

مرور روش‌های مدل‌سازی عوامل مکانی در توصیه مکان جذاب با استفاده از اطلاعات شبکه‌های اجتماعی مکان مبنا

مهدی رضائیان^{۱*}، رحیم علی عباسپور^۲، زهرا بهرامیان^۳

^۱ دانشجوی کارشناسی ارشد سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی -

پردیس دانشکده‌های فنی - دانشگاه تهران

mahdi.rezaeyan@ut.ac.ir

^۲ استادیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - پردیس دانشکده‌های فنی - دانشگاه تهران

abaspour@ut.ac.ir

^۳ دکتری سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - پردیس دانشکده‌های فنی

- دانشگاه تهران

zbahramian@ut.ac.ir

(تاریخ دریافت بهمن ۱۳۹۸، تاریخ تصویب اسفند ۱۳۹۸)

چکیده

رشد سریع فناوری‌های ساخت تلفن همراه و ترکیب آن با فناوری‌های مختلف، باعث اضافه شدن بعد مکان به شبکه‌های اجتماعی و تشکیل شبکه‌های اجتماعی مکان مبنا شده است. در شبکه‌های اجتماعی مکان مبنا از سیستم‌های توصیه‌گر برای توصیه مکان‌های جذاب به کاربران استفاده می‌شود. سیستم‌های توصیه‌گر سنتی مثل توصیه فیلم و کتاب سابقه دیرینه‌ای دارند. با وجود این، به دلیل وجود مولفه مکانی و ارتباط فیزیکی کاربران با دنیای بیرون در شبکه‌های اجتماعی مکان مبنا، یک سری ویژگی‌های خاص مانند عوامل مکانی، زمانی و اجتماعی، به منظور بهبود توصیه‌ها در نظر گرفته می‌شود. در میان ویژگی‌های خاص داده‌های شبکه‌های اجتماعی مکان مبنا، عوامل مکانی نقش مهمی را در بهبود توصیه‌ها ایفا می‌کنند. چرا که تمایل افراد برای بازدید از مکان‌ها به نسبت زیادی متأثر از فاصله بین شخص و مکان مورد نظر است. همچنین توزیع مکان‌های جذاب در منطقه باعث تغییر در الگوی بازدید کاربران می‌شود. در این تحقیق ابتدا با مقایسه سیستم‌های توصیه‌گر مکانی با سیستم‌های توصیه‌گر سنتی چالش‌هایی که شبکه‌های اجتماعی مکان مبنا با آن روبه‌رو هستند مورد بحث و بررسی قرار می‌گیرد. در ادامه عوامل موثر بر توصیه مکان جذاب به تفصیل بررسی شده و سیستم‌های توصیه‌گر مکان مبنا که از اطلاعات شبکه‌های اجتماعی مکان مبنا برای ارائه توصیه استفاده می‌کنند، بررسی می‌شود. در نهایت انواع روش‌های مدل‌سازی عوامل مکانی که یکی از مهمترین عوامل موثر در توصیه مکان جذاب به کاربران با استفاده از داده‌های شبکه‌های اجتماعی مکان مبنا است، بحث و بررسی می‌گردد.

واژگان کلیدی: توصیه مکان جذاب، شبکه‌های اجتماعی مکان مبنا، مدل‌سازی عوامل مکانی، سیستم‌های توصیه‌گر

* نویسنده رابط

۱- مقدمه

رشد سریع فناوری‌های ساخت تلفن همراه و ترکیب آن با تکنولوژی‌های مختلف مانند 3G، 4G و GPS، باعث آسانتر شدن اخذ داده‌های مکانی شد. همین امر موجب شد تا بعد مکانی نیز به عنوان بعدی جدید در ساختار شبکه‌های اجتماعی مطرح شود [۱]. با این تغییر شبکه‌های اجتماعی مکان مبنا^۱ متولد شدند. شبکه‌های اجتماعی مکان مبنا مثل Foursquare دسته‌ای از شبکه‌های اجتماعی هستند که کاربران در آن می‌توانند به فعالیت‌های خود برچسب مکانی بزنند و در مود فعالیتی که انجام می‌دهند توضیحاتی بنویسند [۲]. در طی سالیان اخیر شبکه‌های اجتماعی مکان مبنا هم از نظر کمی و هم از نظر محبوبیت رشد فراوانی داشتند و باعث تغییرات عمده‌ای در دنیای وب شدند.

برای بهبود تجربیات کاربران در شبکه‌های اجتماعی مکان مبنا با استفاده از پیشنهاد مکان‌های جذاب^۲، سیستم‌های توصیه‌گر به کار گرفته شدند [۲]. سیستم‌های توصیه‌گر^۳ تکنیک‌ها و ابزارهای نرم‌افزاری هستند که برای یک کاربر به خصوص، پیشنهادی مناسب با علایق وی فراهم می‌کنند [۱]. این پیشنهادها در حوزه‌های مختلفی که نیاز به انتخاب از بین چندین گزینه وجود دارد، صورت می‌گیرد. انتخاب یک فیلم برای تماشا، انتخاب یک موسیقی برای گوش دادن و انتخاب یک رستوران از بین رستوران‌های مختلف نمونه‌ای از مواردی هستند که در آنها سیستم‌های توصیه‌گر می‌تواند به کاربران کمک کند.

استفاده از سیستم‌های توصیه‌گر در شبکه‌های اجتماعی با توجه به این که این شبکه‌ها داده‌های مختلفی را به صورت صریح و ضمنی از کاربران می‌گیرند امری مفید و معمول است. در این بین شبکه‌های اجتماعی مکان مبنا با اخذ داده‌های مکانی از کاربران در کنار سایر داده‌ها، از قابلیت بیشتری برای توصیه اقلام به کاربران به خصوص اقلام مکانی (مانند رستوران‌ها) برخوردارند. به طور کلی سیستم‌های توصیه‌گر در شبکه‌های اجتماعی مکان مبنا دارای چهارچوب مشترک با سیستم‌های توصیه‌گر در شبکه‌های اجتماعی سنتی هستند. چرا که در هر دو مورد

یک سری گزینه وجود دارد که از بین آنها قرار است با توجه به سلايق کاربر، چند مورد به او توصیه شود. تفاوت اصلی بین این دو مورد استفاده از تاریخچه‌ی مکانی کاربران موجود در شبکه‌های اجتماعی مکان مبنا می‌باشد [۲].

سیستم‌های توصیه‌گر در شبکه‌های اجتماعی مکان مبنا را که به توصیه مکان جذاب می‌پردازند به اختصار، توصیه‌گر مکان مبنا مینامیم. در توصیه‌گر مکان مبنا هدف پیشنهاد چند نمونه مکان جذاب به کاربران باتوجه به پیشینه‌ی بازدید و همچنین زمینه‌ی کاربران و مکان‌های جذاب است. با توجه به این که توصیه‌ی مکان‌های (مثل رستوران و سینما) به کاربر یکی از شاخه‌های سیستم‌های توصیه‌گر است، بنابراین تکنیک‌های مرسوم مورد استفاده در سیستم‌های توصیه‌گر مانند فیلترینگ مشارکتی^۴ می‌توانند در این زمینه به کار گرفته شوند. اما وجود بعد مکان باعث ایجاد چالش‌هایی در تکنیک‌های سیستم‌های توصیه‌گر شده است [۲، ۳]

چالش‌هایی که در سیستم‌های توصیه‌گر مکان مبنا نسبت به توصیه‌گرهای غیر مکانی وجود دارد را می‌توان در موارد زیر خلاصه کرد:

۱- محدودیت‌های فیزیکی: بازدید از یک مکان فعالیت فیزیکی است. بنابراین بر خلاف تماشای فیلم یا خرید آنلاین دارای محدودیت‌های خاص خود هست. به عنوان مثال مردم دوست دارند از مکان‌های جذابی که در اطراف خانه یا محل کارشان قرار دارد بازدید کنند. همچنین مکان‌های جذاب مثل رستوران‌ها در ساعات خاصی از شبانه روز ممکن است خدمات ارائه بدهند که این نیز منجر به محدودیت‌های زمانی می‌شود [۲].

۲- پراکندگی گسترده اطلاعات: هر کاربر در شبکه‌های اجتماعی مکان مبنا ممکن است در سال، از حدود چند صد مکان جذاب بازدید کند در صورتی که میلیون‌های مکان جذاب در این شبکه‌ها وجود دارد. بنابراین توصیه چندین مکان جذاب در این حالت برای کاربر کار سختی است و نیازمند تدابیر ویژه هست. این مساله به خصوص در حالتی که کاربر تازه به شبکه وارد شده و اطلاعاتی از او در دست نیست (مساله شروع سرد) مشکل بزرگی است [۱، ۲]. در سیستم‌های توصیه‌گر سنتی مثل توصیه فیلم نیز این مشکل کم و بیش وجود دارد. اما با توجه به این که در

^۱ LocationBased Social Networks LBSNs

^۲ Point Of Interest (POI)

^۳ Recommender System (RS)

^۴ Collaborative Filtering (CF)

شبکه‌های اجتماعی مکان مبنا تعداد مکان‌هایی که کاربران می‌توانند از آنها بازدید کنند بسیار بیشتر از تعداد کل فیلم‌های ساخته شده است، توجه به این مساله در سیستم‌های توصیه‌گر مکانی ضروری‌تر است [۲].

۳- ناهمگونی اطلاعات: در شبکه‌های اجتماعی مکان مبنا در کنار داده‌های بازدید از مکان‌ها، اطلاعات مربوط به روابط اجتماعی کاربران و نظرات کاربران نسبت به مکان‌ها وجود دارد. وجود این اطلاعات ناهمگون باعث ایجاد سیستم‌های توصیه‌گر مختلف شده است که هر کدام از یک جنبه به توصیه مکان پرداخته اند [۱، ۲].

علاوه بر سه مورد فوق، موارد دیگری نیز به عنوان چالش در توصیه‌گرهای مکان مبنا مطرح شده اند. از جمله این موارد می‌توان به نیاز به رابط گرافیکی فشرده تر، توانایی مقیاس پذیری، تامین امنیت کاربران و نیاز به پاسخ سریع اشاره کرد [۱]. بیشتر این چالش‌ها با رشد تکنولوژی کمرنگ شده‌اند.

در این تحقیق با توجه به این که هدف اصلی بررسی روش‌های مختلف اخذ عوامل مکانی است لذا سیستم‌های ارائه شده بیشتر برای حل چالش محدودیت‌های فیزیکی به کار گرفته شده‌اند. از طرف دیگر، چالش ناهمگونی اطلاعات بیشتر از این که نکته‌ی منفی در سیستم‌های توصیه‌گر محسوب شود، تبدیل به عاملی برای بهبود توصیه‌ها با استفاده از اطلاعات مختلف مثل زمان بازدید از مکان یا نظر کاربران شده است. علاوه بر این، در سیستم‌های توصیه‌گر مکانی برای غلبه بر گستردگی اطلاعات از تکنیک‌هایی استفاده میشود که در نبود اطلاعات کافی از کاربر نیز بتواند به توصیه‌ی مکان جذاب بپردازد.

مساله‌ی اصلی این تحقیق در وهله اول بررسی عوامل موثر بر توصیه‌ی مکان جذاب به کاربران با استفاده از اطلاعات شبکه‌های اجتماعی مکان مبنا است. علاوه بر این، با توجه به این که عوامل مکانی یکی از مهم‌ترین عوامل تاثیرگذار در توصیه‌ی مکان جذاب به کاربران با استفاده از داده‌ی شبکه اجتماعی مکان مبنا هستند، مرور این عوامل امری ضروری است. بنابراین در وهله دوم به بررسی و ارائه‌ی روش‌های مختلف مدل‌سازی عوامل مکانی در آن‌ها پرداخته می‌شود. هدف از این کار جمع‌بندی روش‌های مختلف مدل‌سازی عوامل مکانی به منظور استفاده از روش و یا روش‌های بهینه در توصیه‌گرهای مکانی می‌باشد.

این مقاله در ۵ بخش کلی تنظیم شده است. در بخش دوم عوامل موثر بر توصیه مکان‌های جذاب با توجه به اطلاعات موجود در شبکه‌های اجتماعی مکان مبنا مورد بحث قرار می‌گیرد. بخش سوم به تقسیم بندی سیستم‌های توصیه‌گر مکانی با توجه به نحوه برخورد سیستم با عوامل موثر بر آن در شبکه‌های اجتماعی مکان مبنا می‌پردازد. سپس در بخش چهارم به بررسی روش‌های مختلف مدل‌سازی عوامل مکانی در سیستم‌های توصیه‌گر مکانی با استفاده از اطلاعات شبکه‌های اجتماعی مکان مبنا پرداخته می‌شود. در نهایت بخش پنجم به جمع‌بندی مطالب ارائه شده می‌پردازد.

۲- عوامل موثر بر توصیه‌ی مکان جذاب

با توجه به ویژگی‌های زمانی و مکانی ناشی از محدودیت‌های فیزیکی، وجود اطلاعات متنوع و گسترده ناشی از ناهمگونی اطلاعات و همچنین نیاز به عوامل محدود کننده‌ی انتخاب کاربران به دلیل پراکندگی اطلاعات در شبکه‌های اجتماعی مکان مبنا، فعالیت بازدید کاربر از یک مکان ترکیبی از چندین عامل متنوع می‌باشد. در مقالات مختلف عوامل مختلفی معرفی شده‌اند. در ادامه شش عامل اصلی و عمده شامل عوامل مکانی، اجتماعی^۱، محتوایی^۲، دسته‌بندی^۳، توالی^۴ و زمانی که در اکثر سیستم‌های توصیه‌گر مکانی در شبکه‌های اجتماعی مکان مبنا از آنها استفاده شده ذکر می‌گردد.

۲-۱- عامل مکانی

عامل مکانی عامل مهمی است که توصیه‌ی مکان جذاب را از دیگر روش‌های توصیه جدا می‌کند، چرا که رفتار بازدیدی کاربران وابسته به ویژگی‌های مکانی مکان‌های جذاب و محل کاربر است [۴]. تحلیل داده‌های بازدید کاربران نشان می‌دهد که کاربران در مناطق با قیود جغرافیایی خاص فعالیت می‌کنند. به عنوان مثال، مردم تمایل دارند به مکان‌های جذاب نزدیک به خانه‌ها و یا دفاتر کار خود مراجعه کنند. همچنین ممکن است آنها علاقمند به کاوش در مکان‌های نزدیک مکان‌های جذاب بازدید شده

^۱ Social

^۲ Content

^۳ Categorical

^۴ Sequential

خود باشند. عوامل مکانی می‌توانند کاندیداهای مکان‌های جذاب را کاهش داده و باعث کمتر شدن اثرات منفی گسترده‌ی مکان‌های جذاب شوند [۲، ۴-۱۷].

۲-۲- عامل اجتماعی

با توجه به این که، دوستان در شبکه‌های اجتماعی علایق مشترک بیشتری نسبت به سایر افراد دارند، عامل اجتماعی به عنوان یکی از عوامل تاثیرگذار در بهبود توصیه‌ها در نظر گرفته می‌شود. در دنیای واقعی، افراد معمولاً همراه دوستان خود به مکان‌های جذاب مثل رستوران و سینما می‌روند. حتی در صورتی که امکان بازدید همزمان وجود نداشته باشد، معمولاً افراد به جاهایی که دوستانشان از آن بازدید کرده‌اند می‌روند. بنابراین، اولویت فرد در انتخاب مکان‌های جذاب می‌تواند تحت تاثیر دوستان نزدیک خود یا یک گروه از دوستان با علائق مشترک باشد. لازم به ذکر است که دوستان در شبکه‌های اجتماعی مکان مبنا لزوماً تنها کسانی نیستند که افراد در دنیای واقعی با آنها در ارتباط هستند. به عنوان مثال افرادی که تعداد زیادی بازدید مشترک دارند، به نوعی با هم دیگر ارتباط دوستی دارند که از علایق مشترک آنها نشأت می‌گیرد [۲، ۴، ۵، ۸، ۹، ۱۴].

۲-۳- عامل محتوایی

در شبکه‌های اجتماعی مکان مبنا کاربران اطلاعاتی شامل عکس، ویدیو و نظر خود را در مورد مکان‌های جذاب منتشر می‌کنند که تحت عنوان عوامل محتوایی شناخته می‌شوند. با وجود این که افراد در همه‌ی بازدیدهای خود این اطلاعات را تولید نمی‌کنند، اما اطلاعات و محتوای کم موجود می‌تواند در توصیه به کاربران بسیار مفید واقع شود. استفاده از این محتوا از این جهت مهم است که این اطلاعات به صورت واضح‌تری ترجیحات کاربر را به نمایش می‌گذارند. به عنوان نمونه بازدید از یک رستوران لزوماً دلیل بر پسند کردن آن نیست. ممکن است کاربری به این دلیل که زمان زیادی برای جستجوی رستوران ندارد مجبور باشد به تنها رستوران موجود در محل کار خود رجوع کند [۲، ۱۸-۲۱].

۲-۴- عامل دسته‌بندی

دسته‌بندی هر مکان جذاب نشان‌دهنده‌ی ماهیت فعالیت‌های آن مکان است. بنابراین می‌توان مجموعه‌ای از

مکان‌های جذاب را در دسته‌های مرتبط به هم قرار داد. به عنوان مثال، بازدید فرد از یک رستوران نشان می‌دهد که او مایل است غذا بخورد و بازدید او از سینما بدین معنی است که او در حال تماشای یک فیلم در آنجا است. در عمل واکنش افراد نسبت به دسته‌های مختلف مکان‌های جذاب با همدیگر متفاوت است. به عنوان نمونه افرادی که علاقمند به مسافرت و آشنایی با افراد و فرهنگ‌های مختلف هستند معمولاً برای بازدید از جاذبه‌های توریستی در سراسر جهان اقدام می‌کنند. این در حالی است که ممکن است برای شخصی که به انواع غذاها اهمیت می‌دهد بازدید از رستوران‌ها در اولویت باشد [۳، ۹].

۲-۵- عامل توالی

معمولاً بین کارهای متوالی که انسان‌ها انجام می‌دهند رابطه معنایی وجود دارد که این رابطه عامل توالی نامیده می‌شود. در واقع، رفتار حرکتی انسان‌ها الگوهای متوالی را نشان می‌دهند. به عنوان مثال، افراد سینماها را معمولاً پس از رستوران‌ها بازدید می‌کنند، زیرا کاربران پس از صرف غذا می‌خواهند استراحت کرده و به آرامش برسند. همچنین بازدید از استادیوم و سپس رفتن به رستوران، معمولاً تر از معکوس آن است، زیرا بعد از خوردن یک وعده غذایی، ورزش کردن و فعالیت فیزیکی زیاد رفتار سالمی نیست. بنابراین، تأثیر الگوهای توالی نیز بر رفتارهای بازدید کاربر مهم است [۳، ۸].

۲-۶- عامل زمانی

زمان عامل مهمی در توصیه‌ی مکان جذاب است و در نظر گرفتن زمان‌های مختلف مانند روزهای وسط هفته و تعطیلات آخر هفته تأثیر به سزایی در انتخاب فعالیت‌های انسانی می‌گذارد. به عنوان مثال، کاربران اغلب در ظهر روزهای وسط هفته و همچنین در نیمه شب روزهای آخر هفته به رستوران‌ها می‌روند. این الگوهای روز هفته و آخر هفته، نشان‌دهنده تنظیمات زمان بازدید کاربران از مکان‌های جذاب می‌باشد. این قابلیت می‌تواند به توصیه مکان جذاب مناسب با توجه به زمانی که کاربر درخواست کرده منجر شود [۲، ۳، ۸، ۱۷، ۲۲-۲۴].

شکل ۱ نشان‌دهنده عوامل موثر بر توصیه مکان جذاب با استفاده از اطلاعات شبکه‌های اجتماعی مکان مبنا است.

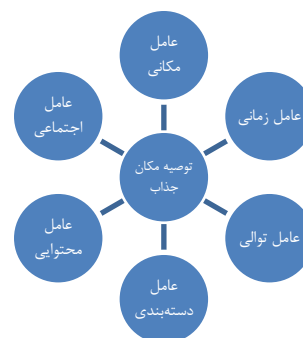
اخذ عوامل اجتماعی و از توزیع قانون توان^۵ برای اخذ عوامل مکانی استفاده کرده‌اند. مطالعه Cheng و همکاران [۴] نمونه‌ای از روش‌های مدل مبنا است که در آن Cheng و همکاران با استفاده از فاکتورگیری ماتریس احتمالاتی^۶، مدل فاکتور احتمالاتی^۷ را برای یادگیری ترجیحات کاربر ارائه کردند. علاوه بر این مدل، عوامل مکانی با مدل گاوسی چندمرکزی^۸ اخذ شد و در نهایت نتیجه این دو مدل برای ارائه‌ی توصیه با هم ترکیب شدند.

۳-۲- مدل اتصال

بر خلاف مدل‌های ترکیبی مدل‌های اتصال چندین عامل موثر بر توصیه‌ی مکان جذاب را با هم یاد گرفته و به توصیه مکان جذاب می‌پردازند [۲]. این مدل، رفتار بازدید کاربران را به عنوان تصمیمی حاصل از چندین فاکتور موثر به نمایش می‌گذارد که سناریو واقعی تری نسبت به مدل‌های ترکیبی است. به همین دلیل توجه بیشتری نسبت به این مدل‌ها شده است [۲].

مدل‌های اتصال ارائه شده را می‌توان در سه دسته طبقه‌بندی کرد [۲]:

- مدل اتصال بر مبنای فاکتورگیری ماتریسی^۹ [۱، ۲، ۴، ۵، ۷، ۹، ۱۲، ۲۵، ۲۸، ۲۹] که عواملی مثل عامل زمانی و عامل مکانی را در مدل فیلترینگ مشارکتی سنتی مثل فاکتورگیری ماتریس و فاکتورگیری تنسور^{۱۰} ادغام می‌کند. مدل ارائه شده توسط Lian و همکاران [۱۲] یک نمونه از این دسته است که عوامل مکانی را با مدل فاکتورگیری ماتریس ادغام کرده و به توصیه مکان جذاب می‌پردازد.
- مدل گرافیکی مولد^{۱۱} [۲] از مدل مولد به همراه مدل گرافیکی برای مدل‌سازی داده‌های بازدید و عوامل دیگر مثل عوامل مکانی استفاده می‌کند. مدل گرافیکی یک مدل احتمالاتی است که در آن ساختار وابستگی شرطی بین متغیرهای تصادفی با یک نمودار بیان می‌شود. مدل مولد یکی از رویکردهای اصلی در طبقه‌بندی آماری است که با داشتن متغیر قابل مشاهده



شکل ۱- عوامل موثر بر توصیه مکان جذاب

با توجه به تاثیر عوامل مکانی در توصیه مکان جذاب، در بخش‌های آتی تقسیم‌بندی سیستم‌های توصیه‌گر مکانی بیان می‌شود. سپس بر مبنای آن روش‌های مختلف مدل‌سازی عوامل مکانی بحث خواهد شد.

۳- تقسیم‌بندی سیستم‌های توصیه‌گر مکانی

بر مبنای تحلیل‌های انجام شده روی داده‌های بازدید کاربران در شبکه‌های اجتماعی مکان مبنا، محققان سیستم‌های توصیه‌گر مکانی فراوانی ارائه داده‌اند که می‌توان آنها را در دو دسته‌ی کلی مدل ترکیبی^۱ و مدل اتصال^۲ بررسی کرد [۴، ۶، ۱۱، ۲۵، ۲۶].

۳-۱- مدل ترکیبی

در این مدل برای هر عامل موثر در توصیه یک مدل مستقل ایجاد می‌شود و در نهایت توصیه‌های حاصل از آنها با توصیه‌های فیلترینگ مشارکتی که ترجیحات کاربران را در مکان‌های جذاب مختلف اخذ می‌کند، ترکیب می‌نماید [۲].

عامل مکانی و عامل زمانی با توجه به این که بیشترین تاثیر را در توصیه‌ها دارند در اکثر مدل‌ها به کار گرفته شده‌اند [۲].

استفاده از روش‌های فیلترینگ مشارکتی برای اخذ ترجیحات کاربر در دو دسته‌ی حافظه مبنا^۳ و مدل مبنا^۴ [۱، ۲] صورت می‌گیرد که تحقیقات مختلفی در هر دو دسته انجام شده است. به عنوان نمونه، مطالعه Ye و همکاران [۲۷] یک نمونه از روش‌های حافظه مبنا است که از فیلترینگ مشارکتی کاربر مبنا برای مدل کردن ترجیحات کاربر، از فیلترینگ مشارکتی دوست مبنا برای

^۵ Power law

^۶ Probabilistic Matrix Factorization (PMF)

^۷ Probabilistic Factor Models (PFM)

^۸ Multi-center Gaussian model

^۹ Matrix factorization

^{۱۰} Tensor factorization

^{۱۱} Generative graphical model

^۱ Fused model

^۲ Joint model

^۳ Memory-based

^۴ Model-based

و متغیر هدف، یک مدل آماری از توزیع احتمالاتی مشترک بین آنها ایجاد می‌کند. مدل ارائه شده توسط Liu و همکاران [۲۶] چهارچوب آنالیز فاکتور احتمالاتی جغرافیایی است که عوامل متنوعی مثل ترجیحات کاربر و عوامل مکانی و الگوی حرکتی کاربران را در برمی‌گیرد و از این دسته هست. فرایند تصمیم یک کاربر برای بازدید از یک مکان جذاب می‌تواند به صورت مولد مشاهده شود. بدین صورت که پارامترهای مختلف تاثیرگذار مثل محل کاربر، موقعیت مکان‌های جذاب و فاصله بین کاربر و مکان‌ها به صورت مولد مدل میشوند و رابطه نهایی منجر به توصیه مکان می‌شود.

- مدل شبکه‌ی عصبی^۱ عوامل موثر را به صورت اتصالی در شبکه عصبی مدل می‌کند. مدل ارائه شده توسط Liu و همکاران [۱۱] برای تخمین مکان جذاب بعدی از شبکه‌ی عصبی بازگشتی^۲ استفاده می‌کند. بیشتر مدل‌هایی که از شبکه‌ی عصبی استفاده کرده‌اند، ابتدا با استفاده از روش‌هایی مانند فاکتورگیری ماتریسی اطلاعات پنهان موجود درد داده‌ها را اخذ نموده و سپس این اطلاعات را به عنوان ورودی شبکه استفاده نموده‌اند. در مقاله [۳۰] Dimitrios و همکاران ابتدا از فاکتورگیری ماتریسی برای به دست آوردن عوامل پنهان مربوط به کاربران و داده‌ها استفاده می‌کنند. سپس، با داشتن ماتریس‌های پنهان کاربر و مکان‌های جذاب از یک شبکه‌ی عصبی با دو لایه‌ی پنهان گزینه‌هایی نهایی برای توصیه به کاربر را اخذ می‌کنند.

با توجه به تقسیم بندی‌های انجام شده در سیستم‌های توصیه‌گر مکانی و همچنین مقالاتی که از عوامل مکانی برای توصیه استفاده کرده‌اند، اخذ عوامل مکانی در سه دسته‌ی کلی بحث می‌شود [۲] که در بخش بعدی به جزئیات آن پرداخته می‌شود.

۴- اخذ عوامل مکانی در شبکه‌های اجتماعی مکان مبنا

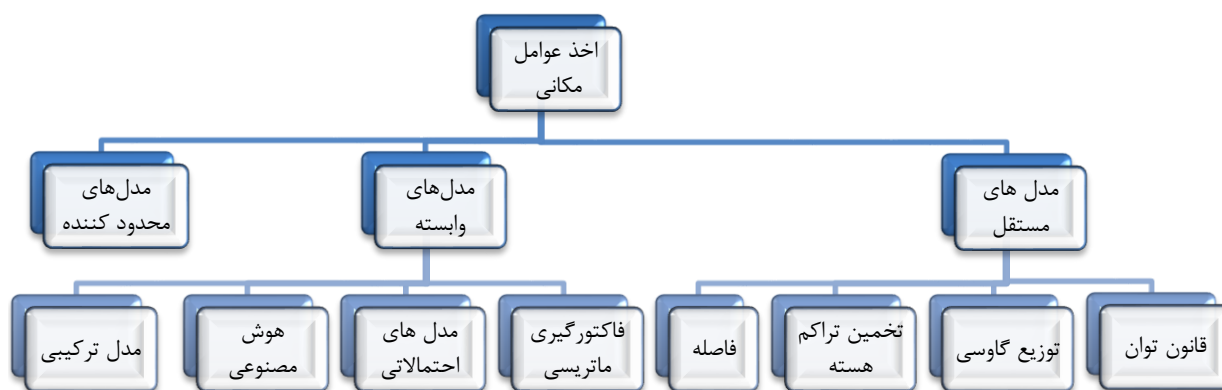
همانطور که در بخش‌های قبل اشاره شد، عوامل مکانی یکی از مهمترین عامل‌ها در توصیه مکان‌های جذاب به کاربران با استفاده از داده‌های شبکه‌های

اجتماعی مکان مبنا است. در طی سالیان گذشته مطالعات زیادی روی استفاده از عوامل مکانی برای بهبود سیستم‌های توصیه‌گر مکانی انجام شده و روش‌ها و رویکردهای مختلفی برای استفاده از این ویژگی به شرح زیر ارائه شده است [۲-۱۸، ۲۲-۲۹، ۳۱-۴۰].

- مدل‌های مستقل [۲] عوامل مکانی را به صورت کاملاً مستقل از دیگر عوامل و ترجیحات کاربران مدل‌سازی می‌کنند. اگرچه ممکن است برای توصیه‌ی مکان، با مدل‌های مستقل از عوامل دیگر ترکیب شوند. این حالت در سیستم‌های توصیه‌گر مدل ترکیبی قابل بحث است.
 - مدل‌های وابسته [۲] که در آنها فرمول یا رابطه‌ی خاصی برای اخذ عوامل مکانی وجود ندارد و این فاکتور در کنار عوامل دیگر در نظر گرفته می‌شود تا در نهایت یک رابطه یا فرمول ایجاد شود که بتواند همه این فاکتورها را در توصیه‌ی مکان به کار گیرد. در صورتی که تمام عوامل به صورت وابسته به هم مدل شوند، مدل اتصالی برای سیستم توصیه‌گر مکانی ارائه خواهد شد. اما توجه به این نکته مهم است که ممکن است عامل مکانی به همراه چند عامل به صورت وابسته به هم مدل شوند و با عوامل دیگر ترکیب گردند. در این صورت با وجود این که مدل اخذ داده مکانی ما وابسته هست اما مدل توصیه‌گر اتصالی نبوده و ترکیبی است.
 - مدل سوم به آن دسته از مقالاتی مربوط می‌شود که عامل مکانی را به عنوان یک عامل محدود کننده در نظر گرفته اند [۸]. به این صورت که به جای بررسی عوامل مکانی یک فاصله معین به عنوان حد آستانه‌ی توصیه برای کاربران در نظر گرفته می‌شود و توصیه‌های ارائه شده برای کاربران، فقط در داخل آن محدوده انجام می‌شود.
- در ادامه‌ی این بخش به بررسی سیستم‌های توصیه‌گر مکانی که عوامل مکانی را به صورت مستقل و وابسته مدل‌سازی کرده اند پرداخته می‌شود. مدل‌های محدود کننده با توجه به این که در اصل کارشان مدل‌سازی نبوده و فقط به عنوان عامل محدود کننده برای توصیه بهتر استفاده می‌شوند، مورد بررسی قرار نمی‌گیرد. شکل ۲ روش‌های ارائه شده برای مدل‌سازی عوامل مکانی به همراه زیر بخش‌های مربوطه را نشان می‌دهد.

^۱ Neural network

^۲ Recurrent Neural Network (RNN)



شکل ۲- اخذ عوامل مکانی در سیستم های توصیه گر مکانی با استفاده از شبکه های اجتماعی مکان مبنا

۴-۱- مدل های مستقل

همانطور که در قسمت قبلی اشاره شد، مدل های مستقل عوامل مکانی را مستقل از دیگر عوامل و ترجیح کاربر مدل سازی می کنند. چندین مدل مختلف مستقل در طی سالیان گذشته برای اخذ عوامل مکانی ارائه شده است که در ادامه به تشریح چهار مدل از مهمترین آنها می پردازیم.

۴-۱-۱- مدل قانون توان

محققان برای مدل کردن عوامل مکانی پس از تجزیه و تحلیل روابط بین مکان های جذاب بازدید شده دریافتند که فاصله بین هر جفت مکان جذاب بازدید شده از توزیع قانون توان پیروی می کند [۲۲، ۲۷].

قانون توان رابطه کاربردی بین دو کمیت است که در آن تغییر نسبی در یک کمیت باعث تغییر در کمیت دیگر میشود و این تغییر مستقل از مقادیر اولیه دو کمیت است. در این توزیع یک کمیت به صورت توانی از کمیت دیگر تغییر میکند [۲۷].

$$y = a * x^b \quad (1)$$

به عنوان مثال رابطه ی بین مساحت مربع در ارتباط با اندازه ضلع آن از قانون توان تبعیت می کند.

برای فهم بهتر تاثیر عوامل مکانی روی رفتارهای بازدید کاربران، Ye و همکاران [۲۷] آنالیزهای مکانی روی مجموعه داده های واقعی بازدید کاربران که از دو شبکه ی اجتماعی مکان مبنای معروف Whrrl و Foursquare گرفته شده بود، انجام دادند.

مشخصا هدف آنها این بود که اثر فاصله در رفتار بازدید کاربران را با اندازه گیری این که چند درصد از مکان های

بازدید شده توسط کاربر در یک محدوده ی مشخص قرار دارد، مطالعه کنند. بدین منظور آنها فاصله ی بین تمام جفت بازدیدهای کاربران را محاسبه و هیستوگرام آن را با توجه به فاصله ی بین مکان های جذابی که توسط کاربر اندازه گیری شده بود رسم کردند. شکل به دست آمده نشان داد که درصد قابل توجهی از جفت مکان های جذاب که توسط یک کاربر بازدید شده اند در فاصله ی کوتاهی از هم قرار گرفته اند که نشان دهنده مولفه ی خوشه بندی مکانی رفتار بازدید کاربران است. Ye و همکاران [۲۷] با توجه به نتیجه به دست آمده تصمیم گرفتند مولفه ی مکانی رفتار بازدید کاربران را برای توصیه مکان جذاب اخذ و استفاده کنند.

برای رسیدن به این هدف آنها ابتدا احتمال بازدید کاربر U_i از مکان های L_k ، L_j را محاسبه کرده به این نتیجه رسیدند که احتمال بازدید کاربر از توزیع قانون توان پیروی می کند. البته این نتیجه گیری کاملا درست هم نبود و توزیع در جاهایی غیرخطی بود. دلیل این اتفاق این است کاربران ممکن است به نقاط مختلفی سفر کنند. بنابراین خوشه های مکانی چندگانه ایجاد می کنند.

آنها محاسبه کردند که نسبت خطی نمودار به دست آمده شامل اکثریت جفت ها (۹۰ درصد) می شود. بنابراین به این نتیجه رسیدند که می توانند از توزیع قانون توان برای مدل کردن احتمال بازدید کاربران در فاصله ی بین دو مکان جذاب که توسط یک کاربر بازدید شده است، استفاده کنند. فرض کنید a و b پارامترهای توزیع، x فاصله بین دو مکان جذاب بازدید شده توسط یک کاربر و y احتمال بازدید است برای خطی سازی از طرفین رابطه ۱ لگاریتم گرفته می شود که رابطه ی نهایی به شکل زیر است:

$$\text{Log}y = w_0 + w_1 * \text{Log}x \quad (2)$$

$y' = a = 2^{w_0}$. $b = w_1$ اگر در نظر گرفته شود که $x' = \text{Log}x$ و $\text{Log}y$ آنگاه:

$$y'(x'.w) = w_0 + w_1 * x' \quad (3)$$

که در آن w_0 و w_1 ضرایب خطی هستند که در کل با w نمایش داده می‌شوند. وزن ضرایب با روش کمترین مربعات محاسبه می‌شود.

بر اساس مدل عامل مکانی که از طریق توزیع قانون توان نشان داده شده است، مکان‌های جدید با استفاده از رابطه‌ی زیر می‌توانند پیشنهاد داده شوند. اگر مجموعه مکان‌های بازدید شده L_i باشد، احتمال بازدید کاربر U_i از مکان L_j با فرمول زیر بدست می‌آید:

$$pr(l_j|L_i) = \frac{pr(l_j \cup L_i)}{pr(L_i)} = \prod_{l_y \in L_i} pr(d(l_j, l_y)) \quad (4)$$

که در آن $d(l_j, l_y)$ فاصله‌ی بین l_j و l_y و $pr(d(l_j, l_y)) = a * d(l_j, l_y)^b$ می‌باشد.

در تحقیق دیگری Yuan و همکاران [۲۲]، بازدیدهای هر کاربر را با زمان مرتب نموده و برای بازدیدهای ساخته شده در یک روز، فاصله‌ی بین دو مکان جذاب مجاور را محاسبه کردند. سپس آنها تعداد بازدیدها را به صورت تابعی از فاصله رسم نمودند. با توجه به شکل ۳ که داده‌های دو منطقه مختلف در آن به نمایش گذاشته شده است، آنها نتیجه گرفتند که توزیع مقادیر احتمالی از توزیع قانون توان پیروی می‌کند.

با فرض این که یک کاربر در حال حاضر در مکان جذاب L_i است و مکان جذاب L_j یک مکان جذاب نامزد برای بررسی در آینده است که در فاصله $dis(l_i, l_j)$ از L_i قرار دارد، Yuan و همکاران [۲۲] احتمال بازدید کاربر در L_j را نسبت به تمایل کاربر برای بازدید در مکان جذابی که در فاصله $dis(l_i, l_j)$ قرار دارد با هم مقایسه می‌کنند. احتمال شرطی با استفاده از رابطه‌ی زیر محاسبه می‌شود.

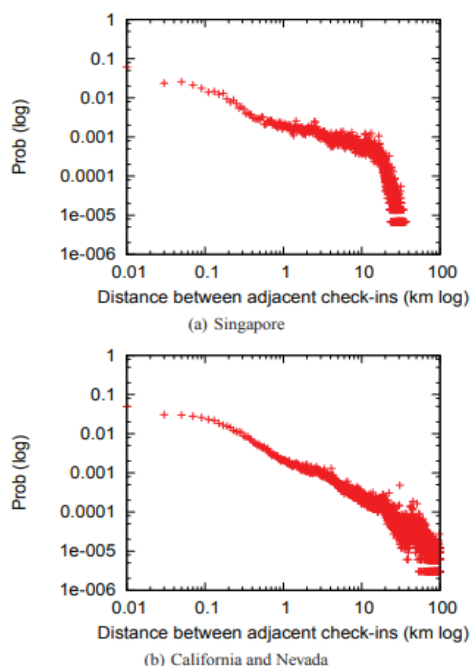
$$p(l_j|l_i) = \frac{wi(dis(l_i, l_j))}{\sum_{l_k \in L, l_k \neq l_i} wi(dis(l_i, l_k))} \quad 5$$

مشاهده می‌شود که با افزایش فاصله احتمال شرطی کاهش می‌یابد. اگر کاربر u مکان‌های جذاب تاریخی آن L_u داده شده باشد، $P(l|L_u)$ به عنوان امتیاز رتبه بندی برای هر مکان نامزد l محاسبه می‌شود. سپس مکان‌های

جذاب برتر را به کاربر توصیه می‌کنیم. بر اساس قانون بویس^۱ [۲۲]، این نمره به صورت زیر محاسبه می‌شود:

$$\hat{C}_{u,l}^{(s)} = p(l|l_u) \propto p(l)p(L_u|l) = p(l) \prod_{l' \in L_u} p(l'|l) \quad (6)$$

بنابراین تفاوت این مدل با مدل Ye و همکاران [۲۷] در استفاده از قانون بویس برای یادگیری پارامترها است [۲].



شکل ۳- نمودار احتمال توزیع بازدیدهای کاربران بر اساس فاصله بین بازدیدهای مجاور در سنگاپور و نوادا [۲۲]

۴-۱-۲- مدل‌های مبتنی بر توزیع گاوسی

علی رغم این که الگوهای حرکتی انسان‌ها دارای درجه آزادی و تنوع زیادی است اما این حرکت‌ها الگوهای ساختاری مشخصی را با توجه به قیدهای مکانی و اجتماعی به نمایش می‌گذارند [۴، ۶، ۴۰]. با توجه به این که اکثر اوقات فعالیت افراد در اطراف مکان‌های خاصی شکل می‌گیرد ایده مدل کردن عوامل مکانی بر مبنای توزیع گاوسی شکل گرفت.

توزیع گاوسی یا توزیع نرمال یکی از مهمترین توزیع‌های احتمال پیوسته در نظریه‌ی احتمالات است. تغییرات بسیاری از کمیت‌های طبیعی حول یک مقدار ثابت از این توزیع پیروی می‌کند. به عنوان مثال توزیع قد و وزن از این توزیع پیروی می‌کند. به زبان ساده در قضیه حد مرکزی نمایش داده می‌شود که مجموع متغیرهای تصادفی مستقل که هر

^۱ Boyce

$$P[x_u(t) = x_i | c_u(t)] = \begin{cases} \sim \mathcal{N}(\mu_H, \sum H) & \text{if } c_u(t) = H \\ \sim \mathcal{N}(\mu_W, \sum W) & \text{if } c_u(t) = W \end{cases} \quad (8)$$

Cheng و همکاران [۴] با مدل کردن احتمال بازدید کاربر از یک مکان جذاب، عوامل مکانی را با مدل چند مرکزی گاوسی اخذ کردند. این روش کمی متفاوت با روش دو حالت خانه و محل کار برای رفتار بازدید کاربران [۴۰] می باشد. در این روش علاوه بر مراکز خانه و محل کار که شامل بیش از پنجاه درصد بازدیدهای کاربران هست، مراکز دیگر حداقل ده درصد از بازدیدها را شامل می شود. این مکان ها می توانند مکان های مسافرت کاری معمول شخص باشند.

در این حالت احتمال بازدید کاربر u از مکان l با توجه به مجموعه مراکز C_u با رابطه زیر مشخص می شود:

$$p(l|C_u) = \sum_{C_u=1}^{|C_u|} p(l \in C_u) \frac{f_{C_u}^\alpha}{\sum_{l \in C_u} f_l^\alpha} \frac{\mathcal{N}(l | \mu_{C_u}, \Sigma_{C_u})}{\sum_{l \in C_u} \mathcal{N}(l | \mu_l, \Sigma_l)} \quad (9)$$

که در آن l بیانگر طول و عرض جغرافیایی موقعیت، C_u مجموعه مراکز برای کاربر u است. برای هر مرکز محاسبه رابطه بالا از ضرب سه قسمت تشکیل شده است:

رابطه $p(l \in C_u) \propto 1/\text{dist}(l, C_u)$ احتمال تعلق مکان l به مرکز C_u را مشخص می کند که با فاصله مکان l از مرکز C_u رابطه معکوس دارد. تاکید قسمت دوم بر روی اثر نرمال فرکانس بازدید که با f_{C_u} نمایش داده می شود روی مرکز C_u است. پارامتر $\alpha \in (0.1)$ ویژگی ناسازگاری فرکانس را نگهداری می کند تا فرکانس بالای بازدید اثر خیلی بزرگی نداشته باشد. قسمت سوم نشانگر احتمال نرمال تعلق مکان به مرکز C_u هست که در آن $(l | \mu_{C_u}, \Sigma_{C_u})$ تابع چگالی احتمال توزیع گاوسی، μ_{C_u} و Σ_{C_u} به ترتیب نشانگر میانگین و کوواریانس نواحی اطراف مرکز C_u هستند

با تعاریف داده شده مساله تبدیل به مساله یافتن مراکز شد. Cheng و همکاران [۴] به منظور یافتن مراکز از الگوریتم خوشه بندی حریص^۱ روی بازدیدها بر اساس قاعده پارتو^۲ استفاده نمودند. آنها اسکن کردن را از پربازدیدترین مکان جذاب شروع کردند و تمام دیگر مکان های بازدید شده، که فاصله ی آنها از مکان جذاب انتخاب شده کمتر از d کیلومتر

کدام میانگین و واریانس متناهی دارند، با افزایش تعداد متغیرها دارای توزیع بسیار نزدیک به توزیع گاوسی است.

Choe و همکاران [۴۰] مشاهده کردند که کاربران در شبکه های اجتماعی مکان مبنا حول مراکز فعالیتی خاصی مثل خانه و محل کار فعالیت می کنند. این مشاهده باعث ارائه ی روش هایی مبتنی بر توزیع گاوسی برای مدل کردن عامل مکانی شد.

آنها مشاهده کردند که کاربران معمولاً در طی روز در شعاع صد کیلومتری حرکت می کنند بنابراین برای تعیین موقعیت خانه ی کاربران که به صورت واضح مشخص نبود، آنها تصمیم گرفتند میانگین بازدیدهای کاربران را در منطقه ۲۵ کیلومتری که بازدیدها بیشترین تراکم را دارند به عنوان موقعیت خانه در نظر بگیرند. طبق بررسی های میدانی انجام شده، این موقعیت در ۸۵ درصد مواقع نزدیک به موقعیت خانه کاربران بود.

برای راحتی آنها دو مکان مرکزی شامل خانه و محل کار را برای مدل در نظر گرفتند. بسته به ساعت روز حرکات افراد با مرکزیت خانه یا محل کار و در جایی بین این دو خواهد بود. آنها برای هر کاربر مراکز مکانی در اینجا خانه و محل کار را مشخص کرده و با توزیع گاوسی آنها را مدل کردند. همچنین احتمال این که کاربر در خانه و یا محل کار هست را با توجه به تابع زمان روز مدل سازی نمودند.

فرض کنید t ساعت روزانه فعلی کاربر و $x_u(t)$ نشانگر موقعیت مکانی کاربر u در زمان t باشد. $C_u(t)$ نشانگر حالت مرکزیت است که اگر خانه باشد برابر با H و اگر محل کار باشد با W نمایش داده می شود. بنابراین موقعیت بازدید کاربران با توزیع مناسب مکانی کاربران اداره می شود. $P(x(t)=x|C_u(t))$ توزیع احتمال روی موقعیت کاربران در زمان t به صورت ساده ترکیبی از توزیع مکانی خانه و محل کار می باشد که ضرایب ترکیب با توجه به مدل زمانی تعیین می شود.

$$P[x(t) = x] = P[x_u(t) = x | c_u(t) = H]. P[c_u(t) = H] + P[x_u(t) = x | c_u(t) = W]. P[c_u(t) = W] \quad (7)$$

توجه داشته باشید که موقعیت حالت های اصلی نامعلوم هستند. Cho و همکاران [۴۰] برای راحتی مکان هایی که بیشترین بازدید را داشتند به عنوان مراکز انتخاب نمودند و در نهایت برای مدل کردن مولفه مکانی از توزیع گاوسی دو بعدی مستقل از زمان استفاده کردند:

^۱ Greedy clustering algorithm

^۲ Pareto

هست را برای منطقه در نظر گرفته و ترکیب کردند. نحوه یافتن مراکز بدین شکل بود که اگر نسبت تعداد بازدیدهای این منطقه به تعداد مجموع بازدیدهای کل کاربران از حد آستانه θ بیشتر بود، موقعیت‌های بازدیدها را یک منطقه در نظر گرفته و مرکز آن مشخص می‌شود.

با وجود ارائه‌ی روش‌هایی مبتنی بر توزیع گاوسی، Zhao و همکاران [۶] به این نتیجه رسیدند که هنوز روش‌های کشف مراکز فعالیت کاربران راضی‌کننده نیستند. آنها دو روش برای اخذ عوامل مکانی ارائه کردند: مدل ترکیبی گاوسی^۱ و مدل ترکیبی گاوسی بر مبنای الگوریتم ژنتیک^۲. به طور خاص هدف آنها استفاده از GMM برای یادگیری اتوماتیک مراکز فعالیت کاربران و استفاده از GA-GMM برای بهبود GMM به وسیله‌ی حذف داده‌های خارج از محدوده بود.

مدل ترکیبی گاوسی در واقع یک مدل خوشه‌بندی است که می‌تواند داده‌هایی که در چند خوشه مختلف هستند اما باهم در جاهایی کاملاً ترکیب شده اند را از هم جدا کند [۶]. در این حالت فرض می‌شود که خوشه‌ها از توزیع گاوسی پیروی می‌کند و سپس به جدا کردن آنها از هم می‌پردازند. می‌توان آن را به شکل زیر فرموله کرد:

$$p(x_i) = \sum_{k=1}^k \pi_k \mathcal{N}(x_i | \mu_k, \Sigma_k) \quad (10)$$

$p(x_i)$ نشان‌دهنده توزیع چگالی احتمال داده x_i و μ_k و Σ_k نشان‌دهنده مقدار میانگین و نشاندهنده ماتریس کوواریانس برای توزیع پایه و k نشان‌دهنده تعداد مولفه‌های پایه‌ای است. π_k ضریب ترکیب می‌باشد. هر مولفه توزیع گاوسی در GMM یک ناحیه فعالیت را نشان می‌دهد و مقدار میانگین نشان‌دهنده طول و عرض جغرافیایی ناحیه‌ی مرکز است. مراکز می‌توانند خانه، محل کار و یا مراکز تفریحی باشند. توصیه مکان جذاب با گام‌های زیر انجام می‌شود:

- یادگیری پارامترهای GMM
- محاسبه احتمال مکان‌های کاندید مناسب برای مدل آزمایشی
- مرتب کردن مکان‌های کاندید و توصیه k مکان بالا

مدل ترکیبی گاوسی بر مبنای الگوریتم ژنتیک به منظور حذف داده‌های خارج از محدوده ارائه شده است. به صورت عمومی از الگوریتم ماکزیمم سازی انتظارات^۳ برای یادگیری GMM استفاده می‌شود. اگر θ نشان‌دهنده پارامترها و x_i داده‌ها باشند. تابع احتمال به صورت زیر خواهد بود:

$$p(X|\theta)_{ML} = \prod_{i=1}^n p(x_i|\theta) \quad (11)$$

علاوه بر این اگر فرم لگاریتم استفاده شود، می‌توان هدف الگوریتم EM را به صورت زیر داشت:

$$\begin{aligned} \hat{\theta}_{ML} &= \operatorname{argmax}_{\theta} \log p(X|\theta)_{ML} \\ &= \operatorname{argmax}_{\theta} \sum_{i=1}^n \log p(x_i|\theta) \end{aligned} \quad (12)$$

این رابطه شامل تمام داده‌های آزمایشی است. تخمین احتمالات درستی^۴ هدفش انتخاب زیرمجموعه‌ای از داده‌ها با جمع ماکزیمم مقادیر احتمال است که برای حذف داده‌های خارج از محدوده به کار می‌رود. می‌توان از الگوریتم ژنتیک برای پیدا کردن زیرمجموعه‌ی بهینه استفاده کرده و الگوریتم EM را برای یادگیری پارامترهای GMM به کار گرفت. در این حالت تابع هدف به شکل زیر خواهد بود:

$$\log p_{TLE}(X|\theta) = \sum_{i=1}^n \omega_i \log p(x_i|\theta) \quad (13)$$

که به ازای $i = 1, 2, 3, \dots, n$ و ω_i عضو مجموعه صفر و یک و $\sum_{i=1}^n \omega_i = m$ که m تعداد داده‌های قابل قبول را نشان می‌دهد. رابطه $\omega_i = m$ نشان می‌دهد که داده‌ی متناظر آن برای زیرمجموعه انتخاب شده است. در غیر این صورت داده خارج از محدوده است و باید حذف شود.

GA-GMM نیز ویژگی‌های الگوریتم ژنتیک را دارد و شامل کدگذاری، تابع سازگاری و عملگرهای تولید مثل^۵، جهش^۶ و انتخاب^۷ هست. راه استاندارد برای پیاده سازی تولید مثل و انتخاب را Zhao و همکاران [۶] ارائه کردند. کدگذاری، تابع سازگاری و جهش راهنمایی شده به شکل زیر تعریف می‌شوند:

۱. کروموزوم با رشته باینری کدگذاری می‌شود و هر بیت وجود داده مشاهداتی متناظر را نشان می‌دهد. هر

^۳ Expectation Maximization (EM)

^۴ Truth Likelihood Estimation (TLE)

^۵ Crossover

^۶ Mutation

^۷ Selection

^۱ Gaussian Mixture Model (GMM)

^۲ Genetic Algorithm – Gaussian Mixture Model (GA-GMM)

است که می‌تواند به کاهش اطلاعات زیاد، کمک کند و یکی از عوامل مهم موفقیت تجارت الکترونیک است.

برای مشاهده رفتارهای منحصر به فرد کاربران، آنها یک تجزیه و تحلیل مکانی بر روی دو مجموعه داده‌ی واقعی موجود که از Gowalla و Foursquare جمع‌آوری شده بود، برای سه کاربر انجام دادند. در هر دو مجموعه داده مشاهده شد که تاثیر مکانی مکان‌های جذاب بر رفتارهای بازدید سه کاربر منحصر به فرد است.

بنابراین آنها یک روش جدید برای شخصی‌سازی مدل سازی عامل مکانی مکان‌های جذاب به عنوان توزیع فاصله‌ی فردی برای هر کاربر به جای توزیع مشترک برای همه‌ی کاربران طراحی کردند.

به طور مشخص، آنها توزیع شخصی از فاصله‌ی هر جفت مکان بازدید شده توسط کاربر را با استفاده از تخمین تراکم هسته مدل می‌کنند، زیرا می‌توان آن را با توزیع‌های دلخواه مورد استفاده قرار داد بدون این فرض که شکل توزیع فاصله، شناخته شده باشد. فرایند تخمین چگالی هسته شامل دو مرحله جمع‌آوری نمونه فاصله و تخمین توزیع فاصله است.

مرحله ۱: جمع‌آوری نمونه فاصله. می‌توان یک نمونه را برای یک کاربر با محاسبه فاصله‌ی بین هر جفت مکان که توسط کاربر بازدید شده است، به دست آورد، زیرا هر بازدید در شبکه‌های اجتماعی مکان مبنا با هویت و موقعیت کاربر آن (به عنوان مثال، مختصات طول و عرض جغرافیایی) مرتبط است. به طور خاص، برای یک کاربر با مشکل شروع سرد (با تنها یک بازدید)، به جای فاصله دو مکان بازدیدی، فاصله بین تنها مکان بازدید شده و محل اقامت شخص به عنوان نمونه استفاده می‌شود.

مرحله ۲: تخمین توزیع فاصله. D یک نمونه فاصله برای یک کاربر خاص است که از یک توزیع با تراکم نامعلوم f گرفته شده است. برآوردگر تراکم هسته $\hat{f}$ برای فاصله d با استفاده از D توسط رابطه زیر بدست می‌آید:

$$\hat{f}(d) = \frac{1}{|D|h} \sum_{d' \in D} K\left(\frac{d-d'}{h}\right) \quad (15)$$

در این رابطه $K(\cdot)$ برابر با تابع هسته و h پارامتر نرم کننده یا عرض دامنه است. Zhang و همکاران [۱۰] از محبوبترین تابع هسته که هسته نرمال است، استفاده می‌شود:

کروموزوم و مدل ترکیبی متناظر آن راه حل ممکن برای مساله خواهد بود.

۲. تابع امتیازدهی سازگاری مجموعه ای است مثل لگاریتم تقسیم احتمال GMM متناظر با کروموزوم یعنی

$$\log_{TLE}(X|\theta) \quad (14)$$

جهش راهنمایی شده به کروموزوم در جمعیت اطمینان می‌دهد که به سمت ماکزیمم شدن امتیاز سازگاری حرکت کند. این بدین معنی است که کروموزومی انتخاب می‌شود که دارای مقدار بالایی در سازگاری با GMM آموزش داده شده باشد.

۴-۱-۳- مدل‌های مبتنی بر تخمین تراکم هسته

مطالعات قبلی انجام شده در مدل‌های توزیع قانون توان و گاوسی، عوامل مکانی را با استفاده از تمامی رفتارهای بازدید کاربران و در یک روش جهانی مدل می‌کنند. در حالی که اثر مکانی برای کاربران باید شخصی باشد [۱۰]. توجه به این نکته منجر به ایجاد روش‌هایی بر مبنای تخمین تراکم هسته^۱ شد.

در آمار، تخمین تراکم هسته روش غیر پارامتریک برای تخمین تابع چگالی احتمال یک متغیر تصادفی است. برآورد تراکم هسته یک مسئله‌ی نرم کردن داده‌های اساسی است که در آن، بر اساس یک نمونه داده‌ی محدود، نتیجه‌گیری‌های مربوط به جمعیت صورت می‌گیرد [۱۰]. Zhang و همکاران [۱۰] به این نتیجه رسیدند که تاثیر مکانی در رفتارهای بازدید کاربران باید منحصر به فرد و شخصی باشد. به عنوان مثال، افرادی که تمایلی به جابجایی زیاد ندارند^۲، بازدید از مکان‌های جذاب در اطراف مناطق زندگی خود را ترجیح می‌دهند در حالی که افرادی که بیشتر دوست دارند از مناطق مختلف بازدید کنند^۳ ترجیح می‌دهند در سراسر جهان برای کشف مکان‌های جذاب جدید سفر کنند. بنابراین، آنها استدلال کردند که تاثیر جغرافیایی مکان‌ها برای رفتارهای فردی بازدید کاربران باید در هنگام توصیه‌ی مکان‌ها برای کاربران شخصی سازی شود. در واقع، شخصی‌سازی یکی از ضروری‌ترین نیازها

^۱ Kernel Density Estimation (KDE)

^۲ indoorsy

^۳ outdoorsy

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (۱۶)$$

و بهینه‌ترین عرض دامنه برابر است با:

$$h = \left(\frac{4\delta^5}{3n} \right)^{\frac{1}{5}} \approx 1.06 \delta n^{\frac{-1}{5}} \quad (۱۷)$$

در این رابطه δ برابر با انحراف معیار نمونه D و n تعداد جفت مکان‌های بازدید شده توسط کاربر است.

پس از پیدا کردن توزیع فاصله براساس تخمین چگالی هسته، یک روش برای ارزیابی احتمال بازدید کاربر U_i از یک مکان جدید L_j ، با توجه به مجموعه مکان‌های بازدید شده توسط این کاربر یعنی:

$$L_i = \{l_1, l_2, \dots, l_n\} \quad (۱۸)$$

طراحی می‌شود. برای این منظور ابتدا، فاصله هر جفت مکان در L_i و L_j را به صورت زیر محاسبه می‌شود:

$$d_{ij} = \text{distance}(l_i, l_j), \forall l_i \in L_i \quad (۱۹)$$

سپس هر d_{ij} برای استخراج احتمال در رابطه (۱۵) قرارداده می‌شود. در نهایت احتمال بازدید کاربر u_i از مکان جدید l_j را می‌توان با احتمال میانگین به دست آورد:

$$p(l_j | L_i) = \frac{1}{n} \sum_{i=1}^n f(d_{ij}) \quad (۲۰)$$

در نهایت، تاثیر مکانی شخصی مکان‌های جذاب برای پیشنهاد مکان برای کاربر U_i با بازگرداندن مکان‌های $top - l_j$ با بیشترین احتمال $p(l_j, L_i)$ اعمال می‌شود. با وجود این، باید آن را با سایر اطلاعات از جمله اولویت کاربر، دوستی اجتماعی و اطلاعات کاربران برای به دست آوردن کیفیت بهتر توصیه‌ی محل، ادغام کرد. برای این منظور، یک چارچوب یکپارچه که تمام این عوامل را با هم ادغام کند، ارائه شده است.

Zhang و همکاران [۸] به منظور اخذ عوامل مکانی به این نتیجه رسیدند که به جای استفاده از فاصله با توزیع یک بعدی، توزیع احتمال بازدید کاربران را به صورت دو بعدی مدل کنند.

به طور خاص، آنها یک توزیع احتمال دو بعدی در راستای مختصات طول و عرض جغرافیایی برای هر کاربر به

جای تخمین توزیع احتمال فاصله یک بعدی که در کارهای قبلی موجود هست، مدل می‌کنند. توزیع احتمال بازدید دو بعدی منطقی‌تر و شهودی‌تر از توزیع فاصله یک بعدی است. دلیل اصلی این است که، محاسبه دقیق احتمال بازدید از یک مکان جدید براساس توزیع احتمال فاصله سخت است، زیرا پیدا کردن یک مکان مرجع برای ایجاد فاصله مناسب برای محل جدید نسبت به محل مرجع در اولین مکان سخت است. برعکس، به طور قابل ملاحظه‌ای استفاده از توزیع احتمال دو بعدی برای محاسبه احتمال بازدید از هر مکان با مختصات طول و عرض جغرافیایی منطقی است. S_u مجموعه مکان‌های بازدید شده توسط کاربر u است.

$$S_u = \{l_1, l_2, \dots, l_n\} \quad (۲۱)$$

در تخمین چگالی هسته، فرم توزیع ثابت به صورت پیش فرض مورد قبول واقع نمی‌شود اما به جای آن فرم‌های توزیع از داده‌های تاریخی یاد گرفته می‌شود. با استفاده از آن احتمال جغرافیایی کاربر u برای مکان جدید l_{n+1} برابر با $(p^{geo}(l_{n+1} | S_u))$ است که به صورت زیر محاسبه می‌شود:

$$p^{geo}(l_{n+1} | S_u) = \frac{1}{n\sigma^2} \sum_{i=1}^n k\left(\frac{l_{n+1} - l_i}{\sigma}\right) \quad (۲۲)$$

جایی که هر مکان $l_i = (lat_i, lon_i)^T$ بردار ستونی دو بعدی با مختصات طول و عرض جغرافیایی داده شده و $K(.)$ تابع هسته و σ پارامتر نرم کننده یا عرض دامنه است. Zhang و همکاران از هسته عادی دو بعدی استاندارد استفاده کرده‌اند:

$$K(x) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}x^T x\right) \quad (۲۳)$$

همچنین عرض بهینه باند به صورت زیر تعریف می‌شود:

$$\sigma = n^{\frac{-1}{6}} \sqrt{\frac{1}{2} \hat{\sigma}^T \hat{\sigma}} \quad (۲۴)$$

که در آن $\hat{\sigma}$ بردار انحراف معیار حاشیه‌ای طول و عرض جغرافیایی در S_u هست که با رابطه زیر مشخص می‌شود:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n l_i, \hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n (l_i - \hat{\mu})^2} \quad (۲۵)$$

با توجه به این واقعیت که توزیع بازدید کاربران از یکدیگر متفاوت است، Zhang و همکاران [۸] فرم توزیع را

به جای استفاده مستقیم از تخمین پایلوت $\hat{f}_{Geo}(l|u)$ برای پیش‌بینی میزان امتیاز ارتباط کاربر u به l ، تخمین پایلوت برای تعیین پهنای باند محلی برای هر مکان بازدید l_i با رابطه‌ی زیر محاسبه می‌شود:

$$h_i = (g^{-1} \cdot \hat{f}_{Geo}(l|u))^{-\alpha} \quad (29)$$

در این رابطه α پارامتر حساسیت بین صفر و یک هست که هر چه این مقدار بیشتر باشد عرض دامنه محلی h_i نسبت به تخمین پایلوت $\hat{f}_{Geo}(l|u)$ حساس‌تر می‌شود و g میانگین هندسی است:

$$g = \sqrt[n]{\prod_{i=1}^n \hat{f}_{Geo}(l|u)} \quad (30)$$

که محدودیتی را به وجود می‌آورد که میانگین هندسی آن $i = 1, 2, \dots$ برابر با یک است. مرحله سوم: ارزیابی کرنل مناسب برای امتیاز ارتباط جغرافیایی.

در نهایت، با پهنای باند جهانی $H = (H_1 \text{ و } H_2)$ روابط قبلی و پهنای باند محلی سازگار h_i ، تخمین انطباق هسته $\hat{f}_{Geo}(l|u)$ توزیع بازدیدهای کاربر u در مکان بازدید نشده l از رابطه‌ی زیر محاسبه می‌شود:

$$\hat{f}_{Geo}(l|u) = \frac{1}{N} \sum_{i=1}^n (R_{u,li} \cdot k_{Hh_i}(l - l_i)) \quad (31)$$

که در آن:

$$k_{Hh_i}(l - l_i) = \frac{1}{2\pi H_1 H_2 h_i^2} \exp\left(-\frac{(x - x_i)^2}{2H_1^2 h_i^2} - \frac{(y - y_i)^2}{2H_2^2 h_i^2}\right) \quad (32)$$

مهم است در نظر داشته باشید که وقتی مکان بازدید l_i در یک منطقه با چگالی بازدید بالا قرار می‌گیرد تخمین پایلوت $\hat{f}_{Geo}(l|u)$ در رابطه ۲۵ مقدار بالایی می‌گیرد و یا عرض دامنه محلی h_i کمتری در رابطه ۲۹ می‌گیرد که یک قله در تخمین هسته انطباقی $\hat{f}_{Geo}(l|u)$ حول l_i در رابطه ۳۱ ایجاد می‌کند. در مقابل وقتی مکان بازدید l_i در یک منطقه با چگالی پایین قرار می‌گیرد تخمین پایلوت $\hat{f}_{Geo}(l|u)$ در رابطه ۲۵ مقدار پایینی می‌گیرد و یا عرض دامنه محلی h_i مقدار بالایی در رابطه ۲۹ می‌گیرد که نرمی در تخمین هسته انطباقی $\hat{f}_{Geo}(l|u)$ حول l_i در رابطه ۳۱ ایجاد می‌کند. بنابراین، روش برآورد هسته سازگار ارائه شده

از مکان‌های جذاب بازدید شده کاربر بر اساس روش برآورد غیرپارامتری، یعنی تخمینی تراکم هسته با پهنای باند ثابت، اخذ می‌کند. با وجود این، پهنای باند ثابت، واقعیات را در داده‌های بازدید کاربر نشان نمی‌دهد: مناطق شهری متراکم چگالی بالای بازدیدها را تجربه می‌کنند و مناطق روستایی دارای مناطق کم جمعیت، چگالی بازدید کمتری را تجربه می‌کنند. برای این منظور، Zhang و همکاران [۹] پهنای باند هسته را با هر نقطه داده بازدید، تطبیق می‌دهد و خود پهنای باند سازگار نیز از داده‌های بازدید پایه آموخته می‌شود. به طور کلی، روش برآورد هسته سازگار شامل سه مرحله است: ارزیابی پایلوت، تعیین پهنای باند محلی و ارزیابی کرنل مناسب برای امتیاز ارتباط جغرافیایی.

مرحله اول: ارزیابی پایلوت

ابتدا، برآورد پایلوت بر اساس تخمینی تراکم هسته با پهنای باند ثابت انجام می‌شود. $L_u = \{l_1, l_2, \dots, l_n\}$ مجموعه مکان‌های بازدید شده توسط کاربر u است که هرکدام از آنها l_i متناظر با مختصات جغرافیایی (x_i, y_i) است. به طور خاص فرکانس یا امتیاز کاربر u در مکان l_i را که با $R_{u,li}$ نشان داده می‌شود، به عنوان وزن l_i استفاده می‌کنیم، چرا که فرکانس یا امتیاز بالا در یک مکان نشان می‌دهد که آن مکان برای کاربر مهمتر است. تخمین و ارزیابی پایلوت $\hat{f}_{Geo}(l|u)$ برای توزیع بازدیدهای کاربر u در مکان بازدید نشده l با رابطه زیر حساب می‌شود:

$$\hat{f}_{Geo}(l|u) = \frac{1}{N} \sum_{i=1}^n (R_{u,li} \cdot k_H(l - l_i)) \quad N = \sum_{i=1}^n R_{u,li} \quad (25)$$

$$k_H(l - l_i) = \frac{1}{2H_1 H_2} \exp\left(-\frac{(x - x_i)^2}{2H_1^2} - \frac{(y - y_i)^2}{2H_2^2}\right) \quad (26)$$

در روابط بالا $k_H(l - l_i)$ تابع هسته‌ی نرمال با عرض دامنه‌ی ثابت H است که از دو عرض دامنه‌ی عمومی H_1 و H_2 برای طول و عرض جغرافیایی تشکیل شده و به صورت زیر محاسبه می‌شوند:

$$H_1 = 1.06n^{-\frac{1}{5}} \sqrt{\frac{1}{N} \sum_{i=1}^n \left(R_{u,li} \cdot x_i - \frac{1}{N} \sum_{j=1}^n R_{u,lj} \cdot x_j \right)^2} \quad (27)$$

$$H_2 = 1.06n^{-\frac{1}{5}} \sqrt{\frac{1}{N} \sum_{i=1}^n \left(R_{u,li} \cdot y_i - \frac{1}{N} \sum_{j=1}^n R_{u,lj} \cdot y_j \right)^2} \quad (28)$$

در این روابط H_1 و H_2 به ترتیب انحراف معیار مقادیر طول و عرض جغرافیایی داده‌های بازدید شده کاربر u می‌باشند.

مرحله دوم: تعیین پهنای باند محلی

در این مقاله می‌تواند پیش‌بینی توزیع برآورد شده بازدیدها را برای امتیاز ارتباط جغرافیایی کاربر u بر روی مکان بازدید نشده l_i بهبود دهد.

۴-۱-۴ مدل بر مبنای فاصله

Zhang و همکاران [۷] عوامل اجتماعی، مکانی و تگ‌ها را به صورت مستقل مدل‌سازی کردند تا در نهایت با ترکیب آنها در چهارچوب فاکتورگیری ماتریسی به توصیه مکان بپردازند.

برای مدل‌سازی عوامل مکانی آنها عنوان کردند که در شبکه‌های اجتماعی مکان مبنای فضای مکانی می‌تواند به چندین منطقه بر اساس طول و عرض جغرافیایی مکان‌های جذاب تقسیم شود. معمولاً، هر منطقه دارای یک مرکز منحصر به فرد است و کاربران تمایل دارند تا اطراف مرکز را بازدید کنند. با فرض این که کاربر i در اطراف ناحیه k است، هدف برآورد امتیاز کاربر در مکان جذاب j جایی که ضرورتاً در ناحیه k نیست، می‌باشد [۷]. زمانی که کاربر i از مکان جذاب j به صورت واقعی بازدید کرده است، تأثیرات مکانی بین i و j با فرمول فاصله‌های نرمال شده زیر تعریف می‌شود:

$$g(i, j) = \begin{cases} 1 & i, j \text{ in same region} \\ 1 + \frac{\text{dis}(i, j) + \text{dis}(c(i), c(j))}{2 \cdot \min} & \text{otherwise} \end{cases} \quad (33)$$

جایی که $c(i)$ و $c(j)$ هر دو مرکز منطقه‌ی کاربر i و مکان جذاب j هستند $\text{dis}(i, j)$ مربوط به فاصله‌ی بین کاربر i و مکان جذاب j است و $\min$ معادل حداقل فاصله بین مراکز مختلف منطقه می‌باشد.

۴-۲ مدل‌های وابسته

مدل‌های وابسته مدل‌هایی هستند که در آنها عوامل مکانی به صورت مستقل مدل‌سازی نشده است. در این حالت این عامل به همراه دیگر پارامترهای تأثیرگذار مثل عوامل اجتماعی، ترجیحات کاربر و عوامل دسته‌بندی مدل‌سازی می‌شود. به نوعی می‌توان گفت در این مدل‌ها روش استفاده شده به صورت اتصالی عمل می‌کند. البته همیشه به این صورت نیست که مدل‌های وابسته همه پارامترها را با هم به صورت اتصالی حل کنند. ممکن است روش ارائه شده یکی از عوامل را به همراه عامل مکانی با مدل اتصالی حل نماید و در کنار دیگر عوامل به صورت ترکیبی حل گردد. به منظور آسانی کار، هر یک از

تکنیک‌های توصیه‌ای که در آن عوامل مختلف باهم محاسبه شده اند به همراه چند نمونه از این سیستم‌ها که از عوامل مکانی نیز استفاده کرده‌اند را در ادامه می‌آوریم.

۴-۲-۱ مدل‌های فاکتورگیری ماتریسی

فاکتورگیری ماتریسی، ماتریس بازدید $C \in R^{M \times N}$ را به دو ماتریس کاربر $U \in R^{M \times K}$ و ماتریس مکان جذاب $L \in R^{N \times K}$ تجزیه می‌کند که در آن M ، N و K به ترتیب برابر با تعداد کاربران، مکان‌های جذاب و عوامل پنهان است [۳۵]. ویژگی‌های پنهان هر کاربر i و مکان جذاب j به ترتیب با u_i و l_j نمایش داده می‌شود. بنابراین امتیاز توصیه‌ای کاربر i به مکان جذاب j به صورت ضرب داخلی مدل‌سازی می‌شود.

$$\hat{C}_{ij} = u_i l_j^T \quad (34)$$

و تابع هدف به صورت زیر فرموله می‌شود.

$$\min_{U, L} \|C - UL^T\|_F^2 + \lambda_1 \|U\|_F^2 + \lambda_2 \|L\|_F^2, \quad (35)$$

در رابطه بالا $\|\cdot\|_F$ نرم فوربینیوس ماتریس و λ_1 و λ_2 پارامترهای تنظیم‌کننده هستند. نرم فوربینیوس یک ماتریس، جذر مجموع توان دوم تمام درایه‌های ماتریس است. با استفاده از این تکنیک مدل‌های زیادی برای توصیه مکان جذاب ارائه شده است.

مدل [۳۴] توسط Liu و همکاران براساس فاکتورگیری ماتریسی وزندار ارائه شده است. مبنای ارائه این روش دو مورد است. اول این که در مکان‌های جذاب همسایگی، کاربر ترجیحات مشابه دارد (تأثیر سطح مکان). دوم این که مکان‌های جذاب در یک منطقه‌ی جغرافیایی ترجیحات مشابه کاربر را به اشتراک می‌گذارند (تأثیر سطح منطقه). این به این معنی است که مکان‌هایی که در یک منطقه‌ی جغرافیایی واقع شده اند کاربرانی را به خود جذب می‌کنند که دارای ترجیحات مشابهی هستند.

برای مدل‌سازی تأثیر سطح مکان، نمره توصیه‌ای کاربر i در مکان جذاب j بیشتر شامل اثرات مکان‌های جذاب در همسایگی $N(l_j)$ می‌شود. این مدل نمره توصیه‌ای را با رابطه زیر مشخص می‌کند:

$$\hat{C}_{ij} = \alpha u_i l_j^T + (1 - \alpha) \frac{1}{z(l_j)} \sum_{l_k \in N(l_j)} \text{Sim}(l_j, l_k) U_i l_k^T \quad (36)$$

مکان جذاب z به وسیله $X_i Y_j^T$ تخمین زده می‌شود. بنابراین نمره توصیه ای نهایی برابر است با:

$$\hat{C}_{ij} = u_i l_j^T + X_i Y_j^T \quad (39)$$

مدل ارائه شده توسط L_i و همکاران یک مدل فاکتورگیری ماتریس بر مبنای رتبه بندی هست [۲۹] که اولاً رتبه‌بندی ترجیحات کاربران در مکان‌های جذاب را یاد می‌گیرد. ثانیاً شامل تاثیرات جغرافیایی مکان‌های جذاب در همسایگی می‌شود.

این مدل علاوه بر ماتریس ترجیحات کاربر $U^{(1)}$ ماتریس پنهان دیگری برای نمایش ترجیحات مکانی کاربران استفاده می‌کند $U^{(2)}$. بنابراین رتبه توصیه ای با رابطه زیر محاسبه می‌شود.

$$\hat{C}_{ij} = U_i^{(1)} l_j^T + U_i^{(2)} \cdot \sum_{l_k \in N(l_j)} \omega_{jk} \cdot l_k^T \quad (40)$$

قسمت اول نمره ترجیح کاربر را مدل‌سازی کرده و قسمت دوم نمره تاثیر مکانی را مدل می‌کند که کاربر به دلیل مکان‌های جذاب در همسایگی یک مکان جذاب دیگر، از آن مکان جذاب خوشش می‌آید. $N(l_j)$ به مکان‌های جذاب در همسایگی j اشاره می‌کند و ω_{jk} وزن فاصله مبنا هست که به k مکان جذاب نسبت داده می‌شود.

مدل [۱۴] که توسط L_i و همکاران ارائه شده یک چهارچوب توصیه‌ی مکان جذاب دو مرحله‌ای است. در مرحله اول مکان‌های جذابی که پتانسیل انتخاب شدن دارند را با توجه به دوستان کاربر یاد می‌گیرد. در مرحله دوم برای غلبه بر مشکل شروع سرد مکان‌های جذاب پتانسیل دار را با روش فاکتورگیری ماتریس وزندار^۳ ترکیب می‌کند. عوامل اجتماعی، دسته‌بندی و مکانی در این مدل استفاده شده‌اند.

امتیاز مکانی فاصله مبنا p_{ij}^G به منظور اعمال اثر عوامل مکانی با نتیجه روش فاکتورگیری ماتریس وزندار ترکیب می‌شود تا امتیاز نهایی به دست آید که p_{ij}^G با استفاده از توزیع فاصله بین خانه و مکان‌های بازدید شده کاربران محاسبه می‌شود.

که در آن $\alpha \in [0.1]$ برای کنترل اثر مکان‌های جذاب همسایگی استفاده می‌شود و $\text{Sim}(l_j, l_k)$ وزن فاصله مبنا بین l_k و l_j است $z(l_j)$ ترم نرمالیزاسیون هست. برای مدل‌سازی تاثیر در سطح منطقه، آنها ابتدا تمامی مکان‌های جذاب را به خوشه‌هایی در G منطقه براساس موقعیت جغرافیایی آنها تقسیم می‌کند. با فرض این که عوامل پنهان در مکان‌های جذاب یک منطقه الگوی پراکندگی مشترکی به اشتراک می‌گذارد آنها یک خطای لاسو^۱ روی تابع هدف به صورت زیر اضافه می‌کنند.

$$\sum_{g=1}^G \sum_{k=1}^k \omega_g \|L_{(g)}^k\|_2 \quad (37)$$

که در آن $L_{(g)}^k$ شامل k تا از عوامل پنهان از تمامی مکان‌های جذاب که در منطقه g هستند و ω_g وزن اختصاصی داده شده به $L_{(g)}^k$ است.

مدل ارائه شده توسط Lian و همکاران یک مدل جغرافیایی دیگر از فاکتورگیری ماتریس وزن‌دار است [۱۲]. این مدل که با نام فاکتورگیری ماتریسی جغرافیایی^۲ معرفی شده است، به منظور اخذ مولفه خوشه‌بندی مکانی (به عنوان نمونه مکان‌های جذابی که توسط کاربران مشابه بازدید شده اند ممکن است در یک منطقه باشند) با مدل‌سازی مناطق فعالیت کاربران و انتشار تاثیر مکانی آنها در فضای جغرافیایی، عامل مکانی را ادغام می‌کند.

GeoMF تمام فضای جغرافیایی را به R شبکه تقسیم می‌کند که هر کدام از آنها یک منطقه جغرافیایی را نشان می‌دهند. برای هر مکان جذاب اثر آن به مناطق اطراف پخش می‌شود تا کاربران نزدیک را برای بازدید جذب کند. به طور خاص دو ماتریس ناحیه فعالیت کاربر $X \in R^{M \times R}$ و ناحیه نفوذ مکان جذاب $Y \in R^{N \times R}$ معرفی شده است. رابطه ۳۸ نشان‌دهنده تاثیر مکان جذاب z در منطقه r است.

$$Y_{jr} = \frac{1}{\sigma} k \left(\frac{d(r, j)}{\sigma} \right) \quad (38)$$

در رابطه بالا $d(r, j)$ فاصله بین مکان جذاب j و منطقه r ، $k(\cdot)$ توزیع نرمال استاندارد و σ انحراف معیار است. X_{ir} نشان‌دهنده این است که چقدر احتمال دارد کاربر i در منطقه r ظاهر شود. ترجیح جغرافیایی کاربر i بر روی

^۳ Weighted Matrix Factorization (WMF)

^۱ lasso penalty

^۲ Geographical Matrix Factorization (GeoMF)

۴-۲-۲- مدل‌های احتمالاتی بر مبنای روش فاکتور پواسون

مدل فاکتور پواسون یک مدل احتمالاتی است که ماتریس بازدید کاربر-مکان جذاب C را به صورت

$$C \sim \text{poisson}(UL^T) \quad (41)$$

تجزیه می‌کند [۳۵]. در ادامه مدل‌هایی که بر مبنای این متدولوژی هستند و از عوامل مکانی نیز استفاده کرده‌اند را معرفی می‌کنیم.

ایده Liu و همکاران که با نام مدل جغرافیایی فاکتور پواسون^۱ معرفی شده، این است که اولویت مکانی و ترجیح کاربر از یکدیگر متاثر می‌شوند و ترجیح کاربر مربوط به هر دو آنها است [۳۶]. از این رو، آنها عوامل مکانی و ترجیح کاربر را با هم و به طور همزمان برای کاربران یاد می‌گیرد. روش استفاده شده در این مدل برای توصیه‌ی مکان جذاب به صورت اتصالی است.

برای مدل‌سازی عوامل مکانی، مناطق پنهان در PFM، توسط توزیع‌های گاوسی دو بعدی در فضای مکانی یکپارچه شده‌اند و هر کاربر دارای توزیع چندتایی در مناطق است.

مدل ارائه شده توسط Cheng و همکاران یک مدل تلفیقی از خروجی‌های PFM و روش مدل‌سازی جغرافیایی مدل گاوسی چند مرکزی است [۴]. روش اتخاذ شده برای اخذ عوامل مکانی قبلا در بخش مدل‌های مستقل بررسی شده است و در اینجا صرفاً برای اشاره به تکنیک آورده شده است.

بر اساس مشاهدات که بازدیدهای کاربر معمولاً حول چندین مرکز، مانند خانه و محل کار توزیع می‌شود، MGM مناطق فعالیت هر کاربر را با استفاده از توزیع‌های گاوسی مختلف یاد می‌گیرد. امتیاز توصیه ای با رابطه

$$P_{ij} = P(C_{ij}) \cdot P(l_j | R_i) \quad (42)$$

مشخص می‌شود که در آن $P(C_{ij}) \propto u_i l_j^T$ خروجی PFM و $P(l_j | R_i)$ به وسیله MGM محاسبه می‌شود و R_i برابر با مناطق کاربر i است.

۴-۲-۳- مدل‌های بر مبنای هوش مصنوعی

از هوش مصنوعی برای مدل‌سازی و اتصال عوامل مهم و تاثیرگذار در شبکه‌های اجتماعی مکان مبنای می‌توان

استفاده کرد. هوش مصنوعی خود شامل روش‌های مختلفی است که ما در اینجا دو نمونه از شبکه عصبی و یادگیری عمیق را بررسی می‌کنیم.

۴-۲-۳-۱- شبکه عصبی

LIU و همکاران [۱۱]، یک روش جدید به نام شبکه‌های عصبی بازگشتی مکانی و زمانی^۲ برای توصیه مکان‌های جذاب با اخذ عوامل مکانی و زمانی ارائه نمودند. ST-RNN می‌تواند زمینه‌های زمانی و مکانی محلی را در هر لایه با ماتریس‌های انتقال زمان خاص برای فواصل زمانی مختلف و ماتریس‌های انتقال مکان خاص برای فاصله‌های مختلف جغرافیایی مدل‌سازی کند.

با توجه به این که اطلاعات فاصله، عامل مهمی برای پیش‌بینی مکان است، Liu و همکاران [۱۱] همانند ماتریس‌های انتقال زمان خاص، ماتریس‌های انتقال مکان خاص را برای فاصله‌های مختلف جغرافیایی بین مکان‌ها در نظر گرفتند. ماتریس‌های انتقال فاصله خاص، خصوصیات جغرافیایی را که بر رفتار انسان تاثیر می‌گذارد اخذ می‌کنند. در ST-RNN، برای کاربر u، نمایش آن در زمان t را می‌توان با فرمول زیر محاسبه کرد:

$$h_{t,q_t^u}^u = f \left(\sum_{q_{t_i}^u \in Q^u, t-w < t_i < t} S_{q_{t_i}^u - q_t^u} T_{t-t_i} q_{t_i}^u + Ch_{t-\hat{w}, q_{t-\hat{w}}^u} \right) \quad (43)$$

که در آن $S_{q_{t_i}^u - q_t^u}$ ماتریس انتقال مکان خاص برای فاصله جغرافیایی $q_{t_i}^u - q_t^u$ با توجه به مختصات فعلی و q_t^u مختصات کاربر u در زمان t است. فاصله جغرافیایی نیز با فاصله اقلیدسی محاسبه می‌شود.

$$q_t^u - q_{t_i}^u = \left\| x_t^u - x_{t_i}^u, y_t^u - y_{t_i}^u \right\|^2 \quad (44)$$

معمولاً مکان q_{t-w}^u که نشانگر مکانی است که کاربر u در زمان t-w از آن بازدید می‌کند، در تاریخچه بازدیدهای Q^u وجود ندارد. می‌توان از مقدار تقریبی $\hat{w}$ به عنوان عرض پنجره محلی استفاده کرد. بر اساس لیست بازدید Q^u و نقطه زمان t-w، مقدار $\hat{w}$ تعیین می‌شود تا اطمینان حاصل شود که $\hat{w}$ نزدیکترین مقدار بسته به w و $q_{t-\hat{w}}^u \in Q^u$ است.

علاوه بر این، زمانی که تاریخچه داده‌ها به مدت کافی نیست یا موقعیت پیش‌بینی شده در اولین قسمت‌های تاریخچه داده‌ها است، $t < w$ می‌باشد. پس رابطه ۴۳ باید بازنویسی شود:

^۲ Spatial Temporal Recurrent Neural Networks (ST-RNN)

^۱ Geographical poisson factor model

RBM معمولی تنها دارای ارتباط بین دو لایه است، اما یک نیمه RBM می‌تواند ارتباطاتی در داخل مکان‌های جذاب لایه ورودی برای تشخیص شباهت مکانی داشته باشد [۳۸]. مطالعات قبلی به ندرت از یک نیمه RBM استفاده می‌کنند، اما برای مدل‌سازی همبستگی بین مکان‌های جذاب این روش مناسب است. به دلیل ارتباط بین واحدهای ورودی، یک نیمه RBM برای مدل‌سازی نزدیکی مکانی یک مکان جذاب به هر مکان جذاب دیگر مناسب است.

به استثنای تأثیرات مکانی و اجتماعی، بسیاری از عوامل دیگر بر رفتار بازدید کاربران، تأثیر می‌گذارد. تقریباً در نظر گرفتن همه عوامل به صورت یکجا غیرممکن است، اما روش یادگیری عمیق به طور ضمنی می‌تواند با این عوامل در سطوح مختلفی از ویژگی‌های مختلف برخورد کند [۱۱]. Junpeng و همکاران [۳۹] از یک نیمه RBM، برای تشخیص همبستگی مکانی و از یک رمزگذار خودکار عمیق^۳ با یک لایه‌ی شرطی برای استخراج همبستگی اجتماعی استفاده کردند

۴-۲-۴- مدل‌های ترکیبی

مدل ترکیبی خروجی‌های دو یا چند روش توصیه‌ای را ترکیب می‌کند و هر کدام از روش‌ها ترجیح کاربر یا یک نوع اطلاعات زمینه‌ای را مدل می‌کند. در اکثر مطالعات مربوط به این بخش با توجه به این که تکنیک‌های مختلفی برای عوامل موثر استفاده می‌شود، عامل مکانی نیز به صورت مستقل بررسی شده است. هدف ما در این بخش بیشتر تمرکز بر استفاده از تکنیک‌های مختلف در کنار هم می‌باشد. در ادامه مدل‌های شامل عوامل مکانی را ارائه می‌کنیم.

مدل ارائه شده توسط Ye و همکاران [۲۷] ترجیح کاربر، عامل اجتماعی و عامل مکانی را همزمان برای توصیه‌های مکان جذاب مدل می‌کند.

در این مدل، از فیلترینگ مشارکتی کاربر مینا برای مدل کردن ترجیح کاربر استفاده می‌شود، جایی که شباهت بین دو کاربر براساس بازدیدهای مشترک آنها محاسبه می‌شود.

برای بهره برداری از عامل اجتماعی، مدل ارائه شده پیشنهاد فیلترینگ مشارکتی مبتنی بر دوست را برای ارائه توصیه‌های مکان جذاب بر اساس دوستان مشابه انجام می‌دهد. شباهت بین دوستان مبتنی بر مکان‌های جذاب بازدید شده مشترک و دوستان مشترک است.

$$h_{t,q_t}^u = f\left(\sum_{q_{t_i}^u \in Q^{u,0} < t_i < t} S_{q_{t_i}^u - q_{t_i}^u} T_{t-t_i} q_{t_i}^u + Ch_0^u\right) \quad (45)$$

$h_0^u = h_0$ نشان‌دهنده وضعیت اولیه است. وضعیت اولیه همه‌ی کاربران باید یکسان باشد. زیرا هیچ اطلاعات رفتاری برای پیش‌بینی شخصی در چنین شرایطی وجود ندارد. در نهایت، پیش‌بینی ST-RNN را می‌توان از طریق محاسبه‌ی ضرب داخلی نمایه کاربر و گزینه‌ها انجام داد. پیش‌بینی این که آیا کاربر u زمان t به محل v می‌رود، می‌تواند به صورت زیر محاسبه شود:

$$O_{u,t,v} = (h_{t,q_v}^u + p_u)^T q_v \quad (46)$$

جایی که در آن p_u نماینده دائمی کاربر u و نشان‌دهنده علاقه و محدوده فعالیت اوست، و h_{t,q_v}^u علائق دینامیک آن را تحت شرایط خاص مکانی و زمانی مشخص می‌کند.

۴-۲-۳-۲- یادگیری عمیق

یادگیری عمیق^۱ یک روش یادگیری نمایش با سطوح مختلف بازنمایی است که با ساختن ماژول‌های غیرخطی ساده بدست می‌آید. هر ماژول سطح جدایی سطوح بالا را از داده‌های سطح پایین ترسیم می‌کند. به عنوان مثال، در دید کامپیوتر، یک تصویر را می‌توان از ویژگی‌های پیکسل سطح پایین به ویژگی‌های سطح بالاتر نشان‌دهنده رنگ، بافت و مورفولوژی در نظر گرفت. از این رو، مدل یادگیری عمیق در کشف ویژگی‌های درونی سطح بالا، بسیار مفید است و به همین دلیل برای سیستم‌های توصیه‌گر مناسب است. Junpeng و همکاران [۳۹] با تحلیل رفتار کاربران، یک مدل یادگیری عمیق برای ادغام تأثیرات مکانی و اجتماعی برای توصیه مکان جذاب ایجاد کردند. به همین منظور از ماشین بولتزمن نیمه محدود^۲ [۳۸] برای مدل‌سازی شباهت مکانی و از یک لایه‌ی شرطی برای مدل‌سازی تأثیر اجتماعی استفاده کردند.

Junpeng و همکاران [۳۹] یک روش جدید یادگیری عمیق را بر اساس یک ماشین محدود بولتزمن برای مدل‌سازی شباهت مکانی پیشنهاد دادند. یک RBM یک شبکه‌ی دو لایه است که می‌تواند توزیع احتمال را بر روی مجموعه‌ای از ورودی‌ها بیاموزد و به طور گسترده در یادگیری عمیق استفاده می‌شود. از لایه‌ی ورودی برای نمایش مکان‌های جذاب و توزیع احتمال خروجی نشان‌دهنده لایه‌ی ویژگی آنها است. یک

^۳ Deep Auto-Encoder (DAE)

^۱ Deep Learning

^۲ Restricted Boltzmann Machine (RBM)

برای کاربر i و مکان جذاب z داده شده، عامل مکانی با احتمال بازدید از z بر مبنای مکان‌های جذاب بازدید شده توسط کاربر تخمین زده می‌شود.

$$p_{ij}^g = \prod_{l_k \in L_i} \text{pr}(d(l_j, l_k)) \quad (47)$$

در این رابطه L_i مکان‌های جذاب بازدید شده توسط کاربر و $d(l_j, l_k)$ برابر با فاصله بین l_k و l_j است. $\text{Pr}(\cdot)$ احتمال این که کاربران مسافت $d(l_j, l_k)$ طی کنند را تخمین می‌زند.

مدل ارائه شده توسط Zhang و همکاران که با نام جغرافیایی-اجتماعی-دسته‌بندی معرفی شده است، یکی دیگر از مدل‌های ترکیبی است که سه دسته از زمینه‌ها شامل عوامل اجتماعی، مکانی و دسته بندی را مدل می‌کند [۹].

GeoSoCa مثل [۸] از KDE دو بعدی برای مدل‌سازی مکانی استفاده می‌کند. تفاوت آن‌ها در این است که GeoSoCa یک پهنای باند محلی (وابسته به کاربر) را اضافه می‌کند تا مدل‌سازی مکانی شخصی‌تر شود.

Zhang و همکاران [۹] توزیع قانون توان را (که به عنوان $fS(xij)$ نامیده می‌شود) از فرکانس بازدیدهای اجتماعی کاربران برآورد می‌کند. فرکانس بازدید اجتماعی xij به فرکانس بازدید در مکان جذاب مربوط به j توسط دوستان کاربر i اشاره می‌کند. GeoSoCa از توزیع تجمعی $fS(xij)$ به عنوان تأثیر اجتماعی در توصیه‌ها استفاده می‌کند.

همانند مدل سازی تأثیر اجتماعی، GeoSoCa توزیع قانون توان (که به عنوان $fC(yic)$ نامیده می‌شود) را برای فرکانس بازدید دسته‌بندی کاربران تخمین می‌زند. فرکانس بازدید دسته‌بندی نشان‌دهنده فرکانس بازدید کاربر i در تمام مکان‌های جذاب با رده c است. توزیع تجمعی $fC(yic)$ به عنوان تأثیر دسته‌بندی در توصیه‌ها مورد استفاده قرار می‌گیرد.

جدول ۱ خلاصه‌ای از روش‌های اخذ عوامل مکانی به همراه ویژگی‌های خاص هریک را نشان می‌دهد.

جدول ۲- ویژگی‌های روش‌های اخذ عوامل مکانی

مدل	تکنیک	ویژگی‌ها	منابع
مدل‌های مستقل	قانون توان	* مدل‌سازی عامل مکانی با استفاده از تمامی داده‌های بازدید * استفاده از فاصله‌ی بین جفت مکان‌های بازدید شده برای تعیین پارامترها * عملکرد مناسب در تعداد داده‌های کم	[۲۷، ۲۲]
	توزیع گاوسی	* مدل‌سازی عامل مکانی با استفاده از تمامی داده‌های بازدید * تعیین مراکز فعالیت برای کاربران * استفاده از الگوریتم ژنتیک برای حذف داده‌های خارج از محدوده * عملکرد مناسب در داده‌های با تعداد بالا	[۴۰، ۴، ۴۰]
	تخمین تراکم هسته	* مدل‌سازی عامل مکانی به صورت شخصی‌سازی شده با استفاده از داده‌های مربوط به هر کاربر * استفاده از مختصات دوبعدی به جای فاصله‌ی یک بعدی * استفاده از پهنای باند متغیر	[۸-۱۰]
	مدل‌های بر مبنای فاصله	* تقسیم منطقه به چندین منطقه بر اساس مختصات مکان‌های جذاب * تعیین تأثیر عامل مکانی با توجه به منطقه‌ای که کاربر در آن قرار گرفته و فاصله او با مکان‌های جذاب	[۷]
مدل‌های وابسته	فاکتورگیری ماتریس	* تجزیه ماتریس بازدید کاربران به ماتریس‌های با بعد کمتر به منظور اخذ ویژگی‌های پنهان کاربران و مکان‌های جذاب	[۲۹، ۱۴، ۱۲، ۳۵، ۳۴]
	مدل‌های احتمالاتی	* تجزیه ماتریس بازدید کاربر-مکان جذاب با مدل فاکتور پواسون و استفاده از آن برای توصیه	[۳۶، ۳۵، ۴]
	هوش مصنوعی	* استفاده از داده‌های صریح و ضمنی در شبکه‌های عصبی برای توصیه‌ی مکان جذاب به کاربر * استفاده از نتایج روش‌های فاکتورگیری ماتریسی در شبکه عصبی برای توصیه‌ی مکان جذاب به کاربر * استفاده از یادگیری عمیق برای اخذ اطلاعات جزئی و توصیه دقیق	[۳۹، ۱۱]
	مدل‌های ترکیبی	* اخذ عوامل مکانی به همراه دیگر عوامل تأثیرگذار در توصیه‌ی مکان جذاب	[۲۷، ۹، ۸]
مدل‌های محدودکننده	* به جای بررسی عوامل مکانی، یک فاصله معین به عنوان حد آستانه‌ی توصیه برای کاربران در نظر گرفته می‌شود.		[۸]

۵- نتیجه گیری

هدف از نگارش این مقاله بررسی انواع سیستم‌های توصیه‌گر در حوزه‌ی توصیه‌ی مکان‌های جذاب با استفاده از اطلاعات حاصل از شبکه‌های اجتماعی مکان مبنا می‌باشد. در این مقاله سعی شد تا با بررسی مقالات مختلف، روش‌هایی که برای توصیه‌ی مکان‌های جذاب از عوامل مکانی استفاده کرده‌اند، جمع‌آوری و دسته‌بندی شوند. دو رویکرد اصلی در سیستم‌های توصیه‌گر مکان مبنا وجود دارد و تقریباً همین دو رویکرد برای اخذ عوامل مکانی به کار رفته است که شامل مدل‌های مستقل و مدل‌های وابسته می‌باشد.

مدل‌های مستقل شامل چهار روش متفاوت است که این چهار روش عبارتند از توزیع قانون توان، توزیع گاوسی، تخمین چگالی هسته و مدل بر مبنای فاصله.

در روش توزیع قانون توان از تمامی داده‌های بازدید موجود و فاصله دوبه‌دو آن‌ها برای مدل‌سازی عوامل مکانی استفاده شده است. برخلاف مدل‌های بر مبنای توزیع قانون توان در مدل‌های ارائه شده با توزیع گاوسی توزیع بازدیدهای کاربر حول چندین مرکز فرض شده است. مدل‌های مبتنی بر تخمین تراکم هسته با توجه به این امر که توزیع بازدیدهای هر شخص با دیگر اشخاص متفاوت است به مدل‌سازی شخصی عوامل مکانی پرداخته است. علاوه بر این، این روش به جز فاصله برای مختصات دو بعدی نیز ارائه شده است. در مدل‌های بر مبنای فاصله مدل‌سازی عوامل مکانی به صورت شخصی برای هر کاربر و با توجه به داده‌های بازدید شخص صورت می‌گیرد.

مراجع

در میان روش‌های مستقل ذکر شده روش توزیع قانون توان از همه روش‌ها عملکرد بهتری دارد [۳۵]. روش‌هایی که از مختصات دو بعدی استفاده می‌کنند مانند تخمین چگالی هسته دو بعدی و مدل چند مرکزی گوسین تنها درحالتی از توزیع قانون توان بهتر عمل می‌کنند که کاربران حداقل ۱۰۰ بازدید از مکان‌های جذاب را ثبت کرده باشند. در میان روش‌های تخمین چگالی هسته مدلی که از پهنای باند متغیر استفاده می‌کند از دیگر مدل‌ها بهتر عمل می‌کند و عملکرد تخمین چگالی هسته یک بعدی از همه روش‌های دیگر بدتر است [۳۵]. مدل‌های بر مبنای فاصله در داده‌های با تعداد کم عملکرد مناسبی نسبت به دیگر روش‌ها دارند. زیرا بیشتر کاربران بیشتر بازدیدهای خود را در یک محدوده خاصی ثبت می‌کنند و در صورت نبودن داده‌ی کافی در نظر گرفتن محدودیت فاصله می‌تواند مدل بهتری برای بازدیدها ارائه کند. مدل‌های وابسته با توجه به تکنیک توصیه‌ای که در آن استفاده شده بود در چهار دسته فاکتورگیری ماتریسی، مدل احتمالی بر مبنای فاکتور پواسون، هوش مصنوعی و مدل‌های ترکیبی بررسی شدند. با توجه به این که عملکرد مدل‌های وابسته به صورت کلی بررسی می‌شود و عامل مکانی به صورت مستقل در آن قابل بررسی نیست لذا در مورد دقت آنها مقایسه‌ای صورت نمی‌گیرد. اما با توجه به دقت بالای سیستم‌های توصیه‌گر مکانی اتصالی نسبت به ترکیبی [۲] امروزه سیستم‌های اتصالی بیشتر مورد توجه قرار گرفته‌اند. لذا مدل‌سازی عامل مکانی نیز به صورت وابسته و در کنار دیگر عوامل موثر انجام می‌شود.

- [1] Ricci, F., L. Rokach, and B. Shapira, Introduction to recommender systems handbook. Recommender systems handbook. 2011: Springer.
- [2] Zhao, S., Lyu, Michael R., King, Irwin, Point-of-Interest Recommendation in Location-Based Social Networks. 2018: Springer Singapore.
- [3] Zhang, J.-D. and C.-Y. Chow, Point-of-interest recommendations in location-based social networks. Sigspatial Special, 2016.
- [4] Cheng, C., et al. Fused matrix factorization with geographical and social influence in location-based social networks. in Twenty-Sixth AAAI Conference on Artificial Intelligence. 2012.
- [5] Zhao, S., I. King, and M.R. Lyu, A survey of point-of-interest recommendation in location-based social networks. arXiv preprint arXiv:1607.00647, 2016.
- [6] Zhao, S., I. King, and M.R. Lyu. Capturing geographical influence in POI recommendations. in International Conference on Neural Information Processing. 2013. Springer.
- [7] Zhang, Z., et al., Fused matrix factorization with multi-tag, social and geographical influences for poi recommendation. World Wide Web, 2019. 22(3): p. 1135-1150.

- [8] Zhang, J.-D., C.-Y. Chow, and Y. Li. Lore: Exploiting sequential influence for location recommendations. in Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. 2014. ACM.
- [9] Zhang, J.-D. and C.-Y. Chow. 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. 2015. ACM.
- [10] Zhang, J.-D. and C.-Y. Chow. 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. 2013. ACM.
- [11] Liu, Q., et al. Predicting the next location: A recurrent model with spatial and temporal contexts. in Thirtieth AAAI Conference on Artificial Intelligence. 2016.
- [12] Lian, D., et al. GeoMF :joint geographical modeling and matrix factorization for point-of-interest recommendation. in Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 2014. ACM.
- [13] Li, H., et al. Point-of-interest recommender systems: A separate-space perspective. in 2015 IEEE International Conference on Data Mining. 2015. IEEE.
- [14] Li, H., et al. 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. 2016. ACM.
- [15] Levandoski, J.J., et al. Lars: A location-aware recommender system. in 2012 IEEE 28th international conference on data engineering. 2012. IEEE.
- [16] Ference, G., M. Ye, and W.-C. Lee. Location recommendation for out-of-town users in location-based social networks. in Proceedings of the 22nd ACM international conference on Information & Knowledge Management. 2013. ACM.
- [17] Chen, X., et al. On information coverage for location category based point-of-interest recommendation. in Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015.
- [18] Yin, H., et al., LCARS: A spatial item recommender system. ACM Transactions on Information Systems (TOIS), 2014. 32(3): p. 11.
- [19] Lian, D., et al. Content-aware collaborative filtering for location recommendation based on human mobility data. in 2015 IEEE International Conference on Data Mining. 2015. IEEE.
- [20] Hu, B. and M. Ester. Social topic modeling for point-of-interest recommendation in location-based social networks. in 2014 IEEE international conference on data mining. 2014. IEEE.
- [21] Gao, H., et al. Content-aware point of interest recommendation on location-based social networks. in Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015.
- [22] Yuan, Q., et al. Time-aware point-of-interest recommendation. in Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval. 2013. ACM.
- [23] Feng, S., et al. Personalized ranking metric embedding for next new POI recommendation. in Twenty-Fourth International Joint Conference on Artificial Intelligence. 2015.
- [24] Cheng, C., et al. Where you like to go next: Successive point-of-interest recommendation. in Twenty-Third International Joint Conference on Artificial Intelligence. 2013.
- [25] Zhao, S., I. King, and M.R. Lyu. Geo-pairwise ranking matrix factorization model for point-of-interest recommendation. in International Conference on Neural Information Processing. 2017. Springer.
- [26] Liu, B., et al. Learning geographical preferences for point-of-interest recommendation. in Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. 2013. ACM.
- [27] Ye, M., et al. Exploiting geographical influence for collaborative point-of-interest recommendation. in Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval. 2011. ACM.
- [28] Zhang, Z., et al., Fused matrix factorization with multi-tag, social and geographical influences for poi recommendation. World Wide Web, 2018: p. 1-16.
- [29] Li, X., et al. Rank-geofm: A ranking based geographical factorization method for point of interest recommendation. in Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2015. ACM.
- [30] Rafailidis, D. and F. Crestani. GeoDCF: Deep Collaborative Filtering with Multifaceted Contextual Information in Location-Based Social Networks. in Joint European Conference on Machine Learning and Knowledge Discovery in Databases. 2018. Springer.

- [31] Zhao, S., et al. Geo-teaser: Geo-temporal sequential embedding rank for point-of-interest recommendation. in Proceedings of the 26th international conference on world wide web companion. 2017. International World Wide Web Conferences Steering Committee.
- [32] Yin, H., et al., Spatial-aware hierarchical collaborative deep learning for POI recommendation. IEEE Transactions on Knowledge and Data Engineering, 2017. 29(11): p. 2537-2551.
- [33] Wang, W., et al. ,ST-SAGE: A spatial-temporal sparse additive generative model for spatial item recommendation. ACM Transactions on Intelligent Systems and Technology (TIST), 2017. 8(3): p. 48.
- [34] Liu, Y., et al. Exploiting geographical neighborhood characteristics for location recommendation. in Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management. 2014. ACM.
- [35] Liu, Y., et al., An experimental evaluation of point-of-interest recommendation in location-based social networks. Proceedings of the VLDB Endowment, 2017. 10(10): p. 1010-1021.
- [36] Liu, B., et al., A general geographical probabilistic factor model for point of interest recommendation. IEEE Transactions on Knowledge and Data Engineering, 2014. 27(5): p. 116. ۱۱۷۹-۷
- [37] Kurashima, T., et al. Geo topic model: joint modeling of user's activity area and interests for location recommendation. in Proceedings of the sixth ACM international conference on Web search and data mining. 2013. ACM.
- [38] Hinton, G.E., A practical guide to training restricted Boltzmann machines, in Neural networks: Tricks of the trade. 2012, Springer. p. 599-619.
- [39] Guo, J., et al., Combining Geographical and Social Influences with Deep Learning for Personalized Point-of-Interest Recommendation. Journal of Management Information Systems, 2018. 35(4): p. 1121-1153.
- [40] Cho, E., S.A. Myers, and J. Leskovec. 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. 2011. ACM