Journal of Information Technology Management
ISSN: 2008-5893
EISSN: 2423-5059
Summer 2018
Vol. 10, No. 2, PP. 283-308
DOI: 10.22059/jitm.2017.223145.1896

University of Tehran
Faculty of Management

Citation: Baradaran, V., Hosseinian, A.H., Derakhshani, R. (2018). A New Meta-heuristic Algorithm based on Multi-criteria Decision Making to Solve Community Detection Problem. *Journal of Information Technology Management,* 10(2), 283-308.

A New Meta-heuristic Algorithm based on Multi-criteria Decision Making to Solve Community Detection Problem

Vahid Baradaran [1], Amir Hossein Hosseinian [2], Reza Derakhshani [3]

Abstract: Community detection is one of the most significant issues in the field of social networks. The main purpose of community detection is to partition the network in such a way that the relations between components of the network are dense. Because of the strong relations among network members in these partitions, you can consider them as a community. Many researchers have developed several algorithms to solve such a problem. Therefore, we present a genetic algorithm based on Topsis which is a multi-criteria decision making method (MCDM). The proposed algorithm uses Topsis to rank solutions based on modularity and modularity density which are two of the most well-known criteria in community detection problem. Thereafter, crossover and mutation operators are only applied on solutions ranked by Topsis. The performance of the proposed algorithm has been evaluated through comparing it against classical genetic algorithm and a greedy one. The results showed that the proposed algorithm outperforms the other two methods. Since the application of MCDM approach has not been reported in the related literature, this paper can be considered as a basis for future studies.

Key words: ***Community detection, Genetic algorithm, Optimization, Social networks, TOPSIS.***

1. Assistant Prof. of Industrial Engineering, Islamic Azad University, North Tehran Branch, Iran
2. Ph.D. Candidate of Industrial Engineering, Islamic Azad University, North Tehran Branch, Iran
3. Ph.D. Candidate of Industrial Engineering, Islamic Azad University North Tehran Branch, Iran

Submitted: 21 / December / 2016
Accepted: 24 / September / 2017
Corresponding Author: Vahid Baradaran
Email: v_baradaran@iau-tnb.ac.ir

مدیریت فناوری اطلاعات
دانشکدهٔ مدیریت دانشگاه تهران
دورهٔ ۱۰، شمارهٔ ۲
تابستان ۱۳۹۷
صص. ۳۰۸- ۲۸۳

Journal of Information Technology Management
DOI: 10.22059/jitm.2017.223145.1896

ارائهٔ روش فراابتکاری مبتنی بر تصمیم‌گیری چندمعیاره در حل مسئلهٔ اجتماع‌یابی

وحید برادران[۱]، امیرحسین حسینیان[۲]، رضا درخشانی[۳]

چکیده: یکی از مسائل مهم در زمینهٔ شبکه‌های اجتماعی، مسئلهٔ اجتماع‌یابی است. هدف اجتماع‌یابی، افرازبندی شبکه به بخش‌هایی است که ارتباط میان اعضای شبکه در این نواحی متراکم است. به‌دلیل ارتباط پررنگ اعضای شبکه در این بخش‌ها، می‌توان این اعضا را متعلق به یک اجتماع دانست. تحقیقات بسیاری به توسعهٔ الگوریتم‌های متعدد در حل این مسائل پرداخته‌اند. در این پژوهش نیز یک الگوریتم ژنتیک مبتنی بر روش تصمیم‌گیری چندمعیارهٔ تاپسیس ارائه شده است. در الگوریتم پیشنهادی، عملگرهای تقاطع و جهش بر جواب‌هایی اعمال می‌شوند که روش تاپسیس به‌دست می‌دهد. برازندگی جواب‌ها بر اساس شاخص‌های پودمانگی و چگالی پودمانگی مشخص می‌شود. در ادامهٔ این مقاله، کارایی الگوریتم پیشنهادی نسبت به الگوریتم ژنتیک کلاسیک و یک الگوریتم حریصانه از طریق انجام آزمایش‌های عددی سنجیده شد که نتایج نشان از برتری الگوریتم پیشنهادی داشت.

واژه‌های کلیدی: اجتماع‌یابی، الگوریتم ژنتیک، بهینه‌سازی، تاپسیس، شبکه‌های اجتماعی.

۱. استادیار گروه مهندسی صنایع، دانشکدهٔ فنی و مهندسی، دانشگاه آزاد اسلامی واحد تهران شمال، تهران، ایران
۲. دانشجوی دکتری مهندسی صنایع، دانشکدهٔ فنی و مهندسی، دانشگاه آزاد اسلامی واحد تهران شمال، تهران، ایران
۳. دانشجوی دکتری مهندسی صنایع، دانشکدهٔ فنی و مهندسی، دانشگاه آزاد اسلامی واحد تهران شمال، تهران، ایران

تاریخ دریافت مقاله: ۱۳۹۵/۱۰/۰۱
تاریخ پذیرش نهایی مقاله: ۱۳۹۶/۰۷/۰۲
نویسندهٔ مسئول مقاله: وحید برادران
E-mail: v_baradaran@iau-tnb.ac.ir

مقدمه

علوم نوین مرتبط با شبکه‌ها، به پیشرفت‌های شایان توجهی در درک سیستم‌های پیچیده منجر شده است. سیستم‌های پیچیده را می‌توان به شکل شبکه‌های پیچیده و گراف مدل‌سازی کرد (شقاقی و تیمورپور، ۲۰۱۵). شبکهٔ مسیرها و شبکهٔ استناد مقالات و اینترنت، از جملهٔ این شبکه‌هاست. گره‌های شبکه و یال‌های متصل‌کنندهٔ آنها، دو عنصر اصلی در گراف‌ها به‌شمار می‌روند. یکی از ویژگی‌های حائز اهمیت گراف‌ها که نشان‌دهندهٔ سیستم‌های دنیای واقعی هستند، ساختار اجتماع است. مطالعه و بررسی ساختار درونی شبکه‌های پیچیده، به درک نحوهٔ عملکرد آنها کمک می‌کند. ساختار اجتماع بیان‌کنندهٔ این است که ارتباط داخلی گره‌های یک خوشه یا اجتماع، به مراتب چگال‌تر از ارتباط آنها با سایر گره‌های شبکه است.

ساختارهای اجتماعی حاوی اطلاعات مفید و ارزشمندی برای اهداف تجاری و علمی هستند (فورتوناتو، ۲۰۱۰). شناسایی ساختارها در سیستم‌هایی که می‌توان آنها را به‌صورت گراف نمایش داد، کاربرد فراوانی دارد. از این حیث، شناسایی اجتماعات در علوم اجتماعی، زیست‌شناسی و رایانه کاربرد پیدا می‌کند و با عنوان مسئلهٔ اجتماع‌یابی شناخته می‌شود (ژانگ، کیوت، جیلز، فُلِی و بِن، ۲۰۰۷). مسئلهٔ اجتماع‌یابی می‌تواند در زمینهٔ داده‌کاوی، بازیابی اطلاعات و تحلیل شبکه‌های آنلاین فروش، کاربرد داشته باشد (چِن، کازمین، بولِسلا و سیزمانسکی، ۲۰۱۴). مسئلهٔ اجتماع‌یابی از جمله مباحث پایه‌ای در تحلیل شبکه‌های اجتماعی است که زمینهٔ پژوهش بسیاری از محققان در سال‌های اخیر بوده است (پان، ژانگ، وُو و لِی ۲۰۱۴). یک اجتماع زیرمجموعه‌ای از گره‌های شبکه است که این گره‌ها نسبت به سایر رئوس شبکه، ارتباط بیشتری با هم دارند (چادهِری و پاول، ۲۰۱۳). گره‌های یک اجتماع از خصوصیات مشترکی برخوردارند. برای نمونه، در شبکه‌های اجتماعی این خصوصیات مشترک می‌تواند شامل علایق مشابه یا منطقهٔ جغرافیایی یکسان باشد (آگاروال و کمپ، ۲۰۰۸). به‌طور کلی در مسئلهٔ اجتماع‌یابی به‌دنبال یافتن افرازِ با کیفیت از رئوس هستیم که این رئوس از نظر ویژگی‌های مختلف با یکدیگر اشتراک دارند. کیفیت افرازهای گوناگون از شبکه توسط معیارهای مختلفی سنجیده می‌شود که از آن جمله می‌توان به شاخص پودمانگی[1] اشاره کرد. این شاخص نخستین‌بار توسط نیومن ارائه شد (فورتوناتو، ۲۰۱۰). پودمانگی، کیفیت افراز گره‌های یک شبکه به اجتماعات مختلف را تعیین می‌کند (شقاقی و تیمورپور، ۲۰۱۵). به این ترتیب از این معیار به‌صورت گسترده‌ای برای آزمون کیفیت ساختارهای اجتماعی به‌دست آمده از الگوریتم‌های اجتماع‌یابی استفاده می‌شود. هرچه پودمانگی اجتماعات کشف‌شده توسط الگوریتم‌های اجتماع‌یابی زیادتر باشد، کیفیت اجتماعات

---

1. Modularity

شناسایی شده بیشتر است. بنابراین از آنجا که در فرایند اجتماع‌یابی به‌دنبال یافتن افرازی از شبکه هستیم که مقدار بهینه‌ای از پودمانگی دارد، مسئلهٔ یاد شده از جمله مسائل NP-Hard محسوب می‌شود (برَندِس، دلینگ و گاتلر، ۲۰۰۸). اما با توجه به زمان‌بر بودن جست‌وجوی تمام افرازهای ممکن از شبکه، روش‌های ابتکاری و فراابتکاری متعددی همچون الگوریتم‌های حریصانه[1]، روش‌های نمونه‌برداری، شبیه‌سازی تبرید[2] و بهینه‌سازی اکسترمال[3]، برای حل مسئله به‌کار گرفته شده‌اند (گوئیمِرا و آمارال، ۲۰۰۵).

ساختار اجتماعی در مسئلهٔ یاد شده می‌تواند در تحلیل شبکه‌های اجتماعی، داده‌کاوی، بازیابی اطلاعات، شبکه‌های مشارکت علمی میان پژوهشگران و شبکه‌های برخط[4] خرده‌فروشی متشکل از مشتریانی که سود و منفعت مشترکی دارند، کاربرد داشته باشد (چِن و همکاران، ۲۰۱۴). خوشه‌بندی[5] مشتریانی که علایق مشترک دارند و از نظر موقعیت جغرافیایی نزدیک به یکدیگرند، می‌تواند به بهبود ارائهٔ خدمات منجر شود. شناسایی اجتماعاتی از مشتریان که علایق مشابه دارند، امکان پیشنهاد محصولاتی که متناسب با علایق آنهاست را فراهم می‌آورد و از این جنبه، فرصت‌های تجاری مناسبی به‌وجود می‌آید. در علم رایانه نیز شناسایی اجتماعات یکی از موضوعات حائز اهمیت است. برای نمونه، در مبحث پردازش موازی[6]، دانستن بهترین روش برای تخصیص عملیات به پردازنده‌ها به‌گونه‌ای که ارتباط میان آنها کمینه شود و فرایند محاسبات به سرعت انجام شود، اهمیت زیادی دارد. این امر زمانی محقق می‌شود که یک اجتماع از رایانه‌ها به گروه‌هایی هم اندازه از پردازشگرها تقسیم شوند، به‌طوری که تعداد ارتباطات فیزیکی میان پردازنده‌ها در گروه‌های مختلف کمینه باشد (فورتوناتو ، ۲۰۱۰). اجتماع‌یابی در شکل‌دهی اجتماعات در شبکه‌های استنادی تحقیقات نیز می‌تواند موارد استفادهٔ فراوان داشته باشد. تشکیل اجتماعات در شبکه‌های استنادی بر پایهٔ موضوع تحقیق صورت می‌پذیرد (چِن و همکاران، ۲۰۱۴).

با توجه به پیچیدگی مسئلهٔ اجتماع‌یابی، هدف از این مطالعه، ارائهٔ روش فراابتکاری مبتنی بر الگوریتم ژنتیک است که به‌واسطهٔ آن بتوان اجتماعات متمایزی از رئوس گراف‌های یک شبکهٔ پیچیده را در زمان اندک شکل داد.

---

1. Greedy Algorithms
2. Simulated Annealing (SA)
3. Extremal Optimization (EA)
4. Online
6. Clustering
7. Parallel Computing

در این پژوهش، شاخص پودمانگی و نیز شاخص چگالی پودمانگی[1] که از جمله معیارهای ارزیابی کیفیت اجتماعات است، به‌عنوان تابع هدف مسئله در نظر گرفته شده است و به‌منظور بیشینه‌کردن این دو شاخص، یک الگوریتم ژنتیک مبتنی بر روش تصمیم‌گیری چندمعیارهٔ تاپسیس ارائه می‌شود. با توجه آن که تا کنون از روش‌های تصمیم‌گیری چندمعیاره در حل مسئله اجتماع‌یابی استفاده نشده است، این نوشتار می‌تواند پایه‌ای برای مطالعات آتی در این زمینه باشد. ساختار مقاله در ادامه به این ترتیب است؛ به‌منظور شناخت و آشنایی بیشتر با مسئلهٔ اجتماع‌یابی، در بخش دوم به مرور تحقیقات مرتبط پرداخته می‌شود. بخش سوم به تشریح کامل مسئله اختصاص دارد. بخش چهارم به معرفی الگوریتم پیشنهادی می‌پردازد و روش اجرای پژوهش در بخش پنجم مطرح می‌شود. در بخش ششم، عملکرد الگوریتم پیشنهادی با تعریف معیارهای متفاوت سنجیده شده است. بخش هفتم نیز به جمع‌بندی مطالب و ارائهٔ پیشنهادهایی برای پژوهش‌های بعدی اختصاص یافته است.

پیشینهٔ پژوهش

شبکه‌های اجتماعی از جمله شبکه‌های پیچیده‌ای هستند که شامل تعداد زیادی از اعضا و ارتباطات بین آنها می‌شوند. یک شبکهٔ اجتماعی می‌تواند در قالب گرافی مدل شود که رئوس آن نمایندهٔ افراد و یال‌های آن نشان‌دهندهٔ ارتباطات موجود بین افراد است (کی‌پور، براری و شیرازی، ۱۳۹۳). تا کنون مطالعات متعددی در زمینهٔ تحلیل شبکه‌های اجتماعی صورت گرفته است. حقیقی و منتظر (۱۳۹۴) در پژوهش خود به ارائهٔ مدلی از عوامل مؤثر بر اعتماد و رضایت کاربران در شبکه‌های اجتماعی برخط پرداختند. در این مطالعه، از روش الکترهٔ فازی برای رتبه‌بندی عوامل اعتماد کاربران استفاده شده است. در مطالعهٔ دیگری، ارتباطات نویسندگان و محققان به‌واسطهٔ همکاری در انتشار مقاله در پایگاه اطلاعات علمی جهاد دانشگاهی، پایگاه استنادی علوم جهان اسلام، پایگاه اطلاعات علمی سیویلیکا[2] و پایگاه اطلاعات علمی ساینس دایرکت[3]، به‌صورت یک شبکهٔ اجتماعی بزرگ مدل شده است (رضایی نور، لسانی، زکی‌زاده و صفامجید، ۱۳۹۳). اطلاعات بر مبنای همکاری محققان طی ۱۰ سال گذشته گردآوری شده است. در این تحقیق، بررسی ارتباطات میان پژوهشگران توسط نرم‌افزار تحلیل شبکه‌های اجتماعی پاجک[4] از نظر دو دیدگاه متنی و کلی صورت پذیرفته است. همچنین ارتباطات ایجادشده میان دانشگاه‌ها و

---

1. Modularity Density
2. Civilica
3. Science Direct
4. Pajek

صنعت که از طریق همکاری محققان با آنها به‌وجود آمده، تجزیه و تحلیل شده است (رضایی نور و همکاران، ۱۳۹۳). فتحیان و حسینی (۱۳۹۳) به بررسی چگونگی سودرسانی اجتماعات مجازی به کسب‌وکارها و بنگاه‌های تجاری پرداختند. اجتماعات مجازی در این پژوهش در قالب یک مدل کسب‌وکار در نظر گرفته شده است که می‌تواند نقش شایان توجهی در فرایند تصمیم‌گیری خرید مشتریان داشته باشد. کی‌پور و همکارانش (۱۳۹۳) روشی برای پیش‌بینی ارتباط بین رئوس موجود در شبکه‌های اجتماعی ارائه کردند. روش پیش‌بینی ارتباطات در این مقاله، مبتنی بر رویکرد محلی بوده و سرعت قابل قبولی دارد. کارایی روش پیشنهادی در این پژوهش، بر مجموعهٔ داده‌های شبکه‌های اجتماعی Facebook و Epinions سنجیده شده است.

حجم فراوان داده‌های موجود در شبکه‌های دنیای واقعی، ضرورت کارایی و دقت زیاد الگوریتم‌های خوشه‌بندی را نمایان می‌کند. پیچیدگی محاسباتی یک الگوریتم، با برآورد میزان منابع مورد نیاز آن در انجام یک فعالیت برآورد می‌شود (فورتوناتو، ۲۰۱۰). در این برآورد، تعداد گام‌های محاسباتی لازم و نیز، تعداد واحدهای حافظهٔ لازم که به‌صورت همزمان باید به انجام محاسبات تخصیص داده شوند، لحاظ می‌شود. این نیازها معمولاً متناسب با اندازهٔ سیستمِ در دست مطالعه تنظیم می‌شوند. اندازهٔ یک گراف با تعداد رئوس (N) یا تعداد یال‌های آن (E) مشخص می‌شود. اگر α و β به‌ترتیب نشان‌دهندهٔ توان‌های تعداد رئوس و یال‌های یک گراف باشند، $O(N^{\alpha}E^{\beta})$ نشان می‌دهد که زمان لازم برای انجام محاسبات با افزایش تعداد رئوس و یال‌ها افزایش خواهد یافت. هرچه میزان α و β کمتر باشد، زمان محاسبات کاهش خواهد یافت. برای مثال، نمی‌توان با استفاده از الگوریتم‌هایی که زمان اجرای آنها بیشتر از $O(N)$ یا $O(E)$ است، گراف‌های شبکهٔ وب جهانی که شامل میلیون‌ها گره و میلیاردها یال هستند را بررسی و تحلیل کرد. بسیاری از مسائل مرتبط با خوشه‌بندی، در دسته مسائل NP-hard قرار می‌گیرند (فورتوناتو، ۲۰۱۰). از این رو، استفاده از الگوریتم‌های دقیق که تنها در سیستم‌های بسیار کوچک کاربرد دارند، بیهوده است. همچنین، حتی اگر الگوریتمی دارای پیچیدگی چندجمله‌ای باشد، ممکن است در تحلیل و بررسی شبکه‌های بزرگ، بسیار کند عمل کند. در چنین مواردی، استفاده از الگوریتم‌هایی که در زمان مناسب جواب‌های تقریبی می‌دهند، بسیار رایج است. الگوریتم‌های تقریبی معمولاً برای مسائل بهینه‌سازی با ابعاد بزرگ به‌کار می‌روند که در این مسائل، یافتن بیشترین یا کمترین مقدار یک تابع ارزیابی حائز اهمیت است. بنابراین از آنجا که شبکه‌های دنیای واقعی اغلب ابعاد وسیعی دارند و مسئلهٔ اجتماع‌یابی نیز در دسته مسائل NP-hard قرار می‌گیرد (بِرَندِس و همکاران، ۲۰۰۸)، استفاده از هوش محاسباتی در تجزیه و تحلیل این شبکه‌ها می‌تواند نقش کلیدی داشته باشد. به بیان دیگر، اهمیت دقت الگوریتم‌های اجتماع‌یابی و سرعت

عملکرد آنها در شناسایی ساختارهای اجتماعی در شبکه‌های پیچیدهٔ دنیای واقعی، محققان را بر آن داشته که از هوش محاسباتی بهره ببرند.

با توجه به اهمیت مسئلهٔ اجتماع‌یابی در شبکه‌های اجتماعی، تا کنون روش‌های فراوانی برای تحلیل ساختار شبکه‌های پیچیده معرفی شده و توسعه یافته‌اند. الگوریتم گیروان ـ نیومن از ابتدایی‌ترین روش‌های معرفی شده در ادبیات موضوع است. استفاده از این روش در بهینه‌کردن شاخص پودمانگی گزارش شده است (نیومن و گیروان، ۲۰۰۴). پودمانگی نخستین‌بار توسط نیومن و گیروان (۲۰۰۴) معرفی شد و مسئلهٔ اجتماع‌یابی در قالب یک مسئلهٔ بهینه‌سازی پودمانگی مد نظر قرار گرفت. شاخص پودمانگی کیفیت افراز خاصی از شبکه را مشخص می‌کند. هرچه میزان پودمانگی اجتماع به‌دست آمده توسط روش حل نزدیک به ۱ باشد، ساختار اجتماع کشف شده کیفیت بیشتری خواهد داشت (گریشیچ و پلوهار، ۲۰۱۱). شقاقی و تیمورپور (۲۰۱۵) به‌منظور بهینه‌سازی پودمانگی، یک الگوریتم حریصانهٔ قدرتمند و ساده‌ای را ارائه کردند. رویکرد حریصانهٔ دیگری توسط بلوندل و همکاران (۲۰۰۸) برای گراف‌های موزون ارائه شده است. این رویکرد ابتدا گره‌های گراف را در اجتماعات مختلف قرار داده و میزان پودمانگی را محاسبه می‌کند. در گام بعد، بررسی می‌شود که در صورت انتقال گره‌ها به اجتماعات همسایگی، پودمانگی به چه میزان افزایش می‌یابد. انتقالی که به بیشترین افزایش پودمانگی بینجامد، پذیرفته می‌شود. روند تغییر اجتماعاتی که گره‌ها به آنها تعلق دارند، در جهت بیشینه‌سازی پودمانگی ادامه می‌یابد. گوئیمرا و آمارال (۲۰۰۵) نخستین‌بار از الگوریتم شبیه‌سازی تبرید به‌منظور بهینه‌سازی پودمانگی استفاده کردند. این محققان در روند جست‌وجوی همسایگی الگوریتم پیشنهادی خود، دو حرکت محلی[1] و سراسری[2] را گنجاندند. در حرکت محلی، هر گره به‌صورت تصادفی از یک اجتماع به اجتماعی دیگر منتقل می‌شود. در حرکت سراسری، اجتماعات کشف شده با یکدیگر ادغام شده یا به اجتماعات کوچک‌تر دسته‌بندی می‌شوند. ماسِن و دوی (۲۰۰۵) در تحقیق خود نشان دادند استفاده از حرکت سراسری، ریسک گرفتارشدن الگوریتم شبیه‌سازی تبرید در بهینهٔ محلی را کاهش می‌دهد. همچنین اثبات کردند که استفاده از حرکت سراسری سبب دستیابی به جواب‌های بهتری نسبت به حرکت محلی می‌شود. بهینه‌سازی اکسترمال 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 نیازمند زمان محاسبات کمتری است (بوئتچِر و پِرکوس، ۲۰۰۱). دوچ و آرِناس (۲۰۰۵) این روش را برای بهینه‌سازی پودمانگی به‌کار برده‌اند.

---

1. Local
2. Global

کاربرد الگوریتم ژنتیک نیز به‌عنوان روشی برای بهینه‌سازی توابع هدف مختلف از مسئله اجتماع‌یابی در ادبیات موضوع گزارش شده است. مزیت ویژهٔ این الگوریتم، به‌دست آمدن تعداد اجتماعات طی روند تکاملی حل مسئله است، در نتیجه دانستن تعداد اجتماعات پیش از روند حل مسئله الزامی ندارد (حافظ، قالی، حسنین و فهمی، ۲۰۱۲). پیزوتی (۲۰۰۸) الگوریتمی با عنوان GA-Net ارائه کرد که شاخص امتیاز اجتماعات[1] را بهینه می‌کرد. ساختار نمایش جواب در این تحقیق از نوع ژنتیکی لوکاسی است. قربانیان و شقاقی (۲۰۱۵) نیز، نوعی ساختار نمایش جواب مبتنی بر ماتریس ارائه کردند و مسئلهٔ اجتماع‌یابی را با هدف بیشینه‌سازی چگالی پودمانگی مد نظر قرار دادند. شی، یان، وانگ، کای و وُو (۲۰۱۰) یک الگوریتم ژنتیک همراه با یک برنامهٔ کدگذاری لوکاسی ویژه[2] معرفی کردند که بر اساس آن فرایند اجتماع‌یابی با انعطاف بیشتری انجام می‌شود. ترکیب الگوریتم ژنتیک با سایر روش‌های ابتکاری و فراابتکاری نیز برای حل مسئلهٔ اجتماع‌یابی مطالعه شده است. شانگ، باو، جیاو و جین (۲۰۱۳)، الگوریتمی متشکل از ژنتیک و شبیه‌سازی تبرید برای افزایش اثربخشی و کارایی جست‌وجو در فضای جواب ارائه کردند.

علاوه بر معرفی الگوریتم‌های متنوع در مسئلهٔ اجتماع‌یابی، توابع هدف گوناگونی نیز برای درک ساختار اجتماعات در شبکه پیشنهاد شده است. از جمله توابع هدف مختلفی که در کنار شاخص پودمانگی در تحقیقات گذشته مطالعه شده‌اند، می‌توان به چگالی پودمانگی، شاخص امتیاز اجتماعات، برازش اجتماعات، معیار وسعت[3] و رسانایی[4] شبکه اشاره کرد. لِسکوویک، لانگ و ماهونی (۲۰۱۰) برخی از شناخته‌شده‌ترین توابع هدف این مسئله را در پژوهشی به‌صورت اجمالی معرفی کردند.

اگرچه استفاده از توابع هدف متنوع در ادبیات موضوع گزارش شده، شاخص پودمانگی اصلی‌ترین هدف در تحقیقات مرتبط با اجتماع‌یابی است. با این که توجه گسترده‌ای به شاخص پودمانگی می‌شود، این معیار عاری از نقص نیست و در ادبیات موضوع به دو اشکال شایان توجه در ارتباط با پودمانگی اشاره شده است (بِرَندِس و همکاران، ۲۰۰۸):

۱. در برخی موارد این معیار تمایل به تقسیم اجتماع‌های بزرگ به دو یا چند اجتماع کوچک دارد.
۲. در موارد دیگری این شاخص تمایل دارد با ترکیب اجتماع‌هایی که از یک حد آستانه‌ای کوچک‌ترند، اجتماع‌های بزرگ شکل دهد.

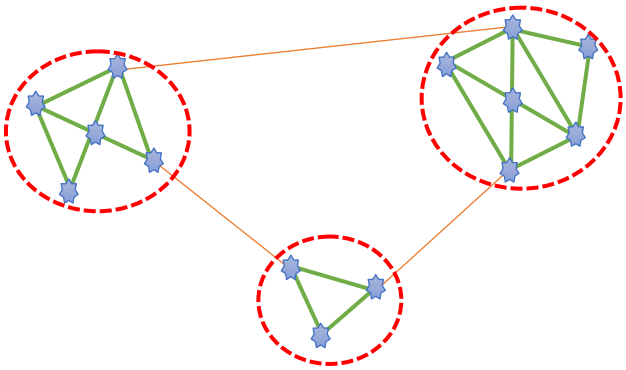
---

1. Community Score
2. Locus-based Solution Representation
3. Expansion
4. Conductance

فورتوناتو و بارتلِمی (۲۰۰۷) اثبات کردند که پودمانگی ممکن است در تشخیص اجتماعاتی که کوچک‌تر از یک مقیاس مشخص هستند، ناموفق باشد. این مقیاس به اندازهٔ کل شبکه و به درجهٔ اتصال اجتماعات بستگی دارد. لِی، ژانگ، وانگ، ژانگ و چِن (۲۰۰۸) تابع کمّی دیگری به نام چگالی پودمانگی ارائه کردند. آنها نشان دادند این شاخص به‌صورت همزمان دچار دو مشکل یاد شده نمی‌شود. بنابراین با توجه به توضیحاتی که بیان شد، در این مقاله شاخص‌های پودمانگی و چگالی پودمانگی که از جمله شاخص‌های معتبر در شکل‌دهی اجتماعات هستند، به‌عنوان معیارهای تشکیل اجتماعات مد نظر قرار گرفته‌اند. همان‌طور که در ادبیات موضوع اشاره شد، بهینه‌کردن این دو شاخص می‌تواند به ایجاد اجتماعاتی همگن و با کیفیت منجر شود. از سوی دیگر، تا کنون در ادبیات موضوع روشی برای بهینه‌کردن همزمان این دو شاخص در مدت زمان معقول پیشنهاد نشده است. بدین سبب، در این نوشتار تلاش شده است با ارائهٔ روشی فراابتکاری برای دستیابی به جواب‌هایی با کیفیت مناسب و در زمانی معقول گام برداشته شود. از سوی دیگر، با وجود ارائه و توسعهٔ الگوریتم‌های مختلف در حل مسئلهٔ اجتماع‌یابی، تاکنون استفاده از رویکردی مبتنی بر تصمیم‌گیری چندگانه برای این مسئله گزارش نشده است. از این حیث، نیاز به بررسی کارایی الگوریتم‌های مبتنی بر روش‌های تصمیم‌گیری چندگانه در حل مسئله اجتماع‌یابی احساس می‌شود. بنابراین با توجه به توضیحاتی که داده شد، در این پژوهش مسئلهٔ اجتماع‌یابی با شاخص‌های بیشینه‌سازی پودمانگی و چگالی پودمانگی در نظر گرفته می‌شود و یک الگوریتم ژنتیک مبتنی بر روش تاپسیس معرفی خواهد شد.

تعریف مسئله

یک شبکهٔ اجتماعی را می‌توان به‌صورت گراف $G(N,E,W)$ نشان داد که در آن N رئوس گراف را نشان می‌دهد و متناسب با هر فرد در شبکه است. E معرف یال‌هایی است که ارتباطات دو طرفه بین افراد را بر قرار می‌کند و W ماتریس وزن شبکه است. در شبکه‌های غیروزنی مؤلفه‌های ماتریس W برابر ۱ است (قربانیان و شقاقی، ۲۰۱۵). همان‌طور که در بخش‌های قبل بیان شد، یک اجتماع یا خوشه، مجموعه‌ای از گره‌های شبکه را شامل می‌شود که تراکم یال‌های متصل‌کنندهٔ آنها زیاد است. در عین حال، چگالی یال‌ها میان اعضای اجتماع و رئوسی که در داخل اجتماع قرار ندارند، پایین است. همان‌طور که در شکل ۱ مشاهده می‌شود، هدف اجتماع‌یابی را می‌توان شناسایی مجموعه‌ای از گره‌های شبکه دانست که به‌صورت نسبی ارتباط آنها در درون اجتماع، چگال‌تر از گره‌های بیرون اجتماع است.



شکل ۱. ارتباط متراکم اعضای یک اجتماع با یکدیگر

الگوریتم پیشنهادی GA-Topsis

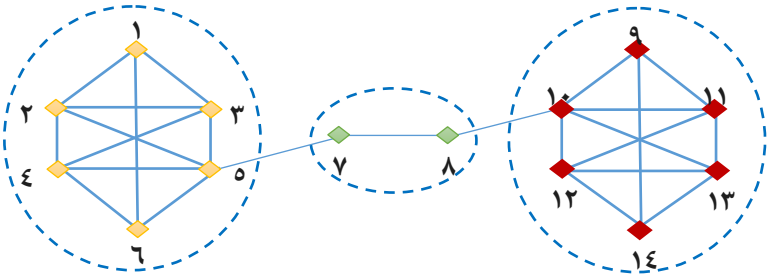
روش پیشنهادی در این نوشتار توسعه‌ای بر الگوریتم فراابتکاری ژنتیک است. الگوریتم ژنتیک از جمله الگوریتم‌های تکاملی است که به‌منظور یافتن راه حل تقریبی در مسائل بهینه‌سازی استفاده می‌شود (بهشتی‌نیا و فرازمند، ۱۳۹۴؛ اصغری اسکویی و قاسم‌زاده، ۱۳۹۵). این روش بر اساس اصول ژنتیک و تکامل توسعه داده شده است و مبتنی بر جمعیتی از جواب‌ها عمل می‌کند (بهشتی‌نیا و قهرمانی، ۱۳۹۵؛ تقوی‌فرد، سادات حسینی و خان‌بابایی، ۱۳۹۲). این الگوریتم فرایند حل مسئله را با جمعیتی اولیه از جواب‌ها آغاز می‌کند و به مرور تکامل داده می‌شوند (فتحیان و حسینی، ۱۳۹۳). فرایند تکامل جواب‌ها در الگوریتم ژنتیک، الهام گرفته از انتقال خصوصیات موروثی از طریق تکنیک‌های زیست‌شناسی مانند جهش و وراثت است (کاباران‌زاده قدیم و آستانه، ۱۳۸۸؛ البرزی، پورزرندری و خان‌بابایی، ۱۳۸۹). اجزای مختلف الگوریتم پیشنهادی در بخش‌های بعدی به تفصیل شرح داده شده است.

ساختار نمایش جواب

نحوهٔ نمایش جواب‌ها در الگوریتم‌های فراابتکاری نقش بسزایی در کارایی و کیفیت عملکرد آنها دارد (الغزالی، ۲۰۰۹). در این پژوهش از ساختار نمایش ژنتیکی لوکاسی[1] استفاده شده است. در این ساختارِ جواب که مبتنی بر گراف است، هر کروموزوم (g)، تعداد n ژن را به‌ترتیب $g_1, \ldots, g_n$ شامل می‌شود. در این حالت، هر ژن g_i می‌تواند هر یک از گره‌های همسایهٔ گرهٔ iام را اختیار کند. این انتخاب می‌تواند دربرگیرندهٔ خود گرهٔ iام نیز باشد. بنابراین، هر مقدار از j که به ژن iام تخصیص داده می‌شود، به‌عنوان ارتباطی بین گرهٔ i و j تفسیر می‌شود؛ در نتیجه،

1. Genetic Locus-based representation

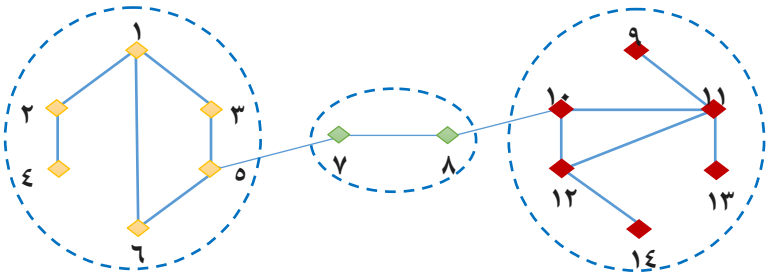
این گره‌ها در اجتماع یکسانی قرار می‌گیرند. کدگشایی این ساختار جواب به شناسایی تمام رئوس مرتبط با یکدیگر نیاز دارد. گام کدگشایی با استفاده از یک برنامهٔ پیمایش معکوس[1] انجام‌پذیر است. از آنجا که در گام کدگشایی تعداد اجتماعات تعیین می‌شوند، دانستن تعداد آنها از قبل الزامی ندارد (شی و همکاران، ۲۰۱۰). برای نمونه، شبکه‌ای را به‌صورت شکل ۲ فرض کنید. این گراف شامل ۱۴ گره است و می‌توان آن را به سه اجتماع دسته‌بندی کرد. نمونه‌ای از یک کروموزوم امکان‌پذیر[2] در شکل ۳ به تصویر کشیده شده است. برای مثال در شکل ۳، اولین ژن بیان‌کنندهٔ ارتباط گرهٔ ۱ با گرهٔ ۲ است. ساختار گراف این کروموزوم در قالب شکل ۴ مشاهده می‌شود. با توجه به شکل ۴، هر اجتماع زیرمجموعه‌ای از رئوس شبکه است.



شکل ۲. ساختار اصلی گراف

۲	۴	۱	۲	۶	۱	۸	۷	۱۱	۱۲	۱۳	۱۴	۱۱	۱۲

شکل ۳. نمونه‌ای از یک کروموزوم شدنی



شکل ٤. ساختار شبکه متناظر با کروموزوم شکل ۳

1. Backtracking
2. Feasible

تابع ارزیابی

پودمانگی Q معیاری کمّی است که نشان می‌دهد در شبکه‌هایی که یال‌ها به‌صورت تصادفی ارتباط بین گره‌ها را برقرار می‌کنند، ساختار اجتماعی وجود ندارد. مقدار شاخص پودمانگی با توجه به رابطهٔ ۱ تعریف می‌شود (پان و همکاران، ۲۰۱۴).

$$Q = \frac{1}{2l}\sum_{ij}\left(a_{ij} - \frac{k_i k_j}{2l}\right)\delta(C_i, C_j)$$ رابطهٔ ۱)

با فرض آن که A نشان‌دهندهٔ ماتریس مجاورت است، در رابطهٔ ۱ اگر گرهٔ i و j با یکدیگر در ارتباط باشند، درایهٔ a_{ij} برابر ۱ و در غیر این صورت مقدار آن صفر است. l برابر تعداد کل یال‌های شبکه است. اگر دو گرهٔ i و j متعلق به یک اجتماع باشند، مقدار تابع δ برابر ۱ و در غیر این صورت صفر است. C_i و C_j به‌ترتیب نشان‌دهندهٔ اجتماعاتی است که گره‌های i و j به آنها تعلق دارند. شاخص دیگری که به‌عنوان ارزش برازندگی جواب‌ها در نظر گرفته شده است، معیار چگالی پودمانگی D است. چگالی پودمانگی از طریق رابطهٔ ۲ محاسبه می‌شود (قربانیان و شقاقی، ۲۰۱۵).

$$D = \sum_{m=1}^{M}\left[\frac{L_m - EX_m}{\sum_i Y_{im}}\right]$$ رابطهٔ ۲)

در رابطهٔ ۲، تعداد کل اجتماعات با M نشان داده شده است. مقدار Y_{im} در صورتی که گرهٔ i در اجتماع m قرار گیرد، برابر ۱ و در غیر این صورت صفر است. L_m دو برابر یال‌های اجتماعِ m و EX_m نشان‌دهندهٔ تعداد یال‌های بیرونی این اجتماع است. برای هر یک از جواب‌های مسئله دو معیار پودمانگی و چگالی پودمانگی محاسبه می‌شود، سپس با استفاده از روش وزن‌دهی به اهداف، یک ارزش مطلوبیت به‌دست می‌آید. در روش وزن‌دهی اهداف، تصمیم‌گیرنده با توجه به مطلوبیت هر هدف، به اهداف مختلف وزن اختصاص می‌دهد و با ضرب کردن توابع هدف در وزن‌های تخصیص داده شده، در نهایت تابع هدف یگانه‌ای به‌دست خواهد آمد. در این مطالعه با توجه به اهمیت هر دو معیار مورد بررسی، وزن‌ها به‌صورت یکسان لحاظ شده است.

انتخاب والدین برای تولید جواب‌های جدید با استفاده از روش تاپسیس

فرایند انتخاب والدین برای تولید فرزندان در این مطالعه از طریق روش تاپسیس صورت می‌پذیرد. این روش یکی از پُرکاربردترین مدل‌های جبرانی است که برای نخستین‌بار توسط یون و هوآنگ ارائه شد. روش تاپسیس هر گزینه را نقطه‌ای در فضا در نظر می‌گیرد و فاصلهٔ اقلیدسی

هر نقطه را نسبت به جواب ایده‌آل مثبت و جواب ایده‌آل منفی محاسبه می‌کند. در نهایت گزینه‌ای انتخاب می‌شود که دارای کمترین فاصله از ایده‌آل مثبت و بیشترین فاصله از ایده‌آل منفی باشد (یویی، ۲۰۱۲). تاپسیس، ماتریس تصمیم را که شامل i جواب و j معیار است، ارزیابی می‌کند. سطرها در ماتریس تصمیم نشان‌دهندهٔ جواب‌های به‌دست آمده از الگوریتم ژنتیک و ستون‌ها برابر میزان پودمانگی و چگالی پودمانگی جواب‌هاست که به‌عنوان معیارهای مسئله در نظر گرفته می‌شود. نحوهٔ به‌کارگیری روش تاپسیس در الگوریتم پیشنهادی به این شکل است که در هر تکرار قبل از انجام تقاطع و جهش، جواب‌های به‌دست آمده توسط الگوریتم ژنتیک به‌کمک روش تاپسیس و با توجه به ارزش پودمانگی و چگالی پودمانگی‌ای که دارند، رتبه‌بندی می‌شوند. در نتیجه دو جوابی که رتبه‌های اول و دوم را میان سایر جواب‌های جمعیت کسب می‌کنند، والد در نظر گرفته شده و عملگر تقاطع روی آنها اعمال می‌شود. همچنین جوابی که رتبهٔ اول را دارد، برای انجام عملگر جهش انتخاب خواهد شد. به این ترتیب در هر تکرار از الگوریتم، روش تاپسیس برای جواب‌های جمعیت اجرا شده و رتبه‌بندی پاسخ‌ها انجام خواهد شد.

عملگر تقاطع[1]

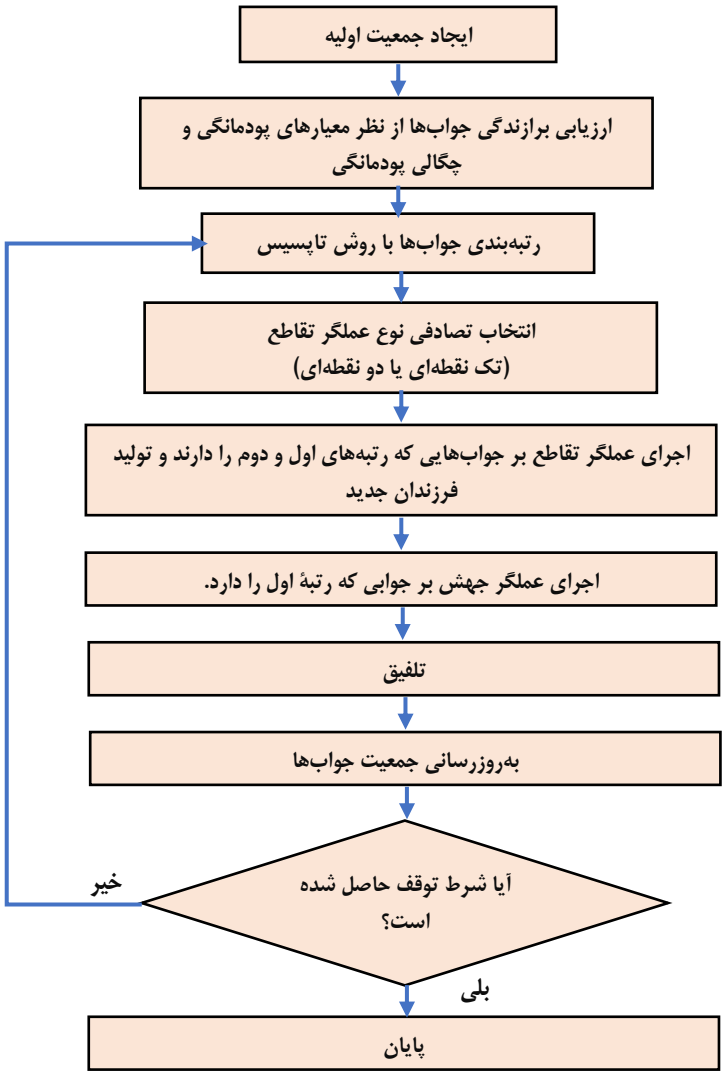
در هر مرحله از روند حل مسئله، یکی از عملگرهای تقاطع تک نقطه‌ای یا دو نقطه‌ای به تصادف انتخاب می‌شود، سپس محل تقاطع در جواب‌های والد به‌صورت تصادفی تعیین می‌گردد. به این ترتیب با تعویض رئوسی که توسط محل‌های تقاطع جدا شده‌اند، فرزندان جدید به‌وجود می‌آیند (گائوکیانگ و ژایوفنگ، ۲۰۱۰).

عملگر جهش[2]

در این مسئله اگر عملگر جهش به‌صورت تصادفی مقدار j از iاُمین ژن را به مقدار k تغییر دهد، ممکن است جست‌وجوی بی‌ثمری در فضای جواب انجام شود و دلیل آن، امکان نبود ارتباط میان گره‌های i و k در شبکهٔ اصلی است. بنابراین، مقادیر امکان‌پذیر آلل‌ها[3] محدود به گره‌های همسایهٔ ژن i می‌شود. این نوع جهش، تولید فرزندان شدنی در فضای جواب را تضمین می‌کند و بدین ترتیب هر گره تنها با یکی از همسایه‌های خود در ارتباط است. از آنجا که پاسخ‌های شدنی فضای جواب‌های ممکن را محدود می‌کند، همگرایی روش را بهبود می‌بخشند (گائوکیانگ و ژایوفنگ، ۲۰۱۰).

---

1. Crossover
2. Mutation
3. Allele



شکل ۵. فلوچارت الگوریتم پیشنهادی

ساختار کلی الگوریتم پیشنهادی

ساختار کلی الگوریتم پیشنهادی در شکل ۵ آورده شده است. همان‌طور که در شکل ۵ مشاهده می‌شود، به‌منظور افزایش کارایی الگوریتم ژنتیک و بهبود کیفیت جواب‌های به‌دست آمده، قبل از اجرای عملگرهای تقاطع و جهش، رتبه‌بندی جواب‌ها توسط روش تاپسیس انجام می‌گیرد. به این ترتیب پاسخ‌های والد برای تولید فرزندان جدید از میان جواب‌هایی که رتبه‌های اول و دوم را

در جمعیت دارند، انتخاب می‌شوند. به‌علاوه، جوابی که رتبهٔ نخست را دارد به‌منظور اعمال عملگر جهش به‌کار گرفته خواهد شد. وزن معیارها در این نوشتار یکسان در نظر گرفته شده است.

روش‌شناسی پژوهش

هدف این پژوهش، ارائهٔ یک الگوریتم فراابتکاری مبتنی بر روش تاپسیس است که از جمله روش‌های تصمیم‌گیری چندمعیاره به حساب می‌آید. استفاده از روش تاپسیس در الگوریتم ژنتیک با هدف بهبود جواب‌های مسئلهٔ اجتماع‌یابی انجام می‌شود. با توجه به مباحث مطرح شده در بخش‌های قبل، جواب‌های به‌دست آمده توسط الگوریتم پیشنهادی از نظر دو معیار پودمانگی و چگالی پودمانگی ارزیابی می‌شوند و هدف ما در این تحقیق یافتن اجتماعاتی با مقدار بیشینهٔ پودمانگی و چگالی پودمانگی است. بدین منظور عملکرد الگوریتم پیشنهادی در مقایسه با دو الگوریتم ژنتیک کلاسیک و یک الگوریتم حریصانه سنجیده می‌شود. آزمایش‌ها روی پنج شبکه در دنیای واقعی صورت پذیرفته است. از این جنبه می‌توان پژوهش پیش رو را پژوهشی آزمایشی دانست. در ادامه توضیحاتی در ارتباط با مسائل نمونه، تنظیم پارامترهای الگوریتم پیشنهادی، تجزیه و تحلیل مسائل نمونه و مقایسهٔ کارایی الگوریتم‌ها آورده شده است.

مقایسهٔ عملکرد الگوریتم‌ها

مسائل نمونه

در این بخش، پنج شبکه‌ای که در دنیای واقعی برای سنجش کارایی الگوریتم پیشنهادی مد نظر قرار گرفته‌اند، به‌صورت اجمالی معرفی می‌شوند:

الف) شبکهٔ کلوپ کاراتهٔ زکریا[1]: این شبکه ۳۴ گره دارد که نشان‌دهندهٔ اعضای یک کلوپ کاراته در ایالات متحده است. ارتباط میان اعضای این کلوپ توسط یال‌ها به‌وجود می‌آید. این شبکه را می‌توان از معتبرترین نمونه در آزمون کارایی الگوریتم‌های اجتماع‌یابی دانست (نیومن، ۲۰۰۶).

ب) شبکهٔ کتاب‌های سیاسی آمریکا[2]: گره‌ها در این شبکه نشان‌دهندهٔ ۱۱۵ کتاب در ارتباط با سیاست‌های آمریکاست که از یک فروشگاه آنلاین خریداری شده‌اند. یال‌های شبکه معرف جفت کتاب‌هایی است که توسط یک خریدار از این فروشگاه خریداری شده است (نیومن، ۲۰۰۶).

---

1. Zakhary Karate Club
2. United States Political Books (PolBooks)

ج) شبکهٔ شخصیت‌های رمان بینوایان[1]: این شبکه دربرگیرندهٔ کاراکترهای رمان بینوایان از نویسندهٔ مشهور ویکتور هوگو است. هر گره در این شبکه نمایندهٔ شخصیتی در رمان است و هر یال بین دو گره نشان‌دهندهٔ این است که دو شخصیت در فصل یکسانی از کتاب قرار دارند. شبکهٔ یاد شده ۷۷ گره و ۲۵۴ یال دارد (نیومن، ۲۰۰۶).

د) شبکهٔ دلفین‌ها[2]: در گراف مربوط به این شبکه ۶۲ دلفین در نظر گرفته شده که هر دلفین نمایندهٔ یک گره است. ارتباط و رفتار این دلفین‌ها از سال‌های ۱۹۹۴ تا ۲۰۰۱ مشاهده شده است. شبکهٔ یاد شده تا حد زیادی به شبکه‌های اجتماعی انسانی شباهت دارد (نیومن، ۲۰۰۶).

ه) شبکهٔ تیم‌های فوتبال دانشگاهی[3]: این شبکه نشان‌دهندهٔ برگزاری مسابقهٔ فوتبال میان دانشگاه‌های یک ایالت در آمریکاست. مسابقات در سال ۲۰۰۰ میلادی برگزار شده‌اند (گیروان و نیومن، ۲۰۰۲).

تنظیم پارامتر الگوریتم‌ها

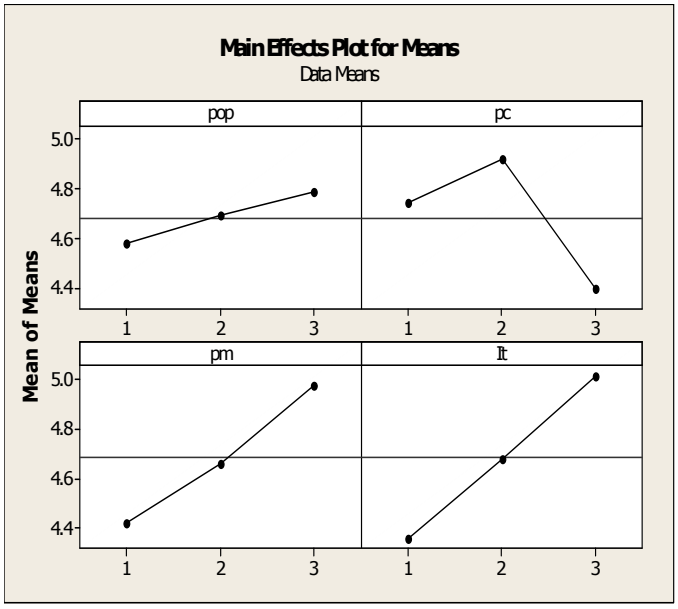
از آنجا که پارامترهای هر الگوریتم بر نتایج به‌دست آمده تأثیر شایان توجهی می‌گذارد، تنظیم آنها می‌تواند گام مهمی در به‌دست آوردن جواب‌های با کیفیت باشد. یکی از روش‌های شناخته شده و کارا در حیطهٔ طراحی آزمایش‌ها، روش تاگوچی[4] است که در این مطالعه از آن استفاده شده است. بر اساس این روش، سطوح متفاوتی به ازای هر پارامتر لحاظ می‌شود. پارامترهای مرتبط با الگوریتم ژنتیک در جدول ۱ درج شده است.

جدول ۱. سطوح پارامترهای ورودی الگوریتم‌ها

پارامترها	نمادها	سطوح
احتمال تقاطع	p_c	۰/۴ – ۰/۶ – ۰/۸
احتمال جهش	p_m	۰/۱ – ۰/۳ – ۰/۵
اندازهٔ جمعیت	pop	۵۰ – ۱۰۰ –۲۰۰
تعداد تکرار	$Iteration$	۱۰۰ – ۲۰۰ –۳۰۰

1. Les Misérables
2. Dolphins Network
3. Football Club
4. Taguchi Method

مطابق آنچه گفته شد، آزمایش‌هایی به روش تاگوچی به‌منظور تعیین انـدازهٔ پارامترهـا بـرای رسیدن به بیشینهٔ مقدار جمع وزنی معیارهای پودمانگی و چگالی پودمانگی طراحی شـد. طراحـی آزمایش‌ها در نرم‌افزار Minitab صورت گرفت. از میان پنج شبکهٔ معرفی شده، یکـی از آنهـا بـه تصادف انتخاب شد و آزمایش‌هایی روی آن انجام گرفت. نتایج به‌دست آمـده پـس از ۲۰ مرتبـه اجرای برنامه و میانگین‌گیری جمع وزنی معیارها در جدول ۲ گزارش شده است. مقدار پودمانگی، چگالی پودمانگی و جمع وزنی این دو معیار به‌ترتیب با نمادهای Q، D و Z نشان داده شده است. مقادیر مناسب برای پارامترها پس از انجام آزمایش‌هـای طراحـی شـده و بـا اسـتفاده از نمـودار میانگین‌های شاخص Z در نرم‌افزار Minitab قابـل دسترسـی اسـت. ایـن نمـودار در شـکل ۶ مشاهده می‌شود. در این نمودار میانگین شاخص Z به ازای اجراهای مختلف آزمایش‌ها با سطوح تعیین شده در جدول ۱ نشان داده شده است.



شکل ٦. نمودار میانگین‌های Z برای پارامترهای الگوریتم ژنتیک

همان‌طور که از شکل ۶ مشخص است با تغییـر احتمـال تقـاطع از ۰/۴ بـه ۰/۶، مقـدار Z افزایش می‌یابد و از ۰/۶ به ۰/۸ مقدار Z کاهش یافته اسـت. از ایـن رو بهتـرین مقـدار احتمـال تقاطع ۰/۶ است. بر همین اساس بهترین مقدار پارامترهای الگوریتم ژنتیک در جدول ۳ گـزارش شده است.

جدول ۲. نتایج آزمایش‌های تاگوچی برای تنظیم پارامترهای الگوریتم ژنتیک

شمارهٔ آزمایش	pop	p_c	p_m	Iteration	بهترین Q	میانگین Q	بهترین D	میانگین D	بهترین Z	میانگین زمان (ثانیه)
۱	۵۰	۰/۴	۰/۱	۱۰۰	۰/۴۳۹۵	۰/۴۰۶۲	۷/۶۵۰	۷/۱۲۵	۴/۰۴۴۷	۰/۹۹۹۹
۲	۵۰	۰/۶	۰/۳	۲۰۰	۰/۴۸۴۰	۰/۴۶۷۰	۹/۰۸۳۳	۸/۹۸۵	۴/۷۸۳۷	۳/۷۵۶۴
۳	۵۰	۰/۸	۰/۵	۳۰۰	۰/۴۹۶۴	۰/۴۸۳۲	۹/۳۲۱۴	۹/۲۱۵	۴/۹۰۸۹	۷/۷۷۸۵
۴	۱۰۰	۰/۴	۰/۳	۳۰۰	۰/۴۵۰۸	۰/۴۴۲۸	۹/۶۵۰	۹/۵۷۸	۵/۰۵۰۴	۸/۵۴۷۳
۵	۱۰۰	۰/۶	۰/۵	۱۰۰	۰/۴۶۲۸	۰/۴۵۲۰	۹/۳۰۷۱	۹/۲۵۱۴	۴/۸۸۴۹	۴/۴۳۷۶
۶	۱۰۰	۰/۸	۰/۱	۲۰۰	۰/۴۶۵۲	۰/۴۵۷۳	۷/۷۹۵۵	۷/۶۵۲۱	۴/۱۳۰۳	۶/۳۱۶۵
۷	۲۰۰	۰/۴	۰/۵	۲۰۰	۰/۴۹۷۴	۰/۴۸۵۳	۹/۷۵۰	۹/۷۲۴	۵/۱۲۳۷	۱۵/۰۴۵
۸	۲۰۰	۰/۶	۰/۱	۳۰۰	۰/۴۷۵	۰/۴۴۶	۹/۶۹۴۴	۹/۵۴۷	۵/۰۸۴۷	۱۴/۸۸۰
۹	۲۰۰	۰/۸	۰/۳	۱۰۰	۰/۴۴۰۹	۰/۴۲۱۵	۷/۸۵۰	۷/۴۲۱	۴/۱۴۵۴	۸/۱۴۱۵

جدول ۳. مقادیر پارامترهای الگوریتم ژنتیک پس از آزمایشات

پارامترها	مقدار پارامتر	پارامترها	مقدار پارامتر
احتمال تقاطع	۰/۶	اندازهٔ جمعیت	۲۰۰
احتمال جهش	۰/۵	تعداد تکرار	۳۰۰

تجزیه و تحلیل مسائل نمونه

در این بخش، عملکرد الگوریتم ژنتیک پیشنهادی با دو الگوریتم ژنتیک کلاسیک و الگوریتم حریصانهٔ معرفی شده در (شقاقی و تیمورپور، ۲۰۱۴) مقایسه می‌شود. معیار مقایسهٔ عملکرد الگوریتم‌ها برابر جمع وزنی پودمانگی و چگالی پودمانگی جواب‌هایی است که توسط هر روش ارائه می‌شود. برنامه‌نویسی مسئله در محیط Matlab انجام شده و اجرای برنامه نیز توسط یک رایانهٔ شخصی با مشخصات QuadCore8200, CPU 2.00 GHz, RAM 4 GB صورت گرفته است. نتایج به‌دست آمده در جدول ۴ مشاهده می‌شود. با توجه به این جدول، نام شبکه، بهترین مقدار پودمانگی، بهترین مقدار چگالی پودمانگی و میانگین این مقادیر نیز پس از ۲۰ مرتبه اجرای برنامه گزارش شده است. همچنین، بهترین مقدار Z و تعداد اجتماعاتی که توسط هر یک از این سه روش به‌دست آمده را نیز می‌توان در جدول ۵ مشاهده کرد.

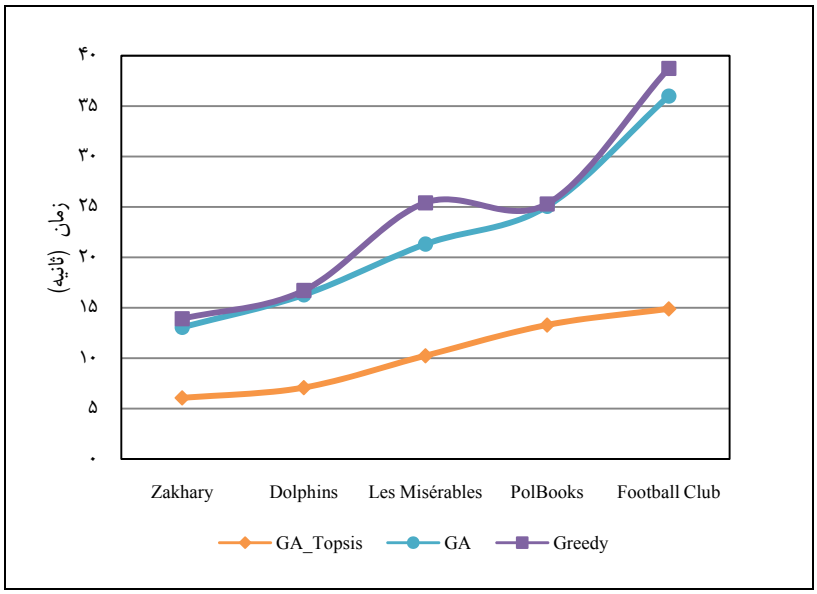
جدول ٤. نتایج به‌دست آمده از حل پنج شبکه در دنیای واقعی

شبکه	الگوریتم	بهترین Q	میانگین Q	بهترین D	میانگین D
کلوپ کاراتهٔ زکریا	GA-Topsis	۰/۴۸۲۴	۰/۴۷۵۹	۹/۹۱۶۷	۹/۵۰۱۴
	GA	۰/۲۴۰۸	۰/۲۲۱۲	۷/۴۴۴	۷/۱۲۶
	Greedy	۰/۲۴۱۸	۰/۲۳۰۷	۸/۶۱۵۴	۸/۵۲۱
دلفین‌ها	GA-Topsis	۰/۶۱۲۵	۰/۵۷۴۳	۲۴/۷۵۰	۲۲/۵۶۱
	GA	۰/۴۰۶۱	۰/۳۹۵۱	۱۸/۵۰	۱۷/۷۸۴
	Greedy	۰/۳۳۹۴	۰/۳۲۵۹	۷/۲۰۱	۶/۹۹۸
تیم‌های فوتبال دانشگاهی	GA-Topsis	۰/۴۸۸۳	۰/۴۱۲۷	۲۷/۷۵۰	۲۷/۱۴۲
	GA	۰/۳۶۷۱	۰/۳۴۷۱	۲۰/۱۷۶۹	۱۹/۱۵۴
	Greedy	۰/۳۲۸۷	۰/۳۱۲۹	۱۸/۸۶۸	۱۷/۵۴۸
شخصیت‌های بینوایان	GA-Topsis	۰/۴۹۲۵	۰/۴۹۰۱	۱۹/۲۵۰	۱۸/۷۵۹
	GA	۰/۴۰۳۸	۰/۳۵۴	۱۷/۷۳۳	۱۶/۳۵۲
	Greedy	۰/۲۸۳۱	۰/۲۵۶۱	۱۲/۳۲۰	۱۱/۹۴۷
کتاب‌های سیاسی آمریکا	GA-Topsis	۰/۴۸۲۳	۰/۴۷۵۲	۲۴/۶۱۳	۲۳/۵۱۷
	GA	۰/۲۸۳۶	۰/۲۶۱۷	۱۹/۴۶۶	۱۸/۵۷۴
	Greedy	۰/۲۳۲۳	۰/۲۰۱۵	۱۶/۶۴۵	۱۵/۲۴۱

جدول ۵. بهترین مقدار Z و تعداد اجتماعات برای پنج شبکه

شبکه	الگوریتم	بهترین Z	تعداد اجتماعات
کلوپ کاراتهٔ زکریا	GA-Topsis	۵/۱۹۹۵	۲
	GA	۳/۸۴۲۶	۴
	Greedy	۴/۴۲۸	۶
دلفین‌ها	GA-Topsis	۱۲/۶۸۱۲	۲
	GA	۹/۴۵۳۰	۳
	Greedy	۳/۶۶۹	۵
تیم‌های فوتبال دانشگاهی	GA-Topsis	۱۴/۱۱۹	۱۰
	GA	۱۰/۲۷۲	۱۲
	Greedy	۹/۵۹۸۴	۱۲
شخصیت‌های بینوایان	GA-Topsis	۹/۸۷۱۲	۶
	GA	۹/۰۶۸۴	۶
	Greedy	۶/۳۰۱	۶
کتاب‌های سیاسی آمریکا	GA-Topsis	۱۲/۵۴۸	۲
	GA	۹/۸۷۴۸	۶
	Greedy	۸/۴۳۸	۵

همان‌طور که در جدول‌های ۴ و ۵ نیز مشاهده می‌شود، الگوریتم پیشنهادی GA-Topsis توانسته است از نظر معیارهای پودمانگی (Q)، چگالی پودمانگی (D) و جمع وزنی این دو شاخص (Z) بهتر عمل کند. شکل ۷ به وضوح نشان می‌دهد الگوریتم پیشنهادی از نظر زمان همگرایی نیز نسبت به دو الگوریتم ژنتیک کلاسیک و حریصانه به مراتب عملکرد مناسب‌تری داشته است.

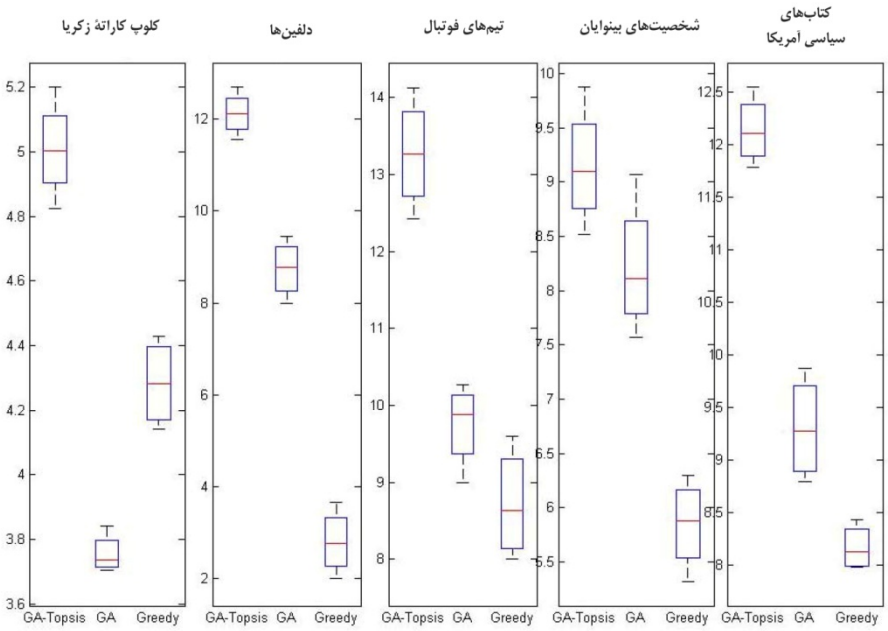


شکل ۷. زمان همگرایی الگوریتم‌ها در حل پنج شبکه

اکنون به‌منظور سنجش این که جواب‌های به‌دست آمده توسط الگوریتم پیشنهادی بهبود شایان توجه و معناداری بیشتری نسبت به جواب‌های سایر الگوریتم‌ها داده است، از آزمون آماری t استفاده می‌شود. بدین صورت می‌توان اطمینان حاصل کرد که میانگین جواب‌های الگوریتم GA-Topsis از نظر جمع وزنی دو معیار پودمانگی و چگالی پودمانگی تفاوت معناداری با جواب‌های دو روش دیگر دارد. از این جنبه، کارایی الگوریتم GA-Topsis در فاصلهٔ اطمینان ۹۵ درصد به‌صورت دو به دو با سایر الگوریتم‌ها آزمون می‌شود. فرض صفر در این آزمون به این معناست که تفاوت معناداری میان میانگین جواب‌های دو الگوریتم مقایسه شده وجود ندارد. آزمون‌ها در نرم‌افزار Minitab انجام شد. نتایج به‌دست آمده از انجام این آزمون‌های آماری در جدول ۶ آورده شده است.

جدول ٦. نتایج آزمون فرض آماری بر عملکرد الگوریتم‌ها از نظر مجموع وزنی معیارها

شبکه	مقایسهٔ الگوریتم‌ها	حد پایین	حد بالا	مقدار آمارهٔ t	مقدار p-value
کلوپ کاراتهٔ زکریا	GATopsis--GA	۱/۵۵۶۶	۱/۶۴۴۲	۷۱/۸۲	۰/۰۰۰
	GATopsis--Greedy	۰/۹۲۹۹	۰/۹۹۴۲	۵۸/۸۰	۰/۰۰۰
دلفین‌ها	GATopsis--GA	۳/۰۸۲	۳/۱۲۹	۲۶۱/۱۸	۰/۰۰۰
	GATopsis--Greedy	۸/۷۸۷۹	۸/۸۲۹۸	۸۲۸/۷۱	۰/۰۰۰
تیم‌های فوتبال دانشگاهی	GATopsis--GA	۳/۸۶۴۷	۳/۹۰۶۹	۳۶۲/۱۳	۰/۰۰۰
	GATopsis--Greedy	۴/۵۶۲۰	۴/۶۲۲۳	۲۹۸/۹۶	۰/۰۰۰
شخصیت‌های بینوایان	GATopsis--GA	۰/۶۰۷۱	۰/۶۶۴۲	۴۳/۷۶	۰/۰۰۰
	GATopsis--Greedy	۳/۴۳۱۱	۳/۴۸۹	۲۳۴/۸۷	۰/۰۰۰
کتاب‌های سیاسی آمریکا	GATopsis--GA	۲/۶۵۳۳	۲/۶۹۳۲	۲۶۲/۸۲	۰/۰۰۰
	GATopsis--Greedy	۴/۱۹۰۲	۴/۲۴۶۵	۲۹۴/۱۷	۰/۰۰۰



شکل ۸. نمودار جعبه‌ای از عملکرد سه الگوریتم به ازای پنج شبکه

با توجه به این که برای اغلب آزمون‌ها مقدار p-value کمتر از ۰/۰۵ گزارش شده است، فرض صفر مبنی بر عدم وجود تفاوت معنادار میان میانگین جواب‌ها رد خواهد شد. بنابراین می‌توان اطمینان گفت که الگوریتم پیشنهادشده عملکرد بهتری در یافتن جواب‌های باکیفیت داشته است. در شکل ۸، نمودارهای جعبه‌ای مرتبط با عملکرد سه الگوریتم به ازای پنج شبکه به تصویر کشیده شده است. همان‌طور که مشخص است، الگوریتم پیشنهادشده توانسته است از نظر کیفیت مجموع وزنی شاخص‌ها، با اختلاف شایان توجهی بر سایر روش‌ها غلبه کند.

نتیجه‌گیری و پیشنهادها

در این مقاله ضمن تشریح کامل مسئلهٔ اجتماع‌یابی، موارد کاربرد آن در مسائل دنیای واقعی بیان شد. از آنجا که رویکردهای سنتی حل مسئلهٔ اجتماع‌یابی تنها یک معیار خاص را در شکل‌دهی اجتماعات یک شبکه در نظر می‌گیرند، ارائهٔ رویکردی که با بهینه‌کردن همزمان چند شاخص بتواند مجموعه‌ای از اجتماعات در شبکه را بیابد، راهگشا خواهد بود. به همین انگیزه، در این نوشتار یک الگوریتم ژنتیک مبتنی بر روش تصمیم‌گیری چندمعیارهٔ تاپسیس ارائه شد. از روش پیشنهادشده به‌منظور بیشینه‌سازی شاخص‌های پودمانگی و چگالی پودمانگی که از توابع هدف معتبر در مسئلهٔ اجتماع‌یابی محسوب می‌شوند، استفاده شد. مقدار این دو شاخص به‌عنوان ملاک برازندگی هر جواب مد نظر قرار گرفت. به‌منظور انتخاب جواب‌های کاندید برای عملگرهای تقاطع و جهش، روش تاپسیس به‌کار رفت؛ سپس عملکرد الگوریتم پیشنهاد شده با دو الگوریتم ژنتیک کلاسیک و حریصانه روی پنج شبکهٔ دنیای واقعی مقایسه شد. نتایج به‌دست آمده گویای برتری الگوریتم ارائه شده بر دو الگوریتم دیگر از نظر توانایی در شناسایی اجتماعات با کیفیت بود. این الگوریتم می‌تواند در کشف اجتماعات دقیق از مشتریان شبکه‌های برخط فروش که علایق شبیه به یکدیگر دارند، به‌کار رود. به این ترتیب فرصت‌های تجاری مناسبی از طریق پیشنهاد محصولات متناسب با علایق مشتریان به‌وجود خواهد آمد. شکل‌دهی اجتماعات علمی در شبکه‌های استنادی تحقیقات، بر پایهٔ موضوعات مورد مطالعه، از کاربردهای دیگر روش پیشنهاد شده است. شناسایی دقیق اجتماعات توسط الگوریتم ارائه شده می‌تواند در تشخیص فعالیت‌های کلاه‌برداری و تقلب در شبکه‌های مخابراتی استفاده شود. پیش‌بینی ارتباطات میان گروه‌های تروریستی در شبکه‌های اجتماعی برخط از کاربردهای دیگر روش پیشنهاد شده است. از این جنبه، الگوریتم معرفی شده قادر است در کشف ساختارهای اجتماعی در بسیاری از شبکه‌های دنیای واقعی به‌کار برده شود. با توجه به این که در الگوریتم پیشنهاد شده از روش تصمیم‌گیری چندمعیارهٔ تاپسیس استفاده شده است، می‌توان سایر معیارهای رایج در شناسایی

ساختارهای اجتماعی، نظیر امتیاز و برازش اجتماعات، وسعت و رسانایی شبکه را نیز در روش پیشنهاد شده گنجاند و به این ترتیب دقت آن را بالا برد. به‌طورکلی الگوریتم پیشنهاد شده با توجه به ویژگی‌هایی که دارد، در اغلب مسائل بهینه‌سازی چندمعیاره کاربرد دارد.

از جمله موضوعات شایان توجه برای مطالعات بعدی، استفاده از الگوریتم پیشنهاد شده در شناسایی اجتماعات در شبکه‌های پویاست. در این صورت، می‌توان ساختارهای اجتماعی در شبکه‌های پیچیده را با توجه به تغییر ارتباطات بین رئوس در گذر زمان کشف کرد. شاخص‌های شکل‌دهی اجتماعات در این پژوهش برای شناسایی اجتماعاتی که با یکدیگر همپوشانی دارند نیز کاربرد دارد. از این جنبه، در مطالعات آینده می‌توان از روش پیشنهاد شده برای کشف اجتماعاتی که همپوشانی دارند نیز استفاده کرد. از سوی دیگر، رویکرد ارائه شده می‌تواند به‌واسطهٔ استفاده از روش تصمیم‌گیری چندمعیارهٔ تاپسیس، در بسیاری از مسائل بهینه‌سازی چندمعیاره مانند بهینه‌سازی همزمان شاخص‌های وسعت، رسانایی، امتیاز اجتماعات و غیره موفق عمل کند. بدین سبب، افزودن سایر معیارهای شکل‌دهی اجتماعات در کنار شاخص‌های پودمانگی و چگالی پودمانگی در الگوریتم پیشنهادشده و مقایسه نتایج به‌دست آمده با یافته‌های این پژوهش، از سایر موضوعاتی است که می‌تواند مبنای مطالعات آتی قرار گیرد. مطالعهٔ مشابه دیگری در ارتباط با بررسی کارایی تلفیق سایر روش‌های تصمیم‌گیری چندمعیاره با الگوریتم ژنتیک پیشنهاد شده می‌تواند زمینهٔ جالب دیگری برای تحقیقات بعدی باشد.

منابع

اصغری اسکویی، محمدرضا؛ قاسم‌زاده، محمد (۱۳۹۵). کاربرد قواعد کشفی و الگوریتم ژنتیک در ساخت مدل ARMA برای پیش‌بینی سری زمانی. *فصلنامهٔ مدیریت فناوری اطلاعات*، ۸(۱)، ۲۶-۱.

البرزی، محمود؛ پورزرندی، محمدابراهیم؛ خان‌بابایی، محمد (۱۳۸۹). به‌کارگیری الگوریتم ژنتیک در بهینه‌سازی درختان تصمیم‌گیری برای اعتبارسنجی مشتریان بانک. *فصلنامهٔ مدیریت فناوری اطلاعات*، ۲(۴)، ۳۸- ۲۳.

بهشتی‌نیا، محمدعلی؛ فرازمند، ناهید (۱۳۹۴). ارائهٔ سیستم پشتیبانی تصمیم نوین به‌منظور موازنهٔ هزینه ـ انتشار دی‌اکسیدکربن گسسته در پروژه‌های ساخت: کاربردی از الگوریتم ژنتیک الگوبرداری. *فصلنامهٔ مدیریت فناوری اطلاعات*، ۷(۱)، ۴۸- ۲۳.

بهشتی‌نیا، محمدعلی؛ قهرمانی، مانی (۱۳۹۵). ارائهٔ سیستم پشتیبانی تصمیم مبتنی بر الگوریتم ژنتیک (مطالعهٔ موردی: زمان‌بندی در زنجیرهٔ تأمین). *فصلنامهٔ مدیریت فناوری اطلاعات*، ۸(۳)، ۴۷۶-۴۵۵.

تقوی‌فرد، محمدتقی؛ سادات‌حسینی، فریبا؛ خان‌بابایی، محمد (۱۳۹۳). مدل رتبه‌بندی اعتباری هیبریدی با استفاده از الگوریتم‌های ژنتیک و سیستم‌های خبرهٔ فازی (مطالعهٔ موردی: موسسهٔ مالی و اعتباری قوامین). *فصلنامهٔ مدیریت فناوری اطلاعات*، ۶(۱)، ۴۶-۳۱.

حقیقی، الهام؛ منتظر، غلامعلی (۱۳۹۴). شناسایی عوامل مؤثر بر اعتمادسازی در شبکه‌های اجتماعی برخط به‌کمک روش الکترهٔ فازی. *فصلنامهٔ مدیریت فناوری اطلاعات*، ۷(۴)، ۷۴۰- ۷۱۵.

رضایی‌نور، جلال؛ لسانی، رضوان؛ زکی‌زاده، عاطفه؛ صفامجید، غدیر. (۱۳۹۳). بررسی شبکه‌های اجتماعی همکاری نویسندگی در حوزهٔ فناوری اطلاعات با استفاده از تکنیک‌های شبکه‌های اجتماعی. *فصلنامهٔ مدیریت فناوری اطلاعات*، ۶(۲)، ۲۵۰- ۲۲۹.

فتحیان، محمد؛ حسینی، محمد (۱۳۹۳). بررسی تأثیر اجتماعات در تقویت رفتار خرید مشتریان. *فصلنامهٔ مدیریت فناوری اطلاعات*، ۶(۳)، ۴۵۴- ۴۳۵.

کاباران‌زاده قدیم، محمدرضا؛ رفوگر آستانه، حسین (۱۳۸۸). طراحی یک سیستم پشتیبان تصمیم‌گیری (DSS) در مدیریت برای حل مسئلهٔ تسطیح منابع در مدیریت پروژه با رویکرد الگوریتم ژنتیک (GA). *فصلنامهٔ مدیریت فناوری اطلاعات*، ۱(۳)، ۸۸- ۶۹.

کی‌پور، اعظم؛ براری، مرتضی؛ شیرازی، حسین. (۱۳۹۳). ارائهٔ روشی جدید برای پیشگویی پیوند بین رأس‌های موجود در شبکه‌های اجتماعی. *فصلنامهٔ مدیریت فناوری اطلاعات*، ۶(۳)، ۴۸۶- ۴۷۵.

References

Agarwal, G., & Kempe, D. (2008). Modularity-maximizing graph communities via mathematical programming. *The European Physical Journal B*, 66(3), 409-418.

Alborzi, M., Pourzarandi, M., & Khanbabaei, M. (2011). Using Genetic Algorithm in Optimizing Decision Trees for Credit Scoring of Banks Customers. *Journal of Information Technology Management*, 2(4), 23-38. (*in Persian*)

Al-Ghazzali, T. (2009). *Meta heuristics: from design to implementation*. Chichester: John Wiley and Sons Inc.

Asghari Oskoei, M., & Ghasemmzade, M. (2016). Application of Heuristic Rules and Genetic Algorithm in ARMA Model Estimation for Time Series Prediction. *Journal of Information Technology Management*, 8(1), 1-26. (*in Persian*)

Beheshtinia, M. A., & Farazmand, N. (2015). A Novel Decision Support System for Discrete Cost-CO2 Emission Trade-off in Construction Projects: The Usage

of Imitate Genetic Algorithm. *Journal of Information Technology Management*, 7 (1), 23-48. (*in Persian*)

Beheshtinia, M. A., & Ghahremani, M. (2016). A Decision Support System Based on Genetic Algorithm (Case Study: Scheduling in Supply Chain). *Journal of Information Technology Management,* 8(3), 455-476. (*in Persian*)

Blondel, V.D., Guillaume, J.L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment,* P10008.

Boettcher, S., & Percus, A.G. (2001). Optimization with extremal dynamics. *Physical Review Letters,* 86(1), 5211–5214.

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. *IEEE Transactions on Computational Social Systems,* 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.

Duch, J., & Arenas, A. (2005). Community detection in complex networks using extremal optimization. *Physical Review E*, 72(2), 027104.

Fathian, M. & Hosseini, M. (2014). Investigating the Impact of Virtual Communities on Furtherance of Customers' Buying Behavior. *Journal of Information Technology Management,* 6(3), 435-454. (*in Persian*)

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. & Shaqaqi, B. (2015). A Genetic Algorithm for Modularity Density Optimization in Community Detection. *International Journal of Economy, Management and Social Sciences,* 4(1), 117-122.

Girvan, M., & Newman, M. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12), 7821–7826.

Griechisch, E. & Pluhar, A. (2011). Community Detection by using the Extended Modularity. *Acta Cybernetica*, 20(1), 69-85.

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.

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.

Haghighi, E., & Montazer, G., (2016). Identifying the Effective Factors in Making Trust in Online Social Networks on the perspective of Iranian experts Using Fuzzy ELECTRE. *Journal of Information Technology Management*, 7(4), 715-470. (*in Persian*)

Kabaranzadeh Ghadim, M.R., & Astaneh, H.R. (2009). Designing a Decision Support System (DSS) schema with Applying Genetic Algorithm for Survey of Resource Leveling Problem-(Vehicles). *Journal of Information Technology Management*, 2(3), 69-88 . (*in Persian*)

Keypour, A., Barari, M, & Shirazi, H. (2014). Presenting a New Method for Link Prediction in Social Networks. *Journal of Information Technology Management*, 6(3), 475-486. (*in Persian*)

Leskovec, J., Lang, K. & Mahoney, M. (2010). Empirical comparison of algorithms for network community detection. *ACM,* 20(16), 631-640.

Li, Z., Zhang, S., Wang, R., Zhang, X., & Chen, L. (2008). Quantitative function for community detection, *Physical Review E*, 77(3), 157-178.

Massen, C.P., & Doye, J., (2005). Identifying communities within energy landscapes, *Physical Review E*, 71(4), 0461011-04610112.

Newman, M. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical Review,*1(3), 12-34.

Newman, M., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review,* 69(2), 22-38.

Pan, G., Zhang, W., Wu, Z., & Li., S. (2014). Online Community Detection for Large Complex Networks. *PLoS ONE,* 9(7), 168-188.

Pizzuti, C. (2008). GA-Net: A Genetic Algorithm for Community Detection in Social Networks. *Computer Science*, 5199(1), 1081-1090.

Rezaeenour, J., Lesani, R., Zakizadeh, A., & Safamajid, G. (2014). Evaluating Authorship Collaboration Networks in the Field of Information Technology

Using Social Netwowk Techniques. *Journal of Information Technology Management*, 6(2), 229-250. (*in Persian*)

Shang, R., Bai, J., Jiao, L., & Jin, C., (2013) Community detection based on modularity and an improved genetic algorithm, *Physica A: Statistical Mechanics and its Applications*, 392(5), 1215-1231.

Shaqaqi, B., & Teymorpour, B., (2015). A new heuristic algorithm for modularity optimization in complex networks community detection. *11th International Industrial Engineering Conference,* 7-8 January.

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.

Taghavifard, M., Hosseini, F., & Khanbabaei, M. (2014). Hybrid credit scoring model using genetic algorithms and fuzzy expert systems Case study: Ghavvamin financial and credit institution. *Journal of Information Technology Management*, 6(1), 31-46. (*in Persian*)

Yue, Z., (2012). Extension of TOPSIS to determine weight of decision maker for group decision making problems with uncertain information. *Expert Systems with Applications*, 39(7), 6343-6350.

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.