

ارزیابی شباهت هندسی اطلاعات مکانی داوطلبانه در شبکه معابر درون شهری

علیرضا چهرقان، دانش آموخته دکتری، دانشکده مهندسی معدن، دانشگاه صنعتی سهند، تبریز، ایران

رحیم علی عباسپور (مسئول مکاتبات)، استادیار، دانشکده مهندسی نقشه برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی، دانشگاه تهران، تهران، ایران

Email: abaspour@ut.ac.ir

دریافت: ۱۳۹۶/۰۱/۱۹ پذیرش: ۱۳۹۶/۰۸/۰۱

چکیده

دسترسی سریع و به‌روز به اطلاعات مرتبط با شبکه راه‌های درون شهری یکی از ضروریات است که در کاربردهای مختلف در حوزه حمل‌ونقل احساس می‌شود. از این رو در این تحقیق تلاش شده است با ارزیابی شباهت هندسی اطلاعات مکانی داوطلبانه (VGI) از طریق مقایسه آنها با مجموعه داده‌های رسمی، جایگزینی مطمئن به محققین ارائه داد. برای این منظور راهکاری نوین ارائه می‌گردد که ضمن شناسایی عوارض متناظر در دو مجموعه داده VGI و رسمی، هر یک از این عوارض از منظر معیارهای طول، فاصله، جهت، حریم مشترک، پیچیدگی، اندازه و اعوجاج با یکدیگر مورد مقایسه قرار می‌گیرند. از این رو منطقه شش تهران به عنوان منطقه مورد مطالعه و مجموعه داده با مقیاس ۱:۲۰۰۰ (تهیه شده توسط سازمان نقشه‌برداری کشور) به عنوان مجموعه داده رسمی انتخاب گردید. نتایج نشان داد که عوارض موجود در مجموعه داده OSM (نمونه‌ای از اطلاعات مکانی داوطلبانه) به طور میانگین ۷۸ درصد شباهت مکانی با مجموعه داده رسمی دارند. همچنین جهت ارزیابی سیر بهبود دقت هندسی مجموعه داده‌های VGI، شبکه راه‌های درون شهری در سال‌های ۲۰۱۴ تا ۲۰۱۷ مورد ارزیابی قرار گرفت. نتایج نشان داد که که ضمن افزایش ۲۷/۸ درصدی مشارکت مردم در ترسیم عوارض، میانگین درصد شباهت مکانی نیز به مقدار ۵/۷۷ درصد بهبود یافته است. از این رو محققین در بسیاری از کاربردهای مرتبط با حوزه حمل و نقل می‌توانند از مجموعه داده‌های VGI به عنوان جایگزینی برای مجموعه داده رسمی استفاده کنند.

واژه‌های کلیدی: ارزیابی شباهت هندسی، اطلاعات مکانی داوطلبانه، شبکه راه شهری، مجموعه داده رسمی، مجموعه داده OSM.

۱. مقدمه

اطلاعات مرتبط با شبکه راه‌های درون شهری به عنوان یکی از اصلی‌ترین مولفه‌های زیرساخت اطلاعات مکانی مطرح است که کاربردهای بسیار وسیعی در حوزه مدیریت حمل و نقل شهری دارد. از جمله کاربردهای داده شبکه راه‌های درون شهری می‌توان به تعیین و ارزیابی نقاط حادثه‌خیز شهری Beheshtinia and [Bargbid, 2017; Rasafi et al. 2012], مدیریت حمل و نقل شهری [Alenouri et al. 2015; Mirbaha et al. 2014], یافتن کوتاه‌ترین و مناسب‌ترین مسیر [Khammar et al. 2016; Zolfaghari and Karke Abadi, 2013], طراحی خطوط تاکسیرانی و اتوبوسرانی [Afandizadeh et al. 2010; Teymourian et al. 2016] اشاره داشت. از این‌رو، در نظر گرفتن داده‌های مکانی به‌روز و دارای دقت مکانی قابل قبول، می‌تواند نتایج حاصل از تحقیقات مورد اشاره را به واقعیت نزدیکتر گرداند. تولید اطلاعات مکانی بصورت سنتی، دارای فرایندی طولانی شامل جمع‌آوری داده، پردازش و انتشار داده است. این فرآیند نیازمند تجهیزات پرهزینه و نیروهای متخصص بوده که تولید اطلاعات مکانی را محدود به مراکز خاص می‌کند. این رویکرد دارای مزایایی مانند اطمینان، دقت و کامل بودن محتوای اطلاعات بوده [Flanagin and Metzger, 2008]، ولی به دلیل پرهزینه بودن فرایند تولید اطلاعات مکانی، استفاده از آن نیز محدود است. با گسترش فناوری‌هایی نظیر وب ۲ و با افزایش امکانات در حوزه شبکه‌های ارتباطی و نیز کاهش قیمت تجهیزات همراه مانند گوشی‌های تلفن همراه مجهز به GPS، تمایل و نیز امکان مشارکت عموم مردم در فعالیت‌های مربوط به اطلاعات مکانی افزایش یافته است [Goodchild, 2007]. این مشارکت باعث شده است که ساختار تولید داده از حالت بالا به پایین (از سمت دولت به سمت مردم) به حالت پائین به بالا که در آن نقش اصلی را مردم ایفا می‌کنند، تغییر کند [Craglia, 2007; Goodchild, 2007; Muki Haklay et al. 2008]. این نوع از داده‌ها که توسط عموم مردم تولید می‌شود، مفهومی را تحت عنوان داده‌های مکانی داوطلبانه ایجاد کرده است [Goodchild, 2007]. تعداد پروژه‌هایی که پایه و اساس گسترش آن بر مبنای همکاری جمعی کاربران است روزبه‌روز در حال افزایش هستند، از جمله آنها می‌توان به

Wikimapia، Google Map Maker و OpenStreetMap اشاره نمود. در این بین OpenStreetMap به دلیل گستردگی استفاده و همچنین به دلیل اجازه دریافت اطلاعات به عنوان داده VGI مورد استفاده در این تحقیق در نظر گرفته شده است. VGI با نمایش سریع و به روز تغییرات می‌تواند تاثیر بسیاری در نتایج حاصل از کاربردها و آنالیزهای مختلف داشته باشد، اما نیازمند این است که از منظر دقت مورد ارزیابی قرار گیرد.

تا به امروز روش‌های مختلفی جهت ارزیابی دقت داده‌های VGI صورت پذیرفته است که بسیاری از آنها تنها به مقایسه طول و فاصله عوارض در دو مجموعه داده VGI و رسمی پرداخته‌اند و از این طریق داده‌های VGI مورد ارزیابی قرار گرفته است. برای مثال Hakaly در سال ۲۰۱۰ به مقایسه داده‌های داوطلبانه OSM با مجموعه داده رسمی راه‌های انگلستان پرداخت [Mordechai Haklay, 2010]. تحقیق وی بیشتر بر روی دو جنبه صحت موقعیتی و میزان کامل بودن داده‌ها تمرکز داشت. Hakaly برای بررسی صحت موقعیتی داده‌های OSM از روش بافر معرفی شده توسط Goodchild و Hunter در سال ۱۹۹۷ استفاده نمود [Goodchild and Hunter, 1997]. در این روش بافری با شعاع در نظر گرفته شده اطراف عارضه با دقت بالاتر زده می‌شود، آنگاه درصدی از عارضه مورد بررسی که در داخل بافر زده شده قرار گرفته، به عنوان معیار کیفیت مورد ارزیابی قرار می‌گیرد. Girres و Touya در سال ۲۰۱۰ با مقایسه فاصله اقلیدسی نقاط به ارزیابی داده‌های VGI پرداختند [Girres and Touya, 2010]. همچنین فرقانی و دلاور در سال ۲۰۱۴ با در نظر گرفتن معیار طول [Forghani and Delavar, 2014] و محمدی و ملک در سال ۲۰۱۵ با در نظر گرفتن معیارهای فاصله و جهت [Mohammadi and Malek, 2015] به ارزیابی داده‌های OSM تهران پرداختند.

همچنین تعدادی از تحقیقات به بررسی سایر مولفه‌های ارزیابی کیفیت داده‌های داوطلبانه و همچنین رسمی پرداخته‌اند. Brunz و Egenhofer در سال ۱۹۹۶ و Sheeren و همکاران [Bruns and Egenhofer, 2009] برای ارزیابی ناسازگاری [Egenhofer, 1996; Sheeren et al. 2009] و Abdolmajidi و همکاران در سال ۲۰۱۴، Abdolmajidi

Du et al. 2015]. و همکاران در سال ۲۰۱۶ یک روشی را مبتنی بر موقعیت گره و همچنین اطلاعات توصیفی برای تناظرایی عوارض در بین داده‌های OpenStreetMap و داده‌های رسمی ارائه دادند. آنها اثر حد آستانه بافر رابر روی نتیجه تناظرایی نهایی بررسی کردند بصورتیکه ۱۲ سطح مختلف از حد آستانه را بین ۱ متر تا ۸۰ متر در نظر گرفتند. نتایج آنها نشان داد که فاصله بافر بین ۱۵ متر تا ۸۰ متر دارای نتایج پایدار و تقریباً یکسانی است، بطوریکه بهترین نتیجه برای فاصله بافر ۳۰ متر با $\text{Recall}=85\%$ و $\text{precision}=89\%$ بدست آمد. [Du et al. 2016]

برخلاف تحقیقات پیشین، در این تحقیق تلاش شده است که عوارض از منظر بسیاری از مولفه‌های هندسس با یکدیگر مورد مقایسه قرار گیرند تا ارزیابی‌ها به واقعیت نزدیک‌تر گردد. از این رو دقت هندسی داده‌های VGI از طریق مقایسه آنها با داده‌های رسمی انجام گرفته و نتایج ارائه می‌گردد. جهت بررسی دقت هندسی، راهکاری نوین ارائه می‌گردد که ضمن شناسایی عوارض متناظر در دو مجموعه داده VGI و رسمی، هر یک از این عوارض از منظر معیارهای طول، فاصله، جهت، حریم مشترک، پیچیدگی، اندازه و اعوجاج با یکدیگر مقایسه شدند. ادامه این مقاله به این شرح است: پس از مقدمه در بخش ۲ چارچوب پیشنهادی ارائه و جزئیات آن تشریح می‌گردد. در بخش ۳ ضمن تشریح منطقه مورد مطالعه، چارچوب پیشنهادی بر روی این منطقه اعمال و نتایج مورد بررسی قرار می‌گیرد. در انتها نیز در بخش ۴ نتیجه گیری و پیشنهادات ارائه می‌گردد.

۲. چارچوب پیشنهادی

جهت ارزیابی هندسی داده‌های VGI در مقایسه با داده‌های رسمی، راهکاری مطابق شکل (۱) پیشنهاد می‌گردد. در این راهکار، ابتدا پیش پردازش بر روی مجموعه داده‌ها صورت می‌گیرد، سپس هر یک از عوارض در مجموعه داده‌ها VGI با عوارض کاندید در مجموعه داده‌های رسمی مورد مقایسه قرار می‌گیرد. هنگامی که تمامی عوارض در مجموعه داده‌ها VGI مقایسه و دقت هندسی هر یک از آنها محاسبه گردید، خروجی نهایی از طریق یک نقشه که نشان دهنده وضعیت دقت هندسی داده‌ها در منطقه است، نشان داده می‌شود.

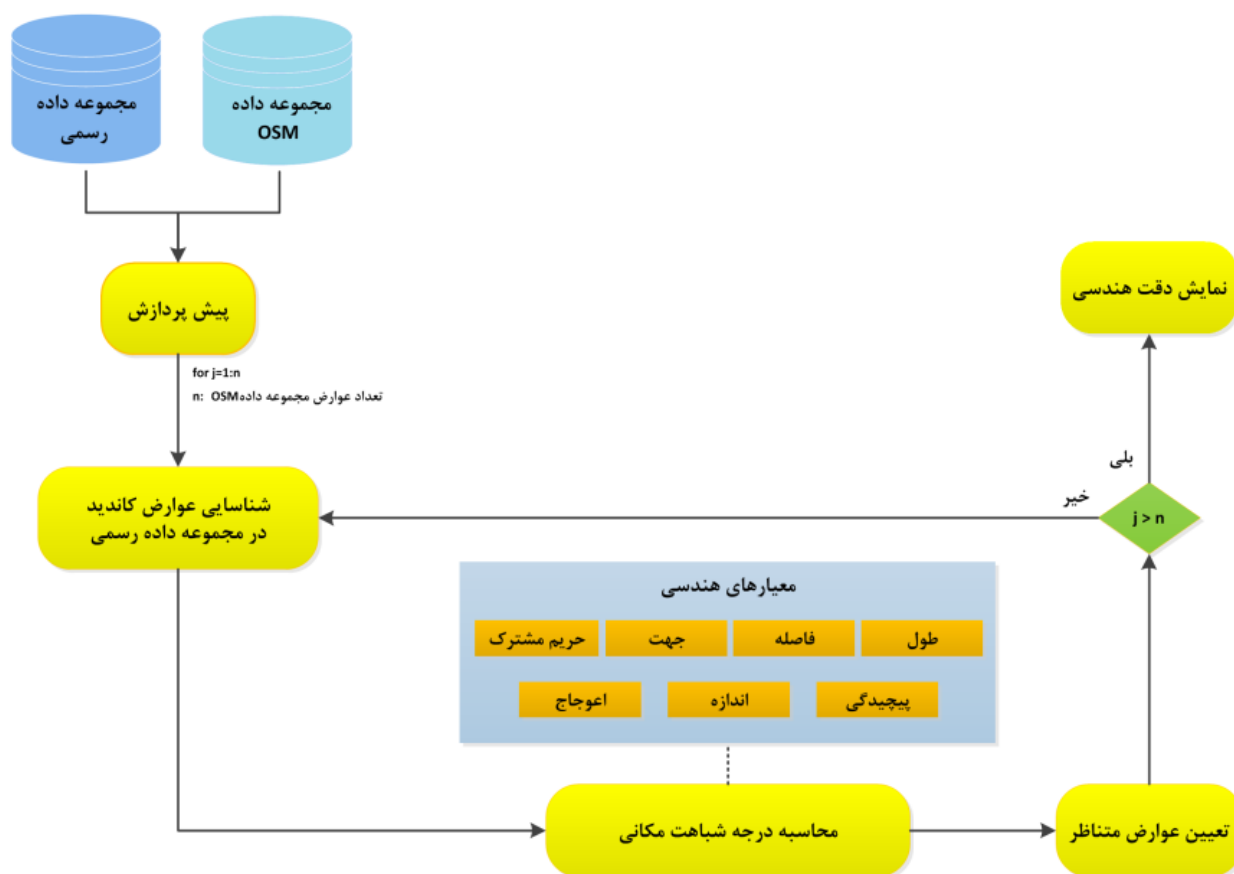
همکاران در سال ۲۰۱۵ و Fan و همکاران در سال ۲۰۱۶ برای بررسی دقت مکانی [E Abdolmajidi et al. 2014; Ehsan Abdolmajidi et al. 2015; Fan et al. 2016] Ludwig و همکاران در سال ۲۰۱۱، Koukolrtsos و همکاران در سال ۲۰۱۵ برای ارزیابی کامل بودن [Koukoletsos et al. 2012; Ludwig et al. 2011] و Fan و همکاران در سال ۲۰۱۴ و Cao و همکاران در سال ۲۰۱۶ برای ارزیابی دقت اطلاعات توصیفی از مولفه‌های مرتبط با تناظرایی عوارض استفاده نمودند [Cao et al. 2016; Fan et al. 2014].

از منظر معیارهای مورد استفاده Dunkars در سال ۲۰۰۳ با استفاده از معیار فاصله به تناظرایی عوارض در مجموعه داده‌های چند نمایشی پرداخت [Dunkars, 2003]. Gosseln و Sester در سال ۲۰۰۴ با ارائه یک تابع مبتنی بر فاصله اقلیدسی عوارض متناظر را در مجموعه داده‌های مختلف شناسایی نمود [Gösseln and Sester, 2004]. همچنین در تحقیقات Min و همکاران در سال ۲۰۰۷، Li و Goodchild در سال ۲۰۱۱ و Tong و همکاران در سال ۲۰۱۴ از فاصله هاسدورف جهت مقایسه دو عارضه و شناسایی عوارض متناظر استفاده شده است [Li and Goodchild, 2011; Min Fritsch و Walter et al. 2007; Tong et al. 2014] در سال ۱۹۹۹ با استفاده از معیارهای هندسی طول، زاویه و موقعیت نقاط و همچنین با معرفی روشی که در آن امکان رشد حریم‌گیری وجود داشت سعی در بهبود دقت تناظرایی عوارض در دو مجموعه داده نمودند [Volker Walter and Fritsch, 1999]. Mustiere و Devogele در سال ۲۰۰۸ فرایندی مبتنی بر خصوصیات هندسی، توپولوژیکی و معنایی را برای تناظرایی شبکه راه‌ها در سطوح مختلف جزئیات ارائه کردند [Mustière and Devogele, 2008]. Yang و همکاران در سال ۲۰۱۳ با استفاده از رویکرد نوین موسوم به رویکرد احتمالاتی به تناظرایی عوارض پرداختند [Yang et al. 2013]. از جمله تحقیقات اخیر می‌توان Olteanu-Raimond و همکاران در سال ۲۰۱۵ را نام برد، در این مقاله یک روش چند معیاره برای تناظرایی عوارض خطی در مجموعه داده‌های چند نمایشی با استفاده از معیارهای موقعیت گره، جهت و اطلاعات توصیفی ارائه گردید [Olteanu-Raimond et al. 2015].

۲-۱ پیش پردازش

قبل از شروع فرایند ارزیابی هندسی، نیاز است تا بر روی مجموعه داده‌های ورودی پیش پردازش صورت گیرد، این عملیات شامل چهار مرحله است. در مرحله اول چنانچه دو مجموعه داده دارای قالب متفاوت باشند، هر دو مجموعه داده به قالب یکسان تبدیل می‌شوند. در مرحله دوم ممکن است دو مجموعه داده دارای سیستم مختصات‌های متفاوتی از یکدیگر باشند، از این رو نیاز است تا هر دو مجموعه به یک سیستم مختصات واحد تبدیل گردند. در مرحله سوم نیز برای کاهش خطاهای سیستماتیک در مجموعه داده‌ها، خطاهای توپولوژیک (شامل خطای ردشدگی و نرسیدگی) حذف می‌گردند. در مرحله چهارم، برای حذف ابهام و عدم قطعیت در شناسایی عوارض متناظر، مجموعه داده‌های ورودی تبدیل به ساختار گراف پیشنهادی می‌گردند. در این تحقیق از نظریه گراف‌ها برای توصیف شبکه راه‌ها به عنوان یکسری نقاط و خطوط متصل به هم استفاده شده است.

در ریاضیات، گراف مرتبط با خطوط شبکه، می‌تواند بصورت زوج مرتب $G = (V, E)$ نمایش داده شود که V شامل مجموعه رئوس شبکه و E شامل مجموعه یال‌های شبکه است. هر یال توسط یک جفت از رئوس قابل شناسایی است که درجه هر یک از این رئوس، تعداد یال‌های متصل به آن است [Mackness and Mackechnie, 1999]. در شبکه راه‌های شهری هر چندخطی شامل چندین رأس و یال است. چندخطی مفروض PL_i شامل نقاط $P_{i,1}, P_{i,2}, \dots, P_{i,n}$ است که هر دو نقطه $P_{i,j}$ و $P_{i,j+1}$ یک یال از چندخطی را تشکیل می‌دهند. همچنین نقاط $P_{i,1}$ و $P_{i,n}$ نقاط ابتدایی و انتهایی چندخطی PL_i هستند. در این تحقیق عوارض بصورتی تعریف می‌گردد که در ساختار گراف این چندخطی‌ها، درجه ابتدایی و انتهایی رئوس آن یک و یا بیشتر از دو باشد. در نتیجه در فرایند تناظریابی هر یک از این چندخطی‌ها بصورت یک عارضه در نظر گرفته می‌شود. بنابراین هر چندخطی ممکن است با یک تقاطع شروع و یا پایان یابد، ولی هرگز شامل یک تقاطع در نقاط میانی نخواهد بود.



شکل ۱. راهکار پیشنهادی جهت ارزیابی هندسی مجموعه داده VGI از طریق داده رسمی

۲-۲ درجه شباهت مکانی

پس از انجام پیش پردازش بر روی مجموعه داده‌ها، عوارض متناظر در مجموعه داده دوم برای هر یک از عوارض موجود در مجموعه داده اول شناسایی می‌گردد. عوارض کاندید در مجموعه داده دوم از طریق الگوریتم BG ارائه شده توسط Walter در سال ۱۹۹۵ شناسایی می‌گردند [Walter and V Fritsch, 1995]، در این الگوریتم برای عارضه‌ای نظیر PL حریمی به فاصله β (به اندازه دقت مکانی مجموعه داده رسمی) در نظر گرفته می‌شود و در صورتی که عوارض کاندیدا در این محدوده شناسایی نگردند فاصله β افزایش می‌یابد [Volker and Fritsch, 1999]. پس از اینکه برای عارضه مرجع PL عوارض کاندید در مجموعه داده دوم مشخص گردید، جدولی به نام جدول مقایسه معادل جدول (۱) تشکیل می‌گردد. در این تحقیق عارضه مرجع مجموعه داده OSM و مجموعه کاندید مجموعه داده رسمی است.

جدول ۱. جدول مقایسه ایجاد شده برای عارضه مرجع و عوارض

کاندید

		اختلاف عارضه مرجع و عوارض کاندید از منظر معیارها		
		PL - PL ₁	PL - PL ₂	PL - PL _n
معیارهای استخراج شده از ویژگی هندسی	C ₁	$\overline{\text{Sim}}_{\text{PL}, \text{PL}_1}^{C_1}$		$\overline{\text{Sim}}_{\text{PL}, \text{PL}_n}^{C_1}$
	C ₂	$\overline{\text{Sim}}_{\text{PL}, \text{PL}_1}^{C_2}$		$\overline{\text{Sim}}_{\text{PL}, \text{PL}_n}^{C_2}$
	C ₃	$\overline{\text{Sim}}_{\text{PL}, \text{PL}_1}^{C_3}$		$\overline{\text{Sim}}_{\text{PL}, \text{PL}_n}^{C_3}$
	C ₄	$\overline{\text{Sim}}_{\text{PL}, \text{PL}_1}^{C_4}$		$\overline{\text{Sim}}_{\text{PL}, \text{PL}_n}^{C_4}$
	C ₅	$\overline{\text{Sim}}_{\text{PL}, \text{PL}_1}^{C_5}$		$\overline{\text{Sim}}_{\text{PL}, \text{PL}_n}^{C_5}$
	C ₆	$\overline{\text{Sim}}_{\text{PL}, \text{PL}_1}^{C_6}$		$\overline{\text{Sim}}_{\text{PL}, \text{PL}_n}^{C_6}$
	C ₇	$\overline{\text{Sim}}_{\text{PL}, \text{PL}_1}^{C_7}$		$\overline{\text{Sim}}_{\text{PL}, \text{PL}_n}^{C_7}$
درجه شباهت		Sim(PL, PL ₁)		Sim(PL, PL _n)

در جدول (۱) PL₁ تا PL_n عوارض کاندید شناسایی شده در مجموعه داده رسمی برای عارضه PL در مجموعه داده VGI،

مقدار محاسبه شده معیار هندسی $\overline{\text{Sim}}_{\text{PL}, \text{PL}_j}^{C_i}$ ام برای عارضه کاندید PL_j و عارضه مرجع PL و C₁ تا C₇ معیارهای مورد استفاده در این تحقیق شامل طول، فاصله، جهت، حریم مشترک، پیچیدگی، اندازه و اعوجاج است. در انتها نیز از طریق رابطه (۱) درجه شباهت مکانی عارضه مرجع PL و عارضه کاندید PL_j محاسبه می‌گردد [Chehreghan and Ali Abbaspour, 2017].

$$\text{Sim}(\text{PL}, \text{PL}_j) = \frac{\sum_{i=1}^m \overline{\text{Sim}}_{\text{PL}, \text{PL}_j}^{C_i}}{m} \times 100 \quad (1)$$

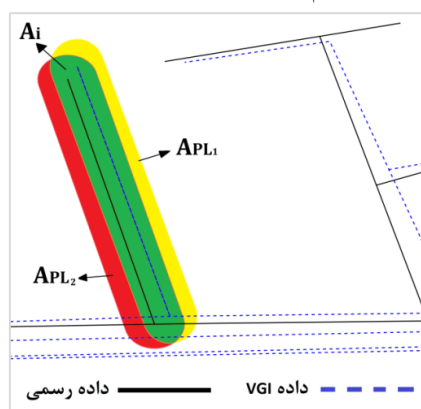
در این رابطه Sim(PL₁, PL₂) درصد شباهت مکانی بین عارضه PL و PL_j است که دارای مقادیری بین ۰ تا ۱۰۰ است. $\overline{\text{Sim}}_{\text{PL}, \text{PL}_j}^{C_i}$ مقادیر نرمال شده $\overline{\text{Sim}}_{\text{PL}, \text{PL}_j}^{C_i}$ است. در ادامه هر یک از معیارهای طول، فاصله، جهت، حریم مشترک، پیچیدگی، اندازه و اعوجاج برای هر یک از عوارض ارائه می‌گردد.

طول: یکی از معیار هندسی که در بسیاری از تحقیقات مورد استفاده قرار گرفته است، در نظر گرفتن طول عوارض است. از این رو اختلاف طول دو عارضه خطی PL₁ و PL₂ به عنوان معیاری جهت مقایسه دو عارضه خطی مورد استفاده قرار می‌گیرد [Wang et al. 2015].

فاصله: فاصله مکانی بین دو عارضه، در حقیقت یک پارامتری برای اندازه‌گیری موقعیت نسبی این عوارض از یکدیگر است که می‌تواند به عنوان معیاری جهت محاسبه درجه شباهت مکانی عوارض مورد استفاده قرار گیرد [Mascret et al. 2006]. فاصله هاسدورف یکی از پرکاربردترین روش‌های محاسبه فاصله در علوم مکانی است [Tong et al. 2014]. این فاصله به عنوان بیشترین فاصله بین کوتاهترین فاصله موجود بین هر نقطه از عارضه اول و مجموعه نقاط عارضه دوم تعریف شده است [Li and Goodchild, 2011]. فاصله هاسدورف (D_H) با توجه به تمامی رئوس موجود در عوارض تعریف می‌شود و برای دو مجموعه نقاط PL₁ و PL₂ بصورت رابطه (۲) محاسبه می‌گردد [Min et al. 2007].

$$D_H(\text{PL}_1, \text{PL}_2) = \max\{h(\text{PL}_1, \text{PL}_2), h(\text{PL}_2, \text{PL}_1)\} \quad (2)$$

مساحت حریم ایجاد شده برای عارضه دوم و A_1 مساحت منطقه مشترک بین دو حریم عارضه خطی PL_1 و PL_2 است.



شکل ۲. حریم مشترک دو عارضه خطی

پیچیدگی: پیچیدگی خطوط از جمله معیارهای دیگری است که می توان از آن جهت محاسبه درجه شباهت مکانی عوارض استفاده نمود. پیچیدگی خطوط را می توان از طریق در نظر گرفتن میانگین وزن دار فاصله رئوس عارضه (h_i) از خط واصل ایجاد شده (D) بین گره ابتدایی و انتهایی محاسبه نمود (شکل ۳) [Anderson et al. 2014].

اندازه: اندازه در یک عارضه خطی به فاصله اقلیدسی بین گره های ابتدایی و انتهایی گفته می شود، از این رو از اختلاف این پارامتر در دو عارضه خطی می توان به عنوان معیاری جهت مقایسه استفاده نمود [Anderson et al. 2014].

اعوجاج: یکی دیگر از معیارهای هندسی مورد استفاده در تناظرایی اعوجاج عوارض می باشد که دو معیار اندازه و طول را شامل می شود و از طریق تقسیم طول بر اندازه محاسبه می شود [Anderson et al. 2014].

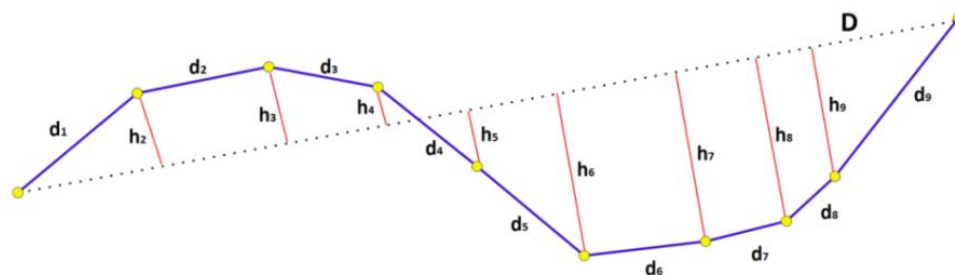
در این رابطه $h(PL_1, PL_2)$ و $h(PL_2, PL_1)$ به عنوان فاصله هاسدورف مستقیم شناخته می شوند و از طریق روابط (۳) و (۴) محاسبه می شوند.

$$h(PL_1, PL_2) = \max_{P_a \in PL_1} \left\{ \min_{P_b \in PL_2} \{d(P_a, P_b)\} \right\} \quad (3)$$

$$h(PL_2, PL_1) = \max_{P_b \in PL_2} \left\{ \min_{P_a \in PL_1} \{d(P_b, P_a)\} \right\} \quad (4)$$

جهت: یکی دیگر از معیارهای هندسی که می توان برای مسأله تناظرایی عوارض خطی مور استفاده قرار داد، اختلاف جهت عوارض خطی از یکدیگر است که می تواند نقش مهمی را در مسأله ایفا کند [Olteanu Raimond and Mustière, 2008]. برای یک عارضه خطی، جهت به زاویه بین خط تشکیل شده از گره های ابتدایی و انتهایی عارضه و محور افقی گفته می شود. در این صورت برای دو عارضه خطی با جهت های α و β ، اختلاف جهت برابر $|\alpha - \beta|$ خواهد بود. در صورتی که مقدار $|\alpha - \beta|$ برابر صفر باشد، دو عارضه تقریباً هم راستا هستند و در صورتیکه مقدار $|\alpha - \beta|$ برابر π باشد، راستای دو عارضه خطی در خلاف جهت یکدیگر است [Zhang, 2009].

حریم مشترک: از جمله معیارهای دیگر مورد استفاده در تحقیقات مختلف، در نظر گرفتن مساحت موجود بین منطقه مشترک بوجود آمده از حریم ایجاد شده بین عوارض است [Fan et al. 2014; Fan et al. 2016]. برای مثال در شکل (۲) مساحت منطقه با رنگ تیره تر می تواند معیاری برای سنجش میزان شباهت بین دو عارضه خطی PL_1 و PL_2 باشد. در این شکل APL_1 مساحت حریم ایجاد شده برای عارضه اول، APL_2



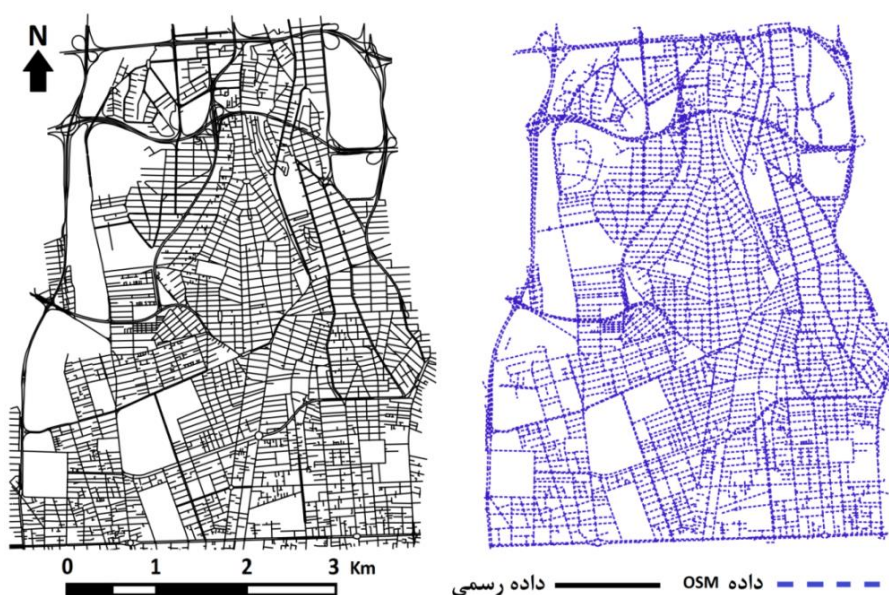
شکل ۳. پارامترهای مورد استفاده در محاسبه پیچیدگی خطوط

۳. پیاده‌سازی و ارزیابی نتایج

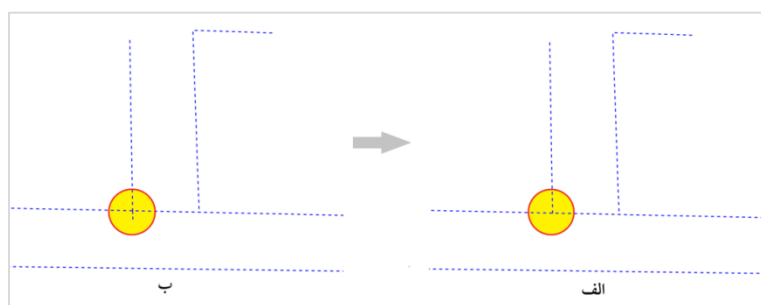
دو مجموعه به سیستم تصویر UTM و بیضوی مرجع به WGS84 تبدیل می‌شود و در ادامه خطاهای توپولوژیکی موجود در هر دو مجموعه داده حذف می‌شوند. به عنوان نمونه شکل (۵) نمونه‌ای از خطای ردشدگی را قبل و پس از تصحیح نشان می‌دهد.

آخرین قسمت از مرحله پیش پردازش شامل تعریف عوارض براساس ساختار گراف تعریف شده در هر دو مجموعه داده است. عوارض بصورتی تعریف می‌گردد که شامل هیچ تقاطعی در گره‌های میانی نباشد. برای مثال در شکل (۶) شکل عارضه PL با گره‌های ۱ تا ۱۱ تبدیل به عوارض PL1 با گره‌های ۱ تا ۵ و عارضه PL2 با گره‌های ۵ تا ۱۱ شده است. جدول (۲) تعداد عوارض هر دو مجموعه داده را قبل و بعد از انجام مرحله پیش پردازش نشان می‌دهد.

در این تحقیق برای ارزیابی دقت هندسی داده‌های VGI از داده‌های OSM استفاده گردید. برای این منظور داده‌های منطقه شش تهران به عنوان منطقه مورد مطالعه در تاریخ ۱۳۹۵/۱۲/۱۵ از سایت WWW.OSM.ORG دانلود گردید. همچنین جهت ارزیابی، این داده‌ها با داده‌های رسمی در مقیاس ۱:۲۰۰۰ که توسط سازمان نقشه‌برداری کشور تهیه شده بود، مورد مقایسه قرار می‌گیرد. شکل (۴) داده‌های OSM و همچنین داده‌های رسمی را در منطقه مورد مطالعه نشان می‌دهد. همانطور که در راهکار پیشنهادی بیان گردید، اولین مرحله شامل پیش پردازش مجموعه داده‌ها است. از آنجائیکه داده OSM در هنگام دانلود دارای فرمت *.pbf است، در ابتدا هر دو مجموعه داده به قالب *.shp تبدیل می‌گردند. سپس سیستم تصویر هر

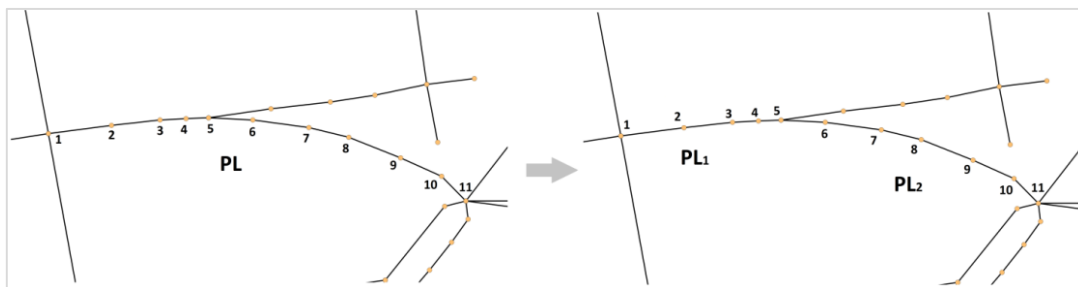


شکل ۴. داده‌های OSM و رسمی در منطقه مورد مطالعه



شکل ۵. نمونه‌ای از خطای توپولوژیکی قبل (الف) و بعد (ب) از تصحیح

علیرضا چهرقان، رحیم علی عباسپور



شکل ۶. نمونه‌ای از تعریف عوارض با استفاده از ساختار گراف تعریف شده

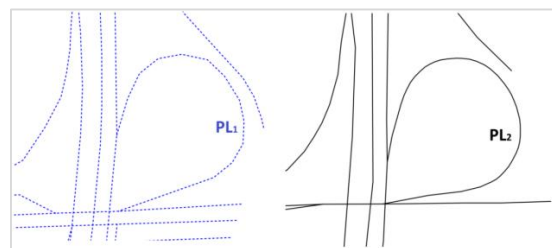
جدول ۲. تعداد عوارض قبل و بعد از انجام مرحله پیش پردازش

مجموعه داده	قبل از پیش پردازش		بعد از پیش پردازش	
	تعداد عوارض	مجموع طول (متر)	تعداد عوارض	مجموع طول (متر)
OSM	۶۳۰۲	۴۶۲/۳۲۲	۵۱۸۶	۴۶۲/۳۲۳
۱:۲۰۰۰	۶۷۰۱	۵۰۱/۲۳۸	۶۴۶۹	۵۰۱/۲۳۸

محاسبه می‌شود. پس از محاسبه درجه شباهت مکانی برای تمامی عوارض نتایج نشان داد که عوارض موجود در مجموعه داده OSM به طور میانگین ۷۸ درصد شباهت مکانی با مجموعه داده رسمی دارند. همچنین ۸۸/۸ درصد از عوارض مجموعه داده OSM دارای شباهت مکانی بالای ۷۰ درصد با عوارض متناظر آنها در مجموعه داده رسمی می‌باشند. شکل (۸) توزیع درصد شباهت مکانی بین مجموعه داده‌های OSM و رسمی را نشان می‌دهد. در این شکل نمایان است که برای ۱۱/۲ درصد از عوارض موجود در مجموعه داده OSM، هیچ تناظری در مجموعه داده رسمی یافت نشد. دلیل این امر می‌تواند به دو موضوع برگردد، اول اینکه ممکن است هندسه عارضه ترسیم شده بوسیله کاربر، کاملاً متفاوت از عارضه نظیر در مجموعه داده رسمی باشد. دوم اینکه، از آنجائیکه مجموعه داده VGI به سرعت تغییرات موجود در دنیای واقعی را از طریق ترسیم آن توسط کاربران نشان می‌دهند، ممکن است این عوارض در زمان ایجاد مجموعه داده رسمی وجود نداشته باشند.

برای نمایش وضعیت توزیع درصد شباهت مکانی بین عوارض در منطقه مورد مطالعه شکل (۹) ارائه می‌گردد. همانطور که در این شکل نمایان است، توزیعی با الگویی مشخص برای پراکندگی مقدار درصد شباهت مکانی وجود ندارد. از آنجائیکه بیش از ۸۸ درصد از عوارض مجموعه داده OSM دارای شباهت مکانی بالای ۷۰ درصد با مجموعه داده رسمی می‌باشند، می‌توان در بسیاری از کاربردهای مرتبط با حوزه حمل و نقل از

در ادامه برای هر یک از عوارض در مجموعه داده OSM، عارضه‌های متناظر در مجموعه داده رسمی شناسایی و از این طریق درصد شباهت مکانی آنها محاسبه می‌گردد. برای مثال شکل (۷) دو عارضه PL_1 و PL_2 را در دو مجموعه داده‌های OSM و رسمی نشان می‌دهد. همچنین جدول (۳) مقادیر محاسبه شده برای هر یک از معیارهای C_1 تا C_7 و درصد شباهت مکانی نهایی را نشان می‌دهد.



شکل ۷. یک عارضه در دو مجموعه داده OSM و رسمی

جدول ۳. مقادیر هر یک از مولفه‌های هندسی در مقایسه دو

عارضه PL_1 و PL_2

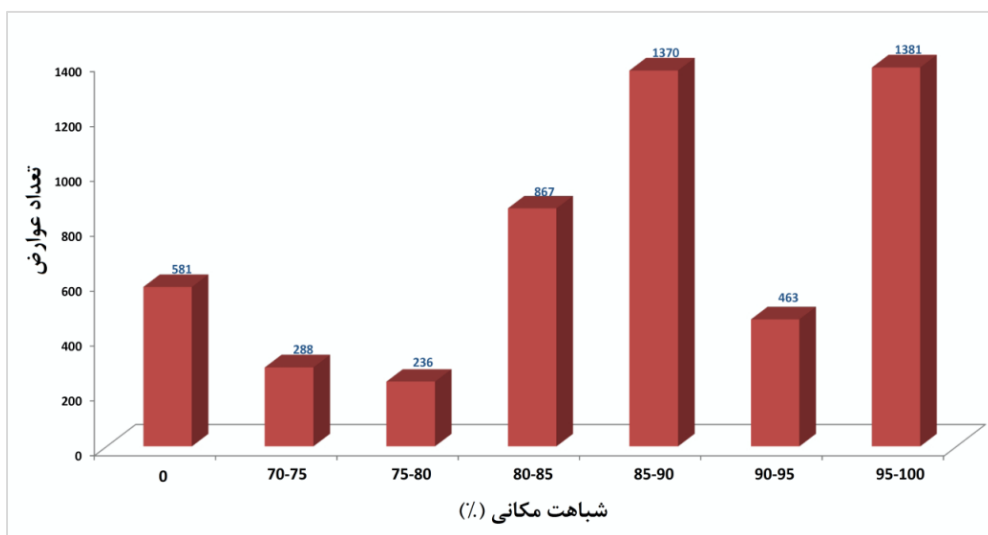
عوارض	مقادیر معیارها						
	C_7	C_6	C_5	C_4	C_3	C_2	C_1
PL_1-PL_2	۹۲/۹۹	۹۰/۴	۹۰/۶	۹۲/۳	۹۸/۴	۸۹/۸	۹۷/۱

همانند مثال ارائه شده برای تمامی عوارض در مجموعه داده OSM، درصد شباهت مکانی از طریق معیارهای هندسی طول، فاصله، جهت، حریم مشترک، پیچیدگی، اندازه و اعوجاج

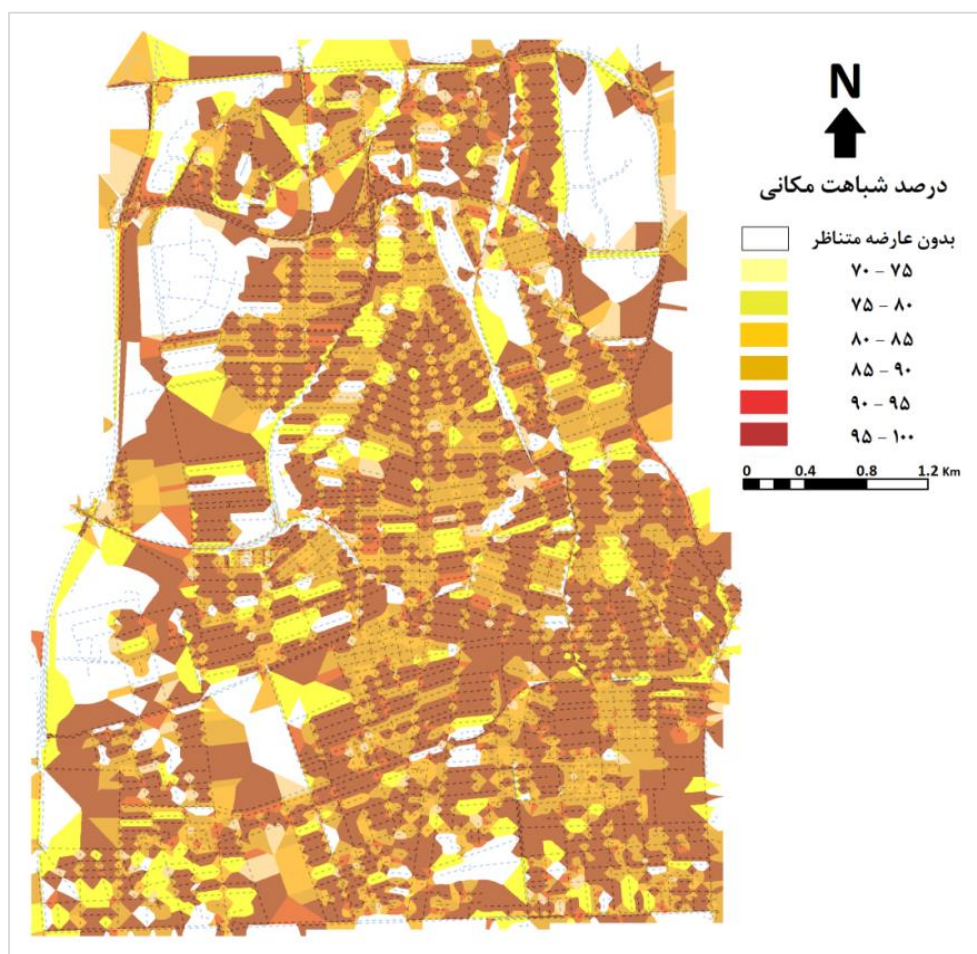
ارزیابی شباهت هندسی اطلاعات مکانی داوطلبانه در شبکه معابر درون‌شهری

بسیاری از جزئیات نظیر کوچه، راه‌های فرعی، راه‌های اصلی و غیره مورد مقایسه قرار گرفته است.

مجموعه داده‌های VGI به جای مجموعه داده‌های رسمی استفاده نمود. شایان ذکر است که در این تحقیق مجموعه داده-های VGI با مجموعه داده رسمی با مقیاس ۱:۲۰۰۰ که دارای



شکل ۸. شباهت مکانی بین مجموعه داده‌های OSM و رسمی



شکل ۹. توزیع درصد شباهت مکانی بین عوارض در منطقه مورد مطالعه

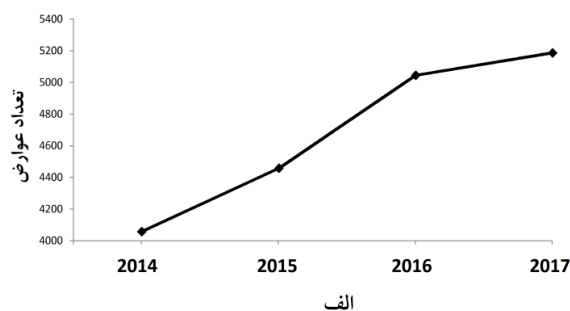
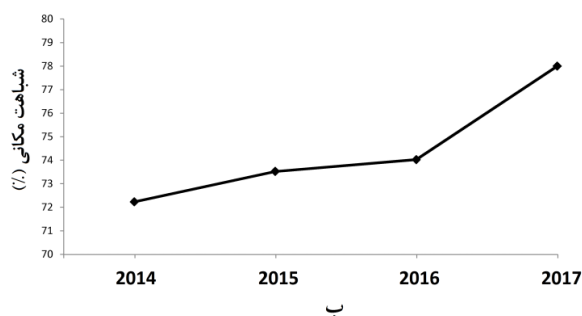
۴. بحث

نشان از دارا بودن ۷۸ درصد شباهت بین این دو مجموعه از منظر مولفه‌های هندسی نظیر طول، فاصله جهت، حریم مشترک، پیچیدگی، اندازه و اعوجاج است. پیش‌بینی می‌شود با روند رو به رشد تعداد مشارکت افراد در تهیه داده و افزایش درصد شباهت هندسی داده‌های OSM با داده‌های رسمی در سال‌های اخیر، منجر به استفاده از داده‌های مکانی داوطلبانه در بسیاری دیگر از کاربردها (که نیاز به دقت‌های بالاتری است) گردد. شایان ذکر است که روش ارائه شده در این تحقیق صرفاً برای ارزیابی شباهت هندسی داده‌های مکانی داوطلبانه با مجموعه داده‌های رسمی ناست و می‌توان از آن برای ارزیابی دو مجموعه داده رسمی مختلف نیز استفاده نمود. هرچند در صورت استفاده برای این منظور می‌توان روش ارائه شده را با بهینه کردن پارامترهای هندسی مختلف بر حسب اختلاف مقیاس دو مجموعه داده، بهبود بخشید.

همچون هر پژوهش دیگری، رویکرد ارائه شده در این تحقیق نیز دارای محدودیت‌هایی است. در صورتیکه منابع داده دارای خطاهای هندسی نظیر اشتباه کاربر در ترسیم عوارض و یا خطای باقی‌مانده از فرایند خلاصه‌سازی در مجموعه داده‌های رسمی باشند، ممکن است این خطاها باعث عدم شناسایی و یا شناسایی اشتباه عوارض متناظر گردد. شناسایی و یا عدم شناسایی عوارض متناظر به میزان تغییر هندسه عوارض در این مجموعه داده‌ها بستگی دارد.

با بررسی میزان پیشرفت مشارکت عموم مردم در تکمیل مجموعه داده‌های VGI در سال‌های اخیر، پیش‌بینی می‌شود که در آینده-ای بسیار نزدیک مجموعه داده‌های VGI تقریباً در بسیاری از کاربردها جایگزین مجموعه داده‌های رسمی گردند. برای مثال شکل (۱۰) نشان دهنده خروجی پیاده‌سازی راهکار ارائه شده در این تحقیق بر روی مجموعه داده‌های OSM در سال‌های ۲۰۱۴ تا ۲۰۱۷ است. همانطور که در شکل (۹-الف) نمایان است تعداد عوارض در منطقه شش تهران از ۴۰۵۷ عارضه در سال ۲۰۱۴ به تعداد ۵۱۸۶ عارضه در سال ۲۰۱۷ رسیده است. این در حالی است که مقدار میانگین درصد شباهت مکانی هر یک از مجموعه داده‌های OSM با مجموعه داده رسمی از مقدار ۷۲/۲۳ درصد در سال ۲۰۱۴ به مقدار ۷۸ درصد در سال ۲۰۱۷ بهبود یافته است. این مقادیر نشان دهنده این موضوع است که ضمن افزایش ۲۷/۸ درصدی مشارکت مردم در ترسیم عوارض، میانگین درصد شباهت مکانی نیز به مقدار ۵/۷۷ درصد بهبود یافته است. از این رو محققین می‌توانند در بسیاری از کاربردها از مجموعه داده‌های VGI به عنوان جایگزینی برای مجموعه داده رسمی استفاده کنند.

نتایج ارزیابی شباهت هندسی داده OSM در منطقه مورد مطالعه با داده تهیه شده توسط سازمان نقشه‌برداری با مقیاس ۱:۲۰۰۰



شکل ۱۰. توزیع درصد شباهت مکانی بین عوارض در منطقه مورد مطالعه

4- Universal Transverse Mercator (UTM)

۵. نتیجه‌گیری و پیشنهادها

در این تحقیق تلاش شده است که داده‌های VGI مرتبط با شبکه راه‌های درون شهری از منظر هندسی با مجموعه داده رسمی با مقیاس ۱:۲۰۰۰ مورد مقایسه قرار گیرد. جهت ارزیابی هندسه داده‌های VGI، راهکاری نوین ارائه گردید که ضمن شناسایی عوارض متناظر در مجموعه داده رسمی، هر یک از این عوارض از منظر معیارهای طول، فاصله، جهت، حریم مشترک، پیچیدگی، اندازه و اعوجاج با یکدیگر مورد مقایسه قرار گرفتند. برای این منظور راهکار ارائه شده بر روی منطقه شش تهران پیاده‌سازی و نتایج ارائه گردید. پس از محاسبه درجه شباهت مکانی برای تمامی عوارض نتایج نشان داد که عوارض موجود در مجموعه داده OSM به طور میانگین ۷۸ درصد شباهت مکانی با مجموعه داده رسمی دارند. همچنین ۸۸/۸ درصد از عوارض مجموعه داده OSM دارای شباهت مکانی بالای ۷۰ درصد با عوارض متناظر آنها در مجموعه داده رسمی می‌باشند. جهت ارزیابی سیر بهبود دقت هندسی مجموعه داده‌های VGI شبکه راه‌های سال ۲۰۱۴ تا ۲۰۱۷ نیز مورد ارزیابی قرار گرفت. نتایج نشان داد که تعداد عوارض در منطقه شش تهران از ۴۰۵۷ عارضه در سال ۲۰۱۴ به تعداد ۵۱۸۶ عارضه در سال ۲۰۱۷ رسیده است. این در حالی است که مقدار میانگین درصد شباهت مکانی هر یک از مجموعه داده‌های OSM با مجموعه داده رسمی از مقدار ۷۲/۲۳ درصد در سال ۲۰۱۴ به مقدار ۷۸ درصد در سال ۲۰۱۷ بهبود یافته است. این مقادیر نشان دهنده این موضوع است که ضمن افزایش ۲۷/۸ درصدی مشارکت مردم در ترسیم عوارض، میانگین درصد شباهت مکانی نیز به مقدار ۵/۷۷ درصد بهبود یافته است. از این رو محققین می‌توانند از مجموعه داده‌های VGI به عنوان جایگزینی برای مجموعه داده رسمی استفاده کنند.

جهت تحقیقات آتی توصیه می‌شود که هندسه سایر عوارض VGI نظیر چندضلعی‌ها نیز مورد ارزیابی قرار گرفته و نتایج ارائه گردد تا مشخص گردد که آیا می‌توان از عوارض چندضلعی ایجاد شده توسط عموم مردم نیز به جای مجموعه داده‌های رسمی استفاده نمود و یا خیر.

۶. پی‌نوشت‌ها

- 1- Volunteer Geographic Information (VGI)
- 2- OpenStreetMap (OSM)
- 3- Buffer Growing (BG)

۷. مراجع

- Afandizadeh, S., Javanshir, H. and Elyasi, R. (2010) "Development of a model for designing urban bus transit network based on tabu search", Quarterly Journal of Transportation Engineering, Vol. 1, No. 4, pp. 13-26.
- Alenouri, H., Meshkani, S. M., Saffarzadeh, M. and sherafatypour, S. (2015) "Locating cameras at the entrances of plate number rationing zone to maximize violations detection", Quarterly Journal of Transportation Engineering, Vol. 6, No. 2, pp. 181-196.
- Anderson, D. L., Ames, D. P. and Yang, P. (2014) "Quantitative methods for comparing different polyline stream network models", Journal of Geographic Information System, Vol. 6, No. 2, pp. 88-98.
- Beheshtinia, M. A. and Bargbid, H. (2017) "A hybrid model to identify effective factors in road transportation hazards and prioritize black spots (case study: Neyshabour-Sabzevar axis)", Quarterly Journal of Transportation Engineering, Online publishing.
- Chehreghan, A. and Abbaspour, R. A. (2017) "An assessment of spatial similarity degree between polylines on multi-scale, multi-source maps", Geocarto International, Vol. 32, No. 5, pp. 471-487.
- Craglia, M. (2007) "Volunteered geographic information and spatial data infrastructures: when do parallel lines converge, paper presented at Position paper for the Specialist Meeting", Volunteered Geographic Information, December 1-13, Santa Barbara.
- Fan, H., Zipf, A., Fu, Q. and Neis, P. (2014) "Quality assessment for building footprints data on OpenStreetMap", International Journal of Geographical Information Science, Vol. 28, No. 4, pp. 700-719.
- Fan, H., Yang, B., Zipf, A. and Rousell, A. (۲۰۱۶) "A polygon-based approach for matching OpenStreetMap road networks with regional transit authority data", International Journal of Geographical Information Science, Vol. 30, No. 4, pp. 748-764.

- Mirbaha, B., Sherafatipour, S. and Mahpour, A. (2015) "Congestion Pricing Model for Urban Congested Roads (Case Study: Sadr Elevated Highway)", *Quarterly Journal of Transportation Engineering*, Vol. 7, No. 2, pp. 353-365. (In Persian)
- Mohammadi, N. and Malek, M. (2015) "VGI and Reference Data Correspondence Based on Location-Orientation Rotary Descriptor and Segment Matching", *Transactions in GIS*, Vol. 19, No. 4, pp. 619-639.
- Olteanu Raimond, A. M. and Mustière, S. (2008) "Data Matching – a Matter of Belief, in Headway in Spatial Data Handling", edited by A. Ruas and C. Gold, pp. 501-519, Springer Berlin Heidelberg.
- Rasafi, R., Momeni, F., Astaraki, Z. and Amini, B. (2012) "The Application of Multi-Criteria Decision-Making Analysis to Identify Hot Spots: Data Envelopment Analysis and Concordance Analysis", *Quarterly Journal of Transportation Engineering*, Vol. 3, No. 4, pp. 301-314.
- Saberian, J., Malek, M. and Hamrah, M. (2014) "Using Dual Graph and Wavelet Transform for Evaluation and Planning Transportation systems", *Quarterly Journal of Transportation Engineering*, Vol. 5, No. 3, pp. 317-328.
- Teymourian, F., Alesheikh, A. A., Alimohammadi, A. and SadeghiNiaraki, A. (2014) "Developing a system for measuring transportation performance and information distribution of urban bus using volunteer geographic information (VGI)", *Quarterly Journal of Transportation Engineering*, Vol. 6, No. 2, pp. 225-236.
- Tong, X., Liang, D. and Jin, Y. (2014) "A linear road object matching method for conflation based on optimization and logistic regression", *International Journal of Geographical Information Science*, Vol. 28, No. 4, pp. 824-846.
- Walter, V. and Fritsch, D. (1995) "Matching techniques for road network data in different data models", paper presented at International Symposium on Automotive Technology & Automation (28th). Proceedings for the dedicated conference on mechatronics: efficient
- Flanagin, A. J. and Metzger, M. J. (2008) "The credibility of volunteered geographic information", *GeoJournal*, Vol. 72, No. 3, pp. 137-148.
- Forghani, M. and Delavar, M. (2014) "A quality study of the openstreetmap dataset for Tehran", *ISPRS International Journal of Geo-Information*, Vol. 3, No. 2, pp.750-760.
- Girres, J. F. and Touya, G. (2010) "Quality assessment of the French OpenStreetMap dataset", *Transactions in GIS*, Vol. 14, No. 4, pp. 435-459.
- Goodchild, M. F. (2007) "Citizens as sensors: the world of volunteered geography", *GeoJournal*, Vol. 69, No. 4, pp. 211-221.
- Goodchild, M. F. and Hunter, G. J. (1997) "A simple positional accuracy measure for linear features", *International Journal of Geographical Information Science*, Vol. 11, No. 3, pp. 299-306.
- 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*, Vol. 37, No. 4, pp. 682-703.
- Haklay, M., Singleton, A. and Parker, C. (2008) "Web mapping 2.0: The neogeography of the GeoWeb", *Geography Compass*, Vol. 2, No. 6, pp. 2011-2039.
- Khammar, G., Pasban, V. and Mojgan, N. (2017) "An application of ant colony algorithm in optimal routing of inter- cities ambulances", *Quarterly Journal of Transportation Engineering*, Vol. 8, No. 3, pp. 389-404.
- Li, L. and Goodchild, M. F. (2011) "An optimisation model for linear feature matching in geographical data conflation", *International Journal of Image and Data Fusion*, Vol. 2, No. 4, pp. 309-328.
- Mackaness, W. A. and Mackechnie, G. A. (1999) "Automating the detection and simplification of junctions in road networks", *GeoInformatica*, Vol. 3, No. 2, pp. 185-200.
- Mascret, A., Devogele, T., Berre, I. L. and Hénaff, A. (2006) "Coastline matching process based on the discrete Fréchet distance", Springer.

-Zhang, M. (2009) "Methods and implementations of road-network matching", Unpublished PhD Dissertation, Technical University of Munich.

-Zolfaghari, A. and KarkeAbadi, Z. (2013) "Intelligent Routing For Rescue Squads using Game Theory Algorithm In Semnan", Quarterly Journal of Transportation Engineering, Vol. 5, No. 1, pp. 19-32.

computer support for engineering manufacturing, testing and reliability.

-Walter, V. and Fritsch, D. (1999) "Matching spatial data sets: a statistical approach", International Journal of Geographical Information Science, Vol. 13, No. 5, pp. 445-473.

-Wang, Y., Chen, D., Zhao, Z., Ren, F. and Du, Q. (2015) "A Back-Propagation Neural Network-Based Approach for Multi-Represented Feature Matching in Update Propagation", Transactions in GIS, Vol. 19, No. 6, pp. 964-993.

علیرضا چهرقان، رحیم علی عباسپور

رحیم علی عباسپور، درجه کارشناسی در رشته مهندسی نقشه برداری را در سال ۱۳۷۹ از دانشگاه تهران و درجه کارشناسی ارشد در رشته مهندسی نقشه برداری - GIS را در سال ۱۳۸۱ از دانشگاه تهران اخذ نمود. در سال ۱۳۸۹ موفق به کسب درجه دکتری در رشته مهندسی نقشه برداری - GIS از دانشگاه تهران گردید. زمینه های پژوهشی مورد علاقه ایشان داده کاوی مکانی-زمانی، بهینه سازی مکانی، اطلاعات مکانی داوطلبانه، خدمات مکان مبنا (LBS) و محاسبات هندسی بوده و در حال حاضر عضو هیات علمی با مرتبه استادیار در پردیس دانشکده های فنی دانشگاه تهران است.



علیرضا چهرقان، درجه کارشناسی در رشته مهندسی نقشه برداری را در سال ۱۳۸۸ از دانشگاه تهران و درجه کارشناسی ارشد در رشته مهندسی نقشه برداری گرایش GIS را در سال ۱۳۹۰ از دانشگاه تهران اخذ نمود. وی همچنین در سال ۱۳۹۶ موفق به کسب درجه دکتری در رشته مهندسی نقشه برداری گرایش GIS از دانشگاه تهران گردید. زمینه های پژوهشی مورد علاقه ایشان داده تناظریابی عوارض برداری، اطلاعات مکانی داوطلبانه، خدمات مکان مبنا (LBS) و محاسبات هندسی است.

