

توسعه یک مدل اعوجاج موقعیت توسط طیف‌های معنایی جوی با هدف بهبود فرآیند معکوس ژئوکدینگ

علی سبزلعلی یمقانی^۱، علی اصغر آل‌شیخ^{۲*}

^۱ دانشجوی دکتری سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه

نصیرالدین طوسی

sabzali.ali@gmail.com

^۲ استاد دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی

alesheikh@kntu.ac.ir

(تاریخ دریافت اسفند ۱۳۹۷، تاریخ تصویب مرداد ۱۳۹۸)

چکیده

معکوس ژئوکدینگ شامل نسبت دادن یک جای‌نام به یک مختصات می‌باشد. روش‌های رایج معکوس ژئوکدینگ موقعیت اشتراک گذاشته شده فرد را به نزدیک‌ترین مکان (بر اساس فاصله اقلیدسی) نسبت می‌دهند. در سال‌های اخیر و به سبب پیشرفت در فناوری‌های موقعیت‌یابی حجم عظیمی از داده‌های مکان‌مبنا در شبکه‌های اجتماعی مکانی مانند Yelp و Swarm تولید شده است. از سوی دیگر سرویس‌های مختلفی امکان ارائه شرایط جوی را در موقعیت و زمان مدنظر به صورت برخط ارائه می‌دهند. این داده‌ها می‌توانند منبعی غنی از الگوهای رفتاری افراد در شرایط جوی مختلف باشند. در این تحقیق تلاش شده تا با کمک این داده‌ها، معکوس ژئوکدینگ بر مبنای فاصله جغرافیایی بهبود داده شود. بدین منظور از داده‌های شرایط جوی برای تولید الگوهای رفتاری و از داده‌های سرویس Swarm برای جمع‌آوری اعلام‌حضورها استفاده شد. داده‌های شرایط جوی در چهار دسته شامل دمای هوا، رطوبت هوا، سرعت باد و میزان ابری بودن برای تولید طیف‌های معنایی جوی گروه‌بندی شدند. در این تحقیق از توابع اعوجاج موقعیت خطی، نسبی و سینوسی جهت برقراری ارتباط شاخص فاصله جغرافیایی با احتمالات جوی در فرآیند معکوس ژئوکدینگ استفاده شد. بعلاوه، از دو مجموعه داده آموزشی و تست در مطالعه موردی (ایالت نیویورک) جهت تعیین پارامترهای مدل و ارزیابی دقت استفاده گشت. نتایج این تحقیق نشان داد که با کمک مدل ترکیب خطی و طیف‌های معنایی جوی می‌توان نتایج معکوس ژئوکدینگ بر مبنای فاصله جغرافیایی را برای شاخص MRR و First Position به ترتیب ۱۸٫۶۴٪ و ۱۱۱٫۴۹٪ بهبود داد.

واژگان کلیدی: معکوس ژئوکدینگ، طیف‌های معنایی جوی، توابع اعوجاج موقعیت، Swarm

* نویسنده رابط

۱- مقدمه

پیشرفت در فناوری‌های موقعیت‌یابی، ناوبری و دنبال‌کننده‌ها^۱ موجب شده تا حجم عظیمی از داده‌های حرکتی در سال‌های اخیر تولید شود (Sharif & Alesheikh, 2017). کاربران شبکه‌های اجتماعی مکان‌مبنا به مانند: Foursquare، Yelp، Badoo و Facebook Places بخش بزرگی از این داده‌ها را تولید و انتشار داده‌اند. داده‌های تولید شده در این شبکه‌ها، مبنای تولید اطلاعات ارزشمندی درباره‌ی الگوهای رفتاری افراد در این شبکه‌ها می‌باشند (Noulas et al., 2011). این الگوها را می‌توان در فرآیندهای مختلف از جمله معکوس‌ژئوکدینگ^۲ به کار گرفت.

مکان^۳ نه تنها حقیقتی است که در چارچوب وسیع‌تری از فضا^۴ توضیح داده می‌شود، بلکه یک واقعیت است که از نقطه نظر افرادی که به آن معنی می‌دهند؛ آشکار و درک می‌شود (McKenzie, 2015). فضا یک ویژگی بنیادین از مکان است که یک بعد یا ترکیبی از سه بعد را به نمایش می‌گذارد. ابعاد دیگر مکان دارای اثر چشمگیری بر روی درک ما از مکان‌ها و انواع آنها دارند. تبدیل نمایش فضایی^۵ به مکانی^۶ یک موقعیت^۷ از جمله نیازهای اساسی در افزایش کارایی سیستم‌ها و کاربردهای مختلف از جمله شبکه‌های اجتماعی مکان‌مبنا می‌باشد. ژئوکدینگ مختصات جغرافیایی را به دیگر انواع اطلاعات کدگذاری شده جغرافیایی مانند آدرس‌ها مرتبط می‌سازد. فرآیند معکوس این موضوع معکوس‌ژئوکدینگ نامیده می‌شود که شامل نسبت دادن یک جای‌نام^۸ به یک مختصات می‌باشد (Kounadi et al., 2013). روش‌های رایج معکوس‌ژئوکدینگ موقعیت اشتراک گذاشته شده فرد را به نزدیک‌ترین مکان و بر اساس فاصله اقلیدسی نسبت می‌دهند. در روش‌های جدید پیشنهادی، محبوبیت مکان (بر مبنای تعداد اعلام-حضورها^۹ و نظرات^{۱۰} کاربران)، دانه‌بندی، پروفایل کاربر، سلسله مراتب عوارض، طیف معنایی زمانی آنها نیز مورد توجه قرار گرفته است (McKenzie & Janowicz, 2015).

Huang & Piemonte, 2013; Lambert et al., 2009; Searight et al., 2010; Zarem et al., 2017). منظور از طیف معنایی جوی در این تحقیق طیف‌هایی است که از تجمیع تعداد اعلام حضورهای کاربران در شرایط جوی مختلف در هر دسته‌بندی تهیه شده است. به عنوان مثال برای پارامتر میزان ابری بودن هوا، بازه ۰ تا ۱ با فواصل ۰٫۲ طیف‌ها را تشکیل داده است. به این ترتیب مثلاً مشخص می‌شود که کاربران مختلف برای هر دسته‌بندی (برای مثال باغ وحش) چه تعداد اعلام حضور در پنج بازه (۰ تا ۰٫۲، ۰٫۲ تا ۰٫۴، ۰٫۴ تا ۰٫۶، ۰٫۶ تا ۰٫۸ و ۰٫۸ تا ۱) شاخص ابری بودن هوا در محدوده زمانی جمع‌آوری داده انجام داده‌اند.

گسترش روزافزون ابزارهای همراه و تجهیز بودن آنها به سیستم‌های تعیین موقعیت موجب شده تا اطلاعات موقعیتی دقیق‌تر و قابل اعتمادتری تولید شود. به این ترتیب نیاز به برقراری ارتباط بین مختصات یک موقعیت با مکان‌های موجود در دنیای واقعی بیش از پیش احساس می‌شود. این اطلاعات را می‌توان در سرویس‌های مختلفی مانند: ناوبری، توصیه‌گری مکان و تبلیغات مکان‌مبنا به کار گرفت. بسیاری از محققان معتقدند که چالش فعلی فقط مربوط به کسب اطلاعات موقعیتی بادقت بالا نیست، بلکه اطلاعات مکانی چیزی بیش از دقت فضایی تنها می‌باشد. به طور مثال اینکه فردی در یک نقطه مورد علاقه^{۱۱} اعلام حضور داشته یا در آن مکان حضور دارد؛ به لحاظ معنایی می‌تواند مهم‌تر از میزان نزدیکی فرد به آنجا باشد (McKenzie & Janowicz, 2015).

امروزه سرویس‌های معکوس‌ژئوکدینگ رایگان متنوعی توسعه داده شده است. از جمله این سرویس‌ها می‌توان به ArcGIS، Nominatim، Google Maps، TravelGIS، Bing، Yahoo، Geonames اشاره نمود. پژوهشگران متعددی نیز به بررسی و مقایسه دقت سرویس‌ها پرداخته‌اند (Kounadi et al., 2013). از سوی دیگر گسترش محتوی کاربر تولید مکانی^{۱۲} منجر به توسعه سرویس‌هایی مانند Foursquare و Yelp شده است. این سرویس‌ها به کاربران این امکان را می‌دهند تا بتوانند موقعیت و توصیفات عوارض را به کمک یک سیستم جمع‌سپاری^{۱۳} بروزرسانی و توسعه دهند.

^۱ Trackers

^۲ Reverse geocoding

^۳ Place

^۴ Space

^۵ Spatial

^۶ Placial

^۷ Location

^۸ Toponym

^۹ Check-ins

^{۱۰} Tips

^{۱۱} Point of interest

^{۱۲} User Generated Spatial Content

^{۱۳} Crowd-Sourced System

تواند موجب بهبود فرآیند معکوس ژئوکدینگ مبتنی بر فاصله شود. در این تحقیق برای ترکیب فاصله جغرافیایی با طیف-های معنایی زمانی از مدل‌های اعوجاج موقعیت^۲ استفاده شد. منظور از مدل‌های اعوجاج، مدل‌های مورد استفاده جهت تغییر موقعیت (نزدیک یا دور کردن) نقاط مورد علاقه بر اساس مقادیر احتمالات طیف‌های معنایی جوی (یا معنایی زمانی) می‌باشد. به عبارت دیگر روش‌های رایج معکوس-ژئوکدینگ ابتدا یک لیست از نزدیک‌ترین عوارض را در اطراف فرد تهیه می‌کنند، سپس نزدیک‌ترین مکان را به وی پیشنهاد می‌دهند. اما برای بهبود فرآیند معکوس ژئوکدینگ، استفاده از اطلاعات بافتی مختلف در کنار شاخص رایج فاصله جغرافیایی منجر به ارائه نتایج دقیق‌تر (از لحاظ شاخص‌های ارزیابی دقت) می‌شود. به عبارت دیگر تلاش می‌شود تا با کمک اطلاعات بافتی فاصله جغرافیایی فرد با محل‌های مختلف تعدیل شود. به این ترتیب لیست نزدیک‌ترین محل به فرد با توجه به اطلاعات بافت بهبود می‌یابد تا جواب نهایی با دقت بهتری پیشنهاد شود. از طرف دیگر Trattner و همکاران (۲۰۱۶) در تحقیقی نشان دادند که با کمک اطلاعات بافتی آب و هوا می‌توان دقت فرآیند توصیه‌گری نقاط مورد علاقه را بهبود داد. جمع‌بندی این موارد موجب شد تا در این تحقیق امکان استفاده از اطلاعات بافتی آب و هوا به عنوان یک منبع غنی تولید طیف‌های معنایی جوی با هدف بهبود فرآیند معکوس ژئوکدینگ مورد بررسی قرار گیرد. این اطلاعات شامل اطلاعات مربوط به دمای هوا، رطوبت هوا، سرعت باد و میزان ابری بودن می‌باشد که در قالب طیف‌های معنایی جوی برای دسته‌بندی‌های مختلف سرویس Swarm تولید می‌شود. در نهایت اهداف این مطالعه بدین صورت انتخاب شد:

- ساختن طیف‌های معنایی جوی دسته‌بندی‌های سرویس Swarm برای پارامترهای دمای هوا، رطوبت هوا، سرعت باد و میزان ابری بودن
- استفاده از داده‌های اعلام‌حضور سرویس Swarm برای ایالت نیویورک در کنار اطلاعات بافتی آب و هوا جهت بررسی امکان بهبود فرآیند معکوس ژئوکدینگ
- توسعه و ارزیابی مدل‌های مختلف خطی، نسبی و سینوسی در فرآیند معکوس ژئوکدینگ جهت ترکیب فاصله جغرافیایی با طیف‌های معنایی جوی

جمع‌آوری برخط اطلاعات مربوط به شرایط جوی برای نقاط مختلف دنیا یکی از چالش‌های بزرگی می‌باشد که امروزه با وجود آمدن سرویس‌های مختلف برخط پیش‌بینی آب و هوایی برطرف شده است. سرویس OpenWeatherMap با آدرس وبسایت openweathermap.org یکی از این سرویس‌ها می‌باشد. با کمک API این سرویس و با استفاده از درخواست‌هایی با فرمت `api.openweathermap.org/data/2.5/weather?lat=num1&lon=num2` می‌توان اطلاعات مربوط به آب و هوای مختصات جغرافیایی موردنظر را استخراج کرد. در این درخواست num1 و num2 به ترتیب عرض جغرافیایی و طول جغرافیایی نقطه مدنظر می‌باشد. برای روشن‌تر شدن موضوع یک نمونه درخواست از این سرویس برای عرض جغرافیایی ۴۰٫۷۳ و طول جغرافیایی ۷۴- به همراه خروجی آن در جدول ۱ آورده شده است.

جدول ۱- درخواست و نتیجه برای یک نمونه پارامتر در

OpenWeatherMap API

پارامترهای درخواست	پاسخ API
coordinate: "longitude":-74, "latitude":40.73	<pre>{ "coord": { "lon": -74, "lat": 40.73 }, "weather": [{ "id": 501, "main": "Rain", "description": "moderate rain", "icon": "10d" }, { "id": 701, "main": "Mist", "description": "mist", "icon": "50d" }], "base": "stations", "main": { "temp": 289.81, "pressure": 1006, "humidity": 93, "temp_min": 288.15, "temp_max": 291.48, "visibility": 6437, "wind": { "speed": 1.5, "deg": 70 }, "rain": { "1h": 3.9 }, "clouds": { "all": 90 }, "dt": 1559260304, "sys": { "type": 1, "id": 3595, "message": 0.0119, "country": "US", "sunrise": 1559208494, "sunset": 1559261948 }, "timezone": -14400, "id": 5128581, "name": "New York", "cod": 200 } }</pre>

انجام فرآیند معکوس ژئوکدینگ بر مبنای طیف‌های معنایی زمانی، می‌تواند به نتایج بهتری نسبت به معکوس-ژئوکدینگ بر مبنای فاصله جغرافیایی تنها منجر شود. این موضوع در تحقیق McKenzie و Janowicz (۲۰۱۵) با کمک حدود ۱ میلیون محل^۱ از سرویس Foursquare به چالش کشیده شد. نتایج این تحقیق نشان داد که استفاده از این طیف‌ها موجب بهبود ۱۴ تا ۲۶ درصدی در فرآیند معکوس-ژئوکدینگ می‌شود. همچنین استفاده از اطلاعات بافتی می-

^۲ Location Distortion Models

^۱ Venue

• ارزیابی نتایج بدست آمده از تحقیق به کمک داده‌های استخراج شده از محدوده زمانی خارج از داده‌های آموزشی مدل

سرویس Swarm، خدمتی جدید از شبکه اجتماعی مکان مبنا Foursquare است که امکان جمع‌آوری اعلام‌حضورها را از طریق Streaming API سرویس Twitter می‌دهد؛ از این‌رو برای این تحقیق از این سرویس استفاده شد. ادامه این نوشتار بر اساس ساختار زیر بنا نهاده شده است. قسمت دوم این نوشتار به بررسی پیشینه تحقیق می‌پردازد. در قسمت سوم به روش‌شناسی تحقیق و داده‌های مورد استفاده پرداخته می‌شود. بررسی مدل‌های اعوجاج بر مبنای طیف‌های معنایی جوی و شاخص‌های ارزیابی دقت در همین قسمت صورت می‌گیرد. قسمت چهارم نیز به پیاده‌سازی تحقیق، ارائه نتایج و ارزیابی مدل می‌پردازد. سرانجام نتیجه‌گیری و ارائه پیشنهادات برای تحقیقات آتی در قسمت پنجم آورده می‌شود.

۲- پیشینه تحقیق

تحقیقات موجود در زمینه خدمات جغرافیایی^۱ مربوط به کاربران و دستگاه‌های موبایل را می‌توان به دو بخش اصلی تقسیم کرد. بخش اول بر روی مباحث فنی مرتبط با تعیین موقعیت افراد (Basiri et al., 2017)، افزایش دقت فناوری‌های مکان مبنا (Falco et al., 2017) و همچنین افزایش کارایی خدمات مکانی در دستگاه موبایل (Kim et al., 2016) متمرکز می‌باشد. هدف اصلی این تحقیقات کاهش عدم قطعیت مرتبط با موقعیت جغرافیایی ابزارهای همراه می‌باشد و کمتر به مباحث غیرفنی توجه می‌کند.

بخش دوم تحقیقات مرتبط با مبحث توصیه‌گری می‌باشد (Bao et al., 2015). بسیاری از این تحقیقات از اطلاعاتی مانند اعلام‌حضورها و نظرات، برای بررسی شباهت کاربران بهره برده‌اند (McKenzie et al., 2013; Yuan et al., 2013; Mohamed & Abdelmoty, 2017). علاوه بر این تحقیقاتی نیز با هدف بررسی الگوهای زمانی رفتار کاربران و کاربرد آنها در توصیه‌گری نقطه مورد علاقه صورت گرفته است (Bannur & Alonso, 2014; Gao et al., 2013). پژوهش‌های متنوعی نیز با هدف بررسی الگوهای حرکتی انسان‌ها توسعه داده شده است (Bottai

et al., 2006; Rhee et al., 2011; Palmer et al., 2013; Yuan et al., 2012).

خیری و همکاران (۱۳۹۵) در تحقیقی به شناسایی الگوهای حرکتی درون‌شهری بر مبنای داده‌های شبکه‌های اجتماعی مکان‌مبنا پرداخته‌اند. در این تحقیق از دو مدل تابشی و PWO^۲ به منظور برآورد ماتریس مبدا-مقصد استفاده شده است. نتایج ارزیابی ماتریس مبدا-مقصد مدل شده نشان می‌دهد که سفرهای پیش‌بینی شده حدود ۷۰٪ سفرهای واقعی را بازسازی می‌نماید. گودرزی و همکاران (۱۳۹۶) در تحقیقی به توسعه الگوریتم توصیه‌گر مکان احتمالاتی بر اساس دسته‌بندی مکان در شبکه‌های اجتماعی مکان‌مبنا پرداخته‌اند. در این پژوهش، الگوریتم توصیه مکان PCLRTGS^۳ با در نظر گرفتن تأثیر مکانی، تأثیر زمانی و تأثیر روابط اجتماعی توسعه داده شده است. نتایج ایشان نشان می‌دهد که الگوریتم پیشنهادی PCLRTGS از نظر دقت و جامعیت، حدود ۱۵-۱۰ درصد عملکرد بهتری نسبت به دو الگوریتم PCLR^۴ و sPCLR^۵ دارد. کریمی پور و همکاران (۱۳۹۷) در تحقیقی به تعیین کاربری اجتماعی در محیط‌های شهری بر مبنای بعد معنایی داده‌های شبکه‌های اجتماعی مکان‌مبنا پرداختند. این تحقیق به خوشه‌بندی محیط‌های شهری بر مبنای کاربری‌های اجتماعی با استفاده از داده‌های مکانی و معنایی شبکه‌های اجتماعی مکان‌مبنا پرداخته است. نتایج این تحقیق، نشان‌دهنده پتانسیل بالای شبکه‌های اجتماعی مکان‌مبنا برای شناخت محیط‌های شهری می‌باشد.

Shaw و همکاران (۲۰۱۳) به بررسی نحوه انتساب تخمین موقعیت کاربران به یک نقطه مورد علاقه مناسب و بامعنی پرداخته‌اند. در این تحقیق تنها به مولفه‌های مکانی و زمانی داده‌های اعلام‌حضور کاربران توجه شده است و به جزئیات دسته‌بندی‌ها (مانند: رستوران ایتالیایی، مشاور املاک و باغ وحش) توجه نشده است. بعلاوه در این تحقیق، به مسئله اعوجاج مکان به کمک تابعی از زمان توجه نشده است. نتایج این تحقیق نشان داد که مدل‌های یادگیری ماشین نسبت به روش‌های رایج بهتر عمل می‌کنند. نوع داده‌های مورد استفاده در این تحقیق مختص به کارکنان

^۲ Population-Weighted Opportunities

^۳ Probabilistic Category-based Location Recommendation Utilizing Temporal Influence, Geographical Influence and Social Influence

^۴ Probabilistic Category-based Location Recommendation

^۵ Probabilistic Category-based Location Recommendation Utilizing Temporal Influence and Geographical Influence

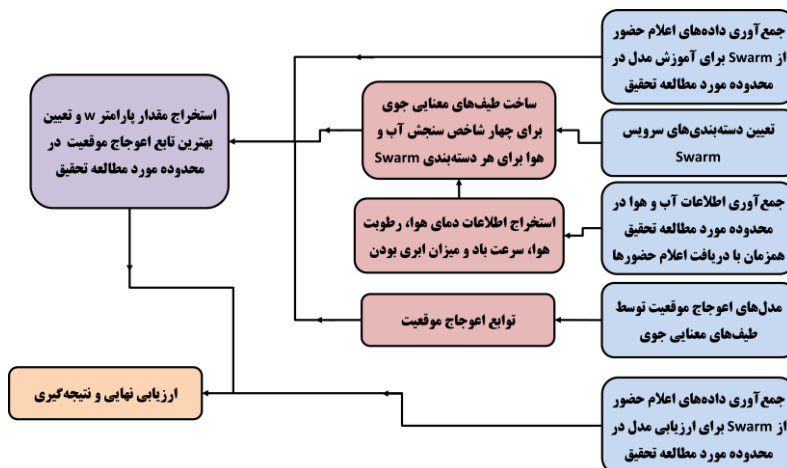
^۱ Geolocation Services

بر اساس منطقه‌ای خارج از محدوده آموزشی مدل صورت گرفت. نتایج این تحقیق نشان داد که با کمک مدل اعوجاج موقعیت خطی و با استفاده از طیف‌های معنایی زمانی می‌توان معکوس‌ژئوکدینگ بر مبنای فاصله جغرافیایی را ۱۵ تا ۳۰ درصد بهبود داد. Trattner و همکاران (۲۰۱۶) در تحقیق خود نشان دادند که با کمک اطلاعات بافتی شرایط جوی می‌توان دقت فرآیند توصیه‌گری نقاط مورد علاقه را بهبود داد. همانطور که مشاهده می‌شود تحقیقات متنوعی در مورد مسائل فنی و غیرفنی شبکه‌های اجتماعی مکان‌مبنا صورت گرفته است. ولی تحقیقات محدودی به بررسی فرآیند معکوس‌ژئوکدینگ پرداخته‌اند. از طرفی نمی‌توان تحقیقی را یافت که به مسئله امکان استفاده از مجموعه داده‌های بافتی مختلف (مانند: داده‌های شرایط آب و هوا) در فرآیند معکوس‌ژئوکدینگ پرداخته باشد. این خلأ به عنوان موضوع اصلی مورد بررسی در تحقیق حاضر می‌باشد.

۳- روش‌شناسی تحقیق

این بخش از تحقیق در ابتدا به بررسی جزئیات مربوط به نحوه جمع‌آوری داده‌های اعلام‌حضور سرویس Swarm به همراه اطلاعات شرایط آب و هوایی برای پارامترهای دمای هوا، رطوبت هوا، سرعت باد و میزان ابری بودن می‌پردازد. در مرحله بعدی جزئیات مربوط به ایجاد طیف‌های معنایی جوی دسته‌بندی‌های سرویس Swarm ارائه می‌شود. سپس مدل‌های اعوجاج موقعیت (خطی، سینوسی و نسبی) و توابع مربوط به آنها معرفی و بررسی می‌شود. در نهایت شاخص‌های بررسی دقت برای آموزش و ارزیابی نهایی نتایج معرفی می‌شوند. شکل ۱ شمای کلی مراحل انجام تحقیق را نمایش می‌دهد.

سرویس Foursquare می‌باشد که افراد دیگر به این داده‌ها دسترسی نداشته و همچنین نتایج این تحقیق محدود به سرویس Foursquare می‌باشد. Lian و Xie (۲۰۱۱) یک روش نام‌گذاری موقعیت را ارائه دادند. این روش بر اساس مشاهده GPS، زمان و تاریخچه اعلام‌حضور افراد نام موقعیتی که فرد حضور دارد را تعیین می‌کند. یکی از محدودیت‌های این تحقیق وابستگی به تاریخچه اعلام‌حضورهای فرد می‌باشد. این مورد برای افراد ناشناس می‌تواند منجر به کاهش دقت در نتایج شود. Bhattacharya و همکاران (۲۰۱۵) روشی را برای شناسایی نقطه مورد علاقه مرتبط از داده‌های غیرقابل اتکا GPS به کمک تخمین تراکم مکانی-زمانی و تقاطع خطوط معرفی نمودند. تعداد دسته‌بندی‌های نقاط مورد علاقه شناسایی شده در این تحقیق محدود به پنج عدد بوده است. مطالعه McKenzie و Janowicz (۲۰۱۵) در راستای رفع نقایص پژوهش‌های پیشین انجام گرفت. در این تحقیق با استفاده از ۴ مدل اعوجاج موقعیت، نحوه ترکیب احتمالات زمانی با فاصله جغرافیایی (بین فرد و محل) مورد بررسی قرار گرفت. سپس از این مدل در بهبود فرآیند معکوس‌ژئوکدینگ استفاده شد. یکی از ضعف‌های این تحقیق ارزیابی مدل با داده‌هایی از مکان مشابه با داده‌های آموزشی می‌باشد. در نتیجه ممکن است این مدل برای مکان‌های دیگر و داده‌های دیگر سرویس‌ها نتایج مناسبی به همراه نداشته باشد. Sabzali Yameqani و Alesheikh (۲۰۱۹) در تحقیقی با ارائه یک الگوریتم تناظریابی دسته‌بندی‌های سرویس Yelp و Swarm بخشی از کاستی‌های تحقیق McKenzie و Janowicz (۲۰۱۵) را برطرف نمودند. در این تحقیق طیف‌های معنایی زمانی سرویس Swarm بر اساس اعلام‌حضورهای سرویس Yelp ایجاد شد. فرآیند ارزیابی نتایج



شکل ۱- شمای کلی مراحل انجام تحقیق

۳-۱- جمع‌آوری داده‌های اعلام‌حضور سرویس Swarm

داده‌ی اعلام‌حضور مورد استفاده در این تحقیق از سرویس Swarm استخراج شد. کاربران Swarm می‌توانند اعلام‌حضورهای خود را به اشتراک بگذارند و یا در محل‌های مختلفی ارسال نظر انجام دهند. در Swarm اطلاعات مربوط به نظرات افراد در مورد محل‌های مختلف به صورت عمومی قابل دسترسی است. ولی اطلاعات مربوط به اعلام‌حضور فقط با اجازه کاربران قابل دسترسی می‌باشد. بنابراین امکان دسترسی به حجم زیادی از اعلام‌حضورها از این طریق تقریباً غیرممکن است؛ چراکه نیاز به مشارکت تعداد زیادی کاربر دارد. خوشبختانه کاربران Swarm می‌توانند اعلام‌حضورهای خود را توسط Twitter به صورت توثیق به اشتراک بگذارند. توثیق‌ها عمومی می‌باشند و دسترسی به آنها فراهم است. اعلام‌حضورها در Twitter به کمک یک URL کوتاه به مانند: "https://www.swarmapp.com/c/Check-in-Id" قابل شناسایی می‌باشند. به این ترتیب با بررسی برخط جریان توثیق‌های Streaming API سرویس Twitter می‌توان به اطلاعات اعلام‌حضور به اشتراک گذاشته شده دسترسی پیدا کرد. بعد از استخراج کد ۱۱ کارکتری هر اعلام‌حضور، می‌توان آن را توسط API سرویس Swarm بازپایی کرد. بعد از استخراج اطلاعات اعلام‌حضور مربوطه، اطلاعات مربوط به شرایط جوی نیز با کمک OpenWeatherMap استخراج شد. در نهایت این اطلاع بصورت برخط در فایل خروجی ذخیره گشت.

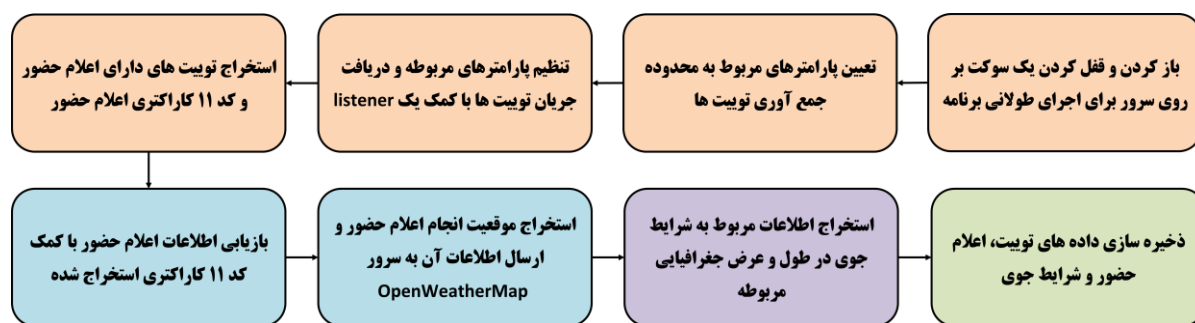
در این تحقیق داده‌های اعلام‌حضور سرویس Swarm شامل اعلام‌حضورهایی می‌باشد که توسط Streaming

API سرویس Twitter برای ایالت نیویورک به مدت یک سال از تاریخ 01/03/2017 تا 01/03/2018 جمع‌آوری شده است. همچنین اطلاعات مربوط به شرایط جوی شامل دمای هوا، رطوبت هوا، سرعت باد و میزان ابری بودن با کمک سرویس OpenWeatherMap به صورت آنی تهیه شد. این اطلاعات به منظور تهیه طیف‌های معنایی جوی مربوط به دمای هوا، رطوبت هوا، سرعت باد و میزان ابری بودن مورد استفاده قرار گرفت. تعداد کل اعلام‌حضورهای مورد استفاده در این تحقیق برای ایالت نیویورک در جدول ۲ قابل مشاهده است.

جدول ۲- تعداد اعلام‌حضورهای موجود در مجموعه داده تحقیق

محدوده مطالعه موردی	تعداد اعلام‌حضورها
ایالت نیویورک	۵۰۳۷۸۹

برای جمع‌آوری اطلاعات مورد نیاز این تحقیق از سرویس‌های رایانش ابری وبسایت PythonAnywhere استفاده شده است. بدین منظور با کمک یک برنامه به زبان پایتون ابتدا جریان توییت‌ها به صورت برخط بررسی شده است. در صورتی که توییتی به همراه اطلاعات اعلام‌حضور (کد ۱۱ کاراکتری) بوده باشد؛ اطلاعات اعلام‌حضور به صورت برخط استخراج شده است. سپس در همان لحظه اطلاعات آب و هوا برای موقعیت آن اعلام‌حضور از سرویس OpenWeatherMap به صورت جدول ۱ تهیه شده است. در نهایت اطلاعات بدست آمده در یک فایل متنی ذخیره شده است تا پردازش‌ها بعدی روی آنها صورت گیرد. شکل ۲ شمای کلی این فرآیند را نمایش می‌دهد.



شکل ۲- شمای کلی فرآیند جمع‌آوری داده‌های مورد نیاز تحقیق

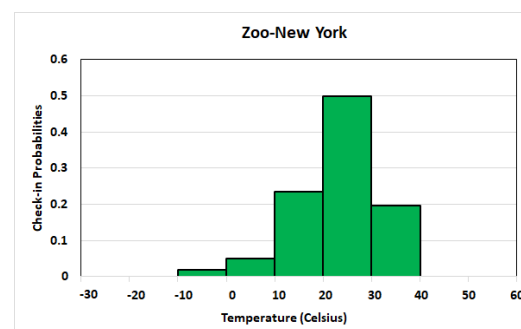
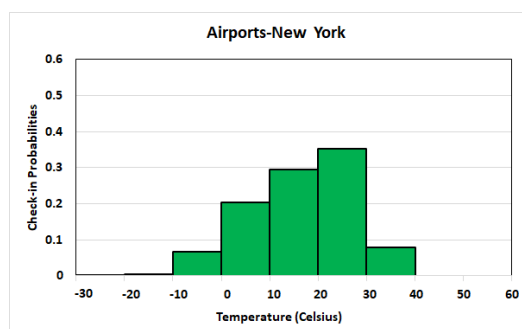
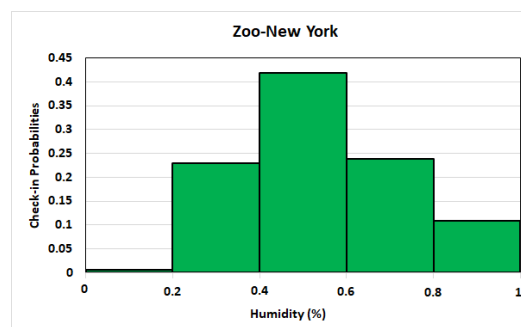
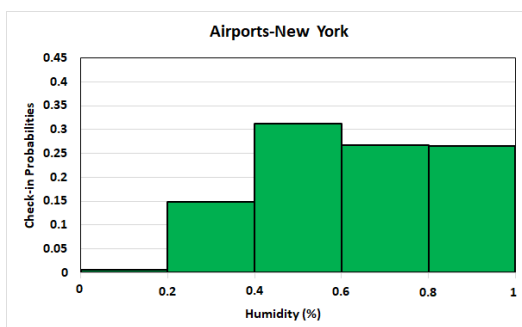
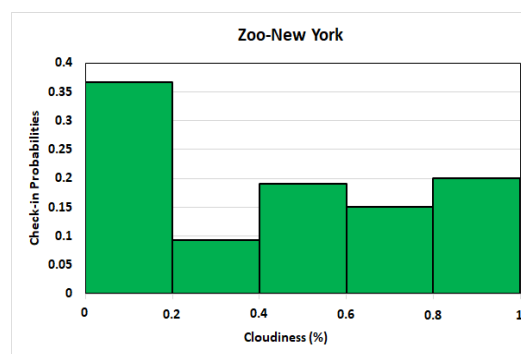
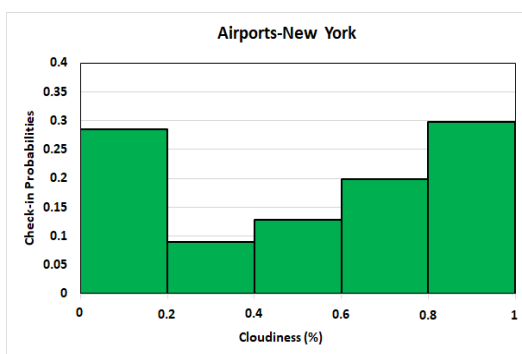
۳-۲- تهیه طیف‌های معنایی جوی

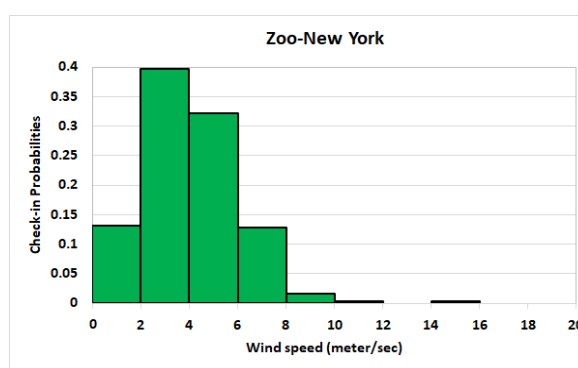
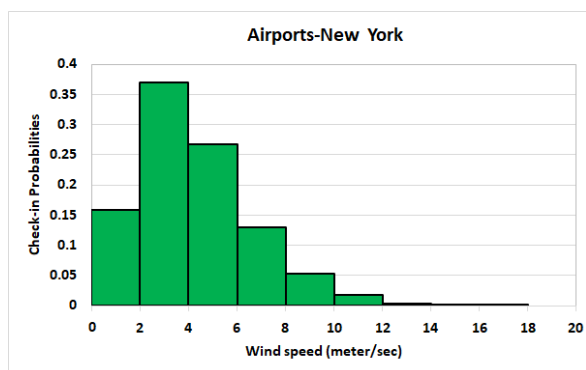
با کمک تمام اعلام‌حضورهای جمع‌آوری شده، طیف‌های معنایی جوی برای دسته‌بندی‌های سرویس Swarm

تهیه شد. برای تعیین تعداد دسته‌بندی‌ها، تمامی داده‌ها مربوط به ایالت نیویورک مورد بررسی قرار گرفت و تمامی دسته‌بندی‌های غیرتکراری استخراج شد. اگر چه تعداد کل دسته‌بندی‌های سرویس Swarm ۹۲۱ عدد می‌باشد؛ تعداد

دسته‌بندی‌های این تحقیق ۵۹۹ عدد شد. مقایسه نتایج نشان داد که تعداد ۳۲۲ عدد از دسته‌بندی‌ها در داده‌های تحقیق هیچ نمونه‌ای نداشته‌اند. این طیف‌ها برای اطلاعات ایالت نیویورک در پایگاه داده ایجاد شد. هر کدام از این طیف‌ها با توجه به نوع تفکیک داده‌ها دارای تعداد کلاس‌های مختلفی می‌باشند. تعداد و محدوده کلاس‌ها در شکل ۳ قابل مشاهده است. از این اطلاعات به عنوان ورودی فرآیند معکوس ژئوکدینگ استفاده شد. جزئیات مربوط به نحوه استفاده از این طیف‌ها در مباحث بعدی اشاره می‌شود. در شکل ۳ تعدادی از این طیف‌ها (برای دسته‌بندی‌های فرودگاه و باغ وحش) ارائه شده است. برای نمایش بهتر و امکان مقایسه، تمامی طیف‌ها بر اساس حاصل جمع اعلام‌حضورهای هر دسته‌بندی نرمال‌سازی شده است. به این ترتیب مجموع اعداد هر طیف مقدار یک می‌باشد. برای آشنایی بهتر با نحوه تهیه طیف‌های معنایی جوی، مراحل ساخت یکی از آنها یعنی میزان رطوبت هوا برای

دسته‌بندی‌های این تحقیق ۵۹۹ عدد شد. مقایسه نتایج نشان داد که تعداد ۳۲۲ عدد از دسته‌بندی‌ها در داده‌های تحقیق هیچ نمونه‌ای نداشته‌اند. این طیف‌ها برای اطلاعات ایالت نیویورک در پایگاه داده ایجاد شد. هر کدام از این طیف‌ها با توجه به نوع تفکیک داده‌ها دارای تعداد کلاس‌های مختلفی می‌باشند. تعداد و محدوده کلاس‌ها در شکل ۳ قابل مشاهده است. از این اطلاعات به عنوان ورودی فرآیند معکوس ژئوکدینگ استفاده شد. جزئیات مربوط به نحوه استفاده از این طیف‌ها در مباحث بعدی اشاره می‌شود. در شکل ۳ تعدادی از این طیف‌ها (برای دسته‌بندی‌های فرودگاه و باغ وحش) ارائه شده است. برای نمایش بهتر و امکان مقایسه، تمامی طیف‌ها بر اساس حاصل جمع اعلام‌حضورهای هر دسته‌بندی نرمال‌سازی شده است. به این ترتیب مجموع اعداد هر طیف مقدار یک می‌باشد. برای آشنایی بهتر با نحوه تهیه طیف‌های معنایی جوی، مراحل ساخت یکی از آنها یعنی میزان رطوبت هوا برای



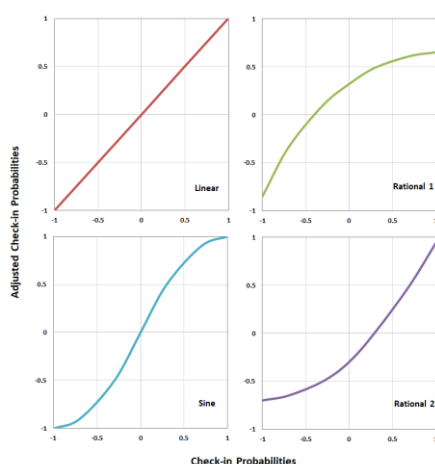


شکل ۳- طیف‌های معنایی جوی برای دو دسته‌بندی فرودگاه و باغ وحش در مجموعه داده‌ی نیویورک

۳-۳- مدل‌های اعوجاج بر مبنای طیف‌های معنایی جوی

در این بخش مدل‌های مورد استفاده جهت اعوجاج مقادیر طیف‌های معنایی جوی معرفی می‌شود. همچنین روابط مربوط به آنها، نحوه انتخاب پارامترهای مدل و شاخص‌های اندازه‌گیری دقت ارائه می‌شود.

۳-۳-۱- مدل‌های اعوجاج



شکل ۴- چهار مدل اعوجاج و نمونه نمایش بصری آنها برای تغییر موقعیت نقاط مورد علاقه بر مبنای طیف‌های معنایی جوی (McKenzie et al., 2015)

استفاده از داده‌های بافتی در فرآیند معکوس-ژئوکدینگ موجب افزایش دقت می‌شود (McKenzie & Janowicz, 2015). ترکیب احتمالات جوی و فاصله فرد از محل‌ها، توسط مدل‌های اعوجاج صورت می‌گیرد. هر یک از این مدل‌ها دارای ویژگی‌هایی می‌باشند که در ادامه به آنها اشاره می‌شود. مدل خطی به صورت متقارن فاصله فرد تا نقاط مورد علاقه را تعدیل می‌کند. به این ترتیب نقاط مورد علاقه با احتمالات جوی پایین از فرد دورتر شده و نقاط مورد علاقه با احتمالات جوی بالاتر با نرخ خطی مشابه به فرد نزدیک می‌شوند. مدل سینوسی اگر چه دارای ساختاری متقارن است ولی به صورت غیرخطی عمل می‌کند. این مدل برای احتمالات نزدیک به میانگین به صورت خطی عمل می‌کند ولی برای احتمالات بالا یا پایین به صورت متفاوت عمل می‌نماید. مدل‌های نسبی به صورت غیرخطی و غیرمتقارن می‌باشند. در حالت نسبی ۱ هرچه احتمال جوی کاربر برای یک نقطه مورد علاقه افزایش یابد، مقداری که با آن فاصله کاهش پیدا می‌کند،

سیستم‌های معکوس ژئوکدینگ رایج بر مبنای موقعیت فرد، لیستی از نقاط مورد علاقه را بر مبنای فاصله آنها بازگشت می‌دهند. حال آنکه پارامترهای متنوع دیگری وجود دارند که می‌توانند ترتیب نتایج را تغییر دهند. احتمالات جوی یکی از این پارامترها می‌باشد. احتمالات جوی از ۴ مولفه تشکیل شده‌اند. این ۴ مولفه شامل دمای هوا، رطوبت هوا، سرعت باد و میزان ابری بودن می‌باشند. به عبارت دیگر به مانند شکل ۳ چهار مقدار مختلف از احتمال اعلام حضور محاسبه می‌شود. سپس این مقادیر با یکدیگر میانگین گرفته می‌شود. این مقدار به عنوان احتمال اعلام حضور جوی در نقطه مورد علاقه مدنظر مورد استفاده قرار می‌گیرد. برای ترکیب فاصله جغرافیایی با احتمالات جوی از مدل‌های اعوجاج استفاده می‌شود. منظور از مدل‌های اعوجاج مدل‌های مورد استفاده جهت تغییر موقعیت نقاط مورد علاقه بر مبنای مقادیر احتمالات جوی در سطح دسته‌بندی می‌باشد. احتمالات جوی برای دسته‌بندی‌های سرویس Swarm محاسبه شده است. در این تحقیق ۴ مدل شامل یک مدل خطی، دو مدل نسبی و یک مدل سینوسی به مانند شکل ۴ مورد استفاده قرار گرفته است.

توابع ترکیب الگوریتم‌های فاصله مبنای موجود با احتمالات جوی می‌باشد. در ابتدا به معرفی متغیرهای مورد استفاده در این توابع پرداخته می‌شود. میزان میانگین احتمالات جوی برای تمامی محل‌های اطراف هر اعلام‌حضور با نماد t'_m نمایش داده می‌شود. برای اینکه بتوان مشخص کرد که یک محل باید نزدیک یا دور شود، میزان احتمال جوی آن (t') از میزان میانگین کسر می‌شود. متغیر حاصل شده ($\tilde{t}'$) میزان و جهت اعوجاج محل-ها را مشخص می‌کند.

$$\tilde{t}' = t'_m - t' \quad (1)$$

متغیر d' برابر با فاصله نرمال شده میان نقطه مورد علاقه مورد نظر و محل اعلام‌حضور می‌باشد. متغیر d_t نیز به عنوان فاصله اعوجاج یافته می‌باشد. در نهایت یک متغیر نیز با نام w جهت تعیین میزان تاثیر فاصله جغرافیایی در مقابل احتمالات جوی در نظر گرفته شده است. با توجه به متغیرهای ذکر شده ۴ تابع اعوجاج به صورت زیر در نظر گرفته شد.

جدول ۳- چهار تابع اعوجاج مورد استفاده در تحقیق

فرمول	نوع تابع
$d_t = d' \cdot w + \tilde{t}' \cdot (1 - w)$	نوع خطی
$d_t = d' - \sin(\tilde{t}') w$	نوع سینوسی
$d_t = d' - (1 - \frac{w}{\tilde{t}' + w})$	نوع نسبی ۱
$d_t = d' - (\frac{w}{-\tilde{t}' + w} - 1)$	نوع نسبی ۲

هر کدام از این توابع خصوصیات منحصر به فردی را دارا می‌باشند که در بخش قبلی به تفصیل ارائه شد. در همه‌ی روابط جدول ۳ پارامتری به نام w وجود دارد که مقدار آن مشخص نیست. این پارامتر میزان تاثیر فاصله جغرافیایی و احتمالات جوی را در فرآیند معکوس-ژئوکدینگ مشخص می‌کند. برای تعیین این پارامتر از داده‌های آموزشی کمک گرفته می‌شود. بدین منظور ابتدا باید شاخص‌های بررسی دقت فرآیند معکوس-ژئوکدینگ مشخص شود. در این تحقیق از دو شاخص MRR^۱ و First Position استفاده شده است. شاخص MRR به صورت فرمول ۲ تعیین می‌شود:

$$MRR = \frac{1}{|N|} \sum_{i=1}^{|N|} \frac{1}{rank_i} \quad (2)$$

^۱ Mean Reciprocal Rank

کاسته می‌شود. متقابلاً، هرچه احتمال جوی کاهش می‌یابد، مقداری که با آن فاصله افزایش می‌یابد، رشد می‌کند. مدل نسبی ۲ نیز به صورت معکوس مدل نسبی ۱ عمل می‌کند. منطقی که مدل‌های نسبی از آن پیروی می‌کنند این است که نزدیک‌شدن یا دورشدن نقطه مورد علاقه باید با نرخ متفاوت انجام پذیرد.

۳-۳-۲- تهیه داده‌های آموزش مدل اعوجاج

برای فرآیند آموزش مدل اعوجاج داده‌های اعلام‌حضور جمع‌آوری شده برای ایالت نیویورک به مدت یک سال از تاریخ 01/03/2017 تا 01/03/2018 (جدول ۲) استفاده شد. در این تحقیق تمامی اعلام‌حضورهای مورد استفاده برای ساخت طیف‌های معنایی جوی جهت آموزش مدل بکار گرفته شد. بدین منظور اعلام‌حضورها به کمک فیلتری پالایش شد تا فقط اعلام‌حضورهای واجد شرایط انتخاب شوند. فرآیند این فیلتر به این صورت بود که ابتدا موقعیت اعلام‌حضور فرد به عنوان موقعیت حضور فرد انتخاب می‌شود. سپس برای شبیه‌سازی خطای GPS موبایل، مقدار خطایی از طریق توزیع نرمالی با میانگین ۳۰ متر و انحراف معیار ۱۰ متر تولید می‌شود. این خطا در یک جهت تصادفی به موقعیت اولیه اعلام‌حضور اعمال می‌شود (McKenzie & Janowicz, 2015). حال موقعیت جدید به عنوان موقعیت فرد ذخیره می‌گردد. سپس به کمک درخواست Venues از API سرویس Swarm تعداد محل‌های اطراف فرد به همراه فاصله آنها دریافت می‌شود. این درخواست با متغیر intent با مقدار match برای سرور ارسال شد تا نتایج دقیقاً منطبق بر موقعیت مورد درخواست باشند. در صورتی که حداقل ۱۵ محل اطراف فرد در فاصله ۱۰۰ متری باشد؛ این اعلام‌حضور به عنوان واجد شرایط انتخاب شده و ذخیره می‌شود. این فرآیند برای انتخاب اعلام‌حضورهای واجد شرایط به منظور آموزش مدل اعوجاج مورد استفاده قرار گرفت.

۳-۳-۳- توابع اعوجاج معنایی جوی و نحوه

پارامتری‌سازی و ارزیابی آنها

در بخش قبل ۴ مدل اعوجاج برای احتمالات جوی ارائه شد. در این بخش به معرفی ۴ تابع اعوجاج بر مبنای مدل‌های اعوجاج ارائه شده، پرداخته می‌شود. هدف این

که در آن rank همان رتبه مربوط به نقطه مورد علاقه مورد نظر در فرآیند معکوس‌ژئوکدینگ می‌باشد. N نیز تعداد کل اعلام‌حضورهای مورد بررسی در هر اجرا می‌باشد. شاخص First Position نیز تعداد دفعاتی را شمارش می‌کند که خروجی فرآیند معکوس‌ژئوکدینگ در رتبه اول قرار می‌گیرد. به عبارت دیگر تعداد نقاط مورد علاقه درست شناسایی شده را مشخص می‌کند.

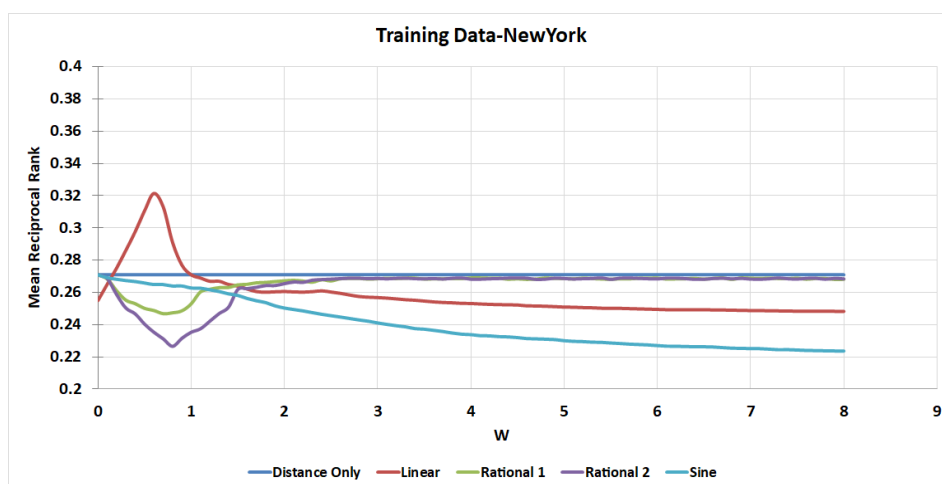
۴- آموزش توابع اوجاج موقعیت و ارزیابی آنها

در این بخش توابع مورد استفاده جهت اوجاج مقادیر طیف‌های معنایی جوی آموزش داده می‌شود. سپس نتایج حاصل از داده‌های آموزشی در قالب نمودار و جدول ارائه

می‌شود. در نهایت نتایج بدست آمده توسط داده‌های تست مورد ارزیابی قرار می‌گیرد.

۴-۱- آموزش توابع اوجاج موقعیت

مدل‌های اوجاج موقعیت و توابع مربوط به آنها در بخش قبل معرفی شد. این توابع همگی دارای متغیری به نام w می‌باشند. برای تعیین بهترین مقدار متغیر w برای هر یک از توابع جدول ۳، نیاز به یک مبنا مقایسه است. در این تحقیق میزان خروجی MRR فرآیند معکوس‌ژئوکدینگ بر اساس فاصله جغرافیایی مبنا قرار گرفت. متغیر w نیز از مقدار صفر به صورت افزایشی ۰٫۱ تغییر یافت. خروجی شاخص MRR برای هر یک از توابع جدول ۳ برای داده‌های آموزشی ایالت نیویورک به مانند شکل ۵ است.



شکل ۵- مقدار شاخص MRR برای توابع مختلف اوجاج برای داده‌های آموزشی در مقایسه با حالت فاصله مبنا در محدوده نیویورک

نتایج شاخص MMR برای داده‌های آموزشی نیویورک برای w با مقدار ۰٫۶ بیشینه می‌باشد. در نتیجه مقدار نهایی w برای ایالت نیویورک برابر با ۰٫۶ در نظر گرفته شد. میزان بهبودها برای دو شاخص MRR و First Position در حالت‌های مختلف در جدول ۴ ثبت شده است. همان‌طور که مشاهده می‌شود میزان بهبود شاخص MRR و

First Position برای مجموع داده نیویورک به ترتیب ۱۸٫۶۴٪ و ۱۱۱٫۴۹٪ می‌باشد. نتایج شکل ۵ نشان می‌دهد که مدل استفاده شده در این تحقیق فقط با توابع اوجاج خطی نتایج مناسبی را تولید می‌کند. به عبارت دیگر استفاده از توابع سینوسی و نسبی موجب بهبود نتایج نمی‌شود.

جدول ۴- نتایج شاخص‌های MRR و First Position برای داده‌های آموزشی

W	First Position	MRR	تابع اعوجاج	مطالعه موردی
	درصد بهبود			
N/A	17531	0.2709	فاصله مبنا	نیویورک
0.6	37077 (+111.49%)	0.3214 (+18.64%)	نوع خطی	
N/A	N/A	N/A	نوع نسبی ۱	
N/A	N/A	N/A	نوع نسبی ۲	
N/A	N/A	N/A	نوع سینوسی	

۴-۲- ارزیابی نتایج

در این تحقیق برای ارزیابی دقت مدل از یک مجموعه داده دوم استفاده شد. بدین منظور داده‌های ایالت نیویورک برای تاریخ 01/03/2018 تا 07/03/2018 مثل قبل جمع‌آوری شده و مانند داده‌های آموزشی فیلتر شدند. استفاده از داده‌های مربوط به محدوده زمانی خارج از داده‌های آموزشی باعث می‌شود تا بتوان میزان عمومیت مدل را سنجش کرد. به عبارت دیگر می‌توان امکان استفاده از مدل ارائه شده در زمانی دیگر در همان منطقه را بررسی نمود. بعد از اجرای مدل مذکور نتایج جدول ۵ حاصل شد. همان‌طور که مشاهده می‌شود میزان بهبود

جدول ۵- نتایج شاخص‌های MRR و First Position برای داده‌های تست

W	First Position	MRR	تابع اعوجاج	مطالعه موردی
	درصد بهبود			
N/A	460	0.3142	فاصله مبنا	نیویورک
0.6	768 (+66.96%)	0.3563 (+13.40%)	نوع خطی	

می‌توان از دیگر اطلاعات بافتی در جهت بهبود فرآیند معکوس‌ژئوکدینگ استفاده کرد. یکی از نقاط قوت این تحقیق استفاده از یک مطالعه موردی وسیع در فرآیند تحقیق می‌باشد. یکی دیگر از نقاط قوت این تحقیق استفاده از داده‌های شرایط جوی شامل: دمای هوا، رطوبت هوا، سرعت باد و میزان ابری بودن برای بهبود فرآیند معکوس-ژئوکدینگ می‌باشد. نتایج این تحقیق نشان داد که می‌توان با کمک اطلاعات بافتی مختلف فرآیند معکوس‌ژئوکدینگ را بهبود داد. نتایج این تحقیق در مقایسه با نتایج McKenzie و Janowicz (۲۰۱۵) نشان داد که استفاده از توابع اعوجاج سینوسی و نسبی نتایج مناسبی را در بر ندارد. ولی استفاده از مدل خطی می‌تواند به بهبود نزدیک ۱۹ درصدی شاخص MRR و بیش از ۱۰۰ درصدی شاخص First Position منجر شود.

یکی از محدودیت‌های اصلی این تحقیق عدم دسترسی به جریان کامل توییت‌ها بوده است. با توجه به اینکه امکان دسترسی به آنها به صورت رایگان فراهم نبوده است؛ در نتیجه از آنها صرف نظر شد. اما دسترسی به آنها می‌تواند موجب شود تا طیف‌های معنایی جوی با تعداد اعلام-حضورها بیشتری تهیه شود. به این ترتیب می‌توان انتظار داشت که نتایج تحقیق نیز بهبود یابد. تهیه طیف‌های فصلی و آزمایش دقت آنها در مقابل طیف‌های مشابه

۵- نتیجه‌گیری و پیشنهادات

توسعه خدمات مکان مبنا در سال‌های اخیر موجب ایجاد فرصت‌های بسیاری شده است. یکی از آنها، تولید و انتشار مجموعه داده‌هایی غنی از رفتارهای کاربران می‌باشد. جمع‌آوری این نوع داده‌ها در مقیاس بالا در گذشته با مشکلات بسیاری همراه بود. با کاهش شدت این مشکلات، تحقیقات متنوعی در مباحث مختلف فنی مانند: بهبود نحوه تعیین موقعیت و یا بهینه کردن مصرف منابع و غیرفنی مانند: توصیه‌گری مکانی، شباهت‌سنجی کاربران و معکوس‌ژئوکدینگ صورت گرفته است. تمامی این تحقیقات با هدف بهبود خدمات مکان مبنا برای کاربران کاهش مشکلات این سیستم‌ها می‌باشد.

در این تحقیق یکی از چالش‌های مهم در بحث خدمات مکانی یعنی فرآیند معکوس‌ژئوکدینگ مورد بررسی قرار گرفت. بدین منظور مدل‌های مختلف اعوجاج معرفی و بررسی شد. مدل ارائه شده در این تحقیق توانست با کمک تابع اعوجاج خطی فاصله بین فرد تا محل‌ها را اصلاح کند. با توجه به اینکه این مدل توانایی خود را برای استفاده به همراه طیف‌های معنایی زمانی و معنایی جوی نشان داده است؛ می‌توان انتظار داشت که اطلاعات بافتی دیگر را نیز بتوان با کمک این مدل مورد استفاده قرار داد. به این ترتیب

تحقیق حاضر می‌تواند موضوع تحقیق دیگری قرار گیرد. البته این موضوع نیازمند دسترسی به جریان کامل توپیت-ها برای تهیه طیف‌های فصلی با دقت مناسب می‌باشد؛ یا مثلاً می‌توان اطلاعات اعلام حضور را در دو یا چند سال تهیه کرد تا امکان تهیه طیف‌های فصلی فراهم شود. برای تحقیقات آتی پیشنهاد می‌شود تا این مدل در محدوده‌های زمانی و مکانی دیگر پیاده‌سازی شود و کارایی

آن سنجیده شود. استفاده از مجموعه داده‌های بافتی به جز موارد مورد استفاده در این تحقیق نیز یکی دیگر از پیشنهادات تحقیقات آتی می‌باشد. در نهایت می‌توان مدل‌های اعوجاج دیگری را در تحقیقات آتی معرفی و مورد بررسی قرار داد.

مراجع

- [1] Kheiri A, Karimipour F, Forghani M. Intra-Urban Movement Patterns Estimation Based on Location Based Social Networking Data. JGST. 2016; 6 (1) :141-158.
- [2] Karimipour F, Tayebi M, Amozande K. Characterization of Social Land use in Urban Environments Based on the Semantic Dimension of Location Based Social Networks' Data. JGST. 2018; 7 (4) :133-145.
- [3] Gudarzi S, Alesheikh A A, Honarparvar S. Devloping a Recommender Location Algorithm Utilizing Temporal Influence, Geographical Influence and Social Influence . JGST. 2017; 7 (2) :215-230.
- [4] Bannur, S., & Alonso, O. (2014, April). Analyzing temporal characteristics of check-in data. In Proceedings of the 23rd International Conference on World Wide Web (pp. 827-832). ACM.
- [5] Bao, J., Zheng, Y., Wilkie, D., & Mokbel, M. (2015). Recommendations in location-based social networks: a survey. Geoinformatica, 19(3), 525-565.
- [6] Basiri, A., Lohan, E. S., Moore, T., Winstanley, A., Peltola, P., Hill, C., ... & e Silva, P. F. (2017). Indoor location based services challenges, requirements and usability of current solutions. Computer Science Review, 24, 1-12.
- [7] Bhattacharya, T., Kulik, L., & Bailey, J. (2015). Automatically recognizing places of interest from unreliable GPS data using spatio-temporal density estimation and line intersections. Pervasive and Mobile Computing, 19, 86-107.
- [8] Bottai, M., Salvati, N., & Orsini, N. (2006). Multilevel models for analyzing people's daily movement behavior. Journal of Geographical Systems, 8(1), 97-108.
- [9] Falco, G., Pini, M., & Marucco, G. (2017). Loose and Tight GNSS/INS Integrations: Comparison of Performance Assessed in Real Urban Scenarios. Sensors, 17(2), 255.
- [10] Gao, H., Tang, J., Hu, X., & Liu, H. (2013, October). Exploring temporal effects for location recommendation on location-based social networks. In Proceedings of the 7th ACM conference on Recommender systems (pp. 93-100). ACM.
- [11] Huang, R. K., & Piemonte, P. (2013). U.S. Patent No. 8,489,127. Washington, DC: U.S. Patent and Trademark Office.
- [12] Kim, D., Lee, S., & Bahn, H. (2016, August). An Energy-Efficient Positioning Scheme for Location-Based Services in a Smartphone. In Embedded and Real-Time Computing Systems and Applications (RTCSA), 2016 IEEE 22nd International Conference on (pp. 139-148). IEEE.
- [13] Kounadi, O., Lampoltshammer, T. J., Leitner, M., & Heistracher, T. (2013). Accuracy and privacy aspects in free online reverse geocoding services. Cartography and Geographic Information Science, 40(2), 140-153.
- [14] Lambert, C., Chu, M., & Seth, R. (2009). U.S. Patent Application No. 12/437,718.
- [15] Lian, D., & Xie, X. (2011, November). Learning location naming from user check-in histories. In Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 112-121). ACM.
- [16] McKenzie, G., Adams, B., & Janowicz, K. (2013). A thematic approach to user similarity built on geosocial check-ins. In Geographic information science at the heart of europe (pp. 39-53). Springer International Publishing.
- [17] McKenzie, G., Janowicz, K., & Adams, B. (2014). A weighted multi-attribute method for matching user-generated points of interest. Cartography and Geographic Information Science, 41(2), 125-137.

- [18] McKenzie, G. D. (2015). A temporal approach to defining place types based on user-contributed geosocial content. University of California, Santa Barbara.
- [19] McKenzie, G., & Janowicz, K. (2015). Where is also about time: A location-distortion model to improve reverse geocoding using behavior-driven temporal semantic signatures. *Computers, Environment and Urban Systems*, 54, 1-13.
- [20] Mohamed, S., & Abdelmoty, A. I. (2017). Spatio-semantic user profiles in location-based social networks. *International Journal of Data Science and Analytics*, 1-16.
- [21] Noulas, A., Scellato, S., Mascolo, C., & Pontil, M. (2011). An Empirical Study of Geographic User Activity Patterns in Foursquare. *ICWSM*, 11, 70-573.
- [22] Palmer, J. R., Espenshade, T. J., Bartumeus, F., Chung, C. Y., Ozgencil, N. E., & Li, K. (2013). New approaches to human mobility: Using mobile phones for demographic research. *Demography*, 50(3), 1105-1128.
- [23] Rhee, I., Shin, M., Hong, S., Lee, K., Kim, S. J., & Chong, S. (2011). On the levy-walk nature of human mobility. *IEEE/ACM transactions on networking (TON)*, 19(3), 630-643.
- [24] Searight, K. R., Logan, D. J., Freddie, J. B. I., Loher, C. J., & Charlton, B. R. (2010). U.S. Patent No. 7,668,651. Washington, DC: U.S. Patent and Trademark Office.
- [25] Sharif, M., & Alesheikh, A. A. (2017). Context-awareness in similarity measures and pattern discoveries of trajectories: a context-based dynamic time warping method. *GIScience & Remote Sensing*, 54(3), 426-452.
- [26] Shaw, B., Shea, J., Sinha, S., & Hogue, A. (2013, February). Learning to rank for spatiotemporal search. In *Proceedings of the sixth ACM international conference on Web search and data mining* (pp. 717-726). ACM.
- [27] Trattner, C., Oberegger, A., Eberhard, L., Parra, D., & Marinho, L. B. (2016, September). Understanding the Impact of Weather for POI Recommendations. In *RecTour@ RecSys* (pp. 16-23).
- [28] Yameqani, A. S., & Alesheikh, A. A. (2019). Evaluating a location distortion model to improve reverse geocoding through temporal semantic signatures. *Computers, Environment and Urban Systems*, 77, 101349.
- [29] Yuan, Z., Jiang, Y., & Gidófalvi, G. (2013, November). Geographical and temporal similarity measurement in location-based social networks. In *Proceedings of the Second ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems* (pp. 30-34). ACM.
- [30] Yuan, Y., Raubal, M., & Liu, Y. (2012). Correlating mobile phone usage and travel behavior—A case study of Harbin, China. *Computers, Environment and Urban Systems*, 36(2), 118-130.
- [31] Zarem, M., Vuillermet, E., & DeAguiar, J. (2017). U.S. Patent No. 9,612,122. Washington, DC: U.S. Patent and Trademark Office.