

توسعه یک روش فرا ابتکاری ترکیبی برای شناسایی اجتماعات در شبکه‌های اجتماعی با هدف چگالی پودمانگی

امیرحسین حسینیان*
بابک تیمورپور**
باقر جمالی هندری***

چکیده

شناسایی ساختارهای موجود در شبکه‌های اجتماعی، فرآیندی حائز اهمیت در تحلیل این شبکه‌ها است. یکی از مسائلی که در سال‌های اخیر در زمینه شناسایی ساختارهای شبکه‌های اجتماعی مطرح شده است، مسئله اجتماع یابی است. با توجه به اهمیت این موضوع، تاکنون روش‌های حل متنوع و گوناگونی برای اجتماع یابی پیشنهاد شده است. در پژوهش حاضر، الگوریتمی ترکیبی از روش‌های بهینه‌سازی علف‌های هرز و ژنتیک پیشنهاد شده است که هدف آن یافتن جواب‌های مناسب و باکیفیت برای مسئله اجتماع یابی است. در این روش ترکیبی، جواب‌های اولیه توسط روش بهینه‌سازی علف‌های هرز تولید می‌شوند و در ادامه جواب‌های یافته شده به وسیله الگوریتم ژنتیک در فرآیند بهینه‌سازی، بهبود می‌یابند. ارزیابی برازندگی جواب‌ها، مبتنی بر معیار چگالی پودمانگی است. چگالی پودمانگی، معیاری با ماهیت بیشینه‌سازی است که میزان کیفیت اجتماعات کشف شده را مشخص می‌کند. به منظور بررسی کیفیت جواب‌های الگوریتم پیشنهادی، نتایج این روش نسبت به چهار الگوریتم علف‌های هرز، ژنتیک، الگوریتم کرم شب‌تاب و یک الگوریتم جستجوی کاملاً تصادفی مقایسه شده‌اند. پارامترهای این الگوریتم‌ها به کمک رویکرد طراحی آزمایش‌ها تنظیم شده‌اند. این مقایسات بر روی شبکه‌های محک گوناگون و با ابعاد متفاوت انجام شده‌اند. با توجه به نتایج به دست آمده، می‌توان دریافت که الگوریتم پیشنهادی قادر به تولید جواب‌هایی باکیفیت بالا است. اعتبارسنجی نتایج الگوریتم‌ها نیز توسط شاخص اطلاعات متقابل نرمال انجام شده است.

کلیدواژه‌گان: اجتماع یابی، محاسبات نرم، بهینه‌سازی علف‌های هرز، چگالی پودمانگی.

* دانش‌آموخته دکتری، مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه آزاد اسلامی واحد تهران شمال، تهران.

** عضو هیئت علمی، گروه مهندسی صنایع و سیستم‌ها، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، تهران. (نویسنده

مسئول): b.teimourpour@modares.ac.ir

*** دانشجوی دکتری، مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه آزاد اسلامی واحد تهران شمال، تهران.

مقدمه

سیستم‌های طبیعی، علمی و اجتماعی را می‌توان در قالب شبکه‌های پیچیده مدل‌سازی کرد؛ بنابراین، به کمک نظریه شبکه‌ها امکان مطالعه این سیستم‌ها به وجود می‌آید (قربانیان و شقاقی^۱، ۲۰۱۵). شبکه اینترنت، شبکه مسیرهای وسایل نقلیه و شبکه استناد تحقیقات علمی از جمله نمونه‌هایی از این شبکه‌های پیچیده هستند. شناخت ساختار درونی گراف شبکه‌ها از این لحاظ، که به شناخت عملکرد، تغییر و پیکربندی آن‌ها منجر می‌شود، بسیار حائز اهمیت است (سپهدوست و صدری، ۱۳۹۶). ساختارهای اجتماعی موجود در شبکه‌ها، حاوی اطلاعات ارزشمند برای اهداف تجاری و علمی هستند (فورچنتو^۲، ۲۰۱۰). ساختار اجتماعی در یک شبکه پیچیده نظیر سیستم حمل و نقل یا اینترنت متشکل از مجموعه‌ای از اعضا است که ارتباط این اعضا با یکدیگر در مقایسه با سایر موجودیت‌ها متراکم‌تر است. با شناخت ساختارهای شبکه در خواهیم یافت که نحوه اثرگذاری توپولوژی شبکه و عملکرد آن‌ها بر یکدیگر چگونه خواهد بود (جعفری و همکاران، ۱۳۹۵). از این حیث، شناسایی و تشخیص ساختارهای اجتماعی از جمله مسائل با اهمیت در مطالعه و بررسی سیستم‌های شبکه است که تحت عنوان مسئله اجتماع یابی شناخته می‌شود (ژانگ و همکاران^۳، ۲۰۰۷). اجتماع یابی، موضوعی کاربردی در زمینه تحلیل شبکه‌های اجتماعی و محاسبات نرم است (ژانگ و همکاران^۴، ۲۰۱۴). در واقع، اجتماع یابی ساختار پنهان و معنادار مستتر در گراف‌ها را روشن می‌سازد (چن و همکاران^۵، ۲۰۱۴). یک اجتماع، به صورت زیرمجموعه‌ای از رئوس شبکه تعریف می‌شود که این رئوس نسبت به سایر گره‌های موجود در شبکه، ارتباطات متراکم‌تر و چگالی‌تری با هم دارند (چادهری و پاول^۶، ۲۰۱۳).

در سال‌های اخیر، روش‌های متعددی برای حل مسئله اجتماع یابی و کشف ساختارهای باکیفیت و مناسب در شبکه‌های پیچیده پیشنهاد شده‌اند. از جمله معیارهای سنجش کیفیت

1. Ghorbanian & Shafaqi
2. Fortunato
3. Zhang et al.
4. Zhang et al.
5. Chen et al.
6. Choudhury & Paul

اجتماعات شبکه، معیار کمی چگالی پودمانگی است که ماهیت بیشینه‌سازی دارد. به عبارت دیگر، هرچه مقدار چگالی پودمانگی اجتماعات کشف شده بیشتر باشد، عملکرد الگوریتم مورد استفاده مطلوب تر بوده است. در این مطالعه نیز به منظور بیشینه‌سازی این شاخص، رویکردی ترکیبی متشکل از دو الگوریتم فرا ابتکاری بهینه‌سازی علف‌های هرز مهاجم^۱ (IWO) و ژنتیک^۲ (GA) ارائه شده است. رویکرد پیشنهادی این پژوهش؛ شامل دو فاز است؛ در فاز اول مسئله اجتماع یابی توسط الگوریتم بهینه‌سازی علف‌های هرز مهاجم حل شده و جواب‌های اولیه با کیفیت مناسب پیدا خواهند شد. در فاز دوم، الگوریتم ژنتیک به کمک جواب‌های اولیه فاز اول، فرآیند بهینه‌سازی را ادامه می‌دهد و به این ترتیب، اجتماعات کشف شده از کیفیت قابل قبول و مطلوبی برخوردار خواهند بود.

ساختار مقاله، در ادامه به این ترتیب است؛ بخش دوم به پژوهش‌های پیشین در زمینه اجتماع یابی مرور شده است. مسئله اجتماع یابی در بخش سوم تعریف شده و الگوریتم پیشنهادی نیز تشریح می‌شود. بخش چهارم به سنجش عملکرد الگوریتم پیشنهادی اختصاص دارد. در بخش پنجم نیز به جمع‌بندی مطالب و ارائه پیشنهادهایی برای مطالعات آتی پرداخته شده است.

پیشینه پژوهش

با افزایش روزافزون کاربران شبکه‌های اجتماعی، پیچیدگی این شبکه‌ها به صورت چشمگیری در حال افزایش است (روشنی و همکاران، ۱۳۹۲؛ بخشی و سمیع‌زاده، ۱۳۹۶). از این رو، با افزایش روش‌های گوناگون برای تحلیل این شبکه‌ها مواجه هستیم. یکی از روش‌های شناخته شده در مسائل اجتماع یابی، الگوریتم گیروان- نیومن^۳ است که برای بیشینه‌سازی شاخص پودمانگی^۴ مورد استفاده قرار گرفته است (نیومن و گیروان^۵، ۲۰۰۴). پودمانگی میزان

1. Invasive Weed Optimization (IWO)

2. Genetic Algorithm (GA)

3. Girvan-Newman(GN)

4. Modularity

5. Newman & Girvan

کیفیت ساختارهای پیدا شده توسط یک روش را به دست می‌دهد که مقدار آن در دامنه بین صفر و یک قرار می‌گیرد (گریشیچ و پلوهار^۱، ۲۰۱۱). بیشینه‌سازی شاخص پودمانگی به منظور یافتن اجتماعات با کیفیت، در دسته مسائل سخت^۲ قرار می‌گیرد (برندیس و همکاران^۳، ۲۰۰۸). با توجه به پیچیدگی این مسئله، الگوریتم‌های ابتکاری و فرا ابتکاری متنوعی برای اجتماع یابی توسعه یافته‌اند (گوئیمرا و آمارال^۴، ۲۰۰۵) که از جمله آن‌ها می‌توان به پژوهش شقاقی و همکاران (۲۰۱۵) اشاره کرد. در این تحقیق، یک روش حریم‌سازانه^۵ برای بهینه کردن مقدار پودمانگی ارائه شده است.

الگوریتم ژنتیک از جمله روش‌های تکاملی قدرتمندی است که مورد توجه پژوهشگران در حیطه مسائل اجتماع یابی قرار گرفته است (حافظ و همکاران^۶، ۲۰۱۲). این الگوریتم توسط بینگول و تاسگین^۷ (۲۰۰۶) (۲۰۰۶) برای بیشینه کردن مقدار پودمانگی بکار گرفته شد. الگوریتم قدرتمند GA-Net توسط پیتزوتی^۸ (۲۰۰۸) پیشنهاد شد که هدف آن بهینه‌سازی معیار امتیاز اجتماع^۹ بود. شی و همکاران^{۱۰} (۲۰۱۰)، یک الگوریتم ژنتیک تطبیقی برای اجتماع یابی توسعه دادند. شقاقی^{۱۱} (۲۰۱۴) به کمک یک ساختار نمایش جواب ماتریسی به بیشینه‌سازی چگالی پودمانگی پرداخت. برادران و همکاران (۱۳۹۷- الف) در پژوهش خود الگوریتم ژنتیک مبتنی بر روش تصمیم‌گیری چندمعیاره تاپسیس^{۱۲} برای مسئله اجتماع یابی پیشنهاد کردند که در هر تکرار از این روش، جواب‌های جمعیت با استفاده از روش تاپسیس رتبه‌بندی شده و بهترین جواب‌ها به‌عنوان کروموزوم‌های والد انتخاب می‌شدند. در مطالعه‌ای دیگر، برادران و همکاران (۱۳۹۷- ب) مسئله اجتماع یابی را با دو شاخص

1. Griechisch & Pluhar
2. NP-hard
3. Brandes et al.
4. Guimera & Amaral
5. Greedy algorithm
6. Hafez et al.
7. Bingol & Tasgin
8. Pizzuti
9. Community Score (CS)
10. Shi et al.
11. Shaqaqi
12. TOPSIS

پودمانگی و امتیاز اجتماعات مدل‌سازی کرده و دو الگوریتم ژنتیک با مرتب‌سازی غیر مغلوب نسخه دوم^۱ (NSGA-II) و الگوریتم ژنتیک با رتبه‌بندی غیر مغلوب^۲ (NRGA) برای بهینه نمودن این دو شاخص پیشنهاد کردند. عملگرهای تقاطع و جهش پیشنهادی در این پژوهش، مبتنی بر ویژگی‌های گره‌ها نظیر میزان نزدیکی و درجات داخلی و خارجی است. در الگوریتم پیشنهادی توسط محمود و اسمال^۳ (۲۰۱۶)، هر گره به صورت ترکیب خطی گره‌های یک اجتماع است. گوئسمی و همکاران^۴ (۲۰۱۶) روشی ابتکاری برای مواجهه با انواع گوناگون اشیا یا موضوعات و روابط مشتق شده از شبکه آن‌ها پیشنهاد داده‌اند. لی و همکاران^۵ (۲۰۱۶)، چارچوبی مبتنی بر ماتریس نامنفی منظم و مشترک برای بهبود دقت اجتماع یابی پیشنهاد کردند. پیل و همکاران^۶ (۲۰۱۷) دو تکنیک آماری برای کمی‌سازی ارتباطات بین فراداده و ساختار اجتماعات پیشنهاد داده‌اند. این تکنیک‌های آماری، برای شبکه‌های واقعی و مصنوعی مورد آزمون قرار گرفتند. در پژوهشی دیگر، حسینیان و برادران^۷ (۲۰۱۸)، سیستمی چندعاملی^۸ برای تعیین ساختارهای شبکه‌های اجتماعی پیشنهاد کردند. در پژوهش انجام شده توسط فورچنتو و بارتلمی^۹ (۲۰۰۷)، نشان داده شد که کشف اجتماعاتی که از یک حد آستانه‌ای مشخص کوچک‌تر باشند، توسط بیشینه‌سازی شاخص پودمانگی دشوار است. از سوی دیگر، شاخص پودمانگی با نقص محدودیت جواب^{۱۰} نیز مواجه است (چن و همکاران، ۲۰۱۴). بدین سبب، ژانگ و لی^{۱۱} (۲۰۰۸)، تابعی کمی تحت عنوان چگالی پودمانگی^{۱۲} برای ارزیابی کیفیت اجتماعات پیشنهاد کردند. این محققان نشان دادند که کارایی الگوریتم آن‌ها با استفاده از شاخص چگالی پودمانگی به مراتب بهتر از سایر

1. Non-dominated Sorting Genetic Algorithm II (NSGA-II)

2. Non-dominated Ranked Generic Algorithm (NRGA)

3. Mahmood & Small

4. Guesmi et al.

5. Li et al.

6. Peel et al.

7. Hosseinian & Baradaran

8. Multi-agent system

9. Fortunato & Barthelemy

10. Resolution Limit Solution

11. Zhang & Li

12. Modularity Density Or D-Value

روش‌ها است؛ بنابراین با توجه به توضیحاتی که داده شد، الگوریتم پیشنهادی در این پژوهش برای یافتن اجتماعات شبکه‌های پیچیده از شاخص چگالی پودمانگی بهره می‌برد.

روش‌شناسی پژوهش

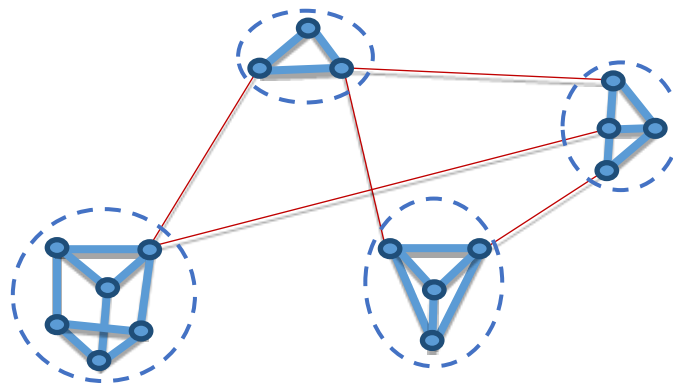
در این پژوهش، الگوریتمی فرا ابتکاری ترکیبی متشکل از دو روش بهینه‌سازی علف‌های هرز و الگوریتم ژنتیک برای بهبود جواب‌های مسئله اجتماع یابی است. بیشینه‌سازی چگالی پودمانگی، به عنوان تابع هدف مسئله در نظر گرفته شده است. کارایی الگوریتم پیشنهادی در مقایسه با چهار الگوریتم ژنتیک کلاسیک، الگوریتم بهینه‌سازی علف‌های هرز، الگوریتم کرم شب‌تاب^۱ (FFA) و روش جستجوی کاملاً تصادفی (RS)^۲ مورد ارزیابی قرار خواهد گرفت. مقایسه عملکرد روش‌ها در ۸ شبکه محک انجام شده است.

تشریح مسئله اجتماع یابی

مسئله اجتماع یابی به افراز یک شبکه به مجموعه‌ای از اجتماعات می‌پردازد که ارتباط بین اعضای اجتماع با یکدیگر قوی بوده و در عین حال ارتباط اعضای یک اجتماع با اعضای اجتماع دیگر کم‌رنگ است؛ بنابراین در این مسئله، هر اجتماع متشکل از مجموعه‌ای از رئوس است. این رئوس دارای ویژگی‌ها یا نقش‌های مشابه هستند (برادران و همکاران، ۱۳۹۷-الف). یک شبکه اجتماعی با گراف $G(H, I)$ را در نظر بگیرد که در آن H نشانگر تعداد کلیه رئوس شبکه است. هر رأس در این شبکه نماینده یک فرد است. I در این گراف نشان‌دهنده تعداد یال‌های شبکه است (قربانیان و شقاقی، ۲۰۱۵). شکل (۱)، نشانگر گراف مربوط به یک شبکه و ارتباطات اعضای آن با یکدیگر است. اجتماعات در این شکل با استفاده از خط‌چین‌ها نمایش داده شده‌اند. همان‌طور که در این شکل نشان داده شده است، هر اجتماع به صورت یک زیرگراف از گراف اصلی شبکه تعریف می‌شود. در این اجتماع،

1. Firefly algorithm (FFA)
2. Random Search (RS)

تراکم یال‌های متصل‌کننده اعضا به یکدیگر زیاد است. در نقطه مقابل، چگالی یال‌ها از گره‌های داخل اجتماع به رئوس خارج از اجتماع پایین است.

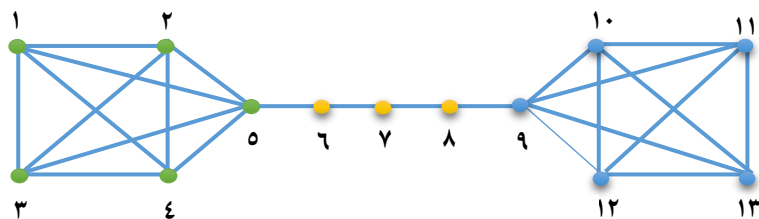


شکل ۱: روابط گره‌های یک شبکه

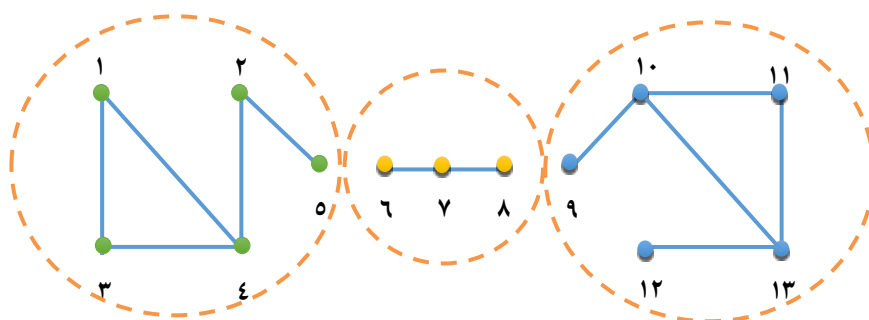
نحوه نمایش جواب‌ها

الگوریتم‌های مورد استفاده در این پژوهش از ساختار نمایش لوکاسی استفاده می‌کنند. با توجه به این ساختار، هر کروموزوم^۱ (جواب) به صورت یک لیست $1 \times n$ تعریف می‌شود، که n نشانگر تعداد ژن‌های این کروموزوم است. مقدار هر ژن با توجه به روابط همسایگی بین گره‌ها تعیین می‌شود. در صورتی که گره‌های i و j در گراف اصلی شبکه با یکدیگر ارتباط داشته باشند، ژن i ام در کروموزوم می‌تواند برابر مقدار j باشد. این امر بدان معنی است که میان گره‌های i و j ارتباطی وجود دارد (شی و همکاران، ۲۰۱۲). شبکه‌ای به صورت شکل (۲) شامل ۱۳ گره و متشکل از سه اجتماع، را در نظر بگیرید. شکل (۳)، یک جواب موجه^۲ را با توجه به گراف شبکه شکل (۲) نشان می‌دهد. ساختار این جواب در شکل (۴) قابل مشاهده است. همان‌طور که در شکل (۴) نشان داده شده است، هر اجتماع متشکل از مجموعه‌ای از گره‌های شبکه اصلی است.

1. Chromosome
2. Feasible



شکل ۲: نمایی از گراف شبکه



شکل ۳: یک جواب موجه با توجه به گراف شکل ۲

موقعیت	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳
جواب	۳	۵	۴	۱	۲	۷	۶	۷	۱۰	۱۱	۱۳	۱۳	۱۰

شکل ۴: اجتماعات کشف شده با توجه به شکل های ۲ و ۳

شاخص ارزیابی جوابها

همان طور که قبلاً اشاره شد، در این پژوهش از شاخص چگالی پودمانگی (D) برای ارزیابی کیفیت جوابها استفاده شده است. چگالی پودمانگی از طریق رابطه (۱) قابل محاسبه است (قربانیا و شقاقی، ۲۰۱۵):

$$D = \sum_{p=1}^P \left(\frac{\theta_p - \gamma_p}{\sum_i \lambda_{ip}} \right) \quad (1)$$

در رابطه (۱)، متغیر λ_{ip} به صورت یک متغیر باینری است، به طوری که اگر گره i در اجتماع p قرار داشته باشد، مقدار این متغیر برابر ۱ خواهد بود و در غیر این صورت ارزش این متغیر صفر است. P نشانگر تعداد کل اجتماعات در شبکه است. θ_p و γ_p به ترتیب نشان‌دهنده دو برابر یال‌های اجتماع p و تعدادی یال‌های خارج شونده از این اجتماع هستند.

فرآیند انتخاب جواب‌های والد

در این پژوهش، فرآیند انتخاب والدین برای تولید فرزندان از طریق انتخاب مسابقه‌ای^۱ صورت می‌پذیرد. انتخاب مسابقه‌ای، روشی برای انتخاب یک جواب از میان جمعیتی از جواب‌ها در الگوریتم ژنتیک است. در روش انتخاب مسابقه‌ای، چندین مسابقه میان تعدادی از جواب‌ها که به صورت تصادفی از جامعه برگزیده شده‌اند، اجرا می‌شود. جوابی که بهترین مقدار برازندگی را داشته باشد، به عنوان برنده هر مسابقه در نظر گرفته شده و عملگر تولید فرزندان بر روی آن اعمال می‌شود. هرچه اندازه مسابقه بزرگ‌تر باشد، جواب‌های ضعیف شانس کمتری برای انتخاب شدن دارند (میلر و گلدبرگ^۲، ۱۹۹۵).

عملگر تقاطع^۳

در این پژوهش از عملگرهای تقاطع تک نقطه‌ای و دونقطه‌ای برای تولید جواب‌های جدید استفاده شده است. در هر تکرار از الگوریتم پیشنهادی، یکی از عملگرهای تک نقطه‌ای یا دونقطه‌ای به تصادف انتخاب خواهند شد. در گام بعد، محل تقاطع در کروموزوم‌های والدین

1. Tournament Selection
2. Miller & Goldberg
3. Crossover operator

به صورت کاملاً تصادفی تعیین شده و کروموزوم‌های فرزند تولید می‌شوند (برادران و همکاران، ۱۳۹۷-الف؛ کوکیانگ و ژایوفنگ^۱، ۲۰۱۰).

عملگر جهش^۲

عملگر جهش مورد استفاده در این پژوهش، مقادیر ژن‌ها در یک جواب را با توجه به گره‌های همسایه آن تغییر می‌دهد. به این ترتیب تعداد مشخصی از ژن‌ها به صورت تصادفی انتخاب شده و مقادیر آن‌ها با توجه به شماره گره‌های همسایگی آن‌ها تغییر خواهد کرد (کوکیانگ و ژایوفنگ، ۲۰۱۰).

الگوریتم بهینه‌سازی علف‌های هرز مهاجم (IWO)

این الگوریتم، روشی فرا ابتکاری است که از فرآیند تکثیر علف‌های هرز الگوبرداری می‌کند. شیوه عملکرد این الگوریتم، مشتمل بر چهار مرحله است (محرابیان و لوکاس^۳، ۲۰۰۶):

۱. **ایجاد جمعیت اولیه:** در ابتدا، تعداد معینی از پاسخ‌ها به صورت تصادفی در فضای جستجوی d بعدی گسترده می‌شوند.

۲. **تکثیر:** هر عضو از جمعیت، دانه‌هایی را در ناحیه مشخصی به مرکزیت خود پراکنده می‌نماید. تعداد دانه‌ها بستگی به میزان برازندگی نسبی هر عضو با توجه به بهترین و بدترین برازندگی دارد. تعداد دانه‌ها به صورت خطی از مقدار min_seed تا max_seed تغییر می‌کند. مقدار min_seed و max_seed به ترتیب بدترین و بهترین عضو جمعیت هستند.

۳. **پراکندگی دانه‌ها:** پراکندگی دانه‌ها در فضای جستجوی d بعدی با توزیع نرمال، میانگین صفر و واریانس σ^2 انجام می‌شود. انحراف استاندارد σ نیز برای کاهش اندازه

1. Guoqiang & Xiaofang

2. Mutation operator

3. Mehrabian & Lucas

نسل‌ها مورد استفاده قرار می‌گیرد. با این گام، احتمال پخش یک دانه در یک ناحیه مشخص به صورت غیر خطی کاهش می‌یابد و الگوریتم به تدریج از رفتار کاوشگرانه به رفتار تبعیت‌جویانه تغییر حالت می‌دهد.

$$\sigma_{iter} = \sigma_{\min} + \left(\frac{iter_{\max} - iter}{iter_{\max}} \right)^{nmi} (\sigma_{\max} - \sigma_{\min}) \quad (2)$$

در رابطه بالا، مقادیر σ_{\max} ، σ_{\min} به ترتیب برابر با بیشترین و کمترین مقدار انحراف استاندارد است. σ_{iter} انحراف استاندارد در تکرار فعلی است. مقدار nmi شاخص تعدیل غیرخطی است. $iter$ و $iter_{\max}$ نیز به ترتیب شماره تکرار فعلی و بیشترین تعداد تکرار را نشان می‌دهند.

۴. حذف رقابتی^۱: نسل گیاهی که فرزندی ندارد، منقرض خواهد شد. در غیر این صورت، فرزندان این گیاه تمامی ناحیه اطراف را فرا خواهند گرفت. در نتیجه، نیاز به وجود نوعی رقابت بین گیاهان است که تعداد آن‌ها در یک جمعیت را محدود نماید. در ابتدا گیاهان سریع تکثیر می‌شوند و تا زمان رسیدن به یک حد معلوم pop_max ، تمامی جواب‌ها در جمعیت جای خواهند گرفت. سپس، تنها گیاهان برازنده باقی می‌مانند. گام‌های ۱ تا ۴ تا زمان رسیدن به بیشینه تعداد تکرار خواهد شد. اندازه جمعیت هر نسل باید کوچک‌تر یا مساوی pop_max باشد.

الگوریتم پیشنهادی GAIWO

با وجود موفقیت الگوریتم IWO در حل بسیاری از مسائل بهینه‌سازی، این الگوریتم معمولاً به همگرایی زودرس می‌رسد و این بیم وجود دارد که در بهینه محلی گرفتار شود؛ بنابراین با

توجه به ویژگی‌های سرآمد روش IWO از جمله توانایی آن در کاوش^۱، تبعیت^۲ و نیز به منظور جلوگیری از گرفتار شدن در بهینه محلی، الگوریتمی ترکیبی از روش IWO و ژنتیک به نام GAIWO ارائه می‌شود تا احتمال گرفتار شدن الگوریتم در بهینه محلی تا حدود زیادی کاهش یافته و کیفیت جواب‌های به دست آمده نیز بهبود پیدا کند.

روش اجرای این الگوریتم به این شرح است که ابتدا مسئله اجتماع یابی با هدف بیشینه‌سازی چگالی پودمانگی توسط الگوریتم IWO حل شده و پس از تعداد تکرار معین، جواب‌هایی که بیشترین میزان برازندگی را دارند، برگزیده خواهند شد. طبیعی است که پاسخ‌های حاصل شده از روش IWO، نسبت به پاسخ‌های تصادفی از شایستگی بالاتری برخوردار خواهند بود. در ادامه جواب‌های تولید شده به عنوان جمعیت اولیه الگوریتم ژنتیک در نظر گرفته می‌شوند و روند حل مسئله به وسیله الگوریتم ژنتیک ادامه پیدا می‌کند و این بار به کمک عملگرهای تقاطع و جهش جواب‌های متنوعی به وجود خواهد آمد.

معرفی شبکه‌های نمونه (محک)

در این بخش، چند شبکه دنیای واقعی که برای مقایسه کارایی الگوریتم‌ها مدنظر قرار گرفته‌اند، در جدول (۱) معرفی می‌گردد.

-
1. Exploration
 2. Exploitation

جدول ۱: معرفی شبکه‌های نمونه برای سنجش کارایی الگوریتم‌ها

نام شبکه	شرح
باشگاه کاراته زکریا ^۱	گره‌های شبکه، اعضای یک باشگاه کاراته هستند. یال‌ها در این شبکه بیانگر روابط بین اعضای باشگاه است (نیومن، ۲۰۰۶).
کتاب‌های سیاست آمریکا ^۲	گره‌های شبکه، کتاب‌هایی مرتبط با سیاست آمریکا است که توسط افراد گوناگون به وسیله یک فروشگاه آنلاین تهیه شده است. یال‌ها، جفت کتاب‌هایی است که یک فرد خریداری کرده است (نیومن، ۲۰۰۶).
شخصیت‌های رمان بی‌نویان ^۳	گره‌ها نماینده شخصیت‌های رمان بی‌نویان است. هر یال نشان از وجود دو شخصیت در یک فصل یکسان دارد (نیومن، ۲۰۰۶).
دلفین‌ها ^۴	دلفین‌ها نماینده هر گره می‌باشند. ارتباط دلفین‌ها با یکدیگر از سال‌های ۱۹۹۴ تا ۲۰۰۱ مورد مطالعه قرار گرفته است (نیومن، ۲۰۰۶).
نوازندگان جاز ^۵	شبکه نشانگر همکاری نوازندگان جاز است. هر گره، یک نوازنده است. هر یال، ارتباط دو نوازنده را که در یک گروه موسیقی فعالیت دارند، نشان می‌دهد. (گلنیز و دانون، ۲۰۰۳).
تیم‌های فوتبال ^۶	در این شبکه برگزاری مسابقات فوتبال بین تیم‌های دانشگاهی در ایالات متحده آمریکا مورد مطالعه قرار گرفته است (نیومن و گیروان، ۲۰۰۴).
مجاورت کلمات ^۸	شبکه شامل صفات و اسامی است که در رمان دیوید کاپرفیلد در مجاورت با یکدیگر قرار می‌گیرند. یال‌ها، جفت کلماتی که در متن کتاب در مجاورت هم قرار دارند را متصل می‌نمایند (نیومن، ۲۰۰۶).
فیس بوک ^۹	این شبکه، متشکل از ۴۰۳۹ گره و ۸۸۲۳۴ یال، از جمله شبکه‌های بزرگی است که به‌منظور آزمون کارایی الگوریتم‌های اجتماع یابی مورد استفاده قرار می‌گیرد (وین و لین، ۲۰۱۸).

یافته‌های پژوهش

پارامترهای ورودی به هر الگوریتم، اثرگذاری ویژه‌ای بر نتایج دارند؛ بنابراین تنظیم پارامترها را می‌توان گامی مهم در به دست آوردن جواب‌های با کیفیت دانست. بدین منظور، الگوریتم‌ها به ازای مقادیر مختلف پارامترها اجرا شده و چگالی پودمانگی متناسب با مسئله نمونه به دست می‌آید. از جمله روش‌های کارآمد در زمینه طراحی آزمایش‌ها، روش

1. Zakhary Karate Club
2. United States Political Books (PolBooks)
3. Les Misérables
4. Dolphins Network
5. Jazz Musicians Network
6. Gleiser & Danon
7. Football Club
8. Word Adjacencies Network(WAN)
9. Facebook
10. Win & Lynn

تاگوچی^۱ است که در این مطالعه مورد استفاده قرار گرفته است. بر این اساس، طراحی آزمایش‌ها با تعریف سطوح متفاوتی برای هر پارامتر صورت می‌پذیرد. جدول (۲) پارامترهای الگوریتم‌های مورد استفاده را نشان می‌دهد. همچنین در این جدول، سطوح مربوط به هر پارامتر نشان داده شده است.

جدول ۲: پارامترهای الگوریتم‌ها، نماد و سطوح آن‌ها

الگوریتم	پارامترها	نمادها	سطوح
GA	نرخ تقاطع	<i>pc</i>	۰/۴-۰/۶-۰/۸
	نرخ جهش	<i>pm</i>	۰/۱-۰/۳-۰/۵
	اندازه جمعیت	<i>pop</i>	۱۰۰-۲۰۰-۳۰۰
	تعداد تکرار	<i>Iteration</i>	۱۰۰-۲۰۰-۳۰۰
IWO	اندازه جمعیت	<i>popI</i>	۶-۸-۱۰
	بیشترین اندازه جمعیت	<i>pop max</i>	۲۰-۵۰-۱۰۰
	حداقل تعداد دانه‌ها	<i>min_seed</i>	۰-۱-۲
	حداکثر تعداد دانه‌ها	<i>max_seed</i>	۳-۴-۵
	توان کاهش واریانس	<i>nmi</i>	۱-۱/۵-۲
	انحراف معیار اولیه	$\sigma_{initial}$	۰/۲۵-۰/۵-۱
	انحراف معیار پایانی	σ_{final}	۰/۰۰۱-۰/۱-۰/۲
FFA	تعداد تکرار	<i>Iteration</i>	۱۰۰-۳۰۰-۵۰۰
	اندازه جمعیت	<i>pop</i>	۱۰۰-۲۰۰-۳۰۰
	اندازه جمعیت فرعی	<i>pop_s</i>	۱۰-۱۵-۲۰
	تعداد جواب‌های نخبه	<i>NE</i>	۵-۱۰-۲۰
	تعداد تکرار	<i>Iteration</i>	۱۰۰-۳۰۰-۵۰۰

بدین ترتیب، آزمایش‌هایی به منظور تعیین اندازه پارامترها برای رسیدن به بیشینه مقدار چگالی پودمانگی طراحی شد. بعد از ده مرتبه اجرای برنامه و میانگین‌گیری، نتایج مرتبط با

پارامترهای الگوریتم GA، IWO و FFA حاصل شده و به ترتیب در جداول (۳)، (۴) و (۵) قابل مشاهده است. مقدار چگالی پودمانگی (مقدار D) نشان داده شده است. مقادیر پارامترها پس از انجام آزمایش‌ها در جدول (۶) آورده شده است. به دلیل آنکه پارامترهای الگوریتم GAIWO متشکل از پارامترهای دو الگوریتم GA و IWO است، از مقادیر جدول (۶) برای الگوریتم پیشنهادی استفاده می‌شود.

جدول ۳: نتایج روش تاگوچی برای الگوریتم ژنتیک

شماره آزمایش	pop	pc	pm	Iteration	میانگین D	میانگین زمان (ثانیه)
۱	۱۰۰	۰/۴	۰/۱	۱۰۰	۷/۳۳۳	۱/۵۳۳۱۲
۲	۱۰۰	۰/۶	۰/۳	۲۰۰	۷/۶۶۷	۵/۲۶۳۶
۳	۱۰۰	۰/۸	۰/۵	۳۰۰	۹/۱۰۸	۱۱/۵۹۶۷
۴	۲۰۰	۰/۴	۰/۳	۳۰۰	۸/۶۶۶۷	۱۲/۵۵۵
۵	۲۰۰	۰/۶	۰/۵	۱۰۰	۸/۴۴۴	۶/۵۷۶۰
۶	۲۰۰	۰/۸	۰/۱	۲۰۰	۸/۰۸۸۹	۱۴/۲۹۹۳
۷	۳۰۰	۰/۴	۰/۵	۲۰۰	۹/۷۵	۱۶/۷۲۴۵
۸	۳۰۰	۰/۶	۰/۱	۳۰۰	۹/۰۷۱۴	۱۶/۸۰۸۹
۹	۳۰۰	۰/۸	۰/۳	۱۰۰	۸/۴۰۰	۹/۱۸۹۸

جدول ۴: نتایج روش تاگوچی برای الگوریتم بهینه‌سازی علف‌های هرز

آزمایش	popI	pop max	Min Seed	Max Seed	nmi	$\sigma_{initial}$	σ_{final}	Iteration	میانگین D	میانگین زمان (ثانیه)
۱	۶	۲۰	۰	۳	۱	۰/۲۵	۰/۰۰۱	۱۰۰	۶/۷۱	۰/۷۳۷۵
۲	۶	۲۰	۰	۳	۱/۵	۰/۵	۰/۱	۳۰۰	۱/۹۶۲۸	۹/۵
۳	۶	۲۰	۰	۳	۲	۱	۰/۲	۵۰۰	۳/۸۵۶	۹/۴۴۴
۴	۶	۵۰	۱	۴	۱	۰/۲۵	۰/۰۰۱	۳۰۰	۷/۲۸۷	۸
۵	۶	۵۰	۱	۴	۱/۵	۰/۵	۰/۱	۵۰۰	۱۲/۹۶۶	۹/۲۶۶
۶	۶	۵۰	۱	۴	۲	۱	۰/۲	۱۰۰	۱۲/۰۶۲	۹/۳۵۷
۷	۶	۱۰۰	۲	۵	۱	۰/۲۵	۰/۰۰۱	۵۰۰	۹/۱۱	۹/۵
۸	۶	۱۰۰	۲	۵	۱/۵	۰/۵	۰/۱	۱۰۰	۸/۶۷	۸/۸۶۶
۹	۶	۱۰۰	۲	۵	۲	۱	۰/۲	۳۰۰	۲۲/۰۲	۹/۵۷۱
۱۰	۸	۲۰	۱	۵	۱	۰/۵	۰/۲	۱۰۰	۱/۰۹	۷
۱۱	۸	۲۰	۱	۵	۱/۵	۱	۰/۰۰۱	۳۰۰	۳/۵۳	۸/۶
۱۲	۸	۲۰	۱	۵	۲	۰/۲۵	۰/۱	۵۰۰	۴/۹۱	۸/۶
۱۳	۸	۵۰	۲	۳	۱	۰/۵	۰/۲	۳۰۰	۹/۷۲۸	۹/۵
۱۴	۸	۵۰	۲	۳	۱/۵	۱	۰/۰۰۱	۵۰۰	۱۴/۷۷	۹/۵۷۱
۱۵	۸	۵۰	۲	۳	۲	۰/۲۵	۰/۱	۱۰۰	۳/۷۷	۷/۷۳۳
۱۶	۸	۱۰۰	۰	۴	۱	۰/۵	۰/۲	۵۰۰	۱۶/۹۷	۹/۴۴
۱۷	۸	۱۰۰	۰	۴	۱/۵	۱	۰/۰۰۱	۱۰۰	۲/۹۲	۸/۴۴
۱۸	۸	۱۰۰	۰	۴	۲	۰/۲۵	۰/۱	۳۰۰	۱۱/۱۴	۹/۵۷
۱۹	۱۰	۲۰	۲	۴	۱	۱	۰/۱	۱۰۰	۱/۸۶۲	۹/۰۷۱
۲۰	۱۰	۲۰	۲	۴	۱/۵	۰/۲۵	۰/۲	۳۰۰	۴/۷۴۹	۹/۵۷
۲۱	۱۰	۲۰	۲	۴	۲	۰/۵	۰/۰۰۱	۵۰۰	۵/۸۷	۹/۳۵
۲۲	۱۰	۵۰	۰	۵	۱	۱	۰/۱	۳۰۰	۷/۷۳	۹/۷۵
۲۳	۱۰	۵۰	۰	۵	۱/۵	۰/۲۵	۰/۲	۵۰۰	۱۰/۶۹	۹/۱۶۶
۲۴	۱۰	۵۰	۰	۵	۲	۰/۵	۰/۰۰۱	۱۰۰	۲/۳۵	۷/۳۳
۲۵	۱۰	۱۰۰	۱	۳	۱	۱	۰/۱	۵۰۰	۲۱/۱۴	۸/۶
۲۶	۱۰	۱۰۰	۱	۳	۱/۵	۰/۲۵	۰/۲	۱۰۰	۴/۸۰۸	۷/۹۰۹
۲۷	۱۰	۱۰۰	۱	۳	۲	۰/۵	۰/۰۰۱	۳۰۰	۱۳/۹۹	۹/۴۴۴

جدول ۵: نتایج روش تاگوچی برای الگوریتم کرم شب تاب

شماره آزمایش	pop	pop_s	NE	$Iteration$	میانگین D	میانگین زمان (ثانیه)
۱	۱۰۰	۱۰	۵	۱۰۰	۹/۲۵۸۹	۲/۶۴۰۹
۲	۱۰۰	۱۵	۱۰	۳۰۰	۹/۶۲۳۲	۷/۳۵۱۴
۳	۱۰۰	۲۰	۲۰	۵۰۰	۶/۵۰۷۹	۱۵/۳۸۱۶
۴	۲۰۰	۱۰	۱۰	۵۰۰	۹/۶۵۳۵	۱۶/۱۲۷۵
۵	۲۰۰	۱۵	۲۰	۱۰۰	۸/۵۲۹۴	۸/۳۸۱۴
۶	۲۰۰	۲۰	۵	۳۰۰	۶/۳۹۰۲	۱۷/۳۸۹۱
۷	۳۰۰	۱۰	۲۰	۳۰۰	۷/۱۱۴۰	۱۸/۶۴۱۳
۸	۳۰۰	۱۵	۵	۵۰۰	۸/۱۸۷۵	۲۵/۷۶۱۴
۹	۳۰۰	۲۰	۱۰	۱۰۰	۹/۸۳۰۰	۱۰/۰۲۵۷

جدول ۶: مقادیر پارامترهای الگوریتم‌ها

الگوریتم	پارامترها	نمادها	مقدار پارامتر
GA	نرخ تقاطع	p_c	۰/۶
	نرخ جهش	p_m	۰/۵
	اندازه جمعیت	pop	۳۰۰
	تعداد تکرار	$Iteration$	۳۰۰
IWO	اندازه جمعیت اولیه	pop_l	۱۰
	بیشترین اندازه جمعیت	pop_max	۱۰۰
	حداقل تعداد دانه‌ها	min_seed	۲
	حداکثر تعداد دانه‌ها	max_seed	۴
	توان کاهش واریانس	nmi	۱/۵
	انحراف معیار اولیه	$\sigma_{initial}$	۱
	انحراف معیار پایانی	σ_{final}	۰/۱
RS	تعداد تکرار	$Iteration$	۳۰۰
	تعداد تکرار	$Iteration$	۳۰۰
FFA	اندازه جمعیت	pop	۳۰۰
	اندازه جمعیت فرعی	pop_s	۱۵
	تعداد جواب‌های نخبه	NE	۱۰
	تعداد تکرار	$Iteration$	۳۰۰

تحلیل نتایج الگوریتم‌ها

برنامه الگوریتم‌ها در نرم‌افزار متلب پیاده‌سازی شده و به وسیله یک رایانه شخصی با ویژگی‌های (Quad Core 8200, CPU 2.00 GHz, RAM 4 GB) اجرا شده‌اند. چگالی پودمانگی و پودمانگی به‌عنوان معیارهای سنجش کارآمدی جواب‌ها، مورد استفاده قرار گرفته‌اند. نتایج حاصل شده در جدول (۷) گزارش شده است. با توجه به این جدول، بهترین مقادیر چگالی پودمانگی و پودمانگی به‌دست آمده توسط الگوریتم‌ها، به ترتیب در ستون‌های Q و D آورده شده است. در ستون‌های میانگین D و میانگین Q نیز می‌توان میانگین نتایج به‌دست آمده در ۳۰۰ تکرار از الگوریتم‌ها را مشاهده کرد.

جدول ۷: نتایج اجرای الگوریتم‌ها

شبکه	گره	یال	الگوریتم	D	میانگین D	Q	میانگین Q	تعداد اجتماعات
Zachary	۳۴	۷۸	GAIWO	۹/۷۵۴	۹/۷۵۴	۰/۳۲۶	۰/۳۲۶	۴
			IWO	۹/۶۶۶	۸/۳۷۰	۰/۳۲۶	۰/۳۰۷	۴
			GA	۹/۴۴۴	۸/۱۳۳	۰/۲۷۶	۰/۲۷۱	۴
			RS	۶/۹۳۳	۶/۶۰۲	۰/۲۰۷	۰/۲۰۱	۳
			FFA	۸/۹۷۸	۸/۶۵۷	۰/۲۶۹	۰/۲۵۴	۴
Dolphins	۶۲	۱۵۹	GAIWO	۲۳/۵۱۴	۲۳/۴۷۰	۰/۴۹۹	۰/۴۹۹	۵
			IWO	۲۱/۲۰۱	۱۷/۰۸۷	۰/۴۱۵	۰/۳۷۶	۵
			GA	۱۸/۶۷۴	۱۷/۱۲۰	۰/۴۴۵	۰/۴۴۷	۵
			RS	۹/۸۸۱	۹/۰۷۵	۰/۳۱۳	۰/۲۸۸	۹
			FFA	۱۸/۵۲۲	۱۶/۹۵۱	۰/۴۱۲	۰/۳۹۲	۵
Football Club	۱۱۵	۶۱۳	GAIWO	۳۲/۳۳۶	۲۹/۳۳۰	۰/۴۹۱	۰/۴۶۵	۶
			IWO	۲۲/۲۵۲	۱۷/۱۹۷	۰/۴۱۲	۰/۳۵۳	۷
			GA	۲۱/۰۵۴	۱۷/۳۹۷	۰/۴۳۵	۰/۳۶۷	۷
			RS	۱۰/۴۵۷	۱۰/۱۱۴	۰/۲۲۸	۰/۲۲۲	۸
			FFA	۲۰/۵۴۱	۱۶/۸۷۴	۰/۴۰۱	۰/۳۸۵	۷
Jazz	۱۹۸	۲۷۴۲	GAIWO	۳۲/۹۰۵	۲۵/۹۰۹	۰/۳۳۲	۰/۲۹۶	۱۰

			IWO	۲۰/۳۲۹	۱۵/۳۳۱	۰/۲۳۳	۰/۲۰۰	۱۰
			GA	۲۲/۴۵۸	۱۸/۴۰۰	۰/۲۶۱	۰/۲۲۹	۱۱
			RS	۸/۲۹۸	۸/۱۸۷	۰/۱۲۹	۰/۱۲۴	۱۵
			FFA	۲۱/۲۴۷	۱۹/۲۷۹	۰/۲۷۸	۰/۲۵۱	۱۰
Les Misérables	۷۷	۲۵۴	GAIWO	۲۲/۴۴۱	۲۲/۲۲۸	۰/۴۷۶	۰/۴۶۸	۶
			IWO	۱۸/۰۱۲	۱۴/۳۰۴	۰/۴۵۷	۰/۴۱۰	۶
			GA	۱۷/۷۴۱	۱۳/۷۰۸	۰/۴۷۵	۰/۴۵۲	۶
			RS	۷/۱۵۲	۶/۳۵۲	۰/۲۸۷	۰/۲۸۰	۸
			FFA	۱۶/۵۲۲	۱۶/۲۲۸	۰/۴۶۶	۰/۴۴۱	۶
PolBooks	۱۰۵	۴۴۱	GAIWO	۳۳/۶۳۳	۲۹/۱۷۶	۰/۴۶۰	۰/۴۳۰	۵
			IWO	۲۱/۲۰۵	۱۶/۱۱۴	۰/۳۸۳	۰/۳۲۴	۵
			GA	۲۹/۳۳۳	۲۴/۵۸۹	۰/۴۱۹	۰/۳۹۹	۵
			RS	۱۰/۸۸۹	۸/۳۱۱	۰/۲۴۳	۰/۲۴۲	۷
			FFA	۲۷/۱۴۴	۲۴/۲۸۷	۰/۳۹۱	۰/۳۷۶	۷
WAN	۱۱۲	۴۲۵	GAIWO	۳۴/۸۴۷	۳۰/۶۵۶	۰/۲۳۰	۰/۲۱۷	۶
			IWO	۲۴/۱۰۵	۱۸/۵۵۷	۰/۱۸۵	۰/۱۶۲	۷
			GA	۲۴/۳۳۶	۲۱/۲۲۰	۰/۱۷۹	۰/۱۶۵	۷
			RS	۹/۵۳۱	۸/۹۴۳	۰/۱۲۸	۰/۱۲۵	۱۰
			FFA	۲۴/۱۸۱	۲۱/۱۸۶	۰/۱۶۱	۰/۱۵۷	۷
Facebook	۴۰۳ ۹	۸۸۲۳ ۴	GAIWO	۴۰/۱۸۵	۳۹/۵۵۷	۰/۴۷۲	۰/۴۵۱	۳۱۴
			IWO	۳۵/۲۲۱	۳۴/۸۸۷	۰/۳۹۱	۰/۳۷۶	۳۱۴
			GA	۳۷/۹۷۱	۳۶/۰۱۴	۰/۳۴۶	۰/۳۳۹	۳۱۴
			RS	۲۹/۶۴۶	۲۸/۳۳۷	۰/۲۹۱	۰/۲۷۵	۲۸۷
			FFA	۳۵/۲۹۸	۳۴/۲۸۳	۰/۴۵۱	۰/۴۴۳	۳۱۴

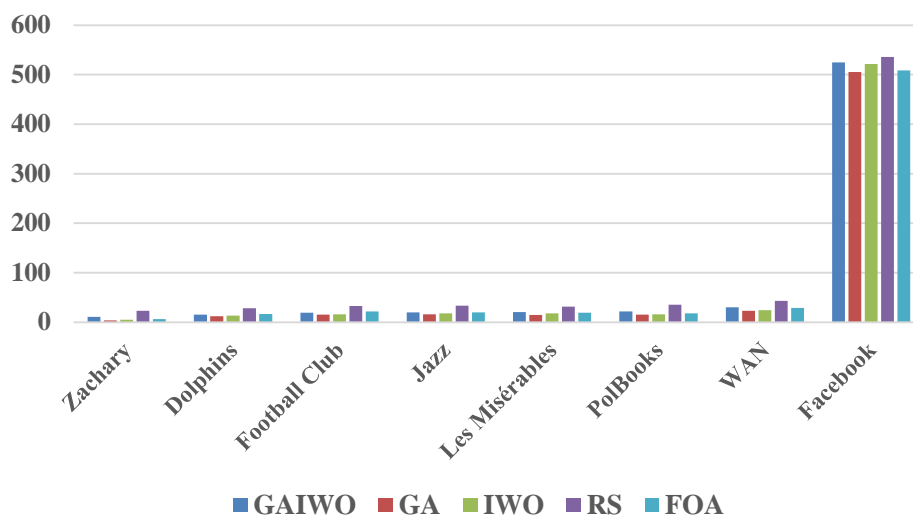
همان‌طور که در جدول (۷) نیز مشخص است، الگوریتم پیشنهادی GAIWO توانسته است از نظر بیشینه‌سازی چگالی پودمانگی و پودمانگی نسبت به الگوریتم‌های دیگر بهتر عمل کند. در ادامه، با استفاده از آزمون آماری t این امر بررسی می‌شود که آیا جواب‌های الگوریتم GAIWO نشانگر بهبود قابل توجه و معناداری نسبت به جواب‌های سایر الگوریتم‌ها هستند یا خیر. فرض صفر این آزمون بیان می‌کند که تفاوت معنادار و قابل توجهی بین میانگین جواب‌های دو الگوریتم وجود ندارد. فرض صفر در صورتی که مقدار P-Value کمتر از

۰/۰۵ باشد، رد خواهد شد. این آزمون فرض‌ها در فاصله اطمینان ۹۵ درصد انجام شده‌اند. جدول (۸)، نتایج این آزمون‌ها را به صورت اجمالی نشان می‌دهد.

جدول ۸: نتایج آزمون آماری t از نظر چگالی پودمانگی

شبکه‌ها	مقایسه الگوریتم‌ها	حد پائین	حد بالا	مقدار آماره t	P- مقدار value
Zachary	GAIWO vs. GA	۱/۴۶۹۶	۱/۷۵۳۲	۲۲/۳۶	۰/۰۰۰
	GAIWO vs. IWO	۱/۲۱۰۵	۱/۵۳۹۲	۱۶/۴۶	۰/۰۰۰
	GAIWO vs. RS	۳/۱۰۳۸	۳/۱۸۲۵	۱۵۷/۱۷	۰/۰۰۰
	GAIWO vs. FFA	۰/۶۴۲	۱/۲۷۷	۱۳/۰۲۳	۰/۰۰۵
Dolphins	GAIWO vs. GA	۶/۱۰۵	۶/۶۱۱	۴۹/۴۱	۰/۰۰۰
	GAIWO vs. IWO	۶/۰۲۵	۶/۷۵۳	۳۴/۵۴	۰/۰۰۰
	GAIWO vs. RS	۱۴/۳۴۵	۱۴/۴۶۷	۴۶۳/۷۷	۰/۰۰۰
	GAIWO vs. FFA	۳/۷۱۱	۵/۸۹۴	۱۸/۹۳۶	۰/۰۰۰
Football Club	GAIWO vs. GA	۱۱/۴۵۵	۱۲/۴۳۶	۴۷/۸۵	۰/۰۰۰
	GAIWO vs. IWO	۱۱/۶۵۰	۱۲/۶۴۲	۴۸/۱۱	۰/۰۰۰
	GAIWO vs. RS	۱۸/۸۹۷	۱۹/۵۶۱	۱۱۳/۹۰	۰/۰۰۰
	GAIWO vs. FFA	۶/۶۷۸	۱۶/۱۸۷	۱۰/۳۴۶	۰/۰۰۹
Jazz	GAIWO vs. GA	۶/۹۹۱	۸/۰۲۶	۲۸/۵۱	۰/۰۰۰
	GAIWO vs. IWO	۱۰/۰۴۱	۱۱/۱۱۹	۳۸/۵۴	۰/۰۰۰
	GAIWO vs. RS	۱۷/۳۸۶	۱۸/۰۵۷	۱۰۳/۸۹	۰/۰۰۰
	GAIWO vs. FFA	۷/۸۵۳	۱۲/۸۷۹	۱۷/۷۴۹	۰/۰۰۰
Les Misérables	GAIWO vs. GA	۸/۱۶۹	۸/۸۴۴	۴۹/۵۴	۰/۰۰۰
	GAIWO vs. IWO	۷/۵۲۸	۸/۲۹۶	۴۰/۵۴	۰/۰۰۰
	GAIWO vs. RS	۱۵/۷۸۰	۱۵/۹۶۶	۳۳۵/۵۴	۰/۰۰۰
	GAIWO vs. FFA	۳/۶۴۹	۷/۱۸۳	۱۳/۱۹۱	۰/۰۰۵
PolBooks	GAIWO vs. GA	۳/۸۸۵	۵/۲۵۹	۱۳/۰۸	۰/۰۰۰
	GAIWO vs. IWO	۱۲/۵۳۵	۱۳/۵۹۱	۴۸/۵۹	۰/۰۰۰
	GAIWO vs. RS	۲۰/۵۱۵	۲۱/۲۱۶	۱۱۷/۰۸	۰/۰۰۰
	GAIWO vs. FFA	۳/۶۴۵	۵/۲۲۵	۱۲/۸۷	۰/۰۰۰
WAN	GAIWO vs. GA	۸/۷۱۷	۱۰/۱۴۹	۲۵/۸۶	۰/۰۰۰
	GAIWO vs. IWO	۱۱/۴۲۶	۱۲/۷۷۲	۳۵/۳۰۱	۰/۰۰۰
	GAIWO vs. RS	۲۱/۲۴۱	۲۲/۱۸۵	۹۰/۵۸۲	۰/۰۰۰
	GAIWO vs. FFA	۸/۶۱۴	۱۱/۵۲۱	۲۶/۸۷۰	۰/۰۰۰
Facebook	GAIWO vs. GA	۰/۹۴۴	۱۳/۷۹۶	۴/۹۳۴	۰/۰۳۸
	GAIWO vs. IWO	۰/۶۶۲	۱۵/۰۷۰	۴/۶۹۸	۰/۰۴۲
	GAIWO vs. RS	۱۷/۲۴۲	۲۵/۸۳۷	۲۱/۵۶۷	۰/۰۰۰
	GAIWO vs. FFA	۲/۶۱۸	۱۰/۴۶۱	۷/۱۷۵	۰/۰۱۸

با توجه به مقادیر P-Value گزارش شده در جدول (۸)، فرض صفر (عدم تفاوت معنادار بین میانگین مقادیر چگالی پودمانگی) در تمامی شبکه‌های نمونه رد شده است. از این حیث، می‌توان اطمینان حاصل کرد که الگوریتم GAIWO توانسته است بهبود معناداری در جواب‌های مسئله اجتماع یابی ایجاد نماید. یکی از معیارهای مقایسه الگوریتم‌ها، زمان مورد نیاز الگوریتم‌ها در حل مسئله است. شکل (۵) میانگین زمان مورد نیاز توسط الگوریتم‌ها را پس از بیست مرتبه اجرای برنامه نشان می‌دهد. با توجه به شکل (۵)، الگوریتم GA توانسته است بهترین نتیجه را کسب کند.



شکل ۵: مقایسه الگوریتم‌ها از نظر زمان محاسبات

اعتبارسنجی الگوریتم‌ها

در این پژوهش برای اعتبارسنجی الگوریتم‌ها، از شاخص اطلاعات متقابل نرمال^۱ (NMI) به عنوان شاخصی که میزان دقت الگوریتم‌ها را در شناسایی اجتماعات با کیفیت و درست از یک شبکه موردسنجش قرار می‌دهد، استفاده شده است. این شاخص میزان تشابه اجتماعات شناسایی شده توسط یک الگوریتم را با اجتماعات حقیقی یک شبکه اندازه‌گیری می‌کند.

1. Normalized Mutual Information (NMI)

فرض می‌کنیم دو افراز γ و λ ، افرازهایی از شبکه باشند. ماتریس C ، یک ماتریس درهم‌ریختگی است که در آن نشانگر تعداد گره‌هایی از اجتماع i در افراز γ است که در اجتماع j از افراز λ نیز حاضر هستند. شاخص اطلاعات متقابل نرمال $NMI(\gamma, \lambda)$ به صورت رابطه (۳) قابل محاسبه است (فورچتو، ۲۰۱۰):

$$NMI(\gamma, \lambda) = \frac{-2 \sum_{i=1}^{c_\gamma} \sum_{j=1}^{c_\lambda} C_{ij} \log(C_{ij}N / C_i.C_j)}{\sum_{i=1}^{c_\gamma} C_i \log(C_i / N) + \sum_{j=1}^{c_\lambda} C_j \log(C_j / N)} \quad (3)$$

در رابطه (۳)، $c_\gamma(c_\lambda)$ نشانگر تعداد گره‌های افراز $\gamma(\lambda)$ است. تعداد گره‌ها با N نمایش داده می‌شود. $C_i.(C_j)$ نشان‌دهنده مجموع درایه‌های ماتریس درهم‌ریختگی C در سطر i (ستون j) است. در صورتی که $\gamma = \lambda$ باشد، آنگاه مقدار $NMI(\gamma, \lambda)$ برابر یک است. اگر $NMI(\gamma, \lambda)$ برابر صفر باشد، آنگاه دو افراز γ و λ کاملاً متفاوت هستند. جدول (۹)، میانگین شاخص NMI برای ۴ الگوریتم پس از ده مرتبه حل مسئله است. با توجه به نتایج گزارش شده در جدول (۹)، الگوریتم پیشنهادی GAIWO توانسته است در تمامی شبکه‌ها، عملکرد بهتری نسبت به سایر روش‌ها داشته باشد.

جدول ۹: شاخص اطلاعات متقابل نرمال الگوریتم‌ها

شبکه‌ها	الگوریتم‌ها				
	GAIWO	GA	IWO	RS	FFA
Zachary	۰/۸۷۳	۰/۷۹۸	۰/۷۵۶	۰/۲۳۵	۰/۸۱۵
Dolphins	۰/۸۶۲	۰/۸۳۳	۰/۷۵	۰/۳۸۸	۰/۷۶۴
Football Club	۰/۹۰۴	۰/۸۵۸	۰/۷۳۶	۰/۵۱۱	۰/۸۰۱
Jazz	۰/۹۶۲	۰/۹۰۷	۰/۸۱۴	۰/۴۲۶	۰/۸۵۵
Les Misérables	۰/۷۳۲	۰/۶۵۷	۰/۶۳۲	۰/۲۷۶	۰/۷۰۶
PolBooks	۰/۸۸۳	۰/۷۱۹	۰/۶۴۶	۰/۳۸۱	۰/۷۵۳
WAN	۰/۹۱۷	۰/۸۴۰	۰/۵۶۷	۰/۲۵۵	۰/۸۳۷
Facebook	۰/۷۹۳	۰/۵۲۸	۰/۴۵۰	۰/۱۳۸	۰/۷۴۸

نتیجه‌گیری و پیشنهادهای پژوهش

در این نوشتار به اهمیت ساختارهای اجتماعی در شبکه‌های پیچیده اشاره و مسئله اجتماع یابی نیز در این شبکه‌های اجتماعی تعریف شد. با توجه به اهمیت کشف ساختارهای اجتماعی در شبکه‌های پیچیده، الگوریتمی فرا ابتکاری مرکب از دو روش بهینه‌سازی علف‌های هرز مهاجم و ژنتیک به منظور بهینه‌سازی هدف چگالی پودمانگی ارائه شد. در ادامه، کارایی الگوریتم پیشنهادی با چهار روش بهینه‌سازی علف‌های هرز مهاجم، ژنتیک، کرم شب‌تاب و یک الگوریتم جستجوی کاملاً تصادفی بر روی ۸ شبکه دنیای واقعی مورد مقایسه قرار گرفت. با توجه به نتایج به دست آمده می‌توان نتیجه گرفت که الگوریتم پیشنهادی توانسته است نسبت به چهار روش دیگر عملکرد مناسب‌تری داشته باشد. مطالعه مشابه دیگر در ارتباط با مسئله اجتماع یابی چندهدفه و یا بررسی کارایی سایر الگوریتم‌های بهینه‌سازی در حل این مسئله، می‌تواند زمینه‌ای برای تحقیقات آتی بشمار آید. استفاده از سایر شاخص‌های سنجش کیفیت اجتماعات در الگوریتم‌های فرا ابتکاری و مقایسه نتایج با این پژوهش از دیگر مواردی است که می‌تواند برای پژوهش‌های آینده مناسب باشد.

منابع

- بخشی، م.؛ سمیع‌زاده، ر. (۱۳۹۶)، مدلی برای پذیرش بانکداری الکترونیکی با در نظر گرفتن عامل اعتماد مشتریان، *مطالعات مدیریت کسب‌وکار هوشمند*، دوره ۵، شماره ۱۹، بهار ۱۳۹۶، صفحه ۵۳-۷۴.
- برادران، و؛ حسینیان، الف. ح؛ درخشانی، ر. (۱۳۹۷-الف)، ارائه روش فرا ابتکاری مبتنی بر تصمیم‌گیری چندمعیاره در حل مسئله اجتماع یابی، *مدیریت فناوری اطلاعات*، دوره ۱۰، شماره ۲، تابستان ۹۷، صفحه ۲۸۳-۳۰۸.
- برادران، و؛ حسینیان، الف. ح؛ درخشانی، ر؛ نیک‌ضمیر، م. (۱۳۹۷-ب)، ارائه یک رویکرد جدید برای حل مسئله اجتماع یابی شبکه‌های اجتماعی با توسعه الگوریتم‌های NSGA-II و NPGA، *مهندسی صنایع و مدیریت شریف*، دوره ۱، شماره ۱/۲، تابستان ۹۷، صفحات ۱۰۱-۱۱۵.
- جعفری، م.ب؛ کریمی، الف؛ ابرقوی‌زاده، ز. (۱۳۹۵)، عوامل تأثیرگذار بر تمایل به ادامه استفاده از وب‌سایت شبکه‌های اجتماعی، *مطالعات مدیریت کسب‌وکار هوشمند*، دوره ۵، شماره ۱۷، پاییز ۱۳۹۵، صفحه ۱۴۷-۱۸۲.
- روشنی، س؛ رضایی نیک، ن؛ شجاعی، م.ح. (۱۳۹۲)، مطالعه مقایسه‌ای قابلیت سازی و جامعه‌پذیری شبکه‌های اجتماعی عمومی و تخصصی، *مطالعات مدیریت کسب‌وکار هوشمند*، دوره ۲، شماره ۵، تابستان ۱۳۹۲، صفحه ۹۷-۱۳۲.
- سپهردوست، ح؛ صدری، ل. (۱۳۹۶)، اثر فناوری اطلاعات و ارتباطات بر رشد بازار سرمایه؛ شواهد تجربی از بورس اوراق بهادار تهران، *مطالعات مدیریت کسب‌وکار هوشمند*، دوره ۵، شماره ۱۹، بهار ۱۳۹۶، صفحه ۱-۲۸.
- Bingol, H. & Tasgin, M. (2006). Community detection in complex networks using genetic algorithms. *Advances in Complex Systems*, 11(4), 1-6.
- Brandes, U., Delling, D. & Gaetler, M. (2008). On Modularity Clustering. *Transactions on Knowledge and Data Engineering*, 20(2), 172-188.
- Chen, M., Kuzmin, K., Boleslaw, K., & Szymanski, F. (2014). Community Detection via Maximization of Modularity and Its Variants. *Trans. Computation Social System*, 1(1), 46-65.

- Choudhury, D. & Paul, A. (2013). Community Detection in Social Networks: An Overview. *International Journal of Research in Engineering and Technology*, 2(2), 6-13.
- Fortunato, S. & Barthelemy, M. (2007). Resolution limit in community detection. *PNAS*, 104(1), 36-41.
- Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3), 1-100.
- Ghorbanian, A. & Shafaqi, B. (2015). A Genetic Algorithm for Modularity Density Optimization in Community Detection. *International Journal of Economy, Management and Social Sciences*, 4(1), 117-122.
- Gleiser, P. & Danon, L. (2003). Community Structure in Jazz. *Advances in Complex Systems*, 6(4), 565-573.
- Guesmi, S., Trabelsi, C., & Latiri, C., (2016). Community detection in multi-relational bibliographic networks, *Database and Expert Systems Applications*, vol. 9828 of Lecture Notes in Computer Science, 11-18, Springer International Publishing, Cham, Switzerland, 2016.
- Guimera, R. & Amaral, L. (2005). Functional Cartography of Complex Metabolic Networks. *Nature*, 433(2), 895-900.
- Guoqiang, C. & Xiaofang, G. (2010). A Genetic Algorithm Based on Modularity Density for Detecting Community Structure in Complex Networks. *Computational Intelligence and Security*, 20(4), 151-154.
- Griechisch, E. & Pluhar, A. (2011). Community Detection by using the Extended Modularity. *Acta Cybernetica*, 20(1), 69-85.
- Hafez, A., Ghali, N., Hassanien, A. & Fahmy, A. (2012). Genetic Algorithms for community detection in social networks. *Intelligent Systems Design and Applications*, 10(2), 460-465.
- Hosseinian, A.H., & Baradaran, V. (2018). A multi-objective multi-agent optimization algorithm for the community detection problem. *Journal of Information Systems and Telecommunication*, 6(3), 166-176.
- Li, Z., Pan, Z., Zhang, Y., Li, G., & Hu, G., (2016). Efficient Community Detection in Heterogeneous Social Networks, *Mathematical Problems in Engineering*, Volume 2016, <http://dx.doi.org/10.1155/2016/5750645>.
- Mahmood, A., & Small, M., (2016). Subspace based network community detection using sparse linear coding, *IEEE Transactions on Knowledge & Data Engineering*, 28(3), 801-812.
- Mehrabian, A. & Lucas, C. (2006). A novel numerical optimization algorithm inspired from weed colonization. *Ecological Informatics*, 1(4), 355-366.
- Miller, B. & Goldberg, D. (1995). Genetic Algorithms, Tournament Selection, and the Effects of Noise. *Complex Systems*, 9(1), 193-212.
- Newman, M. & Girvan, M. (2004). Finding and evaluating community structure in networks. *Phys. Rev.*, 69(2), 22-38.

- Newman, M. (2006). Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev*, 1(3), 12-34.
- Peel, L., Larremore, D.B., & Clauset, A., (2017). The ground truth about mega-data and community detection in networks, *Science Advances*, 3(5), 1-8, DOI: 10.1126/sciadv.1602548.
- Pizzuti, C. (2008). GA-Net: A Genetic Algorithm for Community Detection in Social Networks. *Computer Science*, 5199(1), 1081-1090.
- Shaqqi, B., Teimourpour, B. & Ghorbanian, A. (January, 2015). A new heuristic algorithm for modularity optimization in complex networks community detection. *Proceedings of 11th Industrial Engineering Conference*, Tehran, Iran.
- Shaqqi, B. (2014). A Mathematical Programming Model based on modularity density for community detection. M.Sc dissertation, Tarbiat Modares University, Faculty of Engineering.
- Shi, C., Yan, Z., Wang, Y., Cai, Y. & Wu, B. (2010). A Genetic Algorithm for Detecting Communities in Largescale Complex Networks. *Advance in Complex System*, 13(1), 3-17.
- Shi, C., Yan, Z., Cai, Y. & Wu, B. (2012). Multi-objective community detection in complex networks. *Applied Soft Computing*, 12(2), 850-859.
- Win, H.N., & Lynn, K.T. (2017). Community detection in Facebook with outlier recognition. *18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, Kanazawa, Japan.
- Zhang, H., Qiut, B., Giles, L., Foley, H. & Yen, J. (2007). An LDA-based Community Structure Discovery. *Intelligence and Security Informatics*, 400(2), 200-207.
- Zhang, S. & Li, Z. (2008). Quantitative function for community detection. *Physical Review*, 77(3), 036109.
- Zhang, W., Pan, G., Wu, Z., & Li., S. (2014). Online Community Detection for Large Complex Networks. *PLoS ONE*, 9(7), 168-188.