نشریه علمی- پژوهشی علوم و فنون نقشه برداری، دوره سوم، شماره۴، اردیبهشت ماه ۱۳۹۳

برآورد کمی کیفیت عملی اطلاعات مکانی داوطلبانه به کمک کمیت‌سنج‌های زبانی فازی و عملگر تصمیم‌گیری چند معیاره OWA

بهزاد واحدی طرقبه*[۱]، علی اصغر آل شیخ[۲]، سپهر هنرپرور[۱]

[۱]دانشجوی کارشناسی ارشد سیستم‌های اطلاعات مکانی- دانشکده مهندسی نقشه برداری - دانشگاه صنعتی خواجه نصیر الدین طوسی
{b.vahedi, sep.honarparvar}@ kntu.ac.ir

[۲]استاد گروه مهندسی سیستم‌های اطلاعات مکانی- دانشکده مهندسی نقشه برداری - دانشگاه صنعتی خواجه نصیر الدین طوسی
(عضو قطب علمی مهندسی فناوری اطلاعات مکانی)
alesheikh@kntu.ac.ir

(تاریخ دریافت اسفند ۱۳۹۲، تاریخ تصویب اردیبهشت ۱۳۹۳)

چکیده

با وجود مزایای فراوانی که سیستم‌های اطلاعات مکانی داوطلبانه دارند، کیفیت داده‌های این سیستم‌ها بزرگترین محدودیت آن و محل بحث و تحقیق بسیاری از محققان بوده و تا کنون مقالات گوناگونی در این باره نوشته شده است. اغلب تحقیقات صورت گرفته در زمینه اطلاعات مکانی مردم‌گستر به کیفیت بیرونی[۱] و گاهی هم کیفیت درونی[۲] این داده‌ها پرداخته اند و کیفیت عملی[۳] کمتر مورد توجه واقع شده است. در این تحقیق با استفاده از کمیت‌سنج های زبانی و عملگر تصمیم گیری چندمعیاره OWA[۴] روشی برای ارزیابی کیفیت عملی اطلاعات مکانی داوطلبانه ارائه می‌گردد. در روش ارائه شده برای محاسبه کیفیت عملی ابتدا می‌بایست کیفیت درونی و بیرونی داده‌های تولید شده توسط داوطلبان محاسبه شود. به منظور محاسبه کیفیت درونی و بیرونی برای هر کدام چند شاخص سطح متوسط و سطح پایین در نظر گرفته شده است که با اعمال عملگر OWA روی این شاخص‌ها مقدار کیفیت محاسبه می‌شود. برای ارزیابی روش پیشنهادی میزان فرسایش بناهای تاریخی شهر تهران انتخاب گردید. با توجه به نتایج به دست آمده در این مطالعه موردی، مشخص گردید که تاثیرگذاری کیفیت بیرونی روی کیفیت عملی بیشتر از کیفیت درونی است و همچنین تخصص کاربر روی کیفیت عملی داده های تولید شده توسط وی تاثیر به‌سزایی دارد.

واژگان کلیدی : سیستم‌های اطلاعات مکانی داوطلبانه، عملگر تصمیم گیری چندمعیاره OWA ، کیفیت عملی اطلاعات مکانی، کمیت‌سنج‌های زبانی

* نویسنده رابط

۱ External quality
۲ Internal quality
۳ Pragmatic quality
٤ Ordered weighted averaging

Archive of SID
www.SID.ir

۱- مقدمه

از دیرباز تولید، انتشار و حتی گاهی استفاده از داده‌های مکانی به عنوان یک تخصص و مهارت مطرح بوده که نیاز به آموزش خاص دارد؛ که همواره افراد کمی از این مهارت برخوردار بوده اند. اما در سال‌های اخیر با توجه به پیشرفت‌های حوزه فناوری، و بالا رفتن دانش جوامع نیازهای مردم به داده‌های مکانی نیز تغییر یافته و باعث به وجود آمدن تحولی در الگوی تولید و مصرف داده‌های مکانی شده است. این تحول پایه گذار مفهوم جدیدی با نام اطلاعات مکانی داوطلبانه (مردم‌گستر) در حوزه جغرافیا و علوم مکانی شده است. با وجود مزایای فراوانی که سیستم‌های اطلاعات داوطلبانه دارند، کیفیت داده‌های این سیستم‌ها بزرگترین محدودیت آن و محل بحث و تحقیق بسیاری از محققان بوده و تا کنون مقالات گوناگونی در این زمینه نوشته شده است [۱].

برای کیفیت داده های داوطلبانه می‌توان سه جنبه را در نظر گرفت: کیفیت درونی، بیرونی و عملی[۲]. کیفیت درونی نشان دهنده دقت و صحت اطلاعات است. کیفیت بیرونی به اعتبار منبع تولید داده و کیفیت عملی هم به نیازهای کاربران و هدف در نظر گرفته شده برای پروژه بستگی دارد. هر کدام از این جنبه‌ها خود دارای شاخص های متفاوتی هستند که در بخش ۳-۳ به آنها پرداخته می‌شود.

اغلب تحقیقات صورت گرفته در زمینه اطلاعات مکانی داوطلبانه به کیفیت بیرونی و گاهی هم کیفیت درونی داده‌ها پرداخته اند و کیفیت عملی کمتر مورد توجه واقع شده است. ضمن اینکه نتیجه تحقیقات پیشین اکثرا به صورت کمی بیان شده است که شاید مناسب استفاده کنندگان عادی اطلاعات داوطلبانه نباشد. بنابراین، در این تحقیق به معرفی یک روش زبانی برای بررسی کیفیت اطلاعات داوطلبانه پرداخته شد. حسن روش پیشنهادی در این است که با استفاده از کمیت‌سنج‌های زبانی کثرت شاخص‌های سطح پایین و متوسط را کاهش داده و در نتیجه دشواری وزن‌دهی و تصمیم‌گیری توسط کارشناسان را برطرف می‌سازد. در این روش همچنین تعدادی از شاخص‌های مهم‌تر تعیین و برآورده می‌گردند. تاکید بیشتر این تحقیق بر روی کیفیت عملی اطلاعات داوطلبانه است اما کیفیت درونی و بیرونی نیز مورد بررسی قرار می‌گیرند.

این مقاله در شش بخش ارائه گردیده است: در بخش اول تحقیق، مقدماتی در مورد تعریف مساله، ضرورت و اهداف آن بیان می‌شود. در بخش دوم تحقیقات انجام شده در این زمینه مورد بررسی قرار می‌گیرد. در بخش سوم به بیان تئوری مورد نیاز برای انجام تحقیق پرداخته می‌شود. ابتدا مفهوم اطلاعات مکانی داوطلبانه مورد بررسی قرار می‌گیرد و پس از آن مفهوم کیفیت داده های مکانی به طور عام معرفی می‌شود. سپس کیفیت داده در سیستم های اطلاعات داوطلبانه بررسی شده و در قسمت آخر این بخش به روش ارزیابی چندمعیاره به کمک کمیت‌های زبانی پرداخته می‌شود. در بخش چهارم هم روش انجام کار بیان شده و توضیحاتی در مورد آن ارائه می‌شود. در بخش پنجم نتایج کمی حاصل از تحقیق ارائه می‌شوند و در نهایت در بخش ششم به بیان نتیجه گیری پرداخته می‌شود.

۲- پیشینه تحقیق

تا کنون روشهای متفاوتی برای ارزیابی کیفیت اطلاعات مکانی داوطلبانه استفاده شده اند. این روش‌ها را می‌توان به دو دسته کلی تقسیم کرد: [۳]

-مقایسه داده‌های داوطلبانه با داده‌های مرجع با کیفیت معلوم

-بررسی خود داده‌های داوطلبانه و ارزیابی پارامترهای کیفیت آنها به وسیله آزمون‌ها و روش‌های آماری

اکثر تحقیقات انجام شده در زمینه کیفیت اطلاعات مکانی داوطلبانه در دسته اول قرار می‌گیرند. این نوع ارزیابی، یک ارزیابی کمی است که اغلب برای کیفیت درونی مورد استفاده قرار می‌گیرد و در آن نتایج مستقیما به داده منتسب می‌شوند و به کاربر و یا سایر ویژگی‌های آن ربط نخواهند داشت. به طور مثال Haklay داده‌های سایت OpenStreetMap را با پنج نقشه رستری ۱:۱۰۰۰۰ سازمان نقشه برداری انگلستان مقایسه کرده است [۴]. تاکید وی بر صحت مکانی و تمامیت داده بوده و برای تناظریابی بین عوارض از روشهای بصری استفاده نموده است. Touya و Girres در فرانسه کیفیت داده‌های داوطلبانه OpenStreetMap را با مقایسه آنها با داده‌های سازمان نقشه برداری فرانسه ارزیابی کرده اند [۵]. آنها این مقایسه را برای چند منطقه که به صورت تصادفی

www.SID.ir
Archive of SID

انتخاب شده بودند انجام داده اند و المانهای صحت توصیفی، صحت معنایی، تمامیت، سازگاری منطقی و صحت زمانی را مورد بررسی قرار داده اند. Koukoletsos و همکاران کیفیت داده‌های OpenStreetMap را در انگلستان ارزیابی کرده و برای این منظور از داده‌های سازمان نقشه برداری انگلستان استفاده کرده اند [۶]. در این پژوهش، برای سادگی کار تنها عوارض خطی در نظر گرفته شده اند و پارامترهای تمامیت، صحت مکانی و صحت توصیفی داده‌ها ارزیابی شده اند. محققان دیگری هم مشابه این کار را برای کشورهای خود انجام داده اند [۷و۸و۹].

در دسته دوم تحقیقات در زمینه کیفیت اطلاعات مکانی داوطلبانه، نیازی به مقایسه داده‌ها با داده‌های مرجع نیست بلکه با بررسی خود داده‌ها می‌توان به اطلاعاتی درباره کیفیت آنها دست یافت. این دسته معمولا برای ارزیابی کیفیت بیرونی داده ها مورد استفاده قرار می‌گیرد. به طور مثال Van Exel و همکاران بیان می‌کنند که روش‌های مرسوم به کار رفته برای المان‌های کیفیت، فرض می‌کنند که داده‌های به کار رفته همگن و سازگار هستند در حالی که چنین فرضی در اطلاعات داوطلبانه صادق نیست. آنها در عوض مفهوم کیفیت اجتماع[۱] (مردم) را برای توصیف کمی کیفیت داده‌های داوطلبانه معرفی می‌کنند[۱۰]. این روش نظری دو بعد دارد: کیفیت کاربر و کیفیت عارضه. کیفیت کاربر بر اساس اطلاعات محلی، تجربه و شناخت او تعیین می‌شود و کیفیت عارضه توسط المان‌های رایج کیفیت محاسبه می‌گردد. Goodchild و Li بیان کردند که روش‌های سنتی ارزیابی کیفیت به دلیل پیچیدگی های فنی و هزینه‌های زمانی و مالی مناسب نیستند و ارزیابی جامعی از کیفیت ارائه نمی‌دهند. بنابراین یک روش جدید برای ارزیابی کیفیت ارائه داده اند. در این روش آنها سه جنبه از داده را بررسی می‌کنند: جنبه‌های مردم‌سپاری[۲] ، اجتماعی[۳] و جغرافیایی[۴] [۱۱]. هر کدام از این جنبه‌ها خود مزایا و معایبی دارند. به طور مثال بررسی کیفیت داده از جنبه مردم‌سپاری علی رغم اینکه در علم کامپیوتر کارایی فراوانی دارد مناسب داده‌های مکانی نیست. روش اجتماعی هنگامی کارایی دارد که تعداد مشارکت کنندگان بالا باشد و روش جغرافیایی از آنجا که از قوانین و روابط موجود در علم جغرافیا استفاده می‌کند برای جغرافی دان‌ها مناسب و جذاب است اما برای استفاده عموم کارایی چندانی ندارد.

بر اساس پژوهش‌های نگارندگان، از میان تحقیقاتی که تا کنون در زمینه اطلاعات مکانی داوطلبانه انجام شده، به کیفیت عملی تنها به صورت ضمنی پرداخته شده است. تنها یک مورد به طور مستقیم به بررسی کیفیت عملی پرداخته است [۲]. این مقاله که توسط Bordogna و همکاران نگاشته شده تنها با استفاده از روشهای زبانی به ارزیابی کیفیت عملی داده‌های مکانی داوطلبانه در قالب پروژه مطالعه یخچال‌های طبیعی اروپا پرداخته شده است.

۳- مبانی تئوری تحقیق

در این بخش به بیان تعاریف و مفاهیم به کار رفته در تحقیق پرداخته می‌شود.

۳-۱- اطلاعات مکانی مردم گستر

Goodchild در سال ۲۰۰۷ بیان کرد که زمین خوب نقشه برداری نشده است و بروز رسانی نقشه‌ها و بهبود کیفیت و صحت آنها به طور مداوم انجام نمی‌شود [۱۲]. او مدعی شد که مشکلات موجود در زمینه تولید داده‌های مکانی را می‌توان به کمک یک روش جدید اخذ داده برطرف کرد و این روش جدید را اطلاعات مکانی داوطلبانه نامید. وی عنوان می‌کند که هر یک از انسان‌ها می‌توانند به عنوان یک حسگر عمل کنند. در نظر او دنیا متشکل از شش میلیارد حسگر متحرک است که می‌توانند داده‌های مکانی منحصر به فردی از محیط پیرامون خود تولید کنند و این داده‌ها می‌توانند به عنوان مکمل داده‌های مکانی مرجع که توسط ارگان‌های رسمی تولید می‌شود مورد استفاده قرار گیرند. اگرچه در مقیاس جهانی چنین ادعایی ناممکن به نظر می‌رسد اما در مقیاس‌های خُرد و محلی می‌توان با چنین روشی به نتیجه ای مطلوب دست یافت. تعریف رسمی اطلاعات مکانی داوطلبانه بدین صورت می‌باشد: " بهره برداری از ابزار برای ایجاد، جمع آوری و انتشار داده‌های جغرافیایی که به طور داوطلبانه

---

۱ Crowd quality
۲ Crowd-sourcing
۳ social
٤ geographical

Archive of SID
www.SID.ir

توسط افراد تولید شده است"[۱۳]. به عقیده Goodchild اطلاعات مکانی داوطلبانه نوع خاصی از مفهوم محتوای کاربرتولید است که در دنیای کامپیوتر بسیار مطرح می‌باشد. امروزه پروژه‌های داوطلبانه مختلفی در سطح وب وجود دارند. بخش بزرگی از این پروژه‌ها را سایت‌های به‌اشتراک‌گذاری عکس‌های زمین مرجع شده تشکیل می‌دهند؛ که تعدادی از آنها عبارتند از: سایت‌های Flickr و Panoramio. اما دسته دیگری از پروژه‌های داوطلبانه که حاوی اطلاعات بسیار غنی‌تری هستند، سایت‌های نمایش و ویرایش نقشه هستند که از آن میان می‌توان به سایت های OpenStreetMap، Wikimapia و Google Map Maker اشاره نمود. از این بین سایت OpenStreetMap (به اختصار OSM) به دلیل فراهم آوردن امکان دانلود اصل داده‌ها مورد توجه بسیاری از محققان حوزه اطلاعات مکانی داوطلبانه واقع شده و تاکنون تحقیقات بسیاری در زمینه کیفیت داده‌های این سایت در مناطق مختلفی از دنیا انجام شده است. [۴،۵،۶،۷،۸]

البته تحقیق پیش رو را می‌توان نمونه ای از پروژه های علوم شهروندی[۵] نیز در نظر گرفت. این مفهوم برای توصیف شبکه ای از شهروندان که به عنوان ناظر در علوم مختلف مشارکت می‌کنند استفاده می‌شود. در چنین پروژه‌هایی شرکت کنندگان می‌بایست از معلومات پایه‌ای در حوزه مورد بررسی برخوردار باشند که این عامل باعث محدود شدن تعداد شرکت کنندگان در آنها می‌شود. به طور مثال می‌توان به پروژه سرشماری پرنده "کریسمس"[۶] در آمریکا اشاره کرد که در آن از پرنده شناسان مبتدی برای شمارش چند گونه مختلف از پرندگان استفاده شد [۱۴].

۳-۲- کیفیت داده های مکانی

از دهه ۱۹۸۰ میلادی و به دنبال دو تحول مهم در زمینه علوم مکانی یعنی ظهور سیستم‌های اطلاعات مکانی و افزایش حجم داده‌های حاصل از تصویربرداری ماهواره‌ای، نگرانی و توجه نسبت به کیفیت داده افزایش چشمگیری داشته است. دلیل این نگرانی این است که در نتیجه این تحولات دسترسی، استفاده و تبادل داده های مکانی در حال افزایش چشمگیری است و نیز تعداد استفاده کنندگان داده‌های مکانی که از کیفیت داده بی اطلاعند رو به فزونی است. از همین رو تا کنون منابع مختلفی به تحقیق در زمینه کیفیت داده های مکانی و تبیین استانداردهایی در این زمینه پرداخته اند که از آن جمله می‌توان به Aronoff، انجمن بین المللی کارتوگرافی، کمیته استاندارد تبادل داده مکانی آمریکا و سازمان بین المللی استاندارد (کمیته فنی ۲۱۱) اشاره کرد [۱۵]، [۱۶].

کیفیت بدین صورت تعریف می‌شود: " مجموع ویژگی‌ها و خصوصیات یک محصول که به توانایی آن در برآورده کردن نیازهای روشن و ضمنی مربوطند"[۱۷]. بدون توضیح "ویژگی‌ها"، این تعریف معنی خاصی نخواهد داشت.

خصوصیت یا ویژگی به طور سنتی جزء یا المان نامیده می‌شود بنابراین بایستی درباره المان‌های کیفیت داده بحث شود. تا کنون محققان زیادی از جمله ارگانهایی که پیشتر به آنها اشاره شد به بررسی این المان‌ها پرداخته اند و همه این منابع کم و بیش المان های یکسانی را برای کیفیت داده مکانی معرفی کرده اند. این المان‌ها عبارتند از: تاریخچه داده، صحت مکانی، صحت توصیفی، سازگازی منطقی، تمامیت، صحت معنایی، صحت زمانی، کاربرد-هدف-قیود، تغییرات در کیفیت و قدرت تفکیک. تنها تفاوتی که می‌توان به آن اشاره کرد این است که بعضی از المان‌ها در برخی از منابع به صورت صریح به عنوان یک المان کیفیت داده مکانی در نظر گرفته شده و در برخی دیگر به صورت ضمنی در نظر گرفته شده اند[۱۶].

مهمترین انگیزه برای توصیف کیفیت داده این است که استفاده کننده از یک مجموعه داده مکانی قادر باشد تا تناسب آن داده‌ها را برای اهداف و کاربرد خود برآورد کند. در حقیقت می‌توان گفت "تناسب برای هدف" مهمترین جنبه کیفیت داده و به خصوص داده‌های مکانی است. در واقع تمام المان‌های کیفیت داده در قالب تناسب برای هدف معنا می‌یابند. برای مثال ممکن است برای یک هدف خاص صحت مکانی داده اهمیت چندانی نداشته باشد اما تمامیت آن مهم باشد و برای یک کاربرد دیگر صحت مکانی ارزش بالایی داشته ولی تمامیت بی اهمیت باشد. بنابراین لازم است تولیدکنندگان داده‌های مکانی کیفیت داده‌های خود را در قالب المان‌های کیفیت عرضه

---

۱ Citizen science
۲ Christmas bird

www.SID.ir
Archive of SID

کنند تا کاربران بتوانند با استفاده از آنها تناسب آن مجموعه داده را برای اهداف خود ارزیابی نمایند. این نکته در سیستم‌های اطلاعات مکانی داوطلبانه نمود بیشتری هم پیدا می‌کند. بنابراین می‌بایست ابتدا به مفهوم کیفیت داده در سیستم‌های اطلاعات مکانی داوطلبانه به طور مفصل تری پرداخت.

۳-۳-کیفیت داده در سیستم‌های اطلاعات مکانی داوطلبانه

برای ارزیابی کیفیت داده‌های مکانی به صورت عام تا کنون روشهای مختلفی ارائه شده اند. بطور مثال با مقایسه یک مجموعه داده با داده های مرجع و رسمی که توسط ارگان‌های رسمی نقشه برداری تولید شده اند، می‌توان کیفیت آن را ارزیابی کرد. اما این روش‌ها عموما برای داده هایی که به صورت استاندارد و با توجه به دستورالعمل های تولید داده مکانی ایجاد شده اند مناسب اند و نمی‌توان آنها را برای اطلاعات مکانی داوطلبانه به کار برد. این نوع از داده‌ها تفاوت زیادی با داده‌های رسمی دارند. فقدان استاندارد و نبود فراداده به همراه عدم آگاهی از انگیزه‌های کاربران و میزان اعتبار داده‌های وارد شده توسط آنها باعث می‌شود که داده‌هایی ناهمگن و با کیفیت نامشخص به وجود آیند. در نتیجه روشهای اندازه گیری کیفیتی که برای داده‌های رسمی و با دقت یکسان طراحی شده اند مناسب این نوع از داده‌ها نیستند. دراطلاعات مکانی داوطلبانه ممکن است یک ناحیه پوشش کامل و ناحیه ای دیگر پوشش کم داشته باشد و یا حتی بدون پوشش باشد (تمامیت). ممکن است داده‌های یک منطقه از نظر مکانی دقیق تر از داده‌های یک منطقه دیگر باشند (صحت مکانی). بعضی داده‌ها به خوبی توصیف شده و دارای اقلام توصیفی مناسبی باشند و برخی دیگر خیر (صحت موضوعی). ممکن است توپولوژی موجود نباشد یا دارای خطا باشد (سازگاری منطقی) [۱۶]. از همین رو در اغلب پژوهش‌ها در زمینه کیفیت داده های داوطلبانه المان های تمامیت، صحت مکانی، صحت توصیفی و سازگاری منطقی از بین المان های ده گانه کیفیت به عنوان المان های موثر انتخاب و مطالعه شده‌اند [۶]. با توجه به ماهیت داده های داوطلبانه روشهای ارزیابی کیفیت این گونه داده‌ها ممکن است نه تنها برای پروژه های مختلف بلکه برای یک پروژه و بر اساس فضا، تولید کننده و مصرف کننده داده متفاوت باشد. بنابراین ایجاد روشهایی که مناسب داده‌های داوطلبانه باشند ضروری به نظر می‌رسد.

کیفیت داده های داوطلبانه را می‌توان درسه دسته مختلف مطالعه کرد: کیفیت درونی، بیرونی و عملی. کیفیت درونی به محتوای داده بستگی دارد همچون دقت و صحت متون، تصاویر و اندازه گیری ها. کیفیت بیرونی به مشخصات و اعتبار تولیدکنندگان داده (داوطلبان) مربوط است و کیفیت عملی نیز به هدف در نظر گرفته شده برای آن پروژه بستگی دارد و ممکن است از پروژه ای به پروژه دیگر تغییر کند. برای بیان بهتر هر کدام از این جنبه ها می‌توان برای آنها شاخص هایی در نظر گرفت. برای مثال شاخص های کیفیت بیرونی عبارتند از اعتبار فرد یا سازمان تولید کننده داده، میزان تخصص فرد داوطلب در رابطه با پروژه و یا نقش وی در سازمان مشارکت کننده در پروژه. شاخص های کیفیت درونی هم عبارتند از اعتبار مکانی، زمانی و یا متنی داده، اعتبار اندازه گیری‌های انجام شده و اعتبار دسته بندی. در تحقیقاتی که تا کنون در زمینه کیفیت داده های داوطلبانه انجام شده به کیفیت بیرونی و درونی بسیار پرداخته شده است اما کیفیت عملی یا در واقع همان تناسب برای هدف آن چنان مورد توجه قرار نگرفته است.

۴-۳- روش ارزیابی چندمعیاره‌ی OWA با کمک کمیت‌سنج های زبانی

کیفیت اطلاعات مکانی داوطلبانه تحت تاثیر عوامل مختلفی از جمله صحت مکانی، صحت زمانی و صحت توصیفی داده‌ها و یا میزان اعتماد به داوطلبان قرار دارد. بنابراین برای ارزیابی کیفیت این داده‌ها باید جنبه های مختلفی را مورد بررسی قرار داد و در نتیجه فضای تصمیم در رابطه با کیفیت اطلاعات مکانی داوطلبانه یک فضای چند بعدی است. برای رسیدن به دید جامعی از میزان کیفیت هر داده و توصیف نهایی آن نیاز است فضای چند بعدی تصمیم گیری به فضای تک بعدی تبدیل شود. لذا می‌توان از عملگرهای تصمیم گیری چند معیاره برای این منظور استفاده کرد.

Archive of SID

www.SID.ir

از طرف دیگر در هر کاربردی از سیستم های اطلاعات مکانی داوطلبانه، تعیین میزان اهمیت هر شاخص کیفیت (توصیفی، مکانی، و ...) نیازمند آگاهی از گستره‌ی وسیعی از علوم مختلف همچون سیستم‌های اطلاعات مکانی، علوم اجتماعی، تخصص در زمینه ی برنامه ی کاربردی مورد نظر و ... می‌باشد. پس به عملگری نیاز است که بتواند معیارها و وزن های 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 سناریو های مختلف مورد بررسی قرار دهد. به همین دلیل با توجه به قابلیت سازگاری بالای عملگر OWA[۷] با کمیت‌سنج های زبانی، در این تحقیق از این عملگر برای به دست آوردن مقادیر کیفیت نهایی یا عملی استفاده شده است [۱۸] .

تاکنون روش های مختلفی برای فرآیند تصمیم گیری چندمعیاره ارائه شده است که از جمله‌ی آنها می‌توان به روش عملگرهای صفر و یکی (Boolean) با قوانین ترکیبی غیرجبرانی و روش ترکیب خطی وزنی (WLC)[۸] با قوانین ترکیبی جبرانی اشاره کرد. روش OWA حالت عمومی تری از موارد فوق الذکر می‌باشد [۲۰][۱۹]. به عبارت دیگر این روش از خانواده‌ی فرآیند ترکیب چندمعیاره به حساب می‌آید [۲۱]؛ که در واقع دو سری از وزنها که اولی وزن های 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 از نوع دوم باشد اپراتور OWA به صورت زیر عمل می‌کند:

$$\mathrm{OWA}(a_1, a_2, \dots, a_n) = \sum_{i=1}^{n} w_i\, b_i \qquad (۱)$$

که در آن bi وزن ترتیبی است؛ به صورتی که از بزرگترین مقدار a تا کوچکترین آن به صورت نزولی مرتب شده است. این عملگر هم مثل تمام عملگرهای تصمیم گیری چندمعیاره یک فضای n بعدی را بر روی فضای یک بعدی تصویر می‌کند. به عنوان مثال اگر بتوان مساله را تصمیم گیری چندمعیاره در رابطه با تلیفق نقشه های فاکتور در نظر گرفت، آنگاه وزن های ترتیبی وزن مقادیر پیکسل‌ها در هر نقشه به صورت نرمالیزه و مرتب شده، و امتیاز معیار هر نقشه وزن اهمیت نسبی معیارها خواهد بود. لذا فرمول مقادیر بدست آمده برای هر پیکسل به صورت زیر در می‌آید:

$$OWA_i = \sum_{j=1}^{n} \left(\frac{u_j\, v_j}{\sum_{j=1}^{n} u_j\, v_j} \right) z_{ij} \qquad (۲)$$

که در آن Vj مقادیر وزن های ترتیبی برای هر نقشه و uj وزن هر نقشه‌ی معیار می‌باشد، Zij هم مقادیر مرتب شده‌ی هر پیکسل براساس ترتیب V هاست.

همانگونه که اشاره شد می‌توان عملگرهای بولین مثل MIN و MAX و همچنین عملگرهای ترکیبی خطی وزنی WLC را از OWA استخراج کرد. برای این کار کافی است وزن های V را متناسب با عملگر متناظر مقداردهی کرد، مثل نمونه های زیر:

V=V(1,0,0,...,0)	Maximum
V=V(0,0,0,...,1)	Minimum
V=V(1/n,1/n,...,1/n)	WLC

متاسفانه این روش برای تعداد معیارهای زیاد و داده های با حجم بالا مناسب نیست چرا که برای تصمیم گیران بسیار دشوار و حتی غیرممکن است که برای تعداد زیادی معیار مقدار عددی تخصیص دهند و در فرایند ترکیب برای هر کدام وزنی را مشخص کنند. برای حل این مشکل می‌توان از کمیت‌سنج های زبانی فازی استفاده کرد. به نوعی می‌توان با استفاده از عباراتی مثل حداکثر، حداقل یکی، تمام موارد و ... و اختصاص وزن متناظر با آن عبارت به صورت عددی فرایند تصمیم گیری را ساده نمود. کمیت‌سنج های فازی به دو دسته‌ی مطلق و نسبی تقسیم می‌شوند [۱۸]. کمیت‌سنج های مطلق مثل حدود ۴، و حداقل ۱۰ به صورت مطلق به عدد اشاره می کنند ولی کمیت‌سنج های نسبی عباراتی مثل حداقل، حداکثر، به صورت ناقص به یک عدد اشار دارند [۲۲]. به عنوان مثال اگر Q نمایانگر مفهوم حداکثر باشد عبارت Q(0.95)=0.75 یعنی ۹۵ درصدِ یک عدد، ۰/۷۵ با مفهوم حداکثر تطابق دارد. می‌توان مقادیر وزن های ترتیبی را با استفاده از کمیت‌سنج Q از رابطه‌ی زیر بدست آورد.

۱ Ordered Weighted Averaging
۲ Weighted Linear Combination

www.SID.ir
Archive of SID

$$v_j = Q\left(\frac{j}{n}\right) - Q\left(\frac{j-1}{n}\right) \quad (۳)$$

با تغییر ثابت توانی فرمول ۳ می‌توان طیف وسیعی از عبارات زبانی را پوشش داد.

در نهایت روش کار به این صورت است که در ابتدا با توجه به اهمیت نسبی معیارها، معیارها وزن دهی شده و به صورت نزولی مرتب می شوند (u_j)، سپس مقادیر معیارها هم با توجه به ترتیب وزن های u_j مرتب می شوند (z_{ij}). آنگاه با توجه به توضیحات داده شده در رابطه با کمیت سنج های زبانی برای عبارت زبانی مورد نظر کمیت سنج زبانی فازی مورد نظر (Q) بدست می آید.

حال می توان با استفاده از فرمول ۳ مقدار وزن های ترتیبی (v_j) را به دست آورد. در نهایت با قرار دادن مقادیر به دست آمده از مراحل قبل (v_j، z_{ij} و u_j) با استفاده از فرمول ۲ مقادیر نهایی عملگر OWA محاسبه می‌شوند.

۴- روش انجام کار

در این بخش به توضیح روش ارزیابی کیفیت عملی در قالب مطالعه‌ی موردی بناهای تاریخی شهر تهران پرداخته می شود. برای این منظور، مواردی که توسط گردشگران و سازمان‌ها مقدار دهی شده اند وارد فرآیند تصمیم گیری چندمعیاره می‌شوند و خروجی این فرایند در قالب یک مقدار برای کیفیت عملی برای هر مورد ارسال شده توسط داوطلبان محاسبه می‌شود.

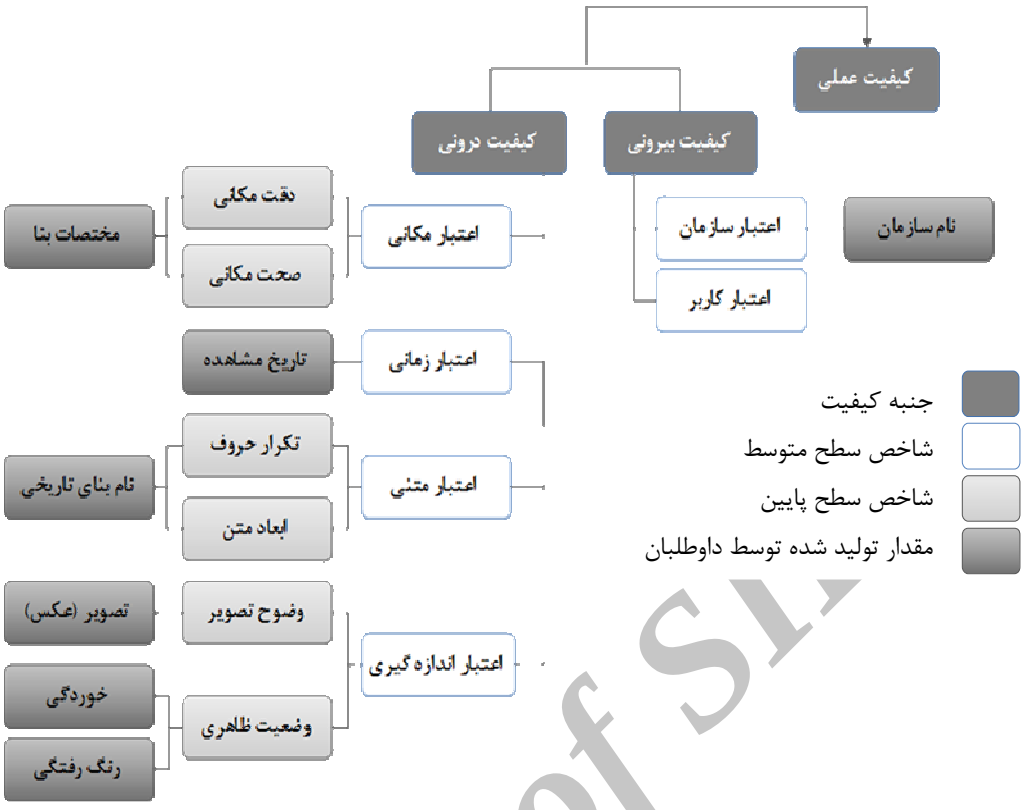
برای به دست آوردن کیفیت عملی می‌بایست ابتدا کیفیت بیرونی و کیفیت درونی را محاسبه کرد[۲]. برای سادگی در محاسبه کیفیت درونی و بیرونی برای هر کدام چند شاخص سطح متوسط و سطح پایین در نظر گرفته شده است؛ کیفیت درونی دارای چهار شاخص سطح متوسط اعتبار مکانی، اعتبار زمانی، اعتبار متنی و اعتبار اندازه‌گیری است و کیفیت بیرونی هم دارای دو شاخص اعتبار سازمان و اعتبار کاربر می‌باشد. شکل ۱ طرح کلی آیتم‌های ورودی و شاخص‌های سطح متوسط و پایین را نشان می‌دهد. این شاخص‌ها بر اساس تعریف کیفیت درونی و بیرونی در نظر گرفته شده اند و همان طور که ملاحظه می‌شود بعضی از آنها دارای یک یا چند شاخص سطح پایین هستند.

ابتدا می بایست با اعمال توابع و روابطی معین (توابع تبدیل) بر روی مقادیر ورودی داوطلبان - با توجه به نوع این داده‌ها- شاخص‌های سطح پایین را به دست آورد. سپس بر اساس ویژگی‌های پروژه به آنها وزن معینی اختصاص داد. به عنوان مثال با استفاده از هیستوگرام می توان میزان وضوح تصویر را که یک شاخص سطح پایین است، به صورت عددی به دست آورد و یا با استفاده از توابع مقایسه رشته‌ای می‌توان میزان صحت نام ورودی برای بنا را تعیین کرد. سپس به هر کدام از شاخص‌های سطح پایین از سه مقدار بالا، متوسط و ضعیف با توجه مقدار عددی این توابع و همچنین نظر کارشناسان یک مقدار تخصیص داده می شود. به طور مثال باتوجه به اینکه مطالعه در مورد بناها و مکان های گردشگری در داخل شهر تهران است، در مناطقی که دارای تراکم بالای بناهای گردشگری هستند، نیاز به دقت مختصاتی بالا است تا مکان های گردشگری با مختصات نزدیک به هم قابل تشخیص باشند؛ بنابراین اعتبار مکانی نسبت به سایر شاخص‌ها وزن بالاتری می‌گیرد. به همین صورت برای تمامی شاخص های سطح پایین مقادیر وزن انتخاب می‌شود و این وزن‌ها به عنوان وزن معیار در روش OWA مورد استفاده قرار می‌گیرد.

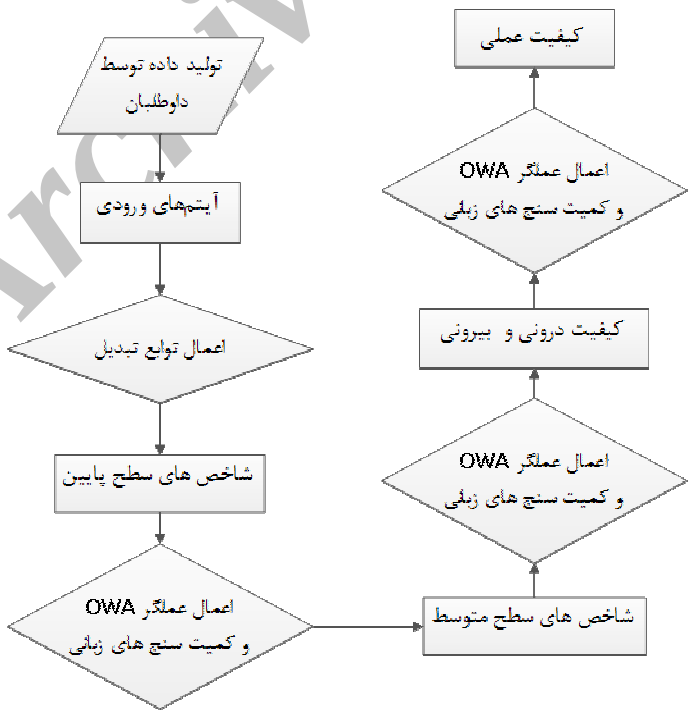
فلوچارت روند کلی کار را در شکل ۲ ملاحظه می‌کنید.

Archive of SID

www.SID.ir



شکل ۱- طرح کلی شاخص‌های پیشنهادی و آیتم‌های ورودی



شکل ۲- فلوچارت مراحل کار

Archive of SID

www.SID.ir

برای وزن دهی به شاخص های سطح پایین از فرمول زیر استفاده شده است. که در آن r_j رتبه‌ی معیار J ام می‌باشد.

$$W_j = \frac{n - r_j + 1}{\sum_{k=1}^{n} n - r_k + 1} \quad (۴)$$

با توجه به اینکه در این مقاله از کمیت‌سنج نسبی استفاده شده است، از یکی از ساده ترین و معروف‌ترین کمیت‌سنج‌های زبانی نسبی با عنوان RIM[۹] بهره گرفته می‌شود که از فرمول $Q(P) = p^{\alpha}$ تبعیت می‌کند. در این فرمول α می‌تواند با توجه به عبارت زبانی مقادیری از نزدیک به صفر تا بینهایت به خود بگیرد. مثلا برای برآورده کردن "همه معیارها" می‌توان مقداری بزرگ نزدیک به ۱۰۰۰ به α اختصاص داد. یعنی اینکه اگر هدف برآورده کردن تمام معیارهای مد نظر باشد از این مقدار برای α استفاده شده است. بنابراین فرمول به دست آمده برای وزن های ترتیبی به صورت زیر در می‌آید:

$$v_j = \left(\sum_{k=1}^{j} u_k\right)^{\alpha} - \left(\sum_{k=1}^{j-1} u_k\right)^{\alpha} \quad (۵)$$

با به دست آمدن وزن های ترتیبی و معیار می‌توان مقادیر عملگر OWA را برای هر شاخص سطح متوسط به دست آورد. با این کار به هر شاخص سطح متوسط، مقداری اختصاص می‌یابد. حال اگر عملگر OWA روی مقادیر شاخص های سطح متوسط اعمال شود می‌توان مقادیر کیفیت درونی و بیرونی و در نهایت عملی را به دست آورد. (شکل ۲)

پس از وزن دهی و طراحی شمای کلی آیتم ها، شاخص‌ها و جنبه های کیفیت، از ۲۵ داوطلب خواسته شد تا مشاهدات خود را از بناهای تاریخی تهران در قالب ۸ آیتم و به کمک پرسشنامه‌ای که در اختیار آنها قرار داده شده بود، بیان کنند. پنج نمونه از مقادیر تولید شده توسط کاربران را در جدول ۱ مشاهده می‌کنید.

۵- نتایج

برای محاسبه کیفیت عملی ابتدا با استفاده از توابع انتقال، مقادیر شاخص‌های سطح پایین مشخص شدند. سپس به هر کدام از آنها بر اساس کاربرد مورد نظر وزنی تخصیص یافت و با اعمال عملگر OWA روی این شاخص ها، شاخص‌های سطح متوسط محاسبه شدند سپس وزن این شاخص‌ها نیز تعیین شده و با اعمال OWA روی آنها میزان کیفیت بیرونی و درونی برای هر مجموعه داده که توسط یک داوطلب تولید شده بود مشخص گردید. در نهایت یک بار دیگر با اعمال عملگر OWA روی مقادیر کیفیت بیرونی و درونی، مقدار کیفیت عملی محاسبه شد. در جدول ۲ نتایج حاصل از ارزیابی کیفیت داده‌های ارائه شده در جدول ۱ مشاهده می‌شوند.

اعداد این جدول همگی نرمال شده اند. بنابراین در مقایسه دو داده، آن که مقدار بالاتری دارد از کیفیت بهتری برخوردار است. هر سطر از این جدول مربوط به یک مجموعه داده در مورد یک بناست که توسط یک فرد خاص تهیه شده است. به طور مثال در مورد داوطلب شماره ۴ که راجع به مسجد نظام الدوله داده تولید کرده است اعتبار اندازه‌گیری برابر ۰/۷ و اعتبار مکانی برابر ۰/۸ است. این بدان معناست که کیفیت مکانی داده تولید شده توسط وی بهتر از کیفیت اندازه‌گیری آن است. همچنین کیفیت بیرونی داده‌ی تولیدی توسط وی ۱، یعنی دارای کیفیت کامل است که با توجه به تخصص وی که باستان شناسی است دور از انتظار نیست و در نهایت کیفیت عملی داده وی ۰/۹ است که مقدار بالایی می‌باشد. این بدان معناست که این داده برای کاربرد مورد نظر در این پروژه تا حد زیادی مناسب است. نکته شایان ذکر دیگر این است که به دلیل بالا بودن صحت داده های متنی، مقادیر تخصیص یافته برابر یک می‌باشند و از آنجا که همه با هم برابرند تاثیری روی تفاوت کیفیت‌های درونی مختلف ندارد.

۱ Regular Increasing Monotone

Archive of SID
www.SID.ir

جدول ۱- مقادیر ورودی تولید شده توسط داوطلبان

نام آیتم / تولیدکننده	عکس	مختصات	رنگ رفتگی	خوردگی	نام بنا	تاریخ مشاهده	تخصص گردشگر	سازمان
داوطلب ۱	1.Jpg	۳۵,۶۹۹۹۴۴ ۵۱,۳۳۷۷۶۸	کم	متوسط	برج آزادی	۹۲/۰۹/۲۰	عکاس	-
داوطلب ۲	2.Jpg	۳۵,۶۹۰۶۱۰ ۵۱,۴۳۴۴۸۹	کم	کم	مجلس شورای ملی	۹۱/۰۷/۰۳	نقشه بردار	-
داوطلب ۳	3.Bmp	۳۵,۶۷۹۶۱۴ ۵۱,۴۲۱۰۵۷	زیاد	متوسط	شمس العماره	۹۰/۰۴/۰۳	کارمند	محیط زیست
داوطلب ۴	4.Jpg	۳۵,۶۸۲۵۰۷ ۵۱,۴۲۳۰۷۷	متوسط	متوسط	مسجد نظام الدوله	۹۲/۰۵/۰۲	باستان شناس	میراث فرهنگی
داوطلب ۵	5.Jpg	۳۵,۶۷۲۷۳۳ ۵۱,۴۰۶۲۹۴	زیاد	زیاد	کلیسای سورپ گورک	۹۱/۱۱/۱۸	آزاد	-

جدول ۲- نتایج ارزیابی کیفیت در سطح شاخص‌ها و جنبه های کیفیت

جنبه یا شاخص کیفیت / تولید کننده	شاخص‌ها کیفت سطح متوسط				جنبه های کیفیت		
	اعتبار اندازه گیری	اعتبار مکانی	اعتبار متنی	اعتبار زمانی	بیرونی	درونی	عملی
داوطلب ۱	۰/۷	۰/۳	۱	۰/۴	۰/۷	۰/۵	۰/۵
داوطلب ۲	۰/۶	۰/۲	۱	۰/۴	۰/۸	۰/۴	۰/۶
داوطلب ۳	۰/۷	۰/۶	۱	۰/۸	۰/۹	۰/۷	۰/۷
داوطلب ۴	۰/۷	۰/۸	۱	۰/۶	۱	۰/۷	۰/۹
داوطلب ۵	۰/۵	۰/۷	۱	۰/۵	۰/۵	۰/۶	۰/۵

۶- نتیجه گیری

با توجه به تحقیقات انجام گرفته توسط نگارندگان مشخص گردید که اغلب پژوهش های صورت گرفته در زمینه کیفیت اطلاعات مکانی داوطلبانه به جنبه های بیرونی و درونی کیفیت پرداخته اند. بنابراین این تحقیق برآن شد تا با استفاده از کمیت‌سنج‌های زبانی و روش OWA روشی را برای ارزیابی کیفیت عملی اطلاعات مکانی داوطلبانه ارائه نماید. با توجه به نتایج به‌دست آمده در این مطالعه مشخص گردید که تاثیرگذاری کیفیت بیرونی روی کیفیت عملی بیشتر از از کیفیت درونی است.

یکی دیگر از نتایج به دست آمده این است که تخصص داوطلب روی کیفیت نهایی تاثیر به‌سزایی دارد. بدین صورت که هر چه تخصص مشارکت‌کننده به تخصص مورد نیاز در یک پروژه نزدیک‌تر باشد، داده هایی که توسط وی تولید می‌شوند از کیفیت درونی و در نهایت عملی بالاتری برخوردار خواهد بود؛ بنابراین می‌توان از ابتدا وزن بیشتری به تخصص داوطلب داد. با توجه به طولانی بودن فرآیند دریافت اطلاعات و محاسبات مکرر مقدار کیفیت عملی در حین این پروسه، مشخص شد که با افزایش تعداد داوطلبان برای هر بنای تاریخی میزان پراکنش مقادیر کیفیت عملی کاهش می‌یابد و به عبارتی مقادیر کیفیت عملی به دست آمده از هر داده‌ی تولید شده همگرا می‌شود. با توجه به توانایی روش فوق الذکر در فرآیند استخراج کیفیت عملی داده های داوطلبانه بدون نیاز به نظارت مستقیم کارشناسان و دخالت مستمر در

www.SID.ir
Archive of SID

خلال فرآیند ارزیابی کیفیت داده‌های داوطلبانه، امکان خودکار سازی فرآیند اخذ داده‌های مذکور و محاسبه کیفیت آنها برای به روز رسانی پایگاه داده های مکانی بسیار مناسب می‌باشند. بنابراین، از این روش می توان برای به روزرسانی اتوماتیک برنامه‌های کاربردی مختلف مکانی نظیر سیستم‌های توصیه گر مکانی، مدیریت بحران، ترافیک و... استفاده کرد.

مراجع

[1] Antoniou, V., Haklay, M., and Morley, J. (2010). "A step towards the improvement of spatial data quality of Web 2.0 geo-applications: the case of OpenStreetMap". In Proceedings from GIS Research UK 18th Annual Conference. p. 197-201. University College London, UK. 14-16 Apr 2010.

[2] Bordogna, G., Carrara, P., Criscuolo, L., Pepe, M., and Rampini, A. (2014). "A linguistic decision making approach to assess the quality of volunteer geographic information for citizen science". Information Sciences, 258, 312-32.

[3] Koukoletsos, T. (2012). "A Framework for Quality Evaluation of VGI linear datasets". Ph.D. thesis, University College London, UK.

[4] 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.

[5] Girres, J-F. and Touya, G. (2010). "Quality Assessment of the French OpenStreetMap Dataset". Transactions in GIS, 14(4):435-459.

[6] Koukoletsos, T., Haklay, M., and Ellul, C. (2012). "Assessing data completeness of VGI through an automated matching procedure for linear data". Transactions in GIS, 16(4), 477-498.

[7] Cipeluch, B., Jacob, R., Winstanley, A., and Mooney, P. (2010). "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, p. 337-340. Leicester, UK.

[8] Ather, A., 2009. "A Quality Analysis Of OpenStreetMap Data". MSc Thesis, University College London, UK.

[9] Ueberschlag, A. (2010). "A first assessment of the OpenStreetMap quality in Switzerland". MSc Thesis, EPFL, Lausanne, Switzerland.

[10] Van Exel, M., Dias, E., and Fruijtier, S. (2010). "The impact of crowdsourcing on spatial data quality indicators". In Sixth international conference on Geographic Information Science. Zurich, Switzerland. 14-18 Sep 2010

[11] Goodchild, M. F., and Li, L. (2012). "Assuring the quality of volunteered geographic information". Spatial statistics, 1, 110-120.

[12] 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

[13] Flanagin, A. J., and Metzger, M. J. (2008). "The credibility of volunteered geographic information". GeoJournal, 72:137-148.

[14] http://birds.audubon.org/christmas-bird-count, accessed 15 Jan 2014

[15] Aronoff, S. (1989). "Geographic information systems: a management perspective". Ottawa: WDL.

[16] Van Oort, P.V. (2006). "Spatial data quality: from description to application". PhD thesis, Netherlands Geodetic Commission, Delft, The Netherlands

Archive of SID
www.SID.ir

[17] ISO (International Standardization Organization), 2002. ISO 19113:2002 Geographic information Quality principles.

[18] 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.

[19] Asproth, V., Holmberg, S.C., Ha°kansson, A. (1999). “Decision Support for spatial planning and management of human settlements”. International Institute for Advanced Studies in Systems Research and Cybernetics. In: Lasker, G.E. (Ed.), Advances in Support Systems Research, vol. 5. Windsor, Ont., Canada, pp. 30–39

[20] Malczewski, J., and Rinner, C. (2005). “Exploring multicriteria decision strategies in GIS with linguistic quantifiers: A case study of residential quality evaluation”. Journal of Geographical Systems, 7(2), 249-268.

[21] Yager, R.R. (1996). “Quantifier guided aggregation using OWA operators”. International Journal of Intelligent Systems 11, 49–73.

[22] Zadeh, L.A. (1983). “A computational approach to fuzzy quantifiers in natural languages”. Computers & Mathematics with Applications. 9, 149–184.

Archive of SID

www.SID.ir