نشریه علمی پژوهشی مهندسی فناوری اطلاعات مکانی

سال پنجم، شماره سوم، پاییز ۱۳۹۶
Vol.5, No.3, Autumn 2017
۵۱-۷۶

تعیین راهبندان‌های ترافیکی بر مبنای خطوط سیر حاصل از اطلاعات مکانی داوطلبانه

وحید شکری [۱]، رحیم علی‌عباسپور [۲] *

۱- دانش‌آموخته کارشناسی ارشد سیستم اطلاعات مکانی، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی، دانشگاه تهران
۲- استادیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی، دانشگاه تهران

تاریخ دریافت مقاله: ۱۳۹۵/۰۵/۰۱ تاریخ پذیرش مقاله: ۱۳۹۵/۱۲/۱۵

چکیده

برخلاف مزایای حسگرهای ثابت در گردآوری داده‌های ترافیکی مانند کیفیت بالا، عواملی مانند وابستگی به تجهیزات، عدم پوشش کامل، هزینه و زمان بالای گردآوری، مدیریت و تحلیل داده‌ها نیاز به توسعه زیرساختی مناسب را آشکار می‌سازد. افزایش رو به رشد اطلاعات مکانی داوطلبانه و توسعه فن‌آوری‌های مکان‌آگاه ابعاد جدیدی برای جمع‌آوری و نشر داده‌های ترافیکی با حجم بالا و هزینه پایین فراهم ساخته است. لذا توسعه سامانه‌ای به‌منظور تسهیل در امر جمع‌آوری، استخراج و نشر اطلاعات ترافیکی با بهره‌گیری از اطلاعات مکانی داوطلبانه مدنظر قرارگرفت، اما با توجه به گستردگی موضوع، در این مقاله تمرکز بر خطوط سیر و ارائه راهکاری جهت استخراج راهبندان‌های ترافیکی است. در ابتدا داده‌های خطوط سیر در مرحله پیش‌پردازش پاک‌سازی و داده‌های نامربوط حذف می‌شوند، سپس با استفاده از الگوریتم تطبیق نقشه مکانی-زمانی بر شبکه معابر منطبق می‌گردند. پس‌ازآن، سرعت متوسط بین هر دونقطه متوالی محاسبه و با مقایسه سرعت جریان آزاد و سرعت شبکه معابر، راهبندان‌های ترافیکی تعیین می‌گردد. به‌منظور ارزیابی سامانه از داده‌های خطوط سیر شهر پکن در بازه زمانی ۲ تا ۸ فوریه سال ۲۰۰۸ میلادی که داوطلبانه توسط تاکسی‌ها جمع‌آوری‌شده استفاده‌گردید. برای نمونه نتایج راهبندان‌های ترافیکی در بازه زمانی ۱۶ تا۱۶:۲۰ اکثراً در معابر اصلی و لینک‌های اتصال مشاهده می‌شود و این امر به علت انتقال جریان ترافیک از مرکز شهر و محل‌های کار به مناطق مسکونی است. بنابراین می‌توان نتیجه گرفت که با استفاده از مشارکت مناسب کاربران در قالب اطلاعات مکانی داوطلبانه با نرخ نفوذ پایین (۳٫۸٪) ترافیک یک شهر بزرگ در حدود ۷۲۷۲۰ قطعه‌راه را با صرف هزینه پایین و بدون نیاز به تجهیزات پیچیده می‌توان تعیین نمود.

کلیدواژه‌ها: اطلاعات ترافیکی، اطلاعات مکانی داوطلبانه *VGI*، خطوط سیر *GPS*، راهبندان‌های ترافیکی.

* نویسنده مکاتبه کننده: گروه *GIS*، دانشکده مهندسی نقشه¬برداری و اطلاعات مکانی ، پردیس دانشکده¬های فنی، دانشگاه تهران تهران، امیرآباد شمالی
تلفن : ۰۹۱۲۳۰۱۲۲۵۵

Email: abaspour@ut.ac.ir

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]

۱- مقدمه

راهبندان‌های ترافیکی[1] در بسیاری از شهرهای بزرگ در سراسر دنیا به‌عنوان مشکلی اساسی به‌شمار می‌رود که میلیون‌ها نفر روزانه با آن درگیرند. افزایش تراکم وسایط نقلیه در شهرهای بزرگ در سال‌های اخیر از یک سو و محدودیت در ایجاد زیرساخت‌های جدید از طرف دیگر، لزوم استفاده از فناوری‌های نوین را آشکار ساخته است. راهبندان‌های ترافیکی پیامدهای متعددی دارد که از میان آنها هزینه‌های گزاف، افزایش زمان سفر و آلودگی هوا به‌عنوان مهمترین آنها به‌شمار می‌آید. دولت‌ها هرساله هزینه‌های هنگفتی را برای تعیین راهبندان‌های ترافیکی در راستای حل این معضل می‌پردازند، اما همچنان به دلیل ماهیت پیچیده راهبندان‌های ترافیکی، مدیریت و حل این مسأله با مشکل مواجه است. یکی از این پیچیدگی‌ها، عدم امکان پیش‌بینی آن در موارد متعدد و ماهیت پویا و وابسته آن است. راهبندان‌های ترافیکی از یک خیابان تا خیابان دیگر توزیع متفاوتی دارد؛ بنابراین طراحی سیستم کاملاً خودکار تعیین مناطق راهبندان‌های ترافیکی مشکل بوده و نیازمند داده‌های مناسب، تجربه و شناخت قوی است[۱].

سیستم‌های گردآوری ضمنی اطلاعات ترافیکی بخشی از سیستم حمل و نقل هوشمند به شمار می‌روند. اخیراً سیستم‌های حمل و نقل هوشمند به‌طور گسترده به‌منظور کاهش مشکلات ترافیکی مورد استفاده قرار گرفته‌اند. مسأله بسیار مهم در سیستم حمل و نقل هوشمند نیاز به جمع‌آوری داده‌های ترافیکی باکیفیت بالا و بی درنگ می‌باشد[۲]. داده‌های بی درنگ ترافیکی عاملی مهم و ضروری در سیستم‌های حمل‌ونقل هوشمند از جمله کنترل ترافیک، هدایت خودروها برای مسیریابی، شناسایی تصادفات و رویدادهای ناگهانی می‌باشد[۳].

زیرساخت‌های نظارت بر ترافیک به‌طور عمده از تجهیزاتی مانند حلقه‌های شناساگر، دوربین و رادارهای ماکروویو تشکیل می‌شود. هزینه بالای نصب، راه‌اندازی و نگهداری این تجهیزات از استقرار آن‌ها در کل شبکه شریانی شهری جلوگیری می‌کند. علاوه بر این، شناساگرهای حلقه القایی مستعد خطا هستند و همیشه به‌درستی عمل نمی‌کنند (به‌عنوان مثال، روزانه در کالیفرنیا، ۳۰٪ از ۲۵۰۰۰ آشکارسازها به‌درستی کار نمی‌کند)[۴]. بنابراین وابستگی به تجهیزات، عدم پوشش همه مناطق شهر، هزینه بالای جمع‌آوری مدیریت و آنالیز حجم وسیع داده‌ها، زمان‌بر بودن این فرآیند و ماهیت دینامیک اطلاعات ترافیکی موجب می‌شود تا توسعه زیرساختی مناسب برای جمع‌آوری اطلاعات ترافیکی با بهره‌گیری از دانش و فناوری‌های روز ضرورت یابد. به همین دلیل، جامعه مهندسی حمل‌ونقل همواره به دنبال راه‌هایی جدید برای جمع‌آوری داده‌های ترافیک برای نظارت بر ترافیک بوده است.

امروزه محبوبیت روبه‌رشد دستگاه‌های مبتنی بر وب ۲ امکان تولید و انتشار مجموعه عظیمی از اطلاعات در دسترس عموم، حاوی المان‌های مکانی را فراهم می‌کند. میزان داده‌ها با توجه به گسترش کاربران اینترنت دائماً در حال افزایش است و توجه پژوهشگران زیادی را به خود جلب کرده است [۵]. همچنین فناوری‌های تعیین موقعیت و ارتباطات بی‌سیم منجر به افزایش اطلاعات مکانی در اینترنت شده‌اند [۶]. با توجه به ویژگی‌های اطلاعات مکانی داوطلبانه[2] ازجمله صحت زمانی بالا، هزینه پایین، قابلیت تولید با سطوح دانش و علایق مختلف، به‌کارگیری آن‌ها به‌عنوان ابزاری مناسب جهت گردآوری، تکمیل و به‌روزرسانی داده‌های پایه‌ای ترافیکی از طریق ردیابی تغییرات محلی به‌موقع توسط

[1] Traffic jam

[2] Volunteered geographic information (VGI)

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]

حجم زیادی از مشاهده‌کنندگان (کاربران) مؤثر به نظر می‌رسد.

تاکنون اقدامات زیادی در این زمینه انجام‌شده است. برای مثال، از سال ۲۰۰۷، شرکت گوگل اقدام به یکپارچه‌سازی اطلاعات بی‌درنگ ترافیکی با سرویس نقشه خود نمود. داده‌های ترافیکی از منابع مختلفی چون حسگرهای جاده‌ای، خودروها، ناوگان تاکسیرانی و اخیراً کاربران موبایل جمع‌آوری می‌شود. لذا در این تحقیق، به بررسی برخی توانایی‌های سیستم مبتنی بر اطلاعات مکانی داوطلبانه که با استفاده از فناوری وب ۲ در محیط موبایل توسعه داده‌شده و امکان جمع‌آوری اطلاعات در فرمتهای مختلف (کمی، کیفی، مکانی، چندرسانه‌ای متن، عکس، خطوط سیر *GPS*)، توسط افراد با سطح دانش مختلف به شیوه‌های مختلف را دارد، پرداخته می‌شود. سامانه توسعه داده‌شده قابلیت‌های زیادی ازجمله تعیین پارامترهای ترافیکی مانند ماتریس مبدأ-مقصد، زمان سفر، سرعت و امکان انتشار اطلاعات ترافیکی، نقشه‌ها و رویدادهای مختلف رخداده در شریان‌های شهری مانند تصادف و خراب‌شدگی چراغ راهنمایی را دارد. اما، در این مقاله تمرکز بر خطوط سیر *GPS* جمع‌آوری‌شده به‌منظور تعیین پارامتر راهبندان‌های ترافیکی می‌باشد. در این راستا در بخش دوم مروری بر مطالعات صورت گرفته در این زمینه بیان می‌شود. در بخش سوم روش تحقیق که شامل مبانی نظری اطلاعات مکانی داوطلبانه، روش مورداستفاده به‌منظور پاک‌سازی داده‌های خطوط سیر *GPS* خام، الگوریتم مکانی-زمانی تطبیق نقشه به‌کاررفته شده با هدف تطبیق خطوط سیر *GPS* با شبکه معابر، روش محاسبه سرعت شبکه و سرعت جریان آزاد[1] و درنهایت تعیین راهبندان‌های ترافیکی تشریح می‌شود. در بخش چهارم به بررسی نتایج حاصل از پیاده‌سازی روش‌های بیان‌شده در قسمت سوم برای داده‌های شهر پکن پرداخته و در بخش پنجم بحث و نتیجه‌گیری در مورد نتایج حاصل از پیاده‌سازی ارائه می‌گردد.

۲- پیشینه تحقیق

رایج‌ترین روش تعیین راهبندان های ترافیکی به‌واسطه حسگرها مانند رادار بوده است[۷]. تحلیل جریان‌های ویدئویی دوربین‌های نصب‌شده در کنار جاده نیز به درک راهبندان‌های ترافیکی کمک می‌کند[۸ و ۹]. با اینحال، هر دوی این روش‌ها نیازمند نصب و نگهداری تجهیزات با هزینه بسیار بالا هستند و تنها در موقعیت‌های ثابتی از شبکه معابر اطلاعات ترافیکی را فراهم می‌کنند.

در حوزه شناسایی راهبندان‌های ترافیکی که یکی از پارامترهای ترافیکی مهم به شمار می‌رود، داده‌های خطوط سیر معمولاً به‌واسطه زمان سفرهای طولانی و یا به‌طور مشابه سرعت‌پایین تعیین می‌شود. بسیاری از الگوریتم‌های تشخیص راهبندان‌های ترافیکی بر مبنای محاسبه سرعت شبکه[۱۰] و یا تشخیص خوشه خودروهای با سرعت‌پایین کار می‌کنند[۱۱ و ۱۲]. برای مثال لیو و همکاران[۱۱] به مطالعه تعامل پارامترها در راهبندان‌های ترافیکی پرداختند. آن‌ها ابتدا یک شهر را به محدوده‌هایی تقسیم کردند و سپس جریان ترافیک لینک‌های بین مناطق را تعیین کردند. راهبندان های ترافیکی به‌صورت درختی در لینکهای بین مناطق شناسایی شدند اما باید توجه داشت که راهبندان های ترافیکی معمولاً در در لینک‌های بین مناطق اتفاق نمی‌افتد بلکه در جاده‌ها رخ می‌دهد.

تاکنون سیستم‌های زیادی به‌منظور پایش ترافیک و فراهم آوردن داده‌های توزیع راهبندان‌های ترافیکی به‌واسطه خطوط سیر *GPS* توسعه‌یافته‌اند، این سیستم‌ها اغلب داده‌ها را از طریق حسگرهای ثابت با هزینه بالا و تنها در نقاط محدودی در اختیار قرار می‌دهند [۷, ۱۳]. به‌علاوه، در تحقیقات اخیر در زمینه تعیین راهبندان‌های ترافیکی با استفاده از داده‌های خطوط سیر *GPS* تنها راهبندان‌های ترافیکی

[1] Free-flow speed

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]

به‌صورت جداگانه تعیین شده‌اند[۱۰ و ۱۴] و ترافیک‌های سنگین به‌صورت پراکنده ارائه شده و امکان تعیین میزان راهبندان‌های ترافیکی یک خیابان خاص وجود ندارد[۱۱ و ۱۵].

ژانگ و همکاران [۱۶] مدلی برای پایش جریان ترافیک، ارزیابی راهبندان‌های ترافیکی بر اساس شبکه‌های حسگر بی‌سیم طراحی نمودند. ایده اصلی این مدل بررسی شرایط راهبندان‌های ترافیکی و تاثیر آن در جریان ترافیک آینده و ترکیب معادلات ترافیک با توابع بهینه‌سازی بود. این امر از طریق روش حل عددی معادلات ترافیک با روش‌های تقریبی و برازش و همبستگی بین داده‌های ترافیکی خوانده شده از حسگر تعریف می‌گردند. نتیجه نشان دهنده عملکرد بهتر و کاهش متوسط تاخیر و حداکثر طول صف است. از معایب این تحقیق محدودیت آن به قطعات ساده با جریان ترافیک پیوسته می‌باشد.

گیانوتی و همکاران [۱۷] به تشخیص راهبندان‌های ترافیکی به وسیله تعیین گروه‌های اتومبیل نزدیک به هم که همه به آرامی در حال حرکت هستند، پرداخته است. این هدف با ارائه یک چارچوب برای تحلیل آثار تحرک خطوط سیر *GPS* با قابلیت‌های پرس‌وجو و ترسیم سطوح فعالیت در طول هفته، طول سفر و توزیع زمان سفر، خوشه بندی مسیرهای به‌/از یک نقطه خاص، و ارتباط بین مناطق مختلف صورت گرفته است. نویسندگان همچنین یک روش برای تشخیص رویدادهای خاص (مانند مسابقات ورزش)، همراه با سفرهای مربوط به این رویداد ارائه نمودند. در واقعف سازوکاری برای انتقال داده‌های خام خطوط سیر به دانش تحرک انسانها با استفاده از روش‌های داده کاوی و آنتولوژی پرداختند.

ناگاراج و همکاران [۱۸] یک سیستم خودتطبیق‌گر با به‌کارگیری روش پردازش تصویر به منظور تشخیص راهبندان‌های ترافیکی توسعه دادند. سیستم ردیابی پیشنهادی امکان تشخیص وسایل نقلیه در شرایط ترافیک سنگین، سایه و شرایط نوری تاثیرگذار در تصاویر ویدئویی که شرایط چالش برانگیزی هستند را دارد. محاسبه تراکم ترافیک با استفاده از عملیات پردازش تصویر بر روی تصاویر دوربین در مقایسه با ابزارهای دیگری ازجمله حسگرها مقرون به صرفه‌تر است. سیستم مورد نظر با استفاده از اطلاعات به دست آمده کاربر قادر خواهد بود تراکم ترافیک را در هر محل انتخابی دریافت نماید.

پنگ و همکاران[۱۹] جریان ترافیک شهری را با ترکیب خطی سه نوع سفر بین محل سکونت و محل کار، مناطق، سفرهای کار و سفرهای تفریحی تقریب زدند. آنها با ارائه روشی پیشنهادی و با بهره‌گیری از روش فاکتور ماتریس غیرمنفی و روش‌های بهینه‌سازی به تعیین این سه ضریب و در نتیجه تولید یک برآورد تقریبی از جریان ترافیک پرداختند.

پن و همکاران [۲۰] رویکردی با استفاده از ترکیب دو نوع داده جمع گستر[1] و داده های حرکت حاصل از شبکه‌های اجتماعی به‌منظور درک یکی از جنبه‌های پویایی شهری ارائه نمودند و به طور خاص، به تشخیص و توصیف ناهنجاریهای ترافیکی با استفاده از یک روش جدید بر اساس رفتار مسیریابی رانندگان پرداختند. در این تحقیق ناهنجاری شناسایی شده توسط یک زیر گراف شبکه جاده‌ای که در آن ها رفتارهای مسیریابی مردم به طور قابل توجهی از الگوهای اصلی خود متفاوت هستند، صورت می‌گیرد. پس از آن، استخراج ناهنجاری به واسطه توصیف شرایط ناهنجاری‌های اعلام شده توسط مردم از رسانه‌های اجتماعی (جمع گستر) انجام می‌شود. رویکرد ارائه شده امکان کشف گراف شبکه جاده مرتبط با ناهنجاریهای ترافیکی را فراهم می‌آورد.

پنگ و همکاران[۲۱] با استفاده از داده‌های *GPS* جمع‌آوری شده توسط تاکسی‌ها به نظارت بر ظهور رفتار غیرمنتظره، که پتانسیل برآورد و بهبود شرایط

[1] Crowd sourcing

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]

ترافیک پیش‌رو را دارد، پرداختند. در این تحقیق از انطباق آزمون آماری نسبت احتمال *(LRT)* برای توصیف الگوهای ترافیک و تشخیص دقیق و سریع رفتار غیر عادی استفاده گردید. همچنین یک روش استخراج الگوی کارآمد برای تهیه داده‌های مکانی-زمانی ترافیک با توانایی تشخیص «اشتباهات مداوم» و «اشتباهات در حال ظهور» ارائه شد. نتیجه نشان می‌دهد عملکرد زمان محاسباتی برای مجموعه داده‌های بزرگ تا حد زیادی بهبود یافت.

دیسای و همکاران [۲۲]به طراحی رویکرد چند عاملی با هدف اجتناب از راهبندان‌های ترافیکی و تخصیص مسیر پرداختند. در این رویکرد از عامل مجازی مذاکره و عاملهای خودرو که به منظور تصمیم گیری و تعیین مسیر در ارتباط محلی با یکدیگر هستند، استفاده شده است. تصمیم گیری مشارکتی تخصیص مسیر در تقاطع‌ها و اتصالات انجام می‌شود. ارتباطات درون خودرو با استفاده از اطلاعات انتشار ترافیک و انجام پردازش توزیع آن صورت می‌گیرد. هر عامل خودرو محاسبه مستقل ترجیح مسیر خود را در تخصیص اولیه مسیرها مبادله می‌کند. عملیات تخصیص با بهره‌گیری از تعداد زیادی عامل مجازی بهبود می‌یابد.

آندریا و مارسلونی[۲۳] یک سیستم خبره برای تشخیص ترافیک و حوادث با بهره‌گیری از داده‌های *GPS* ای که توسط گوشی های هوشمند رانندگان جمع آوری شده است، پرداختند. بسط انجام شده توسط سیستم خبره مستقل از زمینه (شهری و غیر شهری) است و ممکن است به‌طور مستقیم در شبکه های مختلف جاده تقریبا بدون تغییر پارامترهای سیستم و بدون نیاز به فرآیند یادگیری و یا استفاده از داده‌های تاریخی به‌کار گرفته شود. تجزیه و تحلیل تجربی با استفاده از ترکیبی از داده های *GPS* شبیه‌سازی شده و داده های *GPS*واقعی از شهر پیزا انجام شد. این سیستم بر مبنای معماری سرویس گرا و اختصاص سرعت به هر بخش جاده توسعه داده شده است.

پراکاش [۲۴] سیستمی به منظور تحلیل داده‌های *GPS* و شناسایی و برآورد راهبندان‌های ترافیکی توسعه داد. سیستم مذکور داده‌های *GPS* حاصل از تلفن‌های همراه و تبلت‌ها را ساختاردهی و به‌کارمی‌گیرد. در روش مورد استفاده، داده های *GPS* برای اولین بار با استفاده از الگوریتم خوشه‌بندی *K-means* طبقه‌بندی و سپس خوشه‌های به دست آمده فیلتر می‌شوند. در پردازشهای بعدی این خوشه‌ها با استفاده از روش استخراج احتمالی بیز ساده برای تشخیص ترافیک و پیش بینی آن مورد استفاده قرار می‌گیرد.

روش مورداستفاده در این مقاله بر اساس محاسبه سرعت از داده‌های خطوط سیر *GPS*، راهبندان‌های ترافیکی را مشخص می‌کند. این روش داده‌های شرایط ترافیکی را نه‌تنها در راهبندان های ترافیکی به‌صورت جداگانه و در محدوده زمانی کم، بلکه در مناطق وسیع‌تر بازه‌های زمانی بیشتر فراهم می‌آورد. تمایز این روش با روش‌های گذشته در بهره‌گیری از داده‌های مکانی داوطلبانه بدون صرف هزینه و در زمان کوتاه، به‌منظور برآورد مقدار سرعت جریان آزاد از داده‌های واقعی مربوط به هر قطعه از راه است. هم‌چنین در این روش، از برآورد سرعت جریان آزاد در چند منطقه محدود و تعمیم آن به‌کل شبکه راه که ناخودآگاه خطایی را به همراه دارد، خودداری می‌شود. نکته قابل‌توجه دیگر آن است که در این روش به‌طور خودکار و بدون نیاز به دانش محلی اکثراً از خطوط سیری استفاده خواهد شد که در ساعات غیر اوج ترافیکی واقع‌شده‌اند و این امر به دلیل آن است که در الگوریتم مذکور از سریع‌ترین خطوط سیر استفاده می‌شود که این خطوط سیر در ساعات غیر اوج رخ می‌دهد. به علاوه صحت زمانی بالا، هزینه پایین گردآوری، تکمیل و به‌روزرسانی داده‌های پایه‌ای ترافیکی از طریق ردیابی تغییرات محلی به‌موقع توسط حجم زیادی از مشاهده‌کنندگان (کاربران)

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]

همگی از نوآوری‌های این مقاله محسوب می‌شوند.

۳- روش پیشنهادی

در این بخش به توضیح روش پیشنهادی پرداخته می‌شود.

۳-۱-اطلاعات مکانی داوطلبانه

محبوبیت رو به رشد سیستم‌های مبتنی بر وب ۲ مجموعه عظیمی از اطلاعات در دسترس عموم که حاوی اطلاعات مکانی شفاف و اطلاعات مکانی ضمنی است را تولید می‌کند. میزان داده‌ها با توجه به گسترش کاربران اینترنت دائماً در حال افزایش است. برای مثال فناوری‌های تعیین موقعیت و ارتباطات بی‌سیم منجر به افزایش اطلاعات برچسب‌گذاری مکانی در اینترنت شده‌اند [۲۵]. اصطلاح اطلاعات مکانی داوطلبانه که برای اولین بار در سال ۲۰۰۷ توسط گودچایلد معرفی شد [۵] به اطلاعاتی اطلاق می‌شود که توسط افراد مختلف با سطوح دانش متفاوت به‌صورت داوطلبانه و بصورت مکان مبنا تولید می‌شوند. گودچایلد، اطلاعات مکانی داوطلبانه را به‌عنوان بستری توانمند شده به قابلیت‌های وب۲ برای استفاده از ۶ میلیارد سنجنده انسانی با ویژگی‌های خاصی در مقایسه با سنجنده‌های فیزیکی برای جمع‌آوری داده معرفی می‌کند. سنجنده‌های انسانی حواس مختلفی دارند و به این وسیله می‌توانند پارامترهای مختلفی را مشاهده کنند، آن‌ها دانش محلی دارند و به محیط اطرافشان آگاه هستند [۲۶].

منابع تولید داده‌های مکانی به روش سنتی، داده‌هایی باکیفیت بالا که بر اساس روشها و استانداردهای معین تولید و پردازش می‌شوند را فراهم می‌کنند [۲۷]. در مقابل، کیفیت اطلاعات مکانی داوطلبانه اغلب غیرشفاف و نامعلوم است، چراکه توسط کاربران غیرحرفه‌ای با سطح توانایی و علایق مختلف تولید می‌شوند. در حالت ایده‌آل، سازمان‌ها و متخصصین به داده‌های باکیفیت و معتبر (داده‌های سنتی) نیاز دارند، اما با توجه به هزینه و زمان و مسائل دیگری که در جمع‌آوری و تولید این نوع از داده‌های مکانی وجود دارد، نیازی برای تولید یک جایگزین به وجود آمد. به دلیل دسترس‌پذیری بالای اطلاعات مکانی داوطلبانه و ویژگی‌هایی چون فراوانی و به‌روز بودن، این داده‌ها به‌عنوان راه‌حلی برای جایگزینی و یا تکمیل داده‌های سنتی ارائه شدند. از طرفی *VGI* مشکلاتی ازجمله کیفیت داده، چگالی داده، ناسازگاری و توزیع غیریکنواخت را به دنبال دارد.

۳-۲- پاک‌سازی داده‌های خطوط سیر GPS

به دلیل آنکه داده‌های خام خطوط سیر *GPS* به‌دست‌آمده از مشارکت مردمی، خطادار هستند، باید اقداماتی به‌منظور کاهش خطاها در مرحله پیش‌پردازش روی این داده‌ها صورت گیرد. در این مرحله که پاک‌سازی داده‌ها نامیده می‌شود، چهار نوع از داده‌ها شامل داده‌های بی‌ربط، داده‌های نادرست، داده‌ها با نرخ نمونه‌برداری پایین، داده توقف غیرراهبندان‌های ترافیکی (پارک) و داده مسیرهای بسیار کوچک حذف می‌شوند. برای نیل به این هدف از مجموعه‌ای از فیلترها استفاده می‌شود. اطلاعات خارج از محدوده مکانی شبکه معابر بی‌ربط است و توسط فیلتر یک حذف می‌شود. مشکلات مربوط به زمان یا موقعیت مکانی در داده‌های تکراری که معمولاً به‌عنوان دو رکورد با برچسب زمان یکسان هستند و پارک توسط فیلترهای دو و سه حذف خواهند شد. قطعات مسیر با طول کوتاه و تعداد نقاط کم توسط فیلتر چهار حذف می‌شوند و فیلتر پنج نقاط خطوط سیر با نرخ پایین را حذف می‌نماید.

- **فیلتر یک:** حذف مختصات غیرواقعی نقاط نمونه‌برداری یعنی نقاط واقع در خارج از محدوده شهر پکن به مختصات (۱۱۶/۶۷۳و۱۱۶/۱۰۹) درجه شمالی و (۴۰/۱۴و۳۹/۷۰) درجه شرقی. این نقاط خود شامل دودسته‌اند ۱- نقاطی که به‌طور صحیح توسط *GPS* تعیین موقعیت شده،

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]

اما در محدوده جغرافیایی شهر پکن و در محدوده موردمطالعه این تحقیق نبودند. ۲- نقاطی که دارای مختصات (۰۰/۰۰و۰۰/۰۰) ناشی از خطای *GPS* بودند.

- **فیلتر دو:** حذف نقاطی که دارای شناسه خودرو یکسان، زمان ثبت یکسان و موقعیت یکسان (نقاط تکراری) بودند، به‌نحوی‌که تنها اولین نقطه حفظ و نقاط دیگر مسیر با همان برچسب زمانی حذف می‌شوند.
- **فیلتر سه:** چنین فرض می‌شود که قطعاتی از خطوط سیر *GPS* که در آنها خودرویی طول ۱۰۰ متر را در زمان سی دقیقه طی کند، در پارکینگ اقامت دارد و در نتیجه این نقاط حذف می‌گردند. برای این منظور از الگوریتم خوشه‌بندی مبتنی بر چگالی استفاده گردید. از میان روشهای مبتنی بر چگالی، در این تحقیق الگوریتم خوشه‌بندی مبتنی بر چگالی مکانی-زمانی، الگوریتم *ST-DBSCAN* در سه جهت اساسی بهبودیافته است؛ اول، برخلاف الگوریتم‌های خوشه‌بندی مبتنی بر تراکم موجود، الگوریتم موردنظر می‌تواند داده‌های مکانی-زمانی را با توجه به ویژگی‌های مکانی، غیر مکانی مانند زمان خوشه‌بندی کند. دوم اینکه، الگوریتم *DBSCAN* نمی‌تواند برخی از نقاط نویز را در حالتی که خوشه‌ها با تراکم‌های مختلف وجود داشته باشد، تشخیص دهد. الگوریتم مدنظر این مشکل را با اختصاص فاکتور تراکم به هر خوشه حل می‌کند. سوم، اگر مقادیر غیر مکانی اشیاء همسایه دارای تفاوت‌های کوچک و خوشه‌ها مجاور یکدیگر قرار داشته باشند، مقادیر مرز اشیاء در یک خوشه ممکن است بسیار متفاوت از مقادیر مرز اشیاء در طرف مقابل خوشه باشد. الگوریتم مدنظر این مشکل را با مقایسه مقدار متوسط یک خوشه با مقدار جدید خوشه حل می‌کند و بنابراین علاوه بر پارامتر مکان، پارامتر زمان را نیز پوشش دهد و توانایی خوشه بندی با تنظیمات پارامترهای مکانی و زمانی را فراهم آورد[۲۸]. استفاده از روشهای خوشه‌بندی مبتنی بر چگالی به منظور شناسایی مناطق پارکینگ در بسیاری از تحقیقات کارایی این الگوریتم در شناسایی نقاط توقف در میان نقاط خطوط سیر *GPS* اشاره شده است[۲۹]. لذا با توجه به توضیحات فوق حذف نقاط پارکینگ در فیلتر سه با استفاده از الگوریتم خوشه‌بندی مکانی زمانی مبتنی بر تراکم *DBSCAN* صورت می‌گیرد.
- **فیلتر چهار:** حذف خطوط سیر با حداکثر چهار نقطه. پس از اجرای الگوریتم خوشه‌بندی و تجزیه خطوط ممکن است خطوط سیر با کمتر از چهار نقطه ایجاد شوند و یا ممکن است در اثر خطا در برداشت نقاط چنین خطوط سیری حاصل شده باشند.
- **فیلتر پنج:** تقسیم قطعاتی از خطوط سیر که فاصله زمانی بین تعیین‌موقعیت دو نقطه متوالی بیش از ۴ دقیقه باشد و سپس حذف خطوط سیر ایجادشده‌ای که کمتر از ۴ نقطه دارند. به‌بیان‌دیگر، پس از اعمال فیلتر پنج مجدداً فیلتر چهار اعمال می‌شود.

۳-۳- تطبیق نقشه داده‌های خطوط سیر GPS

ازآنجاکه بیشتر گیرنده‌های *GPS* مورداستفاده توسط عموم مردم در گوشی‌های همراه دقت نسبتاً پایینی داشته و با توجه به فاصله نمونه‌برداری کم و همچنین *DOP* پائین (به‌خصوص در مناطق شهری که بلوک ساختمان‌های بلند به‌طورجدی سیگنال *GPS* و درنتیجه دقت تعیین موقعیت را تحت الشعاع قرار می‌دهد) داده‌های مکانی داوطلبانه خام به‌دست‌آمده معمولاً شامل نقاط نمونه ناکافی خطادار است[۳۰]. تطبیق نقشه روند هماهنگی و تطبیق مجموعه‌ای از موقعیت‌های مشاهده‌شده توسط کاربر

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]

بـر روی شـبکه معـابر اسـت. اکثـر الگـوریتم‌هـای تطبیق نقشه بر مبنـای مجموعـه داده‌هـای ردیـابی بـا نرخ بالا هستند. ازآنجاکه اخذ و نگهـداری داده‌هـای بـا نرخ بالا نیازمند فضای ذخیـره‌سـازی و انـرژی مصـرفی بیشتر است، کاربران معمولاً در برداشت داده‌های *GPS*، فرکانس نمونه‌برداری را کـاهش مـی‌دهنـد و درنتیجـه مجموعـه داده‌هـای بسـیاری بـا نـرخ پـایین تولیـد مـی‌شـود[۳۱]. هـدف اصـلی تطبیـق نقشـه بـه‌کـارگیری اطلاعـات شـبکه جـاده راه‌هـا و اطلاعـات مکانی-زمانی *GPS* به‌منظور تصحیح خطاهـا مـی‌باشـد. [۳۲].

۳-۳-۱-انواع الگوریتم‌های تطبیق نقشه

الگوریتم‌های متعددی برای تطبیق نقشه ارائه‌شده است که در سه گروه قرار می‌گیرند: (الف) الگوریتم‌هـایی کـه از اطلاعـات هندسـی بـرای تعیـین موقعیـت استفاده می‌کنند، (ب) الگوریتم‌هـایی کـه از روش‌هـای توپولوژیکی کمک می‌گیرند و (ج) الگوریتم‌هـایی کـه از روش‌هـای یـادگیری ماشـین اسـتفاده مـی‌کننـد[۳۳]. الگوریتم‌های هندسی ازجمله الگوریتم تطبیق نقطـه بـا نقطه، تطبیـق نقطـه بـا منحنـی و تطبیـق منحنـی بـا منحنی تنها از اطلاعـات هندسـی بـرای تطبیـق نقشـه استفاده می‌کنند. به‌طورکلی الگوریتم‌های هندسی تنهـا با توجه به شکل لینک‌های معابر شهری عمل می‌کننـد و اتصال بین یـال‌هـای شـبکه را در نظـر نمـی‌گیرنـد و به‌دلیل حساسیت بالا به داده‌های پرت و خطاهای بـارز، چندان مطلوب نیستند[۳۴] و نیاز به روش‌هـایی اسـت که عمل تطابق را با توجـه بـه نقطـه قبلـی و داده‌هـای ثبت‌شده *GPS* (مانند زمان) و هم‌چنین توجه بیشتر به داده‌هـای نقشـه در دسـترس انجـام‌دهنـد[۳۵]. الگوریتم‌های مبتنی بر روش‌های توپولوژیکی با استفاده از هندسه، اتصال و مجاورت لینک‌ها، مبادرت به تطبیق نقشه می‌کنند[۳۶و ۳۴] . گروه سوم از روش‌هایی نظیر ترکیب روش‌ها و اطلاعات جـانبی *GPS* ماننـد سـرعت، جهت حرکت استفاده می‌کنند و معمولاً نیازمنـد زمـان بیشتری هستند.

با توجه به اینکه داده‌های خطوط سیر مورد استفاده در ایـن تحقیـق بـا نـرخ نمونـه‌بـرداری پـایین (دو تـا شش دقیقه) برداشت شده اند، لذا افـزایش فاصـله بـین دو نقطه مجـاور یـک چـالش بـزرگ در الگـوریتم‌هـای تطبیق نقشه به‌شمار می‌رود؛ چـرا کـه ایـن امـر باعـث کاهش دقت تعیین موقعیـت اشـیا متحـرک مـی‌شـود. این مسئله زمانی که شی در حال حرکت با سرعت بـالا و یا تقاطع‌های زیادی بین دو نقطه مجاور مشاهده شده باشـد، تشـدید مـی‌گـردد. لـذا در ایـن تحقیـق، از الگوریتمی موسوم بـه *ST-matching* کـه از هـردو نـوع اطلاعـات هندسـی و سـاختار توپولـوژیکی شـبکه و محـدودیت‌هـای سـرعت و زمـان خطـوط سـیر *GPS* بهره می‌گیرد، استفاده‌شده است. این الگوریتم با ترکیب تحلیل‌هـای مکـانی و زمـانی و انتخـاب گـراف کاندیـدا الگـوریتم *ST-matching* سـعی بـر آن دارد تـا یـک مسیر کلـی بـا بـالاترین رتبـه را تشـخیص دهـد[۳۲]. شکل (۱) ساختار کلی الگوریتم را نشان می‌دهد.

تـابع تحلیـل مکـانی بـرای محاسـبه شـاخص مکـانی به‌صورت رابطه (۱) می باشد[۳۲]:

رابطه (۱)

$$F_s(c_{i-1}^t \to c_t^i) = N(C_s^i) \times V(c_{i-1}^t \to c_s^i)$$
$$2 \le i \le n$$

کـه در آن $N(c_s^i)$ پـارامتر احتمـال مشـاهده اسـت و میزان شباهت نقطه نمونـه‌بـرداری *GPS* p_i بـا نقطـه کاندیـدا c_j^i مـی‌باشـد کـه بـر اسـاس فاصـله بـین ایـن دونقطـه و از رابطـه (۲) محاسـبه مـی‌گـردد . $V(c_{i-1}^t \to c_s^i)$ پـارامتر احتمـال انتقـال اسـت، دو نقطـه کاندیـدا c_{i-1}^t و c_i^s بـه ترتیـب بـرای دو نقطه همسایه نقطه نمونـه‌بـرداری p_{i-1} و p_i احتمـال انتقال از نقطه c_{i-1}^t به c_i^s به‌صورت شباهت مسـیر واقعی از نقطه p_{i-1} به p_i با کوتاه‌ترین مسـیر c_{i-1}^t

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]

به c_i^s و از رابطه (۳) محاسبه می‌گردد[۳۲]:

$$N(c_j^i)=\frac{1}{\sqrt{2\pi\sigma}}\exp(-\frac{(x_j^i-\mu)^2}{2\sigma^2})$$

رابطه (۲)

رابطه (۳)

$$V(c_{i-1}^t \to c_i^s)=\frac{di-1\to i}{w(i-1,t)\to(i,s)}$$

که در آنها x_j^i فاصله بین دونقطه نمونه برداری p_i و کاندیدا c_j^i است و توزیع نرمال با میانگین صفر و انحراف معیار ۲۰ متر بر اساس ارزیابی‌های تجربی در نظر گرفته‌شده است. $d_{i-1\to i}$ فاصله اقلیدسی بین دو نقطه p_{i-1} و p_i و $w(i-1,t)\to(i,s)$ طول کوتاه‌ترین مسیر از c_{i-1}^t به c_i^s روی شبکه معابر است.

تابع تحلیل زمانی برای محاسبه شاخص زمانی از ضرب دو بردار به‌صورت رابطه (۴) محاسبه می باشد، با به‌کارگیری فاصله کسینوسی میزان شباهت سرعت متوسط واقعی و محدودیت سرعت مسیر از c_{i-1}^t به c_i^s به‌دست می‌آید.

رابطه (۴)

$$F_t(c_{i-1}^t \to c_i^s)=\frac{\sum_{u=1}^{k}(e'_u.\upsilon\times\overline{\upsilon}_{(i-1,t)\to(i,s)})}{\sqrt{\sum_{u=1}^{k}(e'_u.\upsilon)^2}\times\sqrt{\sum_{u=1}^{k}\overline{\upsilon}_{(i-1,t)\to(i,s)}{}^2}}$$

در این رابطه، $e'_u.\upsilon$ و $\overline{\upsilon}_{(i-1,t)\to(i,s)}$ به‌ترتیب محدودیت سرعت و سرعت متوسط محاسبه شده ی، u امین یال از کوتاهترین مسیر روی گراف شبکه از نقطه کاندیدا c_{i-1}^t به c_i^s باشد $G'_T(V'_T,E'_T)$ گراف کاندیدا برای خطوط سیر $T:p_1\to p_2\to....\to p_n.V'_T$ شامل مجموعه‌ای از نقاط کاندیدا برای هر نقطه نمونه‌برداری GPS و E'_T مجموعه یال‌های بیانگر کوتاه‌ترین مسیر بین هر دو نقطه همسایه کاندیدا است.

رابطه (۵)

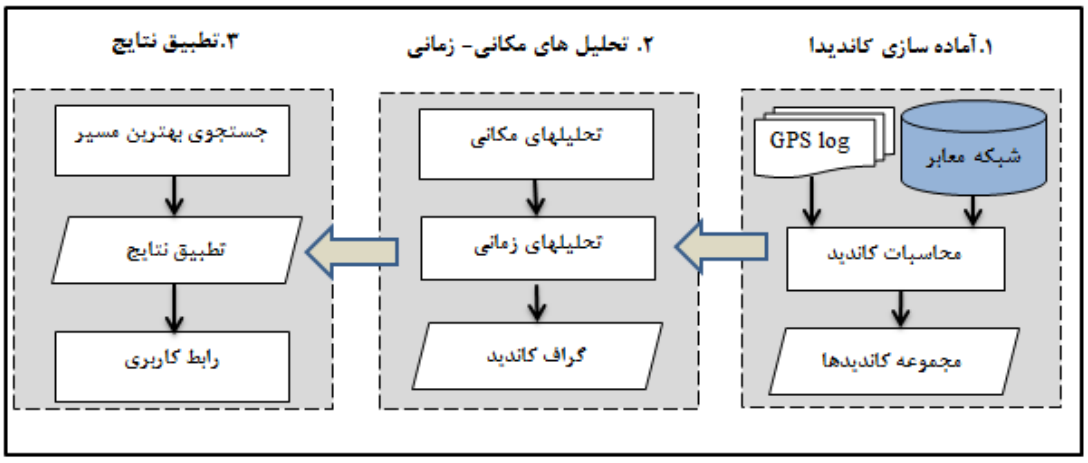
$$F(C_{i-1}^t \to C_i^s)=F_s(C_{i-1}^t \to C_i^s)\times F_t(C_{i-1}^t \to C_i^s)$$
$$2\le i\le n$$

مسیر کاندیدا P_c برای کل خطوط سیر T شامل مسیری در گراف کاندیدا به‌صورت $P_C:c_{c1}^{s_1}\to c_2^{s_2}\to...\to c_n^{s_n}$ است. هدف تعیین مسیر با بالاترین رتبه کلی است که مطابق‌ترین مسیر با خطوط سیر GPS است و از رابطه (۶) به‌دست می‌آید[۳۲]

رابطه (۶)

$$P=\arg\max_{P_c}F(P_c),\forall P_c\in G'_T(V'_T,E'_T)$$

$$F(P_c)=\sum_{i=2}^{n}F(c_{i-1}^{s_{i-1}}\to c_i^{s_i})$$



شکل ۱. ساختار الگوریتم ST-matching [۳۲]

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]

۳-۴- محاسبه سرعت جاده‌ای در شبکه

پس از انجام تطبیق نقشه و تصویر نقاط خطوط سیر روی شبکه، از سرعت خطوط سیر برای محاسبه سرعت راهها می‌توان استفاده نمود. با درنظرگرفتن بازه ۱۵ دقیقه و نقاط خط سیری که تطبیق نقشه روی آن‌ها صورت گرفته، برای هر یال و در این بازه زمانی تمام خطوط سیرهایی که از یال مذکور گذر کرده است را استخراج و با فرض اینکه جابجایی خودروها روی شبکه راه‌ها از نتایج حاصل از الگوریتم تطبیق نقشه پیروی می‌کنند و با فرض اینکه سرعت خودروی متحرک بین دونقطه نمونه‌برداری ثابت است، به تعبیری به بازسازی جابجایی خودرو روی شبکه راه‌ها پرداخته می‌شود؛ بنابراین سرعت متوسط برای هر خودرو بین هر دونقطه نمونه‌برداری محاسبه می‌شود. می‌توان با حذف خودروهای با سرعت بالا به‌عنوان خطا و محاسبه میانگین سرعت‌های متوسط خودروهای باقیمانده، سرعت راه را به‌دست آورد. هنگام میانگین‌گیری سرعت، تعداد خودروهای عبوری از هر یال نیز ثبت می‌شود که بیانگر دقت محاسبه سرعت هر قطعه از قطعات راه مربوطه است، به‌طوری‌که هر چه تعداد خودروهایی که در محاسبه سرعت مشارکت دارند، بیشتر باشد دقت بالاتری را در بردارد و به همین منظور حداقل تعداد خودرو نیز در محاسبات سرعت در نظر گرفته می‌شود و سرعت برآورد شده زمانی معتبر خواهد بود که حداقل ۳ خودرو از یال مذکور در این بازه زمانی عبور کرده باشد.

۳-۵- برآورد سرعت جریان آزاد

سرعت جریان آزاد در کتاب نظام‌نامه ظرفیت راه ۲۰۱۰ این‌گونه تعریف‌شده است: «سرعت نظری خودرو هنگامی‌که تراکم و نرخ جریان در آن بخش از راه هر دو صفر باشند و یا سرعت عمومی در بزرگراه و در جریان بین صفر و ۱۰۰۰ اتومبیل‌سواری در هر ساعت در هر خط از بزرگراه باشد»[۳۷]. درواقع، سرعت خودروها هنگامی‌که شرایط عمومی و غالب طوری باشد که تحت تأثیرهیچ وسیله نقلیه دیگری قرار نداشته باشد، به‌عنوان فضای پایه سرعت میانگین برای محاسبه این ارامتر استفاده می‌شود. نمونه‌هایی از شرایط عمومی که بر سرعت رانندگی تأثیر می‌گذارند عبارت‌اند از عرض خطوط راه، آب‌وهوا، تراز افقی و عمودی مسیر، فاصله دید، سرعت طرح راه، رفتار رانندگان و محیط مجاور[۳۸]. بنابراین، جریان ترافیک باید کم باشد تا بتوان از مشاهدات میدانی و سرعت‌های ثبت‌شده در ساعت‌های غیر اوج ترافیکی سرعت جریان آزاد را تخمین و برآورد کرد.

دیردوف و همکاران [۳۹]پژوهشی جهت یافتن رابطه‌ای میان سرعت جریان آزاد و محدودیت سرعت برمبنای نوع امکانات راه یعنی راه‌های شهری، بزرگراه‌های چندخطی و آزادراه با ثبت سرعت واقعی خودروها در نواحی مختلف و به مدت چندین هفته در شرایط روز و در ساعات غیر اوج ترافیک، دید واضح راننده و سطح خشک جاده انجام دادند. هم‌چنین وانگ و هیوجی [۴۰] مطالعه‌ای برای تعیین سرعت جریان آزاد برمبنای نظام‌نامه ظرفیت بزرگراه در مدل تقاضای سفر منطقه‌ای انجام و روابطی جهت تعیین سرعت جریان آزاد بزرگراه،بزرگراه دوخطه و چند خطه، و راه‌های شهری به‌صورت تابعی از عرض خطوط، تراکم رمپ در طول مسیر، تراکم نقاط دسترسی، سرعت طراحی، نوع میانه، موانع جانبی راه، طول قطعه‌راه، تأخیر کنترل در پایین‌دست تقاطع ارائه دادند. وانگ و همکاران در مطالعه خود به‌منظور تعیین سرعت جریان آزاد تمام سرعت‌های معتبر و مجاز جمع‌آوری‌شده برای هر قطعه‌راه را به ترتیب صعودی مرتب و مقدار سرعت در موقعیت ۸۵ درصد به‌عنوان سرعت جریان آزاد برآورد شده در نظر گرفته‌اند[۱]. همچنین کرو و همکاران [۱۳] از میانگین حداقل ۱۰۰ سرعت نقطه‌ای در ساعات غیر اوج ترافیکی استفاده کرده‌اند. اما سرعت‌های نقطه‌ای دارای ایراداتی هستند اول آنکه وابسته به مکان است و در مکان‌های مختلف تغییر می‌کند، دوم ساعات

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]

غیر اوج تعریف مشخصی ندارد و ممکن است به دلیل فعالیت‌های محلی تغییر کند و بنابراین نیازمند دانش محلی است سوم در اندازه‌گیری سرعت نقطه‌ای تأثیر تقاطع‌های چراغ‌دار در نظر گرفته نمی‌شود.

لذا در این تحقیق جهت برآورد سرعت جریان آزاد از میانگین سرعت مجاز، آلفا درصد (۵ الی ۱۵ درصد) از سریع‌ترین خطوط سیری که از هر قطعه را عبور کرده استفاده‌شده است. از مهم‌ترین مزایای این روش، بهره‌گیری از داده‌های مکانی داوطلبانه بدون صرف هزینه و در زمان کوتاه، به‌منظور برآورد مقدار سرعت جریان آزاد از داده‌های واقعی مربوط به هر قطعه از راه می‌باشد.

هم‌چنین در این روش، از برآورد سرعت جریان آزاد در چند منطقه محدود و تعمیم آن به‌کل شبکه راه که ناخودآگاه خطایی را به‌همراه دارد، خودداری می‌شود. نکته‌ی قابل‌توجه دیگر آن است که در این روش به‌طور خودکار اکثراً از خطوط سیری استفاده خواهد شد که در ساعات غیر اوج ترافیکی واقع‌شده‌اند و بدون نیاز به دانش محلی و این امر به دلیل آن است که در الگوریتم مذکور از سریع‌ترین خطوط سیر استفاده می‌شود که این خطوط سیر در ساعات غیر اوج رخ می‌دهد.

۳-۶- تعیین سرعت نهایی قطعات راه

به‌منظور تعیین سرعت نهایی هر قطعه از جاده ابتدا لازم است برآیند سرعت‌های متوسط محاسبه‌شده در پنجره زمانی تعریف‌شده (در این تحقیق پنجره زمانی ۱۵ دقیقه در نظر گرفته‌شده است) به‌عنوان سرعت نهایی آن قطعه محاسبه شود. جهت تعیین سرعت نهایی هر قطعه با استفاده از سرعت‌های متوسط، در این مقاله از روش میانگین استفاده‌شده است.

اگر کمیت سرعت چندین بار و حتی المقدور تحت شرایط یکسان (با فرض یکسان بودن شرایط ترافیک در بازه‌های زمانی کوتاه) اندازه‌گیری شود، آنگاه آن اندازه‌گیری، یک نمونه از متغییر تصادفی را تشکیل می‌دهد. با فرض اینکه متغییر سرعت از توزیع نرمال پیروی می‌کند، تابع توزیع یک متغییر تصادفی نرمال، با دو پارامتر امید ریاضی μx و واریانس σx^2 بیان می‌شود. بررسی‌های آماری نشان می‌دهد که حدود ۶۸٫۳ درصد از مشاهدات در فاصله $\mu-\sigma$ و $\mu+\sigma$ واقع می‌باشند به‌عبارت‌دیگر ۳۱/۷ درصد از مشاهدات در خارج از این فاصله قرار دارند[۴۱]. دو برآورد کننده نااریب[1] از (μ,σ^2) میانگین نمونه v و واریانس نمونه s^2 می‌باشند که توسط رابطه (۷) تعریف می‌شوند :

رابطه (۷) $\bar{v}=\left(\frac{1}{n}\right)\sum_{i=1}^{n} v_i \quad , \quad s^2=\frac{\sum_{i=1}^{n}(v_i-\bar{v})^2}{n-1}$

در رابطه (۶)، $v_1, v_2, \ldots, v_n$ عناصر نمونه هستند[۴۱]. در این روش ابتدا سرعت‌های نویز حذف و از سرعت‌های باقیمانده میانگین‌گیری می‌شود. برای حذف سرعت‌های نویز از تست آماری مجموعه سرعت‌ها در سطح اطمینان ۹۵ درصد استفاده می‌گردد. فرآیند حذف نویز با میانگین و انحراف معیار جدید تا زمان حذف کامل نویزها تکرار می‌شود.

۳-۷- تعیین مناطق راهبندان های ترافیکی

پس از محاسبه سرعت جاده‌ای، شناسایی رخداد راهبندان‌های ترافیکی بر روی هر یال به‌طور مجزا انجام می‌شود. بدین منظور از اعمال حد آستانه سرعت بر هر یال استفاده می‌شود. بهره‌گیری از محدوده سرعت و حد آستانه مبتنی بر سرعت جریان آزاد می‌تواند برآورد خوبی از راهبندان‌های ترافیکی نتیجه دهد. سرعت جریان آزاد یک معیار بسیار مهم در نظارت ترافیک آنی است و از آن به‌منظور تعیین و طبقه‌بندی سطح ترافیک مسیر استفاده می‌شود[۴۲]. بنابراین از مقایسه سرعت محاسبه‌شده نهایی هر قطعه‌راه و حد آستانه‌های تعریف‌شده، سطح ترافیک قطعه مدنظر در بازه زمانی تعریف‌شده تعیین خواهد شد. در جدول (۱) حدآستانه سرعت برای سطوح مختلف ترافیکی آزادراه ارائه شده است.

[1] unbiased estimators

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]

جدول ۱: طبقه‌بندی سطح ترافیک برمبنای درصد سرعت جریان آزاد

سطح ترافیک	حد آستانه سرعت
ترافیک روان	بین صفر و ۳۵ درصد مقدار سرعت جریان آزاد
ترافیک کند (نیمه سنگین)	بین ۳۵ و ۶۵ درصد مقدار سرعت جریان آزاد
راهبندان های ترافیکی (ترافیک سنگین)	بین ۶۵ الی ۱۰۰ درصد مقدار سرعت جریان آزاد و بالاتر

۴- مطالعه موردی و نمونه‌های محاسباتی

داده‌های مورداستفاده شامل خطوط سیر *GPS* مربوط به ۱۰۳۵۷ تاکسی در بازه زمانی یک هفته ۲ تا ۸ فوریه سال ۲۰۰۸ در شهر پکن (پایتخت چین) می‌باشد. تعداد کل نقاط در این مجموعه داده حدود ۱۵ میلیون نقطه و فاصله کل خطوط سیر به ۹ میلیون کیلومتر می‌رسد. فیلترهای زیر جهت پاک‌سازی داده‌ها و به‌عنوان پیش‌پردازش بر روی‌داده‌های نقطه‌ای خط سیر اعمال گردید که جزییات هر یک در ادامه بیان می‌گردد.

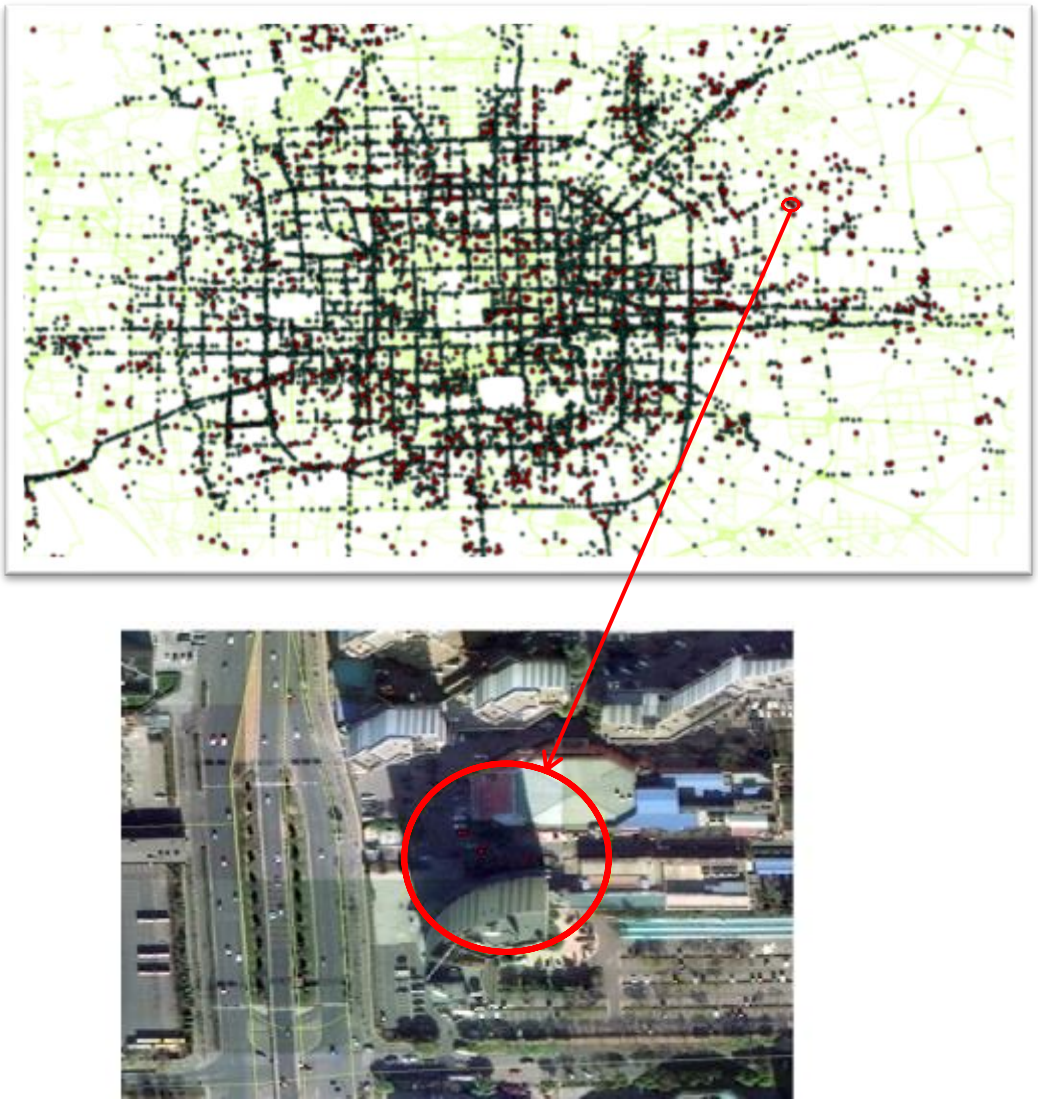
- با اعمال فیلتر اول حذف مختصات غیرواقعی و نقاط نمونه‌برداری خارج از محدوده شهر پکن صورت پذیرفت و ۲۶۸۶۷۷۰ نقطه حذف و تعداد آن به ۱۴۹۷۶۲۱۴ کاهش یافت. همچنین تعداد ۷۹ خودرو حذف گردید و ۱۰۲۵۷ خورد باقی ماند.
- با اعمال فیلتر دوم، ۱۰۴۴۴۵۵ نقطه تکراری حذف و تعداد ۱۳۹۳۱۷۵۹ نقطه باقی ماند.
- در فیلتر سوم نقاطی که در ۳۰ دقیقه، جابجایی کمتر از ۱۰۰ متر داشتند، به‌عنوان نقاط پارکینگ در نظر گرفته‌شده و حذف شدند. شناسایی نقاط پارکینگ با استفاده از الگوریتم خوشه‌بندی *DBSCAN* مکانی-زمانی با تنظیم پارامترهای مکانی و زمانی انجام شد [۲۴]. با اعمال این فیلتر تعداد ۶۶۰۵۷۳۱ نقطه حذف و ۷۳۲۶۰۲۸ نقطه باقی ماند. ضمناً ۲۹۰ خودرو حذف و تعداد خطوط سیر ایجادشده برای ۹۹۶۷ خودروی باقیمانده برابر با ۱۹۲۶۱۸ گردید. شکل (۲) نتایج خوشه‌بندی مکانی زمانی *DBSCAN* بر روی‌داده‌های خطوط سیر *GPS* را نشان می‌دهد. نقاط قرمزرنگ در این تصویر به‌عنوان نقاط پارکینگ شناسایی شده‌اند. به منظور وضوح بهتر از بزرگنمایی یکی از مناطق استفاده شده‌است.

همان‌طور که در شکل (۳) مشاهده می‌شود، نقاط قرمزرنگ شناسایی‌شده در الگوریتم *DBSCAN* مکانی-زمانی، واقع در پارکینگ یک بلوک مسکونی هستند. این امر نشان می‌دهد که الگوریتم مورد استفاده، مناطق پارکینگ (ثبت نقطه با موقعیت یکسان برای مدت‌زمانی طولانی) را به‌خوبی شناسایی کرده است.

- در فیلتر پنجم نقاط خط سیر خودروهایی که دارای رزولوشن زمانی نامناسب بودند (فاصله زمانی دونقطه نمونه‌برداری بیش از چهار دقیقه) از موقعیت مکانی نقطه‌ای که دارای گپ بودند تجزیه و به دو خط سیر مجزا با شناسه فرعی جداگانه تقسیم شدند. سپس تمام خطوط سیری که تعداد نقاطشان کمتر از ۴ نقطه باشد، شناسایی و حذف گردیدند. با اعمال این فیلتر ۲۰۹۵۲۵ نقطه و ۲۲۷ خودرو حذف و ۷۱۱۶۵۰۳ نقطه مربوط به ۹۷۴۰ خودرو باقی ماند.

مجموع خطوط سیر ایجادشده نهایی ۳۴۶۲۷۷ می‌باشد که هر یک از آن‌ها شامل حداقل ۴ نقطه و حداکثر ۴۱۵ نقطه متوالی و با فواصل زمانی هر دونقطه متوالی کمتر از ۴ دقیقه هستند. شکل (۳) تعداد نقاط و خودروهای باقی‌مانده پس از انجام هر فیلتر را نشان می‌دهد.

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]



شکل ۲: خوشه‌بندی مکانی زمانی DBSCAN بر روی کل شبکه شهر پکن و بزرگنمایی یکی از مناطق شناسایی شده جهت حذف پارکینگ

شکل ۳: تعداد نقاط و خودروهای باقی‌مانده پس از اجرای هر فیلتر بر روی‌داده‌های خطوط سیر GPS

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

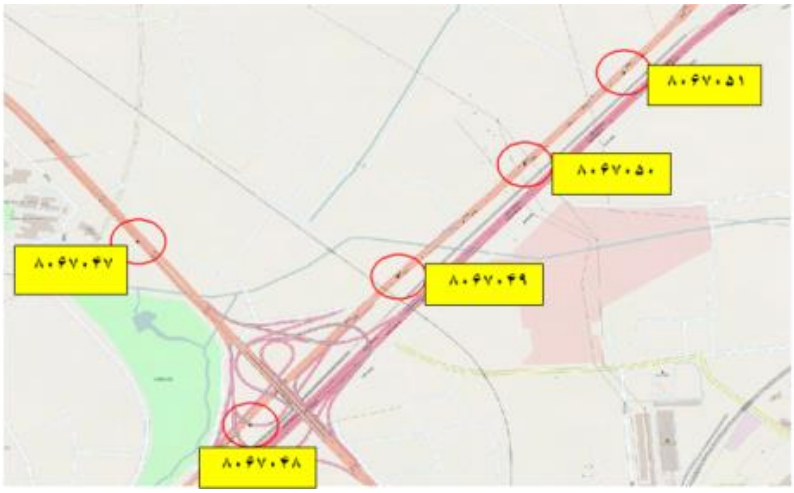
[DOI: 10.29252/jgit.5.3.51]

به‌منظور تشریح نحوه اجرای الگوریتم و محاسبه سرعت متوسط، محاسبات مربوط به تطبیق نقشه مکانی-زمانی و برآورد سرعت متوسط یک نمونه خط سیر با شماره شناسه ۲۸ و شناسه خودروی ۴۸۰۸ شامل ۲۲ نقطه در طول مدت‌زمان تقریبی ۲۱ الی ۲۲ دقیقه که توسط کاربر جمع‌آوری‌شده، بیان می‌شود. به‌دلیل تعداد زیاد نقاط نمونه‌برداری این خط سیر و طولانی شدن جداول مربوطه، جداول و اشکال برای پنج نقطه ابتدایی خط سیر نمونه، نشان داده‌شده است. مشخصات نقاط خط سیر نمونه مطابق جدول (۲) است.

جدول ۲: مشخصات نقاط خط سیر نمونه

شناسه نقطه	برچسب زمانی	عرض جغرافیایی	طول جغرافیایی
۸۰۶۷۰۴۷	"۲۳:۰۰:۱۶ ۰۳-۰۲-۲۰۰۸"	۴۰٫۰۰۴۶۵	۱۱۶٫۴۸۷۱۱
۸۰۶۷۰۴۸	"۲۲:۰۱:۱۶ ۰۳-۰۲-۲۰۰۸"	۳۹٫۹۹۷۱۸	۱۱۶٫۴۹۱۹۱
۸۰۶۷۰۴۹	"۲۲:۰۲:۱۶ ۰۳-۰۲-۲۰۰۸"	۴۰٫۰۰۳۲۲	۱۱۶٫۴۹۸۲۴
۸۰۶۷۰۵۰	"۲۲:۰۳:۱۶ ۰۳-۰۲-۲۰۰۸"	۴۰٫۰۰۷۷۹	۱۱۶٫۵۰۳۶۴
۸۰۶۷۰۵۱	"۲۳:۰۴:۱۶ ۰۳-۰۲-۲۰۰۸"	۴۰٫۰۱۱۵۳	۱۱۶٫۵۰۷۸۷

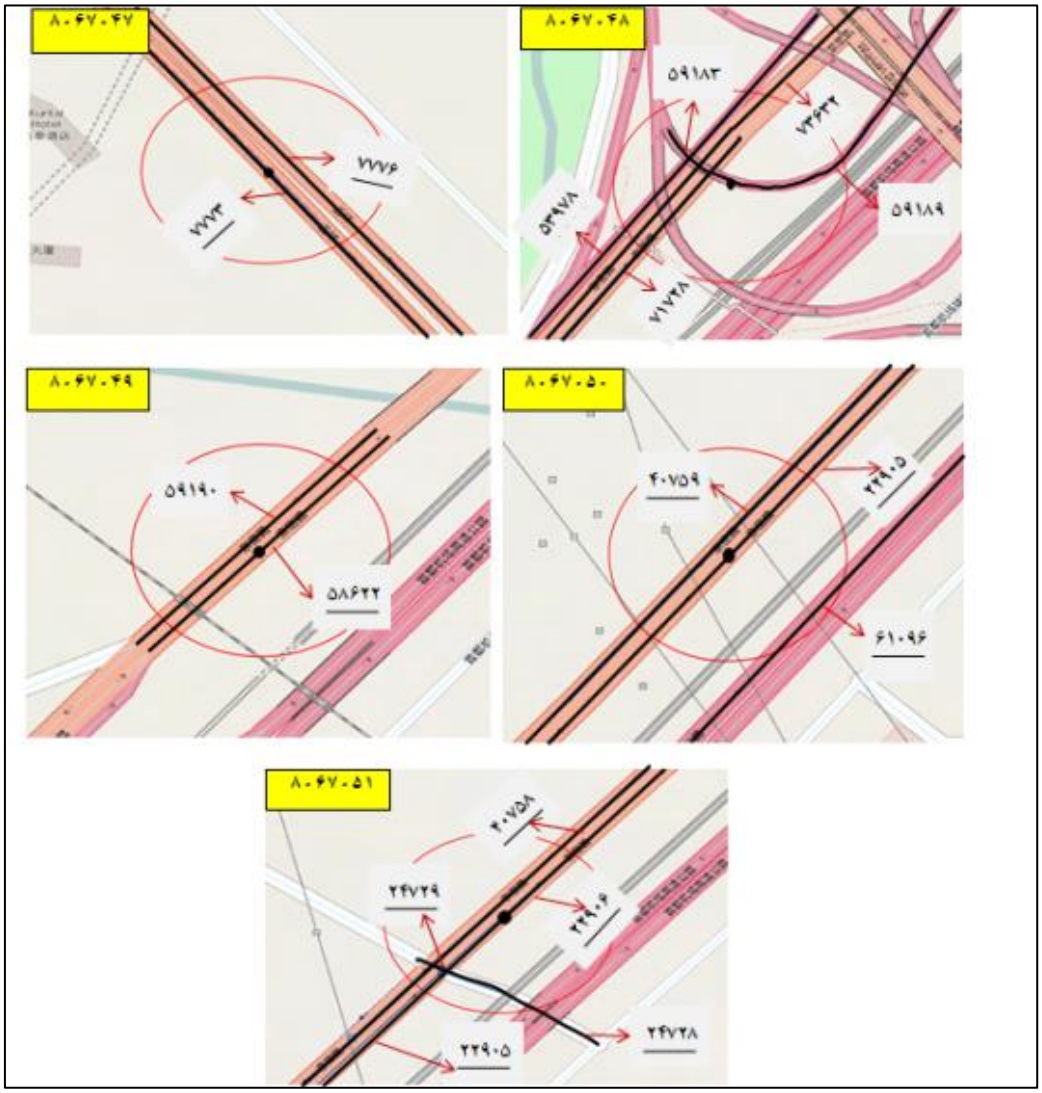
در ابتدا به‌منظور اجرای الگوریتم با در نظر گرفتن یک شعاع حول نقطه نمونه‌برداری، نقاط و یال‌های کاندیدا تعیین می‌شوند. شکل (۴) موقعیت مکانی پنج نقطه ابتدایی خط سیر بر روی شبکه معابر به همراه شناسه آن‌ها و بافر ۱۰۰ متری حول هر یک از نقاط را و شکل(۵) جزییات این نقاط و یال‌ها را نشان می‌دهند. (با توجه به پارامترهای تعریف‌شده در الگوریتم تطبیق نقشه در این تحقیق، شعاع ۱۰۰ متر و حداکثر کاندیدا ۵ در نظر گرفته‌شده است).



شکل ۴: موقعیت مکانی و شناسه پنج نقطه ابتدایی خط سیر GPS نمونه

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]



شکل ۵: یال‌های کاندیدا نقاط نمونه‌برداری متوالی خطوط سیر GPS

پس از تعیین نقاط و یال‌های کاندیدا، برای نقطه اول به دلیل اینکه نقطه‌ای ماقبل آن به‌عنوان مبنای محاسبات وجود ندارد، در میان یال‌های کاندیدا یال با نزدیک‌ترین فاصله به نقطه نمونه‌برداری(نقطه ۷۷۷۳)، انتخاب می‌گردد. در مرحله بعد برای تعیین نقطه کاندیدای برنده از میان پنج نقطه کاندیدا مربوط به نقطه نمونه‌برداری دوم، آنالیز کوتاه‌ترین مسیر بین نقطه برنده اول و تمام نقاط کاندیدای نقطه نمونه‌برداری دوم صورت می‌گیرد. سپس شاخص‌های مکانی و زمانی محاسبه و درنتیجه شاخص نهایی (f) الگوریتم تطبیق نقشه مکانی - زمانی از حاصل‌ضرب شاخص مکانی (fs) و شاخص زمانی (ft) محاسبه می‌گردد. همان‌طور که در جدول (۳) مشاهده می‌شود در میان نقاط کاندیدا نقطه‌ای که دارای بیشترین مقدار شاخص مکانی زمانی است به‌عنوان یال برنده تعیین شد (یال ۵۹۱۸۹). درنتیجه کوتاه‌ترین مسیر بین این دونقطه متوالی به‌عنوان یال‌هایی که خودرو از آن‌ها عبور کرده در نظر گرفته می‌شود. به‌منظور تعیین نقطه کاندیدای برنده برای نقطه نمونه‌برداری بعدی، نقطه برنده قبلی (نقطه ۵۱۱۸۹) مبنای محاسبات قرار می‌گیرد و به‌همین ترتیب فرآیند تطبیق نقشه به‌طور متوالی و

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]

برای تمام زوج نقاط متوالی انجام می‌شود. مسیرها، یال‌های برنده و سرعت متوسط این یال‌ها در جدول‌های (۳ و ۴) آمده‌است. با محاسبه مجموع طول یال‌های پیموده شده در مسیر برنده و همچنین در دست داشتن اختلاف‌زمانی ثبت دونقطه، سرعت متوسط این یال‌ها مطابق جدول (۳) محاسبه گردید.

جدول ۳: مشخصات نقاط و قطعات کاندیدا و پارامترهای تطبیق نقشه مربوطه خط سیر

شناسه نقطه نمونه‌برداری	شناسه یال‌های کاندیدا در شعاع ۱۰۰ متری نقطه نمونه‌برداری	فاصله یال کاندیدا از نقطه نمونه‌برداری (متر)	شاخص تطبیق نقشه از نقطه منتخب قبلی به نقطه کاندیدا	مجموع مسافت از نقطه منتخب قبلی به کاندیدا (متر)	مقدار سرعت متوسط برآوردی برای مسیرهای کاندیدا (کیلومتر بر ساعت)
۸۰۶۷۰۴۷	۷۷۷۳(برنده)	۴۴٫۲			
	۷۷۷۶	۲۲٫۲۲			
۸۰۶۷۰۴۸	۵۹۱۸۹(برنده)	۰۴٫۶	۰٫۰۰۷۱۲۲	۹۶۰٫۲۶	۵۸٫۵۹
	۵۹۱۸۳	۳۷٫۹	۰٫۰۰۷۰۴۱	۹۵۲/۹۱	۵۸٫۱۴
	۷۱۸۲۷	۳۲٫۲۴	۰٫۰۰۲۷۵۲	۲۲۱۳/۹۹	۱۳۵٫۰۹
	۷۳۶۳۲	۶۳٫۳۴	۰٫۰۰۲۷۴۳	۱۹۱۳/۱۴	۱۱۶٫۷۳
	۵۳۹۷۸	۵۱٫۴۶	۰٫۰۰۰۴۱۷	۱۰۴۹۶/۱۶	۶۴۰٫۴۴
۸۰۶۷۰۴۹	۵۸۶۲۲ (برنده)	۷۴٫۰	۰٫۰۰۶۶۰۹	۸۸۸/۱۰	۵۳٫۲۸
	۵۹۱۹۰	۴۳٫۱۱	۰٫۰۰۳۹۲۲	۱۵۸۲/۳۷	۹۴٫۹۴
۸۰۶۷۰۵۰	۲۲۹۰۵(برنده)	۵۳٫۱	۰٫۰۰۷۱۱۱	۶۸۶/۴۰	۴۱٫۱۸
	۴۰۷۵۹	۴۳٫۱۱	۰٫۰۰۲۹۹۲	۱۶۴۶/۰۴۴	۹۸٫۷۶
	۶۱۰۹۶	۴۰٫۹۳	۰٫۰۰۰۲۵۹	۳۴۸۹٫۸۶	۲۰۹٫۳۹
۸۰۶۷۰۵۱	۲۲۹۰۶(برنده)	۶۸٫۲	۰٫۰۰۶۴۹۲	۵۵۰٫۲۵	۳۲٫۴۷
	۲۴۷۲۸	۷۵٫۷۰	۰٫۰۰۲۵۶۷	۵۱۱٫۴۶	۳۰٫۱۸
	۲۴۷۲۹	۵۱٫۷۳	۰٫۰۰۲۵۴۶	۴۷۶٫۰۰	۲۸٫۰۹
	۴۰۷۵۸	۸۰٫۱۴	۰٫۰۰۱۳۴۷	۲۹۸۱٫۲۴	۱۷۵٫۹۴

شکل (۶) سرعت متوسط را برای سه نقطه ابتدایی خط سیر نمونه با فواصل نمونه‌برداری یک دقیقه نشان می‌دهد. بعد از انجام عملیات تطبیق نقشه یال‌های آبی و قرمز که خودرو از آن عبور کرده مشخص گردید. پس از محاسبه فاصله شبکه‌ای بین دو نقطه متوالی، سرعت متوسط بین نقاط نمونه‌برداری با شناسه‌های ۸۰۶۷۰۴۷ و ۸۰۶۷۰۴۸ برابر با ۵۸٫۵۹ کیلومتر بر ساعت و بین نقاط نمونه‌برداری با شناسه‌های ۸۰۶۷۰۴۸ و ۸۰۶۷۰۴۹ برابر با ۵۳٫۲۸ کیلومتر بر ساعت به‌دست آمد. دنباله‌های این دو مطابق سطر اول جدول (۴) می‌باشد.

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]

جدول ۴: دنباله یال‌های گراف شبکه راه از هر نقطه کاندیدا برنده به نقاط کاندیدا بعدی

از یال	به یال	سرعت	دنباله
۷۷۷۳	۵۹۱۸۹ (برنده)	۵۸٫۵۹	(۷۷۷۳,۶۲۰۱۷,۵۴۰۰۲,۲۲۸۹۵,۲۲۸۹۶,۵۳۹۷۴,۵۹۱۸۳)
	۵۹۱۸۳	۵۸٫۱۴	(۷۷۷۳,۶۲۰۱۷,۵۴۰۰۲,۲۲۸۹۵,۲۲۸۹۶,۵۳۹۷۴)
	۷۱۸۲۷	۱۳۵٫۰۹	(۷۷۷۳,۶۲۰۱۷,۵۴۰۰۲,۲۲۸۹۵,۱۰۶۵۶,۱۰۶۵۷,۲۴۷۵۷,۲۴۷۵۸,۳۵۳۴۲)
	۷۳۶۳۲	۱۱۶٫۷۳	(۲۷۶۱۳,۲۷۶۱۲,۲۷۶۱۱,۲۷۶۱۸,۲۷۶۰۹,۲۷۶۰۸,۵۳۹۷۶,۱۰۶۵۱,۱۰۶۵۰,۷۳۶۳۱)
	۵۳۹۷۸	۶۴۰٫۴۴	(۷۷۷۳,۶۲۰۱۷,۵۴۰۰۲,۲۲۸۹۵,۱۰۶۵۶,۵۳۹۷۸)
۵۹۱۸۹	(برنـــــده) ۵۸۶۲۲	۵۳٫۲۸	(۵۹۱۸۹,۲۲۸۹۹)
	۵۹۱۹۰	۹۴٫۹۴	(۵۹۱۸۹,۲۲۸۹۹,۵۸۶۲۲,۵۹۱۷۷,۲۲۹۰۴,۲۶۸۱۵,۴۰۷۶۰,۵۹۱۷۵)
۵۸۶۲۲	۲۲۹۰۵ (برنده)	۴۱٫۱۸	(۵۸۶۲۲,۵۹۱۷۷,۲۲۹۰۴)
	۴۰۷۵۹	۹۸٫۷۶	(۵۸۶۲۲,۵۹۱۷۷,۲۲۹۰۴,۲۲۹۰۵,۲۴۷۲۹)
	۶۱۰۹۶	۲۰۹٫۳۹	(۵۸۶۲۲,۵۹۱۷۷,۲۲۹۰۴,۲۲۹۰۵,۲۲۹۰۶,۲۲۹۰۷,۲۲۸۸۱,۲۲۸۸۲,۲۲۸۷۲,۶۱۱۱۱)
۲۲۹۰۵	۲۲۹۰۶ (برنده)	۳۲٫۴۷	(۲۲۹۰۵)
	۲۴۷۲۸	۳۰٫۱۸	(۲۲۹۰۵,۲۴۷۲۸)
	۲۴۷۲۹	۲۸٫۰۹	(۲۲۹۰۵)
	۴۰۷۵۸	۱۷۵٫۹۴	۲۲۹۰۵,۲۲۹۰۶,۲۲۹۰۷,۲۲۸۸۱,۲۲۸۸۹,۲۲۸۹۰,۲۲۸۹۱,۵۹۱۶۷,) (۵۹۱۶۳,۵۹۱۸۷,۲۲۸۸۰,۲۲۹۱۲,۶۶۵۴۳۴۰۷۵۷
۲۲۹۰۶	۶۶۵۴۴ (برنده)	۴۹٫۱۵	(۲۲۹۰۶,۲۲۹۰۷,۲۲۹۰۸)
	۴۰۷۶۱	۴۹٫۶۶	(۲۲۹۰۶,۲۲۹۰۷,۲۲۹۰۸,۶۶۵۴۴)
	۵۹۱۶۳	۶۶٫۸۶	(۲۲۹۰۶,۲۲۹۰۷,۲۲۸۸۱,۲۲۸۸۹,۲۲۸۹۰,۲۲۸۹۱,۵۹۱۶۷)
	۶۶۵۴۳	۹۲٫۱۷	(۲۲۹۰۶,۲۲۹۰۷,۲۲۸۸۱,۲۲۸۸۹,۲۲۸۹۰,۲۲۸۹۱,۵۹۱۶۷,۵۹۱۶۳,۵۹۱۸۷,۲۲۸۸۰,۲۲۹۱۲)
	۲۲۹۱۲	۹۱٫۴۳	(۲۲۹۰۶,۲۲۹۰۷,۲۲۸۸۱,۲۲۸۸۹,۲۲۸۹۰,۲۲۸۹۱,۵۹۱۶۷,۵۹۱۶۳,۵۹۱۸۷,۲۲۸۸۰)

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]

شکل ۶: سرعت متوسط و دنباله یال‌های گراف شبکه راه برای سه‌نقطه ابتدای خط سیر GPS با شناسه ۲۸

برای نمایش نحوه برآورد سرعت نهایی قطعه یال‌های ۲۲۹۰۵ و ۲۲۹۰۶ در بازه زمانی ۲۰ دقیقه، مجموعه سرعت‌های محاسبه‌شده از خطوط سیر کاربران (خودروها) در جدول (۵) در نظر گرفته شد. میانگین و انحراف استاندارد سرعت‌های قطعه ۲۲۹۰۵ به ترتیب ۶۴/۰۱ و ۲۳/۸۸ کیلومتر بر ساعت و میانگین و انحراف استاندارد سرعت‌های قطعه ۲۲۹۰۶ به ترتیب ۶۲/۷۶ و ۱۵/۰۴ کیلومتر بر ساعت است، بنابراین مقادیر سرعت مجاز برای نمونه اول و نمونه دوم به ترتیب باید در محدوده‌های (۳۲/۷۶-۹۲/۸۵) و (۱۶/۲۴-۱۱۱/۷۶) باشند، لذا سرعت‌هایی که در محدوده قرار ندارند حذف می‌شوند. در مجموعه دوم تمام سرعت‌ها در محدوده مجاز هستند و بنابراین میانگین آن‌ها به‌عنوان سرعت نهایی برآورد می‌شود اما در مجموعه اول مقدار سرعت ۱۱۳/۶۳ خارج از محدوده قرار دارد و حذف می‌شود. میانگین و انحراف استاندارد و محدوده سرعت مجاز جدید به ترتیب ۸۰/۵۷ و ۱۶/۰۰ و (۲۵/۸۰-۸۹/۸۰) برای سرعت‌های باقیمانده محاسبه می‌شود. با مقایسه سرعت‌ها و محدوده سرعت کلیه سرعت‌های باقیمانده در محدوده قرار دارند و بنابراین میانگین سرعت مجموعه باقیمانده ۵۷/۸۰ برآورد و تعیین می‌گردد.

۳۲/۵۶ درصد از کل قطعات شبکه معابر با استفاده از خودروهای شرکت‌کننده تحت عبور قرار گرفته است؛ بنابراین سرعت این قطعات برآورد شده و میزان ۶۷/۴۴ درصد از راه‌های این شهر که اغلب مناطق مسکونی و راه‌های فرعی هستند، عبور کافی جهت تعیین سرعت نداشته‌اند. از میان ۳۲/۵۶ درصد شبکه راه که به‌روز رسانی سرعت روی آن‌ها صورت گرفت ۶۷/۵ درصد آن‌ها راه‌های اصلی شبکه معابر و ۳۲/۵ درصد آن‌ها راه‌های فرعی شبکه معابر شهر هستند. شکل (۷) نقشه راهبندان‌های ترافیکی در شبکه معابر شهر پکن را نمایش می‌دهد. همان‌طور که مشاهده می‌شود معابر این شهر با سه رنگ قرمز، آبی و سبز که به ترتیب نمایانگر راهبندان های ترافیکی (ترافیک سنگین)، ترافیک کند (نیمه سنگین) و ترافیک روان نمایش داده‌شده‌است. با توجه به نقشه به‌دست‌آمده مناطق دارای راهبندان‌های ترافیکی در بازه زمانی ۱۶

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]

الی ۱۶:۲۰ اکثراً در معابر اصلی و مخصوصاً در لینک‌های اتصال بین آن‌ها مشاهده می‌شود و این به علت انتقال جریان از محل‌های کار به محل زندگی می‌باشد لذا نتایج حاصله منطقی است.

جدول ۵: سرعت نهایی برآورد شده دو یال نمونه ۲۲۹۰۵ و ۲۲۹۰۶

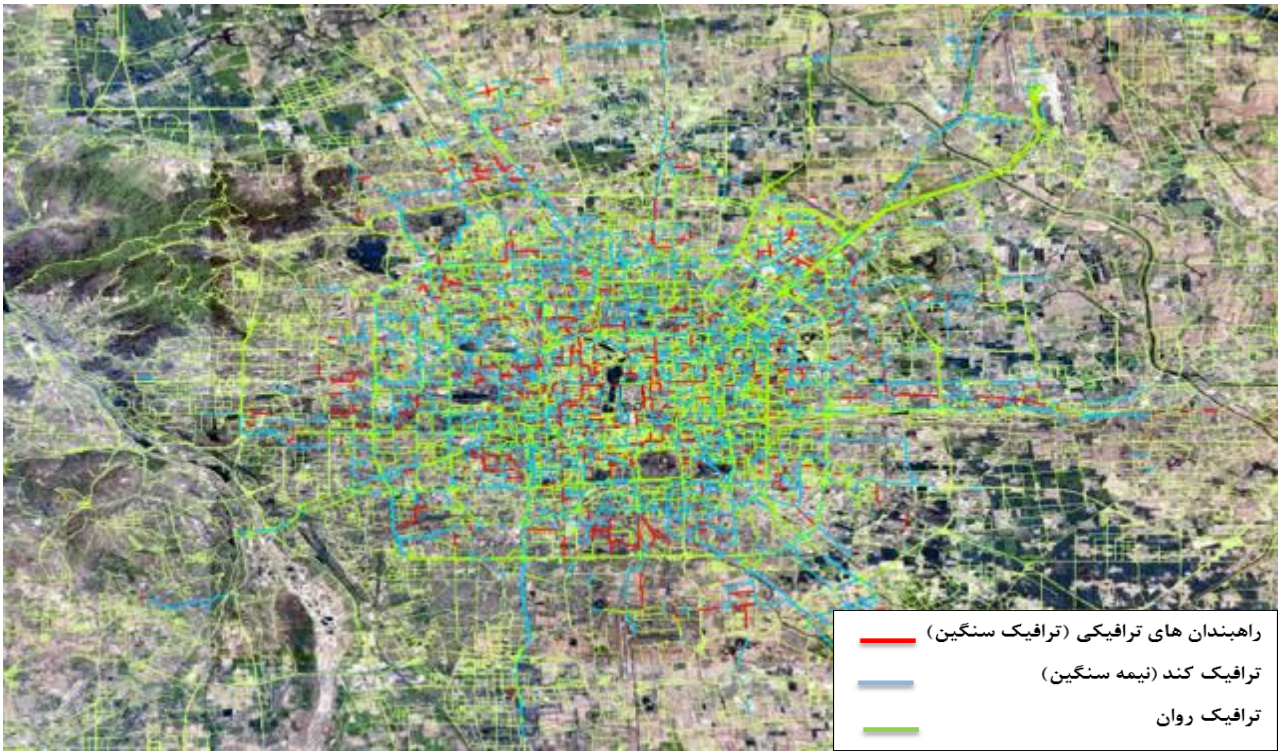
شناسه یال	سرعت جریان آزاد	شناسه خودرو عبور کننده	سرعت متوسط	سرعت نهایی
۲۲۹۰۵	۷۶/۰۷	۸۰۰	۷۷/۴۱	۵۷/۸۰
		۲۷۲۲	۴۸/۵۸	
		۲۹۶۹	۸۹/۷۴	
		۶۰۹۹	۲۳/۳۶	
		۶۷۲۰	۱۵/۷۴	
		۷۲۱۴	۱۵/۴۱	
		۹۵۴۵	(نویز) ۱۱۳/۶۳	
		۱۰۰۲۱	۷۵/۷۱	
		۶۱۸۶	۰۰/۶۴	
۲۲۹۰۶	۷۵/۳۱	۶۱۸۶	۰۰/۶۴	۷۲/۷۶
		۷۲۱۴	۱۵/۴۱	
		۶۷۲۰	۱۵/۷۴	
		۱۰۰۲۱	۷۵/۷۱	

شکل (۸) وضعیت ترافیک در بازه زمانی بین ساعت ۱۶ الی ۱۶:۲۰ را نشان می‌دهد. در این بازه زمانی حدوداً ۳۵۶۰ خودرو مشارکت داشته‌اند که معادل نرخ نفوذ ۳٫۸ درصد کل خودروهای سطح شهر است.

نتایج حاصل از اجرای الگوریتم‌های استخراج سرعت با انجام چندین آزمایش تحت درستی آزمایی و تحلیل‌های آماری قرار گرفتند. به‌منظور ارزیابی نتایج سرعت متوسط از ۱۰۰ خط سیر شامل ۱۴۳۶۶ نقطه نمونه‌برداری و ۱۹۸۲۰ خط شبکه با مجموع طول ۳۵۶۴٫۱۸ کیلومتر مربوط به ۷ کاربر و در مناطق مختلف شبکه معابر با تراکم خطوط مختلف استفاده شد. ۴۴٫۷ درصد خطوط سیر در مناطق بزرگراه‌ها و کمربندی شهر پکن (رینگ سوم و چهارم) و ۵۵٫۳ درصد خطوط سیر در مناطق متراکم شهری (مرکز شهر) با نرخ نمونه‌برداری یک دقیقه در شهر پکن انتخاب و مورد ارزیابی قرار گرفت.

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]



شکل ۷: نقشه راهبندان های ترافیکی شبکه معابر شهر پکن در بازه زمانی بین ساعت ۱۶ الی ۱۶:۲۰

۵- اعتبارسنجی و ارزیابی نتایج

در این بخش به ارائه خطا و ارزیابی نتایج پرداخته می‌شود.

۵-۱-خطای ناشی از تعیین موقعیت GPS بر سرعت

خطایی است که در جابجایی نقطه روی یال انتخاب‌شده ایجاد می‌شود، این خطا موجب تغییر در طول مسافت بین دونقطه نمونه‌برداری و درنتیجه باعث خطا در سرعت متوسط اندازه‌گیری می‌شود. اگر میزان خطا در تمام جهت‌ها یکسان فرض شود (دایره خطا) خطای تعیین موقعیت در راستای خطوط شبکه نیز برابر با مقدار خطا در جهت عمود بر آن خواهد بود، مقدار خطای تعیین موقعیت *GPS* مختلف و متناسب با عواملی از جمله فنّاوری بکار رفته در آن‌ها و تعداد ماهواره‌های دارای دید در لحظه تعیین موقعیت است. به‌منظور ارزیابی این خطا و تأثیر آن بر سرعت، یک نمونه آماری از اندازه‌گیری‌های ۷ کاربر (دستگاه *GPS*) شامل ۲۰۳ نقطه اندازه‌گیری مورد تجزیه و تحلیل قرار گرفت. میانگین خطا برابر با ۱۰٫۱۲ متر است، ضمنا خطای ۹۶ درصد از نقاط نمونه کمتر از ۳۰ متر می باشد. خطایی که تعیین موقعیت *GPS* در سرعت خط سیر ایجاد می کند از رابطه (۸) قابل‌محاسبه است:

$$dV = \frac{\pm dl}{T}$$

رابطه (۸)

که در آن dv خطای سرعت برآورد شده ، dl dl خطای میانگین طول، TT مدت‌زمان بین دو نمونه‌برداری متوالی هستند. با در نظر گرفتن خطای میانگین طول (dl) برابر با ۱۰٫۱۲ متر در فاصله پیموده شده در مدت‌زمان بین دو نمونه‌برداری متوالی (T) یک دقیقه، برابر با ۰٫۶ کیلومتر بر ساعت خواهد بود. مقدار خطا برای ۹۶ درصد سرعت‌ها کمتر از ۱٫۸ کیلومتر بر ساعت است.

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]

۵-۲-ارزیابی خطای ناشی از اجرای الگوریتم تطبیق نقشه

عوامل متعددی دقت تطبیق نقشه را تحت تأثیر قرار می‌دهند، به همین دلیل به‌منظور سنجش تأثیر پارامترهای مختلف بر الگوریتم تطبیق نقشه و تنظیم پارامترها جهت اجرای بهینه (هم ازلحاظ دقت خروجی و هم ازلحاظ زمانی) چندین آزمایش، طراحی و انجام شد. دقت تطبیق نقشه با استفاده از دو پارامتر دقت بر اساس تعداد A_N و دقت بر اساس طول A_L تعیین گردید. دقت بر اساس تعداد (A_N) برابر است با نسبت تعداد خطوطی که به‌درستی تطبیق شده‌اند (N_C) به تعداد کل مسیر ردیابی (N_A) است. دقت بر اساس طول (A_L) برابر است با نسبت مجموع طول قطعاتی که به‌درستی تطبیق شده‌اند (L_C) به‌کل طول مسیر ردیابی (L_A) است و از رابطه(۹) محاسبه می‌شود

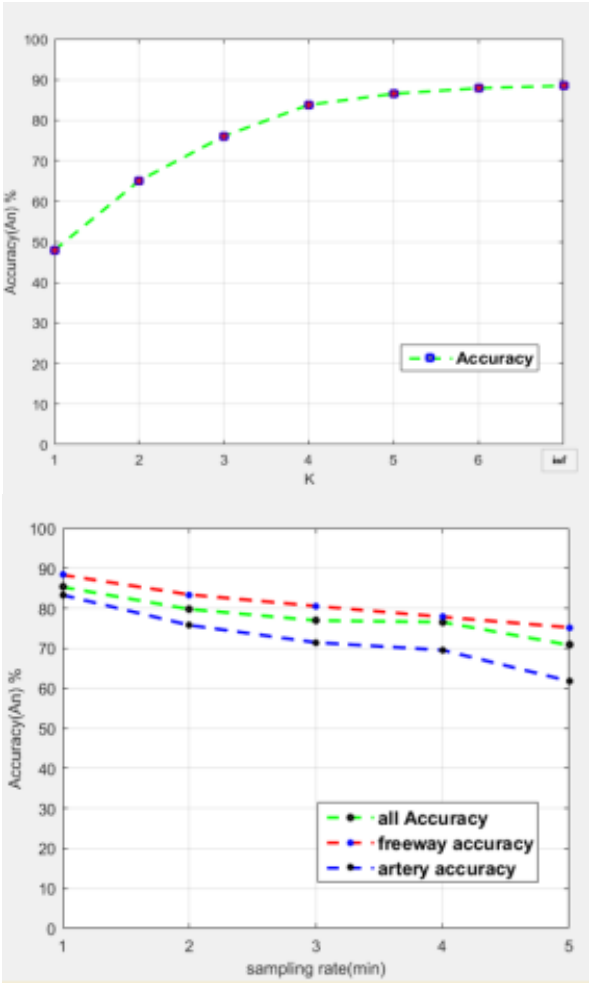
$$A_N = \frac{N_C}{N_A},\ A_L = \frac{L_C}{L_A}$$

رابطه (۹)

با توجه به نمودارهای شکل(۹) مشخص است که در نظر گرفتن ماکزیمم تعداد نقاط کاندیدا ۵ بهینه است. با این فرض مقادیر دقت به‌صورت $A_N = 86.3\%, A_L = 80.12\%$ به‌دست آمد.

دیگر پارامترهای مؤثر بر الگوریتم انتخاب شعاع جستجو کاندیدا و انحراف استاندارد توزیع نرمال است. با انجام آزمایش‌های مختلف و در نظر گرفتن شعاع‌های ۳۰ الی ۱۵۰ متر و انحراف استاندارد ۱۰ الی ۶۰ متر (با توجه به مقدار خطای *GPS* و حداکثر فاصله نقاط از محور راه)، دقت‌هایی در بازه ۷۰٫۴ الی ۸۶٫۳ درصد نتیجه شد. با توجه به تحلیل آزمون‌های انجام‌شده شعاع جستجوی ۵۰ متر و انحراف استاندارد توزیع نرمال ۲۰ متر به‌عنوان بهترین و بهینه‌ترین نتیجه استخراج شد.

یکی دیگر از عوامل مؤثر بر دقت نتایج الگوریتم تطبیق نقشه مکانی - زمانی، قدرت تفکیک یعنی نرخ نمونه‌برداری برداشت *GPS* است. به‌منظور ارزیابی تأثیر نرخ نمونه‌برداری بر دقت نتایج الگوریتم، با حذف نقاط میانی خطوط سیر موردمطالعه (نرخ نمونه‌برداری یک دقیقه) خطوط سیری با قدرت تفکیک زمانی ۲، ۳، ۴ و ۵ دقیقه به‌دست آمد. نتایج اجرای الگوریتم بر خطوط سیر اصلی و خطوط سیر تولیدشده در شکل(۸) بارنگ سبز نمایش داده‌شده است. ارزیابی تأثیر تراکم شبکه در نواحی با تراکم پایین (مانند بزرگراه و راه‌های اصلی) و مناطق شهری متراکم به‌صورت جداگانه مورد ارزیابی قرار گرفت. نتایج نشان می‌دهد تطبیق نقشه خطوط سیر در در نواحی با تراکم پایین نسبت به مناطق پرتراکم شهری کمتر تحت تأثیر نرخ نمونه‌برداری است.



شکل ۸: زمان اجرا و دقت برحسب تعداد نقاط کاندیدا

باتوجه به مطالب مذکور و نیز تنظیم پارامترهای بهینه

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]

می‌توان دریافت دقت نتایج خروجی سرعت دارای ۸۶٫۳ درصد صحت و با دقت میانگین حدود ۱٫۲۲ کیلومتر بر ساعت محاسبه می‌شوند.

۵-۳-ارزیابی سرعت خطوط شبکه

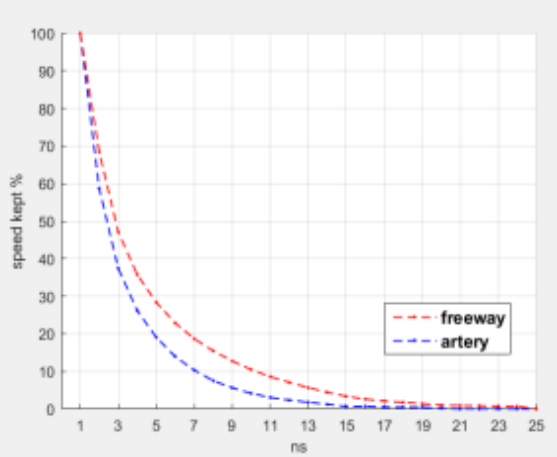
سرعت نهایی هر خط شبکه معابر از میانگین مجموعه سرعت‌های متوسط گذرنده در بازه‌های زمانی مشخص مشروط به در نظر گرفتن حداقل تعداد نمونه سرعت‌های گذرنده از هر خط (ns) در نظر گرفته می‌شود؛ اما سؤال این است که انتخاب ns های مختلف چه تأثیری در دقت نتایج خروجی ایجاد می‌کند؟ برآورد تعداد نمونه‌برداری در تحلیل‌های مهندسی ترافیک پارامتری اساسی و مهم است. قابل‌اطمینان‌ترین روش برآورد تعداد نمونه مناسب فرآیند تکراری نمونه‌برداری و استفاده از انحراف استاندارد است، بنابراین استفاده از آماره – تی (t) به‌دلیل وابستگی به تعداد نمونه و انحراف استاندارد، روشی مناسب است. لذا کمترین تعداد نمونه موردنیاز جهت برآورد سرعت با دقت $\pm\varepsilon$ در سطح اطمینان مشخص با استفاده از رابطه (۱۰) محاسبه می‌شود:

$$n=\left(\frac{t\times s}{\varepsilon}\right)^2 \quad , \quad \varepsilon=\bar{V}-V$$

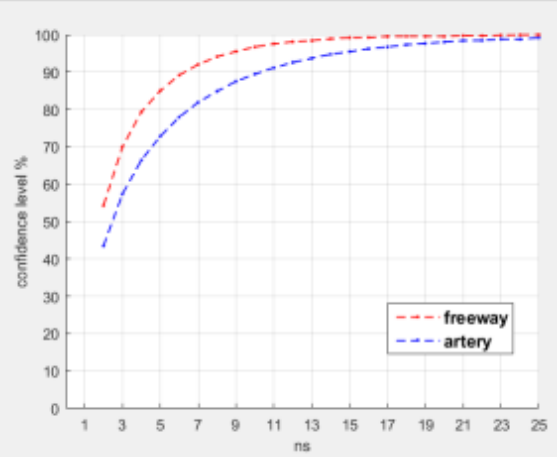
رابطه (۱۰)

در رابطه (۱۰) انحراف استاندارد نمونه (s)، آماره تی برای سطح اطمینان و درجه آزادی مشخص (t)، تعداد نمونه (n)، میانگین سرعت نمونه ($\bar{V}$)، میانگین سرعت جامعه (V)، دقت برآورد (ε) است. با بررسی انحراف استاندارد سرعت‌های برآورد شده در این تحقیق به‌طور پیش‌فرض پارامتر ns تعداد ۴ در نظر گرفته شد. به بیان دیگر حداقل تعداد سرعت‌های متوسطی که در برآورد سرعت نهایی مشارکت دارند ۴ نمونه است.

هر چه مقدار ns بزرگ‌تر در نظر گرفته شود نتایج دقیق‌تری حاصل می‌گردد و از طرفی با انتخاب مقدار بزرگ‌تر ns امکان برآورد خطوط کمتری وجود خواهد داشت و بخشی از اطلاعات از دست می‌رود، شکل الف (۹) تأثیر این پارامتر بر تعداد خطوط قابل برآورد را برای بازه زمانی ۲۰ دقیقه نشان می‌دهد. با توجه به موارد گفته‌شده مطابق شکل ب (۹) و در نظر گرفتن دقت ۶ کیلومتر بر ساعت به‌عنوان دقت مطلوب و مورد انتظار در برآورد سرعت، سطح اطمینان سرعت برآورد شده در بزرگراه بالای ۸۰ درصد و سطح اطمینان درراه‌های شریانی نیز بالای ۶۷ درصد است.



(الف)



(ب)

شکل ۹: تأثیر پارامتر ns بر تعداد خطوط و سطح اطمینان قابل برآورد برای بازه زمانی ۲۰ دقیقه

۶- نتیجه‌گیری

در این مقاله روشی جدید برای جمع‌آوری، تحلیل، استخراج و انتشار اطلاعات ترافیکی با استفاده از

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]

مفاهیم وب ۲ و داده‌های مکانی داوطلبانه توسعه داده‌شده است. این مقاله به‌طور خاص متمرکز بر تعیین پارامتر راهبندان‌های ترافیکی با استفاده از داده‌های خطوط سیر *GPS* است. بدین منظور از داده‌های خطوط سیر *GPS* مربوط به ۱۰۳۵۷ تاکسی در بازه زمانی یک هفته ۲ تا ۸ فوریه سال ۲۰۰۸ در پکن و شبکه معابر ایجادشده از داده‌های *OSM* استفاده شد. در ابتدا داده‌های خام به‌منظور حذف خطاهای موجود وارد مرحله پاک‌سازی‌شده و پس از طی فرآیندی شامل ۵ فیلتر اصلی داده‌های نقاط خارج از محدوده‌ی شهر پکن، نقاط دارای مختصات صفر ناشی از خطای دستگاه *GPS*، نقاط تکراری، نقاط پارکینگ، خطوط سیر با کمتر از ۴ نقطه و نقاطی متوالی با فاصله زمانی بیش از ۴ دقیقه حذف گردید و درنهایت ۷۱۱۶۵۰۳ نقطه مربوط به ۹۷۴۰ خودرو باقی ماند. از طرفی دیگر با انجام تحلیل‌ها و فیلترها شبکه معابر شهر پکن با استفاده از داده‌های *OSM* ایجاد گردید. سپس داده‌های پاک‌سازی‌شده وارد مرحله تطبیق نقشه می‌شوند و به‌منظور منطبق شدن خطوط سیر *GPS* بر شبکه معابر شهری ایجادشده در مرحله قبل با توجه به نرخ نمونه‌برداری پایین داده‌های احتمالی حاصل از سامانه که توسط تاکسی‌ها و شهروندان جمع‌آوری می‌شود از الگوریتم تطبیق نقشه مکانی-زمانی *ST-matching* بهره‌گرفته می‌شود. در مرحله بعد با داده‌های پاک‌سازی‌شده و منطبق بر روی نقشه شبکه معابر، سرعت متوسط بین زوج نقاط متوالی نمونه‌برداری خطوط سیر *GPS* محاسبه می‌شود. پس از حذف سرعت‌های نویز با استفاده از تست آماری برمبنای میانگین و انحراف استاندارد بر روی مجموعه سرعت‌ها در سطح اطمینان ۹۵ درصد از سرعت‌های باقیمانده میانگین‌گیری شود و سرعت نهایی هریال و درنهایت سرعت شبکه تعیین می‌گردد. پس از برآورد سرعت جریان آزاد بر روی هر قطعه از جاده، تشخیص راهبندان‌های ترافیکی در جاده‌ها بر اساس مقایسه سرعت محاسبه‌شده نهایی هر قطعه‌راه و حد آستانه‌های تعریف‌شده بر مبنای سرعت جریان آزاد، در بازه زمانی تعیین خواهد شد. نقشه ترافیک در شبکه معابر شهر پکن برای معابر این شهر به سه رنگ قرمز، آبی و سبز در بازه‌های زمانی مختلف در طول روز به‌دست می‌آید که به ترتیب نمایانگر راهبندان‌های ترافیکی (ترافیک سنگین)، ترافیک کند (نیمه سنگین) و ترافیک روان است.

در این تحقیق جهت برآورد سرعت جریان آزاد از میانگین سرعت مجاز، آلفا درصد (۵ الی ۱۵ درصد) از سریع‌ترین خطوط سیری که از هر قطعه راه عبور کرده استفاده‌شده است. از مهم‌ترین مزایای این روش، بهره‌گیری از داده‌های مکانی داوطلبانه بدون صرف هزینه و در زمان کوتاه، به‌منظور برآورد مقدار سرعت جریان آزاد از داده‌های واقعی مربوط به هر قطعه از راه است. هم‌چنین در این روش، از برآورد سرعت جریان آزاد در چند منطقه محدود و تعمیم آن به‌کل شبکه راه که ناخودآگاه خطایی را به همراه دارد، خودداری می‌شود. نکته‌ی قابل‌توجه دیگر آن است که در این روش به‌طور خودکار اکثراً از خطوط سیری استفاده خواهد شد که در ساعات غیر اوج ترافیکی واقع‌شده‌اند و بدون نیاز به دانش محلی و این امر به دلیل آن است که در الگوریتم مذکور از سریع‌ترین خطوط سیر استفاده می‌شود که این خطوط سیر در ساعات غیر اوج رخ می‌دهد.

برای نمونه نتایج در بازه زمانی ۱۶-۱۶:۲۰ نشان می‌دهد مناطق مرکزی شهر دارای راهبندان‌های ترافیکی، خیابانهای فرعی دارای ترافیک روان و بخش زیادی از بزرگراه‌ها و کمربندی‌های شهر پکن به جز در برخی لینکهای اتصال که وظیفه انتقال ترافیک را دارند نیمه روان بوده‌است. با بررسی وضعیت ترافیک در این بازه زمانی مشخص می‌شود حدودا با نرخ نفوذ ۳٫۸ درصد کل خودروهای سطح شهر مشارکت داشته‌اند و از ۳۲٫۵۶ درصد از کل قطعات شبکه معابر با استفاده از خودروهای شرکت‌کننده عبور انجام پذیرفته است بنابراین می‌توان نتیجه‌گرفت که

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.5.3.51]

با استفاده از مشارکت مناسب کاربران و تولید *VGI* با نرخ نفوذ پایین(۳٫۸٪) ترافیک یک شهر بزرگ در حدود ۷۲۷۲۰ قطعه‌راه را با صرف هزینه پایین و بدون نیاز به تجهیزات با دقت مناسبی تعیین نمود. علی‌رغم مزایای فراوان استراتژی بیان شده برای تعیین راهبندان‌های ترافیکی به‌علت حجم بسیار بالا و زمان بالای محاسبات، استفاده از الگوریتم‌های مبتنی بر رایانش ابری با حفظ حریم خصوصی کاربران پیشنهاد می‌شود که در پژوهش‌های آتی به آن پرداخته می‌شود.

مراجع

[1] Z. Wang, M. Lu, X. Yuan, J. Zhang, and H. Van De Wetering, "Visual traffic jam analysis based on trajectory data", IEEE Transactions on Visualization and Computer Graphics, vol. 19, pp. 2159-2168, 2013.

[2] X. Liu, Z. Wang, Z. Wang, S. Lv, and T. Guan, "A novel real-time traffic information collection system based on smartphone", in China Conference on Wireless Sensor Networks, pp. 291-303, 2013.

[3] Z. Kamoosi, B Rafiee, S. Avalin, M. Charsooghi, M. Navaki, "Collecting traffic real-time data using automotive communication technology", The 17th International Conference on Traffic & Transportation Enineering, Tehran, 2013 (in Persian)

[4] J. C. Herrera, D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, and A. M. Bayen, "Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment", Transportation Research Part C: Emerging Technologies, vol. 18, pp. 568-583, 2010.

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

[6] F. Ostermann and L. Spinsanti, "Context analysis of volunteered geographic information from social media networks to support disaster management: a case study on forest fires," Information Systems for Crisis Response and management, An International Journal, vol. 4, pp. 16-37, 2012.

[7] J. Machay and D. Media. (2013). How does google detect traffic congestion? , http://smallbusiness.chron.com/google-detect-traffic-congestion.

[8] V. Jain, A. Sharma, and L. Subramanian, "Road traffic congestion in the developing world," in Proceedings of the 2nd ACM Symposium on Computing for Development, p. 11,2012.

[9] Y. Yang, Z. Cui, J. Wu, G. Zhang, and X. Xian, "Fuzzy c-means clustering and opposition-based reinforcement learning for traffic congestion identification", Journal of Information and Computer Science, vol. 9, pp. 2441-2450, 2012 .

[10] R. Ong, F. Pinelli, R. Trasarti, M. Nanni, C. Renso and S. Rinzivillo, "Traffic jams detection using flock mining," in Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 650-653, 2011.

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

[12] V. Manolopoulos, S. Tao, S. Rodriguez, M. Ismail, and A. Rusu, "MobiTraS: A mobile application for a smart traffic system", in NEWCAS Conference (NEWCAS), 8th IEEE International, pp. 365-368, 2010 .

[13] B. Krogh, O. Andersen, and K. Torp, "Trajectories for novel and detailed traffic information", in Proceedings of the Third

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]

ACM SIGSPATIAL International Workshop on GeoStreaming, pp. 32-39, 2012.

[14] G. Andrienko, N. Andrienko, C. Hurter, S. Rinzivillo, and S. Wrobel, "From movement tracks through events to places: Extracting and characterizing significant places from mobility data", in Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on, pp. 161-170, 2011.

[15] S. Chawla, Y. Zheng, and J. Hu, "Inferring the root cause in road traffic anomalies", in IEEE 12th International Conference on Data Mining, pp. 141-150, 2012.

[16] W. Zhang, G. Tan, N. Ding, and G. Wang, "Traffic Congestion Evaluation and Signal Timing Optimization Based on Wireless Sensor Networks: Issues, Approaches and Simulation," INFORMATION SCIENCE AND ENGINEERING, vol. 30, pp. 1245-1260, 2014.

[17] F. Giannotti, M. Nanni, D. Pedreschi, F. Pinelli, C. Renso, and S. Rinzivillo, "Unveiling the complexity of human mobility by querying and mining massive trajectory data," The VLDB Journal-The International Journal on Very Large Data Bases, vol. 20(5), pp. 695-719, 2011.

[18] U. Nagaraj, J. Rathod, P. Patil, S. Thakur, and U. Sharma, "Traffic jam detection using image processing," International Journal of Engineering Research and Applications (IJERA), vol. 3, pp. 1087-1091, 2013.

[19] C. Peng, X. Jin, K.-C. Wong, M. Shi, and P. Lio, "Collective Human Mobility Pattern from Taxi Trips in Urban Area," PLOS ONE, vol. 7(8), 2012.

[20] B. Pan, Y. Zheng, D. Wilkie, and C. Shahabi, "Crowd sensing of traffic anomalies based on human mobility and social media," in Proceedings the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Florida,pp. 344-353, 2013.

[21] X. PangLinsey, S. Chawla, W. Liu, and Y. Zheng, "On detection of emerging anomalous traffic patterns using GPS data," Data & Knowledge Engineering, vol. 87, pp. 357-373, 2013.

[22] P. Desai, S. W. Loke, A. Desai, and J. Singh, "CARAVAN: Congestion Avoidance and Route Allocation Using Virtual Agent Negotiation," IEEE Transactions On Intelligent Transportation Systems, vol. 14, pp. 1197-1207, 2013.

[23] E. D'Andrea and F. Marcelloni, "Detection of traffic congestion and incidents from GPS trace analysis," Expert Systems with Applications, vol. 73, pp. 43–56, 2017.

[24] S. Prakash Kaklij, "Mining GPS Data for Traffic Congestion Detection and Prediction " International Journal of Science and Research (IJSR), vol. 47, pp. 876-880, 2015.

[25] F. Osterman and L. Spinsanti, "Context Analysis of Volunteered Geographic Information from Social Media Networks to support disaster management: A Case Study on Forest Fires," Information Systems for Crisis Response and management, An International Journal, vol. 4, pp. 16-37, 2012.

[26] K. Poser and D. Dransch, "Volunteered geographic information for disaster management with application to rapid flood damage estimation", Geomatica, vol. 64, pp. 89-98, 2010.

[27] X. Qian, L. Di, D. Li, P. Li, L. Shi, and L. Cai, "Data cleaning approaches in Web2. 0 VGI application," in 17th International Conference on Geoinformatics, Fairfax, VA, USA, pp. 1-4, 2009 .

[28] D. Birant and A. Kut, "ST-DBSCAN: An algorithm for clustering spatial–temporal data", Data & Knowledge Engineering, vol. 60, pp. 208-221, 2007.

[29] L. Gong, H. Sato, T. Yamamoto, T. Miwa, and T. Morikawa, "Identification of activity

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]

stop locations in GPS trajectories by density-based clustering method combined with support vector machines," Mod. Transport springer, vol. 23(3), pp. 202–213, 2015.

[30] J. Yuan, Y. Zheng, C. Zhang, X. Xie, and G.-Z. Sun, "An interactive-voting based map matching algorithm", in Proceedings of the 2010 Eleventh International Conference on Mobile Data Management, pp. 43-52, 2010 .

[31] [۳۱] D. Wang, Z. Wang, X. Li, and Z. Xiao, "A Map Matching Algorithm to Eliminate Miscalculation Based on Low-Sample-Rate Data", in Proceedings of 2014 International Conference on Computer Science and Service System (CSSS'14), 2014.

[32] [۳۲] Y. Zheng, Y. Lou, C. Zhang, and X. Xie, "Map-Matching for Low-Sampling-Rate GPS Trajectories", in in Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 352-361, 2011 .

[33] S. Syed, "Development of Map-Aided GPS Algorithms for Vehicle Navigation in Urban Canyons," Master's Thesis, Geomatics Engineering, Calgary, 2006.

[34] J J. S. Greenfeld, "Matching GPS observations to locations on a digital map," in Transportation Research Board 81st Annual Meeting, Washington D.C, 2002.

[35] A. Dobash, "Offline Map-Matching for GPS Observations", M.S. Thesis, Sharif University, Tehran, 2014, (in Persian)

[36] R. Frederix, "Dynamic Origin-Destination Matrix Estimation in Large-Scale Congested Networks", Phd, Science and Technology, Catholic University of Leuven, Heverlee, Belgium, 2012.

[37] HCM, "Highway Capacity Manual, Draft Ch. 9/Glossary and Symbols," Washington, D.C: Transportation Research Board, National Research Council, pp. 1-9, 2010.

[38] H. C. Manual, "Transportation research board," National Research Council, Washington, DC, vol. 113, 2000.

[39] M. D. Deardoff, B. N. Wiesner, and J. Fazio, "Estimating free-flow speed from posted speed limit signs", Procedia-Social and Behavioral Sciences, vol. 16, pp. 306-316, 2011.

[40] C. Wang and J. B. Huegy, "Determining the Free-Flow Speeds in a Regional Travel Demand Model Based on the Highway Capacity Manual," in Transportation Research Board 93rd Annual Meeting, Washington DC, pp. 17, 2014.

[41] E. M. Mikhail and F. E. Ackermann, Observations and least squares. Washington, D.C. : University Press of America: University Press of America, 1982 .

[42] M. A. Taylor, J. E. Woolley, and R. Zito, "Integration of the global positioning system and geographical information systems for traffic congestion studies", Transportation Research Part C: Emerging Technologies, vol. 8, pp. 257-285, 2000.

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]

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

Journal of Geospatial Information Technology
Vol.5, No.3, Autumn 2017

Traffic Jams Detection based on the GPS Trajectories Extracted from Volunteered Geographic Information

Vahid Shokri [1], Rahim Ali Abbaspour [2]*

1- MSc of GIS, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran.
2- Assistant Professor, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran

Abstract

Despite the benefits of fixed sensors to collect traffic data such as quality and accessibility, factors such as dependence on equipment, lack of full coverage, the high cost of data collection, management and analysis revealed the need to develop proper infrastructure. Increasing advances in volunteer geographic information (VGI) and location-aware technologies opens new dimension to collect and publish traffic data in high volume with lower costs. Therefore, developing a system to facilitate the collection, extraction, and dissemination of traffic information by utilizing VGI was considered, but given the extent of the issue, this paper focuses on providing strategies to extract traffic jam from GPS trajectories. First in pre-processing stage, trajectories are being cleaned and irrelevant data will be removed. Then ST-matching algorithm GPS trajectories match on road network are being used. After calculating the average speed between each two consecutive points, traffic jams were determined by comparing the free-flow speed and roads network speed. In order to test and evaluate the system, GPS trajectories which is collected voluntarily by taxis between 2nd to 8th of February 2008 in Beijing, is being used. For example, results of traffic jam in the period of 16-16: 20, indicates that most of the main roads and connecting links experiencing traffic jam and this is due to the transfer of traffic from the city center and working areas to residential areas. Results of this paper show that with the appropriate participation of users in producing VGI even with low penetration rate (8/3%) traffic jam could be determined in big cities with a road network about 72720 segments with lower cost and without need of special equipment with an appropriate accuracy.

Key words: traffic data, volunteer geographic information (VGI), GPS trajectories, traffic jam.

Correspondence Address: GIS Department, School of Surveying and Geospatial Engineering,, College of Engineering, University of Tehran, Tehran, Iran.
Tel: +98 21 88008841
Email: abaspour @ut.ac.ir

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.5.3.51]