بیست و دومین کنفرانس ملی سالانه انجمن کامپیوتر ایران
21 الی 23 اسفند 1395

جریان‌سازی ویدیو در شبکه‌های نظیر به نظیر با رویکرد یادگیری تقویتی در انتخاب گره دستیار

مصیب حاجی‌مقصودی[1]، محسن سودخواه محمدی[2]، وصال حکمی[3]، مهدی دهقان تخت‌فولادی[4]

[1] کارشناسی ارشد، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر ، تهران
hajimaghsoodi.m@aut.ac.ir

[2] کارشناسی ارشد، دانشکده ریاضی و علوم کامپیوتر، دانشگاه صنعتی امیرکبیر، تهران
mohsensoodkhah@gmail.com

[3] استادیار، گروه شبکه‌های کامپیوتری، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران
vhakami@iust.ac.ir

[4] استاد، گروه شبکه‌های کامپیوتری، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر، تهران
dehghan@aut.ac.ir

چکیده

با افزایش میزان تقاضای محتوای چندرسانه‌ای، گونه جدیدی از شبکه‌های توزیع محتوا به نام شبکه جریان‌سازی ویدئو نظیربه‌نظیر چندکاناله مطرح شده است. در این شبکه‌ها، گره‌های دستیار، از یک سو میزان ارجاعات نظیرها را به سرور کاهش می‌دهند. از سوی دیگر، به‌عنوان ریز-سرورهای تأمین‌کننده محتوا، نقاط اتصال جدیدی را جهت دریافت محتوای مورد نظر هر نظیر، در اختیار آنها قرار می‌دهند. از طرفی، برای تضمین توازن بار، نیاز به هماهنگ‌سازی میان تصمیمات نظیرها بر سر پیوستن به گره‌های دستیار است. به دلیل نبود نظارت مرکزی در شبکه‌های واقعی، این هماهنگ‌سازی باید به صورت غیرمتمرکز توسط نظیرها و با اتکا به اطلاعات محلی-شان محقق گردد.

در این مقاله، مسألهٔ انتخاب دستیار برای دریافت کانال خاص در یک سیستم نظیر-به-نظیر با جمعیت بالا، به عنوان بازی تکاملی مدل شده است که نقطهٔ تعادل آن به منزلهٔ برقراری هماهنگی میان انتخاب‌های نظیرهاست. تعادل بازی فرمول‌بندی شده، نقطهٔ ثابت معادلهٔ معروف پویای همانندساز است که ما برای محاسبهٔ آن، نظیرها را مجهز به یک الگوریتم یادگیری غیرمتمرکز می‌نماییم. الگوریتم پیشنهادی قابلیت همگرایی برخط به تعادل بازی را با مفروضات اطلاعاتی پایین فراهم می‌آورد. نتایج شبیه‌سازی حاکی از کارایی الگوریتم پیشنهادی در ایجاد هماهنگی میان نظیرها و برقرای توازن بار در سیستم جریان‌سازی است.

کلمات کلیدی

جریان‌سازی ویدیو، شبکه‌های نظیربه نظیر، گره‌دستیار، نظریه بازی، یادگیری نظیرها، تعادل در بازی.

1- مقدمه

با افزایش میزان تقاضا برای محتوای چندرسانه‌ای باکیفیت بالا، گونه جدیدی از شبکه‌های توزیع محتوا به نام شبکه جریان‌سازی ویدئو نظیربه‌نظیر[1] چندکاناله مطرح شده است. سیستم‌های جریان‌سازی ویدئوی نظیربه‌نظیر متشکل از چندین کانال ویدئویی، سیستم‌های چندکاناله نامیده می‌شوند. [1]. در این مدل، کاربران که نظیر نامیده می‌شوند، نه تنها مصرف‌کننده محتوا هستند، بلکه به‌عنوان سرویس‌دهنده به سایر نظیرها نیز عمل می‌کنند. [2].

اگرچه برخی سیستم های P2P VoD موفق و مقیاس‌پذیر در طول چند سال گذشته پیشنهادشده‌اند، عناصر مهم دیگری از کیفیت خدمت در بسیاری از این روش‌هانادیده‌گرفته‌شده‌اند. در سیستم‌های تجاری جریان‌سازی ویدئو، راهکار نظیربه‌نظیر خالص کافی به نظر نمی‌رسد زیرا فاقد ویژگی‌های مهمی همچون تضمین کیفیت سرویس برای کاربران، امنیت، و کنترل توسط تأمین‌کننده محتوا است. برای غلبه بر این محدودیت‌ها، معماری یاری شده–با–دستیار سرورهای محتوا را با معماری نظیربه‌نظیر ترکیب می‌کند[3]. به نظیرهایی که علاوه بر کانال در حال تماشا، به سایر کانال‌های ویدئویی هم کمک می‌کنند، «دستیار» گفته می‌شود. در این مدل، سرورهای محتوا منابع اولیه را برای شروع سرویس عرضه می‌کنند و همزمان، نظیرها منابع محلی خود شامل پهنای باند و فضای ذخیره‌سازی را در اختیار سایر نظیرها قرار می‌دهند [3].

مسئله انتخاب گره دستیار یک مسئله حیاتی و تأثیرگذار در سیستم‌های جریان‌ساز ویدیو به‌خصوص در شبکه‌های نظیر به نظیر می‌باشد و الگوریتم‌های شناخته‌شده و مطرحی در این زمینه ارائه‌شده‌اند[4 و5].

ساختار این شبکه ها به گونه ای است که نمی تواند نظارت مرکزی بر روی انتخاب یک دستیار در بین نظیرها داشته باشد. به دلیل نبود این نظارت، امکان وقوع پدیده پینگ‌پنگی وجود دارد که در آن تمامی نظیرها در هر به‌روزرسانی، دستیار یکسانی را انتخاب می‌کنند و آن را در دور بعدی، به‌عنوان بدترین دستیار از فهرست حذف می‌کنند. تداوم این روند نوسانی، سیستم جریان سازی را ناکارآمد کرده و در نهایت باعث اُفت کیفیت سرویس می‌گردد.

حال با در نظر گرفتن برخی از مسائل در این زمینه مانند ماهیت رقابتی بین نظیرها برای انتخاب پدر، میزان همکاری نظیرها و همچنین میزان تأثیر انتخاب آن‌ها بر روی کیفیت سایرین، می‌توان بهبود بهتری را در این نوع سیستم‌ها شاهد بود.

در سطح سرور، تخصیص پویای منابع خود سرور و منابع کمکی از محل دستیارها مطرح است. این کار بر اساس سیگنال پاداش دریافتی از سرور و بدون نیاز به تبادل پیام، میزان کمبود پهنای باند سرور را تأمین می‌کنند[3]. حال، در این شرایط باتوجه به اینکه میزان ظرفیت هر گره‌دستیار برای کانال‌های مختلف و همچنین میزان هزینه‌ای که هر نظیر باید برای مشاهده کانال خود، به آن گره دستیار بپردازد باهم متفاوت است، در نتیجه گره‌ها در یک شرایط رقابتی برای دستیابی به کانال مورد نظر خود باکیفیت بالاتر به رقابت می‌پردازند.

بیشتر کارهای انجام‌شده، رقابتی را که بین کاربران در انتخاب پدرشان رخ می‌دهد را تحلیل نمی‌کنند و ازدحام‌هایی که باعث کاهش کارایی در این سیستم‌ها می‌شود را نادیده می‌گیرند. در این مقاله، ما سعی داریم این مسئله (انتخاب گره دستیار مناسب) را به‌منظور رسیدن به تعادل بهره در کل سیستم حل کنیم تا گره‌های درخواست‌کننده، کمترین اطلاعات را از دستیارها اتخاذ کنند. این کار باعث می شود تا پهنای باند به صورت بهینه‌تری مصرف شود.

1-1- کارهای مرتبط

برای ساخت یک سیستم جریان‌سازی ویدیو در شبکه‌های نظیر به نظیر، قطعاً یک استراتژی انتخاب پدر (نظیری که کانال را از آن دریافت می‌کنیم) در کیفیت و کارآمد بودن این سیستم موثر است. در بسیاری از سیستم‌ها، این استراتژی انتخاب پدر، مبتنی بر موجود بودن تکه‌های کانال مورد نظر می‌باشد[5 و 4]. در سیستم‌های دیگر، انتخاب پدر مبتنی بر میزان پایداری و یا پهنای باند نظیرهای دیگر[6] میزان تأخیر در دریافت بسته‌ها[7] و یا میزان گم‌شدن بسته‌ها[8] است.

در مقاله[9] نویسنده یک شاخص برتری معرفی کرده است که از حاصل‌ضرب دو پارامتر دیگر به نام مدت‌زمان نشست و پهنای باند بارگذاری برای بدست‌آوردن این پارامتر استفاده کرده است که توسط یک مدل رگرسیون محاسبه می‌شود. همچنین در مقاله [10] یک مدل مبتنی بر خوشه‌سازی با رگرسیون خطی و شبکه عصبی، برای تعیین میزان تخصیص پهنای باند سرورها به‌صورت بهینه و همچنین تشخیص مشکلات سیستم، ارائه شده است. هردو کار بالا توسط مقاله [11] مورد آزمایش قرار گرفت و طبق نتایج آزمایش مشخص می‌شود که متدهای یادگیری، کیفیت جریان‌سازی را بهبود بخشیده‌اند. اگرچه این الگوریتم‌ها می‌توانند تا حدی برای شبکه‌ها نظیر به نظیر کارایی را فراهم کنند، لیکن اگر نظیرها، داده‌های گذشته خود در مورد انتخاب پدر را نگه‌داری و استفاده کنند، می‌تواند باعث افزایش کارایی در کل سیستم شود. جمع آوری اطلاعات محیطی باعث مصرف نمودن بخشی از پهنای باند می شود.

در مقاله [12] از یک روش انتخاب نظیر با استفاده از یادگیری تقویتی[2] برای سیستم‌های اشتراک‌گذاری فایل مانند Bittorent استفاده کرده است. این مدل ارائه‌شده در این مقاله یک بهبود قابل‌توجهی را برای سیستم‌های اشتراک‌گذاری فایل فراهم می‌کند، اما مدل یادگیری آن برای سیستم‌های جریان‌سازی ویدیو در شبکه‌های نظیر به نظیر نمی‌تواند کارآمد باشد. در مقاله [13] نیز یک مدل برای سیستم‌های جریان‌سازی ویدیو ارائه‌شده اما این مقاله به علت بالا بودن پیچیدگی، در شبکه‌هایی با مقیاس بزرگ و تعداد نظیرهای زیاد، نمی‌تواند کارآمد باشد.

در مقاله [14] یک چارچوب مدیریت و تخصیص منابع به‌منظور پشتیبانی از تجمیع چند سرور از نوع‌های با ظرفیت‌های متفاوت معرفی کرده است. همچنین در مقاله [15] یک شمای کلی از مسئله کنترل سطح دسترسی برای کاربران به‌منظور دسترسی به چند سرور متفاوت با تکنولوژی‌ها و ظرفیت‌های مختلف پیشنهاد داده‌شده است. لازم به ذکر است که هیچ یک از

کارهای مورد نظر، مسئله رقابت پویا بین انتخاب چندین شبکه را در نظر نگرفته‌اند .

در اینجا ما با استفاده از برخی از مکانیزم‌ها مانند استفاده از بازی‌های تکاملی پویا[16] و همچنین استفاده از مکانیزم‌های قیمت‌گذاری روی سرویس‌های مختلف به حل این مسئله با محیطی پویا و سرویس‌ها و تکنولوژی‌ها متفاوت با ظرفیت‌های گوناگون می‌پردازیم.

2- رویکرد ارائه شده

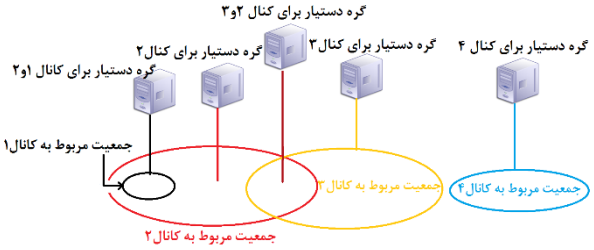
از آنجایی که انتخاب گره دستیار مناسب توسط هر نظیر بر روی سرعت، کیفیت و کارایی شبکه موثر می‌باشد؛ در این پژوهش بازی تکاملی پویا برای انتخاب گره دستیار مناسب، ارائه شده است. در نقطه‌ی تعادل این بازی، ضمانت می‌شود که تمامی افراد در یک گروه همگی میزان پرداختی یکسانی را از کیفیت کانال مدنظر خود دریافت می‌کنند.

نتیجه به دست‌آمده از بازی تکاملی پویا را با نتیجه خروجی یک الگوریتم تمرکز مورد مقایسه قرار می‌دهیم. اگرچه سرعت رسیدن به نتیجه مطلوب (رسیدن به تعادل تکاملی[۳]) در الگوریتم متمرکز سریع‌تر می‌باشد، لیکن در این الگوریتم به علت وجود یک سرور مرکزی برای جمع‌آوری، پردازش و در نهایت بازپخش و کنترل ارتباط نظیرها و اطلاعاتی از قبیل میزان ظرفیت هر یک از این نظیرها، سربار بسیار زیادی را شامل می‌شود. بنابراین، با استفاده از یک روال هماهنگ‌سازی توزیع‌شده مبتنی بر الگوریتم یادگیری چند-عاملهٔ Q-Learning، دیگر نیازی به یک زیرساخت مرکزی مانند یک سرور کنترل‌کننده مرکزی نیست[17].

2-1- فرضیات و مدل سیستم پیشنهادی

یک سیستم جریان‌سازی نظیر-به-نظیر با تعداد N نظیر را در نظر می‌گیریم که هر کدام از آن‌ها تمایل به تماشای یک کانال خاص دارند. مجموعهٔ کلیهٔ کانال‌های مورد نمایش در سیستم که نظیرها خواستار دریافت و تماشای آن‌ها می‌باشند را با نماد $\{a_1, a_2, \ldots, a_n\}$ نشان داده و نماد a_i نیز یک کانال خاص از این مجموعه را نشان می‌دهد. همچنین، مجموعهٔ گره‌های دستیار موجود در سیستم را با نماد $S = \{ s=h_i \mid h_i \in h_1, h_2, \ldots, h_n\}$ نشان می‌دهیم. این گره‌ها مسئول دریافت و بازپخش تمامی کانال‌های موجود برای نظیران می‌باشند. هر گره دستیار بسته‌به ظرفیتش، مسئول پخش یک و یا بیشتر از یک کانال است. هر یک از این نظیرها کانال مدنظر خود را از یک گره دستیار دریافت می‌کند. هر یک از گره‌های دستیار بسته به نیاز کاربران (نظیرهای دیگر) یک تعداد کانال را برایشان فراهم می‌کنند.

همان طور که در شکل(1) قابل مشاهده است، هر نظیر بسته به اینکه چه گره دستیارهایی، کانال مورد نظرش را داراست و می‌تواند از آن گره دستیار آن کانال را ملاحظه کند، در یکی از چند گروه موجود مشخص‌شده (مانند آنچه در شکل آمده) قرار خواهد گرفت. یعنی یک گروه (n_i) شامل تمام کسانی است که تمایل به دیدن یک کانال را دارند.



شکل (1) شکل نمای کلی سیستم

برای مثال، نظیرهایی که تمایل به دیدن (کانال 1) دارند تنها می‌توانند به گره دستیار اولی وصل شوند و کانال را از این طریق دریافت کنند و در گروه مربوط به کانال یک قرار خواهند گرفت. در مقابل کسانی که تمایل به تماشای (کانال 2) را دارند می‌توانند به هر یک از سه گره دستیاری که این کانال را ارائه می‌دهند متصل و برای دریافت کانال مربوطه اقدام کنند و در گروه مربوط به (کانال 2) قرار خواهند گرفت.

هرکدام از گره‌های دستیار، متناسب با ظرفیت پهنای باندی که در اختیاردارند و همچنین تعداد نظیرهایی که برای دریافت کانال‌های موجود بر روی آن گره دستیار، به آن متصل شده‌اند، می‌توانند یک سودمندی ($\pi_i^a(t)$) را برای نظیرها فراهم نمایند. به عبارت دیگر این مقدار مشخص‌کننده سود نظیرهایی است که در زمان t و در گروه a (این گروه شامل کسانی است که می‌خواهند کانال aرا مشاهده کنند) کانال i را انتخاب می‌کنند؛ در این سیستم گره‌های دستیار می‌توانند ظرفیت خود را به‌صورت اختصاص داده‌شده برای هر یک از کانال‌هایی که پشتیبانی می‌کنند در نظر بگیرند و حتی می‌توانند برای کیفیت‌های مختلف سرویس برای کاربران مختلف، ظرفیت‌های متفاوتی را اختصاص دهند. در این مقاله، فرض می‌کنیم که یک گره دستیار تمامی ظرفیت خود را به طور مساوی بین تمامی کانال‌هایی که مسئول پخش آن است و همچنین به طور یکسان بین تمامی نظیرهایی که به آن متصل هستند، پخش می‌کند.

فرض 1 (تابع سودمندی نظیر): میزان سود و یا تابع Net Utilituy برای هر نظیر در گروه a، که گره دستیار i را بعنوان دریافت کانال انتخاب کرده است از رابطه زیر به‌دست می‌آید:

$$U_i(n) = \left(\tau_i(n) - P_i(n)\right) \qquad (1)$$

که در آن n مجموع کل نظیرهایی است که در گروه a قرار دارند و گره دستیار i را انتخاب کرده‌اند. همچنین $P_i(n)$ یک تابع قیمت‌گذاری است که البته در این مسئله یک تابع خطی ثابت در نظر گرفته‌شده است. ما فرض می‌کنیم که تمامی نظیرهایی که یک گره دستیار را انتخاب می‌کنند، یک مقدار مساوی از پهنای باند را دریافت می‌کنند. بنابراین میزان Net Utility می‌توان از رابطه زیر به‌دست آورد:

$$U_i = \left(\frac{C_i^{(a)}}{\sum_{\acute{a}\in A^{(a)}} n_i^a}\right) - P_i * \qquad (2)$$

$\sum_{\acute{a}\in A^{(a)}} n_i^a$

که در آن n_i^a تعداد نظیرهای گروه a که گره دستیار i را انتخاب کرده‌اند، $C_i^{(a)}$ برابر با میزان ظرفیتی که گره i به گروه a اختصاص داده‌است، Pi یک ضریب خطی قیمت‌گذاری است که توسط گره دستیار i

برای دریافت شارژ از کاربران در نظر گرفته‌شده است و در نهایت، A(a) مجموعه‌ای از زیرگروه‌هایی که گروه a آنها را پوشش می‌دهد. همان طور که در رابطه(2) ملاحظه می‌کنید، با افزایش تعداد کاربران در یک گروه برای یک گره دستیار، میزان پرداختی که به هر یک از نظیرها می‌رسد به علت تقسیم میزان کل ظرفیت آن گره دستیار و همچنین افزایش قسمت مربوط به تابع قیمت‌گذاری، کم می‌شود. به همین علت میزان جابجایی و انتخاب یک نظیر در یک گروه بر روی گروه‌های دیگر نیز اثر می‌گذارد.

فرض 2 (میانگین پرداختی هر گره دستیار): خود گره‌های دستیار از رابطه زیر میزان میانگین پرداختی را محاسبه می‌کنند:

$$\bar{\pi}^{(a)} = \left(\frac{\left(\sum \pi_i^{(a)}\right)}{\left(N^{(a)}\right)}\right) \qquad (3)$$

که در آن $\bar{\pi}^{(a)}$ مشخص‌کننده میانگین سود جمعیت‌های دیگر است و $\sum \pi_i^{(a)}$ برابر سود بازیکنانی است که در گروه a (این گروه شامل کسانی است که می‌خواهند کانال a را مشاهده کنند) گره دستیار i را انتخاب می‌کنند که با تقسیم این عدد بر اندازه جمعیت بازی ($N^{(a)}$) میزان میانگین پرداختی هر گره دستیار حساب می‌شود.

فرض 3 (تابع قیمت گذاری): فرض می‌کنیم که گره‌های دستیار، از یک مدل خطی قیمت‌گذاری استفاده کنند که این مدل، یک تابع خطی از مجموع تعداد اتصالاتی از نظیرها در یک گروه است. برای هر یک از گره‌های دستیار، متناسب به شرایطی که در آن قرار دارند مانند نوع کانال و یا کانال‌هایی که توسط آن گره دستیار در حال پخش است و یا حتی شرایط فیزیکی خود گره دستیار مانند پایداری آن در سیستم، این ضریب و یا تابع قیمت‌گذاری برای گره دستیارهای متفاوت، می‌تواند باهم یکسان نباشد. زمانی که ازدحام رخ می‌دهد، گره‌های دستیار در آن گروه، این تابع را بالا می‌برند. بنابراین یک تابع غیر کاهشی، می‌تواند از ازدحام جلوگیری کند. زمانی که یک کاربر دریک گروه می‌تواند چندین گره دستیار را به‌عنوان پدر انتخاب کند، (مانند کسانی که در گروه مربوط به (کانال 2) قرار دارند و می‌توانند به یک از سه گره دستیاری که این کانال را بازپخش می‌کنند متصل شوند) گره دستیاری را انتخاب می‌کنند که بیشترین سود و یا همان میزان پرداختی را از طرف آن گره دستیار دریافت کنند که این مسئله به عواملی مانند ظرفیت گره دستیار، نوع آن و میزان تابع قیمت‌گذاری برای آن و همچنین نظیرهای دیگر که در این گروه و یا گروه‌های دیگر به این گره دستیار متصل هستند بستگی دارد.

میزان سود هر کاربر، توسط میزان ظرفیتی که کاربر از گره دستیار متصل‌شده به آن به‌دست می‌آورد و همچنین میزان هزینه‌ای که بابت این اتصال باید به آن گره دستیار، پرداخت کند($P_i(n)$)، اندازه‌گیری می‌شود. این مقدار هزینه‌ای که هر کاربر باید بابت این اتصال پرداخت کند رابطه مستقیم با تعداد کاربران متصل به آن گره دستیار دارد. همچنین با توجه به اینکه ظرفیت هر گره دستیار محدود است، زمانی که تعداد کاربران افزایش می‌یابد، میزان سود اختصاص داده‌شده به هر کاربر کم شده و در مقابل هزینه اتصال نیز بالا می‌رود. بنابراین، سود هر کاربر با افزایش تعداد کاربران دیگر،کم می‌شود[19].

2-2- فرموله کردن مسئله‌ی انتخاب پدر به شکل بازی‌های تکاملی

در این مقاله سعی داریم، انتخاب بهتر گره دستیار از طرف نظیرها را با استفاده از بازی‌های تکاملی پویا مدل کنیم. کاربران در گروه‌های متفاوت (کانال‌های مختلفی که توسط هر گره دستیار پخش می‌شود) برای دسترسی به کانال مربوطه خود باکیفیت بالاتر، تصمیم‌گیری می‌کنند. 4 عنصر اصلی برای مدل کردن این مسئله به شکل یک بازی تکاملی[16] به شرح زیر می‌باشند:

- ***بازیکنان*** : همه نظیرها (N)
- ***جمعیت*** : جمعیت در این بازی تکاملی، به هر یک از مجموعه‌ای از نظیرها در هر گروه اشاره دارد. (n_i^a)
- ***استراتژی*** : استراتژی برای هر نظیر متناسب است با انتخاب هر یک از گره‌های دستیاری که برای آن گروه سرویس ارائه می‌دهند یعنی کانال موردنظر آن نظیر را پخش می‌کنند.
- ***تابع سودمندی***: میزان سودمندی برای هر نظیر توسط تابع *Net* Utility تعیین می‌گردد($U_i(n)$).

لازم است دو مفهوم اصلی و تکنیکال در بحث بازی‌های تکاملی را در ابتدا بیان کنیم.

Replicator Dynamics : همان طور که گفته‌شده، ما نظیرها را به گروه‌های متفاوت تقسیم کردیم و مسئله اصلی را می‌توان در تعداد کاربرانی که به گره‌های دستیار مختلف برای دریافت ویدیوی خود متصل هستند، بیان کرد، برای این مسئله مشاهده تغییرات تعداد کاربران در هر گروه که به‌مرور زمان صورت می‌پذیرد امری لازم و ضروری است. بنابراین در اینجا ما از مفهوم replicator dynamics برای مدل کردن طریقه تکامل تغییر اندازه گروه‌هایی از نظیرها، استفاده می‌کنیم. در نظریه بازی‌ها، نظیرهای منطقی یقیناً اگر مشاهده کنند که مقدار سود آنها از میانگین سود نظیرهای دیگر در همان گروه کمتر باشد، استراتژی (گره دستیار) خود را عوض می‌کنند. در این جا لازم است که یادآور شویم که میزان سود تمامی کاربرانی که به یک گره دستیار متصل شده‌اند (بدون در نظر گرفتن گروهی که در آن هستند) باهم برابر است. یعنی

$$\pi ia = \pi i \qquad (4)$$

به طور کلی این مفهوم، نرخ تغییر استراتژی را مدل می‌کند که می‌تواند توسط رابطه زیر بیان شود:

$$\dot{x}_I(t) = x_I(t)\left(\pi_I(t) - \bar{\pi}(t)\right) \qquad (5)$$

تعادل تکاملی: این مفهوم به‌عنوان مجموعه‌ای از نقاط پایدار مفهوم قبلی تعریف می‌شود که در آن نقطه و یا نقاط، میزان سود تمامی کاربران باهم برابر است و در نتیجه هیچ کاربری انگیزه‌ای برای تغییر استراتژی خود ندارد. به عبارت دیگر اگر x_i نشان دهنده نسبت بازیکنانی که استراتژی i را انتخاب

کرده‌اند و از عبارت $\frac{n_i}{N}$ به‌دست می‌آید، باشد نرخ تغییر استراتژی در آن نقطه برابر صفر است.($\dot{x}_i = 0$)

ما در این مسئله از بازی تکاملی برای مدل کردن فرآیند پویای تصمیم-سازی کاربران برای انتخاب گره دستیار، استفاده می‌کنیم. در این بازی تکاملی، یک نظیر، زمانی که مشاهده کند که میزان سودمندی به آن از میانگین مابقی هم‌گروهی های($\bar{\pi}$) خود کمتر است، گره دستیار (استراتژی) خود را عوض می‌کند. برای این بازی تکاملی، تعادل تکاملی را به‌عنوان راه‌حل در نظر گرفتیم در این نقطه تعادل، ضمانت می‌شود که تمامی افراد در یک گروه همگی میزان پرداختی یکسانی را از کیفیت کانال مدنظر خود دریافت می‌کنند مطابق با مفهوم replicator dynamics ، زمانی که تمامی بازیکنان سود یکسانی را به‌دست آورند و هیچ کاربری انگیزه و تمایلی برای تغییر نداشته باشد تعادل حاصل می‌شود.

برای محاسبه تعادل تکاملی باید معادله replicator dynamics که قبلا آن را توضیح دادیم را حل کنیم. برای حل این معادله، نیازمند اطلاعات کامل از وضعیت سیستم و نظیرهای مربوط به هر گروه و از این جهت نیازمند یک هماهنگ‌کننده مرکزی برای جمع‌آوری این داده‌ها هستیم.

با توجه به اینکه وجود این هماهنگ‌کننده مرکزی خود دارای سربار زیاد و در برخی از شبکه‌های بزرگ غیر قابل پیاده‌سازی است در نتیجه، ما در الگوریتم پیشنهادی یک روش غیرمتمرکز ارائه می‌دهیم که بدون نیاز به اطلاعات سراسری قادر هست که به صورت تکرارشونده و برخط به نقطه تعادل بازی همگرا شود.

2-3- الگوریتم پیشنهادی

در این قسمت ما به ارائه یک الگوریتم برای بازی‌های تکاملی که قبلاً به آن اشاره‌شده بود، می‌پردازیم که مبتنی بر انتخاب هر نظیر می‌باشد. این الگوریتم، مبتنی بر یادگیری Q-Learning است. در ادامه به معرفی تمامی پارامترهای دخیل در این نوع یادگیری و توضیح خود الگوریتم، می‌پردازیم:

کاربر : یک کاربر در اصل یک تصمیم‌گیرنده است. در اینجا، منظور نظیری است که می‌خواهد بهترین گره‌دستیار که بیشترین سود را در اختیار نظیر قرار می‌دهد، انتخاب کند (منظور از سود، تلفیقی از میزان ظرفیت پهنای باند دریافتی از آن دستیار و میزان هزینه‌ای که بابت اتصال باید به آن بپردازد است).

وضعیت : یک وضعیت شامل موقعیت جاری کاربر و یا انتخاب جاری آن می‌باشد. در مدل مسئله ما، منظور از وضعیت، انتخاب جاری گره نظیر است. مجموعه‌ای از وضعیت‌ها را می‌توان به شکل زیر تعریف کرد:

$S = \{ s=h_i \mid h_i \in h1 , h_2 , \dots , h_n \}$

که h_i معرف هر یک از گره‌های دستیار می‌باشد.

عمل : یک عمل تصمیمی است که توسط کاربر گرفته می‌شود. در مورد مسئله ما، عمل همان انتخاب بهترین گره‌دستیاری که نظیر برای دریافت کانال خود می‌تواند به آن متصل شود، می‌باشد. مجموعه عمل‌ها را می‌توان به شکل زیر تعریف کرد:

$A = \{ a=h_i \mid h_i \in h_1 , h_2 , \dots , h_n \}$

پاداش : منظور از پاداش، یک تابع سودمندی از عمل انجام‌شده است. ما در این مسئله، $Q(s,a)$ به‌عنوان یک سود مورد انتظار برای یک کاربر در وضعیت s که عمل a را انجام می‌دهد، معرفی می‌کنیم.

به علت اینکه نظیرها از انتخاب دیگران و همچنین میزان سود مورد انتظار از گره‌های دستیار مختلف، آگاه نیستند، از تابع $Q_t(s,a)$ برای تخمین $Q(s,a)$ در زمان t استفاده می‌کنند.

1. *initialize Q-value associated with helper i for all users in all groups*
2. *initialize states associated with helper i for all users in all groups*
3. *while (! converged)*
4. *if rand() < ε (exploration step)*
   Select helper i (Randomly take an action a, a∈A)
5. *else*
   Take an action a_{t+1}=argmax $\{Q_t(s, a)\}$
6. *endif*
7. *User Computes the utility of each helper connect with*
8. *Update Q-value* $Q_{t+1} = (1-\lambda) * Q_t + \lambda * (\pi + \beta * maxQ_t)$, *∀s∈S, a∈A*
9. *endWhile*

در ابتدای کار، وضعیت اولیه(گره دستیار اولیه) برای تمامی نظیرها انتخاب می‌شود. یعنی هر نظیر، در ابتدا بعد از انتخاب وضعیت اولیه (Q_0 $(s,a)=0$ برای هر s,a و تمامی نظیرها) هر نظیر به‌صورت حریصانه شروع به انتخاب گره‌دستیار برای خود می‌کند. با توجه به وضعیت جاری خود، گره دستیاری که بیشترین سود (نسبت به دوره قبلی) می‌دهد، را انتخاب می‌کند.

در این الگوریتم، یک کاربر مرحله اکتشاف را با احتمال ε انجام می‌دهد. این مقدار به‌مرور زمان کمتر می‌شود و این بدین معنی است که نظیرها، بیشتر از گذشته خودشان برای انتخاب گره‌دستیار استفاده می‌کنند و در اینجا مفهوم یادگیری رنگ بیشتری به خود می‌گیرد.

همچنین مقدار تابع $Q_t(s,a)$ را می‌توان از رابطه زیر به دست‌آورد:

$$Q_{t+1} = (1-\lambda) * Q_t + \lambda * (\pi + \beta * maxQ_t) \quad (6)$$

که در رابطه بالا، λ نرخ یارگیری است که برای کنترل سرعت تنظیم مقدار Q استفاده می‌شود. همان طور که ملاحظه می‌کنید، مقدار جدید Q که برای زمان بعدی (دوره بعد) است مبتنی بر Q فعلی و میزان سود مشاهده‌شده توسط نظیر، حساب می‌شود.

2-4- پیاده‌سازی و نتایج شبیه‌سازی

در این بخش، میزان کارایی راه‌حل توزیع‌شده پیشنهادی برای هماهنگ‌سازی انتخاب دستیار توسط نظیرها را در سناریوها و پیکربندی‌های مختلف ارزیابی می‌نماییم.

در اینجا لازم است با توجه به رابطه(2) که در (بخش 2-2) بیان شد، به توضیح مفهوم پارامترهای مورد استفاده در پیاده‌سازی پرداخته‌شود.

در رابطه(2)، مفهوم میزان ظرفیت بیان شد که در اینجا ما آن را تنها با یک عدد بیان می‌کنیم و این عدد بیان‌کننده میزان پهنای باند کل گره‌دستیار

است که به‌صورت یکسان بین تمامی نظیرهایی که برای دریافت کانال‌های سرویس داده‌شده توسط گره دستیار به آن متصل هستند، تقسیم می‌شود.

همچنین در رابطه مذکور از تابع قیمت‌گذاری استفاده‌شده است که برای سادگی کار یک مقدار ثابت بین 0 و 1 در نظر گرفته‌ایم. این مقادیر توسط خود گره‌های دستیار به‌صورت مستقل مشخص می‌شود و می‌تواند برای گره‌های دستیار مختلف که حتی کانال یکسان را سرویس می‌دهند باهم متفاوت باشد. در ادامه ما سه نوع سناریو را مورد بررسی قراردمی‌دهیم.

در سناریو اول به مقایسه الگوریتم پیشنهادی برای رسیدن به نقطه تعادل با حالتی که مسئله به صورت غیر برخط با یک کنترل کننده مرکزی حل می‌شود، می‌پردازیم. کنترل‌کننده مرکزی با توجه به میزان پرداختی جاری برای هر نظیر، به محاسبه میانگین پرداختی برای هر گروه می‌پردازد و سپس این مقدار را برای هر کاربر ارسال می‌کند. هر کاربر متناسب با مقدار میزان پرداختی که گرفته و مقایسه آن با مقدار میانگینی که توسط کنترل‌کننده مرکزی دریافت کرده، در مورد تغییر و یا عدم تغییر استراتژی خود تصمیم می‌گیرد.

در سناریو دوم اثر تغییر در پارامترهای موجود در مسئله را مورد بررسی قرار می‌دهیم و در نهایت در سناریو سوم به مقایسه جواب روش پیشنهادی برای حل این مسئله با جواب حل این مسئله با الگوریتم حریصانه می‌پردازیم. در الگوریتم حریصانه، تمامی گره‌ها در هرلحظه با توجه به مقدار بیشینه سودی که در آن لحظه می‌توانند از گره دستیار دریافت کنند، بدون توجه به تاریخچه انتخاب‌های قبلی خود گره دستیاری برای خود انتخاب می‌کند.

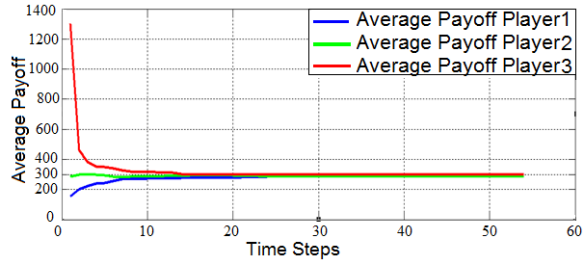
در تمامی سناریوهای مطرح شده، به صورت پیش فرض، سیستم ما دارای سه گره دستیار است که که در سناریو اول و سوم، ظرفیت هر گره‌ها با هم مساوی و برابر با 30000 است (این عدد، بیانگر میزان پهنای باند ارسالی از طرف گره دستیار به گره‌هایی که او را به عنوان پدر انتخاب کرده اند،است.) همچنین تابع قیمت یا همان هزینه‌ای که باید هر گره برای اتصال به گره دستیار پرداخت کند، برای هر گره دستیار مساوی و برابر با 0/1 در نظر گرفته شده است. لازم به ذکر است که این عدد می‌تواند تابعی از پارامترهای محیط مانند زمان و میزان کیفیت ارسال ویدیو از طرف گره دستیار برای هر گروه (کانال) متفاوت در نظر گرفته شود. همچنین فرض شده است 3 کانال توسط این گره‌های دستیار فراهم شده که چگونگی توزیع کانال‌ها در جدول آمده است.

جدول 1 پارامترهای ورودی

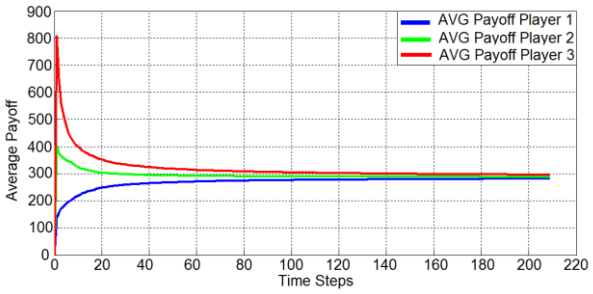
1	تعداد دستیار	3 گره		
2	میزان ظرفیت و تابع قیمت‌گذاری برای هر کدام از گره‌های دستیار		ظرفیت گره	تابع قیمت
		گره دستیار شماره 1	30000	0.1
		گره دستیار شماره 2	30000	0.1
		گره دستیار شماره3	30000	0.1
2	تعداد کانال	3کانال		
3	تعداد نظیرها در گروه	گروه مربوط به کانال 1	100	
		گروه مربوط به کانال 2	100	
		گروه مربوط به کانال 2	100	

4	کانال‌های سرویس داده‌شده توسط گره‌های دستیار		کانال 1	کانال2	کانال3
		گره‌دستیار1	✓	✓	✓
		گره‌دستیار2		✓	✓
		گره‌دستیار3			✓

هدف از این سناریو اول، مقایسه اولیه بین جواب حاصل حل مسئله به صورت غیر برخط و با استفاده از یک کنترل‌کننده مرکزی فرضی با الگوریتم یادگیری که به‌صورت توزیع‌شده این کار را انجام می‌دهد، است. مقدار این میزان سود پرداختی[۴] را در ادامه با مقدار میانگین سود پرداختی که هر گره‌دستیار در الگوریتم توزیع‌شده یادگیری، به گره‌های نظیرشان اختصاص می‌دهند را با نمودارهای حاصل از شبیه‌سازی نمایش می‌دهیم. انتظار می‌رود که این دو مقدار باهم برابر باشند. نتایج حاصل از شبیه‌سازی این سناریو به شکل زیر است.



شکل 2 سود پرداختی هردستیار از الگوریتم متمرکزسناریو1

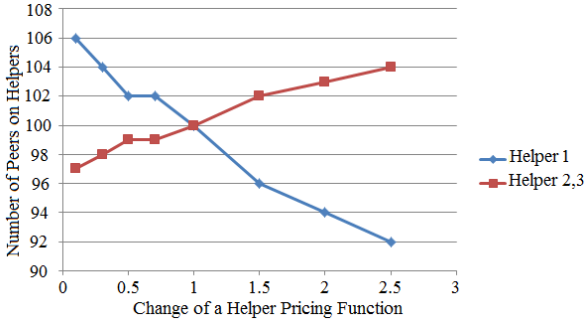


شکل 3 سود پرداختی هردستیار از الگوریتم یادگیری توزیع شده سناریو1

در هر دو روش میزان سود پرداختی که هر گره دستیار در این سیستم با پیکربندی ذکرشده در اختیار نظیرها قرار می‌دهند برابر 300 است. در الگوریتم متمرکز بعد از گذشت 20 گام زمانی سیستم در حالت تعادل پایدار قرار می‌گیرد و در حالت پایدار خود می‌ماند. همان طور که قبلاً هم اشاره شد، در الگوریتم یادگیری این تعداد گام زمانی بسیار بیشتر از الگوریتم اول است و به‌عبارت‌دیگر دیرتر به تعادل پایدار می‌رسد.

هدف از اجرای سناریو دوم، مشاهده اثر تغییر تابع قیمت‌گذاری یکی از گره‌های دستیار در تعداد نظیرهای متصل به آن و گره‌های دستیار موجود در سیستم است. در این سناریو، تابع قیمت‌گذاری یکی از گره‌های دستیار را افزایش می‌دهیم و تابع قیمت‌گذاری سایر گره‌های دستیار دیگر ثابت باقی می‌ماند. در اینجا، سیستم ما دارای سه گره دستیار است که ظرفیت هر گره‌ها با هم مساوی و برابر با 10000 است. همچنین فرض شده است 3 کانال توسط این گره‌های دستیار فراهم شده که در آن کانال شماره 1 توسط گره دستیار اول و گره دستیار سوم، کانال شماره 2 توسط گره‌های دستیار اول و

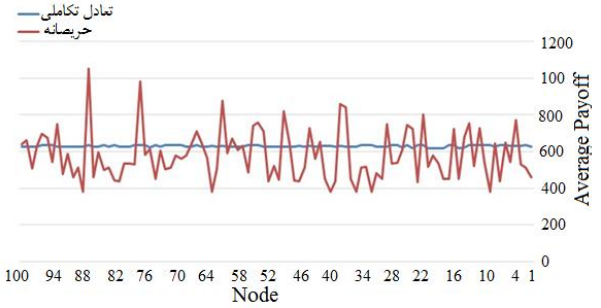
دوم و در نهایت کانال شماره 3 توسط گره‌های دستیاردوم و سوم، فراهم شده و هر کانال 100 گره بازدید کننده دارد.



شکل 4 اثرتغییرتابع قیمت‌گذاری یکی ازگره‌های دستیاردرتعداد نظیرهای متصل به آن

مطابق شکل، با افزایش میزان تابع قیمت‌گذاری گره دستیار اول، تعداد نظیرهایی که برای دریافت کانال مورد نظیر خود به این گره‌دستیار متصل می‌شوند به‌صورتغیرخطی کاهش پیدا می‌کند تا تعادل برقرار باشد.

در سناریو سوم، تمامی پارامترهای این مانند سناریو اول است و هدف از این سناریو، مقایسه نتایج حاصل از الگوریتم پیشنهادی و الگوریتم حریصانه برای انتخاب پدر می‌باشد.



شکل 5 مقایسه نتایج حاصل از الگوریتم پیشنهادی و حریصانه

همانطور که مشاهده می‌شود، حاصل الگوریتم پیشنهادی نتیجه عادلانه‌تری را در سیستم برای نظیرهای مختلف به ارمغان می‌آورد.

3- نتیجه‌گیری

در این مقاله مسألهٔ هماهنگ‌سازی تصمیمات نظیرها برای انتخاب گره‌های دستیار در یک سیستم جریان‌سازی نظیر–به–نظیر ویدئو چند–کاناله را مورد توجه قرار دادیم. در بازی‌های تکاملی به دلیل برخورداری از ماهیت عادلانه‌تر و وجود نوعی مفهوم هماهنگی ذاتی در تعریف، معمولاً به کارایی بهتری در سیستم منجر خواهد شد. حاصل این هماهنگ‌سازی، رسیدن به تعادل تکاملی است که به یکی از دو طریق متمرکز و توزیع‌شده است، صورت پذیرفت.

هرچند که با استفاده از الگوریتم متمرکز که در آن با استفاده از یک کنترل‌کننده مرکزی، به جمع‌آوری اطلاعات میزان سود پرداختی برای نظیرهای مختلف در گروه‌های متفاوت و همچنین محاسبه میانگین برای هر گروه، زودتر به تعادل تکاملی می‌رسیم ولی در عمل وجود یک کنترل‌کننده مرکزی به علت وجود سربار اضافی در شبکه میسر نیست. برای حل این مشکل، با استفاده از الگوریتم‌های یادگیری توزیع‌شده، این امکان را به خود نظیرها می‌دهیم که به عنوان یک عامل یادگیر و با استفاده از گذشته اعمال خود، انتخاب گره دستیار برای خودشان را بدون نیاز به کنترل‌کننده مرکزی، انجام بدهند.. برای ادامه کار در مورد این مقاله، کارهای مختلفی را می‌توان انجام داد مانند ارائه یک راه حل کلی برای میزان اختصاص پهنای باند توسط گره‌های دستیار به‌صورت یادگیری و در نهایت استفاده از مفاهیم دیگر مانند میزان گم‌شدن بسته‌ها برای مفهوم میزان سود پرداختی و همچنین ایجاد یک مکانیزم واحد برای پرداخت به گره‌های دستیار از طرف نظیرها دانست.

مراجع

[1] M.Wang, L.Xu, and B.Ramamurthy, "Linear programing models for multi-channel P2P streaming systems," in INFOCOM, pp. 1-5, 2010 Proceedings IEEE, 2010

[2] S.A Mostafavi, and M. Dehghan. "A stochastic approximation resource allocation approach for HD live streaming." *Telecommunication Systems* (2016): 1-15.

[3] S.A Mostafavi, and M. Dehghan. "Game-theoretic auction design for bandwidth sharing in helper-assisted P2P streaming." *International Journal of Communication Systems* (2015).

[4] S.Xie, B.Li, G. Y. Keung, and X. Zhang, "Coolstreaming: Design, theory, and practice," Multimedia, IEEE Transactions on, vol. 9, pp. 1661-1671, 2007

[5] S. S. Savas, A. M. Tekalp, and C. G. Gurler, "Adaptive multi-view video streaming over P2P networks considering quality of experience," in Proceedings of the 2011 ACM workshop on Social and behavioural networked media access, pp. 53-5, 2011

[6] A. Habib and J. Chuang, "Service differentiated peer selection: an incentive mechanism for peer-to-peer media streaming," IEEE Trans. Multi. vol. 8(3), pp. 610-621, September 2006.

[7] Y.T.H. Li, D. Ren, S.G. Chan, and A.C. Begen, "Low-delay mesh with peer churns for peer-to-peer streaming," ICME, pp. 1546-1547, 2009.

[8] Y.Xu, C.Zhu, W. Zeng, and X.J. Li, "Multiple description coded video streaming in peer-to-peer networks," Image Commun. vol. 27(5), pp. 412-429, May 2012.

[9] Z. Liu, C. Wu, B. Li, and S. Zhao, "Distilling Superior Peers in Large-Scale P2P Streaming Systems," INFOCOM, 2009, pp. 82-90

[10] D. Niu, B. Li, and S. Zhao, "Self-diagnostic peer-assisted video streaming through a learning framework," in Proc. ACM Multimedia, pp. 73-82, 2010

[11] Z. Liu, C. Wu, B. Li, and Shuqiao Zhao, "UUSee: large-scale operational on-demand streaming with random network coding," in Proc. INFOCOM, pp. 2070-2078, NJ, USA, 2010

[12] R. Izhak-Ratzin, H. Park, and M. van der Schaar, "Reinforcement learning in bittorrent systems," in INFOCOM, 2011 Proceedings IEEE, pp. 406-410, 2011

[13] M. Sayit and O. Sonmez, "Reinforcement learning for peer to peer video streaming applications," in Signal

Processing and Communications Applications Conference (SIU), 2012 20th, pp. 1-4, 2012

[14] G. T. Karetsos, S. A. Kyriazakos, E. Groustiotis, F. D. Giandomenico,and I. Mura, “A hierarchical radio resource management framework forintegrating WLANs in cellular networking environments,”IEEE Wireless Commun., vol. 12, no. 6, pp. 11–17, Dec. 2005

[15] F. Yu and Y. Krishnamurthy, “Optimal joint session admission control in integrated WLAN and CDMA cellular networks with vertical handoff,” IEEE Trans. Mobile Comput., vol. 6, no. 1, pp. 126–139, Jan. 2007

[16] D. Niyato and E. Hossain, "Dynamics of network selection in heterogeneous wireless networks: an evolutionary game approach," Vehicular Technology, IEEE Transactions on, vol. 58, pp. 2008-2017, 2009

[17] M. F. Sayit, Y. Kaymak, K. D. Teket, C. Cetinkaya, S. Demirci, and G. Kardas, "Parent selection via reinforcement learning in mesh-based P2P video streaming," in Information Technology: New Generations (ITNG), 2013 Tenth International Conference on, pp. 546-551, 2013

[18] Z. Feng, L. Song, Z. Han, D. Niyato, and X. Zhao, "Cell selection in two-tier femtocell networks with open/closed access using evolutionary game," in Wireless Communications and Networking Conference (WCNC), 2013 IEEE, pp. 860-865, 2013

پانویس‌ها

[1] Peer to Peer
[2] Reinforcement Learning
[3] Evolutionary equilibrium
[4] Payoff