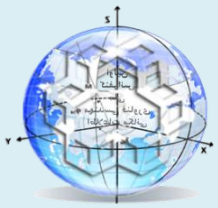
اولین کنفرانس مهندسی فناوری اطلاعات مکانی
دانشگاه صنعتی خواجه نصیرالدین طوسی
دانشکده مهندسی نقشه‌برداری
۲۹ و ۳۰ دی ماه ۱۳۹۴



The 1st National Conference on Geospatial Information Technology

K.N.Toosi University of Technology
Faculty of Geomatics Engineering

19 - 20 January 2016

بازشناسی الگوهای ترافیکی با استفاده از داده کاوی مکانی-زمانی

سجادعزیزی ۱، علی اسماعیلی ۲

۱-دانشجوی کارشناسی ارشد، دانشگاه تحصیلات تکمیلی و صنعتی فناوری پیشرفته کرمان
۲-استادیاردانشکده عمران و نقشه برداری ، دانشگاه تحصیلات تکمیلی و صنعتی فناوری پیشرفته کرمان

چکیده :

افزایش جمعیت بطور طبیعی منجر به افزایش عبور و مرور و در نتیجه رشد ترافیک می شود. اهمیت و تاثیر ترافیک بر جنبه های مختلف زندگی بشری بر هیچ کس پوشیده نیست. با در نظر گرفتن این مسائل بررسی راههای کنترل و مدیریت این معضل بیش از پیش احساس می شود در این تحقیق هدف ارائه سازکار مناسبی جهت استخراج الگوهای ترافیکی از داده های عبور و مرور خودروها می باشد.در این پژوهش ابتدا داده های ترافیکی خودروها پردازش گردیده است و پس از آماده سازی اولیه داده ها، با استفاده از تکنیک های داده کاوی به استخراج الگوهای موجود در این داده ها پرداخته شده است در این تحقیق با توجه به ماهیت مکانی- زمانی داده ها ابتدا داده های ترافیکی، خوشه بندی مکانی گردیده است و سپس به خوشه های زمانی تقسیم بندی شده اند. سعی بر این بوده است تا ترکیب خوشه بندی مکانی-زمانی متفاوت جهت بررسی و مقایسه وضعیت ترافیکی حاصله انجام پذیرد در انتها از الگوهای استخراجی جهت مطالعه وضعیت ترافیکی زمان های آتی استفاده شده است

واژه‌های کلیدی : :ترافیک، الگو ترافیکی، داده کاوی، خوشه بندی، زمانی-مکانی

نویسنده مکاتبه کننده : سجادعزیزی
آدرس پستی : کرمان – کیلومتر ۳۰ جاده هفت باغ- دانشگاه تحصیلات تکمیلی و فناوری پیشرفته کرمان
تلفن : ۰۹۱۷۷۳۲۵۴۸۸
آدرس پست الکترونیک : Sajad.azizi107@gmail.com

۱- مقدمه

انسان همواره به دنبال یافتن الگوها و ارتباطات موجود بین پدیده‌های مختلف بوده است برای این امر با جمع آوری اطلاعات مرتبط با پدیده های مورد مطالعه از منابع مختلف داده و انجام تحلیل های مناسب، سعی در دستیابی به درک بهتری از آنها داشته است. در سالهای اخیر با توسعه روش های جمع آوری داده و مواجعه با حجم عظیمی از داده در دسترس، نیاز به روش هایی که بصورت خودکار قابلیت استخراج الگوهای مختلف را داشته باشند بیش ازپیش حس میگردد با پیشرفت علوم مهندسی، روش های داده کاوی به عنوان جایگزین روش های پیشین جهت استخراج الگوهای مفید، قابل درک و معتبر از داده ها مطرح شده اند.

از سوی دیگر جریان ترافیک و حل معضلات آن یکی از زمینه های کاربردی مهم در سیستم اطلاعات مکانی می باشد. ترافیک به عنوان یکی از معضلات شهری، پدیده ای است که در سالهای اخیر مورد توجه قرارگرفته است. شهرهای بسیاری در جهان وجود دارند که جمعیت آنها در قیاس با وسعتشان بسیار زیاد می باشد بالطبع ساکنین این شهرها نیاز به سیستم حمل ونقل کارآمدی جهت فائق آمدن بر نیاز های روزمره خویش دارند. رشد جمعیت و نیاز به سیستم حمل ونقل از یک سو و هزینه های اضافه کردن تسهیلات حمل و نقل از سوی دیگر منجر به افزایش هزینه های مالی و زمانی می شود. از این رو با توجه به اهمیت و تاثیر ترافیک در جوانب مختلف حیات انسان، در نظر گرفتن سازوکارهای مفید جهت مدیریت آنها، همواره مورد توجه متخصصان حوزه های مختلف بوده است. با توجه به حجم بالای داده در مدیریت ترافیک، داده کاوی مکانی می تواند قابلیت خوبی جهت شناسایی واستخراج الگوهای ترافیک برای کمک به متولیان امر ترافیک در راستای حل مشکلات ترافیکی باشد.

در این تحقیق، ارزیابی توانایی روش داده کاوی مکانی و داده کاوی مکانی_زمانی در مدیریت ترافیک شهری مورد نظر بوده است. همچنین اهمیت و نیاز به استخراج الگوهای ترافیکی منطقه ای برای مدیریت ترافیک شهری شرح داده شد.با استفاده از الگوهای استخراج شده، می توان وضعیت عبور و مرور یک منطقه را به نحو موثری بهبود بخشید. در روش پیشنهادی این تحقیق پس از آماده کردن داده ها و انتقال آنها برای تحلیل، از تکنیک های داده کاوی برای انجام تحلیل استفاده شده است پس از انجام تحلیل های داده کاوی، تراکم ترافیک در قسمت های مختلف خیابان ها بدست می آید.

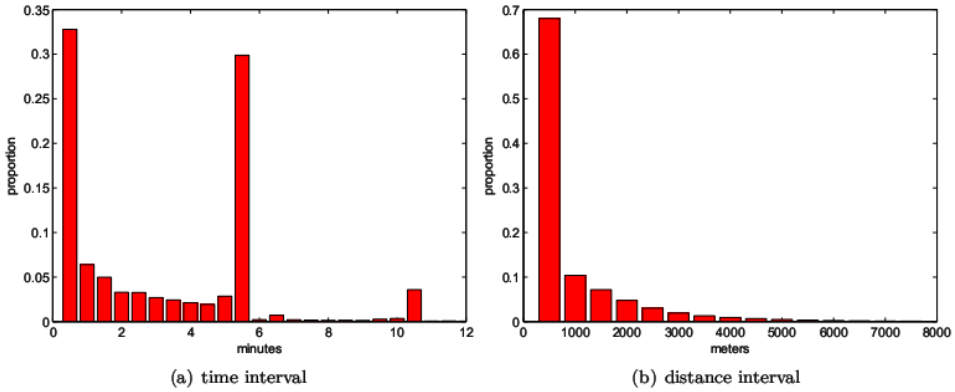
تحلیل مورد استفاده در این تحقیق، خوشه بندی داده های نقطه ای ترافیک بر اساس دو پارامتر مکان و زمان جهت استخراج الگوهای ترافیکی جهت بررسی وضعیت نسبی ترافیک ناحیه ای هر خیابان بوده است. این روش ابزاری کارا به منظور مقایسه وضعیت نسبی ترافیکی نواحی مختلف یک خیابان را در اختیار متخصصان و مدیران شهری قرار می گیرد برای این مهم از داده های نقطه ای مربوط به عبور و مرور وسایل نقلیه به همراه داده های خطی خیابان ها استفاده شده است نتایج عملی نشان می دهدکه روش به کار گرفته شده است بینش خوبی راجع به وضعیت نسبی ترافیک در ناحیه ای از خیابان می دهد.

۲-داده های مورد استفاده

۲-۱- داده های ترافیکی

داده های مورد استفاده در این تحقیق، داده های ترافیکی متعلق به شهر پکن بوده است. این داده ها مربوط به ۱۰۳۵۷تاکسی بوده که دارای حجمی بالغ بر ۱۵ میلیون رکورد داده می باشد این داده ها در بازه زمانی ۲ فوریه ۲۰۰۸ تا ۸ فوریه ۲۰۰۸ جمع آوری شده است که اغلب با نرخ زمانی کمتر از یک دقیقه برداشت شده است در شکل یک نرخ زمانی و نرخ مکانی داده ها آورده شده است، شهر پکن با جمعیتی نزدیک به ۱۲ میلیون نفر به عنوان پایتخت کشور چین شناخته می شود موقعیت جغرافیایی این شهر در محدوده طول جغرافیایی "48 '55 °115 الی "12 '45 °116 و

عرض جغرافیایی "34 ’12 °34 الی "41 ’54 °36 می باشد. در شکل ۱ توزیع نرخ مکانی و زمانی داده ها مشاهده می-گردد همچنین در شکل ۲ پراکندگی داده ها در شبکه راههای شهر پکن نمایش داده شده است.



(a) time interval (b) distance interval

شکل ۱: نمودار نرخ مکانی و زمانی داده ها

شکل۲-پراکندگی داده ها در شبکه راههای شهر پکن

۲-۲- انتخاب داده های خطی منطقه مطالعاتی

امروزه سرویس های زیادی وجود دارند که می توان به کمک آنها به منابع داده ای دست یافت یکی از قدرتمند ترین این سرویس ها موتورجستجوگر Google و تمامی پیوست های آن از قبیل Google Earth می باشد. یکی از نیازهای واضح این تحقیق داده های خطی خیابان های منطقه مورد مطالعه می باشد که با استفاده از قابلیت Google Earth این داده ها در فرمت kmz. استخراج گردیده است. لازم به ذکر است که این نرم افزارداده ها را در قالب فرمت kmz. به عنوان خروجی می دهد.

۳. پردازش داده ها

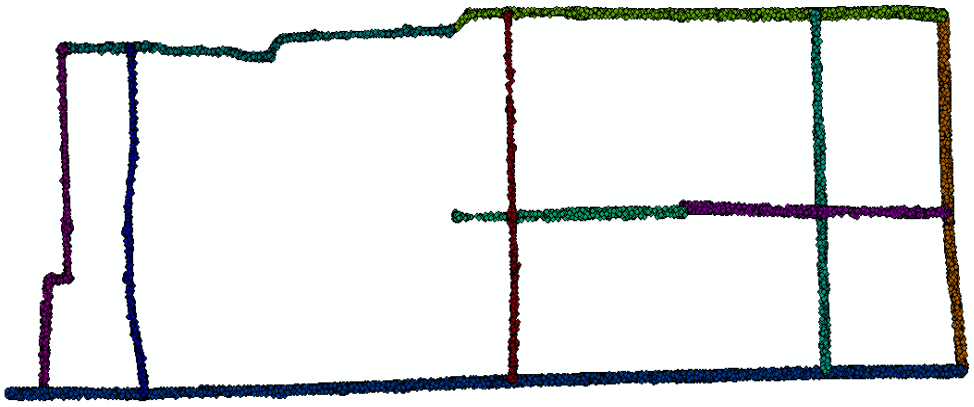
۳-۱- انتخاب داده ای ترافیکی محدوده مورد مطالعه از داده های اولیه

با توجه به اینکه داده های ترافیکی بدست آمده از طریق سامانه تعیین موقعیت جهانی، متعلق بـه تمـامی سـطح شـهر پکن بوده است، و منطقه مورد مطالعه تنها بخشی معدود از شهر پکن بوده است در ایـن مرحلـه بایسـتی داده هـایی را استخراج کرد که متعلق به این معدوده باشد.

شکل۳- داده های کلی محدوده مورد مطالعه

۳-۲- تمیزسازی داده های منطقه مطالعاتی

با توجه به شکل ۳، همانطور که مشاهده می‌گردد در بین خیابان‌های مورد مطالعه داده‌هایی وجـود داردکـه نمـی‌تـوان متعلق به عبور و مرور در هیچ یک از خیابان‌های مورد مطالعه قرار داد بنابرین این داده‌هـا مـی‌تواننـد نتـایج حاصـله را دوراز واقعیت نشان دهند بنابرین جهت انجام فرایند استخراج الگو باید این داده ها حذف گردد تا بتوان به نتایج واقعی-تری رسید. داده هایی که پس از این مرحله حاصل میگردد داده هایی می باشند که جهت انجام تحلیل های داده کاوی مکانی_زمانی مفید می باشند این داده ها در شکل۴ نمایش داده شده است.



شکل۳- داده مفید جهت فرایند داده کاوی

۴- روش انجام کار

تقسیم بندی های عمده در داده کاوی مشابه با حالت معمول بوده با این تفاوت که نه تنها در تعریف هر یک از روش ها بازنگری شده بلکه الگوریتم های خاصی نیز برای تحلیل های مکانی توسعه داده شده اند. ذکر این نکته ضروری است که تقسیم بندی های ذکر شده در جدول ۱ حالت عمومی روش های مورد استفاده در داده کاوی مکانی بوده و هر یک از این روشها خود می توانند دارای تقسیم بندی های مختلفی باشند.

جدول ۱- تحلیل های مورد استفاده در داده کاوی مکانی

تحلیل	توضیح
خوشه بندی مکانی	قرار دادن داده های مشابه مکانی در گروه های از قبل نامعلوم
طبقه بندی مکانی	استفاده از دادههای مکانی آموزشی جهت تعیین کلاس کل داده ها
رگرسیون مکانی	نمایش تغییرات مولفه های داده های مکانی بر اساس دیگر مولفه ها
قواعد هم مکانی	استفاده از گزاره های مکانی جهت تعیین قوانین وابسته به هم
تشخیص روند مکانی	تعیین تغییرات پارمترهای توصیفی براساس تغییرات مکانی
کشف انحرافات مکانی	تشخیص اشیاء مکانی با مقادیر توصیفی متفاوت با نقاط مجاور

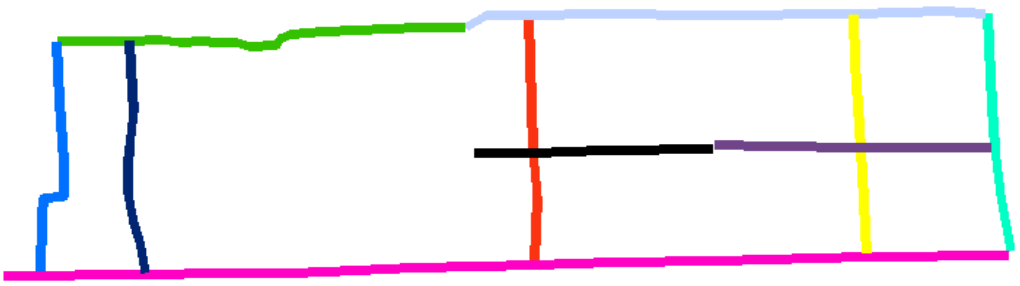
۴-۱ -خوشه بندی مکانی

خوشه بندی یکی از روش های یادگیری غیر نظارت شده است که در این دسته از تحلیل ها بر چسب کلاس ها از ابتدا معلوم نبود و هدف یافتن گروه های مشابه در داده ها به گونه ای است که اشیاء درون هر خوشه تا حد امکان به هم شبیه بوده و در عین حال کمترین شباهت را به اشیائ دیگر خوشه ها داشته باشند. یافتن معیار مناسب جهت سنجیدن تشابه اشیاء مکانی، بر اساس متریک مورد استفاده در فضای داده ها می باشد.

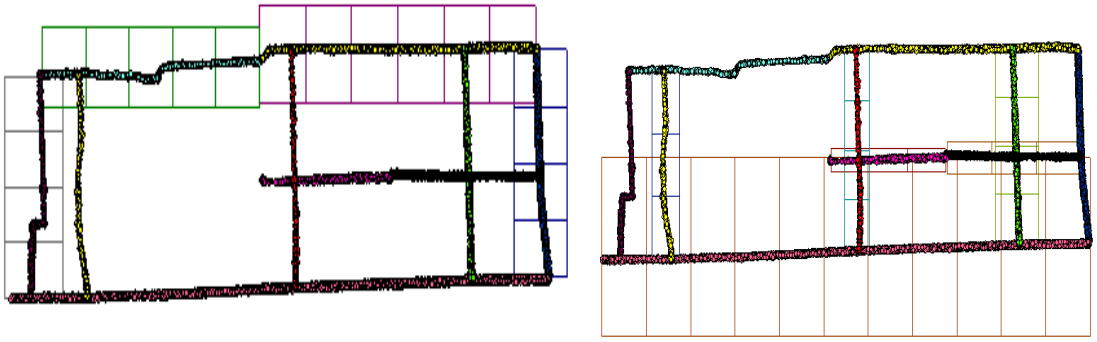
در خوشه بندی مکانی سعی بر این بوده است که داده های هم مکان با هم در یک خوشه قرار گیرند بدین جهت از دونوع خوشه بندی مکانی استفاده گردیده است. نوع اول هر خیابان را یک خوشه مکانی در نظر گرفته شده است و در نوع دوم هر خیابان را به خوشه های هایی با طول مساوی تقسیم شده است. در اشکال زیر هر دو نوع خوشه بندی مشاهده می گردد.

۵-پیاده سازی

در این تحقیق از روش خوشه‌بندی مکانی استفاده گردیده است بدین صورت که با توجه به بعد مکانی و زمانی داده‌ها ابتدا باید بر مبنای یکی از این دومولفه خوشه‌بندی صورت گیرد و سپس براساس مولفه دیگر خوشه بندی صورت گیرد. در این تحقیق ابتدا داده ها خوشه بندی مکانی شده است و سپس خوشه های مکانی براساس پارامتر زمان خوشه بندی شده است و در نهایت خوشه های حاصل خوشه های مکانی_زمانی می باشد.خوشه‌های حاصل شده از این روند تنها وضعیت نسبی ترافیک در بخش های مختلف خیابان ها نشان می دهد و قابلیت بیان وضعیت واقعی ترافیک راندارد.



شکل ۴- نحوه خوشه بندی در روش اول



شکل -۵ نحوه خوشه بندی در روش دوم

۵-۱- خوشه بندی زمانی

با توجه به اینکه معضل ترافیک یک پدیده مکانی-زمانی می‌باشد در تحلیل‌های داده‌کاوی جهت پی بردن به الگوهای ترافیکی نیاز به انجام تحلیل‌های زمانی نیز حس می‌گردد در این تحقیق جهت انجام تحلیل زمانی همانند تحلیل‌های مکانی از تحلیل خوشه بندی استفاده شده است بدین صورت که خوشه های زمانی را در دوره های یک روزه و یک ساعته استخراج شده است وسپس با ترکیب خوشه بندی های زمانی و خوشه بندی های مکانی به یک سری خوشه بندی مکانی-زمانی حاصل می شود.

۶- نتایج بدست آمده از خوشه بندی مکانی-زمانی (تحلیل داده کاوی مکانی-زمانی)

همانطورکه در قسمت ۳-۱و۳-۲ ذکر گردید در بخش بندی مکانی، حوزه مکان را به دو صورت خوشه بندی گردیده است در روش اول هر خیابان را یک خوشه مکانی در نظر گرفته شده است و در روش دوم تمامی خیابان ها را به خوشه های مساوی تقسیم شده اند. در جداول زیر حجم داده های موجود در هر خوشه مکانی نشان داده شده است. این روش یکی از روش های رایج در بررسی و محاسبه حجم ترافیک در تحقیقات پیشین بوده است که دارای یک مشکل اساسی می باشد بدین صورت که با توجه به اینکه خیابان ها عموما از یک نوع نیستند و دارای طول و عرض های مختلفی می باشند بر همین اساس قطعا حجم داده ترافیکی موجود در خوشه‌ها نمی تواند بیانگر حجم و وضعیت واقعی حجم ترافیک در هر خیابان باشد. بنابرین نیاز به پارامتر دیگری برای ارزیابی وضعیت ترافیکی هر خوشه احساس می‌گردد در این پژوهش حجم داده در هر متر مربع را بعنوان پارامتر موثر در نظر گرفته شده است. حال به کمک این پارامترها می توان وضعیت نسبی ترافیکی هر خوشه را نسبت به خوشه های دیگر بررسی کرد.

۶-۱- نتایج حاصل از خوشه بندی مکانی-زمانی روش اول

همانطور که در بخش قبل ذکر گردید در روش اول محدوده مکانی به ۱۰ خوشه مکانی تقسیم شده است به طوری که هر خیابان یک خوشه مکانی فرض شده است همچنین بازه زمانی موجود نیز به ۷ خوشه زمانی تقسیم شده است در خوشه بندی زمانی هر روز یک خوشه زمانی فرض شده است. بنابرین در روش اول ۷۰ خوشه مکانی-زمانی حاصل می‌شود.

جدول۲- حجم داده در هرخیابان،هرروز(تحلیل نوع اول)

	روزاول	روزدوم	روزسوم	روزچهارم	روزپنجم	روزششم	روزهفتم
خیابان ۱	۸۳۲	۸۰۵	۱۱۶۹	۶۷۰	۳۸۲	۳۸۰	۲۲۱
خیابان۲	۵۴۸	۱۴۴۹	۱۶۲۱	۸۶۸	۱۳۲۲	۸۴۹	۱۰۹۶
خیابان۳	۱۹۵	۱۰۲۷	766	521	562	733	733
خیابان۴	1184	677	337	542	224	333	673
خیابان۵	1912	7416	11665	2299	728	1141	775
خیابان۶	374	993	1179	1120	449	500	679
خیابان۷	247	688	707	589	438	373	393
خیابان۸	103	425	432	502	391	292	337
خیابان۹	62	466	286	307	191	107	133
خیابان۱۰	123	386	283	445	336	197	467

جدول۳- چگالی در هر خیابان هرروز

	روز اول	روز دوم	روز سوم	روز چهارم	روز پنجم	روز ششم	روز هفتم
خیابان ۱	0.013797	0.013349	0.019385	0.01111	0.006334	0.006301	0.003665
خیابان ۲	0.006192	0.016373	0.018316	0.009808	0.014938	0.009593	0.012384
خیابان ۳	0.002636	0.013882	0.010354	0.007042	0.007597	0.009908	0.009908
خیابان ۴	0.049309	0.028194	0.014035	0.022572	0.009329	0.013868	0.028028
خیابان ۵	0.008449	0.032769	0.051545	0.010159	0.003217	0.005042	0.003425
خیابان ۶	0.006068	0.016111	0.019129	0.018171	0.007285	0.008112	0.011016
خیابان ۷	0.005574	0.015527	0.015956	0.013293	0.009885	0.008418	0.008869
خیابان ۸	0.002721	0.011226	0.01141	0.013259	0.010328	0.007713	0.008901
خیابان ۹	0.002903	0.021816	0.01339	0.014373	0.008942	0.005009	0.006227
خیابان ۱۰	0.006065	0.019034	0.013955	0.021943	0.016568	0.009714	0.023028

همانطور که در جدول شماره ۲و ۳ مشاهده می گردد حجم داده موجود در هر خیابان نمی تواند بیانگر وضعیت نسبی ترافیک ان خیابان را نسبت به خیابان های دیگر باشدبنابرین در این تحقیق چگالی داده ها در هر متر مربع را به عنوان پارامتر موثر در نظر گرفته شده است به عنوان مثال مطابق با جدول۲ خیابان ۷ در روزپنجم حجم ترافیکی بالاتری نسبت به خیابان ۸ دارد ولی مطابق با جدول ۳ دارای چگالی کمتری نسبت به این خیابان دارد.

رابطه ۱

$$\text{چگالی} = \frac{\text{حجم داده}}{\text{مساحت}}$$

مطابق با رابطه ۱، بدلیل اینکه مساحت یک خیابان ثابت است بنابرین در یک خیابان جهت بررسی وضعیت نسبی ترافیک می توان از پارامتر حجم داده به جای پارامتر چگالی استفاده نمود ولی هنگام بررسی وضعیت نسبی ترافیکی خیابان های متفاوت نسبت به یکدیگر نمی توان از پارامتر حجم استفاده کرد و باید از پارمتری استفاده شود که در

خیابان های مختلف وضعیت یکسانی داشته باشد براین اساس پارامتر چگالی که بیانگر میـزان حجـم داده ترافیکـی در متر مربع می باشد معرفی گردیده و برای خوشه های مکانی-زمانی مختلـف محاسـبه شـده اسـت. حـال براسـاس ایـن پارامتر، خوشه مکانی-زمانی با چگالی بالاتر دارای وضعیت ترافیکی سنگین تری نسبت به دیگر خوشه ها دارد.

۶-۲- نتایج حاصل از خوشه بندی مکانی-زمانی روش دوم

در روش دوم جهت مطالعه دقیق تر و موضعی تر از خوشه های کوچکتر استفاده شده است در ایـن روش خوشـه هـای مکانی و زمانی به خوشه های کوچکتری تقسیم شده است و هدف بررسی وضعیت ترافیکی بخـش هـای مختلـف یـک خیابان نسبت به یکدیگر و خوشه های خیابان های دیگر می باشد در این روش خوشه های مکانی طوری در نظر گرفته شده است در طول هر خیابان به خوشه های مکانی نزدیک به ۵۰۰ متر نزدیک باشد که کل محدوده مورد مطالعه ایـن تحقیق به ۴۷ خوشه مکانی تقسیم شده است در خوشه بندی زمانی جهت این روش نیز خوشه هر روز بـه ۲۴ خوشـه زمانی تقسیم شده است که باتوجه اینکه زمان برداشت اطلاعات ۱۴۷ ساعت بوده است این بازه زمانی بـه ۱۴۷ خوشـه زمانی تقسیم شده است بنابرین ۶۹۰۹خوشه‌مکانی-زمانی در روش دوم وجود دارد. در این روش نیز از پـارامتر چگـالی به عنوان پارامتر معرف وضعیت نسبی ترافیک استفاده شده است.نتایج این روش در جداول زیر مشاهده می‌گردد.

	روز اول	روز دوم	روز سوم	روز چهارم	روز پنجم	روز ششم	روز هفتم
..	.	.	.	.			.
خوشه۶۹۰۷	–	–	–	–	–	–	–
خوشه۶۹۰۸	–	–	–	–	–	–	–
خوشه۶۹۰۹	–	–	–	–	–	–	–

۷- نتایج

تحقیق حاضر با هدف استخراج الگو های ترافیکی یک محیط شهری انجام پذیرفته است کـه از پایگـاه داده عظـیم کـه شامل ۱۵میلیون رکورد داده ترافیک متعلق به شهر پکن بوده و استفاده شده است. در این تحقیـق از دو روش خوشـه بندی استفاده شده است در این دو روش که از فضاهای مختلف مکانی و زمانی به عنوان خوشه استفاده شـده اسـت در این تحقیق نشان داده شد که هر چه فضای مورد مطالعه جهت خوشه بندی کوچکتر باشد می تـوان بـه الگـوی دقیـق تری از وضعیت نسبی ترافیک در بعد مکـان و زمـان رسـید در ایـن تحقیـق در روش دوم دیـد کلـی تـری از توزیـع و پراکندگی ترافیک در نقاط مختلف یک خیابان و در زمان های مختلف می دهد. همچنین از نتایج حاصل شـده از روش دوم باتوجه به وجود توالی در خوشه های زمانی و دوره تکرار ۲۴ ساعته می توان به عنوان یک الگوی ثابتی از وضـعیت نسبی ترافیک دانست.

همانطور که در بخش های قبل ذکر گردید حجم موجود در خوشه را نمی توان به عنوان پارامتر معرف وضعیت ترافیک دانست زیرا خیابان ها از انواع مختلفی تشکیل شده و متناسب با نوع آنهـا سـرعت مجـاز متفـاوت مـی باشـد از طرفـی خیابان ها دارای طول و عرض مختلفی هستند بدین جهت پارامتر چگالی تعریف گردید کـه نتـایج حاصـل از وضـعیت ترافیکی با استفاده از این پارامترمتفاوت از نتایج از حجم داده ترافیکی می باشد بنابرین در این تحقیق در نظـر گـرفتن چگالی به عنوان پارامتر معرف وضعیت نسبی ترافیک بسیاری ازنارسایی های مدل قبل را برطرف کرده است.

منابع

1. Goodchild M F, Citizens as Sensor: The World of Volunteered Geography. Journal of Geography, Vol.69, No. 4, PP. 211-221,2007.

2. Miller, H., Han, J., Geographic Data Mining and Knowledge Discovery: An Overveiw. Miller, H., Han, J., Geographic Data Mining and Knowledge Discovery.CRC Press,Taylor and Francis Group. PP.1-26, 2009.

3. Andrienko, G., Andrienko, N., Data Mining With C4.5 and Interactive cartographic visualization. In N. W. G. T. Paton (Ed), User Interfaces to Data Intensive Seystem.Los Alamitos, CA: IEEE Computer Society. PP.162-165, 1999.

4. Chawla, S., Shekhar, S., Wu, W., Ozesmi, U., Extending Data Mining for Spatial APP : A Case Study In Predicting Nest Locations. In ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, Dallas, TX, 2000.

5. Gahegan, M., Is Inductive Machine Learning Just Another Wild Goose (or might it lay the golden egg)? International Journal of Geographical Information Science, Vol. 17, No. 1, PP. 69-92, 2003.

6. Guo, D.,Peuquet, D., Gahegen, M., ICEAGE: Interactive Clustering and Exploration of Large and High-Dimensional Geodata. Geoinformatica, Vol. 7, No. 3, PP. 229-253, 2003.

7. Guo, D.S., Chen, J., MacEachern, A.M., Liao, k., A Visualization System for Space-Time and Multivariate Patterns (VIS-STAMP). IEEE Transaction on Visualization and Computer Graphics, Vol. 12, No. 6, PP .1461-1474, 2006.

8. Maimon, O., Rokach, L., Data Mining and Knowledge Discovery Handbook, 2nd edition, Springer Science Business Media, 2010.

9. Miller, H,J., Han, J., Geographic Data Mining and Knowledge Discovery, An Overview, in Geographic Data Mining and Knowledge Discovery, Miller, H.J., Han, J., Taylor Francis .ISBN 0-415-23369-0, 2001.

10. Ntoutsi, I., Mitsou, N., Marketos, G., Traffic Mining in a Road-Network: How does the Traffic Fow?, Int .J. Business Intelligence and Data Mining, Vol. 3, No. 1, PP. 82-98,2008.