\*\*عنوان:\*\* مرور روش‌های هیوریستیک و متاهیوریستیک برای تشخیص جوامع در شبکه‌های پیچیده: استفاده‌ی کنونی، توسعه‌ی نوظهور و جهت‌گیری‌های آینده

\*\*نویسندگان:\*\* براء عبدالله عطیه، امنه داود عبود، عمار عبدالله حسن، کلارا پیزوتی، میاداه العانی، سعاد اوزدمیر، روا داوود الدباغ

\*\*چکیده:\*\*

شناسایی ساختارهای پنهان در بسیاری از شبکه‌های دنیای واقعی به طور فزاینده‌ای مورد توجه قرار گرفته و مجموعه‌ای گسترده از تکنیک‌ها را در رابطه با مسئله‌ی شناسایی جوامع (Community Detection) امروزی به وجود آورده است. متاهیوریستیک‌های غیرقطعی ثابت کرده‌اند که به طور رقابتی از محدودیت‌های هیوریستیک‌های قطعی در حل مسئله‌ی شناسایی جوامع فراتر می‌روند. با وجود افزایش علاقه، اکثر الگوریتم‌های موجود مبتنی بر متاهیوریستیک برای تشخیص جوامع (MCD) به شیوه‌ای سنتی عمل می‌کنند. به طور کلی، آن‌ها تمایل دارند برخی ویژگی‌های واقعی جوامع را به طور صریح در تعریف‌های مختلف از توابع بهینه‌سازی تک‌هدفه یا چند‌هدفه پیش‌بینی کنند. با این حال، طراحی سایر اپراتورها به صورت سنتی باقی مانده و علاقه‌ی شدیدی برای بازتاب دانش دامنه‌ای در آن‌ها دیده نمی‌شود. علاوه بر این، تمام مرور‌های منتشر شده تا کنون تلاش مستقیمی برای ارتباط دادن روش‌های مبتنی بر هیوریستیک و متاهیوریستیک در شناسایی جوامع انجام نداده‌اند و این روش‌ها را به طور جداگانه بیان کرده‌اند. مروری که در این مقاله معرفی شده است، تلاش می‌کند تا این مسئله را حل کند. به طور عمده، ما به بررسی الگوریتم‌های اصلی مبتنی بر هیوریستیک و متاهیوریستیک در تشخیص جوامع پرداخته‌ایم. سپس دو طبقه‌بندی جدید برای الگوریتم‌های تشخیص جوامع معرفی می‌کنیم: هیبرید متاهیوریستیک و هیوریستیک فراگیر که می‌تواند به عنوان مبنایی برای طراحی مجموعه‌ای از الگوریتم‌های جدید و موثرتر برای تشخیص جوامع به کار گرفته شود.

بله، حتماً. در ادامه، ترجمه بخش‌های بعدی مقاله به زبان فارسی ارائه شده است:

---

\*\*1. مقدمه\*\*

ظهور بسیاری از شبکه‌های پیچیده برای مدل‌سازی پدیده‌های مختلف در زندگی روزمره ما، نیاز به رویکردهای قابل اطمینان برای درک و تحلیل آن‌ها را نشان می‌دهد. شبکه‌های اجتماعی، شبکه‌های همکاری علمی، شبکه‌های تعامل پروتئین-پروتئین (PPI)، شبکه‌های متابولیک، شبکه‌های ارتباطی و حمل و نقل، شبکه‌های فناوری، شبکه‌های اکولوژیکی، شبکه‌های غذایی و شبکه‌های اطلاعاتی نمونه‌های معمولی از این دست هستند.

جوامع در چنین شبکه‌هایی می‌توانند گروه‌هایی از افراد، صفحات وب، مسیرهای بیوشیمیایی و غیره باشند. ارزش قابل استفاده ساختار جوامع در این شبکه‌های پیچیده اغلب به عنوان واحدهای سازمانی تعبیر می‌شود. به عنوان مثال، امروزه نمونه‌ای با اهمیت خاص تقریباً در سراسر جهان شبکه اجتماعی ما و اتصالات متنوع افرادی است که اثبات شده است به بیماری کروناویروس (COVID-19) مبتلا شده یا نشده‌اند (یا حتی بررسی نشده‌اند) تا جوامع سالم را از جوامع مبتلا شناسایی کنیم و به افراد و جوامع در مورد قرنطینه پزشکی اجباری، قرنطینه خانگی، یا آزادی فعالیت مشاوره دهیم. مثال‌های دیگری از سازمان‌های مدولار، شبکه‌های زیستی مانند شبکه‌های تعامل پروتئین-پروتئین، شبکه‌های تنظیم ژن و شبکه‌های متابولیک هستند. به عنوان مثال، ماژول‌های سلولی پروتئین‌ها می‌توانند هم به عنوان مجتمع‌ها و هم به عنوان ماژول‌های عملکردی در نظر گرفته شوند. با توجه به اینکه پروتئین‌ها و کل ماژول‌ها ممکن است عملکردهای زیستی متفاوتی داشته باشند، جوامع آن‌ها می‌توانند هم‌پوشانی داشته باشند. علاوه بر این، جوامع ژن‌ها در شبکه‌های هم‌وقوعی ژن می‌توانند به یافتن ژن‌های مرتبط با یک بیماری خاص کمک کنند. جوامع نویسندگان در شبکه‌های استنادی می‌توانند روابط بین رشته‌های مختلف را نشان دهند. چنین سیستم‌های طبیعی و مصنوعی را می‌توان به طور دقیق به عنوان گره‌ها (اشیای سیستم) و لبه‌هایی که این اشیا را در ویژگی‌های توپولوژیکی غیر معمول، یعنی پیچیده، متصل می‌کنند، نشان داد.

اگرچه پیچیده است، یک ویژگی کلیدی سیستم‌های شبکه‌ای تمایل کلی به گروه‌بندی نهادهای خود در ساختارهای پنهان است که معمولاً به عنوان جوامع شناخته می‌شوند.

\*\*مسئله‌ی تشخیص جامعه (Community Detection)\*\*

مسئله‌ی تشخیص جامعه یک مسئله‌ی بهینه‌سازی است که به دنبال یک تکنیک تحلیل داده است تا ساختار پنهان مجموعه داده‌های بزرگ شبکه‌ای را به خوشه‌های مجزا و فشرده کشف کند. تعداد و اندازه‌ی زیرگروه‌ها ناشناخته است.

\*\*مسئله‌ی تشخیص جامعه به عنوان یک مسئله‌ی NP-hard\*\*

مسئله‌ی تشخیص جامعه ثابت شده است که یک مسئله‌ی NP-hard (زمانی چندجمله‌ای غیردقیق) است. به طور کلی، کار بر روی این مسئله شامل طبقه‌بندی رویکردهای موجود به دو جریان اصلی است: الگوریتم‌های مبتنی بر هیوریستیک (HCD) و الگوریتم‌های مبتنی بر متاهیوریستیک (MCD). به طور کلی، الگوریتم‌های هیوریستیک می‌توانند به عنوان یک جریان کنترلی از قوانین به خوبی تعریف شده‌ای که به صورت سیستماتیک فضای جستجو برای مسئله را بررسی می‌کند تا به یک راه حل قابل قبول با کمترین مصرف زمان برسد، تعریف شوند. از طرف دیگر، متاهیوریستیک‌ها الگوریتم‌های جستجوی جهانی هستند که پایه‌های مشترکی دارند. در حالی که تأکید بر نقش نمایش راه حل است، این الگوریتم‌ها فضای جستجوی مسئله را به صورت ضمنی به نواحی مختلف تقسیم می‌کنند. آن‌ها با استفاده از توابع هدف تک‌هدفه یا چند‌هدفه، نواحی و نمونه‌های مختلف فضای جستجو را ارزیابی می‌کنند و سعی می‌کنند بر اساس کیفیت نمونه‌های جستجو شده، به نواحی مورد علاقه نزدیک‌تر شوند. نمونه‌های جدید از فضای جستجو می‌توانند با استفاده از عملگرهای تغییر (تعامل/ترکیب نمونه‌ها و جهش نمونه‌ها) نیز بازدید شوند. به طور کلی، پایه‌گذاری این متاهیوریستیک‌ها بر دو ستون مکمل استوار است: سرعت همگرایی (یعنی توانایی جستجوی خطی یا مسیری) و قابلیت اطمینان همگرایی (یعنی توانایی جستجوی حجمی).

\*\*بازنگری‌های موجود\*\*

به دلیل رشد نمایی فضای جستجو با توجه به اندازه‌ی شبکه، الگوریتم‌های دقیق/قطعی برای تشخیص جوامع در شبکه‌های بزرگ به اندازه‌ی کافی امکان‌پذیر نیستند. این مسئله منجر به توسعه‌ی الگوریتم‌های جدید بهینه‌سازی غیردقیق با فرمولاسیون‌های مختلف شده است. در میان رویکردهای پیشنهادی، الگوریتم‌های متاهیوریستیک، به ویژه الگوریتم‌های تکاملی (EA)، در حالی که از عملکرد قدرتمندی برخوردارند، توجه بیشتری به خود جلب می‌کنند.

اولین تلاش‌های غنی برای مرور رویکردهای تشخیص جامعه توسط Fortunato ارائه شده است تا الگوریتم‌های هیوریستیک مختلف را به تجزیه‌ی گراف مرتبط سازد. این مرور نسبتاً کامل با دنباله‌ای از مرور‌ها دنبال شده است که به طور گسترده مسئله را از دیدگاه‌های مختلف مورد بحث قرار می‌دهند، اما عمدتاً به الگوریتم‌های تشخیص جامعه مبتنی بر هیوریستیک و متاهیوریستیک تقسیم شده‌اند.

\*\*مشارکت‌ها\*\*

در این مقاله، سعی می‌کنیم محققان را تشویق کنیم تا طراحی الگوریتم‌های MCD را به دقت بازنگری کنند و آن‌ها را به گونه‌ای انعطاف دهند که مفاهیم اساسی IDL را پوشش دهند. برای این منظور:

- ما مرور جامعی از پیشرفت‌های اصلی و به روز تحقیقات در زمینه‌ی کارهای هیوریستیک و متاهیوریستیک که در ادبیات برای حل مسئله‌ی تشخیص جامعه توسعه یافته‌اند، ارائه می‌دهیم. در طول بحث، اشاره می‌کنیم که جنبه‌ی مهمی از ویژگی‌های طراحی در الگوریتم‌های MCD تقریباً در ادبیات به طور کامل نادیده گرفته شده است.

- به جای ترویج یک هیوریستیک یا متاهیوریستیک خاص از میان الگوریتم‌های مختلف موجود در مجموعه‌ی ادبیات، هدف این مقاله بررسی روش‌های بالقوه در حل مسئله‌ی تشخیص جامعه با استفاده از دیدگاه‌های مختلف همکاری بین این دو نوع الگوریتم است.

بسیار خوب، در ادامه ترجمه‌ی بخش‌های بیشتری از مقاله را ارائه می‌دهم:

---

\*\*1.4 سازماندهی مقاله\*\*

در ادامه‌ی این مرور، ابتدا یک توضیح مختصر در مورد مدل‌سازی شبکه‌های پیچیده و یک طبقه‌بندی مرتبط، مجموعه داده‌ها و معیارهای ارزیابی ارائه می‌شود. سپس به مرور تعدادی از الگوریتم‌های شناخته شده مبتنی بر هیوریستیک و متاهیوریستیک برای تشخیص جامعه پرداخته می‌شود. در ادامه، چهارچوب‌های جدیدی پیشنهاد می‌شود تا طراحی الگوریتم‌های تشخیص جامعه به گونه‌ای سازنده‌تر از محدودیت‌های طراحی الگوریتم‌های موجود فراتر رود. درک نیاز به چهارچوب‌های معنادارتر نیز با چالش‌هایی که باید به دقت مورد توجه قرار گیرد، همراه است. در نهایت، به جمع‌بندی نکات اصلی توصیه شده در این مقاله می‌پردازیم.

\*\*2. یک طبقه‌بندی برای مسئله‌ی تشخیص جامعه و مجموعه داده‌ها\*\*

در شکل 2، یک طبقه‌بندی کلی برای انواع مختلف شبکه‌های پیچیده‌ای که در ادبیات تشخیص جامعه به آن‌ها پرداخته شده است، ارائه شده است. هدف از این طبقه‌بندی، که در شکل 2 نشان داده شده است، این است که تمامی جنبه‌هایی را که در توصیف شبکه‌های پیچیده ضروری هستند و سطوح مختلفی از دشواری را بیان می‌کنند، در بر بگیرد. چهار مؤلفه شبکه‌های پیچیده را توصیف می‌کنند: نوع داده‌های شبکه، جوامع، گره‌ها و ارتباطات. این مؤلفه‌ها به طور فردی و/یا مشترک برای تنظیم سطوح مختلف دشواری برای مسئله‌ی تشخیص جامعه عمل می‌کنند و به نوبه خود، ما را به سمت طراحی مدل‌های تشخیص مناسب هدایت می‌کنند.

\*\*3. چارچوب‌های جدید برای تشخیص جامعه\*\*

در این بخش، به معرفی چهارچوب‌های جدیدی می‌پردازیم که می‌توانند برای طراحی الگوریتم‌های تشخیص جامعه به صورت سیستماتیک و با توجه به کنترل عناصر IDL به کار روند. این چهارچوب‌ها شامل الگوریتم‌های ترکیبی متاهیوریستیک (HMCD) و الگوریتم‌های هیوریستیک فراگیر (HHCD) هستند. این چهارچوب‌های جدید می‌توانند به عنوان پایه‌ای برای طراحی الگوریتم‌های جدید و مؤثرتر تشخیص جامعه عمل کنند.

---

ترجمه به پایان رسید، اگر نیاز به بخش‌های دیگری دارید یا اگر موردی خاصی را مد نظر دارید که نیاز به ترجمه آن دارید، لطفاً به من اطلاع دهید.