی باز و قرارگیری توابع مجازی شبکه را بیان نمودیم تا مروری بر چالشهای مطرح شده نسل پنجم مخابرات کرده و حل این چالشها را مورد بررسی قرار دادیم. در فصل سوم صورت مسئله‌ای در زمینه‌ی برش شبکه در شبکه‌های دسترسی باز، معرفی کرده و با روش ابتکاری، آن را حل نمودیم و نتایج را با مقدار بهینه مقایسه کردیم. در فصل چهارم، دو مسئله‌ی بیان شده در فصل سوم را به صورت کاملاً ساده با روش یادگیری عمیق تقویتی به صورت دینامیکی و در هر بازه‌ی زمان حل نمودیم. این دو مسئله، MDP^۱ بوده و قابل حل با این روش هستند. حال در این فصل در مورد مزایا و معایب کارهای انجام شده در فصل سوم و چهارم صحبت کرده و کارهای آتی و پیشنهادات را بیان می‌کنیم.

۲.۵ نتیجه‌گیری

در اینجا، مسئله‌ی برش شبکه در بخش رادیویی و قرارگیری توابع مجازی شبکه بر روی مراکز داده باهم مورد بررسی قرار گرفته شد. برای حل این مسئله، ابتدا مسئله به دو بخش مخلف شکسته شد که در بخش اول، تخصیص

¹Markov Decision Process

برش شبکه به کاربران سرویسها و تخصیص توان حل شده و پس از آن، برشهایی از شبکه که به سرویس اختصاص داده شده را به مراکز داده نگاشت می‌دهیم. در این مسئله، تاخیر و نرخ هر کاربر در سرویس مورد بررسی قرار گرفته شده و چالش تخصیص منابع که شامل برش بخش رادیویی به هر سرویس است و جاگیری توابع شبکه حل می‌شود. الگوریتم ارائه شده سرعت بسیار بیشتری از الگوریتم بهینه که با MOSEK و CVX بدست می‌آید، دارد. سپس مسئله به صورت ساده‌تر برای حالت دینامیکی با روش یادگیری تقویتی حل گردیده است.

۱.۲.۵ مزایای این چالش و حل آن

در مسئله‌ی بیان شده‌ی فصل سوم، مدل سیستم به صورت دقیق بیان شده و نرخ کاربر، ظرفیت لینک fronthaul و تاخیر به طور دقیق مورد بررسی قرار گرفته شده است. همچنین مسئله به واقعیت نزدیکی زیادی دارد. همچنین الگوریتم ابتکاری تعریف شده در فصل سوم برای حالتی که تداخل به نسبت کم باشد به حالت بهینه بسیار نزدیک است. در فصل چهارم همین مسئله با فرض اینکه سرویسها نیازمند تاخیر کم یا نرخ بالا هستند به صورت پارامتریک در هر لحظه از زمان حل می‌گردند. در بخش بعدی چالشهای قرارگیری توابع مجازی بر روی مراکز داده به طور دقیق بررسی شده و در فصل چهارم این مسئله به صورت دینامیکی در هر لحظه حل گردیده است. در حل مسئله در حالت دینامیکی سعی بر این است که مراکز داده کمترین انرژی را مصرف نموده و از هدر رفت انرژی بپرهیزیم.

۲.۲.۵ معایب پروژه انجام شده

در فصل سوم از الگوریتم ابتکاری در این کار استفاده شده است. زمانی که تعداد بلوکهای منابع فیزیکی به نسبت کاربران بسیار کم باشد و تداخل به شدت زیاد گردد، الگوریتم مسئله‌ی اول به خوبی قادر به پاسخ‌گویی نیست و از حالت بهینه فاصله می‌گردد. همچنین در فصل چهارم صورت مسئله بسیار ساده‌تر از واقعیت است و مسئله در حالت دینامیکی برای این حالت حل گردیده است. در مسئله‌ی دوم، زمانی که تعداد مراکز داده زیاد گردد فاصله‌ی حالت بهینه از الگوریتم ابتکاری زیاد شده است.

۳.۲.۵ نوآوری‌های این پروژه

در این پروژه، تخصیص توان و برش شبکه در شبکه‌های دسترسی باز مورد بررسی قرار گرفته است. ما مسئله‌ی اختصاص UE به خدمات، خدمات به برشها و منابع فیزیکی بی سیم و همچنین مرکز داده به برشها را به عنوان

یک مشکل بهینه سازی فرمول‌بندی کرده‌ایم. سپس با ارائه‌ی روشهای ابتکاری، به حل آنها پرداختیم. در نهایت مسئله‌ی ساده شده را در حالت دینامیکی و متغیر با زمان حل کردیم.

۳.۵ پیشنهادات

کتاب نامه

- [1] (2020) Standard of etsi. [Online]. Available: <https://www.etsi.org/technologies/5g?jjj=1590472780060>
- [2] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, “Cloud ran for mobile networks—a technology overview,” *IEEE Communications surveys & tutorials*, vol. 17, no. 1, pp. 405–426, 2015.
- [3] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, “Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies,” *IEEE Wireless Communications*, vol. 21, no. 6, pp. 126–135, 2014.
- [4] M. Peng, S. Yan, K. Zhang, and C. Wang, “Fog-computing-based radio access networks: issues and challenges,” *IEEE Network*, vol. 30, no. 4, pp. 46–53, 2016.
- [5] C. C. W. D. R. D. S. G. Chih-Lin I, Sachin Katti, “O-ran white paper: Towards an open and smart ran,” O-RAN Alliance, Tech. Rep., 2018.
- [6] J. G. Herrera and J. F. Botero, “Resource allocation in nfv: A comprehensive survey,” *IEEE Transactions on Network and Service Management*, vol. 13, no. 3, pp. 518–532, 2016.
- [7] Y. Li and M. Chen, “Software-defined network function virtualization: A survey,” *IEEE Access*, vol. 3, pp. 2542–2553, 2015.
- [8] X. Zhou, R. Li, T. Chen, and H. Zhang, “Network slicing as a service: enabling enterprises’ own software-defined cellular networks,” *IEEE Communications Magazine*, vol. 54, no. 7, pp. 146–153, 2016.
- [9] Y. L. Lee, J. Loo, T. C. Chuah, and L.-C. Wang, “Dynamic network slicing for multitenant heterogeneous cloud radio access networks,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2146–2161, 2018.

- [10] J. Tang, W. P. Tay, T. Q. Quek, and B. Liang, "System cost minimization in cloud ran with limited fronthaul capacity," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 3371–3384, 2017.
- [11] Compare different generation. [Online]. Available: <http://net-informations.com/q/diff/generations.html>
- [12] S.-H. Park, O. Simeone, and S. S. Shitz, "Joint optimization of cloud and edge processing for fog radio access networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7621–7632, 2016.
- [13] (2018) xran forum merges with c-ran alliance to form oran alliance. [Online]. Available: <https://www.businesswire.com/news/home/20180227005673/en/xRAN-Forum-Merges-C-RAN-Alliance-Form-ORAN>
- [14] Next generation ran architecture. [Online]. Available: <http://rod-stuhlmuller-nydh.squarespace.com/>
- [15] R. Mijumbi, J. Serrat, J.-L. Gorricho, N. Bouten, F. De Turck, and R. Boutaba, "Network function virtualization: State-of-the-art and research challenges," *IEEE Communications surveys & tutorials*, vol. 18, no. 1, pp. 236–262, 2015.
- [16] W. Xia, Y. Wen, C. H. Foh, D. Niyato, and H. Xie, "A survey on software-defined networking," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 27–51, 2014.
- [17] V.-G. Nguyen, A. Brunstrom, K.-J. Grinnemo, and J. Taheri, "Sdn/nfv-based mobile packet core network architectures: A survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1567–1602, 2017.
- [18] H. Zhang, N. Liu, X. Chu, K. Long, A.-H. Aghvami, and V. C. Leung, "Network slicing based 5g and future mobile networks: mobility, resource management, and challenges," *IEEE communications magazine*, vol. 55, no. 8, pp. 138–145, 2017.
- [19] M. Peng, Y. Li, Z. Zhao, and C. Wang, "System architecture and key technologies for 5g heterogeneous cloud radio access networks," *IEEE network*, vol. 29, no. 2, pp. 6–14, 2015.
- [20] L. M. Larsen, M. S. Berger, and H. L. Christiansen, "Fronthaul for cloud-ran enabling network slicing in 5g mobile networks," *Wireless Communications and Mobile Computing*, vol. 2018, 2018.
- [21] S. Costanzo, I. Fajjari, N. Aitsaadi, and R. Langar, "A network slicing prototype for a flexible cloud radio access network," in *2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 2018, pp. 1–4.

- [22] H. Xiang, S. Yan, and M. Peng, "A realization of fog-ran slicing via deep reinforcement learning," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2515–2527, 2020.
- [23] S. E. Elayoubi, S. B. Jemaa, Z. Altman, and A. Galindo-Serrano, "5g ran slicing for verticals: Enablers and challenges," *IEEE Communications Magazine*, vol. 57, no. 1, pp. 28–34, 2019.
- [24] S. D'Oro, F. Restuccia, and T. Melodia, "Toward operator-to-waveform 5g radio access network slicing," *IEEE Communications Magazine*, vol. 58, no. 4, pp. 18–23, 2020.
- [25] L. Feng, Y. Zi, W. Li, F. Zhou, P. Yu, and M. Kadoch, "Dynamic resource allocation with ran slicing and scheduling for urllc and embb hybrid services," *IEEE Access*, vol. 8, pp. 34 538–34 551, 2020.
- [26] V. Sciancalepore, L. Zanzi, X. Costa-Perez, and A. Capone, "Onets: online network slice broker from theory to practice," *arXiv preprint arXiv:1801.03484*, 2018.
- [27] W. Li, Y. Zi, L. Feng, F. Zhou, P. Yu, and X. Qiu, "Latency-optimal virtual network functions resource allocation for 5g backhaul transport network slicing," *Applied Sciences*, vol. 9, no. 4, p. 701, 2019.
- [28] D. Sattar and A. Matrawy, "Towards secure slicing: Using slice isolation to mitigate ddos attacks on 5g core network slices," in *2019 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 2019, pp. 82–90.
- [29] —, "Optimal slice allocation in 5g core networks," *IEEE Networking Letters*, vol. 1, no. 2, pp. 48–51, 2019.
- [30] P. Luong, C. Despins, F. Gagnon, and L.-N. Tran, "A novel energy-efficient resource allocation approach in limited fronthaul virtualized c-rans," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*. IEEE, 2018, pp. 1–6.
- [31] P. Luong, F. Gagnon, C. Despins, and L.-N. Tran, "Joint virtual computing and radio resource allocation in limited fronthaul green c-rans," *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2602–2617, 2018.
- [32] K. Guo, M. Sheng, J. Tang, T. Q. Quek, and Z. Qiu, "Exploiting hybrid clustering and computation provisioning for green c-ran," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 12, pp. 4063–4076, 2016.
- [33] I. Scales. (2018) The open ran (oran) alliance formed to lever open 5g for 'other' technologies? and much more... [Online]. Available: <https://www.o-ran.org/resources>

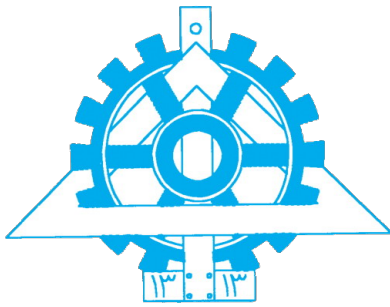
- [34] S. Niknam, A. Roy, H. S. Dhillon, S. Singh, R. Banerji, J. H. Reed, N. Saxena, and S. Yoon, "Intelligent o-ran for beyond 5g and 6g wireless networks," *arXiv preprint arXiv:2005.08374*, 2020.
- [35] L. Gavrilovska, V. Rakovic, and D. Denkovski, "From cloud ran to open ran," *Wireless Personal Communications*, pp. 1–17, 2020.
- [36] S. Sree and S. Ponnekanti, "Open ran deployment using advanced radio link manager framework to support mission critical services in 5g," *EAI Endorsed Transactions on Cloud Systems*, vol. 5, no. 14, 2019.
- [37] S. A. T. Kawahara and A. U. R. Matsukawa, "O-ran alliance standardization trends," 2019.
- [38] X. Wang, C. Wu, F. Le, A. Liu, Z. Li, and F. Lau, "Online vnf scaling in datacenters," in *2016 IEEE 9th International Conference on Cloud Computing (CLOUD)*. IEEE, 2016, pp. 140–147.
- [39] Y. Jia, C. Wu, Z. Li, F. Le, and A. Liu, "Online scaling of nfv service chains across geo-distributed datacenters," *IEEE/ACM Transactions on Networking*, vol. 26, no. 2, pp. 699–710, 2018.
- [40] Z. Luo and C. Wu, "An online algorithm for vnf service chain scaling in datacenters," *IEEE/ACM Transactions on Networking*, 2020.
- [41] R. Cziva, C. Anagnostopoulos, and D. P. Pazaros, "Dynamic, latency-optimal vnf placement at the network edge," pp. 693–701, 2018.
- [42] J. Pei, P. Hong, M. Pan, J. Liu, and J. Zhou, "Optimal vnf placement via deep reinforcement learning in sdn/nfv-enabled networks," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 2, pp. 263–278, 2019.
- [43] B. Ren, S. Gu, D. Guo, G. Tang, and X. Lin, "Joint optimization of vnf placement and flow scheduling in mobile core network," *IEEE Transactions on Cloud Computing*, 2020.
- [44] V. Sciancalepore, K. Samdanis, X. Costa-Perez, D. Bega, M. Gramaglia, and A. Banchs, "Mobile traffic forecasting for maximizing 5g network slicing resource utilization," in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*. IEEE, 2017, pp. 1–9.
- [45] K. Wang, W. Zhou, and S. Mao, "On joint bbu/rrh resource allocation in heterogeneous cloud-rans," *IEEE Internet of Things Journal*, vol. 4, no. 3, pp. 749–759, 2017.
- [46] A. De Domenico, Y.-F. Liu, and W. Yu, "Optimal virtual network function deployment for 5g network slicing in a hybrid cloud infrastructure," *arXiv preprint arXiv:2006.08774*, 2020.

- [47] Y. Hua, R. Li, Z. Zhao, X. Chen, and H. Zhang, "Gan-powered deep distributional reinforcement learning for resource management in network slicing," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 2, pp. 334–349, 2019.
- [48] Y. Hua, R. Li, Z. Zhao, H. Zhang, and X. Chen, "Gan-based deep distributional reinforcement learning for resource management in network slicing," in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.
- [49] T. Li, X. Zhu, and X. Liu, "An end-to-end network slicing algorithm based on deep q-learning for 5g network," *IEEE Access*, vol. 8, pp. 122 229–122 240, 2020.
- [50] M. K. Motalleb, A. Kabiri, and M. J. Emadi, "Optimal power allocation for distributed mimo c-ran system with limited fronthaul capacity," in *2017 Iranian Conference on Electrical Engineering (ICEE)*. IEEE, 2017, pp. 1978–1982.
- [51] P.-R. Li, T.-S. Chang, and K.-T. Feng, "Energy-efficient power allocation for distributed large-scale mimo cloud radio access networks," in *2014 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2014, pp. 1856–1861.
- [52] O. Simeone, J. Kang, J. Kang, and S. Shamai, "Cloud radio access networks: Uplink channel estimation and downlink precoding," *arXiv preprint arXiv:1608.07358*, 2016.
- [53] S.-H. Park, O. Simeone, O. Sahin, and S. S. Shitz, "Fronthaul compression for cloud radio access networks: Signal processing advances inspired by network information theory," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 69–79, 2014.
- [54] M. Gabay and S. Zaourar, "Vector bin packing with heterogeneous bins: application to the machine reassignment problem," *Annals of Operations Research*, vol. 242, no. 1, pp. 161–194, 2016.
- [55] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation for secure ofdma systems," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 6, pp. 2572–2585, 2012.
- [56] P. R. Montague, "Reinforcement learning: an introduction, by sutton, rs and barto, ag," *Trends in cognitive sciences*, vol. 3, no. 9, p. 360, 1999.

Abstract

This thesis studies on writing projects, theses and dissertations using tehran-thesis class.
It ...

Keywords Writing Thesis, Template, \LaTeX , \XeTeX Persian



University of Tehran
College of Engineering
Faculty of Electrical and
Computer Engineering
Network department



Joint Power Allocation and Network Slicing in an End-to-End ORAN System

A Thesis submitted to the Graduate Studies Office
In partial fulfillment of the requirements for
The degree of Doctor of Philosophy
in Electrical Engineering - Communication and Network

By:

Mojdeh Karbalaee Motalleb

Supervisor:

Dr. Shahmansouri

December 2019