

روشی ترکیبی در شبکه‌های اجتماعی جهت تشخیص اجتماع با استفاده از الگوریتم‌های خوشه‌بندی فراابتکاری

سجاد حاجی

دانشگاه آزاد واحد خرم آباد.

نام نویسنده مسئول:

سجاد حاجی

چکیده

مسئله تشخیص خوشه در شبکه‌های اجتماعی یکی از مسائل بسیار مهم در تحقیقات این روزهای شبکه‌های اجتماعی است. شناسایی جامعه در شبکه‌های اجتماعی به درک ساختار درست شبکه، توزیع عامل‌ها، فعالیت‌های انجام شده و کمک به مسئولان جهت انجام تصمیم‌گیری می‌انجامد و ساختار جوامع، روش منتشر شدن اطلاعات و رفتار اشخاص را نشان می‌دهد. بنابراین شناسایی و تشخیص جامعه یکی از اهداف این پژوهش بوده است که به کمک علم داده‌کاوی و الگوریتم‌های فراابتکاری صورت می‌پذیرد. یکی دیگر از فواید تحلیل شبکه‌های اجتماعی این است که ساختار و روابط میان اشخاص و اعضا را توضیح می‌دهد و می‌تواند برای پیش‌بینی رفتارهای اجتماعی استفاده شود، در این پژوهش این فرایند با استفاده از ترکیب دو الگوریتم فراابتکاری ازدحام ذرات و تبرید شبیه‌سازی شده انجام شده است. برای ارزیابی الگوریتم پیشنهادی از یک شبکه‌های اجتماعی باشگاه کاراته و یک شبکه اجتماعی بازی‌های فوتبال بهره گرفته شده است که براساس نتایج شبیه‌سازی مشخص گردید، دقت و کارایی الگوریتم پیشنهادی در مقایسه با سایر الگوریتم‌های ارائه شده در این زمینه بهبود ویژه‌ای داشته است و عمل تشکیل خوشه و بهینه‌سازی را نسبت به الگوریتم‌های ABC و PESD بهینه‌تر انجام می‌دهد.

واژگان کلیدی: تشخیص اجتماعات، خوشه‌بندی، روشی ترکیبی، الگوریتم، بهینه‌سازی، فراابتکاری.

مقدمه

دنیای روبه‌رشد شبکه‌های اجتماعی، شبکه اینترنت را به یک فضای تعاملی و رقابتی تبدیل کرده است، سایت‌های اجتماعی با رشد کاربران خود در صدد هستند تا نرم‌افزارها و مولفه‌های خود را مطابق با سلیقه کاربران به هنگام‌سازی نمایند و شرکت‌های بازرگانی نیز می‌توانند به بزرگترین اجتماعات انسانی، خدمات خود را ارائه دهند. باید گفت در این فضای بزرگ و پیچیده شناسایی افراد و تمایلات آنها بدون استفاده از علوم کامپیوتری ناممکن است. در شبکه‌های اجتماعی جوامع‌ای به صورت پنهان وجود دارند که با استفاده از تکنیک‌های استخراج جوامع شناسایی می‌شوند و پس از استخراج جوامع می‌توان هر گروه را متناسب با ویژگی‌هایش به صورت جزئی‌تر تحلیل نمود. تشخیص ساختارهای جامعه در شبکه‌های اجتماعی وظیفه مهمی در تحلیل شبکه‌های اجتماعی است و می‌تواند برای پیش‌بینی رفتارهای اجتماعی استفاده شود [۱].

رابطه میان زیر جوامع در هر جامعه می‌تواند بیشتر به عنوان ساختارهای جامعه سلسله مراتبی شناخته شود که در آن هر ابرگه در هر سطح سلسله مراتب یک ساختار تودرتو از جوامع یا گره‌ها را نشان می‌دهد. بیشتر مطالعات قبلی که به تشخیص ساختارهای جامعه سلسله مراتبی توجه کرده‌اند، بر روی روش‌های فراابتکاری تمرکز داشته‌اند که از نظر محاسباتی کارآمد هستند، اما تفکیک بهینه‌ای از جامعه را تضمین نمی‌کند. در روش پیشنهادی یک رویکرد برنامه‌نویسی مبتنی بر الگوریتم‌های بهینه‌سازی ازدحام ذرات و تبرید شبیه‌سازی شده برای تشخیص ساختارهای جامعه سلسله مراتبی در شبکه‌های اجتماعی اعمال شده است. الگوریتم پیشنهادی می‌تواند به عنوان یک راه حل نرم‌افزاری برای تشخیص ساختارهای سلسله مراتبی جامعه استفاده شود [۲].

از میان روش‌های خوشه‌بندی، خوشه‌بندی سلسله مراتبی اطلاعات مفیدی در مورد ارتباطات اعضای یک شبکه اجتماعی ارائه می‌دهد. اهمیت این موضوع زمانی آشکار می‌شود که بخواهیم در مورد قسمت اعظمی از جامعه مطالعه‌ای انجام دهیم. چرا که با خوشه‌بندی سلسله مراتبی می‌توان علاوه بر گروه‌بندی اعضا، جزئیات گروه‌ها را، و زیرگروه‌ها را نیز تعیین کرد و تجزیه و تحلیل جداگانه بر روی هر کدام انجام داد [۳]. از این رو در مطالعه شبکه‌های اجتماعی، خوشه‌بندی سلسله مراتبی عملکرد بهتری نسبت به خوشه‌بندی معمولی خواهد داشت. برای تشخیص جوامع سلسله مراتبی، الگوریتم‌های متعددی پیشنهاد شده است. از بین آنها الگوریتم‌های فراابتکاری که امروزه مانند شاه کلید برای اکثر مسائل قابل استفاده است، در این پژوهش انتخاب شده است. این الگوریتم‌ها می‌تواند به خوبی یک مساله بهینه‌سازی را در زمان کوتاهی حل کنند و نسبت به زمان اندک یک جواب نسبتاً بهینه به ارائه دهند.

در تشکیل جوامع سلسله مراتبی لازم است که در هر یک از شبکه اجتماعی، اعضای درون یک جامعه با هم ارتباط نزدیکی داشته و در عین حال ارتباط ضعیف‌تری با جوامع دیگر نسبت به جامعه خود داشته باشند. برای تشخیص این مسائل الگوریتم‌های فراابتکاری به خوبی عمل کرده و یک از بهینه‌ترین جوامع را تشکیل خواهند داد [۲، ۳].

استفاده از چند روش به صورت ترکیبی در شبکه‌های اجتماعی موضوعی است که اخیراً تحقیقات زیادی به آن پرداخته شده است. علاوه بر این استفاده از روش‌های ترکیبی استخراج دانش می‌تواند جوامع یکنواخت‌تری را حاصل کند، گاهی مجبور هستیم به دلیل این که منابع موجود از انواع مختلفی هستند از روش متناسب با هر نوع داده استفاده کنیم و سپس با راهکار مناسبی آنها را ترکیب نماییم. در این پژوهش با استفاده از ترکیب دو الگوریتم فراابتکاری بهینه‌سازی ازدحام ذرات و تبرید شبیه‌سازی شده روشی جدید برای تشخیص اجتماع در شبکه‌های اجتماعی ارائه شده است که دارای جنبه‌های مختلف نوآوری است.

پیشینه تحقیق

به هر ساختار اجتماعی از افراد که بر اساس یک رابطه اجتماعی ایجاد می‌شود، یک شبکه اجتماعی می‌گویند؛ بنابراین هر شبکه اجتماعی شامل مجموعه‌ای از انسان‌ها و روابط اجتماعی میان آنها است. لذا هر شبکه اجتماعی از دو عنصر موجودیت‌های شرکت کننده در ارتباط و ارتباط میان این موجودیت‌ها تشکیل شده است. شبکه‌های اجتماعی به دو نوع آنلاین و آفلاین تقسیم می‌شود. از شبکه‌های آفلاین می‌توان به شبکه دوستان، شبکه همکاران و شبکه همکلاسی‌ها و... اشاره کرد و از شبکه‌های آنلاین می‌توان به شبکه‌های اجتماعی نظیر فیس‌بوک، توئیتر و گوگل پلاس و... اشاره نمود [۴].

شبکه اجتماعی، ساختاری اجتماعی است که از گره‌هایی (عموماً فردی یا سازمانی) تشکیل شده است که توسط یک یا چند نوع خاص از وابستگی‌ها مانند تبادلات مالی، دوست‌ها، خویشاوندی، لینک‌های وب و سرایت بیماری‌ها به یکدیگر متصل‌اند. تحلیل شبکه‌های اجتماعی روابط اجتماعی را با اصطلاحات رأس و یال می‌نگرد. رأس‌ها بازیگران فردی درون شبکه‌ها هستند و یال‌ها روابط میان این بازیگران هستند. انواع زیادی از یال‌ها می‌تواند میان رأس‌ها وجود داشته باشد. نتایج تحقیقات مختلف بیانگر آن است که می‌توان از ظرفیت شبکه‌های اجتماعی در بسیاری از سطوح فردی و اجتماعی به منظور شناسایی مسائل و تعیین راه‌حل آنها، برقراری روابط اجتماعی، اداره امور تشکیلاتی، سیاست‌گذاری و رهنمون‌سازی افراد در مسیر دستیابی به اهداف استفاده نمود. در ساده‌ترین شکل ممکن یک شبکه اجتماعی

نگاشتی از تمام یال‌های مربوط میان رأس‌ها است. شبکه اجتماعی هم چنین می‌تواند برای تشخیص موقعیت اجتماعی هر یک از بازیگران مورد استفاده قرار گیرد. این مفاهیم غالباً در یک نمودار شبکه اجتماعی نشان داده می‌شوند که در آن نقطه‌ها، رأس‌ها را نشان می‌دهند و خط‌ها نشانگر یال‌ها می‌باشند و به‌طور عمده دو شیوه ماتریس و گراف برای نمایش شبکه‌های اجتماعی مورد استفاده قرار می‌گیرد [۴].

نمایش ماتریس گونه

برای نمایش شبکه‌های اجتماعی می‌توانیم از آرایش مربع‌ای (ماتریس) استفاده نماییم. در ماتریس هر سطر متعلق به یک رأس و به همان ترتیب هر ستون نیز به یک رأس متعلق است. ماتریس‌های نشان‌دهنده شبکه‌های اجتماعی لزوماً متقارن نیستند، یعنی روابطی که توسط این ماتریس‌ها بیان می‌شود، می‌تواند یک‌طرفه یا دوطرفه باشد. در صورتی که ماتریس ارائه شده توسط یک شبکه اجتماعی، متقارن باشد، آن شبکه اجتماعی دوطرفه در نظر گرفته می‌شود. می‌توان این موضوع را به صورت واضح در شبکه‌های اجتماعی مشاهده کرد. برای مثال شبکه اجتماعی فیس‌بوک شبکه اجتماعی است که دو نوع ارتباط را امکان‌پذیر ساخته است. اما به خاطر نسبت بسیار بالای روابط دو طرفه (دوست بودن) به رابطه یک طرفه (دنبال کردن) این شبکه اجتماعی، بیشتر به عنوان شبکه اجتماعی دو طرفه در نظر گرفته می‌شود. اما در شبکه‌ها نظیر توئیتر یا گوگل پلاس نیاز به ارتباط رابطه دو طرفه نیست و دنبال کردن به صورت یک طرفه نیز در نظر گرفته می‌شود. هرچند که با توجه به اینکه امروزه اغلب ارتباطات در شبکه‌های اجتماعی دو طرفه هستند در تحلیل این شبکه‌های اجتماعی با توجه به نسبت کم ارتباطات یک طرفه به ارتباطات دو طرفه می‌توان همه ارتباطات را دو طرفه در نظر گرفت. در ماتریس مربوط به یک شبکه اجتماعی مؤلفه $ij\alpha$ برابر یک است، اگر از راس i به راس j یک یال (ارتباط) وجود داشته باشد. در غیر اینصورت مؤلفه $ij\alpha$ برابر صفر قرار داده می‌شود [۵، ۶].

نمایش گراف گونه

گراف‌ها به واسطه نمایش دادن بصری شبکه‌های اجتماعی به عنوان دیگر ابزار ریاضی برای نشان دادن و تجزیه و تحلیل شبکه اجتماعی به کار می‌روند. در این گراف‌ها، بازیگران به‌عنوان رأس‌های گراف در نظر گرفته می‌شوند و ارتباط میان آن‌ها، توسط یال‌های گراف به نمایش گذاشته می‌شود. گراف‌ها نیز همچون ماتریس‌ها برای نمایش هر دو نوع شبکه‌های اجتماعی (یک طرفه و دو طرفه) به کار برده می‌شوند. از گراف‌های جهت‌دار برای نمایش شبکه‌های اجتماعی یک طرفه و از گراف‌های بدون جهت برای نمایش شبکه‌های اجتماعی دو طرفه استفاده می‌گردد [۷]. مؤلفه‌های گراف و ماتریس به راحتی قابل تبدیل به یکدیگر هستند.

آنالیز شبکه‌های اجتماعی

منظور از آنالیز شبکه‌های اجتماعی، مطالعه و تجزیه و تحلیل آن‌هاست. برای تحلیل شبکه‌های اجتماعی روش‌های متفاوتی وجود دارد و این روش‌ها در ابتدا ریشه در جامعه‌شناسی و ریاضیات (نظریه گراف) داشتند، اما امروزه علاوه بر این دو رشته در سایر علوم نیز مورد استفاده قرار می‌گیرند. با استفاده از تحلیل شبکه‌های اجتماعی می‌توانیم به اطلاعات مفیدی از قبیل میانگین دوستان یک فرد در یک شبکه دوستی یا میانگین فاصله هر دو نفر در یک شبکه همکاران دست یابیم. مفاهیم بسیاری وجود دارد که در تحلیل شبکه‌های اجتماعی به کار می‌روند. از جمله می‌توان به شعاع، قطر، توزیع درجات، توزیع درجات گره‌های متصل، ضریب خوشه‌بندی، مرکزیت ارتباط و مؤلفه همبندی اشاره کرد، که در ادامه این بخش به تعریف برخی از این مفاهیم خواهیم پرداخت.

شعاع و قطر: برای تعریف شعاع و قطر، باید نخست مفهوم گریز از مرکز یک رأس را تعریف کرد. گریز از مرکز بیشترین فاصله‌ای است که یک رأس از رأس‌های دیگر می‌تواند داشته باشد. فاصله دو رأس از هم برابر با طول کوتاه‌ترین مسیر بین آن‌هاست. شعاع یک گراف کمترین مقدار گریز از مرکز در کل گراف است. قطر یک گراف برابر با بیشترین مقدار گریز از مرکز در کل گراف است [۸].

توزیع درجه: توزیع درجه در یک گراف، تابع $P(K)$ می‌باشد که نشان دهنده نسبت تعداد رأس‌ها با درجه K ، به تعداد کل رأس‌های گراف است. توزیع درجه یکی از ویژگی‌های مهم یک گراف است. گراف‌های مربوط به شبکه‌های اجتماعی، ویژگی‌های مشابه و خاصی را از خود بروز می‌دهند. از جمله ویژگی‌های متناظر با شبکه‌های اجتماعی، توزیع درجه توانی است.

توزیع درجه راس‌های متصل: تابع توزیع درجه راس‌های متصل با $J(K, M)$ نشان داده می‌شود. در گراف‌های جهت‌دار $J(K, M)$ نشان دهنده نسبت یال‌های جهت‌دار از یک راس با درجه K به راس دیگر با درجه M به کل یال‌های گراف است. تابع توزیع درجه راس‌های گراف که به اختصار با JDD نیز نشان داده می‌شود، معیاری است، برای نشان دادن اینکه تا چه اندازه راس‌ها به راس‌های مشابه خود وصل شده‌اند. این تابع نشان دهنده نسبت تعداد یال‌های بین راس‌هایی با درجه K و M به تعداد کل یال‌های گراف است.

ضریب خوشه‌بندی: معیاری است که درجه گرهای که در یک گراف تمایل به ایجاد یک خوشه با هم دارند را اندازه می‌گیرد. شواهد حاکی از آن است که در اکثر شبکه‌های دنیای واقعی و به خصوص در شبکه‌های اجتماعی، گره‌ها تمایل به ایجاد گروه‌های بافتی که توسط ارتباط نسبتاً پرتراکم مشخص می‌شوند دارد و این احتمال بیش از احتمال میانگین احتمال اتصال‌های تصادفی تشکیل شده بین دو گره است. ضریب خوشه‌بندی مقیاسی است برای بررسی اینکه به چه نسبتی همسایه‌های یک راس با هم در ارتباط هستند. برای یک گراف، ضریب خوشه‌بندی، احتمال همسایه بودن دو راسی را نشان می‌دهد که یک همسایه مشترک دارند. به صورت رسمی می‌توان ضریب خوشه‌بندی را به این گونه تعریف کرد، ضریب خوشه‌بندی راس i که با $C(i)$ نشان داده می‌شود، نسبت تعداد یال‌هایی که بین راس‌های همسایه i وجود دارد به تعداد کل یال‌هایی که می‌تواند بین کل همسایه‌های i موجود باشد، نشان داده می‌شود [۹].

روش پیشنهادی

شناسایی جوامع یا خوشه‌های موجود در یک شبکه به درک ساختار کلی یک شبکه، توزیع موجود بر روی عامل‌ها و فعالیت‌هایی که در شبکه صورت می‌پذیرند، می‌انجامد. همچنین، ساختار جوامع در یک شبکه، روشی که اطلاعات در آن پخش می‌شوند را تحت تاثیر قرار می‌دهند و حتی بر روی رفتار اشخاص نیز تاثیر دارند. تجزیه و تحلیل شبکه‌های اجتماعی منجر به استخراج اطلاعات مفید می‌گردد. در این روش، با بهره‌گیری از مزایای الگوریتم بهینه‌سازی ازدحام ذرات و تبرید شبیه‌سازی شده به ارائه یک الگوریتم جهت بررسی ساختار شبکه‌های اجتماعی پرداخته می‌شود.

بهینه‌سازی ازدحام ذرات از دسته الگوریتم‌های بهینه‌سازی است که بر مبنای تولید تصادفی جمعیت اولیه عمل می‌کنند. در این الگوریتم با الگوبرداری و شبیه‌سازی رفتار پرواز دسته‌جمعی (گروهی) پرندگان یا حرکت دسته‌جمعی (گروهی) ماهی‌ها بنا نهاده شده است. بهینه‌سازی ازدحام ذرات یک الگوریتم بهینه‌سازی مبتنی بر جمعیت است که می‌تواند به راحتی اجرا شود و به راحتی حل شود تا توابع مختلفی از مشکلات بهینه‌سازی حل شوند. به‌عنوان یک الگوریتم، قدرت اصلی بهینه‌سازی ازدحام ذرات، همگرایی سریع آن است [۱۰، ۱۱].

در این الگوریتم بردار سرعت i امین ذره نیز با بردار V_i به شکل زیر تعریف می‌گردد.

$$V_i = (v_{i1}, v_{i2}, v_{i3}, \dots, v_{in}), I = 1, 2, \dots, n \quad (1)$$

بهترین موقعیتی که ذره i ام پیدا کرده است را با P_{id_i} تعریف می‌کنیم. که اندازه ازدحام ذرات است.

$$P_{id_i} = (pid_{i1}, pid_{i2}, \dots, pid_{in}) \quad (2)$$

بهترین موقعیتی که بهترین ذره در بین کل ذرات پیدا کرده است را با P_{gd} به صورت زیر تعریف می‌کنیم.

$$P_{gd} = (p_{g1}, p_{g2}, p_{g3}, \dots, p_{gd}) \quad (3)$$

سپس، این ذرات بر طبق معادلات زیر به کار گرفته می‌شوند.

$$V_i(t) = W * V_i(t-1) + c_1 * rand_1 * (P_{id_i} - X_i(t-1)) + c_2 * rand_2 * (P_{gd} - X_i(t-1)) \quad (4)$$

$$X_i = X_i(t-1) + V_i(t) \quad (5)$$

که W ضریب وزنی اینرسی (حرکت در مسیر خودی) نشان دهنده میزان تأثیر برداری سرعت تکرار قبل $(V_i(t))$ بر روی بردار سرعت در تکرار فعلی $(V_i(t+1))$ است. c_1 ضریب ثابت آموزش (حرکت در مسیر بهترین مقدار ذره مورد بررسی). c_2 ضریب ثابت آموزش (حرکت در مسیر بهترین ذره یافت شده در بین کل جمعیت). $rand_1, rand_2$ دو عدد تصادفی با توزیع یکنواخت در بازه 0 تا 1 است. $V_i(t-1)$ بردار سرعت در تکرار $(t-1)$ ام می‌باشد. $X_i(t-1)$ بردار موقعیت در تکرار $(t-1)$ ام است. برای جلوگیری از افزایش بیش از حد سرعت حرکت یک ذره در حرکت از یک محل به محل دیگر (واگرا شدن بردار سرعت)، تغییرات سرعت را به رنج V_{min} تا V_{max} محدود می‌کنیم و حد بالا و پایین سرعت با توجه به نوع مسئله تعیین می‌گردد.

مراحل اجرای الگوریتم بهینه‌سازی ازدحام ذرات به صورت زیر است. گام اول تولید تصادفی جمعیت اولیه ذرات است. گام دوم انتخاب تعداد ذرات اولیه است. گام سوم ارزیابی تابع هدف (محاسبه هزینه یا برازندگی) ذرات است. گام چهارم ثبت بهترین موقعیت، برای هر ذره و بهترین موقعیت در بین کل ذره‌ها است. گام پنجم به روز رسانی بردار سرعت در تمامی ذرات است و در نهایت در گام ششم تست همگرایی انجام می‌شود [۱۰، ۱۱].

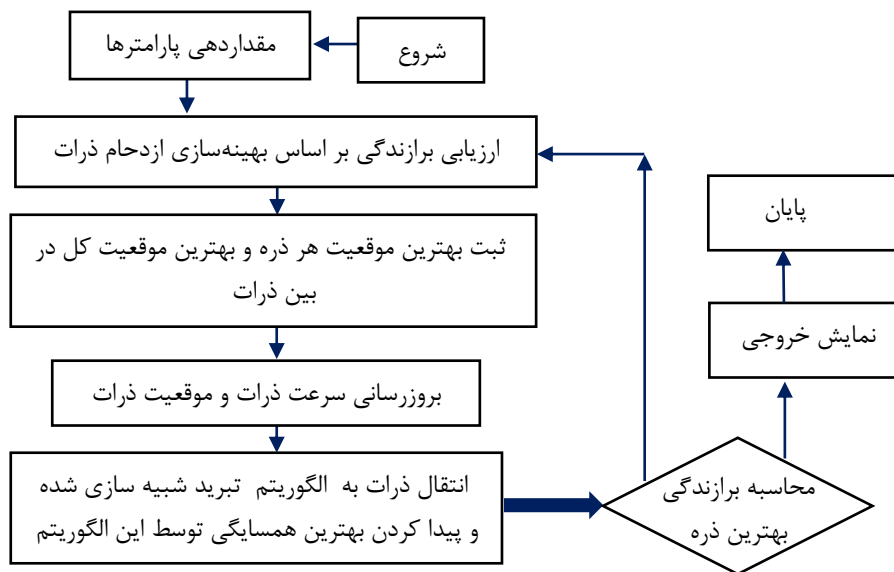
تبرید شبیه سازی شده به معنای شبیه سازی گداخت یا شبیه سازی حرارتی می باشد که برای آن از عبارات شبیه سازی بازپخت فلزات، شبیه سازی آب دادن فولاد و الگوریتم تبرید نیز استفاده شده است. برخی مسائل بهینه سازی صنعتی در ابعاد واقعی غالباً پیچیده و بزرگ می باشند. این ایده ابتدا توسط شخصی که در صنعت نشر فعالیت داشت به نام متروپلیس در سال ۱۹۵۳ بیان شد. الگوریتم متروپلیس شبیه سازی است از فرآیند سرد شدن مواد به وسیله کاهش آهسته دمای سیستم (ماده) تا زمانی که به یک حالت ثابت منجمد تبدیل شود. این روش با ایجاد و ارزیابی جواب های متوالی به صورت گام به گام به سمت جواب بهینه حرکت می کند. برای حرکت، یک همسایگی جدید به صورت تصادفی ایجاد و ارزیابی می شود. در این روش به بررسی نقاط نزدیک در فضای جستجو پرداخته می شود. در صورتی که نقطه جدید، نقطه بهتری باشد (تابع هزینه را کاهش دهد) به عنوان یک نقطه جدید در فضای جستجو انتخاب می شود و اگر بدتر باشد (تابع هزینه را افزایش دهد) و بر اساس یک تابع احتمالی انتخاب می شود. به عبارتی دیگر برای کمینه سازی تابع هزینه، جستجو همیشه در جهت کمتر شدن مقدار تابع هزینه صورت می گیرد، اما این امکان وجود دارد که گاه حرکت در جهت افزایش تابع هزینه باشد. معمولاً برای پذیرفتن نقطه بعدی از معیاری به نام معیار متروپلیس که به صورت زیر است استفاده می شود [۱۳، ۱۲].

$$p(\text{accept}) = \begin{cases} 1, \Delta f \leq 0 \\ e^{-\frac{\Delta f}{c}}, \Delta f > 0 \end{cases} \quad (6)$$

در رابطه بالا P احتمال پذیرش نقطه بعدی است و C یک پارامتر کنترلی می باشد. متغیر Δf برای تغییر هزینه جهت کنترل بهینگی در نظر گرفته شده است. پارامتر کنترل در تبرید شبیه سازی شده، همان نقش دما را در پدیده فیزیکی ایفا می کند. ابتدا ذره (که نمایش دهنده نقطه فعلی در فضای جستجو است) با مقدار انرژی بسیار زیادی (که نشان دهنده مقدار بالای پارامتر کنترلی C است) نشان داده شده است. این انرژی زیاد، به ذره اجازه فرار از یک بهینه محلی را می دهد. همچنان که جستجو ادامه می یابد، انرژی ذره کاهش می یابد (مقدار C کم می شود) و در نهایت جستجو به بهینه کلی میل خواهد نمود. البته باید توجه داشت که در دمای پایین امکان فرار الگوریتم از بهینه محلی کاهش می یابد، به همین دلیل هر چه انرژی آغازین بالاتر، امکان رسیدن به بهینه کلی هم بیشتر است. پ. مراحل اجرای الگوریتم تبرید شبیه سازی شده به صورت زیر است. گام اول تولید تصادفی جمعیت که است. گام دوم مولد هزینه است که این مولد وظیفه تولید حالات بعدی را به عهده دارد. گام سوم ارزیابی تابع هدف است، که با توجه به محاسبه هزینه نقطه فعلی نقطه بعدی می توان مشخص شود. در گام چهارم برنامه سرد کردن اجرا می شود، که مشخص می کند دما چند وقت به چند وقت و به چه میزان کاهش می یابد و دماهای شروع و پایان چقدر باشند [۱۳، ۱۲].

الگوریتم

- بر اساس آنچه که بیان شد، بهینه سازی ازدحام ذرات با استفاده از عملگر ترکیب (مبتنی بر هوش جمعی) و الگوریتم تبرید شبیه سازی شده بر اساس موقعیت فعلی فرد و بر اساس همسایگان، ذرات جدید را تولید می کنند. بر همین اساس بهینه سازی ازدحام ذرات بیشتر خاصیت کشف کنندگی دارد و در مقابل تبرید شبیه سازی شده بیشتر خاصیت استخراج کنندگی دارد. لذا در این پروژه سعی شده است با ترکیب این دو الگوریتم، از خاصیت های گفته شده به صورت همزمان بهره برد. مراحل الگوریتم را می توان به شکل زیر توصیف کرد.
- ۱- مقدار دهی پارامترها (تعداد ذرات، جمعیت و تعداد تکرار).
 - ۲- بر اساس الگوریتم بهینه سازی ازدحام ذرات، بهترین موقعیت هر فرد و بهترین ذره در جمعیت مشخص می شود و مراحل زیر را تا رسیدن به شرط توقف (حداکثر تعداد تکرار) ادامه می یابد.
 - ۳- بر اساس الگوریتم بهینه سازی ازدحام ذرات سرعت و موقعیت ذرات محاسبه می شود.
 - ۴- بر اساس الگوریتم بهینه سازی ازدحام ذرات به روزرسانی سرعت و موقعیت ذرات بر اساس روابط موجود انجام می پذیرد.
 - ۵- الگوریتم تبرید شبیه سازی در همسایگی فرد بهترین موقعیت را جستجو می نماید.
 - ۶- برازندگی هر ذره محاسبه و اگر برازندگی فعلی بهتر از برازندگی قبلی بود، بردار موقعیت قبلی توسط بردار موقعیت فعلی جایگزین می شود.
 - ۷- بهترین برازندگی از بین تمام ذرات انتخاب شده و آن با بهترین ذره قبلی مقایسه می شود، اگر بهتر باشد با بهترین ذره جایگزین می شود و در آخر نتایج نشان داده می شود.



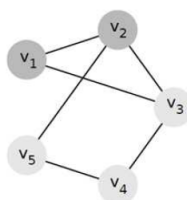
شکل: فلوچارت روش پیشنهادی

تابع هدف

انتخاب یک تابع هدف مناسب برای یک الگوریتم تشخیص جامعه بسیار مهم است. یک تابع هدف خوب باعث می‌شود ذرات با سرعت بیشتری راه‌حل را پیدا کنند. در الگوریتم بهینه‌سازی ازدحام ذرات هر ذره با یک بخش (خوشه) از جامعه مطابقت دارد، بنابراین کیفیت، جامعه تابع تناسب ذرات است. روش‌های زیادی برای ارزیابی کیفیت خوشه‌ها (بخش‌ها) ارائه شده است. اما رایج‌ترین روش ماژولاریتی (پیمانگی) است. روش ماژولاریتی ابتدا به عنوان معیاری جهت تعیین مرحله توقف الگوریتم گیروان و نیومن مورد توجه قرار گرفت، ولی به سرعت به یکی از با اهمیت‌ترین الگوریتم‌های تشخیص انجمن تبدیل شد. این معیار فرمولی برای محاسبه کیفیت تقسیم نودها به انجمن‌های مختلف ارائه می‌کند که به دلیل ساده و موثر بودن آن به پرکاربردترین معیار کمی جهت اندازه‌گیری کیفیت الگوریتم‌های تشخیص انجمن شد.

از آنجایی که لزوماً الگوریتم‌های مختلف تعداد خوشه‌های یکسانی را تولید نمی‌کنند بسیاری از معیارهای موجود جواب‌گوی مقایسه روش‌های مختلف خوشه‌بندی نیستند و امکان استفاده از آنها برای ارزیابی این روش‌ها وجود ندارد. از طرفی این خصوصیت پیمانگی، این امکان را می‌دهد که بتوان از آن برای تعیین تعداد خوشه‌ها نیز استفاده نمود. بنابراین وقتی که از روش‌های سلسله مراتبی بالا به پایین یا پایین به بالا استفاده می‌کنیم، می‌توان روند تغییرات پیمانگی هنگام تقسیم یا ترکیب خوشه‌ها را تحت نظر گرفت و خوشه‌بندی که این کمیت را به حداکثر ممکن می‌رساند به عنوان بهترین خوشه‌بندی گراف در نظر گرفت.

	C ₁		C ₂		
	v ₁	v ₂	v ₃	v ₄	v ₅
C ₁	v ₁	0 1	1 0 0		
	v ₂	1 0	1 0 1		
C ₂			v ₃	1 1 1	0 1 0
			v ₄	0 0	1 0 1
			v ₅	0 1	0 1 0



شکل: مثالی از ماژولاریتی Q

گراف ارتباط c یک ماتریس c^*c است که به عنوان e تعریف می‌گردد و هر e_{ij} کسری از لینک‌هایی که از i به j می‌رود را نشان می‌دهد و هر $a_i = \sum_j e_{ij}$ نشان دهنده لینک‌ها متصل به i می‌باشد. ماژولاریتی Q^2 به صورت زیر تعریف می‌گردد.

$$Q = \sum_i (e_i - e_i^2) \quad (7)$$

اگر تعداد یال‌های درون خوشه‌ای بهتر از حالت گراف تصادفی نباشد، پیمانگی برابر صفر است. حداکثر مقدار پیمانگی زمانی به دست می‌آید که تمام رئوس هر خوشه به هم متصل باشند و هیچ گونه یالی خوشه‌ها را به همدیگر متصل نکند. تجربه نشان می‌دهد که در گراف‌های متعلق به شبکه‌های اجتماعی این مقدار بین ۰/۳ تا ۰/۷ است. مقادیر بالای Q نشان دهنده نشان دهنده این است که راس‌های بیشتری بهم متصل هستند. روش شناسایی جامعه مبتنی بر بهینه‌سازی ازدحام ذرات در ابتدا دو پارتیشن دریافت می‌کند. پس از دو بخش،

هر پارتیشن به خودی خود یک جامعه در نظر گرفته می‌شود. این عمل تکرار می‌شود تا زمانی که ماژولاریتی نمی‌تواند برای جوامع فعلی موجود یافت شود.

اگر در آخرین تکرارها، بهترین راه حل سراسری وجود نداشته باشد، به این معنی است که الگوریتم در یک نقطه بهینه محلی به دام افتاده است. برای فرار از بهینه محلی، الگوریتم تبرید شبیه‌سازی شده را به بهترین راه حل سراسری اعمال می‌کنیم (عملکرد تبرید شبیه‌سازی شده به تعریف چندین پارامتر کنترل بستگی دارد). بعد از محاسبه بهترین همسایگی برازندگی هر ذره محاسبه و اگر برازندگی فعلی بهتر از برازندگی قبلی باشد، بردار موقعیت قبلی توسط بردار موقعیت فعلی جایگزین شود. در نهایت نتایج بعد از رسیدن به حداکثر تکرار در خروجی نمایش داده می‌شود.

شبیه‌سازی و نتایج

برای ارزیابی روش پیشنهادی، نیاز است که کارکرد آن را با شبکه‌های اجتماعی واقعی بررسی نماییم. بنابراین در این پژوهش از یک شبکه اجتماعی واقعی بهره گرفته شده است. شبکه اجتماعی ارائه شده شامل شبکه دو باشگاه کاراته و بازی‌های فوتبال دانشگاهی است.

جدول: توصیف مجموعه داده‌های مورد استفاده

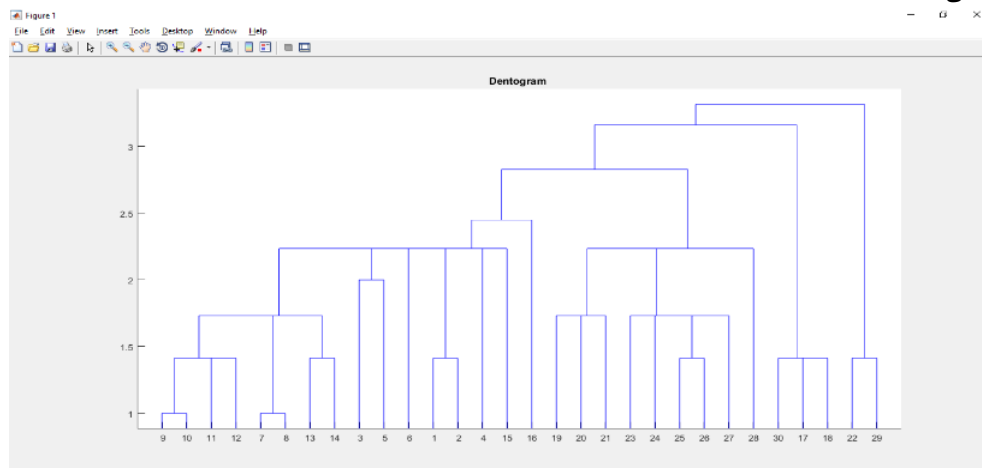
مجموعه داده	تعداد گره‌ها	تعداد یال‌ها	جمعیت
باشگاه کاراته	۳۴	۷۸	۳۳
انجمن فوتبال	۱۱۵	۶۱۶	۵۰

برای ارزیابی و بررسی شیوه‌ی کارکرد روش پیشنهادی، این روش را روی دو شبکه اجتماعی سرشناس باشگاه کاراته و بازی‌های فوتبال که در بخش پیشین شرح داده شدند، آزموده‌ایم. همچنین مقادیر نخستین برای پارامترهای الگوریتم پیشنهادی در جدول زیر آمده است. همانگونه که گفته شد روش پیشنهادی را با ارزش‌های گوناگون برای این پارامترها آزمودیم و ارزش‌هایی که در جدول زیر آمده است، بهترین برآیند را به همراه داشته‌اند.

جدول: سنجش میان الگوریتم‌های گوناگون در خوشه‌بندی شبکه اجتماعی باشگاه کاراته

میزان تابع برازندگی ماژولاریتی	شبکه اجتماعی باشگاه کاراته
۰,۴۷,۵	الگوریتم PESD
۰,۶۰	الگوریتم ABC
۰,۶۲,۵	الگوریتم پیشنهادی

همانگونه که در جدول بالا نشان داده شده است، الگوریتم پیشنهادی در سنجش با دیگر الگوریتم‌ها توانسته است به شیوه مناسب-تری باشگاه کاراته را خوشه‌بندی نماید. در شکل زیر نیز نمودار دندروگرام پایانی از خوشه‌بندی که توسط الگوریتم پیشنهادی به دست آمده است را نمایش می‌دهد.



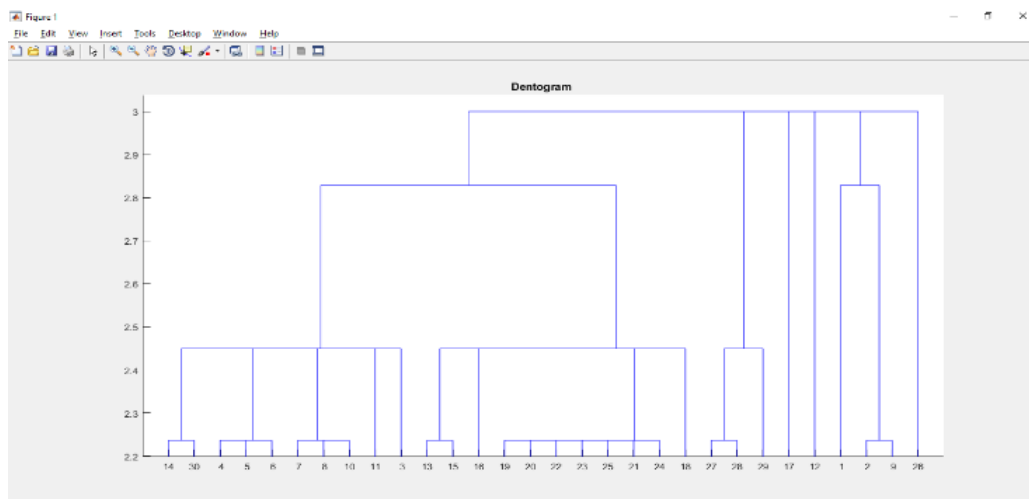
شکل: نمودار دندروگرام از خوشه‌بندی شبکه اجتماعی باشگاه کاراته توسط الگوریتم پیشنهادی

شبکه اجتماعی دیگری که برای ارزیابی راهکار پیشنهادی به کار گرفته شده است، شبکه اجتماعی بازی‌های فوتبال است. از این رو آزمایش‌ها بر روی این شبکه نیز انجام شده است. در این شبکه که دارای ۱۱۵ گره است، هر یال ارتباط بین دو گره، که به این یال پیوند دارند، را برقرار می‌سازد.

جدول: سنجش میان الگوریتم‌های گوناگون در خوشه‌بندی شبکه اجتماعی بازی‌های فوتبال

میزان تابع برازندگی مازولاریتی	شبکه اجتماعی باشگاه کاراته
۰,۵۴,۳	الگوریتم PESD
۰,۶۱,۵	الگوریتم ABC
۰,۶۳	الگوریتم پیشنهادی

همانگونه که در جدول بالا نشان داده شده است، الگوریتم پیشنهادی در سنجش با دیگر الگوریتم‌ها توانسته است، به شیوه مناسب-تری شبکه بازی‌های فوتبال دانشگاهی را خوشه بندی نماید. همچنین در شکل زیر نمودار دندروگرام پایانی از خوشه‌بندی این شبکه توسط الگوریتم پیشنهادی نمایش داده می‌شود.



شکل: بالا نمودار دندروگرام از خوشه‌بندی شبکه بازی‌های فوتبال توسط الگوریتم پیشنهادی را نشان می‌دهد

نتیجه‌گیری

با توسعه روز افزون اینترنت به عنوان شبکه‌ای از شبکه‌ها، ارتباطات بین افراد نسبت به گذشته بزرگتر و روز به روز این روابط پیچیده‌تر می‌شود، در این میان شبکه‌های اجتماعی نیز به عنوان پدیده‌ای جدید دارای مخاطبان خود بوده و روزبه‌روز پیچیده‌تر از گذشته خواند شد. کاربران در شبکه‌های اجتماعی بر یکدیگر تاثیر می‌گذارند و این تأثیر معمولاً از طریق مشاهده پست‌های یکدیگر رخ می‌دهد. به دست آوردن روابط در شبکه‌های اجتماعی و تحلیل آن‌ها بخش‌های مختلفی از زوایای پنهان ارتباطات را روشن خواهد نمود. از طرفی حجم اطلاعات و محتوای تولید شده در شبکه‌های اجتماعی و توانایی دسته‌بندی و مدیریت آن اطلاعات به شکلی که بتوان در راستای تبلیغات، معرفی محصولات جدید و یا کاربردهای فراوان دیگر به کار گرفت مطرح می‌گردد. یکی از راه‌های مدیریت و دسته‌بندی اطلاعات به منظور استفاده بهینه از اطلاعات تولید شده دسته‌بندی دوستان و ایجاد گروه‌های مختلف است. این عمل به عنوان مکانیزمی کارآمد می‌تواند به جداسازی جوامع مختلف برای تولید و به کارگیری اطلاعات هر گروه بسیار کارآمد باشد. از طرف مقابل این گروه‌ها و انجمن‌های ایجاد شده در شبکه‌های اجتماعی برای گروه دیگری از افراد که به دنبال استفاده از ظرفیت ایجاد شده به عنوان گروه‌های اجتماعی هستند اهمیت می‌یابد. دلیل رسیدن به این مفهوم شاید در ابتدا خیلی مهم جلوه نکند، اما در بررسی‌های بعدی و بررسی‌های بر پایه مطالعات ارتباطات بین جوامع کاربردهای بسیاری دارد. با مطالعه شبکه‌های اجتماعی و ارتباطات موجود در آنها می‌توان به درکی از علایق کاربران رسید، همچنین می‌توان عقاید و رفتارهای آنها را بررسی و تحلیل نمود.

منابع و مراجع

- [1] Ibáñez, P and Díaz, C and Perez, a. 2017. Social Networks and Childhood. New Agents of Socialization, 7th International Conference on Intercultural Education, Health and ICT for a Transcultural World, EDUHEM, Almeria, Spain.
- [2] Yin, C and Zhu, S and Chen, h. 2015. A Method for Community Detection of Complex Networks Based on Hierarchical Clustering, Hindawi Publishing Corporation International Journal of Distributed Sensor Networks, Article ID 849140.
- [3] Yuruk, N and Mete, M and Xu, X. 2009. AHSCAN: Agglomerative Hierarchical Structural Clustering Algorithm for Networks, International Conference on Advances in Social Network Analysis and Mining, INSPEC Accession Number 10903164, ISBN 978-0-7695-3689-7, Athens, Greece.
- [4] Batallas, D.A and Yassine, A. 2006. Information Leaders in Product Development Organizational Networks: Social Network Analysis of the Design Structure Matrix. IEEE Transactions on Engineering Management, Volume 53, Issue 4.
- [5] Cai, Q and Gong, M and Ma, l and Ruan, S. 2015. Greedy discrete particle swarm optimization for large-scale social network clustering. Nature-Inspired Algorithms for Large Scale Global Optimization, Volume 316.
- [6] Yao, X. 2007. A new simulated annealing algorithm. International Journal of Computer Mathematics, Volume 56, 1995 - Issue 3-4.
- [۷] فرقانی، م.م و مهجری، ر. ۱۳۹۷، رابطه بین میزان استفاده از شبکه‌های اجتماعی مجازی و تغییر درسبک زندگی جوانان، مطالعات رسانه‌های نوین، سال چهارم، شماره سیزدهم.
- [۸] باستانی، س و رئیسی، م. ۱۳۹۰، روش تحلیل شبکه: استفاده از رویکرد شبکه‌های کل در مطالعه اجتماعات متن باز، مجله مطالعات اجتماعی ایران، دوره پنجم، شماره دوم.
- [۹] ساعد، ن و صادق‌زاده، م. ۱۳۹۳، ارائه الگوریتمی جهت خوشه‌بندی گراف شبکه‌های اجتماعی مبتنی بر مرکزیت گره‌ها، اولین همایش ملی مهندسی برق و کامپیوتر در شمال کشور، موسسه آموزش عالی موج، بندر انزلی.
- [۱۰] یارحمادی، ز و مرادی، س. ۱۳۹۳، مرکز و پیرامون گراف‌های حاصل ضربی، مجله ریاضی شیمی ایران، دوره پنج، شماره اول.
- [۱۱] عرفان‌منش، م.ا و بصریان چهرمی، ر. ۱۳۹۱، شبکه هم تالیفی مقالات منتشر شده در فصلنامه مطالعات ملی کتابداری و سازماندهی اطلاعات با استفاده از شاخص‌های تحلیل شبکه‌های اجتماعی، فصلنامه مطالعات ملی کتابداری و سازماندهی اطلاعات، شماره نود هفتم.
- [۱۲] عزیزی، ع و جعفریان، ا. ۱۳۹۵، یک الگوریتم بهینه سازی جدید ترکیبی با استفاده از روش‌های ازدحام ذرات و خفاش، دومین کنفرانس بین المللی یافته‌های نوین علوم و تکنولوژی، مرکز مطالعات و تحقیقات اسلامی سروش حکمت مرتضوی.
- [۱۳] قدوسی، س و تهرانی، ر و بشیری، م. ۱۳۹۴، بهینه سازی سبد سهام با استفاده از روش تبرید شبیه سازی شده، نشریه تحقیقات مالی، دوره هفدهم، شماره اول.