

ارزیابی مبتنی بر روابط تشابه مکانی خروجی جنرالیزاسیون شبکه‌های راه با استفاده از تکنیک تناظریابی

داود طالبی*^۱، محمد کریمی^۲

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

نصیرالدین طوسی

davoodtalebi16@yahoo.com

^۲ استادیار دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه‌نصیرالدین طوسی

mkarimi@kntu.ac.ir

(تاریخ دریافت شهریور ۱۳۹۶، تاریخ تصویب تیر ۱۳۹۷)

چکیده

با توجه به اهمیت جنرالیزاسیون در ساخت پایگاه داده‌های چندمقیاسی از یک طرف و امکان تولید خروجی‌های مختلف از طریق الگوریتم‌ها و عملگرهای مختلف جنرالیزاسیون از طرف دیگر، انتخاب بهترین و شبیه‌ترین خروجی تولید شده ضروری می‌باشد. هدف این تحقیق، ارائه یک رویکرد مناسب جهت محاسبه میزان تشابه مکانی بین دو شبکه راه اصلی و جنرالیزاسیون شده با استفاده از تناظریابی و روابط تشابه مکانی می‌باشد. برای رسیدن به این مهم یک رویکرد تناظریابی با استفاده از پنج معیار جهت، فاصله هاسدروف میانه بر مبنای طول، مساحت مشترک حریم، طول و تابع تانژانت پیشنهاد شده است. در این تحقیق همچنین جهت محاسبه میزان تشابه مکانی بین دو شبکه راه تابعی از ده معیار طول، اندازه، اعوجاج، پیچیدگی، مساحت شی خطی، انحنا، تعداد پاره خط‌ها، تابع تانژانت، میانگین زاویه-داری و تغییر جهت تعریف شده است. در این تحقیق از دو شبکه راه ۱:۲۵۰۰۰ و ۱:۵۰۰۰۰ استفاده شده و جهت انجام جنرالیزاسیون از الگوریتم داگلاس-پوکر استفاده شده است. میزان دقت تناظریابی با رویکرد پیشنهادی بر اساس مقدار F-Score برابر ۹۸,۳۶۱ درصد بدست آمد. نتایج نشان می‌دهد که شبکه راه تولید شده حاصل از جنرالیزاسیون شبکه راه ۱:۲۵۰۰۰ با الگوریتم داگلاس-پوکر و تلورانس ۳ متر بیشترین میزان تشابه مکانی را به نقشه ۱:۵۰۰۰۰ دارا می‌باشد.

واژگان کلیدی: روابط تشابه مکانی، تناظریابی، جنرالیزاسیون، شبکه راه

* نویسنده رابط

۱- مقدمه

پایگاه‌های داده مکانی در مقیاس‌های مختلف، یکی از اجزای اساسی در زیر ساخت‌های داده‌های مکانی محسوب می‌شوند و نقش بسیار مهمی را در ساخت و ساز شهرهای دیجیتالی^۱ ایفا می‌کنند [۱]. به صورت رایج و سنتی، یک پایگاه داده نقشه چند مقیاسی از یک منطقه، به صورت دستی یا نیمه خودکار از طریق روش‌هایی که به اصطلاح "روش چند نسخه"^۲ نامیده می‌شود، ساخته می‌شوند [۲]. در این روش برای ساخت یک پایگاه داده نقشه چند مقیاسی از یک منطقه، چندین نقشه در مقیاس‌های مختلف از منطقه مورد نظر تهیه شده و پس از پردازش، این پایگاه داده‌ها در یک پایگاه داده بزرگ‌تر با یکدیگر ادغام شده و نگهداری می‌شوند. اگرچه در دهه‌های اخیر تقریباً تمام پایگاه داده‌های نقشه برداری چند مقیاسی در بسیاری از کشورها با استفاده از روش چند نسخه تولید شده است؛ اما پایگاه داده‌های نقشه چند مقیاسی تولید شده با استفاده از این روش با ضعف‌ها و کاستی‌هایی همراه هستند مانند غیر اقتصادی بودن و اشغال فضای زیاد حافظه [۳، ۲].

به منظور رفع ضعف‌ها و کاستی‌های روش چند نسخه، از جنرالیزاسیون برای تولید پایگاه داده مکانی استفاده می‌شود [۴، ۵]. به انتخاب و نمایش ساده سازی شده از جزئیات، متناسب با مقیاس و یا اهداف تولید نقشه، جنرالیزاسیون نقشه گفته می‌شود [۶]. اطلاعات موجود در نقشه بسته به اهداف و مقیاس نقشه جدید تولید شده باید به گونه‌ای خلاصه‌سازی و پردازش شوند که هم ویژگی‌های اصلی و اساسی اطلاعات حفظ شود و هم به لحاظ تصویری و نمایشی اطلاعات موجود خوانا و معنی‌دار باشند [۷]. امروزه استفاده از جنرالیزاسیون برای تولید پایگاه داده چند مقیاسی، به عنوان یک روش ایده‌آل مطرح می‌باشد و در آژانس ملی نقشه‌برداری^۳ از جنرالیزاسیون نیمه خودکار و خودکار به منظور تولید و کوتاه‌تر کردن فرایند به روز رسانی داده‌های جغرافیایی و نقشه‌ها در مقیاس‌های مختلف استفاده می‌شود [۸].

به علت دخالت سلیقه در انتخاب و تنظیم پارامترها، قوانین، قیدها و الگوریتم‌های جنرالیزاسیون معمولاً پایگاه داده‌های مختلفی تولید می‌شود که لازم است رویکردی مناسب جهت مقایسه و ارزیابی خروجی‌های مختلف به منظور انتخاب بهترین خروجی ارائه شود [۱]. به طور کلی برای ارزیابی و مقایسه میزان تشابه عوارض خطی گروهی مانند شبکه راه‌ها، شبکه رودخانه‌ها و بلوک‌های ساختمانی کارهای اندکی انجام شده است [۹، ۱۰] و عمده تحقیقات انجام شده بر روی عوارض خطی فردی مانند یک خط و خروجی‌های حاصل از جنرالیزاسیون آن، انجام شده است [۱۱، ۱۲] که این امر غالباً با استفاده از روابط تشابه مکانی^۴ مدل‌سازی شده است [۶، ۱۳-۱۸].

در تحقیقات پیشین در زمینه تناظرایی با توجه به روش‌های پیشنهادی برای انجام فرایند تناظرایی، از معیارهای هندسی و غیر هندسی مختلفی استفاده شده است. معیارهای هندسی مورد استفاده مانند طول، فاصله، مساحت مشترک حریم، جهت، اعوجاج و پیچیدگی می‌باشد. همچنین معیارهای غیر هندسی مورد استفاده مانند نام و عرض عارضه می‌باشد. Volz در سال ۲۰۰۶ با استفاده از معیارهای طول، جهت، فاصله و درجه گره‌ها راهکاری را جهت شناسایی عوارض متناظر بین دو مجموعه داده ATKIS و GDF ارائه کرده است [۱۹]. Zhang و Meng (۲۰۰۷) با استفاده از معیارهای طول، مساحت شی خطی، فاصله، موقعیت گره‌ها و درجه گره‌ها راهکاری را جهت شناسایی عوارض متناظر بین دو مجموعه داده با هدف یکپارچه‌سازی آن‌ها ارائه داده‌اند [۲۰]. Tong و همکاران (۲۰۰۹) با استفاده از معیارهای جهت، شکل، فاصله و درجه گره‌ها راهکاری را جهت تلفیق دو مجموعه داده ارائه داده‌اند. روش ارائه شده توسط آن‌ها بر مبنای احتمالات بوده است [۲۱]. Yang و همکاران (۲۰۱۳) با استفاده از معیارهای فاصله، جهت و طول راهکاری را جهت شناسایی عوارض متناظر بین دو مجموعه داده ارائه داده‌اند [۲۲]. Olteanu-Raimond و همکاران (۲۰۱۵) با استفاده از معیارهای جهت، موقعیت گره‌ها و اطلاعات غیر هندسی راهکاری را جهت شناسایی عوارض متناظر بین دو مجموعه داده ارائه داده‌اند [۲۳]. Zhang و همکاران (۲۰۱۶) با استفاده از معیارهای طول،

^۱ Digital cities

^۲ Multiple-version method

^۳ National Mapping Agencies (NMAs)

^۴ spatial similarity relations

مساحت شی خطی، فاصله، موقعیت گره‌ها و درجه گره‌ها و اطلاعات غیر هندسی راهکاری را جهت شناسایی عوارض متناظر مجموعه داده‌های مختلف با هدف یکپارچه‌سازی آن‌ها ارائه داده‌اند [۲۴].

برخی از تحقیقات انجام شده در زمینه جنرالیزاسیون و معیارهای مورد استفاده در آن‌ها برای محاسبه میزان تشابه مکانی در ادامه ارائه شده است. McMaster در سال ۱۹۸۶ با استفاده از معیارهای جابه‌جایی بردار، جابه‌جایی مساحت و جابه‌جایی محیط، تعداد پاره خط‌ها، انحنا، میانگین تغییرات در مختصات‌ها و میانگین زاویه‌داری رویکردی را برای ارزیابی الگوریتم‌های ساده‌سازی بر روی عوارض خطی فردی ارائه کرده است [۱۱]. در سال ۱۹۹۹ پروژه Agent به طبقه‌بندی معیارهای که در تحقیقاتی که تا آن زمان در زمینه فرایند جنرالیزاسیون مورد استفاده قرار گرفته بود، پرداخت. در این تحقیق معیارها برای هر سه حالت عوارض خطی شامل نقطه، خط و سطح طبقه‌بندی و مورد تجزیه و تحلیل قرار گرفته‌اند [۲۵]. Mozas و Ariza (۲۰۱۱) با استفاده از معیارهای فاصله هاسدروف، تعداد پاره خط‌ها، طول، طول کوتاه‌ترین ضلع، طول بلندترین ضلع و میانگین طول ضلع‌ها رویکردی را برای ارزیابی دقت مکانی الگوریتم‌های ساده‌سازی بر روی عوارض خطی فردی ارائه کرده است [۲۶].

هدف اصلی این مقاله ارائه یک رویکرد مناسب جهت محاسبه میزان تشابه مکانی بین شبکه‌های راه در فضای چندمقیاسی می‌باشد که برای این منظور از تکنیک تناظریابی و روابط تشابه مکانی استفاده شده است. برای رسیدن به این هدف ابتدا با استفاده از رویکرد پیشنهادی جهت انجام فرایند تناظریابی، خطوط متناظر در دو شبکه راه مورد نظر شناسایی می‌شود. در این رویکرد تناظریابی از پنج معیار جهت، فاصله هاسدروف میانه بر مبنای طول، مساحت مشترک حریم، طول و تابع تانژانت استفاده شده است. سپس میزان تشابه بین خطوط متناظر شناسایی شده با استفاده از معیارهای محاسبه میزان تشابه خطوط فردی شامل طول، اندازه، اعوجاج، پیچیدگی، مساحت شی خطی، انحنا، تعداد پاره خط‌ها، تابع تانژانت، میانگین زاویه‌داری و تغییر جهت محاسبه می‌گردد. در پایان از میزان تشابه مکانی خطوط متناظر، میانگین‌گیری شده و میزان تشابه مکانی کل بین دو شبکه راه محاسبه می‌شود.

برای دسترسی به اهداف تحقیق ساختار مقاله به این صورت می‌باشد. در بخش ۲ مفهوم تناظریابی و همچنین رویکرد پیشنهادی برای تناظریابی و محاسبه میزان تشابه مکانی ارائه شده است. در بخش ۳ رویکرد پیشنهادی در بخش ۲، پیاده‌سازی شده است. در بخش ۴ نتایج حاصل از پیاده‌سازی رویکرد پیشنهادی ارائه شده و مورد بحث قرار گرفته است. در بخش ۵ نتیجه‌گیری و پیشنهادات جهت انجام تحقیقات آتی ارائه شده است.

۲- مواد و روش‌ها

در این بخش مبانی نظری تناظریابی به همراه رویکرد پیشنهادی ارائه شده است.

۲-۱- تناظریابی

به شناسایی عوارض مجموعه داده‌های مختلف که دارای ماهیت یکسانی هستند، تناظریابی گفته می‌شود [۲۷]. در تحقیقات گذشته تناظریابی با اهداف مختلفی انجام شده که مهم‌ترین آن یکپارچه‌سازی مجموعه داده‌های مختلف می‌باشد [۲۰، ۲۸-۳۱]. علاوه بر یکپارچه‌سازی، سایر اهداف مهم تناظریابی عبارتند از غنی‌سازی مجموعه داده‌های مختلف با هدف تلفیق مجموعه داده‌ها [۲۴، ۳۲]، ارزیابی صحت داده‌های مکانی [۳۳-۳۵]، ارزیابی صحت داده‌های توصیفی [۳۶، ۳۷]، به روز رسانی داده‌ها [۳۸-۴۰] و شناسایی و ارزیابی تغییرات [۴۱]. در این تحقیق تناظریابی با هدف شناسایی عوارض متناظر جهت محاسبه میزان تشابه مکانی انجام می‌گیرد که تاکنون به آن پرداخته نشده است.

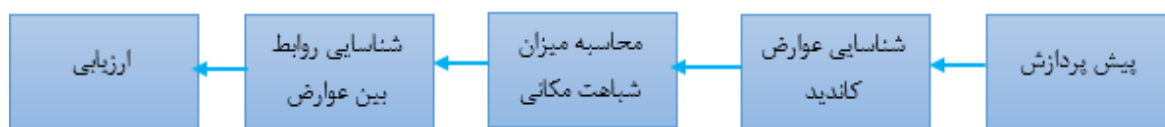
با توجه به گسترش روش‌های تولید داده‌های مکانی مانند VGI^۱ و در نتیجه افزایش داده‌های مکانی، تناظریابی پایگاه داده‌های مکانی به صورت دستی بسیار زمان‌بر و پرهزینه بوده و عملاً غیر قابل انجام می‌باشد [۴۲-۴۴]. به همین جهت، تناظریابی باید به صورت خودکار انجام شود که در این تحقیق نیز تلاش شده است که این مهم صورت پذیرد.

در سال‌های اخیر روش‌های مختلفی برای انجام فرایند تناظریابی ارائه شده است [۲۴، ۳۰، ۴۰، ۴۵]؛ اما اکثر این

^۱ Volunteered Geospatial Information

روش‌های دارای یک چهارچوب یکسانی هستند. مراحل اصلی انجام فرایند تناظریابی در شکل ۱ نشان داده شده است. قبل از ارائه جزئی بخش‌های اصلی رویکرد پیشنهادی جهت انجام فرایند تناظریابی لازم است تا دو مفهوم شبکه راه مرجع و شبکه راه کاندید معرفی شود. برای دو شبکه راه که دارای مقیاس‌های $S1$ و $S2$ هستند ($S1 > S2$)، معمولا شبکه راه با مقیاس کوچکتر ($S2$) به

عنوان شبکه راه مرجع در نظر گرفته شده و هدف شناسایی خطوط متناظر تشکیل دهنده این شبکه راه با خطوط تشکیل دهنده در شبکه راه با مقیاس بزرگتر (شبکه راه کاندید) می‌باشد. خطوط کاندید نیز به تعدادی از خطوط در شبکه راه کاندید گفته می‌شود که هدف شناسایی خطوط متناظر با خطوط شبکه راه مرجع از بین آنهاست.



شکل ۱- مراحل اصلی فرایند تناظریابی

۲-۱-۱- پیش‌پردازش

از آنجایی که ممکن است دو شبکه راه مورد نظر دارای فرمت مختلف باشند و یا سیستم مختصات‌های متفاوتی داشته باشند، ابتدا فرمت و سیستم مختصات آنها یکسان‌سازی می‌شود. سپس خطاهای توپولوژیکی مانند ردشدگی^۱ و نرسیدگی^۲ برطرف می‌شود [۴۶].

صحت و دقت داده‌ها تعیین شده است [۲۰، ۲۱، ۴۷، ۴۸]. همچنین در تعداد کمی از تحقیقات نیز با توجه به دقت مکانی دو داده مورد تناظریابی مقدار L تعیین گردیده است [۵۰، ۵۱]. در این تحقیق مقدار L به صورت تجربی و با توجه به وضعیت دو شبکه راه مورد تناظریابی، نسبت به یکدیگر، تعیین شده است.

۲-۱-۲- محاسبه میزان تشابه مکانی

پس از شناسایی عوارض کاندید در شبکه راه دوم برای هر یک از خطوط شبکه راه مرجع، لازم است تا میزان تشابه مکانی هر یک از عوارض کاندید با عارضه مرجع محاسبه شود. در تحقیقات پیشین برای محاسبه میزان تشابه مکانی از معیارهای هندسی [۲۲، ۵۲، ۵۳] و معیارهای غیر هندسی [۳۹، ۵۴] و یا ترکیبی از آنها [۲۰، ۳۸، ۴۸] استفاده شده است.

۲-۱-۲- شناسایی عوارض کاندید

به منظور انجام شناسایی عوارض کاندید روش‌های مختلفی ارائه شده است که پرکاربردترین آنها روش رشد حریم^۳ [۲۰، ۴۷، ۴۸] و روش تکراری نزدیک‌ترین نقطه^۴ [۱۹، ۴۹] می‌باشد. در این تحقیق از روش رشد حریم استفاده شده است. در این روش ابتدا برای هر یک از خطوط موجود در شبکه راه مرجع یک حریم به فاصله L ترسیم می‌شود و خطوطی از شبکه راه کاندید که در این حریم قرار می‌گیرند، به عنوان خطوط کاندید برای خط مورد نظر در شبکه راه مرجع در نظر گرفته می‌شوند. چنانچه برای هر یک از خطوط شبکه راه مرجع در حریم ترسیم شده خط کاندیدی شناسایی نشود فاصله حریم (L) افزایش می‌یابد [۴۷]. مسئله مهم در این روش تعیین مقدار L است که نقش مهمی در دقت تناظریابی و همچنین زمان محاسبات دارد. این مقدار در بسیاری از تحقیقات گذشته، به صورت تجربی و با توجه به میزان

۲-۱-۴- شناسایی روابط بین عوارض

پس از محاسبه معیارهای در نظر گرفته شده، باید با توجه به میزان تشابه مکانی بین هر یک از عوارض مرجع و عوارض کاندید، روابط بین عوارض شناسایی شوند. به طور کلی ۶ رابطه بین عوارض در تحقیقات گذشته در نظر گرفته شده است [۵۱، ۵۲، ۵۵] که در ادامه هر یک از آنها توضیح داده می‌شوند.

رابطه یک به هیچ: در این حالت برای عارضه در مجموعه داده مرجع، هیچ عارضه متناظری در مجموعه داده کاندید وجود ندارد.

۱ Overshoot
 ۲ Undershoot
 ۳ Buffer Growing
 ۴ Iterative Closest Point

۲-۱-۵- ارزیابی

در تحقیقات گذشته برای ارزیابی نتیجه بدست آمده از فرایند تناظریابی، عمدتاً از دو روش کلی مبتنی بر طول عوارض تناظریابی شده [۵۶] و مبتنی بر تعداد عوارض تناظریابی شده [۲۰، ۲۴، ۴۸، ۵۵] استفاده شده است.

۳- رویکرد پیشنهادی

رویکرد پیشنهادی جهت محاسبه میزان تشابه مکانی دو شبکه راه، به دو بخش کلی رویکرد پیشنهادی برای تناظریابی و رویکرد پیشنهادی برای محاسبه میزان تشابه مکانی تقسیم می‌شود. شکل ۲ فلوجارت کلی از رویکرد پیشنهادی را نشان می‌دهد.

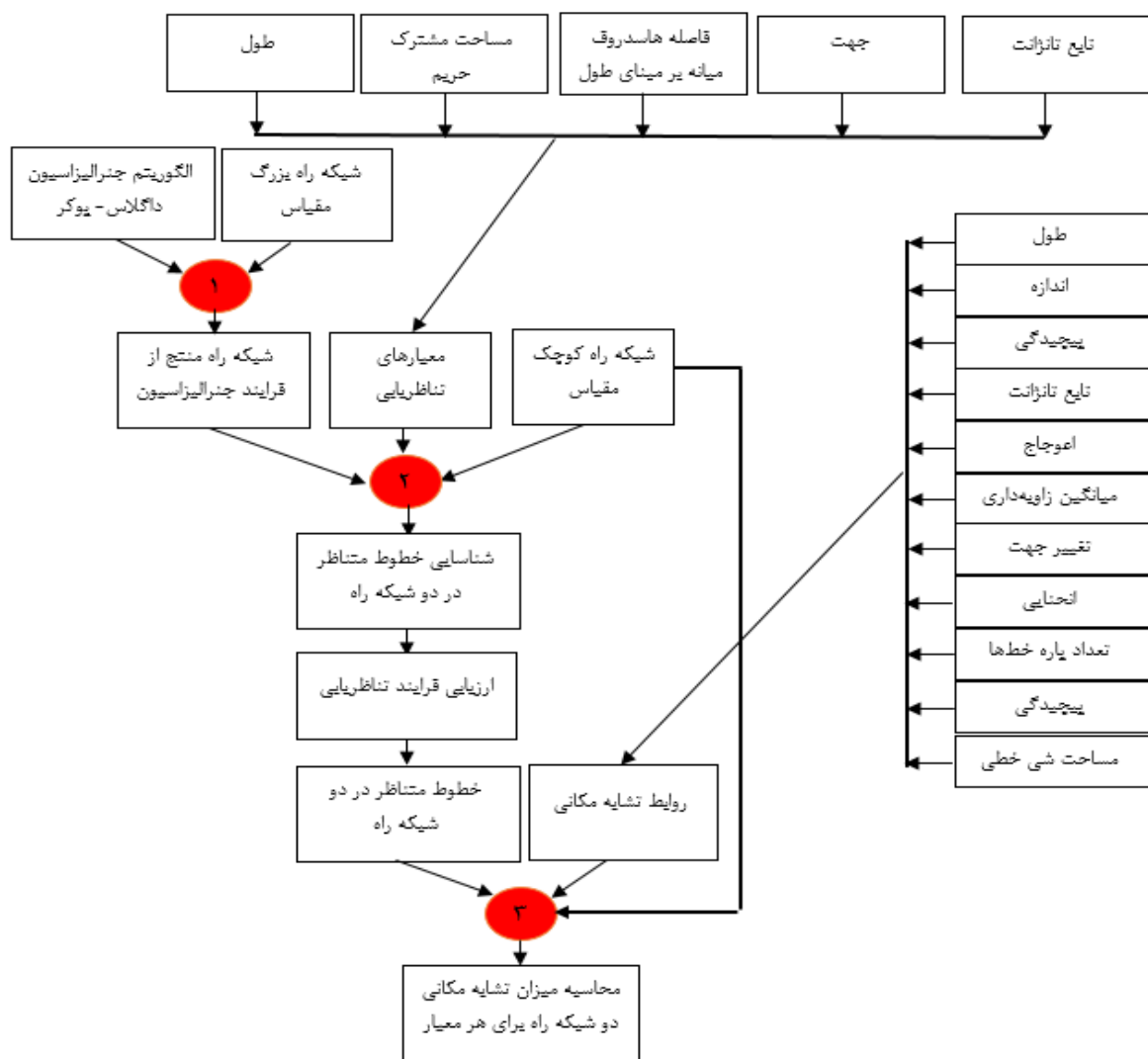
رابطه هیچ به یک: در این حالت برای عارضه در مجموعه داده کاندید، هیچ عارضه متناظری در مجموعه داده مرجع وجود ندارد.

رابطه یک به یک: در این حالت برای عارضه در مجموعه داده مرجع، یک عارضه متناظر در مجموعه داده کاندید وجود دارد.

رابطه یک به چند: در این حالت برای عارضه در مجموعه داده مرجع، بیش از یک عارضه متناظر در مجموعه داده کاندید وجود دارد.

رابطه چند به یک: در این حالت برای چند عارضه در مجموعه داده کاندید، یک عارضه متناظر در مجموعه داده مرجع وجود دارد.

رابطه چند به چند: در این حالت برای چند عارضه در مجموعه داده کاندید، چند عارضه متناظر در مجموعه داده مرجع وجود دارد.



شکل ۲- فلوجارت رویکرد پیشنهادی

۳-۱- رویکرد پیشنهادی برای تناظریابی

رویکرد پیشنهادی برای انجام فرایند تناظریابی به این صورت است که ابتدا پیش پردازش‌های لازم انجام شده و سپس با استفاده از روش رشد حریم برای شبکه راه کوچک مقیاس‌تر، خطوط کاندید در شبکه راه بزرگ مقیاس‌تر شناسایی می‌شوند. سپس پنج معیار تابع تناژانت، جهت، مساحت مشترک حریم، فاصله هاسدروف میانه بر مبنای طول و طول به منظور شناسایی خط یا خطوطی که بین خطوط کاندید بیشترین شباهت را به خط مرجع دارند، محاسبه می‌شوند. دلیل انتخاب پنج معیار تابع تناژانت، جهت، مساحت مشترک حریم، فاصله هاسدروف میانه بر مبنای طول و طول در ادامه ارائه شده است.

با توجه به اینکه در شبکه راه‌ها همیشه امکان دسترسی به داده‌های غیر هندسی وجود ندارد؛ بنابراین در رویکرد پیشنهادی تنها از معیارهای هندسی استفاده می‌شود. با بررسی تحقیقات گذشته در زمینه تناظریابی که برخی از آن‌ها در بخش ۱ ارائه شد، مشخص می‌شود که معیارهای هندسی طول، جهت، شکل و فاصله بیش از سایر معیارها مورد استفاده قرار گرفته و نتایج حاصل از تناظریابی با استفاده از آن‌ها دارای دقت قابل قبولی بوده است. برای معیار فاصله شاخص‌های مختلفی مانند فاصله اقلیدسی و مشتقات آن [۵۵، ۵۷]، فاصله فریسه [۳۸]، فاصله هاسدروف و مشتقات آن [۵۲، ۵۳] ارائه شده است. Tong و همکاران (۲۰۱۴) نشان دادند که معیار فاصله هاسدروف میانه بر مبنای طول نسبت به سایر فواصل هاسدروف و مشتقات آن دارای دقت بهتری بوده و نسبت به تغییر مختصات رئوس خطوط تغییر منطقی‌تری را از خود نشان می‌دهد [۵۳]. بنابراین در این تحقیق از فاصله هاسدروف میانه بر مبنای طول استفاده می‌شود. برای معیار شکل نیز شاخص‌های مختلفی مانند تابع Radius-Vector [۵۸]، تابع Signature [۵۹] و تابع تناژانت [۹، ۱۶، ۵۵] مورد استفاده قرار گرفته که کارکرد هر سه آن‌ها بر اساس اندازه‌گیری طول، زاویه و اعمال انتگرال بر روی آن‌ها می‌باشد. تابع تناژانت بیش از دو تابع دیگر در تحقیقات پیشین مورد استفاده قرار گرفته است؛ بنابراین در این تحقیق از تابع تناژانت استفاده می‌شود. علاوه بر چهار معیار طول، جهت، تابع تناژانت و فاصله هاسدروف میانه بر مبنای طول، معیار مساحت مشترک حریم توسط

Fan و همکاران در سال‌های ۲۰۱۴ و ۲۰۱۶ ارائه شده که باعث افزایش چشم‌گیر دقت تناظریابی شده است [۳۵، ۳۶]. بنابراین در این تحقیق از معیار مساحت مشترک حریم نیز به همراه معیارهای طول، جهت، تابع تناژانت و فاصله هاسدروف میانه بر مبنای طول در انجام فرایند تناظریابی استفاده می‌شود. نحوه محاسبه هر یک از پنج معیارها مذکور در ادامه ارائه شده است.

۳-۱-۱- تابع تناژانت

تابع تناژانت توصیف‌کننده محیط یک شی در مقابل زاویه شیب می‌باشد. محور x نشان دهنده فاصله در طول محیط خط و محور y نشان دهنده زاویه شیب است [۵۵]. رابطه ۱ نحوه محاسبه میزان تشابه مکانی خط کاندید و خط مرجع از طریق معیار تابع تناژانت را نشان می‌دهد.

$$Sim_{O,C}^{FT} = 1 - \frac{|R_O^{FT} - R_C^{FT}|}{\max(R_O^{FT}, R_C^{FT})} \quad (1)$$

در رابطه فوق، R_O^{FT} مساحت زیر نمودار تابع تناژانت خط مرجع، R_C^{FT} مساحت زیر نمودار تابع تناژانت خط کاندید، | | نشان دهنده قدر مطلق و $Sim_{O,C}^{FT}$ میزان تشابه مکانی بین خط مرجع و خط کاندید با توجه به معیار تابع تناژانت است.

۳-۱-۲- جهت

جهت یک خط بیان‌کننده زاویه بین خط واصل بین نقاط ابتدایی و انتهایی آن خط و خط افقی می‌باشد [۶۰]. رابطه ۲ نحوه محاسبه میزان تشابه مکانی خط کاندید و خط مرجع از طریق معیار جهت را نشان می‌دهد.

$$Sim_{O,C}^{Dir} = 1 - \frac{|R_O^{Dir} - R_C^{Dir}|}{\max(R_O^{Dir}, R_C^{Dir})} \quad (2)$$

در رابطه فوق، R_O^{Dir} جهت خط مرجع، R_C^{Dir} جهت خط کاندید، | | نشان دهنده قدر مطلق و $Sim_{O,C}^{Dir}$ میزان تشابه مکانی بین خط مرجع و خط کاندید با توجه به معیار جهت است.

۳-۱-۳- فاصله هاسدروف میانه بر مبنای طول

رابطه ۳ نحوه محاسبه فاصله هاسدروف میانه بر مبنای طول بین خط کاندید و خط مرجع را نشان می‌دهد [۵۳].

مکانی بین خط مرجع و خط کاندید با توجه به معیار مساحت مشترک حریم است.

همانطور که در بخش قبل بیان شد به طور کلی ۶ رابطه یک به هیچ، هیچ به یک، یک به یک، یک به چند، چند به یک و چند به چند بین عوارض خطی در تحقیقات گذشته در نظر گرفته شده است. در این تحقیق رابطه چند به چند در نظر گرفته نمی‌شود؛ زیرا معمولا در تناظریابی بین دو شبکه راه رابطه چند به چند کمتر رخ می‌دهد و همچنین معیارهای در نظر گرفته شده جهت محاسبه میزان تشابه مکانی بین دو خط که در بخش‌های ۱-۲-۳ تا ۱۰-۲-۳ ارائه شده است، برای محاسبه میزان تشابه مکانی عوارضی که دارای روابط چند به چند هستند، مناسب نمی‌باشد. پس از محاسبه معیارهای تابع تانژانت، جهت، مساحت مشترک حریم، فاصله هاسدروف میانه بر مبنای طول و طول برای هر یک خطوط شبکه راه مرجع و خطوط کاندید مربوط به آن‌ها، با استفاده از هشت مرحله زیر خط و یا خطوطی که بین خطوط کاندید بیشترین شباهت را به خط مرجع دارند، شناسایی می‌شوند:

مرحله اول: از بین خطوط کاندید برای یک خط مرجع، خطوطی که مقدار میانگین میزان تشابه مکانی آن‌ها از نظر معیارهای تابع تانژانت و جهت کمتر از ۳۰ درصد باشد، از بین مجموعه خطوط کاندید حذف می‌شوند.

مرحله دوم: از بین خطوط کاندید باقی مانده، خطوطی که فاصله هاسدروف میانه بر مبنای طول آن‌ها بیش از ۱،۲ برابر فاصله در نظر گرفته شده برای تعیین حریم در روش رشد حریم است، حذف می‌شوند.

مرحله سوم: در صورتی که پس از مراحل اول و دوم برای خط مرجع، خط کاندیدی باقی نمانده باشد، رابطه یک به هیچ اتفاق افتاده است.

مرحله چهارم: در صورتی که پس از مراحل اول و دوم برای خط مرجع، تنها یک خط کاندید باقی مانده باشد و میزان تشابه مکانی آن از نظر معیار طول بیش از ۷۰ درصد باشد، رابطه یک به یک اتفاق افتاده است.

مرحله پنجم: در صورتی که پس از مراحل اول و دوم برای خط مرجع، چندین خط کاندید باقی مانده باشد و مجموع میزان تشابه مکانی آن‌ها از نظر معیار طول (مجموع میزان تشابه مکانی دو خط کاندید و یا بیش‌تر) بین ۹۵ تا ۱۰۵ درصد نباشد، خط کاندیدی که میزان تشابه مکانی آن از نظر معیار طول بیش از ۷۰ درصد

$$D_{SMH}(O, C) = \begin{cases} h(O, C), & \text{if length}(O) < \text{length}(C) \\ h(C, O), & \text{if length}(O) \geq \text{length}(C) \end{cases} \quad (3)$$

در رابطه فوق $\text{length}(A)$ و $\text{length}(B)$ به ترتیب نشان دهنده طول شی A و B می‌باشند و $h(A, B)$ و $h(B, A)$ از طریق روابط ۴ محاسبه می‌شوند.

$$h(O, C) = \text{mean}_{o \in O} \left\{ \min_{c \in C} \|o - l_c\| \right\} \quad (4)$$

$$h(C, O) = \text{mean}_{c \in C} \left\{ \min_{o \in O} \|c - l_o\| \right\}$$

در رابطه‌های فوق l_o و l_c به ترتیب دو پاره خط دلخواه از خطی مرجع و خط کاندید، $\|o - l_c\|$ فاصله عمودی یکی از نقاط خط مرجع (به عنوان مثال o) از یکی از پاره خط-های خط کاندید (یعنی l_c) و $\|c - l_o\|$ فاصله عمودی یکی از نقاط خط کاندید (به عنوان مثال c) از یکی از پاره خط-های خط مرجع (یعنی l_o) را نشان می‌دهد.

۳-۱-۴- طول

روابط ۵ نحوه محاسبه طول بین خط مرجع و خط کاندید را نشان می‌دهد [۴۰].

$$R_m^{Len} = \sum_{i=1}^{j-1} \left(\sqrt{(X_m^{i+1} - X_m^i)^2 + (Y_m^{i+1} - Y_m^i)^2} \right) \quad (5)$$

$$Sim_{O,C}^{Len} = 1 - \frac{|R_O^{Len} - R_C^{Len}|}{\max(R_O^{Len}, R_C^{Len})}$$

در رابطه فوق R_m^{Len} طول خط m ام، j تعداد نقاط تشکیل دهنده خط m ام، X_i و Y_i مختصات نقطه i ام از خط m ام، $| \cdot |$ نشان دهنده قدر مطلق و $Sim_{O,C}^{Len}$ میزان تشابه مکانی بین خط مرجع و خط کاندید با توجه به معیار طول است.

۳-۱-۵- مساحت مشترک حریم

رابطه ۶ نحوه محاسبه مساحت مشترک حریم بین خط مرجع و خط کاندید را نشان می‌دهد [۳۵].

$$Sim_{O,C}^{BCA} = \frac{2 * A_c}{A_{Bo} + A_{BC}} \quad (6)$$

در رابطه فوق A_c مساحت مشترک بین دو حریم خط مرجع و خط کاندید، A_{Bo} مساحت حریم خط کاندید، A_{BC} مساحت حریم خط مرجع و $Sim_{O,C}^{BCA}$ میزان تشابه

باشد، به عنوان خط متناظر شناسایی شده و رابطه یک به یک اتفاق افتاده است.

مرحله ششم: در صورتی که پس از مراحل اول و دوم برای خط مرجع، چند خط کاندید باقی مانده باشد و مجموع میزان تشابه مکانی آن‌ها از نظر معیار طول (مجموع میزان تشابه مکانی دو خط کاندید و یا بیش‌تر) بین ۹۵ تا ۱۰۵ درصد باشد، رابطه یک به چند رخ داده است.

مرحله هفتم: در صورتی که پس از انجام مراحل ۱ تا ۶، مشخص شد که برای بیش از دو خط مرجع، یک خط یکسان به عنوان خط متناظر شناسایی شده است و مجموعه میزان تشابه مکانی آن‌ها از نظر معیار طول (مجموعه میزان تشابه مکانی دو خط کاندید و یا بیش‌تر) بین ۹۵ تا ۱۰۵ درصد باشد، رابطه چند به یک رخ داده است.

مرحله هشتم: در صورتی که پس از انجام مراحل فوق، هیچ یک از موارد بالا رخ ندهد رابطه هیچ به یک اتفاق افتاده است.

پس از شناسایی روابط بین خطوط شبکه راه و مشخص شدن خطوط متناظر، با استفاده از روابط زیر میزان دقت فرایند تناظریابی بین دو شبکه راه محاسبه می‌شود [۲۴].

$$Precision = \frac{TP}{TP + FP} * 100$$

$$Recall = \frac{TP}{TP + FN} * 100 \quad (7)$$

$$F - Score = \frac{Precision * Recall}{Precision + Recall} * 100$$

در روابط فوق TP تعداد روابطی هستند که در فرایند تناظریابی به درستی شناسایی شده‌اند. FP تعداد روابطی هستند که در فرایند تناظریابی به اشتباه شناسایی شده‌اند. FN تعداد روابطی هستند که در فرایند تناظریابی شناسایی نشده‌اند. $F - Score$ نشان دهنده ارزیابی دقت نهایی فرایند تناظریابی است با استفاده از $Precision$ و $Recall$ محاسبه می‌شود.

۲-۳- رویکرد پیشنهادی برای محاسبه میزان تشابه مکانی

به طور کلی، عمده معیارهایی که می‌توان از آن‌ها برای محاسبه میزان تشابه مکانی بین دو شی خطی استفاده نمود در زمینه‌های تناظریابی و جنرالیزاسیون ارائه شده و مورد استفاده قرار گرفته‌اند. با بررسی تحقیقات گذشته

می‌توان از معیارهای طول [۴۰]، طول کوتاه‌ترین ضلع، طول بلندترین ضلع، میانگین طول ضلع‌ها [۲۶]، میانگین تغییرات مختصات‌ها [۱۱]، اندازه، اعوجاج [۶۱]، ضریب تغییرات میانگین طول پاره خط‌ها [۶۲]، جابه‌جایی بردار، جابه‌جایی مساحت [۱۱، ۲۶]، جابه‌جایی محیط [۱۱]، انحنایی [۱۱، ۲۵]، پیچیدگی [۶۱]، شکل (شامل تابع تانژانت [۵۵]، تابع Radius-Vector [۵۸] و مشتقات Signature [۵۹]) فاصله (شامل فاصله اقلیدسی و مشتقات آن [۵۵، ۵۷]، فاصله فریسه [۳۸]، فاصله هاسدروف و مشتقات آن [۵۲، ۵۳]، مساحت شی خطی [۵۵]، تعداد پاره خط‌ها [۲۶]، میانگین زاویه‌داری [۱۱، ۲۵] و تغییر جهت [۶۰] برای بیان میزان تشابه مکانی بین دو شی خطی استفاده نمود.

از بین شش معیارهای طول، طول کوتاه‌ترین ضلع، طول بلندترین ضلع، میانگین طول ضلع‌ها، ضریب تغییرات میانگین طول پاره خط‌ها و جابه‌جایی محیط، تنها معیار طول؛ زیرا این شش معیار با یکدیگر همبستگی دارند. از بین معیار فاصله و جابه‌جایی بردار، تنها معیار فاصله انتخاب می‌شود؛ زیرا این دو معیار با یکدیگر همبستگی دارند. از بین انواع معیارهای فاصله، معیار فاصله هاسدروف میانه بر مبنای طول به دلایل گفته شده در بخش ۲-۲-۱ انتخاب می‌گردد. از بین معیارهای جابه‌جایی مساحت و مساحت شی خطی، تنها معیار مساحت شی خطی انتخاب می‌گردد؛ زیرا این دو معیار با یکدیگر همبستگی دارند. از بین انواع معیارهای شکل، تابع تانژانت به دلایل گفته شده در بخش ۲-۲-۱ انتخاب می‌گردد. به دلیل اینکه در یک شبکه راه دو خط متناظر از نظر معیار میانگین تغییرات مختصات‌ها دارای اختلاف بسیار ناچیزی هستند؛ این معیار در نظر گرفته نمی‌شود. بنابراین ده معیار طول، اندازه، اعوجاج، پیچیدگی، مساحت شی خطی، انحنایی، تعداد پاره خط‌ها، تابع تانژانت، میانگین زاویه‌داری و تغییر جهت برای محاسبه میزان تشابه مکانی انتخاب می‌گردند. نحوه محاسبه میزان تشابه مکانی از طریق هر یک از ده معیار مذکور در ادامه ارائه شده است.

۳-۲-۱- طول

روابط ۸ نحوه محاسبه میزان تشابه بین دو شبکه راه مرجع و کاندید از نظر معیار طول را نشان می‌دهد.

$$Sim_{0,C}^{DisSIN} = \frac{\sum_{i=1}^n 1 - \frac{|SIN_0^i - SIN_C^i|}{\max(SIN_0^i, SIN_C^i)}}{n} \quad (10)$$

$$SIN_m = \frac{SI_m}{L_m}$$

در رابطه فوق، SIN_0^i اعوجاج خط i ام از شبکه راه مرجع، SIN_C^i اعوجاج خط متناظر با خط i ام در شبکه راه کاندید، n تعداد خطوط متناظر در شبکه راه مرجع، $|.$ نشان دهنده قدر مطلق، $Sim_{0,C}^{DisSIN}$ میزان تشابه مکانی حاصل از معیار اعوجاج، بین دو شبکه راه مرجع و کاندید، SIN_m اعوجاج خط m ، SI_m اندازه خط m ، L_m طولی خط m است. $Sim_{0,C}^{DisSIN}$ دارای یک ارزش بین صفر و یک است. هر چقدر مقدار $Sim_{0,C}^{DisSIN}$ به یک نزدیک تر باشد، نشان دهنده میزان تشابه بیشتر بین دو شبکه راه می باشد.

۳-۲-۴- پیچیدگی

روابط ۱۱ نحوه محاسبه میزان تشابه بین دو شبکه راه مرجع و کاندید از نظر معیار پیچیدگی را نشان می دهد.

$$Sim_{0,C}^{DisCOM} = \frac{\sum_{i=1}^n 1 - \frac{|COM_0^i - COM_C^i|}{\max(COM_0^i, COM_C^i)}}{n} \quad (11)$$

$$COM_m = \sum_{v=1}^{u-1} \left(\left(\frac{h_m^v + h_m^{v+1}}{2} \right) \cdot \left(\frac{L_m^v}{SI_m} \right) \right)$$

در رابطه فوق، COM_0^i پیچیدگی خط i ام از شبکه راه مرجع، COM_C^i اعوجاج خط متناظر با خط i ام در شبکه راه کاندید، n تعداد خطوط متناظر در شبکه راه مرجع، $|.$ نشان دهنده قدر مطلق، $Sim_{0,C}^{DisCOM}$ میزان تشابه مکانی حاصل از معیار پیچیدگی، بین دو شبکه راه مرجع و کاندید، COM_m پیچیدگی خط m ، L_m^v طول ضلع i ام خط m ، SI_m اندازه خط m ، h_i فاصله عمودی بین راس i ام از خط واصل بین نقاط ابتدایی و انتهایی خط m ، u تعداد رئوس خط m است. $Sim_{0,C}^{DisCOM}$ دارای یک ارزش بین صفر و یک است. هر چقدر مقدار $Sim_{0,C}^{DisCOM}$ به یک نزدیک تر باشد، نشان دهنده میزان تشابه بیشتر بین دو شبکه راه می باشد.

۳-۲-۵- مساحت شیء خطی

روابط ۱۲ نحوه محاسبه میزان تشابه بین دو شبکه راه مرجع و کاندید از نظر معیار مساحت شیء خطی را نشان می دهد.

$$Sim_{0,C}^{DisLen} = \frac{\sum_{i=1}^n 1 - \frac{|L_0^i - L_C^i|}{\max(L_0^i, L_C^i)}}{n} \quad (8)$$

$$L_m = \sum_{v=1}^{u-1} \left(\sqrt{(X_m^{v+1} - X_m^v)^2 + (Y_m^{v+1} - Y_m^v)^2} \right)$$

در رابطه فوق، L_0^i طول خط i ام از شبکه راه مرجع، L_C^i طول خط متناظر با خط i ام در شبکه راه کاندید، n تعداد خطوط متناظر در شبکه راه مرجع، $|.$ نشان دهنده قدر مطلق، $Sim_{0,C}^{DisLen}$ میزان تشابه مکانی حاصل از معیار طول، بین دو شبکه راه مرجع و کاندید، L_m طول خط m ، X_m^v و Y_m^v مختصات نقطه v ام از شی m ام و u تعداد نقاط تشکیل دهنده شی m است. $Sim_{0,C}^{DisLen}$ دارای یک ارزش بین صفر و یک است. هر چقدر مقدار $Sim_{0,C}^{DisLen}$ به یک نزدیک تر باشد، نشان دهنده میزان تشابه بیشتر بین دو شبکه راه می باشد.

۳-۲-۲- اندازه

روابط ۹ نحوه محاسبه میزان تشابه بین دو شبکه راه مرجع و کاندید از نظر معیار اندازه را نشان می دهد.

$$Sim_{0,C}^{DisSI} = \frac{\sum_{i=1}^n 1 - \frac{|SI_0^i - SI_C^i|}{\max(SI_0^i, SI_C^i)}}{n} \quad (9)$$

$$SI_m = \sqrt{(X_m^e - X_m^s)^2 + (Y_m^e - Y_m^s)^2}$$

در رابطه فوق، SI_0^i اندازه خط i ام از شبکه راه مرجع، SI_C^i اندازه خط متناظر با خط i ام در شبکه راه کاندید، n تعداد خطوط متناظر در شبکه راه مرجع، $|.$ نشان دهنده قدر مطلق، $Sim_{0,C}^{DisSI}$ میزان تشابه مکانی حاصل از معیار طول، بین دو شبکه راه مرجع و کاندید، SI_m اندازه خط m ، X_m^e و Y_m^e مختصات نقطه x و y نقطه انتهایی خط m ، X_m^s و Y_m^s مختصات نقطه x و y نقطه ابتدایی خط m است. $Sim_{0,C}^{DisSI}$ دارای یک ارزش بین صفر و یک است. هر چقدر مقدار $Sim_{0,C}^{DisSI}$ به یک نزدیک تر باشد، نشان دهنده میزان تشابه بیشتر بین دو شبکه راه می باشد.

۳-۲-۳- اعوجاج

روابط ۱۰ نحوه محاسبه میزان تشابه بین دو شبکه راه مرجع و کاندید از نظر معیار اعوجاج را نشان می دهد.

$$Sim_{O,C}^{DisNOS} = \frac{\sum_{i=1}^n 1 - \frac{|NOS_O^i - NOS_C^i|}{\max(NOS_O^i, NOS_C^i)}}{n} \quad (14)$$

در رابطه فوق، NOS_O^i تعداد پاره خطها تشکیل دهنده خط i ام از شبکه راه مرجع، NOS_C^i تعداد پاره خطهای تشکیل دهنده خط متناظر با خط i ام در شبکه راه کاندید، n تعداد خطوط متناظر در شبکه راه مرجع، $|.$ نشان دهنده قدر مطلق، $Sim_{O,C}^{DisNOS}$ میزان تشابه مکانی حاصل از معیار انحنایی است. $Sim_{O,C}^{DisNOS}$ دارای یک ارزش بین صفر و یک است. هر چقدر مقدار $Sim_{O,C}^{DisNOS}$ به یک نزدیک تر باشد، نشان دهنده میزان تشابه بیشتر بین دو شبکه راه می‌باشد.

۳-۲-۸- تابع تانژانت

روابط ۱۵ نحوه محاسبه میزان تشابه بین دو شبکه راه مرجع و کاندید با استفاده از معیار تابع تانژانت را نشان می‌دهد.

$$Sim_{O,C}^{DisFT} = \frac{\sum_{v=1}^u (FT_{O,C}^v)}{u} \quad (15)$$

$$FT_{O,C}^i = 1 - \frac{|AFT_O^i - AFT_C^i|}{\max(AFT_O^i, AFT_C^i)}$$

در رابطه فوق، AFT_O^i مساحت زیر نمودار تابع تانژانت خط i ام از شبکه راه مرجع، AFT_C^i مساحت زیر نمودار تابع تانژانت خط متناظر با خط i ام در شبکه راه کاندید، u تعداد خطوط متناظر در شبکه راه مرجع، $|.$ نشان دهنده قدر مطلق، $FT_{O,C}^i$ میزان تشابه مکانی حاصل از معیار تابع تانژانت بین خط i ام از شبکه راه مرجع و خط متناظرش در شبکه راه کاندید و $Sim_{O,C}^{DisFT}$ میزان تشابه مکانی حاصل از معیار تابع تانژانت، بین دو شبکه راه مرجع و کاندید، است. $Sim_{O,C}^{DisFT}$ دارای یک ارزش بین صفر و یک است. هر چقدر مقدار $Sim_{O,C}^{DisFT}$ به یک نزدیک تر باشد، نشان دهنده میزان تشابه بیشتر بین دو شبکه راه می‌باشد.

۳-۲-۹- میانگین زاویه‌داری

روابط ۱۶ نحوه محاسبه میزان تشابه بین دو شبکه راه مرجع و کاندید از نظر معیار میانگین زاویه‌داری را نشان می‌دهد.

$$Sim_{O,C}^{DisALO} = \frac{\sum_{i=1}^n 1 - \frac{|ALO_O^i - ALO_C^i|}{\max(ALO_O^i, ALO_C^i)}}{n} \quad (12)$$

$$ALO_m = \frac{A_m}{SI_m}$$

در رابطه فوق، ALO_O^i مساحت شی خطی خط i ام از شبکه راه مرجع، ALO_C^i مساحت شی خطی خط متناظر با خط i ام در شبکه راه کاندید، n تعداد خطوط متناظر در شبکه راه مرجع، $|.$ نشان دهنده قدر مطلق، $Sim_{O,C}^{DisALO}$ میزان تشابه مکانی حاصل از معیار مساحت شی خطی، بین دو شبکه راه مرجع و کاندید، ALO_m مساحت شی خطی خط m ، A_m مساحت شکل ایجاد حاصل از اتصال نقاط ابتدایی و انتهایی خط m و SI_m اندازه خط m است. هر چقدر مقدار $Sim_{O,C}^{DisALO}$ دارای یک ارزش بین صفر و یک است. هر چقدر مقدار $Sim_{O,C}^{DisALO}$ به یک نزدیک تر باشد، نشان دهنده میزان تشابه بیشتر بین دو شبکه راه می‌باشد.

۳-۲-۶- انحنایی

رابطه ۱۳ نحوه محاسبه میزان تشابه بین دو شبکه راه مرجع و کاندید از نظر معیار انحنایی را نشان می‌دهد.

$$Sim_{O,C}^{DisCurve} = \frac{\sum_{i=1}^n 1 - \frac{|Curve_O^i - Curve_C^i|}{\max(Curve_O^i, Curve_C^i)}}{n} \quad (13)$$

در رابطه فوق، $Curve_O^i$ تعداد دفعات تغییر جهت خط i ام از شبکه راه مرجع از جهت پادساعتگرد به ساعتگرد و بالعکس، $Curve_C^i$ تعداد دفعات تغییر جهت خط متناظر با خط i ام در شبکه راه کاندید از جهت پادساعتگرد به ساعتگرد و بالعکس، n تعداد خطوط متناظر در شبکه راه مرجع، $|.$ نشان دهنده قدر مطلق، $Sim_{O,C}^{DisCurve}$ میزان تشابه مکانی حاصل از معیار انحنایی است. $Sim_{O,C}^{DisCurve}$ دارای یک ارزش بین صفر و یک است. هر چقدر مقدار $Sim_{O,C}^{DisCurve}$ به یک نزدیک تر باشد، نشان دهنده میزان تشابه بیشتر بین دو شبکه راه می‌باشد.

۳-۲-۷- تعداد پاره خطها

رابطه ۱۴ نحوه محاسبه میزان تشابه بین دو شبکه راه مرجع و کاندید از نظر معیار تعداد پاره خطها را نشان می‌دهد.

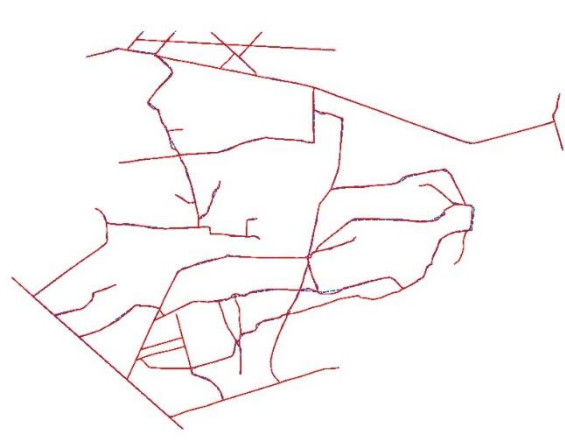
$$Sim_{O,G}^{Total} = \frac{\sum_{i=1}^9 Sim_{O,G}^i}{9} \quad (18)$$

در رابطه فوق $Sim_{O,G}^i$ میزان تشابه مکانی بین شبکه راه مرجع و شبکه راه کاندید از نظر معیار i ام و $Sim_{O,G}^{Total}$ میزان میزان تشابه مکانی کل بین شبکه راه مرجع و شبکه راه کاندید است.

۴- پیاده سازی

در این تحقیق یک شبکه راه ۱:۲۵۰۰۰ متشکل از ۹۶ خط به عنوان شبکه راه بزرگ مقیاس و یک شبکه راه ۱:۵۰۰۰۰ متشکل از ۶۲ خط به عنوان شبکه راه کوچک مقیاس از یک منطقه یکسان در نظر گرفته شده که در شکل ۳ نشان داده شده است.

با استفاده از الگوریتم داگلاس پوکر، چندین شبکه راه در تلورانس‌های مختلف از شبکه راه ۱:۲۵۰۰۰ تولید شده که تعدادی از آن‌ها در شکل ۴ ارائه شده است.



شکل ۳- شبکه راه در دو مقیاس ۱:۲۵۰۰۰ و ۱:۵۰۰۰۰

$$Sim_{O,C}^{DirAV} = \frac{\sum_{i=1}^n 1 - \frac{|AV_O^i - AV_C^i|}{\max(AV_O^i, AV_C^i)}}{n} \quad (16)$$

$$AV_m = \frac{\sum_{v_m=1}^{u_m-1} 1 - \sin(v_m/2)}{u_m - 2}$$

در رابطه فوق، AV_O^i میانگین زاویه‌داری خط i ام از شبکه راه مرجع، AV_C^i میانگین زاویه‌داری خط متناظر با خط i ام در شبکه راه کاندید، n تعداد خطوط متناظر در شبکه راه مرجع، $|.$ نشان دهنده قدر مطلق، $Sim_{O,C}^{DirAV}$ میزان تشابه مکانی حاصل از معیار میانگین زاویه‌داری بین دو شبکه راه مرجع و کاندید، AV_m میانگین زاویه‌داری خط m ، V_m مساحت زاویه بین پاره خط v ام و پاره خط $v+1$ ام از خط m و n تعداد نقاط تشکیل دهنده خط m است. $Sim_{O,C}^{DirAV}$ دارای یک ارزش بین صفر و یک است. هر چقدر مقدار $Sim_{O,C}^{DirAV}$ به یک نزدیک‌تر باشد، نشان دهنده میزان تشابه بیشتر بین دو شبکه راه می‌باشد.

۳-۲-۱۰- تغییر جهت

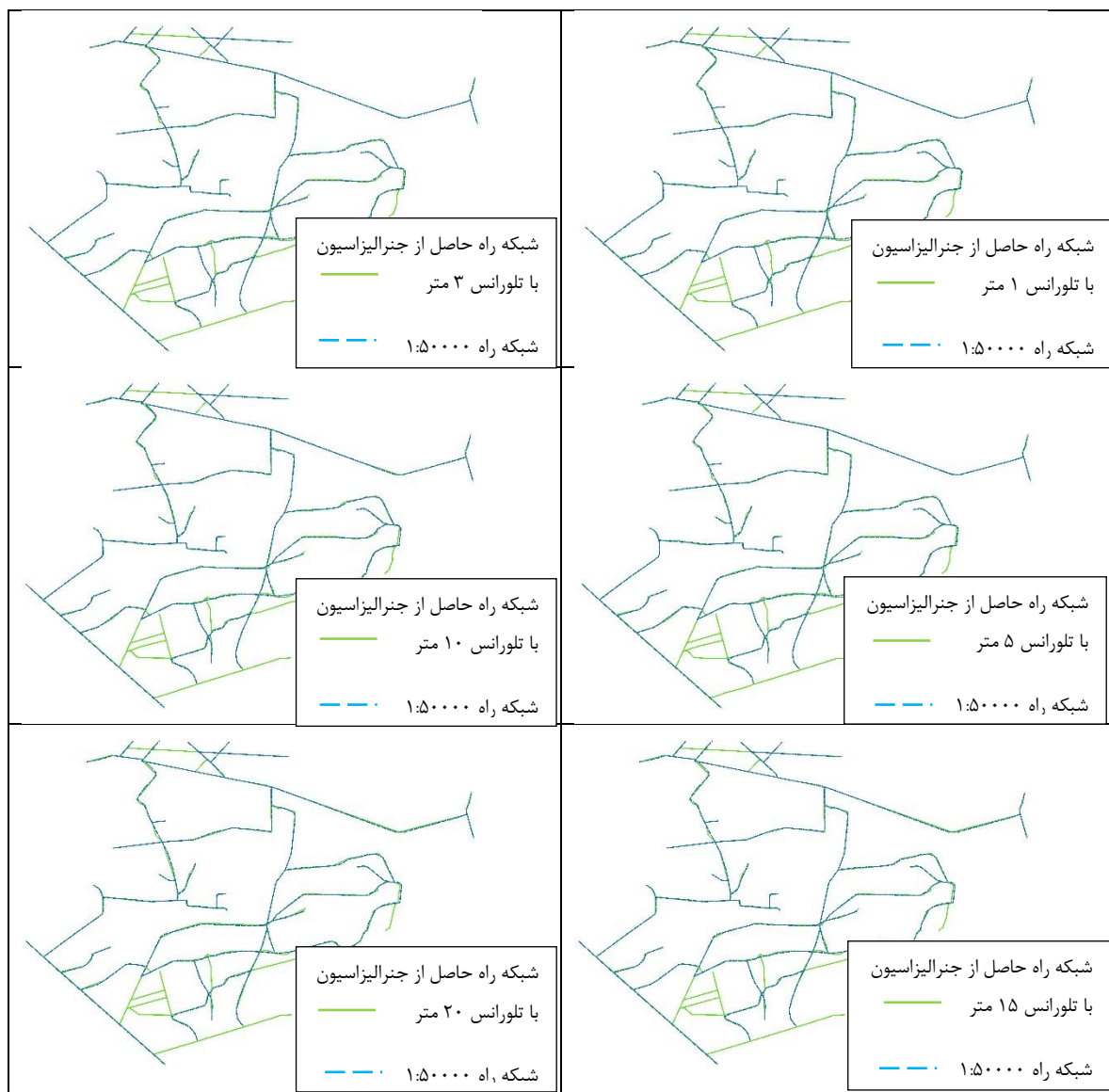
رابطه ۱۷ نحوه محاسبه میزان تشابه بین دو شبکه راه مرجع و کاندید با استفاده از معیار تغییر جهت را نشان می‌دهد.

$$Sim_{O,C}^{DirCD} = \frac{\sum_{i=1}^n 1 - \frac{|CD_O^i - CD_C^i|}{\max(CD_O^i, CD_C^i)}}{n} \quad (17)$$

در رابطه فوق، CD_O^i زاویه بین خط واصل نقاط ابتدایی و انتهایی خط i ام از شبکه راه مرجع و محور افقی، CD_C^i زاویه بین خط واصل نقاط ابتدایی و انتهایی خط متناظر با خط i ام در شبکه راه کاندید و محور افقی، n تعداد خطوط متناظر در شبکه راه مرجع، $|.$ نشان دهنده قدر مطلق، $Sim_{O,C}^{DirCD}$ میزان تشابه مکانی حاصل از معیار تغییر جهت، بین دو شبکه راه مرجع و حاصل از جنرالیزاسیون است. $Sim_{O,C}^{DirCD}$ دارای یک ارزش بین صفر و یک است. هر چقدر مقدار $Sim_{O,C}^{DirCD}$ به یک نزدیک‌تر باشد نشان دهنده میزان تشابه بیشتر بین دو شبکه راه می‌باشد.

۳-۲-۱۱- محاسبه میزان تشابه مکانی کل

رابطه ۱۸ نحوه محاسبه میزان تشابه مکانی کل بین دو شبکه راه مرجع و شبکه راه کاندید را نشان می‌دهد.



شکل ۴- شبکه‌های راه جنرالیزاسیون شده در تلورانس‌های مختلف

از شبکه‌های راه حاصل از جنرالیزاسیون محاسبه شد. نتایج حاصل از محاسبه میزان تشابه مکانی بین شبکه‌های راه مورد مقایسه در جدول ۲ ارائه شده است.

۵- نتایج و بحث

میزان دقت و صحت فرایند تناظریابی بر اساس معیار F-score همانطور که در جدول ۱ ارائه شده است برای ۱۷ جفت شبکه راه مورد تناظریابی ۹۸,۳۶۱ درصد بدست آمده است که نشان دهنده دقت بالا در انجام فرایند تناظریابی می‌باشد. همچنین همانطور که در جدول ۲ قابل مشاهده است شبکه راه حاصل از جنرالیزاسیون با الگوریتم داگلاس-پوکر و تلورانس ۳ متر دارای بیشترین میزان شباهت مکانی به شبکه راه ۱:۵۰۰۰۰ از بین شبکه

سپس فرایند تناظریابی همانطور که در رویکرد پیشنهادی برای تناظریابی در بخش ۳-۱ بیان شد؛ بین شبکه راه ۱:۵۰۰۰۰ به عنوان شبکه راه مرجع و هر یک از شبکه‌های راه تولید شده با استفاده از الگوریتم جنرالیزاسیون داگلاس-پوکر به عنوان شبکه راه کاندید، انجام گردید و دقت و صحت آن با استفاده از رابطه ۷ مورد ارزیابی قرار گرفت که نتایج حاصل از آن در جدول ۱ ارائه شده است. سپس مقدار ده معیار در نظر گرفته شده برای بیان میزان تشابه مکانی که در بخش‌های ۳-۱ تا ۳-۲-۱۰ ارائه گردید، بین شبکه راه ۱:۵۰۰۰۰ و هر یک از شبکه‌های راه حاصل از جنرالیزاسیون با استفاده از روابط ۸ تا ۱۸ محاسبه گردیده و در پایان با استفاده از رابطه ۱۸ میزان تشابه مکانی کل بین شبکه راه ۱:۵۰۰۰۰ و هر یک

واقعی را نتیجه می‌دهد. همانطور که از جدول ۲ قابل مشاهده است با افزایش مقدار تلورانس الگوریتم جنرالیزاسیون از ۳ متر، میزان تشابه مکانی بین شبکه راه حاصله و شبکه راه ۱:۵۰۰۰۰ به صورت تقریباً یکنواخت کاهش می‌یابد.

راه‌های مختلف می‌باشد. این موضوع به این مفهوم است که برای تولید شبکه راه ۱:۵۰۰۰۰ از طریق شبکه راه ۱:۲۵۰۰۰ به صورت خودکار و با استفاده از الگوریتم جنرالیزاسیون داگلاس- پوکر، تلورانس ۳ متر از این الگوریتم شبیه‌ترین شبکه راه به شبکه راه ۱:۵۰۰۰۰

جدول ۲- نتایج حاصل از محاسبه میزان تشابه مکانی

شماره	تلورانس جنرالیزاسیون (متر)	طول	تعداد پاره خط-ها	اندازه	اعوجاج	پیچیدگی	مساحت شی خطی	انحنایی	تابع تانژانت	تغییر جهت	میانگین زاویه داری	میزان تشابه مکانی کل (درصد)
۱	۱	۰,۹۴۵۳	۰,۶۱۴۵	۰,۹۵۵۱	۰,۹۸۲۵	۰,۵۶۵۳	۰,۵۲۵۴	۰,۵۳۲۴	۰,۸۱۰۱	۰,۹۶۰۷	۰,۵۰۶۴	۷۳,۹۷۷۰
۲	۲	۰,۹۴۵۳	۰,۶۰۸۷	۰,۹۵۵۱	۰,۹۸۲۵	۰,۵۶۰۵	۰,۵۲۱۸	۰,۵۲۴۰	۰,۸۱۰۱	۰,۹۶۰۷	۰,۵۰۱۰	۷۳,۶۹۷۰
۳	۳	۰,۹۴۵۴	۰,۶۰۷۳	۰,۹۵۵۱	۰,۹۸۲۶	۰,۵۷۳۷	۰,۵۴۳۴	۰,۵۳۷۱	۰,۸۰۷۸	۰,۹۶۰۷	۰,۵۰۴۰	۷۴,۱۷۱۳
۴	۴	۰,۹۴۵۵	۰,۵۹۱۵	۰,۹۵۵۱	۰,۹۸۲۷	۰,۵۷۲۸	۰,۵۳۹۵	۰,۵۱۶۵	۰,۸۰۶۸	۰,۹۶۰۷	۰,۵۰۴۳	۷۳,۷۵۴۴
۵	۵	۰,۹۴۵۶	۰,۵۶۳۲	۰,۹۵۵۱	۰,۹۸۲۹	۰,۵۵۹۹	۰,۵۲۶۸	۰,۴۸۰۱	۰,۸۰۶۵	۰,۹۶۰۷	۰,۴۹۷۱	۷۲,۷۷۹۴
۶	۶	۰,۹۴۵۷	۰,۵۴۱۱	۰,۹۵۵۱	۰,۹۸۳۰	۰,۵۶۰۱	۰,۵۳۱۸	۰,۴۷۶۷	۰,۸۰۵۷	۰,۹۶۰۷	۰,۴۸۴۸	۷۲,۴۴۷۸
۷	۷	۰,۹۴۵۷	۰,۵۲۲۰	۰,۹۵۵۱	۰,۹۸۳۳	۰,۵۴۳۱	۰,۵۲۰۹	۰,۴۷۱۲	۰,۷۹۶۳	۰,۹۶۰۷	۰,۴۹۶۸	۷۱,۹۵۱۵
۸	۸	۰,۹۴۵۷	۰,۴۸۱۰	۰,۹۵۵۱	۰,۹۸۳۵	۰,۵۴۱۴	۰,۵۳۵۶	۰,۴۴۰۲	۰,۷۹۶۳	۰,۹۶۰۷	۰,۵۰۷۹	۷۱,۴۷۳۴
۹	۹	۰,۹۴۴۸	۰,۴۵۰۷	۰,۹۵۵۱	۰,۹۸۱۴	۰,۵۲۶۰	۰,۵۱۶۹	۰,۴۳۸۲	۰,۷۹۶۰	۰,۹۶۰۷	۰,۵۰۶۰	۷۰,۷۳۸۳
۱۰	۱۰	۰,۹۴۴۸	۰,۴۴۶۱	۰,۹۵۵۱	۰,۹۸۱۴	۰,۵۲۶۰	۰,۵۱۷۱	۰,۴۳۴۸	۰,۷۹۲۳	۰,۹۶۰۷	۰,۵۰۳۲	۷۰,۶۱۶۳
۱۱	۱۱	۰,۹۴۴۸	۰,۴۳۸۳	۰,۹۵۵۱	۰,۹۸۱۵	۰,۵۲۸۸	۰,۵۲۲۵	۰,۴۱۸۷	۰,۷۹۲۳	۰,۹۶۰۷	۰,۵۰۱۰	۷۰,۴۳۷۸
۱۲	۱۲	۰,۹۴۴۹	۰,۴۲۶۷	۰,۹۵۵۱	۰,۹۸۱۵	۰,۵۲۳۰	۰,۵۱۷۵	۰,۴۱۶۳	۰,۷۹۲۵	۰,۹۶۰۷	۰,۵۰۹۴	۷۰,۲۷۷۲
۱۳	۱۳	۰,۹۴۴۸	۰,۴۱۹۳	۰,۹۵۵۱	۰,۹۸۱۶	۰,۵۱۸۳	۰,۵۱۲۷	۰,۴۱۶۳	۰,۷۹۲۶	۰,۹۶۰۷	۰,۵۲۳۱	۷۰,۲۴۵۲
۱۴	۱۴	۰,۹۴۴۸	۰,۴۱۲۱	۰,۹۵۵۱	۰,۹۸۱۷	۰,۵۲۰۱	۰,۵۲۱۴	۰,۴۰۷۶	۰,۷۹۲۶	۰,۹۶۰۷	۰,۵۱۶۵	۷۰,۱۲۶۱
۱۵	۱۵	۰,۹۴۵۳	۰,۴۰۱۵	۰,۹۵۵۱	۰,۹۸۲۲	۰,۵۱۳۸	۰,۵۲۰۶	۰,۴۱۳۸	۰,۸۰۰۴	۰,۹۶۰۷	۰,۵۲۸۷	۷۰,۲۲۱۵
۱۶	۲۰	۰,۹۴۷۰	۰,۳۷۷۹	۰,۹۵۵۱	۰,۹۸۴۲	۰,۴۷۴۵	۰,۴۹۷۳	۰,۴۲۰۵	۰,۸۰۰۴	۰,۹۶۰۷	۰,۵۷۳۷	۶۹,۹۱۲۹
۱۷	۲۵	۰,۹۴۷۶	۰,۳۵۷۷	۰,۹۵۵۱	۰,۹۸۳۹	۰,۴۶۰۹	۰,۴۷۹۵	۰,۳۹۶۳	۰,۷۹۹۸	۰,۹۶۰۷	۰,۵۷۴۷	۶۹,۱۶۳۷

۹۸,۳۶۱ درصد بدست آمد که نشان دهنده دقت بالا در انجام فرایند تناظریابی می‌باشد. نکته قابل توجه در این است که این مقدار بدون در نظر گرفتن معیارهای توپولوژیکی و معیارهای غیر هندسی (معنایی) بدست آمده است. در بخش محاسبه میزان تشابه مکانی دو شبکه راه، میزان تشابه مکانی کل بین شبکه راه ۱:۵۰۰۰۰ و هر یک از شبکه‌های راه حاصل از جنرالیزاسیون با استفاده از میانگین‌گیری مقادیر میزان تشابه بدست آمده از ده معیار طول، تعداد پاره خط‌ها، اندازه، اعوجاج، پیچیدگی، مساحت شی خطی، انحنایی، تابع تانژانت، تغییر جهت و میانگین زاویه‌داری محاسبه گردید. نتایج بدست آمده نشان دادند که شبکه راه حاصل از جنرالیزاسیون شبکه راه ۱:۲۵۰۰۰ با الگوریتم داگلاس- پوکر و تلورانس ۳ متر نسبت به سایر شبکه‌های راه بدست آمده در تلورانس‌های دیگر با استفاده از این الگوریتم، به شبکه راه ۱:۵۰۰۰۰، بیشترین میزان تشابه مکانی را دارا می‌باشد.

۶- نتیجه‌گیری و پیشنهادات

هدف اصلی این مقاله ارائه یک رویکرد مناسب جهت محاسبه میزان تشابه مکانی بین شبکه‌های راه در فضای چندمقیاسی می‌باشد که برای این مهم از تکنیک تناظریابی و روابط تشابه مکانی استفاده شده است. برای این منظور از دو شبکه راه در مقیاس ۱:۲۵۰۰۰ و ۱:۵۰۰۰۰ استفاده شده است. این تحقیق شامل دو بخش کلی تناظریابی و محاسبه میزان تشابه مکانی دو شبکه راه می‌باشد. در بخش تناظریابی، ابتدا پیش پردازش‌های لازم بر روی شبکه‌های راه انجام شد و سپس با استفاده از پنج معیار تابع تانژانت، جهت، فاصله هاسدروف میانه بر مبنای طول، مساحت مشترک حریم خطوط متناظر در هر یک از دو شبکه راه مورد مقایسه شناسایی شدند. در پایان مقدار دقت و صحت فرایند تناظریابی برای هر یک از آن‌ها با توجه به مقدار F- Score محاسبه شده که برابر

مانند عرض راه‌ها و درجه راه‌ها مورد استفاده قرار گیرند. همچنین پیشنهاد می‌شود علاوه بر الگوریتم داگلاس-پوکر، سایر الگوریتم‌های جنرالیزاسیون به همراه دستورالعمل‌های جنرالیزاسیون جهت تولید شبکه‌های راه در مقیاس‌های مختلف مورد استفاده قرار گیرد تا از این طریق بتوان شبیه‌ترین شبکه‌های راه در مقیاس‌های کوچک‌تر را با استفاده از شبکه راه بزرگ مقیاس‌تر شناسایی نمود.

در تحقیقات آینده پیشنهاد می‌شود جهت افزایش بیشتر دقت تناظریابی علاوه بر معیارهای هندسی مورد استفاده در این تحقیق، از معیارهای توپولوژیکی و معیارهای معنایی نیز در فرایند تناظریابی استفاده شود. همچنین توصیه می‌شود علاوه بر معیارهای مورد استفاده در این تحقیق جهت محاسبه میزان تشابه مکانی بین دو شبکه راه، معیارهای توپولوژیکی مربوط به چگونگی روابط بین خطوط در دو شبکه راه و همچنین معیارهای توصیفی

مراجع

- [1] H. Yan, "Theory of Spatial Similarity Relations and Its Applications in Automated Map Generalization," 2014.
- [2] J. Wang, "Principles of general map generalization," Beijing: Surveying and Mapping Press. (in Chinese), 1993.
- [3] A. Ruas, "Automating the generalisation of geographical data: the age of maturity," in Proceedings of the 20th international cartographic conference. Beijing, 2001, pp. 1943-1953.
- [4] A. Ruas, "A method for building displacement in automated map generalisation," International Journal of Geographical Information Science, vol. 12, pp. 789-803, 1998.
- [5] R. Weibel and C. B. Jones, "Computational perspectives on map generalization," Geoinformatica, vol. 2, pp. 307-314, 1998.
- [6] S. Schmid, "Automated constraint-based evaluation of cartographic generalization solutions," Geographisches Institut der Universität Zürich, 2008.
- [7] M. Modiri, M. Mohebbi, M. Masoumi, H. Khanlu, and A. Eftekhari, "Planimetric Features Generalization for the Production of Small-Scale Map by Using Base Maps and the Existing Algorithms," The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 40, p. 197, 2014.
- [8] T. Foerster, J. Stoter, and M.-J. Kraak, "Challenges for automated generalisation at European mapping agencies: a qualitative and quantitative analysis," The Cartographic Journal, vol. 47, pp. 41-54, 2010.
- [9] R. Frank and M. Ester, "A quantitative similarity measure for maps," Progress in Spatial Data Handling, pp. 435-450, 2006.
- [10] T. Gökgöz, A. Sen, A. Memduhoglu, and M. Hacı, "A new algorithm for cartographic simplification of streams and lakes using deviation angles and error bands," ISPRS International Journal of Geo-Information, vol. 4, pp. 2185-2204, 2015.
- [11] R. B. McMaster, "A statistical analysis of mathematical measures for linear simplification," The American Cartographer, vol. 13, pp. 103-116, 1986.
- [12] W. Shi and C. Cheung, "Performance evaluation of line simplification algorithms for vector generalization," The Cartographic Journal, vol. 43, pp. 27-44, 2006.
- [13] R. B. McMaster, "Automated line generalization," Cartographica: The International Journal for Geographic Information and Geovisualization, vol. 24, pp. 74-111, 1987.
- [14] S. Bard, "Quality assessment of cartographic generalisation," Transactions in GIS, vol. 8, pp. 63-81, 2004.
- [15] Y. Filippovska, V. Walter, and D. Fritsch, "Quality evaluation of generalization algorithms," ISPRS Beijing, 2008.
- [16] E. S. Podolskaya, K.-H. Anders, J.-H. Haurert, and M. Sester, "16 Quality Assessment for Polygon Generalization," Quality Aspects in Spatial Data Mining, p. 211, 2009.
- [17] W. Park and K. Yu, "Hybrid line simplification for cartographic generalization," Pattern Recognition Letters, vol. 32, pp. 1267-1273, 2011.
- [18] X. Zhang, "Automated evaluation of generalized topographic maps," Doktora Tezi, University of Twente, Enschede: Faculty of Geo-information Science and Earth Observation, Twente, 2012.
- [19] S. Volz, "An iterative approach for matching multiple representations of street data," International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 36, pp. 101-110, 2006.

- [20] M. Zhang and L. Meng, "An iterative road-matching approach for the integration of postal data," *Computers, Environment and Urban Systems*, vol. 31, pp. 597-615, 2007.
- [21] X. Tong, W. Shi, and S. Deng, "A probability-based multi-measure feature matching method in map conflation," *International Journal of Remote Sensing*, vol. 30, pp. 5453-5472, 2009.
- [22] B. Yang, Y. Zhang, and X. Luan, "A probabilistic relaxation approach for matching road networks," *International Journal of Geographical Information Science*, vol. 27, pp. 319-338, 2013.
- [23] A.-M. Olteanu-Raimond, S. Mustiere, and A. Ruas, "Knowledge formalization for vector data matching using belief theory," *Journal of Spatial Information Science*, vol. 2015, pp. 21-46, 2015.
- [24] M. Zhang, W. Yao, and L. Meng, "Automatic and accurate conflation of different road-network vector data towards multi-modal navigation," *ISPRS International Journal of Geo-Information*, vol. 5, p. 68, 2016.
- [25] D. AGENT, "Selection of Basic Measures," ed, 1999.
- [26] A. T. Mozas and F. J. Ariza, "New method for positional quality control in cartography based on lines. A comparative study of methodologies," *International Journal of Geographical Information Science*, vol. 25, pp. 1681-1695, 2011.
- [27] A. Samal, S. Seth, and K. Cueto 1, "A feature-based approach to conflation of geospatial sources," *International Journal of Geographical Information Science*, vol. 18, pp. 459-489, 2004.
- [28] D. Sheeren, S. Mustière, and J.-D. Zucker, "How to integrate heterogeneous spatial databases in a consistent way?," in *ADBIS*, 2004, pp. 364-378.
- [29] J. O. Kim, K. Yu, J. Heo, and W. H. Lee, "A new method for matching objects in two different geospatial datasets based on the geographic context," *Computers & Geosciences*, vol. 36, pp. 1115-1122, 2010.
- [30] M. Sester, J. J. Arsanjani, R. Klammer, D. Burghardt, and J.-H. Haunert, "Integrating and generalising volunteered geographic information," in *Abstracting geographic information in a data rich world*, ed: Springer, 2014, pp. 119-155.
- [31] Y. Wang, H. Lv, X. Chen, and Q. Du, "A PSO-Neural Network-Based Feature Matching Approach in Data Integration," in *Cartography-Maps Connecting the World*, ed: Springer, 2015, pp. 189-219.
- [32] B. Yang, Y. Zhang, and F. Lu, "Geometric-based approach for integrating VGI POIs and road networks," *International Journal of Geographical Information Science*, vol. 28, pp. 126-147, 2014.
- [33] J. J. Ruiz-Lendínez, F. J. Ariza-López, and M. A. Ureña-Cámara, "Automatic positional accuracy assessment of geospatial databases using line-based methods," *Survey Review*, vol. 45, pp. 332-342, 2013.
- [34] E. Abdolmajidi, J. Will, L. Harrie, and A. Mansourian, "Comparison of matching methods of user generated and authoritative geographic data," in *The 17th ICA Generalization Workshop*, 2014.
- [35] H. Fan, B. Yang, A. Zipf, and A. Rousell, "A polygon-based approach for matching OpenStreetMap road networks with regional transit authority data," *International Journal of Geographical Information Science*, vol. 30, pp. 748-764, 2016.
- [36] H. Fan, A. Zipf, Q. Fu, and P. Neis, "Quality assessment for building footprints data on OpenStreetMap," *International Journal of Geographical Information Science*, vol. 28, pp. 700-719, 2014.
- [37] L. Cao, N. C. Codella, G. Hua, G. Leiguang, A. I. Natsev, and J. R. Smith, "Techniques for spatial semantic attribute matching for location identification," ed: Google Patents, 2016.
- [38] S. Mustière and T. Devogele, "Matching networks with different levels of detail," *Geoinformatica*, vol. 12, pp. 435-453, 2008.
- [39] Y.-h. Wang and F.-Y. Wei, "A schema-matching-based approach to propagating updates between heterogeneous spatial databases," in *Proceedings of SPIE: Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Advanced Spatial Data Models and Analyses*, 2008, p. 714605.
- [40] Y. Wang, D. Chen, Z. Zhao, F. Ren, and Q. Du, "A Back-Propagation Neural Network-Based Approach for Multi-Represented Feature Matching in Update Propagation," *Transactions in GIS*, vol. 19, pp. 964-993, 2015.
- [41] J. McIntosh and M. Yuan, "Assessing similarity of geographic processes and events," *Transactions in GIS*, vol. 9, pp. 223-245, 2005.
- [42] J. W. Crampton, M. Graham, A. Poorthuis, T. Shelton, M. Stephens, M. W. Wilson, et al., "Beyond the geotag: situating 'big data' and leveraging the potential of the geoweb," *Cartography and geographic information science*, vol. 40, pp. 130-139, 2013.
- [43] B. Kang, J. Y. Scully, O. Stewart, P. M. Hurvitz, and A. V. Moudon, "Split-Match-Aggregate (SMA) algorithm: integrating sidewalk data with transportation network data in GIS," *International Journal of Geographical Information Science*, vol. 29, pp. 440-453, 2015.

- [44] E. Xavier, F. J. Ariza-López, and M. A. Ureña-Cámara, "A survey of measures and methods for matching geospatial vector datasets," *ACM Computing Surveys (CSUR)*, vol. 49, p. 39, 2016.
- [45] B. Yang, X. Luan, and Y. Zhang, "A Pattern-Based Approach for Matching Nodes in Heterogeneous Urban Road Networks," *Transactions in GIS*, vol. 18, pp. 718-739, 2014.
- [46] X. Tong, D. Liang, G. Xu, and S. Zhang, "Positional accuracy improvement: a comparative study in Shanghai, China," *International Journal of Geographical Information Science*, vol. 25, pp. 1147-1171, 2011.
- [47] V. Walter and D. Fritsch, "Matching spatial data sets: a statistical approach," *International Journal of geographical information science*, vol. 13, pp. 445-473, 1999.
- [48] M. Zhang and L. Meng, "Delimited stroke oriented algorithm-working principle and implementation for the matching of road networks," *Geographic Information Sciences*, vol. 14, pp. 44-53, 2008.
- [49] G. v. Gösseln and M. Sester, "Integration of geoscientific data sets and the german digital map using a matching approach," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 35, 2004.
- [50] C. Beerli, Y. Doytsher, Y. Kanza, E. Safra, and Y. Sagiv, "Finding corresponding objects when integrating several geo-spatial datasets," in *Proceedings of the 13th annual ACM international workshop on Geographic information systems*, 2005, pp. 87-96.
- [51] E. Safra, Y. Kanza, Y. Sagiv, and Y. Doytsher, "Ad hoc matching of vectorial road networks," *International Journal of Geographical Information Science*, vol. 27, pp. 114-153, 2013.
- [52] L. Li and M. F. Goodchild, "An optimisation model for linear feature matching in geographical data conflation," *International Journal of Image and Data Fusion*, vol. 2, pp. 309-328, 2011.
- [53] X. Tong, D. Liang, and Y. Jin, "A linear road object matching method for conflation based on optimization and logistic regression," *International Journal of Geographical Information Science*, vol. 28, pp. 824-846, 2014.
- [54] M. Al-Bakri and D. Fairbairn, "Assessing similarity matching for possible integration of feature classifications of geospatial data from official and informal sources," *International Journal of Geographical Information Science*, vol. 26, pp. 1437-1456, 2012.
- [55] M. Zhang, "Methods and implementations of road-network matching," Unpublished PhD Dissertation, Technical University of Munich, 2009.
- [56] T. Koukoletsos, M. Haklay, and C. Ellul, "Assessing data completeness of VGI through an automated matching procedure for linear data," *Transactions in GIS*, vol. 16, pp. 477-498, 2012.
- [57] D. Min, L. Zhilin, and C. Xiaoyong, "Extended Hausdorff distance for spatial objects in GIS," *International Journal of Geographical Information Science*, vol. 21, pp. 459-475, 2007.
- [58] D. Stoyan and H. Stoyan, *Fractals, random shapes, and point fields: methods of geometrical statistics* vol. 302: John Wiley & Sons Inc, 1994.
- [59] R. C. Veltkamp and M. Hagedoorn, "State of the art in shape matching," in *Principles of visual information retrieval*, ed: Springer, 2001, pp. 87-119.
- [60] A.-M. O. Raimond and S. Mustière, "Data matching—a matter of belief," in *Headway in spatial data handling*, ed: Springer, 2008, pp. 501-519.
- [61] D. L. Anderson, D. P. Ames, and P. Yang, "Quantitative methods for comparing different polyline stream network models," *Journal of Geographic Information System*, vol. 6, p. 88, 2014.
- [62] M. J. Jasinski, *The comparison of complexity measures for cartographic lines*: University of California, National Center for Geographic Information & Analysis, 1990