

الگوهای آلودگی هوای منطقه شهری با استفاده از مدل توسعه یافته روش داده کاوی هم‌مکان

محمد اکبری*^۱، فرهاد صمدزادگان^۲

^۱ استادیار گروه مهندسی عمران - دانشگاه بیرجند
moakbari@birjand.ac.ir

^۲ استاد دانشکده مهندسی نقشه برداری و اطلاعات مکانی - پردیس دانشکده های فنی - دانشگاه تهران
samadz@ut.ac.ir

(تاریخ دریافت خرداد ۱۳۹۳، تاریخ تصویب آذر ۱۳۹۴)

چکیده

آلودگی هوا در شهر یکی از مهمترین مسائلی است که روی محیط زیست، سلامت جامعه، اقتصاد، مدیریت مناطق شهری و ... تاثیر می گذارد. شناسایی پارامترهای موثر بر آلودگی هوا از طریق کاوش الگوی هم-مکانی می تواند به حل معضل آلودگی هوا کمک نماید. الگوی هم-مکانی بیانگر زیر مجموعه ای از اشیای مکانی است که نمونه های آنها غالباً در یک همسایگی مکانی قرار می گیرند. مدل های موجود کاوش هم-مکانی بدلیل مشکلاتی همچون در نظر گرفتن تنها نوع داده نقطه‌ای، لحاظ کردن ارتباطات همسایگی داده ها بصورت ورودی مساله و توجه نکردن به الگوهای هم-مکانی با محوریت یک داده خاص، توانایی استخراج الگوهای مفید برای کاربردهایی مثل آلودگی هوا را ندارند. لذا در این تحقیق مدلی جدید در حوزه کاوش الگوی هم-مکانی توسعه یافت که: اول، همه نوع داده های مکانی (نقطه‌ای، خطی و سطحی) را در یک ترکیب مشترک در نظر می گیرد. دوم، ارتباطات همسایگی بین اشیاء بصورت غیر صریح در نظر گرفته شده و در فرآیند کاوش الگوها استخراج می گردد. سوم، شناسایی الگوها با محوریت یک عنصر مرکزی انجام می گیرد که در بسیاری از کاربردها همچون کاوش الگوهای آلودگی هوا مناسب می باشد و چهارم، با استفاده از تقسیم بندی فضا (محل سازی) با روش دیاگرام ورونوی ضمن سرعت بخشیدن به فرآیند کاوش الگو، باعث شناسایی الگوهای معنادار بر مبنای مفاهیم مکانی می گردد. مدل توسعه یافته روی داده های بخشی از شهر تهران پیاده سازی و ارزیابی گردید. بررسی الگوهای بدست آمده اولاً نشان دهنده صحت الگوهای استخراج شده و عملکرد مدل بود و ثانیاً بیانگر این موضوع بود که الگوهای استخراجی حول آلودگی های متوسط به بالا رخ داد و این آلودگی ها با ترافیک کم، سرعت باد کم و توپوگرافی کم الگوهای معناداری ایجاد کردند. همچنین گرایش الگوهای حاصل به سمت مناطق مرکزی مورد مطالعه یعنی منطقه ۶ تهران بود.

واژگان کلیدی: داده کاوی مکانی، الگوی هم-مکانی، آلودگی هوا، تهران

۱- مقدمه

مکانی - زمانی آلودگی هوا در مناطق شهری یک موضوع مهم و چالشی است که محورهای تحقیقاتی از مدلسازی های علمی تا زمین آمار و داده کاوی را شامل می شود [۴]. این تحقیق در نظر دارد تا از طریق روش داده کاوی هم-مکانی به شناسایی الگوهای هم-مکانی مهم آلودگی هوای مناطق شهری و پارامترهای موثر بر آن بپردازد.

۲- مروری بر تحقیقات مرتبط

از نقطه نظر مکانی وضعیت آلودگی هوا در یک منطقه می تواند حاصل برهم کنش مجموعه ای از پارامترهای مکانی باشد که این ارتباطات مکانی بین پارامترها می تواند نشان دهنده یکسری الگوی مکانی باشد. شناسایی این ارتباطات و الگوها می تواند در تصمیم گیری هرچه بهتر درخصوص مدیریت آلودگی هوا موثر باشد. تحقیقات زیادی به بررسی رابطه بین آلودگی هوا و پارامترهای مختلف همچون پارامترهای هواشناسی، ترافیکی، توپوگرافی و ... پرداخته اند [۱۶، ۱۷، ۱۸، ۱۹، ۲۰]. براساس بررسی تحقیقات انجام گرفته می توان بطور کلی پارامترهای موثر بر آلودگی هوا را به دو بخش پارامترهای ثابت و پویا مطابق شکل ۱ طبقه بندی نمود. پارامترهای ثابت شامل پوشش گیاهی، شبکه راه ها، توپوگرافی زمین و ساختمان ها می گردند. وجود پوشش گیاهی در یک منطقه می تواند باعث کاهش آلودگی هوا گردیده، نوع و تراکم شبکه راه ها، پستی و بلندی سطح زمین در مناطق مختلف به همراه چگونگی ساختمان های موجود از لحاظ ارتفاع و چینش در منطقه می توانند هریک تاثیراتی در وضعیت آلودگی هوای یک منطقه به عنوان پارامترهای ثابت داشته باشند. پارامترهای پویا شامل ترافیک خیابان ها و پارامترهای هواشناسی می شود. حجم و سرعت وسایل نقلیه به همراه ترکیب نوع وسایل نقلیه که از شاخص های ترافیک می باشند تاثیر مستقیمی بر آلودگی هوا به عنوان یکی از اصلی ترین منابع آلودگی هوا دارند. بعلاوه پارامترهای هواشناسی همچون جهت و سرعت باد، درجه حرارت و میزان نورخورشید به عنوان یکی دیگر از پارامترهای پویا هستند. این پارامترها از طریق تاثیر بر جابجایی آلاینده ها و همچنین عاملی بر واکنش های ثانویه آلاینده ها بر همدیگر می توانند در میزان آلودگی هوای یک منطقه موثر باشند.

آلودگی هوا در محیط های شهری اثرات جدی ای بر روی سلامت و کیفیت زندگی ساکنین آن دارد. طیف گسترده ای از منابع انسانی و طبیعی، سطوح غلظت آلودگی هوای محیط را افزایش می دهند که منجر به بدتر شدن کیفیت هوای پیرامون می شود [۱، ۲، ۳]. تحقیقات اخیر نشان داده است که مساله آلودگی هوا به علت توزیع منابع آلودگی و ریخت شناسی شهر، دارای یک طبیعت چند متغیره و تغییر پذیری مکانی - زمانی محلی بالا می باشد [۴]. در این راستا سازمان محیط زیست لندن در تحقیقی تخمین زده است که ۹۷٪ از CO و ۷۵٪ از NOx بخاطر ترافیک جاده ای است [۵]. دیگر بخش های تاثیر گذار در این آلودگی، مناطق صنعتی، مصرف انرژی خانگی و فعالیت های ساختمانی می باشند.

در شهری مثل تهران، منبع اصلی آلودگی هوا، آلودگی ناشی از ترافیک خیابان ها می باشد [۶، ۷]. از آنجایی که آلاینده CO یکی از اصلی ترین آلاینده های شهر تهران می باشد [۸] و در عین حال مضرترین و خطرناک ترین آلاینده هوای شهری است [۹]؛ لذا آلاینده CO جهت بررسی آلودگی هوای شهر تهران در تحقیق حاضر انتخاب گردید.

مونوکسید کربن یک گاز بی رنگ و بی بو است که در واقع محصول جانبی سوختن و احتراق می باشد. بیشترین تاثیر ناشی از مونوکسید کربن کاهش انتقال اکسیژن به دستگاه های حیاتی مثل قلب و مغز می باشد. تماس با مونوکسید کربن در سطوح بالا و کاهش انتقال اکسیژن به سیستم اعصاب مرکزی آسیب رساننده و توانایی پاسخگویی به محرک های خارجی را کاهش می دهد [۱۰].

داده کاوی مکانی با عنوان روشی جهت کشف الگوهای ارزشمند و مهم اما از قبل نامعلوم از مجموعه ای از داده ها معرفی گردیده است [۱۱، ۱۲]. کاوش الگوی هم-مکانی یکی از مهمترین روش های کاوش داده های مکانی است که امروزه برای یافتن وابستگی و ارتباط بین مجموعه ای از اشیای مختلف استفاده گردیده است [۱۳، ۱۴]. الگوهای هم-مکانی به عنوان خروجی روش های داده کاوی مکانی، ارائه دهنده زیر مجموعه هایی از اشیای مکانی است که نمونه هایشان غالباً در همسایگی مکانی یکدیگر قرار می گیرند [۱۵]. شناخت، مدلسازی و پیش بینی الگوهای



شکل ۱- طبقه بندی پارامترهای موثر بر آلودگی هوا

ثانیاً آنها از یک روش تکراری جستجوی ترکیب و اصلاح الگوها برای شناسایی الگوهای مهم درجات بالا استفاده می کنند که هزینه محاسبات بالایی دارد.

در نهایت بطور خلاصه می توان اشاره نمود که با بررسی تحقیقات موجود در این زمینه کمبودهای زیر در مدل های موجود مشاهده می گردند:

- روش های موجود فقط با یک نوع داده مکانی که اکثراً از نوع نقطه ای است کار می کنند.
- روابط همسایگی بین اشیا به عنوان ورودی در الگوریتم های موجود کاوش هم مکانی در نظر گرفته می شوند.
- روش های موجود در کاوش الگوهای هم مکانی توجهی به یک پارامتر خاص ندارند و عموماً الگوهای کلی از بین مجموعه داده های ورودی را استخراج می نمایند.
- روش های محلی سازی استفاده شده در تحقیقات موجود می توانند به طریق بهینه تری انتخاب گردند که کارایی بیشتری ارائه نمایند.

بر این اساس تحقیق حاضر از طریق ارائه یک مدل جدید داده کاوی هم مکانی؛ اول، همه نوع داده های مکانی (نقطه ای، خطی و سطحی) را در یک ترکیب مشترک در نظر می گیرد. دوم، ارتباطات همسایگی بین اشیا بصورت غیر صریح (یعنی اینکه ارتباطات همسایگی به عنوان اطلاعات اضافی به همراه داده ها ارائه نمی گردد) در نظر گرفته شده و در فرآیند کاوش الگوها استخراج می گردد. سوم، شناسایی الگوها با محوریت یک عنصر مرکزی انجام می گیرد که در بسیاری از کاربردها همچون کاوش الگوهای آلودگی هوا مناسب می باشد و چهارم، با استفاده از تقسیم بندی فضا (محلی سازی) با روش نمودار ورونوی ضمن سرعت بخشیدن به فرآیند کاوش الگو باعث شناسایی الگوهای معنادار بر مبنای مفاهیم مکانی می گردد.

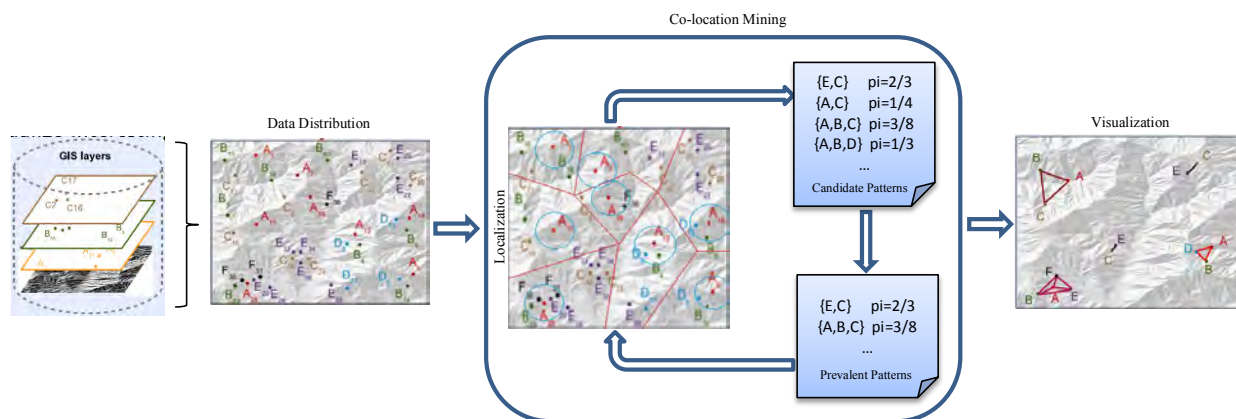
پس از شناسایی پارامترهای موثر بر وضعیت آلودگی هوای مناطق شهری بایستی از طریق روش های داده کاوی مناسب به شناسایی نحوه ارتباطات این پارامترها پرداخت. روش های داده کاوی هم مکانی به منظور استخراج زیر مجموعه هایی از نوع داده هایی که نمونه های آنها غالباً در یک مکان کنار هم دیگر هستند استفاده می شوند. محققان مختلفی روی توسعه و بکارگیری روش های مختلف کاوش الگوهای هم مکانی برای حوزه های کاربردی متفاوت فعالیت کرده اند. یکی از موضوعات مورد بررسی شناسایی شاخص هایی جهت کشف الگوها بوده است. نسبت مشارکت^۱ و ضریب مشارکت^۲ شاخص هایی هستند که توسط محققان مختلف [۲۱،۲۲،۲۳،۲۴،۲۵] به منظور شناسایی الگوهای هم مکانی استفاده شده اند. تحقیقات اشاره شده تنها بر اساس نوع داده نقطه ای به شناسایی الگوهای هم مکانی پرداخته اند. همچنین در کلیه این تحقیقات رابطه همسایگی بین اشیا به صورت ورودی مساله در فرآیند کاوش الگوها در نظر گرفته شده است. Celik و همکاران [۲۲] روش محلی سازی quadtree را در کاوش هم مکانی بکار گرفته اند. Manikandan و همکاران [۲۶] به منظور کاهش زمان اسکن کردن پایگاه داده، از ساختار داده مکانی R-tree به منظور کاوش الگوی هم مکانی استفاده کردند. Venkatesan و همکاران [۲۷] از آمار مکانی و روش های داده کاوی برای شناسایی الگوهای هم مکانی از مجموعه داده های مکانی استفاده کردند. در [۲۸] محققین مدلی بر مبنای بافر برای کاوش الگوهای هم مکانی اشیا مکانی توسعه یافته ارائه دادند اما اولاً این روش تنها یک نوع داده مثلاً خطی را در فرآیند کاوش در نظر می گیرد،

۱ Participation Ratio
۲ Participation Index

۳- روش پیشنهادی

تحقیق حاضر بدنبال استخراج الگوهای هم-مکانی آلودگی هوا و سایر پارامترهای مکانی پویا و ثابت موثر بر آن می باشد. بر اساس آنچه در بالا اشاره گردید، اول اینکه بایستی روابط همسایگی بین اشیای مورد بررسی به صورت غیر صریح در نظر گرفته شود و این ارتباط همسایگی در فرآیند کاوش الگوها شناسایی و استخراج گردد؛ در حالیکه در تحقیقات پیشین روابط همسایگی اشیا به عنوان ورودی مساله در نظر گرفته شده است. دوم، مدلی باید ارائه گردد که الگوها بر مبنای یک عنصر مرکزی کاوش گردند زیرا در خیلی از کاربردها همچون کاوش الگوی آلودگی هوا، الگوی

جرم و الگوی تصادف اتومبیل، یافتن الگوها بدون توجه به یک عنصر مرکزی مفید نخواهد بود. سوم، پارامترهای مورد بررسی دارای نوع عارضه های متفاوتی هستند بعضی دارای نوع نقطه ای، بعضی حالت خطی و بعضی دیگر به صورت سطحی می باشند. لذا استفاده از شاخص های مرسوم مورد استفاده برای داده های نقطه ای در این تحقیق کافی نخواهد بود و ما بایستی از طریق تعریف شاخص های جدید بر این مشکل فائق آییم. بدین منظور در ادامه ابتدا مفاهیمی که بر اساس آن تحقیق شکل گرفته است تشریح و سپس بر مبنای آنها مدل پیشنهادی توسعه می یابد. شکل ۲ فرآیند عملکردی تحقیق حاضر را بطور شماتیک به تصویر می کشد.



شکل ۲- فرآیند عملکردی تحقیق حاضر

۳-۱- مفاهیم تحقیق

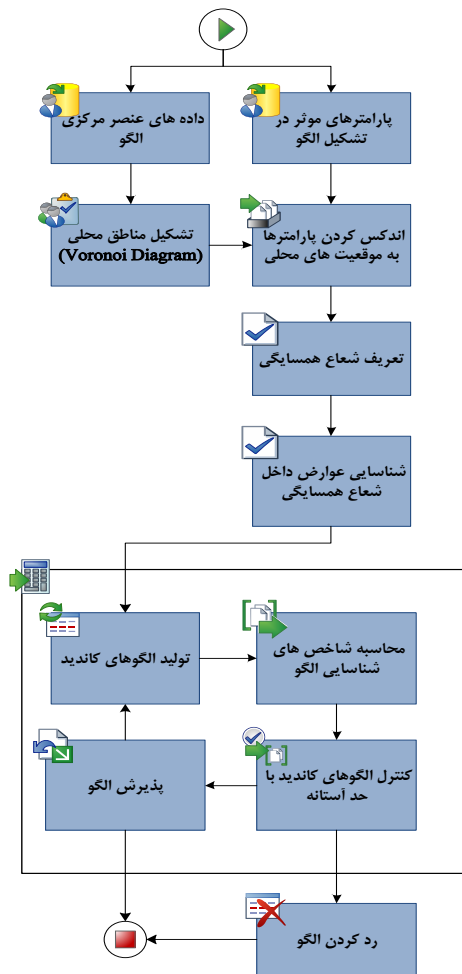
به منظور توسعه الگوریتم های موجود نیاز به ارائه تعاریفی وجود دارد تا بتوان بر مبنای آنها روش پیشنهادی این تحقیق را تشریح کرد.

تعریف ۱: اولین قانون جغرافیای Tobler بیان می کند که "اشیای مکانی با همدیگر در ارتباط هستند و هرچه اشیا به هم نزدیک ترند ارتباط بیشتری دارند و برعکس" [۲۹]. بنابراین در کاوش الگوهای هم-مکانی بهتر است رابطه اشیای مکانی بر اساس روابط محلی کاوش گردند. با توجه به تعریف ۱، فضای مکانی مورد نیاز برای کاوش الگوهای هم-مکانی مشخص گردیده است. فرض کنید چارچوب مکانی D باشد آنگاه SD_i به نحوی که $SD_i \subset D, (1 \leq i \leq n)$ ، یک منطقه مکانی خواهد بود که روابط مکانی داخل آن همگون می باشد و $D = \{SD_1, SD_2, \dots, SD_n\}$ به گونه ای است که این مناطق مکانی همپوشانی ندارند.

تعریف ۲: در تحقیق حاضر با توجه به اینکه هدف یافتن الگوهای هم-مکانی پارامترهای مختلف با آلودگی هوا می باشد، لذا با توجه به نوع مساله تحقیق، هم-مکانی سایر پارامترها بایستی با یک پارامتر مرکزی (در این مورد آلودگی هوا) سنجیده شود. از اینرو این عنصر به عنوان عنصر مرکزی الگو (PCE)^۱ تعریف گردید.

تعریف ۳: علاوه بر آنچه قبلا به عنوان اهداف تحقیق حاضر ذکر شد، این تحقیق به موضوع دیگری نیز توجه داشته است. اول اینکه در اکثر تحقیقات بدنبال یافتن الگوهای هم-مکانی بین همه نوع داده های موجود در چارچوب مکانی مساله می باشند [۲۲، ۲۳، ۳۰، ۳۱] و در این بین هیچ تأکیدی بر یافتن هم-مکانی بر اساس یک رابطه همسایگی بین اشیا در کاربرد مورد نظر به صورت ورودی مساله داده می شود [۲۲، ۲۳، ۳۰، ۳۱] در حالیکه

^۱ Pattern Core Element



شکل ۳- فلوچارت کاوش هم-مکانی

در مدل پیشنهادی RST-CoM، ورود داده ها با دو عنوان انجام می گیرد، داده های موثر در تشکیل الگو و داده های عنصر مرکزی الگو. بر مبنای عنصر مرکزی الگو محلی سازی با استفاده از نمودار ورونوی انجام می پذیرد و سایر داده ها به آن اندکس می گردند تا اولاً سرعت کاوش الگو افزایش یافته و ثانیاً الگوها بصورت محلی کاوش گردند. سپس شعاع همسایگی عنصر مرکزی الگو با توجه به کاربرد مورد بررسی تعیین و بر آن اساس عناصر داخل شعاع همسایگی شناسایی و الگوریتم وارد فرآیند کاوش الگوهای هم-مکانی می گردد. این کار از طریق تشکیل کاندیدهای الگوی هم-مکانی و ارزیابی آنها بر اساس حدآستانه تعریف شده که تابعی از کاربرد مورد بررسی می باشد انجام می گیرد. شناسایی الگوهای هم-مکانی در مساله این تحقیق نیاز به گسترش شاخص های موجود همچون نسبت مشارکت و ضریب مشارکت استفاده شده در تحقیقات پیشین [۲۱، ۲۲، ۲۳]، ۲۴، ۲۵] دارد. بر این اساس در این تحقیق با در نظر گرفتن انواع داده ها، شاخص های پیشین به شکل زیر توسعه یافته اند:

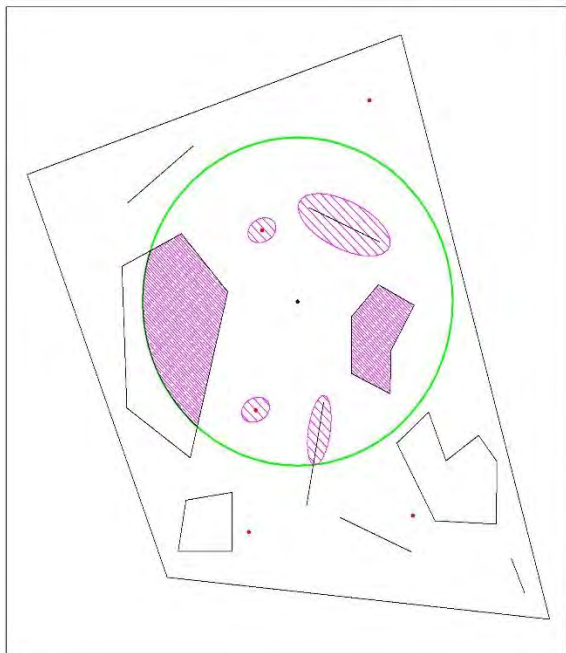
در حوزه علوم مکانی، روابط همسایگی یکسری روابط غیر صریح می باشند و نایستی به عنوان ورودی مساله داده شوند بلکه این رابطه همسایگی در حین حل مساله بایستی شناسایی و استخراج گردد. از اینرو، در این تحقیق الگوی هم-مکانی مطابق تعریف ۳ کاوش می گردد.

تعریف ۴: بر مبنای تعاریف ۱، ۲ و ۳ این تحقیق الگوهای هم-مکانی را در زیرمجموعه های چارچوب مکانی با عنوان مناطق کاوش خواهد نمود. در [۲۲] نیز کاوش هم-مکانی به صورت منطقه ای انجام گرفته است که آنها از روش quad-tree برای تقسیم بندی فضا استفاده نمودند. اما بدلیل مشکلاتی که این روش داشته است همچون پیچیدگی محاسبات و همپوشانی مناطق تعریف شده، اینجا از روشی دیگر برای تقسیم بندی فضا استفاده شده که مشکلات فوق را ندارد. روش مورد استفاده این تحقیق نمودار ورونوی می باشد. برای یک مجموعه سایت و یک اندازه فاصله داده شده در صفحه، نمودار ورونوی صفحه را به مناطقی برای هر سایت تقسیم بندی می کند که همه نقاط داخل یک منطقه به مرکز منطقه (سایت) از هر یک از مراکز دیگر نزدیکترند.

۱- ۲-۳- توسعه الگوریتم

با توجه به آنچه در تعریف مساله آمده است، این تحقیق بدنبال یافتن الگوهای هم-مکانی است. این کار از طریق مدلی جدید انجام می گیرد که با لحاظ کردن شرایطی توسعه یافته است. مدل توسعه یافته نوع داده های مختلف را که بیان کننده یک مساله واقعی هستند در نظر می گیرد. کاوش را در مناطق محلی انجام می دهد چون اشیای مکانی رفتار محلی دارند؛ و بر اساس یک عنصر مرکزی این کاوش انجام می پذیرد چون کاربردهایی همچون الگوی آلودگی هوا و صدا، الگوی جرم و الگوی تصادفات اتومبیل وجود دارند که الگوهای بر مبنای یک عارضه خاص اهمیت دارند. دستیابی به این هدف مستلزم توسعه مفاهیم و معیارهای سنجش الگوها می باشد که مفاهیم مورد نیاز در قسمت قبل تعریف گردید. بر مبنای تعاریف ارائه شده مدل پیشنهادی این تحقیق با نام RST-CoM (Regional Spatio-Temporal Co-location Pattern Mining) توسعه یافت که شکل ۳ فلوچارت آن را نمایش می دهد.

• برای داده های نقطه ای :
 شده در شناسایی الگوهای هم-مکانی با هاشور صورتی رنگ مشخص گردیده است.



شکل ۴- نمایش شماتیک همسایگی عنصر مرکزی الگو

• داده های خطی : برای این نوع داده دو حالت را می توان متصور شد. اول اینکه عارضه خطی بطور کامل در شعاع همسایگی عنصر مرکزی الگوی هم-مکانی قرار بگیرد. در این حالت می توان این عارضه خطی را جزو داده های همسایگی در نظر گرفت. در حالت دوم اینکه عارضه خطی به صورت بخشی در شعاع همسایگی عنصر مرکزی الگوی هم-مکانی قرار بگیرد. در این حالت همانطور که در شکل ۴ نمایش داده شده است بخشی از عارضه خطی که داخل شعاع همسایگی عنصر مرکزی الگو قرار می گیرد شناسایی گردیده و در محاسبات معادله (۲) استفاده می گردد.

• داده های سطحی : برای این نوع داده نیز دو حالت در نظر گرفته شده و تحلیل می گردد. اول اینکه عارضه سطحی بطور کامل در شعاع همسایگی عنصر مرکزی الگوی هم-مکانی قرار بگیرد. در این حالت می توان این عارضه سطحی را جزو داده های همسایگی در نظر گرفت. در حالت دوم اینکه عارضه سطحی به صورت بخشی در شعاع همسایگی عنصر مرکزی الگوی هم-مکانی قرار بگیرد. در این حالت همانطور که در شکل ۴ نمایش داده شده است بخشی از عارضه سطحی که

$$\Pr(C, f_i) = \frac{NI(f_i)}{N(f_i)} \quad (1)$$

جاییکه $NI(f_i)$ تعداد اشیای مشخص f_i در همسایگی نمونه های C و $N(f_i)$ کل تعداد اشیای f_i است.

• برای داده های خطی :

$$\Pr(C, f_i) = \frac{NI(f_i)}{N(f_i)} + \frac{(l_1 w_1 + l_2 w_2 + \dots)}{N(f_i) \times (w_1 + w_2 + \dots)} \quad (2)$$

جاییکه $NI(f_i)$ تعداد اشیای مشخص f_i که کاملاً در همسایگی نمونه های C قرار دارند و $N(f_i)$ کل تعداد اشیای f_i است. همچنین l_i تعداد عارضه های خطی است که w_i درصد آنها داخل همسایگی نمونه های C قرار دارند.

• برای داده های سطحی :

$$\Pr(C, f_i) = \frac{NI(f_i)}{N(f_i)} + \frac{(l_1 s_1 + l_2 s_2 + \dots)}{N(f_i) \times (s_1 + s_2 + \dots)} \quad (3)$$

جاییکه $NI(f_i)$ تعداد اشیای مشخص f_i که کاملاً در همسایگی نمونه های C قرار دارند و $N(f_i)$ کل تعداد اشیای f_i است. همچنین l_i تعداد عارضه های سطحی است که مقدار s_i درصد از مساحت آنها داخل همسایگی نمونه های C قرار دارند.

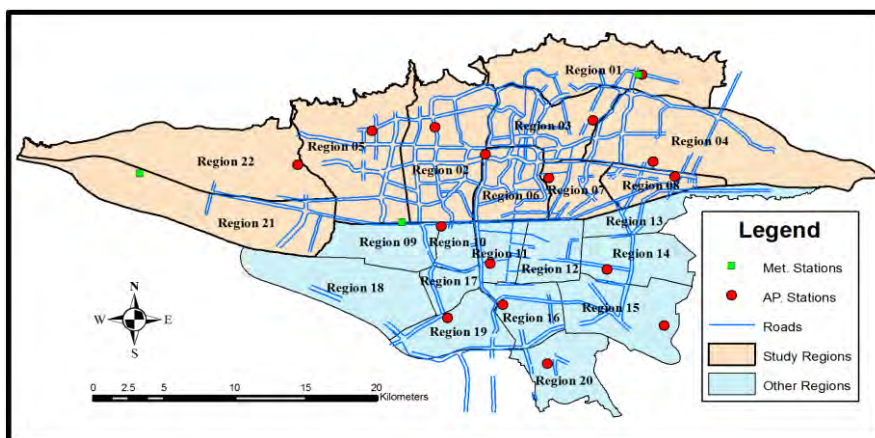
اجرای الگوریتم شناسایی الگوهای هم-مکانی در حالتی که تنها نوع داده مساله نقطه ای باشد یک حالت عادی است زیرا در این وضعیت رابطه همسایگی تنها از طریق کنترل قرارگیری در شعاع همسایگی می تواند تعیین گردد؛ در صورتی که اگر مساله حالت کلی تری را در نظر بگیرد که داده های خطی و سطحی نیز جزو داده های مساله باشند، در این حالت رابطه همسایگی تنها از طریق کنترل شعاع همسایگی تعیین نمی شود. از اینرو با توجه به فرض این تحقیق که بایستی انواع داده ها در فرآیند داده کاوی مکانی منظور گردند، لذا جهت فائق آمدن بر این موضوع تعاریف زیر برای مدیریت نوع داده های خطی و سطحی در فرآیند کاوش هم-مکانی ارائه گردیده است. شکل ۴ نمایش شماتیکی از یک منطقه همسایگی (دایره سبز رنگ) و عنصر مرکزی الگو (نقطه مشکی رنگ) را نشان می دهد که اشیای مختلف در همسایگی آن قرار گرفته اند و قسمت های در نظر گرفته

الگوها با دید محلی، این تحقیق با توجه به اهمیت بحث آلودگی هوا خصوصا در کلان شهری مثل تهران، الگوهای هم-مکانی آلودگی هوا با سایر پارامترهای موثر کاوش گردید. شکل ۵ منطقه مورد مطالعه را نمایش می دهد. محل مورد بررسی مناطق ۱-۸، ۲۱ و ۲۲ از شهر تهران می باشد که اولاً داده‌های آن در دسترس بود و ثانياً داده‌ها بروز و دارای همپوشانی زمانی بود. با توجه شرایط ذکر شده از مجموعه داده های طبقه بندی شده موثر بر بحث آلودگی هوا، مطابق شکل ۶ مجموعه ای از داده ها انتخاب گردید.

داخل شعاع همسایگی عنصر مرکزی الگو قرار می گیرد شناسایی گردیده و در محاسبات معادله (۳) استفاده می گردد.

۴- نتیجه عملی و ارزیابی

به منظور ارزیابی کارایی روش پیشنهادی، مورد مطالعاتی بر روی کاوش الگوی هم-مکانی آلودگی هوای بخشی از شهر تهران انجام گرفت. همانطور که قبلا اشاره گردید با در نظر گرفتن نوع داده های مختلف و کاوش



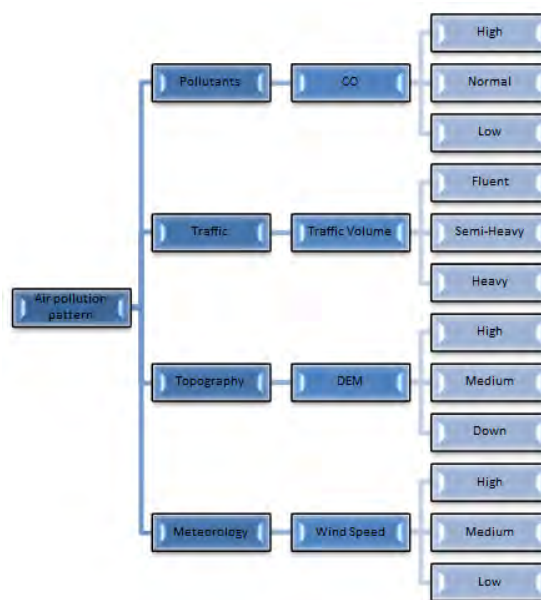
شکل ۵- نمایش منطقه مورد مطالعه

جدول ۱ نحوه انجام این طبقه بندی برای ورود به فرآیند داده کاوی را نشان داده است. این جدول طبقه بندی چهار پارامتر مورد استفاده در این تحقیق را نشان می دهد و در ستون برچسب از نمادهای اختصاری Ap (Air pollutant)، Tp (Topography)، Tr (Traffic) و Wn (Wind) جهت سهولت استفاده در مدل استفاده شده است.

در اجرای مدل پیشنهادی از داده های ۴ روز در ماه های فروردین، تیر، مهر و دی ۱۳۹۰ استفاده گردیده است. ۴ زمان برای اجرای مدل توسعه یافته انتخاب گردید تا اولاً تکرار پذیری اجرای برنامه تست گردد و ثانياً با توجه به تغییرات فصلی اعتبار سنجی نتایج و بالطبع اعتبارسنجی مدل پیشنهادی کنترل گردد. مقادیر ورودی این تحقیق عبارتند از:

- شعاع همسایگی: ۱۵۰۰ متر
- حدآستانه اهمیت: ۰.۵

در انتخاب مقدار شعاع همسایگی با توجه به فاصله بین ایستگاه های سنجش آلودگی هوا و نواحی ورونوی ایجاد شده از



شکل ۶- نمایش داده های مورد استفاده در تحقیق حاضر

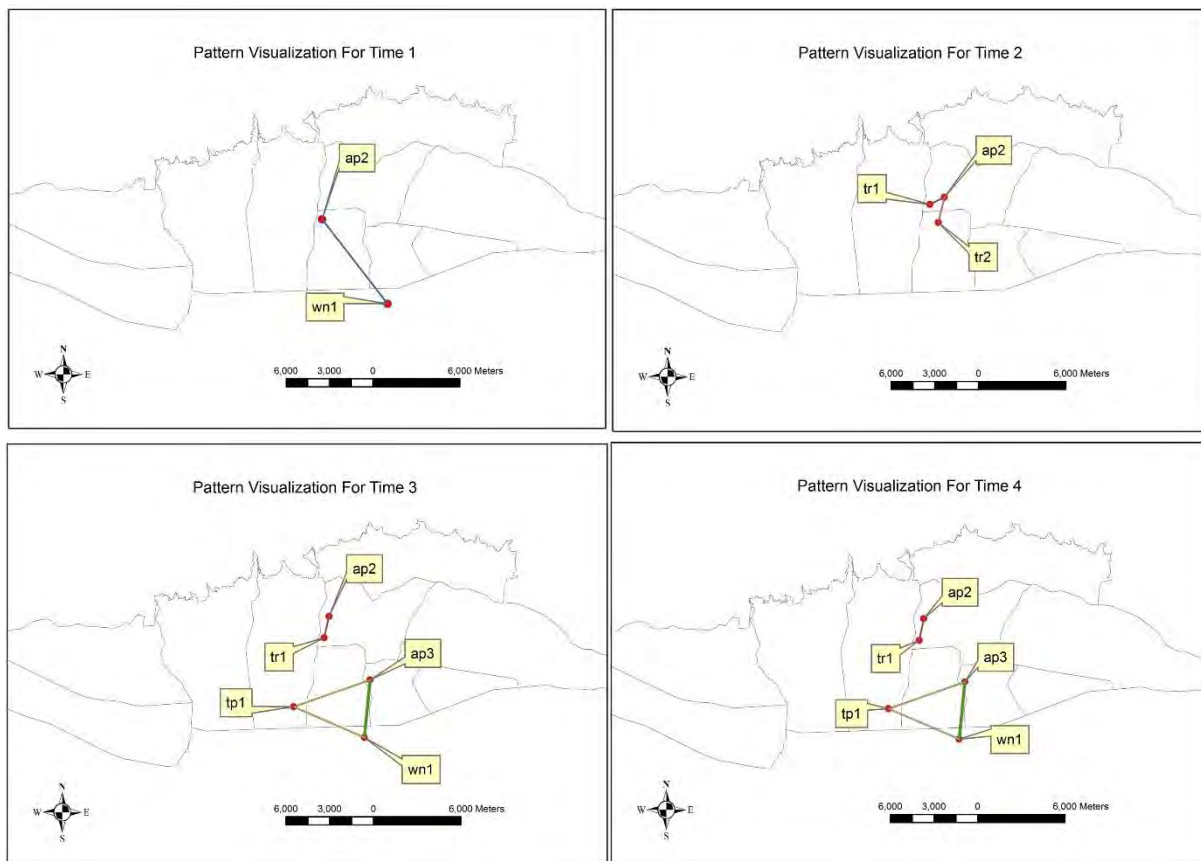
همانطور که در شکل ۶ نشان داده شده است به منظور بکارگیری داده های مدنظر در فرآیند داده کاوی هم-مکانی، هر یک از داده ها در سه دسته طبقه بندی گردیدند.

یک طرف و معنی دار بودن شعاع همسایگی جهت تغییرات پارامترها، مقدار فوق انتخاب گردید. در انتخاب حدآستانه اهمیت مقدار ۰.۵ از این جهت انتخاب گردید که الگویی استخراج گردد که حداقل بیش از نیمی از عوارض آن در الگوی مد نظر مشارکت داشته باشند. لازم به ذکر است که این تحقیق بدنبال یافتن مقادیر بهینه برای شعاع همسایگی و حد آستانه شناسایی الگوی غالب نمی‌باشد و این موارد خارج از بحث این تحقیق است. شکل ۷ نمایش مکانی الگوهای هم-مکانی استخراج شده می‌باشد. به منظور نمایش الگوهای هم-مکانی بدین ترتیب عمل گردید که برای هر الگوی شناسایی شده میانگین هندسی

عوارض آن الگو از بین تمام عوارض مشارکت کننده در تشکیل آن الگو استخراج گردیده است و سپس این عوارض با خطوطی به هم متصل گردیدند که ضخامت آن متناسب با اهمیت الگو و رنگ های متفاوت جهت تشخیص الگوها اختصاص یافته است. الگوهای استخراج شده در این فرایند شامل یک الگوی $\{(ap2;wn1)\}$ در زمان اول (فروردین)، دو الگوی $\{(ap2;tr1), (ap2;tr2)\}$ در زمان دوم (تیر)، ۴ الگوی $\{(ap2;tr1), (ap3;tp1), (ap3;wn1), (ap3;tp1;wn1)\}$ در زمان سوم (مهر) و ۴ الگوی $\{(ap2;tr1), (ap3;tp1), (ap3;wn1), (ap3;tp1;wn1)\}$ در زمان چهارم (دی) می‌باشند.

جدول ۱- طبقه بندی داده های مورد استفاده

ابرکلاس	زیرکلاس	دامنه کمی	برجسب	ابرکلاس	زیرکلاس	دامنه کمی	برجسب
ترافیک	سنگین	ترافیک سنگین و بسیار سنگین	Tr3	ابرکلاس	مرتفع	>1500	Tp3
	نیمه روان	اختلال در حرکت و سنگین در حرکت	Tr2	زیرکلاس	متوسط	1300-1500	Tp2
	روان	ترافیک روان	Tr1	زیرکلاس	پست	<1300	Tp1
سرعت باد	شدید	Wn3	Wn3	آلاینده (CO)	زیاد	>5	Ap3
	متوسط	Wn2	Wn2	آلاینده (CO)	متوسط	1-5	Ap2
	ضعیف	Wn1	Wn1	آلاینده (CO)	کم	<1	Ap1



شکل ۷- نمایش توزیع مکانی الگوهای هم - مکانی استخراج شده در چهار زمان متفاوت

یک همبستگی نزدیک با آلودگی هوا دارند، بخاطر این حقیقت که آن پارامترها منبع آلودگی هوا نیستند و در عوض تنها پارامترهای مهم تاثیرگذار بر رخداد آلودگی هوا هستند، این انتظار وجود ندارد که مراکز هندسی سرعت باد و توپوگرافی مشابه ترافیک نزدیک آلودگی هوا واقع شوند. از دیگر نقطه نظر می توان به این موضوع اشاره کرد که محوریت الگوهای استخراج شده در مناطق مرکزی شهر بویژه منطقه ۶ تهران می باشد و این با وجود طرح ترافیک در این محدوده ها قابل درک می باشد.

۵- نتیجه گیری

بحث آلودگی هوا در مناطق شهری یکی از مسائل بگرنج مدیریت شهری است. مدیران شهری جهت فائق آمدن بر این مشکل بایستی پارامترهای موثر و همچنین نوع عملکرد آنها بر آلودگی هوا را شناخته تا بتوانند برنامه ریزی لازم در خصوص آن را انجام دهند. تحقیقات متعددی در خصوص بررسی اثر پارامترهای مختلف همچون هواشناسی، ترافیک و توپوگرافی بر روی آلودگی هوا انجام گرفته است. در این تحقیق سعی گردیده است تا بطور همزمان اثر هر سه مولفه فوق الذکر بر روی آلودگی هوا در مناطق شهری بررسی گردد و اثرات آنها بر آلودگی هوا از طریق شناسایی الگوهای غالب تعیین گردد. این کار از طریق استخراج الگوهای هم-مکانی آلودگی هوا و سایر پارامترها برای شهر تهران انجام گرفت. بکارگیری این روش مستلزم توسعه الگوریتم های موجود کاوش الگوی هم-مکانی بود که این کار در تحقیق حاضر انجام گرفت.

نتایج حاصل از پیاده سازی این مدل کارایی مدل توسعه یافته را مشخص نمود و همچنین الگوهای بدست آمده نشان داد که آلودگی هوا با ترافیک نوع ۱، سرعت باد نوع ۱ و توپوگرافی نوع ۱ بیشترین ارتباط را دارد و الگوهای غالب عمدتاً گرایش به سمت منطقه مرکزی مورد مطالعه یعنی منطقه ۶ شهر تهران دارند. نکته ای که بایستی به آن اشاره کرد اینکه بحث آلودگی هوا یک موضوع زمانی - مکانی است و علاوه بر در نظر گرفتن بعد مکانی بایستی به بعد زمانی مساله نیز توجه گردد. اما در این تحقیق تنها تمرکز بر توسعه مکانی مدل جهت پاسخگویی به انواع داده بوده است. توسعه مدل موجود به نحوی که کاوش الگوهای مکانی و زمانی به صورت توأمان انجام گیرد جزو اهداف آینده محققین می باشد.

با مشاهده شکل ۷ مشخص می گردد که در همه الگوهای بدست آمده آلودگی هوا از نوع ۲ و ۳ می باشند که با توجه به شکل ۶ نشان دهنده آلودگی متوسط به بالا هستند و همانگونه که [۶،۸،۳۲] به آلودگی هوای زیاد به عنوان مشکل اصلی شهر تهران اشاره کرده اند، رخداد الگوهای با سطوح بالای آلودگی هوا قابل انتظار بود. از دیگر موارد مشهود اینکه در ۳ زمان از ۴ مورد بررسی شده الگویی بین آلودگی هوا و ترافیک وجود دارد. از آنجایی که در این تحقیق از CO به عنوان آلاینده هوا استفاده شده است و همچنین با توجه به این حقیقت که منابع اصلی این آلاینده وسایل نقلیه موتوری هستند، CO به عنوان آلاینده مرتبط با ترافیک شناخته می شود [۶،۷] و در نتیجه وجود الگوهای غالب بین آلودگی هوا و ترافیک مورد انتظار می باشد. از طرفی دیگر پارامترهای مشارکت کننده در الگوهای استخراج شده سرعت باد نوع ۱ می باشد که بیانگر سرعت پایین باد است و این خود با توجه به وضعیت توپوگرافی شهر تهران که در میان رشته کوه البرز از شمال و کوه های بی بی شهربانو از شرق محبوس گردیده است یکی از معضلات شهر تهران و باعث باقی ماندن آلودگی ها در سطح شهر می گردد [۷] و مشابه [۳۲] می توان نتیجه گرفت که سرعت پایین باد یکی از مهمترین پارامترهای هواشناسی است که منجر به سطوح بالای غلظت CO می شود. علاوه توپوگرافی نوع ۱ بیشتر در مناطق مرکزی و جنوبی شهر قرار می گیرد که مناطق دارای آلودگی بیشتر می باشند و همانگونه که در [۹،۳۲] اشاره گردیده است الگوی بین توپوگرافی کم و آلودگی هوای زیاد یک الگوی معتبر است. همچنین با توجه به تغییرات فصلی میزان الگوهای استخراج شده تغییر می کند به نحوی که در فصول پاییز و زمستان تعداد الگوهای بدست آمده در خصوص آلودگی هوا بیشتر است که مطابق با نظر محققین در [۷،۹] با توجه به رخداد پدیده هایی همچون وارونگی دمایی در این فصول این اتفاق قابل انتظار و لذا وجود الگوهای بیشتر و همچنین در ارتباط با آلودگی بیشتر هوا معتبر می باشد. همانگونه که در شکل ۷ دیده می شود در بازه های زمانی ای که تنها الگوها بین آلودگی هوا و ترافیک وجود دارند، طول خطوط کوتاهتر از بازه های زمانی با الگوهای دارای پارامترهای بیشتر می باشد. این تایید کننده این حقیقت است که مهمترین منبع آلودگی هوا ترافیک است [۶،۷] جاییکه مراکز هندسی داده های ترافیک و آلودگی هوا نزدیک همدیگر قرار می گیرند. از طرف دیگر اگرچه سرعت باد و توپوگرافی کم

مراجع

- [1] D. Deligiorgi & K. Philippopoulos, (2011), "Spatial Interpolation Methodologies in Urban Air Pollution Modeling: Application for the Greater Area of Metropolitan Athens", Greece, Advanced Air Pollution, Dr. Farhad Nejadkoorki (Ed.), ISBN: 978-953-307-511-2.
- [2] Y. Fang, et al. (2013), "Air pollution and associated human mortality: the role of air pollutant emissions, climate change and methane concentration increases from the preindustrial period to present." *Atmospheric Chemistry and Physics*, 13.3: 1377-1394.
- [3] Gif-sur-Yvette, France. (2013), "Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change", *Environmental Research Letters*, doi:10.1088/1748-9326/8/3/034005.
- [4] A. Champendal, M. Kanevski, P. E. Huguenot, & J. Golay, (2013), "Exploratory Analysis of Spatial-Temporal Patterns of Air Pollution in the City", *Geophysical Research*, Vol. 15, EGU2013-5197.
- [5] Y. Ma, M. Richards, M. Ghanem, Y. Guo & J. Hassard, (2008), "Air Pollution Monitoring and Mining Based on Sensor Grid in London", *Sensors*, 8, 3601-3623; DOI: 10.3390/s8063601.
- [6] M. H. Azizi, (2011), "Impact of traffic-related air pollution on public health: a real challenge". *Archives of Iranian Medicine*, 14(2), 139.
- [7] S. Y. Safavi, & B. Alijani, (2006). "Evaluation of Geographical parameters in Tehran air pollution". *Geographical Researches Journal*, No. 58, pp.: 99-112. (In Persian)
- [8] A. Kavousi, R. Sefidkar, H. Alavimajd, Y. Rashidi, & Z. A. Khonbi, (2013), "Spatial analysis of CO and PM10 pollutants in Tehran city". *Journal of Paramedical Sciences (JPS)*, Summer, 4(3), 2008-4978.
- [9] N. Rahimi Ghoroghi, (2012), "Evaluation of geographical factors on Tehran air pollution and its relation with temperature inversion". First Conference of air and noise pollution management. Tehran, Iran. http://www.civilica.com/Paper-CANPM01-CANPM01_039.html . (In Persian)
- [10] N. Sheng, & U. W. Tang, (2011), "A building-based data capture and data mining technique for air quality assessment", *Front. Environ. Sci. Engin., China*, 5(4): 543-551, DOI 10.1007/s11783-011-0369-4.
- [11] H. J. Miller, & J. Han, (2009), *Geographic Data Mining and Knowledge Discovery*, 2nd edition, London: CRC Press, published.
- [12] S. Shekhar, & S. Chawala, (2003), *Spatial Databases: A Tour*. Prentice Hall, ISBN: 013-017480-7.
- [13] W. Ding, R. Jiamthaphaksin, R. Parmar, D. Jiang, T. F. Stepinski, & C. F. Eick, (2008), "Towards Region Discovery in Spatial Datasets", In Proc. of Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD).
- [14] X. Xiao, X. Xie, Q. Luo, & W. Y. Ma, (2008), "Density based co-location pattern discovery", In Proc. of ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM-GIS).
- [15] G. Manikandan, & S. Srinivasan, (2012), Mining Spatially Co-Located Objects from Vehicle Moving Data, *European Journal of Scientific Research*, ISSN 1450-216X, Vol. 68, No. 3, pp. 352-366.
- [16] G. Wang, F. H. M. van den Bosch, & M. Kuffer, (2008), "Modeling Urban Traffic Air Pollution Dispersion", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVII. Part B8. Beijing.
- [17] U. W. Tang, & Z. S. Wang, (2007), "Influences of urban forms on traffic-induced noise and air pollution: Results from a modeling system", *Environ. Model. Soft.*
- [18] E. Sahafzadeh, & E. Ahmadi, (2009), "Prediction of Air Pollution of Boushehr City Using Data Mining", Second International Conference on Environmental and Computer Science, DOI 10.1109/ICECS.2009.18.

- [19] E. Parvinnia, (2007) "The Application of Association Rule Mining a Case Study: The Effect of Atmospheric Parameters on Air Pollution", IADIS International Conference Applied Computing.
- [20] S. B. Kim, C. Temiyasathit, S. K. Park, & V. C. P. Chen, (2009), "Spatio-Temporal Data Mining for Air Pollution Problems", IGI Global.
- [21] J. S. Yoo, & M. Bow, (2011), "Mining Top-k Closed Co-location Patterns", In Proc. of IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services (ICSDM), pp. 100-105.
- [22] M. Celik, J. M. Kang, & S. Shekhar, (2007), "Zonal Co-location Pattern Discovery with Dynamic Parameters", In Proc. of Seventh IEEE International Conference on Data Mining,.
- [23] Y., Huang, S. Shekhar, & H. Xiong, (2004), "Discovering co-location patterns from spatial datasets: A general approach", TKDE, Vol. 16, No. 12, pp. 1472-1485.
- [24] J. S. Yoo, and M. Bow, (2011), "Mining top-k closed co-location patterns." Spatial Data Mining and Geographical Knowledge Services (ICSDM), 2011 IEEE International Conference on. IEEE.
- [25] F. Qian, et al., (2012), "Spatial co-location pattern discovery without thresholds". Knowledge and Information Systems, 33.2: 419-445.
- [26] G. Manikandan, & S. Srinivasan, (2012), "Mining of Spatial Co-location Pattern Implementation by FP Growth". Indian Journal of Computer Science and Engineering (IJCSE). Vol. 3, No. 2, pp.: 344-348.
- [27] M. S. Venkatesan, Th. Arunkumar, & P. Prabhavathy, (2012), "Discovering Co-location Patterns from Spatial Domain using a Delaunay Approach". Procedia Engineering, INTERNATIONAL CONFERENCE ON MODELLING OPTIMIZATION AND COMPUTING, Vol. 38, pp.: 2832-2845.
- [28] H. Xiong, S. Shekhar, Y. Huang, V. Kumar, X. Ma, & J. S. Yoo, (2004). A framework for discovering colocation patterns in data sets with Extended spatial objects, In Proceeding of the 2004 SIAM international conference on data mining (SDM'04), Lake Buena Vista, FL, pp.: 78-89.
- [29] M. F. Goodchild, (2003), "The fundamental laws of giscience", Keynote address, Annual Assembly, University Consortium for GIScience, Monterey, CA..
- [30] M. S. Venkatesan, T. Arunkumar, & P. Prabhavathy, (2012), "Discovering Co-location Patterns from Spatial Domain using a Delaunay Approach" , Procedia Engineering, INTERNATIONAL CONFERENCE ON MODELLING OPTIMIZATION AND COMPUTING., Vol. 38, pp. 2832-2845.
- [31] G. Priya, N. Jaisankar, & M. Venkatesan, (2011), "Mining Co-location Patterns from Spatial Data using Rulebased Approach", In Journal of Global Research in Computer Science, Vol. 2, No. 7.
- [32] A. R. Saadatabadi, L. Mohammadian, & A. Vazifeh, (2012), "Controls on air pollution over a semi-enclosed basin, Tehran: A synoptic climatological approach", Iranian Journal of Science & Technology (IJST), Vol.4, pp.501-510.