نشریه علمی- پژوهشی علوم و فنون نقشه برداری، دوره چهارم، شماره ۲، آبان ماه ۱۳۹۳

تحلیل عوامل مکانی مؤثر بر تمرکز تصادفات در راه‌های برون‌شهری با استفاده از GIS و داده‌کاوی

میثم عفتی*[۱]، محمدعلی رجبی[۲]، فرشاد حکیم‌پور[۲]، شاهین شعبانی[۳]

[۱]استادیار گروه مهندسی عمران- دانشکده فنی- دانشگاه گیلان
meysameffati@guilan.ac.ir

[۲]استادیار گروه مهندسی نقشه‌برداری- پردیس دانشکده‌های فنی- دانشگاه تهران
{marajabi, fhakimpour}@ut.ac.ir

[۳]استادیار گروه مهندسی عمران راه و ترابری- دانشگاه پیام نور تهران
shabani@iust.ac.ir

(تاریخ دریافت اسفند۱۳۹۱، تاریخ تصویب مرداد ۱۳۹۳)

Archive of SID

چکیده

یکی از مشکلات اصلی حمل‌ونقل جاده‌ای در کشور، تصادفات است. جهت مدیریت و کاهش تصادفات در یک محور برون‌شهری لازم است که متولیان ایمنی راه اطلاعات کاملی از عوامل تأثیرگذار بر تصادفات آن محور در اختیار داشته باشند. با توجه به تحقیقات صورت گرفته، شناسایی انواع تصادفات یک محور و عوامل مؤثر بر وقوع آن‌ها یکی از راهکارهای مؤثر در کاهش تصادفات جاده‌ای برون‌شهری است. این مطالعه یک استراتژی تلفیقی را جهت بررسی انواع تصادفات و عوامل مکانی مؤثر بر تمرکز تصادفات در راه‌های دوخطه دوطرفه برون‌شهری با استفاده از تحلیل‌های مکانی و روش خوشه‌بندی تجمعی سلسه‌مراتبی بر پایه K-Mean، ارائه می‌نماید. روش پیشنهادی با ایجاد یک پایگاه داده مکانمند متشکل از داده‌های تصادفات و اطلاعات راه و محیط مجاور آن در محور قزوین- رشت (ایران) مورد بررسی و ارزیابی قرار می‌گیرد. نتایج نشان می‌دهد که دقت خوشه‌بندی داده‌های تصادفات به روش پیشنهادی، ۶/۷ درصد بیش از روش تجمعی سلسه‌مراتبی و حدوداً ۱۰ درصد بیشتر از روش K-Mean است. تحلیل خروجی روش پیشنهادی با استفاده از توابع مکانی و الگوریتم‌های استخراج الگوی داده‌کاوی ضمن بیان نوع و خصوصیات تصادفات محور در نواحی تمرکز تصادفات، تأثیرات طرح هندسی راه، ترافیک، توسعه شهری، کاربری‌ها و عوارض اطراف راه را بر تصادفات آشکار می‌سازد.

واژگان کلیدی : ایمنی راه، تصادف، سیستم‌های اطلاعات مکانی (GIS)، خوشه‌بندی تجمعی سلسه‌مراتبی بر پایه K-Mean، الگوریتم Separate-and-Conquer

* نویسنده رابط

www.SID.ir

۱- مقدمه

امروزه تصادفات رانندگی و تلفات و هزینه‌های بالای اقتصادی و اجتماعی ناشی از آن از اساسی‌ترین مشکلاتی است که متخصصان و متولیان امور حمل‌ونقل و ترافیک را به چالش کشانده است. براساس گزارش سازمان بهداشت جهانی با کشته شدن بیش از ۱/۳ میلیون و جراحت ۲۰ تا ۵۰ میلیون نفر در هر سال، تصادفات جاده‌ای به‌عنوان یکی از علل اصلی مرگ‌ومیر در جهان شناخته می‌شود [۳۲]. برآوردهای آماری نشان می‌دهد که در ایران هزینه‌های اقتصادی مرگ‌ومیر و صدمات جانی ناشی از تصادفات ۴-۱/۸ درصد تولید ناخالص داخلی است که این رقم در مقایسه با کشورهای درحال‌توسعه که در آن‌ها هزینه تصادفات حدود ۱/۵ درصد از تولید ناخالص داخلی را تشکیل می‌دهد، بالا است [۳۴]. کاهش تصادفات یکی از اهداف اصلی متولیان حمل‌ونقلی در سراسر جهان محسوب می‌گردد. با توجه به رشد فزاینده تصادفات در راه‌های اصلی دوخطه دوطرفه در این تحقیق به مطالعه‌ی عوامل مکانی مؤثر بر تمرکز تصادفات در این نوع از راه‌های برون‌شهری پرداخته می‌شود که قابل تعمیم به جاده‌های فرعی نیز می‌باشد.

بنا به تحقیقات مجمع جهانی راه[۱]، سهم تجمعی عامل انسان در وقوع تصادف ۹۳ درصد، عامل وسیله‌نقلیه ۱۳ درصد و عامل محیط راه ۳۴ درصد می‌باشد [۲۱]. لذا، تحقیقات صورت گرفته در راستای ارتقای ایمنی راه‌ها و کاهش تصادفات، اکثراً داده‌های تصادفات را از جنبه عوامل انسانی مورد بررسی و تحلیل قرار داده‌اند. Siddiqui و همکاران [۲۴] در تحقیقات خود به این نتیجه رسیدند که یکی از کاستی‌های مطالعات صورت گرفته در تحلیل تصادفات جاده‌ای عدم توجه به تأثیرات مکانی مجاور راه است. آن‌ها معتقدند که این تأثیرات با استفاده از روش‌های غیرپارامتریکی همچون داده‌کاوی[۲] قابل مدل‌سازی هستند.

رویدادهای تصادف در دامنه مکان و زمان بندرت به‌صورت اتفاقی رخ می‌دهند، بلکه ممکن است خوشه‌هایی را تشکیل دهند که همان محل‌های تمرکز تصادفات در فضای جغرافیایی می‌باشند. لذا، فرض اصلی تحقیق این است که تصادفات به‌صورت اتفاقی در طول محور رخ نداده باشند. عوامل ترافیکی، طرح هندسی، توسعه شهری، کاربری‌ها و عوارض مجاور راه می‌توانند نقش مؤثری در تمرکز تصادفات در راه‌های برون‌شهری داشته باشند. لذا، بررسی تحلیلی این تأثیرات و کشف فاکتورهای مکانی مؤثر بر تمرکز تصادفات، می‌تواند سهم قابل‌توجهی در کاهش تصادفات و بهبود ایمنی راه‌های برون‌شهری داشته باشد [۳۳]. بدین منظور تلفیق قابلیت‌های سیستم‌های اطلاعات مکانی با روش‌های استخراج دانش، مفید به نظر می‌رسد. داده‌کاوی با کشف الگوهای پنهان موجود در داده‌های تصادفات می‌تواند نقش مؤثری در کاهش تلفات در راه‌های برون‌شهری داشته باشد [۱]. هدف این تحقیق ارائه روشی مکانمند مبتنی بر الگوریتم‌های خوشه‌بندی[۳] و استخراج الگوی[۴] داده‌کاوی جهت شناسایی نوع تصادفات و فاکتورهای مکانی مؤثر بر تمرکز تصادفات در راه‌های دوخطه دوطرفه برون‌شهری می‌باشد. بدین منظور یک روش خوشه‌بندی تجمعی سلسه‌مراتبی بر پایه K-Mean[۵] ارائه شد که بر پایگاه داده تصادفات محور مورد مطالعه اعمال گردید. سپس با مدل نمودن تأثیر مکانی طرح هندسی راه، تسهیلات و کاربری‌های مجاور، توسعه شهری، و ترافیک، فاکتورهای مکانی مؤثر بر تمرکز تصادفات در هر یک از خوشه‌های تصادف شناسایی شدند. تعیین نوع تصادفات و فاکتورهای مکانی مؤثر بر تمرکز تصادفات در یک بخش از راه می‌تواند نقش مؤثری در پیشگیری از وقوع تصادفات آتی آن بخش از راه داشته باشد.

بخش بعدی مقاله به بررسی تحقیقات پیشین در زمینه شناسایی فاکتورهای مؤثر بر وقوع تصادفات در راه‌های برون‌شهری می‌پردازد. در بخش سوم روش‌شناسی تحقیق ارائه می‌گردد. بخش چهارم پیاده‌سازی روش پیشنهادی را در محور مورد مطالعه مورد بحث قرار می‌دهد. در بخش پنجم ارزیابی روش پیشنهادی انجام می‌گیرد و نتایج پژوهش بحث می‌شوند. نهایتاً در بخش پایانی مقاله راهکارهایی جهت ادامه تحقیق ارائه می‌گردد.

۲- پیشینه تحقیق

در حوزه بررسی فاکتورهای طرح هندسی راه، وسیله نقلیه، و محیط مؤثر بر تصادفات، مطالعات زیادی با

---

۱ Permanent International Association of Road Congresses (PIARC)
۲ Data Mining
۳ Clustering
۴ Discovery of Patterns and Rules
۵ Hierarchical Agglomerative Clustering (HAC)

www.SID.ir
Archive of SID

استفاده از روش‌های رگرسیون صورت گرفته است. مدل‌های رگرسیون خطی، پواسون و توزیع دوجمله‌ای منفی از جمله روش‌هایی هستند که به‌مراتب در مطالعات ایمنی راه و تحلیل تصادفات مورد استفاده قرار گرفته‌اند [۱۶، ۱۹، ۲۳، ۱۸، ۱۷]. مدل‌های رگرسیون جهت تحلیل تصادفات نیاز به فرضیات اولیه و تعریف روابط از پیش تعریف‌شده‌ی بین متغیرهای ورودی دارند. در صورت تغییر یا وجود ابهام در این روابط، نتایج و خروجی تحلیل‌های تصادفات نیز تحت تأثیر قرار خواهند گرفت.

به دلیل اهمیت استخراج دانش از داده‌های تصادفات، روش‌های مبتنی بر داده‌کاوی همچون خوشه‌بندی و طبقه‌بندی نیز این امکان را در اختیار متخصصین قرار داده‌اند که بتوان حجم زیادی از داده‌های مرتبط با تصادفات را تحلیل نمود. در تحقیقات پیشین تکنیک‌هایی از داده‌کاوی از جمله شبکه‌های عصبی[۱] [۵،۲۵] و درخت دسته‌بندی و رگرسیون[۲] [۲،۳] جهت دسته‌بندی، پیش‌بینی و همچنین تعیین فاکتورهای مؤثر بر تصادفات مورد استفاده قرار گرفته‌اند. از جمله می‌توان به تحقیق Montella و همکاران [۲۰] اشاره نمود که با استفاده از تحلیل داده‌کاوی درخت تصمیم بروش CART و همچنین استخراج الگو، ۲۵۴۵۷۵ رکورد از تصادفات موتورسیکلت‌ها را در کشور ایتالیا مورد بررسی قرار دادند. خروجی روش مورد استفاده در بیان خصوصیات و الگوهای تصادفات با سایر تکنیک‌های مورد استفاده در تحلیل تصادفات همچون مدل‌های احتمالاتی برآورد شدت تصادفات سازگار بود. در این مطالعه ازآنجاکه الگوهای زیادی از تصادفات در داده‌های مورد بررسی وجود داشت امکان بروز خطا طی فرآیند استخراج الگو و همچنین انتخاب تصادفی الگوها وجود داشت.

الگوریتم‌های داده‌کاوی به‌تنهایی جهت تحلیل روابط پیچیده موجود در داده‌های تصادفات کافی نیستند. لذا دسته‌ای دیگر از مطالعات از تحلیل‌های مکانی در تلفیق با روش‌های محاسباتی و داده‌کاوی به‌منظور بررسی تصادفات راه‌های برون‌شهری استفاده کردند [۶، ۷، ۹، ۱۱، ۱۲، ۱۴، ۱۵، ۲۶]. در این تحقیقات از GIS در تلفیق با روش‌های داده‌کاوی و یا روش‌های آماری به‌منظور نمایش موقعیت‌های تصادف بر روی نقشه‌های رقومی، برآورد شدت تصادفات و تشخیص نواحی حادثه‌خیز راه استفاده شده است. Ha و Thill [۱۲] با این استراتژی نقش مؤثر بعد مکانی را در تصادفات عابرین در محیط شهری مورد بررسی قرار دادند. Delmelle و همکاران [۴] نیز از تحلیل‌های مکانی و آماری جهت مطالعه فاکتورهای تأثیرگذار بر تصادفات دوچرخه و عابر در شهر بافلو نیویورک آمریکا استفاده نمودند. آن‌ها در مطالعه خود با استفاده از تحلیل‌های همسایگی مکانی تأثیر عوارض فیزیکی راه، تقاطع‌ها، جمعیت و نوع کاربری را بر ریسک تصادف مورد بررسی قرار دادند. نتیجه این مطالعه حکایت از تأثیر بسزای فاکتورهای مکانی چون عوارض فیزیکی راه و نوع کاربری‌های مجاور بر تصادفات دوچرخه و عابر در منطقه مورد مطالعه داشت.

Steenberghen و همکاران [۲۶] کارایی استفاده از تحلیل‌های مکانی و روش‌های الگویی نقطه‌ای را به‌منظور تعیین محل‌های تمرکز تصادفات مورد بررسی قرار دادند. یکی از مشابهت‌های این مطالعه با تحقیق حاضر در این است که در آن کاربری‌ها و تراکم شهری اطراف به‌عنوان یکی از پارامترهای مؤثر جهت تعیین محل‌های تمرکز تصادف مورد بررسی قرار گرفت. این مطالعه نشان داد که در راه‌های با الگوی ترافیکی پراکنده، کاربری‌های مجاور بشکل قابل‌توجهی محل‌های تمرکز تصادفات را تحت تأثیر قرار خواهد داد و با روش‌های خوشه‌بندی خطی و دوبعدی قابل مدل‌سازی می‌باشد. این درحالی بود که آن‌ها در مطالعات خود تأثیرات ترافیکی و طرح هندسی راه را بر تمرکز تصادفات مورد بررسی قرار ندادند.

Sun [۲۷] با استفاده از تحلیل‌های مکانی ابزاری را ارائه نمود که با تکیه بر روند زمانی و مکانی، نوع بزرگراه و خصوصیات هندسی راه، نوع تصادف و شرایط محیطی، تصادفات جاده‌ای را تحلیل می‌نمود. از دیدگاه ایشان بدون استفاده از تحلیل‌های مکانی و تنها تکیه بر روش‌های سنتی و یا داده‌کاوی، امکان انجام تحلیل‌های مختلفی که وابسته به مکان هستند، مثل بررسی تعداد وقوع یک نوع خاص از تصادفات در قطعه‌ای از راه، علل تمرکز تصادفات در یک بخش خاص از راه، و .... به‌سختی امکان‌پذیر خواهد بود.

بررسی‌های انجام شده همچنین نشان می‌دهد که معمولاً تصادفات خوشه‌هایی را در طول راه تشکیل می‌دهند که محل تمرکز آن‌هاست. در این بخش‌های راه تصادفات با شدت یا نرخ بیشتری روی می‌دهند و

---

۱ Artificial Neural Network
۲ Classification and Regression Tree (CART)

Archive of SID
www.SID.ir

فاکتورهای مختلفی بر محل‌های تمرکز تصادفات تأثیرگذار هستند. در مطالعات صورت گرفته در حوزه تصادفات جاده‌ای، تحقیقات کمی فاکتورهای مکانی مؤثر بر تمرکز و نوع تصادفات خودروها را در راه‌های برون‌شهری مورد بررسی قرار داده‌اند. محل تمرکز تصادفات نه‌تنها می‌تواند متأثر از عوامل انسانی و نقص خودرو باشد، بلکه طرح هندسی راه، کاربری‌ها و تسهیلات رفاهی اطراف، فراوانی و نوع تصادفات را تحت تأثیر قرار خواهند داد. کشف این تأثیرات می‌تواند در برنامه‌ریزی‌های آتی ایمنی راه توسط متولیان حمل‌ونقل جاده‌ای مورد استفاده قرار گیرد.

۳- روش پیشنهادی

رویکرد اصلی تحقیق بر تلفیق تحلیل‌های مکانی و داده‌کاوی جهت تحلیل نوع تصادفات و فاکتورهای مکانی مؤثر بر تمرکز تصادفات استوار است. شکل ۱ روش کلی تحقیق را نشان می‌دهد. روش پیشنهادی تحقیق با ارائه یک روش خوشه‌بندی تجمعی سلسه‌مراتبی بر پایه K-Mean و تحلیل نتایج حاصله با توابع مکانی و الگوریتم‌های داده‌کاوی استخراج الگو، یک استراتژی تلفیقی را جهت بررسی تأثیرات مکانی طرح هندسی راه، ترافیک، توسعه شهری و کاربری‌های مجاور راه بر تصادفات برون‌شهری ارائه می‌نماید.

داده‌کاوی بر دو نوع هدایت‌شده و غیرهدایت شده است. داده‌کاوی هدایت‌شده، دارای متغیر خروجی خاص و از پیش تعیین شده است و به دنبال کشف الگوهایی خاص از داده‌های ورودی است؛ درحالی‌که هدف داده‌کاوی غیر هدایت‌شده، یافتن الگوها یا تشابهات بین گروه‌هایی از اطلاعات، بدون داشتن متغیر خروجی خاص و یا مجموعه‌ای از دسته‌ها و الگوهای از پیش تعیین شده می‌باشد. خوشه‌بندی، به‌عنوان یکی از روش‌های اصلی داده‌کاوی، تکنیکی توصیفی جهت تحلیل داده‌ها می‌باشد که در دسته روش‌های کلاسه‌بندی نظارت‌نشده قرار می‌گیرد [۲۸]. خوشه‌بندی به عمل تقسیم مجموعه‌ای از ورودی‌های ناهمگن به تعدادی از زیرمجموعه‌ها یا خوشه‌های همگن گفته می‌شود. در خوشه‌بندی هیچ دسته‌ی از پیش تعیین شده‌ای وجود ندارد و داده‌ها صرفاً براساس تشابه گروه‌بندی می‌شوند و عناوین هر گروه نیز توسط کاربر تعیین می‌گردد.

روش‌های خوشه‌بندی به دو دسته کلی تقسیم می‌گردند: خوشه‌بندی سلسله مراتبی[۱] و خوشه‌بندی افرازی[۲]. روش‌های دیگر همچون خوشه‌بندی بر مبنای تراکم[۳] [۸] از دیگر روش‌هایی هستند که مابین دو روش فوق قرار می‌گیرند.

در روش خوشه‌بندی سلسله مراتبی به خوشه‌ها بر اساس میزان عمومیت آن‌ها، ساختاری سلسله مراتبی معمولاً به‌صورت درختی نسبت داده می‌شود. به این درخت سلسله مراتبی دندوگرام[۴] می‌گویند. روش کار تکنیک‌های خوشه‌بندی سلسله‌مراتبی معمولاً بر اساس الگوریتم‌های حریصانه[۵] و بهینگی مرحله‌ای[۶] است. این روش‌ها بر اساس ساختار سلسله مراتبی تولیدی توسط آن‌ها نیز معمولاً به دو دستهٔ تقسیم‌کننده[۷] و تجمعی[۸] تقسیم می‌شوند.

روش‌های خوشه‌بندی سلسله مراتبی نمی‌توانند خوشه‌های مجزایی از داده‌های تصادف با الگوهای یکسان ایجاد کنند. بعلاوه، در این روش‌ها هر چه اندازه داده‌ها در خوشه‌ها افزایش یابد، الگوهای هر خوشه وابستگی کمتری خواهند داشت [۲۸].

روش‌های خوشه‌بندی افرازی همچون K-Means روش‌های مؤثری هستند که دارای محدودیت‌های روش‌های سلسله مراتبی نبوده و در تحلیل داده‌های با حجم بالا کارا می‌باشند. K-Means یکی از روش‌های معمول و معتبر خوشه‌بندی است که بر اساس کمترین فاصله‌های داده از مرکز خوشه عمل خوشه‌بندی را انجام می‌دهد. در واقع این روش، از داده‌ها مجموعه‌هایی مجزا ایجاد می‌کند طوری که در هر مجموعه، داده‌ها به مرکز خوشه نزدیک باشند. در این الگوریتم پارامتر K نماینده تعداد خوشه‌های مدنظر است. لذا، در ابتدا K نقطه به‌عنوان مراکز اولیه خوشه‌ها در نظر گرفته می‌شوند، سپس طی یک فرآیند تکراری، خوشه‌ها با تخصیص داده‌ها به نزدیک‌ترین مراکز تشکیل می‌شوند. در ادامه مراکز هر خوشه دوباره محاسبه می‌شوند و این پروسه

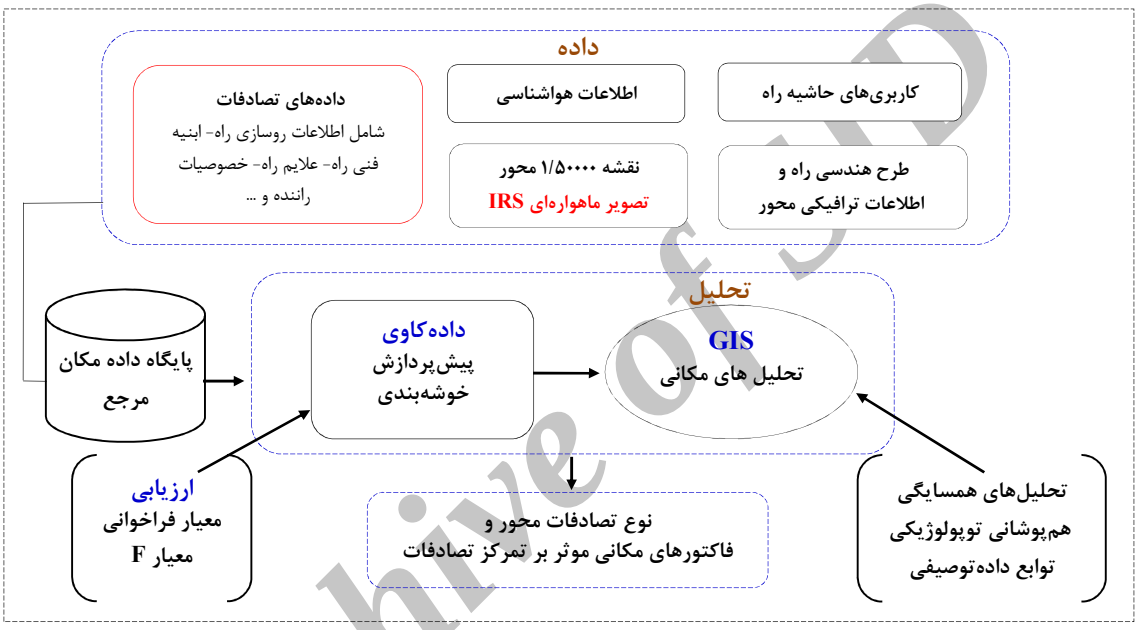
---

۱ Hierarchical
۲ Partitioning
۳ Density-based
۴ Dendogram
۵ Greedy Algorithms
۶ Stepwise-optimal
۷ Divisive
۸ Agglomerative

www.SID.ir
Archive of SID

تکرار می‌گردد تا مراکز خوشه‌ها تغییر نکنند. روش خوشه‌بندی K-Means دارای معایبی نیز می‌باشد. این روش علاوه بر این‌که مراکز اولیه خوشه‌ها را به‌طور تصادفی انتخاب می‌کند، نسبت به خطا در داده‌های ورودی بسیار حساس می‌باشد.

بدین منظور در این تحقیق یک روش تلفیقی از خوشه‌بندی تجمعی سلسله‌مراتبی و K-Means جهت دسته‌بندی داده‌های تصادف ارائه گردید که معایب مذکور را در روش‌های خوشه‌بندی سلسله‌مراتبی و همچنین روش K-Means پوشش خواهد داد. لذا تعیین محل‌های تمرکز تصادفات در دو مرحله صورت می‌پذیرد: ابتدا خوشه‌های اولیه‌ای از داده‌های تصادفات بروش K-Means تشکیل شده، سپس الگوریتم خوشه‌بندی تجمعی سلسله مراتبی بر روی این خوشه‌ها اعمال و نمودار دندروگرام جهت تعیین نواحی تمرکز تصادفات ایجاد می‌گردد. در این روش ابتدا داده‌ها به‌عنوان خوشه‌ای مجزا در نظر گرفته می‌شود و طی فرایندی تکراری در هر مرحله خوشه‌هایی که شباهت بیشتری با یکدیگر دارند ترکیب می‌شوند تا در نهایت یک خوشه و یا تعداد مشخصی خوشه حاصل گردد.



شکل۱- روش پیشنهادی تحقیق

۴- پیاده‌سازی

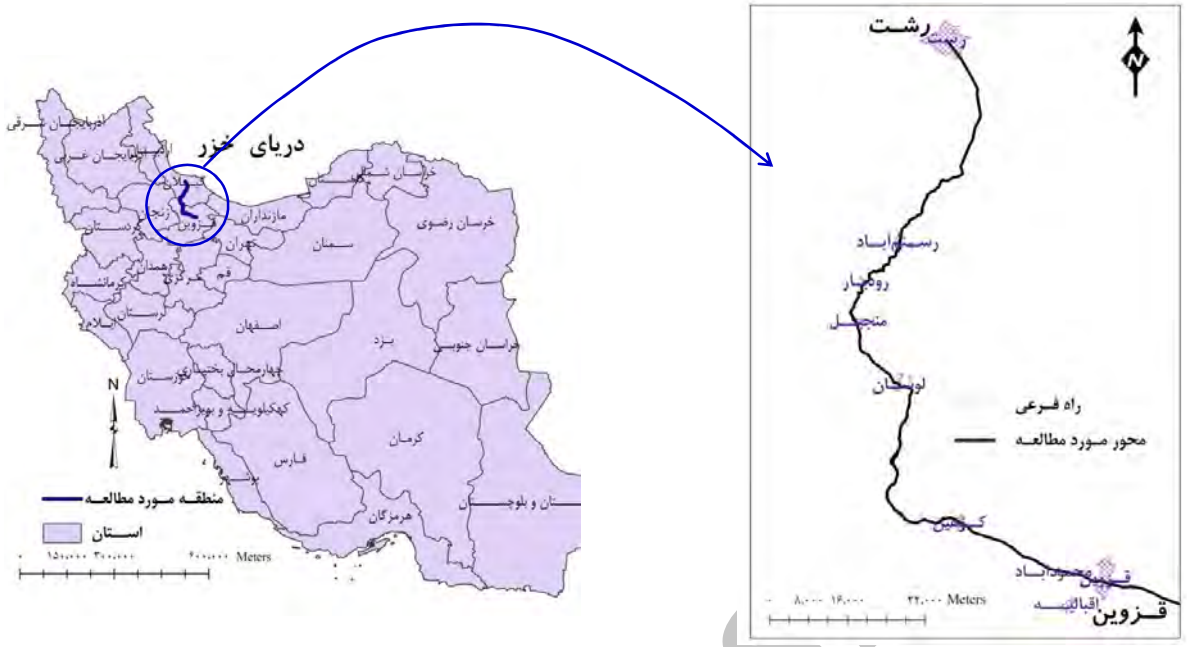
در این بخش ضمن توصیف داده‌های مورد استفاده و معرفی محور مورد مطالعه، نحوه پیاده‌سازی روش پیشنهادی تشریح می‌گردد.

۴-۱- داده و منطقه مورد مطالعه

منطقه مورد مطالعه محور قدیم قزوین-رشت می‌باشد که تهران را به شمال ایران متصل می‌کند. این منطقه کوهستانی بوده و محدوده ارتفاعی آن بین ۳۰۰ تا ۲۳۹۴ متر می‌باشد. محور مورد مطالعه از مختصات جغرافیایی N "۴۵ '۱۹ °۳۶ و E "۵۷ '۳۷ °۴۹ در شهر قزوین شروع شده و به مختصات جغرافیایی "N ۳۴ '۳۷ °۳۶ و E "۲۹ '۳۰ °۴۹ در شهر رشت منتهی می‌گردد. این مسیر از راه‌های اصلی پرخطر کشور بوده که بسیار مستعد تصادف می‌باشد. شکل ۲ منطقه مورد مطالعه را نمایش می‌دهد.

Archive of SID

www.SID.ir



شکل۲- محور مورد مطالعه

این پژوهش از چندین منبع داده در فرمت‌های مختلف استفاده نمود. این داده‌ها به لایه‌هایی در محیط GIS تبدیل شدند. همچنین پیش‌پردازش‌های لازم بر روی داده‌های تصادفات و لایه‌های مکانی قبل از فرآیند داده‌کاوی انجام پذیرفت تا از ایجاد خوشه‌ها و الگوهای نامناسب تصادفات جلوگیری گردد. این پیش‌پردازش‌ها شامل نرمال‌سازی متغیرهای ورودی، هم مقیاس نمودن، و حذف داده‌های تکراری و ناقص می‌باشند. علاوه بر این، برخی ویژگی‌های داده‌های تصادفات با هم ترکیب شدند تا حافظه و حجم پردازش‌های لازم به‌منظور تحلیل تصادفات کاهش یابد. جدول ۱ داده‌های مورد استفاده تحقیق و خصوصیات آن‌ها را توصیف می‌کند.

داده‌های تصادفات شامل بیش از ۶۰ ویژگی هستند که خصوصیات هر رویداد تصادف را بیان می‌کنند. ۷۱۲۱ رویداد تصادف در فرمت متن، عدد، تاریخ و زمان در روش پیشنهادی تحقیق مورد بررسی قرار گرفتند. متناسب با اهداف و روش تحقیق که تحلیل نوع و خصوصیات مکانی تصادفات برون‌شهری می‌باشد و نیز در دسترس بودن داده‌ها، برخی از ویژگی‌های تصادف انتخاب و در محیط GIS با داده‌های طرح هندسی راه و داده‌های محیطی تلفیق شدند.

جدول۱- توصیف داده‌های مورد استفاده

داده	منبع	مقیاس/قدرت تفکیک	توضیح
نقشه توپوگرافی	سازمان نقشه‌برداری کشور	۱/۵۰۰۰۰	رقومی
مدل رقومی ارتفاعی	سازمان نقشه‌برداری کشور	۱۰ متر	رقومی
طرح هندسی راه	وزارت راه و شهرسازی	-	توصیفی
اطلاعات هواشناسی	سازمان هواشناسی کشور	-	ایستگاه‌های هواشناسی
داده‌های تصادف	پلیس‌راه کشور	-	توصیفی
تصویر ماهواره‌ای	سازمان جغرافیای نیروهای مسلح	-	IRS-LISS III-P6

جدول ۲ متغیرهای مورد استفاده تحقیق و خصوصیات آن‌ها را توصیف می‌کند.

www.SID.ir

جدول۲- فاکتورهای مورد بررسی

متغیر ورودی	نوع	توضیح
مکان تصادف	پیوسته	فاصله از مبدأ
شدت تصادف	گسسته	خسارتی، جرحی، فوتی، جرحی و فوتی
سن راننده	پیوسته	عددی بین ۱۵ تا ۸۰
کاربری‌ها و مراکز جمعیتی	گسسته	مسکونی، تجاری، صنعتی، کشاورزی
میزان خسارت	گسسته	عددی بین ۰ تا ۱۱۰۰[1]
حجم ترافیک	گسسته	روان، نیمه‌سنگین، سنگین
هندسه راه	گسسته	تقاطع، پیچ، شیب و ..
ماه تصادف	پیوسته	عددی بین ۱ تا ۱۲
روز تصادف	پیوسته	عددی بین ۱ تا ۳۱
عرض جاده	پیوسته	عددی بین ۱ تا ۳۲
نوع تصادف	گسسته	رودررو، پهلوبه‌پهلو، ...

۴-۲-خوشه‌بندی تصادفات

در این مطالعه جهت استخراج دانش از داده‌های تصادفات، خوشه‌بندی تجمعی سلسله مراتبی با روش K-Mean تلفیق گردید و رویدادهای تصادفات بر اساس خصوصیات آن‌ها در دسته‌های مجزا قرار گرفتند. ساختار دندروگرام شکل ۳ نشان می‌دهد که چگونه رویدادهای تصادفاتی که طی فرآیند خوشه‌بندی K-Means به بیست خوشه تقسیم شده‌اند، مرحله به مرحله جهت نمایش در خوشه‌بندی تجمعی سلسله مراتبی گروه‌بندی شده‌اند. جدول ۳ تعیین تعداد خوشه‌های تصادفات را طی یک پروسه تکراری توسط معیارهای BSS[2] و Gap نمایش می‌دهد. نسبت BSS عدم‌تجانس بین خوشه‌ها را اندازه‌گیری می‌کند [۱۳] و Gap به‌عنوان فضای بین خوشه‌ها تعریف می‌گردد [۲۹]. در یک فرآیند خوشه‌بندی ایده‌آل، دسته‌هایی ایجاد می‌گردند که نسبت BSS و مقدار Gap بالایی دارند. جدول ۳ بیان می‌کند که با انتخاب دو، چهار یا هفت خوشه، مقادیر BSS و Gapمقادیر قابل قبولی خواهند داشت. با انتخاب چهار خوشه، ایده‌ال‌ترین خوشه‌ها با حداکثر جدایی بین داده‌های تصادفات طی فرایند خوشه‌بندی تلفیقی پیشنهادی ایجاد گردید.

۱ میزان خسارت هر رویداد تصادف توسط افسر راهنمایی و رانندگی برآورد شده و در فرم مربوط به مشخصات تصادف وارد شده است.

۲ Between-cluster Sum-of-Squares

جدول۳- انتخاب تعداد خوشه‌ها

تعداد خوشه	Gap	نسبت BSS
۱	0/0000	0/0000
۲	0/4209	0/1559
۳	0/0994	0/2593
۴	0/1894	0/3502
۵	0/0124	0/4174
۶	0/0829	0/4831
۷	0/0640	0/5384
۸	0/0081	0/5732
۹	0/0897	0/6070
۱۰	0/0514	0/6296
۱۱	0/0034	0/6457
۱۲	0/0101	0/6615
۱۳	0/0189	0/6759
۱۴	0/0281	0/6880
۱۵	0/0062	0/6966
۱۶	0/0001	0/7044
۱۷	0/0022	0/7122
۱۸	0/0069	0/7198
۱۹	0/0162	0/7265
۲۰	0/0372	0/7311

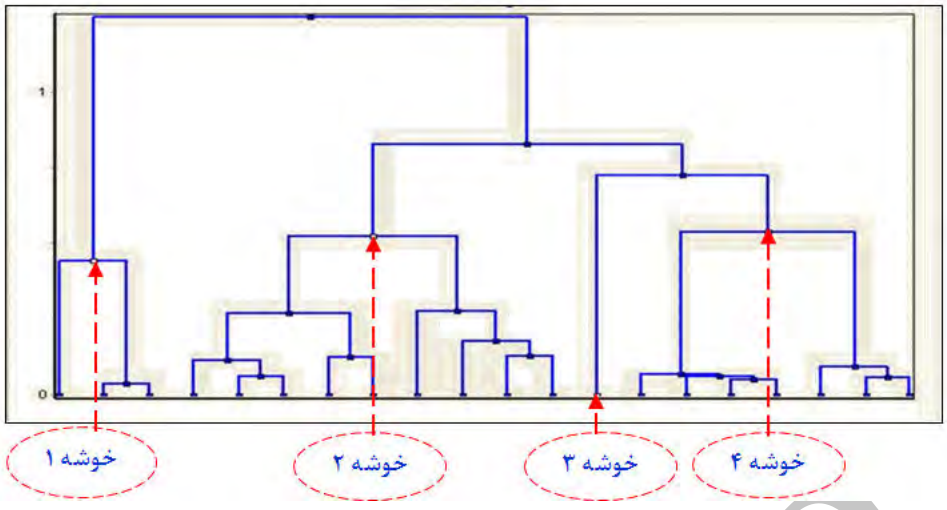
جدول ۴ مقادیر هشت معیار عددی ورودی به مدل را در هر یک از خوشه‌های مربوطه نمایش می‌دهد.

جدول ۴- توصیف خوشه‌های تصادف

خصوصیت تصادف	خوشه ۱	خوشه ۲	خوشه ۳	خوشه ۴
ماه	10	6	7	6
روز	24	11	17	16
عرض جاده (m)	32	6	8	7
مکان تصادف (km)	101/76	94/57	71/78	107/20
سن راننده	18	35	26	29
میزان خسارت	51/83	0/97	26/65	66/8
جرحی	0/07	0/13	1/61	1/24
فوتی	0/004	0/005	2/695	0/022

در این جدول مکان تصادف یک خصوصیت مکانی می‌باشد که محل تمرکز تصادفات را در هر خوشه بر اساس فاصله از مبدأ بر روی محور نمایش می‌دهد. خصوصیات روز تصادف و ماه تصادف، به ترتیب روز و ماه سال را که بیشترین تعداد تصادف در هر خوشه رخ داده است، نمایش می‌دهند. سن راننده خصوصیتی است که متوسط سن رانندگان خاطی را در هر خوشه بیان می‌کند. میزان خسارت نیز متوسط خسارات ایجاد شده در هر خوشه بر اساس شدت وقوع رویدادهای تصادف است. دو سطر انتهایی جدول نیز میانگین تعداد صدمات جرحی و فوتی را در هر خوشه بیان می‌کنند.

Archive of SID
www.SID.ir



شکل۳- نمایش دندروگرام خوشه‌بندی تجمعی سلسه‌مراتبی بر پایه K-Mean تصادفات (محور عمودی بیانگر شباهت بین خوشه‌ها می‌باشد)

۵- ارزیابی نتایج و بحث

همان‌طور که در بخش قبل بیان شد، روش خوشه‌بندی پیشنهادی بر داده‌های تصادفات محور مورد مطالعه که با فاکتورهای مکانی غنی شده‌اند اعمال گردید و نتایج به‌دست‌آمده نشان داد که رویدادهای تصادفات خوشه‌هایی را در طول محور ایجاد می‌کنند. به دلیل اینکه خوشه‌های تصادفات از قبل مشخص نبودند این امکان وجود ندارد که با شاخص‌های بیرونی[۱] به ارزیابی دقت فرآیند خوشه‌بندی پرداخته شود. لذا، ارزیابی روش و اعتبارسنجی خوشه‌ها با شاخص‌های درونی[۲] چون مقادیر فراخوانی[۳] و معیار F[۴] و همچنین یافتن خوشه‌هایی از تصادف که بهترین تناسب را با داده‌های موردنظر داشته باشند، انجام پذیرفت. جدول ۶ مقادیر فراخوانی را برای هر نوع از تصادفات در خوشه‌های ایجاد شده نمایش می‌دهد. مقادیر فراخوانی بالا بدین معنی است که اکثر داده‌های تصادف در خوشه مربوطه و به‌درستی گروه‌بندی شده‌اند. مقایسه مقادیر فراخوانی خوشه‌بندی بروش K-Means، خوشه‌بندی تجمعی سلسله‌مراتبی و خوشه‌بندی تجمعی سلسه‌مراتبی بر پایه K-Mean در جدول ۶ نشان می‌دهد که با اعمال خوشه‌بندی تجمعی سلسه‌مراتبی به فرآیند خوشه‌بندی تصادفات بروش K-Means، دقت فرایند خوشه‌بندی افزایش یافته است. همان‌گونه که پیش‌تر اشاره گردید، به‌منظور اعتبارسنجی خوشه‌ها از معیار F نیز استفاده شد. این معیار در واقع تلفیقی از مقادیر دقت[۵] و فراخوانی می‌باشد که در تحلیل درستی خوشه‌بندی مورد استفاده قرار گرفت. اگر کلاس‌های از پیش تعریف شده نوع تصادفات با i نمایش داده شوند و n_i تعداد رویدادهای تصادف کلاس i باشد؛ همچنین n_j تعداد رویدادهای تصادف خوشه j، و n_{ij} تعداد رویدادهای تصادف کلاس i باشد که در خوشه j قرار دارند، آنگاه مقادیر دقت P(i,j)، فراخوانی R(i,j) و معیار F از روابط زیر تعیین می‌گردند [۳۰]:

$$P(i,j)=\frac{n_{ij}}{n_j} \quad (۱)$$

$$R(i,j)=\frac{n_{ij}}{n_i} \quad (۲)$$

$$F(i,j)=2\frac{P(i,j)R(i,j)}{P(i,j)+R(i,j)} \quad (۳)$$

مقادیر معیار F برای روش‌های K-Means، خوشه‌بندی تجمعی سلسله‌مراتبی و خوشه‌بندی تجمعی سلسه‌مراتبی بر پایه K-Mean محاسبه و در جدول ۵ مقایسه شده‌اند. این مقایسه بیان می‌کند که دقت خوشه‌بندی داده‌های تصادف بروش پیشنهادی ۶/۷ درصد بیش از روش تجمعی سلسه‌مراتبی و بیش از ۱۰ درصد بیشتر از روش K-Mean است.

۱ External indexes
۲ Internal indexes
۳ Recall
۴ F-measure
۵ Precision

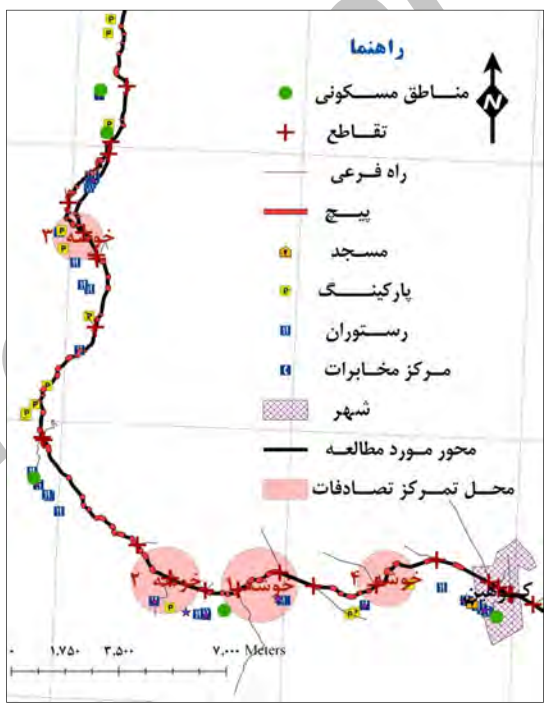
Archive of SID
www.SID.ir

جدول۵- اعتبارسنجی روش‌های خوشه‌بندی با معیار F

روش	دقت
روش K-Means	0/631
روش تجمعی سلسه‌مراتبی	0/659
روش تجمعی سلسه‌مراتبی بر پایه K-Mean	0/726

نتایج جداول ۴ و ۶ همچنین مشخص می‌کند که در خوشه یک بیشتر تصادفات در بخش‌های عریض جاده و با حداکثر میزان خسارت توسط رانندگان جوان (سن ۱۸-۱۷) و اکثراً در روزهای پایانی هر ماه رخ داده است. خوشه سه شامل تصادفات فوتی است که بیشتر از نوع تصادف جلوبه‌جلو (۳/۵۴ درصد) می‌باشد. ازآنجاکه در بخش‌هایی از محور مورد مطالعه مسیرهای رفت‌وبرگشت از یکدیگر مجزا نیستند، تصادفات از نوع جلوبه‌جلو ۱۰/۶ درصد کل تصادف محور را تشکیل می‌دهد. این نوع تصادف به ترتیب ۵۴/۳، ۱۸/۸، ۸/۸ و ۵/۸ درصد تصادفات را در خوشه‌های سه، چهار، دو و یک تشکیل می‌دهد. خوشه دو نیز بیشتر شامل تصادفات از نوع جلوبه‌عقب و جلوبه‌پهلو (به ترتیب ۳۵/۶ و ۲۵/۶ درصد) با حداقل میزان خسارت است که توسط رانندگان میان‌سال رخ داده است.

نتایج بدست آمده نشان داد که روش‌های داده‌کاوی مبتنی بر خوشه‌بندی قادر به کشف گروه‌هایی از داده‌های مشابه در درون مجموعه‌ای از داده‌های ورودی، بدون هیچ‌گونه اطلاعات قبلی از کلاس‌های مربوط به داده‌ها، هستند. در ادامه نتایج حاصله از خوشه‌بندی با روش‌های آماری و تحلیل‌های مکانی تحلیل می‌گردند تا اطلاعات بیشتری از خصوصیات مکانی داده‌های تصادفات استخراج گردد. لذا به‌منظور بررسی اثر طرح هندسی راه و همچنین کاربری‌ها و عوارض مکانی اطراف بر هر یک از خوشه‌های تصادفات با استفاده از هم‌پوشانی مکانی، محل‌های تمرکز تصادفات (ویژگی مکان تصادف در جدول ۴) بر روی محور تصویر گردید. شکل‌های ۴ و ۵ نواحی تمرکز تصادفات را بر مبنای فاصله از ابتدای محور نمایش می‌دهند.



شکل ۴- پراکندگی محل‌های تمرکز تصادفات در طول محور

www.SID.ir

جدول۶- توصیف دقت خوشه‌های تصادفات در روش ۱ (خوشه‌بندی تجمعی سلسه‌مراتبی بر پایه K-Mean)، روش ۲ (خوشه‌بندی K-Mean) و روش ۳ (خوشه‌بندی تجمعی سلسه‌مراتبی)

خوشه یک				
نوع تصادف	درصد از کل تصادفات	روش ۱ درصد در خوشه [درصد فراخوانی]	روش ۲ درصد در خوشه [درصد فراخوانی]	روش ۳ درصد در خوشه [درصد فراخوانی]
نامشخص	5/2 %	46/5 % [84/5 %]	47/5 % [84/5 %]	47 % [84/5 %]
جلوبه‌جلو_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌پهلو_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
عقب‌به‌پهلو_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌عقب_جلوبه‌پهلو_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌جلو_دیگر	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌عقب_دیگر	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌جلو_عقب‌به‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌جلو_جلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌پهلو_دیگر	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌عقب_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌جلو_جلوبه‌عقب	0/1 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌عقب_جلوبه‌پهلو	0/1 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
عقب‌به‌پهلو	0/3 %	0/1 % [4/2 %]	0/2 % [4/2 %]	0/2 % [4/1 %]
عقب‌به‌پهلو	1/0 %	0/6 % [5/4 %]	0/5 % [5/4 %]	0/6 % [5/4 %]
جلوبه‌جلو	10/6 %	5/8 % [5/2 %]	5/6 % [4/9 %]	5/8 % [5/1 %]
جلوبه‌عقب	30/5 %	22/6 % [7/0 %]	22/5 % [6/9 %]	22/5 % [7 %]
پهلوبه‌پهلو	11/8 %	5/2 % [4/2 %]	4/8 % [3/8 %]	4/9 % [4/1 %]
دیگر	16/2 %	7/7 % [4/5 %]	7/7 % [4/4 %]	7/7 % [4/4 %]
جلوبه‌پهلو	23/9 %	11/5 % [4/6 %]	11/2 % [4/4 %]	11/3 % [4/5 %]

خوشه دو				
نوع تصادف	درصد از کل تصادفات	روش ۱ درصد در خوشه [درصد فراخوانی]	روش ۲ درصد در خوشه [درصد فراخوانی]	روش ۳ درصد در خوشه [درصد فراخوانی]
جلوبه‌عقب	30/5 %	35/6 % [83/4 %]	36/1 % [83/1 %]	36/0% [83/3 %]
پهلوبه‌پهلو	11/8 %	13/7 % [83/0 %]	13/8 % [82/3 %]	13/8 % [82/5 %]
جلوبه‌پهلو	23/9 %	25/6 % [76/6 %]	25/5 % [75/1 %]	25/6 % [75/4 %]
عقب‌به‌پهلو	0/3 %	0/4 % [91/7 %]	0/4 % [87/5 %]	0/4 % [89/5 %]
عقب‌به‌پهلو	1/0 %	1/2 % [81/1 %]	1/2 % [79/7 %]	1/2 % [79/9 %]
جلوبه‌عقب_دیگر	0/0 %	0/0 % [100/0 %]	0/0 % [100/0 %]	0/0 % [100/0 %]
جلوبه‌جلو_عقب‌به‌پهلو	0/0 %	0/0 % [100/0 %]	0/0 % [100/0 %]	0/0 % [100/0 %]
جلوبه‌جلو_پهلوبه‌پهلو	0/0 %	0/0 % [100/0 %]	0/0 % [100/0 %]	0/0 % [100/0 %]
جلوبه‌پهلو_پهلوبه‌پهلو	0/0 %	0/0 % [100/0 %]	0/1 % [80/0 %]	0/2 % [85/6 %]
جلوبه‌عقب_جلوبه‌پهلو	0/1 %	0/1 % [80/0 %]	0/0 % [66/7 %]	0/0 % [68/2 %]
جلوبه‌عقب_پهلوبه‌پهلو	0/0 %	0/0 % [66/7 %]	0/0 % [50/0 %]	0/0 % [55/0 %]
جلوبه‌پهلو_دیگر	0/0 %	0/0 % [50/0 %]	0/0 % [50/0 %]	0/0 % [50/0 %]
جلوبه‌جلو_جلوبه‌پهلو	0/0 %	0/0 % [50/0 %]	0/0 % [50/0 %]	0/0 % [50/0 %]
جلوبه‌عقب_جلوبه‌پهلو_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
عقب‌به‌پهلو_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌جلو_جلوبه‌عقب	0/1 %	0/0 % [25/0 %]	0/0 % [25/0 %]	0/0 % [25/0 %]
جلوبه‌جلو_دیگر	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌جلو	10/6 %	9/0 % [59/0 %]	8/7 % [56/5 %]	7/7 % [57/5 %]
دیگر	16/2 %	13/5 % [59/4 %]	13/3 % [57/8 %]	14/2 % [58/5 %]
نامشخص	5/2 %	0/9 % [12/3 %]	0/9 % [11/8 %]	0/9 % [12/0%]

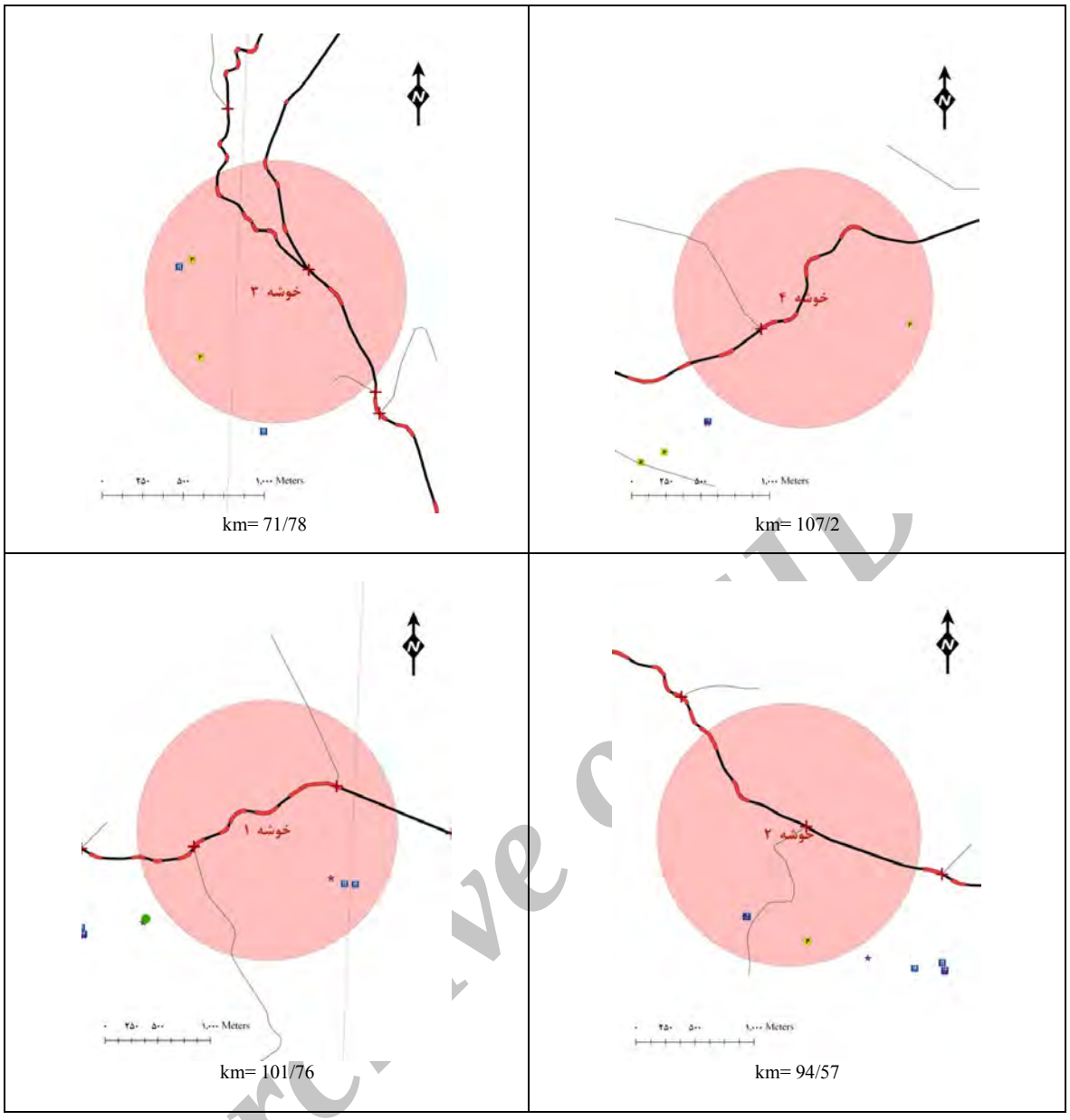
Archive of SID
www.SID.ir

خوشه سه				
نوع تصادف	درصد از کل تصادفات	روش ۱ درصد در خوشه [درصد فراخوانی]	روش ۲ درصد در خوشه [درصد فراخوانی]	روش ۳ درصد در خوشه [درصد فراخوانی]
جلوبه‌جلو	10/6 %	54/3 % [3/3 %]	44/8 % [3/2 %]	44/9 % [3/3 %]
عقب‌به‌پهلو	1/0 %	2/2 % [1/4 %]	28/6 % [1/2 %]	28/0 % [1/3 %]
جلوبه‌پهلو	23/9 %	23/9 % [0/6 %]	20/0 % [0/8 %]	18/5 % [0/7 %]
جلوبه‌جلو_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	1/0 % [0/4 %]	1/0 % [0/1 %]
جلوبه‌پهلو_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
عقب‌به‌پهلو_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌عقب_جلوبه‌پهلو_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌پهلو_دیگر	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌عقب_دیگر	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌جلو_عقب‌به‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌جلو_جلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌جلو_دیگر	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌عقب_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌جلو_جلوبه‌عقب	0/1 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌عقب_جلوبه‌پهلو	0/1 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
عقب‌به‌پهلو	0/3 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
دیگر	16/2 %	13/0 % [0/5 %]	0/0 % [0/0 %]	2/0 % [0/3 %]
نامشخص	5/2 %	0/0 % [0/0 %]	1/0 % [0/1 %]	0/0 % [0/1 %]
پهلوبه‌پهلو	11/8 %	2/3 % [0/1 %]	2/9 % [0/1 %]	3/1 % [0/1 %]
جلوبه‌عقب	30/5 %	4/3 % [0/1 %]	1/9 % [0/1 %]	2/5 % [0/1 %]

خوشه چهار				
نوع تصادف	درصد از کل تصادفات	روش ۱ درصد در خوشه [درصد فراخوانی]	روش ۲ درصد در خوشه [درصد فراخوانی]	روش ۳ درصد در خوشه [درصد فراخوانی]
دیگر	16/2 %	31/3 % [35/5 %]	30/8 % [32/0 %]	30/۰ % [33/5 %]
جلوبه‌جلو	10/6 %	18/8 % [32/5 %]	18/2 % [30/4 %]	۱۹/0 % [31/7 %]
جلوبه‌جلو_دیگر	0/0 %	0/2 % [100/0 %]	0/1 % [100/0 %]	0/1 % [98/0 %]
جلوبه‌جلو_جلوبه‌عقب	0/1 %	0/2 % [75/0 %]	0/2 % [75/0 %]	0/2 % [75/0 %]
جلوبه‌عقب_جلوبه‌پهلو_پهلوبه‌پهلو	0/0 %	0/1 % [100/0 %]	0/1 % [100/0 %]	0/1 % [100/0 %]
عقب‌به‌پهلو_پهلوبه‌پهلو	0/0 %	0/1 % [100/0 %]	0/1 % [100/0 %]	0/1 % [100/0 %]
جلوبه‌جلو_جلوبه‌پهلو	0/0 %	0/1 % [50/0 %]	0/1 % [50/0 %]	0/1 % [50/0 %]
جلوبه‌پهلو_دیگر	0/0 %	0/1 % [50/0 %]	0/1 % [50/0 %]	0/1 % [50/0 %]
جلوبه‌عقب_پهلوبه‌پهلو	0/0 %	0/1 % [33/3 %]	0/1 % [30/0 %]	0/1 % [29/5 %]
جلوبه‌عقب_جلوبه‌پهلو	0/1 %	0/1 % [20/0 %]	0/1 % [19/3 %]	0/1 % [19/5 %]
جلوبه‌پهلو	23/9 %	23/6 % [18/2 %]	0/1 % [17/0 %]	22/6 % [18/0 %]
جلوبه‌پهلو_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	23/7 % [0/8 %]	1/3 % [0/1 %]
جلوبه‌جلو_پهلوبه‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌عقب_دیگر	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
جلوبه‌جلو_عقب‌به‌پهلو	0/0 %	0/0 % [0/0 %]	0/0 % [0/0 %]	0/0 % [0/0 %]
عقب‌به‌پهلو	1/0 %	0/7 % [12/2 %]	0/7 % [11/5 %]	0/7 % [11/6 %]
عقب‌به‌پهلو	0/3 %	0/1 % [4/2 %]	0/1 % [3/3 %]	0/1 % [3/7 %]
پهلوبه‌پهلو	11/8 %	8/0 % [12/7 %]	8/5 % [12/6 %]	8/5 % [12/7 %]
نامشخص	5/2 %	0/9 % [3/2 %]	1/0 % [3/1 %]	1/8 % [3/0 %]
جلوبه‌عقب	30/5 %	15/6 % [9/4 %]	16/0 % [9/2 %]	15/۱ % [9/4 %]

Archive of SID

www.SID.ir



شکل ۵- موقعیت مکانی خوشه‌های تصادف بر مبنای فاصله از ابتدای محور

در ادامه با استفاده از الگوریتم‌های داده‌کاوی استخراج الگو و همچنین تحلیل‌های مکانی مجاورت[۱] هم‌پوشانی[۲] و تحلیل‌های یکپارچه داده‌های توصیفی و مکانی[۳] در محیط GIS، فاکتورهای مکانی تأثیرگذار بر تصادفات هر خوشه تحلیل شدند.

خروجی تحلیل مکانی مجاورت حاکی از آن است که ناحیه تمرکز تصادفات در خوشه سه نزدیک به مراکز تجاری و استراحتگاه‌های کنار جاده است. خوشه چهار در مجاورت مراکز صنعتی شهر کوهین واقع شده است.

همچنین داده‌های تصادفات خوشه‌های یک و دو با استفاده از الگوریتم داده‌کاوی استخراج الگوی Separate-and-Conquer مورد بررسی قرار گرفتند. در این مرحله الگویی کلی با در نظر گرفتن مجموعه‌ای از رویدادهای تصادف خوشه موردنظر ایجاد گردید، رویدادهای بررسی شده از داده‌های آموزشی حذف شدند (بخش separate الگوریتم)، و سپس الگوهای دیگری که مابقی رویدادها را شامل شوند ایجاد شدند (بخش conquer الگوریتم). این فرایند تکرار یافت تا همه رویدادهای تصادف جهت استخراج الگوهای تصادفات خوشه مورد بررسی قرار گرفتند. نتایج تحلیل استخراج الگو دلالت بر این موضوع داشت که طرح هندسی راه شامل پیچ و تقاطع و همچنین

۱ Proximity Function
۲ Overlay
۳ Integrated Analysis of Spatial and Attribute Data

www.SID.ir

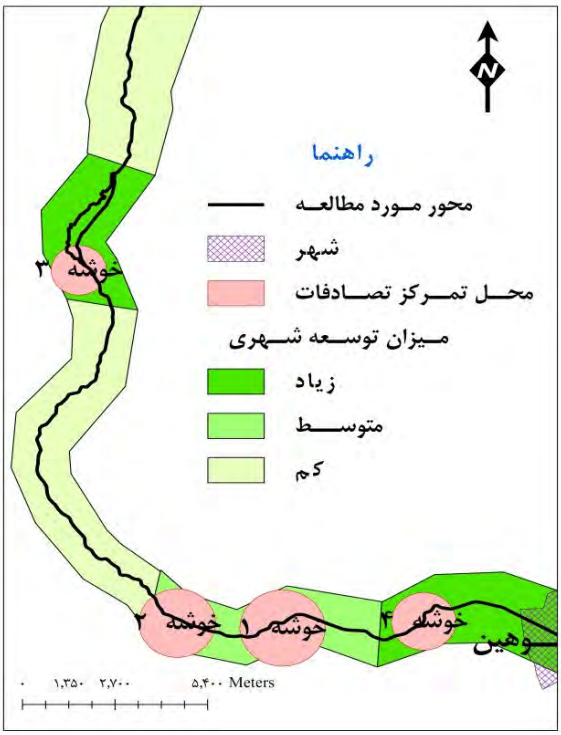
همسایگی به تسهیلات تجاری و مسکونی از علل غالب تصادفات در خوشه‌های یک و دو می‌باشند.

تحلیل‌های مکانی همچنین نشان دادند که خوشه‌های تصادف در مبادی شهرها یا در نزدیکی مراکز جمعیتی واقع شده‌اند.

علاوه بر کاربری‌ها و تسهیلات رفاهی، توسعه شهری مجاور راه نیز می‌تواند بر نرخ و شدت تصادفات هر خوشه از تصادفات تأثیرگذار باشد. لذا به‌منظور بررسی تأثیرات مکانی توسعه شهری بر رکوردهای تصادفات هر خوشه، متغیری با عنوان توسعه شهری که تراکم شهری را در محدوده ۵۰۰ متری مجاور راه مدل می‌کند، ایجاد گردید. بدین منظور مطابق شکل ۶ با استفاده از نقشه‌های رقومی موجود، تصویر ماهواره‌ایIRS، و نقشه‌های Google، تراکم شهری مجاور راه در محدوده ۵۰۰ متری راه و در سه کلاس (کم، متوسط و زیاد) استخراج شده و بر رویدادهای تصادف و مراکز خوشه‌ها هم‌پوشانی گردید.

تحلیل مکانی خوشه‌های تصادف نسبت به میزان توسعه شهری مجاور راه نشان داد که درخوشه‌های ۱ و ۲ اکثر تصادفات در نواحی با توسعه شهری متوسط رخ داده‌اند. همچنین در خوشه ۳، بیش از ۶۵ درصد و در خوشه ۴، نزدیک به ۸۰ درصد رکوردهای تصادف در نواحی با توسعه شهری زیاد رخ داده‌اند. نتایج نشان می‌دهد که وجود توسعه شهری در مجاورت راه‌های اصلی دوخطه دوطرفه یکی از فاکتورهای تأثیرگذار بر افزایش نرخ تصادفات است.

با مقایسه زمان وقوع رویدادهای تصادف خوشه‌ها با ساعات اوج ترافیکی محور[1] مورد مطالعه، که در جدول ۷ نمایش داده شده است، مشخص گردید که به ترتیب ۵۴، ۴۲، ۳۸ و ۶۲ درصد رویدادهای تصادفات خوشه‌های یک، دو، سه و چهار در ساعات اوج ترافیکی رخ داده‌اند. لذا، حجم بالای تردد وسایل نقلیه در ساعات اوج ترافیکی یکی از علل اصلی وقوع تصادف در مجاورت شهرها و مراکز جمعیتی در راه‌های اصلی دوخطه دوطرفه می‌باشد که می‌بایست در مطالعات ایمنی راه در راه‌های برون‌شهری مورد توجه و بررسی بیشتری قرار گیرد.



شکل ۶- خوشه‌های تصادف و میزان توسعه شهری در مجاورت ۵۰۰ متری راه در سه سطح کم، متوسط و زیاد

این یافته، نتایج تحقیقات پیشین [۱۰، ۲۲] را که در آن ترافیک یکی از دلایل اصلی تصادفات در راه‌های برون‌شهری ذکر شده است، تأیید می‌کند. هرچند برخی از مطالعات [۳۱] نیز حاکی از آن است که ترافیک بر تعداد رویدادهای تصادفات (چه تصادفات فوتی و جرحی شدید یا تصادفات با جراحت کم) تأثیر مستقیمی ندارد و در تلفیق با فاکتورهای دیگر است که به‌عنوان یکی از علل تصادفات در محورهای برون‌شهری تلقی می‌گردد.

جدول۷- ساعات اوج ترافیکی در محور قزوین-رشت در یک روز آخر هفته

نام محور	وضعیت ترافیک	میانگین تردد در ساعت	میانگین سرعت	ساعت
قزوین-لوشان (جاده قدیم)	نیمه-سنگین	۱۴۸	۶۰	۱۷-۲۱
لوشان-قزوین (جاده قدیم)	روان	۸۸	۶۴	۱۷-۲۱
رودبار- رستم آباد (جاده قدیم)	سنگین	۲۸۷	۵۱	۱۷-۲۱
رستم آباد- رودبار (جاده قدیم)	روان	۹۶	۵۶	۱۷-۲۱

[1] مرجع اخذ داده http://irantraffic.ir

Archive of SID
www.SID.ir

۶- نتیجه‌گیری

رویکرد اصلی این تحقیق، تلفیق تحلیل‌های مکانی و داده‌کاوی جهت تحلیل نوع تصادفات و فاکتورهای مکانی مؤثر بر تمرکز تصادفات در راه‌های برون‌شهری بود. بدین منظور یک روش خوشه‌بندی تجمعی سلسه‌مراتبی بر پایه K-Mean ارائه گردید. با به‌کارگیری این روش در محور مورد مطالعه، داده‌های تصادفات به چهار خوشه اصلی تقسیم شدند و نتایج با استفاده از معیارهای درونی اعتبارسنجی مورد ارزیابی قرار گرفتند. نتایج نشان داد که با اعمال خوشه‌بندی تجمعی سلسله مراتبی بر K-Means، دقت خوشه‌بندی داده‌های تصادفات بهبود می‌یابد. علاوه بر این، تحلیل‌های مکانی جهت بررسی خصوصیات مکانی خوشه‌های تصادفات و اثر کاربری‌ها، توسعه شهری و عوارض اطراف بر وقوع تصادفات مورد استفاده قرار گرفت. تحلیل‌های مکانی همسایگی و مجاورت نشان داد که علاوه بر توسعه شهری زیاد، طرح هندسی راه و مجاورت به کاربری‌های و تسهیلات رفاهی از دلایل عمده تمرکز تصادف در بخش‌هایی از محور مورد مطالعه بوده‌اند. علاوه بر این، مشخص گردید که حجم بالای تردد وسایل نقلیه در ساعات اوج ترافیکی بخصوص در مجاورت شهرها و مراکز جمعیتی بر افزایش نرخ تصادفات در راه‌های اصلی دوخطه دوطرفه مؤثر است.

یافته‌های این تحقیق می‌توانند به‌منظور کاهش تصادفات جاده‌ای و اصلاح زیرساخت‌های راه توسط متخصصین حوزه ایمنی راه بکار گرفته شوند. لازم به ذکر است که روش پیشنهادی تحقیق در تلفیق تحلیل‌های مکانی و داده‌کاوی به‌منظور تحلیل تصادفات جاده‌ای می‌تواند به‌عنوان روشی مکمل در کنار سایر روش‌های بهبود ایمنی راه مورد استفاده قرار گیرد. در نهایت در راستای تکمیل پژوهش چندین افق تحقیقاتی به شرح زیر پیشنهاد می‌گردد:

- به‌کارگیری روش تحقیق جهت تعیین محل تمرکز تصادفات راه‌های درون‌شهری
- استفاده از روش‌های استنتاجی جهت برآورد شدت تصادفات در محل‌های تمرکز تصادفات

سپاسگزاری

بدین‌وسیله نویسندگان مقاله مراتب قدردانی خود را از سازمان نقشه‌برداری کشور، پلیس راه و پژوهشکده حمل‌ونقل دانشگاه علم و صنعت جهت در اختیار قرار دادن داده‌های مورد استفاده در این تحقیق اعلام می‌دارند.

مراجع

[1] Anderson, T.K. (2009) "Kernel density estimation and K-means clustering to profile road accident hotspots" Accident Analysis & Prevention, Vol. 41, No. 3, pp.359–364.

[2] Clarke, R., Forsyth, R. and Wright, R. (1998) "Machine learning in road accident research: decision trees describing road-accidents during cross-flow turns" Ergonomics, Vol.41, No.7, pp.1060–1079.

[3] Chang, L. and Chen W. (2005) "Data mining of tree-based models to analyze freeway accident frequency" Safety Research, Vol.36, No.4, pp.365-375.

[4] Delmelle, E.C., Thill, J.C. and Ha, H.H. (2012) "Spatial epidemiologic analysis of relative crash risk factors among urban bicyclists and pedestrians" Transportation, Vol.39, No.2, pp. 433-448.

[5] Delen, D., Sharda, R. and Bessonov, M. (2006) "Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks" Accident Analysis & Prevention, Vol.38, No.3, pp. 434-444.

[6] Effati, M., Rajabi, M.A., Samadzadegan, F. and Blais, J.A. (2012) "Developing a Novel Method for Road Hazardous Segment Identification Based on Fuzzy Reasoning and GIS" Journal of Transportation Technologies, Vol.2, pp.32-40.

[7] Erdogan, S., Yilmaz, I., Baybura, T. and Gullu, M. (2008) "Geographical information systems aided traffic accident analysis system case study: city of Afyonkarahisar" Accident Analysis and Prevention, Vol.40, No.1, pp.174-181.

Archive of SID
www.SID.ir

[8] Ester, M., Kriegel, H.P., Sander, J. and Xu, X. (1996) "A density-based algorithm for discovering clusters in large spatial databases with noise" Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), E. Simoudis, J. Han, and U.M. Fayyad, eds., AAAI Press, pp. 226–231.

[9] Flahaut, B., Mouchart, M., Martin, E.S. and Thomas, I. (2002) "The local spatial autocorrelation and kernel method for identifying black zones a comparative approach" Accident Analysis and Prevention, Vol.35, pp. 991-1004.

[10] Garnowski M. and Manner H. (2011) "On factors related to car accidents on German Autobahn connectors" Accident Analysis and Preventation, Vol. 43, No. 5, pp.1864-71.

[11] Gundogdu, I.B. (2010) "Applying linear analysis methods to GIS-supported procedures for preventing traffic accidents: case study of Konya" Safety Science, Vol. 48**,** pp.763-769.

[12] Ha, H.H. and Thill, J.C. (2011) "Analysis of traffic hazard intensity: A spatial epidemiology case study of urban pedestrians" Computers. Environment and Urban Systems, Vol.35, pp. 230-240.

[13] Jain, A.K. and Dubes, R.C. (1988) "Algorithms for clustering data" PrenticeHall, New Jersey.

[14] Jha, M., McCall, C. and Schonfeld, P. (2001) "Using GIS, genetic algorithms and visualization in highway development" Journal of Computer - aided Civil and Infrastructure Engineering, Vol. 16, No.6, pp. 399 - 414.

[15] Jing Wang and Xin Wang (2011) "An Ontology-based traffic accident risk mapping framework" Advances in Spatial and Temporal Databases, Lecture Notes in Computer Science, Springer-Verlag Berlin Heidelberg 2011, Volume, 6849/2011, pp. 21-38.

[16] Joshua, S. and Garber, N. (1990) "Estimating truck accident rate and involvement using linear and poisson regression models" Transportation Planning and Technology, Vol.15, pp.41-58.

[17] Martin, J.L. (2002) "Relationship between crash rate and hourly traffic flow on interurban motorways" Accident Analysis and Prevention, Vol.34, No.5, pp.619 – 629.

[18] McCarthy, P.S. (1999) "Public policy and highway safety: A citywide perspective" Regional Science and Urban Economics, Vol.29, No.3, pp.231 – 244.

[19] Miaou, S. and Lum, H. (1993) "Modeling vehicle, accidents and highway geometric design relationships" Accident Analysis and Prevention, Vol. 25, No.6, pp.689-709.

[20] Montella A., Aria M., D'ambrosio A. and Mauriello F. (2012) "Analysis of powered two-wheeler crashes in Italy by classification trees and rules discovery" Accident Analysis & Prevention, Vol. 49, pp. 58-72.

[21] PIARC Road Safety Manual, (2003)

[22] Ramírez, B.A., Izquierdo, F.A., Fernández, C.G. and Méndez A.G. (2009) "The influence of heavy goods vehicle traffic on accidents on different types of Spanish interurban roads" Accident Analysis and Prevention, Vol. 41, No.1, pp.15-24.

[23] Shankar, V., Mannering, F. and Barfield, W. (1995) "Effect of roadway geometric and environment factors on rural freeway accident frequencies" Accident Analysis and Prevention, Vol. 27, pp. 11-23.

[24] Siddiqui, C., Abdel-Aty, M. and Huang, H. (2012) "Aggregate nonparametric safety analysis of traffic zones" Accident Analysis and Prevention, Vol. 45, pp. 317-325.

[25] Sohn, S. and Hyungwon, S. (2001) "Pattern recognition for a road traffic accident severity in Korea" Ergonomics, Vol. 44, No. 1, pp. 101-117.

[26] Steenberghen, T., Dufays, T., Thomas, I. and Flahaut, B. (2004) "Intra-urban location and clustering of road accidents using GIS: a Belgian Case" International Journal of Geographic Information Science, Vol. 18, No.2, pp.169–181.

Archive of SID

www.SID.ir

[27] Sun, X. (2003) "Identifying highway safety Patterns and trends with a crash data analysis program" the International Forum on Traffic Records and Highway Information Systems, Denvor CO.

[28] Tan, P.N., Steinbach, M. and Kumar, V. (2006) "Introduction to Data Mining" Pearson Addison Wesley, Boston.

[29] Tibshirani, R., Walther, G. and Hastie, T. (2000) "Estimating the number of clusters in a dataset via the gap statistics" Technical report Stanford University, Stanford.

[30] Van Rijsbergen, C.J. (1979) "Information Retrieval" Butterworth-Heinemann. USA.

[31] Wang, C., Quddus, M.A. and Ison S.G. (2009) "Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England" Accident Analysis and Prevention, Vol.41, No.4, pp.798-808.

[32] World Health Organization (WHO), (2009)

[33] Yamada, I. and Thill, J.C. (2004) "Comparison of planar and network K-functions in traffic accident analysis" Journal of Transport Geography 12, pp.149-158.

[۳۴] شفیعی‌مقدم، پ.، ۱۳۸۵، بررسی روند حوادث ترافیک ایران و جهان و عملکرد هلال‌احمر بر اساس مدل ماتریس ADDON، دو ماهنامه علمی- تخصصی پلیس راهور ناجا، سال سوم، شماره پانزدهم، صفحات ۴۵۳-۴۶۲.

Archive of SID

www.SID.ir