

بهبود تناظریابی گذرگاههای درون شهری بر اساس هندسه عوارض در ساختار داده‌های چندمقیاسی

علیرضا چهرقان^۱، رحیم علی عباسپور^{۲*}

^۱ دانشجوی دکتری سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - پردیس

دانشکده‌های فنی - دانشگاه تهران

achehregghan@ut.ac.ir

^۲ استادیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - پردیس دانشکده‌های فنی - دانشگاه تهران

abaspour@ut.ac.ir

(تاریخ دریافت بهمن ۱۳۹۴، تاریخ تصویب تیر ۱۳۹۵)

چکیده

هدف از تناظریابی عوارض، شناسایی عوارض متناظر در مجموعه داده‌های مختلف در دنیای واقعی می‌باشد. تحقیق حاضر چارچوبی را برای بهبود تناظریابی گذرگاه‌های شهری در مجموعه داده‌های با مقیاس و منابع متفاوت ارائه می‌دهد. چارچوب پیشنهادی بر اساس نظریه گراف و در نظر گرفتن معیارهای مرتبط با هندسه عوارض، جهت تناظریابی مجموعه داده‌های شبکه راه‌های درون شهری در مقیاسها و منابع مختلف می‌باشد. در راهکار ارائه شده سعی شده است با در نظر گرفتن معیارهایی هندسی نظیر فاصله، جهت، مساحت، شکل و ناحیه مشترک بین حریم عوارض و در نظر گرفتن ادراک مکانی کارشناسان در تعیین درجه اهمیت هریک از معیارها، درجه شباهت عوارض در منابع داده مختلف تعیین گردد. علاوه بر موارد فوق، در راهکار ارائه شده ضمن حذف ابهام در تعریف عوارض، تمام روابط موجود بین عوارض (شامل یک به هیچ، هیچ به یک، یک به یک، چند به چند، چند به یک و چند به چند) در نظر گرفته می‌شود. همچنین چارچوب پیشنهادی برخلاف بسیاری از روش‌های پیشین حالت کاربردی بودن خود را در مواجهه با داده‌های با مقیاسها و منابع متفاوت از دست نمی‌دهد. نتایج نشان می‌دهد که روش پیشنهادی در مسئله تناظریابی به خوبی عوارض متناظر را در مجموعه داده‌های با مقیاس و منابع مختلف شناسایی می‌کند. برای ارزیابی نتایج از ضریب F-Score استفاده گردید که شامل هر دو پارامتر Precision و Recall می‌باشد. با در نظر گرفتن معیارهای هندسی مختلف مقدار F-Score از ۰/۳۵ تا ۸۲/۶۷ درصد بهبود می‌یابد.

واژگان کلیدی: تناظریابی عوارض خطی، مجموعه داده‌های با مقیاس و منابع متفاوت، تئوری گراف، هندسه عوارض، نظرات کارشناسان

* نویسنده رابط

۱- مقدمه

با پیشرفت سریع روش‌های گردآوری داده‌های مکانی و همچنین تکنیک‌های پردازش آنها، حجم عظیمی از داده‌ها از طریق سازمان‌های دولتی و خصوصی به سهولت در دسترس کاربران قرار می‌گیرد. مجموعه داده‌های با منابع و مقیاس متفاوت در کنار اطلاعات مکانی داوطلبانه^۱ از جمله اطلاعات ارزشمندی است که ممکن است در پایگاه‌های داده ذخیره شده باشند. این داده‌ها ممکن است محدوده جغرافیایی یکسانی را پوشش دهند و در عین حال در مقیاس، موضوع و محتوا، دقت، صحت و روش گردآوری متفاوت باشند. یک مجموعه داده می‌تواند از برخی جنبه‌ها نسبت به دیگر مجموعه داده‌ها متمایز بوده و برتری داشته باشد، اما قطعاً این برتری در همه جنبه‌ها امکان‌پذیر نمی‌باشد. بنابراین به منظور بهره‌مندی از حداکثر مزایای ممکن، بهتر است مجموعه داده‌های متفاوت با هم مورد استفاده قرار گیرند تا کارآمدتر بوده و کاربردهای گوناگونی را در بر گیرند [۱].

در علوم اطلاعات مکانی، تلفیق^۲ مجموعه‌های مختلف داده در حال گسترش می‌باشد، ولی با توجه به پیچیدگی مجموعه منابع داده به همراه سطوح مختلفی از دقت و صحت در المانهای مکانی و توصیفی، به عنوان یک کار چالش برانگیز مطرح می‌باشد [۲]. تلفیق در اصل فرایند ترکیب اطلاعات از دو یا چندین مجموعه داده برای جمع‌آوری دانشی است که در هریک از مجموعه داده‌ها به تنها وجود ندارد. از این رو نتیجه تلفیق داده‌های مکانی یک مجموعه داده جدید می‌باشد که دارای دقت مکانی و همچنین داده‌های توصیفی غنی‌تری می‌باشد [۳]. در فرایند تلفیق اصلی‌ترین گام شناسایی عوارض متناظر بین مجموعه داده‌ها می‌باشد که تحت عنوان تناظریاتی عوارض^۳ شناخته می‌شود، هرچند در برخی از موارد، مراحل کلی تناظریاتی و تلفیق یکی در نظر گرفته می‌شود [۴، ۵]. در فرایند تناظریاتی با توجه به ماهیت داده‌ها که سه نوع نقطه، خط و چند ضلعی می‌باشد، تناظریاتی نیز برای سه نوع از داده‌های نقطه‌ای، خطی و چندضلعی صورت می‌گیرد که هریک از سه نوع عوارض در هر دو

مجموعه می‌تواند متناظر یکدیگر باشند. در حالت کلی جهت مقایسه عوارض در هر یک از مجموعه داده‌ها و تشکیل یک مجموعه جدید، سه خصوصیت هندسی، توپولوژیکی و معنایی^۴ در نظر گرفته می‌شود که بر اساس آنها میزان مشابهت عوارض به یکدیگر سنجیده می‌شود [۶] اما در برخی مسائل امکان دسترسی به معیارهای معنایی وجود ندارد و یا اینکه در کاربردهای تحت شبکه به دلیل زمان بالای محاسبات، نیاز است تا حداقل معیارها در نظر گرفته شود.

هدف از این تحقیق ارائه چارچوبی بر اساس تئوری گراف‌ها و در نظر گرفتن معیارهای مرتبط با هندسه عوارض جهت تناظریاتی مجموعه داده‌های شبکه راه‌ها در مقیاسها و منابع مختلف می‌باشد. در راهکار ارائه شده سعی شده است با در نظر گرفتن معیارهایی هندسی نظیر فاصله، جهت، مساحت، شکل و ناحیه مشترک بین حریم عوارض و در نظر گرفتن ادراک مکانی کارشناسان در تعیین درجه اهمیت هریک از معیارها، درجه شباهت عوارض در منابع داده مختلف و با مقیاسهای متفاوت تعیین گردد. دلیل استفاده از معیارهای هندسی این است که در تمامی مجموعه داده‌ها امکان دسترسی به این معیارها وجود دارد و نیاز به داده اضافی به جز هندسه عارضه مورد نظر نمی‌باشد.

ادامه این تحقیق به شرح ذیل است: پس از مقدمه در بخش ۲ پیشینه تحقیق مرتبط با مسئله ارائه می‌گردد. در بخش ۳ چارچوب پیشنهادی ارائه و جزئیات مورد استفاده تشریح می‌گردد. در بخش ۴ ضمن تشریح منطقه مورد مطالعه، چارچوب پیشنهادی بر روی این منطقه اعمال و نتایج مورد بررسی قرار می‌گیرد. در انتها نیز در بخش ۵ نتیجه گیری و پیشنهادات ارائه می‌گردد.

۲- پیشینه تحقیق

تحقیقات مختلفی پیرامون مسئله تناظریاتی عوارض صورت گرفته است اما یکی از ابتدایی ترین تحقیقات در این زمینه می‌توان تحقیق [۷] را نام برد. این تحقیق بر اساس پروژه‌ای در ایالات متحده بنا شده بود که هدف آن تلفیق دو گروه داده از سازمان زمین شناسی و اداره سرشماری و تعیین عوارض متناظر در این دو گروه داده

^۱ Volunteered Geospatial Information (VGI)

^۲ Conflation

^۳ Object Matching

^۴ Semantic

بود. در این تحقیق صرفاً روابط یک به یک بین عوارض در نظر گرفته شد. با توجه به در نظر گرفتن روابط بین عوارض تحقیق [۸] با استفاده از شاخصه‌های هندسی و توپولوژیک یک روش تناظر یابی تکراری ارائه داد. این روش تنها روابط یک به یک موجود بین دو مجموعه داده را در نظر می‌گرفت. از تحقیقات دیگری که رابطه یک به یک را بین عوارض در نظر گرفته است می‌توان تحقیقات [۹] و [۱۰] را نام برد. تحقیق [۱۱] جهت یکپارچه‌سازی مجموعه داده‌های مختلف به شناسایی و تعدیل عوارض در مجموعه داده‌های مختلف پرداخت. در این فرایند از روابط یک به یک و یک به چند بین عوارض در مجموعه داده‌ها استفاده گردید. تحقیق [۱۲] یک روش مبتنی بر فاصله بافر ارائه داد که روابط یک به یک، یک به چند و چند به چند را در نظر می‌گرفت. [۱۳] و [۴] از جمله تحقیقات دیگری هستند که سه رابطه یک به یک، یک به چند و چند به چند را برای تناظر یابی عوارض در نظر گرفتند. [۱۴] و [۱۵] از جمله تحقیقاتی هستند که هر چهار حالت یک به هیچ، یک به یک، یک به چند (چند به یک) و چند به چند را برای تناظر یابی در نظر گرفته‌اند.

در تحقیقات مختلف با توجه به نوع مسئله مورد نظر از معیارهای متفاوتی استفاده شده است. در بسیاری از تحقیقات تنها از معیارهای هندسی برای تناظر یابی استفاده کرده‌اند. مهمترین معیار هندسی که بیشتر مورد استفاده قرار گرفته است، اندازه‌گیری نزدیکی عوارض از طریق در نظر گرفتن فاصله بین عوارض می‌باشد. از مهمترین آنها می‌توان فاصله اقلیدسی بین نقاط [۱۶، ۱۷]، فاصله هاسدورف^۱ و مشتقات آن [۴، ۱۸، ۱۹]، فاصله L_2 [۸]، فاصله کمترین مربعات [۲۰، ۲۱] و همچنین فاصله فرشت^۲ [۶، ۲۲، ۲۳] را نام برد. تحقیقات [۲۴] و [۶] با در نظر گرفتن هندسه عوارض و با استفاده از داده‌های توصیفی به تناظر یابی عوارض در مجموعه داده‌های مختلف پرداختند. همچنین تحقیق [۲۱] با استفاده از معیارهای طول، فاصله و آزمون عوارض متناظر بین دو مجموعه داده را یافتند. در این تحقیق از مجموعه داده‌های معابر برون شهری در دو مقیاس متفاوت با منبع تهیه یکسان (مقیاس کوچکتر حاصل خلاصه‌سازی مقیاس بزرگتر) استفاده گردید.

[۹] و [۱۸] برای اولین بار مسئله تناظر یابی عوارض خطی را براساس بهینه‌سازی فاصله هاسدورف بین عوارض مطرح کردند. در این تحقیقات تنها رابطه یک به یک بین عوارض در نظر گرفته شد، اما نکته قابل تأمل این تحقیق در نظر گرفتن معیار فاصله می‌باشد. در مواجهه با داده‌های پیچیده، نظیر شبکه داده‌های درون شهری با مقیاس و منابع مختلف صرفاً معیار فاصله نمی‌تواند بخوبی عوارض متناظر را شناسایی کند. تحقیق [۴] با استفاده از فاصله هاسدورف تغییر یافته و رگرسیون لجستیک رویکرد ارائه شده توسط [۱۸] را بهبود بخشیدند، هرچند در این تحقیق نیز صرفاً از معیار فاصله استفاده شده است.

در تحقیق پیشرو فرض بر این است که یک مجموعه داده دارای دقت مکانی بالاتر و مجموعه دیگر دارای اطلاعات توصیفی غنی‌تری می‌باشد. از این رو تنها می‌توان از معیارهای هندسی جهت تناظر یابی استفاده نمود و امکان انجام تناظر یابی معنایی^۳ (استفاده از اطلاعات توصیفی) وجود ندارد. در اصل سعی شده است که تناظر یابی با حداقل اطلاعات ممکن صورت پذیرد، در نتیجه با استفاده از معیارهای هندسی استخراج شده از ساختار عوارض عمل تناظر یابی صورت می‌گیرد. همچنین تحقیق حاضر دو تفاوت با تحقیقات پیشین دارد. اول اینکه بسیاری از تحقیقاتی که از هندسه عوارض جهت شناسایی عوارض متناظر استفاده کرده‌اند از مجموعه داده‌های با مقیاس متفاوت و منبع تهیه یکسان و یا مجموعه داده‌های با مقیاس یکسان و منبع تهیه متفاوت استفاده کرده‌اند. در این صورت به علت پیچیدگی کمتر تناظر یابی داده‌ها، شناسایی عوارض توسط رویکردهای پیشنهادی با دقت قابل قبولی همراه است. اما وقتی این رویکردها در مجموعه داده‌های معابر درون شهری با مقیاس و منبع متفاوت استفاده می‌شود، از کارایی آنها کاسته می‌شود. در نتیجه رویکرد ارائه شده در تحقیق حاضر برای مجموعه داده‌های با مقیاس و منبع تهیه متفاوت و برای داده‌های معابر درون شهری است. همچنین تحقیق حاضر سعی کرده است که معیارهای هندسی معرفی شده در تحقیقات مختلف را در کنار یکدیگر قرار داده تا عوارض متناظر بیشتری قابل شناسایی باشد.

^۳ Semantic Matching

^۱ Hausdorff Distance

^۲ Frechet Distance

تفاوت دوم در مورد ابهام در تعریف عوارض در مجموعه داده‌ها می‌باشد. در بسیاری از تحقیقات تناظریایی براساس یال و یا گره‌های موجود در عوارض خطی می‌باشد. این امر ضمن افزایش پیچیدگی حل مسئله، مفهوم تناظریایی را با مشکل مواجه می‌کند. زیرا هدف از تناظریایی، شناسایی عوارض متناظر در مجموعه داده‌ها (نه شناسایی یال و یا گره متناظر) می‌باشد. از این رو در این تحقیق ساختار گرافی ارائه می‌گردد که ابهام در شناسایی عوارض خطی را برطرف نماید.

۳- چارچوب پیشنهادی

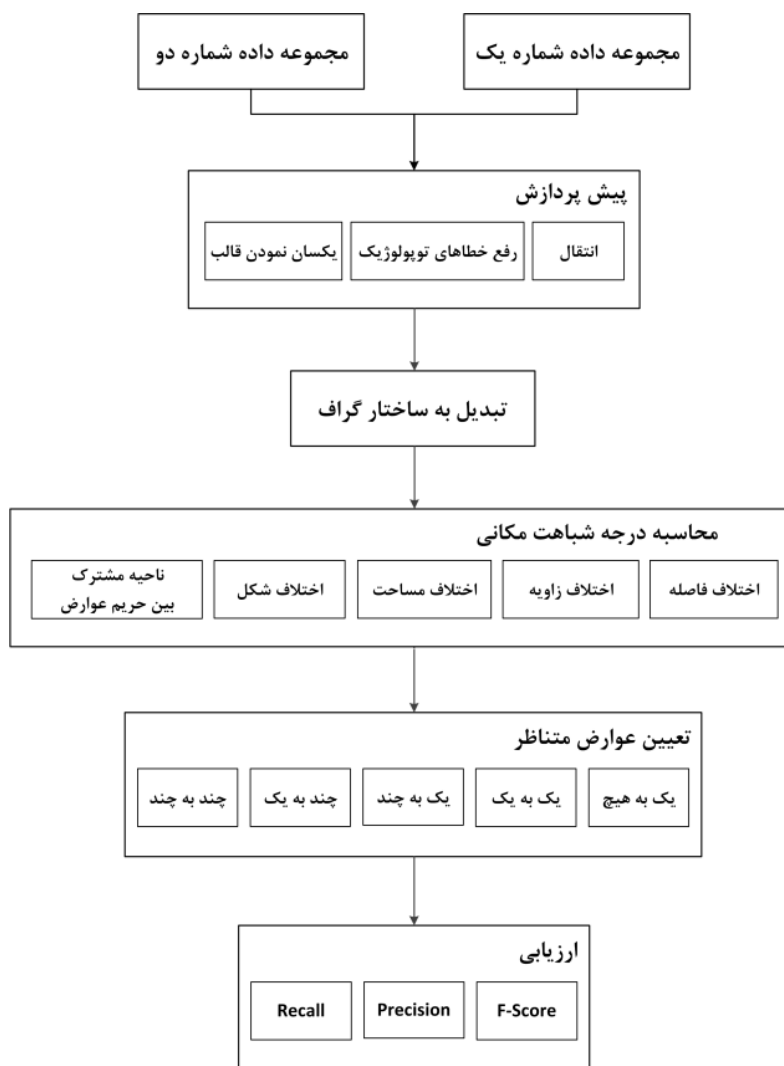
اتصال عوارض در سطوح مختلف جزئیات یکی از مسائل اصلی در ایجاد پایگاه داده چند مقیاسی می‌باشد که از آن به عنوان تناظریایی داده یاد می‌شود. در واقع تناظریایی را می‌توان فرایند جستجوی عوارض متناظر از میان مجموعه داده‌های مختلف که دارای ارتباط مکانی و مشابهت معنایی قوی هستند، دانست [۱]. از این رو در این تحقیق برای پیدا کردن عوارض متناظر در دو مجموعه داده ساختاری مطابق شکل (۱) پیشنهاد می‌گردد که شامل پنج مرحله پیش پردازش، تبدیل به ساختار گراف، محاسبه درجه شباهت مکانی، تناظریایی و ارزیابی می‌باشد.

برای تناظریایی مجموعه داده‌ها با مقیاسها و منابع متفاوت، در ابتدا نیاز است که یک مرحله پیش پردازش بر روی مجموعه داده‌ها صورت گیرد. با توجه به شکل (۱) این پیش پردازش شامل تبدیل به قالب یکسان برای هر دو مجموعه داده‌ها، رفع خطاهای توپولوژیک شبکه راهها و یکسان‌سازی سیستم مختصات می‌باشد که منجر به کاهش خطاهای سیستماتیک در مجموعه داده‌ها می‌شود [۵]. پس از اعمال پیش پردازش، شبکه راهها به عنوان یک ساختار گراف در نظر گرفته می‌شود که دارای خصوصیات از پیش تعریف شده‌ای می‌باشد (بخش ۳-۱). پس از اینکه شبکه راهها و عوارض موجود در آن مطابق خصوصیات ساختار گراف پیشنهادی گردید، برای هریک از عوارض تعیین شده در مجموعه داده مرجع با استفاده از مقدار فاصله بافر (از طریق دقت مکانی مجموعه داده بدست می‌آید) عوارض همسایگی در مجموعه هدف تعیین می‌شود. برای مثال با توجه به شکل (۲) برای عارضه PL

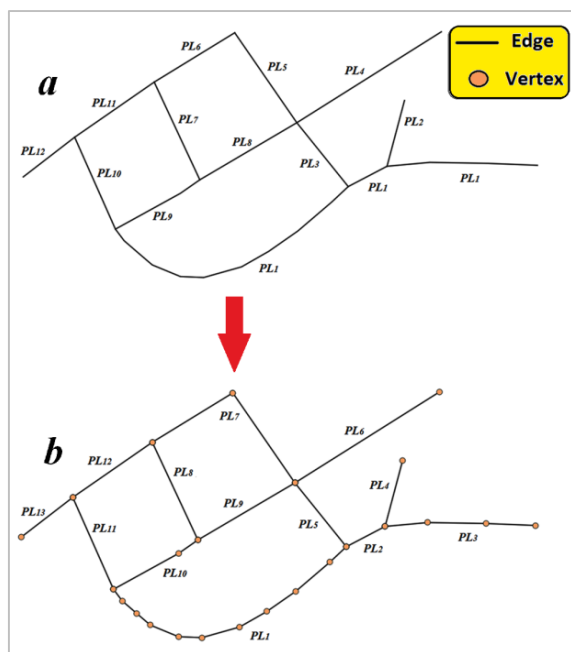
در مجموعه داده مرجع هریک از عوارض PL_1, PL_2, PL_3 , $PL_4, PL_5, PL_6, PL_7, PL_8, PL_9$ و PL_{10} از مجموعه داده هدف انتخاب می‌گردند. در نتیجه ضروری است که مقدار درجه شباهت مکانی هریک از عوارض مجموعه هدف (در محدوده بافر) با عارضه مجموعه مرجع تعیین گردد تا عوارض متناظر مشخص شوند.

۳-۱- ساختار گراف

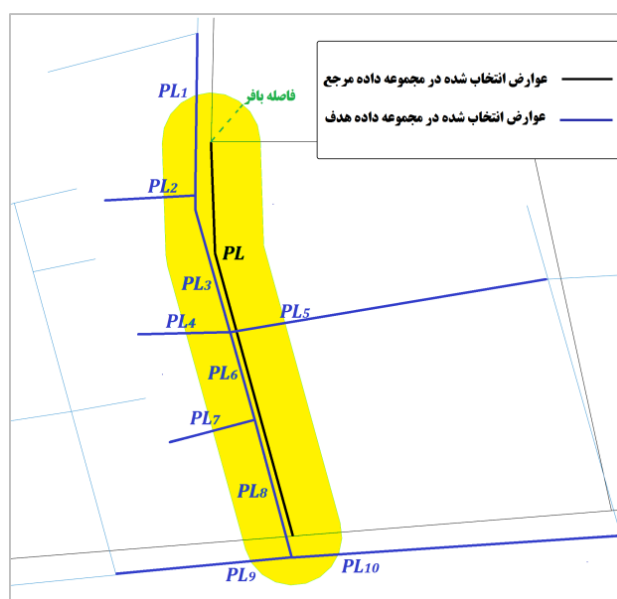
از تئوری گرافها برای توصیف شبکه راهها به عنوان یکسری نقاط و خطوط متصل به هم استفاده شده است [۲۵]. در ریاضیات گراف مرتبط با خطوط شبکه می‌تواند بصورت زوج مرتب $G = (V, E)$ نمایش داده شود که V شامل مجموعه رئوس شبکه و E شامل مجموعه یالهای شبکه می‌باشد. هر یال توسط یک جفت از رئوس قابل شناسایی می‌باشد که درجه هر یک از این رئوس، تعداد یال‌های متصل به آن می‌باشد [۲۶]. در شبکه راههای شهری هر چندخطی شامل چندین رأس و یال می‌باشد. چندخطی مفروض PL_i شامل نقاط $P_{i,1}, P_{i,2}, \dots, P_{i,n}$ می‌باشد که هر دو نقطه $P_{i,j}$ و $P_{i,j+1}$ یک یال از چندخطی را تشکیل می‌دهند که نقاط $P_{i,1}$ و $P_{i,n}$ نقاط ابتدایی و انتهایی چندخطی PL_i می‌باشند [۱۳]. اما در این تحقیق عوارض بصورتی تعریف می‌گردد که در ساختار گراف این چندخطی‌ها درجه ابتدایی و یا انتهایی رئوس آن یک و یا بیشتر از دو باشد. در نتیجه در فرایند تناظریایی هر یک از این چندخطی‌ها بصورت یک عارضه در نظر گرفته می‌شود. بنابراین هر چندخطی ممکن است با یک تقاطع شروع و یا پایان یابد ولی هرگز شامل یک تقاطع در نقاط میانی نخواهد بود [۱۳]. برای مثال در شکل (۳-a) چندخطی‌های PL_5 و PL_6 تبدیل به PL_7 در شکل (۳-b) شده‌اند و یا PL_1 در شکل (۳-a) تبدیل به PL_2, PL_1 و PL_3 در شکل (۳-b) شده‌اند.



شکل ۱- ساختار پیشنهادی جهت تناظریابی عوارض در مجموعه داده‌های چند مقیاسی



شکل ۳- ساختار گراف و نمونه‌ای از پلی‌لاین‌های در نظر گرفته شده



شکل ۲- تعیین عوارض کاندید در مجموعه داده هدف

۳-۲- روابط بین عوارض

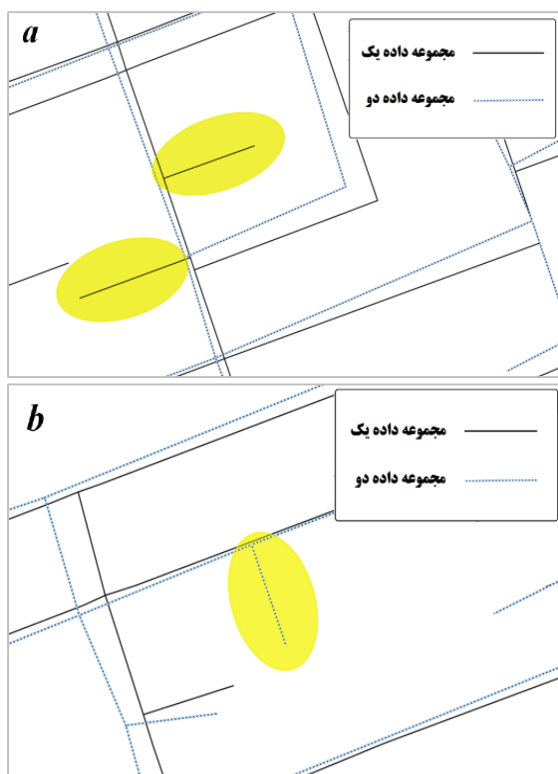
در هنگام تناظریایی مجموعه داده‌های مکانی ممکن است نمایش عارضه‌ای در یکی از مجموعه‌ها نسبت به دیگری متفاوت باشد، مخصوصاً هنگامی که مجموعه داده‌ها از سازمان‌های مختلف، با مقیاس متفاوت و با سطح جزئیات متفاوت تهیه شده باشد، از این رو استراتژی‌های تناظریایی بایستی عوارض متناظر را در این نوع از تلفیق‌ها شناسایی کند [۱]. بنابراین یکی از عوامل مهم در ساختار پیشنهادی، توانایی شناسایی تمامی روابط ممکن بین عوارض متناظر در فرایند تناظریایی می‌باشد [۲۷]. از این رو تمامی حالات دسته بندی ارائه شده برای ارتباطات در تناظریایی در تحقیقات مختلف [۱، ۱۳، ۱۸] در چهار حالت ذیل در نظر گرفته می‌شود.

۱- رابطه یک به هیچ^۱ یا هیچ به یک^۲ (۱:۰ یا ۰:۱): در این حالت برای عارضه‌ای در یک مجموعه داده، عارضه مشابهی در مجموعه دیگر وجود نداشته باشد. شکل (۴-a) دو عارضه در مجموعه داده یک را نشان می‌دهد که هیچ عارضه مشابهی در مجموعه داده شماره دو ندارند. همچنین شکل (۴-b) یک عارضه را در مجموعه شماره دو نشان می‌دهد که هیچ عارضه مشابهی در مجموعه یک ندارد.

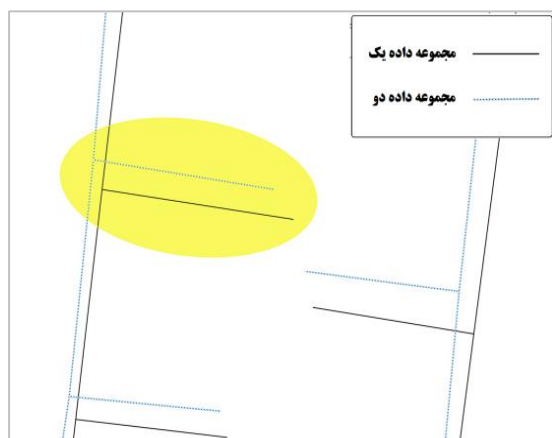
۲- رابطه یک به یک^۳ (۱:۱): برای هر عارضه در یک مجموعه داده یک عارضه مشابه با آن در مجموعه داده دیگر وجود دارد. شکل (۵) نمونه‌ای از این رابطه را در بین دو مجموعه داده نشان می‌دهد.

۳- رابطه یک به چند^۴ یا چند به یک^۵ (۱:N یا N:۱): به ازای هر عارضه در یک مجموعه داده، دو و یا چند عارضه در مجموعه داده دیگر وجود دارد. شکل (۶-a) دو عارضه را در مجموعه یک نشان می‌دهد که معادل آن در مجموعه دوم یک عارضه می‌باشد. همچنین شکل (۶-b) سه عارضه را در مجموعه دو نشان می‌دهد که متناظر آن در مجموعه اول یک عارضه می‌باشد.

۴- رابطه چند به چند^۶ (M:N): در این حالت گروهی از عوارض در یک مجموعه داده متناظر با گروهی دیگر از عوارض در مجموعه داده دوم می‌باشند. شکل (۷) دو عارضه از مجموعه اول را که معادل با دو عارضه از مجموعه دوم می‌باشد، نشان می‌دهد.



شکل ۴- رابطه ۱:۰ و ۰:۱ موجود بین دو مجموعه داده



شکل ۵- رابطه ۱:۱ موجود بین دو مجموعه داده

^۱ One-to-Null

^۲ Null-to-One

^۳ One-to-One

^۴ One-to-Many

^۵ Many-to-One

^۶ Many-to-Many

مشترک بین حریم عوارض می باشد که در ادامه تفسیر می گردد. فرض کنید PL_1 و PL_2 دو چندخطی در دو مجموعه داده با مقیاس و منبع داده متفاوت می باشند که به یک ماهیت در دنیای واقعی اشاره می کنند. در این صورت برای چندخطی PL_2 جدولی به نام جدول مقایسه معادل جدول (۱) وجود دارد، هرچند در عمل این جدول برای گروهی از عوارض دارای چندین سطر خواهد بود.

جدول ۱- جدول مقایسه عوارض

عوارض	معیارهای هندسی				
	C_1	C_2	C_3	C_4	C_5
PL_1, PL_2	$\overline{Sim}_{PL_1, PL_2}^{C_1}$	$\overline{Sim}_{PL_1, PL_2}^{C_2}$	$\overline{Sim}_{PL_1, PL_2}^{C_3}$	$\overline{Sim}_{PL_1, PL_2}^{C_4}$	$\overline{Sim}_{PL_1, PL_2}^{C_5}$

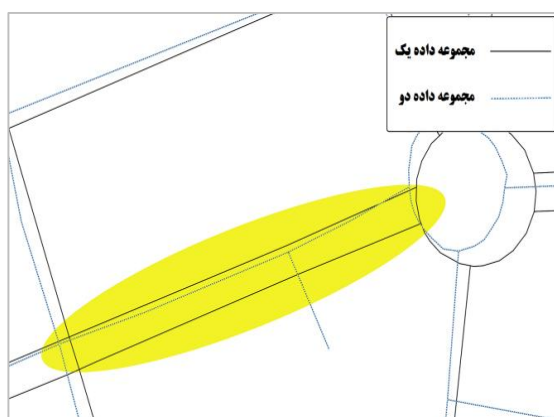
در جدول (۱) $\overline{Sim}_{PL_1, PL_2}^{C_i}$ مقدار بدست آمده برای معیار هندسی i ام می باشد که از طریق روابط (۲)، (۳)، (۴) و (۵) محاسبه می شود. پس از محاسبه مقادیر هریک از معیارهای هندسی نکته ای که قابل اهمیت است مقادیر بدست آمده برای $\overline{Sim}_{PL_1, PL_2}^{C_i}$ می باشد، زیرا این مقادیر نیاز است تا به مقادیر قابل مقایسه تبدیل گردد تا بتوان براساس آنها درجه شباهت چندخطی ها به یکدیگر را بدست آورد. بنابراین در رابطه (۱) $\overline{Sim}_{PL_1, PL_2}^{C_i}$ مقدار نرمال شده $\overline{Sim}_{PL_1, PL_2}^{C_i}$ می باشد. رابطه (۱) نحوه محاسبه درجه شباهت مکانی PL_2 به PL_1 را نشان می دهد [۲۹].

$$Sim(PL_1, PL_2) = \frac{\sum_{i=1}^5 W_i \overline{Sim}_{PL_1, PL_2}^{C_i}}{\sum_{i=1}^5 W_i} \times 100 \quad (1)$$

در رابطه (۱) $Sim(PL_1, PL_2)$ درصد شباهت مکانی بین چندخطی PL_2 و PL_1 می باشد که دارای مقادیری بین ۰ تا ۱۰۰ است. $\overline{Sim}_{PL_1, PL_2}^{C_i}$ مقادیر نرمال شده معیارهای هندسی بدست آمده از روابط (۲)، (۳)، (۴) و (۵) و W_i وزن معیار هندسی i ام در محاسبه درجه شباهت مکانی می باشد. همچنین C_1 تا C_5 به ترتیب معیارهای فاصله، جهت، مساحت، شکل و ناحیه مشترک بین حریم عوارض می باشد. برای تعیین وزن هریک از معیارهای هندسی از منظر ادراک مکانی انسان از طریق پرسشنامه قضاوت کارشناسان مرتبط در مورد اهمیت هریک از معیارهای هندسی در سنجش میزان شباهت دو عارضه در نظر گرفته می شود.



شکل ۶- رابطه $1:N$ و $N:1$ موجود بین دو مجموعه داده.



شکل ۷- رابطه $M:N$ موجود بین دو مجموعه داده

۳-۳- درجه شباهت مکانی عوارض

از جنبه کلی می توان شباهت عوارض به یکدیگر را با در نظر گرفتن معیارهای هندسی، توپولوژیک و معنایی مورد سنجش قرار داد [۱، ۲۸]. با این وجود در این تحقیق از هندسه عوارض جهت تعیین درجه شباهت استفاده می شود، زیرا در تمامی مجموعه داده ها امکان دسترسی به این معیارها وجود دارد و نیاز به داده اضافه به جز هندسه عارضه مورد نظر نمی باشد. معیارهای در نظر گرفته شده شامل فاصله، جهت، مساحت، شکل و ناحیه

۳-۳-۱- فاصله

فاصله مکانی یک پارامتر مهم برای اندازه گیری نسبی موقعیت های عوارض می باشد که می تواند درجه شباهت آن عارضه با عوارض مجاور را نشان دهد [۱۹]. نزدیکی فاصله بین دو عارضه معمولاً به عنوان نوعی اندازه گیری شباهت هندسی در تناظریایی عوارض استفاده می شود. برای مثال فاصله هاسدورف یکی از پرکاربردترین این فواصل در علوم مکانی می باشد [۴]. مناسب بودن فاصله هاسدورف در محاسبه نزدیکی بین عوارض خطی به اثبات رسیده است [۳۰]. این فاصله به عنوان بیشترین فاصله بین کوتاهترین فاصله موجود بین هر نقطه از عارضه اول و مجموعه نقاط عارضه دوم تعریف شده است [۱۸]. با این حال فاصله هاسدورف نسبت به شکل دو عارضه، به خصوص نسبت به قسمت های دور افتاده از مرکز حساس می باشد. تحقیق [۴] با معرفی هاسدورف میانه بر مبنای طول نشان داد که این روش در مقایسه با روش های هاسدورف و هاسدورف میانه دارای واریانس کمتر و عملکرد کارآمدتری در مواجهه با داده های پیچیده جهت اندازه گیری فاصله بین چندخطی ها می باشد. رابطه (۲) فاصله هاسدورف میانه بر مبنای طول بین دو چندخطی مفروض را نشان می دهد [۴].

$$\overline{\text{Sim}}_{PL_1, PL_2}^{C_1} = \begin{cases} m(PL_1, PL_2), & \text{if } L(PL_1) < L(PL_2) \\ m(PL_2, PL_1), & \text{if } L(PL_1) \geq L(PL_2) \end{cases} \quad (2)$$

که در این رابطه $L(PL_1)$ و $L(PL_2)$ طول دو عارضه خطی PL_1 و PL_2 ، $m(PL_1, PL_2)$ و $m(PL_2, PL_1)$ نیز از طریق روابط (۳) و (۴) بدست می آیند.

$$m(PL_1, PL_2) = \text{median}_{P_a \in PL_1} \{ \min_{P_b \in PL_2} \|P_a - L_b\| \} \quad (3)$$

$$m(PL_2, PL_1) = \text{median}_{P_b \in PL_2} \{ \min_{P_a \in PL_1} \|P_b - L_a\| \} \quad (4)$$

در این روابط L_a و L_b دو یال اختیاری از عوارض خطی PL_1 و PL_2 ، $\|P_a - L_b\|$ فاصله عمودی یکی از نقاط عارضه PL_1 (P_a) از یکی از یال های عارضه PL_2 (L_b) و $\|P_b - L_a\|$ فاصله عمودی یکی از نقاط عارضه PL_2 (P_b) از یکی از یال های عارضه PL_1 (L_a) می باشد.

۳-۳-۲- جهت

یکی دیگر از خصوصیات هندسی که می توان برای مسئله تناظریایی عوارض خطی مور استفاده قرار داد،

اختلاف جهت عوارض خطی از یکدیگر می باشد که می تواند نقش مهمی را در مسئله ایفا کند [۳۱]. برای یک عارضه خطی جهت به زاویه بین خط تشکیل شده از گره های ابتدایی و انتهای عارضه و محور افقی گفته می شود. در این صورت برای دو عارضه خطی با جهت های α و β اختلاف جهت برابر $|\alpha - \beta|$ خواهد بود. در صورتیکه مقدار $|\alpha - \beta|$ برابر صفر باشد، دو عارضه تقریباً هم راستا هستند و در صورتیکه مقدار $|\alpha - \beta|$ برابر π باشد، راستای دو عارضه خطی در خلاف جهت یکدیگر می باشد [۱].

۳-۳-۳- مساحت

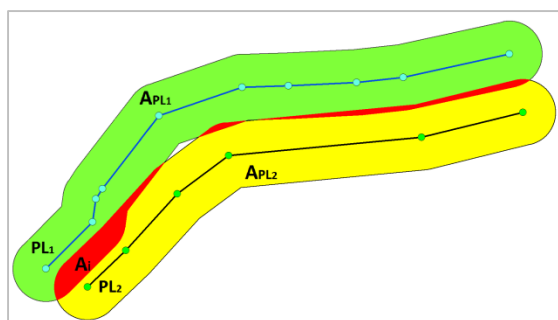
از طریق اتصال گره ابتدایی به انتهای در عوارض خطی یک چند ضلعی ایجاد می شود که می توان اختلاف بین مساحت های آن را به عنوان مبنایی جهت مقایسه عوارض مختلف با یکدیگر مورد استفاده قرار داد. در حقیقت این معیار بر مبنای محاسبه اختلاف مساحت شکل می گیرد و صرفاً اختلاف مساحت دو عارضه به عنوان مقدار معیار در نظر گرفته نمی شود. فرض کنید PL_1 و PL_2 دو عارضه خطی در مقیاسها و منابع تهیه متفاوت هستند که به یک ماهیت در دنیای واقعی اشاره می کنند. در این صورت رابطه (۶) نحوه محاسبه معیار مساحت بین این دو عارضه خطی را نشان می دهد [۱].

$$\overline{\text{Sim}}_{PL_1, PL_2}^{C_3} = \left| \frac{S_1}{D_1} - \frac{S_2}{D_2} \right| \quad (6)$$

در این رابطه S_1 و S_2 مساحت های ایجاد شده برای دو عارضه خطی PL_1 و PL_2 می باشد. همچنین D_1 و D_2 فاصله اقلیدسی بین گره های ابتدایی و انتهای در دو عارضه خطی PL_1 و PL_2 می باشد.

۳-۳-۴- شکل

عوارض خطی در مقیاسهای مختلف می توانند از نظر شکل با یکدیگر متفاوت باشند. در نظر گرفتن این تفاوت می تواند به عنوان یک معیار هندسی در تعیین مقدار درجه شباهت عوارض به یکدیگر مورد استفاده قرار گیرد. یکی از شناخته شده ترین و پرکاربردترین توصیفگرهای مرتبط با شکل عوارض، توصیفگر تابع تجمعی زاویه می باشد که با نام تابع پیچش نیز شناخته می شود [۳۲]. برای مثال شکل (۸) تابع پیچش دو عارضه خطی PL_1 و PL_2 را نشان می دهد. در



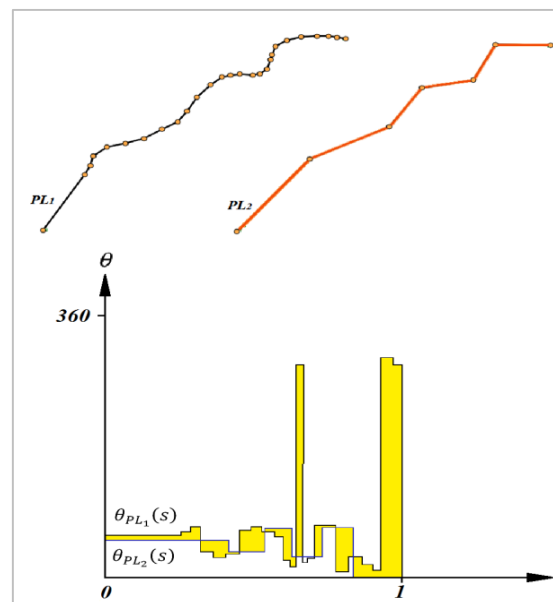
شکل ۹- ناحیه مشترک بین حریم عوارض

۴- پیاده سازی

برای تعیین عوارض متناظر با رویکرد پیشنهادی از دو مجموعه داده با مقیاسها و منابع تهیه متفاوت استفاده گردید تا کارایی رویکرد پیشنهادی به همراه میزان تاثیر هریک از معیارها در تعیین عوارض متناظر در مقیاسها و منابع داده مختلف سنجیده شود. در این تحقیق، از دو مجموعه داده شبکه راههای شهری استفاده می شود که یک مجموعه داده دارای دقت مکانی بالاتر و مجموعه داده دیگر دارای اطلاعات توصیفی غنی تری می باشد. هدف تهیه مجموعه داده جدیدی است که دقت مکانی بالای عوارض به همراه اطلاعات توصیفی غنی در آن به اشتراک گذاشته شده است. از این رو با توجه به شکل (۱۰) شبکه راههای شهری قسمتی از منطقه شش تهران به عنوان منطقه مورد مطالعه انتخاب گردید. در شکل (۱۰) مجموعه اول دارای مقیاس ۱:۵۰۰۰ و در سال ۲۰۱۲ و مجموعه دوم دارای مقیاس ۱:۲۵۰۰۰ و در سال ۲۰۰۱ تهیه شده است. مجموع طول راهها در مجموعه داده اول برابر ۷۴۹۳۳/۰۷ متر و در مجموع داده دوم برابر ۶۳۲۳۱/۸۴ متر می باشد. در ابتدا برای اینکه بتوان داده های متناظر را در هر دو گروه تعیین نمود، نیاز است بر روی داده ها عملیات پیش پردازش صورت گیرد. عملیات پیش پردازش شامل هم قالب نمودن داده ها، بردن به سیستم مختصات یکسان و حذف خطاهای توپولوژیکی در مجموعه داده ها می باشد. پس از انجام عملیات پیش پردازش بر روی داده ها، تعداد عوارض موجود در داده های با مقیاسهای ۱:۵۰۰۰ و ۱:۲۵۰۰۰ به ترتیب برابر ۷۷۹ عارضه و ۵۵۸ عارضه گردید.

با توجه به شکل (۱) در بخش ۳ گام دوم از رویکرد پیشنهادی، تبدیل نمودن مجموعه داده ها به ساختار گراف می باشد که در بخش ۳-۱ تشریح گردید. این مرحله به این

این تابع برای هر کدام از گره ها میزان چرخش یال متصل به گره نسبت به محور افقی در نظر گرفته می شود. در انتها نیز مقدار مساحت محصور بین دو تابع پیچش به عنوان اختلاف شکل دو عارضه خطی محاسبه می گردد [۱].



شکل ۸- تابع پیچش دو عارضه خطی PL₁ و PL₂

۳-۳-۵- ناحیه مشترک بین حریم عوارض

از جمله معیارهای دیگر مورد استفاده در تحقیقات مختلف، در نظر گرفتن مساحت موجود بین منطقه مشترک بوجود آمده از حریم ایجاد شده بین عوارض می باشد. برای مثال در شکل (۹) مساحت منطقه قرمز رنگ می تواند معیاری برای سنجش میزان شباهت بین دو عارضه خطی PL₁ و PL₂ باشد [۳۳]. رابطه (۱۰) نحوه محاسبه اختلاف مساحت بین حریم عوارض را نشان می دهد که هرچه قدر مقدار این مساحت به یک نزدیکتر باشد دو عارضه از نظر هندسی به یکدیگر شبیه تر هستند.

$$C_5 = \frac{2A_i}{A_{PL1} + A_{PL2}} \quad (10)$$

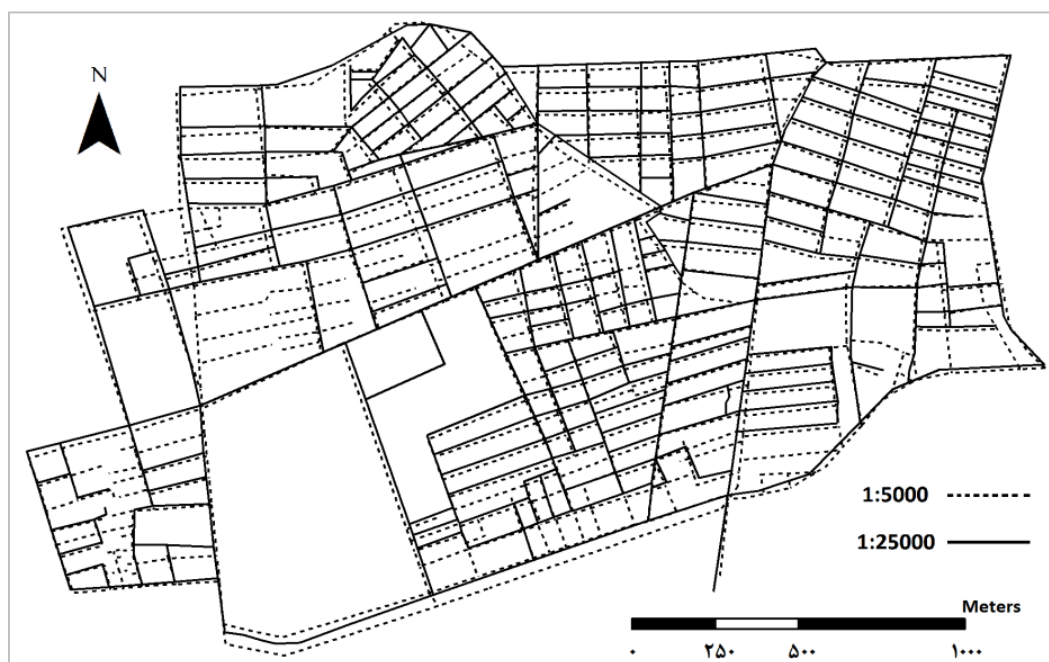
در این رابطه A_{PL1} مساحت حریم ایجاد شده برای عارضه اول، A_{PL2} مساحت حریم ایجاد شده برای عارضه دوم و A_i مساحت منطقه مشترک بین دو حریم ایجاد شده می باشد.

دلیل اعمال می‌گردد که وجود ابهام در تعریف عوارض از بین رفته و تعریف مشخصی از عارضه ارائه گردد. طبق بخش ۱-۳ هر پلی‌لاین ممکن است با یک تقاطع شروع و یا پایان یابد ولی هرگز شامل یک تقاطع در نقاط میانی نخواهد بود. پس از انجام این گام تعداد عوارض در مجموعه اول و دوم به ترتیب برابر ۷۷۳ و ۵۵۳ عارضه می‌گردد.

در ادامه برای تعیین وزن هریک از معیارهای هندسی از منظر ادراک مکانی انسان از طریق پرسشنامه قضاوت کارشناسان مرتبط (دانشجویان و فارغ التحصیلان مقطع کارشناسی (نقشه‌برداری)، کارشناسی ارشد و دکتری

(GIS)) در مورد اهمیت هریک از معیارهای هندسی در سنجش میزان شباهت دو عارضه در نظر گرفته شد، میانگین وزنه‌های اخذ شده به عنوان وزن نهایی هریک از معیارها محاسبه گردید. جدول (۲) وزن نهایی استخراج شده را نشان می‌دهد.

جدول ۲- وزن نهایی هریک از معیارها				
معیارهای هندسی				
C ₁	C ₂	C ₃	C ₄	C ₅
۰/۶۲	۰/۳۱	۰/۴۵	۰/۷۰	۰/۶۵



شکل ۱۰- مجموعه داده‌های مورد استفاده با مقیاس و منابع تهیه متفاوت

همانطور که بیان گردید هدف از این تحقیق یافتن عوارض متناظر مجموعه داده دوم در مجموعه اول می‌باشد. از این رو با توجه به رویکرد پیشنهادی در این مرحله نیاز است که برای عارضه در مجموعه دوم، عوارض کاندید در مجموعه اول انتخاب گردد. برای این امر نیاز است تا فاصله عرض حریم تعیین گردد. برای تعیین عرض حریم از دقت دو مجموعه داده و نشریه مشخصات فنی نقشه‌برداری شماره ۹۵ استفاده می‌شود. از این رو عرض حریم‌گیری با توجه به رابطه‌ی (۱۱) محاسبه می‌گردد [۲۱].

$$W_{\text{buffer}} = \sqrt{\sigma_a^2 + \sigma_b^2} \quad (11)$$

در این رابطه پارامترهای σ_a و σ_b برابر با دقت مجموعه‌داده اول و دوم می‌باشند. از آنجایی که مقیاس

مجموعه داده‌های مورد استفاده در این تحقیق برابر ۱/۲۵۰۰۰ و ۱/۵۰۰۰ می‌باشد و با توجه به بازه اطمینان ۹۰٪ در نشریه مشخصات فنی نقشه‌برداری شماره ۹۵، عرض حریم‌گیری برابر ۳۳/۱۴ متر بدست می‌آید.

پس از اینکه فاصله عرض حریم‌گیری تعیین شد، برای هر کدام از عوارض موجود در مجموعه داده دوم عوارض کاندید در مجموعه اول تعیین می‌شود. برای مثال در شکل (۱۱) عارضه PL در مجموعه دوم دارای کاندیدهای PL₁ تا PL₁₇ می‌باشد. این شکل نشان دهنده یک رابطه یک به چند (یک به شش) در مجموعه داده‌ها است، در اصل عارضه PL در مجموعه دوم با عوارض PL₄، PL₆، PL₉، PL₁₁، PL₁₃ و PL₁₄ در مجموعه اول متناظر می‌باشد. پس از تعیین کاندیدهای تناظریایی، معیارهای پنجگانه هندسی معرفی

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (12)$$

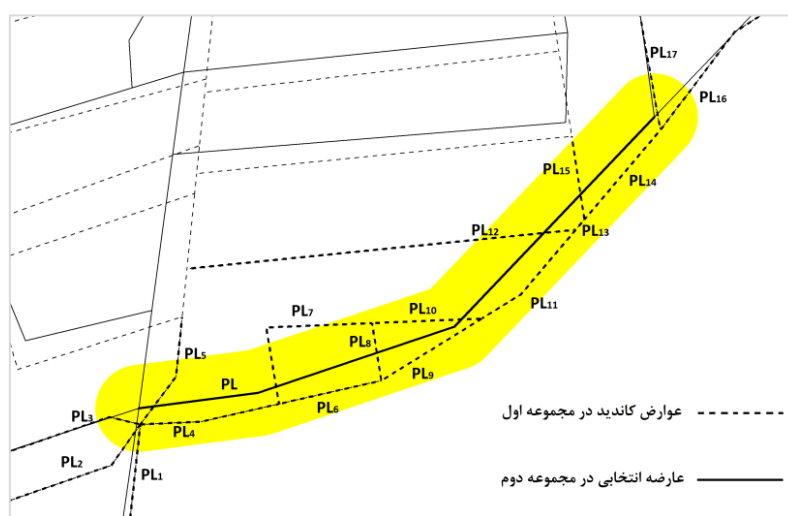
$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (13)$$

$$F - \text{Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (14)$$

در این روابط TP تعداد روابطی هستند که توسط رویکرد پیشنهادی به درستی کشف شده‌اند، FP تعداد روابطی می‌باشد که به اشتباه توسط رویکرد پیشنهادی تعیین شده است، FN تعداد روابطی است که توسط رویکرد پیشنهادی کشف نشده است و F - Score یک کمیتی است که در بر گیرنده هر دو مقدار Precision و Recall می‌باشد و برای ارزیابی نهایی صورت می‌گیرد.

شده برای هریک از عوارض محاسبه می‌شود. جدول (۳) مقادیر محاسبه شده هریک از معیارها را برای هر کدام از عوارض کاندید در مقایسه با عارضه مرجع یعنی عارضه PL نشان می‌دهد. در تعیین عوارض متناظر از مقدار تجربی حد آستانه شباهت برابر ۸۵ درصد استفاده شده است.

در نتیجه با توجه به فرایند تشریح شده برای تمامی عوارض در مجموعه داده دوم، عوارض متناظر در مجموعه داده اول شناسایی می‌گردد. برای ارزیابی رویکرد پیشنهادی از سه پارامتر Precision، Recall و F - Score استفاده می‌گردد که از طریق روابط (۱۲)، (۱۳) و (۱۴) محاسبه می‌شوند.



شکل ۱۱- تعیین عوارض کاندید در مجموعه داده اول برای عارضه منتخب در مجموعه داده دوم

جدول ۳- مقادیر محاسبه شده برای هریک از معیارها برای هر کدام از عوارض کاندید در مقایسه با عارضه مرجع

جفت عوارض	مقادیر معیارهای هندسی					
	فاصله	جهت	مساحت	شکل	ناحیه مشترک حریم‌گیری	درصد شباهت
PL - PL ₁	۱۳/۸۳	۱۲۶/۶۲	۳۵/۵۵	۶۹/۵۸	۰/۱۰	۲۴/۰۴
PL - PL ₂	۷/۵۶	۱۷۱/۳۱	۲۵/۷۱	۱۵/۴۱	۰/۰۵	۵۵/۲۵
PL - PL ₃	۵/۷۲	۱۶۶/۶۴	۲۹/۴۷	۸/۰۱	۰/۱۱	۵۶/۰۶
PL - PL₄	۷/۱۹	۲۱/۰۳	۱۳/۱۷	۵/۲۱	۰/۵۶	۹۳/۷۲
PL - PL ₅	۲۰/۳۴	۳۹/۳۳	۲۹/۳۵	۶۱/۳۶	۰/۲۲	۳۰/۵۱
PL - PL₆	۶/۷۶	۱۶/۶۰	۱۵/۵۵	۵/۴۵	۰/۵۰	۹۱/۱۷
PL - PL ₇	۲۲/۸۵	۱۱/۵۳	۱۰/۰۸	۴۵/۶۹	۰/۳۱	۴۵/۹۹
PL - PL ₈	۱۲/۸۹	۶۹/۷۵	۳۵/۵۵	۸۸/۲۱	۰/۲۱	۲۷/۱۳
PL - PL₉	۱/۸۴	۲/۲۴	۱۵/۵۵	۷/۲۹	۰/۵۴	۹۳/۰۸
PL - PL ₁₀	۱۱/۸۰	۲۷/۰۶	۳۵/۵۵	۵/۰۲	۰/۳۲	۶۷/۳۰
PL - PL₁₁	۷/۲۷	۱۳/۹۲	۱۱/۹۶	۴/۰۲	۰/۵۴	۹۳/۹۵
PL - PL ₁₂	۵۸/۹۱	۱۵۹/۲۸	۱۵/۵۵	۱۳/۹۲	۰/۲۰	۳۰/۲۷
PL - PL₁₃	۱۴/۷۵	۲۰/۲۲	۱۵/۵۵	۱۲/۲۷	۰/۴۷	۸۶/۱۴
PL - PL₁₄	۱۱/۸۱	۲۰/۲۲	۱۵/۵۵	۱۲/۲۷	۰/۵۴	۸۹/۴۸
PL - PL ₁₅	۲۵/۶۶	۶۹/۲۴	۱۵/۵۵	۸۸/۷۱	۰/۲۴	۱۵/۳۰
PL - PL ₁₆	۸/۰۲	۱۷/۸۴	۲۰/۹۴	۱۶/۵۶	۰/۱۴	۶۹/۶۳
PL - PL ₁₇	۷۷/۸۲	۷۰/۶۷	۳۵/۵۵	۸۶/۸۹	۰/۱۴	۴/۸۵

برای نشان داده کارایی رویکرد پیشنهادی بر روی مجموعه داده‌های مورد مطالعه ابتدا میزان شناسایی جفت عوارض متناظر توسط هر یک از معیارهای هندسی مورد بررسی قرار گرفت. جدول (۴) مقدار Precision (P)، Recall (R) و F – Score (F) نهایی بدست آمده توسط هریک از معیارهای پنج‌گانه (هر بار تنها یک معیار در نظر گرفته شده است) را نشان می‌دهد.

جدول ۴- نتیجه تناظریایی با استفاده از هریک از معیارهای هندسی

معیارها	P	R	F
C ₁	۳۱/۱	۳۰/۳	۳۰/۶۹
C ₂	۳۰	۲۹/۴۰	۲۹/۷۰
C ₃	۱۰	۱۱/۸۰	۱۰/۸۳
C ₄	۳۳/۶۰	۳۲/۶۰	۳۳/۰۹
C ₅	۴۱/۲۰	۴۳/۷۰	۴۲/۴۱

با در نظر گرفتن هریک از معیارها (به تنهایی) جهت شناسایی عوارض متناظر بر روی منطقه مورد مطالعه مشخص گردید که معیار پنجم (منطقه مشترک چند ضلعی حاصل از حریم گیری) دارای بیشترین Precision و Recall در شناسایی عوارض متناظر می‌باشد. با این وجود به دلیل مقدار پائین F – Score نهایی، نیاز است تا معیارها در کنار یکدیگر مورد استفاده قرار گیرد.

جدول (۵) Precision، Recall، F – Score و زمان محاسبات (CT) بدست آمده از در نظر گرفتن معیارهای هندسی مختلف در کنار یکدیگر را نشان می‌دهد.

جدول ۵- نتیجه تناظریایی از ترکیب معیارها

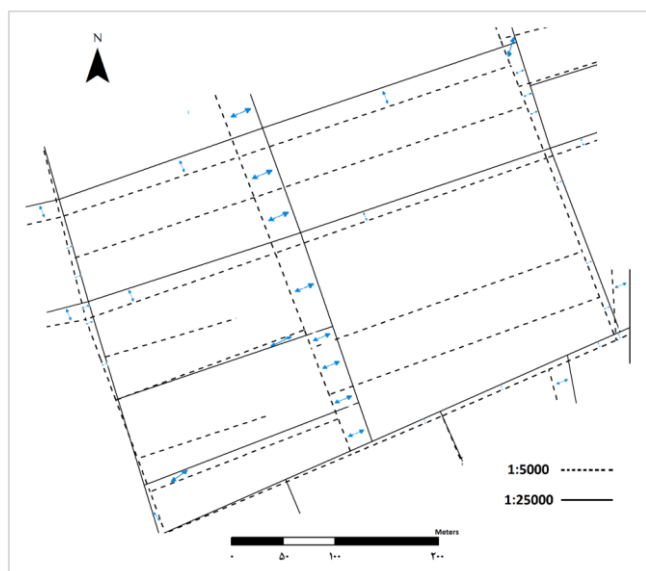
معیارها	P	R	F	CT (Sec)
C ₂ - C ₁	۳۶/۱۰	۳۴	۳۵/۰۲	۱۸۷/۸
C ₃ - C ₂ - C ₁	۴۱/۳۰	۴۰/۷۰	۴۱	۱۹۵/۵
C ₄ - C ₃ - C ₂ - C ₁	۵۵	۶۹/۴۵	۶۱/۳۹	۲۰۵/۳
C ₅ - C ₄ - C ₃ - C ₂ - C ₁	۸۶/۷۰	۷۹	۸۲/۶۷	۳۶۹/۷

RAM: 3 GB
CPU: Intel Core(TM) 2 Duo , T5870@2GHz

در مسائل تناظریایی تعیین عوارض متناظر چالش اصلی در مجموعه داده‌های با مقیاس و منابع تهیه متفاوت می‌باشد. از این رو در این تحقیق با ارائه رویکردی پیشنهادی ضمن رسیدن به مقدار F – Score نهایی برابر ۸۲/۶۷ درصد، تنها از معیارهای هندسی برای این امر استفاده گردید. بدیهی است در صورت اعمال رویکرد پیشنهادی فوق بر روی مجموعه داده‌های با مقیاس

یکسان و منابع متفاوت و یا مقیاس متفاوت و منابع یکسان (یکی حاصل خلاصه‌سازی دیگری) مقدار نهایی F – Score بصورت قابل ملاحظه‌ای افزایش خواهد یافت. نتایج نشان می‌دهد که متدولوژی پیشنهادی در مسئله تناظریایی عوارض بر روی مجموعه داده‌های واقعی شبکه راهها بسیار کاربردی می‌باشد و به خوبی عوارض متناظر را در در مجموعه داده‌های با مقیاسها و منابع مختلف شناسایی می‌کند. شکل (۱۲) نمونه‌ای از تناظریایی عوارض را در منطقه مورد مطالعه نشان می‌دهد.

جدول (۴) و (۵) نشان داد که در صورت استفاده از سایر معیارها به تنهایی امکان رسیدن به نتیجه مطلوب وجود نخواهد داشت و در صورت استفاده معیارها در کنار یکدیگر، مقدار F – Score از ۳۵/۰۲ درصد تا ۸۲/۶۷ بهبود خواهد یافت.



شکل ۱۲- تناظریایی عوارض در قسمتی از منطقه مورد مطالعه

۵- نتیجه‌گیری و پیشنهادها

هدف از این تحقیق ارائه رویکردی بر اساس معیارهای استخراج شده از ساختار هندسی عوارض برای شناسایی عوارض متناظر در مجموعه داده‌های با مقیاس و منبع تهیه متفاوت می‌باشد. برای این امر رویکردی پیشنهادی متشکل از پنج بخش ارائه گردید. در بخش اول ضمن بردن مجموعه داده‌ها به قالب یکسان و یکسان‌سازی سیستم تصویر و مختصات آنها، خطاهای توپولوژیک موجود در مجموعه داده‌ها نیز رفع گردید. در بخش دوم از رویکرد پیشنهادی با در نظر گرفتن یک ساختار گراف

با مقیاس و منابع تهیه متفاوت این امکان وجود ندارد. همچنین در این تحقیق ضمن از بین بردن ابهام در تعریف عوارض، با در نظر گرفتن پنج معیار هندسی استخراج شده از عوارض مقدار F – Score قابل قبول ۸۲/۶۷ درصد بدست آمد. نکته قابل توجه در این است که این مقدار بدون در نظر گرفتن سایر معیارها نظیر معیارهای معنایی و توپولوژیک و همچنین در مجموعه داده ها دارای اختلاف مقیاس بالا و منبع تهیه متفاوت بدست آمده است.

در تحقیقات آتی توصیه می شود که یک ساختار دو مرحله ای ارائه گردد تا در ابتدا عوارضی نظیر میدانها که از طریق تناظریابی خطی قابل شناسایی نیستند، بصورت جداگانه و با رویکردی متفاوت روابط آنها شناسایی شود. سپس سایر عوارض با رویکرد پیشنهادی وارد مرحله تناظریابی شوند.

تعریف شده، ابهام در تعریف عوارض خطی موجود در مجموعه داده ها از بین رفت. در بخش سوم و چهارم از طریق استخراج معیارهای هندسی از ساختار عوارض شامل طول، جهت، مساحت، شکل و منطقه مشترک حاصل از در نظر گرفتن حریم عوارض، مقدار درجه شباهت مکانی عوارض محاسبه و عوارض متناظر تعیین گردید. در بخش پایانی نیز از طریق مقادیر Precision ، Recall و F – Score رویکرد پیشنهادی مورد ارزیابی قرار گرفت.

نتایج این تحقیق نشان داد که برخلاف بسیاری از تحقیقات گذشته که با یک، دو و یا سه معیار هندسی استخراج شده از عوارض در مجموعه داده های با مقیاس متفاوت و منبع تهیه یکسان [۲۱] و یا مقیاس یکسان و منبع تهیه متفاوت [۴، ۹، ۱۸] که تناظریابی با دقت و صحت قابل قبولی صورت می گرفت، در مجموعه داده های

مراجع

- [1] M. Zhang, "Methods and implementations of road-network matching," Unpublished PhD Dissertation, Technical University of Munich, 2009.
- [2] M. F. Goodchild, "CHAPTER FOUR - Attribute accuracy," in Elements of spatial data quality, S. C. Guptill and J. L. Morrison, Eds., ed Amsterdam: Pergamon, 1995, pp. 59-79.
- [3] C.-C. Chen and C. A. Knoblock, "Conflation of geospatial data," in Encyclopedia of GIS, ed: Springer, 2008, pp. 133-140.
- [4] 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.
- [5] 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.
- [6] S. Mustière and T. Devogele, "Matching networks with different levels of detail," GeoInformatica, vol. 12, pp. 435-453, 2008.
- [7] B. Rosen and A. Saalfeld, "Match Criteria for Automatic Alignment," in Proceedings, Auto-Carto VII., 1985.
- [8] A. Saalfeld, "Conflation Automated map compilation," International Journal of Geographical Information Systems, vol. 2, pp. 217-228, 1988/01/01 1988.
- [9] L. Li and M. Goodchild, "Automatically and accurately matching objects in geospatial datasets," in Proceedings of joint international conference on theory, data handling and modelling in geospatial information science, Hong Kong, 2010, pp. 98-103.
- [10] A. J. Saalfeld, "Automated map conflation. Washington, DC: University of Maryland.," 1993.
- [11] M. Sester, G. Gösseln, and B. Kieler, "Identification and adjustment of corresponding objects in data sets of different origin," in 10th AGILE International Conference on Geographic Information Science 2007, 2007.
- [12] V. Walter and D. Fritsch, "Matching spatial data sets: a statistical approach," International Journal of Geographical Information Science, vol. 13, pp. 445-473, 1999.
- [13] E. Safra, Y. Kanza, Y. Sagiv, and Y. Doytshe, "Ad hoc matching of vectorial road networks," International Journal of Geographical Information Science, vol. 27, pp. 114-153, 2013.
- [14] 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.

- [15] 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.
- [16] R. M. Pendyala, *Development of GIS-based conflation tools for data integration and matching*: Florida Department of Transportation, 2002.
- [17] 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.
- [18] 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. ۲۰۱۱-۰۱/۱۲/۲۰۱۱, ۲۲۸-۳۰۹.
- [19] 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, 2007/04/01 2007.
- [20] G. Touya, A. Coupé, J. L. Jollec, O. Dorie, and F. Fuchs, "Conflation optimized by least squares to maintain geographic shapes," *ISPRS International Journal of Geo-Information*, vol. 2, pp. 621-644, 2013.
- [21] A. A. Farahanipooya, R. A. Abbaspour, S. K. Alavipanah, and M. Jafarbeiglu, "Roads matching in a multi-scale spatial database using a least square line," *Journal of Geomatics Science and Technology*, vol. 3, pp. 87-104, 2013.
- [22] A. Nayyeri and A. Sidiropoulos, "Computing the Fréchet Distance Between Polygons with Holes," in *Automata, Languages, and Programming*, ed: Springer, 2015, pp. 997-1009.
- [23] A. Mascaret, T. Devogele, I. Le Berre, and A. Hénaff, *Coastline matching process based on the discrete Fréchet distance*: Springer, 2006.
- [24] 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.
- [25] D. Watanabe, "A study on analyzing the grid road network patterns using relative neighborhood graph," in *The Ninth International Symposium on Operations Research and Its Applications*, 2010, pp. 112-119.
- [26] W. A. Mackaness and G. A. Mackechnie, "Automating the detection and simplification of junctions in road networks," *Geoinformatica*, vol. 3, pp. ۱۹۹۹, ۲۰۰-۱۸۵.
- [27] D. Sheeren, S. Mustière, and J.-D. Zucker, "How to Integrate Heterogeneous Spatial Databases in a Consistent Way?," in *Advances in Databases and Information Systems*. vol. 3255, A. Benczúr, J. Demetrovics, and G. Gottlob, Eds., ed: Springer Berlin Heidelberg, 2004, pp. 364-378.
- [28] H. Yan, "Fundamental theories of spatial similarity relations in multi-scale map spaces," *Chinese Geographical Science*, vol. 20, pp. 18-22, 2010.
- [29] H. Yan and J. Li, *Spatial Similarity Relations in Multi-scale Map Spaces*: Springer, 2014.
- [30] I. Abbas, "Base de données vectorielles et erreur cartographique: problèmes posés par le contrôle ponctuel; une méthode alternative fondée sur la distance de Hausdorff. Computer Science. Paris: Université de Paris VII.," 1994.
- [31] A.-M. Olteanu Raimond and S. Mustière, "Data Matching – a Matter of Belief," in *Headway in Spatial Data Handling*, A. Ruas and C. Gold, Eds., ed: Springer Berlin Heidelberg, 2008, pp. 501-519.
- [32] R. C. Veltkamp, "Shape matching: similarity measures and algorithms," in *Shape Modeling and Applications*, SMI 2001 International Conference on., 2001, pp. 188-197.
- [33] F. Zhonglianga and W. Jianhuaa, "ENTITY MATCHING IN VECTOR SPATIAL DATA," presented at the *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, China, 2008.