علوم رایانش و فناوری اطلاعات
نشریه علمی انجمن کامپیوتر ایران
مجلد ۲۰، شماره ۱، ۱۴۰۱
صفحات ۴۱-۵۳
مقاله پژوهشی

ذخیره‌سازی موقت در شبکه‌های با رسانه مشترک در شرایط محبوبیت غیر یکنواخت و ناهمگن محتوا با استفاده از روش‌های مکاشفه‌ای**

عبدالله غفاری ششجوانی[۱*]، احمد خونساری[۲]، سیدپویا شریعت پناهی[۳]، معصومه مرادیان[۴]

*نویسنده مسئول، دریافت: ۱۴۰۱/۰۳/۱۸، بازنگری: ۱۴۰۱/۰۴/۲۵، پذیرش: ۱۴۰۱/۰۵/۲۲

[۱] دانشجوی دکتری، مهندسی کامپیوتر، دانشگاه تهران، تهران، ایران
[۲] دانشیار، دانشکده مهندسی برق و کامپیوتر ، دانشگاه تهران، تهران، ایران
[۳] استادیار، دانشکده مهندسی برق و کامپیوتر ، دانشگاه تهران، تهران، ایران
[۴] پژوهشگر پسادکتری، پژوهشکده علوم کامپیوتر ، پژوهشگاه دانش‌های بنیادی، تهران، ایران

چکیده

ذخیره‌سازی موقت یکی از راهکارهای جذاب برای کاهش هزینه و تأخیر دستیابی به داده‌های موردنظر در کاربردهای مختلف ازجمله پردازشگرها و شبکه‌های مخابراتی است. در این مقاله، مسئله ذخیره‌سازی موقت با استفاده از روش ترکیبی کدشده-غیرکدشده در شبکه‌های با رسانه مشترک در حالت غیریکنواخت و ناهمگن بودن محبوبیت محتواها مورد توجه قرار گرفته است. در حالت ناهمگن بودن محبوبیت، که در آن محبوبیت محتواها در مکان‌های مختلف متفاوت است، یافتن جانمایی بهینه در ذخیره‌سازهای موقت رام‌نشدنی و از پیچیدگی ترکیبیاتی است. لذا در این مقاله چندین روش مکاشفه‌ای برای ذخیره‌سازی موقت در شرایط ناهمگن بودن محبوبیت محتوا ارائه خواهد شد که مبتنی بر انجام مصالحه بین محبوبیت محلی و محبوبیت سراسری و انجام خوشه‌بندی هستند. روش‌های پیشنهادی از طریق تحلیل عددی و شبیه‌سازی مورد ارزیابی و مقایسه با روش‌های قبلی قرار گرفته است. نتایج ارزیابی‌ها نشان می‌دهد که نرخ ترافیک روی رسانه مشترک با استفاده از روش‌های مکاشفه‌ای مبتنی بر روش ذخیره‌سازی موقت ترکیبی به میزان قابل‌توجهی کمتر از روش‌های قبلی ازجمله روش‌های کدشده محض و غیرکدشده محض است.

کلمات کلیدی: ذخیره سازی موقت، توزیع محبوبیت غیریکنواخت، توزیع محبوبیت ناهمگن، ذخیره سازی موقت ترکیبی کدشده-غیرکدشده، رسانه اشتراکی.

۱- مقدمه

با افزایش روز افزون تعداد ابزارهای بی‌سیم از جمله گوشی‌های هوشمند و تبلت‌ها، درخواست دریافت محتوا از طریق شبکه‌های بی‌سیم و به خصوص شبکه‌های سلولی در سال‌های اخیر افزایش قابل توجهی یافته است. هر چند معرفی فناوری‌های نوین شبکه‌های بی‌سیم از جمله LTE و LTE/A ظرفیت تبادل داده در این شبکه‌ها را افزایش داده است، ولی به دلیل محدود بودن طیف فرکانسی پاسخگوی حجم روز افزون درخواست‌ها به خصوص در مناطق پرتراکم شهری نخواهد بود. یکی از راهکارهایی که برای افزایش ظرفیت در شبکه‌های سلولی، مطرح شده است استفاده از ایستگاه‌های پایه کوچک[۱] (SBS) به‌منظور کاهش اندازه سلول‌ها و لذا ایجاد امکان استفاده مجدد از طیف فرکانسی می‌باشد [۱].

** این مقاله گسترش یافته مقاله با عنوان "Content Caching in Shared Medium Networks with Non-Uniform and User-Dependent Demands" است که در کنفرانس IEEE ICC 2022 ارائه گردیده است.

کوچک کردن اندازه سلول‌ها علاوه بر افزایش ظرفیت شبکه، تأخیر و مصرف باتری کمتر در ابزارهای بی‌سیم کاربران را نیز باعث می‌گردد. با این وجود، به دلیل هزینه‌بر بودن ایجاد اتصالات سیمی برای SBS ها از یک سو و محدودیت‌های پهنای باند بی‌سیم از سوی دیگر، ارتباط بین SBS ها و هسته شبکه یکی از چالش‌های اصلی این راهکار است. برای حل این چالش استفاده از ذخیره‌سازی موقت محتواهای محبوب در SBS ها برای کم کردن نیاز آن‌ها به ارتباطات پرسرعت به هسته شبکه پیشنهاد شده است [۲, ۳, ۴].

ذخیره‌سازی موقت یکی از راهکارهای جذاب برای کاهش هزینه و تأخیر دستیابی به داده‌های موردنظر در کاربردهای مختلف از جمله پردازشگرها و شبکه‌های مخابراتی است. از جمله مزایای استفاده از ذخیره‌سازی موقت در شبکه‌های مخابراتی می‌توان سرعت دریافت بالاتر محتوا، کاهش تأخیر و ازدحام کمتر در ستون فقرات شبکه (با محلی شدن ترافیک) را نام برد.

به‌طورکلی در روش‌های سنتی ذخیره‌سازی موقت در شبکه‌های مخابراتی تلاش می‌گردد محبوب‌ترین محتواها در نزدیکی کاربرانی که به آن‌ها علاقه دارند ذخیره‌سازی شوند. نتیجه این کار باعث کاهش نرخ ترافیک روی شبکه می‌شود که به آن سود محلی ذخیره‌سازی موقت[۲] می‌گویند و بسته به حجم حافظه ذخیره‌سازهای محلی مقدار آن تغییر می‌کند. با ارائه روش ذخیره‌سازی موقت کدشده توسط مداح علی و همکارانش [۵]، نشان داده شد که می‌توان در یک رسانه مشترک نرخ تبادل داده بین ذخیره‌سازهای موقت و سرور محتوا را با استفاده از مزیت همه پخشی و کد نمودن قطعات درخواست شده با یکدیگر به میزان قابل توجهی کاهش داد. استفاده از این روش، علاوه بر سود ذخیره‌سازی محلی منجر به سود دیگری با نام سود سراسری ذخیره‌سازی موقت[۳] به دلیل بهره بردن از فرصت‌های چندپخشی روی رسانه مشترک می‌گردد. میزان سود سراسری ذخیره‌سازی موقت با جمع اندازه حافظه تمامی ذخیره‌سازهای موقت متناسب است.

با توجه به اینکه ارتباط بین SBS ها با مرکز شبکه می‌تواند با استفاده از رسانه مشترک بی‌سیم و از طریق ایستگاه پایه بزرگ[۴] (MBS) برقرار گردد، به این ترتیب شرایط استفاده از روش ذخیره‌سازی موقت کدشده وجود دارد. برای افزایش دادن فرصت‌های چندپخشی در ذخیره‌سازی موقت کدشده می‌بایست قطعات متنوع محتواها در ذخیره‌سازهای مختلف ذخیره گردد که به آن اصل تنوع[۵] گفته می‌شود. ولی از طرف دیگر در شرایطی که محبوبیت محتواها غیریکنواخت است، ذخیره‌سازی موقت محتواهای محبوب در تعداد بیشتری ذخیره‌ساز مختلف، کارایی را بهبود می‌دهد که به آن اصل محبوبیت[۶] گفته می‌شود. در حالی که اصل تنوع تمایل بر این دارد که محتواها در ذخیره‌سازهای مختلف متفاوت باشند، اصل محبوبیت تمایل بر یکسان‌سازی محتواها در ذخیره‌سازهای مختلف دارد. لذا این دو اصل در تناقض با هم بوده و یافتن نقطه بهینه بین آن دو در سناریوهای مختلف و در شرایطی که محبوبیت محتواها غیر یکنواخت است چالش‌برانگیز است.

گرچه روش ذخیره‌سازی موقت کدشده پیشنهاد شده در [۵] در شرایط توزیع یکنواخت محتوا کارایی را به میزان قابل توجهی بهبود می‌دهد، ولی این روش برای توزیع غیریکنواخت محتوا عملکرد مناسبی ندارد. گروه بندی محتواها بر اساس میزان محبوبیت یکی از اولین راهکارهایی بود که برای سازگار نمودن روش ذخیره‌سازی موقت کدشده با توزیع غیر یکنواخت پیشنهاد گردید و پایه روش‌های بعدی قرار گرفت [۶]. کارهای بعدی نشان دادند که حالت بهینه روش گروه بندی ارائه شده در [۶] برای ذخیره‌سازی موقت کدشده، دو قسمت کردن محتوا به دو دسته محبوب و غیر محبوب است. به صورتی‌که محتواهای محبوب با استفاده از روش کدشده ذخیره‌سازی موقت می‌شوند و سایر محتواها اصلاً ذخیره‌سازی موقت نمی‌شوند [۷, ۸, ۹, ۱۰, ۱۱].

در نهایت در [۱۲] و [۱۳] نویسندگان نشان دادند که به خصوص در حالتی که هر کاربر بیش از یک درخواست برای محتوا در فاز تحویل محتوا دارد، تقسیم کردن محتواها به سه قسمت و استفاده از ترکیب ذخیره‌سازی موقت سنتی و کدشده، کارایی را نسبت به روش‌های قبلی بهبود می‌بخشد. در این کارها پیشنهاد شده است که محتواها بر اساس میزان محبوبیت به سه دسته تقسیم‌بندی گردند که عبارت‌اند از:۱-دسته اول شامل محبوب‌ترین محتواها هستند که به صورت کامل و به صورت غیرکدشده ذخیره سازی موقت می‌شوند،۲- دسته دوم محتواهای محبوبی هستند که محبوبیت آن ها از دسته اول کمتر است. این محتواها به صورت کدشده و با استفاده از روش ارائه شده در [۵] ذخیره‌سازی می‌گردند و ۳- دسته سوم محتواهایی که محبوبیت کمی دارند و اصلاً ذخیره‌سازی موقت نمی‌شوند.

کارهای بالا غیریکنواخت بودن محبوبیت محتوا را در نظر گرفته‌اند، ولی اکثر آن‌ها فرض را بر محبوبیت همگن محتوا قرار داده‌اند. درحالی‌که در واقعیت محبوبیت محتواها در مکان‌های مختلف یکسان نیست و محبوبیت محلی محتواها می‌تواند با محبوبیت سراسری آن‌ها بسیار متفاوت باشد. مطالعات مختلف نشان داده که محبوبیت سراسری نمی‌تواند نشانگر خوبی برای محبوبیت محلی باشد. برای مثال، نتایج بررسی نویسندگان در [۱۴] روی ۱۰۰۰ مکان مختلف یکی از شهرهای کشور چین نشان داد که ویدئوها در مکان‌های مختلف دارای محبوبیت‌های کاملاً متفاوتی هستند. ناهمگن بودن توزیع محبوبیت محتوا در این مقاله به این معنی است که توزیع محبوبیت محتوا برای کاربران در محدوده SBS های مختلف می‌تواند یکسان نباشد.

ناهمگنی محبوبیت با تأثیرگذاری بر فرصت‌های چندپخشی بر مصالحه بین اصل محبوبیت و اصل تنوع در ذخیره‌سازی موقت کدشده تأثیر می‌گذارد. با وجود اهمیت توجه به محلی بودن محبوبیت محتواها، تاکنون تعداد کمی از کارها مسئله ناهمگن بودن توزیع محبوبیت را به خصوص برای ذخیره‌سازی موقت کدشده، مورد توجه قرار داده‌اند. از جمله این کارها می‌توان به [۱۵, ۱۶, ۱۷] اشاره کرد. در [۱۶] روش ذخیره‌سازی موقت کدشده [۵] در شرایط توزیع محبوبیت ناهمگن مورد تحلیل قرارگرفته است. با وجود اینکه این کار به ناهمگن بودن توزیع محبوبیت محتوا توجه نموده‌اند، ولی غیریکنواخت بودن محبوبیت محتواها را در نظر نگرفته‌است. در [۱۵] هر کاربر در فاز جانمایی، محتواهای محبوب خود را ذخیره‌سازی موقت می‌کند و سپس در فاز تحویل محتوا سعی می‌کند از فرصت‌های همه‌پخشی سود ببرد. در این کار گرچه غیرهمگن بودن علاقمندی‌های کاربران مورد توجه قرار گرفته است ولی توجهی به انجام بهینه جانمایی محتوا صورت نگرفته است. در نهایت در [۱۷] و [۱۳] شرایط مختلف ناهمگنی و غیریکنواختی برای ذخیره‌سازی موقت در نظر گرفته شده است. در [۱۷] به مسئله در اندازه کوچک وجود فقط دو کاربر و دو محتوا توجه شده است. در [۱۳] نیز نشان داده شده است که یافتن حالت بهینه برای روش‌های ذخیره‌سازی موقت سنتی، کدشده و ترکیبی سنتی-کدشده در شرایط ناهمگن بودن محبوبیت محتواها دارای پیچیدگی ترکیبیاتی است.

لذا نظر به کمبودهای کارهای قبلی، در این مقاله به مسئله ذخیره‌سازی موقت در شبکه‌های با رسانه مشترک با در نظر گرفتن توزیع ناهمگن محبوبیت محتواها پرداخته و سعی در ارائه چندین راهکار مکاشفه‌ای بر پایه روش ترکیبی پیشنهادی در [۱۲] و [۱۳] می‌نماییم که در مقیاس بزرگ قابل استفاده باشند. روش‌های پیشنهادی مبتنی بر انجام مصالحه بین محبوبیت محلی و محبوبیت سراسری و انجام خوشه‌بندی روی SBS ها هستند. از طریق محاسبات عددی و شبیه‌سازی نشان خواهیم داد که روش‌های پیشنهادی مبتنی بر ذخیره‌سازی موقت ترکیبی سنتی-کدشده به میزان قابل توجهی کارایی بهتری در مقایسه با روش‌های قبلی دارند.

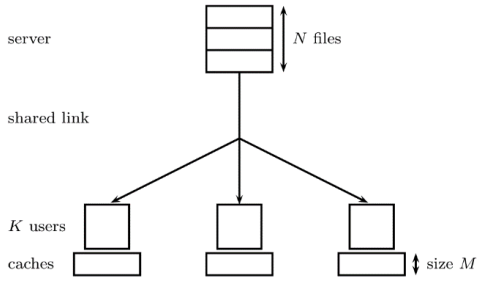
در ادامه این مقاله و در بخش دوم، مفاهیم پایه مورد نیاز را مورد بررسی قرار داده و کارهای پیشین را مرور می‌کنیم. در بخش سوم، ابتدا مدل سیستم را معرفی نموده، سپس تعریف و فرموله‌بندی مسئله ذخیره‌سازی موقت در شرایط ناهمگن بودن محبوبیت را بیان می‌کنیم. در بخش چهارم روش‌های مکاشفه‌ای پیشنهادی را ارائه خواهیم نمود. نتایج تحلیل عددی و شبیه‌سازی‌ها در بخش پنجم گزارش می‌گردند. در نهایت در بخش ششم مقاله جمع‌بندی شده و کارهای آتی را معرفی خواهیم کرد.

۲- پیش‌زمینه و کارهای پیشین

در این بخش ابتدا پیش‌زمینه‌ای در مورد ذخیره سازی موقت کدشده را مطالعه می‌کنیم و سپس به مروری برخی از کارهای پیشین در این حوزه می‌پردازیم.

۲-۱- پیش‌زمینه ذخیره‌سازی موقت کدشده

همان‌گونه که قبلاً گفته شد با ارائه روش ذخیره‌سازی موقت کدشده توسط مداح علی و همکارانش [۵]، نشان داده شد که می‌توان در یک رسانه مشترک نرخ تبادل داده بین ذخیره‌سازهای موقت و سرور محتوا را با استفاده از مزیت همه‌پخشی و کد نمودن قطعات درخواست شده با یکدیگر به میزان قابل‌توجهی کاهش داد. در این روش ذخیره‌سازی موقت کدشده همان‌گونه که در شکل ۱ نشان داده شده، فرض گردیده که یک سرور محتوا وجود دارد که تمامی N محتوا در آن ذخیره شده است. همچنین حجم تمامی محتواها یکسان و برابر F بیت می‌باشد. سرور از طریق یک رسانه مشترک به K کاربر که هرکدام به ذخیره‌ساز موقت به اندازه $M\times F$ بیت مجهز هستند مرتبط است.



شکل ۱- شمایی از سیستم ذخیره‌سازی موقت کدشده مداح علی و همکارانش [۵]

در این روش دو فاز وجود دارد: فاز جانمایی و فاز تحویل محتوا. در فاز جانمایی هر محتوا به $\binom{K}{T}$ تکه که با یکدیگر همپوشانی ندارند تقسیم می‌شوند که در آن مقدار $T=\frac{K\times M}{N}$ است و اندازه هر تکه محتوا برابر $F/\binom{K}{T}$ می‌باشد. تکه‌های محتواها در ذخیره‌ساز موقت کاربران توزیع می‌شوند به‌صورتی که هر کاربر M/N از هر محتوا را ذخیره می‌نماید. بنابراین هر تکه محتوا T کپی در T ذخیره‌ساز موقت مختلف خواهد داشت. در فاز تحویل محتوا هر کاربر یک محتوا را درخواست می‌نماید، سپس سرور تکه‌های موردنیاز کاربران مختلف را با استفاده از راهکار خاصی XOR می‌نماید و بر روی رسانه مشترک برای همه کاربران ارسال می‌نماید.

۲-۲- کارهای پیشین

با توجه به اینکه مسائل مربوط به ذخیره‌سازی موقت در شرایط ناهمگن بودن محبوبیت در بسیاری از سناریوها معمولاً از طریق بهینه‌سازی در پیچیدگی زمانی چندجمله‌ای قابل‌حل نیست، اکثر مقالات در این حوزه به ارائه روش‌های اکتشافی پرداخته و یا از روش‌های یادگیری ماشین استفاده نموده‌اند. در ادامه این قسمت برخی از کارهایی که در حوزه ذخیره‌سازی موقت سنتی و کدشده به محبوبیت ناهمگن محتوا توجه نموده‌اند را مرور خواهیم کرد.

کارهای نسبتاً زیادی ناهمگن بودن محبوبیت محتواها را برای ذخیره‌سازی موقت سنتی مورد توجه قرار داده‌اند. در مقاله [۱۸] به مبحث ذخیره‌سازی موقت با استفاده از روش سنتی در دستگاه‌های کاربران با امکان دریافت محتوا توسط یک کاربر از کاربر دیگر با استفاده از اتصال دستگاه به دستگاه ($D2D$)[۷] پرداخته شده است. با توجه به اینکه ارتباط $D2D$ فقط در فاصله محدودی امکان‌پذیر است، در این مقاله علاقه‌مندی هر کاربر به محتوا به صورت جداگانه مورد توجه قرار گرفته و یک مسئله بهینه‌سازی برای حداقل سازی نرخ ترافیک شبکه تعریف نموده است. نشان داده شده که مسئله تعریف شده NP-Hard بوده و لذا مقاله به سراغ ارائه روش حریصانه رفته است. برای این منظور علاقه‌مندی کاربران با روش تحلیل معنایی نهفته احتمالاتی (pLSA)[۸]، رفتار درخواست محتوا از سوی کاربران مدل گردیده و سپس با استفاده از روش حداکثرسازی امید[۹] (EM) مقدار پارامترها تخمین زده شده است. نتایج شبیه‌سازی در این مقاله نیز نشان می‌دهد که میزان علاقه‌مندی کاربران به‌سرعت قابل تخمین بوده و در مقایسه با روش‌های مبتنی بر محبوبیت محتوا، توانسته سیاست ذخیره‌سازی موقت بهتری یافته و میزان نرخ ترافیک شبکه را به میزان بیشتری کاهش دهد.

مقاله [۱۹] نیز به مبحث ذخیره‌سازی موقت با استفاده از روش سنتی پرداخته و ارزیابی محبوبیت محتواها به‌صورت برخط بر اساس خصوصیات محتواها و علاقه‌مندی کاربران در دو بعد زمان و مکان مورد توجه قرار گرفته است. هدف بهینه‌سازی در این مقاله حداکثر سازی نرخ برخورد ذخیره‌ساز موقت است و از تابع سیگموئید[۱۰] برای تخمین برچسب دسته (موردعلاقه و غیرموردعلاقه برای کاربران آن منطقه) و از رگرسیون برای تخمین علاقه‌مندی کاربران استفاده شده است. نتایج شبیه‌سازی این مقاله نشان می‌دهد که روش پیشنهادی در این مقاله از منظر نرخ برخورد ذخیره‌ساز موقت بهتر از روش‌های مرسوم LRU و LFU بوده است.

در سال‌های اخیر تعداد زیادی از کارها از جمله [۲۰, ۲۱, ۲۲, ۲۳, ۲۴, ۲۵] استفاده از روش‌های یادگیری ماشین را برای حل مسئله جانمایی محتوا در ذخیره-سازی موقت سنتی در شرایط ناهمگن بودن محبوبیت مورد توجه قرار داده‌اند. برخی از کارها از جمله [۲۰, ۲۲] به یادگیری محبوبیت از طریق روش‌های یادگیری تقویتی توجه نموده‌اند. تعداد زیادی از کارها از جمله [۲۳] به یادگیری محبوبیت محتوا با استفاده از شبکه های عصبی LSTM پرداخته‌اند. [۲۴, ۲۵] نیز به یادگیری انتقال[۱۱] در شرایطی که تعداد کاربران کم هستند توجه نموده‌اند. در این مقالات از نتایج یادگیری در دامنه‌های دیگر برای بهبود یادگیری در دامنه هدف استفاده شده است تا کم بودن تعداد کاربران و درنتیجه نمونه‌های آموزش از این طریق جبران شود و یادگیری بهبود یابد.

در [۲۰] فرض شده است که B عدد BS وجود دارد که هرکدام دارای ذخیره‌ساز موقت محدود به اندازه M هستند و محبوبیت محتواها با گذر زمان تغییر کرده و در محدوده BS های مختلف می‌تواند متفاوت باشد. بر اساس این سناریو در این مقاله به حل مسئله در دو حالت مرکزی و توزیع‌شده پرداخته است. در حالت مرکزی نیاز است تا تمامی اطلاعات از تمام BS ها جمع‌آوری شود، لذا با افزایش تعداد BS ها و محتواها، پیچیدگی محاسباتی مسئله بالا می‌رود. لذا برای کاهش سربار ارسال پیام و محاسباتی روش مرکزی، روش توزیع‌شده ارائه شده است. در روش توزیع‌شده هر BS به‌صورت محلی بر اساس درخواست‌های کاربرانش یادگیری را با استفاده از روش MAB[۱۲] انجام می‌دهد. نتایج مقاله نشان می‌دهد که با توجه به تغییر محبوبیت در زمان، تقسیم کردن زمان باعث می‌شود که محبوبیت بهتر ارزیابی گردد، ولی با تقسیم کردن زمان به دلیل کم شدن نمونه‌های آزمایش ممکن است دقت یادگیری کاهش یابد.

مقاله [۲۲] نیز با استفاده از راهکارهای یادگیری تقویتی به دنبال پیش‌بینی محبوبیت زمانی و مکانی درخواست‌ها است. SBS اطلاعات مربوط به محبوبیت محتواهایی که به‌صورت محلی به دست آورده را برای اپراتور شبکه ارسال می‌کند و ارزیابی محبوبیت سراسری را از آن دریافت می‌کند. پس از تبادل اطلاعات، جایگزینی در ذخیره‌ساز موقت با انتخاب بهینه محتواها انجام می‌شود. تابع هزینه در این مقاله از جمع وزن‌دار هزینه جانمایی، هزینه ذخیره نشدن محتواهای محبوب محلی و هزینه ذخیره نشدن محتواهای محبوب سراسری به دست می‌آید. درنهایت مقاله با استفاده از دو الگوریتم از نوع یادگیری تقویتی که یکی از آن‌ها مبتنی بر انجام تقریب برای مقیاس‌پذیری بهتر است به حل مسئله پرداخته است. نتایج نیز نشان می‌دهد که الگوریتم مبتنی بر تقریب ارائه شده سریع‌تر همگرا می‌شود.

گرچه محلی بودن محبوبیت محتواها (ناهمگن بودن محبوبیت) برای روش ذخیره‌سازی موقت سنتی مورد توجه کارهای قبلی بوده است، ولی تاکنون تعداد

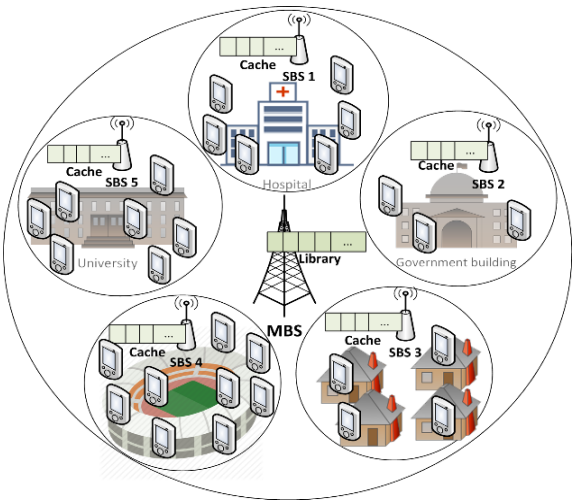
کمی از کارها به ناهمگن بودن محبوبیت برای ذخیره‌سازی موقت کدشده، پرداخته‌اند. در [۱۶] و [۲۶] روش ذخیره‌سازی موقت کدشده مداح علی و همکارانش [۵] در شرایط توزیع محبوبیت ناهمگن مورد تحلیل قرار گرفته است. با وجود اینکه این دو کار به ناهمگن بودن توزیع محبوبیت محتوا توجه نموده‌اند، ولی غیریکنواخت بودن محبوبیت محتواها را در نظر نگرفته‌اند. در [۱۵] هر کاربر در فاز جانمایی محتواهای محبوب خود را ذخیره‌سازی موقت می‌کند و سپس در فاز تحویل محتوا بر اساس درخواست‌های آمده و محتواهایی که در ذخیره‌سازهای موقت قرار دارند، سعی می‌کند با استفاده از یک الگوریتم پوشش‌دسته[۱۳] ساخت‌یافته از فرصت‌های همه‌پخشی سود ببرد. در این کار گرچه ناهمگن بودن علاقه‌مندی‌های کاربران مورد توجه قرار گرفته است ولی توجهی به انجام بهینه جانمایی محتوا صورت نگرفته است. درنهایت در [۱۷] شرایط مختلف ناهمگنی و غیریکنواختی هم در ساختار و هم در محبوبیت شامل ناهمگنی در اندازه محتواها، حجم ذخیره‌سازهای موقت و محبوبیت محتوا را در نظر گرفته‌اند. گرچه این کار ناهمگنی کامل را در نظر گرفته با این حال مسئله را فقط در مقیاس کوچک وجود فقط دو کاربر و دو محتوا تحلیل نموده است.

۳- مدل سیستم، تعریف و فرموله بندی مسئله

در این بخش سعی می‌کنیم روش ذخیره‌سازی موقت ترکیبی کدشده و سنتی [۱۲] و [۱۳] را برای حالت ناهمگن و غیریکنواخت بودن محبوبیت و ناهمگن بودن اندازه ذخیره‌سازهای موقت مورد تحلیل قرار داده و راهکارهای مقیاس‌پذیر ارائه نماییم. لذا در این بخش ابتدا مدل سیستم را معرفی نموده و پس از آن به ارائه راهکار ذخیره‌سازی بهینه برای محبوبیت ناهمگن محتوا مبتنی بر روش ترکیبی پیشنهادی در [۱۲] و [۱۳] خواهیم پرداخت. سپس مسئله را فرموله بندی نموده و در مورد پیچیدگی زمانی یافتن حالت بهینه بحث خواهیم نمود.

۳-۱- مدل سیستم

همان‌طور که در شکل ۲ نشان داده شده است، ما یک شبکه سلولی را در نظر گرفته‌ایم که متشکل از یک MBS است که از طریق یک رسانه مشترک بدون خطا به K عدد SBS متصل است. مجموعه محتواها مشتمل بر $N=\{W_1, W_2, \dots, W_N\}$ محتوای یکتا است که همگی توسط MBS قابل‌دستیابی هستند.



شکل ۲- مدل سیستم با در نظر گرفتن محبوبیت ناهمگن محتواها

بدون اینکه خللی به کلیت موضوع وارد شود، ما فرض کرده‌ایم که تمامی محتواها دارای اندازه یکسان برابر F بیت هستند. هر B نوعی c مسئول خدمت‌رسانی به Z_c کاربر است که $c \in \{1, 2, \dots, K\}$ می‌باشد و همچنین مجهز به ذخیره‌ساز موقت به حجم $M_c \times F$ بیت است که M_c عددی صحیح بوده و $M_c \in [0, N]$ می‌باشد. در این سناریو فرض نمودیم که هر کاربر می‌تواند به MBS و فقط یک SBS متصل و از آن‌ها داده دریافت نماید.

مشابه کارهای پیشین، سیستم پیشنهاد شده در دو فاز عمل می‌کند: ۱- فاز جانمایی محتوا و ۲- فاز تحویل محتوا. فاز جانمایی محتوا در زمان خلوتی شبکه انجام می‌شود. در این فاز، ذخیره‌سازی موقت بر اساس پارامترهایی نظیر میزان محبوبیت محتواها و محدودیت‌های حجم ذخیره‌سازهای موقت انجام می‌گردد.

در فاز تحویل محتوا، فقط MBS به کل محتواها دسترسی مستقیم دارد. در هر فاز تحویل هر کدام از کاربران یک درخواست برای SBS مربوط به خود ارسال می‌دارد، لذا با توجه به اینکه فرض کرده‌ایم به هر SBS نوعی c تعداد Z_c کاربر متصل‌اند و خدمت دریافت می‌دارند، هر SBS نوعی c در هر اسلات تعداد Z_c درخواست برای محتوا دریافت می‌کند. SBS ها نیز با استفاده از داده‌هایی که در ذخیره‌ساز موقت خود دارند ممکن است بتوانند برخی درخواست‌ها را پاسخ دهند و برخی دیگر را برای MBS ارسال نمایند. لازم به ذکر است که با توجه به اینکه فرض کرده‌ایم هر SBS به تعداد دلخواهی کاربر ممکن است متصل باشد، لذا مدل جامع بوده و لذا فرض اینکه هر کاربر یک درخواست را در هر بار فاز تحویل محتوا ارسال می‌کند محدودکننده مدل نیست. بنابراین، تعداد درخواست‌هایی که هر SBS در هر فاز تحویل محتوا دریافت می‌دارد عدد دلخواهی بوده، و هر SBS ممکن است چندین درخواست برای هر محتوا دریافت دارد. امکان دریافت درخواست تکراری برای یک محتوا در SBS وجود دارد. بنابراین، تعداد محتواهای یکتا درخواست شده می‌تواند بین یک تا Z_c باشد (به دلیل امکان درخواست شدن محتواهای یکسان توسط کاربران مختلف). به‌محض دریافت درخواست‌های کاربران در SBS ها، و مشخص شدن محتواهایی که توسط ذخیره‌سازهای موقت قابل ارسال برای کاربران نیست، لیست این محتواها توسط SBS ها به MBS ارسال می‌گردد. MBS نیز فایل‌های مورد نیاز را روی رسانه مشترک برای پاسخ به درخواست‌ها ارسال می‌کند.

توزیع محبوبیت محتواها دلخواه فرض شده است، یعنی محبوبیت محتواها می‌تواند غیر یکنواخت و ناهمگن باشد. به صورتی‌که کوچک‌ترین واحد برای محبوبیت را محدوده یک SBS (که در حقیقت محدوده وجود یک حافظه ذخیره‌ساز موقت است) در نظر می‌گیریم و فرض می‌کنیم توزیع محبوبیت به ازای هر SBS یک توزیع دلخواه است. محبوبیت یک محتوا می‌تواند با دیگر محتواها و همچنین به ازای SBS های مختلف یکسان نباشد. لذا $p_{n,c}$ احتمال درخواست شدن محتوای W_n توسط کاربران حاضر در محدوده SBS نوعی c است (محبوبیت محتوای W_n در محدوده SBS نوعی c) که در آن $n \in \{1, 2, \dots, N\}$ و $c \in \{1, 2, \dots, K\}$ می‌باشد. ممکن است کاربران در محدوده SBS های مختلف علاقمندی‌های مشابهی نیز داشته باشند، لذا ممکن است چندین گروه از SBS های دارای توزیع‌های مشابه محبوبیت در محدوده یک MBS وجود داشته باشد. همچنین ممکن است محتواهایی وجود داشته باشد که تقریباً برای همه کاربران محبوب باشد و از طرف دیگر برخی محتواها فقط برای گروه خاصی از کاربران محبوب باشد. شکل ۲ یک مثال از امکان متفاوت بودن علاقمندی کاربران در محدوده SBS های مختلف را نشان می‌دهد. این شکل، یک MBS که پنج SBS در محدوده آن وجود دارد را نمایش می‌دهد که SBS ها به گروهایی از کاربران با علاقمندی‌های مختلف خدمت‌رسانی می‌کنند.

بر اساس توزیع محبوبیت محتوا تعداد درخواست به ازای هر محتوا و بردار درخواست محتوا در SBS نوعی c مقادیری تصادفی هستند. بر همین اساس، بار ترافیک روی رسانه مشترک که از طرف MBS برای برآورده کردن درخواست‌های تمامی SBS ها در فاز تحویل ارسال می‌گردد، که آن را با نماد R نمایش می‌دهیم، نیز متغیری تصادفی خواهد بود. برخلاف فاز جانمایی، در فاز تحویل، هزینه بار روی شبکه بالا بوده و پهنای باند در دسترس رسانه مشترک گلوگاه سیستم خواهد بود. لذا در این مقاله فقط بار ترافیک در فاز تحویل محتوا مورد توجه است. نمادهای اصلی که در این مقاله مورد استفاده قرار خواهد گرفت در جدول ۱ نمایش داده شده است.

جدول ۱: نمادهای اصلی استفاده شده در مقاله

نماد	توضیح
n	محتوای نوعی
c	SBS نوعی (مجهز به ذخیره‌ساز موقت)
i	شماره یک مرحله نوعی در ارسال محتواها به صورت کدشده
N	تعداد محتواها
F	اندازه هر محتوا (بیت)
K	تعداد SBS ها
M_c	حجم ذخیره‌ساز موقت در SBS ام c
Z_c	تعداد کاربرانی که در محدوده SBS ام c قرار دارند
Z_{max}	$\max(\{Z_c\}_{c=1}^{K})$
$p_{n,c}$	محبوبیت محتوای n ام در SBS ام c
g	گروه نوعی از SBS ها
$Y_{n,c}$	مشخص می‌کند که آیا محتوای n ام در SBS ام c ذخیره‌سازی موقت شده یا نشده
$X_{n,g}$	مشخص می‌کند که آیا محتوای n ام در گروه g ام ذخیره‌سازی موقت شده یا نشده
$S_{c,g}$	مشخص می‌کند که آیا SBS ام c عضو گروه g ام هست یا خیر
r_1	امید ریاضی بار ترافیک روی رسانه مشترک در اثر ارسال محتواهای کدشده
r_2	امید ریاضی بار ترافیک روی رسانه مشترک در اثر ارسال محتواهای غیرکدشده
r	امید ریاضی کل بار ترافیک روی رسانه مشترک $(r_1 + r_2)$

۳-۲- تعریف و فرموله‌بندی مسئله

برای طراحی راهکارهای ذخیره‌سازی موقت توجه به اینکه توزیع محبوبیت در اصل ناهمگن (وابسته به SBS) است، اهمیت بالایی دارد. از طرف دیگر باید توجه نمود که کاربران در محدوده SBS های مختلف می‌توانند علاقمندی‌های مشابهی داشته باشند. به‌طور مثال، در شکل ۲ کاربران محدوده SBS شماره ۴ به محتواهای ورزشی،کاربران محدوده SBS های ۲ و ۵ به محتواهای سیاسی علاقمندی بیشتری دارند و کاربران محدوده SBS شماره ۳ هم محتواهای سیاسی و هم محتواهای ورزشی را دنبال می‌کنند

اگر از روش سنتی ذخیره‌سازی موقت در سناریوی توزیع محبوبیت ناهمگن استفاده شود، در صورت ذخیره‌سازی موقت محبوب‌ترین محتواهای سراسری در همه SBS ها، ممکن است MBS مجبور شود محتواهای محبوب محلی را به دفعات روی رسانه مشترک ارسال کند. از طرف دیگر اگر در هر SBS محبوب‌ترین محتواهای محلی (محبوب‌ترین در محدوده خود آن SBS) ذخیره‌سازی موقت شوند، آنگاه MBS می‌بایست محتواهایی که در SBS ها ذخیره‌سازی موقت نشده را روی رسانه مشترک ارسال کند و البته در هر دو حالت MBS از امکان چندپخشی هم سود می‌برد، به این معنی که در صورت وجود چندین درخواست برای یک محتوا فقط نیاز به یک‌بار ارسال آن روی رسانه مشترک است.

اگر از روش ذخیره‌سازی موقت کدشده در سناریو توزیع محبوبیت ناهمگن استفاده شود، ذخیره‌سازی موقت محتوایی که فقط در محدوده یک SBS محبوب است (و برای سایر کاربران محبوب نیست) باعث عدم استفاده بهینه از حافظه ذخیره‌سازهای موقت در سایر SBS ها و افزایش بار ترافیک روی رسانه مشترک می‌شود. لذا با توجه به موارد بالا، در روش ذخیره‌سازی موقت ترکیبی پیشنهادی در [۱۳]، سعی گردیده تا مصالحه‌ای بین محبوبیت محلی و محبوبیت سراسری با استفاده از هر دو روش‌های کدشده و غیرکدشده ایجاد کند. در ادامه این بخش ابتدا راهکار ترکیبی ارائه شده در [۱۳]، را به صورت مختصر مرور کرده و نشان خواهیم داد که یافتن جواب بهینه برای آن در حالت ناهمگن بودن محبوبیت محتواها دارای پیچیدگی ترکیبیاتی است. لذا در نهایت چند راهکار اکتشافی برای حل این مسئله پیشنهاد می‌دهیم.

۳-۲-۱- پیش‌زمینه ذخیره‌سازی موقت ترکیبی

در فاز جانمایی راهکار ذخیره‌سازی موقت ترکیبی [۱۳]، حافظه ذخیره‌سازی هر کدام از SBS ها به دو بخش تقسیم می‌شود، بخش کدشده و بخش غیرکدشده. مقدار حافظه اختصاص داده شده به بخش کدشده SBS نوعی c را با M_{1c} نمایش می‌دهیم. به این ترتیب تعداد M_{1c} محتوا به صورت کامل در بخش غیر کدشده ذخیره‌ساز موقت SBS نوعی c ذخیره می‌گردند. آرایه دو بعدی $Y_{n,c} \in \{0,1\}$این‌گونه تعریف می‌کنیم: اگر محتوای n ام در SBS نوعی c ام به صورت کامل ذخیره‌سازی موقت شده باشد آنگاه $Y_{n,c} = 1$ و در غیر این صورت $Y_{n,c} = 0$ خواهد بود. بر همین اساس خواهیم داشت: $\sum_{n=1}^{N} Y_{n,c} = M_{1c}$ و مقدار M_c -M_{1c} از حجم ذخیره‌ساز موقت SBS نوعی c ام برای ذخیره‌سازی موقت محتواها به صورت کدشده اختصاص خواهد یافت.

از طرف دیگر، همانطور که قبلاً هم گفته شد، گرچه محبوبیت محتواها برای SBS های مختلف یکسان نیست، ولی ممکن است کاربران در محدوده برخی از SBS ها دارای علاقمندی‌های مشابه باشند. لذا در راهکار ترکیبی امکان وجود گروه‌های مختلف برای تحویل محتوا به صورت کدشده در نظر گرفته شده است و SBS هایی که کاربران حاضر در محدوده آن‌ها علاقمندی‌های مشابه داشته باشند در یک گروه قرار می‌گیرند.

گروه بندی SBS ها در راهکار ترکیبی به این صورت انجام می‌گردد: اگر مجموعه همه SBS ها را با S نشان دهیم، آنگاه $G = \{G_1, G_2, \ldots, G_{|G|}\}$ را پوشانی روی S تعریف می‌کنیم، به‌صورتی‌که هر گروه زیرمجموعه‌ای از S است $(G_i \subset S)$، هر گروه حداقل دو عضو دارد و G یک پوشان روی S است $(\cup_{i=1}^{|G|} G_i = S)$. بر اساس این تعریف، تعداد $|G|$ گروه از SBS ها در پوشان G وجود داشته که هر کدام حداقل دو عضو دارند (برای اینکه بتوان از روش ذخیره‌سازی موقت کدشده استفاده نمود حداقل باید دو ذخیره‌ساز موقت داشته باشد). باید توجه نمود که ممکن است گروه‌ها همپوشانی داشته باشند. یعنی پوشان G لزوماً یک افراز نیست و گروه‌ها می‌توانند عضو مشترک داشته باشند. لذا هر SBS می‌تواند عضو هیچ یا چند گروه باشد، توجه گردد که فضای ذخیره‌ساز موقت SBS ها هم می‌تواند یکسان نباشد. لذا آرایه دو بعدی $S_{c,g} \in \{0,1\}$ را برای مشخص کردن اعضای گروه‌ها این‌گونه تعریف می‌کنیم: اگر SBS نوعی c در گروه نوعی $g \in G$ مشارکت داشته باشد (عضو باشد) آنگاه $S_{c,g} = 1$ و در غیر این صورت $S_{c,g} = 0$ خواهد بود. بر همین اساس تعداد SBS هایی که در گروه نوعی g مشارکت دارند، که آن را با K_g نمایش می‌دهیم، از رابطه $\sum_c S_{c,g} = K_g$ محاسبه می‌شود. همچنین آرایه دو بعدی $X_{n,g} \in \{0,1\}$ را برای مشخص کردن محتواهایی که در SBS ها با استفاده از روش کدشده، ذخیره‌سازی موقت می‌شوند این‌گونه تعریف می‌کنیم: $X_{n,g} = 1$ است اگر محتوای نوعی n با استفاده از روش کدشده در SBS های مشارکت کننده در گروه نوعی g ذخیره‌سازی موقت شده باشد، در غیر این صورت $X_{n,g} = 0$ خواهد بود. بر همین اساس تعداد محتواهایی که با استفاده از روش کدشده در گروه نوعی g ذخیره‌سازی موقت شده‌اند، که آن را با N_g نمایش می‌دهیم، از رابطه $\sum_n X_{n,g} = N_g,\ \forall g$ محاسبه می‌شود.

روش ذخیره‌سازی موقت کدشده مداح علی [۵] به هر گروه به صورت جداگانه اعمال می‌گردد و هر SBS می‌بایست مقداری از فضای ذخیره‌ساز موقت خود را برای مشارکت در هر گروه اختصاص دهد. به‌طور خاص، هر SBS نوعی c که در گروه نوعی g مشارکت دارد (یعنی $S_{c,g} = 1$) می‌بایست مقدار M_g از فضای ذخیره‌سازی خود را برای ذخیره‌سازی موقت محتواهای مربوط به گروه g اختصاص دهد. با توجه به اینکه میزان حافظه ذخیره‌سازی موقت SBS های مختلف نیز می‌تواند یکسان نباشد (SBS نوعی c مقدار M_c حافظه دارد)، اختصاص حافظه در هر SBS نوعی c باید به‌گونه‌ای باشد که عبارت $M_{1c} + \sum_g S_{c,g} \times M_g = M_c$ برقرار باشد.

در فاز تحویل محتوا، هر SBS نوعی c تعداد Z_c درخواست دریافت می‌کند. درخواست‌های مربوط به محتواهای ذخیره‌سازی شده به صورت غیرکدشده (M_{1c} محبوب‌ترین‌ها) از طریق ذخیره‌ساز موقت SBS به صورت محلی و بدون دخالت MBS برای کاربر ارسال می‌گردند. در مقابل درخواست‌های مربوط به محتواهای ذخیره سازی شده به صورت کدشده و ذخیره‌سازی نشده می‌بایست توسط MBS برآورده گردند. با توجه به تعدد کاربران در محدوده هر SBS، در فاز تحویل ممکن

است چندین درخواست محتوا از طرف کاربران در محدوده هر SBS برای MBS ارسال گردد. از Z_c درخواست در محدوده SBS نوعی c، ممکن است برخی مربوط به محتواهای ذخیره‌سازی‌شده به صورت کامل و یا محتواهایی که اصلاً ذخیره‌سازی موقت نشده‌اند باشند. علاوه بر این ممکن است از طرف کاربران محدوده یک SBS درخواست‌های تکراری برای محتواهای ذخیره‌سازی شده به صورت کدشده وجود داشته باشد. لذا تعداد درخواست‌های مربوط به این محتواها می‌تواند کمتر از Z_c باشد. در [۱۳] نشان داده شده که اگر $\Pr\{Q_{i,g}=k\}$ احتمال داشتن درخواست از طرف دقیقا k عدد SBS (در گروه g) برای محتواهای ذخیره‌سازی موقت شده به صورت کدشده در مرحله i ام پاسخ به درخواست‌های کدشده باشد،آنگاه امید ریاضی بار ترافیک برای ارسال محتواهای کدشده (مربوط به همه گروه‌ها) روی رسانه مشترک، که آن را با r_1 نمایش می‌دهیم، برابر است با:

رابطه ۱:
$$r_1=\sum_{g\in\mathcal{G}}\begin{cases}\frac{\sum_{i=1}^{Z_{max}}\binom{K_g}{T_g+1}-\sum_{k=0}^{K_g}\Pr\{Q_{i,g}=k\}\binom{K_g-k}{T_g+1}}{\binom{K_g}{T_g}}, & \text{if } N_g>M_g,\\ 0, & \text{otherwise.}\end{cases}$$

که در آن مقدار $T_g=\frac{K_g\times M_g}{N_g}$ و $Z_{max}=\max_c Z_c$ است. همچنین امید ریاضی نرخ ترافیک برای ارسال محتواهایی که توسط کاربران درخواست می‌شوند ولی در SBS محدوده آن‌ها ذخیره‌سازی موقت نشده‌اند، که آن را با r_2 نمایش می‌دهیم، به صورت زیر محاسبه می‌شود:

رابطه ۲:
$$r_2=\sum_{n=1}^{N}\left(1-\prod_{c=1}^{K}\left(1-\left(p_{n,c}(1-Y_{n,c})\cdot\prod_{g\in\mathcal{G}}(1-X_{n,g}\cdot S_{c,g})\right)\right)^{Z_c}\right),$$

در نهایت، امید ریاضی نرخ ترافیک کل روی رسانه مشترک در فاز تحویل محتوا برابر $r=r_1+r_2$ است. روابط بالا در [۱۳] اثبات شده است و نشان داده شده که یک تخمین مناسب از r در پیچیدگی محاسباتی چند جمله‌ای قابل محاسبه است.

۳-۲-۲- فرموله‌بندی مسئله

اکنون بر اساس موارد گفته‌شده بالا، به فرموله‌بندی مسئله بهینه‌سازی با هدف حداقل‌سازی نرخ ترافیک روی رسانه مشترک می‌پردازیم. اگر $\tilde{G}$ نشان‌دهنده مجموعه تمامی گروه‌های ممکن از SBS ها باشد (که با توجه به اینکه هر گروه می‌بایست حداقل دو عضو داشته باشد داریم: $|\tilde{G}|=2^K-K-k$)، آنگاه مسئله بهینه‌سازی برای حداقل‌سازی نرخ ترافیک روی رسانه مشترک به صورت زیر خواهد شد:

رابطه ۳:
$$\min_{G\subset\tilde{G}}\{\min_{\substack{0\le M_{1c}\le M_c\\ 1\le M_g\le M\\ M_g<N_g\le N\\ X_{n,g},Y_{n,c}\in\{0,1\}}}\{r_1+r_2\}\}$$
$$s.t.$$
$$T_g=\frac{K_g\times M_g}{N_g}\in\mathbb{N},$$
$$\sum_{n=1}^{N}X_{n,g}=N_g,\forall g\in G.$$
$$\sum_{n=1}^{N}Y_{n,c}=M_{1c},\forall c\in\{1,\dots,K\}.$$
$$M_{1c}+\sum_{g\in G}M_g\times S_{c,g}=M_c,\forall c\in\{1,\dots,K\}.$$

در مسئله بالا مقدار r_1 و r_2 از روابط ۱ و ۲ محاسبه می‌شوند. در این مسئله اگر تمامی حافظه هر SBS فقط به بخش غیرکدشده اختصاص یابد (یعنی به ازای $c\in\{1,2,\dots,K\}$ داشته باشیم $M_{1c}=M_c$) آنگاه مسئله به مسئله بهینه‌سازی برای روش ذخیره‌سازی موقت محض غیرکدشده (سنتی) تبدیل می‌شود. به همین ترتیب اگر حافظه‌ای به بخش غیرکدشده اختصاص ندهیم (یعنی به ازای $c\in\{1,2,\dots,K\}$ داشته باشیم $M_{1c}=0$) آنگاه مسئله به مسئله بهینه‌سازی برای روش ذخیره‌سازی موقت محض کدشده تبدیل می‌شود. در حالت بهینه در این مسئله ممکن است محتواهای ذخیره شده در قسمت غیرکدشده برای SBS های مختلف متفاوت باشند. یک حالت خاص از مسئله بالا وقتی است که شباهت معنی‌داری بین محبوبیت محتواها بین برخی SBS ها وجود نداشته باشد (یعنی نتوان بر اساس توزیع محبوبیت SBS ها را به چندین گروه تقسیم بندی نمود). در این حالت، افزایش تعداد گروه‌های تحویل محتوای کدشده، فقط باعث کاهش منفعت بردن از همه پخشی می‌شود، و لذا پوشان بهینه در این حالت داشتن فقط یک گروه ارسال محتوای کدشده است که شامل تمامی SBS ها باشد. توجه گردد که گرچه در این حالت خاص محتواهایی که به صورت کدشده ذخیره‌سازی موقت می‌شوند برای تمامی SBS ها یکسان خواهند بود، ولی ممکن است این محتواها با حالتی که محتواها بر اساس محبوبیت سراسری انتخاب شوند، یکسان نباشند.

۳-۲-۳- بحث پیرامون سختی مسئله

محاسبه امید ریاضی نرخ ترافیک روی رسانه مشترک (r_1+r_2) در [۱۳] با انجام تخمین، استفاده از روش بازگشتی و برنامه‌نویسی پویا (استفاده از حافظه برای نگهداری نتایج میانی در جهت جلوگیری از محاسبات تکراری) در پیچیدگی چندجمله‌ای قابل انجام است. ولی یافتن جواب بهینه برای مسئله بهینه‌سازی مطرح شده در رابطه ۳ رام نشدنی است. حتی اگر پوشان بهینه، چگونگی تقسیم کردن بهینه حافظه (یعنی مقدار M_{1c} برای تمامی SBS ها و M_g برای تمامی گروه‌ها) و تعداد محتواهایی که در هر گروه تحویل محتوای کدشده مشارکت دارند (یعنی N_g به ازای تمامی گروه‌ها) را بدانیم، باز مشخص کردن اینکه کدام محتواها می‌بایست در ذخیره‌سازهای موقت قرار گیرند یک مسئله رام نشدنی بوده و برای یافتن حالت بهینه آن نیاز به $\prod_{c=1}^{K}\binom{N}{M_{1c}}\times\prod_{g=1}^{G}\binom{N}{N_g}$ بار محاسبه مقدار r برای یافتن جانمایی بهینه داریم. حتی اگر فقط حالت خاص وجود یک گروه تحویل محتوای کدشده را در نظر بگیریم، باز برای یافتن حالت بهینه آن نیاز به محاسبه نرخ ترافیک (r) برای $\prod_{c=1}^{K}\binom{N}{M_{1c}}\times\binom{N}{N_g}$ حالت ممکن داریم. همچنین یافتن راه‌حل بهینه در شرایط ناهمگن بودن محبوبیت محتوا برای روش‌های ذخیره‌سازی موقت کدشده محض دو قسمتی [۷, ۸, ۹, ۱۰, ۱۱] (اگر چگونگی تقسیم بندی بهینه محتواها را بدانیم) و غیرکدشده محض (سنتی)، به ترتیب نیاز به محاسبه نرخ بار ترافیک برای $\binom{N}{N_1}$ و $\prod_{c=1}^{K}\binom{N}{M_{1c}}$ حالت ممکن مختلف انتخاب محتواها دارد. لذا همانطور که دیده می‌شود یافتن جانمایی بهینه در شرایط ناهمگن بودن محبوبیت محتواها رام نشدنی و دارای پیچیدگی ترکیبیاتی است. بر این اساس، در ادامه چندین راهکار اکتشافی را برای این مسئله پیشنهاد می‌نماییم.

۴- راهکارهای اکتشافی

در این قسمت، سه راهکار مکاشفه‌ای که هر سه دارای پیچیدگی چندجمله‌ای هستند را برای جانمایی محتوا در شرایط ناهمگن بودن محبوبیت محتوا ارائه می‌نماییم.

۴-۱- راهکارهای اکتشافی اول (روش وابسته-ترکیبی[۱۴])

راهکار پیشنهادی اول را که وابسته-ترکیبی (DH) نامیده‌ایم برای شرایطی پیشنهاد شده است که شباهت معناداری در توزیع محبوبیت محتواها بین برخی

SBS ها وجود ندارد (بر اساس محبوبیت محتواها نتوان SBS ها را به چندین گروه دسته بندی کرد). لذا در این راهکار فقط یک گروه سراسری که شامل همه SBS ها می‌شود، برای تحویل محتوا به صورت کدشده در نظر می‌گیریم (یعنی $\{S_{c,g}\}_{c,g=1}^{K,1}=1$). به‌طور خلاصه، این راهکار سعی می‌کند که مصالحه‌ای بین محبوبیت محلی و محبوب سراسری، از طریق ذخیره‌سازی موقت محتواهای محبوب محلی با استفاده از روش غیرکدشده و محتواهای محبوب سراسری از طریق روش کدشده، برقرار نماید. همانطور که در خط‌های ۲ تا ۵ الگوریتم DH نشان داده شده است، ابتدا فهرست محتواها بر اساس میزان محبوبیت محلی (به ازای هر SBS) و محبوبیت سراسری مرتب‌شده و در دو آرایه ذخیره می‌گردند.

الگوریتم ۱: Dependent Hybrid (DH)

۱: **function** DH $\left(K,N,M,\{Z_c\}_{c=1}^{K},\{p_{n,c}\}_{n,c=1}^{N,K}\right)$
۲: **for all** $c\in\{1,2,\dots,K\}$ **do**
۳: $\{I_{n,c}^{(L)}\}_{n=1}^{N}\leftarrow$ Sort content indexes w.r.t. $\{p_{n,c}\}_{n=1}^{N}$
۴: **end for**
۵: $\{I_{n}^{(G)}\}_{n=1}^{N}\leftarrow$ Sort content indexes w.r.t. $\{p_n=\frac{1}{K}\sum_{c=1}^{K}p_{n,c}\}_{n=1}^{N}$
۶: $\{S_{c,g}\}_{c,g=1}^{K,1}\leftarrow 1$
۷: $r^{\star}\leftarrow -1$
۸: **for** $N_1\in\{M-M_1,\dots,N\}$، where $M_1\in\{0,\dots,M\}$، **do**
۹: $\{X_n\}_{n=1}^{N}\leftarrow 0$
۱۰: $\{Y_{n,c}\}_{n,c=1}^{N,K}\leftarrow 0$
۱۱: **for all** $c\in\{1,\dots,K\}$ and $i\in\{I_{1,c}^{(L)},\dots,I_{M_1,c}^{(L)}\}$ **do**
۱۲: $Y_{i,c}\leftarrow 1$
۱۳: **end for**
۱۴: $\{\hat{I}_n\}_{n=1}^{N}\leftarrow\{I_n^{(G)}\}_{n=1}^{N}-\{I_n^{(G)}|\sum_{c=1}^{K}Y_{I_n^{(G)},c}>0\}_{n=1}^{N}$
۱۵: **for all** $i\in\{\hat{I}_1^{(G)},\dots,\hat{I}_{N_1}^{(G)}\}$ **do**
۱۶: $X_i\leftarrow 1$
۱۷: **end for**
۱۸: محاسبه مقدار r_1 و r_2 را با استفاده از روابط ۱ و ۲
۱۹: $r\leftarrow r_1+r_2$
۲۰: **if** $(r<r^{\star})\,||\,(r^{\star}==-1)$ **then**
۲۱: $r_1^{\star},r^{\star},X_n^{\star},Y_{n,c}^{\star}\leftarrow r_1,r,X_n,Y_{n,c}$
۲۲: **end if**
۲۳: **end for**
۲۴: **end function**
۲۵: **return** $r_1^{\star},r^{\star},X_n^{\star},Y_{n,c}^{\star}$

همانطور که در الگوریتم دیده می‌شود، فهرست مرتب‌شده محتواها بر اساس محبوبیت محلی به ازای هر SBS نوعی c در آرایه دو بعدی $I_{n,c}^{(L)}$ و فهرست مرتب‌شده محتواها بر اساس محبوبیت سراسری در آرایه یک بعدی $I_{n,c}^{(G)}$ ذخیره می‌گردد. سپس هر SBS تعداد M_1 محبوب‌ترین محتواهای محلی خود را (M_1 محتوای اول آرایه $I^{(L)}$) انتخاب کرده و آن‌ها را به صورت کامل (غیر کدشده) ذخیره‌سازی موقت می‌کند (خطوط ۱۱ و ۱۲ الگوریتم DH). سایر فضای ذخیره‌ساز موقت SBS ها M-M_1 برای ذخیره‌سازی موقت کدشده اختصاص می‌یابد. پس از آن تعداد N_1 محبوب‌ترین محتواهای سراسری (N_1 محتوای اول آرایه $I^{(G)}$) که در مرحله قبل در هیچ‌کدام از SBS ها به صورت کامل (غیرکدشده) ذخیره‌سازی موقت نشده بودند برای ذخیره‌سازی موقت با استفاده از روش کدشده مداح علی [۵] انتخاب می‌گردند (خطوط ۱۴ تا ۱۶ الگوریتم DH). در نهایت، امید ریاضی نرخ ترافیک بر اساس جانمایی انجام شده و توزیع محبوبیت محتوا محاسبه می‌گردد (خطوط ۱۸ و ۱۹ الگوریتم DH). بهترین تقسیم بندی حافظه و تعیین تعداد محتواهای کدشده، یعنی مقادیر M_1 و N_1 از طریق جستجوی جامع به دست می‌آید.

۴-۲- راهکارهای اکتشافی دوم (روش خوشه‌بندی شده وابسته-ترکیبی[۱۵])

همانطور که قبلاً هم گفته شد، ممکن است برخی از SBS ها دارای توزیع محبوبیت مشابه باشند و یا به عبارتی گروه‌هایی از SBS ها با توزیع محبوبیت مشابه وجود داشته باشند. گروه بندی این SBS ها به اعمال جداگانه روش ذخیره‌سازی موقت کدشده مداح علی [۵] به هر گروه می‌تواند کارایی ذخیره‌سازی موقت را بهبود بخشد.

روش مکاشفه‌ای دوم، که آن را خوشه‌بندی شده وابسته-ترکیبی (CDH) می‌نامیم، برای حالت وجود این‌چنین گروه‌ها پیشنهاد شده است. الگوریتم ۲ شبه‌کد این روش را نمایش می‌دهد.

الگوریتم ۲: Cluster Dependent Hybrid (CDH)

۱: **function** CDH $\left(K,N,M,\{Z_c\}_{c=1}^{K},\{p_{n,c}\}_{n,c=1}^{N,K}\right)$
۲: $r^{\star}\leftarrow -1$
۳: **for** $j\in\{1,2,\dots,K\}$ **do**
۴: $\mathcal{G}\leftarrow$ k-means$(\{p_{n,c}\}_{n,c=1}^{N,K},j)$
۵: **if** $|\mathcal{G}|<j$ **then** break
۶: **end if**
۷: $\{S_{c,g}\}_{c,g=1}^{K,|\mathcal{G}|}\leftarrow$ Extract clusters from $\mathcal{G}$
۸: $\{X_{n,g}\}_{n,g=1}^{N,|\mathcal{G}|}\leftarrow 0$
۹: $\{Y_{n,c}\}_{n,c=1}^{N,K}\leftarrow 0$
۱۰: **for** $g\in\{1,2,\dots,|\mathcal{G}|\}$ **do**
۱۱: $K_g\leftarrow\sum_{c=1}^{K}S_{c,g}$
۱۲: $\{Z'_{c'}\}_{c'=1}^{K_g}\leftarrow\{Z_c|S_{c,g}=1\}_{c=1}^{K}$
۱۳: $\{p'_{n,c'}\}_{n,c'=1}^{N,K_g}\leftarrow\{\{p_{n,c}\}_{n=1}^{N}|S_{c,g}=1\}_{c=1}^{K}$
۱۴: $r'_1,r',\{X_{n,g}\}_{n=1}^{N},\{Y'_{n,c'}\}_{n,c'=1}^{N,K_g}\leftarrow$ DH $\left(K_g,N,M,\{Z'_{c'}\}_{c'=1}^{K_g},\{p'_{n,c'}\}_{n,c'=1}^{N,K_g}\right)$
۱۵: **for all** c such that $S_{c,g}=1$ **do**
۱۶: $\{Y_{n,c}\}_{n=1}^{N}\leftarrow\{Y'_{n,c'}\}_{n=1}^{N}$
۱۷: **end for**
۱۸: $r_1\leftarrow r_1+r'_1$
۱۹: **end for**
۲۰: محاسبه مقدار r_2 را با استفاده از رابطه ۲
۲۱: $r\leftarrow r_1+r_2$
۲۲: **if** $(r<r^{\star})\,||\,(r^{\star}==-1)$ **then**
۲۳: $r^{\star},S_{c,g}^{\star},X_{n,g}^{\star},Y_{n,c}^{\star}\leftarrow r,S_{c,g},X_{n,g},Y_{n,c}$
۲۴: **end if**
۲۵: **end for**
۲۶: **end function**
۲۷: **return** $r^{\star},S_{c,g}^{\star},X_{n,g}^{\star},Y_{n,c}^{\star}$

راهکار CDH تلاش می‌کند که با خوشه‌بندی روی SBSها ازنظر توزیع محبوبیت محتواها و استفاده از روش DH مصالحه‌ای بین محبوبیت محلی و سراسری انجام دهد. در این روش، همانطور که در الگوریتم CDH نشان داده شده است، با ارجاع آرایه $\{P_{n,c}\}_{n,c=1}^{N,K}$ به عنوان بردار ویژگی و تعداد خوشه‌های مورد نظر به الگوریتم *K-means*، SBSها را در j گروه خوشه‌بندی می‌کنیم (خط چهارم الگوریتم CDH). لذا SBS هایی که دارای توزیع محبوبیت مشابه‌تر هستند، شانس بالاتری برای قرارگیری در یک خوشه دارند. سپس اطلاعات مربوط به SBS های درون هر خوشه (ازجمله Z_c و $\{P_{n,c}\}_{n=1}^{N}$) استخراج (خط ۱۱ تا ۱۳ الگوریتم CDH) و برای الگوریتم DH ارسال می‌گردد. الگوریتم DH روی هر خوشه به صورت جداگانه اعمال و در خروجی جانمایی محتواها و امید ریاضی نرخ ترافیک برای هر خوشه را می‌دهد (خط ۱۴). با اعمال الگوریتم DH روی هر خوشه آرایه مشخص کننده جانمایی محتواهای غیرکدشده در ذخیره‌سازهای موقت (یعنی آرایه Y) و امید ریاضی نرخ ترافیک مربوط به محتواهای کدشده (r_1) بر اساس خروجی الگوریتم DH به روز می‌شود (خطوط ۱۵ تا ۱۸). درنهایت، مقدار امید ریاضی نرخ ترافیک (r_1+r_2) برای کل خوشه‌بندی داده شده محاسبه (خطوط ۲۰ و ۲۱) می‌گردد. اندازه بهینه تعداد خوشه‌ها (یعنی مقدار بهینه برای j) از طریق جستجوی جامع به دست می‌آید.

۴-۳- راهکارهای اکتشافی سوم (خوشه‌بندی‌شده دو قسمتی وابسته-ترکیبی[۱۶])

گرچه محبوبیت محتواها در محدوده SBS های مختلف متفاوت است، ولی ممکن است محتواهایی وجود داشته باشد که تقریباً در همه SBS ها محبوب باشند. ذخیره‌سازی موقت این محتواها در همه SBS ها با استفاده از روش کدشده، ممکن است میزان بهره سراسری ذخیره‌سازی موقت را بهبود بخشد. بر همین اساس، در راهکار پیشنهادی سوم، که آن را **خوشه‌بندی‌شده دو قسمتی وابسته-ترکیبی** (SCDH) می‌نامیم، به محتواهایی که ممکن است برای تقریباً همه SBS ها محبوب باشند توجه گردیده است. الگوریتم ۳ شبه‌کد این روش را نمایش می‌دهد.

الگوریتم۳: Split Cluster Dependent Hybrid (SCDH)

function SCDH $\left(K, N, M, \{Z_c\}_{c=1}^{K}, \{p_{n,c}\}_{n,c=1}^{N,K}, N^{(G)}, \beta\right)$
  $\{Score_n\}_{n=1}^{N} \leftarrow 0$
  $\{p_{n,c}^{(G)}\}_{n,c=1}^{N,K} \leftarrow 0$
  $\{p_{n,c}^{(L)}\}_{n,c=1}^{N,K} \leftarrow 0$
  for all $c \in \{1,2,\ldots,K\}$ and $n \in \{1,2,\ldots,N\}$ **do**
    if n falls in the $2N^{(G)}$ popular contents in SBS c **then**
      $Score_n \leftarrow Score_n + 1$
    end if
  end for
  $\{SI_i\}_{i=1}^{N} \leftarrow$ Sort content indexes w.r.t. $\{Score_n\}_{n=1}^{N}$
  for all $c \in \{1,2,\ldots,K\}$ and $i \in \{SI_1,\ldots,SI_{N^{(G)}}\}$ **do**
    $p_{i,c}^{(G)} \leftarrow p_{i,c}$
  end for
  for all $c \in \{1,2,\ldots,K\}$ and $i \in \{SI_{N^{(G)}+1},\ldots,SI_N\}$ **do**
    $p_{i,c}^{(L)} \leftarrow p_{i,c}$
  end for
  $r^{\star}, \{S_{c,g}^{\star}\}_{c,g=1}^{K,|\mathcal{G}|}, \{X_{n,g}^{(L)\star}\}_{n,g=1}^{N,|\mathcal{G}|}, \{Y_{n,c}^{(L)\star}\}_{n,c=1}^{N,K} \leftarrow$
    CDH $\left(K, N, M, \{Z_c\}_{c=1}^{K}, \{p_{n,c}\}_{n,c=1}^{N,K}\right)$
  $\{X_n^{(G)\star}\}_{n=1}^{N} \leftarrow 0$
  $\{Y_{n,c}^{(G)\star}\}_{n,c=1}^{N,K} \leftarrow 0$
  $M^{(G)\star} \leftarrow 0$
  $M^{(G)} \leftarrow \beta \times M$
  while $M^{(G)} \leq M$ **do**
    $K', r^{(G)}, \{X_n^{(G)}\}_{n=1}^{N}, \{Y_{n,c}^{(G)}\}_{n,c=1}^{N,K} \leftarrow$
      DH $\left(K, N, M^{(G)}, \{Z_c\}_{c=1}^{K}, \{p_{n,c}^{(G)}\}_{n,c=1}^{N,K}\right)$
    $r^{(L)}, \{S_{c,g}\}_{c,g=1}^{K,|\mathcal{G}|}, \{X_{n,g}^{(L)}\}_{n,g=1}^{N,|\mathcal{G}|}, \{Y_{n,c}^{(L)}\}_{n,c=1}^{N,K} \leftarrow$
      CDH $\left(K, N, M - M^{(G)}, \{Z_c\}_{c=1}^{K}, \{p_{n,c}^{(L)}\}_{n,c=1}^{N,K}\right)$
    $r \leftarrow r^{(G)} + r^{(L)}$
    if $r < r^{\star}$ **then**
      $r^{\star}, S_{c,g}^{\star}, X_{n,g}^{(L)\star}, Y_{n,c}^{(L)\star} \leftarrow r, S_{c,g}, X_{n,g}^{(L)}, Y_{n,c}^{(L)}$
      $M^{(G)\star}, X_n^{(G)\star}, Y_{n,c}^{(G)\star} \leftarrow M^{(G)}, X_n^{(G)}, Y_{n,c}^{(G)}$
    end if
    $M^{(G)} \leftarrow M^{(G)} + \beta \times M$
  end while
end function
return $M^{(G)\star}, S_{c,g}^{\star}, X_{n,g}^{(L)\star}, Y_{n,c}^{(L)\star}, X_n^{(G)\star}, Y_{n,c}^{(G)\star}$

گرچه معمولاً محبوب‌ترین محتواها را بر اساس میانگین محبوبیت سراسری، که بر اساس تعداد درخواست‌ها برای یک محتوا در کل سیستم تقسیم بر تعداد کل درخواست‌ها است، معرفی می‌کنند، ولی میانگین محبوبیت سراسری نمی‌تواند شاخص خوبی برای مشخص و انتخاب کردن محبوب بودن یک محتوا در همه SBS ها (محتوای محبوب سراسری) باشد. زیرا محبوبیت بسیار بالای یک محتوا در یک محدوده خاص، می‌تواند شاخص میانگین محبوبیت سراسری را نیز بالا ببرد و این‌گونه برداشت شود که محتوا به صورت سراسری دارای محبوبیت است، در صورتی که چنین نبوده است. شاهد این ادعا، نتایج پژوهش انجام شده در [۱۴] است، که گزارش نموده که در ۶۰ درصد مکان‌ها، میانگین محبوبیت مربوط به ۱۰۰۰ ویدئو محبوب سراسری (۳ دهم درصد کل ویدئوهای تماشا شده) از میانگین محبوبیت ۴۰ درصد از ویدئوهای محبوب محلی کمتر است. یعنی عملاً گرچه این ۱۰۰۰ ویدئو ازنظر محبوبیت سراسری بسیار محبوب ارزیابی شده‌اند ولی حداقل در ۶۰ درصد مکان‌ها محبوب نبوده‌اند.

بر این اساس در راهکار SCDH از یک روش امتیازدهی بر اساس رتبه محبوبیت محتوا برای شناسایی محتواهایی که برای همه محبوب هستند، استفاده می‌کنیم (خط ۲ تا ۱۰ الگوریتم SCDH). لازم به ذکر است که در SCDH از یک پارامتر به نام $N^{(G)}$ در روش امتیازدهی برای تعیین محتواهای محبوب سراسری استفاده می‌کنیم. به این صورت که اگر یک محتوا در بین $2 \times N^{(G)}$ محبوب‌ترین محتواهای یک SBS قرار داشت، امتیاز محبوبیت آن را یک امتیاز افزایش می‌دهیم. پس از بررسی همه محتواها در همه SBS ها، محتواها را بر اساس میزان امتیاز مرتب کرده (خط ۱۰ الگوریتم SCDH) و $N^{(G)}$ محتوا که بیشترین امتیاز را آورده‌اند به عنوان محتواهای محبوب سراسری انتخاب می‌کنیم. برای این منظور میزان محبوبیت $N^{(G)}$ محتوای انتخاب شده به عنوان محتوای محبوب سراسری را در آرایه $\{P_{n,c}^{(G)}\}_{n,c=1}^{N,K}$ و سایر محتواها را در آرایه $\{P_{n,c}^{(L)}\}_{n,c=1}^{N,K}$ ذخیره می‌کنیم (خط ۱۱ تا ۱۶ الگوریتم SCDH).

در راهکار SCDH، حافظه ذخیره‌سازهای موقت SBS ها به دو قسمت ۱-محلی و ۲-سراسری تقسیم بندی می‌شود، که قسمتی که برای ذخیره‌سازی محتواهای سراسری اختصاص می‌یابد با $M^{(G)}$ نشان داده می‌شود. برای اینکه بتوانیم از مزیت همه پخشی برای ارسال محتواهای محبوب سراسری بهره ببریم، سیاست ذخیره‌سازی موقت برای محتواهایی که به عنوان محبوب سراسری انتخاب شده‌اند در همه SBS ها اعمال می‌گردد. برای محتواهای غیر محبوب سراسری از خوشه‌بندی استفاده می‌گردد، لذا SBS هایی که علاقمندی‌های مشابه دارند در یک خوشه قرار گرفته و سیاست ذخیره‌سازی موقت برای محتواهای محبوب محلی به ازای خوشه‌ها اعمال می‌گردد. SCDH این کار را با اعمال الگوریتم‌های DH و CDH به ترتیب روی قسمت‌های سراسری و محلی ذخیره‌سازهای موقت SBS ها انجام می‌دهد (خط‌های ۲۳ و ۲۴ الگوریتم SCDH).

برای بهبود خوانایی شبه‌کد SCDH در خطوط ۱۷ و ۲۴ الگوریتم SCDH، تابع CDH فراخوانی شده و به این ترتیب از نوشتن مراحل تکراری در دو الگوریتم و طولانی شدن شبه‌کد جلوگیری کردیم. ولی لازم به ذکر است که در واقع در کد اصلی الگوریتم SCDH، فقط یک‌بار عمل خوشه‌بندی (مربوط به خط ۴ الگوریتم CDH) قبل از حلقه خط ۲۲ در الگوریتم SCDH انجام می‌شود و لذا هر بار و به ازای مقادیر مختلف $M^{(G)}$ عمل خوشه بندی تکرار نمی‌شود. درنهایت، مقدار امید ریاضی نرخ ترافیک ($r=r^{(G)}+r^{(L)}$) محاسبه شده (خط ۲۵) و پیکربندی بهینه (یعنی مقدار بهینه برای $M^{(G)}$) از طریق جستجوی جامع به دست می‌آید. البته برای یافتن یک سیاست مناسب (ولی نه لزوماً بهینه) برای ذخیره‌سازی موقت با استفاده از SCDH در واقع نیاز به بررسی تمامی مقادیر ممکن برای $M^{(G)}$ نیست. در SCDH امکان مصالحه بین دقت جستجو و سربار محاسباتی با استفاده از تعریف پارامتری با نام β امکان‌پذیر شده است (خطوط ۲۱ و ۳۰ الگوریتمSCDH). می‌توان مقادیر بین $1/M$ تا یک را به β داد، هر چه مقدار β به یک نزدیک‌تر شود دقت و سربار جستجو کمتر و هر چه به $1/M$ نزدیک‌تر شود دقت و سربار جستجو بیشتر می‌شود. همچنین لازم به ذکر است که در روش‌های مکاشفه‌ای ارائه شده در این بخش اگر به‌جای دو قسمت کردن حافظه ذخیره‌سازهای موقت، تمامی حافظه را فقط به بخش کدشده و یا غیرکدشده اختصاص دهیم، آنگاه نسخه‌های کدشده محض و غیرکدشده محض از راهکارهای پیشنهادی خواهیم داشت. تعدادی از این روش‌ها را در فصل بعد برای ارزیابی مورد بررسی قرار خواهیم داد.

۴-۴- تحلیل پیچیدگی

همانطور که دیده شد مسئله اصلی یافتن سیاست بهینه ذخیره‌سازی موقت در شرایط ناهمگن بودن محبوبیت محتواها دارای پیچیدگی ترکیبیاتی بود. راهکارهای اکتشافی ارائه شده در این مقاله همگی دارای پیچیدگی محاسباتی چندجمله‌ای هستند. اصل پیچیدگی محاسباتی روش‌های CDH و SCDH به فراخوانی‌های تابع الگوریتم DH برمی‌گردد. در DH نیز اصل پیچیدگی محاسباتی مربوط به محاسبه امید ریاضی نرخ ترافیک (یعنی r) بر اساس رابطه‌های ۱ و ۲ می‌شود. لذا سربار محاسباتی روش‌های مکاشفه‌ای پیشنهادی متناسب با تعداد دفعات محاسبات r و البته با توجه به اندازه مسئله (بزرگی مقادیر پارامترهای N، K، M) است. لازم به ذکر است که گرچه در روش CDH از خوشه‌بندی استفاده شده و به ازای هر خوشه یک‌بار

DH را فراخوانی می‌کنیم، ولی در هر فراخوانی DH، مسئله کوچک‌تری حل می‌گردد. لذا همانطور که در فصل بعد هم در نتایج نشان داده خواهد شد، پیچیدگی روش CDH با صرف‌نظر از سربار خوشه‌بندی تقریباً با روش DH یکسان می‌شود. در بین روش‌های اکتشافی پیشنهادی SCDH ازنظر محاسباتی بیشترین سربار را دارد. همانطور که در الگوریتم SCDH دیده می‌شود، سربار محاسباتی SCDH از مرتبه $O(CDH)+{}^{1}/_{\beta}\times(O(DH)+O(CDH))$ است، که در آن ${}^{1}/_{M}\le\beta\le 1$ است.

۵- ارزیابی روش‌های پیشنهادی

در این بخش، برای ارزیابی روش‌های مکاشفه‌ای پیشنهادی به مقایسه این روش‌ها با روش‌های پایه‌ای زیر می‌پردازیم:

- غیرکدشده‌محض‌وابسته[۱۷](DPU): در روش ذخیره‌سازی موقت غیرکدشده (سنتی) وابسته به SBS اتخاذ سیاست ذخیره‌سازی موقت توسط هر SBS به‌صورت مستقل از سایر SBS ها و فقط بر اساس محبوبیت محلی محتواها انجام می‌پذیرد. به این معنی که هر SBS فقط علاقه‌مندی کاربران حاضر در محدوده خود را موردتوجه قرار می‌دهد. سیاست ذخیره‌سازی موقت در این روش به این صورت است که هر SBS مجهز به ذخیره‌ساز موقت به‌اندازه M محتوا، تعداد M محبوب‌ترین محتوا در محدوده خود را انتخاب و در ذخیره‌ساز موقت خود قرار می‌دهد.
- کدشده‌محض[۱۸](PC): در روش ذخیره‌سازی موقت کدشده محض دوقسمتی [۷, ۸, ۹, ۱۰, ۱۱] اتخاذ سیاست ذخیره‌سازی موقت که تقسیم‌بندی محتواها به دو گروه محبوب و غیر محبوب است به‌صورت تمرکزی و فقط بر اساس محبوبیت سراسری (یعنی $\{P_n={}^{1}/_{K}\times\sum_{c=1}^{K}P_{n,c}\}_{n=1}^{N}$) انجام می‌پذیرد. در این روش، همان‌طور که قبلاً هم گفته شده، محتواهای گروه محبوب بر اساس روش ذخیره‌سازی موقت کدشده مداح علی و همکارانش [۵] در SBSها ذخیره‌سازی موقت می‌شوند و محتواهای گروه غیر محبوب اصلاً ذخیره‌سازی موقت نمی‌شوند.
- ترکیبی‌کدشده-غیرکدشده‌مستقل[۱۹](IH) : این روش، روش ترکیبی پیشنهادی در [۱۲] است که برای حالت توزیع همگن محتواها پیشنهاد شده است. همان‌طور که قبلاً توضیح داده شده، در این روش حافظه ذخیره‌ساز موقت SBS ها به دو بخش و محتواها به سه بخش تقسیم می‌گردند. این روش نیز مانند روش کدشده محض برای انتخاب سیاست ذخیره‌سازی موقت فقط به محبوبیت سراسری محتواها توجه می‌کند. سیاست نزدیک به بهینه این روش در پیچیدگی زمانی چندجمله‌ای قابل محاسبه است. در زمانی که محبوبیت محتواها همگن است، این روش بر دو روش بالا (DPU و PC) برتری دارد.

درحالی‌که DPU فقط به محبوبیت محلی برای جانمایی محتواها توجه دارد، PC و IH فقط به محبوبیت سراسری توجه دارند. در الگوریتم DH با قرار دادن مقدار $M_1= M$، نتیجه الگوریتم همانند نتیجه روش DPU و با قرار دادن مقدار $M_1 = 0$، نتیجه الگوریتم همانند نتیجه روش PC می‌شود. همچنین اگر به‌جای روش DH در الگوریتم‌های CDH و SCDH روش‌های دیگری مانند PC یا IH فراخوانی گردند، آنگاه نسخه‌های دیگری از روش‌های مکاشفه‌ای به وجود خواهند آمد. به طور مثال روش *SCPC* نسخه‌ی کدشده محض از SCDH است که به جای DH از PC در الگوریتم‌های CDH و SCDH استفاده می‌کند.

۵-۱- نحوه ارزیابی و شبیه‌سازی

برای ارزیابی و شبیه‌سازی روش‌های پیشنهادی برای ذخیره‌سازی موقت در حالت محبوبیت ناهمگن محتواها، فرض شده است که محتواها در چهار دسته قرار دارند (به طور مثال ورزشی، فیلم، اخبار و موسیقی) به‌طوری‌که تعداد محتوای دسته‌ها باهم برابر است. از کل محتواهای هر دسته ۲۰ درصد محتواها برای تمامی کاربران (در تمامی SBS ها) محبوب هستند (یعنی می‌توانند محبوب سراسری باشند) و سایر ۸۰ درصد محتوای دیگر هر دسته فقط برای کاربران در محدوده برخی از SBS ها محبوب هستند (یعنی می‌توانند محبوب محلی باشند). برای هر SBS یکی از دسته‌ها به‌عنوان دسته محبوب محلی آن SBS انتخاب می‌شود که کاربران آن SBS به آن دسته علاقه بیشتری دارند. لذا ممکن است چند SBS دسته محبوب محلی یکسان داشته باشند.

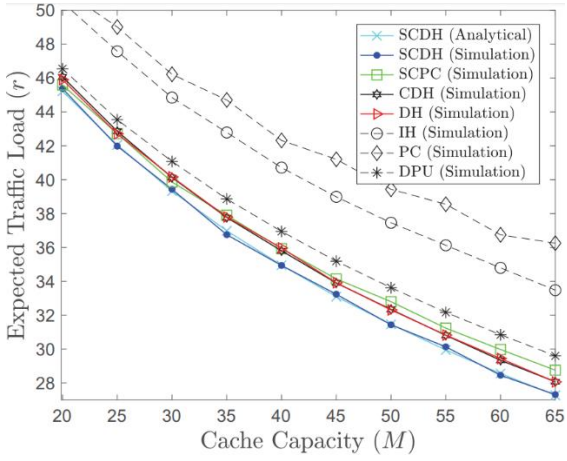
کاربران هر SBS می‌توانند محتواهای محبوب سراسری (که مربوط به همه دسته‌ها هستند)، محتواهای مربوط به دسته محبوب محلی خود و یا سایر محتواهای سایر دسته‌ها را درخواست کنند. ولی با توجه به اینکه فرض کردیم از چهار دست، یک دسته محبوب محلی SBS است، تمام محتواهایی که کاربر از سایر دسته‌ها درخواست می‌کند را محدود به محتواهای محبوب سراسری می‌کنیم. بنابراین، در حقیقت پنج کلاس محتوا در شبیه‌سازی‌ها داریم، یک کلاس محبوب سراسری و چهار کلاس محبوب محلی.

همچنین پارامتری با نام ρ_L تعریف می‌نماییم که درصد درخواست‌هایی که مربوط به کلاس‌های محلی هستند را تعیین می‌کند. به عبارتی دیگر از کل درخواست‌های هر SBS مقدار ρ_L درصد مربوط به کلاس محبوب محلی و $100-\rho_L$ درصد مربوط به کلاس محبوب سراسری خواهند بود. بدون اینکه خللی بر کلیت مسئله وارد شود، توزیع محبوبیت محتواها را روی این پنج کلاس مستقل از هم، یکسان و بر اساس توزیع Zipf با پارامتر α در نظر می‌گیریم. نتایج گزارش شده مربوط به میانگین ۲۰۰۰ بار اجرای فاز تحویل محتوا است. در این بخش به جز مواردی که به‌صراحت در شکل‌ها مشخص می‌گردد، تعداد کل محتواها (N) برابر 1000، تعداد SBS ها (K) برابر 10، پارامتر β برای روش‌های SCDH و SCPC برابر ${}^{1}/_{M}$ و مقدار $\rho_L = 50$ قرار داده شده است.

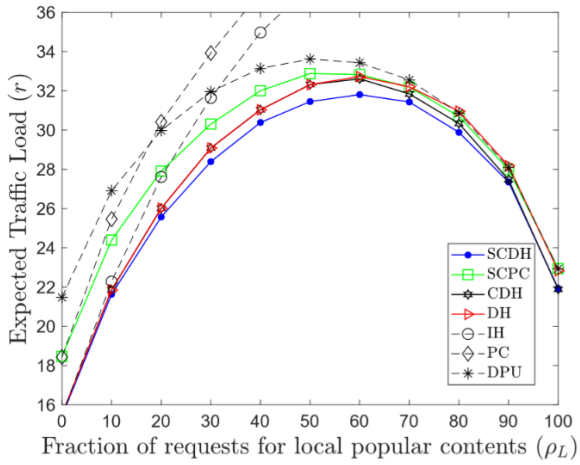
۵-۲- نتایج تحلیل ریاضی و شبیه‌سازی

شکل‌های ۳الف و ۳ب نرخ ترافیک روی رسانه مشترک را به ترتیب به‌ازای حجم مختلف حافظه ذخیره‌سازهای موقت و میزان درصد درخواست‌ها برای محتواهای محبوب محلی (ρ_L) نشان می‌دهند. همچنین جدول ۲ تحلیل بازه اطمینان ۹۵ و ۹۹ درصد را برای شبیه‌سازی‌های انجام شده در حالتی‌که M=50، $\rho_L = 50$ و α=1 است را نشان می‌دهد. همان‌طور که در این شکل‌ها دیده می‌شود، استفاده از روش مکاشفه‌ای SCDH پیشنهادی کمترین نرخ ترافیک روی رسانه مشترک را در بین روش‌های مورد ارزیابی داشته است. همچنین همان‌طور که در جدول ۲ دیده می‌شود، با توجه به اینکه نتایج گزارش شده میانگین ۲۰۰۰ بار اجرای فاز تحویل محتوا است، فاصله‌های اطمینان کوچک هستند. بررسی آزمون‌های آماری نیز نشان داد که اختلاف SCDH با سایر روش‌ها کاملاً معنی‌دار است.

در شکل ۳ب وقتی $\rho_L = 0$ است، تمامی درخواست‌ها برای محتواهای محبوب سراسری است (یعنی معادل حالت همگن بودن محبوبیت محتوا). در این حالت سود سراسری ذخیره‌سازی موقت حداکثر است و همان‌طور که دیده می‌شود تمامی روش‌های ترکیبی پیشنهادی (حتی روش IH) کمترین و روش غیرکدشده (DPU) بیشترین نرخ ترافیک روی رسانه مشترک را دارند. با افزایش مقدار ρ_L به ۵۰ درصد، به دلیل متنوع شدن محتواهای درخواست شده، نرخ ترافیک روی رسانه مشترک نیز افزایش می‌یابد. همان‌طور که در شکل دیده می‌شود با افزایش مقدار ρ_L کارایی روش‌هایی که فقط به محبوبیت سراسری توجه می‌کنند (یعنی روش‌های PC و IH) نسبت به سایر روش‌ها کاهش می‌یابد. به‌طوری‌که برای ρ_L بزرگ‌تر از ۳۰ درصد نرخ ترافیک روی رسانه مشترک برای روش‌های PC و IH از روش غیرکدشده DPU نیز بیشتر می‌شود.



(الف)



(ب)

شکل ۳: نرخ ترافیک روی رسانه مشترک زمانی‌که N= 1000، K =10، $\{Z_c\}_{c=1}^{10} = \mathbf{10}$، و α=۱ به‌ازای (الف) حجم ذخیره‌سازهای موقت SBS ها (M)، (ب) درصد درخواست‌ها برای محتواهای محبوب محلی ($\boldsymbol{\rho}_L$) برای M= 50

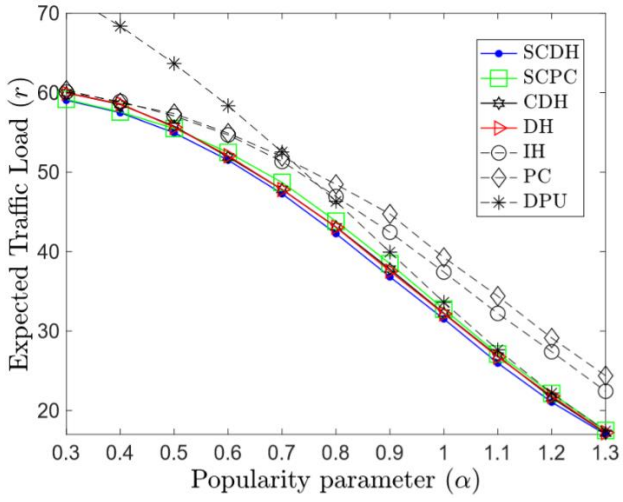
جدول ۲: بازه اطمینان ۹۵ و ۹۹ درصد برای نرخ بار ترافیک روی رسانه مشترک
(N=1000, M=50, $\{Z_c\}_{c=1}^{10} = \mathbf{10}$, K=10, α=1, $\boldsymbol{\rho}_L$=50)

روش	میانگین	انحراف معیار	بازه اطمینان (% ۹۵)	بازه اطمینان (% ۹۹)
SCDH	31.45	4.11	31.45 ∓ 0.151	31.45 ∓ 0.214
SCPC	32.70	4.28	32.70 ∓ 0.157	32.70 ∓ 0.223
CDH	32.28	4.32	32.28 ∓ 0.159	32.28 ∓ 0.225
DH	32.3	4.17	32.3 ∓ 0.153	32.3 ∓ 0.217
IH	37.25	3.81	37.25 ∓ 0.140	37.25 ∓ 0.198
PC	39.34	4.16	39.34 ∓ 0.153	39.34 ∓ 0.216
DPU	33.54	4.55	33.54 ∓ 0.167	33.54 ∓ 0.237

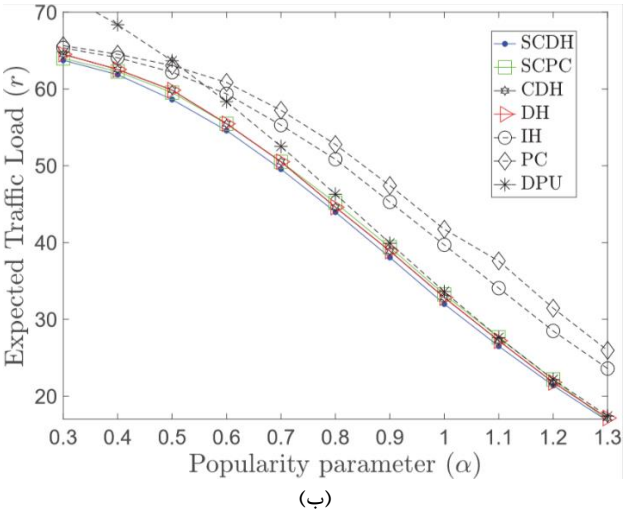
پس از افزایش مقدار ρ_L به بیش از ۵۰ درصد، هر چه مقدار ρ_L بیشتر می‌شود، نرخ ترافیک روی رسانه مشترک کاهش می‌یابد. علت این موضوع این است که در این حالت با افزایش مقدار ρ_L محتواهای محبوب محلی بیشتر از قبل درخواست می‌شوند و به‌تدریج روش‌های ذخیره‌سازی موقت به ذخیره‌سازی محتواهای محبوب محلی تمایل پیدا کرده، در نتیجه از تنوع محتواهای درخواست شده از MBS کاسته می‌شود. با این حال همان‌طور که در شکل دیده می‌شود این کاهش در نمودار متقارن نیست و به مقداری که با افزایش ρ_L از ۰ تا ۵۰ درصد مقدار نرخ ترافیک روی رسانه مشترک افزایش یافته بود با افزایش ρ_L از ۵۰ به ۱۰۰ درصد، مقدار این نرخ کاهش نمی‌یابد. علت این موضوع این است که پس از افزایش ρ_L به مقادیر بیشتر از ۵۰ درصد، کاربران در محدوده SBS های مختلف، محتواهای محلی موردعلاقه خود که ممکن است با سایر SBS ها متفاوت باشد را درخواست خواهند کرد. بنابراین، گرچه هر چه ρ_L از ۵۰ درصد فراتر رود تنوع محتواهای درخواست شده از MBS کمتر می‌شود ولی به نسبت $\rho_L - 50$ درصد این محتواها متنوع‌تر (از دسته‌های مختلف) هستند. لذا نرخ سود سراسری ذخیره‌سازی موقت برای ρ_L های بیشتر از ۵۰ درصد، کمتر از نرخ سود سراسری برای $\rho_L - 50$ درصد است.

زمانی که ρ_L برابر ۱۰۰ درصد است، در هر SBS فقط محتواهای محبوب محلی آن SBS درخواست می‌شوند. همان‌طور که در شکل دیده می‌شود در این حالت نیز روش‌های ترکیبی پیشنهادی SCDH و CDH کارایی بهتری نسبت به روش‌های دیگر دارند. همان‌طور که در شکل دیده می‌شود وقتی ρ_L کمتر از ۶۰ درصد است، کارایی روش‌های DH و CDH تقریباً مثل هم است. ولی وقتی ρ_L به بیش از ۶۰ درصد افزایش می‌یابد، روش مبتنی بر خوشه‌بندی CDH کارایی بهتری نسبت به نسخه مشابه خود که از خوشه‌بندی استفاده نمی‌کند (یعنی DH) دارد.

شکل ۴الف و شکل ۴ب نرخ ترافیک روی رسانه مشترک به‌ازای پارامتر توزیع محبوبیت (α) را به ترتیب برای حالت پراکندگی یکنواخت و یک حالت خاص پراکندگی غیریکنواخت کاربران در SBS ها را نشان می‌دهند. در شکل ۴ب به ترتیب تعداد ۱، ۳، ...، ۱۹ کاربر در محدوده SBS های ۱، ۲، ...، ۱۰ وجود دارد.



(الف)



(ب)

شکل ۴: نرخ ترافیک روی رسانه مشترک به‌ازای پارامتر توزیع محبوبیت (α) برای M=50 و (الف) پراکندگی ۱۰۰ کاربر به‌صورت یکنواخت $\{Z_c\}_{c=1}^{10} = 10$، (ب) پراکندگی غیریکنواخت کاربران $\{Z_c\}_{c=1}^{10} = [1,3,5,7,9,11,13,15,17,19]$.

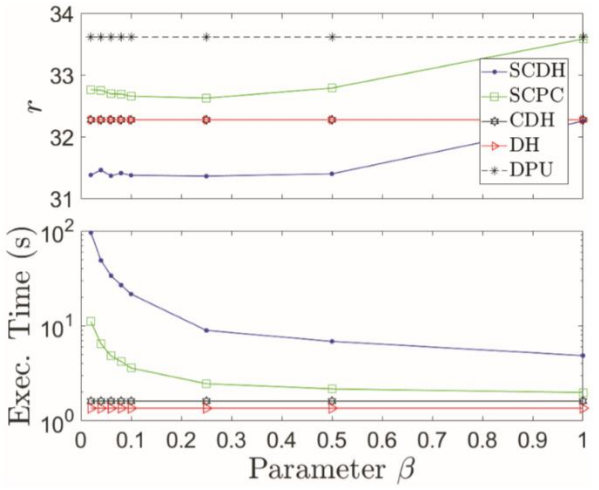
همان‌طور که در شکل دیده می‌شود، گرچه با نامتعادل شدن پراکندگی کاربران بین SBS ها، کارایی روش‌های کدشده کاهش می‌یابد، ولی روش‌های مکاشفه‌ای پیشنهادی در این مقاله برای تمامی مقادیر α کارایی بهتری نسبت به روش‌های پایه دارند. نکته جالب توجه دیگر این است که دیده می‌شود که روش غیرکدشده DPU گرچه به محبوبیت محلی توجه دارد و در شکل ۴ نیز ۵۰ درصد درخواست‌ها مربوط به محتواهای محبوب محلی است، ولی کارایی آن برای مقادیر α کمتر از ۰٫۷ در شکل ۴الف و کمتر از ۰٫۶در شکل ۴ب حتی از روش‌های PC و IH که توجهی به محبوبیت محلی ندارند نیز کمتر است. البته علت این امر، افزایش مزیت روش‌های کدشده برای α کوچک‌تر نسبت به روش‌های غیرکدشده است.

شکل ۵الف نرخ ترافیک روی رسانه مشترک و مدت‌زمان لازم برای اجرا روش‌های SCDH و SCPC پیشنهادی را به‌ازای مقادیر مختلف پارامتر β نشان می‌دهد. انتخاب بهترین مقدار برای پارامتر β به چگونگی توزیع محبوبیت محتوا بستگی دارد. با این وجود، همان‌طور که در شکل ۵الف دیده می‌شود، یک سیاست مناسب ذخیره‌سازی موقت با انتخاب مقدار بزرگ‌تر برای پارامتر β قابل دستیابی است. با انتخاب مقادیر بزرگ‌تر برای پارامتر β نیاز به جستجوی تعداد حالت‌های کمتری از تقسیم حافظه ذخیره‌سازهای موقت به دو قسمت محلی و سراسری داریم و لذا سربار پردازشی کمتری خواهیم داشت.

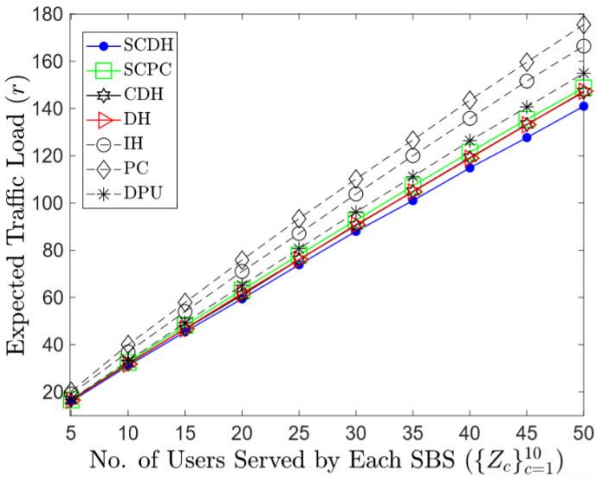
شکل ۵ب مقایسه کارایی روش‌های مکاشفه‌ای پیشنهادی با روش‌های پایه را برای یک مقیاس بزرگ‌تر سیستم و به‌ازای مقادیر مختلف تعداد کاربران در محدوده هر SBS نشان می‌دهد. در این شکل برای مقایسه راحت‌تر اثر تعداد کاربران فرض شده که تعداد کاربران در محدوده همه SBS ها یکسان است. در نمودارهای شکل ۵ب به‌وضوح دیده می‌شود که روش‌های مکاشفه‌ای پیشنهادی که هم به محبوبیت محلی و هم به محبوبیت سراسری توجه دارند، کارایی بهتری نسبت به روش‌های پایه دارند. در واقع، در بین روش‌های پیشنهادی SCDH کمترین نرخ ترافیک روی رسانه مشترک را دارد، ولی این روش هزینه پردازشی بالاتری نیاز دارد. از طرفی، روش‌های DH و CDH با هزینه پردازشی مناسب، کارایی مناسبی داشته و در مقایسه با روش‌های پایه و SCPC در بیشتر مواقع نرخ کمتری روی رسانه مشترک دارند. همچنین دیده می‌شود که اعمال ایده‌های خوشه‌بندی و دو قسمت کردن فضای ذخیره‌سازهای موقت به دو قسمت محلی و سراسری روی روش ذخیره‌سازی موقت کدشده (یعنی اعمال این ایده‌ها روی PC و ایجاد SCPC)، کارایی این روش را بهبود داده است. با این وجود نرخ ترافیک روی رسانه مشترک با استفاده از روش SCPC به میزان معنی‌داری بیشتر از روش‌های مکاشفه‌ای ترکیبی پیشنهادی (SCDH و CDH) است.

۶- جمع بندی و کارهای آتی

در این مقاله مسئله ذخیره‌سازی موقت در شبکه‌های دارای رسانه مشترک در شرایط توزیع ناهمگن محبوبیت مورد توجه قرار گرفت. نشان داده شد که در حالت ناهمگن بودن محبوبیت یافتن جواب بهینه رام نشدنی و از پیچیدگی ترکیبیاتی است. لذا برای این حالت، چندین روش مکاشفه‌ای ارائه نمودیم که مبتنی بر انجام مصالحه بین محبوبیت محلی و محبوبیت سراسری و انجام خوشه‌بندی روی SBS ها هستند. روش‌های اکتشافی پیشنهادی در این مقاله از طریق تحلیل عددی و شبیه‌سازی مورد ارزیابی و مقایسه با روش‌های قبلی قرار گرفت. ارزیابی‌ها با انجام آزمایش‌های متعدد برای شرایط مختلف محبوبیت محتوا، اندازه حجم ذخیره‌سازهای موقت، مقیاس سیستم و چگونگی پراکندگی کاربران انجام گردید. نتایج ارزیابی‌ها نشان داد که در نظر گرفتن محبوبیت محلی در کنار محبوبیت سراسری می‌تواند کارایی ذخیره‌سازی موقت را به میزان قابل‌توجهی بهبود دهد. همچنین دیده شد که روش‌های پیشنهادی در این مقاله کارایی بهتری نسبت به روش‌های پایه قبلی دارند. البته دیده شد که گرچه پیچیدگی زمانی روش‌های مکاشفه‌ای ارائه شده در این مقاله از مرتبه چندجمله‌ای است، ولی این روش‌ها هزینه پردازشی دارند. لذا برای پژوهش‌های آتی برای ذخیره‌سازی موقت (کدشده و ترکیبی) در شرایط ناهمگن بودن محبوبیت محتواها می‌توان استفاده از روش‌های یادگیری ماشین را مورد بررسی قرار داد.



(الف)



(ب)

شکل ۵: نرخ ترافیک روی رسانه مشترک (برای α=1) به‌ازای (الف) پارامتر β در کنار زمان محاسبات مربوطه برای N=1000, M=50 و $\{Z_c\}_{c=1}^{10} = 10$ ، (ب)تعداد کاربران در محدوده هر SBS برای N= 2500, M=100.

۷- مراجع

[1] V. Chandrasekhar, J. G. Andrews and A. Gatherer, "Femtocell networks: a survey," *IEEE Communications magazine,* vol. 46, no. 9, 2008.

[2] K. Shanmugam, N. Golrezaei, A. G. Dimakis, A. F. Molisch and G. Caire, "Femtocaching: Wireless content delivery through distributed caching helpers," *IEEE Transactions on Information Theory,* vol. 59, p. 8402–8413, 2013.

[3] G. Paschos, E. Bastug, I. Land, G. Caire and M. Debbah, "Wireless caching: Technical misconceptions and business barriers," *IEEE Communications Magazine,* vol. 54, p. 16–22, 2016.

[4] Y. Chen, M. Ding, J. Li, Z. Lin, G. Mao and L. Hanzo, "Probabilistic small-cell caching: Performance analysis and optimization," *IEEE Transactions on Vehicular Technology,* vol. 66, p. 4341–4354, 2017.

[5] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Transactions on Information Theory,* vol. 60, no. 5, pp. 2856-2867, 2014.

[6] U. Niesen and M. A. Maddah-Ali, "Coded caching with nonuniform demands," *IEEE Transactions on Information Theory,* vol. 63, p. 1146–1158, 2017.

[7] J. Hachem, N. Karamchandani and S. N. Diggavi, "Coded Caching for Multi-level Popularity and Access," *IEEE Transactions on Information Theory,* vol. 63, pp. 3108-3141, May 2017.

[8] T. Li, M. Ashraphijuo, X. Wang and P. Fan, "Traffic Off-Loading With Energy-Harvesting Small Cells and Coded Content Caching," *IEEE Transactions on Communications,* vol. 65, pp. 906-917, February 2017.

[9] M. Ji, A. M. Tulino, J. Llorca and G. Caire, "Order-optimal rate of caching and coded multicasting with random demands," *IEEE Transactions on Information Theory,* vol. 63, p. 3923–3949, 2017.

[10] J. Zhang, X. Lin and X. Wang, "Coded caching under arbitrary popularity distributions," *IEEE Transactions on Information Theory,* vol. 64, p. 349–366, 2018.

[11] S. A. Saberali, L. Lampe and I. F. Blake, "Full Characterization of Optimal Uncoded Placement for the Structured Clique Cover Delivery of Nonuniform Demands," *IEEE Transactions on Information Theory,* vol. 66, pp. 633-648, January 2020.

[12] A. G. Sheshjavani, A. Khonsari, S. P. Shariatpanahi, M. Moradian and A. Dadlani, "Coded Caching Under Non-Uniform Content Popularity Distributions with Multiple Requests," in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, 2020.

[13] A. G. Sheshjavani, A. Khonsari, S. P. Shariatpanahi and M. Moradian, "Content caching for shared medium networks under heterogeneous users' behaviors," *Computer Networks,* vol. 199, p. 108454, 2021.

[14] G. Ma, Z. Wang, M. Zhang, J. Ye, M. Chen and W. Zhu, "Understanding Performance of Edge Content Caching for Mobile Video Streaming," *IEEE Journal on Selected Areas in Communications,* vol. 35, pp. 1076-1089, 2017.

[15] Y. Lu, W. Chen and H. V. Poor, "Coded Caching Under Heterogeneous User Preferences: An Effective Throughput Perspective," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, 2019.

[16] C. Zhang, S. Wang, V. Aggarwal and B. Peleato, "Coded Caching with Heterogeneous User Profiles," *IEEE Transactions on Information Theory,* pp. 1-1, 2022.

[17] C.-H. Chang, B. Peleato and C.-C. Wang, "Coded Caching with Full Heterogeneity: Exact Capacity of The Two-User/Two-File Case," *IEEE Transactions on Information Theory,* pp. 1-1, 2022.

[18] B. Chen and C. Yang, "Caching Policy Optimization for D2D Communications by Learning User Preference," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, 2017.

[19] Y. Jiang, M. Ma, M. Bennis, F. Zheng and X. You, "User Preference Learning-Based Edge Caching for Fog Radio Access Network," *IEEE Transactions on Communications,* vol. 67, pp. 1268-1283, February 2019.

[20] J. Song, M. Sheng, T. Q. S. Quek, C. Xu and X. Wang, "Learning-based content caching and sharing for wireless networks," *IEEE Transactions on Communications,* vol. 65, p. 4309–4324, 2017.

[21] Z. Chang, L. Lei, Z. Zhou, S. Mao and T. Ristaniemi, "Learn to Cache: Machine Learning for Network Edge Caching in the Big Data Era," *IEEE Wireless Communications,* vol. 25, pp. 28-35, June 2018.

[22] A. Sadeghi, F. Sheikholeslami and G. B. Giannakis, "Optimal and Scalable Caching for 5G Using Reinforcement Learning of Space-Time Popularities," *IEEE Journal of Selected Topics in Signal Processing,* vol. 12, pp. 180-190, February 2018.

[23] H. Pang, J. Liu, X. Fan and L. Sun, "Toward Smart and Cooperative Edge Caching for 5G Networks: A Deep Learning Based Approach," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, 2018.

[24] E. Baştuğ, M. Bennis and M. Debbah, "A transfer learning approach for cache-enabled wireless networks," in *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2015 13th International Symposium on*, 2015.

[25] M. Leconte, G. Paschos, L. Gkatzikis, M. Draief, S. Vassilaras and S. Chouvardas, "Placing dynamic content in caches with small population," in *INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications, IEEE*, 2016.

[26] C. Zhang and B. Peleato, "On the Average Rate for Coded Caching with Heterogeneous User Profiles," in *Proc. IEEE International Conference on Communications (ICC)*, 2020.

عبدالله غفاری ششجوانی تحصیلات خود را در مقاطع کارشناسی و کارشناسی ارشد کامپیوتر گرایش نرم افزار به ترتیب در سال‌های ۱۳۸۷ و ۱۳۹۱ از دانشگاه‌های علوم و فنون هوایی شهید ستاری و تربیت مدرس به پایان رسانده است. ایشان در حال حاضر دانشجوی دکتری کامپیوتر گرایش نرم افزار در دانشگاه تهران است. زمینه‌های تحقیقاتی موردعلاقه ایشان عبارت‌اند از: ذخیره‌سازی موقت در شبکه‌های مخابراتی، شبکه‌های تحویل محتوا، شبکه‌های نظیر به نظیر، شبیه‌سازی و ارزیابی شبکه‌های کامپیوتری. آدرس پست الکترونیکی ایشان عبارت است از:

abdollah.ghaffari@ut.ac.ir

احمد خونساری تحصیلات خود را در مقطع کارشناسی مهندسی برق و کامپیوتر در سال۱۳۷۰ از دانشگاه شهید بهشتی، در مقطع کارشناسی ارشد مهندسی کامپیوتر در سال ۱۳۷۵ از دانشگاه علم و صنعت و در مقطع دکتری علوم کامپیوتر در سال ۲۰۰۳ از دانشگاه گلاسکو انگلستان به پایان رسانده است. وی در حال حاضر دانشیار گروه مهندسی برق و کامپیوتر دانشگاه تهران، و پژوهشگر علوم کامپیوتر در پژوهشگاه دانش‌های بنیادی ایران (IPM) است . علایق پژوهشی وی شامل شبیه سازی و تجزیه و تحلیل داده ها، مدل سازی/ارزیابی عملکرد، شبکه های سیمی/بی سیم، شبکه های لمسی، سیستم های ابری و توزیع شده، پردازش اطلاعات کوانتومی و معماری های کامپیوتری با کارایی بالا می باشد. آدرس پست الکترونیکی ایشان عبارت است از:

a_khonsari@ut.ac.ir و ak@ipm.ir

سید پویا شریعت پناهی مدارک کارشناسی، کارشناسی ارشد و دکتری را در رشته مهندسی برق از دانشگاه صنعتی شریف، تهران، ایران، به ترتیب در سال های ۱۳۸۵، ۱۳۸۷ و ۱۳۹۲ اخذ نموده است. ایشان در حال حاضر استادیار دانشکده مهندسی برق و کامپیوتر دانشگاه تهران، ایران است. وی پیش از ورود به دانشگاه تهران، پژوهشگر پژوهشگاه دانش‌های بنیادی (IPM) ایران بوده است. زمینه های تحقیقاتی ایشان شامل نظریه اطلاعات، علوم شبکه، ارتباطات بی سیم و سیستم های پیچیده است. وی در سال ۱۳۸۰ برنده مدال طلای المپیاد ملی فیزیک ایران شد. آدرس پست الکترونیکی ایشان عبارت است از:

p.shariatpanahi@ut.ac.ir

معصومه مرادیان مدارک کارشناسی، کارشناسی ارشد و دکتری را در رشته مهندسی برق از دانشگاه صنعتی شریف، تهران، ایران، به ترتیب در سال های ۱۳۸۵، ۱۳۸۷ و ۱۳۹۲ اخذ نموده است. وی در سال ۲۰۱۵ دانشجوی مدعو دانشگاه چینی هنگ-کنگ بود. او در حال حاضر پژوهشگر پسا دکتری علوم کامپیوتر در پژوهشگاه دانش‌های بنیادی (IPM) ایران است. علایق پژوهشی ایشان شامل شبکه های ارتباطی برداشت انرژی، تئوری صف و بهینه سازی تصادفی شبکه است. آدرس پست الکترونیکی ایشان عبارت است از:

mmoradian@ipm.ir

[1] Small Base Stations (SBSs)
[2] Local cache gain
[3] Global cache gain
[4] Macro Base Station (MBS)
[5] Diversity principle
[6] Popularity principle
[7] Device to Device
[8] probabilistic Latent Semantic Analysis
[9] Expectation Maximization
[10] Sigmoid
[11] Transfer Learning

[12] Multi-Armed Bandit
[13] clique cover
[14] Dependent Hybrid (DH)
[15] Clustered Dependent Hybrid (CDH)
[16] Split Clustered Dependent Hybrid (SCDH)
[17] Dependent pure uncoded (DPU)
[18] Pure Coded (PC)
[19] Independent Hybrid coded-uncoded (IH)

The CSI Journal on Computing Science and Information Technology

The CSI Journal on
Computing Science and
Information Technology
Vol. 20, No.1, 2022

Content Caching in Shared Medium Networks with Non-Uniform and User-Dependent Demands

Abdollah Ghaffari Sheshjavani[1], Ahmad Khonsari[1,2], Seyed Pooya Shariatpanahi[1], and Masoumeh Moradian[2]

[1]School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran
[2]School of Computer Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

Abstract

Caching is one of the attractive solutions to reduce the cost and delay of accessing the desired data in various applications such as processors and communication networks. In this paper, we investigate the content caching problem under non-uniform and user-dependent demands. Finding the optimal solution in the case of user-dependent demands, in which the popularity of content is different in different places (heterogeneity of popularity), is an intractable and non-trivial combinatorial problem. Therefore, in this paper, we propose several heuristic methods by leveraging hybrid coded-uncoded caching, clustering, and the trade-off between the local and global popularity of contents.

Finally, the proposed methods have been evaluated and compared with previous methods through numerical analysis and simulation. The results show that the traffic rate on the shared medium using the proposed heuristic methods based on the hybrid caching scheme is significantly lower than the previous methods, including the baseline pure coded and pure uncoded schemes.

Keywords: Content caching, shared medium, non-uniform and heterogeneous popularity, the hybrid coded-uncoded caching