ارائه روشی برای تشخیص اجتماع در شبکه‌های پیچیده با الگوریتم بهینه‌سازی خرگوش‌های مصنوعی

آرش هدایتی [۱]، محسن محمودی [۱]

[۱] پژوهشگر پژوهشگاه علوم انتظامی و مطالعات اجتماعی فراجا، تبریز
mohsen.nahmoudi98@email.com ،waro.info@gmail.com

چکیده

شبکه پیچیده نگاهی جدید به پدیده‌های می‌باشد که با توجه به ارتباط اجزای آن و همچنین ارتباط با دیگر پدیده‌ها، دارای پیچیدگی بالایی بوده و رفتار جمعی متفاوتی از خود نشان می‌دهند. تشخیص اجتماع یک چالش مهم در این شبکه‌ها می‌باشد. در این مقاله برای تشخیص اجتماع در شبکه‌های پیچیده از رویکرد فرا اکتشافی و الگوریتم بهینه‌سازی خرگوش مصنوعی استفاده شد. روش پیشنهادی در محیط نرم‌افزاری متلب پیاده‌سازی شده و کارایی آن در چهار مجموعه داده حقیقی (باشگاه کاراته زاخاری، شبکه فوتبال کالج آمریکایی، شبکه دلفین‌های، شبکه پول بوکز) با شاخص‌های پیمانگی و NMI مورد تجزیه‌وتحلیل قرار گرفت. تجزیه‌وتحلیل نتایج نشان داد که در مجموعه داده‌های حقیقی بالاترین مقدار پیمانگی به‌صورت میانگین برای مجموعه داده Polbooks و پایین‌ترین مقدار آن برای مجموعه داده Karate می‌باشد. این در حالی است که در مجموعه داده حقیقی بالاترین میزان NMI برای مجموعه داده Dolphins و کمترین مقدار برای مجموعه داده Karate به‌دست‌آمده است.

کلمات کلیدی: شبکه پیچیده، شبکه اجتماعی، تشخیص اجتماع، الگوریتم بهینه‌سازی خرگوش مصنوعی.

۱ مقدمه

شبکه‌های پیچیده را می‌توان در قالب گراف نمایش داد که در این گراف، گره‌ها معادل با موجودیت‌های شبکه پیچیده بوده و یال‌های گراف نشان‌دهنده ارتباط بین این موجودیت‌ها می‌باشد. در این گراف‌ها توزیع یال‌ها به‌صورت ناهمگن می‌باشد. به این ویژگی شبکه‌های پیچیده ساختار اجتماعی می‌گویند [۱]. اجتماع‌ها اطلاعات ارزشمندی در مورد نوع ارتباط بین موجودیت‌ها، نحوه انتقال اطلاعات بین آن‌ها و نحوه‌ی توزیع موجودیت‌ها در شبکه‌های پیچیده ارائه می‌کنند [۲]. یکی از زمینه‌های تحقیقاتی که در تحلیل شبکه‌های پیچیده بسیار موردتوجه قرارگرفته است، شناسایی اجتماع‌ها می‌باشد. شناسایی اجتماع، روشی برای پیدا کردن ساختار اجتماعی آن شبکه می‌باشد [۳]. شناسایی اجتماع یکی از چالش‌های مهم در شبکه‌های پیچیده

است [۴].

رویکردهای مختلفی برای تشخیص اجتماع ابداع‌شده است. اما با توجه به اینکه شناسایی اجتماع همواره در قالب یک مسئله NP-Hard مطرح‌شده است؛ بهترین رویکرد در این زمینه استفاده از الگوریتم‌های فرا اکتشافی می‌باشد. الگوریتم‌های متعددی در این زمینه مطرح‌شده است. هرکدام از این الگوریتم‌ها در کنار محاسنی که دارند؛ دارای معایبی نیز هستند. یکی از معایب اصلی آن‌ها گرفتار شدن در بهینه‌های محلی است که بی‌شک در کارکرد الگوریتم تأثیر منفی خواهد گذاشت. در این مقاله تلاش خواهد شد از الگوریتم بهینه‌سازی خرگوش مصنوعی برای تشخیص اجتماع در شبکه‌های اجتماعی استفاده شود. این الگوریتم کمتر در بهینه‌های محلی گرفتار می‌شود. در الگوریتم خرگوش مصنوعی، فرآیندهای جستجو با یک ضریب انرژی (A) کنترل می‌شود. روند کاهشی A به الگوریتم کمک می‌کند تا به‌آرامی از جستجوی سراسری به جستجوی محلی تغییر مکان دهد. خرگوش مصنوعی با استفاده از A ابتدا عملیات اکتشاف و سپس عملیات بهره‌برداری را انجام می‌دهد. ضریب انرژی A باعث می‌شود که الگوریتم حتی در مراحل پایانی تکرارها نیز کاوش داشته باشد. از دیگر مزیت‌های الگوریتم خرگوش مصنوعی می‌توان به تعداد کم پارامترها برای تنظیم اشاره کرد. با توجه به مطلب عنوان‌شده مهم‌ترین نوآوری مقاله را می‌توان در موارد زیر خلاصه کرد:

- استفاده از الگوریتم بهینه‌سازی خرگوش مصنوعی برای تشخیص اجتماع
- بهبود دقت تشخص اجتماع در شبکه‌های پیچیده

این مقاله مشتمل بر شش بخش می‌باشد که در بخش دوم آن کارهای پیشین بررسی شده و سپس در بخش سوم روش پیشنهادی شرح داده می‌شود. در بخش چهارم محیط شبیه‌سازی و مجموعه داده معرفی می‌گردد. بخش پنجم نیز به تجزیه‌وتحلیل نتایج می‌پردازد. نهایتاً در بخش ششم، پیرامون نتایج بحث و نتیجه‌گیری شده است.

۲ پیشینه تحقیق

در [۵] پژوهشگران استفاده از الگوریتم‌های مبتنی بر جمعیت را برای تشخیص اجتماع در شبکه‌های اجتماعی موردبررسی قرار داده‌اند. آن‌ها در این راستا چند الگوریتم بهینه‌سازی را انتخاب کرده و فرآیند بیشینه‌سازی را انجام داده‌اند. در [۶] نیز از الگوریتم بهینه‌سازی شعله پروانه گسسته برای یافتن راه‌حل تشخیص اجتماع در شبکه‌های اجتماعی استفاده‌شده است. همچنین در این مقاله پژوهشگران باارزش گذاری گره‌ها و شناسایی گره‌های مؤثر با رویکرد فرا ابتکاری نسبت به بیشینه‌سازی انتشار اقدام کرده‌اند. در [۷] از الگوریتم بهینه‌سازی کلونی مورچه برای تشخیص اجتماع در شبکه اجتماعی استفاده‌شده است. در این مقاله فرآیند تشخیص اجتماع با بهره‌گیری از گره‌های مؤثر انجام می‌گیرد.

در [۸] استفاده از الگوریتم بهینه‌سازی جهش قورباغه بهبودیافته برای شناسایی بهینه اجتماع و همچنین بیشینه‌سازی انتشار در شبکه‌های اجتماعی پیشنهادشده است. در این پژوهش برای حل مسئله ابتدا کاربران در شبکه اجتماعی خوشه‌بندی‌شده و سپس در هر خوشه کاربران بانفوذ شناسایی می‌شوند؛ در ادامه فرایند

انتشار در شبکه با بهره‌گیری از این گره‌های بانفوذ انجام می‌گیرد. در [۹] پژوهشگران با استفاده از الگوریتم بهینه‌سازی علف‌های هرز راه‌حلی را برای شناسایی اجتماع و بیشینه‌سازی انتشار در شبکه‌های اجتماعی پیشنهاد داده‌اند. در این مقاله نیز تمرکز پژوهشگران بر شناسایی گره‌های مؤثر شبکه و انتشار پیام به‌وسیله آن‌ها می‌باشد. در [۱۰] از ترکیب رویکرد مبتنی بر مرکزیت و الگوریتم بهینه‌سازی گرگ خاکستری برای شناسایی اجتماعی در شبکه‌های اجتماعی استفاده‌شده است؛ در این رویکرد الگوریتم بهینه‌سازی گرگ خاکستری با بهبود فرآیند نظریه مرکزیت راه‌حلی را برای تشخیص اجتماع در شبکه‌های اجتماعی ارائه می‌دهد.

در [۱۱] پژوهشگران استفاده از الگوریتم بهینه‌سازی خفاش گسسته را برای شناسایی اجتماع در شبکه‌های اجتماعی، پیشنهاد کرده‌اند. در [۱۲] رویکرد بهینه‌سازی چندهدفه برای بیشینه‌سازی انتشار در شبکه اجتماعی پیشنهادشده است.

بررسی کارهای پیشین گویای آن است که پژوهشگران برای تشخیص اجتماع اغلب از الگوریتم‌های فرااکتشافی استفاده می‌کنند. ولی در استفاده از چنین مدل‌هایی علیرغم نتایج مطلوب باید دو چالش اساسی گرفتار شدن در بهینه‌های محلی و همگرایی زودرس را در نظر گرفت.

۳ روش پیشنهادی

درروش پیشنهادی، جمعیت اولیه خرگوش‌های مصنوعی به‌صورت تصادفی و کاملاً پویا تولیدشده و پس از محاسبه خرگوش‌های مصنوعی (ترکیبی از توابع پیمانگی و پراکندگی)، اعضای جمعیت اولیه بر اساس برازندگی‌های محاسبه‌شده و خرگوش برتر (خرگوشی که بالاترین مقدار برازش را دارد) به‌عنوان راه‌حل انتخاب می‌گردد. لازم به ذکر است در این رویکرد هر خرگوش به‌صورت یک بردار مدل‌سازی شده و هر بردار بیانگر یک راه‌حل برای تشخیص اجتماع در شبکه‌های پیچیده می‌باشد. در ادامه خرگوش‌های مصنوعی با کمک عملگرهای جستجوی انحرافی، پنهان شدن مصنوعی و به‌روزرسانی راه‌حل‌های جدیدی برای تشخیص جوامع در شبکه‌های پیچیده تولید می‌کنند. در هر نسل از الگوریتم راه‌حل‌های جدید جایگزین راه‌حل‌های مرحله قبل شده و مجدد برازندگی آن‌ها محاسبه می‌شود. اگر در بین راه‌حل‌های جدید، راه‌حلی یافت شود که برازندگی آن از برازندگی خرگوش برتر بهتر باشد؛ آن راه‌حل با راه‌حل برتر قبلی جایگزین شده و عملگرهای الگوریتم اجرا می‌شود. این فرآیند تا برقراری شرط خاتمه ادامه یافته و در نسل آخر خرگوش مصنوعی برتر به‌عنوان یک راه‌حل برای تشخیص جوامع در شبکه‌های پیچیده انتخاب می‌شود. شکل ۱ بلوک دیاگرام روش پیشنهادی را نشان می‌دهد.

۱.۳ جمعیت اولیه خرگوش‌های مصنوعی

درروش پیشنهادی هر راه‌حل (عضو جمعیت اولیه) یک جواب برای مسئله تشخیص اجتماع خواهد بود. لذا عضو iام در جمعیت اولیه خرگوش‌های مصنوعی برداری از x_i است که طبق رابطه (۱) مدل‌سازی می‌گردد.

$$x_i = [x_i^1, x_i^2, , x_i^{nv}] \quad (۱)$$

شکل ۱: بلوک دیاگرام روش پیشنهادی

که x_i^j نشان‌دهندهٔ جامعه jاُمین عضو با i اُمین خرگوش مصنوعی و nv نشان‌دهنده تعداد گره‌ها در هر مجموعه داده (شبکه پیچیده) است. شکل ۲ بردار یک عضو از جمعیت اولیه خرگوش‌های مصنوعی را برای شبکه پیچیده مدل‌سازی کرده است.

$W_{1,1}$	$W_{1,2}$	...						...	$W_{1,n}$

شکل ۲: نمایش یک عضو جمعیت اولیه در طرح پیشنهادی برای شبکه تک لایه

۲.۳ محاسبه برازندگی خرگوش‌های مصنوعی

در تشخیص جوامع با الگوریتم‌های فرا اکتشافی می‌توان شاخص‌های مختلفی را برای ارزیابی اعضای جمعیت اولیه استفاده کرد. در این مقاله از شاخص پیمانگی استفاده شده است. رابطه (۲) این شاخص را معرفی

می‌کند.

$$Q = \sum_{k=۱}^{S} \left[\frac{l_k}{L} - \left(\frac{d_k}{۲L} \right)^{۲} \right] \quad (۲)$$

این شاخص، ملاکی برای سنجش امکان تقسیم‌بندی شبکه به اجتماعات است. بر اساس این شاخص هر چه تعداد پیوندها بین گره‌های اجتماع بیشتر باشد و تعداد پیوندهای بین اجتماعات کمتر باشد؛ آن شبکه پیمانگی بالاتری دارد. در این رابطه پارامتر S نشان‌دهنده تعداد کل اجتماعات است. همچنین پارامتر L تعداد کل یال‌های شبکه را نشان می‌دهد. پارامتر l_k و d_k نیز به ترتیب معرف تعداد یال‌های داخل اجتماع، مجموع درجات تمام گره‌های داخل اجتماع k هستند.

۳.۳ جستجوی غذا

فرض بر این است که در الگوریتم پیشنهادی، در مناطق تحت پوشش هر خرگوش، مقداری علف و حفره وجود دارد و خرگوش‌ها همیشه به‌طور تصادفی از موقعیت یکدیگر برای جستجوی غذا بازدید می‌کنند. رفتار جستجوی غذا برای خرگوش نشان می‌دهد؛ هر خرگوش تمایل دارد موقعیت خود را نسبت به خرگوش دیگر که به‌طور تصادفی در منطقه‌ی که تحت پوشش آن قرار دارد، به‌روز نماید. مدل ریاضی جستجوی غذا توسط خرگوش‌ها به‌صورت زیر فرموله می‌شود:

$$\vec{v}_i(t+۱) = \vec{x}_j(t) + R \cdot (\vec{x}_i(t) - \vec{x}_j(t)) + \text{round}(۰/۵ \cdot (۰/۰۵ + r_۱)) \quad (۳)$$

$$R = L \cdot c \quad (۴)$$

$$L = (e - e^{(\frac{t-۱}{T})^{۲}}) \cdot \sin(۲\pi r_۲) \quad (۵)$$

$$c(k) = \begin{cases} ۱ & \text{if } k = g(l) \\ ۰ & \text{else} \end{cases} \quad k = ۱, \ldots, d \text{ and } l = ۱, \ldots, \lceil r_۳ \cdot d \rceil \quad (۶)$$

$$g = \text{randperm}(d) \quad (۷)$$

$$n_۱ \sim N(۰, ۱) \quad (۸)$$

در رابطه (۳) مقدار $\text{round}(۰/۵ \cdot (۰/۰۵ + r_۱))$ به‌عنوان دلتا شناخته می‌شود که به جمعیت اولیه اضافه می‌گردد. هدف از این کار قرار نگرفتن الگوریتم در بهینگی محلی می‌باشد و الگوریتم بتواند به سمت بهینگی سراسری حرکت نماید. L بیانگر میزان دویدن خرگوش می‌باشد که در طول تکرار الگوریتم تغییر می‌نماید. L در طول تکرار از بیشترین مقدار به سمت کمترین مقدار حرکت می‌کند. همچنین C یک بردار نگاشت است که می‌تواند به الگوریتم پیشنهادی کمک کرده تابه صورت تصادفی تعدادی از عناصر جمعیت را برای انجام عملیات جهش انتخاب نماید. نهایتاً R نشان‌دهنده متغیر حرکتی الگوریتم می‌باشد که در شبیه‌سازی تعبیه گردیده است. برای روابط (۳) الی (۸) توضیحات زیر را باید در نظر گرفت.

- $\vec{v}_i(t+1)$ نشان‌دهنده موقعیت خرگوش در لحظه‌ی $t+1$
- $\vec{v}_i(t)$ بیانگر موقعیت خرگوش در لحظه‌ی t
- n بیانگر ابعاد جمعیت اولیه (تعداد خرگوش‌ها)
- d بیانگر ابعاد مسئله‌ی بهینه‌سازی
- T بیانگر حداکثر تکرار الگوریتم بهینه‌ساز خرگوش
- $\lceil . \rceil$ بیانگر تابعی جهت محاسبه‌ی حد بالا (سقف)
- تابع round() گرد کننده عدد به نزدیک‌ترین عدد صحیح
- تابع randperm() بیانگر تولید عدد صحیح در بازه‌ی $[1..d]$
- r_1، r_2 و r_3 سه عدد تصادفی در بازه $(0, 1)$
- L نشان‌دهنده سرعت حرکت در هنگام جستجو
- n_1 تابع توزیع نرمال

معادله (۳) نشان می‌دهد که خرگوش‌ها با توجه به موقعیت خرگوش‌های دیگر، عملیات جستجو را برای یافتن غذا انجام می‌دهند.

۴.۳ پنهان شدن تصادفی

در هر تکرار، خرگوش‌ها d حفره در اطراف خود (هر بعد از فضای جستجو) ایجاد می‌کند؛ تا از دست شکارچیان در امان باشند.در این میان هر خرگوش فقط یکی از لانه‌ها را برای مخفی شدن انتخاب می‌کند (به‌طور تصادفی). روابط زیر فرآیند پنهان شدن تصادفی را نشان می‌دهد.

$$\vec{b}_{i,j}(t) = \vec{x}_i(t) + H \cdot g \cdot \vec{x}_i(t) \quad (۹)$$

$$H = \frac{T - t + 1}{T} \cdot r_4 \quad (۱۰)$$

$$n_2 \sim N(0, 1) \quad (۱۱)$$

$$g(k) = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{else} \end{cases} \quad k = 1, \ldots, d \quad (۱۲)$$

در معادله (۹)، H پارامتر پنهان شدن است که به‌صورت خطی از ۱ به $1/T$ با یک تغییر تصادفی در طول تکرارها کاهش می‌یابد. این پارامتر، در ابتدا، حفره‌ها را در فاصله‌ی دورتری از موقعیت خرگوش ایجاد

می‌کند. با افزایش تکرارها، این فاصله نیز کاهش می‌یابد. برای مدل‌سازی ریاضی استراتژی پنهان شدن تصادفی، معادلات زیر پیشنهادشده‌اند.

$$\vec{v}_i(t+1) = \vec{x}_i(t) + R \cdot (r_4 \cdot \vec{b}_{i,r}(t) - \vec{x}_i(t)) \quad (13)$$

$$g_r(k) = \begin{cases} 1 & \text{if } k = \lceil r_5 \cdot d \rceil \\ 0 & \text{else} \end{cases} \quad k = 1, \ldots, d \quad (14)$$

$$\vec{b}_{i,r}(t) = \vec{x}_i(t) + H \cdot g_r \cdot \vec{x}_i(t) \quad (15)$$

که در آن $\vec{b}_i$ نشان‌دهنده یک لانه تصادفی انتخاب‌شده از میان لانه‌های d و r_4 و r_5 دو عدد تصادفی در بازه $(0, 1)$ هستند.

۵.۳ به‌روزرسانی موقعیت خرگوش

بر اساس معادله (۱۳)، خرگوش سعی خواهد کرد موقعیت خود را نسبت به لانه‌ای که به‌طور تصادفی از میان لانه‌های موجود (d) انتخاب کرده است، به‌روزرسانی نماید. به‌روزرسانی موقعیت خرگوش iام بر اساس دو استراتژی جستجوی علوفه و پنهان شدن تصادفی به‌صورت زیر خواهد بود:

$$\vec{x}_i(t+1) = \begin{cases} \vec{x}_i(t) & f(\vec{x}_i(t)) \leq f(\vec{v}_i(t+1)) \\ \vec{v}_i(t+1) & f(\vec{x}_i(t)) > f(\vec{v}_i(t+1)) \end{cases} \quad (16)$$

این معادله نشان می‌دهد که اگر تناسب موقعیت خرگوش iام بهتر از موقعیت فعلی باشد، خرگوش موقعیت فعلی را رها کرده و در موقعیت کاندید ایجادشده توسط معادله (۱۳) قرار می‌گیرد.

۶.۳ کاهش انرژی

در الگوریتم بهینه‌ساز خرگوش مصنوعی، خرگوش‌ها همیشه تمایل دارند که اغلب در مرحله اولیه تکرارها، جستجوی اکتشافی انجام دهند درحالی‌که اغلب در مرحله بعدی تکرارها مخفی سازی تصادفی انجام می‌دهند. این مکانیسم جستجو، ناشی از انرژی یک خرگوش است که باگذشت زمان به‌تدریج تضعیف می‌شود. بنابراین، احتمال علوفه یابی در فرآیند تکراری حدود ۰٫۵ است. که این فاکتور نشان‌دهنده‌ی تعادل بین مرحله‌ی اکتشاف و بهره‌برداری می‌باشد.

۷.۳ بررسی شرط خاتمه

رویکردهای مختلفی برای خاتمه الگوریتم‌ها مطرح می‌باشد. در این پژوهش از تعداد دفعات اجرا (۱۰۰ دور) برای خاتمه روش پیشنهادی استفاده‌شده است. پس از برقراری شرط خاتمه برترین خرگوش به لحاظ برازندگی، بیانگر بهترین راه‌حل برای تشخیص اجتماع در شبکه‌های پیچیده خواهد بود.

جدول ۱: مجموعه داده‌های تک لایه و واقعی در طرح پیشنهادی

Networks	n	m	C	K	d
Karate	34	78	0.58	4.58	0.139
Dolphins	62	159	0.3	5.12	0.084
Football	115	613	0.4	10.66	0.093
Polbooks	105	441	0.48	8.4	0.068

۴ پیاده‌سازی و مجموعه داده

برای شبیه‌سازی روش پیشنهادی مقاله از نرم‌افزار متلب ۲۰۱۶ استفاده‌شده است. همچنین مجموعه داده پیشنهادی برای این مقاله، مجموعه داده‌های حقیقی شامل داده‌های مربوط به باشگاه کاراته زاخاری، شبکه فوتبال کالج آمریکایی، شبکه دلفین‌ها، شبکه پول بوکز هستند. تمامی این شبکه‌ها دارای ساختار اجتماعی واقعی هستند.

در جدول ۱ پارامتر n نشان‌دهنده تعداد کل گره‌های هر شبکه و پارامتر m نیز بیانگر تعداد کل لبه‌های متصل در هر شبکه است. همچنین پارامتر C معرف میانگین ضریب خوشه‌بندی در هر شبکه و پارامتر K میانگین درجه برای هر خوشه می‌باشد. نهایتاً پارامتر d نیز بیانگر میانگین درجه هر خوشه می‌باشد. روابط (۱۷)، (۱۸) و (۱۹) چولگی محاسبه پارامترهای C، K و d در جدول ۳ را نشان می‌دهند.

$$C = \frac{\sum_i^n C_i}{n} \tag{۱۷}$$

$$K = \frac{\sum_i^n K_i}{n} \tag{۱۸}$$

$$d = \frac{۲m}{n(n-۱)} \tag{۱۹}$$

۵ تجزیه‌وتحلیل

در این بخش، عملکرد روش پیشنهادی مقاله در تشخیص اجتماعات شبکه‌های پیچیده با دو شاخص اطلاعات متقابل نرمال شده و پیمانگی مورد تجزیه‌وتحلیل قرارگرفته است. روابط (۲۰) و (۲۱) این شاخص‌ها را معرفی می‌کنند.

$$\mathrm{NMI}(A,B) = \frac{-۲\sum_{i=۱}^{c_A}\sum_{j=۱}^{c_B} c_{ij}\log\left(\frac{c_{ij}N}{c_{i\cdot}\cdot c_{\cdot j}}\right)}{\sum_{i=۱}^{c_A} c_{i\cdot}\log\left(\frac{c_{i\cdot}}{N}\right)+\sum_{j=۱}^{c_B} c_{\cdot j}\log\left(\frac{c_{\cdot j}}{N}\right)} \tag{۲۰}$$

$$Q = \sum_{k=۱}^{S}\left[\frac{l_k}{L}-\left(\frac{d_k}{۲L}\right)^{۲}\right] \tag{۲۱}$$

جدول ۲: ارزیابی مدل پیشنهادی با شاخص پیمانگی

	Karate	Dolphins	Football	Pollbooks
std	0.07	0.08	0.04	0.05
min	**0.33**	0.38	0.51	0.50
max	0.42	0.52	**0.63**	0.55
mean	0.34	0.44	0.57	0.52

جدول ۳: مقایسه عملکرد طرح پیشنهادی با طرح‌های مشابه در شاخص پیمانگی

	Ref	Karate	Dolphins	Football	Polbooks
MOEA/D-TS	[13]	0.42	0.53	0.60	0.53
MOACO	[14]	0.41	0.52	0.60	0.52
ARO	Our	0.42	0.52	0.63	0.55

یکی از مطرح‌ترین و معتبرترین معیارها در شبکه‌های پیچیده NMI می‌باشد؛ این معیار اطلاعات متقابل نرمال شده نیز نامیده می‌شود؛ یک شاخص ارزیابی خارجی برای شبکه‌های پیچیده بوده و زمانی استفاده می‌شود که ساختار اجتماعِ درست در دست باشد. مقدار NMI در بازه‌ی $[۰, ۱]$ قرار دارد و مقدار بزرگ‌تر برای آن نشان‌دهنده آن است که اجتماعات یافت شده با اجتماعات درست مطابقت بیشتری دارد. در رابطه (۲۰) پارامتر c_{ij} نشان‌دهنده تعداد گره‌های مشترک در اجتماع درست i و اجتماع یافت شده j می‌باشد. با فرض اینکه A و B دو افراز مختلف از شبکه‌ای با N گره باشند و c_A و c_B به ترتیب نشان‌دهنده تعداد افرازهای اجتماعات A و B خواهند بود. همچنین پارامترهای $c_{i.}$ و $c_{.j}$ به ترتیب نشان‌دهنده حاصل جمع عناصر سطر iام و ستون jام هستند. شاخص دیگری که در این مقاله برای ارزیابی مدل پیشنهادی استفاده‌شده است؛ شاخص پیمانگی خواهد بود. توضیحات این شاخص در رابطه (۲) آمده است.

۱.۵ شاخص پیمانگی

در این آزمایش، روش پیشنهادی به تعداد ۳۰ مرتبه بر روی مجموعه داده‌های معرفی‌شده در جدول ۲ اجرا و نتایج ثبت شد. بررسی نتایج نشان می‌دهد که بالاترین مقدار پیمانگی در این آزمایش به‌صورت میانگین برای مجموعه داده Polbooks و پایین‌ترین مقدار آن برای مجموعه داده Karate به‌دست‌آمده است. همان‌طور که در جدول ۲ مشخص‌شده است؛ بالاترین نرخ پیمانگی با روش پیشنهادی مقاله در طول آزمایش‌های مختلف برای مجموعه داده Football با ۰٫۶۳ درصد شده است. این در حالی است که پایین‌ترین نرخ آن نیز برای مجموعه داده Karate با ۰٫۳۳ به‌دست‌آمده است.

مقایسه نتایج در بهترین حالت با الگوریتم‌های مشابه نشان می‌دهد که پیمانگی مدل پیشنهادی در همه داده حقیقی نسبت به طرح‌های مشابه بهبودیافته است. جدول ۳ نتایج این بررسی را نشان می‌دهد.

جدول ۴: بررسی کارایی مدل پیشنهادی با شاخص NMI

	Karate	Dolphins	Football	Polbooks
std	0.05	0.04	0.04	0.06
min	**0.83**	0.88	0.82	0.85
max	**1.00**	**1.00**	0.97	**1.00**
mean	0.90	0.94	0.89	0.91

جدول ۵: مقایسه نتایج مدل پیشنهادی با طرح‌های مشابه در شاخص NMI

	Ref	Karate	Dolphins	Football	Polbooks
MOEA/D-TS	[13]	1.00	1.00	0.97	1.00
MOACO	[14]	1.00	1.00	0.92	0.60
ARO	Our	1.00	1.00	0.97	1.00

۲.۵ شاخص اطلاعات متقابل نرمال شده

در این بررسی روش پیشنهادی به ازای هر یک از مجموعه داده‌های جدول ۳ به تعداد ۳۰ مرتبه اجرا و نتایج ثبت گردید. بررسی نتایج برای نشان می‌دهد که بالاترین نرخ NMI به‌صورت میانگین برای مجموعه داده‌های Polbooks و Dolphins حاصل‌شده است. همچنین پایین‌ترین نرخ آن نیز برای مجموعه داده‌های Football و Polbooks به‌دست‌آمده است. جدول ۴ نتایج این بررسی را نمایش می‌دهد.

بر اساس نتایج به‌دست‌آمده در جدول ۴، بالاترین نرخ NMI به‌دست‌آمده با روش پیشنهادی مقاله در طول آزمایش‌های مختلف برای مجموعه داده‌های Polbooks، Dolphins و Karate حاصل‌شده که مقدار آن برابر با ۱ است. این در حالی است که پایین‌ترین مقدار آن نیز برای مجموعه داده‌های Football و Polbooks با ۰٫۸۳ به‌دست‌آمده است.

مقایسه نتایج طرح پیشنهادی مقاله که یک رویکرد مبتنی بر الگوریتم بهینه‌سازی خرگوش‌های مصنوعی می‌باشد؛ در بهترین حالت بیانگر آن است که نرخ NMI با آن در مجموعه داده‌های حقیقی نسبت به الگوریتم‌ها و مدل‌های مشابه بهبودیافته است. جدول ۵ نتایج روش پیشنهادی و روش‌های مشابه در مجموعه داده‌های حقیقی را با شاخص NMI مقایسه کرده است. بررسی نتایج درج‌شده در جدول ۵ نشان می‌دهد که روش‌های MOACO و MOEA/D-TS نتایج نزدیکی به روش پیشنهادی داشته‌اند. به‌نحوی‌که در مجموعه داده‌های Karate و Dolphins این طرح‌ها همانند روش پیشنهادی از NMI برابر ۱ برخوردند. این در حالی است که الگوریتم MOEA/D-TS در مجموعه داده Polbooks نیز NMI برابر ۱ به دست آورده است. اما در مجموعه داده Football روش پیشنهادی مقاله که یک روش مبتنی بر الگوریتم بهینه‌سازی شاهین طلایی است؛ نسبت به آن نتایج بهتری ارائه می‌دهد.

۶ بحث و نتیجه‌گیری

شبکه پیچیده معرف پارادایم پیچیدگی است که عناصر سازنده آن تشکیل شبکه‌ای را می‌دهند که اجزاء شبکه دارای برهم‌کنش هستند و از اندیشه کل‌نگر بهره می‌گیرند. این شبکه‌ها با چالش‌های مختلفی مواجه هستند که یکی از مطرح‌ترین چالش‌ها شناسایی اجتماع است. روش‌های مختلفی برای تشخیص اجتماع در شبکه‌های پیچیده پیشنهادشده است. اما با توجه به اینکه شناسایی اجتماع همواره در قالب یک مسئله NP-Hard مطرح بوده است؛ بهترین الگو و روش در این زمینه استفاده از الگوریتم‌های فرا اکتشافی می‌باشد. در این مقاله برای برطرف کردن چالش شناسایی اجتماع در شبکه پیچیده از الگوریتم بهینه‌سازی خرگوش‌های مصنوعی استفاده شد. روش پیشنهادی مقاله برای تشخیص جوامع شبکه‌های پیچیده در محیط نرم‌افزاری متلب پیاده‌سازی شده و کارایی آن در چهار مجموعه داده حقیقی با شاخص‌های پیمانگی و NMI مورد تجزیه‌وتحلیل قرار گرفت. تجزیه‌وتحلیل داده‌ها برتری روش پیشنهادی در مقایسه با طرح‌های مشابه را اثبات می‌کند.

۷ کارهای آینده

در این مقاله برای تشخیص اجتماع در شبکه های پیچیده از الگوریتم بهینه سازی خرگوش مصنوعی استفاده شد. این مدل با تکیه بر شاخص پیمانگی در محاسبه برازندگی جمعیت اولیه، نتایج بهتری در مقایسه با مدل های مشابه ارائه می دهد؛ ولی باید در نظر داشت که در تشخیص اجتماع علاوه بر شاخص پیمانگی، شاخص های دیگری نیز برای انتخاب راه حل های مطلوب وجود دارد؛ که می تواند در خروجی مدل موثر باشد. بر همین اساس پیشنهاد می شود در کارهای آینده در تابع برازندگی، تاثیر شاخص های دیگر نیز مورد بررسی قرار گیرد. همچنین می توان مدل را با الگوریتم های بهینه سازی دیگری مانند گورکن عسلخوار، عقاب آکیلا، عقاب طلایی، مرغ های دریایی و ... نیز مورد بررسی قرار داد.

مراجع

[1] C. Li, H. Chen, T. Li, and X. Yang, “A stable community detection approach for complex network based on density peak clustering and label propagation”, Applied Intelligence, vol. 52, no. 2, pp. 1188-1208, 2022.

[2] T. Shaik, V. Ravi, and K. Deb, “Evolutionary multi-objective optimization algorithm for community detection in complex social networks”, SN Computer Science, vol. 2, no. 1, pp. 1-25, 2021.

[3] S. T. Shishavan and F. S. Gharehchopogh, “An improved cuckoo search optimization algorithm with genetic algorithm for community detection in complex networks”, Multimedia Tools and Applications, pp. 1-27, 2022.

[4] S. Taheri and A. Bouyer, “Community detection in social networks using affinity propagation with adaptive similarity matrix”, Big data, vol. 8, no. 3, pp. 189-202, 2020.

[5] A. ŞİMŞEK and K. Resul, "Using swarm intelligence algorithms to detect influential individuals for influence maximization in social networks", Expert Systems with Applications, vol. 114, pp. 224-236, 2018.

[6] L. Wang, L. Ma, C. Wang, N.-g. Xie, J. M. Koh, and K. H. Cheong, "Identifying Influential Spreaders in Social Networks through Discrete Moth-Flame Optimization", IEEE Transactions on Evolutionary Computation, 2021.

[7] S. S. Singh, K. Singh, A. Kumar, and B. Biswas, "ACO-IM: maximizing influence in social networks using ant colony optimization", Soft Computing, vol. 24, no. 13, pp. 10181-10203, 2020.

[8] B. Chatterjee, T. Bhattacharyya, K. K. Ghosh, A. Chatterjee, and R. Sarkar, "A novel meta-heuristic approach for influence maximization in social networks", Expert Systems, p. e12676, 2021.

[9] Y. Zhang, Y. Yong, S. Yang, and T. Zhang, "A New Discrete Grid-Based Bacterial Foraging Optimizer to Solve Complex Influence Maximization of Social Networks", Discrete Dynamics in Nature and Society, vol. 2021, 2021.

[10] A. Zareie, A. Sheikhahmadi, and M. Jalili, "Identification of influential users in social network using gray wolf optimization algorithm", Expert Systems with Applications, vol. 142, p. 112971, 2020.

[11] L. Han, K.-C. Li, A. Castiglione, J. Tang, H. Huang, and Q. Zhou, "A clique-based discrete bat algorithm for influence maximization in identifying top-k influential nodes of social networks", Soft Computing, pp. 1-18, 2021.

[12] D. Bucur, G. Iacca, A. Marcelli, G. Squillero, and A. Tonda, "Multi-objective evolutionary algorithms for influence maximization in social networks", in European conference on the applications of evolutionary computation, 2017, pp. 221-233: Springer.

[13] S. Lotfi and F. Karimi, "A Hybrid MOEA/D-TS for solving multi-objective problems", Journal of AI and Data Mining, vol. 5, no. 2, pp. 183-195, 2017.

[14] C. Mu, J. Zhang, Y. Liu, R. Qu, and T. Huang, "Multi-objective ant colony optimization algorithm based on decomposition for community detection in complex networks", Soft Computing, vol. 23, no. 23, pp. 12683-12709, 2019.