

اندازه‌گیری تشابه خطوط سیر با استفاده از تئوری فازی

فراز برومند^{۱*}، علی اصغر آل‌شیخ^۲، مهدی فرنقی^۳

^۱ دانشجوی کارشناسی ارشد سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه

نصیرالدین طوسی

f.boroumand@email.kntu.ac.ir

^۲ استاد دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی

alesheikh@kntu.ac.ir

^۳ استادیار دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی

farnaghi@kntu.ac.ir

(تاریخ دریافت آبان ۱۳۹۸، تاریخ تصویب بهمن ۱۳۹۸)

چکیده

با پیشرفت انواع سامانه‌های تعیین موقعیت، امکان دسترسی به حجم وسیعی از داده‌های حرکتی فراهم شده است. از جمله روش‌های کشف دانش از این نوع داده‌ها، اندازه‌گیری تشابه خطوط سیر حاصل از حرکت اشیاء است. همچنین اندازه‌گیری تشابه، بطور مستقیم و غیرمستقیم در سایر روش‌های داده‌کاوی مثل خوشه‌بندی و طبقه‌بندی کاربرد دارد و در حال حاضر به‌عنوان یک موضوع چالش‌برانگیز مورد توجه بسیاری از محققان در حوزه سیستم‌های اطلاعات مکانی قرار گرفته است. باوجود اینکه عدم قطعیت یک مسئله اجتناب‌ناپذیر است، تاکنون توجه کمی به این موضوع در زمینه اندازه‌گیری تشابه خطوط سیر شده است. یک راه مقابله با عدم قطعیت در مشاهدات و تعاریف مسئله، استفاده از تئوری فازی است. در این تحقیق، دو روش $sim1$ و $sim2$ به ترتیب بر پایه روش‌های LCSS و EDR ارائه شده که برای مقابله با عدم قطعیت در اندازه‌گیری تشابه خطوط سیر و بهبود کارایی آن‌ها از تئوری فازی استفاده شده است. روش‌های پیشنهادی با استفاده از یک تابع عضویت فازی و بر اساس فاصله میان نقاط دو خط سیر، درجه انطباق هر دو نقطه از دو خط سیر را تعیین می‌کنند که بر اساس آن تشابه دو خط سیر تعیین می‌شود. به‌منظور ارزیابی این دو روش، دو سری آزمایش بر روی خطوط سیر واقعی و مصنوعی خودروهای شخصی انجام شده است. نتایج آزمایش‌ها نشان می‌دهند که روش‌های $sim1$ و $sim2$ از لحاظ حساسیت نسبت به نویز، کاهش و افزایش نرخ نمونه‌برداری عملکردی مشابه روش‌های LCSS و EDR و از لحاظ حساسیت نسبت به جابجایی، عملکرد بهتری نسبت به آن‌ها دارند. به‌طوری‌که برای مثال میانگین درصد تغییرات تشابه نسبت به تغییرات فاصله برای چهار حد آستانه ۵، ۱۰، ۲۵ و ۵۰ متر برای روش LCSS برابر ۰،۰۲، ۰،۹۷، ۰،۶۶ و ۰،۲۳ است و برای روش‌های $sim1$ و $sim2$ برابر ۰،۴۱ درصد است.

واژگان کلیدی: تشابه خطوط سیر، سیستم اطلاعات مکانی، داده‌کاوی، عدم قطعیت، تئوری فازی

* نویسنده رابط

۱- مقدمه

حرکت از مؤلفه‌های اصلی زندگی انسان‌ها و موجودات می‌باشد. مطالعه حرکت پدیده‌های پویا و الگوهای حرکتی آن‌ها، نقش پررنگی در کشف دانش^۱ و شناخت و تجزیه و تحلیل آن‌ها داشته و در زمینه‌های مختلف مانند کنترل ترافیک، حمل و نقل، زیست‌شناسی و هواشناسی اهمیت بسزایی دارد. یکی از اهداف سیستم‌های اطلاعات مکانی^۲ (GIS) بهبود توانایی‌ها و قابلیت‌های بشر برای درک پدیده‌های جغرافیایی و حل مسائل جغرافیایی است [۲۴]. همچنین مطالعه حرکت و پدیده‌های پویا یک زمینه پژوهشی جدید در سیستم‌های اطلاعات مکانی بشمار می‌آید که می‌توان از ترکیب آن با روش‌های کشف دانش، به عنوان ابزاری قدرتمند برای کاوش داده‌های حرکتی و مطالعه پدیده‌های پویا استفاده کرد [۱۱].

در سال‌های اخیر با پیشرفت سامانه‌های تعیین موقعیت و فناوری‌های همراه امکان جمع‌آوری داده‌های مربوط به حرکت اشیاء و مطالعه خطوط سیر حاصل از آن‌ها بیش‌ازپیش فراهم شده است. اندازه‌گیری تشابه خطوط سیر، یک مرحله مهم در فرآیند کاوش داده‌های حرکتی است؛ با شناسایی خطوط سیر متشابه می‌توان روش‌های داده کاوی متنوعی مثل خوشه‌بندی و طبقه‌بندی را برای کشف الگوهای حرکتی مورد استفاده قرار داد [۳۳]. اندازه‌گیری تشابه در حوزه مطالعات حرکت به دلیل نقش آن در درک و پیش‌بینی رفتار اشیاء و پدیده‌های پویا و استخراج الگوهای حرکتی آن‌ها، به‌عنوان یک موضوع مهم و چالش‌برانگیز موردتوجه بسیاری از محققان قرار گرفته و روش‌های مختلفی برای اندازه‌گیری تشابه خطوط سیر در ابعاد مختلف مکانی، زمانی و معنایی ارائه شده است [۲، ۲۷].

در زمینه اندازه‌گیری تشابه خطوط سیر، مسئله‌ای که کمتر به آن پرداخته شده، عدم قطعیت است؛ در حالی که عدم قطعیت یکی از ویژگی‌های اجتناب‌ناپذیر در سامانه‌های اطلاعاتی است. در مجموع، شناخت و اطلاعات انسان راجع به دنیا همواره با ابهاماتی همراه بوده است. این ابهامات منجر به ایجاد عدم قطعیت می‌شود. اهمیت این موضوع در حوزه سیستم‌های اطلاعات مکانی به‌مراتب

پررنگ‌تر نیز می‌شود؛ زیرا اطلاعات مکانی رابطه نزدیکی با مشاهدات از دنیا دارد و این مشاهدات ذاتاً دارای ابهام و عدم قطعیت هستند. بنابراین رویکرد ما در مواجهه با مسائل مختلف باید به‌گونه‌ای باشد که این ابهامات کمترین تأثیر را بر نتایج داشته باشند [۱۴، ۱۸، ۱۹].

عدم قطعیت را از دو منظر می‌توان بررسی کرد. دیدگاه اول به ابهام در تعریف مفاهیم مختلف اشاره دارد [۳۴]. برای مثال در اندازه‌گیری تشابه دو خط سیر، حداقل فاصله برای اینکه دو خط سیر را مشابه در نظر بگیریم و یا حتی اینکه شباهت به چه معناست و بر چه اساس می‌توان دو خط سیر را مشابه در نظر گرفت، از مواردی هستند که نمی‌توان پاسخی دقیق و صریح برای آن‌ها یافت. دیدگاه دوم مربوط به مشاهدات است. یعنی حتی اگر ابهامات موجود در تعریف مفاهیم مسئله را نادیده بگیریم، با توجه به اینکه هیچ‌یک از ابزارهای اندازه‌گیری عاری از خطا نیستند، بنابراین همچنان با مشکل عدم قطعیت در مشاهدات مواجه هستیم [۳۵]. همچنین داده‌های خطوط سیر معمولاً توسط افراد مختلف و با ابزارها و در شرایط متفاوت جمع‌آوری می‌شوند که دقت و نرخ نمونه‌برداری هر یک از آن‌ها متفاوت از دیگری است. یک راهکار مناسب برای حل این مشکل، استفاده از تئوری فازی است [۳].

انسان بسیاری از مفاهیم در جهان واقعی را به‌صورت فازی درک و استفاده می‌کند. مفاهیمی مثل کوتاه و بلند، سرد و گرم و تشابه و عدم تشابه. در تئوری فازی هریک از این مفاهیم را می‌توان بر اساس توابع عضویت تعریف کرد و در بازه پیوسته بین صفر و یک ارزش‌گذاری کرد [۳۴]. یکی از چالش‌هایی که در اندازه‌گیری تشابه خطوط سیر با آن سروکار داریم، تعیین منطبق بودن نقاط دو خط سیر است. روش‌های LCSS^۳ و EDR^۴ از یک حد آستانه برای این منظور استفاده می‌کنند درحالی‌که می‌توان این مفهوم را بر اساس تئوری فازی و با استفاده از توابع عضویت تعریف کرد. در این مقاله دو روش اندازه‌گیری تشابه بر مبنای روش‌های LCSS و EDR ارائه می‌شود و برای این کار از تئوری فازی برای رفع نواقص آن‌ها استفاده خواهد شد.

بنابراین هدف اصلی این مقاله، ارائه دو روش اندازه‌گیری تشابه خطوط سیر بر پایه تئوری فازی برای

^۳ Longest Common SubSequence
^۴ Edit Distance on Real sequence

^۱ Knowledge Discovery
^۲ Geographic Information Science

می‌یابد که منجر به کاهش آلودگی هوا و ترافیک نیز می‌شود [۸]. در تحقیقی دیگر، [۱۵] روشی بر اساس ارتباط میان کاربری زمین و فعالیت‌های انسان برای شناسایی کاربری زمین در مناطق شهری ارائه داده‌اند. این روش، با استفاده از خط سیر افراد و کشف الگوهای حرکتی آن‌ها، نقاط توقف افراد را استخراج و کاربری آن‌ها را تعیین می‌کند. به‌عنوان مثالی دیگر، در حوزه زیست‌شناسی، با اندازه‌گیری تشابه خطوط سیر پرندگان، امکان مطالعه دقیق در مورد تأثیر شرایط جوی مثل سرعت باد بر رفتار آن‌ها، فراهم می‌شود [۴]. همچنین با اندازه‌گیری تشابه مسیر حرکت طوفان‌ها، می‌توان مسیر طوفان‌های بعدی را پیش‌بینی کرده و در نتیجه از بروز خسارات مالی و جانی طوفان‌ها جلوگیری کرد [۱۰].

[۱۹] ارتباط میان عدم قطعیت و مقیاس زمانی را با استفاده از خطوط سیر حیوانات اهلی بررسی کرده‌اند. آن‌ها در این تحقیق نشان داده‌اند که پارامترهای حرکت مثل سرعت و زاویه چرخش، وابسته به نرخ نمونه‌برداری و دقت اندازه‌گیری نقاط خط سیر و همچنین روش مورد استفاده در پردازش این داده‌ها هستند.

همان‌طور که گفته شد، داده‌های خطوط سیر هیچ‌گاه کاملاً دقیق و عاری از خطا نیستند. به همین دلیل، گروهی از محققان روش‌هایی برای پیش‌پردازش داده‌های خطوط سیر به‌منظور کاهش عدم قطعیت پیشنهاد داده‌اند. برای مثال، [۳۶] الگوریتم ^۲HRIS را معرفی کرده‌اند. این الگوریتم حرکت اشیاء را محدود به فضای شبکه فرض می‌کند و از اطلاعات استخراج‌شده از شبکه راه‌ها برای تخمین موقعیت شی در هر لحظه و بهبود نرخ نمونه‌برداری استفاده می‌کند. بنابراین فقط برای اشیاء متحرکی که در فضای محدود شبکه حرکت می‌کنند (مثل خودروها) و تنها برای مقابله با عدم قطعیت ناشی از نرخ نمونه‌برداری پایین مناسب است. در تحقیقی دیگر [۲۰] الگوریتم TrajEstU را برای تخمین خط سیر واقعی یک شی متحرک با در نظر گرفتن عدم قطعیت ناشی از خطای اندازه‌گیری و نرخ نمونه‌برداری پایین معرفی کرده‌اند. از نقاط مثبت این الگوریتم می‌توان به تعیین نرخ نمونه‌برداری و تعداد نقاط خط سیر تخمینی توسط کاربر اشاره کرد. محدودیت اصلی این روش اما نیاز به یک پایگاه داده خطوط سیر برای تعیین

حل مشکلات مربوط به ناهمگونی و عدم قطعیت موجود در داده‌ها است. روش‌های پیشنهادی بر اساس فاصله نقطه به نقطه المان‌های دو خط سیر است و ضمن تعریف یک تابع عضویت فازی بر مبنای فاصله میان نقاط و تعیین درجه انطباق آن‌ها، یک ماتریس تشابه ایجاد می‌شود که بر اساس آن تشابه دو خط سیر تعیین می‌شود. به‌منظور ارزیابی و مقایسه روش‌های پیشنهادی با سایر روش‌ها، از دو مجموعه داده خطوط سیر واقعی و مصنوعی استفاده شده است.

ساختار مقاله در ادامه به این شرح است. در بخش ۲ به مروری بر تحقیقات انجام‌شده در زمینه اندازه‌گیری تشابه خطوط سیر و راه‌های مقابله با عدم قطعیت در اندازه‌گیری تشابه خطوط سیر پرداخته می‌شود. در بخش ۳ مفاهیم پایه در اندازه‌گیری تشابه خطوط سیر بیان می‌شود و برخی از روش‌های موجود اندازه‌گیری تشابه خطوط سیر معرفی و به بررسی نقاط قوت و ضعف هریک از آن‌ها پرداخته می‌شود. در بخش ۴، روش‌های پیشنهادی توسعه داده‌شده بر مبنای تئوری فازی ارائه می‌شوند. پیاده‌سازی و ارزیابی روش‌های توسعه داده‌شده در بخش ۵ انجام می‌پذیرد و در نهایت در بخش ۶ نتیجه‌گیری و پیشنهاد برای تحقیقات آتی ارائه می‌گردد.

۲- مروری بر تحقیقات پیشین

در این بخش، به مروری بر برخی از تحقیقات مرتبط با موضوع این مقاله پرداخته می‌شود. ابتدا چند نمونه از کاربردهای عملی اندازه‌گیری تشابه خطوط سیر در زمینه‌های مختلف و به‌ویژه GIS معرفی می‌شوند سپس تحقیقاتی که به بررسی نقش عدم قطعیت در اندازه‌گیری تشابه خطوط سیر و راه‌های مقابله با آن پرداخته‌اند را مرور می‌کنیم.

موارد زیادی از کاربرد مطالعات حرکت و اندازه‌گیری تشابه خطوط سیر در زمینه‌های مختلف وجود دارد. برای مثال در حوزه حمل‌ونقل، افراد با خط سیر متشابه می‌توانند از یک سیستم سواری اشتراکی^۱ برای به اشتراک‌گذاری وسیله نقلیه شخصی استفاده کنند. به این ترتیب هزینه سفر برای افراد کاهش می‌یابد، علاوه بر این تعداد خودروهای تک‌سرنشین در سطح شهر کاهش

^۲ History based Route Inference System

^۱ Carpooling/Ride-Sharing

الگوی حرکت و سپس تخمین موقعیت نقاط با استفاده از مدل رگرسیون خطی است.

روش‌های مختلفی برای پیش‌پردازش خطوط سیر و کاهش عدم قطعیت ناشی از خطای اندازه‌گیری و نرخ نمونه‌برداری پایین در مقالات وجود دارد. اما با استفاده از این روش‌ها نمی‌توان عدم قطعیت را به‌طور کامل حذف کرد. بنابراین، علاوه بر استفاده از روش‌های پیش‌پردازش خطوط سیر، ضروری است که در اندازه‌گیری تشابه خطوط سیر نیز عدم قطعیت را در نظر گرفت.

برای این منظور [۳۲] با ارائه روش فاصله اقلیدسی مورد انتظار^۱ (EE-distance)، فاصله میان خطوط سیر را اندازه‌گیری می‌کنند. این روش از واریانس و امید ریاضی برای مدل‌سازی عدم قطعیت استفاده می‌کند. [۲۶] نیز روش TIDE^۲ را برای اندازه‌گیری تشابه خطوط سیر با استفاده از حداکثر سرعت تخمینی پیشنهاد کرده‌اند. ویژگی این روش این است که تشابه دو خط سیر را بجای یک عدد، با استفاده از یک بازه توصیف می‌کند. روش UMS^۳ نیز که توسط [۱۲] معرفی شده است، از نمایش بیضوی خطوط سیر برای تعیین حد آستانه پویا، میان نقاط استفاده می‌کند. در نهایت اندازه تشابه بر اساس سه معیار شکل هندسی^۴، محدوده مشترک^۵ و پیوستگی^۶ بیضوی‌های دو خط سیر محاسبه می‌شود.

جنبه مشترک روش‌های معرفی شده در بالا، استفاده از پارامترهایی مثل واریانس، میانگین و سرعت در مدل‌سازی عدم قطعیت است. این در حالی است که این پارامترها خود تحت تأثیر عدم قطعیت هستند [۱۹] و لذا نمی‌توانند به‌طور کامل پاسخگوی این مسئله باشند. بنابراین ما در این مقاله، از تئوری فازی برای حل مسئله عدم قطعیت در اندازه‌گیری تشابه خطوط سیر استفاده خواهیم کرد.

۳- مفاهیم پایه

سری‌های زمانی ساده‌ترین روش برای نمایش تغییرات یک پدیده در طول زمان هستند. یک خط سیر را می‌توان به‌عنوان یک سری زمانی شامل دنباله‌ای متناهی از

موقعیت‌های مکانی برای هر گام زمانی در نظر گرفت [۵]. خط سیر A شامل مجموعه‌ای از m نقطه دارای k بعد به‌صورت زیر نمایش داده می‌شود.

$$A = \{a_1, a_2, \dots, a_m\} \quad (1)$$

$$a_i = (t_i, x_i, y_i, \dots)$$

که t_i برچسب زمانی هر نقطه و $t_i < t_{i+1}$ است. در این مقاله برای پرهیز از پیچیدگی‌های غیرضروری، خطوط سیر را در فضای دوبعدی در نظر می‌گیریم و مختصات هر نقطه a_i را با $(x_{a,i}, y_{a,i})$ نشان می‌دهیم. همچنین برای خط سیر A دارای m نقطه، تابع $\text{Head}(A)$ به‌صورت زیر تعریف می‌شود.

$$\text{Head}(A) = \{a_1, a_2, \dots, a_{m-1}\} \quad (2)$$

همان‌طور که گفته شد اندازه‌گیری تشابه به‌طور مستقیم و غیرمستقیم در پیش‌بینی، تشخیص ناهنجاری، خوشه‌بندی و طبقه‌بندی خطوط سیر کاربرد دارد و مطالعات زیادی در این زمینه انجام شده و روش‌های زیادی برای اندازه‌گیری تشابه میان خطوط سیر ارائه شده است. بنابراین برخی از روش‌های موجود را در بخش‌های بعدی مرور و مزیت‌ها و معایب هر یک را بیان می‌کنیم.

۳-۱- فاصله مینکوفسکی^۷

یکی از توابع مورد استفاده در اندازه‌گیری تشابه میان دو خط سیر فاصله مینکوفسکی است. فاصله مینکوفسکی از درجه p ($p \geq 0$) میان دو خط سیر A و B با طول n به‌صورت زیر تعریف می‌شود:

$$D(A, B) = \sum_{i=1}^n [(x_{a,i} - x_{b,i})^p + (y_{a,i} - y_{b,i})^p]^{1/p} \quad (3)$$

در تجزیه و تحلیل خطوط سیر معمولاً از $p = 1$ یعنی فاصله منهتن^۸ (MaD) یا $p = 2$ یعنی فاصله اقلیدسی^۹ (EuD) استفاده می‌کنند. محاسبه فاصله مینکوفسکی بسیار ساده است. همچنین این تابع متریک است و نیاز به تعیین پارامتر خاصی برای انجام محاسبات ندارد. از معایب این روش می‌توان به حساسیت بالا نسبت به نویز اشاره کرد. همچنین نرخ نمونه‌برداری و تعداد نقاط دو خط سیر باید باهم برابر باشد [۲۳، ۲۱].

^۱ Expected Euclidean distance
^۲ Trajectory Interval Distance Estimation
^۳ Uncertain Movement Similarity
^۴ A likeness
^۵ Shareness
^۶ Continuity

^۷ Minkowski Distance
^۸ Manhattan Distance
^۹ Euclidean Distance

۲-۳- DTW

این تابع برای پیدا کردن کوتاه‌ترین فاصله میان دو خط سیر از انتقال زمان محلی^۲ استفاده می‌کند. یعنی نقاط دو خط سیر را به‌گونه‌ای در طول محور زمان جابجا می‌کند تا مجموع فواصل میان نقاط متناظر کمینه شود [۲۲]. فاصله DTW برای دو خط سیر A و B با طول m و n از رابطه ۴ محاسبه می‌شود.

$$DTW(A, B) = \begin{cases} 0 & m = n = 0, \\ \infty & m = 0, n > 0 \text{ or } m > 0, n = 0, \\ \|a_m - b_n\| + \min \begin{cases} DTW(\text{Head}(A), \text{Head}(B)), \\ DTW(\text{Head}(A), B), \\ DTW(A, \text{Head}(B)). \end{cases} & \end{cases} \quad (۴)$$

که منظور از $\|a_m - b_n\|$ فاصله اقلیدسی میان نقطه mام خط سیر A و نقطه nام خط سیر B است. در این روش یک نقطه از یک خط سیر، می‌تواند با چندین نقطه از خط سیر دیگری منطبق شود. بنابراین این روش برای مقایسه خطوط سیر با طول و نرخ نمونه‌برداری متفاوت مناسب است. اما با توجه به اینکه اندازه آن، از مجموع فاصله‌های نقاط منطبق بر هم محاسبه می‌شود و هر نقطه از خط سیر اول، باید حداقل بر یک نقطه از خط سیر دوم منطبق شود، حساسیت بالایی نسبت به نویز دارد. همچنین این تابع متریک نیست [۱۶، ۱۷].

۳-۳- LCSS

همان‌طور که از نام تابع نیز پیداست، این تابع تشابه میان دو خط سیر را بر اساس طول بزرگ‌ترین زیر دنباله منطبق بر هم محاسبه می‌کند. در این روش همانند روش DTW خطوط سیر می‌توانند در طول محور زمان کشیده شوند؛ درحالی‌که برخلاف آن، همه نقاط دو خط سیر لزوماً بر هم منطبق نمی‌شوند. این ویژگی باعث مقاومت تابع در برابر نویز می‌شود. همچنین این تابع برای خطوط سیر با طول و نرخ نمونه‌برداری متفاوت مناسب است [۲۹]. اندازه بزرگ‌ترین زیر دنباله مشترک میان دو خط سیر A و B با طول m و n و با حد آستانه ε از رابطه ۵ محاسبه می‌شود.

$$LCSS_{\varepsilon}(A, B) = \begin{cases} 0 & m = 0 \text{ or } n = 0, \\ 1 + LCSS_{\varepsilon}(\text{Head}(A), \text{Head}(B)) & \|a_m - b_n\| \leq \varepsilon, \\ \max \begin{cases} LCSS_{\varepsilon}(\text{Head}(A), B) \\ LCSS_{\varepsilon}(A, \text{Head}(B)) \end{cases} & \|a_m - b_n\| > \varepsilon. \end{cases} \quad (۵)$$

^۱ Dynamic Time Wrapping
^۲ Local Time Shifting

چالش اصلی در این روش، پیدا کردن حد آستانه مناسب برای تعیین نقاط منطبق است. زیرا تغییر حد آستانه می‌تواند باعث تغییر در نتایج اندازه‌گیری تشابه شود [۱، ۳۰]. علاوه بر این فواصل میان زیر دنباله‌های مشابه از دو خط سیر، تأثیری در اندازه تشابه کلی دو خط سیر ندارد [۷].

۳-۴- EDR

برای دو خط سیر A و B فاصله EDR، معادل تعداد عملیات موردنیاز برای تبدیل A به B است. این عملیات شامل حذف، اضافه و جایگزینی نقاط است. فاصله EDR میان دو خط سیر A و B به‌صورت زیر تعریف می‌شود.

$$EDR_{\varepsilon}(A, B) = \begin{cases} n = 0, & m = 0, \\ \min \begin{cases} subcost(a_m, b_n) + EDR_{\varepsilon}(\text{Head}(A), \text{Head}(B)) \\ 1 + EDR_{\varepsilon}(\text{Head}(A), B) \\ 1 + EDR_{\varepsilon}(A, \text{Head}(B)) \end{cases} & \end{cases} \quad (۶)$$

$$subcost(a_m, b_n) = \begin{cases} 0 & \|a_m - b_n\| \leq \varepsilon, \\ 1 & \|a_m - b_n\| > \varepsilon. \end{cases} \quad (۷)$$

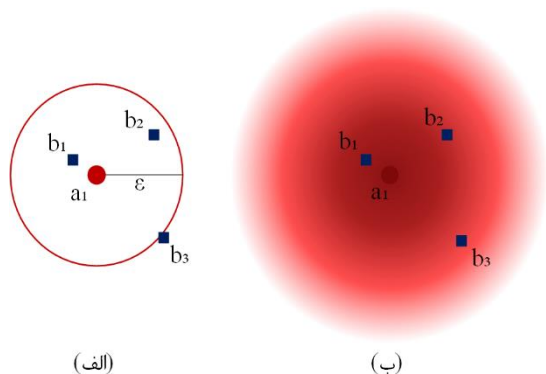
این روش نیز مانند روش LCSS از یک حد آستانه ε برای تعیین منطبق بودن نقاط استفاده می‌کند. بنابراین در مقایسه با روش‌های DTW و فاصله اقلیدسی حساسیت کمتری نسبت به نویز دارد. روش EDR هزینه‌ای برای نقاط غیر منطبق بر مبنای فاصله آن‌ها و حد آستانه تعیین‌شده در نظر می‌گیرد (رابطه ۷) که باعث بهبود دقت آن در مقایسه با روش LCSS می‌شود. هرچند همانند روش LCSS وابستگی نتایج نسبت به حد آستانه از معایب این روش است [۷، ۲۹].

۳-۵- ERP

روش ERP از ترکیب روش‌های EDR و فاصله منهتن ارائه‌شده است [۹] تا از مزیت‌های هر دو روش بهره‌برد. روش EDR از انتقال زمان محلی پشتیبانی می‌کند اما متریک نیست. در مقابل فاصله منهتن متریک است ولی از انتقال زمان محلی پشتیبانی نمی‌کند. به‌این ترتیب ERP روشی متریک است و علاوه بر آن از انتقال زمان محلی نیز

^۳ Edit Distance with Real Penalty

مشاهدات تأثیر کمی بر درجه انطباق نقاط و درنهایت اندازه کلی تشابه دو خط سیر خواهد داشت.

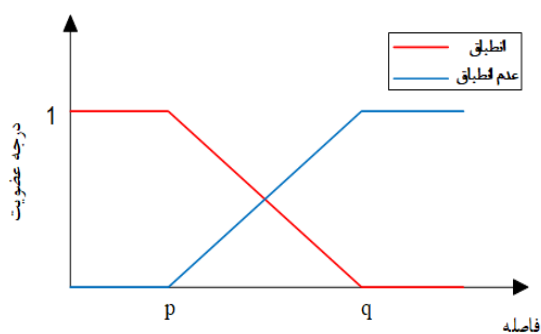


شکل ۱- نمایش وضعیت انطباق نقاط دو خط سیر الف) با استفاده از حد آستانه ب) با استفاده از تابع عضویت فازی

مزیت دیگر استفاده از تابع عضویت در تعریف واژه انطباق است. موضوع بحث‌برانگیز در اینجا تعیین رابطه میان فاصله نقاط و درجه انطباق است که با توجه به نوع خطوط سیر و کاربرد آن تغییر می‌کند؛ اما به‌طور کلی می‌توان گفت که هرچه فاصله میان نقاط کمتر باشد، درجه انطباق نقاط بیشتر است و برعکس. بر این اساس توابع مختلفی را می‌توان تعریف کرد. به‌عنوان مثال تابع عضویت برای دو نقطه a_i و b_j از دو خط سیر A و B بر اساس فاصله میان آن‌ها $(Dist(a_i, b_j))$ را می‌توان با استفاده از رابطه ۹ تعریف کرد. همچنین تابع عدم انطباق با استفاده از رابطه ۱۰ تعریف می‌شود. شکل ۲ نمودار دو تابع موردنظر را نشان می‌دهد.

$$f_{\text{matching}}(a_i, b_j) = \begin{cases} 0 & \text{if } Dist(a_i, b_j) \geq q, \\ 1 & \text{if } Dist(a_i, b_j) \leq p, \\ \frac{q - Dist(a_i, b_j)}{q - p} & p < Dist(a_i, b_j) < q \text{ otherwise.} \end{cases} \quad (9)$$

$$f_{\text{mismatching}}(a_i, b_j) = 1 - f_{\text{matching}}(a_i, b_j) \quad (10)$$



شکل ۲- تابع عضویت انطباق و عدم انطباق

پشتیبانی می‌کند. از معایب این روش حساسیت بالا نسبت به نویز است. فاصله ERP برای دو خط سیر A و B با طول m و n با استفاده از رابطه ۸ محاسبه می‌شود.

$$ERP(A, B) = \begin{cases} \sum_{i=1}^n |b_i - g| & m = 0, \\ \sum_{i=1}^m |a_i - g| & n = 0, \\ \min \begin{cases} \|a_m - b_n\| + ERP(Head(A), Head(B)) \\ \|a_m - g\| + ERP(Head(A), B) \\ \|b_n - g\| + ERP(A, Head(B)) \end{cases} & \text{otherwise} \end{cases} \quad (8)$$

در روش EDR هزینه برای نقاطی که بر هم منطبق نیستند برابر مقدار ثابت ۱ است اما در روش ERP این هزینه بر اساس فاصله تا نقطه g با مقدار پیش‌فرض $g = (0,0)$ محاسبه می‌شود [۶].

۴- روش پیشنهادی

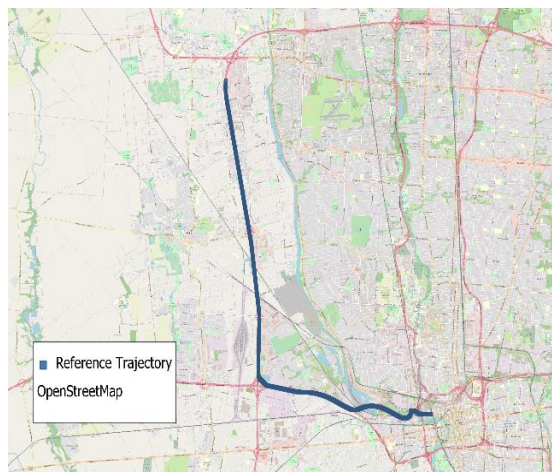
همان‌طور که گفته شد، هر خط سیر از دنباله‌ای از نقاط پشت سر هم تشکیل شده است. در روش‌های پیشنهادی، همه نقاط تشکیل‌دهنده دو خط سیر به‌صورت دوبه‌دو با یکدیگر مقایسه می‌شوند و میزان انطباق و عدم انطباق هر دو نقطه با استفاده از توابع عضویت تعریف شده محاسبه می‌شود. درواقع بجای استفاده از حد آستانه برای تعیین منطبق بودن یا نبودن نقاط، از این توابع برای تعیین درجه منطبق بودن آن‌ها استفاده می‌کنیم.

شکل ۱ وضعیت سه نقطه $b_1, b_2, b_3 \in B$ را نسبت به نقطه $a_1 \in A$ نشان می‌دهد. در روش‌هایی که از حد آستانه استفاده می‌کنند (شکل ۱- الف) وضعیت دو نقطه b_1 و b_2 نسبت به نقطه a_1 یکسان در نظر گرفته می‌شود و تأثیر آن‌ها در اندازه کلی تشابه یکسان خواهد بود. اگرچه فاصله میان a_1 و b_1 کمتر از فاصله میان a_1 و b_2 است. همچنین نقطه b_3 نسبت به نقطه a_1 وضعیت مبهمی دارد زیرا فاصله میان آن‌ها تقریباً برابر حد آستانه (ϵ) است و با در نظر گرفتن خطای مشاهدات، نمی‌توان گفت که دو نقطه a_1 و b_3 بر هم منطبق هستند یا خیر. مزیت اصلی استفاده از تابع عضویت (شکل ۱- ب) این است که می‌توان میزان انطباق نقاط را بر اساس فاصله میان آن‌ها تعیین کرد. در این حالت به دلیل ماهیت پیوسته تابع عضویت، خطای

خطوط سیر جدید بر مبنای خط سیر مرجع استفاده کرد [۳۱]. از آنجایی که میزان تغییرات اعمال شده بر خط سیر مرجع با استفاده از پارامترهای این تبدیلات قابل کنترل هستند، بنابراین می توان مقایسه بهتر و مؤثرتری را نسبت به استفاده از یک مجموعه داده خط سیر واقعی داشت. بنابراین، در این تحقیق به منظور ارزیابی روش های ارائه شده و بررسی جنبه های مختلف آن ها در اندازه گیری تشابه خطوط سیر، دو سری آزمایش بر روی داده های واقعی و مصنوعی، طراحی و پیاده سازی شده است. سری اول این آزمایش ها که با استفاده از داده های مصنوعی انجام می شود، روش های مدنظر را از جنبه های حساسیت نسبت به نویز، تغییر نرخ نمونه برداری و جابجایی، مورد ارزیابی قرار می دهند. سری دوم، با استفاده از داده های واقعی انجام می شود و نتایج اندازه گیری تشابه با استفاده از روش های معرفی شده در این تحقیق، با نتایج حاصل از روش های LCSS و EDR مقایسه می شوند. برای این کار نتایج حاصل از این چهار روش با استفاده از طول خط سیر کوچک تر در بازه بین صفر تا یک نرمال سازی شده اند [۲۹].

۵-۱- داده های مورد نیاز

در این تحقیق از یک مجموعه داده خط سیر خودروهای شخصی استفاده شده است [۲۵] که شامل ۴۹ خط سیر با نرخ نمونه برداری یک ثانیه می باشد. همچنین به منظور ایجاد خطوط سیر مصنوعی، یکی از این خطوط به عنوان خط سیر مرجع در نظر گرفته شده است. این خط سیر دارای ۱۱۶۸ نقطه است که در شکل ۳ نشان داده شده است. به منظور ایجاد خطوط سیر جدید چهار تبدیل در نظر گرفته شده است که در ادامه این تبدیلات، پارامترهای مورد نیاز و نحوه عملکرد آن ها توضیح داده خواهد شد.



شکل ۳- خط سیر مرجع

بنابراین در روش اول که بر مبنای روش LCSS توسعه داده شده است، اندازه تشابه دو خط سیر A و B با استفاده از رابطه ۱۰ محاسبه می شود.

$$\text{sim1}(A, B) = \begin{cases} 0 & \text{if } m = 0 \text{ or } n = 0, \\ \max \begin{cases} f_{\text{matching}}(a_i, b_j) + \text{sim1}(\text{Head}(A), \text{Head}(B)), \\ \text{sim1}(\text{Head}(A), B), \\ \text{sim1}(A, \text{Head}(B)). \end{cases} & \text{otherwise.} \end{cases} \quad (10)$$

این روش مشابه روش LCSS، از یک ساختار بازگشتی استفاده می کند و همه نقاط دو خط سیر، به صورت دوجه دو با یکدیگر مقایسه می شوند. به این ترتیب درجه انطباق هر دو نقطه با هم، تعیین می شود. هر نقطه از یک خط سیر می تواند درجه انطباق یکسانی با چند نقطه از خط سیر دیگری داشته باشد. بنابراین با توجه به اینکه هدف تعیین بیشترین تشابه دو خط سیر است، دو نقطه ای منطبق در نظر گرفته می شوند که اندازه کلی تشابه بیشینه شود. و همچنین در روش دوم که بر اساس روش EDR توسعه داده شده است، فاصله دو خط سیر با استفاده از رابطه ۱۱ به دست می آید.

$$\text{sim2}(A, B) = \begin{cases} m & \text{if } m = 0, \\ n & \text{if } n = 0, \\ \min \begin{cases} f_{\text{mismatching}}(a_i, b_j) + \text{sim2}(\text{Head}(A), \text{Head}(B)), \\ \text{sim2}(\text{Head}(A), B) + 1, \\ \text{sim2}(A, \text{Head}(B)) + 1. \end{cases} & \text{otherwise.} \end{cases} \quad (11)$$

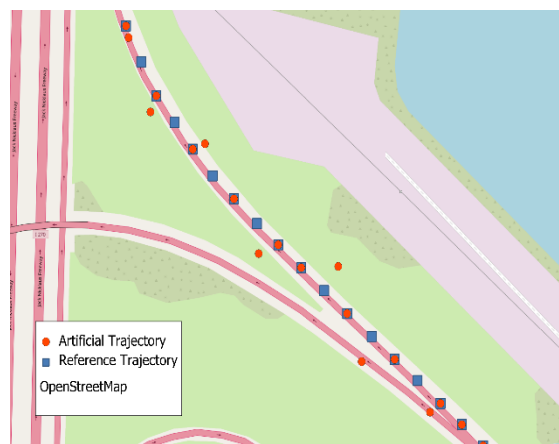
در این روش نیز، از یک ساختار بازگشتی استفاده می شود با این تفاوت که در این روش، درجه عدم انطباق نقاط دو خط سیر محاسبه می شود و هدف تعیین کمترین فاصله میان دو خط سیر است. بنابراین دو نقطه ای منطبق در نظر گرفته می شوند که فاصله کلی میان دو خط سیر کمینه شود.

۵- ارزیابی تجربی

با توجه به اینکه یک مجموعه داده معیار برای ارزیابی و مقایسه روش های اندازه گیری تشابه به صورت کنترل شده وجود ندارد، تنها می توان مقیاس پذیری این توابع را ارزیابی کرد و تعیین عوامل تأثیرگذار بر اندازه تشابه بسیار دشوار خواهد بود [۱۳]. برای حل این مشکل و بررسی اثربخشی توابع مورد بحث در کنار مقیاس پذیری، می توان از یک خط سیر مرجع و یک سری از تبدیلات برای ایجاد

۵-۱-۱- جابجایی نقاط

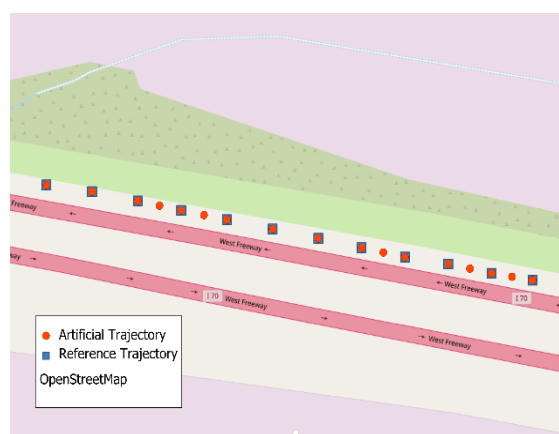
در این تبدیل تعدادی از نقاط خط سیر مرجع با استفاده از نرخ تغییرات موردنیاز به صورت تصادفی انتخاب شده و هرکدام به اندازه‌ای که توسط پارامتر فاصله مشخص می‌شود، در یک جهت تصادفی جابجا می‌شوند. شکل ۴ نحوه جابجا کردن نقاط با استفاده از این تبدیل را نشان می‌دهد.



شکل ۴- عملکرد تبدیل جابجایی نقاط

۵-۱-۲- افزایش نرخ نمونه‌برداری

در این تبدیل با استفاده از پارامتر نرخ تغییرات، تعداد نقاطی که باید به خط سیر مرجع اضافه شوند، تعیین می‌شود و به طور تصادفی میان نقاط خط سیر مرجع قرار می‌گیرند. شکل ۵ نحوه عملکرد این تبدیل را نشان می‌دهد.

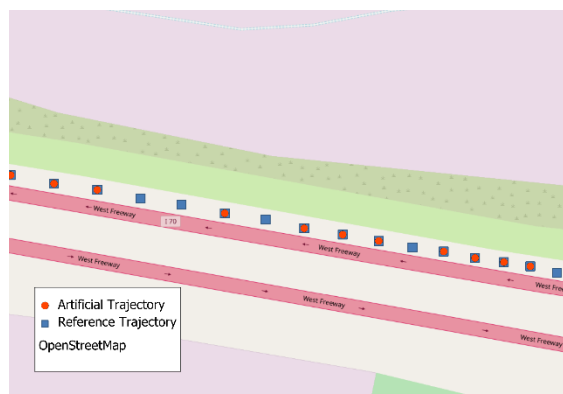


شکل ۵- عملکرد تبدیل افزایش نرخ نمونه‌برداری

۵-۱-۳- کاهش نرخ نمونه‌برداری

این تبدیل مشابه تبدیل افزایش نرخ نمونه‌برداری است، با این تفاوت که در این تبدیل، نقاط خط سیر

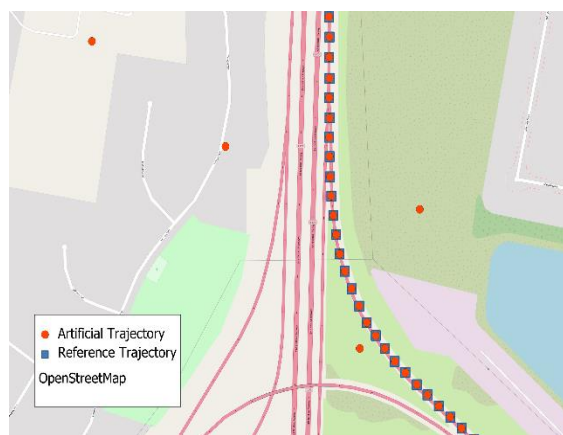
مرجع به تعداد موردنظر و به طور تصادفی حذف می‌شوند. شکل ۶ مثالی از نحوه عملکرد این تبدیل است.



شکل ۶- عملکرد تبدیل کاهش نرخ نمونه‌برداری

۵-۱-۴- افزایش نویز

نقاط جدیدی که تعداد آن‌ها با استفاده از نرخ تغییرات تعیین می‌شود به خط سیر مرجع اضافه می‌شوند. شکاف میان این نقاط با نقاط خط سیر مرجع با استفاده از پارامتر فاصله تعیین می‌شود. مثالی از این تبدیل در شکل ۷ نشان داده شده است.

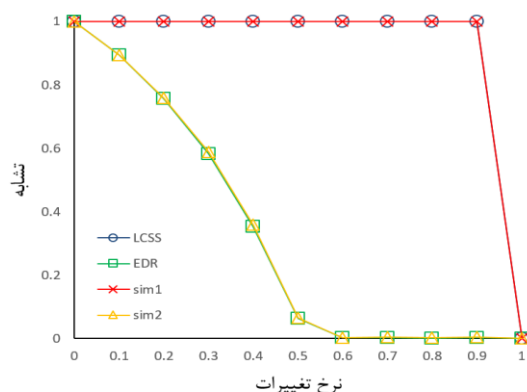


شکل ۷- عملکرد تبدیل افزایش نویز

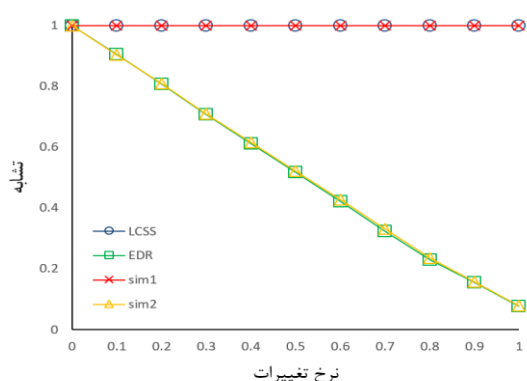
۵-۲- پیاده‌سازی و تحلیل نتایج

۵-۲-۱- آزمایش شماره ۱

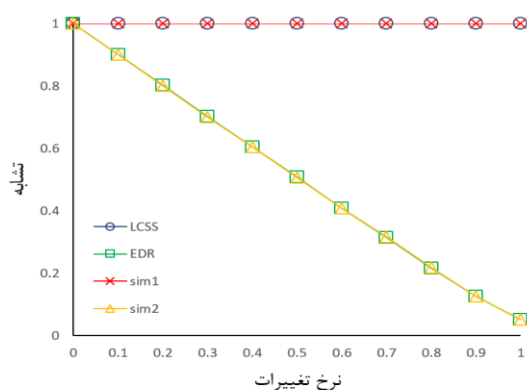
در این قسمت، تبدیلات معرفی شده در بخش ۵-۱ بر خط سیر مرجع اعمال می‌شوند و برای هر یک از تبدیلات و به ازای هر سری از پارامترها ۱۰ خط سیر ایجاد می‌شود. تابع عضویتی مشابه تابع تعریف شده در رابطه ۹ با مقادیر $p=5$ و $q=100$ متر در نظر گرفته شده است. با توجه به اینکه بهترین حد آستانه برای روش‌های LCSS و EDR برابر انحراف معیار



شکل ۱۲- نتایج عملکرد روش‌های LCSS, EDR, sim1, sim2 برای تبدیل کاهش نرخ نمونه‌برداری



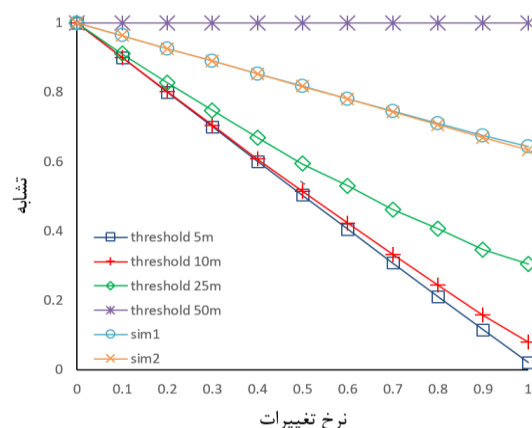
شکل ۱۳- نتایج عملکرد روش‌های LCSS, EDR, sim1, sim2 برای تبدیل افزایش نرخ نمونه‌برداری



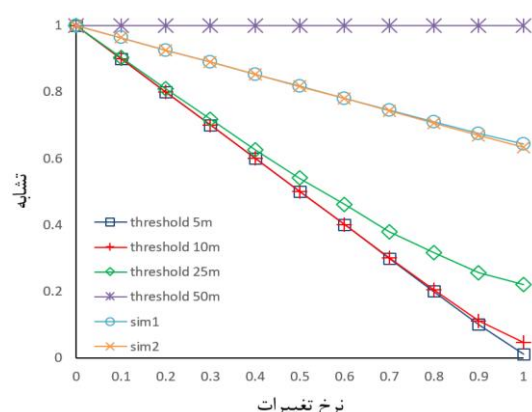
شکل ۱۴- نتایج روش‌های LCSS, EDR, sim1, sim2 نسبت به افزایش نویز

۵-۲-۲- آزمایش شماره ۲

در این قسمت، مقایسه میان روش‌های مدنظر با استفاده از داده‌های واقعی انجام می‌شود. برای این کار چهار حد آستانه ۵، ۱۰، ۲۵ و ۵۰ متر برای روش‌های LCSS و EDR در نظر گرفته شده است. همچنین سه نوع تابع عضویت فازی نیز برای sim1 و sim2 لحاظ شده است (شکل ۱۵).



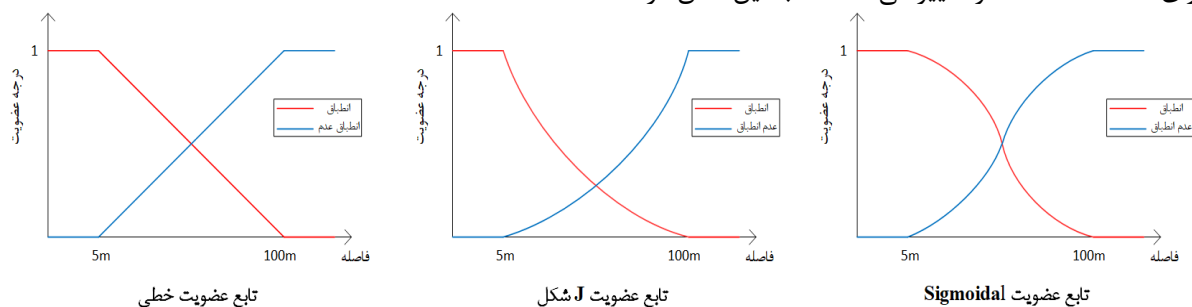
شکل ۱۰- نتایج حاصل از جابجایی نقاط با فاصله ثابت ۳۰ متر و نرخ‌های مختلف برای روش‌های LCSS و sim1 و sim2



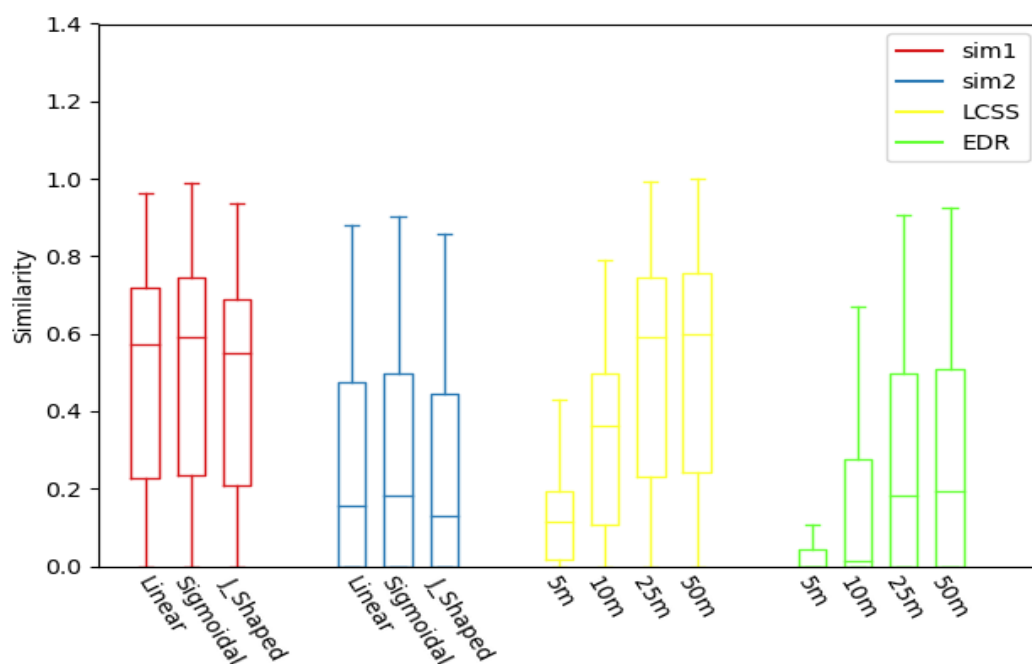
شکل ۱۱- نتایج حاصل از جابجایی نقاط با فاصله ثابت ۳۰ متر و نرخ‌های مختلف برای روش‌های EDR و sim1 و sim2

شکل‌های ۱۲، ۱۳ و ۱۴ نتایج روش‌های موردبحث را برای تبدیلات کاهش نرخ نمونه‌برداری، افزایش نرخ نمونه‌برداری و افزایش نویز نشان می‌دهند. عملکرد روش‌های sim1 و sim2 در این سه تبدیل، به ترتیب مشابه عملکرد روش‌های LCSS و EDR است. همچنین می‌توان دید که در روش‌های EDR و sim2 حساسیت نسبت به کاهش نرخ نمونه‌برداری، بیشتر از حساسیت نسبت به افزایش نرخ نمونه‌برداری است. روش‌های LCSS و sim1 نیز به دلیل اینکه فقط نقاطی که بر هم منطبق هستند را در نظر می‌گیرند و نقاط غیر منطبق را در نظر نمی‌گیرند و به دلیل نرمال‌سازی نتایج، هیچ حساسیتی نسبت به تغییر نرخ نمونه‌برداری و افزایش نویز از خود نشان نمی‌دهند.

روش EDR نیز رخ می‌دهد و دامنه شباهت از ۰,۳ تا حدود ۰,۹ تغییر می‌کند. درحالی‌که در روش‌های sim1 و sim2 باوجود اینکه سه نوع تابع عضویت مختلف استفاده شده است، تغییرات دامنه شباهت کمتر از ۰,۱ است. این بدین معنی است که حساسیت روش‌های sim1 و sim2 نسبت به تابع عضویت، کمتر از حساسیت روش‌های LCSS و EDR نسبت به حد آستانه است.



شکل ۱۵- توابع عضویت فازی مورد استفاده در آزمایش ۲



شکل ۱۶- نمودار جعبه‌ای میزان شباهت‌های خطوط سیر با استفاده از روش‌های مختلف

جایابی و مقایسه آن‌ها با روش‌های LCSS و EDR آزمایش‌های متنوعی طراحی و پیاده‌سازی شد. نتایج این آزمایش‌ها نشان می‌دهند که sim1 و sim2 از جنبه‌های حساسیت به نویز و حساسیت نسبت به تغییرات نرخ نمونه‌برداری که از نقاط قوت روش‌های LCSS و EDR هستند، عملکردی مشابه آن‌ها دارند. همچنین این نتایج نشان می‌دهند که از لحاظ حساسیت نسبت به جایابی، روش‌های sim1 و sim2 عملکرد بهتری نسبت به

اندازه‌گیری شباهت خطوط سیر با روش‌های مختلف به صورت نمودار جعبه‌ای (Box and Whisker plot) در شکل ۱۶ نشان داده شده است. به این ترتیب، اطلاعات متنوعی مثل بیشینه و کمینه مقدار شباهت، دامنه شباهت‌ها و میانگین آن‌ها را می‌توان از این نوع نمودار استخراج کرد.

همان‌طور که در شکل ۱۶ می‌توان دید، دامنه شباهت در روش LCSS از ۰,۴ برای حد آستانه ۵ متر تا حدود ۱ برای حد آستانه ۵۰ متر تغییر می‌کند. مشابه این اتفاق در

۶- نتیجه‌گیری و پیشنهادها

در این تحقیق دو روش sim1 و sim2 مبتنی بر روش‌های LCSS و EDR به منظور اندازه‌گیری شباهت خطوط سیر توسعه داده شد. برای این کار از تئوری فازی برای مقابله با عدم قطعیت و ناهمگونی داده‌ها استفاده شده است. به منظور بررسی اثربخشی روش‌های توسعه داده شده از جنبه‌های حساسیت به نویز، تغییرات نرخ نمونه‌برداری و

روش‌های LCSS و EDR دارند. علاوه بر این، با توجه به نتایج آزمایش دوم، می‌توان گفت که حساسیت روش‌های sim1 و sim2 نسبت به تابع عضویت، کمتر از حساسیت روش‌های LCSS و EDR نسبت به حد آستانه است.

این نتایج نشان‌دهنده مزایای به‌کارگیری تئوری فازی در اندازه‌گیری تشابه خطوط سیر است؛ در شرایطی که خطای اندازه‌گیری و ناهمگونی در داده‌های خطوط سیر وجود دارد و همچنین تعریف دقیقی از تشابه وجود ندارد، نتایج حاصل از روش‌های LCSS و EDR تحت تأثیر عدم قطعیت هستند. چراکه در این دو روش فقط منطبق بودن

یا نبودن نقاط تعیین می‌شود. در مقابل، روش‌های sim1 و sim2 به دلیل اینکه از توابع عضویت فازی برای تعریف تشابه و تعیین میزان انطباق نقاط استفاده می‌کنند، نتایج قابل‌اطمینان‌تری نسبت به روش‌های LCSS و EDR دارند.

در نهایت، با توجه به اینکه حرکت اشیاء در بافت‌های جغرافیایی صورت می‌پذیرد و این بافت‌ها به صورت مستقیم و غیرمستقیم بر فرآیند حرکت تأثیر می‌گذارند [۲۸، ۳۰]، به عنوان پیشنهادی برای تحقیقات آتی، می‌توان با غنی‌سازی خطوط سیر خام با انواع بافت‌ها و ایجاد خطوط سیر معنایی، روش‌های فعلی را توسعه داد.

مراجع

- [1] Abbaspour, R. A., Shaeri, M., & Chehreghani, A. (2017). A method for similarity measurement in spatial trajectories. *Spatial Information Research*, 25(3), 491-500.
- [2] Alvares, L. O., Bogorny, V., Kuijpers, B., de Macedo, J. A. F., Moelans, B., & Vaisman, A. (2007). A model for enriching trajectories with semantic geographical information. Paper presented at the Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems.
- [3] Brimicombe, A. (2009). *GIS, environmental modeling and engineering*: CRC Press.
- [4] Buchin, M., Dodge, S., & Speckmann, B. (2014). Similarity of trajectories taking into account geographic context. *Journal of Spatial Information Science*, 2014(9), 101-124.
- [5] Cai, Y., & Ng, R. (2004). Indexing spatio-temporal trajectories with Chebyshev polynomials. Paper presented at the Proceedings of the 2004 ACM SIGMOD international conference on Management of data.
- [6] Chen, L., & Ng, R. (2004). On the marriage of lp-norms and edit distance. Paper presented at the Proceedings of the Thirtieth international conference on Very large data bases-Volume 30.
- [7] Chen, L., Özsu, M. T., & Oria, V. (2005). Robust and fast similarity search for moving object trajectories. Paper presented at the Proceedings of the 2005 ACM SIGMOD international conference on Management of data.
- [8] Cruz, M. O., Macedo, H., & Guimaraes, A. (2015). Grouping similar trajectories for carpooling purposes. Paper presented at the 2015 Brazilian Conference on Intelligent Systems (BRACIS).
- [9] Ding, H., Trajcevski, G., Scheuermann, P., Wang, X., & Keogh, E. (2008). Querying and mining of time series data: experimental comparison of representations and distance measures. *Proceedings of the VLDB Endowment*, 1(2), 1542-1552.
- [10] Dodge, S., Laube, P., & Weibel, R. (2012). Movement similarity assessment using symbolic representation of trajectories. *International Journal of Geographical Information Science*, 26(9), 1563-1588.
- [11] Frihida, A., Marceau, D. J., & Thériault, M. (2004). Extracting and visualizing individual space-time paths: an integration of GIS and KDD in transport demand modeling. *Cartography and Geographic Information Science*, 31(1), 19-28.
- [12] Furtado, A. S., Alvares, L. O. C., Pelekis, N., Theodoridis, Y., & Bogorny, V. (2018). Unveiling movement uncertainty for robust trajectory similarity analysis. *International Journal of Geographical Information Science*, 32(1), 140-168.
- [13] Furtado, A. S., Kopanaki, D., Alvares, L. O., & Bogorny, V. (2016). Multidimensional similarity measuring for semantic trajectories. *Transactions in GIS*, 20(2), 280-298.
- [14] Goodchild, M. F., & Gopal, S. (1989). *The accuracy of spatial databases*: CRC Press.
- [15] Jalili, M., Hakimpour, F., & Van der Spek, S. C. (2018). Extraction of Usage Patterns for Land-Use Types by Pedestrian Trajectory Analysis. Paper presented at the International Symposium on Web and Wireless Geographical Information Systems.
- [16] Keogh, E., & Ratanamahatana, C. A. (2005). Exact indexing of dynamic time warping. *Knowledge and information systems*, 7(3), 358-386.
- [17] Kurbalija, V., Radovanović, M., Geler, Z., & Ivanović, M. (2014). The influence of global constraints on similarity measures for time-series databases. *Knowledge-Based Systems*, 56, 49-67.

- [18] Laube, P. (2014). Computational movement analysis: Springer.
- [19] Laube, P., & Purves, R. S. (2011). How fast is a cow? Cross-scale analysis of movement data. *Transactions in GIS*, 15(3), 401-418.
- [20] Leal, E., Gruenwald, L., & Zhang, J. (2016). Handling uncertainty in trajectories of moving objects in unconstrained outdoor spaces. Paper presented at the 2016 IEEE International Conference on Big Data (Big Data).
- [21] Lin, B., & Su, J. (2005). Shapes based trajectory queries for moving objects. Paper presented at the Proceedings of the 13th annual ACM international workshop on Geographic information systems.
- [22] M. Shaeri, & R. A. Abbaspour. (2015). Comparison of Distance Functions for Similarity Measurement in Spatial Trajectories. *Journal of Geomatics Science and Technology*, 4(3), 201-212. Retrieved from <http://jgst.issge.ir/article-1-66-fa.html>
- [23] M. Sharif, & A. A. Alesheikh. (2016). A Review on the Process of Point Objects' Movement and Their Trajectory Similarity Measurement Approaches. *Geospatial Engineering Journal*, 7(1), 41-54. Retrieved from <http://gej.issge.ir/article-1-151-fa.html>
- [24] Miller, H. J., & Wentz, E. A. (2003). Representation and spatial analysis in geographic information systems. *Annals of the Association of American Geographers*, 93(3), 574-594.
- [25] Moosavi, S., Omidvar-Tehrani, B., & Ramnath, R. (2017). Trajectory Annotation by Discovering Driving Patterns. Paper presented at the Proceedings of the 3rd ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics.
- [26] Naderivesal, S., Kulik, L., & Bailey, J. (2019). An effective and versatile distance measure for spatiotemporal trajectories. *Data Mining and Knowledge Discovery*, 33(3), 577-606.
- [27] Parent, C., Spaccapietra, S., Renso, C., Andrienko, G., Andrienko, N., Bogorny, V., . . . Pelekis, N. (2013). Semantic trajectories modeling and analysis. *ACM Computing Surveys (CSUR)*, 45(4), 42.
- [28] Sharif, M., & Alesheikh, A. A. (2017). Context-awareness in similarity measures and pattern discoveries of trajectories: a context-based dynamic time warping method. *GIScience & Remote Sensing*, 54(3), 426-452.
- [29] Toohey, K., & Duckham, M. (2015). Trajectory similarity measures. *Sigspatial Special*, 7(1), 43-50.
- [30] Vlachos, M., Gunopoulos, D., & Kollios, G. (2002). Discovering similar multidimensional trajectories. Paper presented at the icde.
- [31] Wang, H., Su, H., Zheng, K., Sadiq, S., & Zhou, X. (2013). An effectiveness study on trajectory similarity measures. Paper presented at the Proceedings of the Twenty-Fourth Australasian Database Conference-Volume 137.
- [32] Wang, S., Wu, L., Zhou, F., Zheng, C., & Wang, H. (2015). Group Pattern Mining on Moving Objects' Uncertain Trajectories. *International Journal of Computers Communications & Control*, 10(3), 428-440.
- [33] Xia, Y., Wang, G.-Y., Zhang, X., Kim, G.-B., & Bae, H.-Y. (2011). Spatio-temporal similarity measure for network constrained trajectory data. *International Journal of Computational Intelligence Systems*, 4(5), 1070-1079.
- [34] Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), 338-353.
- [35] Zhang, J., & Goodchild, M. F. (2002). Uncertainty in geographical information: CRC press.
- [36] Zheng, K., Zheng, Y., Xie, X., & Zhou, X. (2012). Reducing uncertainty of low-sampling-rate trajectories. Paper presented at the 2012 IEEE 28th International Conference on Data Engineering.