نشریه علمی پژوهشی مهندسی فناوری اطلاعات مکانی

سال ششم، شماره سوم، پاییز ۱۳۹۷

Vol.6, No.3, Autumn 2018

۱ -۱۳

شناسایی کاربری در محیط‌های شهری بر مبنای رفتار کاربران در شبکه‌های اجتماعی مکان‌مبنا

مرتضی طیبی[۱]، فرید کریمی‌پور[۲*]

۱- دانشجوی دکتری سیستم‌های اطلاعات مکانی– دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی– پردیس دانشکده‌های فنی - دانشگاه تهران

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

تاریخ دریافت مقاله: ۱۳۹۶/۰۹/۲۵ تاریخ پذیرش مقاله: ۱۳۹۶/۱۱/۲۸

چکیده

امروزه پس از گذشت زمان نسبتاً کوتاهی از ظهور شبکه‌های اجتماعی مکان‌مبنا، این رسانه‌های اجتماعی به یک عنصر جدایی‌ناپذیر از زندگی اجتماعی به‌خصوص در شهرهای بزرگ تبدیل شده‌اند. افزایش محبوبیت شبکه‌های اجتماعی، پیشرفت فناوری‌های تعیین موقعیت مکانی از جمله سیستم تعیین موقعیت جهانی(GPS) و هم‌چنین فراگیر شدن گوشی‌های تلفن هوشمند که علاوه بر مجهز بودن به *GPS* به‌راحتی به شبکه جهانی وب متصل می‌شوند، باعث تولید انبوه داده‌های مکانی به‌روز، کم‌هزینه و ارزشمند شده است. در کنار کاربردهای متعدد، از این داده‌ها می‌توان در راستای برطرف نمودن چالش‌های شهری سود جست. به عنوان نمونه‌ای از این کاربردها، در این تحقیق از رفتار کاربران در شبکه‌های اجتماعی مکان‌مبنا جهت تعیین کاربری اجتماعی در محیط‌های شهری استفاده شده است. در این راستا ابتدا با خوشه‌بندی داده‌های مکانی کاربران به تشخیص مناطق در محیط‌های شهری پرداخته شده، و سپس ضمن استخراج الگوهای تغییرات تعداد داده‌های کاربران و جابجایی شهروندان در مناطق شهری در شبکه‌های اجتماعی در طول شبانه‌روز، از این الگوها به عنوان شاخص تعیین کاربری استفاده شده و به هر منطقه، یک نوع کاربری اجتماعی نسبت داده می‌شود. به منظور ارزیابی روش پیشنهادی ماتریس در هم گسیختگی بر مبنای حضور مکان‌های ثبت شده با کاربری معین در مناطق شناسایی شده تشکیل شد و مقدار دقت کلی و شاخص کاپا به ترتیب ۷۹ و ۷۱ درصد محاسبه گردید. همچنین مقایسه نتایج حاصل با تصاویر ماهواره‌ای نشان‌دهنده پتانسیل بالای شبکه‌های اجتماعی مکان‌مبنا برای شناخت محیط‌های شهری و حل چالش‌های موجود است.

کلید واژه‌ها : شبکه اجتماعی مکان‌مبنا، خوشه‌بندی، کاربری اجتماعی.

* نویسنده مکاتبه کننده: خیابان کارگر شمالی، بالاتر از جلال آل احمد، پردیس دانشکده‌های فنی دانشگاه تهران، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی

تلفن: ۶۱۱۱۴۳۷۶-۰۲۱

Email: fkarimipour@ut.ac.ir

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.6.3.1]

۱- مقدمه

شهرها دستگاه‌های پیچیده‌ای هستند که برای چندین دهه، ساختار فضایی آن‌ها توسط محققان شهری به‌منظور توصیف ساختار، نقش و کارکرد آن‌ها در زندگی مردم مورد مطالعه قرارگرفته است. یکی از ارکان مهم در شکل‌گیری ساختار شهری، چگونگی تقسیم‌بندی منطقه‌ای داخل شهرها و مشخص کردن کاربری هر منطقه است. منطقه‌بندی جغرافیایی شهرها به واحدهای ارضی منسجم و پیوسته و تعیین کاربری هر یک نه‌تنها برای مدیریت و توزیع مسئولیت‌های اداری و تخصیص منابع عمومی، ضروری است بلکه به‌عنوان یک چهارچوب مهم مرجع برای درک انواع پدیده‌های مربوط به فعالیت‌های انسانی به‌کار می‌رود. ازاین‌رو امروزه یکی از موضوعات مهم برای سیاست‌گذاران و برنامه‌ریزان شهری، تعریف کاربری شهرهاست. تعیین کاربری شهری در کنار موضوعات دیگری نظیر مسائل ترافیکی، آلودگی و غیره به‌عنوان هدف یک مفهوم کلی‌تر به نام محاسبات شهری[1] قرار می‌گیرد، که به‌صورت دریافت، تجمیع و پردازش داده‌های عظیم و گوناگون تولید شده از منابع مختلف موجود در فضای شهری به‌منظور حل اصلی‌ترین معضلات موجود در محیط‌های شهری تعریف می‌شود [۱]. داده‌های عظیم مورد استفاده در محاسبات شهری در نتیجه استفاده انبوه شهروندان از امکاناتی نظیر سیستم تعیین موقعیت جهانی (*GPS*) خودروها، تاکسی‌ها، اتوبوس‌ها، کارت‌های اعتباری، تلفن‌های همراه، شبکه‌های حمل‌ونقل زیر زمینی و غیره حاصل می‌شود. از میان این مشاهدات حرکتی نوین، داده‌های شبکه‌های اجتماعی مکان‌مبنا علاوه بر داشتن حجم بالا، توزیع مکانی و زمانی مناسب، دقت کافی و هزینه پایین برخلاف داده‌های گوشی همراه مشکلات امنیت و حفظ اطلاعات شخصی کاربران را نیز ندارند و مجموعه‌ای قابل اعتماد، گران‌بها و عظیم از داده‌های خام را در اختیار قرار می‌دهند. در این تحقیق به دنبال تعیین کاربری محیط‌های شهری بر مبنای داده‌های شبکه‌های اجتماعی مکان‌مبنا هستیم. در این راستا، محاسبات شهری را می‌توان به‌عنوان یکی از زمینه‌های کاربردی شبکه‌های اجتماعی مکان‌مبنا مطرح کرد [۲].

به‌صورت کلی یک شبکه‌ی اجتماعی، بازسازی‌کننده‌ی روابط اجتماعی افراد در محیط‌هایی مانند یک وب سایت بوده و راه‌هایی را برای کاربران برای به اشتراک‌گذاری ایده‌ها، فعالیت‌ها، اتفاقات و علائق بر روی بستر اینترنت فراهم می‌کند. ظهور و فراگیر شدن گوشی‌های هوشمند و مجهز به *GPS* که به شبکه جهانی وب دسترسی دارند، موجب پیدایش شبکه‌های اجتماعی مکان‌مبنا شد. شبکه‌های اجتماعی مکان‌مبنا، افراد را قادر ساخته است تا بعد مکانی را نیز به روش‌های مختلف، به شبکه‌های اجتماعی موجود اضافه کنند. برای مثال کاربرها می‌توانند عکس‌های خود را با مکان نشانه‌گذاری کرده و در شبکه‌های اجتماعی منتشر کنند(فلیکر[2])، یا نظرشان را در مورد یک رویداد و زمانی که در محل آن رخداد هستند بیان کنند(توییتر[3]). یا هر لحظه دوستانشان را از مکان خود مطلع کنند (فوراسکوئر[4]) [۳].

باوجود آنکه زمان زیادی از پیدایش مفهوم شبکه‌های اجتماعی مکان‌مبنا نمی‌گذرد، اما هم در زمینه تجاری و هم در زمینه تحقیقاتی مورد توجه زیادی قرارگرفته و رشد چشمگیری داشته است. یکی از جامع‌ترین تحقیقات را در زمینه‌ی شبکه‌های اجتماعی مکان‌مبنا ژنگ و همکاران در بخش تحقیقات شرکت مایکروسافت انجام داده‌اند. از اولین تحقیقات این گروه که در سال ۲۰۰۸ ارائه شد می‌توان به

[1] Urban Computing

[2] Flicker
[3] Twitter
[4] Foursquar

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.6.3.1]

محاسبه میزان مشابهت کاربران بر اساس پیشینه مکانی اشاره کرد [۴]؛ اما به‌عنوان نقطه‌ی عطف تحقیقات انجام شده توسط این گروه می‌توان به‌پروژه‌ی ژئولایف[1] در سال ۲۰۰۹ اشاره کرد. هدف این پروژه ارائه یک شبکه اجتماعی مکان‌مبنا بود که بر مبنای داده‌های خط السیر *GPS* که در شبکه به اشتراک گذاشته می‌شد عمل نماید. از مهم‌ترین بخش‌های این سیستم می‌توان به مدل کردن داده‌های خط السیر کاربران در یک چارچوب سلسله مراتبی، کاوش جذاب‌ترین مکان‌ها در یک محدوده جغرافیایی و برنامه‌ریزی سفر برای کاربران و ارائه سرویس توصیه‌گر فردگرا اشاره کرد [۵]. یکی از اولین تحقیقات بزرگ مقیاس انجام شده بر روی شبکه‌های اجتماعی مکان‌مبنا تحقیق لی و همکاران در سال ۲۰۰۹ است که در آن برای مطالعه چگونگی به اشتراک‌گذاری داده‌های مکانی توسط کاربران، داده‌های زیادی از شبکه‌ی اجتماعی مکان‌مبنای برایت کیت جمع‌آوری‌شده و نتایج تحلیل این داده‌ها از نظر ویژگی‌های کاربران، مشخصات حرکتی کاربران و ارتباط آن‌ها ارائه شده است [۶]. در تحقیق دیگری اسکلاتو و همکاران در سال ۲۰۱۰ با تحلیل داده چهار شبکه اجتماعی مکان‌مبنای مختلف و با استفاده از یک معیار ابتکاری فرضیه وابستگی ساختار اجتماعی کاربران به وضعیت جغرافیایی آن‌ها را بررسی کردند [۷].

در یکی دیگر از تحقیقات انجام شده، کرانشو و همکاران در سال ۲۰۱۲ با استفاده از داده‌های شبکه‌های اجتماعی مکان‌مبنا و نیز با استفاده از الگوریتم خوشه‌بندی به شناسایی مناطق شهری تحت عنوان لایوهودز[2] پرداخته‌اند [۸]. همچنین در سال ۲۰۱۷ لیو و همکاران با تلفیق داده‌های سنجش از دور و داده‌های شبکه‌های اجتماعی و با به‌کارگیری الگوریتم طبقه‌بندی *SVM*[3] موفق به استخراج مناطق با سطوح ترافیکی مختلف شدند [۹].

در این تحقیق با بهره‌گیری از داده‌های مکانی کاربران در شبکه‌های اجتماعی، به طبقه‌بندی کاربری در مناطق شهری می‌پردازیم. تحقیقاتی پیش از این در این زمینه صورت گرفته است. به‌عنوان مثال، در سال ۲۰۱۱، نوآلاس و همکاران با استفاده از داده‌های شبکه فوراسکوئر برای دو شهر لندن و نیویورک کاربری شهری را تعیین کردند. در این تحقیق از داده‌های معنایی موجود در داده‌های این شبکه که نوع فعالیت کاربران را مشخص می‌کند، بهره‌گیری شد [۱۰]. همچنین در سال ۲۰۱۳، هاسان و همکاران با استفاده از داده‌های شبکه توییتر و با بهره‌گیری از بعد معنایی داده‌ها به تعیین کاربری محیط شهری پرداختند. آن‌ها با استفاده از تابع چگالی کرنل و برای شش کاربری مختلف و در بازه‌های سه ساعته شبانه‌روز، چگالی هر کاربری در سطح محیط شهری برآورد کردند [۱۱]. از جمله نقاط ضعفی که مطالعات گذشته داشته‌اند؛ استفاده از افراز منظم محیط شهری توسط شبکه سلولی است که با توجه به توزیع خوشه‌ای داده‌های شبکه‌های اجتماعی نمی‌تواند مناسب باشد. هم چنین رزولوشن زمانی پایین و ارزیابی غیر کامل روش‌های پیشنهادی از دیگر مشکلات تحقیقات گذشته است که در این تحقیق برطرف می‌شود.

در ادامه این مقاله، در بخش دوم داده‌های مورد استفاده معرفی می‌شود. در بخش سوم به بیان روش انجام تحقیق و ارائه مدل پیشنهادی برای تعیین کاربری‌ها در محیط‌های شهری می‌پردازیم. در بخش چهارم، مدل پیشنهادی برای داده‌های مکانی کاربران شبکه اجتماعی توییتر در شهر توکیو پیاده‌سازی شده و نتایج آن در بخش پنجم مورد ارزیابی قرار خواهد گرفت. درنهایت، در بخش ششم به بحث، نتیجه‌گیری و

[1] Geolife
[2] Livehoods
[3] Support Vector Machine

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.6.3.1]

ارائه پیشنهادهایی برای تحقیقات آینده پرداخته خواهد شد.

۲- مواد و روش‌ها

در این تحقیق از داده‌های شبکه اجتماعی توییتر استفاده‌شده است .در این شبکه کاربران می‌توانند پیام‌هایی متنی با حداکثر ۱۴۰ کاراکتر و یا پیام‌هایی در قالب عکس را به اشتراک بگذارند. هم‌چنین در صورت تمایل می‌توانند موقعیت مکانی خود را نیز به پیام ضمیمه کنند. به هریک از این پیام‌ها اصطلاح توییت[1] اختصاص داده شده است.

در ژوئن ۲۰۱۱ تعداد کاربران این وب‌گاه از مرز ۲۰۰ میلیون نفر گذشت و تعداد ۲۰۰ میلیون توییت در روز به ثبت رسید. این عدد برابر با نوشتن ده میلیون صفحه کتاب در روز است، خواندن این تعداد از توییت حدود ۳۱ سال طول می‌کشد. در سه‌ماهه اول سال ۲۰۰۷، ۴۰۰,۰۰۰ توییت ارسال شد، این رقم در سه‌ماهه نخست سال ۲۰۰۸ به ۱۰۰ میلیون توییت رسید و در فوریه ۲۰۱۰ کاربران توییتر روزانه ۵۰ میلیون توییت در روز به این وب سایت فرستادند. در ژوئن ۲۰۱۰ در حدود ۶۵ میلیون توییت در روز که برابر است با فرستادن ۷۵۰ توییت در هر ثانیه به توییتر، ارسال شد. طبق آخرین آمار موجود در سال ۲۰۱۵ توییتر دارای بیش از ۲۸۲ میلیون کاربر فعال است. به‌تعبیر جالب دیگری اگر توییتر با تعداد کاربران موجود خود به‌عنوان کشوری اعلام استقلال کند، دوازدهمین کشور پر جمعیت جهان خواهد بود. تعداد توییت‌های ارسالی توسط این کاربران در روز حدود ۵۰۰ میلیون توییت است. ۶۳٪ از برندهای دنیا حساب کاربری توییتر گوناگون دارند و حدود ۹۲٪ این تعداد بیش از یک توییت در روز ارسال می‌کنند که نشان‌دهنده‌ی فعالیت آن‌ها و موفقیت توییتر در جذب بازاریابان و فعالان حوزه بازاریابی و فروش است.

آمارهای فوق نشان‌دهنده‌ی اهمیت و پتانسیل بالای داده‌های شبکه اجتماعی به‌ویژه شبکه توییتر در راستای حل معضلات محیط‌های شهری و کمک به مدیران شهری برای پیش برد اهداف خود می‌باشد. به توییت‌هایی که دارای بعد مکانی باشند توییت مکان-مبنا[2] گفته می‌شود. جمع‌آوری این داده‌ها از طریق سرویس باز *API*[3] توییتر [۱۲] امکان‌پذیر است. مجموعه داده گردآوری شده می‌تواند شامل شناسه کاربری*(ID)*، موقعیت مکانی برحسب طول و عرض جغرافیایی، زمان ارسال پیام شامل روز و ساعت، پیام متنی و نام کاربر باشد. به دلیل محدودیت‌های موجود برای دریافت داده، تنها می‌توان داده‌هایی در یک دایره به مرکز معین و شعاع مشخص‌شده دریافت کرد.

۳- روش اجرایی

روش پیشنهادی این تحقیق برای تعیین کاربری در محیط‌های شهری بر مبنای داده‌های شبکه‌های اجتماعی مکان‌مبنا توییتر در فلوچارت شکل (۱) نشان داده شده است. به‌طور کلی روش پیشنهادی شامل سه مرحله است: ۱) مراحل پیش‌پردازش و حذف نویز ۲) تفکیک مناطق شهری ۳) تعیین کاربری هر منطقه.

۳-۱- پیش‌پردازش

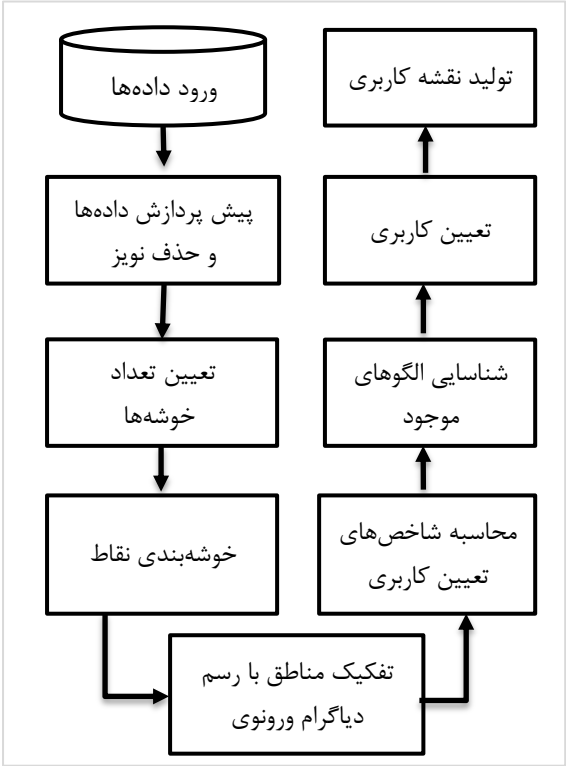
داده‌های جمع‌آوری شده از رابط برنامه‌نویسی برنامه (*API*[4]) توییتر ممکن است دارای نواقص و مشکلاتی باشد. برای رسیدن به داده‌های آماده برای پردازش در اولین قدم با تعریف یک مرز برای شهر یا منطقه موردنظر، داده‌های خارج از مرز حذف می‌شوند. هم‌چنین می‌بایست از داده‌های فاقد بعد مکانی صرف‌نظر شود. در اقدام بعدی ممکن است برخی از

[1] Twitt

[2] Geo-tagged Twitt
[3] Twitter open API
[4] Application Programming Interface

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.6.3.1]

داده‌ها در مناطقی باشند که دارای بافت شهری نیست، به‌عنوان مثال داده‌های موجود در حاشیه شهرها. این داده‌ها اغلب کم تعداد بوده و چگالی ناچیزی دارند. به‌منظور حذف داده‌های کم‌چگالی که در خارج از مناطق شهری هستند می‌توان از روش خوشه‌بندی *DBSCAN*[1] که یک روش چگالی مبنا با کاربرد شناسایی نویز است، استفاده کرد.



شکل ۱: فلوچارت روش پیشنهادی تحقیق

الگوریتم *DBSCAN* به دنبال اشیائی است که دارای همسایه‌های متراکم هستند و این اشیا را با نام اشیا هسته می‌شناسد. پس‌ازآن با اتصال اشیا هسته و همسایه‌های آن‌ها مناطق متراکم به‌عنوان خوشه معرفی می‌شود. این روش دو پارامتر حداقل نقطه و شعاع همسایگی را به‌عنوان ورودی دریافت می‌کند. سپس داده‌هایی را که تعداد همسایه‌های آن در شعاع همسایگی موردنظر بیشتر از حداقل نقطه تعریف‌شده است به‌عنوان شی هسته در نظر می‌گیرد؛ آنگاه نقاط همسایه و سایر داده‌هایی که با پارامترهای مذکور در حالت اتصال به این شی هسته هستند را در یک خوشه قرار می‌دهد. در پایان الگوریتم داده‌هایی که در هیچ خوشه‌ای قرار نگرفته‌اند به‌عنوان نویز شناسایی می‌شود [۱۳]. بنابراین با استفاده از این روش خوشه‌بندی می‌توان داده‌هایی که به شکل معناداری دورتر از سایرین هستند را به‌عنوان نویز شناسایی و حذف کرد.

۳-۲- تفکیک مناطق

به‌منظور شناخت محیط‌های شهری و سنجش اجتماعی، در مرحله اول باید محیط شهری به بخش‌هایی افراز شود تا مطالعات بعدی بر روی هر کدام از این بخش‌ها انجام شود [۱۴]. به‌منظور افراز محیط‌های شهری در سنجش اجتماعی می‌توان از روش‌های مختلفی استفاده کرد. الف) می‌توان از شبکه اصلی راه‌های منطقه شهری به‌عنوان مرز بخش‌های مختلف استفاده کرد [۱۵]. ب) همچنین می‌توان محیط شهری را با یک شبکه منظم سلولی پوشش داد و از این سلول‌ها به‌عنوان واحد مکانی در سنجش اجتماعی بهره گرفت [۱۶]. ج) رویکرد بعدی می‌تواند استفاده از تقسیمات سیاسی و اداری شهرها باشد [۱۷]. د) در روشی دیگر ابتدا داده‌های مکانی به تعداد معینی خوشه تقسیم می‌شوند و سپس با رسم دیاگرام ورونوی[2] بر روی مراکز این خوشه‌ها محیط شهری به بخش‌هایی افراز خواهد شد و از هر چندضلعی ورونوی می‌توان به‌عنوان واحد مکانی در سنجش اجتماعی بهره گرفت [۱۸].

هر کدام از روش‌های مذکور می‌تواند در کاربرد خاصی موفق باشند، در این تحقیق از دیاگرام ورونوی به‌منظور افراز محیط شهری استفاده شده است. چرا که مکان‌های مورد علاقه کاربران در شبکه‌های اجتماعی

[1] Density-Based Spatial Clustering of Applications with Noise

[2] Voronoi Diagram

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.6.3.1]

مکان‌مبنا مانند رستوران ها، محل‌های تفریحی، تجاری و غیره دارای تراکم مکانی هستند و بنابراین انتظار می‌رود که محل‌های با کاربری یکسان در یک خوشه قرار بگیرند و هر خوشه دارای نقش اجتماعی مستقلی باشد.

برای این منظور پس از تعیین تعداد خوشه‌ها، با استفاده از روش خوشه‌بندی *K-means* داده‌های مکانی کاربران به چند خوشه تقسیم می‌شود. این روش از جمله روش‌های پایه خوشه‌بندی است و نقاطی را که به لحاظ مکانی به یکدیگر نزدیک‌تر هستند در یک خوشه قرار می‌دهد [۱۹]. پس از اعمال الگوریتم خوشه‌بندی می‌توان از مراکز خوشه به‌عنوان نقاط نمونه استفاده کرد و با ترسیم دیاگرام ورونوی برای این نقاط منطقه مد نظر را به چند زیر منطقه افراز کرد.

نکته اصلی در این مرحله یافتن تعداد خوشه‌ها است. به این منظور در این تحقیق از روش رایجی به نام نمودار البو (*elbow*) استفاده شده است. روش *elbow* بر این واقعیت بنا نهاده شده است که افزایش تعداد خوشه‌ها می‌تواند به کاهش مجموع واریانس درون خوشه‌ای کمک کند[۲۰]. موضوع از این قرار است که تعداد خوشه‌های بیشتر باعث می‌شود تا خوشه‌ها کوچک‌تر شوند و در نتیجه عضوهای هر خوشه به یکدیگر نزدیک‌تر و شبیه‌تر می‌شوند. اما اگر ما با تعداد خوشه‌های زیادی روبرو باشیم، به دلیل شکستن یک خوشه‌ی منسجم به دو خوشه، تنها کاهش اندکی را در واریانس خواهیم داشت و به همین دلیل کاهش مجموع واریانس‌ها محسوس نیست، در نتیجه می‌توان منحنی مجموع واریانس‌های درون خوشه‌ای نسبت به تعداد خوشه‌ها را رسم کرد و نقطه‌ای از منحنی که پس از آن تغییر محسوسی وجود ندارد را به‌عنوان تعداد خوشه‌های مناسب در نظر بگیریم.

۳-۳- تعیین کاربری

برای تعیین کاربری، داده‌های موجود در هر منطقه را در بازه‌های ۶ ساعته از شبانه‌روز (۰-۶ و ۶-۱۲ و ۱۲-۱۸ و ۱۸-۲۴) تقسیم نموده و دو شاخص زیر را معرفی می‌کنیم:

الف) تغییرات صورت گرفته در تعداد داده‌ها در هر بازه نسبت به بازه قبلی: اگر f_1, f_2, f_3, f_4 به ترتیب بیانگر تعداد پیام‌های توییتر در ربع اول تا ربع چهارم شبانه‌روز باشند، می‌توان از اختلاف آن‌ها (یعنی f_2-f_1 و f_3-f_2 و f_4-f_3) به‌عنوان شاخص‌هایی که نشان دهنده نوع کاربری اجتماعی مناطق شهری است، سود جست.

ب) حرکت شهروندان بین مناطق شناسایی شده: به‌طور کلی سه نوع حرکت برای کاربران قابل تعریف است: ۱) جابجایی کاربر داخل منطقه ۲) ورود کاربر به منطقه ۳) خروج کاربر از منطقه. در راستای یافتن کاربری اجتماعی منطقه و هم‌چنین ساده‌سازی تحلیل‌ها از حرکت نوع اول چشم‌پوشی شد. سپس نسبت تعداد شهروندان وارد شده به تعداد شهروندان خارج شده به عنوان شاخص دوم کاربری اجتماعی طبق رابطه (۱) معرفی می‌شود.

$$movement_index = \frac{Incomings}{Outgoings} \quad \text{رابطه (۱)}$$

۴- مورد مطالعه

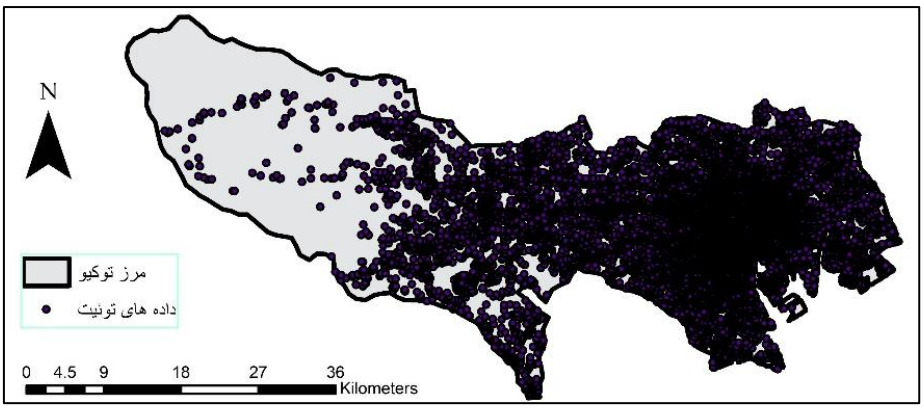
از آنجایی که شبکه‌های اجتماعی در کشورهای توسعه‌یافته مانند ژاپن بیشتر فراگیر هستند و از طرفی شهر توکیو یکی از پرجمعیت‌ترین و متراکم‌ترین شهرهای دنیاست، لذا داده‌های این منطقه می‌تواند کمیت و کیفیت مطلوبی داشته باشد. از همین رو روش پیشنهادی این تحقیق بر روی داده‌های شهر توکیو پیاده‌سازی شد که در ادامه جزئیات آن به‌تفصیل بیان خواهد شد.

۴-۱- جمع‌آوری داده و پیش‌پردازش

در این تحقیق از داده‌هایی که در یک دایره به شعاع ۵۰ کیلومتر و به مرکز شهر توکیو و در طی ۹ روز جمع‌آوری شده است استفاده گردیده است. داده‌های دریافت شده مربوط به اواسط ماه آوریل سال ۲۰۱۶

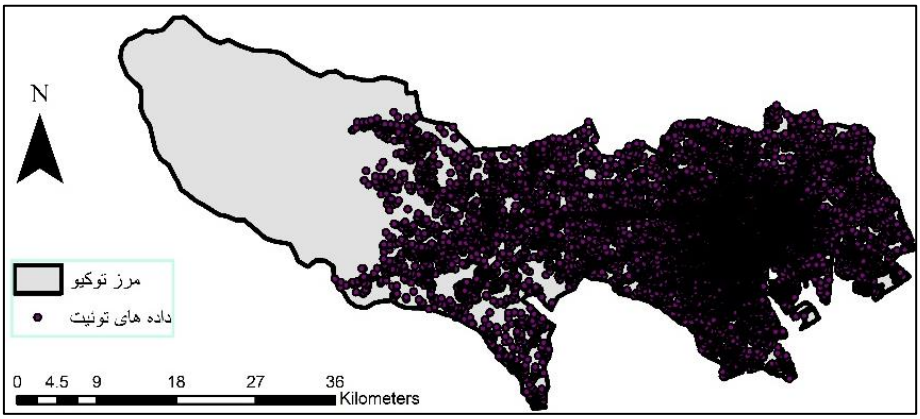
[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.6.3.1]

است و به شکل میانگین بیش از ۲۰۰۰۰ توییت در هر روز جمع‌آوری شده است. در اولین قدم با تعریف یک مرز برای شهر توکیو داده‌های خارج از توکیو حذف شدند و هم‌چنین از داده‌های فاقد بعد مکانی صرف‌نظر شد و در پایان بیش از ۱۳۰۰۰ داده برای هر روز در شهر توکیو به دست آمد. در شکل (۲) توزیع تمام داده‌های موجود در شهر توکیو نشان داده شده است. از آنجا که قسمت‌های غربی شهر توکیو غیرمسکونی بوده و دارای پوشش جنگلی و مرتع هستند، در شکل (۱) دارای توزیع بسیار ناچیز داده هستند.



شکل ۲: توزیع داده‌های مکانی توییتر در شهر توکیو

در شکل (۳) نیز داده‌های آماده جهت پردازش بعد از حذف نویزها که غالباً در قسمت‌های غربی بودند نمایش داده شده است. همان‌طور که پیش از این ذکر شد به‌منظور حذف نویزها از روش خوشه‌بندی *DBSCAN* استفاده گردید.

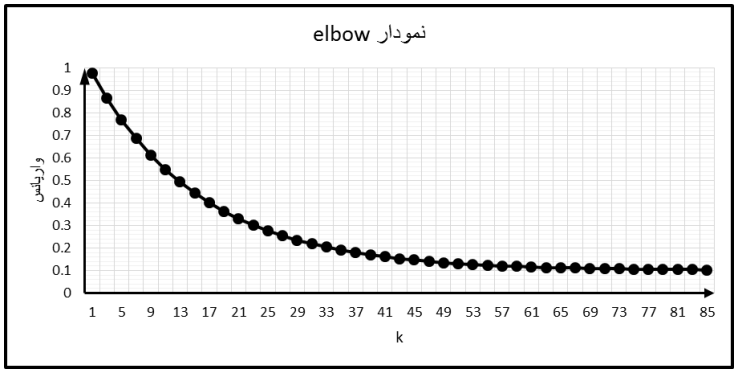


شکل ۳: توزیع داده‌های مکانی توییتر در شهر توکیو پس از حذف نویز

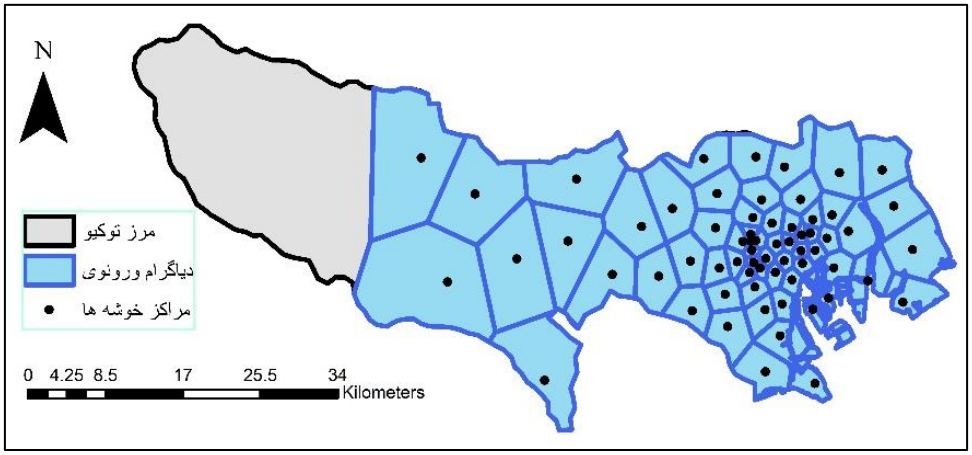
۴-۲- تفکیک مناطق

به‌منظور تفکیک مناطق شهری بایستی در ابتدا تعداد خوشه‌ها معین شود. به این منظور همان‌طور که در شکل (۴) نشان داده‌شده است به کمک نمودار *elbow* تعداد خوشه‌ها برابر ۶۵ مشخص شد، با اعمال کردن الگوریتم خوشه‌بندی *K-means* بر روی مختصات جغرافیایی داده‌های پیش‌پردازش شده، داده‌ها به ۶۵ خوشه تقسیم‌بندی شدند. سپس با به‌کارگیری مراکز خوشه به‌عنوان نقاط نمونه و ترسیم دیاگرام ورونوی، مناطق شهری ترسیم شدند. در شکل (۵) مراکز خوشه‌ها و دیاگرام ورونوی متناظر با آن‌ها نشان داده شده است.

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.6.3.1]



شکل ۴: نمودار elbow برای داده های مورد مطالعه



شکل ۵: مناطق شهری تفکیک‌شده بر مبنای دیاگرام ورونوی مراکز خوشه‌های توزیع مکانی داده‌های کاربران

۴-۳- تعیین کاربری

در این مرحله به‌منظور تعیین کاربری هر یک از مناطق تشخیص داده‌شده در مرحله قبل، دو شاخص معرفی شده در بخش ۳-۳ استخراج گردید:

الف) تعداد داده‌ها در هر یک از ربع‌های شبانه‌روز (۶-۰ = f_1 و ۱۲-۶= f_2 و ۱۸-۱۲= f_3 و ۲۴-۱۸= f_4) محاسبه شد و سپس با محاسبه تغییرات تعداد داده‌ها در ربع‌های مختلف، الگوهای ذکرشده در جدول (۱) حاصل شدند. به‌عنوان مثال الگوی ۱ از جدول (۱) بدان معناست که تعداد داده‌های کاربران در ربع دوم (۱۲-۶) نسبت به ربع اول (۶-۰) افزایش پیدا کرده است و به همین دلیل مقدار $f_2 - f_1$ بزرگ‌تر از صفر است. به شکل مشابه چون تعداد این داده‌ها در ربع سوم (۱۲-۱۸) نسبت به ربع دوم کاهش پیداکرده است مقدار $f_3 - f_2$ منفی است.

هم‌چنین از منفی بودن مقدار $f_4 - f_3$ در الگوی ۱ می‌توان دریافت که تعداد داده‌ها در ربع چهارم (۱۸-۲۴) نسبت به ربع سوم کاهش پیداکرده است.

جدول۱: الگوهای تغییر زمانی فرکانس داده‌های توییتر در طول شبانه‌روز

شماره الگو	$f_2 - f_1$	$f_3 - f_2$	$f_4 - f_3$
۱	+	-	-
۲	+	-	+
۳	+	+	-
۴	+	+	+

ب) شاخص $movement_index$ برای مناطق شناسایی شده، در ربع‌های شبانه‌روز محاسبه گردید. بر اساس دو شاخص معرفی شده به هر منطقه می‌توان یک بردار ۷ مؤلفه‌ای نسبت داد که سه مؤلفه مربوط به تغییرات تعداد داده‌ها و ۴ مؤلفه مربوط به شاخص

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.6.3.1]

$movement_index$ در چهار ربع شبانه‌روز است. از بین حالت زیادی که ممکن بود، حالات محدودی رخ داد که در جدول (۲) نشان داده شده است.

جدول۲: الگوهای کشف شده بر اساس تغییرات تعداد داده‌ها و جابجایی شهروندان

الگو	$movement_index$				تغییر تعداد داده‌ها			تعداد مناطق
	(۶-۰)	(۱۲-۶)	(۱۸-۱۲)	(۲۴-۱۸)	f_2-f_1	f_3-f_2	f_4-f_3	
۱	۱٫۰۲	۲٫۲۷	۰٫۷۴	۰٫۹۰	+	−	−	۳
۲	۰٫۶۷	۰.۹۰	۰٫۹۷	۲٫۴۱	+	−	+	۶
۳	۰٫۹۵	۰٫۶۸	۱٫۷۴	۱٫۲۵	+	+	−	۳۷
۴	۱٫۰۳	۲٫۰۲	۱٫۴۷	۱٫۰۶	+	+	+	۱۹

بر اساس الگوهای کشف شده، چهار نوع کاربری برای مناطق شهری به شرح ذیل تعریف شد:

۱) **اداری (الگوی ۱)**: همان‌طور که در جدول (۲) نشان داده شده است، تعداد داده‌های توییتری در این مناطق پس از افزایش در ربع دوم (۶-۱۲) شبانه‌روز، در ربع سوم (۱۲-۱۸) روند کاهشی پیدا می‌کنند و این روند تا پایان روز ادامه پیدا می‌کند. در مورد جابجایی نیز در ربع اول جابجایی معناداری وجود ندارد، در ربع دوم ورودی‌ها به شکل معناداری بیشتر از خروجی‌ها بوده سپس با پایان ساعت اداری به‌تدریج این مناطق خالی می‌شوند.

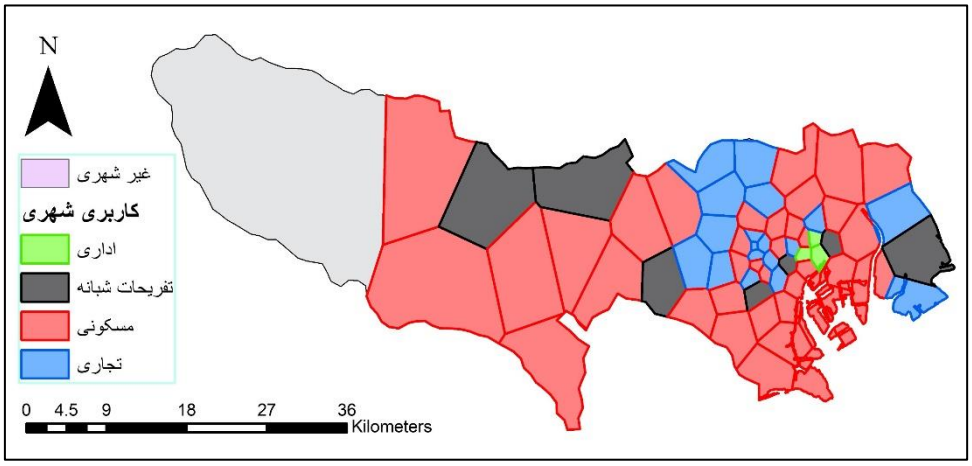
۲) **تفریحات شبانه (الگوی ۲)**: مطابق جدول در این مناطق پس از افزایش ابتدای روز در تعداد داده‌های توییتر، تعداد داده در ربع سوم کاهش پیداکرده و سپس در انتهای روز و در ساعات شب (۱۸-۲۴) افزایش پیدا می‌کند. درواقع مردم در ساعات صبح تا بعدازظهر در این مناطق حضور کمتری دارند و در انتهای روز برای تفریحات شبانه مانند باشگاه، رستوران، شهربازی، کنسرت، تئاتر و غیره به این مناطق سفر می‌کنند و در بازه (۶-۰) این منطقه را تخلیه می‌کنند.

۳) **مسکونی (الگوی ۳)**: در این کاربری نیز مانند سایر موارد ابتدا تعداد داده‌ها در ربع دوم (۶-۱۲) نسبت به ربع اول (۰-۶) افزایش پیدا می‌کند، ربع سوم (۱۲-۱۸) روند افزایشی دارد و با رسیدن به ربع انتهایی شبانه‌روز تعداد داده‌ها کم می‌شود. درواقع مردم در بعد از پایان ساعت کار یا مدرسه و دانشگاه به منازل خود برمی‌گردند و باعث افزایش تعداد داده‌ها می‌شوند و با نزدیک شدن به ساعات پایانی شب و زمان خواب تعداد پیام‌های توییتری کاهش پیدا می‌کند اما بازگشت مردم به مناطق مسکونی همچنان ادامه پیدا می‌کند.

۴) **تجاری و تفریحی (الگوی ۴)**: در این کاربری در طی روز تعداد داده‌ها در ساعات اولیه صبح افزایش پیدا می‌کند و این روند افزایش تا انتهای شبانه‌روز ادامه پیدا می‌کند و در ساعات بامداد (۰-۶) مجدداً کاهش پیدا می‌کند. به‌بیان‌دیگر مردم برای رسیدگی به اهداف مختلف خود مانند خرید، تفریح و غیره وارد این مناطق پراهمیت می‌شوند و حضور آن‌ها تا پایان روز ادامه پیدا می‌کند.

همان‌طور که توضیح داده شد همخوانی مطلوبی بین تغییرات تعداد داده‌ها، جابجایی شهروندان و کاربری‌های تخصیص داده شده وجود دارد. در شکل (۶) نتایج حاصل نهایی از انتساب کاربری‌های فوق به هر یک از مناطق شهری که پیش از این با دیاگرام ورونوی تفکیک‌شده‌اند، نمایش داده شده است.

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.6.3.1]



شکل ۶: کاربری‌های کشف شده برای مناطق شهری

۵- ارزیابی

به‌منظور ارزیابی نتایج حاصل‌شده از بعد معنایی موجود در شبکه‌های اجتماعی استفاده شده است. به این منظور از یک مجموعه داده شامل مشخصات مکانی و گروه معنایی مکان[1] های ثبت شده شهر توکیو در شبکه اجتماعی فراسکوئر استفاده کردیم. شبکه اجتماعی فوراسکوئر، مکان‌های موجود در محیط شهری را در ۴۰۰ گروه طبقه‌بندی کرده است. این ۴۰۰ گروه به شکل سلسله مراتبی در زیرگروه ۹ بخش اصلی قرار دارند که عبارت‌اند از: ۱- هنر و سرگرمی ۲- دانشگاه‌ها و مراکز آموزشی ۳- غذا ۴- تفریح ۵- مراکز تفریحی شبانه ۶- اداره‌ها ۷- مسکونی ۸- فروشگاه و خدمات (تجاری) ۹- حمل‌ونقل و سفر. در شکل (۷) بخشی از این طبقه‌بندی نشان داده شده است. قابل ذکر است که این مکان‌های ثبت شده تأیید شده هستند و صحت آن‌ها قطعی است. جدول (۳) میزان حضور مکان‌های متعلق به گروه‌های معنایی را در مناطق شناسایی شده متناظر نشان می‌دهد.

جدول (۳) نشان دهنده ماتریس ابهام[2] برای روش پیشنهادی است. برای ارزیابی روش پیشنهادی دو شاخص کاپا[3] و دقت کلی[4] محاسبه گردید. دقت کلی عبارت است از نسبت تعداد مکان‌هایی‌هایی که درست شناسایی شده‌اند (مجموع عناصر قطر اصلی) به تعداد کل مکان‌ها. شاخص کاپا نیز که معرف میزان موافقت دو طبقه‌بندی است از رابطه (۲) محاسبه می‌گردد:

$$\hat{\kappa} = \frac{N\sum_{i=1}^{k} x_{ii} - \sum_{i=1}^{k}(x_{i+}\times x_{+i})}{N^2 - \sum_{i=1}^{k}(x_{i+}\times x_{+i})} \qquad \text{رابطه (۲)}$$

که در آن N تعداد کل مکان‌هاست و $(x_{i+}\times x_{+i})$ به معنای حاصل ضرب مجموع سطر i ام و ستون i ام در ماتریس در هم گسیختگی است. دقت کلی و شاخص کاپا به ترتیب برابر ۷۹ و ۷۱ درصد محاسبه گردید که نشان دهنده دقت خوب روش پیشنهادی این تحقیق است.

[1] Venue
[2] Confusion matrix
[3] Kappa index
[4] Overall accuracy

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.6.3.1]



Category	Checkins	Venues
Subways	256	30
Malls	39	13
Cafés	25	12
Кофехаус	3	
Кофе Хауз	3	
Shokoladnitsa	1	
Кофе Хауз	1	
Starbucks	1	
Coffee House	1	
Шоколадница	9	
Саквояж	1	
Кофехауз	2	
Фасолька, европейское кафе	1	
Столовая #1, "Копейка"	1	
Мама Раша	1	
Japanese Restaurants	17	12
Highways or Roads	91	11

شکل۷: بخشی از گروه‌بندی مکان‌ها در شبکه اجتماعی فوراسکوئر

جدول۳: ارزیابی روش پیشنهادی بر مبنای حضور مکان‌های ثبت شده در شبکه فراسکوئر

کاربری / گروه معنایی	اداری	تفریحات شبانه	مسکونی	تجاری و تفریحی
اداره‌ها	۳۹۳	۴۵	۸۴	۲۹
مراکز تفریحی شبانه	۶	۱۵۱	۲۰	۱۴
مسکونی	۰	۳۵	۲۸۴	۲۹
تجاری، غذا و تفریح	۴	۲۲	۱۰۱	۶۴۴

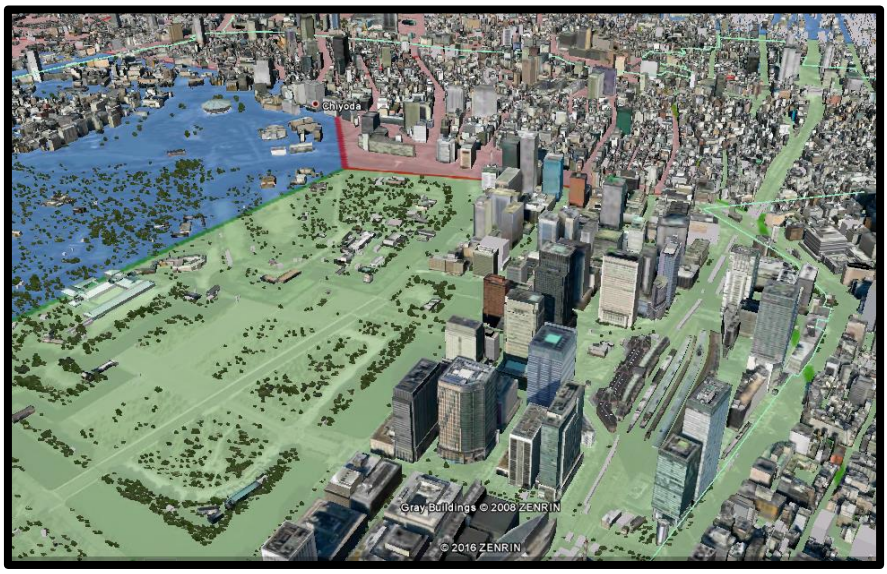
علاوه بر این همان‌طور که در شکل (۷) نشان داده شده است؛ مناطق با کاربری یکسان دارای پیوستگی مکانی هستند. هم‌چنین به منظور اطمینان از نتایج حاصل‌شده، آن‌ها را با تصاویر ماهواره‌ای مقایسه کردیم؛ که یک نمونه از آن برای یکی از مناطقی که با کاربری اداری شناسایی شده‌اند، در شکل (۸) نمایش داده شده است (منطقه نشان داده شده با زمینه سبز رنگ). این مقایسه به‌خوبی نشان می‌دهد که نتایج به‌دست‌آمده تا حد قابل قبولی با واقعیت سازگار هستند، به‌طوری‌که در مناطق شناسایی شده با کاربری اداری، تعداد زیادی ساختمان بلند با نام‌های مرتبط وجود دارند. هم‌چنین در مناطق مسکونی می‌توان وجود خانه‌ها، آپارتمان‌ها و مدارس را مشاهده کرد.

۶- نتیجه‌گیری

در این تحقیق، از رفتار اجتماعی شهروندان در شبکه‌های اجتماعی به عنوان یک شاخص برای تعیین نوع کاربری مناطق مختلف شهری استفاده گردید. برای این منظور پس از افراز محیط شهری به چندین زیر بخش، بر مبنای الگوی رفتار کاربران در شبکه‌های

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]
[DOI: 10.29252/jgit.6.3.1]

اجتماعی در طول ۲۴ ساعت شبانه‌روز، به هر منطقه یک کاربری (اداری، مسکونی، تجاری، تفریحی و کاربری تفریحات شبانه) نسبت داده شد. برای ارزیابی روش پیشنهادی، با استفاده از داده‌های شبکه اجتماعی فراسکوئر به بررسی میزان تعلق مکان‌های با کاربری اجتماعی مشخص در مناطق متناظر پرداخته شد. همچنین نتایج حاصل از این تحقیق با مؤلفه‌های شهری موجود در عکس‌های ماهواره‌ای مقایسه گردید، که نتایج آن بیانگر قابل قبول بودن نتایج به دست‌آمده بود. نتایج این تحقیق نشان می‌دهد که ساختار فیزیکی و نقش اجتماعی هر منطقه بر روی رفتار شهروندان تأثیر می‌گذارد و این تأثیر به شبکه‌های اجتماعی نیز سرایت می‌کند، بنابراین می‌توان از پتانسیل بالای این داده‌ها در راستای شناخت هر چه بیشتر محیط‌های شهری بهره گرفت. دسترسی به داده‌های بزرگ‌تر و فراهم کردن زیرساخت محاسباتی برای اعمال روش های لازم می‌تواند نتایج تحقیق را بهبود بخشد. برای تحقیقات آینده می‌توان با تحلیل متن در پست‌های کاربران در شبکه‌های اجتماعی به نتایج تازه‌ای دست یافت، هم‌چنین می‌توان از داده‌های حرکتی سایر منابع حرکتی نظیر *GPS* خودروها، تاکسی‌ها، تلفن‌های همراه، کارت‌های اعتباری و غیره استفاده نموده و نتایج حاصل را با هم مقایسه کرد. هم‌چنین استفاده از داده‌های شبکه‌های اجتماعی مکان‌مبنا در راستای حل معضلات شهری نظیر ترافیک، آلودگی صوتی، آلودگی هوا و غیره نیز می‌تواند چشم‌انداز جدیدی چه در زمینه پژوهشی و چه در زمینه تجاری باشد.



شکل ۸: مقایسه نمونه‌ای از کاربری اداری تشخیص داده (رنگ سبز) شده بر اساس روش پیشنهادی با تصویر ماهواره‌ای

مراجع

[1] microsoft_research. (4/11/2016). Available: https://www.microsoft.com/en-us/research/project/urban-computing/

[2] M. Hosseinpour and M. Malek, "Location Based Social Networks: Opportunities and Challenges," Geospatial Engineering Journal, vol. 6, pp. 51-62, 2015.

[3] A. Kheiri, F. Karimipour, and M. Forghani, "Intra-Urban Movement Patterns Estimation Based on Location Based Social Networking Data," Journal of Geomatics Science and Technology, vol. 6, pp. 141-158, 2016.

[4] Q. Li, Y. Zheng, X. Xie, Y. Chen, W. Liu, and W.-Y. Ma, "Mining user similarity based on location history," in Proceedings of the 16th ACM SIGSPATIAL

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.6.3.1]

international conference on Advances in geographic information systems, 2008, p. 34.

[5] Y. Zheng, Y. Chen, X. Xie, and W.-Y. Ma, "GeoLife2. 0: a location-based social networking service," in 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware, 2009, pp. 357-358.

[6] N. Li and G. Chen, "Analysis of a location-based social network," in Computational Science and Engineering, 2009. CSE'09. International Conference on, 2009, pp. 263-270.

[7] S. Scellato, C. Mascolo, M. Musolesi, and V. Latora, "Distance Matters: Geo-social Metrics for Online Social Networks," in WOSN, 2010.

[8] J. Cranshaw, R. Schwartz, J. I. Hong, and N. Sadeh, "The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City," in Proceedings of the Sixth International AAAI Conference on Weblogs an Social Media (ICWSM 2012), Trinity College in Dublin, Ireland, 2012.

[9] X. Liu, J. He, Y. Yao, J. Zhang, H. Liang, H. Wang, et al., "Classifying urban land use by integrating remote sensing and social media data," International Journal of Geographical Information Science, pp. 1-22, 2017.

[10] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil, "Exploiting Semantic Annotations for Clustering Geographic Areas and Users in Location-based Social Networks," in Fifth International AAAI Conference on Weblogs and Social Media, 2011.

[11] S. Hasan, X. Zhan, and S. V. Ukkusuri, "Understanding urban human activity and mobility patterns using large-scale location-based data from online social media," in Proceedings of the 2nd ACM SIGKDD international workshop on urban computing, 2013, p. 6.

[12] TwitterAPI. (4/11/2016). Available: http://apiwiki.twitter.com/

[13] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in Kdd, 1996, pp. 226-231.

[14] V. Frias-Martinez and E. Frias-Martinez, "Spectral clustering for sensing urban land use using Twitter activity," Engineering Applications of Artificial Intelligence, vol. 35, pp. 237-245, 2014.

[15] J. Yuan, Y. Zheng, and X. Xie, "Discovering regions of different functions in a city using human mobility and POIs," in Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, 2012, pp. 186-194.

[16] Y. Liu, X. Liu, S. Gao, L. Gong, C. Kang, Y. Zhi, et al., "Social sensing: A new approach to understanding our socioeconomic environments," Annals of the Association of American Geographers, vol. 105, pp. 512-530, 2015.

[17] S. Rinzivillo, S. Mainardi, F. Pezzoni, M. Coscia, D. Pedreschi, and F. Giannotti, "Discovering the geographical borders of human mobility," KI-Künstliche Intelligenz, vol. 26, pp. 253-260, 2012.

[18] S. Wakamiya, R. Lee, and K. Sumiya, "Social-urban neighborhood search based on crowd footprints network," in International Conference on Social Informatics, 2013, pp. 429-442.

[19] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm," Journal of the Royal Statistical Society. Series C (Applied Statistics), vol. 28, pp. 100-108, 1979.

[20] T. M. Kodinariya and P. R. Makwana, "Review on determining number of Cluster in K-Means Clustering," International Journal, vol. 1, pp. 90-95, 2013.

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.6.3.1]

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

Journal of Geospatial Information Technology
Vol.6, No.3, Autumn 2018

Land use characterization in urban environments based on the behaviors of location based social networks' users

Morteza Tayebi[1], Farid Karimipour*[2]

1- PhD Student, Remote sensing Eng. Dep., Faculty of Geomatics, KN Toosi University of Technology
2- Assistant professor, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran

Abstract

Nowadays, after a relatively short time from the emergence of location-based social networks, these social media have become an inseparable element of social life, especially in big cities. The increased popularity of social networks, the development of location-based technologies including Global Positioning System (GPS) as well as the pervasive use of smart-phones that besides being equipped with GPS can connect to the world-wide web easily have led to the generation of update, low-cost and valuable spatial data. Beside various applications, these data can be used to resolve urban challenges. As an example, in this paper, the behaviors of the location-based social networks' users were assessed to classify the land use in urban environments. In this regard, initially, we attempted to recognize regions in urban environments by clustering the users' spatial data, then the variation patterns of users' social behaviors through the day and night were extracted and used as an index to land use determination and a type of social land use was designated to each region. Evaluating the obtained results and comparing them with satellite images show the high potential of location-based social networks to recognize urban environments and resolve existing challenges.

Key words: location-based social network, clustering, social land use.

Correspondence Address : GIS Group, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran.
Tel : +98 21 61114376
Email: fkarimipour@ut.ac.ir

[Downloaded from jgit.kntu.ac.ir on 2026-11-16]

[DOI: 10.29252/jgit.6.3.1]