علوم و مهندسی کامپیوتر
نشریه علمی پژوهشی انجمن کامپیوتر ایران
مجلد ۹، شماره ۱ و ۳ (الف)، ۱۳۹۰
صفحات ۸۰-۸۸
مقاله عادی

یک مکانیزم انگیزشی برای شبکه‌های همکاری با استفاده از بازی نظیریابی تصادفی

فائزه بهرامیان حمیدرضا شهریاری

دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر، تهران، ایران

چکیده

تامین هزینه و انرژی موردنیاز برای ارائه سرویس در شبکه‌های همکاری به عهده کاربران شبکه است. از سوی دیگر، کاربران شبکه‌های همکاری عامل‌هایی مستقل بوده و هدفشان بیشینه کردن بهره‌ای است که از شبکه به دست می‌آورند. بنابراین، در حالت عادی، کاربران شبکه انگیزه کافی برای همکاری با سایر اعضا ندارند. استفاده از مکانیزم‌های انگیزشی روشی مرسوم برای اعمال همکاری در چنین شبکه‌هایی است.

در این مقاله، یک مکانیزم انگیزشی برای شبکه‌های همکاری با استفاده از سیستم شهرت مرکزی و مبتنی بر نظریه بازی‌ها ارائه می‌شود. به نظر می‌رسد بازی نظیریابی تصادفی برای مدل کردن خصوصیات پایه‌ای شبکه‌های همکاری مناسب باشد. زیرا فرض تصادفی بودن نظیریابی بازیکنان در این مدل، سازگاری زیادی با ماهیت پویا و در حال تغییر شبکه‌های همکاری دارد. از این روی، ابتدا تعاملات بین گره‌ها به عنوان یک بازی نظیریابی تصادفی مدل شده و سپس سه پروفایل استراتژی تنبیه، توبه و توبه بهینه که در نحوه تنبیه بازیکن متخلف با یکدیگر تفاوت دارند و به ترتیب، افزایش کارایی شبکه را منجر می‌شوند، پیشنهاد می‌کنیم. اثبات می‌شود که این پروفایل‌های استراتژی تعادل زیربازی-کامل هستند. به این ترتیب، این مکانیزم در تشویق گره‌ها به همکاری موفق خواهد بود.

کلمات کلیدی: شبکه همکاری، مکانیزم انگیزشی، سیستم شهرت، بازی نظیریابی تصادفی.

۱- مقدمه

نسل جدید شبکه‌های ارتباطی به کاربران خود اجازه می‌دهند که تصمیم بگیرند و عمل خود را آزادانه و مستقل اجرا کنند. هیچ واحد مرکزی وجود ندارد که عملکرد کل شبکه را مدیریت کند. برای هر کاربر، مجموعه اعمالی تعریف شده است که می‌تواند بر اساس آنها عمل خود را در هر لحظه انتخاب کند. از بین این مجموعه اعمال، رفتار مشخصی وجود دارد که از کاربران انتظار می‌رود برای بقا و کارایی بیشتر شبکه آن را اجرا نمایند. این رفتار همکاری نامیده می‌شود. یک شبکه همکاری [۱] از موجودیت‌هایی (سازمان‌ها و انسان‌ها) تشکیل شده است که کاملا خودمختار و از لحاظ جغرافیایی توزیع شده‌اند و در برخی ویژگی‌هایشان مانند محیط عملیاتی، فرهنگ، هنجارهای اجتماعی و اهداف متفاوت و ناهمگن هستند. ارتباط بین این موجودیت‌ها از طریق شبکه‌های کامپیوتری صورت می‌گیرد. شبکه‌های نظیربه‌نظیر، موردی و سیار موردی مشهورترین انواع شبکه‌های همکاری هستند.

امروزه شبکه‌های همکاری نقش مهمی در دنیای مجازی و اینترنت بازی می‌کنند. بر اساس آماری که آی‌پوک[1] در مورد ترافیک اینترنت در هشت منطقه مختلف از جهان در دوره زمانی ۲۰۰۸-۲۰۰۹ منتشر کرده است، شبکه‌های نظیربه‌نظیر بیشترین ترافیک را در همه مناطق بررسی شده تولید کرده‌اند [۲]. همچنین این آمار نشان می‌دهد که در بین پروتکل‌های نظیربه‌نظیر، بخش اعظم ترافیک را پروتکل‌های بدون ساختار متمرکز، مانند بیت‌تورِنت[2] به خود اختصاص داده‌اند. همکاری در این نوع شبکه‌ها نیازمند صرف مقداری هزینه است که تامین آن به عهده کاربران شبکه است. از سوی دیگر، معمولا کاربران شبکه‌های همکاری افرادی خودخواه فرض می شوند که هدفشان رسیدن به بیشترین بهره از شبکه، بدون ارائه خدمات به دیگران است. بنابراین، در حالت عادی، کاربران شبکه انگیزه کافی برای همکاری با سایر اعضا و دادن هزینه تامین منابع درخواستی آنها

را ندارند. استفاده از مکانیزم‌های انگیزشی روشی مرسوم برای اعمال همکاری در چنین شبکه‌هایی است. مکانیزم‌های انگیزشی به دو دسته مبتنی بر پول و مبتنی بر شهرت تقسیم می‌شوند. در مکانیزم‌های مبتنی بر پول، گره‌ها از پرداخت الکترونیکی خرد برای تبادل و به اشتراک‌گذاری منابع استفاده می‌کنند. به عنوان مثال، یک فایل هنگامی به اشتراک گذاشته می شود که مبلغ تعیین شده برای آن، قبل از بارگذاری، پرداخت شود. مشکل مهم این روش نیاز به یک بانک مرکزی برای مدیریت پرداخت و مقدار زیادی سربار برای هزینه‌های ارتباطاتی و محاسباتی است.

از سوی دیگر، مکانیزم‌های مبتنی بر شهرت به هر گره شهرتی نسبت می‌دهند که رفتار گذشته او را در شبکه منعکس می‌کند. بر اساس این شهرت، گره‌ها درمورد چگونگی تعاملاتشان با یکدیگر تصمیم می‌گیرند. بنابراین میزان شهرت می‌تواند بر خدماتی که یک گره از شبکه دریافت می‌کند و در نتیجه بهره او مؤثر باشد. عدم همکاری باعث کاهش شهرت و در نتیجه عدم امکان دریافت خدمات از شبکه خواهد شد.

شهرت درکی است که یک عامل از عملکرد گذشته‌اش نسبت به نیت و مدل رفتاری خود می‌سازد[۳]. به عبارت دیگر می‌توانیم از شهرت به عنوان وسیله‌ای برای ایجاد اعتماد و یا معیاری برای ارزیابی قابلیت اعتماد استفاده کنیم. منظور از اعتماد در اینجا تصمیم‌گیری در مورد عامل دیگری در شبکه است که می‌تواند برای یک عملیات خاص مورد اعتماد باشد یا خیر؟

بنابر آنچه گفته شد، هدف مکانیزم‌های انگیزشی اِعمال همکاری در شبکه است. برای دستیابی به این هدف، اینگونه سیستم‌ها می‌بایست به دقت طراحی شوند به طوری که کاربران عاقل و خودمختار شبکه رفتار مورد انتظار طراح سیستم را انجام دهند.

براساس اصل پایه‌ای نظریه بازی‌ها، بازیکنان افرادی خودخواه و مستقل‌اند و بهره‌ای که از بازی به دست می‌آورند به عمل خود و رقیبشان بستگی دارد. هدف اصلی آنها بیشینه کردن درآمدشان از بازی است. گره‌های یک شبکه همکاری نیز ویژگی‌هایی مشابه بازیکنان دارند. آنها تصمیم‌گیرندگانی خودمختارند که همه تلاششان برای به دست آوردن سود بیشتر است. این شباهت باعث می‌شود بتوانیم نگاشت صحیحی بین مولفه‌های معمولی نظریه بازی‌ها و اجزای شبکه همکاری برقرار کنیم.

به نظر می‌رسد بازی نظیریابی تصادفی[۳] برای مدل کردن شبکه‌های همکاری مناسب و سازگار با خصوصیات پایه‌ای آنها باشد. در مقالات ارائه شده در زمینه مکانیزم های انگیزشی، کمتر از این نوع مدل بازی استفاده شده است و غالبا تعاملات گره‌های شبکه با بازی معروف معمای زندانی[۴] تکراری مدل شده است. در بازی نظیریابی تصادفی فرض می‌شود در هر دور بازی، بازیکنان در برابر فرد جدیدی قرار می‌گیرند و رقیب هر بازیکن در هر دور به طور تصادفی انتخاب می‌شود که این شرایط از ویژگی‌های اساسی شبکه‌های همکاری به حساب می‌آید.

مرجع [۱۷] یکی از مهمترین کارهایی است که از بازی نظیریابی تصادفی برای حل مساله مسیریابی در شبکه‌های همکاری، با فرض وجود یک سیستم شهرت، استفاده می‌کند. اما استراتژی پیشنهاد شده برخلاف ادعای مقاله، یک استراتژی تعادل زیربازی ـکامل نیست.

در این مقاله، مدلی براساس بازی نظیریابی تصادفی ارائه می‌کنیم که تا حد امکان ساده و در عین حال به واقعیت نزدیک باشد. تحلیل دقیقی از دوره تنبیه گره متخلف ارائه خواهد شد که منجر به یافتن استراتژیهای زیربازی ـکامل می‌شود. سه پروفایل استراتژی تنبیه، توبه و توبه بهینه که در نحوه تنبیه بازیکن متخلف با یکدیگر تفاوت دارند و به ترتیب، افزایش کارایی شبکه را منجر می‌شوند، پیشنهاد می‌کنیم. بدین ترتیب، مدل ارائه شده در طراحی مکانیزم های انگیزشی برای به کار گیری در انواعی از شبکه‌های همکاری مفید خواهد بود. این موضوع نشان دهنده قابلیت استفاده عملی از مکانیزم ارائه شده است.

ادامه مطالب این مقاله به صورت زیر سازماندهی شده است. در بخش ۲ بازی نظیریابی تصادفی توضیح داده می‌شود. بخش ۳ مدل و فرضیات بازی گره‌ها در شبکه‌های همکاری را شرح می‌دهد. استراتژی‌های پیشنهادی و اثبات اینکه زیر بازی ـکامل هستند در بخش ۴ آمده است. بخش ۵ به تحلیل و ارزیابی مدل و استراتژیها می‌پردازد. در بخش ۶ برخی از مهمترین مقالات مرتبط با این حوزه شرح داده می‌شود و بخش ۷ مقاله را به پایان می‌رساند.

۲- بازی نظیریابی تصادفی

در بازی نظیریابی تصادفی، فرض بر این است که تعداد زیادی بازیکن، بی‌شمار بار بازی می‌کنند. در هر دور بازی، هر بازیکن در برابر بازیکن دیگری قرار می‌گیرد که به صورت تصادفی انتخاب می‌شود. تابع انتخاب را می‌توان به دلخواه تعریف کرد. در هر مرحله، انتخاب رقبا مستقل از مراحل قبلی است [۴].

اولین تحلیل از مدل نظیریابی تصادفی را روزنتال[۵] و لَندو[۶] در [۵، ۶] ارائه کردند. آنها فرض کردند که اطلاعات هر بازیکن در هر دور تنها شامل آخرین عملی است که خود و بازیکن رقیبش در دور قبل انجام داده اند. بنابراین، اگر بازی مرحله[۷]، یک بازی معمای زندانی باشد چهار حالت (C,C)، (D,D)، (C,D) و (D,C) برای سابقه بازی یک جفت از بازیکنان وجود دارد که در آن C به معنی همکاری و D به معنی تخلف است. به طور مثال، (C,D) نشان می‌دهد که در مرحله قبل بازیکن اول C و بازیکن دوم D بازی کرده است.

کَندوری[۸] تحلیل دیگری از این بازی را در [۷] ارائه کرد. او فرض کرد که اطلاعات بازیکنان در هر دور شامل نتیجه آخرین بازی است که توسط خودش و بازیکن رقیبش در دور قبل انجام شده است. یعنی علاوه بر اینکه از عمل طرف مقابلش در آخرین دور اطلاع دارد عمل رقیب طرف مقابلش را نیز می‌داند. طبق تعریف کَندوری از این مدل، ساختار اولیه بازی مطابق زیر توصیف می‌شود. مجموعه بازیکنان $\mathrm{N} = \{1,..,2\mathrm{n}\}$ که به دو مجموعه با تعداد افراد مساوی $\mathrm{N}_1 = 1,..,\mathrm{n}$ و $\mathrm{N}_2 = \mathrm{n}+1,..,2\mathrm{n}$ افراز می‌شود. در هر مرحله هر یک از بازیکنان از نوع N_1 با یکی از بازیکنان از نوع N_2 بر طبق یک قانون نظیریابی جفت می‌شوند و یک بازی مرحله دو نفره را بازی می‌کنند. این روند نامتناهی بار تکرار می‌شود و بهره نهایی هر بازیکن برابر است با مجموع مورد انتظار بهره‌های هر مرحله با احتساب مقدار فاکتور تخفیف $\delta \in (0,1)$.

کَندوری همچنین نشان داد که حتی اگر یک بازیکن چیزی غیر از تجربه فردی خودش در مورد دیگران نداشته باشد امکان برقراری همکاری وجود دارد. با اینکه این مدل با شرایط شبکه های همکاری تا حد زیادی مطابقت دارد، امکان ارائه استراتژی‌هایی که در آنها همکاری پایدار باشد وجود ندارد. وی در این مقاله رفتار مطلوب جمعی همراه با قوانین هماهنگ در یک جامعه را نرم اجتماعی تعریف می‌کند و به این مسئله که چگونه در جامعه‌ای با اعضای خودخواه، یک نرم اجتماعی برقرار می‌شود، می‌پردازد. او همچنین چند قضیه عامه[۹] برای بازیهای نظیریابی تصادفی اثبات می‌کند. براساس این قضایا جامعه‌ای که در آن هر بازیکن یک برچسب مانند شهرت، عضویت یا گواهی که به صورتی سیستماتیک اعلان می‌شود، دارد، می‌تواند به هر نتیجه متقابل و امکان‌پذیر بازی در قالب یک استراتژی تعادل دست پیدا کند.

۳- مکانیزم انگیزشی مبتنی بر بازی نظیریابی تصادفی

مکانیزم‌های انگیزشی با رویکرد نظریه بازی‌ها از شهرت به عنوان ابزاری به منظور ایجاد انگیزه در عامل‌ها برای همکاری استفاده کرده‌اند. با تنبیه عامل‌های بد و

تشویق عامل‌های خوب، سیستم را به صورت یک بازی مدل می‌کنند که در این بازی همکاری کردن برای همه عامل‌های عاقل، که می‌خواهند بهره مورد انتظار نهاییشان را بیشینه کنند، یک استراتژی تعادل باشد. مسئله اصلی آن است که استراتژی‌های دریافت بازخورد، تنبیه و تشویق چه طور تعریف شوند که تخطی یک طرفه از آنچه همکاری نامیده می شود سود آور نباشد. فرضیات و مدلی که در این بخش ارائه خواهد شد مشابه مدل [۸] است.

۳-۱- فرضیات

در این بخش به توصیف شبکه همکاری، سیستم شهرت و فرضیاتی که مدل بر پایه آنها بنا شده است، می‌پردازیم.

بازیکنان بازی همان گره‌های شبکه هستند که با تعامل و همکاری با یکدیگر شبکه را تشکیل می‌دهند. فرض می‌کنیم بازیکنان عاقلند و همیشه بهترین تصمیم را با توجه به شرایط موجود اتخاذ می‌کنند. همه گره‌ها از نظر نقش در شبکه یکسانند و واحد مدیریت مرکزی وجود ندارد. امکانات گره‌ها مشابه و محدود است. در هر تعامل، یک بازیکن نقش سرویس‌دهنده و دیگری نقش سرویس‌گیرنده دارد و بازیکنان در مراحل مختلف می‌توانند نقش‌های متفاوتی داشته باشند.

سرویس‌دادن هزینه‌بر و سرویس‌گرفتن سودمند است. بنابراین، گره‌ها ترجیح می‌دهند بیشترین سرویس را از شبکه دریافت کنند در حالی که کمترین خدمات را به سایر اعضای شبکه ارائه نمایند. در صورتی که اگر بیشتر گره‌ها استراتژی عدم همکاری را انتخاب کنند و به یکدیگر سرویس ندهند، شبکه از کار می افتد. در این مقاله، فرض می‌شود هزینه سرویس و سود حاصل از آن برای همه سرویس‌ها و همه بازیکنان یکسان است.

هر بازیکن یک شناسه منحصربه‌فرد، مثلا آدرس IP، دارد که براساس آن در شبکه شناخته می‌شود. علاوه بر این، هر بازیکن برچسبی به نام شهرت دارد که پس از هر مرحله از بازی به‌روز می‌شود. مقدار شهرت صفر یا یک است که به ترتیب نشان دهنده عدم همکاری و همکاری در آخرین مرحله بازی هستند. شهرت اولیه همه بازیکنان برابر با یک است. همچنین فرض می‌شود که بازیکنان تبانی نمی‌کنند.

یک سیستم شهرت مرکزی مورد اعتماد در شبکه وجود دارد. هر بازیکن مقدار شهرت دیگران را از این مرکز دریافت می‌کند و پس از انجام هر عملیات نتیجه تعامل با طرف مقابلش را به آن گزارش می‌دهد. در واقع این واحد مرکزی مسئولیت جمع‌آوری، نگهداری و انتشار اطلاعات شهرت را به عهده دارد.

۳-۲- مدل

مدل ارائه شده، بازی نظریابی تصادفی و از نوع بازی‌های تکراری نامتناهی است. بازی به صورت زمان گسسته و تعاملات بین بازیکنان به عنوان چندین بازی مرحله‌ای همزمان مدل می‌شود. مولفه‌های بازی را به صورت زیر تعریف می‌کنیم. مجموعه بازیکنان را با $N = \{1, \dots, 2n\}$ نشان می‌دهیم که به دو مجموعه N_1 و N_2 با اندازه مساوی تقسیم می‌شود. $N_1 = \{1, \dots, n\}$ مجموعه بازیکنان با نقش سرویس‌گیرنده و $N_2 = \{n+1, \dots, 2n\}$ مجموعه بازیکنان با نقش سرویس‌دهنده را شامل می‌شود. بازی مرحله، یک بازی دو نفره نامتقارن است. بازیکن اول از مجموعه N_1 و بازیکن دوم از مجموعه N_2 انتخاب می‌شود. مجموعه اعمال ممکن برای سرویس‌دهنده $A = \{C, D\}$ است. که در آن C به معنی همکاری و منظور از D عدم همکاری است.

هر بازیکن در واحد زمان u، با احتمال p عملی در شبکه انجام می‌دهد. برای اینکه تعامل بین گره‌های شبکه با بازی نظریابی تصادفی سازگار باشد، هر مرحله از بازی را معادل تعامل گره‌ها در بازه زمانی u/p در نظر می‌گیریم. به این ترتیب، همه گره‌ها به طور متوسط در هر مرحله از بازی یکبار حضور خواهند داشت. در یک مرحله، احتمال اینکه گره i درخواست خدمات داشته باشد و یا مورد درخواست قرار گیرد مساوی و برابر $1/2$ درنظر گرفته می‌شود. به عبارت دیگر، هر بازیکن در هر مرحله با احتمال مساوی به مجموعه N_1 یا N_2 تعلق دارد. بنابراین، بازیکنان در مورد ارسال یا دریافت درخواست سرویس، استراتژیک عمل نمی‌کنند. برای پاسخگویی به هر درخواست، بازیکنی از مجموعه N_2 برطبق قانون نظریابی شبکه انتخاب می‌شود.

شهرت بازیکن i در زمان t را با r_i^t نشان می‌دهیم. قبل از انجام بازی مرحله، هر یک از طرفین بازی می‌توانند با درخواست از واحد شهرت مرکزی از شهرت بازیکن مقابلشان مطلع شوند و بر اساس آن تصمیم برای ادامه بازی و نحوه انتخاب عملشان اتخاذ نمایند. پس از اتمام بازی مرحله، مقدار شهرت بازیکنان بر اساس گزارش بازیکن سرویس‌گیرنده و مقدار اولیه شهرت به‌روز می‌شود.

بهره هر یک از بازیکنان براساس نقشی که در بازی مرحله ایفا می‌کنند، بعد از انجام هر تعامل مطابق با جدول ۱ محاسبه می‌شود. عملیات نوشته شده در سطرهای جدول مربوط به بازیکن اول و عملیات نوشته شده در ستون‌های جدول مربوط به بازیکن دوم است. در هر خانه جدول عدد سمت چپ، بهره بازیکن اول و عدد سمت راست، بهره بازیکن دوم را نشان می‌دهد. فرض می‌شود مقدار سودی که هر بازیکن از دریافت خدمت به دست می‌آورد از هزینه‌ای که باید برای ارائه سرویس به دیگران بپردازد بیشتر است. یعنی $s > c$.

جدول ۱- تابع بهره بازی مرحله

	همکاری	عدم همکاری
درخواست سرویس	s , -c	-d , 0

با توجه به فرض مسئله که بازیکنان در هر دور با احتمال مساوی ممکن است سرویس‌دهنده یا سرویس‌گیرنده باشند، تابع بهره بازی مرحله را می‌توانیم به صورت جدول ۲ نیز نشان دهیم. این تابع، بهره مورد انتظار یک مرحله از بازی مکانیزم انگیزشی را محاسبه می‌کند.

جدول ۲- تابع بهره مورد انتظار یک مرحله از بازی مکانیزم انگیزشی

	همکاری	عدم همکاری
همکاری	$\frac{1}{2}(s-c), \frac{1}{2}(s-c)$	$\frac{1}{2}(-c-d), \frac{1}{2}s$
عدم همکاری	$\frac{1}{2}s, \frac{1}{2}(-d-c)$	$-\frac{1}{2}d, -\frac{1}{2}d$

بهره نهایی هر بازیکن از روش جمع مورد انتظار بهره بازی‌های مرحله با فاکتور تخفیف $\delta \in (0,1)$ محاسبه می‌شود. فاکتور δ می‌تواند به عنوان تفاوت در ارزش زمان یا نرخ تورم تفسیر شود. همچنین می‌توان از آن به عنوان احتمال توقف بازی در انتهای هر دور تعبیر نمود. بنابراین بهره کلی بازیکن i مطابق فرمول (۱) محاسبه می‌شود.

$$U_i = \sum_{t=0}^{\infty} \delta^t u_i^t \qquad (۱)$$

که در آن u_i^t برابر است با بهره بازیکن i مربوط به بازی مرحله در زمان t.

سابقه عملکرد بازیکن i در زمان t با h_i^t نشان داده می‌شود. طبق تعریف بازی، سابقه بازیکن i در قالب شهرت او بیان می‌شود. یعنی داریم $h_i^t = r_i^t$. مجموعه همه سوابق ممکن در مورد بازیکن i در زمان t برابر است با $H_i^t = \{0,1\}$. آنچه که یک بازیکن از سابقه بازی می‌داند با آنچه سایر بازیکنان می‌دانند برابر است و به صورت

$$h^t = \cup_{i=1}^{N} h_i^t \,, \tag{۲}$$

نشان داده می‌شود و مجموعه سوابق ممکن برای بازی برابر است با $H^t = \times_{i=1}^{2n} H_i^t$.

۴- استراتژی‌های تعادل زیربازی ـ کامل

یک استراتژی تعادل مطلوب، استراتژی است که نتیجه آن همکاری بوده و همچنین بر روی این نتیجه پایدار باشد. یعنی درصورت عدم همکاری و یا تخلف عده‌ای از بازیکنان از تعادل، عدم همکاری در کل شبکه سرایت نکند و پس از چند مرحله، نتیجه استراتژی مجددا به همکاری همگرا شود.

۴-۱- استراتژی تنبیه

در این بخش ابتدا یک پروفایل استراتژی ارائه می‌شود که نیازمندی‌های استراتژی تعادل مطلوب را برآورده می‌کند. سپس اثبات می‌کنیم استراتژی پیشنهادی یک تعادل زیربازی‌ـکامل است. s^* یک پروفایل استراتژی متقارن است و برای بازیکن سرویس‌دهنده i چنین تجویز می‌کند:

s_i^*: در مرحله t، j را با شناسه بازیکن طرف مقابلت جایگزین کن و به این صورت عمل کن: اگر $r_j^t = 1$ ، همکاری کن در غیر اینصورت اگر $r_j^t = 0$، همکاری نکن. بازیکن متخلف باید به اندازه p دوره زمانی تنبیه شود یعنی از زمان t تا $t + p$ نباید سرویس دریافت کند.

بیان ریاضی و دقیق‌تر استراتژی به صورت زیر است. مجموعه $Z = \{(r, k) | r \in \{0,1\}, k \in \{0,1, . . , p\}\}$ را فضای حالت هر یک از بازیکنان در نظر بگیرید، که در آن r بیان کننده شهرت بازیکن و k تعداد مراحل باقی مانده از تنبیه بازیکن متخلف است. رابطه (۳) تابع انتقال حالت را تعریف می‌کند.

$$t(1,0,a) = \begin{cases}(1,0) & a \in B(s_i^*)\\(0,p) & a \notin B(s_i^*)\end{cases}$$
$$t(0,k) = \begin{cases}(0,k-1) & k\neq 1\\(1,0) & k=1\end{cases} \tag{۳}$$

که $B(s_i^*)$ مجموعه اعمال براساس استراتژی s_i^* است. اکنون با توجه به تابع انتقال t، استراتژی s_i^* را می‌توان به صورت رابطه (۴) بیان نمود.

$$s_i^*(r_i, r_j) = \begin{cases}C & r_i, r_j = 1\\D & \text{otherwise}\end{cases} \tag{۴}$$

که در آن j شناسه بازیکن مقابل i است.

قضیه ۱: پروفایل استراتژی s^* یک تعادل زیربازی ـکامل است، اگر

$$\delta^{p-1} \le \frac{p_e}{-1+\delta+p_e}\left(1 - \frac{c(1-\delta)}{\delta(s-c+d)}\right) \tag{۵}$$

و رابطه

$$\max\left(\frac{c}{s+d}, 1 - p_e\right) < \delta \tag{۶}$$

در شبکه صادق باشد. که در این روابط p بیانگر طول دوره تنبیه است و p_e احتمال روبرو شدن دو بازیکن خاص در یک مرحله از بازی با توجه به قانون نظیریابی شبکه است.

اثبات: از آنجا که مدل ارائه شده از نوع بازی‌های تکراری با فاکتور تخفیف بوده و بهره بازی مرحله محدود است، قضیه تخطی یک مرحله‌ای[10] می‌تواند برای اثبات استفاده شود [۹]. بنابراین، برای اثبات قضیه کافی است نشان دهیم برای هر یک از بازیکنان، یک مرحله تخلف از استراتژی در هیچ یک از زیربازی‌ها سودمندانه نخواهد بود. برای انجام این کار باید همه زیربازی‌های ممکن و همه حالت‌های تخلف‌های یک مرحله‌ای از آن را بررسی کنیم.

براساس رابطه (۲)، سابقه این بازی در مقدار عدد شهرت بازیکنان خلاصه می‌شود و استراتژی s_i^* نیز براساس مقادیر شهرت بازیکنان عمل می‌کند. بنابراین مجموعه زیربازی های ممکن برابر است با $H^t = \times_{i=1}^{2n} H_i^t$. لذا تعداد زیربازی های ممکن برابر است با $(2n)^2$. با در نظر گرفتن شهرت دو بازیکن مقابل هم، چهار حالت مختلف برای سابقه بازی خواهیم داشت. به عنوان مثال، زیربازی $(0,1)$ یعنی تا این مرحله از بازی، $r_1^t = 0$ و $r_2^t = 1$ است. در ضمن در هر یک از این مجموعه زیربازی‌ها اثر شهرت سایر بازیکنان را نیز مورد بررسی قرار خواهیم داد.

زیربازی $(1,1)$: فرض کنید در زمان t، $r_1^t = 1$ و $r_2^t = 1$ باشد. کمترین مقدار بهره برای تبعیت از استراتژی وقتی است که همه بازیکنان غیر از 1 و 2 متخلف باشند. فرض کنید احتمال اینکه در یک مرحله از بازی دو بازیکن مشخص با یکدیگر روبرو شوند برابر p_e باشد. در صورتی که بازیکن 1 با بازیکنان متخلف بازی کند بهره متوسط بازی مرحله او برابر $-\frac{1}{2} * d$ می‌شود و این پیشامد با احتمال $1 - p_e$ رخ می‌دهد.

با فرض اینکه همه بازیکنان پس از زمان t از استراتژی مورد نظر پیروی می‌کنند حداکثر پس از p مرحله دوره تنبیه بازیکنان متخلف به پایان خواهد رسید. سمت چپ نامعادله (۷) بهره مورد انتظار استراتژی تعادل در بازه زمانی $[t, t + p]$ در بدترین حالت و سمت راست آن بهره مورد انتظار تخلف از استراتژی را در این بازه نشان می‌دهد. برای سادگی کار مقادیر مشترکی که در هر دو طرف نامعادله وجود داشتند مانند بهره مربوط به زمان بعد از $t + p$ را حذف کردیم.

$$\left[p_e\left(\frac{1}{2}*s-\frac{1}{2}*c\right)+(1-p_e)\left(-\frac{1}{2}*d\right)\right]\delta^t + \sum_{\tau=t+1}^{t+p-1}\delta^\tau\left[p_e\left(\frac{1}{2}*s-\frac{1}{2}*c\right) + (1-p_e)\left(-\frac{1}{2}*d\right)\right]+\frac{1}{2}*(s-c)\delta^{t+p} > \left[p_e\left(\frac{1}{2}*s\right)+(1-p_e)\left(-\frac{1}{2}*d\right)\right]\delta^t+\sum_{\tau=t+1}^{t+p}\delta^\tau\left(-\frac{1}{2}*d\right) \tag{۷}$$

$$\delta^{p-1} \le \frac{p_e}{-1+\delta+p_e}\left(1-\frac{c(1-\delta)}{\delta(s-c+d)}\right) \tag{۸}$$

همانطور که مشاهده می‌شود نامعادله (۸) با رابطه قضیه ۱ معادل است. از آنجا که δ^p مقداری بزرگتر از صفر است لذا می‌بایست سمت راست نامعادله (۸) > 0 باشد. با توجه به شرط ذکر شده در قضیه داریم:

$$\max\left(\frac{c}{s+d}, 1-p_e\right) < \delta$$

یعنی $1-p_e<\delta$ ، $\frac{c}{s+d}<\delta$ و

$$1 > \frac{c(1-\delta)}{\delta(s-c+d)} \qquad (۹)$$

که این شرایط مثبت بودن سمت راست نامعادله (۸) را تضمین می‌کند.

زیربازی $(1,0)$: فرض کنید در زمان t، $r_1^t = 0$ و $r_2^t = 1$ باشد. در این صورت اگر سرویس‌دهنده بر طبق استراتژی تعادل رفتار کند، باید سرویس‌گیرنده را تنبیه کرده و به او سرویس ندهد که در این صورت هزینه‌ای برای سرویس دادن متحمل نمی‌شود ولی اگر برخلاف استراتژی تعادل عمل کرده و همکاری را انتخاب کند، باید به اندازه c هزینه بدهد. بدیهی است تخلف از استراتژی در این زیربازی تصمیم عاقلانه‌ای نخواهد بود. زیرا $-c < 0$.

زیربازی $(p_1, 0,1)$: فرض کنید در زمان t، $r_1^t = 1$ ، $r_2^t = 0$ و p_1 واحد زمانی از شروع تنبیه او گذشته باشد. توجیه این حالت نیز مانند زیربازی فوق خواهد بود. زیربازی $(p_1, 0,0)$ نیز مشابه حالت $(p_1, 0,1)$ تحلیل می‌شود.

قضیه فوق نشان می‌دهد استراتژیهای مشابه TFT[11] که برای بازی‌های دو نفره تکراری تعادل است در بازی نظیریابی تصادفی با بازی مرحله مشابه بازی تکراری لزوما تعادل نیست. زیرا در بازی نظیریابی تصادفی وضعیت رفتاری سایر بازیکنان در انتخاب استراتژی افراد موثر است که در بازیهای تکراری دو نفره این محدودیت و وابستگی وجود ندارد.

۴-۲- استراتژی توبه

در این بخش استراتژی جدیدی ارائه می‌دهیم که زمان تنبیه آن به ازای هر تخلف کمتر از استراتژی تنبیه باشد. این خاصیت باعث می‌شود نرخ همکاری و کارایی شبکه و بنابراین بهره جمعی افزایش یابد. در استراتژی تنبیه، از نقطه min-max تابع برای تنبیه بازیکن متخلف استفاده شد. بازیکن خوش‌رفتار در صورت مواجه شدن با بازیکن متخلف عملی را انتخاب می‌کرد که بازیکن متخلف در صورت انتخاب بهترین استراتژی کمترین بهره ممکن را به دست بیاورد. در این شرایط با فرض ثابت بودن بازی طرف مقابل، بهترین عمل ممکن برای بازیکن تنبیه شده بازی min-max است.

به این ترتیب بازیکن متخلف در طول دوره تنبیه انگیزه‌ای برای تخلف از استراتژی تنبیه مورد نظر ندارد. زیرا در صورت تخلف از آن جریمه بیشتری باید بپردازد که برایش سودمندانه نخواهد بود.

در این بازی می‌توانیم به جای تنبیه با نقطه min-max تابع از تنبیه دیگری استفاده کنیم که سختگیرانه تر باشد ولی در عوض طول دوره تنبیه کوتاهتر شود. با این استراتژی فرد متخلف در طول دوره تنبیه موظف است با همکاری بدون دریافت خدمت از شبکه جریمه تخلف خود را بپردازد.

نکته‌ای که در این استراتژی باید به آن توجه نمود امکان تخلف سودمندانه فرد متخلف از استراتژی است. از آنجایی که تنبیه در این استراتژی از تنبیه با نقطه min-max سختگیرانه‌تر است، بنابراین فرد متخلف با انتخاب بازی نقطه min-max تابع می‌تواند از گذراندن تنبیه سخت‌تر اجتناب کند و جریمه کمتری بپردازد. برای حل این مسئله باید تنبیه جدیدی برای تخلف در طول دوره تنبیه در نظر گرفته شود تا سود حاصل از این تخلف را خنثی کند.

مجموعه $Z = \{(r,k) | r \in \{0,1\}, k \in \{0,1,..,p\}\}$ را در نظر بگیرید. رابطه (۱۰) تابع انتقال حالت را تعریف می‌کند.

$$t(1,0,a) = \begin{cases} (1,0) & a \in B(s_i') \\ (0,p) & a \notin B(s_i') \end{cases}$$

$$t(0,k,a) = \begin{cases} (0,k-1) & a \in B(s_i') \wedge k \neq 1 \\ (1,0) & a \in B(s_i') \wedge k = 1 \\ (0,p) & a \notin B(s_i') \end{cases} \qquad (۱۰)$$

که $B(s_i')$ مجموعه اعمال براساس استراتژی s'_i است. اکنون با توجه به تابع انتقال t، استراتژی s_i' را می‌توان به صورت رابطه (۱۱) بیان نمود.

$$s_i'(r_i, r_j) = \begin{cases} C & r_j = 1 \\ D & \text{otherwise} \end{cases} \qquad (۱۱)$$

که در آن j شناسه بازیکن مقابل i است. به عبارت دیگر استراتژی متقارن s' برای بازیکن i چنین تجویز می‌کند:

s_i': در مرحله t، j را با شناسه بازیکن طرف مقابلت جایگزین کن و به این صورت عمل کن: اگر $r_j^t = 1$ همکاری کن در غیر اینصورت همکاری نکن. بازیکن متخلف به اندازه p دوره زمانی تنبیه می‌شود. یعنی از زمان t تا $t+p$ نمی‌تواند خدمات دریافت کند. در صورتی که بازیکن متخلف در طول دوره تنبیه از استراتژی سرپیچی کرد دوره تنبیه او تا p مرحله بعد افزایش می‌یابد.

قضیه ۲: پروفایل استراتژی s_i' یک تعادل زیربازی-کامل است، اگر

$$\frac{c}{s+d} \le \delta^p \le \frac{\delta(s+d+c)-c}{\delta(s+d)} \qquad (۱۲)$$

مشروط بر این که فاکتور تخفیف در رابطه

$$\frac{c}{s+d} \le \delta$$

صدق کند. که در این روابط p بیانگر طول دوره تنبیه است.

اثبات: برای بررسی استراتژی، زیربازی‌ها را براساس شهرت بازیکن i به دو دسته زیربازی‌هایی که در آنها $r_i^t = 1$ و زیربازی هایی که در آنها $r_i^t = 0$ است تقسیم می‌کنیم.

زیربازی $r_i^t = 1$: فرض کنید در زمان t، $r_i^t = 1$ باشد. براساس فرض اولیه بازی بازیکن i در این مرحله با احتمال $\frac{1}{2}$ در نقش سرویس‌دهنده و با همین احتمال نیز در نقش سرویس‌گیرنده بازی خواهد کرد.

اگر در زمان t، بازیکن i سرویس‌دهنده باشد و بر طبق استراتژی رفتار نماید، کمترین مقدار بهره‌ای که می‌تواند در این مرحله به دست بیاورد در حالتی است که بازیکن رقیب او j و $r_j^t = 1$ باشد. زیرا در این صورت i همکاری را انتخاب کرده و باید به اندازه c هزینه کند. ولی در غیر اینصورت یعنی اگر $r_j^t = 0$ باشد بازیکن i نباید همکاری کند.

اگر بازیکن i در این مرحله سرویس‌گیرنده باشد، با فرض تبعیت سایر بازیکنان از استراتژی، سرویس مورد نظرش را از بازیکن j دریافت خواهد کرد و به اندازه s واحد سود می‌کند. این رفتار به مقدار شهرت بازیکن j وابسته نیست چون بر اساس استراتژی در هر دو صورت j باید همکاری را انتخاب کند.

اگر احتمال اینکه در زمان τ، $r_j^t = 1$ باشد را با $\pi(\tau)$ نشان دهیم مقدار مورد انتظار بهره‌ای که بازیکن i می‌تواند در این زیربازی با تبعیت از استراتژی به دست بیاورد برابر است با

$$\frac{1}{2}\left(s - c * \pi(0)\right)\delta^t + \sum_{\tau=t+1}^{\infty} \delta^\tau \frac{1}{2}\left(s - c * \pi(\tau)\right).$$

بعد از p دوره تنبیه بازیکنان متخلف با فرض تبعیت از استراتژی تمام شده و شهرت همه بازیکنان برابر با ۱ می‌شود. بنابراین به ازای $\tau \ge t+p$ ، $\pi(\tau) = 1$ است.

اگر بازیکن i تخلف از استراتژی را انتخاب کند. در این صورت مقدار مورد انتظار بهره‌ای که می‌تواند به دست آورد مطابق زیر محاسبه می‌شود.

$$\frac{1}{2}s\delta^t+\sum_{\tau=t+1}^{t+p}\delta^\tau\frac{1}{2}\big(-d-c*\pi(\tau)\big)+\sum_{\tau=t+p+1}^{\infty}\delta^\tau\frac{1}{2}(s-c)$$

بازه تنبیه باید طوری طراحی شود که سود حاصل از تخلف را خنثی نماید. بنابراین باید داشته باشیم.

$$\begin{aligned}&\frac{1}{2}\big(s-c*\pi(0)\big)\delta^t+\sum_{\tau=t+1}^{\infty}\delta^\tau\frac{1}{2}\big(s-c*\pi(\tau)\big)\\&\geq\frac{1}{2}s\delta^t+\sum_{\tau=t+1}^{t+p}\delta^\tau\frac{1}{2}\big(-d-c*\pi(\tau)\big)+\sum_{\tau=t+p+1}^{\infty}\delta^\tau\frac{1}{2}(s-c)\\&\delta^p\leq\max\left(1-\frac{(1-\delta)c*\pi(0)}{(s+d)\delta}\right)=1-\frac{(1-\delta)c}{(s+d)\delta}\end{aligned}\tag{۱۳}$$

همانطور که مشاهده می‌کنید نامعادله (۱۳) با سمت راست رابطه (۱۲) در قضیه ۲ معادل است.

زیربازی0 $r_i^t=$: فرض کنید در زمان t، $r_i^t=0$ باشد. اگر بازیکن بر طبق استراتژی رفتار نماید، مقدار بهره مورد انتظاری که می‌تواند در این مرحله به دست بیاورد برابر است با

$$\textstyle\sum_{\tau=t}^{t+p-1}\delta^\tau\frac{1}{2}\big(-d-c*\pi(\tau)\big)+\sum_{\tau=t+p}^{\infty}\delta^\tau\frac{1}{2}(s-c)\tag{۱۴}$$

و مقدار بهره مورد انتظاری که بازیکن i با تخلف از استراتژی می‌تواند به دست بیاورد برابر است با

$$\begin{aligned}&\tfrac{1}{2}(-d)\delta^t+\textstyle\sum_{\tau=t+1}^{t+p-1}\delta^\tau\frac{1}{2}\big(-d-c*\pi(\tau)\big)\\&+\delta^{t+p}\frac{1}{2}(-c-d)\\&+\sum_{\tau=t+p+1}^{\infty}\delta^\tau\frac{1}{2}(s-c)\end{aligned}\tag{۱۵}$$

بازه تنبیه باید طوری طراحی شود که سود حاصل از تخلف را خنثی نماید. بنابراین باید داشته باشیم (۱۵) $\geq$ (۱۴). به عبارت دیگر داریم.

$$\delta^p\geq\max\left(\tfrac{c*\pi(0)}{s+d}\right)=\tfrac{c}{s+d}\tag{۱۶}$$

مشاهده می‌کنید که نامعادله (۱۶) با سمت چپ رابطه (۱۲) در قضیه ۲ معادل است. برای اینکه حداقل یک جواب معتبر برای دوره تنبیه وجود داشته باشد، باید نامساوی

$$0<\frac{c}{s+d}\leq\delta^p\leq\frac{\delta(s+d+c)-c}{\delta(s+d)}<1$$

نیز برقرار باشد. این نامساوی با توجه به فرض قضیه و فرضیات اولیه بازی همواره صحیح است.

۴-۳- استراتژی توبه بهینه

در این بخش استراتژی دیگری برای بازی مکانیزم انگیزشی ارائه می‌کنیم که مشابه استراتژی توبه است با این تفاوت که این استراتژی افزایش h واحد زمانی به بازه تنبیه به ازای هر واحد تخلف در طول دوره تنبیه را پیشنهاد می‌کند. h مقداری ثابت و مثبت و $h\geq1$ است.

افزایش چند مرحله تنبیه عادلانه تر از تجدید تنبیه است. زیرا در استراتژی توبه مقدار تنبیه برای کسی که در مرحله اول تنبیه از p دوره قرار دارد با کسی که در مرحله pام قرار دارد متفاوت است. در حالی که سود حاصل از تخلف برای هر دو یکی بوده است.

به عبارت دیگر مقدار تنبیه متناسب با مقدار سود تخلف نیست و علاوه بر آن به یک عامل بیرونی دیگر نیز وابسته است. در حالی که در این استراتژی، تنبیه کسی که در حال تنبیه تخلف می‌کند، متناسب با مقدار سودی است که از آن تخلف به دست می‌آورد.

استراتژی جدید را به کمک تعریف فضای حالت Z شرح می‌دهیم. رابطه (۱۷) تابع انتقال حالت را تعریف می‌کند.

$$\begin{aligned}&t(1,0,a)=\begin{cases}(1,0) & a\in B(s_i')\\(0,p) & a\notin B(s_i')\end{cases}\\&t(0,k,a)\\&(0,k,a)=\begin{cases}(0,k-1) & a\in B(s_i')\ \wedge\ k\neq1\\(1,0) & a\in B(s_i')\ \wedge\ k=1\\(0,k+h) & a\notin B(s_i')\end{cases}\end{aligned}\tag{۱۷}$$

اکنون با توجه به تابع انتقال t، استراتژی s_i'' را می‌توان به صورت زیر بیان نمود.

$$s_i''(r_i,r_j)=\begin{cases}C & r_j=1\\D & \text{otherwise}\end{cases}\tag{۱۸}$$

قضیه ۳: پروفایل استراتژی s_i'' یک تعادل زیربازی-کامل است، اگر

$$\tfrac{c}{s+d}\leq\delta^p\leq\tfrac{\delta(s+d+c)-c}{\delta(s+d)}\tag{۱۹}$$

مفروض بر این که فاکتور تخفیف δ در نامساوی

$$\frac{c}{(s+d)}\leq\delta$$

صدق کند. که در این روابط p بیانگر طول دوره تنبیه است.

اثبات: قضیه ۳ نیز مشابه قضیه ۲ اثبات می‌شود. با این تفاوت که مقادیر بهره در زیربازی $r_i^t=0$ و در نتیجه نامعادله (۷) تغییر می‌کند. بنابراین باید حداقل بازه لازم برای تنبیه و شرایط فاکتور تخفیف بر مبنای استراتژی جدید از ابتدا محاسبه و تعیین شوند.

زیربازی $r_i^t=1$: فرض کنید در زمان t، $r_i^t=1$ باشد. مشابه اثبات قضیه ۲ با نوشتن روابط لازم به نامساوی زیر می‌رسیم.

$$\delta^p\leq\tfrac{\delta(s+d+c)-c}{\delta(s+d)}\tag{۲۰}$$

زیربازی $r_i^t=0$: توجیه شهودی و اولیه برای درستی استراتژی در این زیر بازی به این صورت است که در این بازی همواره بیشترین تنبیه ممکن برای یک بازیکن خوش‌رفتار از بیشترین سود حاصل از تخلف بازیکن متخلف از تنبیه بیشتر است. بنابراین اگر فاکتور تخفیف به اندازه کافی بزرگ انتخاب شود یک مرحله تنبیه در انتهای دوره تنبیه فعلی می‌تواند سود حاصل از تخلف در دوره تنبیه را خنثی کند.

فرض کنید در زمان t، $r_i^t=0$ باشد و تعداد p_1 واحد از بازه تنبیه او باقی مانده باشد. ابتدا ثابت می‌کنیم قضیه برای $h=1$ یعنی استراتژی با افزایش یک

واحد تنبیه برای فرد متخلف تعادل است. به طور شهودی برای $h \geq 1$ نیز قضیه برقرار خواهد بود.

نامعادله (۲۱) مقدار مورد انتظار بهره ای که بازیکن i با پیروی از استراتژی می‌تواند به دست بیاورد را نشان می‌دهد.

$$\sum_{\tau=t}^{t+p_1-1} \delta^{\tau}\frac{1}{2}\left(-d - c*\pi(\tau)\right) + \sum_{\tau=t+p_1}^{\infty} \delta^{\tau}\frac{1}{2}\left(s - c*\pi(\tau)\right) \quad (۲۱)$$

اگر بازیکن i تخلف از استراتژی را انتخاب کند. در این صورت بهره مورد انتظار او برابر است با

$$\frac{1}{2}(-d)\delta^{t} + \sum_{\tau=t+1}^{t+p_1} \delta^{\tau}\frac{1}{2}\left(-d - c*\pi(\tau)\right) + \sum_{\tau=t+p_1+1}^{\infty} \delta^{\tau}\frac{1}{2}\left(s - c*\pi(\tau)\right) \quad (۲۲)$$

بازه تنبیه باید طوری طراحی شود که سود حاصل از تخلف را خنثی نماید. بنابراین باید داشته باشیم (۲۲) $\geq$ (۲۱).

$$\begin{aligned}\left(-d - c*\pi(0)\right) &+ \sum_{\tau=1}^{p_1-1} \delta^{\tau}\left(-d - c*\pi(\tau)\right) + \delta^{p_1}\left(s - c*\pi(p_1)\right) \\ &\geq -d + \sum_{\tau=1}^{p_1-1} \delta^{\tau}\left(-d - c*\pi(\tau)\right) + \delta^{p_1}\left(-d - c*\pi(p_1)\right) \\ \delta^{p_1} &\geq \max\left(\frac{c*\pi(0)}{s+d}\right) = \frac{c}{s+d}\end{aligned} \quad (۲۳)$$

با جایگذاری مقادیر ابتدایی و انتهایی بازه $p_1 = \{1,..,p\}$ در نامعادله (۲۳) روابط (۲۴) و (۲۵) بدست می‌آیند.

$$\delta^{p} \geq \frac{c}{s+d} \quad (۲۴)$$

$$\delta \geq \frac{c}{s+d} \quad (۲۵)$$

با ترکیب نامعادلات (۲۴) و (۲۰) نامساوی قضیه ۳ حاصل می‌شود. نامعادله (۲۵) نیز براساس فرض قضیه صحیح است.

۵- تحلیل و ارزیابی

برای مدلسازی سیستم شهرت در شبکه های همکاری، بازی نظیریابی تصادفی نسبت به بازیهای تکراری دو نفره مناسبتر بوده و شباهت و سازگاری بیشتری با محیط سیستم شهرت دارد. همانطور که مشاهده شد، استراتژی تعادلی که در بازی دونفره تکراری صادق است در مدل نظیریابی تصادفی با همان بازی مرحله، تعادل نیست. چون در بازی نظیریابی تصادفی علاوه بر سابقه دو بازیکن رقیب، وضعیت شهرت سایر بازیکنان نیز در بهره و انتخاب استراتژی آنها موثر است.

مدل ارائه شده و استراتژیهای پیشنهادی از نظر پیاده‌سازی ساده هستند و حافظه مصرفی و توان محاسباتی کمی لازم دارند. پیچیدگی حافظه آن از مرتبه n و پیچیدگی محاسبه آن از مرتبه ۱ است. لذا برای پیاده سازی در شبکه های همکاری که ممکن است از نظر سخت‌افزاری امکانات محدودی داشته باشند، مناسب خواهند بود.

در استراتژی تنبیه بازیکن متخلف در طول دوره تنبیه ملزم به همکاری نیست. بنابراین با افزایش تعداد بازیکنان متخلف کارایی شبکه کاهش می‌یابد. که باعث می‌شود بهره بازیکنان خوش‌رفتار نیز کم شود. این مسئله باعث کاهش کارایی شبکه و کاهش بهره جمعی می‌شود. به همین علت این استراتژی در عمل کارایی چندانی نخواهد داشت.

از طرف دیگر استراتژی تنبیه به مقدار p_e و قانون نظیریابی شبکه وابسته است که با زیاد شدن گره‌های شبکه این احتمال کمتر شده و دوره تنبیه طولانی‌تر می‌شود.

در استراتژی توبه به جای تنبیه با نقطه min-max از تنبیه دیگری استفاده می‌کنیم که سختگیرانه‌تر باشد ولی در عوض طول دوره تنبیه کوتاهتر شود. با کوتاهتر شدن طول دوره تنبیه فرد متخلف زودتر می‌تواند وارد تعامل با بازیکنان شده و از خدمات شبکه استفاده کند. بنابراین این استراتژی باعث افزایش کارایی شبکه و در نتیجه افزایش بهره جمعی بازیکنان می‌شود.

استراتژی توبه بهینه با افزایش یک مرحله تنبیه به جای تجدید تنبیه، عادلانه‌تر از استراتژی توبه عمل می‌کند. در این استراتژی، تنبیه کسی که در حال تنبیه تخلف می‌کند، متناسب با مقدار سودی است که از تخلف به دست می‌آورد.

مهمترین مسئله‌ای که می‌تواند پیاده‌سازی این مدل را با چالش روبرو کند، عدم انگیزه کافی در گره‌ها برای ارسال گزارش اطلاعات شهرت به واحد شهرت مرکزی است. ارسال گزارش به مرکز شهرت نیازمند هزینه‌ای هر چند اندک از سمت گره فرستنده خواهد بود. بنابراین گره‌ها بدون یک مکانیزم تشویق و یا تنبیه، انگیزه کافی برای صرف این هزینه و ارسال اطلاعات به مرکز را ندارند. برای حل این مشکل باید در سمت مرکز شهرت، قانونی تنظیم شود که به موجب آن استفاده از اطلاعات آن مرکز تنها برای اعضایی که گزارش شهرت ارسال می‌کنند امکان‌پذیر باشد.

سادگی مدل ارائه شده در تعریف مجموعه مقادیر ممکن برای متغیر شهرت، از ایرادات دیگری که در نگاه اول ممکن است به این مکانیزم وارد شود. شهرت در این مدل یک متغیر گسسته است که تنها مقادیر صفر و یک را می‌تواند بپذیرد.

برای توجیه این مسئله لازم است ابتدا در مورد علت و فایده استفاده از شهرت پیوسته توضیحاتی داده شود. به طور کلی هنگامی استفاده از شهرت پیوسته مفید است که استراتژی مسئله متناسب با شهرت پیوسته باشد. همچنین مسئله نیز طوری باشد که بالابردن شهرت عاقلانه و سودمندانه باشد.

به منظور بیان دقیق‌تر، این مفهوم را در قالب قضیه ۴ بیان کرده و با استفاده از اصول و قضایای نظریه بازیها اثباتی برای آن ارائه می‌کنیم. بازی $G = \{N, A, U\}$ را در نظر بگیرید. مجموعه اعمال ممکن برای هر بازیکن برابر است با $A = [C, D]$ که در آن منظور از C و D به ترتیب همکاری و عدم همکاری است. تابع بهره U را مشابه تابع بهره بازی معمای زندانی در نظر بگیرید. هر بازیکن یک برچسب شهرت r دارد. متغیر شهرت r به بازه پیوسته $R = [a, b]$ تعلق دارد. $f: R * A \rightarrow R$ تابع محاسبه شهرت و $r_{t+1} = f(r_t, a_t)$ است.

استراتژی مطلوب در این بازی C است. از عمل D برای تنبیه بازیکن متخلف استفاده می‌شود. عمل همکاری عامل افزایش شهرت و عدم همکاری باعث کاهش آن می‌شود. همکاری و بنابراین افزایش شهرت هزینه‌بر است.

در روش‌های احتمالاتی برای حل مسئله معمولا یک حد آستانه در بازه شهرت انتخاب کرده و شهرت بالاتر از آن را به شهرت خوب و پایینتر از آن را به شهرت بد تفسیر می‌کنند. استراتژی به این صورت است که با افراد خوب همکاری کن و در

برابر افراد بد عدم همکاری را انتخاب کن. یعنی افرادی که شهرتشان زیر حد آستانه است تنبیه می‌شوند.

در شبکه‌های همکاری دادن سرویس فقط هزینه دارد و گره‌ها تمایلی به دادن سرویس ندارند. بنابراین در شبکه‌های همکاری اگر استراتژی سیستم شهرت، مبتنی بر حد آستانه شهرت عمل کند، گره‌ای که شهرتش خیلی بالا باشد ضرر می‌کند. چون بالا بردن شهرت مستلزم صرف مقداری هزینه است که با توجه به فرضیات بازی این هزینه از جای دیگری تامین نخواهد شد. در این شرایط استراتژی بهینه این است که گره ها شهرتشان را در حد مقدار آستانه نگه دارند که بتوانند از شبکه خدمات دریافت کنند و حداقل هزینه را برای بالا نگه‌داشتن شهرتشان بپردازند. چون در غیر این صورت بیشتر از آنچه از شبکه خدمات دریافت می‌کنند باید به بقیه سرویس بدهند.

قضیه ۴: در بازی G، کارایی استراتژی تعادل مبتنی بر روش حد آستانه در بازه پیوسته شهرت با استراتژی تعادل مبتنی بر شهرت باینری یکسان است.

اثبات: فرض کنید حد آستانه شهرت θ طوری انتخاب شده باشد که خسارت ناشی از کاهش شهرت به مقدار کمتر از θ سود حاصل از تخلف را خنثی کند. اگر داشته باشیم $r_t^i > \theta$، $f(r_t^i, C) > \theta$ و $f(r_t^i, D) > \theta$. در این صورت، بهترین عملی که بازیکن i می‌تواند انتخاب کند D است چون مجبور نیست هزینه لازم برای عمل C را بپردازد و با توجه به مقدار r_{t+1}^i بهره ادامه بازی نیز در دو حالت یکسان است. بنابراین، با این شرایط تخلف از تعادل برای او سودمندتر از پیروی از آن است. برای اینکه این استراتژی نش باشد باید $f(r_t, D) < \theta, \forall r_t \in R$. یعنی به ازای هر تخلف شهرت فرد هر چه باشد به شهرت کمتر از حد آستانه یعنی شهرت بد تبدیل شود. مدت زمانی که یک فرد متخلف لازم دارد تا بتواند شهرتش را به حد آستانه برساند معادل دوره تنبیه در مکانیزمهای باینری است. در حقیقت این استراتژی معادل استراتژی شهرت باینری است و همان کارایی را دارد.

۶- کارهای مرتبط

علاوه بر استفاده نظریه بازیها برای مدلسازی سیستم شهرت مفهوم شهرت به طور مستقل در نظریه بازیها کاربرد دارد. تحقیقات تئوری در زمینه شهرت در نظریه بازیها با کارهای [۱۰-۱۲] شروع شد که نشان دادند چطور مقدار کمی اطلاعات ناکامل برای تولید اثر شهرت در یک بازی معمای زندانی تکراری و محدود، کافی است. و سپس این ایده با کارهای [۱۳، ۱۴] ادامه یافت. در سال‌های اخیر نیز مقاله های زیادی در این رابطه ارائه شده است.

در مقالات معمولا بازی شهرت را از نوع بازیهای تکراری معمولی و بدون درنظر گرفتن ماهیت متغیر بودن بازیکنان رقیب در هر دور در نظر می‌گیرند. به عنوان مثال [۱۵] بازی شهرت در شبکه‌های نظیربه‌نظیر را با بازی تکراری شبیه به معمای زندانی مدل کرده است.

نویسنده ادعا می‌کند در شرایطی که تغییر شناسه برای بازیکنان امکان پذیر باشد و برخی بازیکنان به قصد سود بیشتر از حمله لاپوشانی استفاده کنند استراتژی TFT یک استراتژی تکاملی پایدار[۱۲] نیست. و شرایطی تعیین می‌کند که تحت آن استراتژی TFT در برابر حمله لاپوشانی مقاوم باشد. این استراتژی و اثبات آن در بازی تعریف شده در مقاله صحیح است اما مدل ارائه شده اساسا با سیستم شهرت واقعی اختلاف دارد. در این مقاله نشان دادیم که استراتژی معروف TFT[۱۳] در بازی تکراری از نوع نظیریابی تصادفی استراتژی تعادل نیست.

برخی از مقالاتی هم که این ویژگی را در مدل سیستم شهرت مورد توجه قرار داده‌اند، به جای تحلیل ریاضی و اثبات با استفاده از قضایای نظریه بازیها از شبیه سازی برای برای پیش‌بینی نتیجه بازی استفاده کرده‌اند [۱۶].

مرجع [۱۷] یکی از مهمترین کارهایی است که از بازی نظیریابی تصادفی برای حل مساله مسیریابی در شبکه‌های همکاری، با فرض وجود یک سیستم شهرت، استفاده می کند. اما استراتژی پیشنهاد شده برخلاف ادعای مقاله، یک استراتژی تعادل زیربازی - کامل نیست. زیربازی که در آن بازیکن متخلف در حال تنبیه تقاضایی برای ارائه سرویس دریافت می‌کند مثال نقضی برای این موضوع است.

استراتژی پیشنهادی [۱۷] به این شرح است: اگر بازیکن رقیب خوش‌رفتار بود همکاری کن در غیر اینصورت همکاری نکن. برای تخلف از این استراتژی یک روش تنبیه مبتنی بر زمان ارائه می‌کند. به این صورت که وقتی یک گره از این استراتژی تخطی کرد برای بازه زمانی τ تنبیه می‌شود. اگر هنگامی که در دوره تنبیه قرار دارد دوباره تخطی کند، بازه تنبیه او از ابتدا در نظر گرفته می‌شود. در طول دوره تنبیه هر تقاضایی که از سمت این گره به شبکه ارسال شود بی‌پاسخ می‌ماند اما این گره باید به تقاضای گره‌های خوش‌رفتار در شبکه پاسخ بدهد.

مطابق استراتژی مذکور بازیکن متخلف باید تا انتهای دوره تنبیه خود، بدون انتظار دریافت سرویس از شبکه و سایر گره‌ها، همکاری کند. در حالی که با نوشتن روابط مربوط به بهره بازیکن در این زیربازی، برای تبعیت از استراتژی و تخلف از آن معلوم می‌شود تخلف از استراتژی سودمندتر از تبعیت از آن برای بازیکن متخلف خواهد بود. بنابراین استراتژی بیان شده یک استراتژی تعادل زیربازی ـ کامل نیست.

۷- نتیجه‌گیری

مدل‌های مبتنی بر نظریه بازی‌ها از قابلیت و دقت بیشتری نسبت به سایر مدل‌های احتمالاتی برخوردارند. زیرا اگر سیستم به درستی طراحی شود و همکاری در یک نَش از بازی قرار گیرد، بازیکنان عاقل رفتار مورد انتظار طراح بازی را انجام می‌دهند.

در این مقاله یک مکانیزم انگیزشی با فرض وجود یک سیستم شهرت مرکزی برای شبکه‌های همکاری توزیع شده و با رویکرد نظریه بازی‌ها ارائه شد. مدل ارائه شده کلی، ساده و بنابراین براحتی قابل گسترش برای انواع مختلف شبکه‌های همکاری با جزئیات بیشتر خواهد بود. همچنین به دلیل فرض اولیه تصادفی بودن نظیریابی بازیکنان در بازی نظیریابی تصادفی، سازگاری بیشتری با ماهیت پویا و در حال تغییر شبکه‌های همکاری نسبت به بازی‌های تکراری معمولی دارد. استراتژی‌های پیشنهاد شده بازیکنان را به همکاری در شبکه تشویق می‌کند و در صورت تخلف برخی از بازیکنان از استراتژی نیز، نتیجه تعادل همچنان به همکاری منتهی خواهد شد.

تشکر و قدردانی

تحقیق حاضر توسط مرکز تحقیقات مخابرات ایران تحت قرارداد ۵۰۰/۱۷۴۱۴/ت در تاریخ ۱۳۸۸/۱۲/۱ پشتیبانی شده است.

مراجع

[1] L. Camarinha-Matos, and H. Afsarmanesh, "Collaborative Networks: A New Scientific Discipline," *Intelligent Manufacturing*, vol. 16, pp. 439-452, 2005.

[2] <u>*Http://Www.Ipoque.Com/Resources/Internet-Studies*</u>.

[3] L. Mui, M. Mohtashemi, and A. Halberstadt, *AComputational Model of Trust and Reputation*, pp. 2431-2439, 2002.

[4] D. Fudenberg, and J. Tirole, *Game Theory*. Cambridge: MIT press, 1991.

[5] R. W. Rosenthal, and H. J. Landau, "A Game-Theoretic Analysis of Bargaining with Reputations," *Mathematical Psychology*, vol. 20, pp. 233-255, 1979.

[6] R. W. Rosenthal, "Sequences of Games with Varying Opponents," *Econometrica: Econometric Society*, vol. 47, pp. 1353-1366, 1979.

[7] M. Kandori, "Social Norms and Community Enforcement," *Rev. Economic Studies*, vol. 5, pp. 63-80, 1992.

[۸] ف. بهرامیان, ح . شهریاری، "یک مکانیزم انگیزشی برای شبکه‌های همکاری با استفاده از سیستم شهرت مرکزی و بازی‌های تکراری،" ارائه شده در شانزدهمین کنفرانس ملی سالانه انجمن کامپیوتر ایران، تهران، ۱۳۸۹.

[9] D. Abreu, "On the Theory of Infinitely Repeated Games with Discounting", *Econometrica*, pp. 383-396, 1988.

[10] P. Milgrom, and J. Roberts, "Predation, Reputation and Entry Deterrence," *Journal of economic theory*, vol. 27, pp. 280-312, 1982.

[11] J. Andreoni, and J. Miller, "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence," *The Economic Journal*, pp. 570-585, 1993.

[12] D. Kreps, and R. Wilson, "Reputation and Imperfect Information," *Readings in Industrial Organization*, vol. 27, p. 378, 2000.

[13] K. Schmidt, "Reputation and Equilibrium Characterization in Repeated Games with Conflicting Interests," *Econometrica: Journalof the Econometric Society*, pp. 325-351, 1993.

[14] D. Fudenberg, and D. Levine, "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica: Journal of the Econometric Society*, pp. 759-778, 1989.

[15] J. Ouyang, Y. Wang, X. Hu, and Y. Lin, *An Incentive Mechanism Using Game Theory for P2p Networks*, 2009, pp. 715-718.

[16] M. Feldman, K. Lai, I. Stoica, and J. Chuang, *Robust Incentive Techniques for Peer-to-Peer Networks*, 2004, pp. 102-111.

[17] A. Blanc, Y. Liu, and A. Vahdat, "Designing Incentives for Peer-to-Peer Routing," *in Proc. IEEE INFOCOM*, 2005, pp. 374-385.

فائزه بهرامیان مدرک کارشناسی خود را در رشته مهندسی نرم افزار از دانشگاه صنعتی امیرکبیر در سال ۱۳۸۷ و مدرک کارشناسی ارشد خود را در رشته فناوری اطلاعات گرایش امنیت اطلاعات از دانشگاه صنعتی امیرکبیر در سال ۱۳۸۹ دریافت نموده است. موضوعات پژوهشی مورد علاقه وی امنیت اطلاعات، عقلانیت محدود و نظریه بازی‌ها، ریاضیات و منطق می‌باشد.

آدرس پست‌الکترونیکی ایشان عبارت است از:

bahramiyan@aut.ac.ir

حمیدرضا شهریاری مدرک کارشناسی خود را از دانشگاه صنعتی شریف در سال ۱۳۷۶ و مدرک کارشناسی ارشد خود را از دانشگاه صنعتی امیرکبیر در سال ۱۳۷۹ در رشته مهندسی نرم افزار دریافت کرده است. وی تحصیلات خود را در دکتری مهندسی نرم‌افزار دانشگاه صنعتی شریف در سال ۱۳۸۶به پایان برده است و در حال حاضر استادیار دانشگاه صنعتی امیرکبیر، دانشکده مهندسی کامپیوتر و فناوری اطلاعات می‌باشد. موضوعات پژوهشی مورد علاقه وی امنیت اطلاعات، امنیت پایگاه داده‌ها و تحلیل آسیب‌پذیری‌های سیستم‌های کامپیوتری می‌باشد.

آدرس پست‌الکترونیکی ایشان عبارت است از:

shahriari@aut.ac.ir

اطلاعات بررسی مقاله:

تاریخ ارسال: ۹۰/۲/۸

تاریخ اصلاح: ۹۰/۱۲/۱۲

تاریخ قبول شدن: ۹۰/۱۲/۱۸

نویسنده مرتبط: دکتر حمیدرضا شهریاری، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر، تهران، ایران.

[1] Ipoque
[2] BitTorrent
[3] Random Matching Game
[4] Prisoner`s Dilemma
[5] Rosenthal
[6] Landau
[7] Stage Game
[8] Kandori
[9] Folk Theorem
[10] One Stage Deviation
[11] Tit-for-Tat
[12] Evolutionary Stable Strategy
[13] Tit-for-Tat