



ارائه روشی برای سنجش میزان شباهت مکانی کاربران در سامانه های توصیه گر مکان مبنا با رویکرد امنیت داده های اجتماعی

محمد نظری فرخی^۱

بیژن نهاوندی^{۲*}

حسام زند حسامی^۳

۱. چکیده

برای ارائه خدمات مکان مبنا در شبکه های اجتماعی مرسوم، می بایست خلأ میان این شبکه ها و دنیای فیزیکی را با افزودن بُعد مکان پر کرد. یکی از راهکارهایی که با هدف افزودن بُعد مکان به ساختار شبکه اجتماعی ایجاد شدند، شبکه های اجتماعی مکان مبنا هستند. امروزه استفاده از سامانه های توصیه گر موجب شده است تا کشف و جستجوی مکان های مدنظر افراد در کوتاه ترین زمان ممکن، انجام گیرد. اساس کار سامانه های توصیه گر مکانی، تعیین میزان شباهت بین کاربران است. در این پژوهش دو معیار برای تعیین میزان شباهت مکانی کاربران با در نظر گرفتن ابعاد امنیت داده های اجتماعی ارائه گردید. معیار اول با تمرکز بر مکان های مشترک و تاریخ بازدید، بازدیدهای مشترک را به دو گروه هم زمان و غیرهم زمان تقسیم و میانگین ترکیبی آن ها را محاسبه می کند. معیار پیشنهادی دوم با تلفیق سه معیار کلیدی شباهت سفارشی (تمرکز بر اشتراک های مکانی-زمانی)، شباهت کسینوسی (تحلیل جهت گیری حرکات) و ضریب همبستگی پیرسون (سنجش رابطه خطی رفتاری)، امکان ارزیابی چندبعدی رفتار کاربران را فراهم می کند. معیارهای پیشنهادی با روش های متداول سنجش شباهت، شامل ضریب همبستگی پیرسون، شباهت کسینوسی و ضریب جاکارد، مورد مقایسه قرار گرفتند.

۱. گروه علمی مدیریت فناوری اطلاعات، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران،
M.nazarifarokhi@iau.ac.ir

۲. نویسنده مسئول: گروه علمی مدیریت فناوری اطلاعات، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران (نویسنده مسئول)
Bnahavandi@iau.ac.ir

۳. گروه علمی مدیریت فناوری اطلاعات، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران،
H.zand@iau.ac.ir

نتایج مقایسه نشان می‌دهد که هر دو معیار پیشنهادی، عملکرد بهتری نسبت به ضریب جاکارد دارند، اما در مقایسه با شباهت کسینوسی و ضریب پیرسون عملکرد ضعیف‌تری از خود نشان می‌دهند. با این حال، به منظور سنجش دقت در معیارهای پیشنهادی، با استفاده از معیارهای میانگین خطای مطلق و جذر میانگین مربعات خطا، نتایج حاکی از برتری قابل توجه روش‌های پیشنهادی دارند.

واژگان کلیدی: شبکه‌های اجتماعی، خدمات مکان مبنا، سامانه توصیه‌گر، شباهت مکانی کاربران.

۲. مقدمه

افزودن قابلیت‌های موقعیت‌یابی به اپلیکیشن‌های موبایل، جذابیت و فراگیری آن‌ها را افزایش داده و به ظهور شبکه‌های اجتماعی مکان مبنا انجامیده است (طرفدار و همکاران، ۲۰۲۴: ۱۱۲۹). برخلاف شبکه‌های اجتماعی سنتی، این پلتفرم‌ها اطلاعات جغرافیایی واقعی و نظرات کاربران را با یکدیگر تلفیق کرده و بستری برای مطالعه ترجیحات و الگوهای رفتاری شخصی شده فراهم می‌کنند (چانگ و همکاران، ۲۰۲۳: ۸۵۰). داده‌های مکانی در این شبکه‌ها نقش محوری داشته و نشان‌دهنده میزان آشنایی کاربران با مکان‌هاست (بائو و همکاران، ۲۰۱۴: ۱). سوابق کاربران بازتاب تجربیات آن‌هاست و اطلاعات غنی مکانی-زمانی و اجتماعی را برای بهبود سامانه‌های توصیه‌گر فراهم می‌کند (ژاو و همکاران، ۲۰۲۰: ۱۹۶). ساختار اطلاعاتی این شبکه‌ها عموماً در سه لایه جغرافیایی (تاریخچه حضور)، اجتماعی (روابط) و محتوایی (نظرات) سازمان‌دهی می‌شود (دولیزیا و همکاران، ۲۰۲۱: ۱۹).

از نظر نوع اشتراک‌گذاری داده‌های مکانی، شبکه‌های اجتماعی مکان مبنا به سه دسته اصلی تقسیم می‌شوند. شبکه‌های مبتنی بر نشانه‌گذاری مکانی که امکان پیوند مکان به محتوای متنی، تصویری یا ویدیویی را فراهم می‌کنند (خیری و همکاران، ۲۰۱۶: ۱۴۱). شبکه‌های مبتنی بر داده‌های نقطه مکان که در آن‌ها مکان و شباهت‌های مکانی محور اصلی ایجاد ارتباط بوده و سایر اطلاعات نقش مکمل ایفا می‌کنند (سانچز و همکاران، ۲۰۲۳: ۱۸۸۵) و شبکه‌های اجتماعی دنباله‌محور که علاوه بر مکان‌های

بازدید شده، بر جزئیات مسیرهای حرکتی و الگوهای رفتاری کاربران تمرکز دارند و هدف آن‌ها استخراج الگوهای حرکتی از داده‌های دنباله‌ای در پلتفرم‌هایی مانند فوراسکوئر و توییتر است (ژنگ و همکاران، ۲۰۰۹: ۳۵۷؛ ژانگ، ۲۰۲۴: ۲۴۱۰).

پیشنهاد‌های فردی برای خدمات و مکان‌ها توسط سامانه‌های توصیه‌گر ارائه می‌شوند که به عنوان بخشی ضروری از زندگی روزمره ما تبدیل شده‌اند. این پیشنهاد‌های برگرفته از ترجیحات و اقدامات ما هستند (مینا و همکاران، ۲۰۲۴: ۶۴).

بر مبنای تحلیل‌های انجام شده روی داده‌های بازدید کاربران در شبکه‌های اجتماعی مکان مینا، محققان سامانه‌های توصیه‌گر مکانی فراوانی ارائه داده که می‌توان آن‌ها را در دو دسته کلی مدل ترکیبی^۱ و مدل اتصالی^۲ بررسی کرد (چنگ و همکاران، ۲۰۱۲: ۱۷؛ ژائو و همکاران، ۲۰۱۳: ۵۳۰؛ ژائو و همکاران، ۲۰۱۷: ۳۶۸). همچنین الگوریتم‌های مورد استفاده در سامانه‌های توصیه‌گر مکانی را برای یافتن مکان‌های مناسب به صورت جذابیت مکان‌ها، مشابهت بافت مکانی، تأثیرپذیری از ارتباطات اجتماعی، نزدیکی به محل سکونت و مشابهت و هم‌سلیقگی قابل تفکیک هستند (نولاس و همکاران، ۲۰۱۲: ۱۴۴؛ چو و همکاران، ۲۰۱۱: ۱۰۸۲؛ کوئرسیا و همکاران، ۲۰۱۰: ۹۷۱). از جمله این الگوریتم‌ها می‌توان به تجزیه ماتریس مبتنی بر زمینه^۳ اشاره کرد (لی و همکاران، ۲۰۲۵: ۲۱۵).

مهم‌ترین و پرکاربردترین داده‌های زمینه‌ای که در سامانه‌های توصیه‌گر استفاده می‌شوند عبارتند از اطلاعات مکانی^۴ (ژائو و همکاران، ۲۰۱۸)، اطلاعات اجتماعی^۵ (ژائو و همکاران، ۲۰۱۸: ۱؛ لی و همکاران، ۲۰۱۶: ۹۷۵؛ ژانگ و همکاران، ۲۰۱۵: ۴۴۳)، اطلاعات محتوایی^۶ (گائو و همکاران، ۲۰۱۵: ۱۷۲۱)، اطلاعات دسته‌بندی^۷ (ژانگ و چو، ۲۰۱۶: ۲۶) و اطلاعات زمانی^۸ (چنگ و همکاران، ۲۰۱۳: ۲۶۰۵). مطالعات نشان می‌دهد افراد معمولاً با

1. Fused model
2. Joint model
3. Context-Aware Matrix Factorization
4. Spatial
5. Social
6. Content
7. Categorical
8. Temporal

کسانی که از نظر جغرافیایی در مجاورت آن‌ها قرار دارند، ارتباط آسان‌تری برقرار می‌کنند (گولدنبرگ و لوی، ۲۰۰۹: ۹۶۰؛ اسکالتو و همکاران، ۲۰۱۱: ۱۰۴۶).

یکی از چالش‌های بنیادین در سامانه‌های توصیه‌گر مکان مبنا، تعادل میان دقت توصیه و حفظ حریم خصوصی کاربران است. روش‌های سنتی سنجش شباهت مانند معیارهای کسینوسی و پیرسون برای محاسبه شباهت بین دو کاربر، نیازمند دسترسی به تاریخچه کامل بازدیدهای آن‌ها هستند که با اصول حفظ حریم خصوصی مغایرت دارد. پژوهش حاضر با ارائه دو معیار جدید که بر مکان‌های مشترک تمرکز دارند و با داده‌های تعمیم‌یافته قابل محاسبه هستند، گامی به سوی سنجش شباهت آگاه از امنیت داده‌های اجتماعی برمی‌دارد.

با وجود ماهیت فنی پیاده‌سازی معیارهای شباهت، این پژوهش از نظر ماهیت، یک پژوهش نظری-کاربردی محسوب می‌شود؛ زیرا تمرکز اصلی آن بر ارائه چارچوب‌های مفهومی جدید برای سنجش شباهت مکانی شامل مدل ترکیبی مکان-زمان و تابع ترکیب وزنی خطی محدب است و پیاده‌سازی‌های فنی صرفاً برای اعتبارسنجی تجربی این چارچوب‌های نظری انجام شده‌اند؛ بنابراین، دستاورد اصلی پژوهش، دو سازه نظری قابل تعمیم به حوزه سامانه‌های توصیه‌گر است، نه صرفاً یک کدنویسی خاص برای یک مجموعه داده معین؛ به عبارت دیگر، این پژوهش به دنبال پاسخ به پرسش‌های نظری بنیادین در حوزه سنجش شباهت مکانی است. اینکه چگونه می‌توان ابعاد چندگانه رفتار مکانی-زمانی کاربران را در یک معیار واحد و قابل تنظیم ترکیب کرد؟ و کدام ترکیب از معیارهای موجود، کمترین خطای پیش‌بینی و بیش‌ترین پایداری را در شرایط مختلف داده‌ای به همراه دارد؟

۳. بیان مسئله و ضرورت پژوهش

به‌طورکلی، شناسایی کاربران مشابه از طریق مقایسه داده‌ها و استفاده از معیارهای شباهت‌یابی انجام می‌شود (دی و همکاران، ۲۰۰۱: ۹۷). با این حال، علیرغم نقش مهم معیارهای مکانی (بوگومی و همکاران، ۲۰۱۴: ۶۶) و زمانی (چاکری و همکاران، ۲۰۱۷: ۳۹۹) در تشخیص شباهت رفتاری، همچنان شکاف اساسی بین رویکردهای موجود و نیازهای واقعی کاربران در سامانه‌های توصیه‌گر مکان مبنا وجود دارد. پژوهش‌های پیشین عمدتاً

بر یک یا دو بعد مکانی، زمانی یا اجتماعی تمرکز کرده‌اند؛ درحالی‌که سه چالش اصلی پابرجاست. نخست، تقلیل ابعاد چندگانه رفتار کاربر به یک معیار ساده که جنبه‌هایی چون هم‌زمانی بازدیدها و جهت‌گیری حرکتی را نادیده می‌گیرد (یان و همکاران، ۲۰۱۳: ۱)؛ دوم، غفلت از امنیت داده‌های اجتماعی به‌گونه‌ای که معیارهای سنتی نیازمند دسترسی سطح بالا به داده‌های خام بوده و ملاحظات حریم خصوصی را لحاظ نمی‌کنند؛ سوم، نبود یک معیار ترکیبی با قابلیت تنظیم و تفسیرپذیری که در سناریوهای مختلف بتواند اهمیت متفاوت ابعاد مکانی، زمانی و حرکتی را منعکس سازد. هر فعالیت در یک مکان، نشان‌دهنده هدف یا کاربری خاصی است و غنی‌سازی داده‌های حرکت با ابعاد معنایی برای ایجاد مدل‌های دقیق‌تر ضروری است (ژائو و همکاران، ۲۰۱۰: ۴۴۲).

افزون بر سه چالش بالا، چالش چهارم نادیده‌گرفتن امنیت داده‌های اجتماعی در سنجش شباهت مکانی است. روش‌های سنتی برای محاسبه شباهت، نیازمند دسترسی به تاریخچه کامل بازدیدهای کاربران هستند که ریسک نقض حریم خصوصی را به همراه دارد. یک معیار شباهت مطلوب باید (۱) با حداقل داده‌های خام قابل محاسبه باشد، (۲) قابلیت پیاده‌سازی در معماری‌های غیرمتمرکز را داشته باشد و (۳) وابستگی به اطلاعات شناسایی‌کننده کاربران را به حداقل برساند. معیارهای موجود فاقد این ویژگی‌ها هستند. با توجه به چهار چالش مذکور، ارائه معیارهای جدید سنجش شباهت مکانی که چندبعدی، آگاه از امنیت داده و دارای انعطاف وزنی باشند، ضرورتی انکارناپذیر است. اهمیت این پژوهش از سه منظر قابل تبیین است: از منظر علمی، این پژوهش فراتر از معیارهای خطی و هم‌پوشانی محور رفته و با دو معیار نوین یعنی تفکیک بازدید هم‌زمان / ناهم‌زمان و ترکیب وزنی سه بعد مکمل گامی به سوی مدل‌سازی واقعی‌تر رفتار انسان در فضای مکان مبنا برمی‌دارد. از منظر عملی، افزایش حتی چند درصدی دقت شباهت در پلتفرم‌هایی مثل فوراسکوئر و گوگل میپس، نرخ رضایت کاربر را به‌طور معناداری افزایش می‌دهد و رویکرد آگاه از امنیت داده، اعتماد کاربران را تسهیل می‌کند. از منظر شکاف پژوهشی، هیچ‌یک از مطالعات پیشین هم‌زمان به سه ویژگی چندبعدی بودن، انعطاف وزنی و قابلیت حفظ امنیت داده نپرداخته‌اند؛ بنابراین پژوهش حاضر دقیقاً در نقطه شکاف دانشی قرار دارد.

این پژوهش با هدف پرکردن شکاف مذکور به دو پرسش اصلی پاسخ می‌دهد. اول، آیا ترکیب بازدیدهای هم‌زمان و ناهم‌زمان با ضریب تعدیل زمانی، شباهت مکانی را دقیق‌تر از معیارهای تک‌بعدی مانند جاکارد یا کسینوسی برآورد می‌کند؟ دوم، آیا ترکیب وزنی خطی سه معیار مکمل یعنی شباهت سفارشی مکان-زمان، کسینوسی و پیرسون با قابلیت تنظیم وزن‌ها، دقت و پایداری بیشتری نسبت به هر یک از آن معیارها به‌تنهایی دارد؟ پاسخ به این پرسش‌ها از طریق ارزیابی کمی با معیارهای MAE و RMSE و مقایسه با سه روش پایه انجام می‌شود. موفقیت در این مسیر، گامی ضروری به سوی ارائه چارچوبی کارآمد و دقیق برای بهبود کیفیت توصیه‌های مکانی در شبکه‌های اجتماعی مکان مبنا است.

۴. پیشینه پژوهش

پژوهش‌های پیشین در حوزه سامانه‌های توصیه‌گر مکان مبنا را می‌توان در چهار دسته مفهومی بررسی کرد. دسته نخست، روش‌های مبتنی بر خوشه‌بندی و قوانین انجمنی هستند. قباخلو و همکاران (۱۳۹۲: ۱) با داده‌کاوی تاریخچه مکانی، کاربران مشابه را خوشه‌بندی و با قوانین انجمنی پیش‌بینی انجام دادند. پارک و همکاران (۲۰۰۷: ۱۱۳۰) از پیشینه مکانی برای استنباط ترجیحات کاربران و توصیه مکان استفاده کردند. لی و چن (۲۰۰۹: ۱) یک مدل سه‌لایه‌ای برای تلفیق ابعاد مکانی، اجتماعی و محتوایی ارائه دادند. این پژوهش‌ها برخلاف رویکرد ما که معیارهای شباهت پیوسته را ارائه می‌دهد، عمدتاً مبتنی بر خوشه‌بندی گسسته و قوانین انجمنی هستند. دسته دوم، روش‌های مبتنی بر فاکتورسازی ماتریس و یادگیری ماشین هستند. ژئو و همکاران (۲۰۱۲: ۲۵۵) با مقایسه روش‌های مختلف، برتری روش PLSA را نشان دادند. لئو و همکاران (۲۰۱۴: ۱۱۶۷) روش فاکتورسازی ماتریس همسایگی منطقه-موقعیت را برای بهره‌گیری از ویژگی‌های جغرافیایی معرفی کردند. لی و همکاران (۲۰۱۶: ۹۷۵) با تعریف سه نوع دوست و یک مدل فاکتورگیری دومرحله‌ای، به چالش شروع سرد و پراکندگی داده پرداختند. گو و همکاران (۲۰۱۸: ۸۵) یک چارچوب رتبه‌بندی بیزی با استفاده از ویژگی‌های جغرافیایی همسایگی ارائه کردند. ژانگ و همکاران (۲۰۱۹: ۱۱۳۵) تأثیرات چند برچسبی، اجتماعی و جغرافیایی را در یک مدل فاکتورسازی ماتریس ترکیب کردند. ژاو و همکاران (۲۰۲۳: ۱) روش PGR-ELM را برای توصیه گروهی نقاط موردعلاقه ارائه دادند. کازیو و همکاران (۲۰۲۶: ۱) با

استفاده از تجزیه ماتریس مفرد، یک سامانه توصیه‌گر آگاه از زمینه با دقت بالا ارائه دادند که بر پیش‌بینی رتبه‌بندی کاربر-مکان در بافت‌های مختلف متمرکز است. اگرچه این پژوهش‌ها از تکنیک‌های پیشرفته استفاده کرده‌اند، اما معیار شباهت آن‌ها عمدتاً مبتنی بر روش‌های خطی سنتی است.

دسته سوم، روش‌های مبتنی بر عوامل زمینه‌ای (مکانی، زمانی و اجتماعی) هستند. گودرزی و همکاران (۱۳۹۶: ۲۱۵) الگوریتم PCLRTGS را با ادغام تأثیرات مکانی، زمانی و اجتماعی توسعه دادند. ژنگ و همکاران (۲۰۱۱: ۲۴۳) با تحلیل داده‌های GPS، توصیه توأم مکان و فعالیت ارائه کردند. اسکالتو و همکاران (۲۰۱۱: ۱۰۴۶) نشان دادند بازدید از مکان‌های مشترک، عامل ۳۰٪ از ارتباطات جدید است. بائو و همکاران (۲۰۱۲: ۱۹۹) سامانه دویخشی مبتنی بر اولویت شخصی کاربر و نظرات متخصصان محلی طراحی کردند. ژانگ و چو (۲۰۱۳: ۳۳۴) تأثیر جغرافیایی شخصی شده را با اولویت کاربر و نفوذ اجتماعی یکپارچه کردند. وانگ و همکاران (۲۰۱۳: ۳۷۴) با چهار عامل رفتاری، اجتماعی، مکانی و الگوریتم‌هایی برای توصیه مکان جدید ارائه دادند. ووکوویچ و جتیچ (۲۰۱۵: ۵۱۷) سامانه‌ای برای تحلیل حرکت و پیش‌بینی مکان بعدی افراد ایجاد کردند. ژو و همکاران (۲۰۱۷: ۱) مدل مکان‌یابی توالی مکانی را برای استخراج ویژگی‌های مکانی و ترتیبی ارائه دادند. سی و همکاران (۲۰۱۹: ۲۶۷) با استخراج ویژگی‌های فعالیت سه بعدی و زمانی-مکانی، یک الگوریتم تطبیقی POI پیشنهاد دادند. یانگ و همکاران (۲۰۲۰: ۱) با تلفیق یادگیری عمیق و عوامل جغرافیایی، مدلی برای پیش‌بینی رفتارهای کوتاه و بلندمدت طراحی کردند. چن و همکاران (۲۰۲۱: ۵۴۷) چارچوبی برای پیش‌بینی مقصد با ترکیب فاصله جغرافیایی و روابط معنایی ارائه دادند. وانگ و همکاران (۲۰۲۲: ۳۸۷۲) یک مدل مبتنی بر یادگیری عمیق هندسی برای مدل‌سازی روابط پیچیده کاربر-مکان معرفی کردند. ژانگ و همکاران (۲۰۲۴: ۱۲۱۶۳) مدل MGEL را با بهره‌گیری از تعبیه گراف و رویکردی یکپارچه، تأثیر هم‌زمان عوامل جغرافیایی، زمانی و اجتماعی را در یک فضای معنایی مشترک برای پیش‌بینی دقیق‌تر مکان‌های مورد علاقه کاربران به کار گرفتند. این پژوهش‌ها اهمیت توجه هم‌زمان به ابعاد مکانی، زمانی و اجتماعی را تأیید می‌کنند، اما هیچ‌یک معیاری با قابلیت ترکیب وزنی قابل تنظیم و توجه به امنیت داده ارائه نداده‌اند.

دسته چهارم، روش‌های جدید با تأکید بر حریم خصوصی و معیارهای شباهت نوین هستند. چن و همکاران (۲۰۲۵: ۲۴۵) با بهره‌گیری از الگوریتم‌های ترکیبی، شباهت جغرافیایی و علائق کاربران را تلفیق کردند. وانگ و همکاران (۲۰۲۵: ۵۱۲۵۶) با ترکیب حریم خصوصی تفاضلی، مدل انتشار و شباهت رنی، الگوریتم توصیه‌گری ارائه دادند که با تزریق نویز کنترل شده به فرایند یادگیری، از افشای اطلاعات حساس کاربران جلوگیری می‌کند؛ استفاده از شباهت رنی به جای معیارهای سنتی، امکان اندازه‌گیری انعطاف پذیرتر شباهت بین توزیع‌های داده را فراهم می‌آورد. بتاش و همکاران (۲۰۲۵: ۳۷۹) با انتقاد از معیارهای سنتی شباهت مانند کسینوسی، جاکارد و پیرسون که فاصله فیزیکی بین کاربران را نادیده می‌گیرند، معیار جدیدی به نام IPUMC ارائه دادند که شباهت کسینوسی را با فاصله جغرافیایی کاربران تلفیق می‌کند؛ ارزیابی این روش بر روی مجموعه داده فوراسکوئر نشان داد که این معیار نسبت به روش‌های پایه، افزایش دقت را به همراه دارد. وو و همکاران (۲۰۲۵: ۱۱۱۲۶۴۸۵) با نقد روش‌های مبتنی بر اعتماد صریح (مانند دوستی‌های اعلام شده) به دلیل سوگیری و عدم فعالیت کاربران، رویکرد TC-new را برای استنتاج اعتماد ضمنی از روی بازدیدهای کاربران ارائه دادند؛ این روش بدون نیاز به روابط اجتماعی صریح و تنها با محاسبه شباهت الگوهای بازدید، مکان‌های مورد علاقه کاربران را پیش‌بینی می‌کند. این دسته از پژوهش‌ها به پژوهش حاضر نزدیک‌تر هستند، زیرا هم به معیارهای شباهت جدید و هم به امنیت داده توجه کرده‌اند. با این حال، هیچ‌یک از آن‌ها معیاری با ترکیب هم‌زمان سه بعد مکانی، زمانی و حرکتی همراه با قابلیت تنظیم وزن و تمرکز بر داده‌های تعمیم یافته ارائه نداده‌اند که هدف اصلی پژوهش حاضر است.

۵. روش‌شناسی پژوهش

در پژوهش‌های پیشین، سه معیار اصلی برای سنجش شباهت مکانی کاربران مورد استفاده گسترده قرار گرفته‌اند. ضریب جاکارد که صرفاً بر اساس هم‌پوشانی مکان‌های بازدید شده عمل می‌کند و ابعاد زمانی و حرکتی را نادیده می‌گیرد؛ شباهت کسینوسی که جهت‌گیری حرکت کاربران را تحلیل می‌کند اما از بزرگی داده‌ها و هم‌زمانی بازدیدها غافل است؛ و ضریب همبستگی پیرسون که روابط خطی بین بازدیدها را می‌سنجد، اما به نقاط

پرت حساس بوده و روابط غیرخطی را پوشش نمی‌دهد. نقطه ضعف مشترک هر سه معیار، عدم توجه هم‌زمان به ابعاد چندگانه رفتار کاربر (مکان، زمان، توالی و جهت‌گیری حرکت) و نیز نادیده‌گرفتن ملاحظات امنیت داده‌های اجتماعی و حریم خصوصی است. در مقابل، نوآوری پژوهش حاضر در ارائه دو معیار جدید سنجش شباهت مکانی خلاصه می‌شود که به‌طور مشخص سه شکاف اصلی دانشی را پر می‌کنند. نخست، معیار اول با تفکیک بازدیدهای هم‌زمان و ناهم‌زمان و ضریب تعدیل زمانی، هم‌زمان اشتراک مکانی و هماهنگی زمانی را در یک فرمول واحد ترکیب می‌کند. معیار دوم با تلفیق وزنی سه بعد مکمل شباهت سفارشی مکان-زمان، شباهت کسینوسی و ضریب پیرسون و با قابلیت تنظیم پویای ضرایب α و β ، امکان انعطاف‌پذیری متناسب با سناریوهای مختلف را فراهم می‌آورد. سوم، هر دو معیار پیشنهادی با تمرکز بر مکان‌های مشترک به جای نیاز به داده‌های خام تمام مکان‌های بازدیدشده، گامی عملی به سوی سنجش شباهت آگاه از حریم خصوصی محسوب می‌شوند و قابلیت پیاده‌سازی در معماری‌های حفظ امنیت داده را دارا هستند؛ به عبارت دیگر، برخلاف روش‌های موجود که حداکثر دو بعد از رفتار کاربر را مدل می‌کنند، معیارهای پیشنهادی این پژوهش به‌طور هم‌زمان سه بعد مکانی، زمانی و حرکتی را با رویکردی کارآمد پوشش می‌دهند.

چارچوب پیشنهادی پژوهش جهت ارائه روشی برای تعیین میزان شباهت مکانی کاربران به صورت زیر است:

- استخراج داده‌های مربوط به کاربران از مجموعه داده مانند شناسه کاربر، برچسب زمان و تاریخ بازدید، مختصات مکان، شناسه مکان و انجام عملیات پیش‌پردازش.
- کشف کاربران مشابه بر اساس مکان و زمان بازدید و ارائه معیار شباهت مکانی بین کاربران.
- مقایسه معیار شباهت پیشنهادی با روش‌های دیگر مانند ضریب همبستگی پیرسون، معیار شباهت کسینوسی و ضریب جاکارد.
- سنجش دقت در معیار پیشنهادی با استفاده از دو معیار خطای مطلق میانگین و جذر میانگین مربعات خطا.

۱.۵. معیار پیشنهادی اول برای تعیین شباهت مکانی کاربران

سامانه‌های توصیه‌گر بر اساس علایق شخصی کاربران و با شناخت ویژگی‌های افراد بر اساس پیشینه مکانی آن‌ها تلاش می‌کنند؛ مکانی را به کاربران پیشنهاد دهند که مطابق با سلیقه آن‌ها باشد. این‌گونه سامانه‌ها با در نظر گرفتن تعداد دفعات بازدید و علاقه کاربران به مکان‌های مختلف به برآورد میزان تمایل کاربر به بازدید از مکان‌هایی که با آن‌ها ناآشنا هستند، پرداخته و در نهایت مکان‌هایی که بر اساس برآورد انجام شده شانس بیشتری برای مطابقت با سلیقه کاربر دارند را به او پیشنهاد می‌دهند (ژنگ و همکاران، ۲۰۱۱: ۲۴۳). اهمیت تفکیک بازدیدهای هم‌زمان از ناهم‌زمان در سنجش شباهت مکانی، پیش‌تر توسط (طرفدار و همکاران، ۲۰۲۴: ۱۱۲۹) در چارچوب پیش‌بینی هم‌زمانی مورد تأکید قرار گرفته است. با این حال، پژوهش حاضر برای اولین بار این مفهوم را با ضریب تعدیل زمانی در یک معیار شباهت ترکیبی کمی‌سازی می‌کند.

تحقیقات نشان می‌دهد که همسانی در الگوهای رفتاری مکانی، یک نماینده قوی برای تشابه در ترجیحات و علایق کاربران است و پایه‌ای برای توسعه سرویس‌های نوینی مانند شبکه‌های اجتماعی مکان مبنا فراهم می‌کند (یو و همکاران، ۲۰۱۱: ۳۶۱). با این حال، سنجش دقیق این شباهت فراتر از محاسبه صرف مکان‌های مشترک است و مستلزم در نظرگیری دو بُعد کلیدی توالی مکانی و هم‌زمانی است. توالی مکانی به تشابه در ترتیب بازدید از مکان‌ها اشاره دارد. از سوی دیگر، هم‌زمانی نقش تعیین‌کننده‌تری ایفا می‌کند و خود شامل چند لایه است: فاصله زمانی کوتاه‌تر بین رویدادهای مشابه (مثلاً یک ساعت در مقابل یک هفته)، هم بازدید در ساعات مشابهی از روز (الگوی روزانه) و هم بازدید در روزهای مشابهی از هفته (الگوی هفتگی مانند آخر هفته). از آنجاکه داده‌های مکان محور کاربران ذاتاً پراکنده و با فواصل زمانی متغیر هستند، هرچه هماهنگی در این ابعاد زمانی بیشتر باشد، قرابت رفتاری و در نتیجه احتمال تشابه علایق بین کاربران به طور معناداری افزایش می‌یابد؛ بنابراین، یک معیار سنجش شباهت کارآمد باید فراتر از شمارش مکان‌های مشترک رفته و این ابعاد پویای مکانی-زمانی را به طور یکپارچه مدل‌سازی نماید.

تعیین شباهت و کشف همسایگان کاربران در بین مجموعه‌های داده عظیم، به دلیل ماهیت محاسبات جفت‌شوندگی و افزایش تصاعدی پیچیدگی، با چالش‌های محاسباتی سنگینی مواجه است. این چالش در رویکردهای مبتنی بر الگوریتم‌های تکاملی که نیازمند ارزیابی تکراری تابع برازش برای هر کاربر هستند، تشدید می‌شود؛ بنابراین، طراحی یک تابع سنجش شباهت کارا با هزینه محاسباتی خطی یا نزدیک به خطی، یک ضرورت اساسی برای افزایش مقیاس‌پذیری سامانه‌های توصیه‌گر مبتنی بر همسایگی است. در راستای این هدف، معیار پیشنهادی حاضر، یک مدل ترکیبی مکان-زمان ارائه می‌دهد که از هر دو بعد اشتراک مکانی و هم‌زمانی بهره می‌برد. هسته این مدل، تمرکز بر مجموعه اشتراک مکانی بین دو کاربر است. این مجموعه به دو زیرمجموعه تفکیک می‌شود: بازدیدهای هم‌زمان (تطابق کامل برچسب زمانی) و بازدیدهای ناهم‌زمان. به منظور اجتناب از اثرات نویز ناشی از مکان‌های غیرمشترک، این مدل تنها بر اشتراک‌ها متمرکز شده و از در نظر گرفتن مکان‌های انحصاری هر کاربر صرف نظر می‌کند.

نوآوری کلیدی در پردازش بازدیدهای ناهم‌زمان است که به جای یکسان‌پنداری، با یک ضریب تعدیل زمانی وزن‌دهی می‌شوند. این ضریب به صورت نسبت کمترین فاصله زمانی مشاهده شده بین دو کاربر در یک مکان مشترک، به میانگین فاصله زمانی آن‌ها در کل مکان‌های مشترک محاسبه می‌شود. این مکانیسم، درجه هم‌زمانی را کمی‌سازی کرده و به بازدیدهایی که در الگوهای زمانی مشابه‌تری رخ داده‌اند، امتیاز بالاتری اختصاص می‌دهد. این تابع به صورت رابطه زیر نشان داده می‌شود.

در این رابطه برابر تعداد مکان‌های مشترک بازدید شده توسط دو کاربر به طور هم‌زمان، تعداد کل مکان‌های مشترک کاربران، نسبت کمترین فاصله زمانی بین دو کاربر نسبت به میانگین فاصله زمانی آن‌ها در مکان‌های مشترک و تعداد مکان‌های مشترک بازدید شده توسط دو کاربر به طور ناهم‌زمان، است.

۲.۵. معیار پیشنهادی دوم برای تعیین شباهت مکانی کاربران

امکان ترکیب وزنی ابعاد مختلف در سنجش شباهت کاربران، پیش‌تر توسط (آچاریا و همکاران، ۲۰۲۴: ۱۱۵۸۴) در قالب یک مدل ترکیب زمانی-مکانی بررسی شده است. با این

حال، نوآوری پژوهش حاضر در استفاده از یک ترکیب خطی محدب با ضرایب قابل تنظیم و است که هم‌زمان سه بعد مکانی-زمانی، حرکتی و خطی را پوشش می‌دهد.

تابع شباهت ترکیبی یک روش ترکیب وزنی برای محاسبه شباهت مکانی کاربران ارائه می‌دهد که از سه معیار اصلی شباهت سفارشی شباهت کسینوسی و ضریب همبستگی پیرسون (استفاده می‌کند. این تابع با تخصیص ضرایب وزنی قابل تنظیم و امکان ادغام هوشمندانه معیارهای مختلف را فراهم می‌سازد و می‌تواند هم‌زمان اشتراک‌های مکانی-زمانی، جهت‌گیری حرکتی و الگوهای خطی رفتاری کاربران را در نظر بگیرد و از این طریق، دقت ارزیابی شباهت را افزایش دهد. فرمول ترکیب وزنی ارائه‌شده با هدف ارزیابی جامع شباهت مکانی کاربران طراحی شده است. این فرمول با تلفیق سه معیار کلیدی شباهت سفارشی (تمرکز بر اشتراک‌های مکانی-زمانی)، شباهت کسینوسی (تحلیل جهت‌گیری حرکات) و ضریب همبستگی پیرسون (سنجش رابطه خطی رفتاری)، امکان ارزیابی چندبعدی رفتار کاربران را فراهم می‌کند.

به صورت پیش فرض برابر ۰.۴ در نظر گرفته می‌شود که وزن اختصاص یافته به معیار سفارشی بوده و هر چه مقدار آن بزرگ‌تر باشد، تأثیر معیار سفارشی در امتیاز نهایی بیشتر می‌شود. نیز به صورت پیش فرض برابر ۰.۴ در نظر گرفته می‌شود که وزن اختصاص یافته به معیار کسینوسی بوده و اگر مقدار آن افزایش یابد، شباهت کسینوسی اهمیت بیشتری پیدا می‌کند؛ ه عبارت دیگر درصد امتیاز ناشی از معیار سفارشی است؛ درصد از امتیاز ناشی از معیار کسینوسی بوده و درصد باقیمانده به معیار پیرسون اختصاص می‌یابد. اگر باشد، معیار پیرسون نادیده گرفته می‌شود. همچنین تعریف شده است که مجموع ضرایب و باید کوچک‌تر یا مساوی ۱ باشد تا اطمینان حاصل شود که مجموع وزن‌ها از محدوده مجاز خارج نمی‌شود. این رویکرد دارای چند مزیت اصلی است:

- انعطاف‌پذیری: امکان تنظیم سهم هر معیار بر اساس اهمیت آن در مسئله (مثلاً افزایش برای تأکید بر اشتراک‌های مکانی)، ه عبارت دیگر پارامترهای و امکان تنظیم سهم هر معیار را بر اساس نیاز مسئله فراهم می‌کنند. برای مثال اگر اشتراک مکان‌ها مهم‌تر باشد، افزایش می‌یابد و اگر جهت‌گیری حرکت کاربران کلیدی باشد، بزرگ‌تر انتخاب می‌شود.

- بهبود دقت: ترکیب چندین معیار مکمل، خطای ناشی از اتکا به یک معیار واحد را کاهش می‌دهد.
- مقیاس‌پذیری: قابلیت استفاده در مجموعه داده‌های بزرگ با محاسبات موازی.
- کاهش خطا: کاهش وابستگی به محدودیت‌های یک معیار خاص با ترکیب موزون معیارها.
- پایداری: ایجاد تعادل بین معیارها با شرط این رویکرد نه تنها شرط نرمالیزاسیون (مجموع وزن‌های برابر با ۱) را تضمین می‌کند، بلکه از حذف ناخواسته یکی از معیارها جلوگیری می‌نماید. از منظر آماری، این فرمول یک ترکیب خطی محدب^۱ است که دو ویژگی اساسی حفظ مقادیر در بازه [۰, ۱] و تفسیرپذیری سهم هر معیار را دارا است.

۳/۵. معیارهای مقایسه با معیار پیشنهادی

محاسبه شباهت یکی از مؤلفه‌های اصلی در اکثر روش‌های توصیه است. این روش‌ها عموماً نیازمند دانستن ارتباط مکانی میان کاربران بوده و از این حیث معیارهای محاسبه شباهت ابزار مناسبی به شمار می‌آیند. پژوهش جو و همکاران (۲۰۲۵: ۱۱۲۸۶۷) نشان داده است که محاسبه شباهت مکانی کاربران در معماری یادگیری فدرال بدون نیاز به اشتراک‌گذاری داده‌های خام امکان‌پذیر است؛ هرچند معیار شباهت مورد استفاده در این پژوهش مبتنی بر روش‌های سنتی کسینوسی و جاکارد بوده است. در ادامه به مرور معیارهای شباهتی که برای مقایسه با معیار شباهت پیشنهادی استفاده شده‌اند، می‌پردازیم.

۱/۳/۵. معیار شباهت کسینوسی^۲

داده‌های دوستی موجود در شبکه‌های اجتماعی مکان مینا، نشانگر کاربرانی است که در شبکه یکدیگر را به عنوان دوست قبول کرده‌اند. لزوماً دوست اجتماعی بودن دو کاربر باهم نشانگر این نیست که این دو کاربر حتماً باید علایق و سلیق مشترکی داشته باشند؛ چراکه اکثر این دوستی‌ها به دلیل شناخت این دو کاربر از هم به دلایل خانوادگی، کاری

1. Convex Combination
 2. Cosine Similarity Measure

و ... است. لذا تعریف دیگری از دوستی برای کاربران نیاز است. برای این منظور، رویه رایج محاسبه شباهت بازدیدهای کاربران است. هر چه شباهت بازدید کاربران زیادتر باشد، نشانگر این است که دو کاربر از علایق و ترجیحات مشترک بیشتری بهره می‌برند. لذا از این داده‌ها می‌توان با اطمینان بیشتری برای توصیه بهره برد. یکی از روش‌هایی که برای مقایسه با روش پیشنهادی در این پژوهش استفاده می‌شود، محاسبه شباهت مکانی بین کاربران بر اساس معیار شباهت کسینوسی است. میزان شباهت کسینوسی همواره بین صفر و یک است و هرچه به یک نزدیک‌تر باشد، نشانگر شباهت بیشتر دو کاربر از لحاظ مکانی به هم است (رحمان و همکاران، ۲۰۱۷: ۳۲).

۲.۳.۵. ضریب همبستگی پیرسون^۱

روش دیگری که برای مقایسه با روش پیشنهادی در این پژوهش استفاده می‌شود، ضریب همبستگی پیرسون است. در این ضریب، اگر دو کاربر مکان‌های نزدیک به هم را بازدید کنند، می‌توان گفت که در آینده نزدیک نیز مکان‌های تقریباً یکسانی را بازدید خواهند کرد. به عنوان مثال، اگر دو کاربر الف و ب مکان‌های ۱ و ۲ را بازدید کرده باشند، احتمالاً به مکان ۳ نیز بازدید مشترکی داشته باشند؛ بنابراین اگر بخواهیم مکان احتمالی کاربر الف روی مکان ۳ را پیش‌بینی کنیم، باید از مکان کاربر ب روی همین مکان کمک بگیریم. ضریب همبستگی پیرسون بین ۱+ و ۱- است، اگر مقدار به دست آمده مثبت باشد، به معنی این است که دو کاربر از لحاظ مکانی شباهت بیشتری دارند و اگر مقدار به دست آمده منفی باشد، یعنی دو کاربر شباهت مکانی کمتری دارند. اگر این مقدار صفر شد، یعنی دو کاربر شباهت مکانی باهم ندارند (دره‌شیری و فرنقی، ۱۳۹۴: ۱).

۳.۳.۵. ضریب جاکارد^۲

روش دیگر برای مقایسه با روش پیشنهادی، ضریب جاکارد است. تقریباً ۳۰ درصد از مکان‌های جدید بازدید شده توسط یک کاربر، قبلاً توسط یکی از دوستانش و یا دوستان دوستانش بازدید می‌شود (وانگ و همکاران، ۲۰۱۳)؛ بنابراین به منظور به دست آوردن این معیار از ضریب جاکارد برای محاسبه شباهت بین کاربر و دوستش از نظر تعداد دوستان

1. Pearson Correlation Coefficient

2. Jaccard Coefficient

مشترکی که دارند، استفاده می‌شود. همچنین با استفاده از ضریب جاکارد، شباهت بین کاربر و دوستش، از نظر دسته مکان‌های مشترکی که بازدید کرده‌اند، محاسبه می‌شود. هر چه مقدار این شباهت بیشتر باشد، توصیه‌های دوست کاربر قابل اعتمادتر می‌شود (گودرزی و همکاران، ۱۳۹۶: ۲۱۵).

دسته مکان‌هایی است که دوست کاربر بازدید کرده است و دسته مکان‌هایی است که کاربر بازدید کرده است.

۶. تجزیه و تحلیل یافته‌ها

۱.۶. مجموعه داده مورد استفاده

در این پژوهش ما از داده‌های وبسایت گولا استفاده می‌کنیم. وبسایت گولا یک شبکه اجتماعی مکان مینا است که از سال ۲۰۰۹ شروع به کار کرده است. کاربران این شبکه از طریق اعلام اطلاعات نقطه-مکان خود با گوشی‌های تلفن همراه می‌توانند مکان خود را با دوستانشان به اشتراک بگذارند. در این شبکه ایجاد رابطه بین کاربران مشروط به فرستادن درخواست از سوی یک کاربر و پذیرش آن از سوی کاربر دیگر است. (شکل ۱) شماتیکی از مجموعه داده گولا را نشان می‌دهد که شامل ستون‌های شناسه کاربر، برچسب زمان و تاریخ بازدید، مختصات مکان به صورت ستون‌های X و Y و شناسه مکان است.

Id_user	Date	Time	X_loc	Y_loc	Id_loc
0	20101019	23:55:27	30.2359091167	97.7951395833	22847
0	20101018	22:17:43	30.2691029532	97.7493953705	420315
0	20101017	23:42:03	30.2557399927	97.7633857727	316637
0	20101017	19:26:05	30.2634181234	97.7575966669	16516
0	20101016	18:50:42	30.2742518584	97.7405226231	5535878
0	20101012	23:58:03	30.261599404	97.7585805953	15372
0	20101012	22:02:11	30.2679095833	97.7493124167	21714
0	20101012	19:44:40	30.2691029532	97.7493953705	420315
0	20101012	15:57:20	30.2811204101	97.7452111244	153505
0	20101012	15:19:03	30.2691029532	97.7493953705	420315
0	20101012	00:21:28	40.6438845363	73.7828063965	23261
0	20101011	20:21:20	40.74137425	73.9881052167	16907
0	20101011	20:20:42	40.741388197	73.9894545078	12973
0	20101011	09:06:30	40.7249103345	73.9946207517	341255
0	20101010	22:00:37	40.729768314	73.9985353275	260957
0	20101010	21:17:14	40.7285271242	73.9968681335	1933724
0	20101010	17:47:04	40.7417466987	73.993421425	105068
0	20101009	23:51:10	40.7341635333	74.0041635333	34817
0	20101009	22:27:07	40.7425115937	74.0060305595	27836
0	20101009	21:39:26	40.7423961659	74.0075433254	15079
0	20101009	21:36:05	40.7423961659	74.0075433254	15079
0	20101009	21:05:23	40.7358847426	74.0049684048	22806
0	20101009	20:55:47	40.7275253534	73.9853990078	1365909
0	20101009	01:37:03	40.7568799674	73.9862251282	11844
0	20101008	21:48:37	40.7074172208	74.0113627931	11742
0	20101008	21:45:48	40.7071727167	74.0105454333	19822
0	20101008	21:43:52	40.7070708167	74.0119528667	15169
0	20101008	21:43:02	40.705823135	73.9966964722	11794
0	20101008	19:28:36	40.7693780407	73.9630830288	1567837
0	20101008	17:24:27	40.7808054632	73.9764726162	35513
1	20100725	11:40:18	48.86337875	2.333328717	1500177
1	20100724	12:21:43	45.202483667	0.746830983	1493267
1	20100715	04:09:15	44.724745941	5.03296136	1441698
1	20100714	04:13:35	44.916842649	5.784218561	1436795
1	20100713	03:33:55	45.664003665	6.4229900217	1431949
1	20100711	11:19:19	45.95037652	6.740224373	1423291
1	20100711	03:45:31	46.130260017	6.516775058	1422219
1	20100710	03:33:04	46.300467489	5.740870777	1414779
1	20100708	05:53:44	48.419669031	3.016968297	1404455
1	20100707	07:53:20	49.6100008243	3.999686	1399686
1	20100706	10:46:41	50.381005442	3.430709991	1395104
1	20100705	04:17:34	50.28882455	5.9133318	1389275
1	20101021	00:03:50	34.0430230998	118.2671570778	14637
1	20101020	17:18:59	34.0172234606	118.447508812	59838
1	20101015	16:00:09	34.0172234606	118.447508812	59838
1	20101012	01:35:37	34.0172234606	118.447508812	59838
1	20101008	02:00:05	34.02476572	118.32811021	1474903
1	20101007	19:14:49	34.0308269167	118.4680975833	335197
1	20101007	00:29:04	34.0426141541	118.2696247101	31248
1	20101007	00:17:00	34.0172234606	118.447508812	59838
1	20101006	23:39:38	34.0303804344	118.4662628174	2781866
1	20101006	20:34:07	34.031481915	118.46877458	1327534

(شکل ۱). شماتیکی از مجموعه داده گولا

۲٫۶. پیاده‌سازی معیار پیشنهادی اول

برای پیاده‌سازی مدل خود از زبان برنامه‌نویسی پایتون که یکی از محبوب‌ترین زبان‌های برنامه‌نویسی دهه کنونی است؛ استفاده می‌نماییم. ما با کمک نرم‌افزار پایتون IDLE 3.10 بر روی سامانه ۶۴ بیتی پیاده‌سازی را انجام دادیم. کتابخانه پایتون مجموعه‌ای از کدها و توابع از پیش نوشته شده است که مجموعه‌ای از ابزارهای مفید و قابل استفاده مجدد را برای وظایف و اهداف مختلف ارائه می‌دهد. این کتابخانه‌ها توسط توسعه‌دهندگان ایجاد شده‌اند تا با ارائه توابع، کلاس‌ها و ماژول‌های از پیش پیاده‌سازی شده‌ای که می‌توانند به راحتی در برنامه‌های پایتون شما ادغام شوند، به ساده‌سازی و تسریع فرایند کدنویسی کمک کنند. ما در پیاده‌سازی خود از کتابخانه‌های زیر استفاده نمودیم.

- numpy برای کار با فهرست‌ها و آرایه‌ها
- matplotlib.pyplot برای رسم انواع نمودارها

برای پیاده‌سازی معیار شباهت ابتدا کاربران، زمان‌ها و مکان‌ها را از بین تقریباً ۱۴ هزار داده‌ای که به صورت تصادفی انتخاب شده بودند، فراخوانی می‌کنیم. مکان‌ها را به X و Y تفکیک کرده و کاربران و مکان‌ها را یونیک کرده و در مرحله بعد آرایه‌ای را برای آزمون تعریف می‌کنیم که شامل ۱۰، ۲۰، ۳۰، ... ۲۰۰ کاربر است.

```
file = open(filename, 'r')
data = [line.split() for line in file]
for a in data:
    user_1.append(int(float(a[0])))
    date_1.append(int(float(a[2])))
    local_1.append(int(float(a[1])))
    xx1.append(float(a[3]))
    yy1.append(float(a[4]))
user_2=np.unique(user_1)
loc_2=np.unique(local_1)
u1=[]
u2=[]
sml=[]
cm1=[]
pm1=[]
jm1=[]
loc_uj_sh=[]
d_rand=[10,20,30,40,50,60,70,80,90,100,150,200]
```

(شکل ۲). شبه کد معیار شباهت پیشنهادی و روش‌های مقایسه

حال از بین کاربران یونیک شده؛ با توجه به آرایه‌ای که برای آزمون ساختیم (مثلاً ۱۰ کاربر) را انتخاب سپس زمان‌ها و مکان‌های این ۱۰ کاربر را جدا و به X و Y تفکیک می‌کنیم. در مرحله بعد ۱۴ هزار داده را یکی یکی فراخوانی کرده و هرکدام از این ۱۴ هزار داده که id_user آن‌ها با این ۱۰ کاربر یکی بود را جدا می‌کنیم.. این حلقه برای ۱۰، ۲۰، ۳۰ و ... ۲۰۰ کاربر تکرار می‌شود.

```
for ii in range(len(data)):
    if(id_ui==user_1[ii]):
        l1.append([local_1[ii],date_1[ii]])
        lc1.append(local_1[ii])
        xx=xx1[ii]
        yy=yy1[ii]
    elif (id_uj==user_1[ii]):
        l2.append([local_1[ii],date_1[ii]])
        lc2.append(local_1[ii])
        xx2.append(xx1[ii])
        yy2.append(yy1[ii])
    xx2.append(xx)
    yy2.append(yy)
```

(شکل ۳). شبه کد تفکیک تاریخ و مکان کاربران

در مرحله بعد کاربرانی که از لحاظ مکان و زمان مشترک هستند را جدا و آن‌هایی که از نظر مکانی مشترک و از نظر زمانی مشترک نیستند را نیز جدا می‌کنیم. سپس طول (تعداد) آن‌هایی که از از لحاظ زمانی و مکانی مشترک بودند با طول (تعداد) آن‌هایی که از لحاظ مکانی مشترک ولی از لحاظ زمانی متفاوت بودند را باهم جمع می‌کنیم.

```
####date_i==date_j and local_i==local_j
loc_date_share=[x for x in l1 if x in l2]
####date_i<>date_j and local_i==local_j
loc_share=[x for x in lc1 if x in lc2]
#####local_i==local_j#####
loc_uj_sh.append(len(loc_date_share)+len(loc_share))
#####
```

(شکل ۴). شبه کد کاربران مشترک و غیرمشترک از لحاظ زمانی و مکانی

حال آن‌هایی که مکان مشترک دارند و زمان مشترک ندارند را در داخل آرایه‌ای قرار می‌دهیم. اگر این آرایه خالی نبود، فاصله آن‌ها را از لحاظ مکانی به دست می‌آوریم (به عنوان مثال X های آن‌ها یکی و Y های آن‌ها باهم متفاوت باشد). در اینجا یا باید محور X ها را در نظر بگیریم یا محور Y ها؛ که ما محور X ها را در نظر می‌گیریم. مقدار اولیه آلفا به صورت پیش فرض برابر صفر است. اگر P برابر صفر باشد، مقدار آلفا را یک می‌دهیم. اگر P مخالف صفر

باشد، کمینه اختلاف دو مکان را به دست آورده و بر میانگین اختلاف مکانی آن‌ها تقسیم می‌کنیم. هر چه این اختلاف کمتر باشد، آلفا به یک نزدیک‌تر می‌شود. $Cm1$ ، $Pm1$ ، $Sm1$ و $Jm1$ به ترتیب معیار پیشنهادی، ضریب همبستگی پیرسون، معیار شباهت کسینوسی و ضریب جاکارد هستند که تا ۵ رقم اعشار روند شده‌اند. درواقع هرکدام از روش‌ها یعنی معیار پیشنهادی، ضریب همبستگی پیرسون، معیار شباهت کسینوسی و ضریب جاکارد دارای ۱۲ نتیجه هستند؛ یعنی برای ۱۰، ۲۰، ۳۰ و ... ۲۰۰ کاربر عملیات انجام می‌شود.

```
c=np.shape(loc_date_share)[0]
loc2=np.shape(loc_share)[0]
##get sub date_i-date_j for locals share
date_temp=[]
if(loc2):
    date_temp=[np.absolute(int(y)-int(y1)) for[x,y]
in l1 for [x1,y1] in l2 if(int(x)==int(x1))and(int(y)!=int(y1)) ]
p=np.shape(date_temp)[0]
alpha=0
if(p==0):
    alpha=1
else:
    ulfa=np.min(date_temp)/np.average(date_temp)
#sim=(c/len(l1))+(ulfa*p/len(l1))
sim=c+(ulfa*p)
u1.append(id_ui)
u2.append(id_uj)
sml.append(round(sim,5))
sl=s1+round(sim,5)
pm1.append(pearson_correclation_coef(xx2,yy2))
s2=s2+pearson_correclation_coef(xx2,yy2)
cm1.append(cosine_sim(xx2,yy2))
s3=s3+cosine_sim(xx2,yy2)
jm1.append((c/len(l1)))
s4=s4+(c/len(l1))
j=j+1
```

(شکل ۵). شبه کد مقایسه معیار پیشنهادی با ضریب همبستگی پیرسون، معیار شباهت کسینوسی و ضریب جاکارد

در ادامه کل اشتراکات را منهای ۱۲ نتیجه‌ای که در مرحله قبل برای هرکدام از روش‌ها به دست آمد، می‌کنیم. در مرحله بعد میانگین شباهت‌هایی که در مرحله قبل به دست آمد را محاسبه می‌نماییم.

۳.۶. پیاده‌سازی معیار پیشنهادی دوم

در مرحله اول تعریف تابع با پارامترهای ورودی استاندارد صورت می‌پذیرد. به این صورت که تابع شباهت ترکیبی با پنج پارامتر طراحی می‌شود که هرکدام نماینده یک معیار

شباهت یا وزن مربوطه هستند. در مرحله بعد اعتبارسنجی ریاضی وزن‌ها انجام می‌شود. شرط با استفاده از بررسی می‌شود تا از نرمالیزه بودن ترکیب خطی اطمینان حاصل شود. در ادامه وزن سوم یعنی وزن معیار پیرسون به صورت محاسبه می‌شود تا مجموع ضرایب دقیقاً برابر با ۱ باقی بماند. ترکیب خطی وزنی در قالب فرمول کلی محاسبه شده و برای حفظ بازه خروجی محدود کردن نتیجه نهایی به بازه $[1, 0]$ اعمال می‌شود تا از تفسیرپذیری آماری اطمینان حاصل شود.

```
import numpy as np
from typing import Union
def hybrid_similarity(
    custom_sim: Union[float, np.ndarray],
    cosine_sim: Union[float, np.ndarray],
    pearson_sim: Union[float, np.ndarray],
    alpha: float = 0.4,
    beta: float = 0.4
) -> Union[float, np.ndarray]:
    Computes weighted hybrid similarity score using the formula:
     $\alpha \cdot \text{custom\_sim} + \beta \cdot \text{cosine\_sim} + (1 - \alpha - \beta) \cdot \text{pearson\_sim}$ 
    Parameters:
    custom_sim : float or ndarray
    Custom similarity score (spatio-temporal) in range [0,1]
    cosine_sim : float or ndarray
    Cosine similarity score in range [0,1]
    pearson_sim : float or ndarray
    Pearson correlation coefficient in range [-1,1]
    alpha : float, optional
    Weight for custom similarity (default: 0.4)
    beta : float, optional
    Weight for cosine similarity (default: 0.4)
    float or ndarray
    Normalized hybrid score in range [0,1]
    ValueError
    If weights are invalid or inputs are out of bounds
    # Validate weight constraints
    if not (0 <= alpha + beta <= 1):
        raise ValueError("Sum of alpha and beta must be in [0,1]")
    # Normalize Pearson correlation to [0,1] range
    normalized_pearson = (pearson_sim + 1) / 2
    # Compute weighted combination
    hybrid_score = (
        alpha * custom_sim +
        beta * cosine_sim +
        (1 - alpha - beta) * normalized_pearson
    )
    # Ensure output stays within [0,1] bounds
    return np.clip(hybrid_score, 0, 1)
```

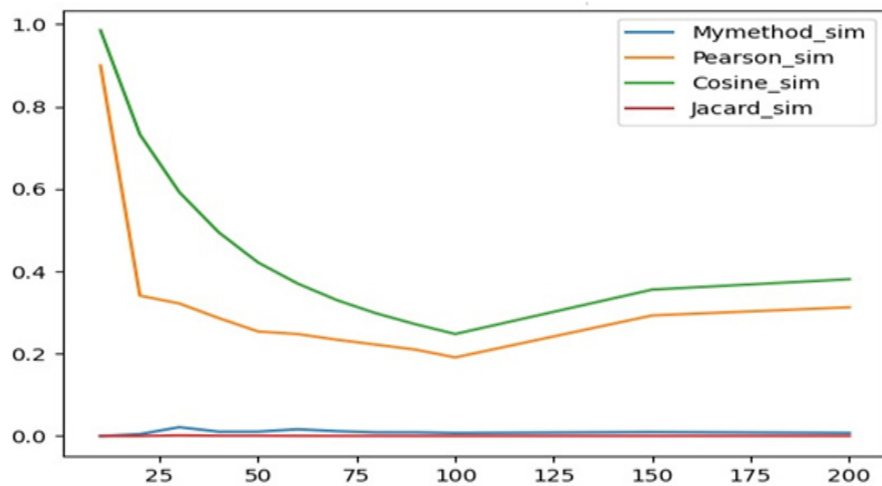
(شکل ۶). شبه کد پیاده‌سازی معیار پیشنهادی دوم

۴٫۶. ارزیابی معیار شباهت پیشنهادی اول و دوم

از آنجا که هسته اصلی سامانه توصیه‌گر، محاسبه شباهت مکانی بین کاربران است؛ ارزیابی سامانه به واسطه ارزیابی کیفیت برآورد شباهت بین کاربران امکان‌پذیر خواهد بود. از آنجا که ساختار ارتباطی کاربران یکی از داده‌های در دسترس در شبکه‌های اجتماعی مکان مبنا محسوب می‌شود و با این فرض که کاربرانی که در شبکه با یکدیگر در ارتباط هستند؛ احتمالاً با یکدیگر شباهت‌هایی نیز دارند، می‌توان با مقایسه کاربرانی که توسط سامانه مشابه تشخیص داده شده و کاربرانی که با یکدیگر در ارتباط هستند، به سنجش عملکرد سامانه پرداخت. مجموعه داده گوالا دارای ۱۲۸۰۹۶۹ موقعیت مکانی و تعداد کاربران ۱۰۷۰۹۲ است که در ارزیابی ما ۱۰۰۰۰۰ کاربر و ۵۰۰۰ موقعیت مکانی را به صورت تصادفی انتخاب کرده و معیار پیشنهادی را با معیارهای شباهت کسینوسی، ضریب همبستگی پیرسون و ضریب جاکارد مورد آزمایش و مقایسه قرار می‌دهیم.

(جدول ۱). جدول میانگین شباهت با توجه به تعداد مکان‌ها بین کاربران معیار پیشنهادی اول

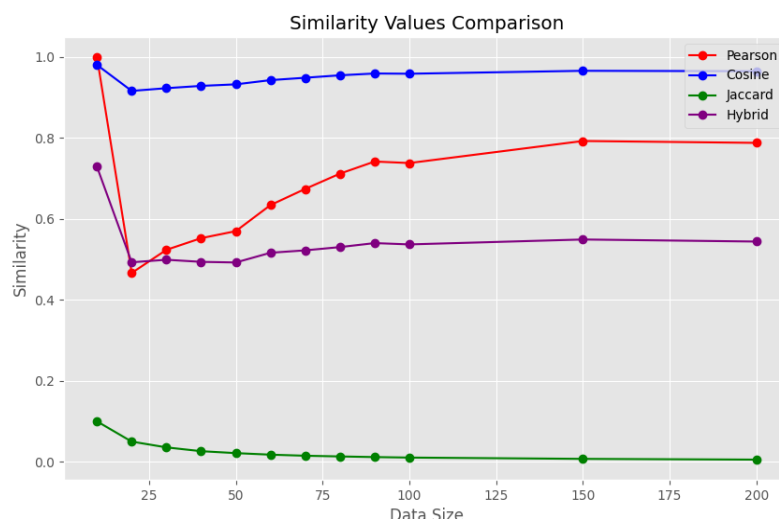
تعداد موقعیت‌ها	روش پیشنهادی	ضریب پیرسون	شباهت کسینوسی	شباهت جاکارد
۱۰	۰٫۰	۰٫۹	۰٫۹۸۵۴۲۴۶۴۶	۰٫۰
۲۰	۰٫۰۰۴۵۴۵۴۵۵	۰٫۳۴۰۹۰۹۱	۰٫۷۳۲۳۰۶۰۰۲	۰٫۰
۳۰	۰٫۰۲۱۴۴۷۲۱۹	۰٫۳۲۱۸۷۵	۰٫۵۹۲۲۰۲۲۸۳	۰٫۰۰۱۵۶۲۵
۴۰	۰٫۰۱۰۵۱۸۷۲۹	۰٫۲۸۶۴۲۸۵۷۱	۰٫۴۹۴۶۶۳۵۷۸	۰٫۰۰۰۷۱۴۲۸۶
۵۰	۰٫۰۱۰۶۲۴۴۳۸	۰٫۲۵۳۸۴۶۱۵۴	۰٫۴۲۱۳۰۸۹۵۰	۰٫۰۰۰۷۰۰۵۴۹
۶۰	۰٫۰۱۶۳۶۷۲۷۶	۰٫۲۴۷۶۹۵۸۵۳	۰٫۳۷۰۳۶۸۲۹۱	۰٫۰۰۰۴۱۹۶۸۴
۷۰	۰٫۰۱۱۷۶۱۰۰۹	۰٫۲۳۳۸۲۹۳۶۵	۰٫۳۲۹۹۲۸۰۸۴	۰٫۰۰۰۲۷۱۰۴۶
۸۰	۰٫۰۰۸۹۲۱۶۳۵	۰٫۲۲۱۸۴۹۵۹۳	۰٫۲۹۷۷۵۶۶۴۲	۰٫۰۰۰۲۶۱۳۲۴
۹۰	۰٫۰۰۸۹۶۷۳۷۸	۰٫۲۰۹۷۵۸۴۵۴	۰٫۲۷۱۲۸۶۴۱۸	۰٫۰۰۰۱۸۶۳۳۵
۱۰۰	۰٫۰۰۷۴۸۳۱۱۱	۰٫۱۹۰۸۶۹۸۷۵	۰٫۲۴۷۷۷۱۵۹۸	۰٫۰۰۰۱۳۷۵۱۰
۱۵۰	۰٫۰۰۹۶۲۹۱۱۹	۰٫۲۹۲۸۴۰۲۶۸	۰٫۳۵۵۶۰۶۲۸۶	۰٫۰۰۰۲۴۶۳۲۳
۲۰۰	۰٫۰۰۷۷۰۵۳۴۳	۰٫۳۱۲۶۴۰۵۹۱	۰٫۳۸۰۸۴۱۱۲۵	۰٫۰۰۰۱۷۵۷۱۴



(نمودار ۱). نمودار میانگین شباهت با توجه به تعداد مکان‌ها

(جدول ۲). جدول میانگین شباهت با توجه به تعداد مکان‌ها بین کاربران معیار پیشنهادی دوم.

تعداد موقعیت‌ها	روش پیشنهادی	ضریب پیرسون	شباهت کسینوسی	شباهت جاکارد
۱۰	۰٫۷۲۹۲	۱٫۰۰۰۰	۰٫۹۸۰۸	۰٫۱۰۰۰
۲۰	۰٫۴۹۳۱	۰٫۴۶۶۷	۰٫۹۱۶۵	۰٫۰۵۵۰
۳۰	۰٫۴۹۹۳	۰٫۵۲۳۸	۰٫۹۲۳۳	۰٫۰۳۵۷
۴۰	۰٫۴۹۴۱	۰٫۵۵۲۶	۰٫۹۲۸۹	۰٫۰۲۶۳
۵۰	۰٫۴۹۲۶	۰٫۵۷۰۰	۰٫۹۳۲۷	۰٫۰۲۱۵
۶۰	۰٫۵۱۶۶	۰٫۶۳۴۵	۰٫۹۴۳۲	۰٫۰۱۷۷
۷۰	۰٫۵۲۲۶	۰٫۶۷۴۴	۰٫۹۴۹۰	۰٫۰۱۵۱
۸۰	۰٫۵۳۰۶	۰٫۷۱۳۱	۰٫۹۵۵۱	۰٫۰۱۳۳
۹۰	۰٫۵۴۰۵	۰٫۷۴۲۰	۰٫۹۵۹۵	۰٫۰۱۱۸
۱۰۰	۰٫۵۳۷۲	۰٫۷۳۸۳	۰٫۹۵۹۰	۰٫۰۱۰۵
۱۵۰	۰٫۵۴۹۴	۰٫۷۹۲۸	۰٫۹۶۶۲	۰٫۰۰۷۳
۲۰۰	۰٫۵۴۴۳	۰٫۷۸۸۱	۰٫۹۶۵۲	۰٫۰۰۵۴



(نمودار ۲). نمودار میانگین شباهت با توجه به تعداد مکان‌ها بین کاربران در معیار پیشنهادی دوم

روش‌های پیشنهادی نسبت به معیارهای شباهت کسینوسی و معیار شباهت پیرسون در تعیین شباهت مکانی بین کاربران عملکرد ضعیف‌تری داشته و نسبت به روش جاکارد عملکرد بهتری دارند. با توجه به اینکه هدف در سامانه‌های توصیه‌گر افزایش دقت در توصیه است، صرفاً تعیین شباهت مکانی نمی‌تواند ملاک عمل قرار گیرد؛ بلکه باید برای سنجش دقت در معیار پیشنهادی، میانگین مطلق خطا و مجذور میانگین مربعات خطا را بررسی کرده و در صورتی که کمترین خطا را داشته باشند، معیارهای ارائه شده می‌توانند معیارهای مناسبی از نظر دقت توصیه باشند.

۵.۶. ملاک‌های ارزیابی

به منظور سنجش دقت در معیار پیشنهادی، معمولاً معیارهای ارزیابی متداولی مورد محاسبه قرار می‌گیرند که در این پژوهش از دو معیار خطای مطلق میانگین^۱ و جذر میانگین مربعات خطا^۲ استفاده می‌کنیم.

1. Mean Absolute Error
2. Root Mean Square Error

۱-۵-۶. میانگین مطلق خطا

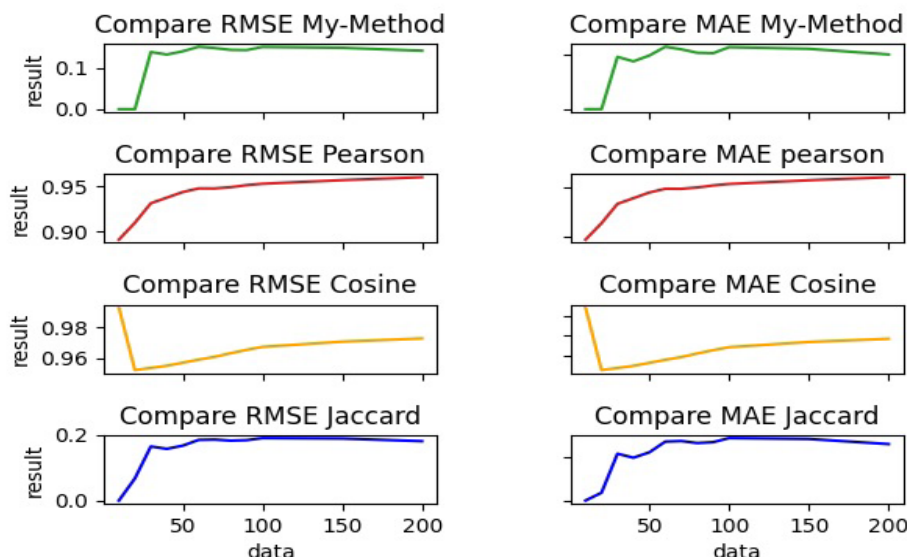
رایج‌ترین معیار برای مقایسه، میانگین مطلق خطا است که برای سنجش توانایی یک سامانه جهت پیش‌بینی علاقه یک کاربر به موردی خاص به کار می‌رود. این معیار، میانگین قدر مطلق تفاضل‌های مقدار واقعی و مقدار پیش‌بینی شده را به صورت زیر به دست می‌آورد:

۲-۵-۶. مجذور میانگین مربعات خطا

مجذور میانگین مربعات خطا نیز یکی از معیارهای ارزیابی در سامانه‌های توصیه‌گر است که به صورت رابطه زیر تعریف می‌گردد:

(جدول ۳) خطای روش پیشنهادی و مقایسه با خطای مطلق میانگین و جذر میانگین مربعات خطا در معیار پیشنهادی اول

User ID Number	RMSE				MAE			
	My- Method	Pearson	Cosine	Jaccard	My- Method	Pearson	Cosine	Jaccard
۱۰	۰	۰٫۸۹۰	۰٫۹۹۲	۰	۰	۰٫۷۹۳	۰٫۹۸۵	۰
۲۰	۰	۰٫۹۰۹	۰٫۹۵۲	۰٫۰۶	۰	۰٫۸۲۷	۰٫۹۰۷	۰٫۰۰۴
۳۰	۰٫۱۳۷	۰٫۹۳۱	۰٫۹۵۳	۰٫۱۶۵	۰٫۰۱۸	۰٫۸۶۷	۰٫۹۰۹	۰٫۰۲۷
۴۰	۰٫۱۳۱	۰٫۹۳۷	۰٫۹۵۵	۰٫۱۵۸	۰٫۰۱۷	۰٫۷۷۹	۰٫۹۱۲	۰٫۰۲۵
۵۰	۰٫۱۳۹	۰٫۹۴۳	۰٫۹۵۷	۰٫۱۶۷	۰٫۰۱۹	۰٫۸۹۰	۰٫۹۱۶	۰٫۰۲۸
۶۰	۰٫۱۵۰	۰٫۹۴۷	۰٫۹۵۹	۰٫۱۸۵	۰٫۰۲۲	۰٫۸۹۸	۰٫۹۱۹	۰٫۰۳۴
۷۰	۰٫۱۴۷	۰٫۹۴۷	۰٫۹۶۰	۰٫۱۸۶	۰٫۰۲۱	۰٫۸۹۸	۰٫۹۲۳	۰٫۰۳۴
۸۰	۰٫۱۴۲	۰٫۹۴۹	۰٫۹۶۲	۰٫۱۸۳	۰٫۰۲۰	۰٫۹۰۱	۰٫۹۲۷	۰٫۰۳۳
۹۰	۰٫۱۴۲	۰٫۹۵۱	۰٫۹۶۵	۰٫۱۸۴	۰٫۰۲۰	۰٫۹۰۵	۰٫۹۳۲	۰٫۰۳۴
۱۰۰	۰٫۱۴۹	۰٫۹۵۳	۰٫۹۶۷	۰٫۱۹۰	۰٫۰۲۲	۰٫۹۰۸	۰٫۹۳۵	۰٫۰۳۶
۱۵۰	۰٫۱۴۷	۰٫۹۵۷	۰٫۹۷۰	۰٫۱۸۹	۰٫۲۱	۰٫۹۱۶	۰٫۹۴۲	۰٫۰۳۵
۲۰۰	۰٫۱۴۰	۰٫۹۶۰	۰٫۹۷۲	۰٫۱۸۱	۰٫۰۱۹	۰٫۹۲۲	۰٫۹۴۶	۰٫۰۳۲



(نمودار ۳). نمودار خطای روش پیشنهادی و مقایسه با خطای مطلق میانگین و جذر میانگین مربعات خطا

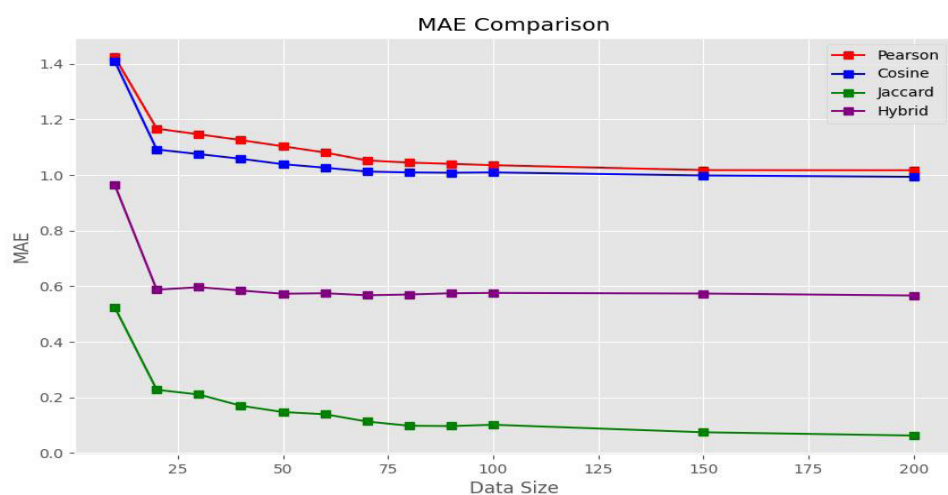
نتایج به وضوح نشان می‌دهد که روش پیشنهادی با جذر میانگین مربعات خطای صفر در داده‌های کم حجم (۱۰-۲۰ موقعیت) و مقادیر نزدیک به صفر (۰.۱۳۷-۰.۱۵۰) در داده‌های بزرگ‌تر، دقیق‌ترین روش است، درحالی‌که خطای مطلق میانگین آن نیز بین ۰.۰۲۲-۰ در نوسان است که نشان‌دهنده دقت خوب این روش است. در مقابل، روش‌های پیرسون با جذر میانگین مربعات خطا ۰.۸۹۰-۰.۹۶۰ و خطای مطلق میانگین ۰.۷۷۹-۰.۹۲۲ و کسینوسی با جذر میانگین مربعات خطا ۰.۹۷۲-۰.۹۵۲ و خطای مطلق میانگین ۰.۹۰۷-۰.۹۴۶ با خطای قابل توجهی مواجه هستند. این نتایج تأیید می‌کند که روش پیشنهادی نه تنها در هر دو معیار جذر میانگین مربعات خطا و خطای مطلق میانگین برتری مطلق دارد، بلکه این برتری با افزایش حجم داده نیز پایدار می‌ماند، درحالی‌که روش‌های سنتی به‌ویژه در داده‌های کم حجم، خطای فاحشی نشان می‌دهند.

(جدول ۴). خطای روش پیشنهادی و مقایسه با خطای مطلق میانگین در معیار پیشنهادی دوم

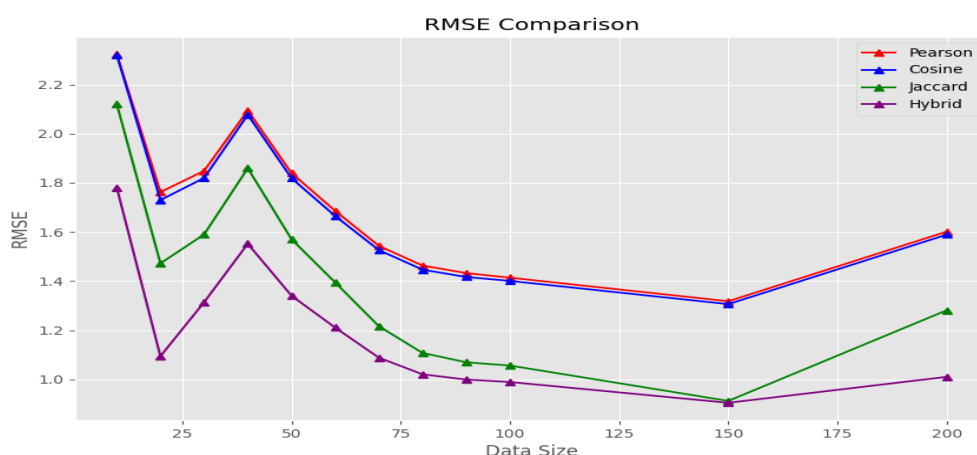
تعداد موقعیت‌ها	۱۰	۲۰	۳۰	۴۰	۵۰	۶۰	۷۰	۸۰	۹۰	۱۰۰	۱۵۰	۲۰۰
ضریب پیرسون	۱,۴۲۵۰	۱,۱۶۶۷	۱,۱۴۶۴	۱,۱۳۵۷	۱,۱۰۳۳	۱,۰۸۱۰	۱,۰۵۳۴	۱,۰۴۴۸	۱,۰۴۰۳	۱,۰۳۵۳	۱,۰۱۷۷	۱,۰۱۷۲
ضریب کسینوسی	۱,۴۰۹۶	۱,۰۹۱۸	۱,۰۷۵۳	۱,۰۵۸۴	۱,۰۳۹۰	۱,۰۲۶۳	۱,۰۱۲۵	۱,۰۰۹۵	۱,۰۰۸۲	۱,۰۰۹۴	۰,۹۹۸۵	۰,۹۹۳۶
ضریب جاکارد	۰,۵۵۵۰	۰,۲۲۷۸	۰,۲۱۰۷	۰,۱۷۰۴	۰,۱۴۷۶	۰,۱۳۹۸	۰,۱۱۳۵	۰,۰۹۸۴	۰,۰۹۷۳	۰,۱۰۱۹	۰,۰۷۴۹	۰,۰۶۳۱
روش ترکیبی	۰,۹۶۵۴	۰,۵۸۷۶	۰,۵۹۶۶	۰,۵۸۴۷	۰,۵۷۳۱	۰,۵۷۴۹	۰,۵۶۷۶	۰,۵۷۰۳	۰,۵۷۴۸	۰,۵۷۶۰	۰,۵۷۳۸	۰,۵۶۶۸

(جدول ۵). خطای روش پیشنهادی و مقایسه با جذر میانگین مربعات خطا در معیار پیشنهادی دوم

تعداد موقعیت‌ها	۱۰	۲۰	۳۰	۴۰	۵۰	۶۰	۷۰	۸۰	۹۰	۱۰۰	۱۵۰	۲۰۰
ضریب پیرسون	۲,۶۲۳۸	۱,۷۶۲۳	۱,۸۴۸۷	۲,۰۹۴۶	۱,۸۴۰۱	۱,۶۸۵۸	۱,۵۴۳۴	۱,۶۶۳۲	۱,۶۳۳۴	۱,۴۱۳۸	۱,۳۱۷۷	۱,۶۰۱۳
ضریب کسینوسی	۲,۳۲۲۰	۱,۷۳۹۷	۱,۸۱۹۸	۲,۰۷۷۲	۱,۸۱۹۰	۱,۶۶۴۹	۱,۵۲۶۱	۱,۴۴۶۴	۱,۴۱۶۵	۱,۴۰۰۶	۱,۳۰۶۰	۱,۵۸۹۰
ضریب جاکارد	۲,۱۲۱۳	۱,۴۷۲۰	۱,۵۸۹۸	۱,۸۵۹۴	۱,۵۷۰۶	۱,۳۹۴۴	۱,۲۱۶۱	۱,۱۰۷۳	۱,۰۶۹۲	۱,۰۵۶۰	۰,۹۱۱۹	۱,۲۸۰۹
روش ترکیبی	۱,۷۸۱۰	۱,۰۹۳۸	۱,۳۱۳۷	۱,۵۵۱۹	۱,۳۴۰۴	۱,۲۱۰۵	۱,۰۸۸۱	۱,۰۳۰۲	۰,۹۹۸۹	۰,۹۸۸۹	۰,۹۰۴۳	۱,۰۰۹۸



(نمودار ۴). نمودار خطای مطلق میانگین در روش پیشنهادی دوم



(نمودار ۵). نمودار خطای جذر میانگین مربعات خطا در روش پیشنهادی دوم

نتایج نمودار خطای مطلق میانگین به وضوح نشان می دهد که روش ترکیبی (هیبرید) در تمامی حجم های داده (چه کم و چه زیاد) کمترین میزان خطا را دارد و از این رو دقیق ترین روش محسوب می شود. این برتری ناشی از ترکیب هوشمندانه معیارهای مختلف است که نقاط ضعف هر روش را پوشش داده و خطای سامانمند را کاهش

می‌دهد. در مقابل، روش‌های تکی مانند جاکارد (با بیش‌ترین خطای مطلق میانگین) به دلیل محدودیت در پردازش داده‌های پیوسته و پیرسون و کسینوسی به ترتیب به دلیل حساسیت به پرت‌ها و نادیده‌گرفتن بزرگی داده‌ها، عملکرد ضعیف‌تری نشان می‌دهند. همچنین، با افزایش حجم داده، روش هیبرید پایداری بیشتری از خود نشان می‌دهد که گواهی بر مقیاس‌پذیری مناسب آن است. در روش ترکیبی محدوده خطای مطلق میانگین ۰.۵۶-۰.۵۹ بوده که از بین روش‌ها دقیق‌ترین روش با خطای پایین است. محدوده خطای مطلق میانگین معیار کسینوسی ۰.۹۹-۱.۴۰ و برای روش پیرسون ۱.۰۱-۱.۴۲ و روش جاکارد ۰.۵۲-۰.۰۶ است.

نتایج نمودار خطای جذر میانگین مربعات خطا به وضوح نشان می‌دهد که روش ترکیبی (هیبرید) با پایین‌ترین مقدار خطای جذر میانگین مربعات خطا (نزدیک به ۱.۰) در تمامی حجم‌های داده، از دقت و پایداری بالاتری نسبت به سایر روش‌ها برخوردار است. این برتری ناشی از ترکیب بهینه معیارهای مختلف است که نقاط ضعف هر روش را جبران می‌کند. در مقابل، روش‌های تکی مانند جاکارد (با بیش‌ترین خطای جذر میانگین مربعات خطا بین ۱.۶ تا ۲.۲) به دلیل محدودیت در پردازش داده‌های پیوسته و روش‌های پیرسون و کسینوسی به ترتیب به دلیل حساسیت به پرت‌ها و نادیده‌گرفتن بزرگی داده‌ها، عملکرد ضعیف‌تری دارند. با افزایش حجم داده، روش هیبرید نه تنها بهبود می‌یابد، بلکه پایداری خود را نیز حفظ می‌کند که نشان‌دهنده قابلیت اطمینان آن در مقیاس‌های بزرگ است. روش ترکیبی در ۱۰ موقعیت ۱.۷۸ و در موقعیت ۲۰۰ تقریباً برابر ۱.۰۱ است. روش جاکارد، پیرسون و کسینوسی به ترتیب در موقعیت ۱۰ برابر ۲.۳۲، ۲.۱۲ و ۲.۳۲ و در موقعیت ۲۰۰ به ترتیب ۱.۲۸، ۱.۵۹ و ۱.۶۰ است. روش ترکیبی با ۴۵ درصد کاهش خطا نسبت به نزدیک‌ترین رقیب (جاکارد) بهترین انتخاب است.

۷. نتیجه و پیشنهادهای پژوهشی

این پژوهش با هدف بهبود سنجش شباهت مکانی کاربران در سامانه‌های توصیه‌گر مکان مینا، دو معیار جدید ارائه داد. معیار اول مبتنی بر تحلیل مکان‌های مشترک و تاریخ بازدید، با تفکیک بازدیدهای هم‌زمان و ناهم‌زمان، امتیازی ترکیبی از مکان و زمان ارائه می‌دهد. معیار دوم یک رویکرد ترکیبی وزنی است که با تخصیص ضرایب قابل تنظیم،

هم‌زمان اشتراک‌های مکانی-زمانی، الگوی حرکتی (کسینوسی) و رابطه خطی رفتاری (پیرسون) را ادغام می‌کند تا ارزیابی چندبعدی و دقیق‌تری فراهم نماید. نتایج ارزیابی نشان می‌دهد که در مقایسه با روش‌های پایه، هر دو معیار پیشنهادی عملکرد بهتری نسبت به ضریب جاکارد دارند، اما در مقایسه با شباهت کسینوسی و ضریب پیرسون به‌تنهایی، نتایج ضعیف‌تری ارائه می‌دهند. با این حال، در سنجش دقت پیش‌بینی با معیارهای کمترین خطای مطلق و پایدارترین جذر خطا، هر دو معیار پیشنهادی برتری قابل توجهی از خود نشان دادند. به‌طور مشخص، معیار اول در داده‌های کم‌حجم به مجذور میانگین مربعات خطا صفر و در داده‌های حجیم‌تر به مقادیر بسیار پایین (۰.۱۳۷-۰.۱۵۰) دست‌یافت و خطای مطلق آن نیز ناچیز (۰-۰.۰۲۲) بود. در همین حال، معیار دوم در تمامی حجم‌های داده، کمترین خطای مطلق و پایین‌ترین و پایدارترین جذر خطا نزدیک به ۱۰۰ را ثبت کرد و از نظر دقت و پایداری برتر از همه روش‌ها شناخته شد. یافته‌ها به‌طور کلی تأیید می‌کنند که رویکرد ترکیبی پیشنهادی، با یکپارچه‌سازی چندین بُعد رفتاری، قادر است دقت و پایداری سامانه توصیه‌گر را به‌طور معناداری افزایش دهد. این نتایج گامی مؤثر در جهت توسعه توابع شباهت کارآمدتر و قوی‌تر برای ارائه خدمات شخصی‌سازی شده در شبکه‌های اجتماعی مکان مبنا محسوب می‌شود. پیشنهادهای برای پژوهش‌های آتی عبارتند از:

۱. استفاده از مدل‌های یادگیری عمیق برای یادگیری پویا و بهینه وزن‌های تلفیق معیارهای مختلف، به‌منظور مدل‌سازی روابط پیچیده‌تر و غیرخطی در رفتار کاربران.
۲. تلفیق ابعاد معناشناسی فعالیت‌ها و ملاحظات امنیتی یا حریم خصوصی در چارچوب سنجش شباهت، با بهره‌گیری از روش‌هایی چون پردازش زبان طبیعی، گراف دانش و محاسبات امن، برای ارائه توصیه‌های هم‌زمان دقیق‌تر و قابل اعتمادتر.

۷-۱. محدودیت‌های پژوهش

مهم‌ترین محدودیت این پژوهش، استفاده از تنها یک مجموعه داده یعنی گوالا و فرض ضمنی همسانی علایق در صورت اشتراک مکانی بوده است. دستیابی به نمرات عالی و پایدار در تمامی معیارها مستلزم آزمایش روی چندین مجموعه داده مستقل با ویژگی‌های جمعیت‌شناختی و مکانی متفاوت مانند فوراسکوئر، یلپ، گوگل مپس و

داده‌های بومی هر منطقه است. انجام چنین فرایندی به دلیل حجم بالای پیش پردازش، تفاوت در ساختار داده‌ها و زمان بر بودن پیاده‌سازی جداگانه برای هر مجموعه داده، نیازمند زمانی طولانی‌تر است؛ به عبارت دیگر، برای حصول اطمینان از تعمیم‌پذیری نتایج، به یک مطالعه فراتحلیل بر روی چندین دیتاست مستقل نیاز است که انجام آن در قالب یک مقاله و در بازه زمانی معمول پژوهش‌های کوتاه مدت امکان‌پذیر نیست. این محدودیت به عنوان یک جهت‌گیری روشن برای پژوهش‌های آتی مدنظر قرار گرفته است.

۸. منابع و مراجع

- دره‌شیری، شکوه و فرنقی، مهدی. (۱۳۹۴). توسعه سامانه درگاه مکانی توصیه‌گر بر اساس روش محتوی مینا و پالایش مشارکتی، همایش ملی ژئوماتیک، ۱-۱۰.
- گودرزی، صدیقه؛ آل شیخ، علی اصغر و هنرپرور، سپهر. (۱۳۹۶). توسعه الگوریتم توصیه‌گر مکان احتمالاتی بر اساس دسته‌بندی مکان در شبکه‌های اجتماعی مکان مینا مبتنی بر تأثیر زمانی، تأثیر مکانی و تأثیر روابط اجتماعی. علوم و فنون نقشه‌برداری، ۲۳۰-۲۱۵.
- قباخلو نسیبه؛ رجبی محمدعلی؛ ناصری فرزین و قباخلو مجتبی. (۱۳۹۲). کاربرد داده‌کاوی مکانی-زمانی برای ایجاد سرویس‌های مکان مینا هوشمند. همایش ملی ژئوماتیک، ۱-۱۰.
- Bao, J., Zheng, Y., Wilkie, D., & Mokbel, M. F. (2013). A survey on recommendations in location-based social networks. *ACM Transaction on Intelligent Systems and Technology*, 1-30.
- Bogorny, V., Renso, C., de Aquino, A. R., de Lucca Siqueira, F., & Alvares, L. O. (2014). Constant-a conceptual data model for semantic trajectories of moving objects. *Transactions in GIS*, 18(1), 66-88.
- Chakri, S., Raghay, S., & El Hadaaj, S. (2017). Semantic Trajectory Knowledge Discovery: A Promising Way to Extract Meaningful Patterns from Spatiotemporal Data. *International Journal of Software Engineering and Knowledge Engineering*, 27(03), 399-421.
- Chang, W., Sun, D., & Du, Q. (2023). Intelligent sensors for POI recommendation model using deep learning in location-based social network big data. *Sensors*, 23(2), 850.
- D'Ulizia, A., Grifoni, P., & Ferri, F. (2021). Query processing of geosocial data in location-based social networks. *ISPRS International Journal of Geo-Information*, 11(1), 19.
- Dey, A. K., Abowd, G. D., & Salber, D. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human-Computer Interaction*, 16(2-4), 97-166.
- Goldenberg, J., & Levy, M. (2009). Distance is not dead: Social interaction and geographical distance in the internet era. *arXiv preprint arXiv:0906.3202*.
- Guo, L., Jiang, H., & Wang, X. (2018). Location regularization-based poi recommendation in location-based social networks. *Information*, 9(4), 85.
- Li, X., Zhang, Y., Chen, H., & Wang, Z. (2025). Multi-factor POI recommendation: Integrating spatial, social, and preference contexts. *Journal of Location-Based Services*, *19*(3), 215-234.
- Liu, B., Xiong, H., Papadimitriou, S., Fu, Y., & Yao, Z. (2014). A general geographical probabilistic factor model for point of interest recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 27(5), 1167-1179.

- Meena, G., Indian, A., Mohbey, K. K., & Jangid, K. (2024). Point of Interest Recommendation System Using Sentiment Analysis. *Journal of Information Science Theory and Practice*, 12(2), 64-78.
- Noulas, A., Scellato, S., Lathia, N., & Mascolo, C. (2012, September). A random walk around the city: new venue recommendation in location-based social networks. In 2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing (pp. 144-153). Ieee.
- Quercia, D., Lathia, N., Calabrese, F., Di Lorenzo, G., & Crowcroft, J. (2010, December). Recommending social events from mobile phone location data. In 2010 IEEE international conference on data mining (pp. 971-976). IEEE.
- Rehman, F., Khalid, O., & Madani, S. A. (2017). A comparative study of location-based recommendation systems. *The Knowledge Engineering Review*, 32, e7.
- Si, Y., Zhang, F., & Liu, W. (2019). An adaptive point-of-interest recommendation method for location-based social networks based on user activity and spatial features. *Knowledge-Based Systems*, 163, 267-282.
- Tarafdar, M., & Minaei-Bidgoli, B. (2024). A Co-occurrence Prediction Framework in Location-Based Social Networks. *New Generation Computing*, 42(5), 1129-1163.
- Vukovic, M. & D. Jevtic (2015). Agent-based Movement Analysis and Location Prediction in Cellular Networks. *Procedia Computer Science*, 60, 517-526.
- Xu, Q., Wang, J., & Xiao, B. (2017, October). Personalized location recommendation for location-based social networks. In 2017 IEEE/CIC International Conference on Communications in China (ICCC) (pp. 1-6). IEEE.
- Zhang, J. D., & Chow, C. Y. (2016). Point-of-interest recommendations in location-based social networks. *Sigspatial Special*, 7(3), 26-33.
- Zhang, Z., Liu, Y., Zhang, Z., & Shen, B. (2019). Fused matrix factorization with multi-tag, social and geographical influences for POI recommendation. *World Wide Web*, 22(3), 1135-1150.
- Zhao, G., Lou, P., Qian, X., & Hou, X. (2020). Personalized location recommendation by fusing sentimental and spatial context. *Knowledge-Based Systems*, 196, 105849.
- Zhao, S., Lyu, M. R., & King, I. (2018). Point-of-Interest Recommendation in Location-Based Social Networks. Springer Singapore. pp 1-28.
- Zhao, X., Zhang, Z., Bi, X., & Sun, Y. (2023). A new point-of-interest group recommendation method in location-based social networks. *Neural Computing and Applications*, 1-12.

- Zheng, Y. (2011). Location-based social networks: Users. In *Computing with spatial trajectories* (pp. 243-276). Springer, New York, NY.
- Zhang, Y., & Liu, Y. (2024). Point-of-interest recommendation based on LBSN with multi-aspect fusion of social and individual features. *Neural Computing and Applications*, 36(20), 12163-12184.
- Ju, C., Zhou, H., Bao, F., & Xu, C. (2025). A multi-spatiotemporal joint next-POI travel sequence recommendation method based on federated learning. *Pattern Recognition*, 112867.
- Zhang, Q., Yang, P., Yu, J., Wang, H., He, X., Yiu, S. M., & Yin, H. (2025). A survey on point-of-interest recommendation: Models, architectures, and security. *IEEE Transactions on Knowledge and Data Engineering*.
- Wang, L., et al. (2022). "Spatial-Social Preference Fusion for POI Recommendation via Geometric Deep Learning". *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 34(8), 3872-3885.
- Chen, H., et al. (2021). "Hybrid Spatial Similarity Modeling for Next-POI Recommendation". *Information Sciences*, 547, 1-18.
- Yang, D., et al. (2020). "Modeling User Spatial Preference for Personalized Recommendation". *ACM Transactions on Intelligent Systems and Technology (TIST)*, 11(3), 1-25.
- Park, M. H., Hong, J. H., & Cho, S. B. (2007, July). Location-based recommendation system using Bayesian user's preference model in mobile devices. In *International conference on ubiquitous intelligence and computing* (pp. 1130-1139). Springer, Berlin, Heidelberg.
- Li, N., & Chen, G. (2009, July). Multi-layered friendship modeling for location-based mobile social networks. In *2009 6th Annual International Mobile and Ubiquitous Systems: Networking & Services, MobiQuitous* (pp. 1-10). IEEE.
- Yu, X., Pan, A., Tang, L. A., Li, Z., & Han, J. (2011, July). Geo-friends' recommendation in gps-based cyber-physical social network. In *2011 International Conference on Advances in Social Networks Analysis and Mining* (pp. 361-368). IEEE.
- Cho, E., Myers, S. A., & Leskovec, J. (2011, August). Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 1082-1090).
- Zhou, D., Wang, B., Rahimi, S. M., & Wang, X. (2012, May). A study of recommending locations on location-based social network by collaborative filtering. In *Canadian Conference on Artificial Intelligence* (pp. 255-266). Springer, Berlin, Heidelberg.

- Cheng, C., Yang, H., King, I., & Lyu, M. (2012, July). Fused matrix factorization with geographical and social influence in location-based social networks. In Proceedings of the AAAI conference on artificial intelligence (Vol. 26, No. 1).
- Bao, J., Zheng, Y., & Mokbel, M. F. (2012, November). Location-based and preference-aware recommendation using sparse geo-social networking data. In Proceedings of the 20th international conference on advances in geographic information systems (pp. 199-208). ACM.
- Wang, H., Terrovitis, M., & Mamoulis, N. (2013, November). Location recommendation in location-based social networks using user check-in data. In Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 374-383). ACM.
- Zhang, J. D., & Chow, C. Y. (2013, November). iGSLR: personalized geo-social location recommendation: a kernel density estimation approach. In Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 334-343).
- Zhao, S., King, I., & Lyu, M. R. (2013, November). Capturing geographical influence in POI recommendations. In International Conference on Neural Information Processing (pp. 530-537). Springer, Berlin, Heidelberg.
- Gao, H., Tang, J., Hu, X., & Liu, H. (2015, February). Content-aware point of interest recommendation on location-based social networks. In Proceedings of the AAAI conference on artificial intelligence (Vol. 29, No. 1). pp. 1721 - 1727.
- Zhang, J. D., & Chow, C. Y. (2015, August). Geosoca: Exploiting geographical, social and categorical correlations for point-of-interest recommendations. In Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval (pp. 443-452).
- Li, H., Ge, Y., Hong, R., & Zhu, H. (2016, August). Point-of-interest recommendations: Learning potential check-ins from friends. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 975-984).
- Zhao, S., King, I., & Lyu, M. R. (2017, November). Geo-pairwise ranking matrix factorization model for point-of-interest recommendation. In International Conference on Neural Information Processing (pp. 368-377). Springer, Cham.
- Kheiri, A., Karimpour, F., & Forghani, M. (2016). Intra-Urban Movement Patterns Estimation Based on Location Based Social Networking Data. Journal of Geomatics Science and Technology, 6(1), 141-158.

- Sánchez, P., Bellogín, A., & Boratto, L. (2023). Bias characterization, assessment, and mitigation in location-based recommender systems. *Data Mining and Knowledge Discovery*, 37(5), 1885-1929.
- Zheng, Y., Chen, Y., Xie, X., & Ma, W. Y. (2009, May). GeoLife2. 0: a location-based social networking service. In 2009 tenth international conference on mobile data management: systems, services and middleware (pp. 357-358). IEEE.
- Zhang, Q., Yang, P., Yu, J., Wang, H., He, X., Yiu, S. M., & Yin, H. (2024). A Survey on Point-of-Interest Recommendation: Models, Architectures, and Security. *arXiv preprint arXiv:2410.02191*.
- Cheng, C., Yang, H., Lyu, M. R., & King, I. (2013, August). Where you like to go next: Successive point-of-interest recommendation. In *IJCAI* (Vol. 13, pp. 2605-2611).
- Scellato, S., Noulas, A., & Mascolo, C. (2011, August). Exploiting place features in link prediction on location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 1046-1054).
- Yan, Z., Chakraborty, D., Parent, C., Spaccapietra, S., & Aberer, K. (2013). Semantic trajectories: Mobility data computation and annotation. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 4(3), 1-38.
- Xiao, X., Zheng, Y., Luo, Q., & Xie, X. (2010, November). Finding similar users using category-based location history. In *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems* (pp. 442-445).
- Park, M. H., Hong, J. H., & Cho, S. B. (2007, July). Location-based recommendation system using Bayesian user's preference model in mobile devices. In *International conference on ubiquitous intelligence and computing* (pp. 1130-1139). Springer, Berlin, Heidelberg.
- Casillo, M., Colace, F., Conte, D., Lombardi, M., Santaniello, D., & Valentino, C. (2026). A context-aware recommender system-based framework for cultural heritage enhancement. **User Modeling and User-Adapted Interaction*, 1-32.
- Wang, Y., Wang, Y., Zhang, Z., Deng, J., Ye, J., & Zhang, L. Y. (2025). Differentially private recommendation algorithm based on diffusion model and Rényi similarity. *Information Sciences*, 122565.
- Bettache, D., Dennouni, N., & Harbouche, A. (2025). Enhancing points of interest recommendation by integrating users' proximity into the calculation of their similarities. *Indonesian Journal of Electrical Engineering and Computer Science*, 40(1), 379-387.

- Wu, et al. (2025). Implicit Trust-Based POI Recommendations Using User Check-Ins. In Proceedings of the International Conference on Electrical Engineering and Computer Applications (ICEECA 2025). IEEE.
- Acharya, M., Mohbey, K. K., & Rajput, D. S. (2024). Long-term preference mining with temporal and spatial fusion for point-of-interest recommendation. IEEE Access, 12, 11584-11596.
- Chen, H., et al. (2025). "User Similarity Measurement in Location-Based Social Networks: A Hybrid Approach." Expert Systems with Applications, Vol. 245.