

ارائه سیستم توصیه گر زمانمند، ترجیح آگاه و مکان آگاه در شبکه اجتماعی مکان مبنا

رویا شورونی^{۱*}، محمدرضا ملک^۲

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

خواجه نصیرالدین طوسی

royashurouni@gmail.com

^۲ دانشیار گروه سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی

(قطب علمی مهندسی فناوری اطلاعات مکانی)

mrmalek@kntu.ac.ir

(تاریخ دریافت دی ۱۳۹۳، تاریخ تصویب خرداد ۱۳۹۴)

چکیده

رشد سریع شبکه‌های اجتماعی مکان مبنا به واسطه‌ی جذب میلیون‌ها کاربر، نشان از محبوبیت زیاد آن‌ها در مدت‌زمانی کوتاه دارد. با تجزیه و تحلیل داده‌های این شبکه‌ها در زمینه‌ی مکانی، زمانی و اجتماعی، می‌توان سرویس‌های متنوع مکان مبنا ارائه کرد. سیستم‌های توصیه گر از جمله سرویس‌های مکان مبنا محبوب این شبکه‌ها محسوب می‌شود. سیستم‌های توصیه گر، با بهره‌گیری از تکنیک‌های آماری و تکنیک‌های کشف دانش، به توصیه مکان‌های جدید و کاهش مشکلات ناشی از حجم زیاد داده‌ها می‌پردازند. هدف پژوهش حاضر، ارائه روش ترکیبی برای توصیه مکان‌های جدید به کاربران بر اساس پیشینه مکانی آن‌ها است. روش ارائه شده، مجموعه‌ای از مکان‌های نزدیک به موقعیت کاربر را با در نظر گرفتن شرایط زمانی، فاصله جغرافیایی بین کاربران و همچنین ترجیحات مکانی کاربر که به‌طور خودکار از پیشینه مکانی او به دست می‌آید، پیشنهاد می‌دهد. در این روش، از پالایش محتوای اطلاعاتی، در راستای برآورد میزان علاقه کاربر به مکان‌های بازدید شده استفاده می‌شود. از پالایش مشارکتی نیز برای محاسبه شباهت بین کاربران و پیش‌بینی میزان علاقه کاربر به مکان جدید، بر اساس تجمیع امتیازات داده شده توسط کاربران مشابه، استفاده می‌شود. در نهایت ترکیب این دو روش، امکان ارائه مکان‌های دیدنی را در نزدیکی کاربر سواره در بازه زمانی فعلی فراهم می‌سازد. در این مقاله از داده‌های نقطه-مکان شبکه Gowalla مربوط به شهر پکن چین در بازه زمانی بین اکتبر ۲۰۱۱ تا نوامبر ۲۰۱۱ استفاده شده است. به منظور بررسی عملکرد روش مورد استفاده، نتایج حاصله از این روش با نتایج دو روش پایه در سیستم توصیه گر مقایسه شد. در نهایت توصیه مکان توسط این روش، نسبت به پالایش مشارکتی و GM-FCF، به ترتیب حدود ۱۵ و ۱۲ درصد افزایش دقت را نشان می‌دهد.

واژگان کلیدی: شبکه‌های اجتماعی مکان مبنا، پیشینه مکانی، سیستم توصیه گر

۱- مقدمه

در سال‌های اخیر شبکه‌های اجتماعی^۱ با هدف ایجاد ارتباطات بین افراد [۱]، جوابگویی به حس گرایش انسان‌ها نسبت به هم و ایجاد زمینه اطلاعاتی [۲] به وجود آمد. با رشد سریع این شبکه‌ها، افراد تمایل زیادی به تبادل داده و به اشتراک‌گذاری متن‌ها و تصاویر با دیگر کاربران پیدا کردند. در این شبکه‌ها هر کاربر نه تنها مصرف‌کننده‌ی اطلاعات، بلکه تولیدکننده داده نیز است. در این شبکه‌ها سعی می‌شود به‌نوعی ارتباطات در دنیای واقعی شبیه‌سازی شود [۲]. با گسترده‌ی ارتباطات در این شبکه‌ها، تحلیلگران نیاز دارند تا مطالعات خود را فراتر از بررسی و شناخت کاربران به‌صورت انفرادی پیش ببرند. آن‌ها با بررسی روابط اجتماعی بین افراد، به کشف الگوهای برجسته مانند ردیابی جریان اطلاعاتی و کاوش تأثیر این شبکه‌ها بر روی روابط بین افراد می‌پردازند [۳].

تجزیه و تحلیل فضا و مکان در دو دهه گذشته، به‌عنوان یکی از اجزای کلیدی پژوهش‌های علوم اجتماعی شده است و این را می‌توان به تحول فضای اجتماعی در سراسر جهان، به همراه تغییرات در نظریه‌پردازی علوم اجتماعی نسبت داد. یکی از این تغییرات، تحت عنوان مفهوم "فشرده‌سازی فضا و مکان" انجام گرفت و نتایج آن باعث تغییرات وسیعی در فن‌آوری اطلاعات و ارتباطات شد [۳]. در گذشته توجه به بعد مکانی در زمینه علوم اجتماعی، نامشهود بوده؛ اما در سال‌های اخیر احیای آن در حال وقوع است. به‌عنوان مثال، بسیاری از دانشمندان علوم اجتماعی، از فن‌آوری‌های تجزیه و تحلیل مکانی جدید مانند سیستم‌های اطلاعات مکانی و سنجش‌ازدور در کارهای تجربی خود بهره می‌برند و در چارچوب نظری توجه زیادی به تعاملات مکانی می‌کنند.

با اضافه شدن بعد مکان به شبکه‌های اجتماعی، شبکه‌های اجتماعی مکان‌مبنا به منصف ظهور رسید. داده‌های مکانی مکان در شبکه‌های اجتماعی مکان‌مبنا، ارتباطات و اطلاعات موجود در فضای مجازی را با ایجاد یک پل ارتباطی با دنیای فیزیکی به واقعیت نزدیک‌تر می‌کند. با توسعه سرویس‌های مبتنی بر شبکه‌های اجتماعی مکان‌مبنا از جمله Foursquare، Loopt

و GeoLife، کاربران به راحتی می‌توانند تجربه‌های زندگی‌شان از جهان فیزیکی را از طریق تجهیزات همراه به اشتراک بگذارند [۴].

نتیجه عملکرد سرویس‌های مبتنی بر شبکه‌های اجتماعی مکان‌مبنا، دستیابی به مجموعه بزرگی از داده است که پیشینه مکانی^۲، ساختار روابط اجتماعی، رفتار حرکتی و ویژگی‌های کاربران از آن قابل استخراج است. این منابع غنی اطلاعاتی، توانایی بالایی در شناخت کاربران دارد. زیرا که پیشینه مکانی یک فرد در دنیای واقعی، تا حدی نشان‌دهنده علایق و رفتار او است. بر این اساس، افرادی که پیشینه مکانی مشابهی دارند، به احتمال زیاد سلاقی و رفتار مشترکی دارند. برآورد میزان مشابهت بین کاربران، منجر به ارائه سرویس‌ها و سیستم‌هایی جدید و ابتکاری می‌شود که یکی از مهم‌ترین آن‌ها سیستم‌های توصیه‌گر است.

سیستم توصیه‌گر یا سامانه پیشنهادگر با تحلیل رفتار کاربر، اقدام به پیشنهاد مناسب‌ترین اقلام مانند داده، مکان یا کالا می‌نماید. این سیستم رویکردی است که برای مواجهه با مشکلات ناشی از حجم فراوان و رو به رشد اطلاعات ارائه‌شده است و به کاربر کمک می‌کند تا در میان حجم عظیم اطلاعات سریع‌تر به هدف خود نزدیک شود. از این رو برخی سیستم توصیه‌گر را پالایش گروهی^۳ می‌دانند [۴]. برای ارائه توصیه‌های شخصی به کاربران با توجه به پیشینه مکانی آن‌ها، روش‌های گوناگونی بر اساس اختلاف در الگوریتم پیشنهادات وجود دارد که می‌توان آن‌ها را به چهار دسته کلی پالایش مشارکتی، پالایش محتوا محور^۴، دانش محور^۵ و بافت آگاه تقسیم کرد.

علاوه بر این با ترکیب این روش‌ها می‌توان تکنیک‌های مختلف دیگری ارائه داد [۵]. در این تحقیق از دو روش پالایش محتوا محور و پالایش مشارکتی به‌عنوان عمومی‌ترین تکنیک‌های توصیه، برای توصیه مکان استفاده شده است.

هدف تحقیق حاضر، ارائه یک سیستم توصیه‌گر مکان، بر مبنای داده‌های نقطه-مکان به اشتراک گذاشته‌شده توسط کاربران در شبکه اجتماعی مکان‌مبنا است. در

^۲ Location History

^۳ Collaborativ Filtering

^۴ Content-Based Filtering

^۵ Knowledge-Base Filtering

^۶ Context-Aware Recommender Systems

^۱ Social Network

راستای رسیدن به این هدف، دو گام اساسی انجام می-پذیرد. گام نخست محاسبه امتیازات داده شده به مکان ها با به کارگیری تکنیک پالایش محتوا محور است. مرحله بعد اجرای پالایش مشارکتی بر مبنای امتیازات محاسبه شده، برای محاسبه مشابهت بین کاربر هدف به دیگر کاربران و پیشنهاد مکان به کاربر هدف است.

ادامه مقاله در چهار بخش به شرح زیر تدوین شده است: در بخش دوم یک مرور کلی بر شبکه های اجتماعی و شبکه های اجتماعی مکان مینا انجام می شود؛ سپس سیستم های توصیه گر مبتنی بر شبکه های اجتماعی مکان- مینا بررسی قرار می شود. در بخش سوم مفاهیم مورد نیاز برای ارائه روش پیشنهادی بیان می شد. بخش چهارم به معرفی داده های مورد استفاده و منطقه مورد مطالعه و همچنین روش اجرای الگوریتم پیشنهادی می پردازد. در نهایت نتیجه گیری و ارزیابی در بخش پنجم ارائه شده است.

۲- پیشینه تحقیق

شبکه اجتماعی یک ساختار اجتماعی از افراد در ارتباط باهم، توسط یک یا چند نوع وابستگی متقابل از جمله دوستی و علائق و سلاقی مشابه است. شبکه اجتماعی، سازنده و منعکس کننده روابط اجتماعی میان مردم در زندگی واقعی و همچنین ارائه دهنده راه کارهایی به کاربران برای به اشتراک گذاری ایده ها، فعالیت ها، رویدادها، و علائق آن ها از طریق اینترنت است [۱].

از طرف دیگر پیشرفت در ارتباطات بی سیم و فن آوری های اخذ مکان قابلیت مناسبی را برای ارائه ی سرویس های مختلف از جمله سرویس های مکان آگاه برای کاربران به وجود آورده است که افراد را قادر به اضافه کردن بعد مکان به شبکه های اجتماعی کرده و نتیجه ی آن ارائه ی سرویس های متنوع مکانی- اجتماعی می باشد. ترکیب اطلاعات عمومی کاربران با اطلاعات مکانی و تحلیل آن ها این امکان را فراهم می کند تا بتوانیم به کشف بهتر فعالیت های روزانه، روابط اجتماعی و الگوهای رفتاری کاربران بپردازیم. به عنوان مثال، کاربر می تواند با یک تلفن همراه نظرات خود را نسبت به یک رستوران در یک شبکه اجتماعی آنلاین به اشتراک بگذارد تا مردم بتوانند در زمان بازدید از رستوران به آن ها رجوع کنند.

از نظر تحقیقاتی بر روی شبکه های اجتماعی مکان مینا و سرویس های مبتنی بر آن مطالعات زیادی انجام شده است و به طور فزاینده ای مورد توجه محققان دانشگاهی و صنعتی برای رسیدن به اهدافشان قرار گرفته است. در سال های اخیر ارائه ی سیستم های توصیه گر مکانی بر مبنای داده های این شبکه ها توجه بسیاری از محققین را به سمت خود جلب کرده و تحقیقات متفاوتی انجام شده است. اساس این تحقیقات استفاده از داده های گروهی کاربران برای ارائه ی پیشنهاد های مکانی است. Bao و همکاران [۶]، طی یک بررسی جامع، به مطالعه چالش ها و فرصت های جدیدی که موقعیت مکانی برای توصیه در شبکه های اجتماعی مکان مینا به ارمغان می آورد، پرداختند. آن ها سیستم های توصیه گر را بر اساس داده ها، الگوریتم های توصیه و هدف توصیه در سه دسته کلی قراردادند. منظور از داده ها پروفایل کاربر، مکان های بازدید شده توسط کاربران و خط سیر^۱ کاربران است. هدف توصیه می تواند شامل مکان ها، کاربران، فعالیت ها و رسانه های اجتماعی باشد و روش های به کار گرفته شده می تواند روش های مبتنی بر محتوا، روش های مبتنی بر تجزیه و تحلیل روابط، و روش های مبتنی بر پالایش مشارکتی باشد. از اولین تحقیقاتی که در زمینه ی استفاده از پیشینه ی مکانی برای ارائه ی پیشنهادات انجام شده، می توان به تحقیق صورت گرفته توسط Park و همکاران در سال ۲۰۰۷، اشاره کرد. هدف آن ها ارائه ی پیشنهاد مکان بر اساس ترجیحات و خصوصیات استنباط شده ی کاربران از پیشینه ی مکانی آن ها بوده است [۷]. احتمالاً جامع ترین تحقیقات را در زمینه ی شبکه های اجتماعی مکان مینا، Zheng و همکاران در بخش تحقیقات شرکت مایکروسافت انجام داده اند [۸]. آن ها بر اساس پیشینه مکانی کاربران با استفاده از خط سیر جمع آوری شده از GPS علاوه بر توصیه مکان، به پیشنهاد فعالیت به کاربران شبکه اجتماعی پرداختند.

می توان گفت در اکثر تحقیقات، پیشینه مکانی کاربران از داده های حرکتی GPS استخراج شده و داده های نقطه^۲ مکان کمتر مورد توجه قرار گرفته است، در حالی که این نوع داده از امنیت بالاتری برخوردار است. داده های

^۱ Trajectory

^۲ Check-in

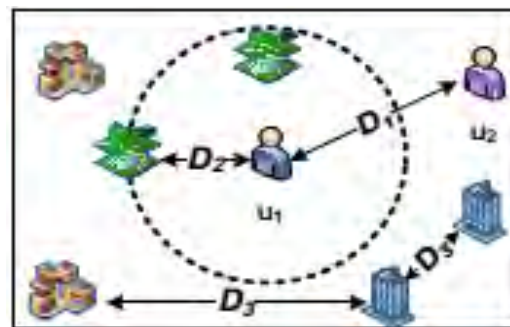
نقطه مکان شبکه‌های اجتماعی مکان مبنا را می‌توان به روش‌های مختلف برای سیستم‌های مختلف مکانی استفاده می‌شود. YE و همکاران [۹]، تعداد دفعات بازدید از مکان‌ها را که به صورت داده نقطه مکان می‌باشد، به یک فرمت دوتایی کاربر و مکان جهت امتیازدهی به مکان‌های مورد بازدید کاربران تبدیل کردند. سپس، این امتیازها را برای محاسبه وزن مشابهت میان کاربران مورد استفاده قرار دادند. در اینجا ارائه‌ی پیشنهاد به مسئله‌ی تخمین امتیاز برای موارد اطلاعاتی که توسط کاربر بازدید نشده است، تقلیل می‌یابد. واضح است این تخمین عموماً بر اساس امتیازات داده‌شده از طرف کاربر به موارد اطلاعاتی دیگر صورت می‌گیرد. وقتی که ما بتوانیم امتیازاتی را از طرف کاربر برای موارد اطلاعاتی برآورد کنیم، می‌توانیم موارد اطلاعاتی با بالاترین امتیاز را که از پیشینه مکانی کاربران مشابه به دست آمده، نیز به کاربر پیشنهاد کنیم. Zheng و همکاران در سال ۲۰۱۱ [۱۰]، با توجه به این واقعیت که دوستان اجتماعی نسبتاً موارد اطلاعاتی مشابهی را به اشتراک می‌گذارند؛ روش پالایش مشارکتی حافظه مبنا را بعنوان یک الگوریتم ابتکاری برای به پیش‌بینی امتیازها بر اساس امتیازهای پیشین کاربران به کار برد.

Wang و همکاران در تحقیق خود مشکل توصیه مکان جدید به کاربران شبکه‌های اجتماعی مکان مبنا را بررسی کردند [۱۱]. آن‌ها با در نظر گرفتن چهار فاکتور رفتار کاربر (مکان‌های بازدید شده)، روابط اجتماعی بین کاربران، شباهت بین کاربران و موقعیت مکان‌ها، الگوریتم‌هایی برای توصیه مکان ارائه دادند. ایده اصلی این مطالعات، کسب علائق مشابه از طریق روابط اجتماعی میان دوستان به منظور ارتقاء توصیه‌های شخصی و قابل اطمینان است. به‌طور کلی می‌توان نتیجه گرفت که تمرکز عمده مطالعات پیشین بر روی توصیه‌های متداول در شبکه‌های اجتماعی مکان مبنا بوده است. تفاوت اصلی فعالیت حاضر با بقیه فعالیت‌های انجام‌شده، بهره‌برداری از هر دو ملاک نزدیکی مکان‌ها و مشابهت کاربران برای توصیه است. بررسی‌ها نشان می‌دهد که افراد نزدیک بهم با احتمال بسیار بالاتری

مکان‌های مشترک را بازدید میکنند. به‌طور خلاصه می‌توان دریافت که فعالیت‌های اجتماعی کاربران در واقعیت تا حد زیادی تحت تأثیر مجاورت مکانی قرار می‌گیرد.

۳- روش پیشنهادی

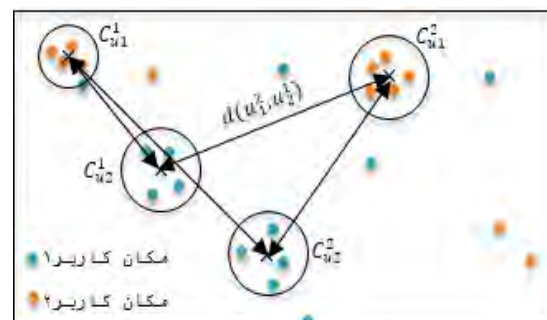
در این فصل، ضمن بیان مبانی نظری روش‌های مورد استفاده، مدل پیشنهادی را با هدف توسعه یک سرویس توصیه‌گر با در نظر گرفتن تأثیر بافت‌های مکانی، زمانی، فردی را شرح می‌دهیم. در مرحله نخست، تأثیر پارامتر فاصله بین کاربران در میزان اشتراک کاربران در بازدید از مکان‌ها بررسی می‌شود. سپس با ایجاد ماتریس کاربر-مکان در چهار بازه زمانی مربوط به یک روز و استفاده از پالایش محتوا محور، میزان علاقه‌ی کاربران به مکان‌ها برآورد و ماتریس کاربر-مکان مجدداً ساخته می‌شود. در مرحله بعد با در اختیار داشتن موقعیت فعلی کاربر، مجموعه‌ی مکان‌ها با توجه به فاصله‌ی آن‌ها از کاربر محدود می‌شود. این کار علاوه بر اینکه سرعت محاسبات و پردازش‌ها را بالا می‌برد، باعث افزایش دقت و صحت پیشنهادها نیز می‌شود. در نهایت با استفاده از یک تابع ابتکاری، مشابهت بین کاربر هدف با دیگر کاربران بر اساس ترکیب فاصله بین آن‌ها و امتیازاتی که به مکان‌ها داده‌اند، برآورد می‌شود. امتیاز مکان‌ها برای کاربر هدف، بر اساس تجمیع امتیازات داده‌شده توسط کاربران مشابه محاسبه می‌گردد. در نهایت مکان‌هایی که دارای بیشترین وزن هستند، به کاربر هدف توصیه می‌شوند. مراحل اجرای مدل مورد استفاده در شکل ۱ مشاهده می‌شود.



شکل ۳- انواع فاصله در شبکه‌های اجتماعی مکان مبنا [۴]

نخست فاصله بین کاربران که شباهت بین آن‌ها را تحت تأثیر قرار می‌دهد. به عنوان مثال کاربرانی که نزدیک به هم زندگی می‌کنند، با احتمال بیشتری از مکان‌های مشابه که در نزدیکی آن‌ها است، بازدید می‌کنند. دومین فاصله، فاصله بین مکان‌ها و کاربران است که احتمال بازدید کاربر از یک مکان را تحت تأثیر قرار می‌دهد. سوم، فاصله بین مکان‌ها که ارتباط بین آن‌ها را تحت تأثیر قرار می‌دهد.

در اینجا برای محاسبه میزان شباهت بین کاربران، نحوه محاسبه فاصله بین کاربران را شرح می‌دهیم. همان‌طور که در شکل ۴ نشان داده شده فاصله بین کاربر u_i و u_j از میانگین فاصله بین مرکز مکان‌های اکتشافی از پیشینه مکانی آن‌ها به دست می‌آید.



شکل ۴- فاصله بین کاربران شبکه اجتماعی مکان مبنا

نقاط قرمز و آبی رنگ شکل، داده‌های ثبت شده توسط کاربران و دواير سیاه رنگ، مکان‌های اکتشافی از خوشه‌بندی است. برای کشف مکان‌ها، نقاط ثبت شده کاربر را با استفاده از الگوریتم Mean-Shift خوشه‌بندی می‌کنیم. تنها پارامتر ورودی الگوریتم Mean-Shift شعاع خوشه‌ها است که در اینجا با مقایسه نقاط مکانی به اشتراک گذاشته شده توسط کاربران با مجموعه داده‌ی واقعی مربوط به مکان‌ها، دریافتیم نقاطی را که در فاصله

۱۰۰ متری از هم قرار دارند، می‌توان درون یک خوشه قرار داد که مرکز خوشه مربوط به یک مکان منحصر به فرد در دنیای فیزیکی است. پس از کشف مکان‌ها، میانگین فاصله بین مکان‌ها را توسط معادله ۱ محاسبه می‌کنیم:

$$D_1(u_i, u_j) = \frac{\sum_l \sum_k d(C_{u_i}^l, C_{u_j}^k)}{l * k} \quad (1)$$

در این معادله $C_{u_i}^l$ و $C_{u_j}^k$ به ترتیب بیانگر مرکز کلاس یا مکان l و k از مجموعه مکان‌های بازدید شده توسط کاربر u_i و u_j تابع d بیانگر فاصله بین مکان‌های اکتشافی و l و k تعداد مکان‌های مورد بازدید توسط کاربر u_i و u_j است. بین هر دو مکان یک فاصله برآورد می‌شود و برای l و k مکان بازدید شده توسط کاربر u_i و u_j تعداد $l * k$ فاصله وجود دارد، و پارامتر $\frac{1}{l * k}$ از مجموع فواصل، میانگین می‌گیرد.

حال فاصله بین مکان‌ها و کاربران را بیان می‌داریم. این فاصله بر روی احتمال بازدید کاربر از مکان تأثیرگذار است. در واقعیت کاربران علاقه‌مند به بازدید از مکان‌هایی هستند که به موقعیت فعلی آنها نزدیک‌تر است. سیستم‌های توصیه‌گر در شبکه‌های مکان مبنا، می‌بایست چگونگی تأثیر موقعیت فعلی کاربر را در ارائه پیشنهادها بررسی کند. امتیازدهی به مکان‌ها بر اساس فاصله کاربر از مکان و کیفیت آن مکان یک امر غیر بدیهی است. موقعیت کاربر علاوه بر اینکه یک محدودیت مکانی برای توصیه ایجاد می‌کند، بلکه ترجیحات کاربر را تحت تأثیر قرار می‌دهد.

بر این اساس، فضای جست‌وجو تنها به مکان‌هایی که در نزدیکی موقعیت فعلی کاربر سواره هستند، کاهش می‌یابد. این فاصله به علت متغیر بودن شرایط ترافیکی، وضعیت کاربر، میزان آشنایی کاربر و دیگرها، توسط کاربر سیستم به طور دلخواه تعیین می‌شود. هم‌چنین با به‌روزرسانی موقعیت کاربر سیستم و تغییر فاصله توسط کاربر، مجموعه مکان‌ها در فضای جست‌وجو مجدداً تغییر می‌کند.

۳-۲- ترجیح آگاهی و زمانمندی

در سیستم‌های توصیه‌گر مکانی متداول بر اساس پالایش مشارکتی، تعداد بازدید از مکان توسط کاربر،

یک مکان برای کاربر در بازه زمانی خاص است که به ترتیب توسط معادله ۲ و ۳ محاسبه می‌شوند. این مقادیر بر روی هم تأثیر می‌گذارند، درواقع اهمیت کاربر برای مکان و اهمیت مکان برای کاربر، یک رابطه دوجانبه است و می‌توان بر اساس تجمیع این دو مقدار، امتیاز کلی داده‌شده از طرف کاربر به مکان را در بازه زمانی خاص برآورد کرد. این امتیاز با تعریف FIF برای سه‌تایی (کاربر، مکان، زمان) محاسبه می‌شود.

$$UFILF(u, l, t) = \frac{n(u, l, t)}{nc_U(l, t)} \times \log \frac{N_L(t)}{n_L(u, t)} \quad (2)$$

$$LFIUF(u, l, t) = \frac{n(u, l, t)}{nc_L(u, t)} \times \log \frac{N_U(t)}{n_U(l, t)} \quad (3)$$

$$FIF = LFIUF + UFILF \quad (4)$$

در معادله مذکور $n(u, l, t)$ بیانگر تعداد دفعات بازدید از مکان l توسط کاربر u در بازه زمانی t ، $nc_U(l, t)$ تعداد داده‌های نقطه‌ای که در بازه زمانی t توسط تمام کاربران در مکان l ثبت شده، $N_L(t)$ تعداد تمامی مکان‌های مورد بازدید و $n_L(u, t)$ نیز تعداد تمامی مکان‌های که در بازه زمانی t توسط کاربر u مورد بازدید قرار گرفته است. این الگوریتم بین تعداد تکرار مکان‌ها و تعداد کاربران تعادل ایجاد می‌کند. تمامی مکان‌ها در این معادله در فاصله D_2 از موقعیت فعلی کاربر هستند.

ابتدا در چهار بازه زمانی مختلف برای هر کاربر FIF را برای تمامی مکان‌های که در فاصله D_2 از موقعیت فعلی کاربر بازدید کرده است، محاسبه و سپس با استفاده از روش نرمال‌سازی به مقادیر بین (۰، ۱) تبدیل می‌کنیم. مقادیر نرمال شده FIF امتیاز کاربرها به مکان‌هاست. هر چه مقدار FIF برای یک مکان بیشتر باشد، می‌توان نتیجه گرفت که کاربر در آنجا و همسایگی آنجا بیشتر حضور پیدا می‌کند و آن منطقه دارای اهمیت بیشتری است. در واقعیت کاربران علاقه‌مند به بازدید از مکان‌هایی هستند که علاوه بر اینکه مورد بازدید کاربران مشابه به آن‌ها قرار گرفته باشد، در نزدیکی آن‌ها نیز باشد.

۳-۳- توصیه مبتنی بر پالایش مشارکتی

الگوریتم‌های مبتنی بر پالایش مشارکتی تا به امروز موفق‌ترین روش پیشنهاد دهی به کاربران در بین سایر

تجویزاً به‌عنوان امتیاز داده‌شده از طرف کاربر به مکان در نظر گرفته می‌شد. اما مشکل اساسی این روش این است که تعداد بازدید از یک مکان (n_{ij}) نمی‌تواند به‌طور صریح میزان علاقه کاربر به آن مکان را نشان دهد. درواقع پیش‌بینی ترجیحات مکانی کاربران نباید تنها بر این اساس باشد که وی چه تعداد دفعات از یک مکان بازدید کرده است، بلکه باید اهمیت هر مکان را برای کاربر نسبت به تمامی مکان‌های بازدید شده توسط آن کاربر بسنجد. چگونگی استفاده از داده‌های نقطه‌ای جهت توصیه مکان یک سؤال چالش‌برانگیز برای محققان است و در حال حاضر مطالعات اندکی با تمرکز بر این مسئله وجود دارد. لذا در ادامه محاسبه امتیازات داده‌شده به مکان‌ها بر اساس یک تکنیک موجود در پالایش محتوا محور توضیح داده می‌شود.

پالایش محتوا محور، آیتم‌های موردعلاقه کاربران را بر اساس آیتم‌های مشابهی که کاربر قبلاً به آن ابراز علاقه نموده‌است، استخراج می‌کند [۱۳]. این روش تلاش می‌کند یک پروفایل کاربری بر پایه مشخصات کاربر بسازد و از آن برای پیش‌بینی نظرات کاربر در مورد آیتم‌هایی که تا به حال مورد بازدید او قرار نگرفته است، استفاده کند. در این تکنیک اهمیت یک آیتم در مجموعه آیتم‌ها با استفاده از معیار وزن گونه‌ای w_{ij} به روش‌های مختلف تعریف می‌شود. یکی از بهترین روش‌ها برای مشخص کردن وزن آیتم‌ها در بازیابی اطلاعات، روش $TF - IDF$ است. این روش برای بازیابی و غربال اطلاعات متنی است [۱۴]. در اطلاعات متنی، محتوای اطلاعاتی با کلیدواژه‌ها بیان می‌شوند و روش $TF - IDF$ با اندازه‌گیری اهمیت کلمات کلیدی در یک فایل متنی، اطلاعات مهم را بازیابی می‌کند. در مقاله حاضر، برای مشخص کردن امتیازات داده‌شده از طرف کاربران به مکان‌ها در بازیابی مکان‌ها، از روش ابتکاری FIF که بر اساس تکنیک $TF - IDF$ است، استفاده می‌کنیم. با توجه به دو نوع محتوای اطلاعاتی کاربر و مکان، روش FIF از دو قسمت $UFILF^3$ و $LFIUF^4$ تشکیل شده است. $UFILF$ بیانگر میزان اهمیت بازدید یک کاربر از مکان و $LFIUF$ بیانگر میزان اهمیت

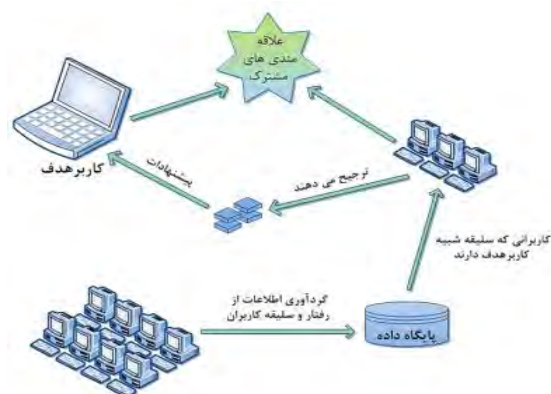
^۱ Term Frequency Inverse Document Frequency

^۲ Frequency - Inverse Frequency

^۳ User Frequency - Inverse Location Frequency

^۴ Location Frequency - Inverse User Frequency

تکنیک‌ها و روش‌های مورد استفاده در سیستم‌های توصیه‌گر دانست. همان‌طور که در شکل ۵ ملاحظه می‌کنید ایده توصیه آیت‌ها به کاربران بر اساس رتبه داده‌شده به سایر کاربران و آیت‌های مشابه است [۱۵].



شکل ۵- توصیه بر مبنای تکنیک پالایش مشارکتی

الگوریتم‌های مبتنی بر پالایش مشارکتی به دو دسته مدل‌مبنا و حافظه‌مبنا تقسیم می‌شوند. روش‌های حافظه‌مبنا از امتیاز مربوط به داده‌ها جهت محاسبه شباهت بین کاربران یا مکان‌ها استفاده می‌کنند و پیش‌بینی مکان‌ها بر اساس این مقادیر صورت می‌گیرد. از مهم‌ترین نقاط قوت این روش ایده شهودی آن است که باعث می‌شود به آسانی قابل درک باشد و زمانی که اطلاعات کافی موجود باشد، توانایی انجام پیش‌بینی‌های دقیقی را دارا است. در کنار آن روش‌های مدل‌مبنا مدل‌های آموزشی را جهت یادگیری از الگوهای پیچیده با استفاده از داده‌های آموزشی توسعه می‌دهند. این مدل‌ها می‌توانند از روش‌های مختلف مانند روش‌های احتمالاتی و یا هوش مصنوعی به وجود بیایند. در اینجا پالایش مشارکتی حافظه‌مبنا، جهت توصیه مکان در سیستم توصیه‌گر مورد استفاده قرار گرفته است.

۳-۳-۱- محاسبه شباهت بین کاربران

پس از محاسبه امتیازات داده‌شده به مکان‌ها از طرف کاربران، با استفاده از روش پیرسون که یک نوع تابع ابتکاری برای محاسبه وابستگی بین متغیرها است، شباهت بین کاربران را به صورت خطی محاسبه می‌کنیم. این تابع برای اولین بار در سال ۱۹۹۴ توسط رسنیک مطرح شد. روش پیرسون، شباهت بین کاربرهای u و v را که با

$sim(u, v)$ بیان می‌شود، توسط معادله ۵ محاسبه می‌کند [۱۶].

$$sim(u, v) = D_1(u_i, u_j) \times \frac{\sum_{i \in I} (r_{u,i} - \bar{r}_u)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i \in I} (r_{u,i} - \bar{r}_u)^2} \sqrt{\sum_{i \in I} (r_{v,i} - \bar{r}_v)^2}} \quad (5)$$

در رابطه‌ی فوق $i \in I$ نشان‌دهنده موارد اطلاعاتی است که هر دو کاربر u و v به آن امتیاز داده و $r_{u,i}$ امتیاز کاربر u به مورد اطلاعاتی i است و $\bar{r}_u$ میانگین امتیازاتی است که کاربر u به آن موارد اطلاعاتی داده است. نتیجه این مرحله تشکیل مجموعه‌ای از مکان‌ها رتبه‌بندی شده به نام مجموعه همسایگان است. این مجموعه بر اساس میزان شباهت بین کاربران به کاربر هدف ایجاد شده است.

۳-۳-۲- تابع تجمع امتیازات در پالایش مشارکتی حافظه‌مبنا

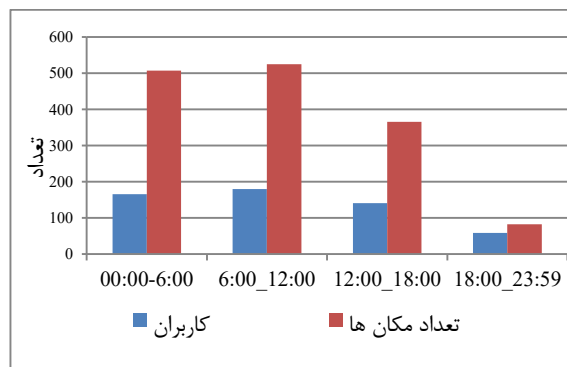
با فرض داشتن امتیازات کاربران به مکان‌ها و ماتریس کاربر-مکان، علاقه کاربر a به مکان z از دو رابطه‌ی ۶ و ۷ که از الگوریتم عمومی پالایش مشارکتی کاربر مبنا استخراج شده، برآورد می‌شود. در این رابطه، شباهت بین کاربر a و دیگر کاربران $sim(u_a, u_i)$ ، به‌عنوان وزن استفاده می‌شود و معنای آن این است که هرچقدر شباهت بین کاربر u_i با کاربر هدف بیشتر باشد، میزان امتیاز کاربر u_i در پیش‌بینی امتیاز مکان z وزن بیشتری دارد. امتیاز مکان z که تا به حال مورد بازدید کاربر a قرار نگرفته، توسط معادله‌های ۶ و ۷ پیش‌بینی می‌شود [۱۷]:

$$P_{u_a, l_j} = \bar{r}_a + \frac{\sum_{u_i \in U_{l_j}} (r_{i,j} - \bar{r}_i) \cdot sim(u_a, u_i)}{\sum_{u_i \in U_{l_j}} sim(u_a, u_i)} \quad (6)$$

$$\bar{r}_a = \frac{\sum_{j \in J_a} r_{a,j}}{|J_a|}, J_a = \{j \in J | r_{a,j} \neq \emptyset\} \quad (7)$$

در این رابطه، ابتدا مقیاس امتیازدهی هر کاربر با محاسبه میانگین امتیازات او ($\bar{r}_a$) لحاظ می‌شود. J_a بیانگر مجموعه مکان‌هایی است که در نزدیکی کاربر هدف قرار گرفته است، U_{l_j} بیانگر مجموعه کاربرانی است که از مکان z بازدید کرده‌اند، $\bar{r}_a$ و $\bar{r}_i$ میانگین امتیازاتی است که از طرف کاربر a و i به مکان‌های مورد بازدید داده شده است، $r_{i,j}$ نیز امتیاز u_i به مکان z است. در نهایت، علاقه

می‌کنند. بر اساس پایگاه داده موجود یک مجموعه سه‌تایی (u_i, l_j, n_{ij}) را استخراج کردیم که n_{ij} نشان‌دهنده تعداد دفعات بازدید کاربر i از مکان j در زمان t_{ij} است. همان‌طور که در شکل ۶ نشان داده‌شده، با دسته‌بندی داده‌ها بر اساس زمان بازدید، توانستیم در چهار بازه زمانی مختلف از یک روز، توصیه کنیم.



شکل ۶- نمودار توزیع داده‌های Gowalla در چهار بازه زمانی

با توجه به استراتژی حفظ حریم خصوصی در سایت‌های شبکه‌های اجتماعی، کشف روابط اجتماعی بین کاربران سخت است. بنابراین، در اینجا به پیشنهاد مکان با تکیه بر داده‌های نقطه-مکان و بدون هیچ اطلاعات اضافی در مورد روابط اجتماعی می‌پردازیم.

کلید پیاده‌سازی‌ها با زبان C# انجام شده است و داده‌ها در پایگاه داده SQL SERVER ذخیره می‌شوند. در نتیجه پیاده‌سازی سرویس توصیه‌گر، مکان‌هایی با توجه به موقعیت فعلی کاربر و ترجیحات شخصی او پیشنهاد می‌شود. برای مثال مطابق شکل ۷، سیستم برای کاربر A که در ساعت ۷ غروب وارد خیابان Haidian پکن چین شده، چهار مکان برای بازدید در نزدیکی او پیشنهاد داده می‌کند.

نام مکان (نوع کاربری)	فاصله تا موقعیت کاربر (متر)	امتیاز نهایی
Quanjude in Hai Dian Da jie (Chinese Restaurant)	۶۳۰ متر	۱
Danling Street (Chinese Restaurant)	۱۷۴۰ متر	۰.۸۶
Gate of Heavenly Peace (Historic Place)	۲۶۸۰ متر	۰.۷۹
Fangshan Restaurant, East Gate (Restaurant)	۴۳۹۰ متر	۰.۵۷

کاربر هدف به یک مکان، با جمع وزن‌دار امتیازات داده‌شده از طرف کاربران مشابه به آن مکان، پیش‌بینی می‌شود. سپس مجموعه‌ای از N مکان اول که دارای بالاترین امتیاز هستند، به کاربر پیشنهاد می‌شود. ایده اصلی الگوریتم‌های توصیه‌گر مانند پالایش مشارکتی استفاده از علایق مشترک بین کاربران مشابه برای پیش‌بینی موارد از دست‌رفته است. بنابراین، به جای اسکن کردن مکان‌های مورد بازدید تمامی کاربران برای محاسبه وزن شباهت، وزن شباهت بین کاربران را بر اساس فاصله بین آن‌ها مدل می‌کنیم و به جای اسکن تمامی مکان‌ها، تنها از مکان‌هایی که در فاصله مشخصی از موقعیت فعلی کاربر هستند، استفاده می‌کنیم. در نهایت مکان‌هایی که دارای وزن بالاتری هستند و تا به حال مورد بازدید کاربر قرار نگرفته‌اند، به کاربر هدف پیشنهاد می‌شود.

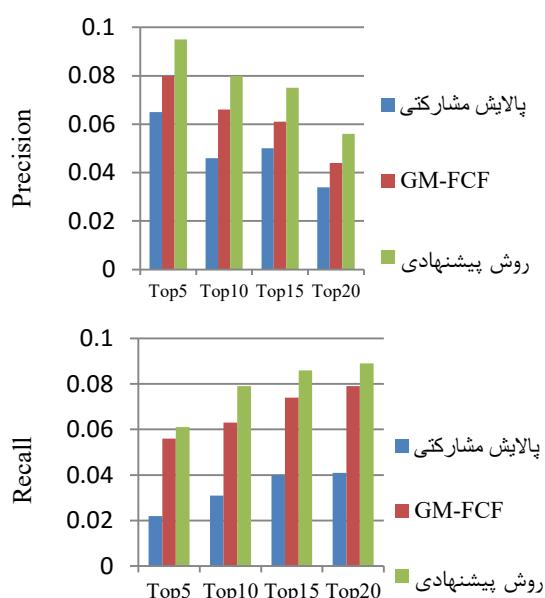
۴- پیاده‌سازی و ارزیابی

به علت عدم وجود داده‌های مربوط به شبکه‌های اجتماعی مکان‌مبنا در ایران، برای انجام آزمایش‌ها و تجزیه و تحلیل داده‌های نقطه-مکان شبکه Gowalla را که به‌عنوان یکی از محبوب‌ترین شبکه‌های اجتماعی مکان‌مبنا در سال‌های اخیر شناخته شده است، از اکتبر ۲۰۱۱ تا نوامبر ۲۰۱۱ جمع‌آوری کردیم. در مجموع پروفایل ۳۳۰ کاربر را در ۱۵۵۴ مکان واقع در شهر پکن چین مورد بررسی قرار گرفت. با صرف‌نظر از کاربرانی که تنها یک‌بار ثبت موقعیت کرده‌اند و مکان‌هایی که فقط یک‌بار مورد بازدید قرار گرفته‌اند، تعداد کل مشاهدات کاربر-مکان به ۵۲۱۶ نقطه می‌رسد. بر اساس تجزیه و تحلیل داده‌های مورد مطالعه، مشاهده می‌کنیم که اکثر کاربران در یکسری مناطق جغرافیایی خاص ثبت موقعیت



شکل ۷- مکان‌های توصیه‌شده توسط سیستم توصیه‌گر توسعه یافته

نشان می‌دهد. برد $Precision$ و $Recall$ بازه [۱،۰] است و هر چه مقدار $Precision$ به ۱ نزدیک‌تر و مقدار $Recall$ به ۰ نزدیک‌تر باشد، حاکی از عملکرد بهتری از سیستم است. شکل ۸ نتایج $Precision$ و $Recall$ را برای $N=5,10,15,20$ برای ۱۰ کاربر سیستم به‌طور تصادفی نشان می‌دهد و N بیانگر تعداد مکان‌های توصیه‌شده است.



شکل ۸- نتایج $Precision$ و $Recall$ برای ارزیابی عملکرد سیستم توصیه‌گر توسعه‌یافته

با توجه به شکل ۸ می‌توان دریافت که الگوریتم پیشنهادی نسبت به پالایش مشارکتی که ماتریس کاربر-مکان را بر اساس تعداد بازدید تولید می‌کند، از دقت بالاتری برخوردار است. علاوه بر این محدود نکردن فضای جست‌وجو به دوستان، انعطاف‌پذیری بالایی را در توصیه دامنه وسیعی از مکان‌ها به وجود می‌آورد و دارای صحت بالاتری نسبت به روش GM-FCF است. همچنین در نظر گرفتن پارامتر زمان و موقعیت فعلی کاربر نیز از ویژگی‌های اصلی روش پیشنهادی محسوب می‌شود. بنا بر آزمون صورت گرفته دریافتیم احراز شرایط مکانی و زمانی در سیستم توصیه‌گر، همیشه باعث بهبود عملکرد سیستم خواهد بود.

۵- نتیجه‌گیری

این مقاله یک سیستم توصیه‌گر زمانمند، مکان آگاه و ترجیح آگاه ارائه می‌دهد که به پیشنهاد مکان در اطراف

با توجه به جدول موجود در شکل ۷ می‌توان دریافت، کاربری رستوران از موارد مورد علاقه کاربر است و سیستم توصیه‌گر با در نظر گرفتن این واقعیت و با توجه به شرایط زمانی، در کنار تنوع مختلفی از مکان‌ها مانند مکان تاریخی، بیشتر مکان‌هایی از نوع رستوران توصیه کرده است. همچنین حداکثر فواصل بین این مکان‌ها تا موقعیت کاربر هم‌چنین حد اکثر فواصل بین این مکان‌ها تا موقعیت کاربر D_2 تعیین شده بود.

برای ارزیابی عملکرد روش ارائه‌شده در توصیه مکان‌ها، نتایج به‌دست‌آمده را با نتایج حاصل از دو تکنیک پالایش مشارکتی کاربرمنا و GM-FCF^۱ [۱۸] مقایسه می‌کنیم. در تکنیک پالایش مشارکتی، تعداد بازدید هر کاربر از هر مکان تلویحاً به‌عنوان امتیاز داده‌شده از طرف کاربر به مکان در نظر گرفته می‌شود. اما مشکل اساسی این روش این است که تعداد بازدید از یک مکان نمی‌تواند به‌طور صریح میزان علاقه کاربر به آن مکان را نشان دهد. درواقع پیش‌بینی ترجیحات مکانی کاربران نباید تنها بر این اساس باشد که وی چه تعداد دفعات از یک مکان بازدید کرده است، بلکه باید اهمیت هر مکان را برای کاربر نسبت به تمامی مکان‌های بازدید شده توسط آن کاربر بسنجد. GM-FCF با محاسبه مشابهت بین کاربر هدف با دوستان کاربر، به توصیه مکان‌ها می‌پردازد. برای مقایسه عملکرد روش پیشنهادی با دو روش مذکور، از شاخص صحت رتبه که رایج‌ترین روش ارزیابی سیستم‌های توصیه‌گر به شمار می‌آید، استفاده می‌کنیم. شاخص صحت، ترتیب ارقام پیشنهادی توسط سیستم توصیه‌گر را موردبررسی قرار می‌دهد که دو معیار $Precision$ و $Recall$ معمول‌ترین معیارهای بررسی خروجی سیستم‌های توصیه‌گر از این منظر هستند. فرمول‌های ۸ تعاریف مذکور را نشان می‌دهند.

$$Precision = \frac{\alpha}{\alpha + \omega} \quad Recall = \frac{\alpha}{\alpha + \beta} \quad (8)$$

α تعداد اقلامی است که در جمع اقلام برتر توصیه‌شده به کاربر، موردتوجه و موردپسند وی قرارگرفته‌اند. β نیز معرف تعداد اقلامی است که در اقلام باقی‌مانده، نظر کاربر را جلب کرده‌اند و ω تعداد اقلامی که در جمع اقلام برتر توصیه‌شده به کاربر مورداستفاده وی قرار نگرفته‌اند، را

^۱ Geo-measured friend-based collaborative Filtering

در زمان‌های مختلف از بدون توجه به موقعیت کاربر یک نتیجه را دربر خواهد داشت. همچنین مزیت روش پیشنهادی نسبت به پالایش مشارکتی، محاسبه امتیازات کاربران به مکان‌ها توسط یک روش داده‌کاوی است که میزان علاقه کاربران به مکان‌ها را صرفاً بر اساس تعداد بازدید برآورد نمی‌کند، بلکه به‌صورت نسبی و وزن‌دار توصیف می‌کند. لذا محاسبه میزان علاقه کاربران به مکان‌ها، محدود نکردن فضای جست‌وجو به دوستان، در نظر گرفتن پارامتر زمان و فاصله از موقعیت فعلی کاربر، از ویژگی‌های اصلی سیستم توسعه‌یافته است.

موقعیت جغرافیایی کاربر بر اساس ترجیحات شخصی کاربر و موقعیت فعلی کاربر در بازه‌های زمانی مختلف می‌پردازد. پیاده‌سازی و ارزیابی بر روی مجموعه‌ای از داده‌های واقعی جمع‌آوری شده از شبکه اجتماعی Gowalla انجام شده است. در روش پیشنهادی برای توصیه مکان، با مرتب‌سازی امتیازات بدست‌آمده برای هر مکان توسط کاربران مشابه به کاربر هدف، N مکان با بالاترین امتیاز به او پیشنهاد می‌شود.

برای ارزیابی توانایی‌های سیستم به مقایسه‌ی آن با دو روش پالایش مشارکتی و GM-FCF پرداخته شد. نتایج توصیه حاصل از هر دو روش زمانمند و مکان‌آگاه نیست و

مراجع

- [1] Boyd, D. M. (2004). Friendster and publicly articulated social networking.
- [2] Counts, S., & Fisher, K. E. (2008, January). Mobile social networking: An information grounds perspective. In Hawaii International Conference on System Sciences, Proceedings of the 41st Annual (pp. 153-153). IEEE.
- [3] Ellison, N. B., & Boyd, D. (2007). Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication, 13(1), 210-230.
- [4] Felfernig, A., Friedrich, G., & Schmidt-Thieme, L. (2007). Recommender systems. IEEE Intelligent Systems, 22(3), 18.
- [5] Burke, R. (2002). Hybrid recommender systems: Survey and experiments. User modeling and user-adapted interaction, 12(4), 331-370.
- [6] Bao, J., Zheng, Y., Wilkie, D., & Mokbel, M. F. (2014). A survey on recommendations in location-based social networks. Submitted to Geoinformatica.
- [7] Park, M. H., Hong, J. H., and Cho, S. B. (2007) Location-based recommendation system using Bayesian user's preference model in mobile devices, Ubiquitous Intelligence and Computing, 1130-1139.
- [8] Zheng, Y., and Zhou, X. (2011) Computing ith spatial trajectories, Springer-Verlag New York Inc.
- [9] YE, M., Y IN, P., L EE, W., AND LEE, D. (2011c). Exploiting geographical influence for collaborative point-of-interest recommendation. In Proceedings of the 34th international ACM SIGIR conference on Research and development in Information. ACM, 325-334.
- [10] Zheng, Y., Zhang, L., Ma, Z., Xie, X., and Ma, W. Y. (2011) Recommending friends and locations based on individual location history, ACM Transactions on the Web (TWEB) 5, 5.
- [11] 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. 364-373). ACM.
- [12] TOBLER, W. (1970). A computer movie simulating urban growth in the detroit region. Economic geography 46, 234-240.
- [13] Balabanović, M., & Shoham, Y. (1997). Fab: content-based, collaborative recommendation. Communications of the ACM, 40(3), 66-72.

- [14] Salton, G. (1989) Automatic Text Processing: The Transformation, Analysis, and Retrieval of, Addison-Wesley.
- [15] Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, 22(1), 5-53.
- [16] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994, October). GroupLens: an open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work* (pp. 175-186). ACM.
- [17] Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, 17(6), 734-749.
- [18] Yu, X., Pan, A., Tang, L. A., Li, Z., and Han, J. (2011) Geo-Friends Recommendation in GPS-Based Cyber-Physical Social Network.