

ارائه روشی جدید برای پیشگویی پیوند بین رأس‌های موجود در شبکه‌های اجتماعی

اعظم کی‌پور^۱، مرتضی براری^۲، حسین شیرازی^۳

چکیده: امروزه شبکه‌های اجتماعی برخط، به دلیل امکان ایجاد ارتباط بین افراد مختلف در سرتاسر دنیا محبوبیت زیادی دارند. این شبکه‌های اجتماعی که از امکاناتی مانند پیشنهاد دوست به کاربران برخوردارند، در اغلب موارد برای بیان پیشنهادهای خود، از ویژگی‌های محلی ساختار گراف شبکه استفاده می‌کنند. برای بیان پیشنهاد در این شبکه‌ها، روش‌های مختلفی با دو رویکرد محلی و سراسری پیمایش گراف شبکه پیشنهاد شده و مورد استفاده قرار می‌گیرد. در این مقاله روشی با رویکرد محلی ارائه شده است که در مقایسه با روش‌های دیگر، دارای کارایی مناسبی است، علاوه بر اینکه به دلیل محلی بودن رویکرد، از سرعت قابل قبولی برخوردار است. بر اساس ارزیابی‌ای که روی مجموعه داده‌های دو شبکه اجتماعی بزرگ Facebook و Epinions انجام شده است، این ویژگی جدید می‌تواند پیشگویی خوبی برای یال‌هایی انجام دهد که قرار است در آینده شکل بگیرند و در نتیجه پیشنهادهای قابل قبولی را ارائه دهد.

واژه‌های کلیدی: پیشگویی پیوند، سیستم‌های پیشنهاددهنده، شبکه‌های اجتماعی.

۱. دانشجوی کارشناسی ارشد هوش مصنوعی، دانشگاه صنعتی مالک اشتر، تهران، ایران

۲. استادیار مخابرات سیستم، دانشگاه صنعتی مالک اشتر، تهران، ایران

۳. دانشیار هوش مصنوعی، دانشگاه صنعتی مالک اشتر، تهران، ایران

تاریخ دریافت مقاله: ۱۳۹۲/۰۸/۲۵

تاریخ پذیرش نهایی مقاله: ۱۳۹۳/۰۴/۳۰

نویسنده مسئول مقاله: اعظم کی‌پور

E-mail: a.keypour@gmail.com

مقدمه

شبکه‌های اجتماعی شبکه‌های پویایی‌اند که همواره اعضا و ارتباطات و پیوندهای بین آنها رو به افزایش است. زنجیره این پیوندها گاهی به دلیل فرآیند ایجاد ناقص یا به دلیل اینکه هنوز در این شبکه‌ها انعکاس نیافته‌اند (برای مثال، دوستان دنیای واقعی که یک ارتباط اجتماعی مجازی ایجاد نکرده‌اند)، پاره شده و از دست می‌روند (فایر و دیگران، ۲۰۱۱)، لذا یکی از مسائل مهم در شبکه‌های اجتماعی، مسئله پیشگویی پیوند است.

مسئله پیشگویی، یکی از جنبه‌های مهم داده‌کاوی است. داده‌کاوی پیشگویی‌کننده، مدلی از سیستم ارائه می‌کند که این مدل را مجموعه‌ای از داده‌های مشخص، پیش‌بینی می‌کند و هدف کلی آن ایجاد الگویی برای طبقه‌بندی، پیش‌بینی و تخمین داده‌ها است (رادفر، نظافتی و یوسفی اصل، ۲۰۱۴).

پیشگویی پیوند، به معنای وجود و عدم وجود یک پیوند یا ارتباط در آینده بین دو رأس یک شبکه اجتماعی است و یک ابزار مهم برای تحلیل شبکه‌های اجتماعی به‌شمار می‌رود که کاربردهای زیادی نیز در حوزه‌های دیگری چون، بازیابی اطلاعات، بایوانفورماتیک و تجارت الکترونیک دارد. به علاوه در حوزه علم وب و اینترنت، می‌تواند در کارهایی مانند ایجاد ابرپیوند خودکار وب و پیش‌بینی ابرپیوند سایت‌های وب کاربرد داشته باشد (الحسن و زکی، ۲۰۱۱).

یکی از مهم‌ترین کاربردهای پیشگویی پیوند در تجارت الکترونیک، ایجاد سیستم‌های پیشنهاد دهنده است. این پیشنهادها می‌توانند شامل پیشنهاد کالا یا پیشنهاد دوست باشند. از جمله فواید پیشگویی پیوند اینکه با پیشنهادها مناسب در زمینه کالا، می‌توان ضریب خرید اینترنتی افراد جامعه را بالا برد. عمل خرید تحت تأثیر بسیاری از خصایص مشتریان، مانند ویژگی‌های شخصیتی، سبک زندگی، دانش و مهارت‌ها، عوامل اجتماعی، عوامل روانی و عوامل جمعیتی قرار دارد (حسنقلی‌پور، امیری، فهیم و قادری عابد، ۲۰۱۴). خرید اینترنتی براساس تجربه واقعی از خرید کالا صورت نمی‌گیرد، بلکه براساس ظواهری مانند تصویر، شکل، اطلاعات کیفی و تبلیغات از کالا استوار است؛ به همین دلیل جلب اعتماد مشتری برای انجام مبادلات از طریق اینترنت، مورد توجه بسیاری از سازمان‌ها و مشتریان قرار گرفته است (ساجدی‌فر، اسفیدانی، وحدت‌زاد و محمودآذر، ۲۰۱۲).

یکی از راه‌های جلب اعتماد مشتری، به واسطه اعتماد به افرادی حاصل می‌شود که با وی در یک شبکه اجتماعی در ارتباط بوده و از کالای مورد نظر استفاده می‌کنند. به همین دلیل،

هم‌پیوندی به معنی تمایل اعضا برای پیوند و ارتباط با یکدیگر، به‌منزلهٔ ویژگی مهمی برای یکپارچه‌سازی اطلاعات یاد شده است که موجب پایداری کسب‌وکارها می‌شود (ایرانی و حقیقی، ۲۰۱۴). در این زمینه پیشگویی پیوند با تحلیل ارتباطات موجود در شبکه‌های اجتماعی، می‌تواند فرآیند پیشنهاد کالا و خدمات را تسهیل کند.

از جنبهٔ دوست‌یابی نیز، شبکه‌های اجتماعی برخطی مانند فیس‌بوک دارای حجم انبوهی از داده هستند که می‌توانند با جست‌وجو در میان آنها و تحلیل اطلاعات موجود، در این مورد که چه کسی ممکن است بخواهد با دیگری دوست شود، پیشگویی کنند و بر اساس این پیشگویی‌ها به کاربران پیشنهادهای مناسبی ارائه دهند.

ساختار شبکه‌های اجتماعی، به‌صورت گرافی مدل می‌شود که رأس‌های گراف همان کاربران شبکه هستند و ارتباطات بین این کاربران، یال‌های گراف را شکل می‌دهند. برای ارائهٔ پیشنهاد در این شبکه‌ها دو رویکرد کلی وجود دارد. رویکرد اول که مبتنی بر ویژگی‌های محلی ساختار گراف شبکه است، روشی است که معمولاً در شبکه‌های اجتماعی برخط مورد استفاده قرار می‌گیرد. این شبکه‌ها معمولاً پیشگویی‌های خود را بر اساس تعداد دوستان مشترکی که دو کاربر دارند، انجام می‌دهند؛ زیرا احتمال بسیار زیادی وجود دارد که دو کاربری که دوستان مشترک زیادی دارند، تمایل داشته باشند با یکدیگر چه در دنیای مجازی و چه در دنیای واقعی دوست شوند. اشکالی که به این روش‌ها وارد است، به در نظر گرفتن حداکثر فاصلهٔ ۲ بین هر دو کاربر در گراف شبکه برمی‌گردد که ممکن است از دقت کافی برخوردار نباشند.

در مقابل روش‌های با رویکرد محلی، مانند روش دوست - دوست یا دوستان مشترک (الحسن و زکی، ۲۰۱۱)، آدامیک - آدار (آدامیک و آدار، ۲۰۰۵)، ضریب جاکارد (الحسن و زکی، ۲۰۱۱) و... روش‌های سراسری نیز وجود دارند که برای یافتن میزان شباهت بین دو فرد، کل گراف شبکه را پیمایش می‌کنند. از آنجایی که این روش‌ها کل اطلاعات شبکه را در نظر می‌گیرند، ممکن است از دقت زیادی برخوردار باشند، ولی بار محاسباتی بالایی دارند و طبیعتاً برای شبکه‌های اجتماعی برخط مناسب نیستند. از جمله روش‌هایی که مبتنی بر این رویکرد هستند، معیار کاتز (لیبن ناول و کلینبرگ، ۲۰۰۳)، گام تصادفی با شروع مجدد^۱ (بکستورم و لسکائوس، ۲۰۱۱) و simrank (جه و ویدوم، ۲۰۰۲) را می‌توان نام برد.

در این نوشتار روشی ارائه خواهد شد که با استفاده از رویکردی محلی، در تهیهٔ یک معیار جدید برای پیشنهاد دوست و پیشگویی پیوند، در کنار بهره‌مندی از مزیت سرعت روش‌های محلی، از کارایی مناسبی نسبت به سایر روش‌های موجود، برخوردار است.

این مقاله به این صورت ساماندهی شده است؛ در ادامه مطالب فعالیت‌ها، کارهای مرتبط و روش‌های موجود پیشگویی پیوند بیان شده است. در بخش بعدی روش پیشنهادی شرح داده می‌شود و پس از آن نتایج تجربی مربوط به روش پیشنهادی بیان خواهد شد و با یک نتیجه‌گیری از مطالب گفته‌شده به پایان می‌رسد.

پیشینه پژوهش

ویژگی‌ها و معیارهای مختلفی برای پیشگویی پیوند بین رأس‌های موجود در شبکه‌های اجتماعی وجود دارد. برخی از این روش‌ها در استخراج اطلاعات از گراف ارتباطات شبکه‌های اجتماعی، رویکردی محلی دارند و برخی از رویکردی سراسری پیروی می‌کنند. در ادامه تعدادی از این معیارها را که معمولاً در شبکه‌های اجتماعی به کار برده می‌شوند، معرفی کرده و سپس به ارزیابی آنها پرداخته خواهد شد.

معیار دوستان مشترک: به منزله یکی از معیارهای محلی می‌توان از روش دوست - دوست^۱ یا دوستان مشترک^۲ (الحسن و زکی، ۲۰۱۱) نام برد که در شبکه‌هایی چون فیس‌بوک^۳ مورد استفاده قرار می‌گیرند. در این روش رأس‌هایی که با طول مسیر ۲ به هم مرتبط هستند، کاندیدای پیشنهاد دوست در نظر گرفته شده و پیشنهاد می‌شوند و آن رأس‌هایی که با تعداد مسیر بیشتری با یکدیگر ارتباط دارند (دوستان مشترک بیشتری دارند)، دارای احتمال بالاتری برای ایجاد پیوند هستند.

معیار آدامیک آدار: معیار آدامیک آدار^۴ (آدامیک و آدار، ۲۰۰۵) از معیار شباهتی استفاده می‌کند که در اصل برای یافتن ارتباطات قوی بین صفحات وب به کار می‌رود و مربوط به تعداد ویژگی‌های مشترکی است که دو صفحه به اشتراک گذاشته‌اند. در مسئله پیشگویی پیوند، این ویژگی مشترک همان همسایه مشترک دو رأس است. میزان شباهت بین دو رأس در این روش از رابطه ۱ به دست می‌آید.

$$score(u, v) = \sum_{z \in \Gamma(u) \cap \Gamma(v)} \frac{1}{\log |\Gamma(z)|} \quad \text{رابطه ۱}$$

در این رابطه، Z : همسایه مشترک دو رأس u و v و $|\Gamma(z)|$: درجه رأس Z است.

-
1. FOAF
 2. Common friends
 3. www.facebook.com
 4. Adamic-Adar

معیار کاتز: معیار کاتز^۱ (لیبن ناول و کلینبرگ، ۲۰۰۳) از دسته معیارهای سراسری به‌شمار می‌رود و مبتنی بر معیار کوتاه‌ترین مسیر است. برای به‌دست آوردن میزان شباهت بین دو رأس، تمام مسیرهای موجود بین آن دو را بررسی می‌کند و به کمک رابطه^۲ محاسبه می‌شود.

$$katz(u, v) = \sum_{l=1}^{l_{max}=\infty} \beta^l |path_{u,v}^l| \quad \text{رابطه ۲}$$

که در آن $|path_{u,v}^l|$ به تعداد مسیرهایی اشاره می‌کند که با طول ۱ بین u و v وجود دارد و β مقداری بین صفر و ۱ دارد. ایرادی که به این معیار وارد است، مرتبه زمانی چندجمله‌ای درجه سه آن است، به همین دلیل نمی‌توان برای شبکه‌های بزرگ از آن استفاده کرد.

ضریب جاکارد: ضریب جاکارد^۲ (الحسن و زکی، ۲۰۱۱) شباهت بین مجموعه نمونه‌ها را اندازه‌گیری می‌کند و در واقع، اندازه ارتباطات تقسیم بر اندازه اجتماع مجموعه‌های نمونه تعریف می‌شود. بنابراین ضریب جاکارد که نسبت بین دوستان مشترک u و v به کل دوستان این دو است، به صورت زیر محاسبه می‌شود (رابطه^۳).

$$jaccard - coef(u, v) = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|} \quad \text{رابطه ۳}$$

در این رابطه $\Gamma(u)$ مجموعه همسایه‌های رأس u است.

امتیاز الحاق ترجیحی: امتیاز الحاق ترجیحی^۳ (کوکیوسکی، هامنر و یانگ، ۲۰۱۱) بر این ایده استوار است که کاربران با دوستان زیاد، به ایجاد ارتباطات بیشتر در آینده تمایل دارند. این معیار که در موارد بسیاری پاسخ‌های مثبتی داشته است، از رابطه^۴ محاسبه می‌شود.

$$preferential - attachment - score(u, v) = |\Gamma(u)| \cdot |\Gamma(v)| \quad \text{رابطه ۴}$$

اندیس ترفیع هاب: اندیس ترفیع هاب (HPI)^۴ (زیو، لو، زنگ و زو، ۲۰۱۲) برای تعیین کیفیت هم‌پوشانی توپولوژیک جفت لایه‌ها در یک شبکه دگرگون‌شونده به کار می‌رود و به صورت رابطه^۵ تعریف می‌شود.

-
1. Katz
 2. Jaccard Coefficient
 3. Preferential attachment
 4. Hub Promoted Index

$$HPI(u, v) = \frac{|\Gamma(u) \cap \Gamma(v)|}{\min\{|\Gamma(u)|, |\Gamma(v)|\}} \quad \text{رابطه ۵}$$

اندیس فشردۀ هاب: اندیس فشردۀ هاب (HDI) (زیو، لو، زنگ و زو، ۲۰۱۲) مشابه اندیس بالاست، با این تفاوت که مقدار حداکثر درجه‌ها را در نظر می‌گیرد (رابطه ۶).

$$HDI(u, v) = \frac{|\Gamma(u) \cap \Gamma(v)|}{\max\{|\Gamma(u)|, |\Gamma(v)|\}} \quad \text{رابطه ۶}$$

معیار دوستان: معیار دوستان (فایر و دیگران، ۲۰۱۱) با در نظر گرفتن همهٔ همسایه‌های دو کاربر، به معنای تعداد اتصالات بین همسایه‌های u و v است که برای دستیابی به این تعداد از رابطه ۷ استفاده می‌شود.

$$friends - measure(u, v) = \sum_{x \in \Gamma(u)} \sum_{y \in \Gamma(v)} \delta(x, y) \quad \text{رابطه ۷}$$

$$\delta(x, y) = \begin{cases} 1 & \text{if } x = y \text{ or } (x, y) \in E \text{ or } (y, x) \in E \\ 0 & \text{otherwise} \end{cases}$$

معیار وزن یال: معیار وزن یال^۲ (کوکیرسکی و همکاران، ۲۰۱۱)، ابتدا به صورت دو ویژگی جداگانه برای هر یک از دو سر لبه محاسبه می‌شود (رابطه ۸).

$$w(v) = \frac{1}{\sqrt{1 + |\Gamma(v)|}} \quad \text{رابطه ۸}$$

$$w(u) = \frac{1}{\sqrt{1 + |\Gamma(u)|}}$$

حال وزن یال بین دو رأس u و v می‌تواند از طرق مختلف محاسبه شده و مورد بهره‌برداری قرار گیرد و دو معیار زیر را شکل دهد.

جمع وزن‌ها: جمع وزن‌ها (کوکیرسکی و همکاران، ۲۰۱۱) برابر است با جمع دو وزن تعریف شده در رابطه ۸:

1. Hub Depressed Index
2. Edge weight

$$w(u, v) = w(u) + w(v) \quad \text{رابطه ۹}$$

ضرب وزن‌ها: ضرب وزن‌ها (کوکیرسی و همکاران، ۲۰۱۱) برابر است با ضرب دو وزن تعریف شده در رابطه ۸:

$$w(u, v) = w(u) \times w(v) \quad \text{رابطه ۱۰}$$

ضریب خوشه‌بندی: ضریب خوشه‌بندی^۱ (فیزا، بیکدش و لبی، ۲۰۱۱) برابر است با نسبت تعداد همسایگان مشترک بین دو رأس بر مقدار حداقل درجه این دو رأس و مقدار آن به کمک رابطه ۱۱ محاسبه می‌شود.

$$\text{رابطه ۱۱}$$

در رابطه ۱۱ a_{uz} درایه ماتریس مجاورت گراف شبکه است.

همبستگی درجه: همبستگی درجه^۲ (فیزا و همکاران، ۲۰۱۱) متناسب با ضریب همبستگی پیرسون است، به شکل زیر تعریف می‌شود:

$$\text{degree - correlation}(u, v)$$

$$= \frac{4 \cdot |\Gamma(u)| \cdot |\Gamma(v)| - |\Gamma(u)| - |\Gamma(v)|}{2 \cdot |\Gamma(u)|^2 + 2 \cdot |\Gamma(v)|^2 - |\Gamma(u)| - |\Gamma(v)|} \quad \text{رابطه ۱۲}$$

روش Friend Link: روش Friend Link (پاپادیمتریو، سیمونیدیس و مانولویوس، ۲۰۱۱) با در نظر گرفتن یک کران، مسیرهای کوتاه‌تر از این کران را بین دو رأس بررسی می‌کند و این گونه از مزیت سرعت روش‌های محلی و دقت روش‌های سراسری بهره می‌برد. رابطه ۱۳ روش محاسبه تعیین میزان شباهت دو رأس در این روش را نشان می‌دهد.

$$\text{friend - link}(u, v) = \sum_{i=2}^l \frac{1}{(i-1)} \cdot \frac{|\text{paths}_{u,v}^i|}{\prod_{j=2}^i (n-j)} \quad \text{رابطه ۱۳}$$

که در آن $|\text{path}_{u,v}^i|$ تعداد مسیرهای با طول i بین دو رأس است.

1. Clustering co-efficient
2. Degree correlation

علاوه بر معیارهای مذکور، معیارهای دیگری نیز وجود دارند که به علت تشابه و نزدیکی نتایج آنها با معیارهای بیان شده، از معرفی آنها خودداری شده است.

روش‌شناسی پژوهش

معیار آدمیک - آدار ایده مناسبی برای تعیین میزان شباهت کاربران شبکه‌های اجتماعی ارائه می‌دهد و توانسته است پیشگویی‌های خوبی را به انجام رساند. ایده موجود در این معیار این است که اگر همسایه مشترک دو کاربر دوستان کمتری داشته باشد (درجه کمتر) معلوم است گزینه‌تر انتخاب می‌کند و احتمال ایجاد ارتباط قوی‌تری بین دو رأس مورد نظر وجود دارد. لذا به درجه هر یک از همسایگان مشترک، وزن معکوسی داده است تا چنانچه همسایه مشترک دارای درجه بالاتری بود، وجود این همسایه واسطه اثر کمتری بر نتیجه پیشنهادها داشته باشد.

نکته گفتنی در این روش، اینکه ممکن است رأس مشترکی که دوستان زیادی دارد، در بسیاری از این دوستان با دو سر ارتباط (دو کاربری که در حال بررسی میزان شباهت آنها هستیم) اشتراک داشته باشد. به بیان دیگر، داشتن دوستان زیاد برای یک رأس واسطه یک امتیاز منفی محسوب می‌شود، اما در صورتی که این دوستان، خود دوستان مشترک یکی از دو کاربری باشند که قصد پیشنهاد دوست به آنها را داریم، این تعداد دوستان باید امتیاز مثبت تلقی شود. در این بخش، ما یک معیار تعیین شباهت بین جفت رأس‌های گراف شبکه پیشنهاد داده‌ایم که با در نظر گرفتن درجه رأس‌های میانی و نیز دوستان مشترک رأس میانی با دو رأس اصلی، این میزان شباهت را با توجه به معادله زیر بیان می‌کند.

$$new - measure(u, v) = \sum_{z \in \Gamma(u) \cap \Gamma(v)} \frac{cn(u, z) + cn(z, v)}{|\Gamma(z)|} \quad \text{رابطه ۱۴}$$

که در آن $Cn(u, z)$ تعداد همسایه‌های مشترک دو رأس u و z است. رابطه ۱۴ برای تعداد بالای دوستان مشترک رأس میانی با دو رأس اصلی، امتیاز مثبت قائل می‌شود و به همین دلیل، مجموع این تعداد در صورت معادله قرار داده شده است تا میزان شباهت را افزایش دهد. از سوی دیگر، برای درجه رأس میانی امتیاز منفی در نظر گرفته شده و در مخرج معادله جای می‌گیرد.

در مقایسه‌هایی که بین این روش و روش‌های معرفی شده صورت گرفته است، مشاهده می‌شود که این معیار کارایی قابل قبولی نسبت به روش‌های دیگر دارد که در ادامه این کارایی اثبات می‌شود.

یافته‌های پژوهش

در این بخش الگوریتم پیشنهادی این پژوهش با سیزده ویژگی شرح داده شده، مقایسه می‌شود و کارایی آن به صورت تجربی اثبات خواهد شد. برای این کار، از دو مجموعه داده واقعی Epinions^۱ و Facebook^۲ استفاده شده است.

مجموعه داده‌های شبکه اجتماعی Epinions که یک شبکه اجتماعی «چه کسی به چه کسی اعتماد می‌کند» به شمار می‌رود، در نظر گرفته شده است و از ۴۹ k کاربر و ۴۸۷ k یال بین جفت رأس‌ها تشکیل شده است. مجموعه داده‌های Facebook نیز که در ۳۰ اکتبر ۲۰۰۹ جمع‌آوری شده (پاپادیمیترو و همکاران، ۲۰۱۱) شامل ۳/۷ k کاربر و ۱۳/۷ k یال است.

برای ارزیابی و مقایسه روش‌ها و الگوریتم‌های مختلف پیشگویی پیوند با روش پیشنهادی، از معیار MAP^۳ استفاده شده است، تا در بررسی کارایی الگوریتم مورد نظر، روی ترتیب دوستان پیشنهاد داده شده با این روش‌ها، تأکید بیشتری شود. معادله MAP (پاپادیمیترو و همکاران، ۲۰۱۱) به صورت رابطه ۱۵ تعریف می‌شود.

$$MAP = \frac{1}{|N|} \sum_{u=1}^{|N|} \frac{1}{r_u} \sum_{k=1}^{r_u} Precision_u@k \quad (\text{رابطه ۱۵})$$

در این رابطه؛ N: تعداد کاربران در مجموعه داده آزمایشی؛ r_u : تعداد کاربران مرتبط با کاربر u در مجموعه آزمایشی و $Precision_u@k$: مقدار precision در kامین موقعیت فهرست پیشنهادها برای u است. توجه کنید که MAP در واقع هر دو مقدار precision و recall را در خود دارد و از نظر هندسی، در زیر منحنی precision-recall جای دارد.

به منظور ارزیابی روش‌ها، از هر مجموعه آموزشی به صورتی کاملاً تصادفی، ۱۰۰۰ رأس انتخاب شد و با در نظر گرفتن همه ارتباطات میان آنها، پس از محاسبه و مقایسه مقادیر معیارها، آزمایش‌های لازم روی آنها انجام گرفت. از آنجایی که با این انتخاب همه رأس‌ها، حتی رأس‌هایی که هنوز دوستی ندارند و به اصطلاح افراد تازه‌وارد این شبکه اجتماعی به شمار می‌روند نیز انتخاب می‌شوند، این امکان وجود دارد که در نتایج همه انواع روش‌های معرفی شده، با دقت پایینی مواجه شویم؛ چراکه این نوع رأس‌ها، با مشکلی به نام آغاز سرد^۴ برخورد می‌کنند که یک

1. <http://www.trustlet.org/wiki/>
 2. <http://delab.csd.auth.gr/~symeon>
 3. Mean Average Precision
 4. Cold Start

مشکل رایج در ارائه پیشنهادها شبکه‌های اجتماعی محسوب می‌شود. در این حالت، اطلاعاتی از فرد تازه‌وارد در دست نیست تا بتوان از این طریق، علایق وی را شناسایی کرد و پیشنهادها مناسب را به او ارائه داد. در شبکه‌های اجتماعی مشهوری چون Facebook که از روش «دوست - دوست» برای ارائه پیشنهادها استفاده می‌کنند، این مشکل از راه بررسی اسامی موجود در پست الکترونیکی فرد و تطابق آن با اعضای شبکه اجتماعی و ارائه اسامی تطابق یافته، از بین می‌رود. در گوگل پلاس نیز برای حل این مشکل، فرد را وادار می‌کنند که در آغاز ورود به شبکه، ده نفر دوست را انتخاب کرده و معرفی کند.

برای تفکیک داده‌های آموزشی و آزمایشی در هر دو مورد از اعتبارسنجی ضربدری ده‌لایه‌ای^۱ استفاده شد و الگوریتم‌ها به کمک زبان نرم‌افزاری جاوا، به اجرا گذاشته شدند. مقایسه بین الگوریتم‌های مختلف در رابطه با مجموعه داده‌های انتخابی و نتایج آن، در جدول ۱ نشان داده شده است.

جدول ۱. مقادیر MAP الگوریتم‌ها برای مجموعه داده‌های واقعی

الگوریتم	Epinions	Facebook
روش جدید	۰/۵۲۷	۰/۳۱۴
امتیاز الحاق تدریجی	۰/۲۳۳	۰/۱۳۶
ضریب جاکارد	۰/۴۸۴	۰/۱۶۹
HDI	۰/۴۸۴	۰/۱۵۶
کاتز	۰/۴۶۳	۰/۲۰۲
دوستان مشترک	۰/۴۳۵	۰/۲۵۱
آدامیک - آدار	۰/۴۷۵	۰/۲۴۸
HPI	۰/۴۹۸	۰/۱۲۲
FriendLink	۰/۴۸۱	۰/۲۶۰
معیار دوستان	۰/۱۹۹	۰/۲۹۳
ضریب خوشه‌بندی	۰/۵۸۷	۰/۱۹۰
ضریب درجه	۰/۰۹۸	۰/۱۳۸
مجموع وزن‌ها	۰/۱۱۹	۰/۰۳۴
ضرب وزن‌ها	۰/۱۱۹	۰/۰۳۴

همان‌طور که در جدول ۱ مشاهده می‌شود، روش جدید در رابطه با مجموعه داده Epinions پس از معیار ضریب خوشه‌بندی، بهترین جواب را داده است و در مقایسه با سایر معیارها بهتر

عمل کرده است. در مورد Facebook نیز به کمک روش پیشنهادی، در مجموع بهترین پاسخ در رابطه با مجموعه داده انتخابی به دست آمده است؛ در حالیکه در این مجموعه داده، ضریب خوشه‌بندی دارای کارایی بسیار پایینی بوده است. بنابراین همان‌گونه که در جدول ۱ مشاهده می‌شود، روش جدید در مورد هر دو مجموعه داده در مقایسه با سایر روش‌ها، پاسخ‌های مناسب و قابل قبولی خواهد داشت و پیش‌بینی‌های خوبی را ارائه می‌دهد.

نتیجه‌گیری

یکی از عوامل مهم در تعیین میزان شباهت بین دو رأس، میزان درجه رأس‌های رابط بین این دو رأس است. هر چه میزان درجه رأس پیونددهنده بین دو رأس کمتر باشد، بدین معناست که فرد رابط گزیده‌تر انتخاب می‌کند و دوستانش شباهت بیشتری به یکدیگر خواهند داشت و در نتیجه می‌تواند رابط خوبی برای ایجاد پیوند بین دو فرد باشد، اما در صورتی که رأس رابط با وجود درجه بالای ارتباطات، دارای دوستان مشترک زیادی با دو رأس باشد، این خاصیت می‌تواند به‌منزله یک امتیاز مثبت تلقی شده و نقش مهمی در قوی کردن رابطه بین دو رأس ایفا کند.

References

- Adamic L. & Adar E. (2005). How to search a social network. *Social Networks*, 27(3):187-203.
- Al Hasan, M., & Zaki, M. J. (2011). A survey of link prediction in social networks. In *Social network data analytics* (pp. 243-275). Springer US.
- Backstrom, L., & Leskovec, J. (2011, February). Supervised random walks: predicting and recommending links in social networks. In *Proceedings of the fourth ACM international conference on Web search and data mining* (pp. 635-644). ACM.
- Cukierski, W., Hamner, B., & Yang, B. (2011, July). Graph-based features for supervised link prediction. In *Neural Networks (IJCNN), The 2011 International Joint Conference on* (pp. 1237-1244). IEEE.
- Feyessa, T., Bikdash, M., & Leby, G. (2011, October). Node-pair feature extraction for link prediction. In *Privacy, security, risk and trust (passat), 2011 IEEE third international conference on and 2011 IEEE third international conference on social computing (socialcom)* (pp. 1421-1424). IEEE.
- Fire, M., Tenenboim, L., Lesser, O., Puzis, R., Rokach, L., & Elovici, Y. (2011). Link prediction in social networks using computationally efficient topological features. In *Privacy, security, risk and trust (passat), 2011 IEEE*

third international conference on and 2011 IEEE third international conference on social computing (socialcom) (pp. 73-80). IEEE.

- Hasangholipour, T., Amiri, M., Fahim, F. & Ghaderi Abed, A. (2014). Effects of Consumer Characteristics on their Acceptance of Online Shopping: A Survey in Faculty of Management, University of Tehran. *Journal of Information Technology Management*, 5(4): 67-84. (in Persian)
- Irani, M. & Haghghi, M. (2014). The Impact of Social Networks on the Internet Business Sustainability (With Emphasis on the Intermediary Role of Entrepreneurial Purpose of Online Branches of Mellat Bank's Portal). *Journal of Information Technology Management*, 5(4): 23-46. (in Persian)
- Jeh, G., & Widom, J. (2002, July). SimRank: a measure of structural-context similarity. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 538-543). ACM.
- Liben-Nowell, D., & Kleinberg, J. (2007). The link-prediction problem for social networks. *Journal of the American society for information science and technology*, 58(7): 1019-1031.
- Papadimitriou, A., Symeonidis, P., & Manolopoulos, Y. (2011). Predicting links in social networks of trust via bounded local path traversal. In *Proceedings 3rd Conference on Computational Aspects of Social Networks (CASON'2011)*, Salamanca, Spain.
- Papadimitriou, A., Symeonidis, P., & Manolopoulos, Y. (2011). Friendlink: link prediction in social networks via bounded local path traversal. In *Computational Aspects of Social Networks (CASoN), 2011 International Conference on* (pp. 66-71). IEEE.
- Radfar, R., Nezafati, N. & Yousefi Aasl, S. (2014). Classification of Internet banking customers using data mining algorithms. *Journal of Information Technology Management*, 6(1): 71-90. (in Persian)
- Sajedifar, A., Asfidani, M., Vahdatzad, M. & Mahmoudi Azar, M. (2012). The Effect of Electronic Services Quality on Trust-Building in Online Customers of Tehran's Brokerage Firms. *Journal of Information Technology Management*, 4(11): 47-68. (in Persian)
- Zhu, Y. X., Lü, L., Zhang, Q. M., & Zhou, T. (2012). Uncovering missing links with cold ends. *Physical A: Statistical Mechanics and its Applications*, 391(22): 5769-5778.