نشریه علمی پژوهشی مهندسی فناوری اطلاعات مکانی

سال پنجم، شماره دوم، تابستان ۱۳۹۶

Vol.5, No.2, Summer 2017

۳۵-۵۵

طراحی و توسعه یک سیستم چندعامله به‌منظور استخراج خودکار هندسه راه توسط عامل‌های کاوشگر

رضا محمدی[۱]، مهدی فرنقی[۲*]

۱- دانشجوی کارشناسی ارشد سیستم‌های اطلاعات مکانی دانشگاه خواجه‌نصیرالدین طوسی

۲- استادیار گروه سیستم‌های اطلاعات مکانی دانشگاه خواجه‌نصیرالدین طوسی

تاریخ دریافت مقاله ۱۳۹۵/۰۱/۳۰ تاریخ پذیرش مقاله: ۱۳۹۵/۰۳/۰۸

چکیده

نقشه‌های رقومی بهنگام نقش بسزایی در ارائه خدمات باکیفیت در کاربردهای مختلف سیستم¬های خدمات رسانی مرتبط با مکان نظیر ناوبری، توریسم، مدیریت ترافیک و سامانه‌های مکان‌مبنا ایفا می¬کنند. امروزه استفاده از داده‌های مردم¬گستر، به‌منظور تولید نقشه‌های رقومی مورد توجه قرارگرفته است. روش‌های ارائه‌شده به‌منظور استخراج نقشه رقومی راه با استفاده از داده‌های مردم¬گستر، عموماً ایستا بوده و به تغییرات پیوسته شبکه راه‌ها که توسط داده‌های خط سیر کاربران گزارش می‌شوند توجه نمی‌کنند. ایستا بودن روش‌های ارائه‌شده به‌منظور استخراج خودکار راه باعث می‌شود که بهنگام رسانی نقشه‌های راه مردم¬گستر به‌طور بهینه صورت نگیرد. استفاده از عامل‌ها به‌دلیل قابلیت حس محیط و توانایی عکس‌العمل در برابر تغییرات محیط، تولید نقشه به‌صورت پویا را امکان¬پذیر می‌کند. این پژوهش قصد دارد روشی پویا به‌منظور بهنگام رسانی نقشه‌های راه بر مبنای عامل‌های کاوشگر ارائه دهد. در این مقاله با الهام از نحوه پایش پیوسته محیط وب توسط عامل¬های کاوشگر وب، روشی پویا به‌منظور بهنگام رسانی نقشه‌های شبکه راه، با در نظر گرفتن تغییرات محیط که توسط خط سیرهای جمع‌آوری‌شده کاربران گزارش می‌شوند، ارائه شد. بدین منظور عامل‌هایی توسعه داده شد که با حرکت بر روی داده‌های خط سیر به استخراج راه می‌پردازند. روش خوشه‌بندیEM به‌منظور تعیین گره¬¬های شبکه راه مورد استفاده قرار گرفته و سپس روشی ابتکاری به‌منظور اتصال گره¬های مرتبط با یکدیگر ارائه شده است. راه‌کار ارائه شده امکان استخراج راه در فضای سه‌بعدی داده‌های خط سیر را فراهم می‌آورد. در پایان نتایج حاصل از الگوریتم ارائه شده به‌صورت عددی با نقشه مبنای OSM مورد ارزیابی قرار گرفته است.

واژه‌های کلیدی: داده‌های مردم¬گستر، خط سیر، استخراج خودکار راه، ¬عامل، عامل کاوشگر.

* نویسنده مکاتبه کننده: : تهران - خیابان ولیعصر - تقاطع میرداماد - روبروی ساختمان اسکان.

تلفن: ۱-۳۸۸۸۷۷۰۷

Email: farnaghi@kntu.ac.ir

۱- مقدمه

بسیاری از کاربردها نظیر ناوبری، توریسم، مدیریت ترافیک، مدیریت ناوگان‌های شهری و بین‌شهری، سامانه‌های حمل‌ونقل هوشمند، سرویس‌های نقشه تحت وب و سامانه‌های مکان‌مبنا [۱، ۲ و ۳] به داده‌های به‌هنگام نیاز دارند. لذا ارائه روشی جایگزین به‌منظور تولید و به‌هنگام رسانی اطلاعات مکانی شبکه راه‌ها به‌صورت خودکار ضروری به‌نظر می‌رسد. از طرفی امروزه تعداد استفاده‌کنندگان از شبکه‌های اجتماعی[1] رو به افزایش است. این شبکه‌ها بر پایه‌ی فناوری وب ۲ و در تعامل پیوسته با کاربران ایجادشده‌اند. در چنین شرایطی هر کاربر، به‌عنوان یک سنجنده[2] عمل کرده، داده‌های مکانی را جمع‌آوری و برای سرور مرکزی ارسال می‌کند [۴]. از این طریق منابع اطلاعات مکانی شبکه‌های اجتماعی به‌صورت پیوسته تغذیه و غنی‌سازی می‌شود.

داده‌های مردم‌گستر می‌توانند به‌عنوان تکمیل‌کننده داده‌های مکانی استاندارد مورداستفاده قرار گیرند [۴]. این داده‌ها حاوی اطلاعاتی فراتر از اطلاعات مکانی موجود در داده‌های استاندارد بوده و می‌توانند به‌منظور تحلیل رفتاری-مکانی کاربران، که بر روی‌داده‌های استاندارد امکان‌پذیر نیست، مورداستفاده قرار گیرند [۵]. همچنین داده‌های مردم‌گستر، به خاطر ماهیت پویایی آن‌ها، معمولاً بهنگام هستند و نسبت به داده‌های مکانی استاندارد بهتر می‌توانند تغییرات دنیای واقعی را نمایش دهند [۵]. این ویژگی‌های مهم، اهمیت استفاده و به‌کارگیری از این منابع ارزشمند را تبیین می‌کنند.

فعالیت‌های زیادی درزمینه استخراج خودکار راه از داده‌های مردم‌گستر صورت گرفته است. محققین بسیاری [۶، ۷، ۸، ۹ و ۱۰] فضای هندسی داده‌های GPS را شبکه‌بندی نموده، سپس توزیع نمونه‌ها در هریک از پیکسل‌ها را شمارش و بر اساس آن یک تصویر ایجاد کرده‌اند. در ادامه با اعمال یک حد آستانه پیکسل‌های مربوط به شبکه راه از سایر پیکسل‌ها جدا شده و درنهایت از روش‌های مختلف پردازش تصویر به‌منظور کشف مرکز راه‌ها استفاده کرده‌اند. کائو و کرام [۱۱] و گائو و همکاران [۱۲] با مدل‌سازی اثر نیروهای داده‌های خط سیر[3] بر یکدیگر، اقدام به استخراج هندسه و جهت راه نمودند. همچنین، به‌منظور کاهش نویز داده‌ها، که می‌تواند باعث ایجاد راه‌های کاذب شود، پالایه[4] ویژه‌ای اعمال نمودند. برخی از محققین [۱۳، ۱۴، ۱۵، ۱۶، ۱۷، ۱۸ و ۱۹] از روش‌های خوشه‌بندی به‌منظور استخراج راه استفاده‌نمودند. به این صورت که مجموعه‌ای از نقاط متعلق به خط سیرهای مختلف که به یک‌راه تعلق دارند، یک خوشه را تشکیل می‌دهند. سپس میانگین موقعیت‌های نقاط تشکیل‌دهنده یک خوشه، به‌عنوان مقادیر اولیه الگوریتم کا_مینز[5] در نظر گرفته می‌شوند. در ادامه از الگوریتم کا_مینز به‌منظور کشف موقعیت مجموعه نقاط به‌گونه‌ای که خط سیرها را به بهترین شکل توصیف کنند، استفاده گردید. نهایتاً به‌منظور استخراج راه مراکز این خوشه‌ها به یکدیگر متصل شدند.

از معایب روش‌های ارائه‌شده درزمینه استخراج راه از داده‌های مردم‌گستر می‌توان به ایستا[6] بودن این روش‌ها اشاره نمود. داده‌های مردم‌گستر به‌صورت پیوسته و پویا جمع‌آوری می‌شوند. این در حالی است که روش‌های ارائه‌شده، تغییرات پیوسته شبکه راه‌ها که توسط داده‌های مردم‌گستر گزارش می‌شوند را در نظر نمی‌گیرند. به‌عبارتی‌دیگر الگوریتم‌های پیشنهادشده قابلیت درک تغییرات محیط (که این تغییرات می‌تواند راهی که اخیراً ساخته شده و

[1] Social networks
[2] Sensor
[3] Trajectory data
[4] Filter
[5] k-means
[6] Static

یا مسدود شده و مسیری جایگزین برای آن در نظر گرفته شده باشد) را ندارند، بلکه در بازه‌های زمانی بخصوصی، به‌منظور به‌هنگام‌رسانی نقشه‌ها، باید دوباره اجرا شوند. لذا عدم توجه به ماهیت پویایی داده‌های خط‌سیر باعث می‌شود تا روش‌های ارائه شده در نگاشت تغییرات محیط واقعی با مشکل مواجه گردند. از طرفی به‌هنگام رسانی پیوسته نقشه‌های راه با دو مشکل حجم بالای محاسبات و نیز ناسازگاری داده‌های خط سیر روبرو است. نحوه مرتفع نمودن این دو مشکل توسط راه‌کار ارائه شده در بخش‌های بعدی این مقاله تشریح خواهد شد.

در این مقاله با الهام از نحوه پایش پیوسته محیط وب توسط عامل‌های کاوشگر وب[1]، با توجه به تغییرات پیوسته داده‌های جمع‌آوری‌شده خط سیر توسط کاربران، روشی پویا به‌منظور به‌هنگام رسانی نقشه‌های شبکه راه، ارائه‌شده است. بدین منظور عامل‌هایی توسعه داده شد که با حرکت بر روی‌داده‌های خط سیر نسبت به تغییرات محیط واکنش نشان داده و به‌بهنگام رسانی نقشه شبکه راه‌ها بپردازند. این عامل‌ها با در نظر گرفتن فضای سه‌بعدی داده‌های خط سیر راه‌حلی کارآمد به‌منظور استخراج راه، پی خواهند گرفت. استفاده از عامل‌های کاوشگر باعث می‌شود تا عمل بهنگام رسانی، با در نظر گرفتن تغییرات محیط که توسط خط سیرهای جمع‌آوری‌شده کاربران گزارش می‌شوند، به‌خوبی صورت گیرد. این عامل‌ها با در نظر گرفتن فضای سه‌بعدی داده‌های مردم‌گستر عملکرد مناسبی در استخراج راه در جداشدگی و یکی شدگی مسیر و تقاطع غیر هم‌سطح دارند.

ساختار مقاله در ادامه به این شرح است: در بخش دوم راه‌کار ارائه‌شده در این مقاله به‌منظور استخراج راه توسط عامل‌های کاوشگر به‌صورت کامل تشریح شده است. بخش سوم با در نظر گرفتن ساختارهای پیچیده شهری نحوه‌ی پاسخگویی راه‌کار ارائه‌شده به مسائل موجود در استخراج راه از ساختارهای پیچیده شهری را توضیح می‌دهد. بخش چهارم به نحوه توسعه راه‌کار ارائه‌شده و تحلیل نتایج می‌پردازد. نتایج و پیشنهادات نیز بخش پایانی این مقاله را شکل داده است.

۲- استخراج راه توسط عامل‌های کاوشگر

همان‌گونه که گفته شد، روش‌های ارائه‌شده درزمینه استخراج راه از داده‌های خط سیر که در این مطالعه مورد بررسی قرار گرفت عموماً ایستا بوده و به تغییرات پیوسته شبکه راه‌ها که توسط داده‌های خط سیر کاربران گزارش می‌شوند، توجه نمی‌کنند. این روش‌ها معمولاً یک‌بار روی‌داده‌های خط اجراشده و شبکه‌ی راه را استخراج می‌نمایند. ایستا بودن روش‌های استخراج راه از داده‌های خط سیر منجر به عدم به‌هنگام‌رسانی نقشه‌های شبکه راه متناسب با تغییرات دنیای واقعی می‌گردد. استفاده از عامل‌ها به‌عنوان مولفه‌های نرم‌افزاری باقابلیت عکس‌العمل نسبت به تغییرات محیط، به‌عنوان ابزار بهنگام رسانی نقشه‌های راه‌ها، می‌تواند سازوکارهای ایستایی ارائه‌شده را به ساز و کارهای پویا تبدیل نماید.

راسل [۲۰] عامل‌ها را کدهای کامپیوتری معرفی می‌کند که از قابلیت درک محیط و عکس‌العمل در برابر تغییرات برخوردار می‌باشند. نوع ویژه‌ای از عامل‌های هوشمند که به‌منظور شناسایی و نمایه‌گذاری[2] صفحات مفید وب توسط موتورهای جست‌وجو[3] مورد استفاده قرار می‌گیرند، عامل‌های کاوشگرهای وب نامیده می‌شوند [۲۱]. عامل‌های کاوشگر وب به‌طور پیوسته و خودکار بر روی صفحات وب (که با لینک به یکدیگر متصل شده‌اند) حرکت نموده و به کاوش محتوی

[1] Web Crawler-Agent

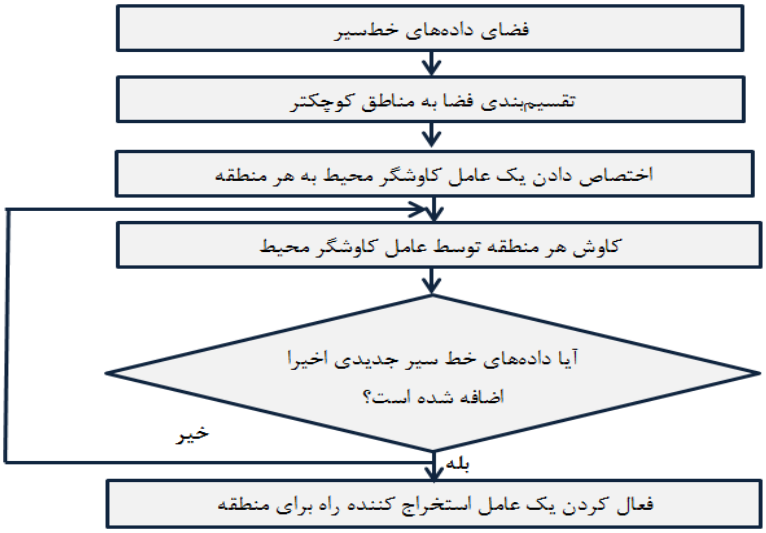
[2] Indexing
[3] Search engine

صفحات وب می‌پردازند.

در این پژوهش با الهام از نحوه پویش صفحات وب توسط عامل‌های کاوشگر وب، طراحی و توسعه عامل‌هایی در دستور کار قرار گرفت که می‌توانند با حرکت در محیط داده‌های مکانی خط سیر نسبت به تغییرات محیط واکنش نشان داده و به‌بهنگام رسانی نقشه شبکه راه‌ها بپردازند. بدین‌منظور سیستم چند عامله‌ای شامل سه دسته عامل طراحی‌شده است:

- عامل کاوشگر محیط باهدف کاوش فضای سه‌بعدی داده‌های مکانی خط سیر
- عامل استخراج‌کننده راه باهدف استخراج راه از طریق حرکت بر روی‌داده‌های خط سیر
- عامل مدیر باهدف مدیریت رفتار عامل‌های کاوشگر محیط و عامل استخراج‌کننده راه

در این پژوهش محیط عمل عامل‌ها، فضای داده‌های مردم‌گستر خط سیر است. شکل (۱) روال کلی استخراج خودکار هندسه راه با استفاده از عامل‌های کاوشگر را نمایش می‌دهد. فضای داده‌های مردم‌گستر توسط عامل مدیر به مناطقی تقسیم‌شده، سپس به هر منطقه یک عامل کاوشگر محیط اختصاص داده خواهد شد. این عامل کاوشگر محیط به‌طور پویا در منطقه به جستجو می‌پردازد و در صورت پیدا نمودن داده‌هایی که بیانگر تغییرات شبکه‌ی راه باشند، گزارشی از موقعیت آن منطقه در اختیار عامل مدیر قرار می‌دهد. عامل مدیر برای آن منطقه مربوطه، در صورت نیاز یک عامل استخراج‌کننده راه تخصیص خواهد داد. هر مسیر تنها توسط یک عامل استخراج‌کننده استخراج خواهد شد. این عامل بر روی‌داده‌های خط سیر حرکت کرده و به استخراج راه می‌پردازد. در ادامه وظیفه هر عامل و نحوه عملکرد هرکدام به‌طور کامل شرح داده خواهد شد.



شکل ۱: روال کلی استخراج خودکار هندسه راه بر مبنای عامل‌های کاوشگر

۲-۱- عامل مدیر

عامل مدیر وظیفه مدیریت عامل های کاوشگر محیط و عامل‌های استخراج‌کننده راه را بر عهده دارد. این عامل به منظور جلوگیری از اجرای الگوریتم استخراج راه برای کل فضای داده‌های خط سیر، فضای داده‌های خط سیر تحت پوشش خود را به مناطق کوچک‌تری تقسیم می‌کند. این کار به‌منظور تعیین محدوده فعالیت عامل‌های کاوشگر محیط در سیستم چندعامله صورت می‌گیرد. برای این‌منظور عامل مدیر سپس مختصات ابعاد هر منطقه را تعیین می‌کند. برای داده‌های خط‌سیر ذخیره شده، زمان ثبت داده‌ها در پایگاه داده وارد شده است. با استفاده از زمان ثبت، امکان تعیین داده‌های خط‌سیر جدید فراهم می‌شود.

به‌منظور تعیین داده‌های خط سیر جدید، عامل مدیر به هر یک از مناطق یک عامل کاوشگر محیط اختصاص می‌دهد. عامل کاوشگر محیط منطقه مربوط به خود را تا انتها کاوش نموده و به‌محض مشاهده تعداد قابل قبول از داده های خط سیری که بیانگر تغییرات آن منطقه باشند، موقعیت داده‌های خط‌سیر جدید را به عامل مدیر گزارش می‌دهد. سپس عامل مدیر در موقعیت کشف شده یک عامل استخراج‌کننده راه فعال خواهد نمود. در مناطقی که در آن‌ها تغییراتی گزارش نشده، عامل استخراج‌کننده‌ای فعال نخواهد شد. ابعاد این مناطق در سیستم چند عامله توسعه داده شده قابل تنظیم است.

به‌منظور تعیین تعداد قابل قبول داده های خط سیر جدید، عامل مدیر پالونه‌ای که شامل دو پارامتر زمانی و تعداد (تعداد قابل قبولی از داده‌های خط سیر) است را برای داده‌های خط سیر جمع‌آوری شده توسط کاربران بررسی می‌کند. در صورت اضافه شدن تعداد قابل قبولی از داده‌های خط سیر جدید، این داده‌ها به محیط عمل سیستم چند عامله اضافه خواهند شد. در پیاده‌سازی انجام شده مقادیر اولیه داده‌های خط سیر دو هفته اخیر به‌عنوان پارامتر زمانی و تعداد ۲۰ داده خط سیر به‌عنوان پارامتر تعداد در نظر گرفته شد. این پارامترها در سیستم چند عامله توسعه داده شده قابل تنظیم می‌باشند و در پیاده سازی عملیاتی می‌توان بر اساس نیاز آن‌ها را، به‌منظور رسیدن به نتایج بهتر، تغییر داد.

ازآنجایی‌که حضور و فعالیت عامل‌های استخراج‌کننده راه مختلف در محیط به منابع پردازشی بالایی نیاز دارند، عامل مدیر سعی در برقراری تعادل میان منابع پردازشی موجود و میزان پردازش هر منطقه دارد. در صورتیکه عامل مدیر تعداد عامل‌های استخراج‌کننده راه را به‌صورت نامحدود افزایش دهد، سیستم عملکرد مناسبی نخواهد داشت و این عمل می‌تواند باعث کندی عملکرد سیستم شود. بررسی‌های تجربی انجام شده در این پژوهش بیان‌گر این مهم بود که سخت‌افزاری که سیستم چندعامله روی آن پیاده‌سازی می‌شود تا زمانی که میزان استفاده از CPU کمتر از ۷۰ درصد است به‌صورت بهینه عمل کرده و پردازش‌ها را با سرعتی مناسب اجرا می‌کند. بر این اساس در پیاده‌سازی عامل مدیر کتابخانه SIGAR[1] به منظور پایش میزان CPU استفاده شده به کار گرفته شد. عامل مدیر با استفاده از این کتابخانه در هرلحظه مقادیر CPU را بررسی نموده و تا زمانی که مقدار CPU کمتر از ۷۰ درصد باشد، در صورت لزوم عامل‌های استخراج کننده راه جدیدی اضافه می‌کند. همچنین درصورتی‌که مقدار CPU بیشتر از ۷۰ درصد باشد، عامل‌هایی را غیرفعال خواهد نمود. بدین‌ترتیب عامل مدیر سعی در ایجاد توازن لازم در کل سیستم چند عامله را دارد.

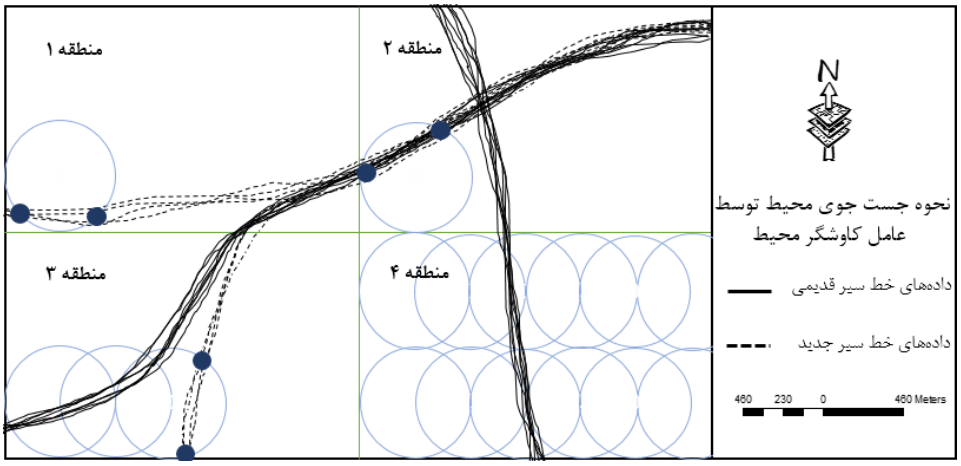
عامل‌های استخراج کننده راه موقعیت خود در محیط را به عامل مدیر گزارش می‌دهد. از آنجایی‌که هر مسیر تنها توسط یک عامل استخراج‌کننده استخراج می‌شود لذا از بین عامل‌هایی که به‌طور مشترک کار استخراج راه برای یک مسیر را انجام می‌دهند، یکی توسط عامل مدیر غیرفعال خواهد شد. همچنین در صورت کشف داده‌های خط سیر جدید توسط عامل کاوشگر محیط، یک عامل استخراج‌کننده جدید توسط عامل مدیر فعال خواهد شد.

۲-۲- عامل‌های کاوشگر محیط

پس از تقسیم فضای داده‌های مردم‌گستر به مناطق و تخصیص یک عامل کاوشگر محیط به هر منطقه، عامل کاوشگر محیط در فضای سه‌بعدی داده‌های مردم‌گستر حرکت کرده و به‌محض مشاهده تعداد کافی از داده‌های خط سیری که بیانگر تغییرات آن منطقه باشند، نتیجه را به عامل مدیر گزارش می‌دهد. به عبارتی برای خط سیرهای جدید جمع‌آوری‌شده از یک مسیر که فراوانی کمی دارند و همچنین خط سیرهای قدیمی کار استخراج راه انجام نخواهد گرفت. شکل (۲) نحوه جستجوی عامل

[1] https://support.hyperic.com/display/SIGAR/Home

کاوشگر محیط در هر منطقه را به‌صورت شماتیک نمایش می‌دهد. خطوط کامل، داده‌های خط سیر قدیمی که پیش‌ازاین به‌منظور ایجاد نقشه مورداستفاده قرارگرفته است را نمایش می‌دهد و داده‌های خط سیر جدید که بیانگر تغییرات محیط هستند با خط‌چین نمایش داده‌شده است. عامل‌های کاوشگر محیط به‌منظور پیدا کردن داده‌های خط سیر موجود در محدوده جستجو، منطقه مربوطه به خود را تا انتها کاوش می‌کنند. در برخی از این مناطق داده‌های خط سیر جدیدی وجود ندارد (به‌عنوان‌مثال منطقه ۴)، لذا برای این مناطق هیچ عامل استخراج‌کننده راهی فعال نخواهد شد. در برخی دیگر از این مناطق ممکن است عامل کاوشگر داده‌های خط سیر جدید را کشف کند (به‌عنوان‌مثال مناطق ۱، ۲ و ۳)، برای این مناطق یک عامل استخراج‌کننده راه فعال خواهد شد. به منظور فعال نمودن یک عامل استخراج کننده، عامل مدیر موقعیت‌های ابعاد منطقه مجاز به منظور فعالیت عامل استخراج کننده و نیز موقعیت داده‌های خط سیر جدید کشف شده را به عامل استخراج کننده معرفی می‌کند.



شکل ۲: جستجوی منطقه توسط عامل‌های کاوشگر محیط

۲-۳- عامل‌های استخراج‌کننده راه

با پیدا شدن خط سیرهای جدید توسط عامل‌های کاوشگر محیط، محدوده این خط سیرها به عامل مدیر گزارش داده می‌شوند. عامل مدیر یک عامل کاوشگر استخراج‌کننده راه در موقعیت گزارش داده‌شده فعال می‌کند. عامل استخراج‌کننده راه بر روی‌داده‌های خط سیر حرکت نموده و به استخراج راه می‌پردازد.

شکل (۳) راه‌حل استخراج راه توسط عامل استخراج‌کننده را به‌صورت شماتیک نمایش می‌دهد. عامل استخراج‌کننده راه جهت‌های مختلفی برای حرکت پیش رو دارد. به‌منظور استخراج راه، این عامل باید در امتداد داده‌های خط سیر حرکت کند. بدین‌منظور این عامل در موقعیت خود صفحه‌ای ایجاد نموده سپس بهترین حالت صفحه، به‌طوریکه عامل استخراج‌کننده با حرکت در راستای گرادیان صفحه می‌تواند بر روی‌داده‌های خط سیر حرکت کند، تعیین می‌شود. بدین منظور، این صفحه حول محور Z سیستم مختصات محلی تعریف‌شده برای آن، دوران داده می‌شود. پس از تعیین بهترین حالت صفحه، به‌منظور تعیین گره‌های[1] شبکه راه، بر روی نقاط حاصل از تقاطع صفحه با خط سیرها خوشه‌بندی صورت می‌گیرد. سپس عامل استخراج‌کننده در امتداد گرادیان صفحه به جلو حرکت نموده و صفحه جدیدی به‌منظور تعیین گره‌های گراف شبکه راه ایجاد می‌کند. پس از تشکیل صفحه جدید و شناسایی گره‌های جدید،

[1] Node

به‌منظور تشکیل اتصال[1] شبکه راه، گره‌های جدید به گره‌های قبلی متصل می‌شوند. درنهایت به‌منظور استخراج شبکه راه، عامل استخراج‌کننده به جلو حرکت نموده و با تکرار مراحل تشکیل صفحه، تعیین گره‌های جدید و اتصال آن‌ها به گره‌های قبلی، به تشکیل اتصالات بیشتر می‌پردازد. شبه کد (۱) نحوه استخراج راه توسط عامل استخراج‌کننده را تشریح می‌کند.

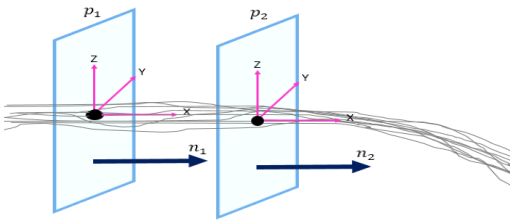
Input: Trajectories of floationg cars

Output: Road network

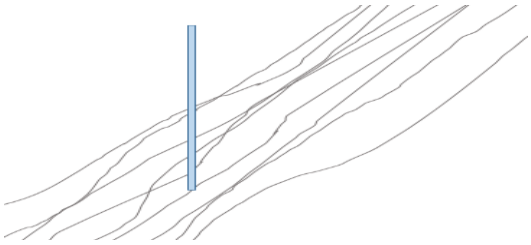
1: number of square = 1;
2: P (A_o (x_o ,y_o ,z_o , N_o)) ← Create square
3: Points ← find all trajectories which has intersected with square
4: Center ← Calculate mean of coordinates (points)
5: Square ← Move square (Center)
6: Road_Extraction_Agent report its location to Manager Agent
7: Points ← find all trajectories which has intersected with square
8: **for** (i=1:9) **do**
9: Rotate Square
10: Baarda test
11: Calculate variance of coordinates
12: **end for**
13: Choose the best square's state
14: Clustering
15: Clusters validation test
16: **If** (number of squares > ۱) **then**
17: Create weight matrix
18: Link relevant clusters
19: **end if**
20: go forward along the squares gradient for 100 meter
21: number of squares ++;
22: go to step 2

شبه کد ۱: نحوه استخراج راه توسط عامل استخراج‌کننده راه

در ادامه مراحل استخراج راه توسط عامل استخراج‌کننده به‌صورت کامل تشریح می‌شود.



شکل ۳: نحوه ایجاد صفحه و حرکت بر روی‌داده‌های خط سیر توسط عامل استخراج‌کننده راه

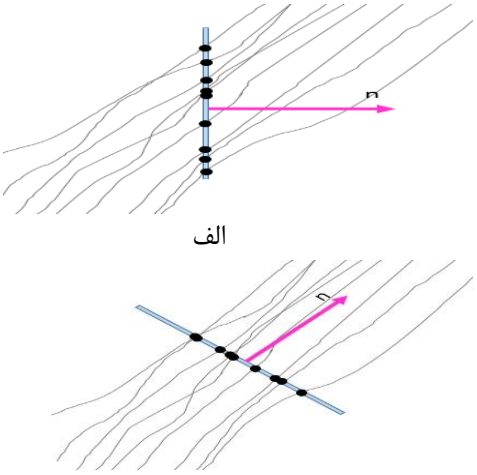


شکل ۴: صفحه ایجادشده توسط عامل استخراج‌کننده در موقعیت خود (پیش از منتقل‌شده به میانگین نقاط حاصل از تقاطع آن با خط سیرها)

۲-۳-۱- انتخاب جهت حرکت مسیر

به‌منظور مشخص نمودن حرکت عامل استخراج‌کننده بر روی‌داده‌های خط سیر، این عامل در موقعیت خود یک صفحه ایجاد نموده (خط ۲ شبه کد (۱) که در آن x_0، y_0 و z_0 مختصات مرکز صفحه و نیز N_0 بردار نرمال صفحه می‌باشند) و تقاطع این صفحه با خط سیرها تعیین می‌شود (خطوط ۳ شبه کد (۱)). شکل (۴) صفحه ایجادشده توسط عامل استخراج‌کننده را نمایش می‌دهد. این صفحه ممکن است برخی از خط سیرهایی که می‌توانند در استخراج راه شرکت داشته باشند را قطع نکند. به‌منظور حل این مشکل صفحه به میانگین مختصات نقاط حاصل از تقاطع این صفحه با خط سیرها منتقل می‌شود (خطوط ۴ و ۵ شبه کد (۱)). عامل استخراج کننده راه موقعیت خود در محیط را به عامل مدیر گزارش می‌دهد (خط ۶ شبه کد (۱)). سپس تقاطع این صفحه با خط سیرها دوباره محاسبه خواهد شد (خط ۷ شبه کد (۱))

[1] Edge



شکل ۵: الف- صفحه منتقل‌شده به میانگین تقاطع آن با خط سیرها و ب- بهترین حالت صفحه به‌طوریکه بردار نرمال آن هم‌جهت با خط سیرها را نمایش می‌دهد.

شکل (۵-الف) صفحه تشکیل‌شده توسط عامل استخراج‌کننده، پس از انتقال به میانگین مختصات نقاط، را نمایش می‌دهد. این صفحه هنوز صفحه ایده‌آل به‌منظور حرکت عامل استخراج‌کننده بر روی‌داده‌های خط سیر نیست. صفحه‌ای که مختصات نقاط حاصل از تقاطع آن با خط سیرها دارای انحراف معیار کمتری هستند، بهترین حالت صفحه (شکل (۵-ب))، به‌منظور حرکت عامل استخراج‌کننده بر روی‌داده‌های خط سیر است. در این حالت بردار نرمال صفحه بر روی‌داده‌های خط سیر منطبق شده و جهت حرکت عامل استخراج‌کننده را به‌خوبی مشخص می‌کند. به‌منظور تعیین بهترین حالت صفحه، این صفحه حول محور Z با زاویه ۲۰ درجه دوران داده می‌شود (خط ۹ شبه کد (۱)). سپس انحراف معیار مختصات نقاط حاصل از تقاطع صفحه با خط سیرها برای هر حالت صفحه محاسبه‌شده و حالتی از صفحه که انحراف معیار کمتری دارد به‌عنوان بهترین حالت صفحه انتخاب می‌شود.

پیش از محاسبه انحراف معیار ، ازآنجایی‌که به‌دلیل نویز ذاتی داده‌های خط سیر ممکن است یک نقطه تغییر زیادی در مقدار انحراف معیار مختصات نقاط حاصل از تقاطع صفحه با خط سیرها داشته باشد، لذا برای جلوگیری از این مشکل و به‌منظور اطمینان از عدم حضور مشاهدات اشتباه از آزمون باردا[1] [۲۲] استفاده می‌شود (خط ۱۰ شبه کد (۱)) و مشاهدات اشتباه حذف می‌شوند. عموما پیش از آزمون باردا، آزمون فاکتور واریانس ثانویه به‌منظور بررسی توزیع (تصادفی و یا نرمال بودن) داده‌ها صورت می‌گیرد. رد شدن آزمون فاکتور واریانس ثانویه به این مفهوم است که داده‌ها از توزیع نرمال پیروی نموده و لذا می‌توان از این داده‌ها در آزمون باردا استفاده نمود. اما در این مطالعه به پیروی از مطالعات پیشین در زمینه استخراج راه از داده‌های خط سیر (مراجعه شود به [۱۱, ۱۵, ۲۳, ۲۴, ۲۵]) فرض شد که موقعیت داده‌های خط‌سیر در سطح راه‌ها از توزیع نرمال پیروی می‌کنند و لذا آزمون باردا بدون در اعمال آزمون فاکتور وریانس روی داده‌ها اعمال شد.

پس از حذف نقاط اشتباه، انحراف معیار تک‌تک مؤلفه‌های مختصات هر یک از نقاط باقیمانده از رابطه (۱) و درنتیجه انحراف معیار مختصات نقاط باقیمانده از رابطه (۲) محاسبه می‌شود (خط ۱۱ شبه کد (۱)).

که در آن $g_i - \bar{g}$ مقدار اختلاف مؤلفه هر نقطه نسبت به مقدار میانگین است. به‌این‌ترتیب حالتی از صفحه که مختصات نقاط حاصل از تقاطع آن با خط سیرها انحراف معیار کمتری داشته باشند، به‌عنوان بهترین حالت صفحه انتخاب می‌شود.

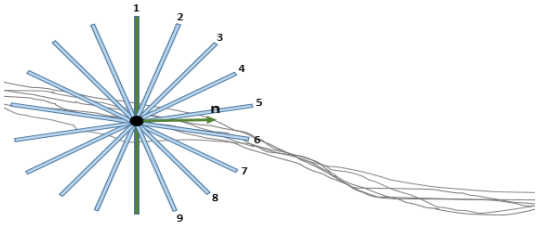
جدول (۱) نمونه‌ای از انحراف معیار مختصات نقاط حاصل از تقاطع حالت‌های مختلف صفحه با داده‌های خط سیر پس از اعمال آزمون باردا و شکل (۶) حالت‌های مختلف هر صفحه پس از دوران تحت زوایای مختلف را نمایش می‌دهد. طبق جدول (۱) و شکل (۶) تعداد نقاط حاصل از تقاطع صفحه با خط سیرها برای حالت‌های ۱- ۵ و ۷-۹ بیشترین و برای حالت ۶

[1] Baarda test

کمترین تعداد است. لذا انتظار می‌رود که انحراف معیار مختصات نقاط برای حالت ۶ صفحه نسبت به سایر حالت‌ها عدد کمتری باشد. درنتیجه این حالت صفحه "به اشتباه" به‌عنوان بهترین حالت صفحه انتخاب شود. این در حالی است که با توجه به شکل (۶) ازآنجایی‌که عامل استخراج‌کننده راه با حرکت در راستای گرادیان صفحه نمی‌تواند بر روی داده‌های خط سیر حرکت کند، لذا این حالت نمی‌تواند به‌عنوان بهترین حالت صفحه در نظر گرفته شود. به‌منظور جلوگیری از این مشکل، بهترین حالت صفحه تنها از بین حالت‌هایی که در آن بیشترین تعداد نقاط حاصل از تقاطع صفحه با خط سیرها را دارا باشد، انتخاب خواهد شد. که در این صورت حالت ۱ صفحه که انحراف معیار کمتری دارد "به‌طور صحیح" به‌عنوان بهترین حالت صفحه انتخاب خواهد شد. پس از تعیین بهترین حالت صفحه، گره‌های شبکه راه تعیین خواهد شد. بخش بعدی این مقاله به نحوه تعیین گره‌های شبکه راه می‌پردازد.

جدول ۱: انحراف معیار مختصات نقاط حاصل از تقاطع حالت‌های مختلف صفحه با داده‌های خط سیر

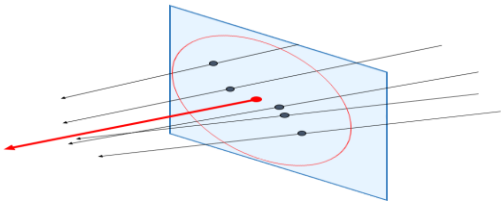
شماره حالت صفحه	زاویه دوران (درجه)	تعداد نقاط	انحراف معیار نقاط
۱	۰	۶	۶۶۱۲٫۵۱۳۶
۲	۲۰	۶	۶۶۷۳٫۸۵۲۹
۳	۴۰	۶	۶۷۵۴٫۴۳۷۳
۴	۶۰	۶	۶۸۲۷٫۶۲۷۵
۵	۸۰	۶	۷۰۳۱٫۳۶۵۸
۶	۱۰۰	۴	۵۷۵۶٫۴۸۲۵
۷	۱۲۰	۶	۷۲۶۴٫۹۵۳۰
۸	۱۴۰	۶	۶۷۱۸٫۵۶۳۳
۹	۱۶۰	۶	۶۶۴۸٫۲۴۴۷



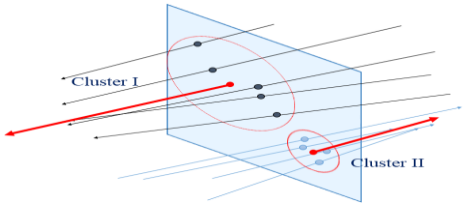
شکل ۶: حالت‌های مختلف هر صفحه و همچنین بهترین حالت صفحه

۲-۳-۲- تعیین گره‌های شبکه راه

فرض اولیه روش‌های استخراج راه از داده‌های مردم گستر خط سیر بر این است که با توجه به اینکه ماشین‌ها بر روی سطح خیابان‌ها حرکت می‌کنند، انتظار می‌رود داده‌های جمع‌آوری‌شده نماینده‌ای از ساختار هندسه راه‌ها باشند. با این فرض می‌توان گفت که داده‌های خط سیر جمع‌آوری‌شده توزیع خوشه‌ای دارند، لذا می‌توان از روش‌های خوشه‌بندی به‌منظور تعیین مراکز راه استفاده نمود [۲۶]. پس از تعیین صفحه بهینه، به‌منظور تعیین گره‌های شبکه راه، بر روی نقاط حاصل از تقاطع آن با خط سیرها، خوشه‌بندی انجام خواهد شد (خط ۱۴ شبه کد ۱).



الف



ب

شکل ۷: الف: تقاطع صفحه با خط سیرها در یک مسیر یک‌طرفه- خطوط سیر با رنگ سیاه و بردار گرادیان صفحه با رنگ قرمز مشخص‌شده است؛ ب: تقاطع صفحه با خط سیرها در یک مسیر دوطرفه، جداشدگی و یا یکی شدگی مسیر

شکل (۷- الف) تقاطع صفحه با داده‌های خط سیرها جمع‌آوری‌شده از سطح یک مسیر یک‌طرفه را نمایش می‌دهد. همان‌طور که از شکل پیداست، در مسیرهای یک‌طرفه تنهای یک خوشه وجود خواهد داشت که مرکز آن به‌عنوان گره شبکه راه تعیین خواهد شد. در مسیرهای دوطرفه، بزرگراه‌ها، جداشدگی دو مسیر، تقاطع‌های هم‌سطح و غیر هم‌سطح، این صفحه ممکن است خط سیرهای مربوط به راه دیگری را نیز در برگیرد. در این حالت بیش از یک مرکز خوشه وجود خواهد داشت. شکل (۷-ب) تقاطع صفحه با داده‌های خط سیر یک مسیر دوطرفه را نمایش می‌دهد. می‌توان نتیجه گرفت که خوشه‌ها در مسیرهای دوطرفه، جداشدگی و جمع شدگی مسیر، گره‌هایی از مسیرهای متفاوت هستند. بنابراین هدف از خوشه‌بندی تعیین تعداد و موقعیت مراکز خوشه یا خوشه‌ها خواهد بود.

ازجمله روش‌های خوشه‌بندی می‌توان به روش‌های کا_مینز ,FCM[2] ,EM[1] و ISODATA اشاره نمود [۲۷]. از میان روش‌های مطرح در زمینه‌ی خوشه‌بندی در این پژوهش از روش EM به‌منظور تعیین گره‌های مسیر استفاده شد.

الگوریتم EM یک روش تکراری و بر مبنای توزیع احتمالی به‌منظور تعیین بیشترین شباهت[3] بین پارامترهای یک مدل آماری است [۲۸]. این الگوریتم فرض می‌کند هر خوشه دارای یک توزیع به همراه یکسری پارامترهای مشخص است. این الگوریتم در یک فرایند تکراری مقدار احتمالی که هر نقطه به یک خوشه بر اساس پارامترهای تعیین‌شده، متعلق است را تعیین می‌کند و نمونه‌ها به خوشه‌هایی با بیشترین احتمال منتسب می‌شوند. در بین روش‌های خوشه‌بندی موجود، EM یکی از روش‌هایی است که نیاز به معرفی مکان اولیه خوشه‌ها ندارد، لذا این روش به‌منظور تعیین تعداد و موقعیت خوشه‌ها مورد استفاده قرار گرفت.

این الگوریتم شامل دو مرحله تکراری است. در مرحله اول مقدار مورد انتظار تابع لگاریتم درستنمایی[4] با در نظر گرفتن شرایط توزیع نمونه‌های مشاهده‌شده (α) و مشاهده نشده (β) تحت پارامترهای شباهت داده‌شده (θ) مطابق با رابطه (۳) تعیین می‌گردد. سپس در مرحله دوم پارامترهایی که شرط رابطه (۴) را برقرار کنند، مشخص خواهند شد[۲۸]. این پارامترها در مرحله اول مورد استفاده قرارگرفته و تا هنگام رسیدن به مقدار مورد انتظار تابع لگاریتم درستنمایی تکرار خواهند شد.

رابطه (۳)
$$Q(\theta \mid \theta^{(t)}) = E_{\beta|\alpha,\theta^{(t)}} [\log L(\theta; \alpha, \beta)]$$

رابطه (۴)
$$\theta^{(t+1)} = \arg\max Q(\theta \mid \theta^{(t)})$$

که در آن α تعداد نمونه‌های مشاهده‌شده و β تعداد نمونه‌های مشاهده نشده و θ برداری از پارامترهای مجهول است.

در این مقاله، در مرحله اول مجموعه نقاط حاصل از تقاطع صفحه با خط سیرها به‌عنوان داده‌های ورودی (داده‌های مشاهده نشده) به الگوریتم معرفی شدند. سپس با معلوم بودن میانگین و کواریانس نقاط مشاهده نشده (β)، برآوردی از تابع شباهت θ تعیین شد. در مرحله دوم به‌منظور بیشنه نمودن این تابع شباهت، رابطه (۴) استفاده شد. سپس بر روی‌داده‌های مشاهده نشده خوشه‌بندی صورت گرفت. این داده‌های خوشه‌بندی شده به‌عنوان داده‌های مشاهده‌شده (α) در نظر گرفته‌شده و در تکرار دوم برآورد جدیدی از تابع شباهت θ به دست آمد. مراحل تا هنگام رسیدن به مقدار مورد انتظار تابع لگاریتم درستنمایی تکرار شد. درنهایت با تکرار مراحل، تعداد و

[1] Expectation Maximization
[2] Fuzzy C-means
[3] Maximum Likelihood

[4] log-likelihood : تابعی است که توزیع طبیعی داده‌ها را نشان می‌دهد.

موقعیت خوشه‌ها به دست آمد.

پس از تعیین خوشه‌ها، به‌منظور جلوگیری از ایجاد خوشه‌های کاذب، آزمون اعتبارسنجی خوشه‌بندی صورت می‌گیرد (خط ۱۵ شبه کد ۱). این آزمون بیان می‌کند که تنها خوشه‌هایی که خط سیرهای تشکیل‌دهنده آن‌ها دارای دو شرط شاخص زمانی و جمعیت باشند، معتبر خواهند بود. شاخص زمان و جمعیت بیان می‌کند که تنها خوشه‌هایی که از زمان t تاکنون جمع‌آوری‌شده‌اند و همچنین جمعیت خط سیرهای تشکیل‌دهنده (گذرنده از) آن‌ها بیشتر از حد آستانه n باشند، خوشه‌های معتبری خواهند بود. به عبارتی خط سیرهای جدید جمع‌آوری‌شده از یک مسیر که فراوانی کمی دارند و همچنین خط سیرهای قدیمی فاقد اعتبارند و به‌منظور استخراج راه استفاده نخواهند شد. این خوشه‌های معتبر به‌عنوان گره‌های مسیر در نظر گرفته می‌شوند. پس از تعیین گره‌های مسیر در موقعیت عامل استخراج‌کننده، این عامل به حرکت در راستای داده‌های خط سیر به‌منظور تعیین گره‌های بیشتر مسیر می‌پردازد.

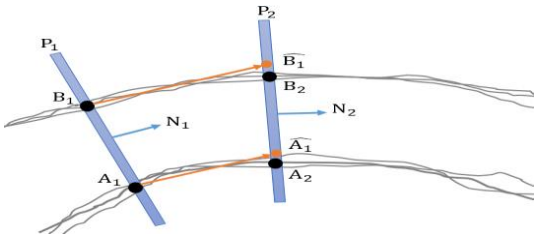
۲-۳-۳- انتخاب گره‌های مسیر در موقعیت جدید

به‌منظور تعیین گره‌های بیشتری از مسیر، تعداد صفحات تشکیل‌شده در خط ۱۶ شبه کد بررسی خواهد شد. اگر تنها یک صفحه تشکیل‌شده باشد، عامل استخراج‌کننده راه در راستای گرادیان صفحه به اندازه ۱۰۰ متر به جلو حرکت کرده (خط ۲۰ شبه کد ۱) و مراحل انتخاب جهت حرکت مسیر (خطوط ۳ الی ۱۳ شبه کد۱)، تعیین گره‌های شبکه راه (خطوط ۱۴ و ۱۵ شبه کد ۱) و اتصال گره‌های جدید به گره‌های قبلی (خطوط ۱۶ الی ۱۹ شبه کد ۱) تکرار خواهد شد.

درصورتی‌که بیش از یک صفحه توسط عامل استخراج‌کننده تشکیل‌شده باشد، ماتریس وزن به‌منظور اتصال گره‌ها به گره‌های قبلی، تشکیل خواهد شد. این موضوع در ادامه به تفضیل توضیح داده خواهد شد.

۲-۳-۴- اتصال گره‌های جدید به گره‌های قبلی

بعد از تعیین گره‌ها، به‌منظور تشکیل گراف راه، گره‌های تعیین‌شده به گره‌های استخراج‌شده در مرحله قبلی متصل خواهند شد. برای این منظور ماتریس وزن با استفاده از سه ویژگی فاصله، اختلاف جهت و جمعیت نمونه‌ها برای هر دو گره متوالی (به‌منظور تشکیل یک اتصال[1]) تشکیل می‌شود (خط ۱۶ شبه کد ۱). شکل (۸) دو صفحه متوالی را نمایش می‌دهد. با فرض اینکه صفحه P_1 دارای نرمال صفحه (N_1) و دو مرکز گره A_1 و B_1 است، هدف تعیین مناسب‌ترین گره به‌منظور اتصال به گره A_1 است. از بین گره‌های موجود در صفحه P_2 (که گره‌های کاندید برای متصل شدن به گره A_1 از صفحه P_1 هستند) مناسب‌ترین گره به‌منظور اتصال به گره A_1، با تشکیل ماتریس وزن، انتخاب خواهد شد.



شکل ۸: اتصال گره‌های متوالی به یکدیگر

در راه‌کار پیشنهادی این مقاله برای اتصال گره‌ها به گره‌های پیشین، مهم‌ترین ویژگی در تشکیل ماتریس وزن، فاصله است. با حرکت از A_1 در امتداد بردار نرمال صفحه P_1 و در فاصله‌ای که دو صفحه از یکدیگر واقع‌شده‌اند، نقطه $\widehat{A_1}$ مشخص خواهد شد که برآوردی از گره A_2 خواهد بود. فاصله این نقطه با گره‌های A_2 و B_2 محاسبه می‌شود. این فواصل به‌عنوان یک شاخص به‌منظور تشکیل ماتریس وزن ذخیره خواهند شد.

جهت اتصالات می‌تواند به‌عنوان یکی دیگر از شاخص‌های تأثیرگذار در تشکیل ماتریس در نظر گرفته شود. بدین منظور گره A_1 را یک‌بار به

[1] Edge

A_2 و بار دیگر به B_2 متصل می‌شوند. اتصال مناسب اتصالی با کمترین اختلاف جهت[1] این اتصال با نرمال صفحه P_1 خواهد بود. این ویژگی به‌عنوان دومین ویژگی تشکیل‌دهنده ماتریس وزن در نظر گرفته می‌شود.

سومین ویژگی، اختلاف جمعیت نقاط دو گره متوالی است. دو گره متوالی A_1 و A_2 در صورتی به یکدیگر متصل می‌شوند که خط سیرهایی که از گره A_1 عبور می‌کنند همان سیرها از گره A_2 نیز عبور کنند [۱۱, ۱۴]. به عبارتی برای دو گره متوالی انتظار می‌رود، تعداد نقاط تشکیل‌دهنده گره با یکدیگر اختلاف کمی داشته باشند.

با در نظر گرفتن سه ویژگی بیان‌شده، ماتریس وزن اتصال بین گره‌های دو صفحه p_k و p_{k+1} را به‌صورت رابطه (۵) می‌توان تعریف نمود:

رابطه (۵)

$$W^{(k,k+1)} = \begin{bmatrix} w_{(i=1)(j=1)} & w_{i(j+1)} & \cdots & w_{i(j+n)} \\ w_{(i+1)j} & w_{(i+1)(j+1)} & \cdots & w_{(i+1)(j+n)} \\ \vdots & \vdots & \vdots & \vdots \\ w_{(i+m-1)j} & w_{(i+m)(j+1)} & \ldots & w_{(i+m)(j+n)} \end{bmatrix}$$

که در آن، i شماره گره از صفحه p_k ، j شماره گره کاندید از صفحه p_{k+1} و درایه‌های این ماتریس مقادیر وزن اتصالات به یکدیگر است. به‌عنوان نمونه $w_{ij} = dist_{ij} \times dDirection_{ij} \times dPopulation_{ij}$ است که در آن $dist_{ij}$ فاصله برآورد گره A_i از گره A_j ، $dDirection_{ij}$ اختلاف جهت اتصال تشکیل‌شده از گره‌های A_i و A_j از نرمال صفحه p_k و $dPopulation_{ij}$ اختلاف جمعیت گره‌های A_i (از صفحه p_k) و A_j (از صفحه p_{k+1}) است. درنهایت از بین گره‌های کاندید از صفحه p_{k+1} برای یک گره از صفحه p_k اتصالی با وزن کمتر انتخاب خواهد شد.

۲-۳-۵- شرط توقف عامل استخراج‌کننده راه

همان‌طور که پیش‌ازاین بیان شد، هر مسیر توسط یک عامل استخراج‌کننده، استخراج می‌شود و در صورت همپوشانی مسیر در حال استخراج توسط دو یا چند عامل استخراج‌کننده، یکی از عامل‌ها توسط عامل مدیر غیرفعال خواهد شد. بدین منظور عامل استخراج‌کننده در هرلحظه گزارشی از موقعیت خود، که در حال حرکت بر روی‌داده‌های خط سیر به‌منظور استخراج راه برای مسیر خاصی است، به عامل مدیر ارائه می‌دهد (خط ۶ شبه کد ۱). عامل مدیر موقعیت این عامل‌های استخراج‌کننده را بررسی نموده و به‌محض همپوشانی راه استخراج‌شده توسط عامل‌های استخراج‌کننده متفاوت، یکی از این عامل‌ها را غیرفعال خواهد نمود. به‌این‌ترتیب کار استخراج راه برای این عامل استخراج‌کننده به پایان می‌رسد.

۲-۳-۶- جایگزینی یک مسیر جدید

همان‌گونه که پیش‌ازاین اشاره شد، داده‌های مردم‌گستر تغییرات محیط را به‌خوبی نمایش می‌دهند. یکی از این تغییرات می‌تواند جایگزینی یک مسیر جدید به‌جای مسیر قبلی شود. شکل (۹) یک حالت جایگزینی مسیر را به‌صورت شماتیک نمایش می‌دهد. مسیر مشخص‌شده با خط کامل، مسیر مسدود شده و مسیر مشخص‌شده با خط‌چین مسیر جایگزین را نمایش می‌دهد. ازآنجایی‌که عامل‌های کاوشگر محیط، در فضای داده‌های خط سیر حضور دارند و به‌محض اضافه شدن تعداد قابل قبولی از خط سیرها در محیط، واکنش نشان می‌دهند، لذا در صورت عدم جمع‌آوری خط سیرهای جدید برای یک مسیر، آن مسیر تحت عنوان یک مسیر مسدود شده برچسب زده می‌شود و مسیر جایگزین به نقشه اضافه خواهد شد.

[1] Direction

شکل ۹: مسیر مسدود شده و مسیر جایگزین

۲-۴- بهنگام رسانی نقشه راه

بعد از یک‌بار تولید نقشه راه، به‌منظور بهنگام رسانی نقشه تولیدشده، عامل‌های کاوشگر محیط، در فضای داده‌های خط سیر حضور دارند و به‌محض اضافه شدن تعداد قابل قبولی از خط سیرها در محیط، برای منطقه موردنظر، عامل استخراج‌کننده راه، فعال خواهد شد. این عامل پس از حرکت بر روی‌داده‌های خط سیر و تعیین گره‌های مسیر، در هر گره وجود گره‌های دیگر در اطراف خود را بررسی خواهد نمود. سپس تابع رابطه ۶ برای هر گره بررسی خواهد شد. درصورتی‌که مقدار این تابع کمتر از حد آستانه k باشد، راه جدید استخراج‌شده به نقشه راه قبلی اضافه نخواهد شد و عامل استخراج‌کننده غیرفعال خواهد شد، در غیر این صورت این راه استخراج‌شده به نقشه قبلی اضافه خواهد شد.

برای این‌کار عامل استخراج‌کننده در هر گره، وجود گره‌های دیگر در اطراف خود را بررسی می‌کند. در صورت یافتن گرهی در این فاصله، فاصله اقلیدسی بین آن‌ها، طبق رابطه ۶، بررسی می‌شود.

رابطه (۶) $$D = \sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2}$$

که در آن D فاصله اقلیدسی بین دو خوشه است. با حرکت عامل کاوشگر محیط در فضای داده‌ها و استخراج راه از داده‌های خط سیر جدید و نقشه راه تولیدشده در گام قبلی، نقشه راه در صورت لزوم تصحیح و بهنگام می‌شود. به‌این‌ترتیب حضور عامل‌ها و عامل‌های کاوشگر، پویایی تولید و بهنگام رسانی نقشه راه را تضمین می‌کند.

۳- استخراج راه در ساختارهای پیچیده شبکه

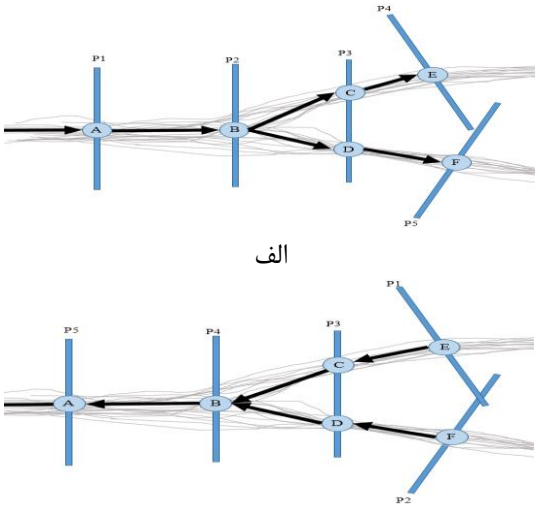
یکی از مزایای استفاده از راه‌کار عامل_مبنای ارائه‌شده برای استخراج راه در این مقاله عملکرد مناسب آن در ساختارهای پیچیده شهری است. عامل استخراج‌کننده راه به‌منظور عملکرد صحیح در ساختارهای مختلف (نظیر چهارراه‌ها، جداشدگی، یکی شدگی)، استراتژی‌های متفاوتی را پی می‌گیرند. برای این منظور در ابتدا باید نوع ساختار راه تعیین و سپس با توجه به آن رویکرد مناسبی اتخاذ گردد.

همان‌طور که پیش‌ازاین بیان شد، عامل استخراج‌کننده به‌منظور استخراج راه بر روی‌داده‌های خط سیر حرکت نموده و در هر موقعیت، به‌منظور تعیین گره‌های مسیر صفحه‌ای ایجاد کرده و سپس از نقاط حاصل از تقاطع صفحه با خط سیرها، گره‌های مسیر تعیین می‌شود. هنگامی‌که اختلاف تعداد خط سیرهای گذرنده از دو خوشه دو صفحه متوالی عدد بزرگی باشد. وجود نوع ساختار یکی شدگی و جداشدگی توسط عامل استخراج‌کننده بررسی خواهد شد. شکل‌های (۱۰-الف و ۱۰-ب) به ترتیب یک جداشدگی و یکی شدگی مسیر را نمایش می‌دهند. با در نظر رابطه (۷) اختلاف تعداد خط سیرها محاسبه می‌شود.

رابطه (۷) $$\begin{cases}\Delta n_{BC} = n_C - n_B \\ \Delta n_{BD} = n_D - n_B\end{cases}$$

که n_B، n_C و n_D به ترتیب تعداد خط سیرهای گذرنده

از خوشه‌های B، C و D است. بر اساس [۱۱, ۱۴] برای دو خوشه متوالی داده‌های خط سیرها گذرنده از خوشه اول از خوشه دوم نیز عبور می‌کنند. لذا می‌توان نتیجه گرفت که مقادیر Δn_{CB} و Δn_{DA} یا هردو مثبت و یا هردو منفی خواهند بود. درصورتی‌که مقادیر Δn_{CB} و Δn_{DA} منفی باشند، جداشدگی مسیر (شکل (۱۰-الف)) و در صورت مثبت بودن این مقادیر، یکی شدگی مسیر (شکل (۱۰-ب)) را خواهیم داشت. در صورت شناسایی جداشدگی مسیر، عامل استخراج‌کننده راه، مسیری با تعداد خط سیرهای بیشتر را ادامه داده و عامل مدیر یک عامل استخراج‌کننده راه جدید برای شاخه دیگر راه فعال می‌کند و در صورت شناسایی یکی شدگی مسیر عامل استخراج‌کننده راه گذرنده از خوشه‌ای با تعداد خط سیرهای کمتر، غیرفعال شد و عامل استخراج کننده راه گذرنده از خوشه‌ای با تعداد خط سیرهای بیشتر ادامه مسیر را پی می‌گیرد

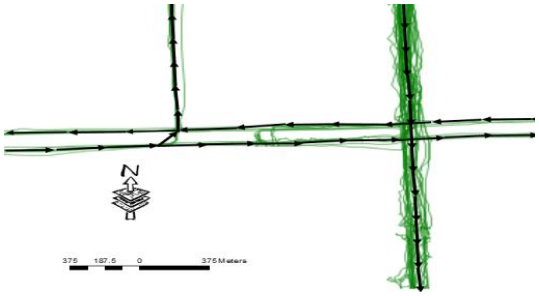


الف

ب

شکل ۱۰: الف- جداشدگی مسیر و ب- یکی شدگی مسیر را نمایش می‌دهد.

ازآنجایی‌که در راه‌کار ارائه‌شده، فضای سه‌بعدی داده‌های مردم‌گستر لحاظ شده است، لذا استخراج راه در تقاطع‌های غیر هم‌سطح نیز به‌صورت قابل قبولی صورت می‌گیرد. شکل (۱۱) یک تقاطع غیر هم‌سطح را نمایش می‌دهد. در تقاطع‌های غیر هم‌سطح برای بخشی از مسیری که زیر مسیر دیگری واقع‌شده است هیچ داده‌ای جمع‌آوری نخواهد شد و با نقص داده مواجه خواهیم بود. عامل استخراج‌کننده بر روی‌داده‌های خط سیر حرکت می‌کند که به‌محض رسیدن به این مناطق قادر به استخراج ادامه مسیر نخواهد بود. لذا در این شرایط این عامل استخراج‌کننده راه توسط عامل مدیر غیرفعال شده و عامل استخراج‌کننده دیگری ادامه مسیر را پی خواهد گرفت.



شکل ۱۱: استخراج راه در تقاطع غیر هم‌سطح

۴- پیاده‌سازی و تحلیل نتایج

در این بخش ابتدا مجموعه داده‌های استفاده‌شده در این مقاله معرفی و پیش‌پردازشی اعمال‌شده به‌منظور کاهش حجم و نویز داده‌ها تشریح می‌شود. سپس نقشه شبکه راه استخراج‌شده با راه‌کار ارائه‌شده توسط عامل‌ها نمایش داده خواهد شد.

محیط چند عامله ارائه‌شده در این پژوهش با استفاده از زبان برنامه‌نویسی جاوا توسعه داده شد. علاوه بر این، به دلیل یکپارچگی بالای جاوا با پایگاه داده متن‌باز PostgreeSQL از این نرم‌افزار و از ماژول PostGIS به‌منظور ذخیره‌سازی داده‌های مکانی و اجرای پرسش و پاسخ‌های مکانی و توصیفی موردنیاز، استفاده شد. کتابخانه متن‌باز Weka [۲۹] تحت مجوز GPL [۱] توسط دانشگاه وایکاتو[۲] منتشرشده است که دارای مجموعه‌ای از الگوریتم‌هایی درزمینه یادگیری ماشین[۳]

[1] General Public Licence
[2] Waikato
[3] Machine Learning

است. این کتابخانه به کاربران این اجازه را می‌دهد تا بتوانند از این الگوریتم‌ها در کد جاوای موردنظر خود استفاده کنند. در این مقاله از این کتابخانه به‌منظور انجام خوشه‌بندی (بخش ۲.۳.۲) استفاده شد.

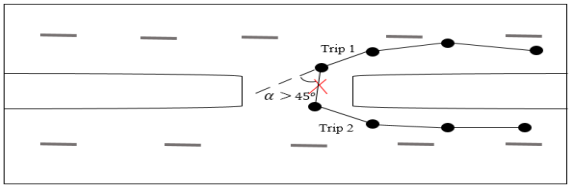
۴-۱- داده‌های مورد استفاده و پیش‌پردازش داده‌ها

داده‌های مورداستفاده در این پژوهش، داده‌های پروژه GeoLife GPS [۳۰] مربوط به شهر پکن است. این مجموعه داده[1] توسط ۱۸۲ کاربر در مدت سه سال (آوریل ۲۰۰۷ تا اوت ۲۰۱۲) جمع‌آوری‌شده است. نرخ نمونه‌برداری این داده‌ها متفاوت است اما ۹۱.۵٪ [۳۰] داده‌ها از توزیع مناسبی برخوردار هستند. به عبارتی این داده‌ها هر ۱-۵ ثانیه و یا بافاصله ۵-۱۰ متر نمونه‌برداری شده‌اند. هر خط سیر در این مجموعه داده با استفاده از مجموعه‌هایی از نقاط بافاصله زمانی مشخص، نمایش داده می‌شود که شامل طول و عرض جغرافیایی، ارتفاع و زمان نمونه‌برداری (x,y,z,timestamp) و در سیستم مختصات سه‌بعدی WGS 84 می‌باشند. از طرفی OSM از وب-مرکاتور[2] به‌عنوان سیستم تصویر استفاده می‌کند. تبدیل داده‌های خط‌سیر از سیستم مختصات سه بعدی WGS84 به سیستم تصویر سه بعدی وب-مرکاتور و برعکس امکان‌پذیر است که در این تبدیل مولفه ارتفاع نیز ترانسفورم می‌شود. علاوه بر داده‌های خط سیر مورد استفاده در این مقاله، داده‌های خط سیری جمع‌آوری شده بصورت سه بعدی، می‌توانند به‌عنوان ورودی روش ارائه شده در این مقاله در نظر گرفته شوند.

نویز ذاتی GPS در دقت و صحت داده‌های جمع‌آوری‌شده تأثیرگذار است. در شرایط خاصی این نویز می‌تواند منجر به ایجاد راه‌های جعلی گردد. همچنین این داده‌ها دارای اشتباه نیز می‌باشند. علاوه بر این، وجود داده‌های اضافی[3] می‌تواند باعث اعمال محاسبات اضافی به سیستم شود. نمونه‌ای از داده‌های ضافی داده‌هایی است که کاربران هنگامی‌که در ترافیک قرارگرفته‌اند، جمع‌آوری می‌کنند.

داده‌های خط سیر زمانی می‌توانند بیانگر راه باشند که توسط خودروهای در حال حرکت جمع‌آوری‌شده باشند [۲۶]. همچنین ازآنجایی‌که داده‌های جمع‌آوری‌شده تشکیل‌دهنده یک خط سیر، بر اساس توالی زمانی به یکدیگر متصل می‌شوند، لذا در برخی از موارد یک اتصال غیرمنطقی بین سفرهای[4] مختلف وجود دارد. بنابراین به‌منظور استفاده از این داده‌ها باید این سفرها از یکدیگر جدا شوند [۳۱]. ازآنجایی‌که هدف ارائه راهکاری باقابلیت کار با داده‌های با حجم بالا است، در ابتدا پالایه‌ای، با در نظر گرفتن قوانین زیر به‌منظور کاهش بار محاسبات اضافی و عملکرد بهتر سیستم، اعمال می‌گردد.

- قانون ۱: داده‌هایی با سرعت کمتر از ۱۰ کیلومتر بر ساعت حذف خواهند شد [۲۶].
- قانون ۲: اگر فاصله بین دونقطه بیشتر از ۳۰۰ متر و یا اینکه تغییر جهت[5] آن‌ها بیشتر از ۴۵ درجه باشد، داده‌های خط سیر دو سفر مجزا در نظر گرفته می‌شوند [۳۱]. شکل (۱۲) این حالت را به‌خوبی نمایش می‌دهد. سپس سفرهایی با تعداد نقاط کمتر از خط سیر ۲۰ عدد از مجموعه داده حذف می‌شوند.



شکل ۱۲: جداسازی سفرهایی با زاویه بیش از ۴۵ درجه

[1] Data set
[2] Web Mercator
[3] Redundant data
[4] Trip: هر یک از فعالیت های داوطلب (کاری، ورزشی، تفریحی و ...) یک سفر نامیده می شود.
[5] Direction

با اعمال پالایه‌های فوق، داده‌های اضافی حذف خواهند شد. همچنین به‌منظور کاهش نویز داده‌ها روش‌های مختلفی نظیر روش پالایه کالمن و پالایه‌های میانگین و میانه وجود دارد که شرح این روش‌ها خارج از حوصله این مقاله است. یکی از مزایای پالایه میانه حساسیت کمتر آن به داده‌های پرت[1] است [۳۲]. در این مقاله از پالایه میانه به‌منظور کاهش نویز داده‌های خط سیر استفاده شد.

پس از اعمال پیش‌پردازش، داده‌های باقیمانده به‌منظور استخراج راه توسط عامل‌ها مورداستفاده قرار گرفت. در ادامه نتایج استخراج راه بر اساس راه‌کار ارائه‌شده در این پژوهش ارائه‌شده است.

۴-۲- نمایش نقشه راه

شکل (۱۳) نقشه شبکه راه برای منطقه موردمطالعه را نمایش می‌دهد. در این شکل، بخش مشخص‌شده ۱ راه استخراج‌شده در جداشدگی و یکی شدگی را نمایش می‌دهد. ازآنجایی‌که هر مسیر توسط یک عامل استخراج می‌شود، لذا استفاده از راه‌کار ارائه‌شده در استخراج راه در جداشدگی و یکی شدگی به‌خوبی صورت گرفته و از اتصال گره‌های غیر مرتبط جلوگیری می‌شود. بخش مشخص‌شده ۲ در شکل (۱۳) یک تقاطع غیر هم‌سطح را نمایش می‌دهد. ازآنجایی‌که در راه‌کار ارائه‌شده ارتفاع داده‌های مردم‌گستر نیز لحاظ شده است، لذا استخراج راه در تقاطع‌های غیر هم‌سطح نیز امکان پذیر است. در تقاطع‌های غیر هم‌سطح برای بخشی از مسیری که زیر مسیر دیگر قرار دارد، هیچ داده‌ای جمع‌آوری نخواهد شد و حرکت عامل استخراج‌کننده بر روی‌داده‌های خط سیر به‌منظور استخراج راه امکان‌پذیر نخواهد بود. لذا در این نوع ساختارها استخراج راه نیازمند پس پردازشی به‌منظور استخراج راه است. بخش مشخص‌شده ۳ در شکل (۱۳) یک چهارراه را نمایش می‌دهد. حجم بالای داده‌های خط سیر جمع‌آوری‌شده در چهارراه‌ها استخراج راه در این ساختارها را با مشکل مواجه می‌کند. به‌طوریکه عملکرد عامل استخراج‌کننده در تعیین گره‌های مسیر در چهارراه‌ها با مشکل مواجه می‌کند. لذا استخراج راه در چهارراه‌ها پس از تعیین چهارراه، نیازمند پس پردازش است. که در این مقاله به این موضوع پرداخته نشده است.

۴-۳- ارزیابی نتایج

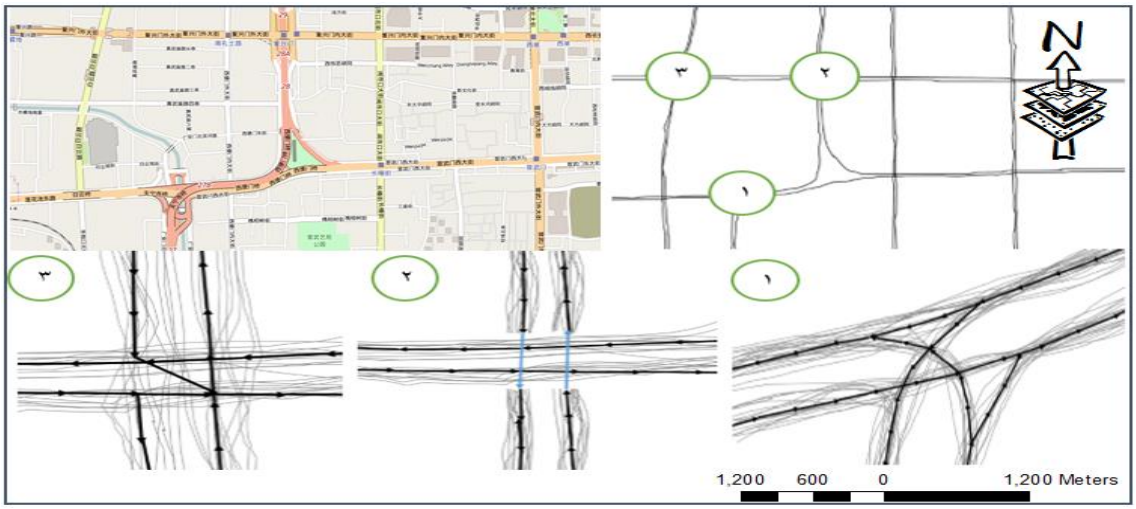
به منظور ارزیابی عملکرد راه‌کار استفاده شده در این پژوهش روش پیشنهاد شده توسط [۳۳] به‌عنوان روش پایه مورد استفاده قرار گرفت. در این ارزیابی نقشه OSM موجود به‌عنوان نقشه مبنا استفاده شد و داده‌های خط سیر موجود برای ۵۶ راه مورد ارزیابی قرار گرفت. شکل (۱۴) نقشه مبنا و نقشه تولید شده را به‌صورت نمایش می‌دهد.

در روش استفاده شده با محاسبه پارامترهای SPURIOUS، MISSING و F_score دقت نقشه تولیدشده نسبت به نقشه مبنا تعیین می‌شود. پارامتر MISSINGبه‌صورت نسبتی از تعداد راه‌های موجود در نقشه مبنا که در نقشه تولید شده وجود ندارند به کل راه‌های موجود در نقشه مبنا و پارامترSPURIOUS به‌صورت نسبتی از تعداد راه‌های موجود در نقشه تولید شده که در نقش مبنا وجود ندارند به کل راه‌های موجود در نقشه تولید شده تعریف می‌شوند. مقدار پارامتر MISSING بیانگر تعداد راه‌هایی از نقشه مبنا که در نقشه تولید شده استخراج نشده‌اند و مقدار پارامتر SPURIOUS بیانگر تعداد راه‌های تولید شده جعلی در نقشه تولید شده است. همچنین پارامتر F_Score که میانگینی از دو پارامتر SPURIOUS و MISSING است، به‌منظور ارائه دقت نهایی نقشه تولید شده محاسبه شد. مقادیر پایین پارامتر SPURIOUS و MISSING و مقدار بالای پارامتر F_score بیانگر دقت بالای نقشه تولید شده نسبت به نقشه مبنا می‌باشد.

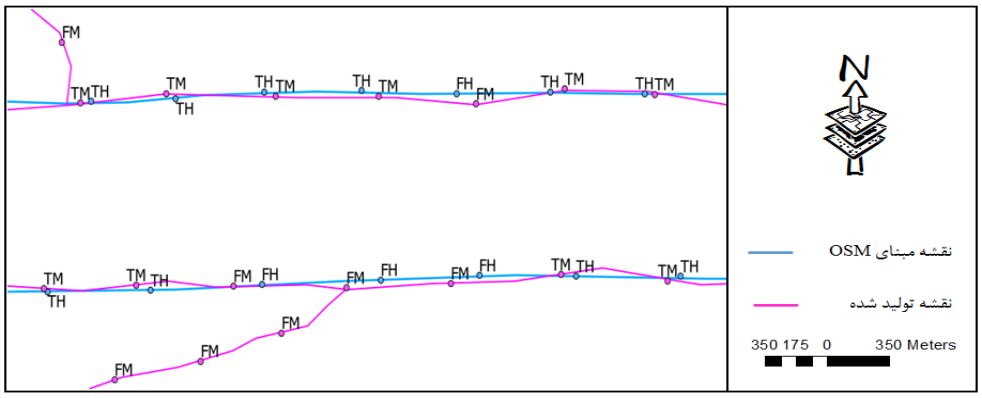
[1] Outlier

به‌منظور تعیین پارامترهای SPURIOUS و MISSING، راه‌های موجود در نقشه مبنا و نقشه تولیدشده به پاره‌خط‌هایی مساوی تقسیم و سپس مراکز هرکدام از این پاره‌خط‌ها تعیین شد. مراکز پاره خطها در نقشه مبنا به‌عنوان نقاط "H" و مراکز پاره‌خطها در نقشه تولیدشده به‌عنوان نقاط "M" در نظر گرفته شد. به‌منظور تعیین پارامتر MISSING، برای هر یک از نقاط H وجود و یا عدم وجود نقاط M در شعاعی مشخص بررسی خواهد شد. درصورتی‌که در این شعاع مشخص شده حداقل یک نقطه M کشف شود، این نقطه به‌عنوان TH و در غیر این صورت این نقطه به‌عنوان FH در نظر گرفته خواهد شد (شکل (۱۴)).

به‌طور مشابه، به منظور تعیین پارامتر SPURIOUS با حرکت بر روی نقاط M در نقشه تولید شده، درصورتی‌که در شعاع مشخص شده حداقل یک نقطه H کشف شود، این نقطه به‌عنوان TM و در غیر این صورت به‌عنوان FM در نظر گرفته خواهد شد (شکل (۱۴)).



شکل ۱۳: نمایش نقشه شبکه راه برای منطقه موردمطالعه. ساختارهای مشخص‌شده ۱، ۲ و ۳ به ترتیب یک جداشدگی و یکی شدگی، تقاطع غیر هم‌سطح و یک تقاطع هم‌سطح را نمایش می‌دهد.



شکل ۱۴: نقاط TH و FH و همچنین نقاط TM و FM را نمایش می‌دهد.

با شمارش تعداد نقاط TH، FH، TM و FM پارامترهای SPURIOUS و MISSING به‌صورت رابطه‌های (۸ و ۹) محاسبه می‌شود.

$$spurious = \frac{\#FM}{\#FM + \#TM} \quad \text{رابطه (۸)}$$

$$\text{missing}=\frac{\#FH}{\#FH+\#TH}$$ رابطه (۹)

پس از تعیین پارامترهای SPURIOUS و MISSING، دقت نقشه تولیدشده نسبت به نقشه مبنا با استفاده از مقدار F_score به‌صورت رابطه (۱۰) قابل‌بیان است.

$$F_score=2\times\frac{(1-spurious)(1-\text{missing})}{(1-spurious)+(1-\text{missing})}$$ رابطه (۱۰)

پس از اجرای روش ارزیابی برای داده‌های خط سیر موجود مربوط به ۵۶ راه، با در نظر گرفتن مقادیر ۹ و ۱۵ به‌ترتیب به‌عنوان طول پاره‌خط‌های تقسیم‌بندی و فاصله شعاعی، مقادیر پارامترهای F_score، MISSING و SPURIOUS به ترتیب برابر با ۰٫۷۰ و ۰٫۳۲ و ۰٫۲۸ حاصل شد. با مقایسه نتایج به‌دست آمده با نتایج مقالات [۱۰, ۳۳] به‌نظر می‌رسد نقشه تولید شده در این مقاله از دقت قابل قبولی برخوردار است. با این حال به دلیلی عدم دسترسی به داده‌های خط‌سیر استفاده شده در مقالات [۱۰, ۳۳] نتایج حاصل از این بخش با نتایج این مقالات بصورت عددی مورد مقایسه قرار نگرفت.

۵- نتایج و پیشنهادات

ایستا بودن روش‌های ارائه‌شده به‌منظور استخراج راه از داده‌های مردم‌گستر باعث می‌شود که علیرغم ماهیت پویایی این داده‌ها که بیانگر تغییرات محیط هستند، عمل بهنگام رسانی به‌صورت گسسته و در بازه‌های زمانی قابل‌توجه صورت گیرد. در این مقاله با الهام از نحوه عملکرد عامل‌های کاوشگر وب در محیط وب، توسعه عامل‌هایی به‌منظور بهنگام رسانی نقشه‌های راه مردم‌گستر در دستور کار قرار گرفت. استفاده از عامل‌ها به‌عنوان مولفه‌های نرم‌افزاری باقابلیت حس محیط و عکس‌العمل در برابر تغییرات لحظه‌ای محیط، می‌تواند سازوکار ایستایی ارائه‌شده را به سازوکارهای پویا تبدیل نماید.

بهنگام‌رسانی پیوسته نقشه‌های را با توجه به داده‌های خط سیر گزارش شده توسط کاربران، با دو مشکل حجم بالای محاسبات و نیز ناسازگاری داده‌های مردم‌گستر روبرو است. مشکل حجم بالای محاسبات با تقسیم وظیفه‌ی مشخص میان عامل‌های طراحی شده در یک سیستم چند عامله صورت گرفت. همچنین به‌منظور جلوگیری از ورود داده‌های خط سیر ناسازگار، در سیستم توسعه داده شده عامل مدیر فیلتری که شامل دو پارامتر زمانی و تعداد است را برای داده‌های خط سیر جمع‌آوری شده توسط کاربران بررسی می‌کند.

سازوکار ارائه‌شده در این مقاله به سیستم محاسبات اضافی اعمال نمی‌کند و محاسبات تنها درزمانی صورت می‌گیرد که تغییر متمایزکننده‌ای در فضای داده‌های خط سیر صورت گیرد. ازآنجایی‌که برای هر مسیر تنها یک عامل استخراج‌کننده راه فعال می‌شود، لذا استخراج راه در جداشدگی‌ها و یکی شدگی‌ها و تقاطع‌های غیر هم‌سطح، به‌خوبی صورت می‌گیرد.

به علت عدم پشتیبانی سیستم‌های تعیین موقعیت در تقاطع‌های غیر هم‌سطح برای بخشی از مسیر که در زیر مسیر دیگری واقع‌شده است و درنتیجه عدم جمع‌آوری داده‌های خط سیر برای این بخش، پس پردازشی به‌منظور استخراج راه موردنیاز خواهد بود. علاوه بر این تراکم بالای داده‌های خط سیر جمع‌آوری‌شده در مرکز چهارراه‌ها به پیچیدگی‌های استخراج راه در آن می‌افزاید، بطوریکه حرکت عامل استخراج‌کننده بر روی‌داده‌های خط سیر در این ساختارها و تعیین گره‌های مسیر در چهارراه‌ها امکان‌پذیر نیست. لذا استخراج راه در این ساختارها نیازمند مطالعه و ارائه روش‌های مناسب‌تری است. همچنین عدم علاقه‌مندی و مشارکت مردم در به اشتراک‌گذاری موقعیت خود به مسائل جامعه‌شناسی و فرهنگی یک اجتماع بستگی دارد و این خود زمینه گسترده‌ای به‌منظور تحقیق است و به‌عنوان کارهای آینده پیشنهاد می‌شود. علاوه بر این پیاده‌سازی و اجرای الگوریتم پیشنهادی در مقیاس بالا با استفاده از محاسبات توزیع‌شده امکان‌پذیر است. لازم به ذکر است که

پیاده‌سازی چنین ساختار توزیع‌شده‌ای نیازمند مطالعات و برنامه‌نویسی ویژه‌ای است و برنامه‌نویسی انجام شده در حال حاضر نمی‌تواند بدون تغییر در بستر توزیع‌شده اجرا شود. اما به نظر می‌رسد ساختار ارائه شده قابلیت توسعه جهت اجرا در بستر توزیع‌شده را دارا می‌باشد. موارد بیان‌شده پیشنهادهایی است که در پژوهش‌های آتی مورد توجه نویسندگان قرار خواهد گرفت. علاوه بر آنچه بیان شد در پژوهش‌های آتی روشی به‌منظور استخراج اطلاعات کیفی راه‌های استخراج‌شده موردبررسی قرار خواهد گرفت.

مراجع

[1] W. Holzapfel, M. Sofsky, and U. Neuschaefer-Rube, "Road profile recognition for autonomous car navigation and Navstar GPS support", Aerospace and Electronic Systems, IEEE Transactions on, Vol. 39, pp. 12, 2003.

[2] B. Wessel, "Road network extraction from SAR imagery supported by context information", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science,Vol. 35, pp. 360-36, 2004.

[3] C. Zhang, "Towards an operational system for automated updating of road databases by integration of imagery and geodata", ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 58, pp. 166-186, 2004.

[4] M.F. Goodchild, "Citizens as sensors: the world of volunteered geography", GeoJournal, Vol. 69, pp. 211-221, 2007.

[5] N. Mohammadi, "Improving Logical Consistency in Volunteered Geospatial Information Environments", Ph.D Thesis, Department of Spatial Information Science and Engineering at K.N.Toosi University of Technology, 2014.

[6] J. Davics, A.R. Beresford, and A. Hopper, "Scalable, distributed, real-time map generation", Pervasive Computing, IEEE, Vol. 5, p p. 47-5, 2006.

[7] C . Chen, and Y. Cheng, "Roads digital map generation with multi-track GPS data", presented at the Education Technology and Training, USA, 2008.

[8] W. Shi, S. Shen, and Y. Liu, "Automatic generation of road network map from massive GPS, vehicle trajectories", presented at the Intelligent Transportation Systems, 12th International IEEE Conference,USA, 2009.

[9] J. Biagioni, T. Gerlich, T. Merrifield, and J. Eriksson, "Easytracker: automatic transit tracking, mapping, and arrival time prediction using smartphones", presented at the proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems, USA, 2011.

[10] J. Biagioni, and J. Eriksson, "Map inference in the face of noise and disparity", presented at the Proceedings of the 20th International Conference on Advances in Geographic Information Systems, USA, 2012.

[11] L. Cao, and J. Krumm, "From GPS traces to a routable road map", presented at the Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, USA, 2009.

[12] F. Chazal, D. Chen, L. Guibas, X. Jiang and C. Sommer, "Data-driven trajectory smoothing", presented at the Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, USA, 2011.

[13] S. Edelkamp, and S. Schrödl, Route planning and map inference with global positioning traces.USA: Springer press, 2003.

[14] S. Schroedl, K. Wagstaff, S. Rogers, P. Langley and C. Wilson, "Mining GPS traces

for map refinement", Data mining and knowledge Discovery, Vol. 9, pp. 59-87, 2004.

[15] T. Guo, K. Iwamura, and M. Koga, " Towards high accuracy road maps generation from massive GPS Traces data", presented at the Geoscience and Remote Sensing Symposium, Spain, 2007.

[16] S.Worrall, and E. Nebot, "Automated process for generating digitised maps through GPS data compression", presented at the Australasian Conference on Robotics and Automation, Australian, 2007.

[17] S. Jang, T. Kim, and S. Lee, "Map generation system with lightweight GPS trace data", presented at the Advanced Communication Technology (ICACT), The 12th International Conference, Denmark, 2010.

[18] G. Agamennoni, J.I. Nieto, and E.M. Nebot, "Robust inference of principal road paths for intelligent transportation systems", Intelligent Transportation Systems, IEEE Transactions on, Vol. 12(1), pp. 298-308, 2011.

[19] Y. Wang, Y. Zhu, Z. He, Y. Yue, and Q. Li, "Challenges and opportunities in exploiting large-scale GPS probe data", HP Laboratories, Vol.21, 2011.

[20] S. Russell, and P. Norvig, Artifitial Intelligence A Modern Approach. New Jersey: pearson press, 2010.

[21] S. Dhenakaran, and K.T. Sambanthan, "Web crawler-an overview", International Journal of Computer Science and Communication, Vol. 2, pp. 265-267, 2011.

[22] P. Vanicek, and E.J. Krakiwsky, Geodesy: the concepts. Netherlands: Elsevier, 2013.

[23] L. Zhang, F. Thiemann, and M. Sester," Integration of GPS traces with road map", presented at the Proceedings of the second international workshop on computational transportation science, USA, 2010.

[24] Y. Chen, and J. Krumm, "Probabilistic modeling of traffic lanes from GPS traces", presented at the Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, USA, 2010.

[25] F.v. Diggelen, "SYSTEM DESIGN & TEST-GNSS Accuracy-Lies, Damn Lies, and Statistics", GPS World, Vol. 18, pp. 26-33, 2007.

[26] J. Li, Q. Qin, C. Xie, and Y. Zhao, "Integrated use of spatial and semantic relationships for extracting road networks from floating car data", International Journal of Applied Earth Observation and Geoinformation, Vol. 19, pp. 238-247, 2012.

[27] R.O. Duda, P.E. Hart, and D.G. Stork, Pattern classification. New York: John Wiley & Sons, 2001.

[28] R.M. Neal, and G.E. Hinton, "A view of the EM algorithm that justifies incremental, sparse, and other variants ,in Learning in graphical models", 1998, Springer. p. 355-368.

[29] Waikato, W.u.o. accessed 5 Jan 2015; Available from: http://www.cs.waikato.ac.nz/ml/weka/downloading.html.

[30] Trajectories, G.G., accessed 5 Jan 2015: p. Available from: http://research.microsoft.com/en-us/downloads/b16d359d-d164-469e-9fd4-daa38f2b2e13./

[31] D. Zhang, N. Li, Z.-H. Zhou, C. Chen, L. Sun and S. Li, "iBAT: detecting anomalous taxi trajectories from GPS traces", presented at the Proceedings of the 13th international conference on Ubiquitous computing, China, 2011.

[32] Y. Zheng, and X. Zhou, Computing with spatial trajectories. USA: Springer Science & Business Media, 2011.

[33] J. Biagioni, and J. Eriksson, "Inferring Road Maps from Global Positioning System

Traces: Survey and Comparative Evaluation", Transportation research record, pp. 61-71, 2012.

K. N. TOOSI UNIVERSITY OF TECHNOLOGY
FACULTY OF GEODESY AND GEOMATICS ENGINEERING

Journal of Geospatial Information Technology
Vol.5, No.2, Summer 2017

Designing and developing a multi-Agent system for automatic extraction of road geometry by the Crawler-Agent

Reza Mohammadi [1], Mehdi Farnaghi*[2]

1- Ms.c in geospatial information system, Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology.
2- Assistant professor in GIS, Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology.

Abstract

Up-to-date digital maps play a significant role as required data in various areas such as navigation, tourism, traffic control, urban and interurban fleet management, ITS, web-based map services and location-based services. Nowadays, VGI is known as an important resource to generate digital maps. In the last few years, the use of these resources to produce low cost, up-to-date, and reliable maps have been considered. Previous studies have utilized static methods for automatic road extraction from volunteered trajectory data. However, they have not considered the continuous changes of the road network which are reported dynamically by the volunteered trajectory data of users. So, due to their static nature they are not efficient to update road maps dynamically. Considering the fact that crawler-agents are capable of sensing and reacting changes in the environment it is possible to dynamically update road maps based on the latest changes presented by the trajectory data. The key idea of this research is to design and develop crawler-agents that dynamically search and update road maps from volunteered trajectory data. In this article, inspired by the continues monitoring of the web by web Crawler-Agents, a dynamic method is presented to updating road maps with considering the collected participatory data that has been gathered and recorded by the GPS sensors on the mobile device of users. To do so, various group of agents that move on trajectory data to extract road map are designed and developed. EM clustering is used to estimation of the network nodes. Also, a heuristic method is developed to link the nodes which are related to each other. Furthermore, an efficient approach is presented to extraction of road networks e by applying three dimensional trajectory data in the case of road's challenge structure.

Key words: Agent, Crawler-Agent, Automatic road extraction, Trajectory data, VGI, EM Clustering.

Correspondence Address No. 1346, ValiAsr Street, Mirdamad cross, Geomatics Engineering Faculty, K.N. Toosi University of Technology, Tehran, Iran.
Tel :
Email: farnaghi@kntu.ac.ir