نشریه علمی پژوهشی مهندسی فناوری اطلاعات مکانی

سال سوم، شماره سوم، پاییز ۱۳۹۴

Vol.3, No.3, autumn 2015

۷۵-۹۳

روشی نوین در تخصیص منابع متحرک بر روی شبکه راه با استفاده از خوشه‌بندی مکانی-زمانی رویدادها

محسن گودرزی [۱]، فرشاد حکیم‌پور [۲] *، پرهام پهلوانی [۳]، سید محمود حاجی میررحیمی [۴]

۱- دانش‌آموخته کارشناسی‌ارشد اطلاعات مکانی، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس فنی دانشگاه تهران، ایران

۲- ۲- استادیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس فنی دانشگاه تهران، ایران

۳- ۳- استادیار، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس فنی دانشگاه تهران، ایران

۴- ۴-دانش‌آموخته کارشناسی ارشد GIS&RS، دانشکده علوم زمین، دانشگاه تبریز، ایران

تاریخ دریافت مقاله: ۹۴/۰۷/۰۱ تاریخ پذیرش مقاله: ۹۵/۰۱/۱۵

چکیده

یک مسئله معمول تخصیص و جایابی منابع *(LA)* از دو بخش جایابی منابع و تخصیص منابع به تقاضاها تشکیل می‌گردد. مدیران شهری از روش‌های متنوعی همچون ناحیه‌بندی قسمت‌های مختلف شهر که در نتیجه‌ی حل یک مسئله تخصیص منابع حاصل می‌شود، با هدف برنامه‌ریزی شهری استفاده می‌نمایند. تغییرات در ناحیه‌بندی مناطق شهری به طور موثر از میزان پویایی تقاضاهای مرتبط نشأت می‌گیرد. تلاش‌هایی به منظور ایجاد ارتباط بین مبحث خوشه‌بندی و تخصیص منابع انجام گرفته است. در اکثر این تحقیقات از خوشه‌بندی به عنوان ابزاری جهت جایابی منابع بهره گرفته شده است. اما روبرو شدن با مسائل پویای تخصیص منابع سبب می‌شود که اهمیت بخش جایابی منبع تا حد زیادی تضعیف گردد و در مقابل، اهمیت بخش تخصیص، فزونی یابد. روش پیشنهادی این مقاله به گونه‌ای است که با در نظر گرفتن موقعیت متحرک برای منابع، با استفاده از روش‌های خوشه‌بندی به تخصیص منابع به تقاضاها می‌پردازد. مزیت روش پیشنهادی، با ترکیب روش آماری کا-فانکشن محلی و سلسله‌مراتبی، در آن است که منطقه مورد مطالعه با استفاده از روش کا-فانکشن محلی به دو بخش همبسته و ناهمبسته تفکیک می‌گردد. به منظور پیاده‌سازی کا-فانکشن محلی از شبیه‌سازی مونته‌کارلو بهره گرفته شده است. از آن‌جا که عملیات تفکیک در مقیاس بزرگ تعیین می‌گردد؛ لازم است با استفاده از روش سلسله‌مراتبی، خوشه‌های کوچک‌تر شکل گیرد. خوشه‌های کوچک باید دو شرط را برآورده نمایند: ۱) تا حد ممکن از ظرفیت کارشناس تصادفات استفاده شود ۲) مدت زمان رسیدن کارشناس به محل حادثه از حد معینی تجاوز ننماید. نتایج ارزیابی‌ها، نشان از بازدهی این روش در روزهای کاری هفته و در ساعات غیر از نیمه‌شب دارد. این در حالی است که در روزهای تعطیل هفته و همچنین در ساعات نیمه‌شب، استفاده از روش سلسله‌مراتبی به تنهایی سبب صرفه‌جویی در تعداد کارشناسان تصادف می‌گردد.

کلیدواژه‌ها: خوشه‌بندی، تخصیص منابع، تصادفات، کا-فانکشن، سلسله‌مراتبی.

* نویسنده مکاتبه‌کننده: تهران، خیابان کارگر شمالی، پردیس دانشکده های فنی دانشگاه تهران، دانشکده مهندسی نقشه برداری و اطلاعات مکانی. کدپستی: ۱۴۳۹۹۵۷۱۳۱

تلفن: ۹۸۲۱۶۱۱۱۴۵۲۰

Email: fhakimpour@ut.ac.ir

۱- مقدمه

نحوه تخصیص منابع موجود، همواره یکی از چالش‌های مدیریت شهری بوده است. به طور کلی در عملیات تخصیص و جایابی[1]، ابتدا جایابی منابع و سپس تخصیص هر کدام از تقاضاها به منابع انجام می‌پذیرد. تخصیص، فرایندی است که طی آن بخش‌هایی از شبکه یا مساحتی از محدوده بر اساس یکسری شرایط از قبل تعیین شده، به یک مرکز خدماتی اختصاص می‌یابد [۱]. به این فرآیند ناحیه‌بندی[2] نیز گفته می‌شود. ناحیه‌بندی آموزشی [۲]، بهداشتی [۳]، پشتیبانی [۴]، سیاسی [۵]، پلیس [۶] و اورژانس [۷] از نمونه‌های کاربردی تخصیص منابع می‌باشند. از دیاگرام ورونوی [۸] که یکی از عملگرهای نزدیکی[3] در سیستم‌های اطلاعات مکانی می‌باشد، به‌عنوان ابزاری مناسب و مرسوم در ناحیه‌بندی استفاده می‌شود. از آن‌جا که، بسیاری از مسائل ناحیه‌بندی شهری بدون لحاظ نمودن شبکه معابر شهری قابل حل نبوده و تجزیه و تحلیل این مسائل با فرض اندازه‌گیری فاصله اقلیدسی پاسخ درست و واقعی نخواهد داد [۹]، ورونوی شبکه توسعه یافته و در زمینه ناحیه‌بندی مورد استفاده قرار گرفته است[۱۰, ۱۱]. فرایند ناحیه‌بندی معمولا با معیارهایی مطابق عرف مقید می‌شود. یکی از این معیارها برقراری تعادل میان عرضه و تقاضا در نواحی مختلف است. علاوه بر آن، نواحی حاصله باید همجوار[4] و از لحاظ مکانی متراکم[5] باشند.

جهت مشخص نمودن رابطه بین تخصیص منابع و خوشه‌بندی لازم است که تعریف خوشه‌بندی ارائه گردد. خوشه‌بندی به فرآیند جداسازی مجموعه‌ای از اشیاء واقعی یا انتزاعی گفته می‌شود به گونه‌ای که نمونه‌های متشابه در خوشه‌های یکسان قرار گیرند [۱۲]. با این دیدگاه که مجموعه نیازمندی‌هایی که در یک ناحیه قرار می‌گیرند تشکیل دهنده‌ی یک خوشه هستند، مسئله تخصیص منابع را می‌توان یک مسئله خاص خوشه‌بندی دانست. تفاوت اصلی تخصیص منابع و خوشه‌بندی در تعیین موقعیت منبع درون خوشه است. بدین معنا که در خوشه‌بندی نیازی به تعیین محل دقیق استقرار منبع وجود ندارد ولی در یک مسئله معمول تخصیص منابع این کار صورت می‌پذیرد [۱۲]. در بحث تخصیص منابع، دو دسته منبع وجود دارد: ایستا و متحرک. منبع ایستا همچون مدرسه و بیمارستان و منبع متحرک مانند کارشناس تصادف و کارشناس پرداخت خسارت سیار بیمه. اکثر سازمان‌ها و شرکت‌های فعال در حوزه اطلاعات مکانی به ارائه راه‌حل‌هایی به‌منظور تخصیص منابع ایستا اقدام نموده‌اند. به‌عنوان مثال شرکت *ESRI* در افزونه آنالیز شبکه[6] امکان جایابی و تخصیص مکانی[7] منابع را پیاده‌سازی نموده است؛ اما برای منابع متحرک راه‌حلی ارائه نکرده است. به‌منظور پوشش این مهم، روش ارائه شده در این تحقیق از خوشه‌بندی در راستای تخصیص منابع استفاده می‌نماید. بدین صورت که ابتدا خوشه‌بندی تقاضاهای مرتبط با منابع متحرک انجام می‌شود. سپس با تعمیم خوشه‌ها به شبکه راه از طریق دیاگرام ورونوی، منطقه مورد مطالعه به طور کامل و با کمترین تعداد منبع ناحیه‌بندی می‌گردد.

میزان وابستگی بین اشیاء، مبنایی جهت تشخیص خوشه‌ها خواهد بود. از دیدگاه علم اطلاعات مکانی، تشخیص میزان وابستگی در بسیاری از رویدادهای دنیای واقعی معمولا تواماً در بُعد مکان و زمان صورت می‌پذیرد. از آن‌جا که تصادفات یکی از انواع پدیده‌هایی است که زمان و مکان

[1] Location-Allocation
[2] Districting
[3] Proximity
[4] Contiguous
[5] Compact

[6] Network Analyst Extension
[7] Location-Allocation

آن نقش بسزایی در تحلیل‌های مربوطه دارد، تمرکز اصلی این تحقیق بر روی این نوع داده قرار گرفته است. به‌طور کلی گروهی از تحلیل‌ها در صدد تشخیص عوامل ایجاد تصادفات هستند تا بتوانند نسبت به برطرف کردن آن‌ها اقدام نمایند و گروهی دیگر به دنبال مدیریت و رسیدگی سریع بعد از وقوع تصادفات می‌باشند. هدف این تحقیق در راستای تحلیل‌های گروه دوم می‌باشد.

راهبرد معرفی شده در این تحقیق بر روی داده تصادفات شهری پیاده‌سازی خواهد شد که خروجی آن نواحی تحت پشتیبانی منابعی همچون کارشناسان تصادفات خواهد بود. نواحی تعیین شده باید به گونه‌ای باشند که کنترل آن‌ها با صرفه‌جویی در تعداد کارشناسان به کارگرفته شده همراه باشد و از حداکثر توان آن‌ها تجاوز ننماید. ضمنا مدت زمان رسیدن منبع به محل تقاضا در هر کدام از نواحی، نباید از حد استانداردی فراتر رود.

۲- پیشینه تحقیق

لی و یانگ [۱۳] در سال ۲۰۰۵ جهت ترکیب روش‌های افراز و سلسله‌مراتبی[1] در خوشه‌بندی، تکنیکی را ارائه کردند که در آن از اطلاعات توپولوژی با استفاده از دیاگرام ورونوی استفاده می‌شد. در تحقیق آنان از بین روش‌های افراز، کا-مینز[2] به دلیل بازدهی[3] خوب آن انتخاب شد، هر چند که عدم توانایی این روش در کشف خوشه‌های غیر دایره‌ای از معایب آن عنوان شده است. روش سلسله‌مراتبی نیز به دلیل داشتن تاثیرگذاری[4] بالا و قابلیت یافتن خوشه‌های غیردایره‌ای دومین روش بکارگرفته شده آنان بود تا از این طریق بتوانند از مزایای هر کدام از دو روش استفاده نمایند.

مراحل این تکنیک بدین صورت است که ابتدا با روش کا-مینز تعداد K نقطه در مراکز خوشه‌ها جایابی می‌شود. سپس همسایه هر کدام از نقاط با استفاده از دیاگرام ورونوی استخراج می‌گردد و از روی دیاگرام ورونوی مقدار جذابیت[5] هر سلول ورونوی بر سلول‌های مجاور محاسبه می‌شود. در پایان با استفاده از روش سلسله‌مراتبی ادغام سلول‌های کاندید انجام می‌گیرد. علت به‌کارگیری اطلاعات توپولوژی در این روش آن است که بتوان خوشه‌هایی با اشکال مورد نظر را استخراج کرد.

تکنیک مورد استفاده آنان دارای تفاوت‌هایی با موضوع مطرح در این مقاله است که می‌توان به آن‌ها اشاره نمود:

- توپولوژی در فضای صفحه با فضای شبکه متفاوت است. لذا شکل خوشه‌های استخراجی از روی شبکه مشابه صفحه نخواهد بود.
- پیاده‌سازی این روش بر روی داده‌های شبیه‌سازی شده انجام گرفته و به‌کارگیری آن بر روی داده‌های واقعی به نظر مشکل می‌رسد.
- تعداد خوشه‌ها قبل از پردازش باید تعیین گردد.
- شرط خاتمه‌ای برای پایان ادغام خوشه‌ها در نظر گرفته نشده است.
- کنترلی بر روی اندازه خوشه‌ها و یا تعداد آن‌ها وجود ندارد.

لیائو و گوئو [۱۴] در سال ۲۰۰۸ مقاله‌ای را به‌منظور حل مسئله تخصیص امکانات با استفاده از روش‌های خوشه‌بندی ارائه نمودند. قید تخصیص منابع در نظر گرفته شده در این مقاله متغیر بودن ظرفیت منابع و احتمال ناکافی بودن ظرفیت منبع جهت پوشش کامل تقاضاها بوده است. روش پیشنهادی آنان از دو فاز تشکیل می‌شد:

۱) تخصیص امکانات با ظرفیت متغیر به تقاضاها

[1] Hierarchical
[2] K-means
[3] Efficiency
[4] Effectiveness
[5] Attractiveness

به‌گونه‌ای که کمترین هزینه را داشته باشد.
۲) بهینه‌سازی تکراری موقعیت امکانات با استفاده از روش خوشه‌بندی کا-مینز. این روش خوشه-مبنا با یک روش هوش مصنوعی با استفاده از تکنیک ژنتیک مقایسه شده است. الگوریتم ژنتیک در قسمت دوم یعنی تعیین بهینه موقعیت امکانات به کار گرفته می‌شود. به عبارت دیگر عملیات تخصیص تقاضاها در هر دو روش یکسان است. تفاوت‌های این روش با مسئله مطرح در مقاله پیش‌رو را می‌توان از چند جنبه بیان نمود:

- تخصیص تنها برای نقاط تقاضا انجام گرفته است. به عبارت دیگر روشی برای تعیین نواحی ارائه نشده است.
- تعداد منابع قبل از پردازش باید تعیین گردد.
- فاصله اقلیدسی در نظر گرفته شده است.
- ممکن است برخی تقاضاها در هیچ کدام از خوشه‌ها جای نگیرند.

لیو و همکارانش [۱۵] در سال ۲۰۱۱ مسئله تخصیص منابع را با استفاده از ترکیب دو روش خوشه‌بندی مبتنی بر گرید[1] و سلسله‌مراتبی حل نمودند. در روش آن‌ها منطقه مورد مطالعه به سلول‌های مربع شکل تقسیم می‌شد و به هر سلول یک بردار سه مولفه‌ای (x, y, z) اختصاص می‌یافت. در این بردار x و y مختصات مرکز سلول و z مقدار مشخصه توصیفی آن است. از آن‌جا که در این روش تخصیص ایستگاه‌های سوخت‌رسانی در محدوده شهری مد نظر بوده است لذا مشخصه توصیفی سلول‌ها مقدار جریان ترافیک در نظر گرفته شده است. در ادامه با استفاده از روش سلسله‌مراتبی، ادغام سلول‌های مشابه انجام می‌گیرد. این کار تا رسیدن به تعداد خاصی خوشه ادامه می‌یابد. در پایان، خوشه‌هایی که حجم ترافیک در آن‌ها از یک مقدار حداقلی کمتر باشد، حذف می‌گردند. تمرکز اصلی تحقیق انجام گرفته بر روی شیوه تشکیل خوشه‌هاست و بر روی نحوه بهینه‌سازی موقعیت منابع بحث نمی‌گردد. مشکلات این روش را می‌توان سه مورد ذکر کرد:

- فاصله اقلیدسی در نظر گرفته شده است.
- در این روش تعادل در خوشه‌های حاصله به وجود نمی‌آید و همان‌طور که گفته شد با حذف خوشه‌های کم ترافیک به نوعی حالت جزیره‌ای در منطقه به وجود می‌آید.
- رسیدن به تعداد خاص خوشه برای شرط پایان ادغام در روش سلسله‌مراتبی با مسئله مطرح در این مقاله تطابق ندارد هر چند که برای مسئله در نظر گرفته شده در مقاله مذکور مناسب به‌نظر می‌رسد.

۳- روش‌های رایج خوشه‌بندی بر روی شبکه

روش‌های رایج خوشه‌بندی را می‌توان به دو دسته آماری و غیرآماری تقسیم نمود. خوشه‌بندی آماری نقاط که قسمتی از حوزه تحلیل الگوی نقاط[2] می‌باشد، در صدد استخراج الگو از بین نقاط با استفاده از آماره‌ها است. در این روش‌ها از فرض صفر استفاده می‌شود. فرض صفر توزیع تصادفی را برای نقاط در نظر می‌گیرد. در صورتی که این فرض رد شود، نشان از میل نقاط به تشکیل خوشه و یا منظم بودن چینش آن‌ها خواهدبود. در تحلیل الگوی نقاط در فضای دو بعدی، دو گروه از فنون خوشه‌بندی ارائه شده است [۱۶, ۱۷]. گروه اول آزمون‌هایی برای تشخیص وجود خوشه و یا اندازه‌گیری درجه آن‌ها است. این گروه از روش‌ها تنها وجود یا عدم وجود تمایل به شکل‌گیری خوشه را در یک الگوی مشاهداتی مشخص می‌کند و قابلیت تعیین موقعیت و اندازه خوشه‌ها را ندارند. به عبارت دیگر این روش‌ها در مقیاس سراسری[3]

[1] Grid-based

[2] Point Pattern Analysis
[3] Global scale

کار می‌کنند. دومین گروه شامل روش‌هایی است که علاوه بر تشخیص تمایل به شکل‌گیری، به تعیین موقعیت خوشه‌ها نیز می‌پردازد. در گروه اول می‌توان به روش *کا-فانکشن*[1] سراسری اشاره کرد [۱۸]. شیوه ارائه شده توسط اودریسکول[2] و اوکابه[3][۱۹, ۲۰]، باعث تعمیم روش کا-فانکشن روی شبکه شد. از اواخر دهه ۸۰ تمرکز محققین به‌سمت روش‌های گروه دوم یعنی آن‌هایی که تمایل به شناسایی خوشه‌ها به همراه موقعیت آن‌ها را دارند منعطف شد و دلیل آن نیز می‌تواند جزئیات بیشتر و اطلاعات عملی‌تری باشد که توسط این گروه از روش‌ها حاصل می‌شود. از آن‌جا که نقاط واقع شده درون هر کدام از خوشه‌های این گروه از لحاظ آماری با یکدیگر وابستگی دارند، به این مناطق، مناطق همبسته[4] نیز گفته می‌شود. نسخه محلی کا-فانکشن در گروه دوم روش‌های خوشه‌بندی جای می‌گیرد.

در روش‌های غیرآماری وضعیت توزیع نقاط به تنهایی و بدون در نظر گرفتن فرض صفر بررسی می‌شود. به عبارت دیگر امکان مقایسه رویدادهای واقعی با نمونه‌های قابل انتظار وجود ندارد. این امر باعث می‌شود که در برخی موارد از دیدگاه غیرآماری وجود چند نقطه در یک منطقه، خوشه تلقی گردد در حالی‌که روش‌های آماری چنین تشخیصی نمی‌دهند. از مزایای خوشه‌بندی غیرآماری نسبت به آماری، در نظر گرفتن ترجیحات کاربر از جمله تعداد و اندازه خوشه‌ها است. روش‌های *کا-مینز*، *کا-مدویدز*[5] و سلسله مراتبی نمونه‌های خوشه‌بندی غیرآماری هستند.

۳-۱- کا-فانکشن محلی شبکه مبنا[6]

نسخه محلی کا-فانکشن روی شبکه توسط یامادا[7] و همکارانش ارائه شد [۲۱]. در این روش با اندازه‌گیری مقدار *کا* محلی (LK) بر روی یکسری نقاط مرجع[8] اقدام به شناسایی مقادیر معنادار می‌شود. نقاط مرجع، نقاطی هستند که به‌صورت تصادفی بر روی شبکه توزیع می‌گردند و تمایل به ایجاد خوشه بر روی آن‌ها اندازه‌گیری می‌شود. در صورتی که P مجموعه رویدادها، E امید ریاضی و m نشان‌دهنده تعداد نقاط مرجع توزیع‌یافته روی شبکه T باشد برای هر نقطه مرجع j $(j=1,...,m)$، کا-فانکشن محلی روی شبکه به صورت فرمول (۱) تعریف می‌شود:

$$LK_j(r) = E\left(\begin{array}{c}\text{تعداد نقاط در } \boldsymbol{P} \text{ در فاصله} \\ \boldsymbol{r} \text{ از نقطه مرجع } \boldsymbol{j}\end{array}\right) \qquad \text{رابطه (۱)}$$

$$r \geq 0 \qquad j = 1,...,m$$

اگر $d(i,j)$ فاصله شبکه‌ای دو نقطه i و j از هم باشد و تابع شناساگر به صورت فرمول (۲) تعریف شود:

$$I_r(d(i,j)) = \begin{cases} 1 & if \quad d(i,j) \leq r \\ 0 & Otherwise \end{cases} \qquad \text{رابطه (۲)}$$

برآوردکننده نقاط مشاهداتی از طریق فرمول (۳) به‌دست می‌آید:

$$\hat{LK}_j(r) = \sum_i I_r(d(i,j)) \qquad \text{رابطه (۳)}$$

مقدار این آماره در مقابل مقدار فرض صفر قرار می‌گیرد. فرض صفر توزیع یکنواخت و مستقل را برای رویدادها انتظار دارد. از آن‌جا که توزیع دوجمله‌ای به‌طور ضمنی فرض یکنواختی و استقلال توزیع نقاط را روی شبکه در بر دارد، توزیع انتخاب‌شده معمولا از نوع دوجمله‌ای خواهد بود[۲۰]. مقدار برآوردکننده نقاط شبیه‌سازی شده را با $LK^{sim}{}_j(r)$ نشان می‌دهند. در صورتی که $\hat{LK}_j(r) > LK^{sim}{}_j(r)$ باشد، آنگاه نقطه مرجع j در شعاع r اطراف آن، یک خوشه را

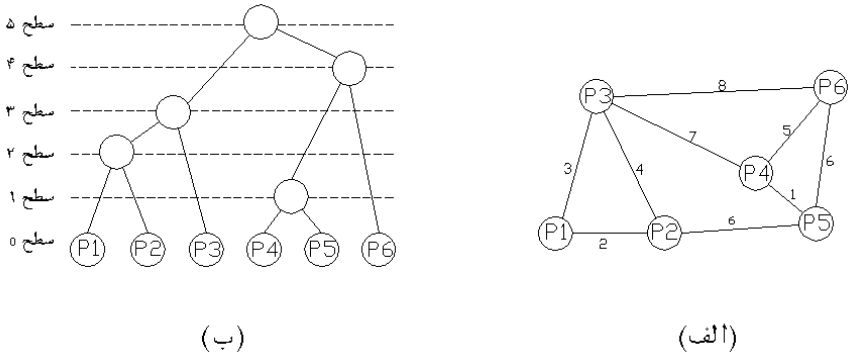
[1] K-function
[2] O'Driscoll
[3] Okabe
[4] Correlated
[5] K-medoids
[6] Local network *K-function*
[7] Yamada
[8] Reference points

تشکیل می‌دهد.
نکته حائز اهمیت در این روش انتخاب مقیاس مناسب است. بدین منظور به‌طور گسسته چندین مقیاس مورد ارزیابی قرار می‌گیرد. اگر r_h مقیاس خوشه‌بندی در نظر گرفته شود و $r_h = lh$ به ازا ی $h = 1, 2, ..., q$ و l طول واحد باشد، منطقه مورد مطالعه با استفاده از کا-فانکشن سراسری و r_h های مختلف مورد ارزیابی قرار می‌گیرد. مقیاسی که ایجاد کننده بیشترین اختلاف مثبت بین $\hat{K}$ و K باشد به‌عنوان مقیاس مناسب انتخاب می‌گردد.

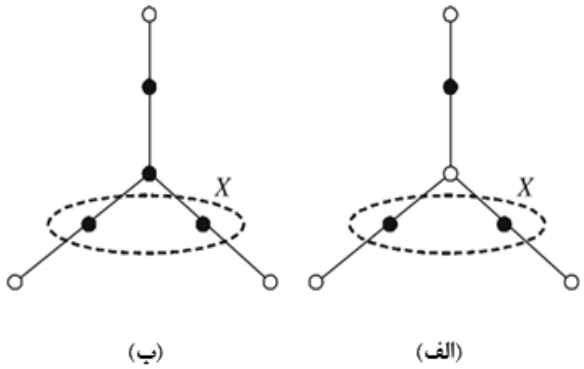
۳-۲- خوشه‌بندی سلسله‌مراتبی[1]

روش سلسله‌مراتبی را می‌توان از طریق یک مثال (شکل۱-(الف)) شرح داد. این مثال از شش راس و اتصال‌های[2] بین آن‌ها تشکیل شده است. فاصله راس‌های مرتبط بر روی اتصال مربوطه نوشته شده است. در شکل۱-(ب) که نمودار دارنگاشت[3] این مثال را نشان می‌دهد، هر گره یک خوشه تک‌عضوی را مشخص می‌کند که خوشه‌های سطح صفر هستند. فاصله بین $\{p_4\}$ و $\{p_5\}$ کوچک‌ترین است. بنابراین با یکدیگر ترکیب شده و $\{p_4, p_5\}$ را می‌سازند و این خوشه به‌عنوان یک گره در مرحله بالاتر نمایش داده می‌شود. همان‌طور که سطح ساختار، بالا می‌رود، تعداد خوشه‌ها یکی یکی کاهش می‌یابد. در سطح $n-1$ تنها یک خوشه از کلیه نقاط P باقی خواهد ماند. به این فرآیند، خوشه‌بندی سلسله مراتبی و ترتیب حاصل از مجموعه خوشه‌ها را ساختار سلسله مراتبی خوشه‌ها[4] گویند.
از آن‌جا که با ساختار شبکه کار می‌شود، انتظار طبیعی آن است که خوشه‌ها، تکه‌های به هم متصل شبکه را نشان دهند و این مهم‌ترین خصوصیت تحلیل خوشه‌بندی روی شبکه در مقایسه با نمونه معمول آن در صفحه است. برای تضمین وجود چنین خصوصیتی، محدودیتی در روش سلسله‌مراتبی، اعمال می‌گردد که در ادامه آورده می‌شود.
اگر P مجموعه رویدادها بر روی شبکه باشد و X زیرمجموعه‌ای از P باشد، به آن متصل درون خوشه‌ای[5] در P گویند به شرطی که هر دو راس در X از طریق مسیری در T به یکدیگر متصل شوند که از هیچ راس دیگری در $P \setminus X$ عبور نکند. به عبارت دیگر همه رئوس رویداد غیر از X و اتصالات آن‌ها در T را حذف کرده و متصل بودن X در بقیه شبکه تعریف می‌شود. به‌عنوان مثال شبکه شکل ۲ را در نظر بگیرید که دوایر توپر، رئوس رویداد یا به عبارتی عناصر P را نشان می‌دهند و دوایر توخالی، رئوس غیر رویدادی را نشان می‌دهند. X زیرمجموعه‌ای از رئوس رویداد است که با منحنی بسته خط‌چین در اطراف آن مشخص شده‌است. دو راس در X از طریق مسیر یکه عبوری از راس مرکزی به هم متصل شده‌اند.
در شکل ۲- (الف) راس مرکزی به P تعلق ندارد و بنابراین X متصل درون خوشه‌ای است و در شکل ۲-(ب) راس مرکزی متعلق به P است و بنابراین X متصل درون خوشه‌ای نمی‌باشد. بر این اساس، قانون ذیل شکل می‌گیرد:
قانون اتصال درون خوشه‌ای: در هر سطحی از خوشه‌بندی، دو خوشه X و Y زمانی می‌توانند با هم ترکیب شوند که $X \cup Y$ متصل درون خوشه‌ای در P باشد.

[1] Hierarchical clustering
[2] Edge
[3] Dendogram
[4] Hierarchical structure of the clusters
[5] Intra-cluster connected



شکل ۱-(الف) گراف متشکل از شش راس رویداد (ب) نمودار دارنگاشت گراف



شکل ۲-(الف) یک زیرمجموعه متصل درون خوشه‌ای (ب) زیرمجموعه غیرمتصل [۲۲]

۳-۳- خوشه‌بندی بر اساس چگالی

عملکرد این روش به طور کلی با بررسی وجود حداقل تعداد مشخصی از نقاط در یک فاصله معلوم از نقطه تصادفی p انجام می‌گیرد. اگر شرط مذکور برقرار باشد یک خوشه برای p و نقاط در آن فاصله شکل می‌گیرد. این کار برای همه نقاط درون خوشه ادامه می‌یابد. الگوریتمی با این روش به نام اپسیلون-لینک[1] در [۲۳] ارائه شده است. در این الگوریتم، نقاط خوشه‌بندی نشده از مجموعه نقاط موجود، انتخاب شده و سنجش وجود حداقل چگالی، روی آن‌ها اجرا می‌گردد. اگر m، نقطه مورد نظر باشد، الگوریتم از آن‌جا شروع می‌شود که ابتدا اتصالی که نقطه m روی آن قرار گرفته بررسی می‌شود. اگر نقطه دیگری روی این اتصال، حداکثر به فاصله ε قرار داشته باشد، در خوشه جاری C قرار می‌گیرد. با ورود یک نقطه گرهی در خوشه C، بررسی حداقل چگالی بر اساس فاصله ε بر روی آن گره نیز اجرا می‌شود و بدین ترتیب خوشه C بزرگ‌تر می‌گردد تا جایی که دیگر امکان ورود هیچ گره دیگری به خوشه C وجود نداشته باشد.

۴-۳- خوشه‌بندی بر اساس افراز[2]

روش‌های افراز، نقاط را به k گروه مختلف تقسیم می‌کنند و تا زمانی که در کیفیت خوشه‌ها بهبودی حاصل نشود، به صورت تکراری این نقاط

[1] $\varepsilon - Link$

[2] Partitioning

بین خوشه‌ها مبادله می‌شوند [۲۳]. *کا-مینز* و *کا-مدویدز* روش‌های موجود در این طبقه هستند. در الگوریتم‌های *کا-مینز*، مرکز خوشه‌ها با یک نقطه متوسط، مثل مرکز ثقل نمایش داده می‌شود و تبادل نقاط تا همگرایی فاصله میانگین نقاط با مقدار متوسط ادامه می‌یابد. در الگوریتم‌های *کا-مدویدز* مثل پم[1]، کلارا[2] و کلارنس[3] مرکز خوشه‌ها با یک نقطه حقیقی از بین نقاط موجود نشان داده می‌شوند. به این نقطه حقیقی که جایگزین نقطه میانگین می‌گردد، *مدویدز* گفته می‌شود. ابتدا k تا نقطه به‌صورت تصادفی به عنوان *مدویدز* انتخاب می‌شود و جمع فاصله نقاط تا نزدیک‌ترین *مدویدز* به‌عنوان تابع ارزیابی، محاسبه می‌شود. سپس یک *مدویدز* با نقطه تصادفی دیگر جایگزین می‌شود. اگر مقدار تابع ارزیابی کوچک شد، تغییر اعمال می‌شود و در غیر این صورت یک نقطه تصادفی دیگر امتحان می‌گردد. این فرایند برای تعداد تکرار مشخصی انجام می‌گیرد و در صورتی که بعد از این تعداد، بهبودی در تابع ارزیابی حاصل نشد الگوریتم قطع شده و آخرین مرحله به‌عنوان خوشه‌های بهینه محلی در نظر گرفته می‌شود. اگر چه در این روش کل نقاط در خوشه‌ها جای می‌گیرند ولی تعیین پارامتر k از محدودیت‌های آن می‌باشد.

۴- روش پیشنهادی

در گام اول روش پیشنهادی، لازم است که در منطقه مورد مطالعه رویدادهایی که با زمان مورد نظر برای برنامه‌ریزی وابستگی دارد، استخراج گردد. به‌طور کلی یکی از شیوه‌های در نظر گرفتن زمان در تحلیل‌ها، استخراج بازه‌ای رویدادهای وابسته قبل از ورود به تحلیل مورد نظر است. شبکه معابر شهری نیز در مقیاس‌های مختلف زمانی تغییرات ترافیکی زیادی را تجربه می‌کند. اهمیت شناخت مقیاس‌های زمانی تغییر ترافیک به این دلیل است که بتوان ارتباط رویدادهای مختلف (در این‌جا تصادفات) را جهت برنامه‌ریزی و تخصیص منابع تشخیص داد. تغییرات ترافیکی به صورت دقیقه‌ای، ساعتی، روزانه، هفتگی، ماهانه و سالانه وجود دارد. شیوه استخراج تصادفات وابسته در این تحقیق با ذکر مثالی مشخص می‌گردد. در این مثال هدف برنامه‌ریزی جهت تخصیص منابع در تاریخ ۱۵ می ۲۰۱۲ از ساعت ۱۶ تا ۱۸ و داده تصادفات برای سال‌های ۲۰۰۰ تا ۲۰۰۹ در دسترس است. تمامی تصادفاتی که در تاریخ‌های ۱ تا ۳۱ می از ساعت ۱۶ تا ۱۸ و در سال‌های ۲۰۰۰ تا ۲۰۰۹ رویداده‌اند، استخراج می‌گردد. لازم به ذکر است از آن‌جا که این روز یک روز کاری محسوب می‌شود، تنها روزهای کاری این بازه در تحلیل شرکت می‌کنند و تعطیلات از آن کسر می‌گردد.

بعد از استخراج رویدادهای وابسته زمانی نوبت به تفکیک منطقه به دو بخش همبسته و ناهمبسته می‌رسد. این کار بدین دلیل انجام می‌گیرد که برای هر کدام از مناطق، پارامترهای مجزایی جهت تخصیص منابع استفاده گردد. کا-فانکشن محلی که در قسمت انواع روش‌های خوشه‌بندی به آن اشاره شد، روشی است که مناطق همبسته را از ناهمبسته جدا می‌کند.

در ادامه لازم است که کوچک‌سازی مناطق همبسته و ناهمبسته تا رسیدن به اندازه مشخص انجام گیرد. کوچک‌سازی مناطق به منظور تخصیص بهینه منابع در بین تقاضاها انجام می‌گیرد. روش‌های غیرآماری خوشه‌بندی بر خلاف روش‌های آماری، دارای این خصوصیت می‌باشند که ترجیحات کاربر را (همچون اندازه خوشه و یا تعداد خوشه) در نظر می‌گیرند. از میان سه روش سلسله مراتبی، افراز و چگالی-مبنا، یکی را باید به عنوان ابزاری برای این منظور انتخاب نمود. از نکاتی که باید مورد توجه

[1] *PAM*
[2] *CLARA*
[3] *CLARANS*

قرار گیرد آن است که در تخصیص منابعی همچون کارشناسان تصادفات، کل محدوده شهر باید پوشش داده شود. یعنی روش به کار گرفته شده نباید نواحی گسسته و از هم جدا ایجاد نماید. در روش چگالی-مبنا نیاز مذکور برآورده نمی‌گردد چرا که این روش تنها به استخراج مناطق پرتراکم می‌پردازد و باقی مناطق را وا می‌گذارد. در روش‌های گروه افراز نیز الزام به مشخص بودن تعداد منابع قبل از شروع پردازش، عدم انطباق با مسئله مطرح در این تحقیق را به دنبال دارد. لازم به یادآوری است که هدف این تحقیق، ناحیه‌بندی شهر با کمترین تعداد منبع و استفاده بهینه از همه آن‌هاست. تنها روشی که با این رویکرد منطبق خواهد بود، روش سلسله‌مراتبی است.

اما مهم‌ترین تصمیم در استفاده از روش سلسله‌مراتبی، نحوه‌ی پایان دادن به ادغام خوشه‌ها است. به دلیل آن‌که هدف، ایجاد خوشه‌هایی با اندازه مشخص است، (به گونه‌ای که هر کارشناس تصادفات بتواند به آن‌ها رسیدگی کند) لذا اندازه خوشه یعنی تعداد نقاطی که در یک خوشه جای می‌گیرند، معیار پایان دادن به عملیات ادغام انتخاب می‌گردد. در کنار قوانین محدودکننده‌ای مانند قانون اتصال درون خوشه‌ای، تعداد رویدادهای واقعی معمولا مضربی از اندازه خوشه نمی‌باشند. بنابراین اندازه هر خوشه همواره نمی‌تواند با آن‌چه که از قبل معین می‌گردد، برابر باشد (ممکن است از آن کمتر باشد). ناگزیر حداکثر اندازه خوشه به عنوان پارامتر پایان ادغام درنظر گرفته می‌شود.

اگر T بازه ساعتی استخراج داده، Y تعداد سال‌های مشارکتی، d تعداد روزهای مشارکتی در هر سال و n اندازه هر خوشه باشد، آنگاه فرمول‌های ۴ و ۵ به عنوان توابع هدف ارائه می‌گردد:

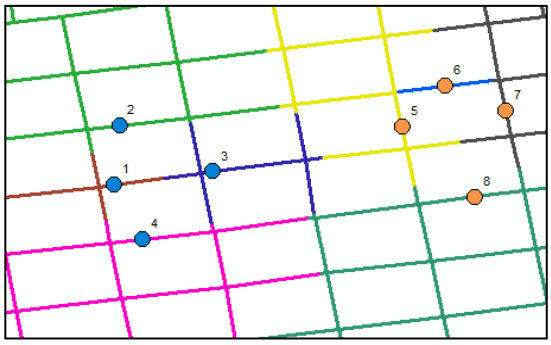
رابطه (۴) $$t_d \le t_s$$

رابطه (۵) $$T \ge \frac{n}{Y \times d}(t_d + t_p):$$
$$n = \max\{x \mid x \in N\} \quad (N = Natural\ numbers)$$

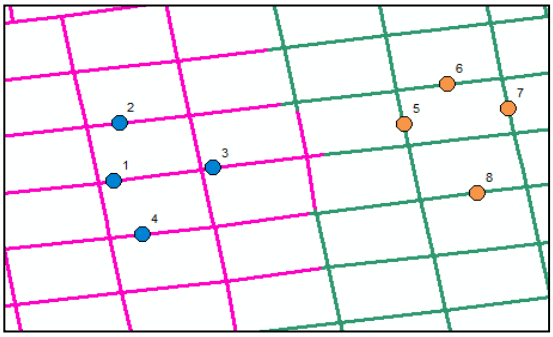
در فرمول‌های ۴ و ۵ ، t_s حداکثر زمان استاندارد رسیدن کارشناس به محل حادثه، t_d زمان واقعی رسیدن کارشناس به محل حادثه و t_p مدت زمان مورد نیاز جهت بررسی صحنه تصادف و دارای مقداری ثابت است. نامساوی ۴ نشان می‌دهد که زمان واقعی رسیدن کارشناس به محل حادثه بایستی کمتر از حداکثر زمان استاندارد باشد. فرمول ۵ نیز نشان می‌دهد که میزان تصادفات تخصیص داده شده به هر کارشناس تصادفات نباید از توان کارشناس بیشتر باشد. هر دو طرف نامساوی‌های ۴ و ۵، بر اساس واحد زمان است. در فرمول ۵ باید مقدار مناسب برای n پیدا شود. برای این منظور، اعداد طبیعی به طور صعودی جای‌گذاری می‌شوند و بزرگ‌ترین آن‌ها که نامساوی ۵ در آن برقرار است، به عنوان اندازه بهینه خوشه انتخاب می‌گردد. توابع هدف برای هر دو منطقه همبسته و ناهمبسته به‌کار گرفته می‌شود. روشن است که t_d در مناطق همبسته به دلیل نزدیکی تصادفات به یکدیگر، کوچک‌تر از مناطق ناهمبسته تصادفاتی خواهد بود. بنابراین، عدد n بزرگ‌تری در مناطق همبسته به دست می‌آید.

بعد از شکل‌گیری خوشه‌ها در دو منطقه همبسته و ناهمبسته با اندازه مشخص، نوبت به تشکیل نواحی با استفاده از دیاگرام ورونوی یال شبکه می‌رسد. در این روش، دیاگرام ورونوی یال شبکه هر کدام از تصادفات ترسیم می‌گردد و اجتماع دیاگرام‌هایی که نقاط مرکزی آن‌ها (تصادفات) در یک خوشه قرار دارد، محاسبه می‌گردد. این دیاگرام‌ها محدوده معابر تحت پوشش هر کارشناس تصادف را تعیین می‌نماید. شکل‌های ۳ و ۴ روند انجام این کار را نمایش می‌دهند. از آن‌جا که در مرحله کوچک‌سازی مناطق همبسته و

ناهمبسته از روش سلسله‌مراتبی استفاده گردیده و در این روش شرط اتصال درون خوشه‌ای در نظر گرفته شده است، پیوسته بودن نواحی تضمین خواهدشد. از سوی دیگر استفاده از دیاگرام ورونوی باعث اطمینان از پوشش کامل منطقه خواهدبود. چرا که دیاگرام ورونوی تا رسیدن به مرز ناحیه مورد مطالعه پیشروی می‌کند. در واقع در تخصیص منابع، مقاومت ظاهری در نظر گرفته نشده و تنها عامل قطع یال ورونوی، رسیدن به مرز ناحیه مورد مطالعه است.



شکل ۳- ترسیم دیاگرام ورونوی یال شبکه برای ۸ نقطه که در دو خوشه قرار گرفته‌اند.



شکل ۴- اجتماع دیاگرام‌های ورونوی یال شبکه مربوط به نقاط هر خوشه

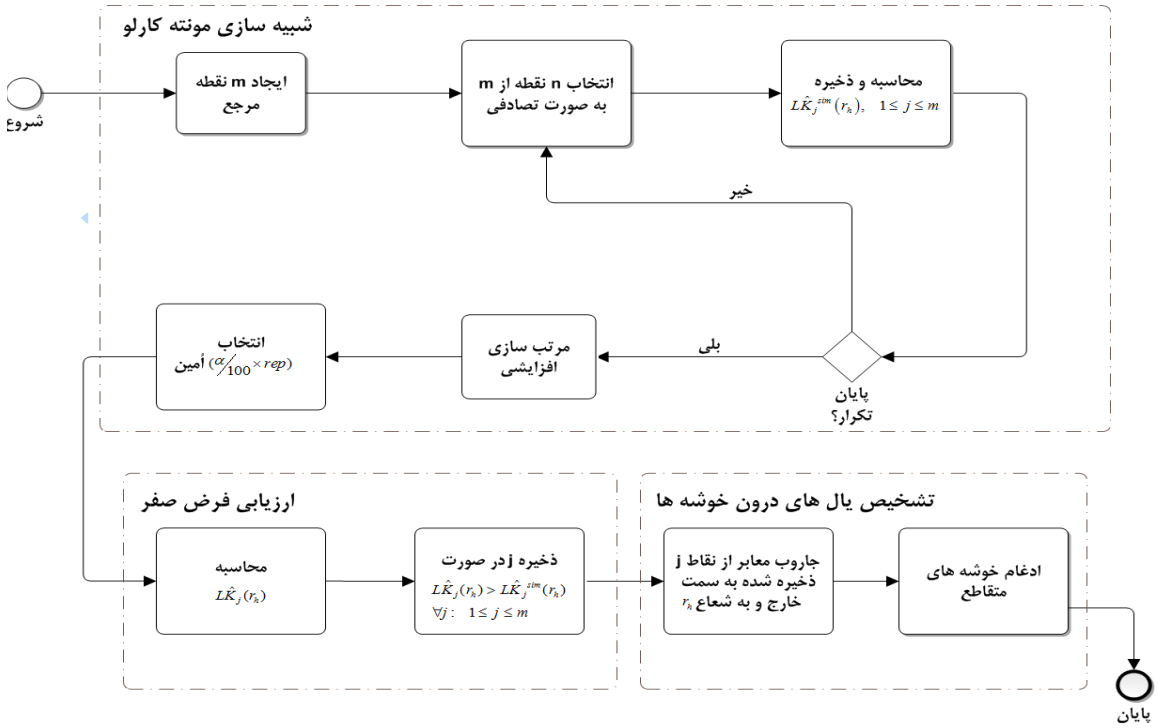
۵- پیاده‌سازی

استخراج داده با استفاده از کلاس‌های پرس‌وجوی توصیفی[1] به زبان SQL[2] انجام می‌گیرد. این استخراج با محدودکردن بازه زمانی سالانه، روزانه و ساعتی در کل منطقه به گونه‌ای که قبلا شرح داده شد، انجام می‌شود. خروجی مرحله استخراج داده، انتخاب رویدادهایی از پایگاه داده مکانی است که در بازه‌های انتخابی توسط کاربر جای گرفته‌اند.

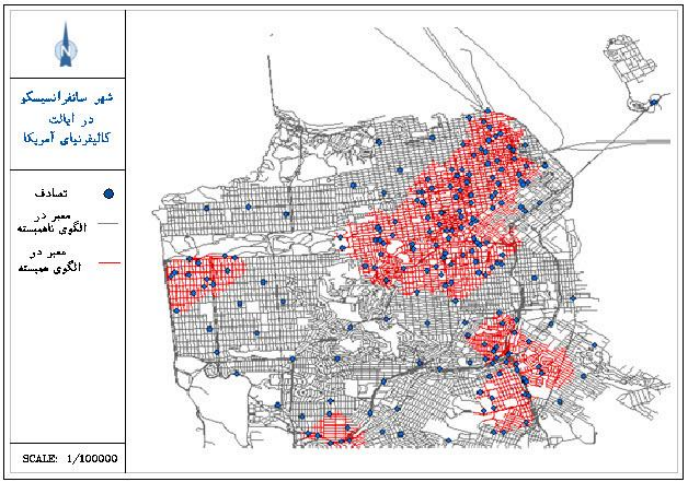
روش پیاده‌سازی مرحله دوم پروژه یعنی تفکیک الگوی نقاط به دو دسته همبسته و ناهمبسته با استفاده از شبیه‌سازی مونته‌کارلو مطابق شکل ۵ اجرا گشته است. این مدل فرآیند با ایجاد m نقطه مرجع که از توزیع دوجمله‌ای یکنواخت تبعیت می‌کند، شروع می‌شود. به‌منظور انجام این ارزیابی، لازم است که نقاط رویداد واقعی با نقاط شبیه‌سازی شده‌ای که به صورت تصادفی بر روی شبکه توزیع شده‌اند، مقایسه گردند. تعداد این نقاط باید با تعداد نقاط واقعی (n) برابر باشد. لذا به منظور تسریع در پردازش، تعداد n نقطه از m نقطه مرجع تولید شده به‌طور تکراری (تعداد تکرار= rep) انتخاب شده (روشن است که m باید بزرگ‌تر از n باشد) و در هر مرحله مقدار $\hat{LK}_j^{sim}(r)$ برای همه نقاط مرجع، محاسبه و ذخیره می‌گردد. پس از پایان تکرار، مجموع مقادیر $\hat{LK}_j^{sim}(r)$ برای همه نقاط مرجع در هر مرحله محاسبه شده و به‌صورت افزایشی مرتب می‌گردد. اگر α سطح اطمینان در نظر گرفته شده توسط کاربر باشد، آنگاه $(\alpha/100 \times rep)$ اُمین سطر از مراحل تکرار به عنوان نماینده مقادیر نقاط شبیه‌سازی شده انتخاب می‌گردد.

از سوی دیگر با محاسبه تابع $\hat{LK}_j(r)$ برای نقاط واقعی تصادفات، امکان بررسی فرض صفر محقق می‌گردد. فرض صفر، توزیع تصادفی را برای نقاط رویداد واقعی در نظر می‌گیرد. در صورتی که مقدار $\hat{LK}_j(r) > \hat{LK}_j^{sim}(r)$، آنگاه نقطه j به‌عنوان مرکز خوشه تلقی می‌گردد و فرض صفر رد می‌شود. در شکل ۶ خوشه‌های تشخیص داده شده، الگوی همبسته نقاط را نشان می‌دهند و باقی مناطق حکم مناطق ناهمبسته را خواهند داشت.

[1] Attribute Query
[2] Structured Query Language



شکل۵- دیاگرام شبیه‌سازی مونته کارلو در مرحله دوم روش پیشنهادی



شکل ۶- خروجی مرحله تفکیک الگوی نقاط

الگوی همبسته نقاط وارد الگوریتم خوشه‌بندی سلسله‌مراتبی می‌گردد تا اندازه خوشه‌های آن تعدیل گردد. به همین منظور ابتدا رویدادهای هر خوشه به‌صورت مجزا در نظر گرفته می‌شود و فاصله هر دو نقطه درون آن از طریق ماتریس OD[1] مشخص می‌گردد. به دلیل جهت‌دار بودن شبکه معابر،

[1] Origin-Destination

لازم است که فاصله نزدیک‌ترین جفت در دو خوشه مختلف در مسیرهای رفت و برگشت محاسبه شده و میانگین آن‌ها مورد استفاده قرار گیرد. فاصله میانگین رفت و برگشت نقاط به همراه شماره هر دو نقطه در یک آرایه ذخیره می‌گردد. سپس این آرایه از فاصله کوچک به بزرگ مرتب می‌گردد. نقاط با فواصل کوچک شروع به ترکیب می‌کنند و به هر کدام از آن‌ها، نام خوشه جدید اختصاص داده می‌شود. دو شرط بیشینه‌ی اندازه خوشه و اتصال درون خوشه‌ای جهت ترکیب دو خوشه در نظر گرفته شده است. این کار تا جایی پیش می‌رود که دیگر امکان ترکیب هیچ دو خوشه‌ای با هم وجود نداشته باشد.

اما قسمت خارج از محدوده همبسته نقاط نیز به‌صورت یک خوشه مجزا در نظر گرفته می‌شود و رویدادهای آن در خوشه‌های کوچک‌تر جای می‌گیرند. شبه‌کد ذیل روند کوچک‌سازی مناطق همبسته و ناهمبسته را نشان می‌دهد.

Cluster Shrinking
Input: Accidents, Link-Clusters, ClusterSize
Output: Point-Clusters
1. For every Link-Clusters do
2. Col ← Take out intersected Accidents with Link-Clusters
3. OD ← Create Origin-Distance matrix
4. Dist[1] ← Number of origin points
5. Dist[2] ← Number of Destination points
6. Dist[3] ← Calculate average distance from OD
7. Sort Dist incrementally based on Dist[3]
8. For every rows in Dist
C ← Cluster(Dist[1]) 9.
C´ ← Cluster(Dist[2]) 10.
L ← ClusterSize (C)+ClusterSize (C´) 11.
If L <= ClusterSize and merge (C,C´) is 12. intra- cluster connected then
Merge (C,C´) 13.

شبه‌کد ذیل نحوه ناحیه‌بندی با استفاده از دیاگرام ورونوی یال شبکه را نشان می‌دهد.

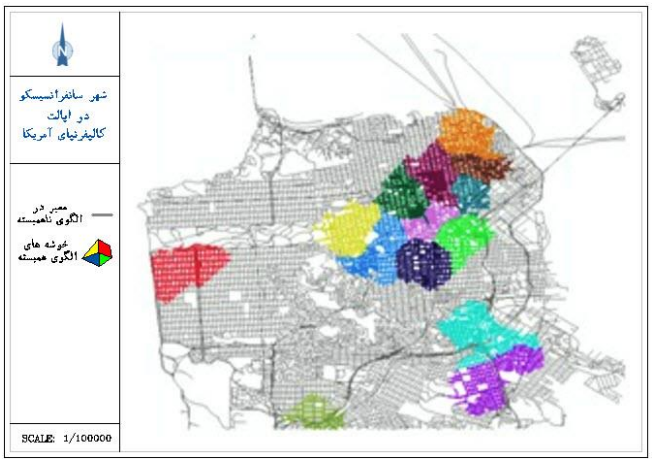
Districts Generation
Input: Point-Clusters
Output: Link-Clusters
1. For every Point-Clusters do
2. P ← take out points within Link-Clusters
3. VD ← Calculate Link Voronoi Diagram of P
Link-Cluster ← Union VD

اشکال ۷ و ۸ به ترتیب خوشه‌های کوچک شده در ناحیه همبسته و ناهمبسته را نشان می‌دهد.

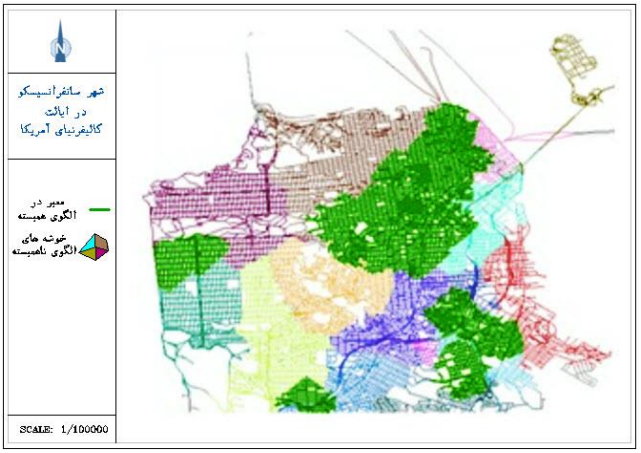
۶- ارزیابی

میزان کارایی روش پیشنهادی، با مطالعه موردی بر روی محدوده شهر سانفرانسیسکو در ایالت کالیفرنیای آمریکا سنجیده شده است. علت انتخاب این منطقه، در دسترس بودن داده تصادفات شهری آن طی سال‌های ۲۰۰۰ تا ۲۰۰۹ و همچنین شبکه کامل معابر آن بوده است. تاریخ ۱۷ اکتبر سال ۲۰۱۲، ساعت ۸ تا ۹ صبح جهت برنامه‌ریزی در نظر گرفته شد. داده تصادفات بین سال‌های ۲۰۰۷ تا ۲۰۰۹ و در بازه دو هفته قبل و بعد از تاریخ مذکور در ساعت ۸ تا ۹ صبح استخراج شد. از آن‌جا که تاریخ مورد نظر یک روز کاری بود، لذا در بازه استخراجی روزهای تعطیل حذف گردید. بنابراین ۱۰ روز قبل و بعد از آن تاریخ در طی سه سال در نظر گرفته شد (جمعا ۶۰ روز کاری).

به‌منظور استخراج الگوی نقاط، چهار مقیاس در نظر گرفته شد و مقادیر کا-فانکشن سراسری برای داده‌های شبیه‌سازی شده و واقعی محاسبه شد. در جدول ۱ که مقادیر مقیاس‌های مختلف را نشان می‌دهد، مقیاس ۱۵۰۰ دارای بیشترین اختلاف بین $\hat{K}$ و $\hat{K}^{sim}$ است لذا این مقیاس به‌عنوان مقیاس مناسب انتخاب گردید.



شکل ۷- تعیین محدوده‌های کوچک شده در ناحیه همبسته



شکل ۸- تعیین محدوده‌های کوچک شده در ناحیه ناهمبسته

جدول ۱- انتخاب مقیاس مناسب جهت ارزیابی

#	مقیاس (متر)	$\hat{K}^{sim}$	$\hat{K}$	$\hat{K}-\hat{K}^{sim}$
۱	۱۰۰۰	۱۹۸۶/۳۵	۱۹۸۹/۲۴	+۲/۸۹
۲	۱۲۵۰	۲۰۱۰/۱۸	۲۰۱۴/۰۹	+۳/۹۱
۳	۱۵۰۰	۲۰۱۴/۶۸	۲۰۱۸/۶۹	+۴/۰۱
۴	۱۷۵۰	۲۰۲۵/۵۶	۲۰۲۸/۳۷	+۲/۸۱

الگوی همبسته نقاط بر اساس پارامتر تعداد تکرار ۱۰۰، سطح اطمینان ۵٪ و مقیاس خوشه‌بندی ۱۵۰۰ تشکیل شد.

به‌منظور تعیین نواحی با اندازه کوچک و قابل کنترل توسط هر کارشناس تصادفات، لازم است که منطقه همبسته با اندازه خوشه‌های مختلف ناحیه‌بندی گردد و بهترین اندازه خوشه انتخاب شود. بهترین اندازه خوشه، اندازه‌ای است که اولا حداکثر استفاده از منبع (کارشناس تصادفات) به عمل آید و ثانیا زمان رسیدن به محل تصادف از مقدار مشخصی تجاوز ننماید. از آن‌جا که محل ثابتی برای استقرار کارشناسان تصادف در نظر گرفته نشده است، لذا زمان طی مسیر تا محل تصادف به صورت کسری از زمان کلی عبور از معابر واقع در هر خوشه در نظر گرفته می‌شود. اگر چه مقدار این کسر به‌طور دقیق مشخص نشده‌است ولی به دلیل آن‌که بر روی همه نواحی به طور یکسان عمل می‌کند، می‌توان آن را جهت ارزیابی استفاده نمود. پس از این مرحله بیشینه توان اسمی هر کارشناس تصادف بر اساس تعداد تصادف در هر ناحیه مشخص می‌گردد. این کار

با استفاده از فرمول ۵ انجام می‌گیرد.

جداول ۲ و ۳ مقادیر به دست آمده برای اندازه خوشه‌های مختلف را در مناطق همبسته و ناهمبسته نشان می‌دهد.

اندازه خوشه مناسب در این جداول، اندازه‌ای است که تا حد ممکن به توان اسمی منبع نزدیک باشد و زمان رسیدن به محل تصادف از t_s تجاوز ننماید. این زمان می‌تواند بر اساس استانداردهای هر کشور و یا شهر مختلف باشد اما در این‌جا ۱۵ دقیقه در نظر گرفته شده است. مقادیر انتخابی از هر جدول با رنگ متمایز مشخص شده‌اند.

جدول ۲- انتخاب اندازه خوشه مناسب جهت منطقه همبسته

#	اندازه خوشه	طول متوسط ± انحراف معیار (کیلومتر)	زمان متوسط ± انحراف معیار (دقیقه)	بیشینه توان اسمی (#)	تعداد خوشه
۱	۳۰	۳۰/۸۱ ± ۶/۴۸	۴/۸۲ ± ۰/۹۶	۱۰۳/۴	۳۹
۲	۳۵	۳۳/۳۸ ± ۶/۹۳	۵/۱۵ ± ۱/۰۵	۱۰۲/۴۲	۳۶
۳	۴۰	۳۶/۴۲ ± ۷/۵۹	۵/۴۶ ± ۱/۱۴	۱۰۱/۵	۳۳
۴	۴۵	۳۸/۷۶ ± ۸/۱۹	۶/۳ ± ۱/۳۸	۹۹/۲	۳۱
۵	۵۰	۵۲/۲۵ ± ۱۱/۴۳	۷/۸ ± ۱/۷۱	۹۵/۲	۲۳
۶	۵۵	۵۷/۲۲ ± ۱۲/۱۸	۸/۸۷ ± ۱/۸۳	۹۲/۶	۲۱
۷	۶۰	۶۶/۷۶ ± ۱۴/۰۴	۱۰ ± ۲/۱	۹۰	۱۸
۸	۶۵	۸۵/۸۳ ± ۱۸/۲۷	۱۲/۸ ± ۲/۷۳	۸۴/۱	۱۴
۹	۷۰	۱۰۰/۱۴ ± ۲۲/۵۹	۱۵/۰۲ ± ۳/۳۹	۸۰	۱۲
۱۰	۷۵	۱۲۰/۱۷ ± ۲۵/۵۶	۱۸/۵۴ ± ۳/۸۴	۷۴/۲	۱۰
۱۱	۸۰	۱۵۰/۲۱ ± ۳۱/۳۵	۲۳/۲۷ ± ۴/۷۱	۶۷/۶	۸

جدول ۳- انتخاب اندازه خوشه مناسب جهت منطقه ناهمبسته

#	اندازه خوشه	طول متوسط ± انحراف معیار (کیلومتر)	زمان متوسط ± انحراف معیار (دقیقه)	بیشینه توان اسمی (#)	تعداد خوشه
۱	۲۰	۶۶/۳۶ ± ۱۱/۰۴	۱۰/۲۳ ± ۱/۶۵	۸۹/۴۸	۱۱
۲	۲۵	۸۱/۱۱ ± ۱۳/۵۳	۱۲/۴۳ ± ۱/۹۲	۸۴/۸۴	۹
۳	۳۰	۱۰۴/۲۸ ± ۱۶/۸	۱۳/۵۲ ± ۲/۵۸	۸۲/۷۲	۷
۴	۳۵	۱۲۱/۶۶ ± ۱۸/۹	۱۶/۱۸ ± ۲/۷۶	۷۷/۹۵	۶
۵	۴۰	۱۲۱/۶۶ ± ۲۰/۲۵	۱۶/۱۸ ± ۳/۰۳	۷۷/۹۵	۶
۶	۴۵	۱۴۶ ± ۲۴/۰۶	۲۳/۴۲ ± ۳/۶۹	۶۷/۳۹	۵
۷	۵۰	۱۴۶ ± ۲۴/۰۶	۲۳/۴۲ ± ۳/۶۹	۶۷/۳۹	۵
۸	۵۵	۱۸۲/۵ ± ۳۰/۳	۳۰/۴۶ ± ۴/۵۴	۵۹/۵۴	۴
۹	۶۰	۱۸۲/۵ ± ۳۰/۳	۳۰/۴۶ ± ۴/۵۴	۵۹/۵۴	۴
۱۰	۶۵	۱۸۲/۵ ± ۳۰/۳	۳۰/۴۶ ± ۴/۵۴	۵۹/۵۴	۴
۱۱	۷۰	۲۴۳/۳۳ ± ۴۰/۵۳	۴۳/۵۱ ± ۶/۲۷	۴۸/۹۷	۳

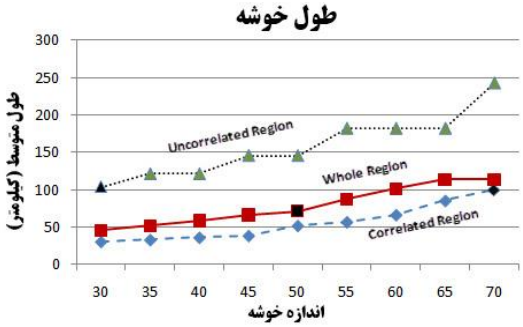
جهت ارزیابی روش پیشنهادی لازم است که نتایج آن را با یک روش جایگزین مقایسه نمود. از آن‌جا که کا-فانکشن محلی نمی‌تواند به تنهایی کل منطقه را ناحیه‌بندی نماید لذا روش سلسله مراتبی به تنهایی انتخاب گردید تا با روش پیشنهادی مورد مقایسه قرار گیرد. در حقیقت در روش جایگزین تفکیک منطقه به دو بخش همبسته و ناهمبسته صورت نمی‌گیرد. جدول ۴ نتایج این روش را برای کل منطقه نشان می‌دهد. اندازه خوشه انتخابی در این روش نیز با رنگ متمایز مشخص شده است.

جدول ۴- جدول انتخاب اندازه خوشه مناسب جهت کل منطقه

#	اندازه خوشه	طول متوسط ± انحراف معیار (کیلومتر)	زمان متوسط ± انحراف معیار (دقیقه)	بیشینه توان اسمی (#)	تعداد خوشه
۱	۲۵	۴۱/۰۸ ± ۱۶/۵۶	۶/۳۶ ± ۲/۴۹	۹۹	۴۷
۲	۳۰	۴۵/۹۷ ± ۱۸/۶	۶/۹۲ ± ۲/۷۹	۹۷/۵۱	۴۲
۳	۳۵	۵۲/۱۸ ± ۲۰/۸۵	۸/۰۶ ± ۳/۱۵	۹۴/۵۸	۳۷
۴	۴۰	۵۸/۵۱ ± ۲۳/۶۱	۹/۲ ± ۳/۵۴	۹۱/۸۳	۳۳
۵	۴۵	۶۶/۵۸ ± ۲۶/۸۵	۱۱/۰۱ ± ۴/۰۲	۸۷/۷۸	۲۹
۶	۵۰	۷۱/۵۱ ± ۲۸/۵۳	۱۳/۲۵ ± ۴/۳۲	۸۳/۲۳	۲۷
۷	۵۵	۸۷/۷۷ ± ۳۵/۵۸	۱۵/۹۸ ± ۵/۳۱	۷۸/۲۹	۲۲
۸	۶۰	۱۰۱/۶۳ ± ۴۰/۸۶	۱۹/۹۶ ± ۶/۱۵	۷۲/۰۵	۱۹
۹	۶۵	۱۱۳/۵۸ ± ۴۵/۸۱	۲۲/۵ ± ۶/۸۷	۶۸/۵۷	۱۷
۱۰	۷۰	۱۱۳/۵۸ ± ۴۵/۸۱	۲۲/۵ ± ۶/۸۷	۶۸/۵۷	۱۷
۱۱	۷۵	۱۲۰/۶۸ ± ۴۸/۹	۲۸/۳۱ ± ۷/۲۹	۶۱/۷۳	۱۶

از مقادیر به دست آمده چنین نتیجه می‌شود که برای تاریخ مذکور به جای استقرار ۲۷ کارشناس تصادف می‌توان از ۱۹ کارشناس استفاده کرد. لذا به میزان تقریبی ۳۰٪ در استفاده از کارشناسان تصادف صرفه‌جویی به عمل آمده است.

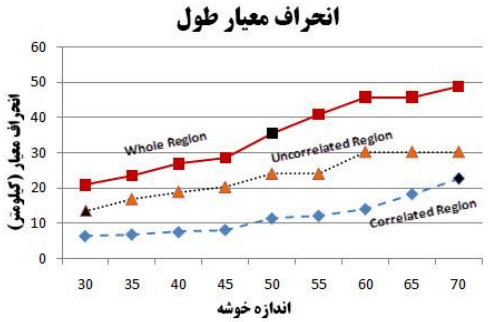
در نمودار شکل ۹ روند تغییرات طول متوسط معابر در نواحی، بر اساس اندازه‌های متفاوت خوشه دیده می‌شود. بر اساس این نمودار، منطقه همبسته دارای کوچک‌ترین طول و ناحیه ناهمبسته دارای بیشترین طول خوشه‌ای هستند و این امر نشان از صحیح بودن اندیشه تفکیک منطقه به دو قسمت همبسته و ناهمبسته دارد.



شکل ۹- نمودار طول متوسط معابر در خوشه‌های حاصله با تغییر اندازه خوشه

نمودار شکل ۱۰ مقادیر انحراف معیار طول متوسط معابر در هر ناحیه را با تغییر اندازه خوشه به تصویر می‌کشد. این نمودار نشان می‌دهد که اجرای روش پیشنهادی می‌تواند تاثیر به‌سزایی در ایجاد

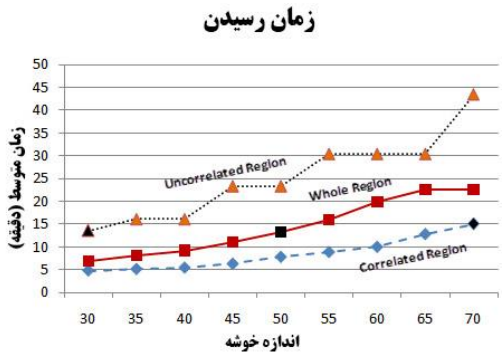
تعادل بین خوشه‌ها از دیدگاه طول معابر داشته باشد.



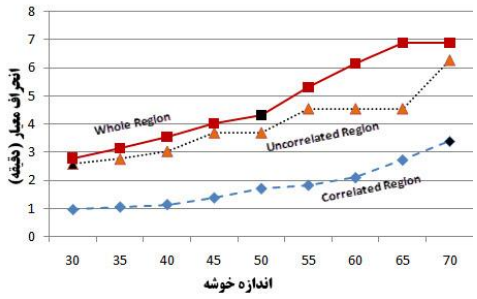
شکل ۱۰-نمودار انحراف معیار طول معابر خوشه‌ها بر اساس اندازه خوشه

در نمودار شکل ۱۱، روند تغییراتی زمان رسیدن کارشناس به محل تصادف بر اساس تغییرات اندازه خوشه آورده شده است. همان‌طور در این نمودار مشخص است، اندازه خوشه در زیر خط ۱۵ دقیقه انتخاب شده است. در این نمودار مقایسه منطقه همبسته و ناهمبسته با یکدیگر ، فرضیه‌ی توان رسیدگی بالاتر کارشناسان در ناحیه همبسته را نسبت به ناحیه ناهمبسته تایید می‌کند.

مقادیر انحراف معیار زمانی خوشه‌ها در کل منطقه که با روش سلسله‌مراتبی به دست آمده (بر اساس شکل ۱۲) از دو منطقه همبسته و ناهمبسته بیشتر است. بنابراین می‌توان چنین نتیجه گرفت که برقراری تعادل بین خوشه‌ها در روش پیشنهادی از دیدگاه زمان رسیدن کارشناس به محل حادثه بهبود یافته است.



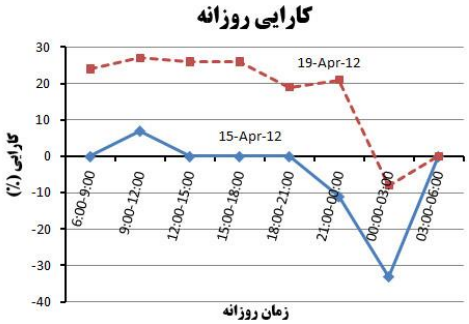
شکل ۱۱- نمودار زمان رسیدن کارشناس به محل حادثه با تغییر اندازه خوشه



شکل ۱۲- نمودار انحراف معیار زمان رسیدن به محل حادثه بر اساس اندازه خوشه

میزان کارایی روش پیشنهادی در استفاده از منابع در طول شبانه‌روز به ازای بازه‌های سه ساعته برای تاریخ ۱۹ آوریل (روز کاری) و تاریخ ۱۵ آوریل (روز تعطیل) مورد ارزیابی قرار گرفت که نتایج آن

در نمودار شکل ۱۳ به تصویر کشیده شده است. بر اساس نتایج حاصله، استفاده از روش پیشنهادی در کلیه ساعات یک روز کاری به غیر از نیمه‌شب می‌تواند مفید و موثر واقع گردد. بر اساس این جدول، میانگین صرفه‌جویی روش پیشنهادی در ساعات غیر از نیمه‌شب به ۲۳/۸٪ می‌رسد. اما بر اساس این نمودار، کارایی قابل توجهی در روز تعطیل مشاهده نمی‌شود و حتی در برخی ساعات اثر معکوس دارد. تاثیر معکوس روش پیشنهادی بدین معنا است که تفکیک منطقه مورد مطالعه به دو قسمت همبسته و ناهمبسته، باعث ایجاد جزیره‌هایی در آن می‌گردد که نه تنها اثر مثبتی ندارد بلکه از هم گسیختگی منطقه ناهمبسته را به دنبال خواهد داشت.



شکل ۱۳- میزان کارایی روش پیشنهادی برای روز کاری و تعطیل

۷- جمع‌بندی و نتیجه‌گیری

در روش‌های مرسوم، به منظور تخصیص منابع شهری به تقاضاها، لازم است که موقعیت ثابتی برای منابع در نظر گرفت. مواجه شدن با منابع متحرک بر روی شبکه معابر شهری همچون کارشناسان تصادفات باعث کاهش اهمیت بخش جایابی منبع در یک مسئله تخصیص مکانی می‌گردد. از سوی دیگر شباهت نزدیک بین مفهوم خوشه‌بندی و تخصیص مکانی باعث گردید که در این تحقیق از خوشه‌بندی به‌عنوان ابزاری جهت رسیدن به مرحله تخصیص و بدون گذر از مرحله جایابی استفاده گردد. در فاز اول روش پیشنهادی، رویدادهای وابسته به زمان مورد نظر جهت برنامه‌ریزی، استخراج گردید و سپس در فاز دوم بر اساس این رویدادها، منطقه شهری با استفاده از روش خوشه‌بندی کا-فانکشن محلی به دو بخش همبسته و ناهمبسته تقسیم شد. وسعت زیاد مناطق همبسته و ناهمبسته باعث شد که در ادامه با بهره‌گیری از روش خوشه‌بندی سلسله‌مراتبی، به کوچک‌سازی این مناطق اقدام گردد. فرآیند کوچک‌سازی با در نظر گرفتن توان هر کدام از کارشناسان تصادفات و البته زمان استاندارد رسیدن آن‌ها به محل حادثه انجام گرفت. در نهایت خوشه‌های تشکیل شده با استفاده از دیاگرام ورونوی به شبکه معابر تعمیم داده شد و محدوده تحت کنترل هر کدام از کارشناسان تصادف مشخص گردید.

ارزیابی‌های انجام گرفته بر روی روش پیشنهادی، صرفه‌جویی قابل ملاحظه‌ای را در به‌کارگیری تعداد نیروی کارشناس تصادفات در روزهای کاری هفته و در ساعات غیر از نیمه شب نشان می‌دهد. این در حالی است که روش پیشنهادی در روزهای تعطیل و ساعات نیمه شب کارایی مورد انتظار و قابل توجهی را از خود بروز نداده است.

۸- تحقیقات آتی

یکی از ویژگی‌های روش پیشنهادی، انتخاب رویدادهای وابسته در بازه‌های کریسپ[1] است. این در حالی است که در نظر گرفتن وزن برای داده‌های تاریخچه‌ای تصادفات باعث نزدیک‌تر شدن خروجی مورد نظر

[1] Crisp

به واقعیت می‌گردد. در نظر گرفتن وزن بدین معناست که با فاصله گرفتن زمانی هر تصادف از زمان مورد نظر، وزن آن کاهش یابد. شیوه انجام این کار می‌تواند یکی از موضوعات تحقیقاتی در آینده باشد.

از دیگر نکاتی که باعث واقعی‌تر شدن خروجی مورد نظر می‌گردد، در نظر گرفتن شرایطی است که مرز نواحی استخراجی به خیابان‌های اصلی شهر منتهی گردد. این کار می‌تواند از طریق به کارگیری ورونوی توانی بر روی شبکه راه و یا استخراج فازی مرزها، انجام پذیرد. پرداختن به این موضوع، به تحقیقات آتی محول شده است.

ایجاد ارتباط محیطی بین زمان مورد نظر برای برنامه‌ریزی و داده‌های تاریخچه‌ای می‌تواند اندیشه‌ای برای علاقه‌مندان جهت ادامه مسیر باشد. به عنوان مثال اگر برنامه‌ریزی برای یک روز بارانی مد نظر باشد، باید تصادفاتی که در شرایط بارانی حادث شده‌اند، دارای وزن بیشتری نسبت به تصادفات حادث شده در شرایط آب و هوایی صاف باشند.

مراجع

[1] S. Fotheringham and M. Wegener, Spatial Models and GIS: New and Potential Models vol. 7: CRC, 1999.

[2] F. Karimi, M. R. Delavar, and M. A. Mostafavi, "Space allocation of educational centers using multiplicatively weighted voronoi diagram," in ISPRS Workshop on Quality, Scale and Analysis Aspects of City Models, Lund,Sweden, 2009, p. 4.

[3] M. Musio, E. A. Sauleau, and N. Augustin, "Resources allocation in healthcare for cancer: a case study using generalised additive mixed models," Geospatial Health, vol. 7, pp. 83-89, 2012.

[4] L. C. Galvão, A. G. N. Novaes, J. E. Souza de Cursi, and J. C. Souza, "A multiplicatively-weighted Voronoi diagram approach to logistics districting," Computers & Operations Research, vol. 33, pp. 93-114, 2006.

[5] F. Ricca and B. Simeone, "Local search algorithms for political districting," European Journal of Operational Research, vol. 189, pp. 1409-1426, 2008.

[6] S. J. D'Amico, S.-J. Wang, R. Batta, and C. M. Rump, "A simulated annealing approach to police district design," Comput. Oper. Res., vol. 29, pp. 667-684, 2002.

[7] M. Fogue, P. Garrido, F. J. Martinez, J.-C. Cano, C. T. Calafate, and P. Manzoni, "A novel approach for traffic accidents sanitary resource allocation based on multi-objective genetic algorithms," Expert Systems with Applications, vol. 40, pp. 323-336, 2013.

[8] G. F. Voronoi, "Nouvelles applications des parametres continus a la theorie des formes quadratiques," J. Reine u. Angew. Math., vol. 34, pp. 198-287, 1908.

[9] B. N. Boots and A. Getis, Point pattern analysis, illustrated ed.: Sage Publications, 1988.

[10] A. Okabe and M. Kitamura, "A Computational Method for Market Area Analysis on a Network," Geographical Analysis, vol. 28, pp. 330-349, 1996.

[11] A. Okabe and K.-i. Okunuki, "A Computational Method for Estimating the Demand of Retail Stores on a Street Network and its Implementation in GIS," Transactions in GIS, vol. 5, pp. 209-220, 2001.

[12] Y. Levin and A. Ben-Israel, "A heuristic method for large-scale multi-facility location problems," Computers & Operations Research, vol. 31, pp. 257-272, 2004.

[13] I. Lee and J. Yang, "Voronoi-based topological information for combining partitioning and hierarchical clustering," in Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on, 2005, pp. 484-489.

[14] K. Liao and D. Guo, "A Clustering-Based Approach to the Capacitated Facility Location Problem1," Transactions in GIS, vol. 12, pp. 323-339, 2008.

[15] L. Liu, S. Fong, and A. Ip, "Grid-based

hierarchical clustering for spatial resource allocation," in IADIS International Conference e-Society 2011, 2011, pp. 513-517.

[16] J. Besag and J. Newell, "The Detection of Clusters in Rare Diseases," Journal of the Royal Statistical Society. Series A (Statistics in Society), vol. 154, pp. 143-155, 1991.

[17] A. B. Lawson, Statistical methods in spatial epidemiology: Wiley, 2006.

[18] B. D. Ripley, "The Second-Order Analysis of Stationary Point Processes," Journal of Applied Probability, vol. 13, pp. 255-266, 1976.

[19] R. L. O'Driscoll, "Description of spatial pattern in seabird distributions along line transects using neighbour K statistics," Marine Ecology Progress Series, vol. 165, pp. 81-94, May 07, 1998 1998.

[20] A. Okabe and I. Yamada, "The K-Function Method on a Network and Its Computational Implementation," Geographical Analysis, vol. 33, pp. 271-290, 2001.

[21] I. Yamada and J.-C. Thill, "Local Indicators of Network-Constrained Clusters in Spatial Point Patterns," Geographical Analysis, vol. 39, pp. 268-292, 2007.

[22] K. Sugihara, A. Okabe, and T. Satoh, "Computational method for the point cluster analysis on networks," GeoInformatica, vol. 15, pp. 167-189, 2011.

[23] M. L. Yiu and N. Mamoulis, "Clustering objects on a spatial network," presented at the Proceedings of the 2004 ACM SIGMOD international conference on Management of data, Paris, France, 2004.

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

Journal of Geospatial Information Technology
Vol.3, No.3, autumn 2015

A New Approach for Allocating Dynamic Resources on Road Network by Applying Spatiotemporal Clustering of Events

Mohsen Goodarzi [1], Farshad Hakimpour [*2], Parham Pahlavani[3], Seyed Mahmood Hajimirrahimi [4]

1- GIS MSc, School of Surveying and Geospatial Engineering, University of Tehran, Iran
2- Assistant professor, School of Surveying and Geospatial Engineering, University of Tehran, Iran
3- Assistant professor, School of Surveying and Geospatial Engineering, University of Tehran, Iran
4- GIS&RS MSc, Earth Science Department, Tabriz University, Iran

Abstract

A location-Allocation problem consists of two parts: locating resources and allocating them due to demands. Urban managers in order to plan metropolises, utilize various methods like districting urban regions, which is the result of solving a LA problem. Changing in urban districts, originates from dynamism rate of demands. Some efforts have been conducted to relate clustering to LA. Most of these researches apply clustering as a tool in locating resources whereas facing with dynamic demands is diminishing the importance of locating rather than allocating part. In this article the proposed method suggested dynamic location for resources in order to allocating them to the plea. The advantage of proposed method, which is a combination of statistical local K-function and non-statistical hierarchical clustering methods, lies in division of urban region into correlated and non-correlated parts. In order to deploy local K-function, Monte Carlo simulation has been applied. The Correlated and non-correlated parts are being constituted by local k-function in wide-extend; in continue they should shrink using hierarchical method. The small clusters should meet two conditions: 1) consuming accident experts' capacity as much as possible 2) not overrunning duration of reaching accident experts to collision places from a standard quantity. Evaluations show a considerable efficiency in weekdays except midnight hours. However, deploying just hierarchical method on weekends and midnight hours leads to better efficiency and also reduction in number of accident experts.

Key words: Clustering, Location-Allocation, Accidents, K-function, Hierarchical

Correspondence Address: School of Surveying and Geomatics Engineering,College of Engineering, University of Tehran,North Kargar, Tehran 1439957131.Tel: +98 21 61114520.
Email: fhakimpour@ut.ac.ir