



ارائه سرویس توصیه مسیر بر اساس رتبه‌بندی خط سیر کاربران محلی

رویا شورونی^{۱*}، محمدرضا ملک^۲

۱- دانشجوی کارشناسی ارشد سیستم اطلاعات مکانی - دانشکده مهندسی ژئودزی و ژئوماتیک - دانشگاه صنعتی خواجه نصیرالدین طوسی
 ۲- دانشیار گروه سیستم اطلاعات مکانی - دانشکده مهندسی ژئودزی و ژئوماتیک - دانشگاه صنعتی خواجه نصیرالدین طوسی

تاریخ دریافت مقاله: ۱۳۹۴/۰۸/۰۲ تاریخ پذیرش مقاله: ۱۳۹۴/۱۲/۰۲

چکیده

امروزه مجهز شدن دستگاه‌های همراه به فناوری‌های مدرن مکان‌یابی، باعث تحولی در عرصه‌ی فناوری اطلاعات مکانی-زمانی و ظهور مجموعه داده خطوط سیر مکانی (Spatial Trajectory) شده است. وجود حجم زیادی از خطوط سیر مکانی، فرصتی برای ارائه سرویس‌های متنوع مکان‌مبنا از طریق تجهیزات همراه فراهم می‌کند. یکی از مهم‌ترین و کاربردی‌ترین سرویس‌های مکان‌مبنا، سرویس‌های توصیه مسیر زمان‌مند به کاربران ناآشنا به منطقه هستند. یافتن بهترین مسیر، یکی از مسائل رایج و مبتلابه زندگی کنونی برای هدایت کاربر جهت رسیدن به مقصد است. در تحقیق حاضر جهت یافتن مسیر بهینه، ابتدا با رتبه‌بندی خط سیر کاربران با استفاده از الگوریتم HITS، کاربران آشنا به منطقه یا به اصطلاح "کاربران محلی" و قطعه خیابان‌های مورد تردد اکثر کاربران محلی یا "قطعه خیابان‌های محلی" استخراج شد. سپس با هدف ارائه یک مسیریابی تقریبی گرافی تحت عنوان گراف محلی ایجاد شد. مرکز قطعه خیابان‌های محلی به‌عنوان گره‌های آن و یال‌های آن بر اساس دنباله‌ی ترتیبی قطعه خیابان‌های محلی از خط سیر کاربران به دست می‌آید. پس‌از آن با به‌کارگیری الگوریتم دایجسترا بر روی شبکه راه اصلی و مسیر تقریبی به‌دست‌آمده، کوتاه‌ترین مسیر بین بر اساس این گراف از دو مرحله مسیریابی برای به دست آوردن مسیر بهینه استفاده شده است. برای پیاده‌سازی و آزمون، داده‌های مورداستفاده از خط سیر ۱۰۰ کاربر متحرک در منطقه شش شهر تهران به مدت سه ماه به‌صورت روزانه با استفاده از تلفن‌های همراه مجهز به GPS جمع‌آوری شده است. برای ارزیابی عملکرد مسیریابی دومرحله‌ای، میزان زمان سفر در مسیر پیشنهادی نسبت به کوتاه‌ترین مسیر دایجسترا برای طول‌های مختلف و کاربرانی با سطح آگاهی‌های مختلف از منطقه، به‌طور تجربی مقایسه شد. بر این اساس زمان سفر در روش پیشنهادی نسبت به کوتاه‌ترین مسیر حدود ۶۲ درصد کاهش یافت.

کلیدواژه‌ها: توصیه مسیر، خط سیر، الگوریتم رتبه‌بندی، کاربران و قطعه خیابان‌های محلی.

* نویسنده مکاتبه کننده: رویا شورونی، تهران - خ ولیعصر - تقاطع میرداماد - دانشکده مهندسی نقشه‌برداری، کدپستی: ۱۵۴۳۳-۱۹۹۶۷

تلفن: ۰۹۱۲۵۶۴۶۶۴۱

۱- مقدمه

متوالی که دارای مؤلفه های مکانی-زمانی هستند، مدل می شود. این داده، یکی از مهم ترین اجزای مربوط به کاربران است که دانش گسترده ای در مورد میزان آشنایی کاربران از راه ها و میزان اهمیت راه ها برای کاربران می تواند بدهد [۱]. به بیان دیگر خط سیر کاربران را می توان به عنوان شرح مختصری از رفتارهای مکرر کاربر، در دو بعد مکان (مکان های بازدید شده در طول حرکت) و زمان (مدت زمان حرکت) تعریف کرد. اخیراً سرویس های توصیه مسیر بر اساس اخذ و به روزرسانی خط سیر کاربران توانسته است راه حل مفید برای افرادی باشد که به دنبال سفر به منطقه ای هستند که با آن آشنایی ندارند. در اینجا سامانه ای ارائه شده که در آن رانندگانی که مورد ردیابی قرار می گیرند، افراد محلی و آشنا به منطقه هستند که در انتخاب مسیرها، بهترین مسیر را برمی گزینند. مقاله حاضر، ارائه مدلی جهت استفاده از داده های خط سیر ثبت شده توسط GPS جهت غنی نمودن داده های معابر برای مسیر یابی در مناطق شهری است. ما در اینجا برای ارائه سرویس توصیه مسیر بر اساس رفتار حرکتی افراد محلی، خط سیر تعدادی از کاربران مجهز به وسیله نقلیه شخصی را بررسی و مدل سازی می کنیم. در ادامه، فصل دو پیشینه تحقیق بررسی می شود. فصل سوم اختصاص به حل مشکلات نظری، مدل سازی و الگوریتم های توسعه داده شده دارد. پیاده سازی و ارزیابی سرویس در فصل چهارم شرح داده می شوند. فصل آخر نیز به نتیجه گیری و جمع بندی اختصاص یافته است.

۲- پیشینه تحقیق

داده های GPS برای اتصال موقعیت جغرافیایی و اطلاعات زمانی به کاربر متحرک استفاده می شوند. بسیاری از برنامه ها و سرویس های مکان مبنای، از طریق داده کاوی و پردازش داده های GPS قادر به استخراج اطلاعات معنی دار در کاربردهای مختلف مانند کشف الگوهای حرکتی [۲ و ۳] و پیش بینی

با توجه به افزایش ترافیک و به دنبال آن زیاد شدن زمان سفر در شهرهای بزرگ مانند تهران، ارائه راه حل برای کاهش این مشکلات امری ضروری است. یکی از مهم ترین عوامل تأثیرگذار در ترافیک شهری، یافتن مسیر بهینه برای رسیدن به مقصد است. یافتن مسیر بهینه یکی از اصلی ترین فعالیت های است که افراد همه روزه در سفرهای درون شهری خود با آن مواجه می شوند. در واقع مسیریابی به یک فعالیت روزانه تبدیل شده و به عنوان یکی از قابلیت های کلیدی در سرویس هایی چون Google Maps در دسترس عموم قرار گرفته است. اما متأسفانه در حال حاضر در تهران هیچ گونه سامانه مدیریت ناوگانی برای کاربران سواره وجود ندارد.

اصولاً مردم برای انتخاب یک مسیر یا عدم انتخاب آن باید دلیل خوبی داشته باشند. برای مثال افراد از انتخاب مسیریابی که در آن احتمال اتلاف زمان سفر وجود دارد، اجتناب می کنند. بر مبنای این فرض می توان مسئله را این گونه تعریف کرد که مسیریابی را که اغلب توسط عموم کاربران آشنا به منطقه مورد استفاده قرار می گیرند، با هدف کشف واقعیت های پنهان از طریق داده کاوی به دست بیاوریم. زیرا افراد محلی علاوه بر کوتاهی طول مسیر، فاکتورهای دیگری چون تقاطع ها، ترافیک، چراغ های راهنمایی، تغییر در جهت حرکت معابر و غیره را در مسیریابی لحاظ می کنند. اما شناخت کاربران محلی و آشنا به منطقه مسئله ای است که جز تحلیل رفتار حرکتی آنها ممکن نیست. امروزه افزایش روزافزون دستگاه های متحرک و فناوری های موقعیت یابی، منجر به حجم انبوه و بی شماری از اطلاعات مکانی-زمانی و در پی آن نیازمند کاربردها و تحلیل های زیادی شده است. خط سیر مکانی یک نوع اطلاعات مکانی-زمانی محسوب می شود که جابجایی مکانی عامل متحرک را در طی زمان ثبت می کند و به صورت نقاط گسسته و

معین، مشابه است و برای کاربران حق انتخاب بین چندین مسیر وجود ندارد. [۱۱] خط سیر کاربران را بر اساس گراف سلسله مراتبی درختی^۱ مدل کردند و توسط مدل استنتاجی HITS مکان‌های جذاب و دنباله ترتیبی آن‌ها را در مقیاس‌های مختلف به دست آوردند. هدف آن‌ها توصیه مسیرهای جذاب، تنها بر اساس دنباله‌ای ترتیبی از مکان‌های جذاب، بدون در نظر گرفتن طول مسیر بود و این مسیر تنها برای توصیه به گردشگران مناسب است.

سیستم‌های توصیه مسیر به‌طور کلی دو دسته تقسیم می‌شوند که عبارتند از: سیستم‌های تعاملی توصیه‌گر که از نظرات کاربر برای توصیه استفاده می‌کنند و سیستم‌های خودکار توصیه‌گر که با هدف کاهش میزان وابستگی سیستم به کاربر ارائه می‌شوند. در زمینه سیستم‌های تعاملی [۱۲] با تحلیل خط سیر افراد خبره، سیستم توصیه‌گر مسیر توسعه دادند. روش آن‌ها نیازمند مدت‌زمان سفر و دو نقطه به‌عنوان مبدأ و مقصد است. آن‌ها با استفاده از تکنیک داده‌کاوی مکان‌های پرتردد متوالی، زمان سفر و زمان اقامت بین مکان‌ها را استخراج کردند و نهایتاً با اطلاعات به‌دست‌آمده برای مکان‌ها و ارتباط بین آن‌ها گرافی را برای مسیریابی زمانمند کاربران توسعه می‌دهند. [۱۳] یک سیستم تعاملی توصیه مسیر سفر ارائه دادند که در آن کاربر قادر است تعیین کند از چه مکان‌هایی بازدید و از توصیه چه مکان‌هایی اجتناب شود. به‌طور مشابه، در سیستم تعاملی طراحی شده توسط [۱۴] امکان تعیین محدودیت‌های عمومی مانند زمان و میزان جذابیت مکان‌ها توسط کاربر در برنامه سفر گنجانده شده است. چنین سیستم‌های تعاملی بیشتر برای کاربران آشنا به منطقه مناسب است و تنها برنامه سفر دقیق تر و جزئی‌تری به کاربر ارائه می‌شود و برای کاربران نا آشنا

الگوی تکراری [۴] هستند. پردازش داده‌های خام GPS می‌تواند به فرم‌های قابل استفاده از جمله تولید نقشه‌های مسیر [۵] تبدیل شوند. همچنین استخراج دنباله‌های متوالی سفر و ساختار سلسله مراتبی فضا‌های مکانی از خط سیر کاربران، منعکس‌کننده شباهت بین کاربران است [۶]. در چند سال اخیر تحقیقات بسیاری در زمینه‌ی برنامه‌ریزی مسیر بر اساس داده‌های GPS یا خط سیر کاربران انجام شده است. در این زمینه اساس کار بسیاری از محققین، استخراج خط سیر کاربران از داده‌های GPS ثبت‌شده توسط آن‌ها و استفاده از آن‌ها به‌عنوان داده ورودی برای کشف دنباله‌ای از قطعه خیابان‌های پرتردد است. روش‌های مختلفی برای بهبود کیفیت سرویس‌های برنامه‌ریزی مسیر با استفاده از خط سیر کاربران معرفی شده‌اند و ایده اصلی آن‌ها ایجاد یک شبکه راه از قطعه خیابان‌های وزن‌داری است که وزن آن‌ها از مجموعه خط سیر کاربران قابل‌بازیابی است [۷]. تولید شبکه راه، بر اساس اولویت‌بندی قطعه خیابان‌هایی است که مورد استفاده کاربران خاص مانند رانندگان تاکسی یا دیگر رانندگان قرار می‌گیرد [۸ و ۹] پس از کشف قطعه خیابان‌های مورد تردد کاربران، ترتیب عبور از قطعه خیابان‌ها از خط سیر کاربران برای برنامه‌ریزی مسیر به دست می‌آید. با ارائه چارچوب برنامه‌ریزی مسیر بر اساس اطلاعات جمع‌آوری شده، می‌توان مسیرهای پرترفدار را با استفاده از خط سیرهای غیرمسلّم و نامعلوم کاربران استخراج کرد. این چارچوب با توجه به یک دنباله مکانی و فاصله زمانی، قادر به تعیین k مسیر پرترفدار است که به‌طور متوالی از تجمیع خط سیرهای نامعلوم عبور می‌کند. در حوزه برنامه‌ریزی مسیر [۱۰] با استفاده از خط سیر رانندگان تاکسی و تولید گراف معابر اصلی، بهترین مسیر را به دست آوردند. روش آن‌ها به افراد خاص یعنی رانندگان تاکسی محدود شده است. درحالی‌که مسیر رانندگان تاکسی، در یک مبدأ و مقصد

¹ tree-based hierarchical graph

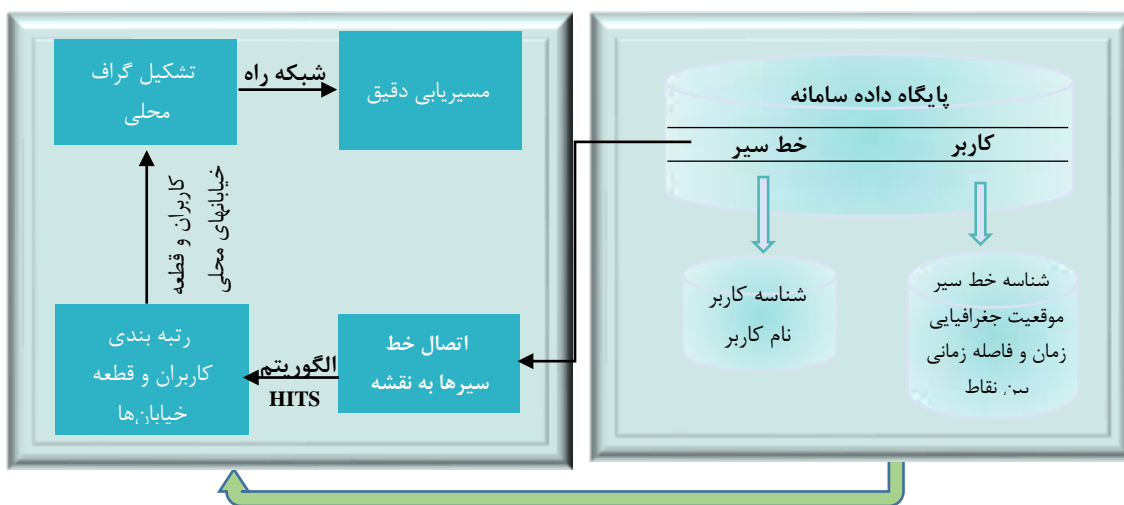
به منطقه، عملی و قابل استفاده نمی‌باشد.

استفاده می‌کنیم.

۳- مدل‌سازی

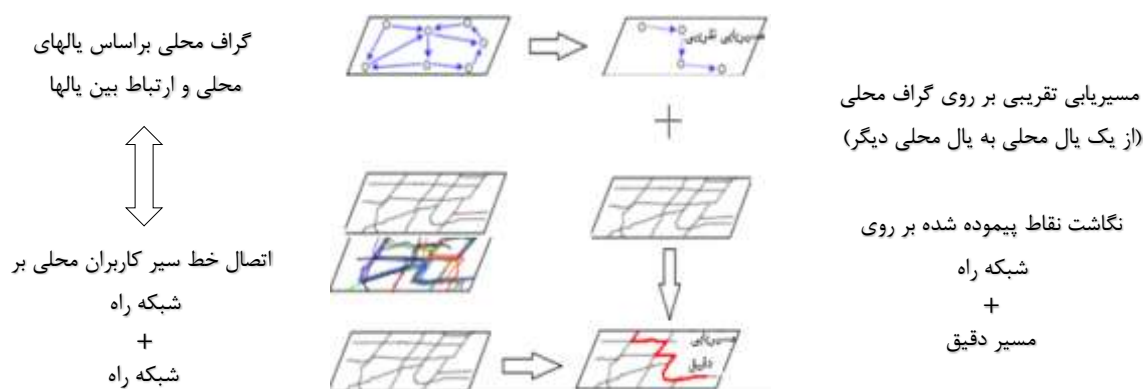
در سیستم‌های توصیه‌مسیر با توجه به شرایط ترافیکی و وضعیت راه‌ها، علاوه بر کوتاهی مسیر، نیاز است فاکتورهای همچون پیچیدگی مسیر، تقاطع‌ها، چراغ‌های راهنمایی، جریان ترافیک و غیره در انتخاب مسیر بهینه رعایت شوند. در هر منطقه افراد آشنا به منطقه یا به اصطلاح "کاربران محلی"، اغلب دارای دانش خوبی از مسیرهای مختلف هستند. رانندگان محلی معمولاً مسیرهایی انتخاب می‌کنند که علاوه بر کوتاهی، دارای کم‌ترین پیچیدگی و تغییر در جهت حرکت معابر، تقاطع‌ها و چراغ‌های راهنمایی باشد. در هر منطقه سرویس‌های توصیه‌مسیر می‌توانند با جمع‌آوری و کاوش خط سیر وسیله نقلیه کاربران و کشف کاربران محلی و راه‌های محلی بر اساس الگوریتم‌های داده‌کاوی، از دانش آن‌ها بهره‌مند شوند و برای مسیر یابی دقیق به کار ببرند. بنابراین در تحقیق حاضر، هدف مدل‌سازی سرویس توصیه‌مسیر بر پایه شناسایی کاربران و قطعه خیابان‌های محلی بر اساس تحلیل خط سیر کاربران است. با توجه به اهداف ذکر شده شکل (۱)، مراحل شکل‌گیری اجزای مورد استفاده مدل را نشان می‌دهد.

در حوزه سیستم‌های خودکار، [۱۵] یک سیستم توصیه‌مسیر بر اساس هستی‌شناسی و برآورد ترجیحات کاربران با استفاده از اطلاعات ناهمگن مکانی کاربران، توسط شبکه‌های بیزین ایجاد کردند. [۱۶] دنباله حرکتی کاربران را از تصاویر فلیکر که حاوی اطلاعات مکان و زمان است، استخراج کردند و مسیر سفر را بر اساس دنباله‌های حرکتی پرترفدار توصیه ارائه دادند. مقاله‌ی [۱۷]، سریعترین مسیر را با استفاده از الگوی رانندگی و سرعت راننده، با به‌کارگیری تاریخچه‌ی حرکت پیش‌بینی می‌کند. روش مورد استفاده در این مقاله از جنبه‌های زیر با [۱۷] متفاوت است: در [۱۷] به طور صریح به شناسایی الگوی رانندگی و سرعت وسایل نقلیه پرداخته است، اما در اینجا از مفهوم گراف محلی برای مدل کردن هوش رانندگان محلی استفاده شده است. گراف محلی از الگوی طبیعی فکری افراد آشنا به منطقه پیروی می‌کند و بهینه‌ترین مسیر را ارائه می‌کند. هدف این مقاله، توصیه خودکار مسیر بر اساس پردازش خط سیر کاربران محلی است. ورودی سیستم تنها دو نقطه به‌عنوان مبدأ و مقصد است. برای ارزیابی روش ارائه شده، از داده‌های GPS ایجاد شده توسط کاربران واقعی به‌منظور بهره‌وری خروجی الگوریتم



شکل ۱: مراحل شکل‌گیری اجزای مورد استفاده تحقیق

مسیر تقریبی به دست آمده از مرحله قبل به صورت آنی محاسبه می‌شود. خروجی این دو مرحله، یک مسیر بر پایه‌ی ترکیب طول مسیر و میزان ترجیحات کاربران محلی است. هرگاه خط سیر افراد بروز شود، گراف محلی مجدداً ساخته می‌شود. شکل (۲) به‌طور خلاصه مراحل مسیریابی دومرحله‌ای را نشان می‌دهد.



شکل ۲: مراحل روش پیشنهادی مسیریابی دومرحله‌ای

در اینجا ما روشی را برای تولید گراف محلی پیشنهاد داده‌ایم. در روش پیشنهادی، قطعه خیابان‌های شبکه‌ی راه، بر اساس خط سیرهای عبوری کاربران بر اساس یک الگوریتم رتبه‌بندی، رتبه‌دهی می‌شوند. قطعه خیابان‌هایی که دارای رتبه بالایی هستند یا به اصطلاح مورد تردد اکثر کاربران محلی قرار می‌گیرند، به عنوان قطعه خیابان‌های محلی تعریف می‌شوند. مرکز قطعه خیابان‌های محلی به عنوان گره‌های گراف محلی در نظر گرفته می‌شوند. هم‌چنین یال‌های گراف محلی که بیانگر ارتباط بین قطعه خیابان‌های محلی است، از خط سیر کاربران قابل بازیابی است. برای توسعه روش پیشنهادی نخست اصطلاحات مورد استفاده را به‌طور مختصر تعریف می‌کنیم، سپس به حل مسئله می‌پردازیم.

برای استفاده از خط سیر کاربران در روش پیشنهادی، باید نقاط GPS به نقشه تصویر شوند. یکی از چالش‌های عمده استفاده از خط سیر کاربران، پراکندگی داده و فاصله زیاد بین نقاط است. زیرا کاربران به علت عدم دسترسی به اینترنت در تمام لحظات و صرفه‌جویی در هزینه مخابراتی، موقعیت خود را در فرکانس پایین ارسال کرده که خود باعث عدم قطعیت می‌شود [۹]. در اینجا برای اتصال نقاط به نقشه از روش $IVMM^1$ که یکی از مناسب‌ترین روش‌ها

در اینجا ما روشی را برای تولید گراف محلی پیشنهاد داده‌ایم. در روش پیشنهادی، قطعه خیابان‌های شبکه‌ی راه، بر اساس خط سیرهای عبوری کاربران بر اساس یک الگوریتم رتبه‌بندی، رتبه‌دهی می‌شوند. قطعه خیابان‌هایی که دارای رتبه بالایی هستند یا به اصطلاح مورد تردد اکثر کاربران محلی قرار می‌گیرند، به عنوان قطعه خیابان‌های محلی تعریف می‌شوند. مرکز قطعه خیابان‌های محلی به عنوان گره‌های گراف محلی در نظر گرفته می‌شوند. هم‌چنین یال‌های گراف محلی که بیانگر ارتباط بین قطعه خیابان‌های محلی است، از خط سیر کاربران قابل بازیابی است. برای توسعه روش پیشنهادی نخست اصطلاحات مورد استفاده را به‌طور مختصر تعریف می‌کنیم، سپس به حل مسئله می‌پردازیم.

شبکه راه، به فرم گراف جهت‌دار $G = (V, E)$ تعریف می‌شود که در آن V شامل تقاطع‌ها و نقاط پایانی راه‌ها و E مجموعه‌ی قطعه خیابان‌ها است. هر قطعه خیابان r در مجموعه‌ی E به صورت $r.s \rightarrow$

¹ Interactive-voting based map matching

صفحات وب پیوند ارتباطی می‌زنند. یک صفحه مهم می‌تواند صفحات معتبری را در مورد یک موضوع خاص پیشنهاد دهد، و یک صفحه معتبر توسط بسیاری از صفحات مهم ارجاع داده شود. این رابطه مدور منجر به تعریف الگوریتم تکراری HITS می‌شود.

در پایگاه داده خط سیر کاربران، دو نوع مورد اطلاعاتی کاربر و قطعه خیابان وجود دارد که تردد هر کاربر از هر قطعه خیابان، به صورت یک یال ارتباطی جهت‌دار از کاربر به قطعه خیابان در نظر گرفته می‌شود. بر پایه مدل HITS هر کاربر به عنوان یک مرجع Hub و هر قطعه خیابان به عنوان یک مرجع Authority عمل می‌کند. با استفاده از این الگوریتم، محلی بودن قطعه خیابان‌ها و کاربران به ترتیب با محاسبه میزان اعتبار قطعه خیابان و اهمیت کاربران به صورت نسبی برآورد می‌شود. در این روش باید توجه داشته باشیم که میزان محلی بودن قطعه خیابان را تنها از تعداد افرادی که از آنجا تردد کرده‌اند، نمی‌توان نتیجه گرفت؛ بلکه می‌بایست بر اساس اینکه چه کسانی از آنجا تردد کرده‌اند، برآورد کرد [۲۰]. دانش و تسلط مکانی افراد، ناحیه‌ای است [۲۱]. بدین معنا که میزان تسلط افراد مختلف به یک منطقه مکانی متفاوت بوده و پر واضح است کاربرانی که آشنایی بهتری با منطقه دارند، مسیرهای بهتری را برای تردد انتخاب می‌کنند. اگر سطح آشنایی کاربران با منطقه در نظر گرفته نشود، نتیجه رتبه‌بندی، قطعه خیابان‌های پرتردد و پرتراфик توصیه می‌شود هم‌چنین میزان محلی بودن کاربران نسبی است و به ناحیه جغرافیایی وابسته است. در واقع ممکن است کاربری که در یک منطقه دانش و تسلط زیادی دارد، از مسیرهای بهینه منطقه دیگری هیچ اطلاعاتی نداشته باشد. بر این پایه، میزان اعتبار قطعه خیابان‌ها و اهمیت کاربران در یک منطقه خاص، با در نظر گرفتن وابستگی این دو مقدار به یکدیگر و وابستگی محلی بودن کاربران به نواحی مقاداردهی می‌شوند.

برای غلبه بر خطاهای GPS است، استفاده می‌شود [۱۸]. در این روش، هر نقطه GPS به نزدیک‌ترین قطعه خیابان در شبکه راه‌ها تصویر می‌شود و خط سیرها مجدداً به صورت دنباله‌ای از قطعه خیابان‌ها به فرم $T: r_1 \rightarrow r_2 \rightarrow \dots \rightarrow r_n$ تعریف می‌شوند.

اولین قدم برای مدل‌سازی ترجیحات کاربران محلی در انتخاب قطعه خیابان‌ها، فرموله کردن وزن قطعه خیابان‌ها و استخراج قطعه خیابان‌های محلی و انتخاب مراکز آن‌ها به عنوان گره‌های گراف است.

۳-۱- کشف کاربران و قطعه خیابان‌های محلی توسط الگوریتم HITS

افراد در حین ورود به مناطق ناآشنا، عموماً نیاز به انتخاب مسیرهایی دارند که آن‌ها را در یافتن سریع‌تر مقصدشان هدایت کند. در هر منطقه، مسیرهایی وجود دارند که برای اکثر افراد آشنا به منطقه، شناخته‌شده و بهینه است. برای کشف این مسیرها که در اینجا به عنوان مسیرهای محلی از آن‌ها یاد شده است، می‌بایست قطعه خیابان‌های شبکه راه بر اساس میزان تردد از آن‌ها رتبه‌دهی و با یکدیگر مقایسه شوند. یکی از روش‌های رتبه‌دهی، استفاده از الگوریتم‌های رتبه‌بندی است. الگوریتم HITS [۱۹]، یک الگوریتم رتبه‌بندی پرکاربرد در موتورهای جستجوی وب برای رتبه‌بندی و بازگرداندن اطلاعات مبتنی بر پرس و جوی انجام شده است که دو امتیاز برای هر کدام از صفحات وب برآورد می‌کند. (۱) میزان اهمیت^۱ (۲) میزان اعتبار^۲. یک صفحه معتبر، حاوی مطالب مورد درخواست و محتویات اطلاعاتی است که توسط دیگران استفاده می‌شود. یک صفحه با اهمیت بالا، کاربران را به جست‌وجوی صفحات معتبر هدایت می‌کنند و به صورت لیستی از منابع عمل می‌کنند و به بسیاری از

¹ Hub score

² Authority score

HITS برای تعیین میزان اعتبار یک صفحه در هر تکرار، تنها میزان اهمیت صفحات در ارتباط با آن را در نظر می‌گیرد. ما به جای ماتریس مجاورت، از یک ماتریس جدید که شامل اطلاعاتی در مورد پیوندهای ارتباطی با طول بزرگ‌تر است، استفاده می‌کنیم. برای در نظر گرفتن ارتباطات با طول بیش از ۱ توجه داشته باشید که ماتریس مجاورت A ، تعداد مسیرهای به طول m از گره i به گره j را در نظر می‌گیرد. در این روش با توجه به این قانون که تحت شرایط مشخصی وقتی $K \rightarrow \infty$ آنگاه $A^k X$ به بردار ویژه متناظر با بزرگ‌ترین مقدار ویژه A از نظر قدر مطلق میل می‌کند در هر مرحله بزرگ‌ترین مقدار ویژه از نظر قدر مطلق و بردار ویژه متناظر با آن به دست می‌آید و این کار تا رسیدن به دقت مورد نظر ادامه می‌یابد.

۳-۲- گراف محلی

پس از تعیین وزن قطعه خیابان‌ها $S(r_i)$ توسط الگوریتم HITS بر اساس میزان تردد کاربران محلی و تشکیل مجموعه‌ی $R = \{r_1, r_2, \dots, r_n\}$ به عنوان مجموعه‌ی قطعه خیابان‌های محلی، گراف جهت‌دار $G = (R, T)$ را به عنوان گراف محلی ایجاد می‌کنیم، به طوری که گره‌های گراف از مرکز قطعه خیابان‌های محلی و یال‌های گراف، از مجموعه خط سیرهای عبوری از بین قطعه خیابان‌های محلی قابل کشف است. با توجه به مجموعه‌ی R و مجموعه‌ی $\{E\}-\{R\}$ که بیانگر مجموعه‌ی قطعه خیابان‌های موجود در شبکه راه که غیر محلی هستند، ارتباط بین قطعه خیابان‌های محلی از طریق کاوش مجموعه‌ی خط سیرها قابل کشف است. به طور مثال با توجه به جدول (۱)، قطعه خیابان‌های محلی r_1 و r_7 از طریق پیمایش T_1 و توسط قطعه خیابان‌های غیر محلی r_2 و r_4 به یکدیگر مرتبط می‌شوند و یال e_{17} را در گراف محلی ایجاد می‌کنند. این یال در واقع ارتباط بین گره‌های v_1 و v_3 را در گراف محلی ایجاد کرده است.

برای اجرای مدل HITS، یک ماتریس همسایگی M بین کاربران (u_i) و قطعه خیابان‌ها (r_i) تابع ترددی که کاربران از قطعه خیابان‌ها داشته‌اند، ساخته می‌شود. در این ماتریس مقدار V_l^k بیانگر تعداد دفعاتی است که کاربر u_k از قطعه خیابان r_l بازدید کرده است. میزان اعتبار قطعه خیابان، به صورت A_l و اهمیت کاربر u_k به صورت H_k تعریف می‌شود [۱۹].

$$M = \begin{matrix} & r_1 & \dots & r_n \\ u_1 & \left[\begin{array}{ccc} V_1^1 & \dots & V_n^1 \\ \vdots & \ddots & \vdots \\ V_1^n & \dots & V_n^n \end{array} \right] \\ u_n & & & \end{matrix}$$

ارتباط متقابل و وابستگی میزان اعتبار قطعه خیابان‌ها و میزان اهمیت کاربران با روابط زیر بیان می‌شود:

$$A_l = \sum_{u_k \in U} H_k \times V_l^k \quad \text{رابطه (۱)}$$

$$H_k = \sum_{u_k \in U} A_l \times V_l^k \quad \text{رابطه (۲)}$$

J به عنوان ستونی است که المان‌های آن بیانگر میزان اعتبار قطعه خیابان‌ها است و E نیز به عنوان سطری که المان‌های آن بیانگر میزان اهمیت برآورده شده برای کاربران است. رابطه‌های (۳) و (۴) می‌تواند به صورت زیر بازنویسی شود [۱۹]:

$$J = M^T \cdot E \quad \text{رابطه (۳)}$$

$$J = E^T \cdot M \quad \text{رابطه (۴)}$$

فرایند تکراری محاسبه اهمیت افراد و اعتبار قطعه خیابان‌ها به فرم زیر است [۱۹]:

$$E_n = M \cdot M^T \cdot E_{n-1} = (M \cdot M^T)^n \cdot E_0 \quad \text{رابطه (۵)}$$

$$J_n = M^T \cdot M \cdot J_{n-1} = (M^T \cdot M)^n \cdot J_0 \quad \text{رابطه (۶)}$$

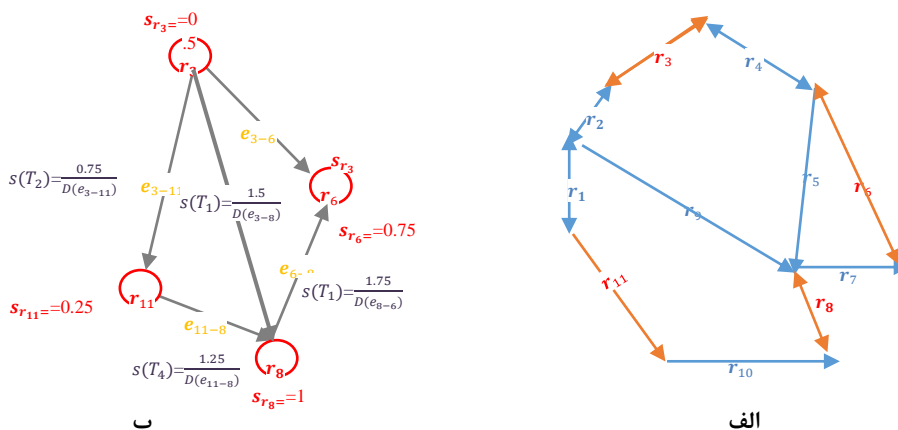
با استفاده از روش توانی تکرارشونده و با فرض $J_0 = E_0 = (1, 1, 1, \dots, 1)$ می‌توان اهمیت کاربران و اعتبار قطعه خیابان‌ها را برآورد کرد و بر اساس این مقادیر، k کاربر محلی و قطعه خیابان بهینه را در ناحیه جست‌وجوی استخراج کرد.

آنها به عنوان گره گراف محلی در نظر گرفته می شود. در شکل (۳-ب) که نشانگر گراف محلی است، دایره قرمز رنگ، گره گراف و خطوط تیره رنگ یال گراف است که وزن آنها از میاگین وزن گره های ورودی و خروجی و فاصله بین گره ها به دست می آید.

همانطور که در شکل (۳-الف) مشاهده می شود، خطوط جهت دار r_i بیانگر قطعه خیابانهای پیموده شده توسط کاربران است، که از خط سیر آنها استخراج شده و خطوط قرمز رنگ قطعه خیابانهای پرتردد هستند (دارای وزن بالاتری هستند) که در مرحله بعد مرکز

جدول ۱: خط سیرها به عنوان دنباله ای از قطعه خیابانهای محلی

شناسه خط سیر	دنباله ای از قطعه خیابانهای محلی	یال گراف محلی
T_1	$r_8 \rightarrow r_7 \rightarrow r_6$	e_{8-6}
T_2	$r_3 \rightarrow r_3 \rightarrow r_1 \rightarrow r_{11}$	e_{3-11}
T_3	$r_3 \rightarrow r_2 \rightarrow r_9 \rightarrow r_8$	e_{3-8}
T_4	$r_1 \rightarrow r_{11} \rightarrow r_{10} \rightarrow r_8$	e_{11-8}



شکل ۳: الف) خط سیر کاربران، ب) گراف محلی

آن قطعه خیابانهای محلی به صورت ترتیبی پیموده شده اند. اگر هیچ خط سیری بین قطعه خیابانهای محلی یافت نشود، روش خوشه بندی نمی تواند مسیری پیدا کند و در نتیجه از کوتاه ترین مسیر مستقیم بین میان مبدأ و مقصد، استفاده می کند.

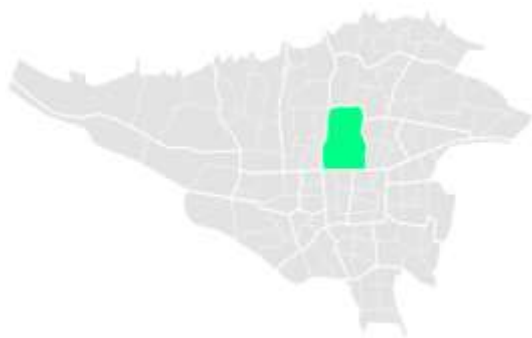
به بیان دیگر وزن یالهای محلی، به صورت ترکیبی از طول مسیر و میزان ترجیحات کاربران محلی (وزن داده شدن توسط کاربران محلی به قطعه خیابانها)، با توجه به معادله (۷) قابل محاسبه است [۲۲].

۴- پیاده سازی

بخش حاضر به پیاده سازی مدل و الگوریتمهای ارائه شده اختصاص دارد. جهت پیاده سازی، ابتدا منطقه مورد مطالعه و داده های مورد نیاز معرفی می شوند، سپس نتایج عملی نمونه پیاده سازی شده، ارائه و ارزیابی می شوند.

رابطه (۷)
$$S(T) = \frac{(S(r_s) + S(r_e)) / 2}{\text{distance}(r_s, r_e)}$$
 وزن گره ورودی گراف محلی، $S(r_e)$ وزن گره خروجی و $\text{distance}(r_s, r_e)$ فاصله اقلیدسی بین گره ها به واحد کیلومتر است. با تعیین مبدأ و مقصد، در مرحله اول یک مسیریابی تقریبی بر روی گراف محلی برای هدایت افراد برای عبور از قطعه خیابانهای محلی، انجام می شود. سپس توسط الگوریتم دایجسترا کوتاه ترین مسیری در شبکه راهها محاسبه شده که در

ابتدا سرویس نقشه Bing Map به‌عنوان نقشه پایه استفاده شده و لایه‌های پایه موردنیاز کاربران که شامل شبکه راه‌ها و نقشه‌ی بلوک‌های ساختمانی است و لایه‌های مربوط به مکان‌های به اشتراک گذاشته شده توسط کاربران، معابر محلی و شاخص‌ها در قالب سرویس مکانی WMS در دسترس کاربران قرار می‌گیرد.



الف



ب

شکل ۴: توزیع مکانی داده‌های شبکه اجتماعی مکان‌مبنا در

منطقه مورد مطالعه

(الف: منطقه مورد مطالعه تحقیق، ب: خط سیر کاربران)

جدول ۲: مشخصات مجموعه داده

سرعت (KM/H)	بازه مکانی برداشت (M)	بازه زمانی برداشت (S)	
۷.۸۵۷۷	۸.۹۸	۳/۳۴۵۷	متوسط
۷۱.۹۴۹۷	۵۸۴	۴۲	بیشینه
۰.۰۷	۵	۱	کمینه

۴-۱- منطقه مورد مطالعه و داده‌های مورد استفاده

در بین مناطق ۲۲ گانه تهران، منطقه ۶ به علت دارا بودن خوابگاه‌های دانشجویی و بخش‌های اداری و تجاری بسیار، از مناطق پررفت و آمد تهران به حساب می‌آید که به‌عنوان منطقه مورد مطالعه تحقیق انتخاب گردید. برای پیاده‌سازی از داده‌های موجود در سیستم‌های ندریت ناوگان که توسط شرکت‌های مختلف^۱ ارائه می‌شود، استفاده کردیم و مجموعه خط سیر ۶۰ کاربر متحرک مجهز به وسیله نقلیه شخصی را به صورت روزانه در بازه زمانی سه‌ماهه از مهرماه تا پایان آذرماه سال ۱۳۹۳ اخذ نمودیم. خط سیر کاربران سواره، توسط تلفن‌های همراه مجهز به GPS در نرم‌افزار *My Tracks* ثبت گردید. مجموعه خط سیر شامل ۱,۰۷۴,۵۱۱ نقطه GPS و ۳۱۲۶ خط سیر منحصربه‌فرد است. ۹۱.۵ درصد خط سیرها به صورت متراکم و هر ۱ تا ۵ ثانیه و یا هر ۵ تا ۱۰ متر نمونه‌برداری شده‌اند. شکل (۴) نمایشی از داده‌های مکانی کاربران را بر روی نقشه نمایش می‌دهد.

با اصلاح و آماده‌سازی داده‌ها بر اساس مشخصات خط سیر، خط سیر که حداقل دارای ۱۰۰ و حداکثر ۱۴۸۹ نقطه برداشت شده و طولی بین ۲۸ تا ۱۲۰۰۰ متر دارند، برای پیاده‌سازی انتخاب شده‌اند که مشخصات آن‌ها در جدول (۲) آمده است.

۴-۲- توسعه سرویس توصیه مسیر

در این تحقیق به منظور پیاده‌سازی سرویس توصیه مسیر از API سیلور لایت ۵ موجود در Visual Studio 2012 برای اجرای آنالیزهای مکانی و از نرم‌افزار ArcGIS Server 10.2 به‌عنوان سرویس‌دهنده نقشه برای تبادل داده‌ها و خدمات GIS استفاده شده است. محیط برنامه در بستر وب پیاده‌سازی شده است. پیاده‌سازی سرویس به صورت گام‌به‌گام انجام می‌گیرد.

^۱ <http://www.gpstrack.ir>

کوتاه‌ترین مسیر بین قطعه خیابان‌های محلی با استفاده از الگوریتم دیکسترا بر روی شبکه راه‌ها نشان داده می‌شود. در سرویس پس از اجرای مسیریابی تقریبی بر روی گراف محلی و ارائه مسیر محلی به صورت غیردقیق، با زدن گزینه‌ای مسیریابی دقیق بر روی شبکه راه‌ها اجرا می‌شود. در شکل (۶)، خطوط قرمز رنگ بیانگر مسیر تقریبی و خطوط آبی‌رنگ بیانگر مسیر دقیق بر روی شبکه راه‌ها است.

۴-۳- ارزیابی

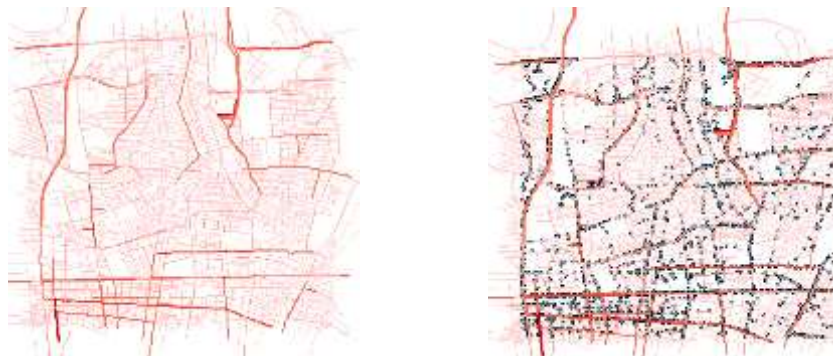
برای ارزیابی عملکرد روش پیشنهادی، نتیجه حاصل از مسیریابی دومرحله‌ای بر روی گراف محلی با نتایج به‌دست آمده از روش کوتاه‌ترین مسیر مقایسه می‌شود. برای مقایسه نتایج حاصل از روش‌های مذکور، از پارامتر زمان سفر استفاده می‌کنیم. این پارامتر تابعی از ویژگی‌های مسیر و خصوصیات کاربر هدف است. برای کسب ویژگی‌های مسیر تمرکز ما در اینجا به‌طور خاص بر روی طول مسیر است. در واقعیت با افزایش طول مسیر، احتمال ورود کاربر به مناطق نا آشنا، تعداد قطعه خیابان‌ها و نقاط تصمیم‌گیری بیشتر می‌شود و می‌توان زمان سفر بیشتری را در طول فرایند مسیریابی انتظار داشت. برای انجام آزمون، از ۱۰ کاربر سواره نا آشنا به منطقه ۶ تهران، درخواست کردیم تا به صورت تجربی مسیرهای به‌دست آمده از دو روش مذکور را بپیمایند و زمان سفر را در هر دو روش اندازه بگیرند (شکل (۷)).

برای تعیین قطعه خیابان‌های محلی یک وب‌سرویس WCF بر اساس مدل HITS، باهدف تشخیص میزان اهمیت کاربران و اعتبار قطعه خیابان‌ها طراحی گردیده است. ورودی وب‌سرویس HITS، ماتریس کاربر-قطعه خیابان است که درایه‌های آن را تعداد خط سیرهای عبوری کاربران از قطعه خیابان‌ها تشکیل می‌دهند و خروجی سرویس، یک عدد بین صفر و یک است. به ازای هر بار اجرا شدن سرویس HITS، مقادیر این ستون برای هر قطعه خیابان و هر کاربر در پایگاه داده مرکزی به‌روزرسانی می‌شود. پس از اجرای سرویس HITS، قطعه خیابان‌ها و کاربرانی که به ترتیب دارای اعتبار و اهمیت بیشتر از ۰.۵ هستند، به‌عنوان قطعه خیابان‌ها و کاربران محلی شناخته می‌شوند. پس از آن فرایند مسیریابی در دو مرحله صورت می‌گیرد. در مرحله اول پس از اجرا و پیاده‌سازی روند مرحله به مرحله گراف محلی، مطابق شکل (۵) گراف محلی در منطقه ۶ تهران تولید شد. در شکل (۵-ب) خطوط قرمز رنگ، نشان‌دهنده قطعه خیابان‌هایی است که بیشتر از همه مورد تردد کاربران محلی قرار گرفته است. با توجه به شکل (۵-ج) نقاط قرمز رنگ، مراکز این قطعه خیابان‌ها بوده که به‌عنوان گره‌های گراف محلی استفاده می‌شوند و خطوط طوسی رنگ نشان‌دهنده یال‌های گراف محلی است.

پس از آن، کاربر سرویس با تعیین مبدأ و مقصد بر روی نقشه درخواست مسیریابی می‌کند، ابتدا یک درخواست به‌طور آنی به سرویس آنالیز شبکه بر روی گراف محلی ارسال می‌گردد. در این مرحله یک مسیر تقریبی،

ارائه سرویس توصیه مسیر بر اساس رتبه‌بندی سیر...

رویا شورونی، محمدرضا ملک

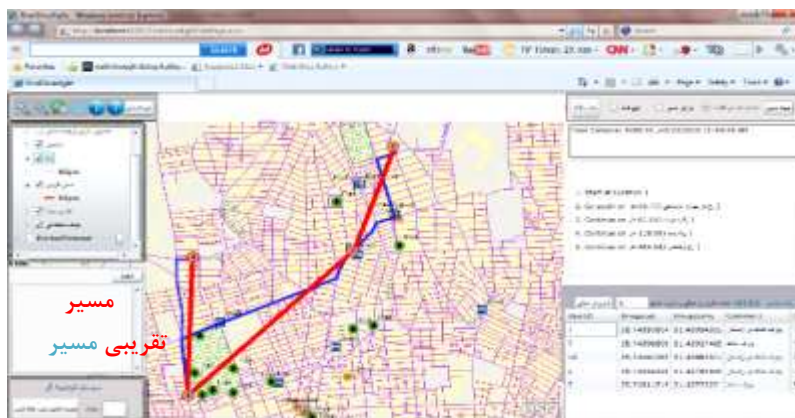


الف) اتصال نقاط GPS بر روی نقشه ب) وزن دهی قطعه خیابان‌ها بر اساس میزان تردد افراد محلی

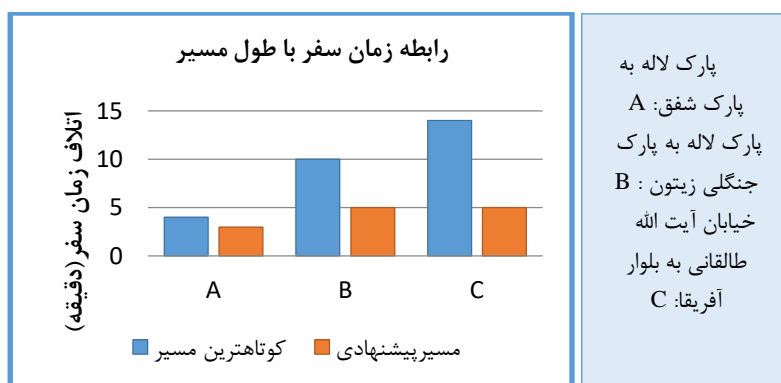


ج) تعیین گره‌ها و بال‌های گراف محلی

شکل ۵: مراحل ایجاد گراف محلی در منطقه ۶ تهران



شکل ۶: نمایی از مسیریابی دومرحله‌ای در سرویس توصیه مسیر



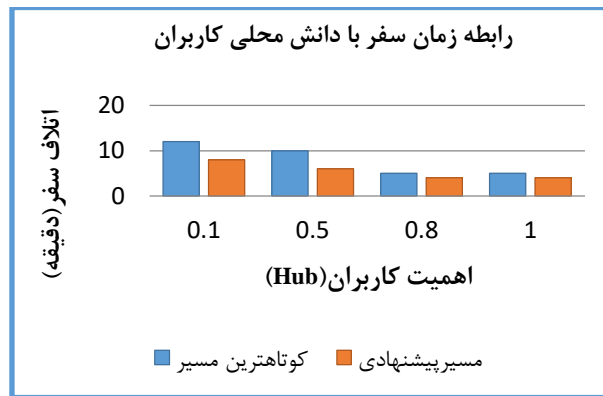
پارک لاله به
 پارک شفق: A
 پارک لاله به پارک
 جنگلی زیتون: B
 خیابان آیت الله
 طالقانی به بلوار
 آفریقا: C

شکل ۷: بررسی رابطه زمان سفر با طول مسیر برای دو روش مسیریابی

مختلف برآورد کرد (شکل ۷)). هر چه میزان محلی بودن کاربر در یک منطقه بیشتر باشد، میزان آشنایی او از قطعه خیابان‌های محلی بیشتر خواهد بود و در نتیجه روش پیشنهادی نسبت به روش کوتاه‌ترین مسیر، از عملکرد بهتری برخوردار خواهد بود. در اینجا میزان محلی بودن فرد در محدوده مورد ارزیابی، به‌واسطه رتبه‌بندی کاربران توسط الگوریتم HITS برآورد می‌شود. محدوده مورد ارزیابی از شمال شرق به بوستان نظامی گنجوی و از جنوب غرب به خیابان جمال‌زاده در محدوده بلوار کشاورز غربی منحصر می‌شود و کوتاه‌ترین مسیر بین این دو مکان حدود ۸۰۰۰ متر است.

از آنجایی که در روش کوتاه‌ترین مسیر، وزن دهی قطعه خیابان‌ها تنها بر اساس طول مسیر است، در مقایسه با روش پیشنهادی که ترکیب طول مسیر و ترجیحات کاربران محلی است زمان سفر بیشتری انتظار می‌رود. در فواصل کوتاه، به علت اندک بودن قطعه خیابان‌های محلی، روش پیشنهادی تفاوت چندانی با کوتاه‌ترین مسیر ندارد. اما با افزایش طول مسیر به علت افزایش قطعه خیابان‌های محلی، روش پیشنهادی نسبت به روش دایجسترا عملکرد بهتری دارد و زمان سفر به‌طور میانگین ۵۳ درصد کاهش می‌یابد.

برای محاسبه رابطه زمان سفر با ویژگی‌های کاربر، می‌توان به میزان محلی بودن کاربران در یک محدوده جغرافیایی اشاره کرد و تأثیر این ویژگی را در مسیرهای



شکل ۸: رابطه زمان سفر با دانش منطقه‌ای کاربران برای دو روش مسیریابی

بتوان با استفاده از مشارکت مردم، حجم کثیری از این داده‌ها را جمع‌آوری کرد. بنابراین کاربران با استفاده از GPS گوشی‌های همراه خود، به‌عنوان سنجنده‌های متحرکی عمل می‌کنند که ترافیک موجود در خیابان‌ها را کاوش می‌کنند. در نتیجه، با دسترسی و تحلیل تاریخچه حرکتی وسایل نقلیه در محدوده‌های جغرافیایی مختلف، تشخیص الگوی حرکتی وسایل نقلیه در راه‌ها و استخراج مسیرهای بهینه حاصل می‌شود. در این سناریو کاربران در ازای جمع‌آوری داده‌های خط سیر، آنالیزهای مهم سیستم اطلاعات مکانی نظیر یافتن بهترین مسیر را به‌صورت رایگان دریافت می‌کنند. استفاده از چنین سیستمی می‌تواند علاوه بر ترغیب کاربران جهت جمع‌آوری خط سیر، نیازهای آنان را به‌منظور هدایت مناسب در شبکه حمل و نقل برطرف نماید.

مقاله حاضر به ارائه یک سرویس توصیه مسیر، بر مبنای ترکیب ترجیحات کاربران محلی و آشنا به منطقه در انتخاب راه‌ها و طول مسیر می‌پردازد. روند کار سرویس مسیریابی به این صورت است که مسیر بهینه بر روی گراف محلی، گراف وزن‌دار به‌دست‌آمده از خط سیر کاربران ارائه می‌شود. بر این اساس زمان سفر در این روش نسبت به روش کوتاه‌ترین مسیر برای طول‌های مختلف و کاربران با سطوح آگاهی مختلف به‌طور میانگین حدود ۶۲ درصد کاهش یافت.

مطابق شکل (۸) برای کاربران آشنا به منطقه به علت وجود تعداد زیادی از قطعه خیابان‌های آشنا، خروجی سرویس توصیه مسیر به علت عبور از قطعه خیابان‌های محلی نسبت به روش کوتاه‌ترین مسیر بهتر عمل می‌کند. برای کاربران ناآشنا به منطقه مسیر به‌دست‌آمده از سرویس توصیه مسیر نسبت به مسیر حاصله از آن روش دایجسترا از زمان سفر کمتری برخوردار است. به‌طور کلی می‌توان گفت با افزایش میزان تسلط کاربران به منطقه، زمان سفر سرویس توصیه مسیر، نسبت به کوتاه‌ترین مسیر حدود ۷۱ درصد کاهش یافت.

۵- نتیجه‌گیری

افزایش وسایل نقلیه در تهران و عدم برنامه‌ریزی مناسب برای به کنترل درآوردن سفرهای درون‌شهری، باعث شده همواره اتلاف زمان سفر و ترافیک در رأس مسائل لاینحل قرار گرفته و از اهم آن به‌شمار برود. از این رو ارائه یک سرویس توصیه مسیر مناسب جهت جلوگیری از اتلاف زمان و سردرگمی مسافران ضروری است. به‌منظور جمع‌آوری داده‌های مربوط به توصیه مسیر بهینه با توجه به شرایط ترافیکی، از داده‌های به‌نگام اخذشده از سنجنده‌های موجود در راه‌ها استفاده می‌شود که مشکل عمده‌ی این روش، پراکندگی و پوشش کم منطقه‌ای سنجنده‌ها است. مجهز شدن تجهیزات همراه به دستگاه گیرنده GPS، نویدبخش شرایطی است که

ممکن برای این مشکل استفاده از سه مسیر برتر (به جای بهترین مسیر) برای متعادل کردن بار و همچنین به روزرسانی داده‌های موجود در پایگاه داد با سرعت بالا می‌باشد. همچنین تلفیق خط سیر کاربران با سنجنده‌های ترافیکی می‌تواند نتایج حاصل از مسیریابی را بهبود بخشد.

با وجود اینکه روش مورد استفاده در این مقاله بر پایه نظرات کاربران آشنا به منطقه است، می‌تواند جزء یکی از بهترین روش‌های یافتن سریع‌تر مسیر باشد، اما همواره مشکلات زیر در تمام روش‌های مسیریابی وجود دارند: مسیری که بیشترین پیشنهاد به کاربر را دارد، مسیری که بیشترین پیشنهاد به کاربران را دارد، بعد از مدتی شلوغ و پر ترافیک می‌شود. راه‌حل‌های

مراجع

- [1] Bao, Jie, Yu Zheng, David Wilkie, and Mohamed F. Mokbel. "A survey on recommendations in location-based social networks." *ACM Transaction on Intelligent Systems and Technology*, 2013.
- [2] Monreale A, Pinelli F, Trasarti R, Giannotti F (2009) WhereNext: a location predictor on trajectory pattern mining. In: *Proceedings of KDD 2009*, pp. 637–646. doi:10.1145/1557019.1557091.
- [3] Giannotti F, Nanni M, Pinelli F, Pedreschi D (2007) Trajectory pattern mining. In: *Proceedings of KDD 2007*, pp 330–339. doi:10.1145/1281192.1281230.
- [4] Krumm J (2010) Where will they turn: predicting turn proportions at intersections. *Pers Ubiquit Comput* 14:591–599. doi:10.1007/s00779-009-0248-1
- [5] Cao L, Krumm J (2009) From GPS traces to a routable road map. In: *Proceedings of GIS 2009*, pp 3–12. doi:10.1145/1653771.1653776
- [6] Li Q, Zheng Y, Xie X, Chen Y, Liu W, Ma WY (2008) Mining user similarity based on location history. In: *Proceedings of GIS 2008*, pp1–10. doi:10.1145/1463434.1463477
- [7] Ceikute, V., & Jensen, C. S., "Routing Service Quality--Local Driver Behavior Versus Routing Services", In *Mobile Data Management (MDM)*, 2013 IEEE 14th International Conference on (Vol. 1, pp. 97-106), 2013.
- [8] Chen, Z., Shen, H. T., & Zhou, X, "Discovering popular routes from trajectories", In *Data Engineering (ICDE)*, IEEE 27th International Conference on (pp. 900-911), 2011.
- [9] Cao L, Krumm J (2009) From GPS traces to a routable road map. In: *Proceedings of GIS 2009*, pp 3–12. doi:10.1145/1653771.1653776
- [10] Yuan, Jing, et al. "T-drive: enhancing driving directions with taxi drivers' intelligence." *Knowledge and Data Engineering, IEEE Transactions on* 25.1 (2013): 220-232.
- [11] Zheng, Y., Zhang, L., Xie, X., Ma, W.: Mining interesting locations and travel sequences from gps trajectories. In: *WWW*, pp. 791–800. ACM (2009).
- [12] Yoon, H., Zheng, Y., Xie, X., & Woo, W. (2012). Social itinerary recommendation from user-generated digital. *Pers Ubiquit Comput*, 469–484.
- [13] Dunstall S, Horn MET, Kilby P, Krishnamoorthy M, Owens B, Sier D, Thiebaut S (2003) An automated itinerary planning system for holiday travel. *Inf Technol Tour*. 6:195–210.
- [14] Ardissono L, Goy A, Petrone G, Segnan M (2005). A multi-agent infrastructure for developing personalized web-based systems. *ACM Trans Internet Tech* 5:47–69. doi:10.1145/1052934.1052936.
- [15] Huang Y, Bian L (2009) A Bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the Internet. *Expert Syst Appl* 36:933–943. doi:

10.1016/j.eswa.2007.10.019

- [16] Chodhury MD, Feldman M, Amer-Yahia S, Golbandi N, Lempel R, Yu C (2010) Automatic construction of travel itineraries using social breadcrumbs. In: Proceedings of HT 2010, pp 35–44. doi: 10.1145/1810617.1810626
- [17] Gonzalez, H., Han, J., Li, X., Myslinska, M., & Sondag, J. P. (2007, September). Adaptive fastest path computation on a road network: a traffic mining approach. In Proceedings of the 33rd international conference on Very large data bases (pp. 794-805). VLDB Endowment.
- [18] Yuan, J., Zheng, Y., Zhang, C., Xie, X., & Sun, G. Z., "An interactive-voting based map matching algorithm", In *Proceedings of the 2010 Eleventh International Conference on Mobile Data Management* (pp. 43-52). IEEE Computer Society, 2010.
- [19] Kleinberg, J. M., "Authoritative sources in a hyperlinked environment", *Journal of the ACM (JACM)*, 46(5), 604-632, 1999.
- [20] Zheng, Y., & Zhou, X, Computing with spatial trajectories. Springer Science & Business Media, 2011.
- [21] Liben-Nowell, D., Novak, J., Kumar, R., Raghavan, P., and Tomkins, A. (2005) Geographic routing in social networks, Proceedings of the National Academy of Sciences of the United States of America 102, 11623.
- [22] Chang, K. P., Wei, L. Y., Yeh, M. Y., & Peng, W. C. (2011, November). Discovering personalized routes from trajectories. In Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks (pp. 33-40). ACM.



Route recommendation based on local users' trajectories

Roya Shurouni^{1*}, Mohammadreza Malek²

1- M.Sc. student of geographic information system in Department of Geomatics, College of Engineering, Khaje Nasir Toosi University.

2- Associate professor in Department of Geomatics, College of Engineering, Khaje Nasir Toosi University.

Abstract

Large amount of users' trajectories, is an emerging source of inexpensive data that can be used to provide an opportunity to present route recommendation service to the unfamiliar users within the area. In this study, with the aim of finding the optimal route, we first extract both local users and local road segments data sets by ranking them via HITS algorithm. In this model, a hub is a user who many time has crossed many road segments of a region, and an authority is a road segment that has been crossed by many users. Therefore, users' travel experiences (hub scores) and the interests of road segments (authority scores) have a mutual reinforcement relation. We also propose a novel approach in which the basic unit of routing is separate road segment instead of GPS trajectory segment. Moreover, to provide the approximate routing, we create a local graph. The center of the local road segments are considered as nodes and are based on local streets sequence arrange the pieces obtained from the trajectory of the user as edges of local graph. According to this graph, two steps of routing are used to obtain the optimal path. Then using Dijkstra's algorithm on the main road network and obtained an approximate route, shortest route between two local road segments based on this graph is used to obtain the optimal route. To implement and test, used data, from the trajectories of moving users in Tehran, has been gathered for 3 months on daily basis. To evaluate performance of the two-step routing, we experimentally compared the travel time in proposed route to Dijkstra's shortest path for different lengths and users with different levels of regional knowledge. The travel time in the proposed method was decreased 60 percent compare to shortest route.

Key words: Recommendation system, Route planning, Trajectory collection, Ranking.

Correspondence Address : GIS Group, Department of Geomatics, College of Engineering, University of Tehran, Tehran, Iran

Tel : +98 912 5646641

Email : Shourouni@mail.kntu.ac.ir