

## طراحی و توسعه یک سیستم چندعامله به منظور استخراج خودکار هندسه راه توسط عامل‌های کاوشگر

رضا محمدی<sup>۱</sup>، مهدی فرنقی<sup>۲\*</sup>

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 سیستم‌های خدمات رسانی مرتبط با مکان نظیر ناوبری، توریسم، مدیریت ترافیک و سامانه‌های مکان‌مبنا ایفا می‌کنند. امروزه استفاده از داده‌های مردم‌گستر، به‌منظور تولید نقشه‌های رقومی مورد توجه قرار گرفته است. روش‌های ارائه‌شده به‌منظور استخراج نقشه رقومی راه با استفاده از داده‌های مردم‌گستر، 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 مردم‌گستر به‌طور بهینه صورت نگیرد. استفاده از عامل‌ها به‌دلیل قابلیت حس محیط و توانایی عکس‌العمل در برابر تغییرات محیط، تولید نقشه به‌صورت پویا را امکان‌پذیر 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 پیش‌پیوسته محیط وب توسط عامل‌های 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 خوشه‌بندی EM به‌منظور تعیین گره‌های شبکه راه مورد استفاده قرار گرفته و سپس روشی ابتکاری به‌منظور اتصال گره‌های 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 مردم‌گستر، خط سیر، استخراج خودکار راه، عامل، عامل کاوشگر.

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

تلفن: ۳۸۸۸۷۷۰۷-۱

## 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<sup>۱</sup> رو به افزایش است. این شبکه‌ها بر پایه‌ی فناوری وب ۲ و در تعامل پیوسته با کاربران ایجاد شده‌اند. در چنین شرایطی هر کاربر، به‌عنوان یک سنجنده<sup>۲</sup> 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 روی داده‌های استاندارد امکان‌پذیر نیست، مورداستفاده قرار گیرند [۵]. همچنین داده‌های مردم‌گستر، به‌خاطر ماهیت پویایی آن‌ها، معمولاً به‌هنگام 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<sup>۳</sup> بر یکدیگر، اقدام به استخراج هندسه و جهت راه نمودند. همچنین، به‌منظور کاهش نویز داده‌ها، که می‌تواند باعث ایجاد راه‌های کاذب شود، پالایه<sup>۴</sup> ویژه‌ای اعمال نمودند. برخی از محققین [۱۳، ۱۴، ۱۵، ۱۶، ۱۷، ۱۸ و ۱۹] از روش‌های خوشه‌بندی به‌منظور استخراج راه استفاده نمودند. 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 ک-مینز<sup>۵</sup> در نظر گرفته می‌شوند. در ادامه از الگوریتم ک-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 ایستاد 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. به‌عبارتی دیگر الگوریتم‌های پیشنهادشده قابلیت درک تغییرات محیط (که این تغییرات می‌تواند راهی که اخیراً ساخته شده و

<sup>3</sup> Trajectory data<sup>4</sup> Filter<sup>5</sup> k-means<sup>6</sup> Static<sup>1</sup> Social networks<sup>2</sup> Sensor

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 در زمینه 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 روی داده‌های 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 به‌هنگام‌رسانی 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<sup>۲</sup> صفحات مفید وب توسط موتورهای جست‌وجو<sup>۳</sup> 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

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 ارائه‌شده در نگاشت تغییرات محیط واقعی با مشکل مواجه گردند. از طرفی به‌هنگام‌رسانی 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 ارائه‌شده در بخش‌های بعدی این مقاله تشریح خواهد شد.

در این مقاله با الهام از نحوه پایش پیوسته محیط وب توسط عامل‌های کاوشگر وب<sup>۱</sup>، با توجه به تغییرات پیوسته داده‌های جمع‌آوری‌شده خط سیر توسط کاربران، روشی پویا به‌منظور به‌هنگام‌رسانی نقشه‌های شبکه راه، ارائه‌شده است. بدین منظور عامل‌هایی توسعه داده شد که با حرکت بر روی داده‌های خط سیر نسبت به تغییرات محیط واکنش نشان داده و به‌هنگام‌رسانی نقشه شبکه راه‌ها پردازند. 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 به‌هنگام‌رسانی، با در نظر گرفتن تغییرات محیط که توسط خط‌سیرهای جمع‌آوری‌شده کاربران گزارش می‌شوند، به‌خوبی صورت گیرد. این عامل‌ها با در نظر گرفتن فضای سه‌بعدی داده‌های مردم‌گستر عملکرد مناسبی در استخراج راه در جداسازی و یکی‌شدگی مسیر و تقاطع غیرهم‌سطح 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

<sup>2</sup> Indexing

<sup>3</sup> Search engine

<sup>1</sup> Web Crawler-Agent

صفحات وب می‌پردازند. در این پژوهش با الهام از نحوه پوشش 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 طراحی شده است:

- عامل کاوشگر محیط باهدف کاوش فضای سه‌بعدی داده‌های مکانی خط سیر
- عامل استخراج‌کننده راه باهدف استخراج راه از طریق حرکت بر روی داده‌های 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



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

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.

## ۲-۱- عامل مدیر

عامل مدیر وظیفه مدیریت عامل‌های کاوشگر محیط و عامل‌های استخراج‌کننده راه را بر عهده دارد. این عامل به‌منظور جلوگیری از اجرای الگوریتم استخراج راه برای کل فضای داده‌های خط سیر، فضای داده‌های خط سیر تحت پوشش خود را به مناطق کوچک‌تری

که میزان استفاده از CPU کمتر از ۷۰ درصد است به صورت بهینه عمل کرده و پردازش‌ها را با سرعتی مناسب اجرا می‌کند. بر این اساس در پیاده‌سازی عامل مدیر کتابخانه<sup>۱</sup> SIGAR به منظور پایش میزان CPU استفاده شده به کار گرفته شد. عامل مدیر با استفاده از این کتابخانه در هر لحظه مقادیر CPU را بررسی نموده و تا زمانی که مقدار CPU کمتر از ۷۰ درصد باشد، در صورت لزوم عامل‌های استخراج‌کننده راه جدیدی اضافه می‌کند. همچنین در صورتی که مقدار CPU بیشتر از ۷۰ درصد باشد، عامل‌هایی را غیرفعال خواهد نمود. بدین ترتیب عامل مدیر سعی در ایجاد توازن لازم در کل سیستم چند عامله را دارد.

عامل‌های استخراج‌کننده راه موقعیت خود در محیط را به عامل مدیر گزارش می‌دهد. از آنجایی که 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 به محض 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 جمع‌آوری شده از یک مسیر که فراوانی کمی دارند و همچنین خط سیرهای قدیمی کار استخراج راه انجام نخواهد گرفت. شکل (۲) نحوه جستجوی عامل

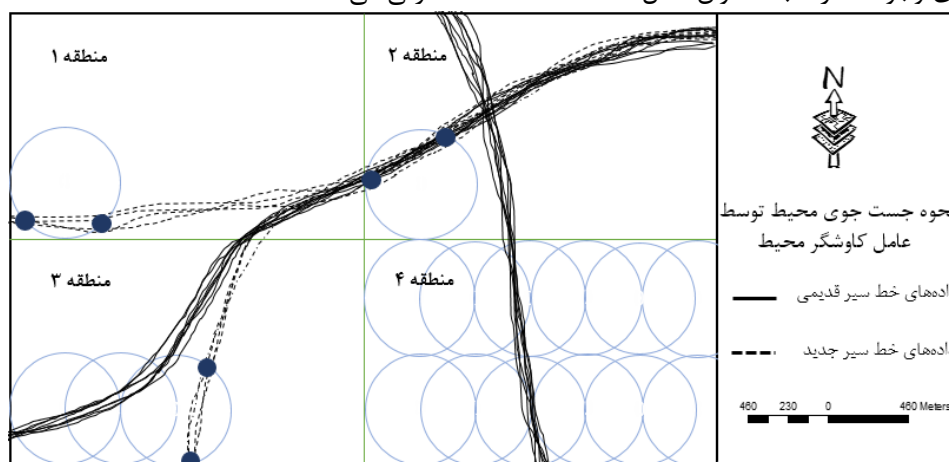
به‌منظور تعیین داده‌های خط سیر جدید، عامل مدیر به هر یک از مناطق یک عامل کاوشگر محیط اختصاص می‌دهد. عامل کاوشگر محیط منطقه مربوط به خود را تا انتها کاوش نموده و به محض مشاهده تعداد قابل قبول از داده‌های خط سیری که بیانگر تغییرات آن منطقه باشند، موقعیت داده‌های خط‌سیر جدید را به عامل مدیر گزارش می‌دهد. 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 داده‌های 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 به عنوان پارامتر زمانی و تعداد ۲۰ داده خط سیر به عنوان پارامتر تعداد در نظر گرفته شد. این پارامترها در سیستم چند عامله توسعه داده شده قابل تنظیم می‌باشند و در پیاده‌سازی عملیاتی می‌توان بر اساس نیاز آن‌ها را، به‌منظور رسیدن به نتایج بهتر، تغییر داد.

از آنجایی که 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 به صورت 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

<sup>1</sup> <https://support.hyperic.com/display/SIGAR/Home>

لذا برای این مناطق هیچ عامل استخراج کننده راهی فعال نخواهد شد. در برخی دیگر از این مناطق ممکن است عامل کاوشگر داده‌های خط سیر جدید را کشف کند (به‌عنوان مثال مناطق ۱، ۲ و ۳)، برای این مناطق یک عامل استخراج کننده 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

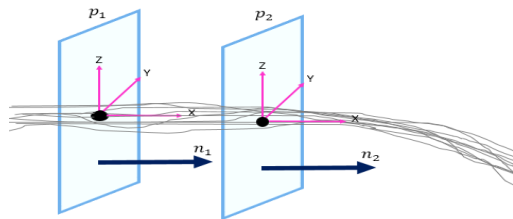
به‌طوریکه عامل استخراج کننده با حرکت در راستای گرادیان صفحه می‌تواند بر روی داده‌های خط سیر حرکت کند، تعیین می‌شود. بدین منظور، این صفحه حول محور  $Z$  سیستم مختصات محلی تعریف شده برای آن، دوران داده می‌شود. پس از تعیین بهترین حالت صفحه، به‌منظور تعیین گره‌های<sup>۱</sup> شبکه راه، بر روی نقاط حاصل از تقاطع صفحه با خط سیرها خوشه‌بندی صورت می‌گیرد. سپس عامل استخراج کننده 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،

<sup>1</sup> Node

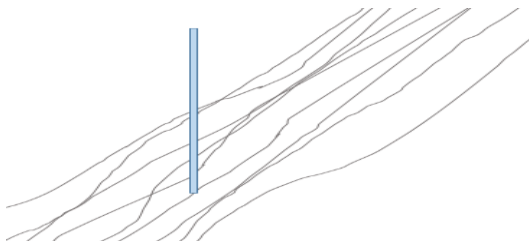
کاوشگر محیط در هر منطقه را به‌صورت شماتیک نمایش می‌دهد. خطوط کامل، داده‌های خط سیر قدیمی که پیش از این به‌منظور ایجاد نقشه مورد استفاده قرار گرفته است را نمایش می‌دهد و داده‌های خط سیر جدید که بیانگر تغییرات محیط هستند با خط چین نمایش داده شده است. عامل‌های کاوشگر محیط به‌منظور پیدا کردن داده‌های خط 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 (به‌عنوان مثال منطقه ۴)،

## ۲-۳- عامل‌های استخراج کننده 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 استخراج کننده راه در موقعیت گزارش داده شده فعال می‌کند. عامل استخراج کننده راه بر روی داده‌های خط سیر حرکت نموده و به استخراج راه می‌پردازد. شکل (۳) راه‌حل استخراج راه توسط عامل استخراج کننده را به‌صورت شماتیک نمایش می‌دهد. عامل استخراج کننده راه جهت‌های مختلفی برای حرکت پیش رو دارد. به‌منظور استخراج راه، این عامل باید در امتداد داده‌های خط سیر حرکت کند. بدین منظور این عامل در موقعیت خود صفحه‌ای ایجاد نموده سپس بهترین حالت صفحه،



شکل ۳: نحوه ایجاد صفحه و حرکت بر روی داده‌های خط سیر توسط عامل استخراج کننده راه



شکل ۴: صفحه ایجاد شده توسط عامل استخراج کننده در موقعیت خود (پیش از منتقل شده به میانگین نقاط حاصل از تقاطع آن با خط سیرها)

### ۲-۳-۱- انتخاب جهت حرکت مسیر

به منظور مشخص نمودن حرکت عامل استخراج کننده بر روی داده‌های 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 ایجاد شده توسط عامل استخراج کننده را نمایش می‌دهد. این صفحه ممکن است برخی از خط سیرهایی که می‌توانند در استخراج راه شرکت داشته باشند را قطع نکنند. به منظور 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))

به منظور تشکیل اتصال<sup>۱</sup> شبکه راه، گره‌های جدید به گره‌های قبلی متصل می‌شوند. در نهایت به منظور استخراج شبکه راه، عامل استخراج کننده 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 استخراج کننده را تشریح می‌کند.

**Input:** Trajectories of floating cars

**Output:** Road network

```

1: number of square = 1;
2: P ( $A_0(x_0, y_0, z_0, N_0)$ ) ← 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 استخراج کننده راه

در ادامه مراحل استخراج راه توسط عامل استخراج کننده به صورت کامل تشریح می‌شود.

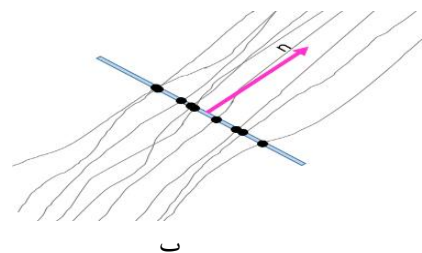
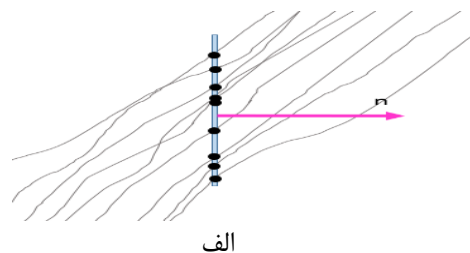
<sup>۱</sup> Edge

تقاطع صفحه با خط سیرها داشته باشد، لذا برای جلوگیری از این مشکل و به منظور اطمینان از عدم حضور مشاهدات اشتباه از آزمون باردا<sup>۱</sup> [۲۲] استفاده می‌شود (خط ۱۰ شبه کد (۱)) و مشاهدات اشتباه حذف می‌شوند. عموماً پیش از آزمون باردا، آزمون فاکتور واریانس ثانویه به منظور 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 واریانس روی داده‌ها اعمال شد.

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

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

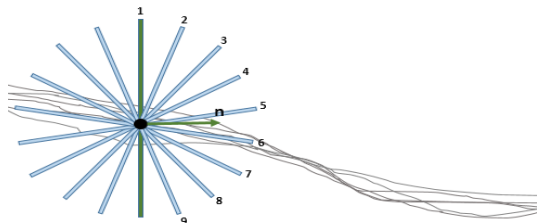


شکل ۵: الف- صفحه منتقل شده به میانگین تقاطع آن با خط سیرها و ب- بهترین حالت صفحه به طوریکه بردار نرمال آن هم جهت با خط سیرها را نمایش می‌دهد.

شکل (۵-الف) صفحه تشکیل شده توسط عامل استخراج‌کننده، پس از انتقال به میانگین مختصات نقاط، را نمایش می‌دهد. این صفحه هنوز صفحه ایده‌آل به منظور حرکت عامل استخراج‌کننده بر روی داده‌های 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)، به منظور حرکت عامل استخراج‌کننده بر روی داده‌های خط سیر است. در این حالت بردار نرمال صفحه بر روی داده‌های خط سیر منطبق شده و جهت حرکت عامل استخراج‌کننده را به خوبی مشخص می‌کند. به منظور 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 محاسبه شده و حالتی از صفحه که انحراف معیار کمتری دارد به عنوان بهترین حالت صفحه انتخاب می‌شود.

پیش از محاسبه انحراف معیار، از آنجایی که به دلیل نوپز ذاتی داده‌های خط سیر ممکن است یک نقطه تغییر زیادی در مقدار انحراف معیار مختصات نقاط حاصل از

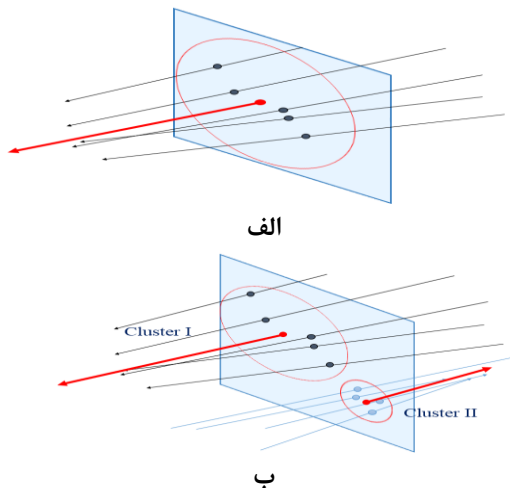
<sup>1</sup> Baarda test



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 جمع‌آوری شده نماینده‌ای از ساختار هندسه راه‌ها باشند. با این فرض می‌توان گفت که داده‌های خط سیر جمع‌آوری شده 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).



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

کمترین تعداد است. لذا انتظار می‌رود که انحراف معیار مختصات نقاط برای حالت ۶ صفحه نسبت به سایر حالت‌ها عدد کمتری باشد. در نتیجه این حالت صفحه "به اشتباه" به‌عنوان بهترین حالت صفحه انتخاب شود. این در حالی است که با توجه به شکل (۶) از آنجایی که 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

به منظور 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<sup>۴</sup> با در نظر گرفتن شرایط توزیع نمونه‌های مشاهده شده ( $\alpha$ ) و مشاهده نشده ( $\beta$ ) تحت پارامترهای شباهت داده شده ( $\theta$ ) 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 قرار گرفته و تا هنگام رسیدن به مقدار مورد انتظار تابع لگاریتم درست‌نمایی تکرار خواهند شد.

$$Q(\theta | \theta^{(t)}) = E_{\beta|\alpha, \theta^{(t)}} \quad \text{رابطه (۳)}$$

$$[\log L(\theta; \alpha, \beta)]$$

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

که در آن  $\alpha$  تعداد نمونه‌های مشاهده شده و  $\beta$  تعداد نمونه‌های مشاهده نشده و  $\theta$  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 ( $\beta$ )، برآوردی از تابع شباهت  $\theta$  تعیین شد. در مرحله دوم به‌منظور بیشینه نمودن این تابع شباهت، رابطه (۴) استفاده شد. سپس بر روی داده‌های مشاهده نشده خوشه‌بندی صورت گرفت. این داده‌های خوشه‌بندی شده به‌عنوان داده‌های مشاهده شده ( $\alpha$ ) در نظر گرفته شده و در تکرار دوم برآورد جدیدی از تابع شباهت  $\theta$  به دست آمد. مراحل تا هنگام رسیدن به مقدار مورد انتظار تابع لگاریتم درست‌نمایی تکرار شد. در نهایت با تکرار مراحل، تعداد و

شکل (۷-الف) تقاطع صفحه با داده‌های خط سیرها جمع‌آوری شده 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، جدانشدگی 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، جدانشدگی و جمع‌شدگی مسیر، گره‌هایی از مسیرهای متفاوت هستند. بنابراین هدف از خوشه‌بندی تعیین تعداد و موقعیت مراکز خوشه یا خوشه‌ها خواهد بود.

از جمله روش‌های خوشه‌بندی می‌توان به روش‌های کا-مینز، FCM، EM، و ISODATA اشاره نمود [۲۷]. از میان روش‌های مطرح در زمینه 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<sup>۱</sup> 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 تعیین شده، 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

<sup>۴</sup> log-likelihood : تابعی است که توزیع طبیعی داده‌ها را نشان می‌دهد.

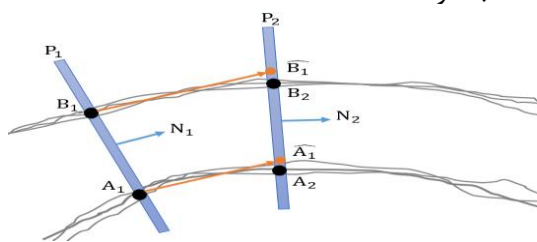
<sup>۱</sup> Expectation Maximization

<sup>۲</sup> Fuzzy C-means

<sup>۳</sup> Maximum Likelihood

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

بعد از تعیین گره‌ها، به منظور تشکیل گراف راه، گره‌های تعیین شده به گره‌های استخراج شده 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 (به منظور تشکیل یک اتصال<sup>۱</sup>) 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 به منظور 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 به منظور 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 واقع شده‌اند، نقطه  $\bar{A}_1$  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

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 تشکیل دهنده 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 جمع‌آوری شده‌اند و همچنین جمعیت خط سیرهای تشکیل دهنده (گذرنده از) آن‌ها بیشتر از حد آستانه  $n$  باشند، خوشه‌های معتبری خواهند بود. به عبارتی خط سیرهای جدید جمع‌آوری شده 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 تشکیل شده در خط ۱۶ شبه کد بررسی خواهد شد. اگر تنها یک صفحه تشکیل شده 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 کد ۱)، 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.

در صورتی که بیش از یک صفحه توسط عامل استخراج‌کننده تشکیل شده باشد، ماتریس وزن به‌منظور اتصال گره‌ها به گره‌های قبلی، تشکیل خواهد شد. این موضوع در ادامه به تفصیل توضیح داده خواهد شد.

<sup>1</sup> Edge

عامل استخراج کننده، استخراج می شود و در صورت همپوشانی مسیر در حال استخراج توسط دو یا چند عامل استخراج کننده، یکی از عامل ها توسط عامل مدیر غیرفعال خواهد شد. بدین منظور عامل استخراج کننده در هر لحظه گزارشی از موقعیت خود، که در حال حرکت بر روی داده های خط سیر به منظور استخراج راه برای مسیر خاصی است، به عامل مدیر ارائه می دهد (خط ۶ شبه کد ۱). عامل مدیر موقعیت این عامل های استخراج کننده را بررسی نموده و به محض همپوشانی راه استخراج شده توسط عامل های استخراج کننده متفاوت، یکی از این عامل ها را غیرفعال خواهد نمود. به این ترتیب کار استخراج راه برای این عامل استخراج کننده به پایان می رسد.

#### ۲-۳-۶- جایگزینی یک مسیر جدید

همان گونه که پیش از این اشاره شد، داده های مردم گستر تغییرات محیط را به خوبی نمایش می دهند. یکی از این تغییرات می تواند جایگزینی یک مسیر جدید به جای مسیر قبلی شود. شکل (۹) یک حالت جایگزینی مسیر را به صورت شماتیک نمایش می دهد. مسیر مشخص شده با خط کامل، مسیر مسدود شده و مسیر مشخص شده با خط چین مسیر جایگزین را نمایش می دهد. از آنجایی که عامل های کاوشگر محیط، در فضای داده های خط سیر حضور دارند و به محض اضافه شدن تعداد قابل قبولی از خط سیرها در محیط، واکنش نشان می دهند، لذا در صورت عدم جمع آوری خط سیرهای جدید برای یک مسیر، آن مسیر تحت عنوان یک مسیر مسدود شده برچسب زده می شود و مسیر جایگزین به نقشه اضافه خواهد شد.

$A_2$  و بار دیگر به  $B_2$  متصل می شوند. اتصال مناسب اتصالی با کمترین اختلاف جهت<sup>۱</sup> این اتصال با نرمال صفحه  $P_1$  خواهد بود. این ویژگی به عنوان دومین ویژگی تشکیل دهنده ماتریس وزن در نظر گرفته می شود.

سومین ویژگی، اختلاف جمعیت نقاط دو گره متوالی است. دو گره متوالی  $A_1$  و  $A_2$  در صورتی به یکدیگر متصل می شوند که خط سیرهایی که از گره  $A_1$  عبور می کنند همان سیرها از گره  $A_2$  نیز عبور کنند [۱۱]، [۱۴]. به عبارتی برای دو گره متوالی انتظار می رود، تعداد نقاط تشکیل دهنده گره با یکدیگر اختلاف کمی داشته باشند.

با در نظر گرفتن سه ویژگی بیان شده، ماتریس وزن اتصال بین گره های دو صفحه  $p_k$  و  $p_{k+1}$  را به صورت رابطه (۵) می توان تعریف نمود:

رابطه (۵)

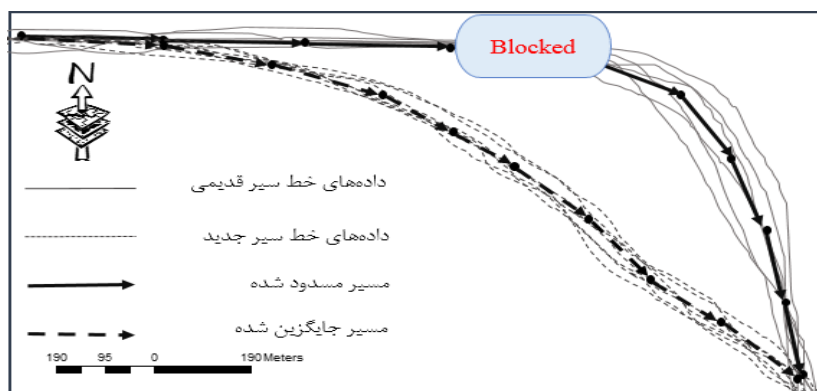
$$W^{(k,k+1)} = \begin{bmatrix} W_{(i=1)(j=1)} & W_{i(j+1)} & \dots & W_{i(j+n)} \\ W_{(i+1)j} & W_{(i+1)(j+1)} & \dots & W_{(i+1)(j+n)} \\ \vdots & \vdots & \vdots & \vdots \\ W_{(i+m-1)j} & W_{(i+m)(j+1)} & \dots & W_{(i+m)(j+n)} \end{bmatrix}$$

که در آن،  $i$  شماره گره از صفحه  $p_k$ ،  $j$  شماره گره کاندید از صفحه  $p_{k+1}$  و درایه های این ماتریس مقادیر وزن اتصالات به یکدیگر است. به عنوان نمونه  $w_{ij} = \text{dist}_{ij} \times d\text{Direction}_{ij} \times d\text{Population}_{ij}$  است که در آن  $\text{dist}_{ij}$  فاصله برآورد گره  $A_i$  ( $\widehat{A}_1$ ) از گره  $A_j$ ،  $d\text{Direction}_{ij}$  اختلاف جهت اتصال تشکیل شده از گره های  $A_i$  و  $A_j$  از نرمال صفحه  $p_k$  و  $d\text{Population}_{ij}$  اختلاف جمعیت گره های  $A_i$  (از صفحه  $p_k$ ) و  $A_j$  (از صفحه  $p_{k+1}$ ) است. در نهایت از بین گره های 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 استخراج کننده راه

همان طور که پیش از این بیان شد، هر مسیر توسط یک

<sup>۱</sup> Direction



شکل ۹: مسیر مسدود شده و مسیر جایگزین

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 روی داده‌های 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. هنگامی که 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.

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

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

### ۲-۴- بهنگام رسانی نقشه راه

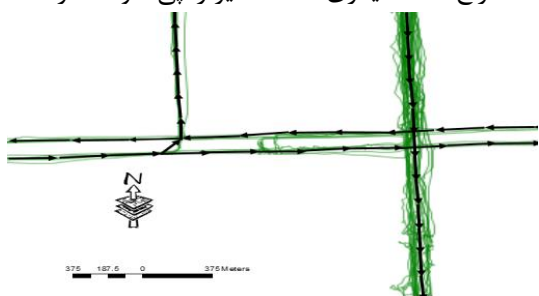
بعد از یک بار تولید نقشه راه، به‌منظور بهنگام رسانی نقشه تولیدشده، عامل‌های کاوشگر محیط، در فضای داده‌های خط سیر حضور دارند و به محض اضافه شدن تعداد قابل قبولی از خط سیرها در محیط، برای منطقه موردنظر، عامل استخراج‌کننده راه، فعال خواهد شد. این عامل پس از حرکت بر روی داده‌های خط سیر و تعیین گره‌های مسیر، در هر گره وجود گره‌های دیگر در اطراف خود را بررسی خواهد نمود. سپس تابع رابطه ۶ برای هر گره بررسی خواهد شد. در صورتی که 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 این کار 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.

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

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. به این ترتیب حضور عامل‌ها و

نمایش می‌دهد. در تقاطع‌های غیر هم‌سطح برای بخشی از مسیری که زیر مسیر دیگری واقع شده است هیچ داده‌ای جمع‌آوری نخواهد شد و با نقص داده مواجه خواهیم بود. عامل استخراج‌کننده بر روی داده‌های خط سیر حرکت می‌کند که به محض 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.



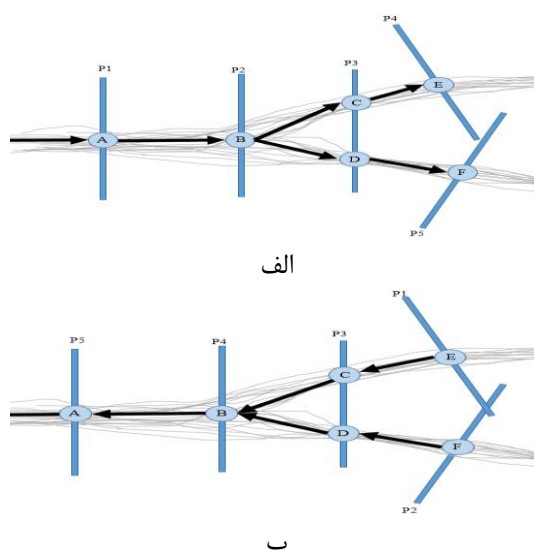
شکل ۱۱: استخراج راه در تقاطع غیر هم‌سطح

#### 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 PostgreSQL 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<sup>۱</sup> توسط دانشگاه وایکاتو<sup>۲</sup> منتشر شده است که دارای مجموعه‌ای از الگوریتم‌هایی در زمینه یادگیری ماشین<sup>۳</sup>

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  $\Delta n_{DA}$  یا هر دو مثبت و یا هر دو منفی خواهند بود. در صورتی که مقادیر  $\Delta n_{CB}$  و  $\Delta n_{DA}$  منفی باشند، جدادگی مسیر (شکل (۱۰-الف)) و در صورت مثبت بودن این مقادیر، یکی شدگی مسیر (شکل (۱۰-ب)) را خواهیم داشت. در صورت شناسایی جدادگی مسیر، عامل استخراج‌کننده راه، مسیری با تعداد خط سیرهای بیشتر را ادامه داده و عامل مدیر فعال می‌کند و در صورت شناسایی یکی شدگی مسیر عامل استخراج‌کننده راه گذرنده از خوشه‌ای با تعداد خط سیرهای کمتر، غیرفعال شد و عامل استخراج‌کننده راه گذرنده از خوشه‌ای با تعداد خط سیرهای بیشتر ادامه مسیر را پی می‌گیرد



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

از آنجایی که 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

<sup>1</sup> General Public Licence

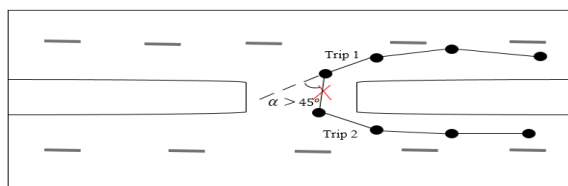
<sup>2</sup> Waikato

<sup>3</sup> Machine Learning

بر این، وجود داده‌های اضافی<sup>۳</sup> می‌تواند باعث اعمال محاسبات اضافی به سیستم شود. نمونه‌ای از داده‌های اضافی داده‌هایی است که کاربران هنگامی که در ترافیک قرار گرفته‌اند، جمع‌آوری می‌کنند.

داده‌های خط سیر زمانی می‌توانند بیانگر راه باشند که توسط خودروهای در حال حرکت جمع‌آوری شده باشند [۲۶]. همچنین از آنجایی که داده‌های جمع‌آوری شده تشکیل‌دهنده یک خط سیر، بر اساس توالی زمانی به یکدیگر متصل می‌شوند، لذا در برخی از موارد یک اتصال غیرمنطقی بین سفرهای<sup>۴</sup> مختلف وجود دارد. بنابراین به منظور استفاده از این داده‌ها باید این سفرها از یکدیگر جدا شوند [۳۱]. از آنجایی که هدف ارائه راهکاری با قابلیت کار با داده‌های با حجم بالا است، در ابتدا پالایه‌ای، با در نظر گرفتن قوانین زیر به منظور 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 دو نقطه بیشتر از ۳۰۰ متر و یا اینکه تغییر جهت<sup>۵</sup> آن‌ها بیشتر از ۴۵ درجه باشد، داده‌های خط سیر دو سفر مجزا در نظر گرفته می‌شوند [۳۱]. (شکل ۱۲) این حالت را به خوبی 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

<sup>3</sup> Redundant data

<sup>۴</sup> Trip: هر یک از فعالیت‌های داوطلب (کاری، ورزشی، تفریحی و ...) یک سفر نامیده می‌شود.

<sup>5</sup> Direction

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 به منظور 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<sup>۱</sup> توسط ۱۸۲ کاربر در مدت سه سال (آوریل ۲۰۰۷ تا اوت ۲۰۱۲) جمع‌آوری شده 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 با فاصله 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<sup>۲</sup> به‌عنوان سیستم تصویر استفاده می‌کند. تبدیل داده‌های خط‌سیر از سیستم مختصات سه‌بعدی WGS84 به سیستم تصویر سه‌بعدی 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 سه‌بعدی، می‌توانند به‌عنوان ورودی روش ارائه شده در این مقاله در نظر گرفته شوند.

نویز ذاتی GPS در دقت و صحت داده‌های جمع‌آوری شده تأثیرگذار است. در شرایط خاصی این نویز می‌تواند منجر به ایجاد راه‌های جعلی گردد. همچنین این داده‌ها دارای اشتباه نیز می‌باشند. علاوه

<sup>1</sup> Data set

<sup>2</sup> Web Mercator

خط سیر جمع‌آوری شده 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 MISSING، SPURIOUS و F\_score دقت نقشه تولید شده نسبت به نقشه مبنا تعیین می‌شود. پارامتر MISSING به‌صورت نسبتی از تعداد راه‌های موجود در نقشه مبنا که در نقشه تولید شده وجود ندارند به کل راه‌های موجود در نقشه مبنا و پارامتر SPURIOUS 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 MISSING و SPURIOUS 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.

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<sup>۱</sup> 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 ارائه شده در این پژوهش ارائه شده است.

#### ۴-۲- نمایش نقشه راه

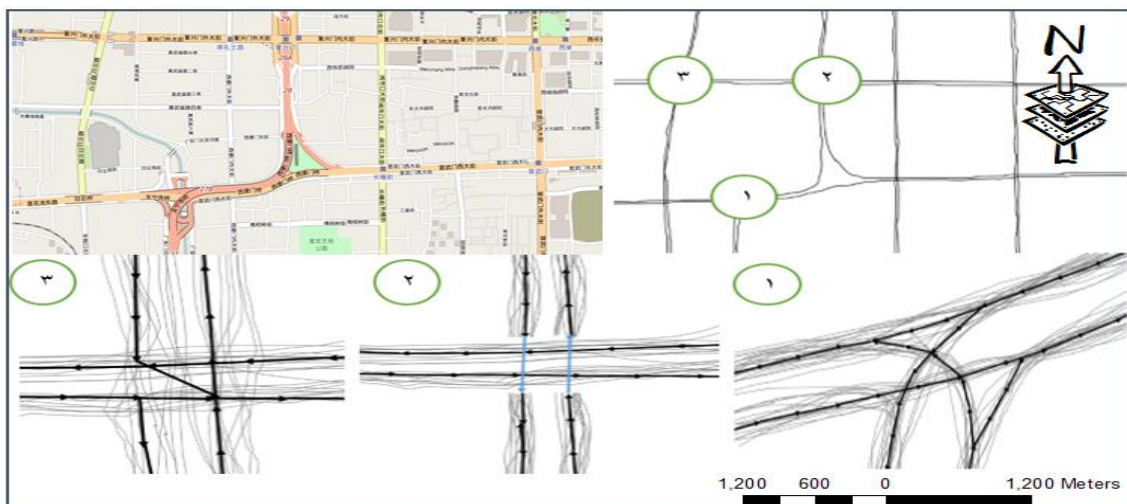
شکل (۱۳) نقشه شبکه راه برای منطقه مورد مطالعه را نمایش می‌دهد. در این شکل، بخش مشخص شده ۱ راه استخراج شده در جدانشدگی و یکی شدگی را نمایش می‌دهد. از آنجایی که هر مسیر توسط یک عامل استخراج می‌شود، لذا استفاده از راه‌کار ارائه شده در استخراج راه در جدانشدگی و یکی شدگی به‌خوبی صورت گرفته و از اتصال گره‌های غیر مرتبط جلوگیری می‌شود. بخش مشخص شده ۲ در شکل (۱۳) یک تقاطع غیر هم‌سطح را نمایش می‌دهد. از آنجایی که در راه‌کار ارائه شده 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 روی داده‌های خط سیر به‌منظور استخراج راه امکان پذیر نخواهد بود. لذا در این نوع ساختارها استخراج راه نیازمند پس پردازشی به‌منظور استخراج راه است. بخش مشخص شده ۳ در شکل (۱۳) یک چهارراه را نمایش می‌دهد. حجم بالای داده‌های

<sup>1</sup> Outlier

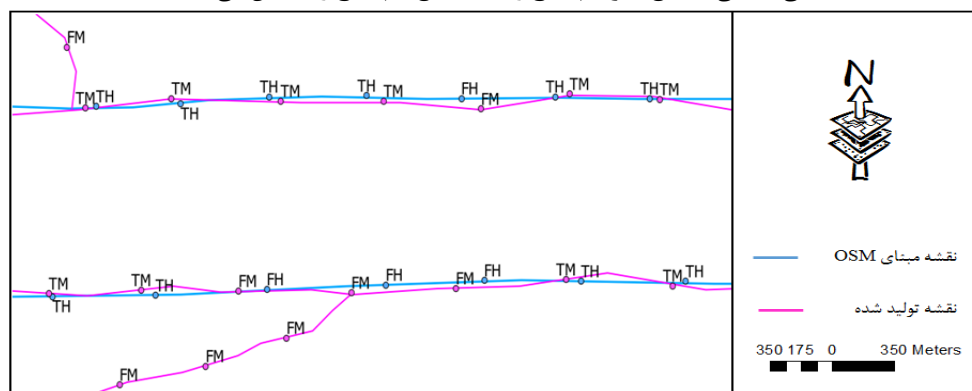
کشف شود، این نقطه به عنوان TH و در غیر این صورت این نقطه به عنوان FH در نظر گرفته خواهد شد (شکل ۱۴)).

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

به منظور تعیین پارامترهای SPURIOUS و MISSING، راه‌های موجود در نقشه مبنا و نقشه تولید شده به پاره‌خطهایی مساوی تقسیم و سپس مراکز هر کدام از این پاره‌خطها تعیین شد. مراکز پاره خطها در نقشه مبنا به عنوان نقاط "H" و مراکز پاره‌خطها در نقشه تولید شده به عنوان نقاط "M" در نظر گرفته شد. به منظور تعیین پارامتر MISSING، برای هر یک از نقاط H وجود و یا عدم وجود نقاط M در شعاع مشخص بررسی خواهد شد. در صورتی که در این شعاع مشخص شده حداقل یک نقطه M



شکل ۱۳: نمایش نقشه شبکه راه برای منطقه مورد مطالعه. ساختارهای مشخص شده 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.

$$spurious = \frac{\# FM}{\# FM + \# TM}$$

رابطه (۸)

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

مردم گستر 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 در زمانی صورت می‌گیرد که تغییر متمایزکننده‌ای در فضای داده‌های خط سیر صورت گیرد. از آنجایی که 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 واقع شده است و در نتیجه 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 روی داده‌های 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

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

پس از تعیین پارامترهای SPURIOUS و MISSING، دقت نقشه تولیدشده نسبت به نقشه مبنا با استفاده از مقدار  $F\_score$  به‌صورت رابطه (۱۰) قابل بیان است.

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

پس از اجرای روش ارزیابی برای داده‌های خط سیر موجود مربوط به ۵۶ راه، با در نظر گرفتن مقادیر ۹ و ۱۵ به ترتیب 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 [۳۳، ۱۰] 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 [۳۳، ۱۰] 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 قابل توجه 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

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.

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

## 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, *Artificial 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.



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

Reza Mohammadi<sup>1</sup>, Mehdi Farnaghi<sup>\*2</sup>

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