

# ارائه‌ی یک رویکرد جدید برای حل مسئله‌ی اجتماعیابی شبکه‌های اجتماعی با توسعه‌ی الگوریتم‌های NSGAI و NRGAI

Research Note

وحید بوادران\* (استادیار)

امیرحسین حسینیان (دانشجوی دکتری)

رضا درخشانی (دانشجوی دکتری)

محمد نیک‌ضمیر (دانشجوی دکتری)

گروه مهندسی صنایع، دانشگاه آزاد اسلامی واحد تهران شمال

مسئله‌ی اجتماعیابی، از جمله مسائل تحلیل شبکه‌های اجتماعی، عبارت است از افزایش‌یابی شبکه به بخش‌هایی که ارتباط میان اعضای شبکه در هر بخش متراکم است. در این مقاله، افزایش‌یابی شبکه در قالب مسئله‌ی بهینه‌سازی چندهدفه با توابع هدف، شاخص پودمانگی و امتیاز اجتماعات مدل‌سازی شده است. به‌منظور حل مدل بهینه‌سازی چندهدفه، دو الگوریتم تکاملی NSGAI و NRGAI بهبود یافته است. الگوریتم‌های پیشنهادی برای ایجاد جواب‌های اولیه از یک روش حریم‌ناهی استفاده می‌کنند. همچنین، عملگرهای تقاطع و جهشی جدید برای الگوریتم‌های پیشنهادی طراحی شده است. عملگر تقاطع ارائه شده، مبتنی بر مقدار نزدیکی گره‌هاست. عملگر جهش نیز مبتنی بر روش تاپسیس عمل می‌کند. کارایی الگوریتم‌های پیشنهادی نسبت به سه الگوریتم کلاسیک NSGAI، NRGAI و MOGA-Net از طریق انجام آزمایش‌های عددی بر روی ۶ شبکه‌ی دنیای واقعی مورد سنجش قرار گرفته است. نتایج، نشان از بهبود قابل توجهی در جواب‌های مسئله‌ی اجتماعیابی توسط الگوریتم‌های پیشنهادی داشت.

v\_baradaran@iau-tnb.ac.ir  
 ah\_hosseinian@iau-tnb.ac.ir  
 r\_derakhshani@iau-tnb.ac.ir  
 m\_nikzamir@iau-tnb.ac.ir

واژگان کلیدی: اجتماعیابی، بهینه‌سازی چندهدفه، شبکه‌های پیچیده، روش تاپسیس.

## ۱. مقدمه

با توسعه‌ی علوم و فناوری، انواع متنوعی از شبکه‌ها نظیر شبکه‌های رایانه‌یی، شبکه‌های سیستم برق، شبکه‌های رسانه‌های اجتماعی و شبکه‌های زیستی پدید آمده‌اند و در طول زندگی روزانه در حال تغییر هستند.<sup>[۱]</sup> بسیاری از شبکه‌های دنیای واقعی پیچیده هستند و توسط حجم زیادی از داده‌ها و ارتباطات درونی توصیف می‌شوند. تجزیه و تحلیل شبکه‌ها ابزاری اثر بخش در کشف و جستجوی خصوصیات بالقوه‌ی شبکه‌هاست و منجر به درک بهتر ما از شبکه‌های پیچیده می‌شود.<sup>[۲]</sup> یکی از ویژگی‌های حائز اهمیت گراف‌ها که نشانگر سیستم‌های دنیای واقعی هستند، ساختار اجتماع<sup>۱</sup> است.<sup>[۳]</sup>

کشف ساختارهای پنهان در شبکه، با عنوان مسئله‌ی اجتماعیابی<sup>۲</sup> شناخته می‌شود. مسئله‌ی اجتماعیابی می‌تواند در زمینه‌ی داده‌کاوی، بازیابی اطلاعات و تحلیل شبکه‌های آنلاین فروش موارد استفاده گوناگون داشته باشد.<sup>[۴]</sup> یک اجتماع<sup>۱</sup>

کشف ساختارهای پنهان در شبکه، با عنوان مسئله‌ی اجتماعیابی<sup>۲</sup> شناخته می‌شود. مسئله‌ی اجتماعیابی می‌تواند در زمینه‌ی داده‌کاوی، بازیابی اطلاعات و تحلیل شبکه‌های آنلاین فروش موارد استفاده گوناگون داشته باشد.<sup>[۴]</sup> یک اجتماع<sup>۱</sup>

\* نویسنده مسئول

تاریخ: دریافت ۱۳۹۴/۱۱/۱۳، اصلاحیه ۱۳۹۵/۱۲/۲۸، پذیرش ۱۳۹۶/۳/۲۸

DOI:10.24200/J65.2018.5611

## ۲. پیشینه‌ی پژوهش

شبکه‌های اجتماعی از جمله شبکه‌های پیچیده‌ی هستند که شامل تعداد زیادی از اعضا و ارتباطات بین آن‌ها می‌شوند. یک شبکه اجتماعی می‌تواند در قالب گرافی مدل شود که رئوس آن نماینده‌ی افراد و یال‌های آن نشان‌دهنده ارتباطات موجود بین افراد است.<sup>[۲]</sup> حجم بالای داده‌های موجود در شبکه‌های دنیای واقعی، ضرورت کارایی و دقت بالای الگوریتم‌های خوشه‌بندی را نمایان می‌کند. پیچیدگی محاسباتی یک الگوریتم، با برآورد میزان منابع مورد نیاز آن در انجام یک فعالیت تخمین زده می‌شود.<sup>[۳]</sup> در این تخمین، تعداد گام‌های محاسباتی مورد نیاز و نیز تعداد واحدهای حافظه‌ی مورد نیاز که به صورت هم‌زمان باید به انجام محاسبات تخصیص داده شوند، لحاظ می‌شود. این نیازمندی‌های معمولاً متناسب با اندازه‌ی سیستم مورد مطالعه تنظیم می‌شود. اندازه‌ی یک گراف با تعداد رئوس آن  $N$  یا تعداد یال‌های آن  $E$  معرفی می‌شود. اگر  $\alpha$  و  $\beta$  به ترتیب نشان‌دهنده‌ی توان‌های تعداد رئوس و یال‌های یک گراف باشد،  $O(N^\alpha E^\beta)$  نشان می‌دهد که زمان مورد نیاز برای انجام محاسبات با افزایش تعداد رئوس و یال‌ها افزایش خواهد یافت. هرچه میزان  $\alpha$  و  $\beta$  کمتر باشد، زمان محاسبات کاهش خواهد یافت. مثلاً نمی‌توان با استفاده از الگوریتم‌هایی که زمان اجرای آن‌ها بیشتر از  $O(N)$  یا  $O(E)$  است، گراف‌های شبکه‌ی وب جهانی را که شامل میلیون‌ها گره و میلیاردها یال هستند، مورد بررسی و تحلیل قرار داد. بسیاری از مسائل مرتبط با خوشه‌بندی، در دسته‌ی مسائل NP-Hard قرار می‌گیرند.<sup>[۴]</sup> از این رو، استفاده از الگوریتم‌های دقیق که فقط قابل به کارگیری در سیستم‌های بسیار کوچک هستند، بیهوده است. همچنین، حتی اگر الگوریتمی دارای پیچیدگی چندجمله‌ی باشد، ممکن است در تحلیل و بررسی شبکه‌های بزرگ بسیار کند عمل کند. در چنین مواردی، استفاده از الگوریتم‌هایی که جواب‌های تقریبی در زمان مناسب به دست می‌دهند، بسیار رایج است. الگوریتم‌های تقریبی معمولاً برای مسائل بهینه‌سازی با ابعاد بزرگ مورد استفاده قرار می‌گیرد که در این مسائل، یافتن بیشترین یا کمترین مقدار یک تابع ارزیابی حائز اهمیت است. بنابراین از آنجایی که شبکه‌های دنیای واقعی غالباً ابعاد وسیعی دارند و مسئله‌ی اجتماع‌یابی نیز در دسته‌ی مسائل NP-Hard قرار می‌گیرد<sup>[۵]</sup>، استفاده از هوش محاسباتی در تجزیه و تحلیل این شبکه‌ها می‌تواند نقش مهمی داشته باشد. به عبارت دیگر اهمیت دقت الگوریتم‌های اجتماع‌یابی و سرعت عملکرد آن‌ها در شناسایی ساختارهای اجتماعی در شبکه‌های پیچیده‌ی دنیای واقعی، محققان را بر آن داشته است که از هوش محاسباتی بهره بگیرند.

با توجه به اهمیت مسئله‌ی اجتماع‌یابی، تاکنون روش‌های فراوانی برای تحلیل ساختار شبکه‌های پیچیده معرفی شده و توسعه یافته است. اساس عملکرد این الگوریتم‌ها بر پایه‌ی اصول هوش مصنوعی، نظریه‌ی گراف و مبانی فیزیک است.<sup>[۶]</sup> از معروف‌ترین این الگوریتم‌ها می‌توان به الگوریتم گبروان - نیومن<sup>[۷]</sup> اشاره کرد. کاربرد این روش در بهینه‌سازی معیاری پایه‌ی به نام پودمانگی گزارش شده است. پودمانگی از معمول‌ترین شاخص‌ها در سنجش میزان کیفیت اجتماع‌های به دست آمده است که نخستین بار توسط نیومن و گبروان<sup>[۸]</sup> در قالب یک مسئله‌ی بهینه‌سازی پودمانگی معرفی شد. پودمانگی شاخصی است که کیفیت یک افزایش ویژه از شبکه را به دست می‌دهد. توجیه این امر این است که ارتباطات اعضای یک اجتماع باید از یک گردهمایی تصادفی از افراد بیشتر باشد.<sup>[۹]</sup> هرچه میزان پودمانگی نزدیک به ۱ باشد، بدان معنی است که ساختار اجتماع کیفیت بالایی دارد.<sup>[۱۰]</sup> بنابراین در پژوهش‌ها، مقدار بزرگ شاخص پودمانگی نشان‌دهنده‌ی ساختار قدرتمند اجتماع است و افزایشی که بیشترین مقدار پودمانگی را دارد، از کیفیت بالاتری برخوردار

یک هدف برای انجام بهینه‌سازی وجود دارد. از طرف دیگر، تک‌هدفه در نظر گرفتن این مسئله، اشکالاتی را در پی خواهد داشت.<sup>[۱۱]</sup> از جمله این معایب این است که با بهینه‌سازی یک هدف خاص، ساختارهای اجتماعی کشف شده از شبکه، تنها با تمرکز بر یک معیار شکل می‌گیرند. بنابراین اگر تابع هدف به درستی انتخاب نشود، الگوریتم حل جواب مناسبی ارائه نخواهد کرد.<sup>[۱۲]</sup> به علاوه، اجتماعاتی که بر مبنای یک معیار خاص ایجاد شده‌اند، برای شبکه‌هایی که ساختارهای بالقوه‌ی چندگانه دارند، مناسب نیستند.<sup>[۱۳]</sup> راهکار مناسب به منظور غلبه بر این معایب، تبدیل مسئله به یک مسئله‌ی بهینه‌سازی چندهدفه است. بنابراین با توجه به معایب ذکر شده، در این نوشتار، اجتماع‌یابی به صورت یک مسئله‌ی بهینه‌سازی چندهدفه در نظر گرفته شده است. اهداف مورد مطالعه در این مقاله، مقدار پودمانگی<sup>۵</sup> و امتیاز اجتماعات<sup>۶</sup> است. این دو هدف از شناخته‌شده‌ترین معیارهای سنجش کیفیت افزایش‌های گوناگون از شبکه هستند.<sup>[۱۴]</sup>

به منظور بهینه‌سازی اهداف پودمانگی و امتیاز اجتماعات و بهبود جواب‌های مسئله‌ی اجتماع‌یابی، دو الگوریتم توسعه یافته با عنوان‌های P\_NSGAII<sup>۷</sup> و P\_NRGA<sup>۸</sup> پیشنهاد شده است. این دو روش با توسعه‌ی الگوریتم‌های کلاسیک NSGAII و NPGA ایجاد شده‌اند. الگوریتم‌های پیشنهادی برای ایجاد جواب‌های اولیه از یک روش حریصانه‌ی معرفی شده در مرجع<sup>[۱۵]</sup> استفاده می‌کنند.

ارائه‌ی یک روش جدید تقاطع<sup>۹</sup> و جهش<sup>۱۰</sup> در الگوریتم‌های پیشنهادی از دیگر جنبه‌های نوآوری در این مطالعه است. رویکرد جدید مورد استفاده در عملگر تقاطع، مبتنی بر مقدار نزدیکی<sup>۱۱</sup> گره‌هاست. از آنجایی که کارایی یک شبکه در مسئله‌ی اجتماع‌یابی بر مبنای میانگین فواصل معکوس گره‌ها نسبت به هم تعریف می‌شود، هرچه گره‌ها به یکدیگر نزدیک‌تر باشند کارایی شبکه بالاتر خواهد بود. نزدیکی گره‌ها به یکدیگر منجر به سهولت در انتقال اطلاعات خواهد شد.<sup>[۱۶]</sup> از این حیث، ارائه‌ی روشی که بتواند شاخص نزدیکی گره‌ها را در شناسایی اجتماعات لحاظ کند و قادر به شکل‌دهی افزایش‌های همگون و در عین حال متمایز از یکدیگر باشد، از جنبه‌های نوآوری پیشنهادی در ساختار مسئله‌ی اجتماع‌یابی است. از طرف دیگر، با توجه به آنکه اجتماع‌یابی با کیفیت مناسب شامل گره‌هایی است که ارتباط بین آن‌ها بسیار متراکم است و در نقطه‌ی مقابل، ارتباط کم‌رنگی بین رئوس اجتماعات متفاوت وجود دارد در نظر گرفتن درجه‌ی درونی و بیرونی رئوس در تشکیل اجتماعات حائز اهمیت است. بدین منظور در این نوشتار، یک عملگر جهش جدید در دو الگوریتم پیشنهادی P\_NRGA و P\_NSGAII ارائه شده است که مبتنی بر روش تصمیم‌گیری چندمعیاره<sup>۱۲</sup> تاپسیس است. عملگر پیشنهادی در تولید جواب‌های جدید، سه معیار مرکزیت نزدیکی، درجه‌ی درونی و درجه‌ی بیرونی را در نظر می‌گیرد. بدین سبب ارائه‌ی عملگر جهش جدید نیز از دیگر موارد نوآوری این نوشتار در ساختار مسئله است؛ زیرا در شناسایی اجتماعات و ساختارهای اجتماعی شبکه، سه معیار فوق لحاظ شده است.

ساختار مقاله، در ادامه به این ترتیب است: به منظور شناخت و آشنایی بیشتر با مسئله‌ی اجتماع‌یابی، در بخش دوم به مرور تحقیقات مرتبط می‌پردازیم. بخش سوم به تشریح کامل مسئله اختصاص یافته است. بخش چهارم در ارتباط با معرفی الگوریتم‌های پیشنهادی است. در بخش پنجم، عملکرد الگوریتم‌های پیشنهادی با تعریف معیارهای متفاوت مورد سنجش قرار گرفته است. بخش ششم نیز به جمع‌بندی مطالب گفته شده و ارائه‌ی پیشنهادهایی برای تحقیقات آتی اختصاص داده شده است.

برازندگی بیشتری دارند، شانس بیشتری برای بقا در نسل بعدی کروموزوم‌ها خواهند داشت.<sup>[۲۷]</sup>

نخستین بار بینگول و تاسگین<sup>[۲۸]</sup> از الگوریتم ژنتیک برای حل مسئله‌ی اجتماع‌یابی با هدف پودمانگی استفاده کردند. در مطالعه‌ی دیگر، تاسگین و بینگول<sup>[۲۹]</sup> افزایش مختلف از شبکه را به عنوان کروموزوم‌های الگوریتم ژنتیک در نظر گرفتند. در این مطالعه نیز تابع برازندگی، پودمانگی اجتماعات است. امیری و همکارانش الگوریتم ژنتیکی را به منظور بهینه‌سازی توأم دو شاخص امتیاز و برازش اجتماعات ارائه کردند.<sup>[۳۰]</sup>

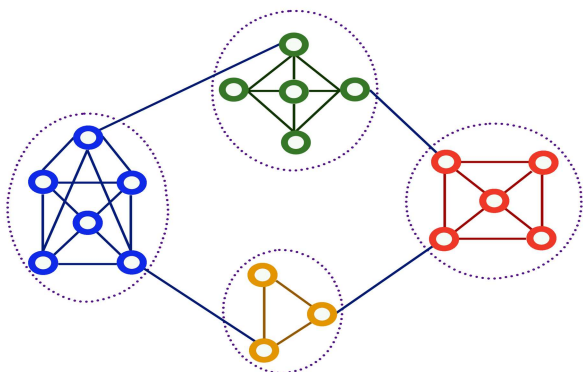
یان و همکاران<sup>[۳۱]</sup> الگوریتم ژنتیکی ارائه کرده‌اند که جمعیت اولیه‌ی آن بر اساس شباهت گره‌ها شکل می‌گیرد. این امر منجر به ایجاد جواب‌های اولیه متنوع می‌شود. پژوهشگران در این مقاله نشان دادند که استفاده از این روند، دقت و کارایی الگوریتم را در جستجوی جواب بهینه بهبود می‌دهد. در این مطالعه از ساختار نمایش ماتریسی<sup>۲۱</sup> استفاده شده است که در این ساختار نیاز به انجام فرایند کدگذاری برای به کارگیری عملگرهای تقاطع و جهش نیست. لی و لیو<sup>[۳۲]</sup> یک الگوریتم ژنتیک چندعاملی<sup>۲۲</sup> برای بهینه‌سازی شاخص پودمانگی ارائه کرده‌اند. هر عامل که نشان‌دهنده‌ی جوابی از مسئله‌ی اجتماع‌یابی است، به عنوان بخشی از شبکه کدگذاری می‌شود. در این پژوهش، چندین عملگر جدید برای تولید جواب‌های همسایگی پیشنهاد شده است. این عملگرها عبارت‌اند از عملگر انشعاب<sup>۲۳</sup>، عملگر ترکیب مبتنی بر رقابت همسایگی<sup>۲۴</sup>، عملگر تقاطع همسایگی ترکیبی<sup>۲۵</sup>، عملگر جهش تطبیقی<sup>۲۶</sup>، و عملگر خودفراگیر<sup>۲۷</sup> شی و همکارانش<sup>[۳۳]</sup> الگوریتم ژنتیک جدیدی به همراه یک برنامه‌ی کدگذاری لوکاسی<sup>۲۸</sup> ویژه پیشنهاد کرده‌اند. ساختار جواب پیشنهادی در این تحقیق، این امکان را فراهم می‌آورد که تعداد اجتماعات به صورت خودکار و تطبیقی تعیین شود. در نتیجه فرایند اجتماع‌یابی با انعطاف بیشتری صورت می‌پذیرد. در مطالعه‌ی دیگر، قربانیان و شقاقی<sup>[۸]</sup> الگوریتم ژنتیکی ارائه کردند که ساختار نمایش جواب آن ماتریسی است و فرایند بهینه‌سازی مبتنی بر معیار چگالی پودمانگی<sup>۲۹</sup> صورت می‌پذیرد. شانگ و همکارانش<sup>[۳۴]</sup> الگوریتم تلفیقی متشکل از ژنتیک و شبیه‌سازی تبرید برای بهینه‌سازی میزان پودمانگی پیشنهاد کردند. علاوه بر موارد ذکر شده، در زمینه‌ی اجتماع‌یابی چندهدفه نیز تحقیقاتی صورت پذیرفته است. از جمله‌ی این تحقیقات می‌توان به پژوهش شی و همکارانش<sup>[۱۹]</sup> اشاره کرد. در این تحقیق، به منظور بهینه‌سازی دو هدف روابط داخلی و خارجی بین گره‌ها، الگوریتمی با عنوان MOCD پیشنهاد شد. کاربرد الگوریتم ژنتیک در مسئله‌ی اجتماع‌یابی چندهدفه نیز در پژوهش‌های پیشین گزارش شده است. پیتزوتی<sup>[۱۱]</sup> الگوریتم MOGA-Net را برای بهینه‌سازی معیارهای پودمانگی و امتیاز (برازش) اجتماعات<sup>۳۰</sup> معرفی کرد. در پژوهشی دیگر، چن و همکاران<sup>[۳۵]</sup> یک روش تکاملی چندهدفه مبتنی بر الگوریتم ژنتیک با مرتب‌سازی غیرمغلوب برای شناسایی اجتماعات موجود در شبکه‌های پویا پیشنهاد کرده‌اند. شناسایی اجتماعات در این الگوریتم با بهینه‌سازی شاخص چگالی پودمانگی صورت پذیرفته است. مازورو و همکارانش<sup>[۳۶]</sup>، دو هدف پودمانگی و امتیاز اجتماعات را به عنوان توابع برازندگی الگوریتم ژنتیک به کار گرفته شده در مطالعه‌ی خود لحاظ کردند. همچنین الگوریتم پیشنهادی در این تحقیق امکان کشف اجتماعات کوچک را، که حتی ممکن است متشکل از یک گره باشند، نیز فراهم آورده است.

علاوه بر معرفی الگوریتم‌های متنوع در مسئله‌ی اجتماع‌یابی، توابع هدف گوناگونی نیز برای درک ساختار اجتماعات در شبکه پیشنهاد شده است. در زمینه‌ی نظریه‌ی گراف<sup>۳۱</sup>، کانان و ومپالا<sup>[۳۷]</sup> دو معیار وسعت<sup>۳۲</sup> و رسانایی<sup>۳۳</sup> را معرفی کردند. معیار وسعت تعداد یال‌هایی را که از هر گره به سمت خارج اجتماع می‌روند، اندازه‌گیری

است.<sup>[۱۵]</sup> بدین سبب، جستجو برای یافتن افزاری از شبکه که مقدار بهینه‌ی پودمانگی را دارد، از جمله مسائل NP-Hard محسوب می‌شود.<sup>[۹]</sup> با توجه به آنکه بهینه‌سازی پودمانگی از جمله مسائل NP-Hard است، جستجوی تمام افزاری‌های ممکن امری بسیار زمان‌بر و پرهزینه است. به این دلیل روش‌های ابتکاری و فرابابتکاری متعددی نظیر الگوریتم‌های حریمانه<sup>۱۴</sup>، روش‌های نمونه برداری، شبیه‌سازی تبرید<sup>۱۵</sup> و بهینه‌سازی اکستریمال<sup>۱۶</sup> برای حل مسئله به کار گرفته شده‌اند.<sup>[۱۶]</sup> مزیت ویژه‌ی این الگوریتم‌ها به دست آوردن تعداد اجتماعات در طی روند تکاملی حل مسئله است.<sup>[۱۷]</sup> از جمله تحقیقاتی که از این الگوریتم‌ها برای یافتن افزاری با کیفیت استفاده کرده‌اند، می‌توان به پژوهش شقاقی و تیمورپور<sup>[۱۳]</sup> اشاره کرد. این محققان یک الگوریتم حریمانه‌ی قدرتمند ولی در عین حال ساده به منظور بهینه‌سازی پودمانگی پیشنهاد داده‌اند. در تحقیقی مگانانان<sup>[۱۸]</sup> یک الگوریتم حریمانه ارائه کرده است که با بهینه‌سازی شاخص پودمانگی به شناسایی اجتماعاتی که همپوشانی دارند، پرداخته است. رویکرد حریمانه‌ی دیگری توسط بلوندل و همکاران<sup>[۱۹]</sup> برای گراف‌های موزون ارائه شده است. این رویکرد در ابتدا گره‌های گراف را در اجتماعات مختلف قرار می‌دهد و سپس میزان پودمانگی را محاسبه می‌کند. در گام بعد، بررسی می‌شود که در صورت انتقال گره‌ها به اجتماعات همسایگی، میزان پودمانگی به چه میزان افزایش می‌یابد. انتقالی که منجر به بیشترین افزایش پودمانگی شود، مورد پذیرش قرار می‌گیرد. روند تغییر اجتماعاتی که گره‌ها به آن‌ها تعلق دارند، در جهت بهینه‌سازی پودمانگی ادامه می‌یابد. گوئیمرا و آمارال<sup>[۱۶]</sup> نخستین بار از الگوریتم شبیه‌سازی تبرید به منظور بهینه‌سازی پودمانگی استفاده کردند. این محققان در روند جستجوی همسایگی الگوریتم پیشنهادی خود، دو حرکت محلی<sup>۱۷</sup> و سراسری<sup>۱۸</sup> را گنجانده‌اند. در حرکت محلی، هر گره به صورت تصادفی از یک اجتماع به اجتماعی دیگر منتقل می‌شود. در حرکت سراسری، اجتماعات کشف شده با یکدیگر ادغام می‌شوند یا به اجتماعات کوچکتر تقسیم می‌شوند. ماسین و دوی<sup>[۲۰]</sup> در تحقیق خود نشان دادند که استفاده از حرکت سراسری ریسک گرفتار شدن الگوریتم شبیه‌سازی تبرید در بهینه‌ی محلی را کاهش می‌دهد. همچنین ثابت کرده‌اند که استفاده از حرکت سراسری منجر به دستیابی به جواب‌های بهتری نسبت به حرکت محلی می‌شود. بهینه‌سازی اکستریمال نیز از جمله روش‌های ابتکاری است که در سال ۲۰۰۱ پیشنهاد شد. این الگوریتم، دقتی نزدیک به دقت شبیه‌سازی تبرید دارد اما در عین حال زمان محاسبات آن کمتر است.<sup>[۲۱]</sup> این روش توسط دوچ و آرناس.<sup>[۲۲]</sup>

برای بهینه‌سازی پودمانگی مورد استفاده قرار گرفته است. در تحقیقی دیگر، یک روش جدید برای کشف اجتماعات بر مبنای نظریه‌ی اطلاعات<sup>۱۹</sup> پیشنهاد شد.<sup>[۲۳]</sup> اولین روشی که بر مبنای الگوریتم نفوذ گروهی توانایی شناسایی اجتماعات هم‌پوشانده را داشت، توسط پالا و همکارانش ارائه شد.<sup>[۲۴]</sup> بر همین مینا الگوریتم‌های SCP و EO پیشنهاد شدند.<sup>[۲۵-۲۲]</sup> کافیری و همکارانش نیز با استفاده از یک الگوریتم قطعی، اقدام به بهینه‌سازی شاخص پودمانگی در شبکه‌ها نمودند.<sup>[۲۶]</sup>

الگوریتم ژنتیک از جمله روش‌های تکاملی است که در بهینه‌سازی توابع هدف گوناگون مسئله‌ی اجتماع‌یابی بسیار مورد استفاده قرار گرفته است. در الگوریتم ژنتیک استاندارد، مجموعه‌ی جواب‌های کاندید برای یک مسئله وجود دارند که به صورت عددی در قالب کروموزوم‌ها کدگذاری می‌شوند. همچنین تابع هدفی وجود دارد که باید با توجه به محدودیت‌های مسئله بهینه شود. این تابع هدف نقش یک برازندگی زیستی<sup>۲۰</sup> را برای کروموزوم‌ها ایفا می‌کند. الگوریتم ژنتیک از مجموعه‌ی تصادفی از جواب‌ها فرایند حل مسئله را شروع می‌کند و به تدریج این مجموعه توسط عملگرهای تقاطع و جهش تغییر می‌یابند. برازندگی کروموزوم‌ها در طی روند حل مسئله محاسبه خواهد شد و کروموزوم‌هایی که



شکل ۱. ارتباط متراکم اعضای یک اجتماع با یکدیگر.

مشترکی دارند کاربرد داشته باشد.<sup>[۳۹]</sup> خوشه‌بندی<sup>۳۷</sup> مشتریانی که علائق مشترک دارند و از نظر موقعیت جغرافیایی نزدیک به یکدیگر هستند، می‌تواند منجر به بهبود ارائه‌ی خدمات شود. شناسایی اجتماعاتی از مشتریان که علائق مشابه دارند، امکان پیشنهاد محصولاتی را که متناسب با علائق آن‌ها هستند، فراهم می‌آورد. از این حیث، فرصت‌های تجاری مناسبی به وجود می‌آید. در علم کامپیوتر نیز شناسایی اجتماعات یکی از موضوعات حائز اهمیت است. مثلاً در مبحث پردازش موزی،<sup>۳۸</sup> دانستن بهترین روش برای تخصیص عملیات به پردازنده‌ها به گونه‌ی که ارتباط میان آن‌ها کمینه شود و فرایند محاسبات به سرعت انجام شود، از اهمیت بالایی برخوردار است. این امر زمانی محقق می‌شود که یک اجتماع از رایانه‌ها به گروه‌هایی هم اندازه از پردازشگرها تقسیم شوند، به طوری که تعداد ارتباطات فیزیکی میان پردازنده‌های گروه‌های مختلف کمینه باشد.<sup>[۴۱]</sup> اجتماع‌یابی در شکل‌دهی اجتماعات در شبکه‌های استنادی تحقیقات نیز می‌تواند موارد استفاده‌ی فراوان داشته باشد. تشکیل اجتماعات در شبکه‌های استنادی بر پایه‌ی موضوع تحقیق صورت می‌پذیرد.<sup>[۳۹]</sup> با توجه به کاربردهای ذکر شده، هر اجتماع یا خوشه را می‌توان مجموعه‌ی از گره‌ها دانست که خصوصیات مشترک دارند یا نقش مشابهی را در شبکه ایفا می‌کنند. فرض می‌کنیم شبکه‌ی  $N_w$  در دست است. این شبکه را می‌توان به صورت گراف  $G(V, E, W)$  مدل کرد که در آن  $V$  نشان‌دهنده‌ی رئوس گراف و متناسب با هر فرد در شبکه است.  $E$  نشانگر یال‌هایی است که ارتباطات دو طرفه بین افراد را برقرار می‌کنند و  $W$  ماتریس وزن شبکه است. همان‌طور که در شکل ۱ قابل مشاهده است، یک اجتماع یا خوشه، زیرگرافی از شبکه است که چگالی یال‌ها بین اعضای اجتماع بالاست و تراکم یال‌ها از اعضای زیرگراف به رئوس که در داخل اجتماع قرار ندارند، پایین است. خط‌چین‌ها در شکل ۱ نشان‌دهنده‌ی اجتماعات است. مسئله‌ی اجتماع‌یابی به شناسایی  $k$  اجتماع در شبکه می‌پردازد، به گونه‌ی که مقدار  $k$  از پیش مشخص نیست. اجتماع‌یابی،  $k$  زیرگراف از گره‌های شبکه را شناسایی می‌کند، به طوری که ارتباط گره‌های داخل این  $k$  زیرگراف بسیار متراکم است و ارتباط گره‌های داخل این  $k$  زیرگراف با گره‌های خارج از اجتماع ضعیف است.<sup>[۴۰]</sup>

می‌کند. شاخص رسانایی نیز بیانگر تعداد کل یال‌هایی است که از یک خوشه خارج می‌شوند. در حیطه‌ی نظریه‌ی اطلاعات، روسوال و برگستروم<sup>[۲۳]</sup> شاخصی را با عنوان امتداد توصیف<sup>۳۴</sup> ارائه کردند که بتوان به تراکم بهینه‌یابی از توپولوژی شبکه دست یافت. برخی از توابع هدف مرسوم این مسئله به صورت اجمالی در مرجع<sup>[۳۸]</sup> معرفی شده‌اند. با این حال همان‌طور که در بخش قبل نیز ذکر شد، اجتماع‌یابی به صورت یک مسئله‌ی بهینه‌سازی تک‌هدفه می‌تواند شامل معایب زیر باشد:<sup>[۱۰]</sup>

۱. الگوریتم‌های اجتماع‌یابی تک‌هدفه، شناسایی اجتماعات را با توجه به یک معیار خاص انجام می‌دهند. بنابراین با توجه به توابع هدف انتخاب شده برای شناسایی اجتماعات در یک شبکه خاص، جواب‌های متفاوتی حاصل خواهد شد.

۲. در صورتی که تابع هدف مناسب انتخاب نشود، الگوریتم نیز جواب‌های مناسبی ارائه نخواهد کرد.

۳. بسیاری از روش‌های پیشنهاد شده در حل مسائل تک‌هدفه، به برخی اطلاعات اولیه از جمله تعداد اجتماعات نیاز دارند. در حالی که این اطلاعات برای شبکه‌های دنیای واقعی عمدتاً مجهول است.

۴. استفاده از مدل‌سازی تک‌هدفه برای برخی از شبکه‌ها از جمله شبکه‌های سلسله‌مراتبی یا شبکه‌هایی که دارای هم‌پوشانی‌اند، مناسب نیست.

به علاوه برای اجتماع‌یابی چندهدفه، مزایایی نسبت به حالت تک‌هدفه در پژوهش‌ها گزارش شده است.<sup>[۱۰]</sup>

۱. مجموعه‌ی جواب‌های غیرمغلوب<sup>۳۵</sup> جبهه‌ی پارتو<sup>۳۶</sup> حاصل از حل مسئله‌ی اجتماع‌یابی چندهدفه، همواره شامل جواب بهینه‌ی حاصل از حل مسئله‌ی اجتماع‌یابی به صورت تک‌هدفه می‌شود. در واقع جواب بهینه‌ی مسئله‌ی اجتماع‌یابی تک‌هدفه همواره عضوی از مجموعه جواب‌های غیرمغلوب این مسئله در حالت چندهدفه است.

۲. استفاده از چندین هدف در مسئله، منجر به شناسایی اجتماعات از مناظر مختلف می‌شود. به این ترتیب، فرایند بهینه‌سازی چندهدفه، با در نظر گرفتن توازن میان اهداف، از قرار گرفتن جواب در بهینه‌ی محلی جلوگیری می‌کند.

۳. تعداد بهینه‌ی اجتماعات حین فرایند بهینه‌سازی چندهدفه به صورت پویا تعیین می‌شود. روابط میان توابع هدف می‌تواند تعداد اجتماعات شناسایی شده را به صورت پویا مشخص کند و از هم‌گرایی شدن الگوریتم به سمت کشف جواب‌های بی‌اهمیت، جلوگیری کند. با توجه به دلایل ذکر شده، در این تحقیق مدل اجتماع‌یابی به صورت چندهدفه و با هدف بیشینه‌سازی شاخص‌های پودمانگی و امتیاز اجتماعات در نظر گرفته شده است. همچنین به دلیل عملکرد موفق الگوریتم ژنتیک در حل این مسئله، روش‌های توسعه‌یافته مبتنی بر الگوریتم ژنتیک هستند.

### ۳. تعریف مسئله و بیان کاربردهای اجتماع‌یابی

مسئله‌ی اجتماع‌یابی می‌تواند در زمینه‌ی جامعه‌شناسی، علوم زیستی، رایانه و به طور کلی علوم که امکان نمایش سیستم‌ها در آن‌ها به صورت گراف وجود دارد، کاربرد قابل توجهی پیدا کند.<sup>[۶]</sup> ساختار اجتماعی در این مسئله می‌تواند در تحلیل شبکه‌های اجتماعی، داده‌کاوی، بازیابی اطلاعات، شبکه‌های مشارکت علمی میان پژوهشگران و شبکه‌های آنلاین خرده‌فروشی متشکل از مشتریانی که سود و منفعت

### ۴. الگوریتم‌های پیشنهادی

#### ۴.۱. ساختار نمایش جواب

در این مطالعه از ساختار نمایش لوکاسی استفاده شده است. در این ساختار که مبتنی بر گراف است، هر ژنوتیپ<sup>۳۹</sup>  $g$ ، تعداد  $n$  ژن را به ترتیب  $g_1, \dots, g_n$  شامل می‌شود. در این حالت، هر ژن  $g_i$  می‌تواند هر کدام از گره‌های همسایه‌ی گره  $i$  را

### ۲.۴. ایجاد جمعیت اولیه

با وجود موفقیت الگوریتم‌های فراابتکاری در حل بسیاری از مسائل بهینه‌سازی، این بیم وجود دارد که این الگوریتم‌ها در بهینه‌ی محلی گرفتار شوند و به همگرایی زودرس برسند. بنابراین در این مطالعه به منظور کاهش احتمال گرفتار شدن الگوریتم‌ها در بهینه‌ی محلی و به دست آمدن جواب‌های با کیفیت، از الگوریتم حریرانه‌ی ارائه شده در تحقیقات قبلی<sup>[۱۳]</sup> برای تولید جواب‌های اولیه استفاده شده است. روش حریرانه از یک نقطه‌ی اولیه شروع می‌کند و در هر تکرار، جواب فعلی با یکی از بهترین جواب‌های همسایه جابه‌جا خواهد شد. هنگامی جابه‌جایی صورت می‌پذیرد که جواب همسایگی از نظر تمام توابع هدف از جواب فعلی بهتر باشد. شرط توقف الگوریتم تعداد تکرار از پیش تعیین شده است. مزیت روش حریرانه‌ی مورد استفاده، سرعت بالا و پیچیدگی پایین آن است.<sup>[۱۳]</sup> بر اساس رویکرد پیشنهادی، مسئله‌ی اجتماع‌یابی در ابتدا با هدف بیشینه‌سازی پودمانگی و امتیاز اجتماعات توسط الگوریتم حریرانه حل می‌شود و پس از تعداد تکرار معین، جواب‌هایی که بیشترین میزان برازندگی را دارند، برگزیده خواهند شد. در ادامه جواب‌های تولید شده به عنوان جمعیت اولیه‌ی الگوریتم‌های NSGAI و NPGA در نظر گرفته می‌شوند و روند حل مسئله به وسیله‌ی این الگوریتم‌ها ادامه پیدا می‌کند.

### ۳.۴. توابع برازندگی

در پژوهش‌های پیشین معیارهای فراوانی برای ارزیابی اجتماعات پیشنهاد شده است و تعداد زیادی از الگوریتم‌های تکاملی که اجتماع‌یابی را به صورت تک‌هدفه یا چندهدفه مورد مطالعه قرار داده‌اند، معرفی شده است. وجه اشتراک این روش‌ها این است که مسئله‌ی اجتماع‌یابی به صورت یک مسئله‌ی بهینه‌سازی مدنظر قرار گرفته است. همان‌طور که در بخش‌های قبلی گفته شد، در این مقاله دو معیار پودمانگی و امتیاز اجتماعات به عنوان معیارهای ارزیابی اجتماعات کشف شده توسط الگوریتم‌های پیشنهادی در نظر گرفته شده‌اند که این بخش به شرح این دو شاخص اختصاص داده شده است.

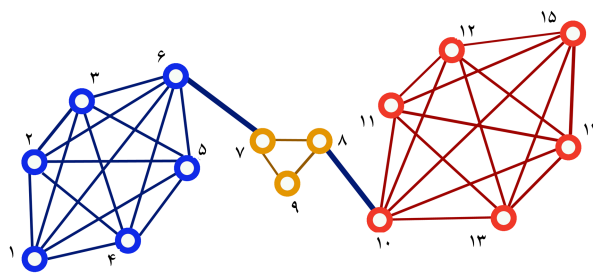
#### ۱.۳.۴. پودمانگی

شاخص پودمانگی  $Q$  معیاری کمی برای افراز شبکه است که مشخص می‌کند رئوس گراف با چه کیفیتی به اجتماعات گوناگون افراز شده‌اند. پودمانگی نشان می‌دهد که در شبکه‌هایی که یال‌ها به صورت تصادفی ارتباط بین گره‌ها را برقرار می‌کنند، ساختار اجتماعی وجود ندارد. مقدار شاخص پودمانگی با توجه به رابطه‌ی (۱) تعریف می‌شود:<sup>[۲۱]</sup>

$$Q = \frac{1}{\sum_{ij} A_{ij}} \sum_{ij} \left( A_{ij} - \frac{k_i k_j}{\sum_{ij} A_{ij}} \right) \delta(C_i, C_j) \quad (1)$$

در رابطه‌ی ۱،  $A$  نشان‌دهنده‌ی ماتریس مجاورت است. اگر گره  $i$  و  $j$  با یکدیگر در ارتباط باشند، درایه‌ی  $A_{ij}$  برابر ۱ و در غیر این صورت مقدار آن صفر است.  $l$  برابر تعداد کل یال‌های شبکه است. اگر دو گره  $i$  و  $j$  متعلق به یک اجتماع باشند، مقدار تابع  $\delta$  برابر یک و در غیر این صورت صفر است.  $C_i$  و  $C_j$  به ترتیب نشان‌دهنده‌ی اجتماعاتی است که گره‌های  $i$  و  $j$  به آن‌ها تعلق دارند.

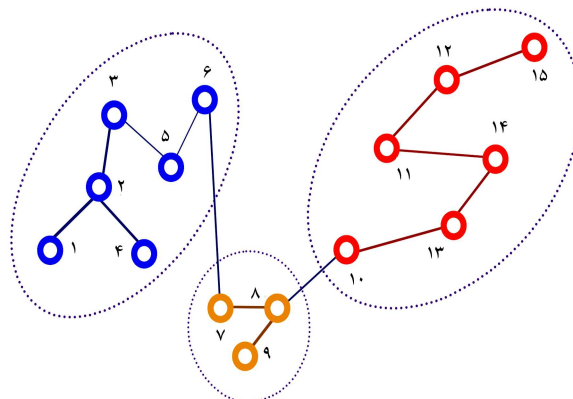
شاخص پودمانگی را می‌توان برای شبکه‌های وزنی نیز مورد استفاده قرار داد.<sup>[۲۱]</sup> در این حالت، مقادیر  $k_i$  و  $k_j$  در رابطه‌ی (۱)، با میزان نیرومندی<sup>۲۲</sup> گره‌ها جایگزین می‌شود. میزان نیرومندی گره‌های  $i$  و  $j$  به ترتیب با  $s_i$  و  $s_j$  نمایش داده می‌شود.



شکل ۲. ساختار اصلی گراف.

محل	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵
کروموزوم	۲	۳	۵	۲	۶	۷	۸	۱۰	۸	۱۳	۱۲	۱۵	۱۴	۱۱	۱۲

شکل ۳. نمونه‌ی یکی از یک کروموزوم امکان‌پذیر (شدنی).



شکل ۴. ساختار اجتماعات و شبکه‌ی متناظر با کروموزوم شکل ۳.

انتخاب می‌تواند در برگزیده‌ی خود گره  $z$ ام نیز باشد. بنابراین، هر مقدار از  $z$  که به ژن  $z$ ام تخصیص داده می‌شود، به عنوان ارتباطی بین گره  $i$  و  $z$  تفسیر می‌شود. در نتیجه، این گره‌ها در اجتماع یکسانی قرار می‌گیرند. کدگشایی این ساختار نیازمند شناسایی تمام رئوس مرتبط با یکدیگر است. گام کدگشایی با استفاده از یک برنامه‌ی پیمایش معکوس<sup>۴۰</sup> قابل انجام است. از سوی دیگر از آنجایی که در گام کدگشایی تعداد اجتماعات تعیین می‌شوند، دانستن تعداد آن‌ها از قبل الزامی ندارد.<sup>[۳۳]</sup> به عنوان نمونه، فرض می‌کنیم که یک شبکه به صورت شکل ۲ مفروض است. این گراف ۱۵ گره دارد و می‌توان آن را به سه اجتماع تقسیم‌بندی کرد. نمونه‌ی یکی از یک کروموزوم امکان‌پذیر<sup>۴۱</sup> در شکل ۳ به تصویر کشیده شده است. ساختار گراف این کروموزوم به صورت شکل ۴ است. با توجه به شکل ۴، هر اجتماع زیر مجموعه‌ی رئوس شبکه است.

ساختار نمایش لوکاسی سه مزیت عمده دارد:<sup>[۱۲]</sup>

- این ساختار جواب از آنجایی که عملگرهای تقاطع و جهش در الگوریتم ژنتیک کروموزوم غیرمعتبر و نشدنی ایجاد نمی‌کند، مناسب است.
- فضای جواب ایجاد شده توسط این ساختار به صورت قابل ملاحظه‌ی کاهش می‌یابد.
- دامنه‌ی اعداد در ساختار نمایش جواب از ۱ تا  $n$  است و نشانگر اجتماعی است که هر گره به آن تعلق دارد. پیچیدگی فضای جواب این ساختار برابر  $n^n$  است.

به‌طور کلی امتیاز اجتماع‌یابی در شبکه‌ی  $k$  شامل اجتماع است، از رابطه‌ی (۹) به دست خواهد آمد:<sup>[۱۲]</sup>

$$CS = \sum_{i=1}^k score(C_i) \quad (9)$$

در رابطه‌ی (۹)، امتیاز تمام اجتماعات شبکه محاسبه می‌شود و در نهایت با یکدیگر جمع می‌شوند تا امتیاز اجتماع‌یابی به دست بیاید.

#### ۴.۴. مرکزیت نزدیکی<sup>۴۳</sup>

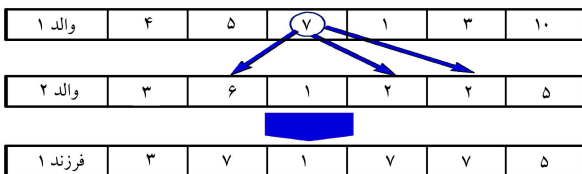
در گراف‌های بسته، یک معیار فاصله‌ی طبیعی میان زوج گره‌ها وجود دارد که توسط طول کوتاه‌ترین مسیرهای بین آن‌ها تعریف می‌شود.<sup>[۴۳]</sup> در این پژوهش به منظور به دست آوردن کوتاه‌ترین مسیرها بین زوج گره‌های گراف از الگوریتم فلوئید - وارشال<sup>۴۴</sup> استفاده شده است. الگوریتم فلوئید - وارشال، روشی است که برای یافتن کوتاه‌ترین مسیرها در شبکه‌های وزنی با اوزان مثبت و منفی یال‌ها به کار گرفته می‌شود. میزان دوری گره  $x$  با مجموع فواصل آن از تمام گره‌ها محاسبه می‌شود. بولاس میزان نزدیکی گره‌ها را به صورت معکوس مقدار دوری گره  $x$  تعریف کرد که مقدار آن از رابطه‌ی (۱۰) به دست می‌آید:<sup>[۴۴]</sup>

$$C(x) = \frac{1}{\sum_y d(y, x)} \quad (10)$$

در این رابطه، میزان نزدیکی گره  $x$  با  $C(x)$  نشان داده می‌شود. مقدار  $d(y, x)$  برابر کوتاه‌ترین فاصله بین دو گره  $x$  و  $y$  است. بنابراین هرچه درجه‌ی مرکزیت یک گره بیشتر باشد، مجموع فواصل این گره از تمام گره‌ها کمتر است.

#### ۵.۴. تقاطع

در این بخش، یک عملگر جدید تقاطع مبتنی بر مقدار نزدیکی گره‌ها معرفی می‌شود. بر این اساس، گره‌ی که در هر والد دارای بالاترین مقدار نزدیکی است، انتخاب خواهد شد. همسایگی‌های این گره با توجه به شبکه‌ی اصلی در نظر گرفته می‌شود و گره انتخابی در محل همسایگی‌های خود در والد دیگر قرار می‌گیرد و فرزند جدید به وجود می‌آید. مثلاً در شکل ۵ فرض می‌کنیم گره شماره‌ی ۷ بالاترین مقدار نزدیکی را در والد یک دارد. با فرض آنکه در شبکه‌ی اصلی، همسایگی‌های گره ۷ برابر رئوس ۲، ۴ و ۵ باشند، این گره در والد ۲ در محل گره‌های همسایگی خود قرار می‌گیرد و فرزند ۱ ایجاد می‌شود. در والد ۲ نیز فرض می‌کنیم گره ۵ بالاترین میزان نزدیکی را دارد. همسایگی‌های این گره برابر رئوس ۱ و ۳ در شبکه هستند. این بار گره ۵ در محل همسایگی‌های خود در والد ۱ قرار می‌گیرد و فرزند دوم با توجه به شکل ۶ حاصل خواهد شد. این روش تقاطع از آنجایی که نزدیکی بین گره‌ها را به عنوان معیاری برای تشکیل فرزندان در نظر می‌گیرد، می‌تواند در تشکیل اجتماعات موفق عمل کند و منجر به بهبود پودمانگی و امتیاز اجتماعات شود. در این مقاله، عملگر تقاطع مذکور برای دو الگوریتم پیشنهادی P\_NRGA و P\_NSAGII مورد استفاده قرار خواهد گرفت.



شکل ۵. نحوه‌ی ایجاد فرزند اول بر اساس مقدار نزدیکی گره‌ها.

مقدار نیرومندی هر گره، برابر مجموع اوزان یال‌هایی است که در مجاورت این گره قرار دارند. تعداد کل یال‌های شبکه‌ی  $l$  در رابطه‌ی (۱)، با مجموع اوزان تمام یال‌های شبکه،  $W$  جایگزین می‌شود. به این ترتیب مقدار پودمانگی در شبکه‌های وزنی از رابطه‌ی (۲) قابل محاسبه است:<sup>[۱]</sup>

$$Q_w = \frac{1}{\sqrt{W}} \sum_{ij} \left( W_{ij} - \frac{s_i s_j}{\sqrt{W}} \right) \delta(C_i, C_j) \quad (2)$$

#### ۲.۳.۴. امتیاز اجتماعات

فرض می‌کنیم یک شبکه به صورت گراف  $G = (V, E)$  وجود دارد که  $V$  مجموعه‌ی گره‌ها و  $E$  مجموعه‌ی یال‌هاست. در این صورت، درجه‌ی گره  $i$  در زیرشبکه  $C$  به گونه‌ی  $C \subset G$  است، از رابطه‌ی (۳) قابل محاسبه خواهد بود:<sup>[۱۲]</sup>

$$k_i(c) = k_i^{in}(c) + k_i^{out}(c) \quad (3)$$

در رابطه‌ی (۳)،  $k_i^{in}(c)$  بیانگر تعداد یال‌هایی است که گره  $i$  را به سایر گره‌ها در زیرشبکه  $C$  متصل می‌کند و  $k_i^{out}(c)$  بیانگر تعداد یال‌هایی است که گره  $i$  را به سایر اعضای شبکه که خارج از اجتماع  $C$  قرار دارند، متصل می‌کند. بر این اساس، مقادیر  $k_i^{in}(c)$  و  $k_i^{out}(c)$  به ترتیب از طریق روابط (۴) و (۵) محاسبه می‌شود:<sup>[۱۲]</sup>

$$k_i^{in}(c) = \sum_{j \in C} A_{ij} \quad (4)$$

$$k_i^{out}(c) = \sum_{j \notin C} A_{ij} \quad (5)$$

زیر شبکه‌ی  $C$ ، یک اجتماع است اگر شرط

$$\sum_{i \in c} k_i^{in}(c) > \sum_{i \in c} k_i^{out}(c)$$

برقرار باشد.<sup>[۴۳]</sup> حال فرض می‌کنیم  $\mu_i$  نشان‌دهنده‌ی نسبت یال‌هایی باشد که گره  $i$  را به سایر گره‌ها وصل می‌کند. بدین سبب  $\mu_i$  از رابطه‌ی (۶) قابل محاسبه است:<sup>[۱۲]</sup>

$$\mu_i = \frac{1}{|C|} k_i^{in}(c) \quad (6)$$

در رابطه‌ی (۶)،  $|C|$  برابر با تعداد اعضای زیرشبکه  $C$  است. میانگین توانی  $C$  از مرتبه‌ی  $r$  را با  $M(C)$  نشان می‌دهند که به صورت رابطه‌ی (۷) تعریف می‌شود:<sup>[۱۲]</sup>

$$M(C) = \frac{\sum_{i \in c} (\mu_i)^r}{|c|} \quad (7)$$

از طرف دیگر حجم یک زیرشبکه برابر تعداد یال‌هایی است که گره‌های زیر شبکه  $C$  را به هم وصل می‌کنند. حجم زیرشبکه‌ی  $C$  با  $v_c$  نشان داده می‌شود. شاخص امتیاز اجتماعات از رابطه‌ی (۸) به دست می‌آید:<sup>[۱۲]</sup>

$$score(C) = M(C) \times v_c \quad (8)$$

همان‌طور که از رابطه‌ی (۸) مشخص است، شاخص امتیاز اجتماعات میانگین توانی  $M(C)$  و حجم روابط داخلی یک اجتماع  $v_c$  را به صورت هم‌زمان در نظر می‌گیرد.

جدول ۱. سطوح پارامترهای ورودی الگوریتم‌ها.

الگوریتم	پارامترها	نمادها	سطوح	سطح بهینه
NSGAI	نرخ تقاطع	Pc	۰/۸-۰/۶-۰/۴	۰/۶
	نرخ جهش	Pm	۰/۵-۰/۳-۰/۱	۰/۵
	اندازه‌ی جمعیت	POP	۱۰۰-۵۰-۲۰	۱۰۰
	تعداد تکرار	It	۳۰۰-۲۰۰-۱۰۰	۳۰۰
NRGA	نرخ تقاطع	Pc	۰/۸-۰/۶-۰/۴	۰/۶
	نرخ جهش	Pm	۰/۵-۰/۳-۰/۱	۰/۳
	اندازه‌ی جمعیت	POP	۱۰۰-۵۰-۲۰	۵۰
	تعداد تکرار	It	۳۰۰-۲۰۰-۱۰۰	۲۰۰

جدول ۲. مشخصات شبکه‌های محک.

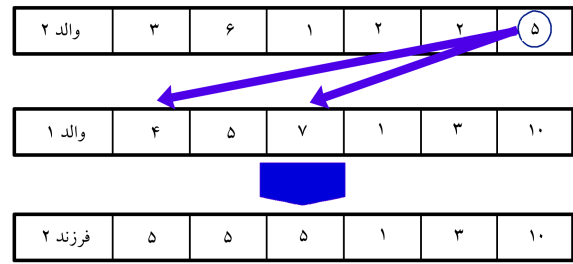
شماره	شبکه	معادل فارسی	گره	یال
۱	Gery Kangaroos	شبکه ارتباط کانگوروها	۱۷	۹۱
۲	Zors	شبکه‌ی ارتباط زورس‌ها	۲۸	۱۱۱
۳	Windsurfers	شبکه‌ی موج‌سواران	۴۳	۳۳۶
۴	Train Bombers	شبکه آموزش بمب‌گذاران	۷۰	۲۴۳
۵	Les Miserables	شخصیت‌های رمان بینوایان	۷۷	۲۵۴
۶	Facebook	شبکه‌ی اجتماعی فیس‌بوک	۴۰۳۹	۸۸۲۳۴

استفاده از روش وزن‌دهی مجموعه جواب پارتو به دست می‌آید که معیار سنجش پارامترهای ورودی است.

مقدار پارامترها پس از انجام آزمایش‌ها، با استفاده از نمودار نسبت سیگنال به نویز<sup>۴۷</sup> در دسترس است. این نمودار برای دو الگوریتم NSGAI و NRGA به ترتیب در اشکال ۱۰ و ۱۱ قابل مشاهده است. مقدار بهینه پارامترهای دو الگوریتم در ستون آخر جدول ۱ گزارش شده است.

## ۲.۵. مقایسه‌ی عملکرد الگوریتم‌ها

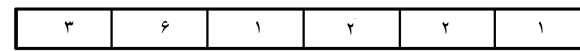
در این بخش، عملکرد دو الگوریتم پیشنهادی P-NSGAI و P-NRGA با الگوریتم‌های NSGAI و NRGA کلاسیک و نیز الگوریتم شناخته شده‌ی MOGA-Net<sup>[۱۲]</sup> که به منظور بهینه‌سازی دو شاخص بودمانگی و امتیازاجتماعات به کارگرفته شده است، مورد سنجش قرار می‌گیرد. مقایسات با حل مسئله‌ی اجتماع‌یابی در ۶ شبکه‌ی دنیای واقعی صورت پذیرفته است. برنامه‌نویسی مسئله در محیط Matlab انجام شده و اجرای برنامه توسط یک کامپیوتر شخصی با مشخصات (Quad Core ۸۲۰۰, CPU ۲٫۰۰ GHz, RAM ۴ GB) صورت گرفته است. مقایسه‌ی عملکرد الگوریتم‌ها در دو بخش انجام می‌شود. در بخش نخست، جواب الگوریتم‌ها از نظر میزان بودمانگی و امتیازاجتماعات مورد مقایسه قرار می‌گیرند. در بخش دوم نیز خروجی الگوریتم‌ها از نظر شاخص‌های مختص روش‌های تکاملی چندهدفه بررسی می‌شود. مشخصات شبکه‌های محک در جدول ۲ آورده شده است.



شکل ۶. نحوه‌ی ایجاد فرزند دوم بر اساس مقدار نزدیکی گره‌ها.



شکل ۷. جواب انتخاب شده برای انجام جهش.



شکل ۸. جواب جهش‌یافته.

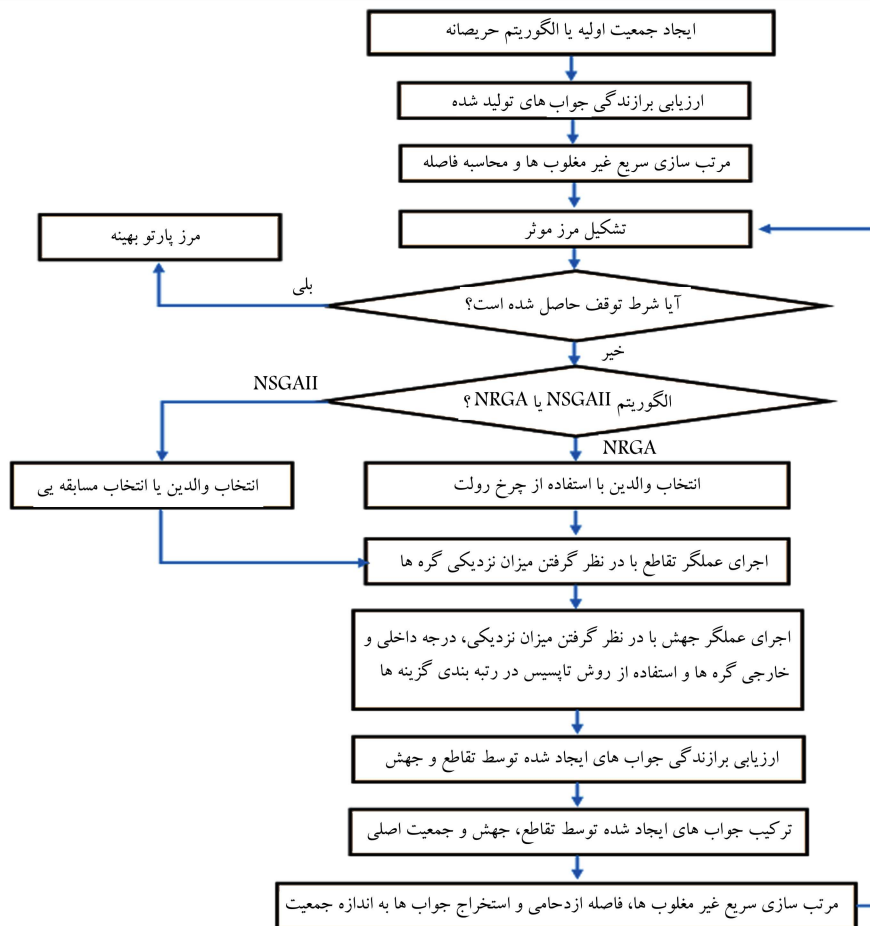
## ۶.۴. جهش

در این بخش، یک روش جهش جدید مبتنی بر تصمیم‌گیری چندمعیاره برای بهبود جواب‌های مسئله‌ی اجتماع‌یابی پیشنهاد می‌شود. در روش جهش پیشنهادی، در ابتدا گرهی به صورت تصادفی از جوابی که برای انجام جهش در نظر گرفته شده است، انتخاب می‌شود. همسایگی‌های گره انتخاب شده بر اساس سه معیار مقدار نزدیکی، درجه‌ی درونی و درجه‌ی بیرونی رئوس با توجه به روش تصمیم‌گیری چندمعیاره‌ی تاپسیس<sup>[۴۵]</sup> رتبه‌بندی می‌شوند. توجه به این نکته ضروری است که دو شاخص میزان نزدیکی و درجه‌ی داخلی گره‌ها از جمله معیارهای مثبت و شاخص درجه‌ی خارجی گره‌ها، از جمله معیارهای منفی است. اوزان معیارها در این پژوهش یکسان در نظر گرفته شده است. پس از انجام روش تاپسیس، گره همسایگی که رتبه‌ی بهتری داشته باشد، جایگزین گره انتخابی خواهد شد. مثلاً در شکل ۷ گره شماره ۵ به تصادف انتخاب شده است. همسایگی‌های این گره طبق شبکه‌ی اصلی برابر رئوس ۱ و ۳ هستند. ماتریس تصمیم‌گیری به ترتیب بر حسب میزان نزدیکی گره‌ها، درجه‌ی داخلی و درجه‌ی خارجی تشکیل می‌شود. سطرها و ستون‌های ماتریس تصمیم‌گیری به ترتیب رئوس همسایگی و معیارها را نشان می‌دهند. در ادامه با توجه به روش تاپسیس رتبه‌بندی گزینه‌ها صورت می‌پذیرد. فرض کنید پس از انجام روش تاپسیس، گره شماره ۱ توانسته است رتبه اول را کسب کند. بنابراین این گره جایگزین گره ۵ خواهد شد. به این ترتیب جواب جهش‌یافته به صورت شکل ۸ تغییر خواهد کرد. ساختار کلی الگوریتم‌های پیشنهادی در شکل ۹ آورده شده است. در این شکل می‌توان به صورت اجمالی رویکرد ارائه شده در این دو الگوریتم را مشاهده کرد.

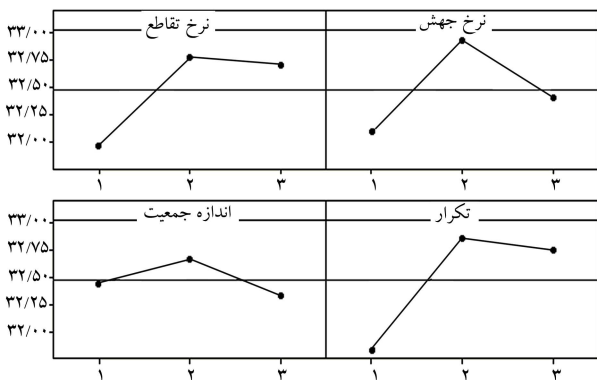
## ۵. تجزیه و تحلیل مسائل نمونه

### ۱.۵. تنظیم پارامتر الگوریتم‌ها

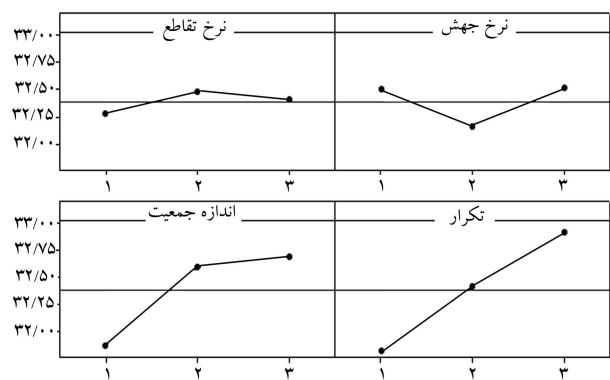
در این مطالعه، به منظور تنظیم پارامترهای الگوریتم‌های پیشنهادی از روش تاگوچی<sup>[۴۶]</sup> استفاده شده است. در این روش سطوح مختلفی برای پارامترها تعیین می‌شود و طراحی آزمایش‌ها با تعریف سطوح متفاوتی برای پارامتر صورت می‌پذیرد. پارامترهای مرتبط با الگوریتم‌های NSGAI و NRGA و نیز مقادیر سطوح مختلف در جدول ۱ آورده شده‌اند. در انتهای هر آزمایش، یک جواب با



شکل ۹. فلوجارت الگوریتم های پیشنهادی.



شکل ۱۱. نمودار نسبت سیگنال به نویز برای پارامترهای الگوریتم ژنتیک با رتبه بندی نامغلوب.



شکل ۱۰. نمودار نسبت سیگنال به نویز برای پارامترهای نسخه دوم الگوریتم ژنتیک با مرتب سازی نامغلوب.

۱.۲.۵. مقایسه ی عملکرد الگوریتم ها از نظر توابع هدف در این قسمت عملکرد الگوریتم ها از نظر مقادیر پودمانگی و امتیاز اجتماعات در شبکه های محک معرفی شده، مورد قیاس قرار گرفته است. نتایج حاصل شده را از نظر شاخص پودمانگی می توان در جدول ۳ مشاهده کرد. در این جدول، بهترین مقادیر پودمانگی  $Q^*$  و میانگین این شاخص پس از بیست مرتبه اجرای برنامه گزارش شده است. همان طور که در جدول ۳ مشخص است، الگوریتم های پیشنهادی از نظر مقدار امتیاز اجتماعات عملکرد بهتری نسبت به سایر روش ها داشته باشند.

۱.۲.۵. مقایسه ی عملکرد الگوریتم ها از نظر توابع هدف در این قسمت عملکرد الگوریتم ها از نظر مقادیر پودمانگی و امتیاز اجتماعات در شبکه های محک معرفی شده، مورد قیاس قرار گرفته است. نتایج حاصل شده را از نظر شاخص پودمانگی می توان در جدول ۳ مشاهده کرد. در این جدول، بهترین مقادیر پودمانگی  $Q^*$  و میانگین این شاخص پس از بیست مرتبه اجرای برنامه گزارش شده است. همان طور که در جدول ۳ مشخص است، الگوریتم های پیشنهادی از نظر مقدار امتیاز اجتماعات عملکرد بهتری نسبت به سایر روش ها داشته باشند.



جدول ۴. عملکرد الگوریتم‌ها از نظر شاخص امتیاز اجتماعات.

میانگین CS	CS*	الگوریتم	شبکه
۸۴,۳۵۵	۹۰,۲۴۳	P_NSGAII	ارتباط گانگورها
۸۰,۸۹۸	۸۷,۲۴۳	P_NRGA	
۷۴,۳۱۱	۸۷,۲۴۳	NSGAII	
۶۷,۵۰۴	۸۷,۲۴۳	NRGA	
۷۵,۱۲۱	۸۷,۲۴۳	MOGA-Net	
۱۱۲,۱۹۱	۱۱۵,۰۸۰	P_NSGAII	ارتباط زورس‌ها
۱۱۲,۵۳۳	۱۱۶,۰۸۰	P_NRGA	
۱۱۰,۶۰۷	۱۱۰,۶۶۰	NSGAII	
۱۱۰,۲۳۵	۱۱۰,۶۶۰	NRGA	
۱۱۰,۸۰۹	۱۱۱,۳۲۶	MOGA-Net	
۱۸۱,۹۵۸	۱۹۴,۶۹۲	P_NSGAII	موج سواران
۱۷۹,۸۲۹	۱۹۲,۰۲۱	P_NRGA	
۱۷۰,۷۲۲	۱۷۵,۰۵۱	NSGAII	
۱۷۲,۸۰۴	۱۷۴,۲۱۵	NRGA	
۱۷۵,۵۵۴	۱۷۷,۲۳۹	MOGA-Net	
۱۰۰,۴۴۵	۱۰۶,۷۷۲	P_NSGAII	آموزش بمب گذاران
۹۹,۹۰۹	۱۰۹,۱۰۳	P_NRGA	
۷۹,۴۷۹	۸۵,۲۷۸	NSGAII	
۸۴,۰۳۷	۸۷,۶۳۰	NRGA	
۱۰۰,۳۳۶	۱۰۶,۷۷۲	MOGA-Net	
۲۲۷,۷۵۱	۲۳۱,۱۲۴	P_NSGAII	شخصیت‌های رمان بینوایان
۲۲۳,۶۱۴	۲۲۵,۸۶۹	P_NRGA	
۱۹۰,۳۴۵	۱۹۵,۲۷۸	NSGAII	
۱۸۸,۴۶۵	۱۸۹,۰۰۶	NRGA	
۲۲۴,۱۱۵	۲۳۱,۱۲۴	MOGA-Net	
۴۲۱,۸۵۷	۴۳۵,۱۲۶	P_NSGAII	فیسبوک
۴۳۴,۹۷۴	۴۴۱,۲۳۸	P_NRGA	
۳۵۰,۶۲۱	۳۵۱,۷۶۹	NSGAII	
۳۳۹,۴۵۹	۳۴۶,۲۳۸	NRGA	
۳۸۶,۴۳۸	۳۹۱,۲۵۸	MOGA-Net	

جدول ۳. عملکرد الگوریتم‌ها از نظر شاخص پودمانگی.

میانگین Q	Q*	الگوریتم	شبکه
۰,۱۲۹۹	۰,۱۵۸۹	P_NSGAII	ارتباط گانگورها
۰,۱۴۱۳	۰,۱۵۸۹	P_NRGA	
۰,۰۴۱۳	۰,۰۵۸۹	NSGAII	
۰,۳۴۲	۰,۰۵۵۵	NRGA	
۰,۰۵۶۸	۰,۰۵۸۹	MOGA-Net	
۰,۳۵۷۲	۰,۳۷۲۲	P_NSGAII	ارتباط زورس‌ها
۰,۴۰۸۶	۰,۴۲۲۲	P_NRGA	
۰,۲۵۵۰	۰,۲۶۴۵	NSGAII	
۰,۲۵۵۷	۰,۲۷۲۲	NRGA	
۰,۲۷۵۲	۰,۲۸۳۲	MOGA-Net	
۰,۴۱۰۷	۰,۴۱۵۳	P_NSGAII	موج سواران
۰,۴۱۴۱	۰,۴۱۵۳	P_NRGA	
۰,۲۵۳۷	۰,۲۵۴۹	NSGAII	
۰,۲۴۵۶	۰,۲۶۵۳	NRGA	
۰,۲۵۵۲	۰,۲۶۵۳	MOGA-Net	
۰,۵۵۴۴	۰,۵۸۷۰	P_NSGAII	آموزش بمب گذاران
۰,۵۵۳۵	۰,۵۸۷۷	P_NRGA	
۰,۴۲۴۳	۰,۴۲۹۶	NSGAII	
۰,۴۲۸۴	۰,۴۳۶۱	NRGA	
۰,۴۹۵۱	۰,۵۰۲۴	MOGA-Net	
۰,۴۶۸۲	۰,۴۷۶۸	P_NSGAII	شخصیت‌های رمان بینوایان
۰,۴۵۲۰	۰,۴۷۵۳	P_NRGA	
۰,۳۳۰۲	۰,۳۴۵۸	NSGAII	
۰,۳۵۶۹	۰,۳۷۰۶	NRGA	
۰,۳۹۲۸	۰,۴۱۰۶	MOGA-Net	
۰,۲۹۵۸	۰,۳۰۱۷	P_NSGAII	فیسبوک
۰,۳۰۲۶	۰,۳۱۲۴	P_NRGA	
۰,۲۴۱۳	۰,۲۵۲۹	NSGAII	
۰,۲۴۱۸	۰,۲۶۱۸	NRGA	
۰,۲۵۴۸	۰,۲۶۳۵	MOGA-Net	

الگوریتم P\_NSGAII و P\_NRGA توانسته‌اند در تمام ۶ شبکه‌ی نمونه، سایر روش‌ها را مغلوب سازند.

یکی از شاخص‌هایی که میزان دقت الگوریتم‌ها را در شناسایی اجتماعات باکیفیت و درست از یک شبکه تعیین می‌کند، معیار اطلاعات متقابل نرمال<sup>۴۸</sup> است. [۱۲] NMI یک شاخص شناخته شده در حوزه‌ی نظریه‌ی اطلاعات است که میزان شباهت خوشه‌های شناسایی شده را در مقایسه با خوشه‌های حقیقی تخمین می‌زند. با فرض آنکه دو افراز A و B، افرازهایی از شبکه باشند، ماتریس C را ماتریس درهم ریختگی می‌نامیم که  $C_{ij}$  در آن برابر تعداد گره‌هایی از اجتماع i در افراز A است که در اجتماع j از افراز B نیز حضور دارند. به این ترتیب، شاخص اطلاعات متقابل نرمال  $I(A, B)$  به صورت زیر تعریف می‌شود:

$$I(A, B) = \frac{-2 \sum_{i=1}^{CA} \sum_{j=1}^{CB} C_{ij} \log(C_{ij}N / C_i \cdot C_j)}{\sum_{i=1}^{CA} C_i \cdot \log(C_i / N) + \sum_{j=1}^{CB} C_j \cdot \log(C_j / N)} \quad (11)$$

در رابطه‌ی ۱۱،  $C_i$  (C.j) برابر تعداد گروه‌های افراز A(B) است.  $C_i$  (C.j) برابر مجموع درایه‌های ماتریس C در سطر i (ستون j) است. N برابر تعداد گره‌هاست. اگر  $A = B$  باشد، آنگاه  $I(A, B) = 1$ . اگر دو افراز A و B کاملاً متفاوت باشند، آنگاه  $I(A, B) = 0$ . میانگین نتایج حاصل از محاسبه‌ی شاخص NMI برای ۵ الگوریتم پس از بیست مرتبه اجرای برنامه‌ها در جدول ۷ گزارش شده است. همان‌طور که در جدول ۷ مشخص است، دو الگوریتم پیشنهادی P\_NSGAII و P\_NRGA توانسته‌اند در ۶ شبکه عملکرد به مراتب بهتری نسبت به سه الگوریتم دیگر داشته باشند.

به منظور سنجش این مسئله که سطح کیفی جواب‌های الگوریتم‌های پیشنهادی از نظر معیارهای GD, IGD, S, NOS, T و NMI اختلاف معناداری نسبت به جواب‌های سایر الگوریتم‌ها دارند، از آزمون آماری t استفاده می‌شود. بدین منظور عملکرد الگوریتم‌های P\_NSGAII و P\_NRGA در فاصله‌ی اطمینان ۹۵ درصد به صورت دو به دو با الگوریتم‌های کلاسیک NSGAII و NRG A و الگوریتم MOGA-Net مورد آزمون قرار می‌گیرد. فرض صفر در این آزمون نشان دهنده‌ی عدم وجود تفاوت معنادار میان جواب‌های دو الگوریتم مورد مقایسه است. نتایج آزمون‌های آماری نشان می‌دهد که عملکرد الگوریتم‌های پیشنهادی از نظر معیارهای GD, S, NOS, T و NMI تفاوت معناداری در مقایسه با الگوریتم‌های کلاسیک NSGAII و NRG A دارند. زیرا مقدار p-value برای اکثر آزمون‌ها کمتر از ۰۰۵ است. بنابراین می‌توان اطمینان حاصل کرد که الگوریتم‌های پیشنهادی در کل عملکرد بهتری در یافتن جواب‌های با کیفیت داشته‌اند.

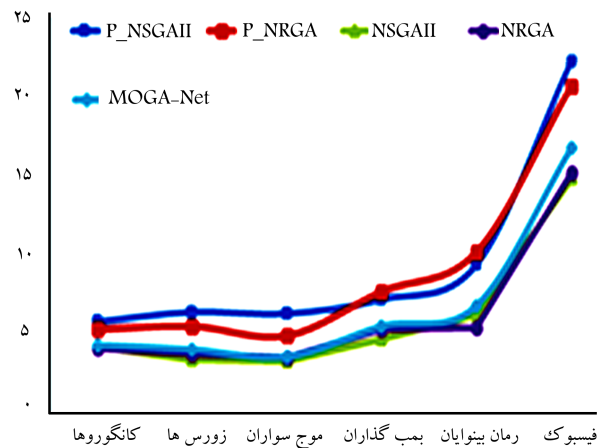
## ۶. نتیجه‌گیری و پیشنهادها

در این مقاله مسئله‌ی اجتماع‌یابی در شبکه‌های دنیای واقعی به تفصیل تشریح شد. به جنبه‌های کاربردی که مدل اجتماع‌یابی می‌تواند در آن‌ها مفید واقع شود نیز اشاره شد. با توجه به آنکه روش‌های سنتی حل مسئله‌ی اجتماع‌یابی، شکل‌دهی اجتماعات را تنها بر مبنای یک معیار صورت می‌دهند ارائه‌ی رویکردی که با بهینه‌سازی هم‌زمان چندشاخص بتواند مجموعه‌ی اجتماعات در شبکه را بیابد، راهگشا خواهد بود. بدین منظور در این پژوهش، دو روش فراابتکاری جدید که توسعه‌ی بی‌روش‌های

۲.۲.۵. مقایسه‌ی عملکرد الگوریتم‌ها از نظر معیارهای چندهدفه در این بخش، عملکرد روش‌ها بر اساس معیارهای سنجش کارایی الگوریتم‌های چندهدفه، مقایسه می‌شود. این معیارها عبارت‌اند از شاخص فاصله‌گذاری (S) [۲۷]، فاصله‌ی عمومی (GD) [۲۸]، فاصله‌ی عمومی معکوس (IGD) [۲۹]، تعداد جواب‌های به دست آمده توسط هر الگوریتم (NOS) [۳۰]، زمان محاسبات T [۵۰]، و پوشش مجموعه [۵۰] هرچه مقدار شاخص‌های S, GD, IGD و T کمتر باشد، عملکرد الگوریتم حل بهتر خواهد بود و هرچه تعداد جواب‌های غیرمغلوب به دست آمده بیشتر باشد، عملکرد الگوریتم حل بهتر است. مقایسه‌ی کارایی الگوریتم‌های پیشنهادی از نظر معیارهای S, GD و IGD در جدول ۵ آمده است. نتایج بعد از بیست مرتبه اجرای برنامه و میانگین‌گیری حاصل شده‌اند. همان‌طور که در این جدول مشهود است، دو الگوریتم P\_NSGAII و P\_NRGA از نظر معیارهای فاصله‌گذاری، فاصله‌ی عمومی و فاصله‌ی عمومی معکوس نسبت به سه الگوریتم دیگر به مراتب عملکرد بهتری داشته‌اند.

زمان محاسبات از دیگر معیارهای تعیین‌کننده در ارزیابی عملکرد روش‌های حل مسئله است. در این نوشتار، فرض شده است که بعد از همگرایی الگوریتم‌ها به مجموعه‌ی از جواب‌های غیرمغلوب و عدم تغییر ترکیب جواب‌های مرز مؤثر پس از سپری شدن ۵۰ تکرار متوالی، اجرای الگوریتم‌ها متوقف خواهد شد. میانگین زمان اجرای الگوریتم‌ها برحسب ثانیه بعد از بیست مرتبه اجرای برنامه برای شناسایی اجتماعات این ۶ شبکه در جدول ۵ آمده است. همان‌طور که در این جدول مشخص است، دو الگوریتم پیشنهادی P\_NSGAII و P\_NRGA توانسته‌اند زودتر از سه الگوریتم دیگر به همگرایی برسند.

یکی دیگر از معیارهای معرفی شده، تعداد جواب‌های مرز مؤثر است. با توجه به این شاخص، هرچه تعداد جواب‌های پیدا شده توسط الگوریتم بیشتر باشد، روش حل موفق‌تر عمل کرده است. در شکل ۱۲ میانگین تعداد جواب‌های کشف شده پس از بیست مرتبه اجرای برنامه قابل مشاهده است. با توجه به آنچه در شکل ۱۲ مشهود است، دو الگوریتم P\_NSGAII و P\_NRGA موفق به کشف تعداد جواب‌های بیشتری نسبت به سه روش حل NSGAII, NRG A و MOGA-Net شده‌اند. معیار دیگری که معرفی شد، معیار پوشش مجموعه است که مشخص می‌کند جواب‌های هر الگوریتم چقدر در مغلوب‌سازی جواب‌های الگوریتم دیگر موفق هستند. جدول ۶ مقایسات دو به دو میان روش‌های پیشنهادی و سه الگوریتم دیگر را در ۶ شبکه نشان می‌دهد. با توجه به آنچه در جدول ۶ مشهود است، دو



شکل ۱۲. میانگین تعداد جواب‌های کشف شده توسط الگوریتم‌ها.

جدول ۵. مقایسه‌ی الگوریتم‌ها از نظر معیارهای فاصله‌گذاری، فاصله‌ی عمومی، فاصله‌ی عمومی معکوس.

زمان محاسبات (s)	S	IGD	GD	الگوریتم	شبکه
۱,۵۲	۱,۱۶۰	۴۳,۲۲۶	۰,۳۷۳	P_NSGAII	ارتباط گانگورها
۱,۲۲	۱,۰۶۶	۳۴,۲۰۷	۰,۲۹۶	P_NRGA	
۱,۹۵	۱۸,۱۹	۱۰۶,۰۷	۱,۶۱۴	NSGAII	
۱,۹۸	۱۵,۲۳	۹۵,۶۱۹	۱,۳۹۳	NRGA	
۱,۹۱	۶,۹۹۸	۶۸,۹۲۵	۱,۲۲۰	MOGA-Net	
۲,۵	۰,۰۰۱	۳,۹۵۲۱	۰,۲۴۶	P_NSGAII	ارتباط زورس‌ها
۲,۸۲	۰,۰۰۳	۳,۴۲۱	۰,۳۰۸	P_NRGA	
۵,۲۲	۴,۸۷۰	۱۶,۰۶۵	۰,۸۸۱۸	NSGAII	
۵,۳۴	۳,۸۹۰	۱۹,۳۴۵	۰,۹۹۱۹	NRGA	
۴,۷۱	۰,۰۱۲۹	۱۴,۴۷۹	۰,۶۷۸۵	MOGA-Net	
۴,۱۴	۰,۰۴۹	۳,۷۱۶۶	۰,۵۳۹۸	P_NSGAII	موج سواران
۳,۹۹	۰,۰۱۷	۳,۲۶۱۳	۰,۵۳۹۸	P_NRGA	
۷,۱	۷,۳۵۷	۱۰,۷۷۲	۱,۴۵۳۸	NSGAII	
۸,۰۱	۱۰,۳۲	۱۱,۶۹۱	۱,۴۹۹۴	NRGA	
۶,۷۸	۰,۰۱۷	۳,۵۵۴۲	۰,۵۳۹۸	MOGA-Net	
۸,۰۲	۰,۰۱۹	۶,۹۱۸۰	۰,۲۶۴۸	P_NSGAII	آموزش بمسب گذاران
۸,۱۱	۰,۰۲۸	۷,۰۸۵۱	۰,۱۳۷۷	P_NRGA	
۱۰,۳۵	۶,۲۳۹	۲۴,۸۷۱	۰,۶۶۷۰	NSGAII	
۱۰,۶۱	۶,۴۱۲	۲۱,۴۹۰	۰,۷۶۹۳	NRGA	
۱۱,۳۵	۶,۰۲۶	۲۴,۱۶۹	۰,۶۶۵۷	MOGA-Net	
۹,۰۱	۰,۰۰۳	۶,۰۲۱	۰,۴۱۲	P_NSGAII	شخصیت‌های رمان بیتویان
۱۰,۴۱	۰,۰۰۲	۷,۲۹۵	۰,۳۷۲	P_NRGA	
۱۷,۹۴	۸,۳۶۹	۱۵,۰۲۴	۲,۸۷۴۳	NSGAII	
۱۷,۰۲	۵,۶۸۷	۱۶,۳۲۹	۲,۰۶۲۵	NRGA	
۱۷,۳۳	۳,۰۰۷	۱۲,۳۸۶	۱,۶۹۸۱	MOGA-Net	
۲۳۶,۱۴	۰,۰۲۴۴	۹۶,۶۴۲	۶,۰۱۵۷	P_NSGAII	فیسپوک
۲۳۴,۷۷	۰,۰۱۳۲	۸۳,۳۵۴	۵,۸۸۴۱	P_NRGA	
	۰,۰۴۱۸	۱۰۵,۶۳۱	۱۰,۱۹۹	NSGAII	
۲۲۶,۶۹	۰,۰۴۷۳	۱۰۱,۴۷۱	۱۱,۴۷۴	NRGA	
	۰,۰۳۲۶	۱۰۱,۲۰۵	۷,۷۰۲	MOGA-Net	

جدول ۶. مقایسه‌ی چهار الگوریتم از نظر معیار پوشش مجموعه.

مقایسه الگوریتم‌ها												
C(MOGA - Net, P_NRGA, )	C(P_NRGA, MOGA - Net)	C(MOGA - Net, P_NSGAII, )	C(P_NSGAII, MOGA - Net)	C(NRGA, P_NRGA)	C(P_NRGA, NRGA)	C(NSGAII, P_NRGA)	C(P_NRGA, NSGAII)	C(NRGAII, P_NSGAII)	C(P_NSAII, NRGA)	C(NSGAII, P_NSGAII)	C(P_NSGAII, NSGAII)	شبکه
۰	۰٫۷	۰	۰٫۹	۰	۰٫۸	۰	۰٫۹	۰	۰٫۸	۰	۰٫۹	ارتباط کانگوروا
۰	۰٫۶	۰	۰٫۹	۰	۰٫۸	۰	۰٫۷	۰	۰٫۸	۰	۰٫۸	زورس‌ها
۰	۰٫۶	۰	۱	۰	۱	۰	۱	۰	۱	۰	۱	موج سواران
۰	۰٫۹	۰	۰٫۷	۰	۰٫۸	۰	۰٫۹	۰	۰٫۹	۰	۰٫۹	آموزش بمب گذاران
۰	۰٫۸	۰	۰٫۸	۰	۰٫۸	۰	۰٫۹	۰	۱	۰	۱	رمان بینوایان
۰	۰٫۷	۰	۰٫۸	۰	۰٫۷	۰	۰٫۹	۰	۰٫۸	۰	۰٫۸	فیسپوک

جدول ۷. شاخص NMI الگوریتم‌ها در ۶ شبکه.

الگوریتم‌ها					
MOGA-Net	NRGA	NSGAII	P_NRGA	P_NSGAII	شبکه
۰٫۴۶۲	۰٫۲۰۴	۰٫۱۹۵	۰٫۸۵۸	۰٫۸۸۴	کانگوروا
۰٫۴۷۷	۰٫۳۶۲	۰٫۵۵۲	۰٫۸۰۱	۰٫۸۳۶	زورس‌ها
۰٫۵۷۷	۰٫۴۳۵	۰٫۴۲۹	۰٫۹۰۸	۰٫۹۱۱	موج سواران
۰٫۶۰۳	۰٫۴۰۲	۰٫۴۱۶	۰٫۹۷۸	۰٫۹۷۱	بمب گذاران
۰٫۳۰۱	۰٫۲۸۸	۰٫۳۳۱	۰٫۶۸۸	۰٫۷۵۵	بینوایان
۰٫۲۰۵	۰٫۱۸۲	۰٫۱۹۷	۰٫۸۵۰	۰٫۸۷۱	فیسپوک

در گام بعد، عملکرد الگوریتم‌های پیشنهادی با سه روش NSGAII و NRGA کلاسیک و نیز MOGA-Net، که از قدرتمندترین الگوریتم‌های پیشنهاد شده در پژوهش‌های پیشین است، در ۶ شبکه‌ی دنیای واقعی مقایسه شد. مقایسات با توجه به مقدار توابع هدف و شاخص‌های شناخته شده در سنجش کارایی الگوریتم‌های چندهدفه انجام شد. نتایج حاصل، نشان از برتری الگوریتم‌های ارائه شده بر سه الگوریتم دیگر داشت.

از جمله موارد کاربرد الگوریتم‌های پیشنهادی می‌توان به کشف اجتماعاتی دقیق از مشتریان شبکه‌های برخط فروش که علائق شبیه به یکدیگر دارند، اشاره کرد. در این صورت با پیشنهاد محصولاتی متناسب با علائق مشتریان، فرصت‌های تجاری مناسبی پدید خواهد آمد. از دیگر موارد کاربرد روش‌های ارائه شده، شکل‌دهی اجتماعات علمی در شبکه‌های استنادی تحقیقات است. این شکل‌دهی برپایه‌ی

کلاسیک NSGAII و NRGA می‌باشند، ارائه شد. این دو روش فضای جواب را با هدف یافتن یک افراز بهینه از شبکه و بیشینه‌سازی پودمانگی و امتیاز اجتماعات جستجو می‌کنند. پودمانگی و امتیاز اجتماعات از جمله توابع هدف معتبر در مسئله‌ی اجتماع‌یابی هستند که مقدار این دو شاخص به عنوان ملاک برابری جواب‌ها در نظر گرفته شد. به منظور بهبود کیفیت پاسخ‌های مسئله، جواب‌های اولیه‌ی این دو الگوریتم توسط یک روش حریرانه به دست آمد. دو الگوریتم پیشنهادی از یک روش جدید تقاطع و جهش استفاده می‌کنند. رویکرد جدید در عملگر تقاطع، مبتنی بر مقدار نزدیکی گره‌هاست. در روش جهش پیشنهادی نیز سه معیار میزان نزدیکی، درجه‌ی داخلی و درجه‌ی خارجی گره‌ها ملاک قرار می‌گیرند و ایجاد جواب جدید با استفاده از روش تصمیم‌گیری چندمعیاره‌ی تاپسیس صورت می‌پذیرد. توجیه استفاده از این معیارها، تشکیل اجتماعات متمایز و بهینه‌سازی توابع هدف مسئله است.

اجتماعی نظیر چگالی پودمانگی، وسعت و رسانایی شبکه را نیز در روش‌های معرفی شده گنجانند و به این ترتیب دقت روش‌های پیشنهادی را بالا برد. از سوی دیگر، به کارگیری روش‌های پیشنهادی در شناسایی اجتماعات شبکه‌های پویا از دیگر موضوعاتی است که می‌تواند مبنای مطالعات آتی قرار گیرد. پیشنهاد مشابه دیگر، بررسی ارتباط سایر خصوصیات گره‌های شبکه در تشکیل اجتماعات است که موضوعی قابل تأمل برای پژوهش‌های آینده است.

موضوعات مورد مطالعه صورت می‌پذیرد. همچنین خروجی این دو الگوریتم می‌تواند در شناسایی دقیق فعالیت‌های کلاه‌برداری، تقلب و تروریستی مورد استفاده قرار گیرد. بنابراین، روش‌های معرفی شده قادر هستند به عنوان ابزاری در کشف ساختارهای اجتماعی در بسیاری از شبکه‌های دنیای واقعی به کار بروند. از آنجایی که در الگوریتم‌های پیشنهادی از روش تصمیم‌گیری چندمعیاره‌ای تاپسیس استفاده شده است، می‌توان سایر معیارهای معمول در شناسایی ساختارهای

## پانوشته‌ها

1. community structure
2. community detection problem
3. cluster
4. evolutionary algorithms (EO)
5. modularity
6. community score (CS)
7. proposed non-dominated sorting genetic algorithm
8. proposed non-dominated ranked genetic algorithm
9. crossover
10. mutation
11. closeness centrality
12. multi-criteria decision making(MCDM)
13. Girvan-Newman (GN)
14. greedy algorithms
15. simulated annealing
16. extremal optimization
17. local
18. global
19. information theory
20. biological fitness
21. matrix representation
22. multi-agent genetic algorithm
23. split
24. merging based neighborhood competition operator
25. hybrid neighborhood crossover
26. adaptive mutation
27. self-learning operator
28. locus-based solution representation
29. modularity density
30. community fitness
31. graph theory
32. expansion
33. conductance
34. description length
35. non-dominated solutions
36. pareto frontier
37. clustering
38. parallel computing
39. genotype
40. backtracking
41. feasible
42. strength
43. closeness centrality
44. floyd-warshall algorithm
45. Topsis
46. Taguchi Method
47. signal to noise ratio (S/N)
48. normalized mutual information (NMI)

## منابع (References)

1. Newman, M.E.J. "Networks: An introduction", Published to Oxford Scholarship Online, 1st Edn., Oxford, UK (2010), DOI: 10.1093/acprof:oso/9780199206650.001.0001.
2. Fortunato, S. "Community detection in graphs", *Physics Reports*, **486**(3), pp. 1-100, (2010).
3. Girvan, M. and Newman, M.E.J. "Community structure in social and biological networks", *Proceedings of the national academy of sciences of the United States of America*, **99**(12), pp. 7821-7826, (2002).
4. Zhang, H., Qiut, B., Giles, L., Foley, H. and Yen, J. "An LDA-based community structure discovery", *Intelligence and Security Informatics*, **400**(2), pp. 200-207 (2007).
5. Pan, G., Zhang, W., Wu, Z. and Li, S. "Online community detection for large complex networks", *Plos One*, **9**(7), pp. 168-188 (2014).
6. Choudhury, D. and Paul, A. "Community detection in social networks: an overview", *International Journal of Engineering and Technology*, **2**(2), pp.6-13 (2013).
7. Agarwal, G. and Kempe, D., "Modularity-maximizing graph communities via mathematical programming", *The European Physical Journal B*, **66**(3), pp. 409-418 (2008).
8. Ghorbanian, A. and Shafaqi, B. "A Genetic Algorithm for Modularity Density Optimization in Community Detection", *International Journal of Economy, Management and Social Sciences*, **4**(1), pp. 117-122 (2015).
9. Brandes, U., Delling, D., Gaetler, M. "On Modularity clustering", *Transactions on Knowledge and Data Engineering*, **20**(2), pp.172-188 (2008).
10. Shi, C., Yan, Z., Cai, Y. and Wu, B. "Multi objective community detection in complex networks", *Applied Soft Computing Journal*, **12**(2), pp. 850-859 (2012).
11. Fortunato, S. and Barthelemy, M. "Resolution limit in community detection", *PNAS*, **104**(1), pp. 36-41. (2007).
12. Pizzuti, C. "GA-Net: A Genetic algorithm for community detection in Social Networks", *Computer Science*, **5199**(1), pp. 1081-1090 (2008).
13. Shafaqi, B. and Teimourpour, B. "A new heuristic algorithm for modularity optimization in complex networks

- community detection”, *11th international industrial engineering conference*. 7-8 January, (2015).
14. Newman, M. and Girvan, M. “Finding and evaluating community structure in networks”, *Physical Review E*, **69**(2) (2004).
  15. Griechisch, E. and Pluhar, A. “Community detection by using the extended modularity”, *Acta Cybernetica*, **20**(1), pp. 69-85 (2011).
  16. Guimera, R. and Nunes Amaral, L. “Functional cartography of complex metabolic networks”, *Nature*, **433**(7028), pp. 895-900 (2005).
  17. Hafez, A., Ghali, N., Hassaniien, A. and Fahmy, A. “Genetic Algorithms for community detection in social networks”, *Intelligent Systems Design and Applications*, **10**(2), pp. 460-465 (2012).
  18. Meghanathan, N. “A Greedy algorithm for neighborhood overlap-based community detection”, *Algorithms*, **9**(1), pp. 1-26 (2016).
  19. Blondel, V.D., Guillaume, J.L. and Lambiotte, R. Lefebvre, E. “Fast unfolding of communities in large networks”, *Journal of Statistical Mechanics: Theory and Experiment*, **8**(3), pp. 12-24 (2008).
  20. Massen, C.P. and Doye, J., “Identifying communities within energy landscapes”, *Physical Review E*, **71**(4), pp.40-58 (2005).
  21. Boettcher, S. and Percus, A.G., “Optimization with extremal dynamics”, *Phys Rev Lett*, **86**(1), pp. 5211-5214 (2001).
  22. Duch, J. and Arenas, A., “Community detection in complex networks using extremal optimization”, *Physical Review E*, **72**(2), pp. 4-8 (2005).
  23. Rosvall, M. and Bergstrom, C. “An information-theoretic framework for resolving community structure in complex networks”, *PNAS*, **104**(18), pp.7327-7331 (2007).
  24. Palla, G., Derney, I. and Farkas, I. “Uncovering the overlapping community structure of complex networks in nature and society”, *Nature*, **435**, pp. 814-818 (2005).
  25. Kumpala, J., Kirea, M. and Kaski, K. “Sequential algorithm for fast clique percolation”, *Physic Review E*, **78**(3), pp. 20-38 (2008).
  26. Cafieri, S., Hansen, P. and Liberti, L. “Loops and multiple edges in modularity maximization of networks”, *Physical Review E*, **81**(4), pp. 1-21 (2010).
  27. Holland, J.H., “An Introductory analysis with applications to biology, control and artificial intelligence”, *In Adaptation in Natural and Artificial Systems*, 1st Edn., MIT Press Cambridge, MA, USA (1992).
  28. Bingol, H. and Tasgin, M. “Community detection in complex networks using genetic algorithms”, *Advances in Complex Systems*, **11**(4), pp.1-6 (2006).
  29. Tasgin, M. and Bingol, H. “Gossip on weighted networks”, *Advances in Complex Systems*, **15**(1), pp. 7-15 (2012).
  30. Amiri, B., Hossain, L., Crawford, J. and Wigand, T. “Community detection in complex networks: multi-objective enhanced fire fly algorithm and genetic algorithm”, *Knowledge based systems*, **46**, pp. 1-11 (2013).
  31. Yun, L., Gang, L. and Yang, L., “A genetic algorithm for community detection in complex networks”, *Journal of Central South University*, **20**(5), 1269-1276, (2013).
  32. Li, Z. and Liu, J., “A multi-agent genetic algorithm for community detection in complex networks”, *Physica A*, **449**(1), pp. 336-347 (2016).
  33. Shi, C., Yan, Z., Wang, Y., Cai, Y. and Wu, B. “A Genetic algorithm for detecting communities in largescale complex networks”, *Advance in Complex System*, **13**(1), pp. 3-17 (2010).
  34. Shang, R., Bai, J., Jiao, L. and Jin, C., “Community detection based on modularity and an improved genetic algorithm”, *Physica A*, **392**, pp. 1215-1231 (2013).
  35. Chen, M., Kuzmin, K., Boleslaw, K. and Szymanski, F. “Community detection via maximization of modularity and its variants”, *Trans. Computation Social System*, **1**(1), pp. 46-65 (2014).
  36. Mazur, P., Zmarzowski, K. and Orłowski, A.J., “Genetic algorithms approach to community detection”, *Acta Physica Polonica A*, **117**(4), pp. 703-706 (2010).
  37. Kannan, R. and Vempala, S. “On clusterings: good, bad and Spectral”, *Journal of the ACM*, **51**(3), pp. 497-515 (2004).
  38. Leskovec, J., Lang, K., Mahoney, M. “Empirical comparison of algorithms for network community detection”, *ACM*, **20**(16), pp. 631-640 (2010).
  39. Liu, R., Feng, S., Shi, R. and Guo, W., “Weighted graph clustering for community detection of large social networks”, *Procedia Computer Science*, **31**(3), pp. 85-94 (2014).
  40. Agrawal, R. “Bi-objective community detection (BOCD) in networks using genetic algorithm”, In *Contemporary Computing*, pp. 5-15, Springer, Berlin, Heidelberg, (2011). DOI: [https://doi.org/10.1007/978-3-642-22606-9\\_5](https://doi.org/10.1007/978-3-642-22606-9_5).
  41. Zhang, W., Pan, G., Wu, Z., Li, S. “Online Community Detection for Large Complex Networks”, *Plos One*, **9**(7), pp. 168-188 (2014).
  42. Radicci, F., Castellano, C., Cecconi, F., Loreto, V. and Parisi, D. “Defining and identifying communities in networks”, *Proceedings of the National Academy of Science of the USA (PNAS O4)*, **101**(9), pp. 2658-2663 (2004).
  43. Bavelas, A. “Communication patterns in task-oriented groups”, *J. Acoust. Soc. Am*, **22**(6), pp. 725-730 (1950).
  44. Sabidussi, G. “The centrality index of a graph”, *Psychometrika*, **31**, pp. 581-603 (1966).
  45. Yue, Z. “Extension of TOPSIS to determine weight of decision maker for group decision making problems with uncertain information”, *Expert Systems with Applications*, **39**(7), pp. 6343-6350 (2012).
  46. Golphalsamy, B., Mondal, B., Ghosh, S. “Taguchi method and ANOVA: an approach for process parameters optimization of hard machining while machining hardened steel”, *Journal of Scientific & Industrial Research*, **68**, pp. 686-695 (2009).
  47. Jiang, S., Ong, Y., Feng, L. “Consistencies and contradictions of performance metrics in multiobjective optimization”, *Cybernetics*, **44**(12), pp. 2391-2404 (2014).

48. Veldhuizen, D.A. "Multiobjective evolutionary algorithms: classifications, analyses, and new innovations, Ph.D. dissertation", Air Force Institute of Technology, Wright-Patterson AFB, Ohio (1999).
49. Czyzzak, P., Jaskiewicz, A. "Pareto simulated annealing- a metaheuristic technique for multiple objective combinatorial optimization", *Journal of Multi-Criteria Decision Analysis*, **7**, pp. 34-47 (1998).
50. Zitzler, E., and Thiele, L. "Multi-objective evolutionary algorithms: a comparative case study and the strength Pareto approach", *IEEE Transactions on Evolutionary Computation*, **3**(4), pp. 257-271, (1999).