Archive of SID

نشریه علمی علوم و فنون نقشه برداری، دوره نهم، شماره ۲، آذر ماه ۱۳۹۸ – مقاله ترویجی

مرور روش‌های شباهت‌یابی در شبکه‌های اجتماعی مکان‌مبنا

فهیمه کرمی[۱]، محمدرضا ملک[*۲]

[۱]دانشجوی کارشناسی ارشد سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی
fahime.karami@email.kntu.ac.ir

[۲]دانشیار دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین‌طوسی
mrmalek@kntu.ac.ir

(تاریخ دریافت بهمن ۱۳۹۷، تاریخ تصویب مهر ۱۳۹۸)

چکیده

شبکه‌های مخابرات بی‌سیم و دستگاه‌های مجهز به فناوری‌های تعیین موقعیت، امکان اشتراک‌گذاری موقعیت کاربران به همراه سایر موارد را برای کاربران شبکه‌های اجتماعی فراهم ساخته‌اند. از این رو شاخصه‌های مکانی را می‌توان به شبکه‌های اجتماعی افزود و از آن به‌عنوان داده مکان‌مبنا استفاده برد. یکی از مباحث پراهمیت در زمینه شبکه‌های اجتماعی مکان‌مبنا، تعیین شباهت میان کاربران است. در این تحقیق، برمبنای تحقیقات صورت پذیرفته راجع به شبکه‌های اجتماعی مکان‌مبنا، سعی شده مرور کاملی بر معیارهای تعیین شباهت کاربران انجام پذیرد. بعلاوه تلاش شده‌است که با طبقه‌بندی حوزه‌های تحقیقاتی موجود، وضعیت روشن‌تری در اختیار کارشناسان و علاقه‌مندان این حوزه گذاشته شود. با توجه به نتایج تحقیق حاضر، می‌توان عنوان کرد که تحقیقات انجام شده در این زمینه هنوز در مراحل ابتدایی بوده و برخی معیارها همانند معیارهای معنایی و محتوایی بیشتر از سایرین نیازمند تحقیقات می‌باشند.

واژگان کلیدی: شبکه‌های اجتماعی مکان‌مبنا، تعیین شباهت کابران، شباهت معنایی

* نویسنده رابط

www.SID.ir

Archive of SID

۱- مقدمه

پیشرفت‌های اخیر در زمینه‌ی فناوری‌های موقعیت مکانی، مانند سیستم تعیین موقعیت جهانی[۱] و شبکه‌ی بی‌سیم محلی[۲]، امکان افزودن بعد مکان را در شبکه‌های برخط[۳] موجود فراهم آورده است. از این‌رو شبکه‌های اجتماعی دارای شاخص مکانی شده و به عنوان داده‌های مکان‌مبنا بکار می‌روند. این شبکه‌ها منبعی غنی برای تعیین الگوهای رفتار مکانی، زمانی و معنایی کاربران در مکان‌های مختلف هستند که بازتابی از زندگی واقعی میان مردم را فراهم می‌آورند.

رشد بالای تعداد کاربران شبکه‌های اجتماعی مکان‌مبنا، باعث بالا رفتن غنای اطلاعاتی این شبکه‌ها شده‌است. به طوری که امروزه این شبکه‌ها، بعنوان منبعی از داده‌های مکان‌مبنا مورد استفاده قرارگیرند. غنای بالای اطلاعات مکانی و اجتماعی این شبکه‌ها کاربردهای گسترده‌ای نظیر استخراج اطلاعات مکانی، تعیین مکان‌های شاخص، اظهار نظر درمورد مکان‌های خاص، کشف شباهت میان افراد، توصیه‌گری و برنامه‌ریزی سفر را به‌همراه داشته‌است [۱-۴]. درواقع این شبکه‌ها به‌عنوان پل ارتباطی میان جهان واقعی و شبکه‌های اجتماعی برخط عمل می‌کنند.

استفاده از داده‌های شبکه‌های اجتماعی مکان‌مبنا با دو چالش بزرگ روبه‌رو است که عبارتند از: ۱) مدل‌سازی موقعیت‌های افراد از تاریخچه شخصی آن‌ها و ۲) شباهت-یابی کاربران، معرفی افراد شبیه به‌هم و برقراری ارتباط میان کاربران مشابه در شبکه‌های اجتماعی.

شباهت‌یابی کاربران، نشان‌دهنده‌ی الگوهای رفتاری مردم است که همواره نقش کلیدی در برنامه‌های شهری نظیر برنامه‌ریزی حمل‌ونقل، پیش‌بینی ترافیک و سرویس‌های توصیه‌گر دارد [۱, ۲, ۵, ۶].

معیارهای شباهت‌یابی به طور کلی به سه دسته مکانی، زمانی و معنایی قابل تقسیم هستند. معیارهای مکانی مختصات جغرافیایی مکان کاربر مثل مهاجرت جمعیت را درنظر می‌گیرند و کاربرانی را که به مکان‌های یکسانی رفته‌باشند مشابه در نظر می‌گیرد. درحالیکه معیارهای زمانی، ارتباط وارد یا خارج شدن از مکان‌ها مثل مهاجرت پرندگان، را بیان می‌کند. در این معیار باتوجه به زمان، کاربران را دسته‌بندی می‌کنند. همچنین معیار معنایی، علت مراجعه به مکان را بیان می‌کند، مثلا فرد برای صرف غذا به رستوران می‌رود.

در این تحقیق برمبنای تحقیقات صورت پذیرفته راجع به شبکه‌های اجتماعی مکان‌مبنا، سعی شده مرور کاملی درمورد معیارهای تعیین شباهت کاربران انجام پذیرد. با توجه به نتایج تحقیق حاضر می‌توان عنوان کرد که تحقیقات انجام شده در این زمینه هنوز در مراحل ابتدایی بوده و برخی معیارها همانند داده‌های معنایی و محتوایی بیشتر از سایرین نیازمند تحقیقات می‌باشند. افرادی که به کاربری‌های مشابه در مکان‌های نزدیک و در زمان‌های مشابه رفته‌باشند، شباهت و علایق مشترک بیشتری خواهند داشت. از این‌رو، اخیرا تحقیقاتی در زمینه‌ی تلفیق معیارهای مکانی، معنایی و زمانی ارائه شده‌است [۷-۱۳] که در ادامه به توضیح آن‌ها خواهیم پرداخت.

۲- شبکه‌های اجتماعی مکان‌مبنا

شبکه اجتماعی[۴]، ساختاری اجتماعی متشکل از ارتباطاتی مانند روابط دوستانه، علاقمندی‌های عمومی و انتقال دانش میان کاربران مختلف است. بطور کلی یک شبکه اجتماعی، بازتابی از زندگی واقعی میان مردم، از طریق سکو-های[۵] آنلاین مانند سایت‌ها را فراهم می‌کند [۱۴].

نتایج حاصل از شبکه‌های اجتماعی، می‌توانند در شناخت الگوهای رفتاری، تردد افراد و ارائه خدمات بهتر به کاربران مورد استفاده قرارگیرند. در حوزه گردشگری، با بررسی حرکت گردشگران حین بازدید از مکان‌های تفریحی یا تاریخی و تجزیه‌وتحلیل مسیرهای عبوری و مکان‌های مشاهده شده، نه تنها امکان خدمات‌رسانی بهتر به بازدیدکنندگان میسر می‌شود، بلکه گردشگران جدید می‌توانند از تجربیات و مسیرهای موردعلاقه افراد قبلی بهره‌ببرند [۱۵, ۱۶]. به این سرویس‌ها، خدمات مکان‌مبنا[۶] گفته‌می‌شود که از داده‌های جغرافیایی موجود در شبکه-های اجتماعی مکان‌مبنا استفاده می‌کنند. شبکه اجتماعی را میتوان وجه اشتراک اینترنت، موبایل و GIS[۷] تلقی‌کرد.

۱ Global Positioning System-GPS
۲ Wireless Fidelity-WiFi
۳ online

٤ Location Based Social Network-LBSN
٥ platform
٦ Location Based Service-LBS
٧ geographic information system

www.SID.ir

Archive of SID

این سرویس‌ها به سه دسته‌ی کلی شناسایی افراد، شناسایی مکان بر اساس افراد بازدیدکننده‌ی آن‌ها و کشف رویداد از رسانه‌های اجتماعی تقسیم می‌شوند [۶, ۱۷].

دسته‌ی اول، شناسایی افراد بر اساس موقعیت مکانی آن‌ها است که خود به سه بخش تقسیم می‌شود. بخش اول، افراد مشابه را بر اساس رفتار و سوابق مکان‌های بازدید شده تعیین می‌کند، درواقع افرادی که از مکان‌های یکسانی بازدید داشته‌اند را مشابه در نظر می‌گیرد [۱۸-۲۱]. بخش دوم، کشف کارشناسان محلی و افراد بومی محل است [۶, ۲۲]. با استفاده از اطلاعات مکانی کاربران، افرادی که اطلاعات بیشتری را در مورد یک منطقه داشته و یا تجربه‌ی مکانی بالایی دارند شناسایی می‌شوند. اطلاعات حاصل از تجربیات مکانی این افراد مثلا مکان‌هایی که بازدید کرده‌اند، برای یک سیستم پیشنهادگر مکانی بسیار ارزشمند می‌باشد. بخش سوم، شناسایی جامعه و گروه‌بندی افراد بر اساس سلایق آن‌ها برای برنامه‌ریزی‌های فرهنگی و شهری می‌باشد [۶, ۲۲-۲۴].

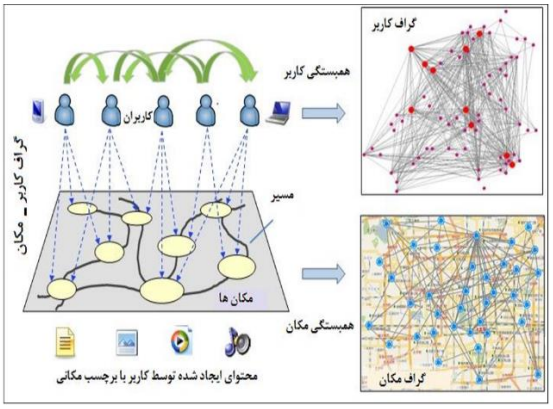
دسته‌ی دوم، شناسایی مکان بر اساس افراد بازدید کننده‌ی آن‌ها می‌باشد که خود شامل چند بخش است. در بخش اول، مکان‌های مورد علاقه گردشگری استخراج می‌شود [۶, ۲۲, ۲۵, ۲۶]. در این حوزه، Taha سیستم همراه Alexplore را در شهر اسکندریه مصر برای بارزسازی پتانسیل شبکه اجتماعی در زمینه گردشگری و برنامه‌ریزی شهری تشریح کرد [۱]. بخش دوم، برنامه‌ریزی سفر می‌باشد. سرویس‌های توصیه مکان و توصیه سفر امروزه بسیار مورد توجه قرار گرفته‌است [۲, ۲۷-۳۰]. با استفاده از داده‌های مکانی به اشتراک گذاشته‌شده در شبکه‌های اجتماعی مکان‌مبنا، سرویس‌های مختلفی برای شناسایی مکان‌های پر بازدید در محدوده مکانی کاربر وجود دارد [۳۱]. درواقع در این سرویس‌ها امکان مشاهده‌ی سابقه‌ی مکانی افراد با سلایق و ویژگی‌های مشابه به کاربر و نیز نظرات آن‌ها وجود دارد. لذا افراد می‌توانند از تجربیات یکدیگر استفاده کنند.

بخش سوم برای توصیه فعالیت‌های مکانی شامل بهترین مکان برای فعالیتی مشخص یا فعالیت‌های مناسب برای مکان خاص انجام می‌شود [۳۲].

دسته‌ی سوم، کشف رویداد از رسانه‌های اجتماعی مکان‌مبنا در اثر تغییرات سریع مثل افزایش اطلاعات از یک مکان خاص در زمانی مشخص است [۳۳-۳۷].

شبکه‌های اجتماعی مکان‌مبنا، تنها سرویس‌های جدید را ایجاد نمی‌کنند، بلکه داده‌های بیشتری برای هر چه بهتر شدن زندگی انسان‌ها را مهیا می‌سازند [۳۸].

شبکه اجتماعی، مجموعه داده‌های ناهمگن[۱] و چند رابطه ای[۲] است که توسط یک گراف نمایش داده می‌شود. ساختار شبکه‌های اجتماعی بر پایه‌ی گراف است، گراف‌هایی که از این شبکه‌ها می‌توان استخراج کرد به سه دسته کاربر با کاربر، کاربر با مکان و مکان با مکان تقسیم می‌شوند [۳۹]. گراف‌های مستخرج از شبکه‌های اجتماعی، حاوی اطلاعات زیادی است که از آن‌ها در زمینه‌های متنوعی نظیر مکانی، زمانی و معنایی می‌توان استفاده‌کرد به گونه‌ای که بیان کننده‌ی رفتار کاربران در جهان واقعی می‌باشند. در این گراف‌ها، همانطور که در شکل ۱ نمایش داده شده، نودها کابر یا مکان و یال‌ها ارتباطات میان نودها است. از این گراف‌ها می‌توان اطلاعات متنوعی نظیر مکان‌های شاخص و یا افراد مهم هر گروه را استخراج کرد. برای مثال از گراف کاربر با کاربر همبستگی کاربران، مثل دوستی‌ها و رابطه خویشاوندی و از گراف کاربر با مکان، مکان‌های مورد علاقه‌ی کاربران و از گراف مکان با مکان، همبستگی مکان‌ها قابل استخراج است.



شکل۱- انواع گراف‌های موجود در شبکه‌های اجتماعی مکان‌مبنا [۳۹]

۳- معیارهای شباهت‌یابی

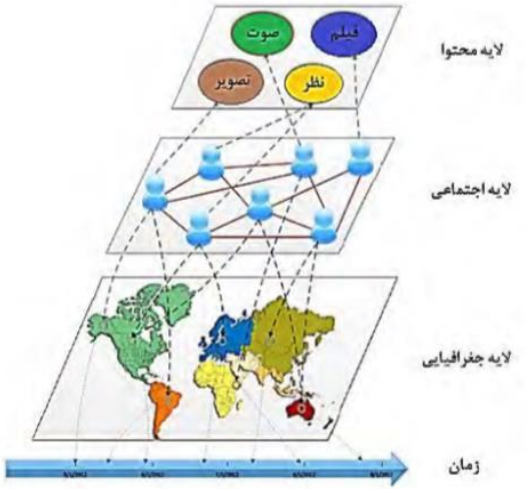
داده‌های شبکه اجتماعی حجم زیادی از اطلاعات مربوط به کاربران را شامل می‌باشند. این اطلاعات، توصیف کننده‌ی رفتار و زندگی روزمره‌ی کاربران آن‌هاست. شبکه‌های اجتماعی مکان‌مبنا دارای چارچوب لایه‌ای مطابق شکل۲ می‌باشند. لایه‌ی جغرافیایی شامل تاریخچه‌ای از

۱ heterogeneous
۲ Multirelational

www.SID.ir

Archive of SID

ورودی‌های کاربر، لایه‌ی اجتماعی دربرگیرنده‌ی اطلاعات مربوط به دوستی‌های اجتماعی و لایه‌ی محتوا دربردارنده‌ی نظرات و بازخورد کاربران درمورد مکان‌های مختلف است. همه‌ی این سه لایه به‌طور مشترک با لایه اطلاعات زمانی مربوط به ورودی‌های کاربر در ارتباط هستند [۴۰].



شکل۲- لایه‌های اطلاعاتی موجود در شبکه‌های اجتماعی مکان‌مبنا [۴۰]

اهمیت اندازه‌گیری و تجزیه‌وتحلیل داده‌های شبکه-های اجتماعی را می‌توان در کشف الگوهای حرکتی و رفتاری، تقسیم‌بندی داده‌های در حال حرکت، پیش‌بینی مکان بعدی کاربر، برنامه ریزی‌های شهری و سرویس‌های توصیه‌گری دانست [۳].

امروزه سرویس‌های توصیه مکان و توصیه سفر بسیار مورد توجه قرار گرفته‌است [۲]. افراد در مناطق جدید علاقمندند که مسیرها و مکان‌های جذاب را بشناسند. مکان‌های جذاب علاوه بر موزه‌ها و آثار تاریخی، رستوران-ها، مراکز خرید و موارد مشابه دیگر را نیز شامل می‌شوند [۴۱]. با استفاده از داده‌های مکانی به اشتراک گذاشته‌شده در شبکه‌های اجتماعی مکان‌مبنا، سرویس‌های مختلفی برای شناسایی مکان‌های پر بازدید در محدوده مکانی کاربر وجود دارد [۳۱]. درواقع در این سرویس‌ها، امکان مشاهده‌ی سابقه‌ی مکانی افراد با سلایق و ویژگی‌های مشابه به کاربر و نظرات آن‌ها وجود دارد. لذا افراد می-توانند از تجربیات یکدیگر استفاده کنند. در این سیستم‌ها، با مقایسه داده‌های کاربر با داده‌های مستخرج از بقیه کاربران و همچنین استفاده از معیارهای شباهت‌یابی، مکان‌های جذاب و پربازدید که به سلایق کاربر نزدیک است را به کاربر پیشنهاد می‌کنند [۴].

موضوع بحث‌برانگیز در این مورد، تعیین شباهت میان کاربران و شناخت معیارهای شباهت‌یابی است. مفهوم کلی از تعیین شباهت اینگونه است که هرچه دو شیء مشترکات بیشتری داشته باشند، شبیه‌ترند و هرچه تفاوت بیشتری با یکدیگر داشته باشند، کمتر به‌هم شبیه هستند [۴۲].

به‌دلیل وجود تعاریف متفاوتی از شباهت و توجه به جنبه‌های مختلف، روش‌های زیادی برای اندازه‌گیری شباهت وجود دارد که می‌توان آن‌ها را به سه دسته‌ی روش‌های مبتنی بر معیارهای مکانی، زمانی و معنایی تقسیم کرد. از این‌رو در ادامه به معرفی هر معیار و روش-های مربوطه می‌پردازیم.

۳-۱- معیار مکانی

طبق آنچه که ذکرگردید، از شبکه‌های اجتماعی مکان‌مبنا بعنوان منبعی از داده‌های مکانی استفاده می-شود. خط‌سیر افراد یکی از داده‌های استخراجی از شبکه-های اجتماعی مکان‌مبنا است. اطلاعات کاربر بصورت رشته‌ای از نقاط زمانی و مکانی ذخیره و مدل‌سازی می-شوند که اصطلاحا خط‌سیر نامیده‌می‌شوند [۴۳].

با توجه به قانون اول جغرافیا "همه چیز به همدیگر وابسته هستند، اماچیزهایی که در فواصل نزدیک به‌هم قرار دارند، نسبت به فواصل دورتر بیشتر باهم در ارتباط-اند[۴۴]." فاصله جغرافیایی در تعیین شباهت افراد نقش موثری دارد. با بررسی فاصله‌ی جغرافیایی افراد و ارتباط اجتماعی میان آن‌ها، محققین به این نتیجه رسیده‌اند که در بیشتر مواقع افراد با افرادی که از نظر مکانی در مجاورتشان هستند، راحت‌تر ارتباط برقرار می‌کنند [۱۸, ۴۵].

در تحقیق [۴۶]، با آنالیز داده‌ی چهار شبکه‌ی اجتماعی مکان‌مبنای مختلف، وابستگی ساختار اجتماعی کاربران به وضعیت جغرافیایی آن‌ها بررسی شد. این تحقیق نشان داد که بخش قابل توجه‌ای از ارتباطات موجود در شبکه‌های اجتماعی، بین کسانی وجود دارد که در دنیای واقعی از نظر جغرافیایی به یکدیگر نزدیک هستند. در این زمینه Liben-Nowell و همکاران، رابطه‌ی میان دوستی افراد و فاصله‌ی جغرافیایی بین آن‌ها را بررسی کردند [۴۷]. همچنین Lubke و همکاران در تحقیق خود، سیستم MobilisGroups را با هدف ایجاد گروه‌هایی در شبکه اجتماعی مبتنی بر مکان کاربر ارائه-کردند [۴۸]. برای نشان دادن تاثیر فاصله‌ی جغرافیایی بر

www.SID.ir

Archive of SID

دوستی‌ها Scellato در تحقیقی چهار ماهه با استفاده از داده‌های شبکه Gowalla نشان داد که سی درصد از ارتباطات جدید می‌تواند میان کاربران بازدیدکننده از مکانی یکسان برقرار شود [۱۸].

از ساده‌ترین راه‌های تعیین شباهت مکانی استفاده از فاصله p-norm است.

$$\| x \|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \quad (۱)$$

در رابطه‌ی۱ به ازای p=2 فاصله اقلیدسی و p=1 منهتن و وقتی p بی‌نهایت، بیشترین فاصله است.

با استفاده از فاصله‌ی اقلیدسی کوتاه‌ترین فاصله بین دو نقطه محاسبه می‌شود. اگر به جای مربع فاصله بین مختصات، از قدر مطلق تفاضل بین مختصات نقاط استفاده شود، تابع فاصله را منهتن می‌نامند. البته این فاصله گاهی به نام "فاصله تاکسی" یا "بلوک شهری" نیز خوانده می‌شود.

اندازه‌گیری فاصله اقلیدسی مزایای مختلفی دارد. فهم آن آسان، محاسبه‌ی آن ساده و برای حل مشکلات دیگر مانند شاخص‌گذاری و خوشه‌بندی سری‌های زمانی قابل استفاده است. هرچند معایبی نیز دارد که مهمترین آن‌ها عدم کاربرد برای سری‌های با خط‌مبنای متفاوت است. با استفاده از نرمال‌سازی می‌توان معایب p-norm را در اندازه‌گیری تشابه برطرف کرد. در رابطه‌ی۲ اگر μ(X) میانگین و var(X) واریانس دنباله باشد دنباله X را با دنباله نرمال X' جایگزین می‌کنیم:

$$X' = (x_i - \mu(X))/\mathrm{var}(X) \quad (۲)$$

بعد از نرمال‌سازی، هردو دارای یک خط‌مبنا شده چون میانگین هردو با نرمال‌سازی یکسان می‌شود. به علاوه، به علت اینکه واریانس‌ها یکسان گردیده اند، دامنه نیز یکسان خواهد بود. البته فرایند نرمال‌سازی هم معایب خود را دارد، مثلا به تاخیر و تقدم فاز در زمان بسیار حساس است.

علاوه بر روش p-norm، برای شباهت مکانی از روش شبه‌فاصله‌ی هاسدورف نیز می‌توان استفاده کرد [۴۹]. روش هاسدورف ماکزیمم کمترین فاصله بین دو دسته داده را در نظر می‌گیرد. به این صورت که، ابتدا فاصله‌ی بین تمامی داده‌های دو گروه را بدست آورده و سپس کمترین آن‌را در نظر می‌گیرد [۵۰].

$$H(M,I) = \max\left(h(M,I), h(I,M)\right) \quad (۳)$$

$$h(M,I) = \max_{M\in m} \min_{I\in i} \| m - i \| \quad (۴)$$

در رابطه‌ی۳،m و i به ترتیب اعضا دو خطسیر M و I بوده و h(M,I) فاصله‌ی هر نقطه ازM تا نزدیکترین نقطه از خطسیر I است. علامت ‖.‖ بیانگر فاصله‌ی اقلیدسی است.

روش پیرسون، که یک نوع تابع ابتکاری برای محاسبه وابستگی بین متغیرها است، شباهت بین کاربران را به‌صورت خطی محاسبه می‌کند. این تابع شباهت کاربر u وv را توسط رابطه‌ی ۵ محاسبه می‌کند.

$$Sim(u,v) = D_1(u_i,u_j) \times \frac{\sum_{i\in I}(r_{u,i}-\bar{r}_u)(r_{v,i}-\bar{r}_v)}{\sqrt{\sum_{i\in I}(r_{u,i}-\bar{r}_u)^2}\sqrt{\sum_{i\in I}(r_{v,i}-\bar{r}_v)^2}} \quad (۵)$$

در رابطه‌ی۵، $i \in I$ نشان‌دهنده موارد اطلاعاتی است که هردوکاربر u و v به آن امتیاز داده و $r_{u,i}$ امتیاز کاربر u به مورد اطلاعاتی i بوده و $\bar{r}_u$ میانگین امتیازاتی است که کاربر u به آن موارد اطلاعاتی داده است.

فاصله‌ی فرشه بصورت منحنی $\alpha,\beta : [0,1] \to R^2$ که در رابطه‌ی۶ نشان داده شده است، یک معیار اندازه‌گیری تشابه بین دو منحنی است [۷, ۴۹, ۵۱]. مزیت این اندازه‌گیری نسبت به اندازه‌گیری‌های دیگر مانند فاصله هاسدورف این است که، این فاصله ترتیب نقاط در امتداد منحنی‌ها را در نظر می‌گیرد.

$$D(A,B) = inf \max_{t\in[0,1]} \|\alpha(\sigma(t)) - \beta(\tau(t))\| \quad (۶)$$

در رابطه‌ی۶، علامت ‖.‖ بیانگر فاصله‌ی اقلیدسی و $\tau,\sigma: [0,1] \to [0,1]$ در طول توابع پیوسته و افزایشی متغیر هستند.

علاوه بر روش‌های ذکر شده برای تعیین شباهت مکانی، خطوط‌سیر می‌توانند مشابه فرض شوند اگر ازنظر کشیدگی یا مقدار تراکم دارای شکل مشابه بوده، بیانگر ترتیب خاص یا روند یکنواختی باشند [۵۲]، تشابهاتی در پارامترهای حرکتی استخراج شده مثل سرعت میانگین مشابه، به اشتراک گذارند [۵۱, ۵۳] یا از مکان‌های مشابه عبور کرده باشند. اندازه‌گیری تشابه خطوط‌سیر با در نظر گرفتن پارامترهای حرکت [۵۳] می‌تواند به کشف قوانین و

www.SID.ir

Archive of SID

الگوهای رفتاری‌حرکتی جدید، ارزیابی نظم حرکت مثل تغییر جهت پی‌درپی، پیشنهاد مسیر جایگزین بهینه از نقطه‌نظر کمترین هزینه و بهترین دسترسی، مدیریت منابع و پیش‌بینی ترافیک کمک کند.

از پارامترهای حرکت می‌توان به سرعت اشاره کرد. سرعت، مسافت طی شده در واحد زمان است. یعنی متحرک با چه سرعت ثابتی می‌تواند حرکت کند تا بتواند در مدت زمان معین، مسافت مورد نظر را بپیماید. رابطه-ی۷، نحوه‌ی محاسبه سرعت را نشان می‌دهد.

$$\bar{v} = \frac{p_2 - p_1}{t_2 - t_1} \qquad (۷)$$

در رابطه‌ی شماره‌ی۷، p_1 و p_2 موقعیت‌های ابتدایی و انتهایی بوده و t_1 و t_2 زمان‌های حضور در آن موقعیت‌ها می‌باشند. موقعیت متحرک در این رابطه، با استفاده از فاصله‌ی اقلیدسی میان دو نقطه بدست می‌آید.

Pelekis و همکاران [۵۱]، با معرفی مجموعه‌ای از عملگرهای فاصله برمبنای مختصات مکانی-زمانی خطوط-سیر و همچنین پارامترهای حرکتی سرعت و جهت، به اندازه‌گیری تشابه خطوط‌سیر پرداختند. در این تحقیق از روش‌های فاصله فرشه و مساحت‌مبنا استفاده گردید که در فضای اقلیدسی برای ناوگان کامیون مورد ارزیابی قرار گرفت. علی‌رغم در نظر گرفتن پارامترهای حرکت، اما به شکل هندسی خطوط‌سیر و یافتن نقاط تلاقی آن‌ها حساس است.

۳-۲- معیار زمانی

در پژوهش‌های انجام شده، بیشتر توجه برروی معیار مکانی بازدید کاربرها است. در حالیکه معیار زمان نیز نقش مهم و موثری در تعیین شباهت افراد دارد. بعنوان مثال فردی که صبح‌ها به پارک می‌رود از نظر سلایق فردی می‌تواند با فردی که شب به پارک می‌رود متفاوت باشد. همچنین فردی که ناهار را بعد از کار در شرکت، در رستوران میل می‌کند با فردی که اخر هفته‌ها به رستوران می‌رود، متفاوت هستند، چه بسا هدف یک فرد می‌تواند صرف ناهار یا تفریح باشد. پس می‌توان گفت پارامتر زمان فاکتور قابل اهمیتی خواهد بود. هم‌چنین از داده‌های زمانی می‌توان اطلاعات معنایی را استخراج کرد. بعنوان نمونه فردی که ساعت ۴ به کافی‌شاپ می‌رود، احتمالا قهوه می‌خورد و فردی که صبح به کافی‌شاپ می‌رود احتمال صرف صبحانه بیشتر است.

به‌عنوان یک تحقیق پیش‌قدم در زمینه تجزیه‌وتحلیل مکانی-زمانی، Sinha و Mark برای اندازه‌گیری شباهت از فاصله اقلیدسی در فضای حرکتی آکواریوم مکان-زمان[۱] استفاده کردند [۵۴]. با استفاده از معیار زمانی، wang و همکاران در مقاله‌ی خود از ترکیب معیارهای شخصی، عمومی و زمانی برای پیش‌بینی دوستی‌ها استفاده کردند [۵۵].

در مقاله‌ی [۵۶]، برای حل مشکل تعیین شباهت کاربران در طولانی مدت، پس از استفاده از الگوریتم کلاسه‌بندی سلسله‌مراتبی سه‌لایه، داده‌های GPS را به داده‌های واقعی تبدیل کرده و بر اساس شباهت ضریب همبستگی کسینوسی، فعالیت روزانه کاربران را گروه‌بندی کردند. در این تحقیق شباهت کاربران با در نظر گرفتن دفعات بازدید کاربر از مکان مشخص، بعنوان وزن محاسبه شد [۵۶].

یکی از معمول‌ترین کارهایی که با سری‌های زمانی انجام می‌دهند، مقایسه‌ی یک دنباله با دنباله‌ی دیگر است. گاهی حالتی پیش می‌آید که دو تابع با حالتی تقریبا یکسان در قسمتی نسبت به هم کشیده‌تر هستند. در چنین دنباله‌هایی یک یا هردو دنباله را روی محور زمان می‌تابانیم، روش تاباندن محور زمان به صورت پویا (DTW) در این مورد کارآمد می‌باشد(رابطه ۸).

$$\mathrm{DTW}(Q, C) = min \sqrt{\sum_{k=1}^{k} w_k / K} \qquad (۸)$$

در رابطه‌ی۸، Q و C دو سری زمانی و W ماتریسی است که عناصر آن فاصله‌ی بین دو نقطه‌ی qi و ci است و k عضو دارد.

تحقیقات حاکی از آن است که DTW می‌تواند تغییرات در محور yها را با تاباندن روی محور xها توضیح دهد. این کار می‌تواند باعث تطابق‌های غیر شهودی هنگام تصویر یک نقطه منفرد از یک سری زمانی، روی یک زیربخش بزرگ از سری زمانی دیگر شود که به این رفتار مقادیر منفرد یا تکین[۲] می‌گویند.

---

۱ Space-time Aquarium
۲ singularities

www.SID.ir

Archive of SID

در روش شباهت بزرگترین زیر دنباله‌ی مشترک(lcss) ایده‌ی اصلی، تطبیق دو دنباله بر یکدیگر است، حتی اگر چند جز آن‌ها منطبق نباشد[۵۷, ۵۸]. روش lcss دو مزیت دارد: ۱-بعضی از اجزا مثل نقاط پرت می‌توانند منطبق نباشند. ۲-کارایی محاسبات تقریبی را بیشتر می‌کند.

Michail Vlachos در تحقیق [۵۹]، کارایی بهتر Lcss[۱] را نسبت به DTW نشان داد. DTW بعلت خطاهای زیاد، نمی‌تواند تفاوت کلاس‌ها را به‌ویژه در ابتدا و انتهای خط‌سیرها تشخیص دهد. روش اقلیدسی نیز در این زمینه مطلوب نیست. از ایرادات روش فاصله اقلیدسی حساسیت زیاد نسبت به تغییر زمان و قابلیت استفاده تنها برای توالی‌های مساوی است. ایراد DTW نیز بررسی تمام المانت‌ها می‌باشد. توالی‌های با طول نامساوی نیازمند مقایسه درحالتی است که تفاوت در نسبت نمونه‌ها و نیز تفاوت نقطه شروع می‌باشد، هستند. روش DTW از آنجاییکه بجای تمرکز روی المان‌های تکی الگوی نوسانات آن‌ها، در امتداد زمان را بررسی می‌کند، مناسب است. اما این روش تساوی مثلثی را در نظر نمی‌گیرد. بنابراین در بسیاری از برنامه‌هایی که خطا را حذف نمی‌کنند کارایی ندارد [۵۹].

همچنین از معیار زمانی برای تعیین شباهت در مقاله‌های [۶۰, ۶۱]، از بازه‌ی زمانی ماندن در یک مکان برای شباهت استفاده شده است. علاوه بر روش‌های بیان شده، تحقیقات [۲۰, ۶۲, ۶۳] زمان را برای در نظر گرفتن توالی نقاط طی شده کاربر در محاسبه‌ی شباهت در روش MTM[۲] در نظر گرفتند.

با اضافه شدن معیار زمانی، هرچند که دقت شباهت‌یابی بیشتر از پیش شده است، اما با چالش‌هایی نیز روبرو گردیده است. تکنیک‌های داده‌کاوی مورد استفاده در داده‌های مکانی-زمانی، به دو دلیل چالش‌های محاسباتی را درپی دارند. ۱) پیچیدگی ناشی از جاسازی داده‌ها در فضای پیوسته‌ای از زمان و مکان ۲) تاثیرات همسایگی مکانی-زمانی آن‌ها است. داده‌های مکانی-زمانی بصورت ضمنی تعریف شده‌اند، بصراحت در مجموعه داده‌های ما وجود ندارند، بصورت فرضیات در نظر گرفته می‌شوند و یا در صورت نیاز از داده‌ها استخراج می‌شوند. همچنین داده‌های مکانی ناهمگون هستند و داده‌های زمانی بطور دائمی در حال حرکت هستند. برای داده‌های مکانی-زمانی چهار نوع الگوریتم آنالیز داده زمانی، آنالیز داده مکانی، آنالیز مکانی-زمانی پویا و آنالیز مکانی-زمانی ایستا را دارا هستیم [۶۴].

در آنالیز داده‌های زمانی بعد مکانی را ثابت در نظر می‌گیرند و تغییرات زمانی مثل مهاجرت را مطالعه می‌کنند. اما در آنالیز داده‌های مکانی، تجزیه‌وتحلیل چگونگی تغییر داده‌های ویژگی‌های موضوعی با توجه به فاصله از یک مرجع فضایی در یک زمان خاص را مطالعه می‌کنند. و در آنالیز داده‌های مکانی-زمانی پویا بعد ویژگی‌های موضوعی را ثابت در نظر می‌گیرند و رابطه میان زمان و مکان مثل حرکت ماشین‌ها و عابر پیاده را شرح می‌دهند. همچنین در آنالیز داده‌های مکانی-زمانی ایستا، بعد ویژگی موضوعی و زمان را ثابت در نظر می‌گیرند و به مطالعه‌ی بعد مکانی مثل تعیین مناطق شاخص می‌پردازند.

۳-۳- معیار معنایی

تحقیقات انجام شده در زمینه تعیین شباهت میان کاربران، بیشتر داده‌های خام را مورد توجه قرارداده و اطلاعات زمینه[۳] که توصیف کننده‌ی دانش معنایی[۴] در مورد حرکت است را نادیده می‌گیرند [۶۵]. این در حالیست که فعالیت افراد در یک مکان و تصمیمات آن‌ها، نشان‌دهنده‌ی یک اتفاق یا یک کاربری خاص است.

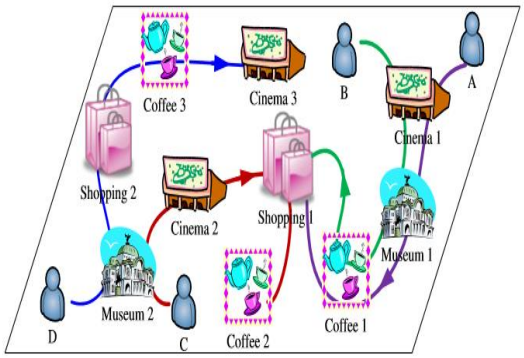
اضافه کردن معنا به داده‌های خام خط‌سیر، به منظور ایجاد یک مدل داده از حرکت شئ متحرک است که قادر به نمایش معانی سطح بالاتری از داده‌های حرکت باشد. غنی‌سازی معنایی خط‌سیر، باعث بهبود مدل‌سازی و تحلیل داده‌های شئ متحرک می‌شود. برای مثال ممکن است دو کاربر به مکان‌هایی که از نظر معنایی شبیه‌اند رفته باشند ولی از نظر مکانی شبیه نباشند. بعنوان مثال خط‌سیر کاربر D و C در شکل ۳ نشان داده شده است. خط‌سیر این دو کاربر از نظر معنایی شبیه هستند ولی از نظر مکانی شبیه نیستند. درحالیکه دو کاربر A و B هم از نظر مکانی و هم از نظر معنایی شبیه هستند. پس با در نظر گرفتن معیار معنایی، دو کاربر A و B شبیه‌تر از دو کاربر D و C هستند [۲۴].

---

۱ Longest commen sub sequense
۲ Maximal Travel Match
۳ background information
۴ semantic information

www.SID.ir

Archive of SID



شکل۳- خط‌سیر کاربران و نقاط شاخص [۲۴]

در مورد اهمیت معیار معنایی، محققین در تحقیق [۶۶] نشان دادند که ۸۷,۱۴٪ از کاربران از نظر معنایی یا زمانی-مکانی شبیه هستند؛ درحالیکه کاربرانی که از هر دو جنبه معنایی و مکانی-زمانی شبیه باشند به ۲,۷۲٪ می‌رسد. در واقع اضافه‌کردن معنا، باعث کاهش ۵۶,۴۱٪ شباهت میان افراد می‌شود.

برای معنادهی به داده‌ها از داده‌های خام حرکت شامل نقطه شروع، داده‌های ضمنی مانند شرایط جوی و همچنین استنتاج از داده‌ها مانند سرعت و شتاب می‌توان استفاده‌کرد [۶۷-۷۲]. یکی از مدل‌های اولیه در زمینه غنی‌سازی معنایی، مدل توقف-حرکت است. این مدل در حقیقت مکان‌هایی که کاربر حداقل برای یک بازه‌ی زمانی مشخص آنجا بوده است را مشخص می‌کند [۸, ۷۳]. مدل SMOT[۱] نیز برای غنی‌سازی معنایی از الگوریتم تقاطع‌مبنا استفاده می‌کند. در این مدل، کاندیدای توقف توسط کاربر براساس کاربرد به مدل معرفی می‌شود، علی‌رغم کارایی مشارکت کاربر، احتمال اینکه توقفی از دید کاربر دور بماند نیز وجود دارد [۷۴].

Rocha و همکاران [۷۴]، شناسایی مناطق توقف و حرکت در خط‌سیر و غنی‌سازی معنایی را بر اساس تغییر جهت نقاط تعریف کردند. خوشه‌هایی از نقاط که تغییر جهت بیشتری نسبت به نقاط دیگر داشته باشند، مناطق مهم خط‌سیر هستند. در روش پیشنهادی که روی داده‌های خط‌سیر کشتی‌های ماهیگیری پیاده‌سازی شد، هدف یافتن مناطقی است که فعالیت‌های ماهیگیری در آن بیشتر است.

تعیین شباهت کاربران براساس معیار معنایی با دو چالش روبه‌رو است، که عبارت است از: ۱) همه‌ی مکان‌ها دارای معنا نیستند. بعنوان مثال خانه و محل کار کاربران اطلاعاتی است که، از روی نقشه نمی‌توان معنای آن را استخراج کرد. ۲)معنای مکان‌ها برای افراد مختلف متفاوت است. به عنوان مثال رستوران برای آشپز محل کار است، در حالیکه برای کارمند محل غذاخوردن است.

یکی از اولین رویکردهای استخراج اطلاعات جغرافیایی از داده‌های حجیم، GeoMiner است که برای کشف قوانین انجمن فضایی تلاش می‌کند [۳۷]. در روش SPIN باترکیب الگوریتم‌های مختلفی مانند multirelational و sub-group spatial cluster analysis سکویی برای استخراج اطلاعات مکانی را فراهم می‌کند [۷۵]. همچنین Cao و همکاران در مقاله‌ی خود الگوریتم‌های داده‌کاوی برای داده‌های خط‌سیر را ایجاد کرده‌اند [۷۶, ۷۲].

Yuan و همکاران در مقاله‌ی خود، به کمک نحوه جابجایی انسان‌ها بین نقاط مختلف و شاخصه‌های زمینی (مثل هتل، پارک، مرکز خرید) شهر را به بخش‌های گوناگون تقسیم می‌کنند. درواقع هر ناحیه را بعنوان یک سند و عملکرد آن بعنوان موضوع سند در نظرگرفته و از روش DMR[۲] عملکرد هر قسمت را تعیین می‌کنند [۵].

Bogorny و همکاران در مطالعات خود، مدل مفهومی CONSTANT را در چارچوب مدل Seek اجرا کرده و با تمرکز بر روی مدل‌سازی، نمایش، آنالیز و کشف دانش از داده‌های خط‌سیر، نتایج خود را ارائه کردند. در این مدل با استفاده از الگوریتم‌های داده‌کاوی متفاوت استخراج داده‌های معنایی صورت گرفته است [۴۳].

Li و Chen اعتقاد دارند برای پیشنهاد افراد مشابه در شبکه‌های اجتماعی، می‌بایست تاثیرات هر سه لایه‌ی گراف مکانی، اجتماعی و محتوا در نظر گرفته‌شود. این محققین بر اساس اعتقاد خود یک مدل سه لایه برای شباهت یابی ارائه کردند [۷۷].

Ying و همکاران [۹] در مقاله‌ی خود، شباهت معنایی کاربران را برای توصیه‌گری استخراج کردند. ابتدا داده‌های GPS را به خط‌سیر معنایی، که هر مکان را با کاربری متناظر با آن نشان می‌دهد، تبدیل کردند. سپس با استفاده از الگوریتم تشخیص الگوهای مکرر، prefix-span، مسیرهای مکرر معنایی را تشخیص داده و از روش Maximal Semantic Trajectory Pattern Similarity

[۱] Stop & Move of Trajectories

[۲] Dirichlet Multinomial Regression

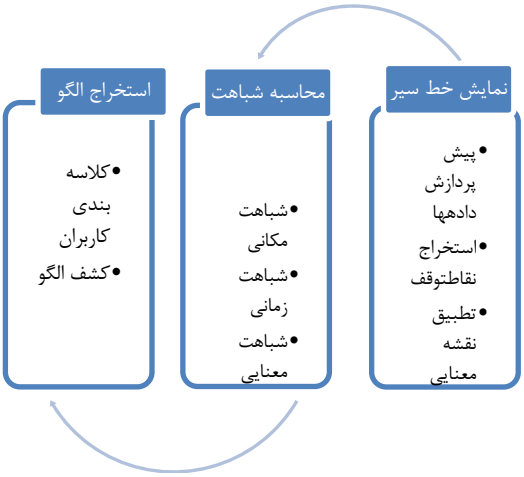
www.SID.ir

Archive of SID

(MSTP Similarity) برای شباهت‌یابی میان دو کاربر استفاده کردند. آن‌ها فاصله‌ی اقلیدسی میان کاربران را محاسبه و نرمال کردند. درنهایت هرکاربر شامل لیستی از فاصله اقلیدسی تا کاربران دیگر است که از ترکیب آن‌ها شباهت بدست می‌آید.

درزمینه استفاده از اطلاعات معنایی در تحقیق [۶۴]، با غنی‌سازی معنایی داده‌های زمانی-مکانی، به کشف دانش معنایی از داده‌های مکانی پرداخته‌است. مسیری که داده‌های GPS را برای استخراج داده‌های معنایی آماده می‌کند یک فرایند سلسله‌مراتبی شامل سه مرحله‌ی اصلی ساخت خط‌سیر خام از داده‌های GPS، غنی‌سازی خط‌سیر معنایی و استخراج داده‌های معنایی است [۶۴].

همانطور که در شکل ۴ نشان داده شده، تبدیل داده-های خام به الگوهای حرکتی شامل سه مرحله‌ی اساسی نمایش خط‌سیر، محاسبه‌ی شباهت و استخراج الگو است. مهمترین قدم در تبدیل داده‌های خام به الگوهای حرکتی محاسبه‌ی شباهت می‌باشد.



شکل۴- مراحل تبدیل داده‌های خام به الگوهای حرکتی

دانشمندان در [۷۴]، ابتدا نقاط توقف در حرکت کاربران که کاربر حداقل برای یک بازه‌ی زمانی مشخص در یک مکان بوده، را مشخص کردند. سپس با استفاده از نقشه‌ای شامل اطلاعات معنایی، معانی نقاط توقف را استخراج کردند. این محققین، برای تعیین مناطق مهم، از الگوریتم تشخیص الگوهای مکرر استفاده کردند. علی‌رغم کارا بودن روش پیشنهادی، بدلیل رفتارهای تصادفی کاربران و عدم پیوستگی مکانی آن‌ها، برخی از مکان‌ها فقط در صورتی ظاهر می‌شوند که حدآستانه تکرار الگو بسیار کم باشد.

محققین در [۴۳]، با در نظر گرفتن خاصیت سلسله-مراتبی مکان، از دو روش IB SMOT[1] و CB SMOT برای ترکیب معنا و مکان با قابلیت تعیین بخش‌های مهم استفاده کردند.

در روش ارائه شده در مقاله‌ی [۹]، ابتدا از الگوریتم تشخیص الگوهای مکرر برای تعیین الگوهایی با بیشترین شباهت معنایی، استفاده گردید. سپس شباهت میان هر جفت از این الگوها با روش MSTP-Similarity measurement برای توصیه پتانسیل دوستی به افراد، ارائه شده است.

اخیرا در [۷۸]، مدل درختی زمانی-مکانی-معنایی را برای مدل‌سازی تاریخچه خط‌سیر کاربران ارایه‌کرده است. در این تحقیق، شباهت میان کاربران با جمع سه بعد شباهت با وزن‌های متفاوت محاسبه شده‌است. این روش دارای اشکالاتی از جمله فقدان دانش قبلی برای اختصاص وزن مناسب به هر بعد، و تعیین شباهات به صورت مستقل در هر بعد است، به گونه‌ای که دو فرد را شبیه فرض می‌کند در حالیکه ممکن است بر اساس یکی از معیارها شبیه نباشند. به علاوه، روش مذکور، برای پیش‌بینی مکان بعدی کاربر طراحی شده و ممکن است برای کاربردهای دیگر مانند آنالیز کلاسه‌بندی خط‌سیر کارایی نداشته‌باشد.

محققین در مقاله‌ی[۷۹]، یک تکنیک نام‌گذاری معنایی ارائه‌کرده‌اند. بسیاری از مکان‌ها، دارای برچسب مشخص مانند رستوران، پارک و بیمارستان می‌باشند، ولی بعضی مکان‌ها فاقد برچسب معنایی هستند. در این روش بر اساس سوابق مکانی کاربر، الگوی آن‌ها استخراج و سپس احتمال متعلق بودن هر برچسب به هر طبقه از کاربری‌ها مشخص می‌شود.

نتیجه‌ی مطالعات انجام شده توسط [۲۱]، ارائه‌ی یک سیستم توصیه‌گر شخصی است. این محققین با بررسی دقیق رفتار حرکتی کاربران، نقاط توقف را استخراج و سپس از الگوریتم کلاسه‌بندی تراکم‌مبنا برای سازماندهی نقاط توقف بصورت یک گراف سلسله‌مراتبی استفاده کردند. در این روش، برای هر فرد یک گراف شخصی تشکیل داده شد. به علاوه، برای هر سطح از گراف سلسله-مراتبی، خط‌سیر کاربر بصورت توالی‌ای از نقاط توقف تبدیل گردید. برای شباهت یابی میان کاربران، در این روش، توالی‌های مشابه از نقاط توقف استخراج و برای آن-ها ارزش IDF-TF محاسبه می‌شود.

[۳] intersection-based stops and moves of trajectories

www.SID.ir

Archive of SID

TF حداقل فرکانسی که دو کاربر مکان را بازدید کرده‌اند را نمایش داده و IDF تعداد کاربرانی که از آن نقطه بازدید کرده‌اند را مشخص می‌کند. درنهایت بر اساس ارزش TF-IDF شباهت میان کاربران استخراج می‌شود.

$$W_{ji} = tf_{i,j} \times \log_2\left(\frac{N}{df_i}\right) \quad (۹)$$

در رابطه‌ی۹، tf تعداد دفعاتی است که یک گروه POI در خط‌سیر معنایی ظاهر می‌شود و df تعداد خط‌سیرهای شامل POI است. tf اهمیت POIها را در خط‌سیر نشان می‌دهد و IDF چگونگی توزیع POI را نمایش می‌دهد.

در پژوهش [۸۰]، شباهت زمانی-مکانی و معنایی را با استفاده از روش Topic model تحلیل کرده‌است. در این پژوهش، شباهت مکانی-زمانی از روشی بنام Kullback Leibler برآورد شده‌است. محققین این تحقیق، برای هر فرد شباهت معنایی و مکانی را تحلیل کرده و سپس بین کل افراد شباهت را مقایسه می‌کند. ایراد روش ارائه شده این است که دو معیار مکانی و معنایی را ترکیب نمی‌کند بلکه مقایسه و تحلیل می‌کند [۶۰, ۸۰].

در مقاله [۶۰]، شباهت را با ترکیب معیارهای مکانی، زمانی و فعالیت محاسبه کرده است. برای تعیین شباهت فعالیت‌هایی که در مکان‌های مختلف هستند، از شباهت معنایی Wpath similarity استفاده کرده است. روش‌های محاسبه‌ی شباهت معنایی بطور کلی دو گروه syntactic similarity و conceptually هستند. روش Wpath similarity جزو گروه دوم می‌باشد.

محققین در مقاله‌ی [۷]، ابتدا تمامی داده‌ها را در پنج گروه کلی شامل: ساختمان آموزشی، هتل و رستوران، سرگرمی داخلی، فعالیت‌های بیرونی و حمل و نقل تقسیم‌بندی کردند. در این تحقیق، سپس از یک الگوریتم کلاسه‌بندی تراکم‌مبنا انعطاف پذیر، برای تعیین شباهت تراکم مکانی-معنایی استفاده شده‌است. این محققین، شباهت معنایی را از روش کسینوسی استخراج کرده که برای وزن‌دهی از روش TF-IDF استفاده می‌کند. لازم به ذکر است که روش وزن‌دهی مذکور، علارغم بقیه‌ی روش-های موجود، به دانش قبلی نیاز ندارد و مستقیماً از داده‌های اولیه بدست می‌آید. همچنین در این تحقیق، برای بررسی شباهت مکانی از شبه‌فاصله‌ی هاسدورف استفاده شده‌است. در مقاله مذکور، شباهت مکانی در مکان‌هایی که از نظر معنایی شبیه هستند محاسبه می‌شود. در نتیجه مکان‌هایی که از نظر معنایی شبیه نباشند حذف می‌شوند. You Wan و همکارانش [۷]، شباهت را با ترکیب معیارهای مکانی و معنایی بیان می-کنند و رابطه‌ی۱۰ را برای ترکیب شباهت‌ها ارائه کردند.

$$\text{semantic intensity}(\text{user}_1, \text{user}_2) = \frac{semantic\ similarity\ (STrace_1, STrace_2)}{2 - geoghraphic\ similarty\ (GTrace_1, GTrace_1)} \quad (۱۰)$$

در رابطه‌ی۱۰، STrace خط‌سیری شامل نقاط شاخص، زمان بازدید و تعداد دفعات بازدید است. CTrace، مجموعه‌ای از POI ها که شامل مکان جغرافیایی و پارامتر معنایی هستند، به‌همراه زمان بازدید است.

در این مقاله، شباهت های مکانی، مکانی-زمانی، تراکم مکانی-معنایی و GTS[۱] با هم مقایسه شده و کارایی بهتر روش تراکم مکانی-معنایی نشان داده شده‌است. از مشکلات روش تراکم مکانی-معنایی این است که وقتی فاصله‌ی جغرافیایی کمتر از ۱ کیلومتر باشد، تفاوتی میان فاصله‌ها در نظر نمی‌گیرد. اگر چه درصد کاربران شبیه در این روش کاهش می‌یابد، ولی دقت بهتری را نسبت به روش GTS دارا می‌باشد.

پژوهشگران در تحقیق [۱۱]، شباهت کاربران را بر مبنای ترکیب معیارهای مکانی، زمانی و معنایی بنام MSM[۲] معرفی کردند. در این تحقیق شباهت مکانی با فاصله‌ی اقلیدسی و شباهت زمانی با فرمول ۱۰ محاسبه می‌شود، اگر شباهت در هر بعد از حدآستانه_ای مشخص کمتر باشد امتیاز شباهت ۱ و در غیر اینصورت ۰ می‌شود، سپس از جمع شباهت‌ها در هر سه بعد امتیاز شباهت را بدست می‌آورند.

$$Dis(A,B) = 1 - \frac{duration\left([a_{t1}, a_{t2}] \cap [[b_{t1}, b_{t2}]]\right)}{duration\left(min[a_{t1}, b_{t1}], max[[a_{t2}, b_{t2}]]\right)} \quad (۱۱)$$

در فرمول ۱۰، a_{t1} زمان شروع در خط‌سیر a و a_{t2} زمان پایان در خط سیر a می‌باشد. سپس با تعریف parity و میانگین‌گیری از $parity(a,b)$ و $parity(b,a)$ شباهت را بدست می‌آورند.

۱ Geographic-temporal-semantic
۲ Multidimensional Similarity Measuring

www.SID.ir

Archive of SID

$$parity(A,B) = \sum_{a \in A} max \begin{Bmatrix} score(a,b): \\ b \in B \end{Bmatrix} \qquad (۱۲)$$

در این تحقیق، وزن هر سه معیار یکسان در نظر گرفته شده است . علی‌رغم اینکه هر سه معیار را باهم ترکیب می‌کند اما در این روش نقاط توقف را مستقل از هم در نظر گرفته و توالی را در نظر نمی‌گیرد.

به منظور بهبود روش MSM در پژوهش [۱۲]، روش SMSM[۱] معرفی شده‌است. در روش ارائه شده، اگر نقاط توقف شبیه بودند، شباهت نقاط حرکت بین هر جفت نقطه توقف را محاسبه کرده و با جمع وزن‌دار آن‌ها، شباهت خط‌سیر را بدست می‌آید. مولفه‌ی معنایی در شباهت نقاط حرکت، نام خیابان است که نشان دهنده‌ی مسیر عبوری کاربران می‌باشد. در نظر گرفتن نقاط حرکت به گردشگرها برای انتخاب مسیر و برنامه ریزی سفر کمک می‌کند. روش‌های محاسبه‌ی شباهت هر معیار همانند روش MSM و فرمول‌های معرفی شده می‌باشد.

اخیرا در [۱۳]، شباهت کاربران براساس شباهت خط-سیر آن‌ها و شباهت مکان‌هایی که کاربران شبکه‌ی اجتماعی بصورت متوالی می‌روند، محاسبه شده‌است. این محققین، مکان‌هایی که کاربران بصورت متوالی می روند را با روش DBSCAN مشخص کرده و سپس شباهت خط-سیرها با روش lcss بدست آوردند. خروجی این الگوریتم شامل lcss بین خط‌سیر کاربران و طول lcss بین خط-سیرهای کاربران در لایه‌های زمانی ۲۴ ، ۱۲و ۶ ساعته است که از جمع وزن‌دار طولlcss در هرلایه امتیاز کاربر بدست می‌آید.

۴- جمع بندی و پیشنهادها

امروزه با گسترش استفاده از شبکه‌های اجتماعی مکان‌مبنا بعنوان داده‌های مکانی بدلیل اطلاعات بالای این شبکه‌ها، استفاده از آن‌ها در زمینه‌های مختلف مانند شباهت‌یابی کاربران با استفاده از داده‌های خط‌سیر، که پایه سرویس‌های توصیه‌گری است، گسترش چشم‌گیری داشته‌است. محققان این حوزه، در ابتدا شباهت میان کاربران شبکه‌های اجتماعی را تنها با در نظر گرفتن معیار مکانی کاربر، که افراد نزدیک به‌هم را شبیه می‌داند، تعیین می‌کردند. عنصر مکان بعنوان مهمترین مولفه شبکه‌های اجتماعی، از معیارهای اصلی تعیین شباهت میان کاربران می‌باشد. لذا باید قبل از استفاده از صحت آن اطمینان حاصل کرد. در پژوهش‌های بعدی با اضافه کردن بعد زمان اطلاعاتی نظیر صرف صبحانه یا شام و همچنین محل کار کاربر با توجه به بازه‌ی زمانی کاربر در هر مکان قابل استخراج خواهد بود. امروزه برای تشخیص دقیق‌تر شباهت افراد بازدید کننده از مکان‌های نزدیک به‌هم، که ممکن است نوع کاربری‌ها کاملا متفاوت باشند، از داده های معنایی خط‌سیر استفاده می‌شود.

به عنوان جمع‌بندی بایستی گفت که هر یک از معیار-های تعیین شباهت دارای نقاط قوت و ضعف هستند. ولی با ترکیب این معیارها با یکدیگر می‌توان بر نقاط ضعف آن‌ها غلبه نمود. البته بدون شک یکی از چالش‌های اصلی، چگونگی ترکیب این معیارها است. در جدول ۱، روش‌های شباهت یابی استفاده شده در این حوزه آورده شده‌است.

از چالش‌های پیش رو در در زمینه‌ی شباهت یابی کاربران اینکه در شبکه‌های اجتماعی ممکن است کاربر اطلاعات کافی در زمینه‌ی نحوه اشتراک‌گذاری موقعیت خود نداشته و یا تمایلی به این کار نداشته باشد. همچنین سرویس‌های شبکه‌های اجتماعی مکان‌مبنا با سنسورهای زمانی-مکانی متنوعی در ارتباط بوده ولی همه‌ی آن‌ها در همه جا کاربرد ندارند، بعنوان نمونه فناوری‌های تعیین موقعیت مکانی مانند GPS در فضاهای بسته کار نمی‌کنند و روش‌های دیگر نیز دقت خوبی ندارند. لذا ما در بعضی موارد با اطلاعات ناقص مواجه هستیم.

مسئله‌ی حریم شخصی کاربر از دیگر چالش‌های پیش‌روی در این شبکه‌ها می‌باشد که نیاز است کاربر از امنیت و استفاده‌ی درست اطلاعات خود مطمئن باشد. برای مثال هنگامی که کاربر بیرون از خانه موقعیت خود را وارد کند این اطلاعات می‌تواند برای سواستفاده‌گران فرصت با ارزشی باشد. بایستی به مرور زمان برای حریم شخصی و امنیت تدابیری اندیشیده شود تا کاربران نسبت به عدم سواستفاده از اطلاعات خود اطمینان یابند.

[۱] Stop and Moves Similarity Measure

www.SID.ir

Archive of SID

جدول ۱- روش‌های شباهت‌یابی

نام روش	معیار پشتیبانی کننده	ویژگی‌های مثبت و منفی	تحقیقات مرتبط
کسینوسی	مکانی و معنایی	بین صفر و یک، ساده است.	[۷, ۱۳, ۲۰, ۲۱, ۲۴, ۵۶, ۶۲, ۷۸, ۸۱, ۸۲]
اقلیدسی	مکانی و زمانی	نامنفی، در نامساوی مثلثی صدق می‌کند، فهم آسان، محاسبه‌ی ساده، برای حل مشکلات دیگر مثل شاخص-گذاری و خوشه‌بندی سری‌های زمانی قابل استفاده است. اما فقط برای سری‌های زمانی با خط مبنای یکسان مناسب است.	[۹, ۱۱, ۱۲, ۵۴]
هاسدورف	مکانی	مقایسه هر دو خط سیر بصورت یکجا صورت می‌گیرد.	[۷, ۴۹, ۸۱, ۸۳-۸۵]
تاباندن زمانی پویا	زمانی و مکانی	یک انتقال است که به‌صورت زمانی یک دنباله‌ای را پیچ و تاب می‌دهد تا فاصله بین آن‌هارا حداقل کند. مستقل از انتقال زمان محلی و کشیدگی خطوط‌سیر را بررسی می-کند ولی برای محاسبه‌ی فاصله شباهت حاصل از اختلاف زمانی نقاط پیشنهاد نمی‌شود.	[۱۰, ۵۹, ۸۶, ۸۷]
بزرگترین زیر دنباله‌ی مشترک	زمانی و معنایی	بزرگی زیر دنباله‌ها با یک حدآستانه مقیاس‌بندی می‌شوند و از چینش دوباره‌ی المان‌ها جلوگیری می‌کند. نسبت به نویز مقاوم است، محاسبات تخمینی کارآمدتری نسبت به DTW دارد.	[۹, ۱۳, ۵۷-۵۹]
فرشه	مکانی	به دلیل وابستگی به شکل و در نظر نگرفتن جنبه زمانی برای تجزیه‌وتحلیل خط‌سیر نامناسب است. این فاصله ترتیب نقاط در امتداد منحنی ها را در نظر می گیرد.	[۴۹, ۸۳, ۸۴, ۸۶]
تشخیص الگوهای مکرر	زمانی و معنایی	برای کشف مناطق مهم شهر قابل استفاده است.	[۹, ۱۳, ۶۱, ۶۲, ۷۴]
الگوریتم کلاسه‌بندی تراکم-مبنا(DBSCAN[۱])	مکانی و زمانی	در حذف داده‌های پرت مفید هستند، نیاز به تعریف شعاع همسایگی وحداقل تعداد نقاط در شعاع مشخص توسط کاربر	[۷, ۱۳, ۲۱, ۵۶, ۶۲, ۸۱]
فراوانی عبارت (TF)	معنایی	درصورت طولانی بودن یک متن نسبت به دیگری بردار بسمت متن طولانی‌تر بایاس می‌شود. راه حل نرمال کردن هر کلمه	[۷, ۲۱, ۸۱]
معکوس فراوانی سند(IDF)	معنایی	اگر یک کلمه در تعداد کمی سندها وجود داشته باشد تفکیک بیشتری ایجاد می‌کند، بیانگر کمیاب بودن	[۷, ۲۱, ۸۱]

مراجع

[1] D. S. Taha, "The influence of social networks in visiting, planning and living in cities. Alexplore: A pilot project in Alexandria," Alexandria Engineering Journal, vol. 52, no. 3, pp. 479-488, 2013.

[2] K. Kodama, Y. Iijima, X. Guo, and Y. Ishikawa" ,Skyline queries based on user locations and preferences for making location-based recommendations," in Proceedings of the 2009 International Workshop on Location Based Social Networks, 2009: ACM, pp. 9-16.

[3] S. Dodge, "Exploring movement using similarity analysis," Verlag nicht ermittelbar, 2011.

[4] A. K. Dey, G. D. Abowd, and D. J. H. C. I. Salber, "A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications," vol. 16, no. 2-4, pp. 97-166, 2001.

[5] J. Yuan ,Y. Zheng, and X. Xie, "Discovering regions of different functions in a city using human mobility and POIs," in Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, 2012: ACM, pp. 186-194.

۱ Density-based spatial clustering of applications with Noise

www.SID.ir

Archive of SID

[6] J. Bao, Y. Zheng, and M. F. Mokbel, "Location-based and preference-aware recommendation using sparse geo-social networking data," in Proceedings of the 20th international conference on advances in geographic information systems, 2012: ACM, pp. 199-208.

[7] Y. Wan, C. Zhou, and T. J. I. I. J. o. G.-I. Pei, "Semantic-geographic trajectory pattern mining based on a new similarity measurement," vol. 6, no. 7, p. 212, 2017.

[8] S. Spaccapietra, C. Parent, M. L. Damiani, J. A. de Macedo, F. Porto, and C. Vangenot, "A conceptual view on trajectories," Data & knowledge engineering, vol. 65, no. 1, pp. 126-146, 2008.

[9] J. J.-C. Ying, E. H.-C. Lu, W.-C. Lee, T.-C. Weng, and V. S. Tseng, "Mining user similarity from semantic trajectories," in Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks, 2010: ACM, pp. 19-26.

[10] M. Shokoohi-Yekta, B. Hu, H. Jin, J. Wang, and E. Keogh, "Generalizing DTW to the multi-dimensional case requires an adaptive approach," Data mining and knowledge discovery, vol. 31, no. 1, pp. 1-31, 2017.

[11] A. S. Furtado, D. Kopanaki, L. O. Alvares, and V. Bogorny, "Multidimensional similarity measuring for semantic trajectories," Transactions in GIS, vol. 20, no. 2, pp. 280-298, 2016.

[12] A. L. Lehmann, L. O. Alvares, and V. Bogorny, "SMSM: a similarity measure for trajectory stops and moves," International Journal of Geographical Information Science, pp. 1-26, 2019.

[13] Z.-l. Xiong, C.-h. Xia, B. Sun, and M.-j. Hou, "Mining Similarity of Users in Location-Based Social Networks for Discovering Overlapping Communities," in Recent Developments in Intelligent Computing, Communication and Devices: Springer, 2019, pp. 417-428.

[14] D. M. Boyd, "Friendster and publicly articulated social networking," 2004.

[15] Y. Zheng, X. J. A. T. o. I. S. Xie, and Technology, "Learning travel recommendations from user-generated GPS traces," vol. 2, no. 1, p. 2, 2011.

[16] T. Zhu, R. Greiner, and G. Häubl, "Learning a model of a web user's interests," in International Conference on User Modeling, 2003: Springer, pp. 65-75.

[17] P. Symeonidis, A. Krinis, and Y. Manolopoulos, "Geosocialrec: Explaining recommendations in location-based social networks," in East European Conference on Advances in Databases and Information Systems, 2013 :Springer, pp. 84-97.

[18] S. Scellato, A. Noulas, and C. Mascolo, "Exploiting place features in link prediction on location-based social networks," in Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, 201 :۱ACM, pp. 1046-1054.

[19] J. Cranshaw, E. Toch, J. Hong, A. Kittur, and N. Sadeh, "Bridging the gap between physical location and online social networks," in Proceedings of the 12th ACM international conference on Ubiquitous computing, 2010: ACM, pp. 11.۱۲۸-۹

[20] Q. Li, Y. Zheng, X. Xie, Y. Chen, W. Liu, and W.-Y. Ma, "Mining user similarity based on location history," in Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems, 2008: ACM, p. 34.

[21] Y. Zheng, L. Zhang, Z. Ma, X. Xie, and W.-Y. Ma, "Recommending friends and locations based on individual location history," ACM Transactions on the Web (TWEB), vol. 5, no. 1, p. 5, 2011.

[22] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, "Mining interesting locations and travel sequences from GPS trajectories," in Proceedings of the 18th international conference on World wide web, 2009: ACM, pp. 791-800.

[23] C.-C. Hung, C.-W. Chang, and W.-C. Peng, "Mining trajectory profiles for discovering user communities," in Proceedings of the 2009 International Workshop on Location Based Social Networks, 2009: ACM, pp. 1-8.

[24] X. Xiao, Y. Zheng, Q. Luo, and X. Xie, "Finding similar users using category-based location history," in Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems, 2010: ACM, pp. 442-445.

[25] H. Zhang, Y. Yang, and Z. Zhang, "Cts: combine temporal influence and spatial influence for time-aware poi recommendation," in International Conference of Pioneering Computer Scientists, Engineers and Educators, 2016: Springer, pp. 272-286.

[26] M. Li, G. Sagl, L. Mburu, and H. Fan, "A contextualized and personalized model to predict user interest using location-based social networks," Computers, Environment and Urban Systems, vol. 58, pp. 97-106, 2016.

www.SID.ir

Archive of SID

[27] D. D'Agostino, F. Gasparetti, A. Micarelli, and G. Sansonetti, "A social context-aware recommender of itineraries between relevant points of interest," in International Conference on Human-Computer Interaction, 2016: Springer, pp. 354-359.

[28] Y. Zheng and X. Xie, "Learning travel recommendations from user-generated GPS traces," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 2, no. 1, p. 2, 2011.

[29] H. Yoon, Y. Zheng, X. Xie, and W. Woo, "Smart itinerary recommendation based on user-generated GPS trajectories," in International Conference on Ubiquitous Intelligence and Computing, 2010: Springer, pp. 19-34.

[30] H. Yoon, Y. Zheng, X. Xie, and W. Woo, "Social itinerary recommendation from user-generated digital trails," Personal and Ubiquitous Computing, vol. 16, no. 5, pp. 469-484, 2012.

[31] H. A. Karimi, B. Zimmerman, A. Ozcelik, and D. Roongpiboonsopit, "SoNavNet: a framework for social navigation networks," in Proceedings of the 2009 International Workshop on Location Based Social Networks, 2009: ACM, pp. 81-87.

[32] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with GPS history data," in Proceedings of the 19th international conference on World wide web, 2010: ACM, pp. 1029-1038.

[33] A. Aldhaheri and J. Lee, "Event detection on large social media using temporal analysis," in 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC), 2017: IEEE, pp. 1-6.

[34] M. Hasanuzzaman and A. Way, "Local Event Discovery from Tweets Metadata," in Proceedings of the Knowledge Capture Conference, 2017: ACM, p. 39.

[35] G. Stilo and P. Velardi, "Efficient temporal mining of micro-blog texts and its application to event discovery," Data Mining and Knowledge Discovery, vol. 30, no. 2, pp. 372-402, 2016.

[36] R. Lee and K. Sumiya, "Measuring geographical regularities of crowd behaviors for Twitter-based geo-social event detection," in Proceedings of the 2nd ACM SIGSPATIAL international workshop on location based social networks, 2010: ACM, pp. 1-10.

[37] W. Liu, Y. Zheng, S. Chawla, J. Yuan, and X. Xing, "Discovering spatio-temporal causal interactions in traffic data streams," in Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, 2011: ACM, pp. 1010-1018.

[38] G. Xu-Rui, W. Li, and W. Wei-Li, "Using multi-features to recommend friends on location-based social networks," Peer-to-Peer Networking and Applications, vol. 10, no. 6, pp. 1323-1330, 2017.

[39] J. Bao, Y. Zheng, D. Wilkie, and M. F. Mokbel, "A survey on recommendations in location-based social networks," ACM Transaction on Intelligent Systems and Technology, 2013.

[40] H. Gao, J. Tang, and H. Liu, "Exploring social-historical ties on location-based social networks," in Icwsm, 2012.

[41] C. Emmanouilidis, R.-A. Koutsiamanis, A. J. J. o. n. Tasidou, and c. applications, "Mobile guides: Taxonomy of architectures, context awareness, technologies and applications," vol .36 ,no. 1, pp. 103-125, 2013.

[42] D. Lin, "An information-theoretic definition of similarity," in Icml, 1998, vol. 98, no. 1998: Citeseer, pp. 296-304.

[43] V. Bogorny, C. Renso, A. R. de Aquino, F. de Lucca Siqueira, and L. O. Alvares, "Constant–a conceptual data model for semantic trajectories of moving objects," Transactions in GIS, vol. 18, no. 1, pp. 66-88, 2014.

[44] W. R. Tobler, "A computer movie simulating urban growth in the Detroit region," Economic geography, vol. 46, no. sup1, pp. 234-240 ,1970.

[45] J. Goldenberg and M. J. a. p. a. Levy, "Distance is not dead: Social interaction and geographical distance in the internet era," 2009.

[46] S. Scellato, C. Mascolo, M. Musolesi, and V. Latora, "Distance Matters: Geo-social Metrics for Online Social Networks," in WOSN, 2010.

[47] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins, "Geographic routing in social networks," Proceedings of the National Academy of Sciences, vol. 102, no. 33, pp. 11623-11628, 2005.

[48] R. Lübke, D. Schuster, and A. Schill, "Mobilisgroups: Location-based group formation in mobile social networks," in Pervasive Computing and Communications Workshops (PERCOM Workshops), 2011 IEEE International Conference on, 2011: IEEE, pp. 502-507.

[49] M. Buchin, S. Dodge, and B. Speckmann, "Similarity of trajectories taking into account geographic context," Journal of Spatial Information Science, vol. 2014, no. 9, pp. 101-124, 2014.

www.SID.ir

Archive of SID

[50] H. Alt, M. J. I. J. o. C. G. Godau, and Applications, "Computing the Fréchet distance between two polygonal curves," vol. 5, no. 01n02, pp. 75-91, 1995.

[51] N. Pelekis, I. Kopanakis, G. Marketos, I. Ntoutsi, G. Andrienko, and Y. Theodoridis, "Similarity search in trajectory databases," in 14th International Symposium on Temporal Representation and Reasoning (TIME'07), 2007: IEEE, pp. 129-140.

[52] J. Gudmundsson, P. Laube, and T. Wolle, "Computational movement analysis," in Springer handbook of geographic information: Springer, 2011, pp. 423-438.

[53] S. Dodge, P. Laube, and R. Weibel" ,Movement similarity assessment using symbolic representation of trajectories," International Journal of Geographical Information Science, vol. 26, no. 9, pp. 1563-1588, 2012.

[54] G. Sinha and D. M. J. J. o. G. S. Mark, "Measuring similarity between geospatial lifelines in studies of environmental health," vol. 7, no. 1, pp. 115-136, 2005.

[55] H. Wang, Z. Li, and W.-C. Lee, "PGT: Measuring mobility relationship using personal, global and temporal factors," in Data Mining (ICDM), 2014 IEEE International Conference on, 2014: IEEE, pp. 570-579.

[56] M. Lv, L. Chen, and G. J. I. S. Chen, "Mining user similarity based on routine activities," vol. 236, pp. 17-32, 2013.

[57] M. Vlachos, D. Gunopoulos, and G. Kollios, "Discovering similar multidimensional trajectories," in icde, 2002: IEEE, p. 0673.

[58] Z. Shou and X. Di, "Similarity analysis of frequent sequential activity pattern mining," Transportation Research Part C: Emerging Technologies, vol. 96, pp. 122-143, 2018.

[59] M. Vlachos, G. Kollios, and D .Gunopulos, "Discovering similar multidimensional trajectories," in Data Engineering, 2002. Proceedings. 18th International Conference on, 2002: IEEE, pp. 673-684.

[60] A. G. Prabono, S.-L. Lee, and B. N. Yahya, "Context-based similarity measure on human behavior pattern analysis," Soft Computing, pp. 1-13, 2018.

[61] A. G. Prabono, S.-L. Lee, and B. N. Yahya, "Context-based similarity measure on human behavior pattern analysis," Soft Computing, vol. 23, no. 14, pp. 5455-5467, 2019.

[62] M. Lv, L. Chen, and G. Chen, "Mining user similarity based on routine activities," Information Sciences, vol. 236, pp. 17-32, 2013.

[63] X. Xiao, Y. Zheng, Q. Luo, and X. Xie, "Inferring social ties between users with human location history," Journal of Ambient Intelligence and Humanized Computing, vol. 5, no. 1, pp. 3-19, 2014.

[64] S. Chakri, S. Raghay, and S. El Hadaj, "Semantic Trajectory Knowledge Discovery: A Promising Way to Extract Meaningful Patterns from Spatiotemporal Data," International Journal of Software Engineering and Knowledge Engineering, vol. 27, no. 03, pp. 399-421, 2017.

[65] Z. Yan, D. Chakraborty, C. Parent, S. Spaccapietra, and K. Aberer, "Semantic trajectories: Mobility data computation and annotation," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 4, no. 3, p. 49, 2013.

[66] W. Huang and S. Li, "Understanding human activity patterns based on space-time-semantics," ISPRS journal of photogrammetry and remote sensing, vol. 121, pp. 1-10, 2016.

[67] B. Guc, M. May, Y. Saygin, and C. Körner, "Semantic annotation of gps trajectories," in 11th AGILE international conference on geographic information science, 2008, vol. 38, no. 6, pp. 1-9.

[68] K.-F. Richter, F. Schmid, and P. Laube, "Semantic trajectory compression: Representing urban movement in a nutshell," Journal of Spatial Information Science, vol. 2012, no. 4, pp. 3-30, 2012.

[69] J. A. M. Rocha, V. C. Times, G. Oliveira, L. O. Alvares, and V. Bogorny, "DB-SMoT: A direction-based spatio-temporal clustering method," in Intelligent systems (IS), 2010 5th IEEE international conference, 2010: IEEE, pp. 114-119.

[70] Z. Yan, D. Chakraborty, C. Parent, S. Spaccapietra, and K. Aberer, "SeMiTri: a framework for semantic annotation of heterogeneous trajectories," in Proceedings of the١٤ th international conference on extending database technology, 2011: ACM, pp. 259-270.

[71] E. Camossi, P. Villa, and L. Mazzola, "Semantic-based anomalous pattern discovery in moving object trajectories," arXiv preprint arXiv:1305.1946, 2013.

[72] Z. Li et al., "Movemine: Mining moving object data for discovery of animal movement patterns," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 2, no. 4, p. 37, 2011.

[73] C. Parent et al., "Semantic trajectories modeling and analysis," ACM Computing Surveys (CSUR), vol. 45, no. 4, p. 42, 2013.

www.SID.ir

Archive of SID

[74] L. O. Alvares, V. Bogorny, B. Kuijpers, J. A. F. de Macedo, B. Moelans, and A. Vaisman, "A model for enriching trajectories with semantic geographical information," in Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems, 2007: ACM, p. 22.

[75] M. May and A. Savinov, "Spin!-an enterprise architecture for spatial data mining," in International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, 2003: Springer, pp. 510-517.

[76] H. Cao, N. Mamoulis, and D. W. Cheung, "Discovery of periodic patterns in spatiotemporal sequences," IEEE Transactions on Knowledge and Data Engineering, vol. 19, no. 4, pp. 453-467, 2007.

[77] N. Li and G. Chen, "Multi-layered friendship modeling for location-based mobile social networks," in Mobile and Ubiquitous Systems: Networking & Services, MobiQuitous, 2009. MobiQuitous' 09. 6th Annual International, 2009: IEEE, pp. 1-10.

[78] J. J.-C .Ying, W.-C. Lee, and V. S. Tseng, "Mining geographic-temporal-semantic patterns in trajectories for location prediction," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 5, no. 1, p. 2, 2013.

[79] M. Ye, D. Shou, W.-C. Lee, P. Yin, and K. Janowicz, "On the semantic annotation of places in location-based social networks," in Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, 2011: ACM, pp. 520-528.

[80] W. Huang, S. J. I. J. o. P. Li, and R. Sensing, "Understanding human activity patterns based on space-time-semantics," vol. 121, pp. 1-10, 2016.

[81] Y. Wan, C. Zhou, and T. Pei, "Semantic-geographic trajectory pattern mining based on a new similarity measurement," ISPRS International Journal of Geo-Information, vol. 6, no. 7, p. 212, 2017.

[82] Y. Cao, Y. Si, Z. Cai, and Z. Ding, "Mining Spatio-Temporal Semantic Trajectory for Groups Identification," in 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2018: IEEE, pp. 308-313.

[83] G. Trajcevski, H. Ding, P. Scheuermann, R. Tamassia, and D. Vaccaro, "Dynamics-aware similarity of moving objects trajectories," in Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems, 2007: ACM, p. 11.

[84] K. Buchin, M. Buchin, J. Gudmundsson, M. Löffler, and J. Luo, "Detecting commuting patterns by clustering subtrajectories," International Journal of Computational Geometry & Applications, vol. 21, no . ,۰۳pp. 253-282, 2011.

[85] H. Alt, "The computational geometry of comparing shapes," in Efficient Algorithms: Springer, 2009, pp. 235-248.

[86] H. Ding, G. Trajcevski, and P. Scheuermann, "Efficient similarity join of large sets of moving object trajectories," in 2008 15th International Symposium on Temporal Representation and Reasoning, 2008: IEEE, pp. 79-87.

[87] B.-K. Yi, H. Jagadish, and C. Faloutsos, "Efficient retrieval of similar time sequences under time warping," in Proceedings 14th International Conference on Data Engineering, 1998: IEEE, pp. 201-208.

www.SID.ir