

توسعه سامانه به‌روزرسانی خودکار پایگاه داده مکانی با استفاده از داده‌های مکانی مردم گستر (مطالعه موردی: رستوران‌ها)

سپهر هنرپرور^{۱*}، علی اصغر آل شیخ^۲

^۱ دانشجوی کارشناسی ارشد سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی
sep.honarparvar@kntu.ac.ir

^۲ استاد گروه سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی
alesheikh@kntu.ac.ir

(تاریخ دریافت بهمن ۱۳۹۳، تاریخ تصویب اردیبهشت ۱۳۹۴)

چکیده

در کلان‌شهرهایی مثل تهران، کاربری‌های اراضی به علل مختلف از جمله کنترل ترافیک و رفع نیازهای نواحی مختلف شهری، در حال تغییر می‌باشند. مراکز ارائه‌دهنده خدمات گردشگری مثل رستوران‌ها هم تحت تأثیر این تغییرات قرار می‌گیرند. در بعضی از این مراکز حتی کیفیت ارائه خدمات هم دست‌خوش تغییر می‌شود. بنابراین به منظور به‌روزرسانی سریع اطلاعات خدمات گردشگری در هر کاربرد از جمله گردشگری به یک سامانه احتیاج است. از آنجایی که به‌روزرسانی پایگاه داده توسط سازمان‌ها و مؤسسات مستلزم صرف زمان و هزینه نسبتاً زیادی است، به سامانه‌ای احتیاج است که بتواند داده‌های مکانی را به صورت خودکار به‌روز کند، به صورتی که فرآیند به‌روزرسانی اطلاعات احتیاج به طی پروسه‌های زمان‌بر و هزینه‌بر نداشته باشد. با توجه به اینکه اطلاعات مکانی مردم گستر از کاربران معمولی دریافت می‌شوند، نیازی به زمان یا هزینه بالا برای جمع‌آوری ندارند. بنابراین با استخراج این داده‌ها می‌توان به صورت خودکار پایگاه داده مراکز خدماتی را به‌روزرسانی کرد. هدف مقاله این است که قابلیت اطلاعات مکانی داوطلبانه در به‌روزرسانی خودکار پایگاه داده آزمایش شود. برای جمع‌آوری داده‌های VGI^۱ از یک برنامه کاربردی موبایل استفاده شده است، و داده‌های مرجع رستوران‌ها هم از GoogleMap و سایت Fidio استخراج شده‌اند. داده‌های VGI مناسب برای ورود به پایگاه داده مکانی، بایستی ابتدا کیفیت سنجی شوند. مسئله این است که کیفیت VGI ابعاد و المان‌های مختلفی دارد ولی برای به‌روزرسانی نیاز به مقدار واحد است. لذا از روش OWA^۲ و کمیت سنج‌های فازی برای تبدیل کیفیت چندبعدی به یک‌بعدی استفاده شد، و داده‌های با کیفیت بالا برای ورود به پایگاه داده آماده شدند. نتایج حاکی از آن بود که سامانه در تکمیل اطلاعات ۶۰ درصد از رستوران‌های ناحیه موفق بوده است. در نهایت، و پس از بررسی نقشه به دست آمده از این سامانه و نقشه مرجع مشخص شد که برای دریافت اطلاعات رستوران‌ها، لزوم بالا بودن صحت توصیفی «نام رستوران‌ها» به عنوان عامل کنترل‌کننده، مانع از بروز خطا در تخصیص صحیح اطلاعات به رستوران‌های مورد نظر در نواحی متراکم می‌گردد.

واژگان کلیدی: اطلاعات مکانی مردم گستر، به‌روزرسانی خودکار پایگاه داده، OWA و کمیت سنج‌های فازی، رستوران‌ها، کیفیت VGI، برنامه کاربردی موبایل

* نویسنده رابط

^۱ Volunteered Geographic Information

^۲ Ordered Weighted Averaging

۱- مقدمه

در شهرهای بزرگ مثل تهران کاربری‌های اراضی، به علل مختلفی از جمله کنترل ترافیک و رفع نیازهای نواحی مختلف شهری، در حال تغییرند. مراکز ارائه‌دهنده خدمات گردشگری مثل رستوران‌ها هم تحت تأثیر این تغییرات قرار می‌گیرند. در بعضی از این مراکز حتی کیفیت ارائه خدمات هم دست‌خوش تغییر می‌شوند. بنابراین به منظور به‌روزرسانی سریع این موارد در هر کاربرد از جمله گردشگری، به یک سامانه احتیاج است؛ زیرا در صورت عدم توجه به این مسئله از یک سو اعتماد کاربران به هر برنامه کاربردی وابسته به اطلاعات اولیه مراکز خدماتی (مانند رستوران‌ها) پایین می‌آید، و از طرفی باعث اتلاف هزینه و زمان کاربران برای یافتن نتیجه صحیح می‌شود. از سوی دیگر اطلاعات به دست آمده از مردم درباره‌ی سلیقه‌شان و کیفیت خدمات گردشگری (مانند رستوران‌ها) همیشه بهترین منبع برای برنامه‌ریزی و سرمایه‌گذاری در زمینه‌ی ارائه‌ی خدمات در آینده است چرا که در واقع منبع اصلی تولید سرمایه در تجارت گردشگری همین گردشگران می‌باشد. در نهایت اضافه کردن مراکز خدماتی جدید توسط مردم می‌تواند به صاحبان خدمات جدید (که در پایگاه داده‌ی برنامه کاربردی وجود نداشتند) کمک کند تا بتوانند خود را به کاربرانی که اطلاعی از وجود این رستوران ندارند، معرفی کنند. بنابراین نیاز به سامانه‌ای است تا بتواند پایگاه داده خدمات گردشگری را به‌روزرسانی کند.

از آنجایی که به‌روزرسانی پایگاه داده توسط سازمان‌ها و مؤسسات مستلزم صرف زمان و هزینه نسبتاً زیادی هست احتیاج به سامانه‌ای است که بتواند داده‌های مکانی را به صورت خودکار به‌روز کند، به صورتی که فرآیند به‌روزرسانی اطلاعات، احتیاج به طی پروسه‌های زمان‌بر و گران اداری نداشته باشد [۱]. علت این امر این است که در برخی کاربردها مثل سامانه‌های توصیه گر گردشگری به اطلاعات به‌روز نیاز است و دسترسی به اماکن گردشگری جدید اعم از خدمات و یا راه‌های جدید نیازمند اطلاعات صحیح و جدید است. از طرفی نیاز مبرم به برخی اطلاعات کیفی مثل کیفیت غذا، کیفیت خدمات‌رسانی، میزان ازدحام در منطقه‌ی گردشگری مورد نظر، زیبایی و میزان

جذابیت اماکن و غیره بدون استفاده از نظر کاربران غیر خبره و عموم ممکن نیست.

اطلاعات مکانی مردم گستر که از کاربران معمولی دریافت می‌شوند، نیازی به زمان یا هزینه بالا برای جمع‌آوری ندارد. بنابراین با استخراج این داده‌ها می‌توان به صورت خودکار پایگاه داده مراکز خدماتی مثل رستوران‌ها را به‌روزرسانی کرد. برای جمع‌آوری داده‌های مردمی از یک برنامه کاربردی موبایل استفاده شده است. در نهایت هدف مقاله توسعه سامانه‌ای است که داده‌های مردمی را از طریق سکوی موبایل جمع‌آوری کرده، و به صورت آنی برای به‌روزرسانی پایگاه داده آماده کند. مسئله اصلی این تحقیق تشخیص و ورود داده‌های با کیفیت است. با توجه به این که کیفیت VGI ماهیت چندبعدی دارد، تصمیم‌گیری برای انتخاب داده‌های باکیفیت کاری مشکل است. پس باید فضای تصمیم را به یک بعد کاهش داد. در این حالت نمی‌توان از وزن‌دهی و اولویت‌بندی صحیح ابعاد یا المان‌های کیفیت اطمینان حاصل کرد. لذا به روشی نیاز است که میزان وابستگی تصمیم‌گیری به اولویت‌بندی المان‌ها را تا حد ممکن بکاهد، و با بررسی حالات موجود بهترین تصمیم را برای تناسب کیفیت بگیرد. OWA به همراه کمیت‌سنج‌های زبانی می‌تواند سناریوهای مختلف زبانی را در اختیار تصمیم‌گیرنده قرار داده، و میزان اطمینان را در اتخاذ تصمیم نهایی بالا برد. در بخش دوم مبانی نظری تحقیق شامل به‌روزرسانی خودکار پایگاه داده، اطلاعات مکانی مردم گستر و کیفیت اطلاعات مکانی مردم گستر بررسی می‌شود. در بخش سوم به روش پیاده‌سازی و ارزیابی کیفیت داده‌های VGI برای ورود به پایگاه داده تشریح می‌شود. در بخش چهارم به بررسی نتایج و در بخش پنجم به پرداخته می‌شود.

۲- مبانی نظری تحقیق

در این بخش ابتدا توضیح مختصری در رابطه با تعریف و ضرورت استفاده از به‌روزرسانی خودکار پایگاه داده خواهد شد و سپس به بررسی اطلاعات مکانی مردم گستر و کیفیت VGI، پژوهش‌های صورت گرفته در این زمینه‌ها پرداخته خواهد شد.

۲-۱- به روزرسانی خودکار پایگاه داده

امروزه حجم انتقال داده‌های مکانی بین سازمان‌های دولتی و خصوصی و مردم افزایش چشمگیری داشته است. بنابراین مدیریت و کنترل این حجم عظیم از داده‌های مکانی امری حیاتی است [۲]. در این میان طراحی یک پایگاه داده مکانی مناسب می‌تواند فرآیند انتقال، ذخیره‌سازی، حذف یا افزودن داده‌های مکانی را بهبود بخشد. یکی از گران‌ترین و زمان‌برترین مراحل در پیاده‌سازی سامانه‌های اطلاعات مکانی، در آماده‌سازی اطلاعات ورودی مکانی صرف می‌شود، به طریقی که ۶۰ تا ۸۰ درصد از هزینه‌های پیاده‌سازی را در بر می‌گیرد [۳]. به روزرسانی داده‌های مکانی هم از آن جهت ضرورت دارد که عوارض مکانی مخصوصاً در مناطق شهری در حال تغییرند و اطلاعات به روز شده عوارض مکانی دارای صحت بیشتری هستند. بنابراین مسئله‌ی ورود و به روزرسانی پایگاه داده مکانی دارای اهمیت ویژه‌ای است. در رابطه با لزوم استفاده از به روزرسانی خودکار می‌توان دو دلیل عمده زیر را ذکر کرد:

- به روزرسانی داده‌های مکانی شامل سه مرحله آشکارسازی اشیاء در حال تغییر، مشخص کردن نوع تغییر و بازنگری و به روزرسانی داده مکانی است [۴]. ولی هر یک از این مراحل نیازمند صرف هزینه و زمان زیاد است. به روزرسانی پایگاه داده خودکار زمانی که مراحل جمع‌آوری و به روزرسانی همزمان انجام شود، بیشتر مورد استفاده قرار می‌گیرد [۵].

- دلیل دیگر استفاده از به روزرسانی خودکار برتری آن نسبت به روش‌های دیگر است. از میان روش‌های به روزرسانی، روش دستی بسیار پرهزینه و زمان‌بر است [۶]. به عبارتی دخالت عامل انسانی در پروسه به روزرسانی بر کیفیت فراداده تأثیر می‌گذارد [۷]. رویکردهای نیمه خودکار هم محدوده‌ی مشخصی از فرمت‌ها را پوشش می‌دهد، و از این لحاظ برای پایگاه داده حجیم با داده‌های ورودی زیاد مناسب به نظر نمی‌رسد [۸]. از طرفی در روش دستی و نیمه خودکار، به روزرسانی پایگاه داده و فراداده در دو مدل داده مختلف قرار می‌گیرد [۹] و این نوع ذخیره‌سازی باعث می‌شود که در مواردی به روزرسانی فراداده از داده زودتر یا دیرتر صورت گیرد که این امر باعث کاهش اعتمادپذیری داده‌های مکانی در پروسه به روزرسانی می‌شود. بنابراین هر

چه پروسه‌های به روزرسانی داده‌های مکانی ارتباط کمتری با عامل انسانی داشته باشد، مناسب تر است [۸]. لذا به روزرسانی پایگاه داده خودکار می‌تواند برای خیلی از کاربردها مناسب‌ترین گزینه باشد.

برای به روزرسانی خودکار پایگاه داده باید به نوع داده‌های ورودی (منابع داده‌ای که وارد پروسه به روزرسانی می‌شود) توجه کرد. به طور کلی، منابع به روزرسانی پایگاه داده از روش‌های فتوگرامتری، سنجش از دور و GPS^۱ تأمین می‌شود. و در این بین جمع‌آوری داده‌ها از طریق GPS برای به روزرسانی خودکار پایگاه داده مکانی عملی‌تر به نظر می‌رسد، چرا که مختصات را می‌توان بدون انجام تحلیل و محاسبه، به طور مستقیم به دست آورد. لذا در این پژوهش به منظور به روزرسانی خودکار از مختصات ورودی GPS تلفن‌های هوشمند استفاده شده است.

در به روزرسانی پایگاه داده، اطمینان از فراداده‌های مکانی اهمیت ویژه‌ای دارد. به خصوص فراداده‌هایی مثل کیفیت و تاریخ تولید داده‌های مکانی. پایگاه داده مکانی که دارای داده‌های با کیفیت بیشتر بسیار قابل اعتمادتر است و به همین دلیل داده‌ها قبل از ورود به پایگاه داده باید کیفیت سنجی شوند و داده‌های با کیفیت بالاتر جدا شوند. لذا در این مقاله مبحث کنترل کیفیت داده‌های مکانی ورودی نیز مورد بررسی قرار خواهد گرفت.

۲-۲- اطلاعات مکانی مردم گستر

در سایه پیشرفت حوزه فناوری و افزایش آگاهی جوامع نسبت به داده‌های مکانی تحولی در الگوی تولید و مصرف داده‌های مکانی توسط مردم به وجود آمده است. گودچایلد [۱۰] این تحول را با معرفی مفهوم جدیدی به نام اطلاعات مکانی داوطلبانه تفسیر می‌کند. تعریف رسمی اطلاعات مکانی داوطلبانه بدین صورت می‌باشد: « بهره‌برداری از ابزار برای ایجاد، جمع‌آوری و انتشار داده‌های جغرافیایی که به طور داوطلبانه توسط افراد تولید شده است » [۱۱]. با انتشار این مفهوم در مجامع علمی، محققان حوزه‌های فناوری اطلاعات و سامانه اطلاعات مکانی بر آن شدند تا از اطلاعات مکانی داوطلبانه را در کاربردهای مختلف مثل تولید نقشه‌های گردشگری، مدیریت بحران و به روزرسانی نقشه‌های شبکه راه‌های شهری استفاده کنند.

^۱ Global Positioning System

^۲ Goodchild

۲-۳- کیفیت داده‌های مکانی مردم گستر

کیفیت VGI، ابعاد و المان‌های مختلفی دارد. در اغلب پژوهش‌ها در زمینه کیفیت داده‌های داوطلبانه المان‌های تمامیت، صحت مکانی، صحت توصیفی و سازگاری منطقی از بین المان‌های مرسوم در منابع علمی مختلف کیفیت به عنوان المان‌های موثر انتخاب و مطالعه شده‌اند [۱۵]. با توجه به ماهیت داده‌های داوطلبانه روش‌های ارزیابی کیفیت این‌گونه داده‌ها ممکن است نه تنها برای پروژه‌های مختلف، بلکه برای یک پروژه و بر اساس فضا، تولیدکننده و مصرف‌کننده داده متفاوت باشد. تاکنون روش‌های متفاوتی برای ارزیابی کیفیت اطلاعات مکانی داوطلبانه استفاده شده‌اند. این روش‌ها را می‌توان به دو دسته کلی تقسیم کرد [۱۶]:

- دسته اول از پروژه‌های صورت گرفته در ارزیابی کیفیت اطلاعات مکانی مردم گستر به این مسئله اشاره دارد که برای ارزیابی باید نقشه حاوی داده‌های مردمی با نقشه مرجع مقایسه و نقاط و عوارض موجود در آن‌ها تناظر یابی شوند. به عبارتی تناظر یابی اولین مرحله از فرآیند ارزیابی کیفیت در روش اول می‌باشد، که به دو صورت دستی و خودکار امکان‌پذیر است. از پروژه‌های صورت گرفته در این زمینه [۱۶] و [۲۸] تناظر یابی را به صورت خودکار انجام داده‌اند و بقیه به صورت دستی انجام شده‌اند [۱۷]، [۱۸]، [۱۹]، [۱۳].

- در دسته دوم داده‌های VGI با داده‌های مرجع مقایسه نمی‌شوند و می‌توان از بررسی مستقیم خود داده‌ها درباره‌ی کیفیت آن‌ها اطلاعاتی به دست آورد. از میان این دسته از پژوهش‌ها می‌توان به پروژه شناسایی یخچال‌ها توسط داوطلبان توسط [۲۰] اشاره کرد. این مقاله کیفیت اطلاعات مکانی داوطلبانه را به سه دسته کیفیت ذاتی^۱، بیرونی^۲ و عملی^۳ تقسیم کرده است. کیفیت ذاتی به محتوای داده بستگی دارد همچون دقت و صحت متون، تصاویر و اندازه‌گیری‌ها. کیفیت بیرونی به مشخصات و اعتبار تولیدکنندگان داده (داوطلبان) مربوط است و کیفیت عملی نیز به هدف در نظر گرفته‌شده برای آن پروژه بستگی دارد و ممکن است از پروژه‌ای به پروژه دیگر تغییر کند [۲۱]. در واقع در این دسته از تحقیقات کیفیت اطلاعات مکانی به صورت مستقیم محاسبه می‌شود و از

اطلاعات مکانی داوطلبانه را می‌توان نوع خاصی از مفهومی رایج در حوزه فناوری اطلاعات به نام محتوای کاربرتولید در نظر گرفت. از پروژه‌هایی که در راستای استفاده از محتوای کاربرتولید در محیط وب است می‌توان به Flickr و Panaramo اشاره کرد. کاربران در این دو سایت امکان به اشتراک‌گذاری عکس‌های زمین مرجع شده را در محیط وب دارند. از پروژه‌های دیگر در این زمینه می‌توان به سایت Wikipedia اشاره کرد که کاربران را قادر می‌سازد تا محتوای علمی سایت را تکمیل کنند. سایت مشابه Wikimapia را می‌توان نمونه‌ای از پروژه‌های مورد استفاده در زمینه اطلاعات مکانی داوطلبانه نام برد که با بهره‌گیری از فناوری مورد استفاده از این سایت Wikipedia کاربران را در تکمیل نقشه‌های شهری کمک می‌کند. از پروژه‌های دیگر در این زمینه می‌توان به Google Map Maker و OpenStreetMap اشاره کرد.

بسیاری از تحقیقات صورت گرفته در زمینه اطلاعات مکانی داوطلبانه در رابطه با بررسی میزان کیفیت داده‌های جمع‌آوری شده توسط این سایت‌هاست. در این میان سایت OpenStreetMap به علت امکان مستقیم داده‌های مکانی بیشتر مورد توجه محققان قرار گرفته است، و تحقیقات بیشتری در مناطق مختلف دنیا روی آن صورت گرفته است [۱۲]، [۱۳]. تحقیقات دیگری که در زمینه VGI صورت می‌گیرد، بیشتر مرتبط با حوزه علوم شهروندی می‌باشند. نمونه‌ای از این تحقیقات پروژه سرشماری پرنده‌ی «کریسمس» در آمریکا است که پرنده‌شناسان مبتدی در شمارش گونه‌های مختلف پرنده‌گان محققان را یاری می‌دهند. اصولاً در این نوع تحقیقات شبکه‌ای از شهروندان که از دانش محدودی در حوزه مورد تحقیق برخوردارند، مورد استفاده قرار می‌گیرد. قاعدتاً حجم مشارکت عمومی در این پروژه‌ها پایین‌تر از سایر پروژه‌هاست [۱۴].

برای به‌روزرسانی خودکار پایگاه داده مکانی، نیاز به داده‌های قابل‌اعتماد و باکیفیت است، لذا باید داده‌های ورودی مردمی را تا حد ممکن پالایش کرد. برای این منظور باید شناخت کافی از کیفیت داده‌های مکانی و اطلاعات مکانی مردم گستر داشت. لذا در بخش‌های بعدی به توضیح مختصری در رابطه باکیفیت داده‌های مکانی و کیفیت اطلاعات مکانی مردم گستر پرداخته خواهد شد.

^۱ Intrinsic

^۲ Extrinsic

^۳ Pragmatic



شکل ۱- صفحه‌ی رابط کاربری موبایل برای ورود مختصات رستوران‌ها



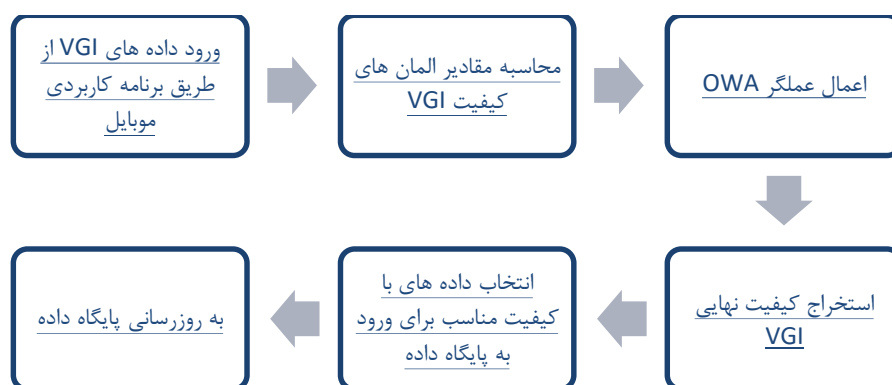
شکل ۲- صفحه‌ی رابط کاربری موبایل برای ورود اطلاعات رستوران‌ها

پس از تخلیه اطلاعات از برنامه کاربردی موبایل کاربران به پایگاه داده SQL Server، داده‌ها برای آنالیز به برنامه ارزیابی کیفیت که با زبان C# نوشته شده است فرستاده می‌شوند. پس از ارزیابی، داده‌های باکیفیت مشخص می‌شوند و برای به‌روزرسانی به پایگاه داده ارسال می‌شوند. روند کلی به‌روزرسانی خودکار توسط سامانه در (شکل ۳) تشریح شده است.

این رو به فرآیند ارزیابی کیفیت صورت خودکار و با سرعت بالا قابل انجام است. بنابراین معمولاً در پروژه‌هایی که به‌روزرسانی خودکار پایگاه داده مکانی با استفاده از داده‌های مردمی اهمیت ویژه‌ای دارد، از این‌گونه روش‌ها استفاده می‌شود. لذا با توجه به اینکه هدف به‌روزرسانی خودکار یک پایگاه داده مکانی است، روش مورد استفاده در این پژوهش در این گروه از تحقیقات قرار می‌گیرد.

۳- پیاده‌سازی

داده‌های VGI مربوط به اطلاعات رستوران‌ها از طریق یک برنامه کاربردی موبایل از کاربران دریافت می‌شود. داده‌های مرجع هم از سایت Fidio که حاوی اطلاعات رستوران‌هاست، و سایت GoogleMaps که حاوی مختصات است، به دست آمده است. برنامه کاربردی مورد نظر در محیط برنامه‌نویسی اندروید نوشته شده است، و طراحی آن به گونه‌ای است که انتخاب نقاط رستوران‌ها با استفاده از ابزارهای رابط کاربری با ذخیره مختصات GPS و بدون نیاز به کلیک روی نقشه صورت می‌گیرد. با توجه به این که احتمال خطای انسانی در زمان کلیک روی نقطه مورد نظر بالاست، استفاده از سنسور GPS تلفن همراه هوشمند می‌تواند احتمال خطای انسانی را پایین آورد. در برنامه کاربردی موبایل، کاربران ابتدا مختصات رستوران را از GPS تلفن همراه هوشمند وارد می‌کنند (شکل ۱) و سپس اطلاعات آن رستوران را وارد می‌کنند (شکل ۲).



شکل ۳- روند کلی سامانه به‌روزرسانی خودکار

در روند استخراج کیفیت اطلاعات مکانی داوطلبانه، اولین مرحله مشخص کردن و مقداردهی به درایه‌های بردار کیفیت می‌باشد. بردار کیفیت، برداری است که هر درایه آن حاوی یک المان کیفیت (مثل صحت مکانی، توصیفی و غیره) است. روند تعیین مقادیر کیفیت بستگی به نوع پروژه و دانش موجود در رابطه با اطلاعات مورد نیاز درباره پروژه دارد. به عبارتی برای مشخص کردن مقادیر المان‌های کیفیت، با توجه به تعریف، نیاز به اطلاعاتی است که میزان تطابق کیفیت به دست آمده از اندازه‌گیری‌ها را با معیارهای کاربرد مورد نظر را مشخص کند. در این تحقیق برای مشخص کردن این مقادیر، از توابع و قیود فازی استفاده شده است. چگونگی استفاده از این توابع در ادامه مقاله آمده است. خروجی این توابع تطابق کیفیت داده‌ها با مقادیر مورد انتظار پروژه برای هر المان کیفیت اطلاعات مکانی مردم گستر را در اختیار می‌گذارد.

اما هدف در این پروژه به دست آوردن خودکار کیفیت داده‌های مکانی داوطلبانه و استفاده از آن‌ها در پایگاه داده مکانی سامانه توصیه‌گر است. برای رسیدن به این هدف احتیاج به یک دید کلی در رابطه با کیفیت داده‌هاست که با تبدیل فضای چندبعدی المان‌های کیفیت به یک مقدار کیفیت تک بعدی میسر می‌شود. از روش‌های موجود برای این تبدیل می‌توان به روش‌های مختلف تصمیم‌گیری چندمعیاره اشاره کرد که از میان آن‌ها روش OWA و کمیت سنج‌های فازی برای این تحقیق انتخاب شده است. این روش در به دست آوردن کیفیت اطلاعات مکانی مردم گستر مربوط به یخچال‌های قطبی توسط [۲۰] و کیفیت اطلاعات مکانی مردم گستر مربوط به میزان تخریب بناهای تاریخی توسط [۲۰] آورده شده است. در این دو مقاله صرفاً به ارزیابی کیفیت عملی با استفاده از OWA اشاره شده است ولی برای آن‌ها کاربردی تعریف نشده است و به چگونگی استفاده از این روش برای به‌روزرسانی پایگاه داده اشاره‌ای نشده است.

تعیین میزان اهمیت هر المان کیفیت اطلاعات مکانی داوطلبانه (توصیفی، مکانی، و غیره) نسبت به المان‌های دیگر به دلیل متفاوت بودن المان‌ها از یکدیگر کار مشکلی است. به عنوان مثال اهمیت المان توصیفی به مکانی در کاربردهای مختلف را نمی‌توان با اطمینان کامل مشخص کرد، زیرا ماهیت دو المان کاملاً متفاوت بوده، و تأثیر آن‌ها

بر کیفیت نهایی به صورت قطعی قابل پیش‌بینی نیست. پس به عملکردی نیاز است که بتواند معیارها و وزن‌های مقداردهی شده را تحت سناریوهای مختلف زبانی مورد بررسی قرار دهد تا اولاً از میزان وابستگی به کارشناسان کم کند، و ثانیاً میزان اهمیت نظر کارشناسان علوم مختلف نسبت به یکدیگر را در فرایندی با سناریوهای مختلف مورد بررسی قرار دهد [۲۱]. به همین دلیل، با توجه به قابلیت سازگاری بالای عملگر OWA با کمیت‌سنج‌های زبانی، در این تحقیق از این عملگر برای به دست آوردن مقادیر کیفیت نهایی یا عملی استفاده شده است [۲۲].

OWA از خانواده‌ی فرآیند ترکیب چندمعیاره به حساب می‌آید [۲۳]؛ که در واقع دو سری از وزن‌ها که اولی وزن‌های ترتیبی و دومی اهمیت نسبی معیارها نسبت به یکدیگر است، را در خود دارد. اگر a برداری n بعدی در نظر گرفته شود، و W هم وزن‌های اهمیت نسبی معیارها باشد، عملگر OWA به صورت زیر عمل می‌کند:

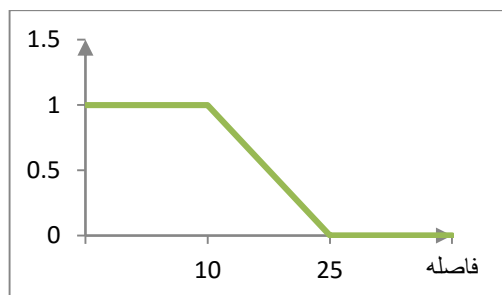
$$OWA(a_1, a_2, \dots, a_n) = \sum_{i=1}^n w_i b_i \quad (1)$$

که در آن b_i وزن ترتیبی است؛ به صورتی که از بزرگ‌ترین مقدار a تا کوچک‌ترین آن به صورت نزولی مرتب شده است. به حالت کلی فرمول محاسبه OWA مطابق زیر می‌باشد.

$$OWA_i = \sum_{j=1}^n \left(\frac{u_j v_j}{\sum_{j=1}^n u_j v_j} \right) z_{ij} \quad (2)$$

که در آن v_j مقادیر وزن‌های ترتیبی و u_j وزن هر المان می‌باشد، z_{ij} هم مقادیر مرتب شده‌ی هر المان کیفیت بر اساس ترتیب v هاست.

در هر پروژه بسته به کاربرد برای ارزیابی کیفیت می‌توان المان‌های مختلفی را برای کیفیت VGI در نظر گرفت. المان‌های صحت مکانی، صحت زمانی و صحت توصیفی در اکثر منابع تحقیقاتی به عنوان المان‌های کیفیت در نظر گرفته می‌شوند. با توجه به میزان اثرگذاری پایین و حتی عدم اثرگذاری المان‌های دیگر در کیفیت داده‌های مکانی داوطلبانه (مثل سازگاری منطقی که معمولاً در رابطه با داده‌های خطی و پلیگونی مطرح



شکل ۴- نمودار تابع عضویت فازی المان صحت مکانی برای فواصل مختلف

۴-۲- صحت توصیفی

این المان به صورت متنی یا رشته‌ای است. به عبارتی به زبان انسان وارد مجموعه داده‌ها می‌شود. اما برای برآورد صحت آن نیاز به معیاری عددی است که برای این کار باید آن را به زبان محاسباتی ترجمه کرد. صحت توصیفی در این تحقیق میزان صحت نام رستوران‌ها را بررسی می‌کند. صحت توصیفی از آن جهت اهمیت دارد که اصلی‌ترین عامل تعریف‌کننده هر رستوران است و مردم رستوران‌ها را بیشتر با نام‌هایشان می‌شناسند. در مکان‌هایی که رستوران‌ها نزدیک به یکدیگر می‌باشند هم به علت امکان بروز خطای تعیین موقعیت نیاز به عاملی کنترل‌کننده است، که در این مورد داده‌های با صحت بالای توصیفی می‌تواند کمک‌رسان باشد. صحت توصیفی بر اساس نسبت تعداد حروف مشابه داده‌های مرجع و داده‌های VGI به تعداد کل حروف رشته موجود در پایگاه داده محاسبه می‌شود.

۴-۳- صحت زمانی

المان صحت زمانی به طور مستقیم از اختلاف زمانی محاسبه می‌شود. این المان اختلاف بین دو آیتم ورودی زمان تولید داده و زمان به‌روزرسانی را مشخص می‌کند. برای محاسبه میزان این اختلاف تعداد روزها شمرده می‌شود. هر چه میزان این اختلاف کمتر باشد، صحت زمانی بیشتر است. صحت زمانی از آنجا اهمیت دارد که برخی رستوران‌ها تغییر کاربری و یا تغییر عنوان می‌دهند یا ممکن است نوع سرویس‌دهی و کادر خدمات و کیفیت غذای خود را تغییر دهند. لذا میزان صحت زمانی داده‌ها در اینجا از اهمیت ویژه‌ای برخوردار است.

می‌شود و به هندسه عارضه وابسته است)، در مطالعه‌ی موردی رستوران‌ها از آن‌ها استفاده نشده است. المان اعتماد پذیری معمولاً به عنوان المانی برای تعیین کیفیت بیرونی مد نظر قرار می‌گیرد [۲۴]، [۲۰]. در این تحقیق با توجه به روش به کار گرفته‌شده، بایستی داده‌های ورودی رستوران‌ها که توسط مردم ایجاد می‌شوند، از لحاظ میزان قابلیت اطمینان سنجیده شوند، لذا این المان به عنوان المانی برای ارزیابی کیفیت در نظر گرفته می‌شود. برای محاسبه المان‌های کیفیت به صورت زیر عمل می‌شود.

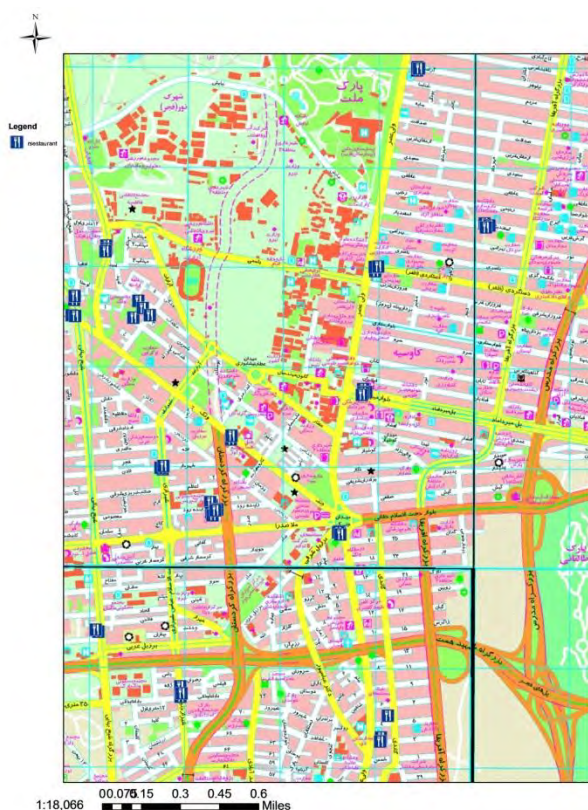
۴- المان‌های کیفیت

۴-۱- صحت مکانی

صحت موقعیت می‌تواند با محاسبه جذر کمترین مربعات یا تابع توزیع نرمال گوسی روی دو نقطه متناظر در دو نقشه مورد نظر محاسبه شود [۲۵]، [۲۶]، [۲۷]. برای تناظرایی نقاط رستوران‌ها بین نقشه مرجع و نقشه VGI از روش مورد استفاده در [۱۶] استفاده شده است؛ با این تفاوت که در این پژوهش از عوارض خطی استفاده شده است. بنابراین به جای استفاده از فرآیند تناظر یابی هفت مرحله‌ای مورد استفاده در آن تحقیق از تناظر یابی تک‌مرحله‌ای با شعاع بافر ۲۵ متر استفاده شده است. این میزان شعاع با بررسی نقشه تراکم توزیع رستوران‌ها به دست آمده است. در نهایت برای تعیین میزان این زیر المان تابع عضویت فازی برای فواصل اقلیدسی از نقطه‌ای توسط کاربر وارد شده، مشخص شده است. تابع فازی مورد نظر در شکل ارائه شده است. اگر خط را در دو جهت y و x برابر فرض کرد، بیضی خطا به دایره تبدیل می‌شود. بنابراین خطای مکانی، رابطه مستقیم با شعاع بافر حول این نقطه را دارد. به عبارتی صحت مکانی با دور شدن از مرکز به صورت خطی کاهش می‌یابد. با توجه به اینکه حداقل فاصله دو رستوران در نقشه ۱۰ متر است، بافرهای با شعاع زیر ۱۰ متر همگی دارای کیفیت کامل می‌باشند. از طرفی شعاع بیشتر از ۲۵ متر هم در تناظر یابی وارد نمی‌شود و به صورت خودکار صفر در نظر گرفته می‌شود.

۵- نتایج

پس از آزمایش سامانه به روزرسانی خودکار رستوران‌ها با استفاده از اطلاعات مکانی مردم گستر نقشه‌ای حاوی اطلاعات به اشتراک گذاشته شده توسط مردم به دست



(ب)



(الف)

شکل ۵-الف) نقشه مرجع رستوران‌های شهر تهران (ب) نقشه رستوران‌های اطلاعات مکانی مردم گستر

۶- بحث و نتیجه‌گیری

با بررسی کلی نتایج مشخص شد که توزیع نامناسب رستوران‌ها در نقشه VGI نشان‌دهنده نایقینی از لحاظ تمامیت می‌باشد و این المان کیفیت مقدار کمی را به خود اختصاص می‌دهد. علت این امر آن است که کثرت داده‌های تولید شده توسط مردم تأثیر مستقیمی بر مقدار المان تمامیت^۱ کیفیت اطلاعات مکانی مردم گستر دارد. و به علت زمان کم برای آزمایش سامانه به روزرسانی خودکار پایگاه داده، داده زیادی در دست نیست و به همین علت در مناطقی از نقشه کمبود نقاط مشاهده می‌شود.

از طرفی رستوران‌هایی که در ناحیه با تراکم نقطه بالا قرار دارند، احتمال بروز خطا را بالا برده و با توجه به دقت

همان‌گونه که در شکل‌های (۵-الف) و (۵-ب) قابل مشاهده است توزیع و تعداد رستوران‌ها در نقشه VGI کمتر از نقشه مرجع می‌باشد، ولی در عوض در برخی مناطق تراکم رستوران‌های نقشه VGI بیشتر است. با بررسی داده‌های مناطق پر تراکم دریافتیم که رستوران‌های این مناطق از رستوران‌های مشهور ناحیه مثل هاید و اسکان می‌باشند و سامانه در به روزرسانی خودکار داده‌های رستوران‌های مشهور ناحیه توفیق نسبی داشته است. نهایتاً با در نظر گرفتن نسبت تعداد رستوران‌هایی که کل منطقه را پوشش می‌دهند، مشخص می‌شود که اطلاعات مکانی مردم گستر توانسته حدود ۶۰ درصد از رستوران‌های ناحیه را پوشش دهد.

^۱ Completeness

پایین GPS گوشی‌های هوشمند، مکان بعضی از رستوران‌ها با رستوران‌های مجاور جابه‌جا می‌شود. اما برای دریافت اطلاعات رستوران‌ها، لزوم بالا بودن صحت توصیفی نام رستوران‌ها به عنوان عامل کنترل‌کننده مانع از بروز خطا در تشخیص صحیح اطلاعات به رستوران مورد نظر می‌گردد. به عنوان مثال بیشینه شعاعی که احتمال دارد رستوران‌های مجاور باهم اشتباه شوند ۱۴ متر می‌باشد. در این شعاع مطمئناً هیچ دو رستوران متشابهی از لحاظ اسمی یافت نمی‌شود. لذا ممکن است رستورانی از لحاظ مختصاتی به اشتباه به جای رستوران مجاورش در نظر گرفته شود، لیکن لزوم صحت بالای توصیفی داده‌های ورودی باعث می‌شود اسم رستوران درست تشخیص داده شود و در نتیجه اطلاعات در نظر گرفته‌شده برای رستوران مورد نظر (مثل کیفیت غذا، نوع سرویس‌دهی، امکانات نرم‌افزاری و ...) به درستی در پایگاه داده ذخیره شود. پس از انجام پروژه به نظر می‌رسد، به علت تراکم بالای رستوران‌ها در بعضی مناطق احتمال بروز خطا در فرآیند تناظریابی نقاط در دو نقشه مرجع و مردمی بالاست. لذا به منظور مرتفع کردن این مشکل می‌توان به

جای روش پیشنهادی در این مقاله (تناظر یابی سه مرحله‌ای مکانی و اسمی) از روش رستری هم برای تناظریابی رستوران‌ها در دو نقشه مرجع و مردمی به عنوان مرحله‌ای دیگر از تناظریابی به منظور کنترل بیشتر در جلوگیری از بروز خطا استفاده کرد. از طرفی برای کاهش حجم محاسبات در مناطق با تراکم بالا با حجم داده بالا در تناظریابی می‌توان از خوشه‌بندی رستوران‌ها با در نظر گرفتن میزان تراکمشان استفاده کرد. با این عمل می‌توان حوزه و تعداد جستجوها را کاهش داده و نهایتاً میزان محاسبات را کم کرد. به عنوان پیشنهاد نهایی، به کار بردن روش مذکور برای ورود عوارض پلی گونی به جای نقطه‌ای به پایگاه داده می‌تواند اطلاعات مفیدی در رابطه با کارایی روش مذکور در رابطه با سطوح بالاتر عوارض در اختیار قرار دهد. در این حالت مسلماً آلمان جدید صحت هندسی به مجموعه آلمان‌ها اضافه‌شده و لزوم محاسبه میزان دقت روابط توپولوژی اهمیت ویژه‌ای می‌یابد.

مراجع

- [1] Howe, J., 2006. The Rise of Crowdsourcing. Wired magazine June 2006, [Online]. Available at: <http://www.wired.com/wired/archive/14.06/crowds.html#Replay>
- [2] Göbel, S., Lutze, K. (1998). Development of meta databases for geospatial data in the WWW, In: Proceedings of the Sixth ACM International Symposium on Advances in Geographic Information Systems 1998, ACM, Washington DC, USA, pp. 94–99.
- [3] Ebadi, H., et al., 2004, Report Sets of Master Plan of Geographic Information System for Electrical Industry-Distribution Networks (Production, Editing and Updating Instructions), Faculty of Geodesy and Geomatics, K.N.Toosi University of Technology, Tehran, Iran.
- [4] Haj-Yehia, B., & Peled, A. (2004). Rule-Based System for Updating Spatial Data-Base. In ISPRS-XXth Congress, Istanbul, Turkey.
- [5] Ahmadi, F. F., & Ebadi, H. (2009). An integrated photogrammetric and spatial database management system for producing fully structured data using aerial and remote sensing images. Sensors, 9(4), 2320-2333.
- [6] Najar, C., Rajabifard, A., Williamson, I. P., & Giger, C. (2007). A framework for comparing Spatial Data Infrastructures on the basis of web services and metadata management: an Australian-Swiss case study.
- [7] Hatala, M., & Forth, S. (2003, May). A comprehensive system for computer-aided metadata generation. In Proceedings of 12th International Conference of The World Wide Web Consortium (WWW2003), Budapest, May 20 (Vol. 24).
- [8] Olfat, H., Kalantari, M., Rajabifard, A., Senot, H., & Williamson, I. P. (2013). A GML-based approach to automate spatial metadata updating. International Journal of Geographical Information Science, 27(2), 231-250.

- [9] Kalantari, M., Rajabifard, A., & Olfat, H. (2009, September). Spatial metadata automation. In Surveying and Spatial Sciences Institute Biennial International Conference.
- [10] Goodchild, M.F. (2007). "Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0." *International Journal of Spatial Data Infrastructures Research*, 2:24-32
- [11] Flanagan, A. J., & Metzger, M. J. (2008). The credibility of volunteered geographic information. *GeoJournal*, 72(3-4), 137-148.
- [12] Ather, A. (2009). A quality analysis of openstreetmap data. ME Thesis, University College London, London, UK.
- [13] Ciepluch, B., Jacob, R., Mooney, P., & Winstanley, A. (2010, July). Comparison of the accuracy of OpenStreetMap for Ireland with Google Maps and Bing Maps. In *Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences 20-23rd July 2010* (p. 337). University of Leicester.
- [14] <http://birds.audubon.org/christmas-bird-count> Retrieved October 14, 15,
- [15] Koukoletsos, T., Haklay, M., & Ellul, C. (2012). Assessing data completeness of VGI through an automated matching procedure for linear data. *Transactions in GIS*, 16(4), 477-498.
- [16] Koukoletsos, T. (2012). A Framework for Quality Evaluation of VGI linear datasets (Doctoral dissertation, UCL (University College London)).
- [17] Haklay, M. (2010). "How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets". *Environment and planning. B, Planning and design*, 37(4), 682.
- [18] Girres, J-F. and Touya, G. (2010). "Quality Assessment of the French OpenStreetMap Dataset". *Transactions in GIS*, 14(4):435-459.
- [19] Zielstra, D. and Zipf, A., 2010. A Comparative Study of Proprietary Geodata and VGI for Germany. In 13th AGILE International Conference on Geographic Information Science. Guimaraes, Portugal. 10-14 May 2010.
- [20] Bordogna, G., Carrara, P., Criscuolo, L., Pepe, M., & Rampini, A. (2014). A linguistic decision making approach to assess the quality of volunteer geographic information for citizen science. *Information Sciences*, 258, 312-327.
- [21] B. Vahedi , A. A. Alesheikh , S. Honarparvar . Quantitative Assessment of Pragmatic Quality of Volunteered Geographic Information Using Fuzzy Linguistic Quantifiers and OWA Operator. *JGST*. 2014; 3 (4) :65-76
- [22] Malczewski, J. (2006). Ordered weighted averaging with fuzzy quantifiers: GIS-based multicriteria evaluation for land-use suitability analysis. *International Journal of Applied Earth Observation and Geoinformation*, 8(4), 270-277.
- [23] Yager, R. R. (1996). Quantifier guided aggregation using OWA operators. *International Journal of Intelligent Systems*, 11(1), 49-73.
- [24] Devillers, R., & Jeansoulin, R. (Eds.). (2006). *Fundamentals of spatial data quality*. London: ISTE.
- [25] FGDC, 1998a. FGDC-STD-007.1-1998 Geospatial positioning accuracy standards. Part 1: Committee,[Online].Availableat:http://www.fgdc.gov/standards/standards_publications/index_html
- [26] FGDC, 1998b. FGDC-STD-007.3-1998 Geospatial positioning accuracy standards. Part 3: Available at: http://www.fgdc.gov/standards/standards_publications/index_html
- [27] Tveite, H. (1999). An accuracy assessment method for geographical line data sets based on buffering. *International journal of geographical information science*, 13(1), 27-47.
- [28] Ludwig, I., Voss, A., & Krause-Traudes, M. (2011). A Comparison of the Street Networks of Navteq and OSM in Germany. In *Advancing Geoinformation Science for a Changing World* (pp. 65-84). Springer Berlin Heidelberg.