الگو رفتاری مسافران اتوبوس و مترو در اصفهان براساس اطلاعات ثبت شده کارت بلیط

عرفان محمدی*[۱] ، مریم جامی[۲] ، مریم مصلحی[۳] و مهدی شیروانی[۴]

[۱]شرکت آریاراد فناوران راه نوین، rfnmhmd@gmail.com

[۲] شرکت آریاراد فناوران راه نوین ، maryamj777@gmail.com

[۳] شرکت آریاراد فناوران راه نوین ، mmaria.moslehi@gmail.com

[۴] شرکت آریاراد فناوران راه نوین ، mahdishirvanea@gmail.com

*سرپرست تیم

چکیده - ماتریس مبدأ-مقصد یک ابزار بسیار مهم در شبکه حمل‌ونقل شهری است که الگوی جریان مسافران وسایل نقلیه مختلف را در شبکه نشان می‌دهد. به‌عبارت‌دیگر، این ماتریس دوبعدی، الگوی جابجایی افراد از مبدأ به مقصد را نمایش می‌دهد. با این مفهوم که چه تعداد افراد با چه نوع وسایل نقلیه از یک منطقه به منطقه دیگری سفر می‌کنند. برای به دست آوردن این ماتریس می‌توان از داده‌های کارت‌های هوشمند نقلیه استفاده کرد. در این چالش از داده‌های شماره خط، شماره کارت، تاریخ و زمان استفاده از کارت برای داده‌های مترو و اتوبوس، پلاک اتوبوس و ایستگاه‌های مترو استفاده‌شده است. هدف به دست آوردن ایستگاه‌های سوارشدن و پیاده شدن است. به این منظور، .ما در این پروژه با توجه به داده‌های موجود به‌وسیله‌ی روش‌های خوشه‌بندی ایستگاه‌های سوارشدن را به دست آوردیم و پس‌ازآن به‌وسیله این ایستگاه‌ها و همچنین با استفاده از ارتباط مکانی و زمانی بین ایستگاه‌ها ماتریس مزبور را به دست آوردیم.

کلیدواژه – خوشه‌بندی، ماتریس مبدأ-مقصد ، DBSCAN, KMEANS, HPC.

۱- مقدمه و بیان مسئله

ماتریس مبدأ-مقصد یا OD (Origin-Destination Matrix) به‌عنوان یک ابزار تحلیلی مهم در حوزه حمل‌ونقل و برنامه‌ریزی شهری، به مدیران و برنامه‌ریزان کمک می‌کند تا تصمیم‌های بهتر و اثربخش‌تری را برای بهبود و بهینه‌سازی سیستم حمل‌ونقل در یک شهر یا منطقه اتخاذ کنند. این ماتریس به ما امکان می‌دهد تا الگوی جریان مسافران و وسایل نقلیه بین مناطق مختلف یک شهر یا یک منطقه را درک کنیم.

اهمیت و کاربردهای اصلی این ماتریس عبارت‌اند از:

۱. برنامه‌ریزی حمل‌ونقل: ماتریس OD به مدیران حمل‌ونقل کمک می‌کند تا با درک الگوی جریان مسافران و وسایل نقلیه در سطح شهر یا منطقه، برنامه‌ریزی‌های مؤثرتری برای بهبود سیستم حمل‌ونقل و کاهش ترافیک ایجاد کنند.

۲. توسعه سیستم حمل‌ونقل عمومی: با تحلیل این ماتریس، نقاط ضعف و قوت در سیستم حمل‌ونقل عمومی شناسایی‌شده و بهینه‌سازی‌های لازم اعمال می‌شود.

۳. مدیریت ترافیک: داده‌های ماتریس مبدا-مقصد کمک می‌کند تا روند ترافیک در نقاط مختلف شهر یا منطقه برنامه‌ریزی و بهبود داده شود.

۴. برنامه‌ریزی شهری: اطلاعات به‌دست‌آمده از این ماتریس برای برنامه‌ریزی و توسعه شهری، ازجمله توسعه محله‌ها، ارائه خدمات عمومی و توسعه زیرساخت‌های شهری، بسیار مفید است.

به دست آوردن ماتریس OD در برخی موارد ممکن است به دلیل عدم دسترسی مستقیم به محل پیاده شدن مسافران یا نقص در داده‌های موردنیاز سختی‌هایی داشته باشد. این چالش‌ها ممکن است بر تحلیل دقیق و کامل الگوهای جریان مسافران و وسایل نقلیه در یک شبکه حمل‌ونقل تأثیر بگذارند. چند چالش مهم که ممکن است باعث سختی به دست آوردن ماتریس OD شوند عبارت‌اند از:

۱. نواقص در داده‌ها: دقت در داده‌های مربوط به حمل‌ونقل بسیار حائز اهمیت است. اگر داده‌هایی که جهت به دست آوردن ماتریس OD استفاده می‌شوند، ناقص یا نادقیق باشند، می‌تواند به تحلیل غیردقیق و اشتباه در تصمیم‌گیری‌ها منجر شود. به‌عنوان‌مثال باید به شیوه‌ای از ایستگاه‌های سوارشدن افراد مطمئن شد.

۲.عدم دسترسی به داده‌های دقیق پیاده شدن مسافران: در اکثر شبکه‌های حمل‌ونقل، داده‌های دقیق و جامع از محل پیاده شدن مسافران در دسترس نیست. این موضوع می‌تواند باعث ایجاد پوشش ناکافی برای تعیین ماتریس OD شود. برای حل این مشکل می‌توان از روش‌های تخمین ایستگاه پیاده شدن استفاده کرد یا اطلاعات واقعی ثبت‌شده توسط افراد را به‌عنوان نمونه بررسی نمود.

ما در این چالش سعی داریم به‌وسیله‌ی الگوریتم‌های خوشه‌بندی و مدل‌های تحلیل داده، اطلاعات ایستگاه‌های سوارشدن و پیاده شدن را به‌درستی به دست می‌آوریم تا با استفاده از آن ماتریس OD را به دست آوریم. بنابراین در

ادامه در این گزارش به تفضیل در مورد کارهایی که انجام‌گرفته است توضیحاتی را ارائه می‌دهیم .

این گزارش شامل قسمت‌های زیر است .

۱. ابتدا مروری بر کارهای گذشته داریم .

۲. پس‌ازآن مراحل انجام کار را با توجه به داده‌هایی که در اختیار قرار داده شد موردبررسی قرار می‌دهیم .

۳. در آخر نیز به بررسی نتایج به‌دست‌آمده می‌پردازیم. ذکر این نکته لازم است که ما برای نمایش داده و خروجی‌ها یک برنامه کاربردی تحت وب نیز طراحی کرده‌ایم

۲- مرور ادبیات

داده‌های کارت هوشمند به‌عنوان یک منبع داده جدید و قابل‌اعتماد برای جمع‌آوری و تجزیه‌وتحلیل اطلاعات سفر مسافران موردتوجه فزاینده‌ای قرارگرفته است[1]. عملکرد آماری داده‌های کارت هوشمند از داده‌های بررسی میدانی برای ارائه اطلاعات مکانی-زمانی جامع در مورد سیستم حمل‌ونقل و ثبت پویایی سفرهای مسافری بهتر است[2] .

جی و همکاران یک مدل بیزی برای تخمین ماتریس‌های جریان OD سطح سفر با استفاده از داده‌های جمع‌آوری‌شده توسط حسگرهای Wi-Fi و داده‌های سوارشدن توسط سیستم‌های مکان‌یابی خودکار خودرو (AVL) پیشنهاد می‌کند. بااین‌حال، در مقایسه با داده‌های کارت هوشمند، پایداری سنسورهای Wi-Fi و قابلیت اطمینان داده‌های جمع‌آوری‌شده نسبتاً پایین است. [3] . ما و همکاران تراکنش‌های کارت هوشمند و داده‌های GPS را بررسی کردند. نتایج دقت نسبتاً بالایی را در اطلاعات مبدأ مسافر نشان می‌دهد.[4]

برای شناسایی ایستگاه سوارشدن، دو نوع تکنیک عمدتاً اتخاذ می‌شود، تجزیه‌وتحلیل خوشه‌بندی و تطبیق داده‌های GPS. ژانگ و همکاران سفرهای رفت‌وآمد مسافران را با استفاده از داده‌های کارت هوشمند شناسایی می‌کند و از روش تحلیل خوشه‌بندی برای شناسایی ایستگاه‌های سوارشدن مسافران با توجه به زمان ثبت‌شده در داده‌های کارت هوشمند استفاده می‌کند[5] .

فرزین از تطبیق زمان GPS برای شناسایی مکان سوارشدن از طریق تطبیق همبستگی اطلاعات حرکت کارت آی سی اتوبوس، اطلاعات موقعیت‌یابی GPS و مکان‌های ایستگاه اتوبوس در سائوپائولو، برزیل استفاده می‌کند.[6]

برای به دست آوردن مقصد مسافرها نیز روش‌های متفاوتی وجود دارد.

در روش اول از یک تابع احتمال توزیعی فاصله سفر و میزان اهمیت آن ایستگاه، مکان پیاده شدن آن ایستگاه حدس زده می‌شود. برای به دست آوردن تخمین‌های قابل‌اتکاتر نیاز است تا اطلاعات کارت مسافرها را با اطلاعات دیگری مثل سامانه خودروباب ترکیب کنیم.

برای به دست آوردن احتمال پیاده شدن مسافران در ایستگاه‌های خاص می‌توان از مفهوم زنجیره سفر استفاده کرد. این عبارت برای توصیف سفر روزانه یک فرد ایجاد می‌شود که از چندین سفر اتوبوس تشکیل‌شده است با این فرض که ایستگاه سوارشدن مسافر در یک سفر نزدیک به ایستگاه پیاده شدن آخرین سفر است. میزان شناسایی ایستگاه پیاده شدن مسافر با استفاده از روش زنجیره سفر بین ۶۷ تا ۷۱ درصد است[7]

جعفری و همکاران برای به دست آوردن ماتریس مبدأ مقصد اتوبوس‌های تندرو تهران از داده‌های GPS یک سال اتوبوس‌ها برای تخمین زمان سفر بین ایستگاه‌ها استفاده شد و سپس با با داده‌های GIS خطوط و ایستگاه‌های BRT برای محاسبه فاصله پیاده‌روی بین ایستگاه‌های انتقال ترکیب گردید. با انجام این کار محدودیت‌های زمانی و مکانی‌ای در نظر گرفته شد که بررسی می‌کند آیا انتقال در یک دوره و فاصله منطقی انجام‌شده است یا خیر. در این پژوهش نیز ۵۲ درصد سفرها با استفاده از پیگیری زنجیره سفرها قابل‌بررسی بودند.[8]

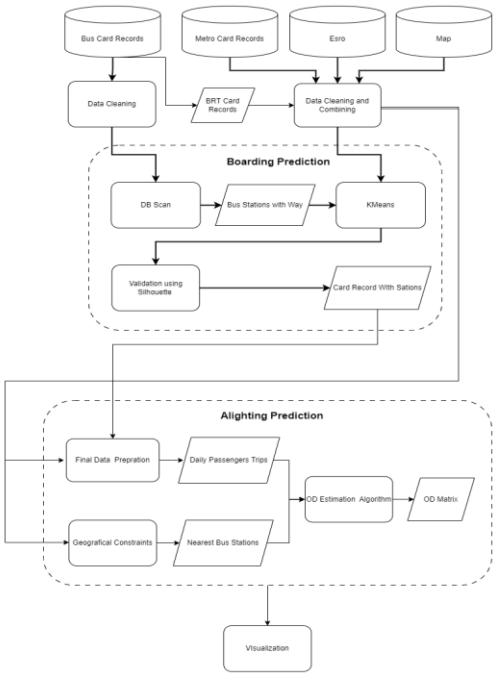
در پروژه حاضر با استفاده از اطلاعات کارت‌های اتوبوس و همچنین اطلاعات مسیر خط‌های اتوبوس‌رانی که از برنامه اصرو به‌دست‌آمده است و شامل مکان ایستگاه‌ها و همچنین ترتیب ایستگاه‌ها است استفاده‌شده و همچنین با استفاده از مفاهیم خوشه‌بندی و زنجیره سفر بدون نیاز به داده‌های GPS اطلاعات مبدأ مقصد مسافران اتوبوس را به دست آورده و ماتریس مبدأ مقصد را تشکیل داده‌ایم.

۳- مراحل انجام پروژه و حل چالش

در این بخش روش‌ها و مراحلی که برای حل چالش انجام‌گرفته است را توضیح داده می‌دهیم و در حین این توضیحات به بررسی داده نیز می‌پردازیم.

روند انجام این پژوهش در شکل ۳-۱نمایش داده‌شده است.

همان‌طور که در شکل ۳-۱دیده می‌شود مراحل انجام پروژه از چهار قسمت تشکیل‌شده است . ۱- آماده‌سازی داده‌ها ۲- به دست آوردن ایستگاه‌های سوارشدن ۳ – به دست آوردن ایستگاه‌های پیاده شدن و تشکیل ماتریس OD ۴- مصورسازی داده‌ها در بستر نرم‌افزار تحت وب . در این بخش ۳ مرحله اول توضیح داده‌شده است. مرحله آخر در بخش نتایج مفصل توضیح داده‌شده است.



شکل ۳-۱روند اجرای پروژه

۳-۱- آماده‌سازی داده‌ها

در مرحله اول داده‌ها مورد تجزیه‌وتحلیل اکتشافی(EDA) قرارگرفته و باید تمام داده‌ها را متناسب با اهداف تحلیل آماده کرد. این بخش خود شامل دو زیر بخش است. ۱-پاک‌سازی داده‌ها ۲- به دست آوردن داده‌های اطلاعات جغرافیایی ایستگاه‌ها

پاک‌سازی داده‌ها

داده‌های موجود شامل سه دسته مختلف هستند.

۱.داده‌های اتوبوس‌های شهری

۲.داده‌های خطوط تندرو یا BRT

۳.داده‌های خطوط مترو

شامل ستون‌های

- خطوط اتوبوس
- شماره پلاک اتوبوس
- تاریخ
- ساعت
- شماره کارت مسافران
- ایستگاه‌های مترو

در بخش پاک‌سازی داده‌های اتوبوس شهری موارد زیر انجام‌شده است

۱. پاک‌سازی داده‌های تکراری که در هر دو داده مترو و اتوبوس تعداد قابل‌توجهی بود. (حدوداً نیمی از داده‌ها در مترو)

۲. پاک‌سازی داده‌هایی که زمان آن‌ها در محدوده زمانی ساعت ۰۵-۲۳ قرار نداشت.

۳. پاک‌سازی شماره خط‌هایی که تعداد رکوردهای آن‌ها از ۱۰۰۰ کم‌تر بود. لازم به ذکر است که این موارد در اطلاعات خطوط اصرو نیز موجود نیستند که بتوان از اطلاعات جغرافیایی ایستگاه‌های آن‌ها استفاده نمود.

۴. پاک‌سازی شماره خط‌هایی که در هر دو نقشه راهنمای اتوبوس‌رانی سال ۱۳۹۸ و برنامه کاربردی اصرو نبود.

۵. در مرحله EDA (تجزیه‌وتحلیل اکتشافی داده) متوجه حجم زیاد رکوردهایی شدیم با شماره کارت یکسان که با بررسی بیشتر به این نتیجه رسیدیم که این رکوردها متعلق به رانندگان اتوبوس هستند و درنتیجه آن‌ها را از داده اصلی حذف کردیم. لازم به ذکر است که داده‌ها در تاریخ ۲۳-۲۸ ام شهریورماه (شنبه-پنجشنبه) و ۰۶-۱۱ مهرماه (شنبه-پنجشنبه) برای ۱۲ روز است که در خوش‌بینانه‌ترین حالت اگر مسافری ۷ بار در روز هم کارت زده باشد تعداد رکوردهای بیشتر از ۸۴ بار با شماره کارت یکسان به‌احتمال بسیار زیاد متعلق به رانندگان اتوبوس‌هاست. از طرفی بیشتر رانندگان در یک شماره خط یکسان کارت زده‌اند که به حدس ما صحت بیشتری می‌بخشد.

۶. پاک‌سازی ایستگاه‌های خط BRT که در فایل راهنما ایستگاه نظیری نداشت و تعداد رکوردهای آن‌ها بسیار کم بود که حدس زدیم نویز باشد.

۷. پاک‌سازی ایستگاه‌های تکراری شمالی جنوبی مترو و تجمیع آن در یک ایستگاه:

به دست آوردن داده‌های اطلاعات جغرافیایی ایستگاه‌ها

در مرحله اول اطلاعات جغرافیایی اولیه‌ای از سایت اصرو پایش شد که حاوی اطلاعات جغرافیایی ایستگاه‌های اتوبوس است. این داده‌ها ازآنجایی‌که ازنظر زمانی با نقشه قدیم اتوبوس‌رانی اختلاف دارند. بعد از ترسیم نقشه خطوط اصرو و مقایسه با نقشه اتوبوس‌رانی قدیمی به‌طور کامل تطبیق داده شدند. همچنین خط BRT با داده‌های اصرو و اطلاعات حقیقی به‌طور کامل تطبیق و اطلاعات مترو اضافه شد تا بتوان از اطلاعات جغرافیایی تمام ایستگاه‌ها استفاده نمود. برای اضافه کردن ایستگاه‌های مترو به این داده‌ها اطلاعات جغرافیایی آن‌ها از وب‌سایت نشان به‌دست‌آمده و به داده‌های جغرافیای دیگر اضافه شد.

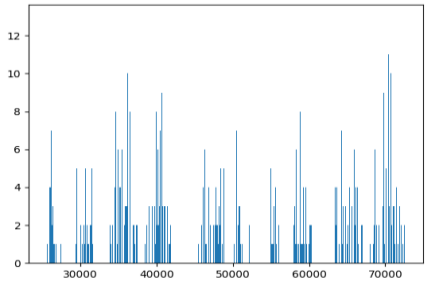
۳-۲- به دست آوردن ایستگاه‌های سوارشدن

یکی از اهداف اولیه برای حل این چالش به دست آوردن ایستگاهی است که هر مسافر در آن کارت زده است. برای رسیدن به این هدف ۳ مرحله به ترتیب انجام شد. ۱- بررسی تصویری داده‌ها برای انتخاب الگوریتم‌های خوشه‌بندی مناسب ۲-پیاده‌سازی الگوریتم خوشه‌بندیDBSCAN ۳- پیاده‌سازی الگوریتم خوشه‌بندی K-means

بررسی تصویری داده‌ها برای انتخاب الگوریتم‌های خوشه‌بندی مناسب

در این بخش داده‌ها با توجه به خط ، پلاک و روز گروه‌بندی شدند. تعبیر این گروه‌ها تمام کارت‌هایی است که در یک روز در یک اتوبوس با یک شماره پلاک بخصوص و متعلق به یک خط خاص جمع استفاده‌شده است.

برای مصورسازی هر یک از این گروه‌ها آن‌ها را بر روی نمودار هیستوگرام مصور کردیم. محور افقی این نمودار زمان زدن کارت و محور عمودی آن تعداد کارت‌هایی است که در هر زمان زده‌شده. در نمودار هیستوگرام می‌توان داده‌ها را دسته‌بندی کرد. در اینجا ما داده‌ها را در گروه‌های یک‌دقیقه‌ای تقسیم کردیم. به دلیل اینکه میدانیم اتوبوس در هر ایستگاه حدود یک دقیقه می‌ایستد. نمایش یک نمونه از این داده‌ها به‌صورت شکل ۳-۲شد. این داده‌ها مربوط به خط شماره یک در تاریخ ۲۴ شهریور و پلاک ۱۴۵A است. همان‌طور که در شکل مشخص است، داده‌ها در دسته‌های متمایز قرارگرفته‌اند. باکمی دقت می‌توان دریافت که هرکدام از این دسته‌ها مربوط به‌یک‌بار حرکت اتوبوس از ایستگاه اول به آخر (و یا از آخر به اول است).



شکل ۳-۲ نمودار هیستوگرام برای یک پلاک مشخص در یک روز مشخص

پیاده‌سازی الگوریتم خوشه‌بندی DBSCAN

با توجه به اینکه داده‌های موجود از نوع سری زمانی هستند برای خوشه‌بندی آن‌ها از الگوریتم DBSCAN استفاده کردیم. این الگوریتم خوشه‌بندی بر پایه چگالی داده‌ها عمل می‌کند و بدون نیاز به دانستن تعداد خوشه‌ها و فقط با توجه به مکان‌هایی که داده در آن تجمع کرده است سعی در شناسایی خوشه‌ها و جداسازی داده‌ها از یکدیگر دارد.

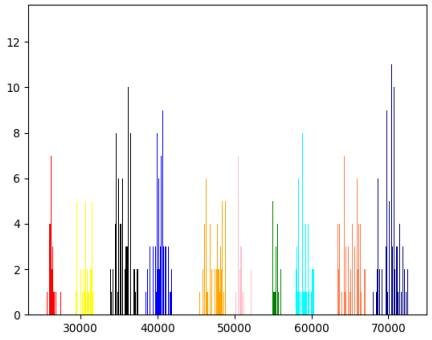
موارد زیر از مزیت‌های مهم این الگوریتم است.

تشخیص خوشه‌هایی با اندازه‌های متفاوت

مقاومت در برابر نویز

قابلیت کار با سری‌های زمانی

این الگوریتم دو پارامتر به‌عنوان ورودی می‌گیرد، eps که شعاعی است که برای همسایگی هر دونقطه در نظر گرفته می‌شود به این معنی که هر دونقطه‌ای که فاصله‌ای کمتر از esp دارند با یکدیگر همسایه‌اند و minpts که حداقل تعداد نقاطی است که باید در همسایگی یک نقطه قرار گیرد. یکی از چالش‌های استفاده از DBSCAN پیدا کردن این پارامترها است. ما با توجه به نوع داده متوجه شدیم بهترین eps (زمان بین دو ایستگاه) احتمالاً مقداری بین ۲۰۰ تا ۱۰۰۰ دارد و بهترین minpts (حداقل نفراتی که در یک ایستگاه سوار می‌شوند) مقداری بین ۴ تا ۱۰ دارد. با استفاده از grid search تعدادی از این گزینه‌ها را امتحان کردیم و بهترین آن را در نظر گرفتیم. معیار ما مقدار silhouette بود که برای سنجش میزان درستی خوشه‌بندی داده‌ها، میزان فواصل خوشه‌ها از یکدیگر و فواصل نقاط داخل هر خوشه را ارزیابی می‌کند. شکل زیر نتیجه بهینه الگوریتم DBSCAN است که هر رنگ نمایانگر یک خوشه است.



شکل ۳-۲ نمودار هیستوگرام بعد از اجرای DBSCAN

درنتیجه این خوشه‌بندی یک ستون way به داده اولیه اضافه شد. هرکدام از way ها نشان‌دهنده یک مسیر اتوبوس (رفت یا برگشت) در مسیر خط مرتبط است.

در این مرحله هدف پیدا کردن تعداد ایستگاه‌های هر خط اتوبوس و ایستگاه‌های سوارشدن مسافران است.

پیاده‌سازی الگوریتم خوشه‌بندی K-MEANS

با استفاده از نتایج به‌دست‌آمده در مرحله‌ی قبل و همچنین تعداد ایستگاه‌های هر مسیر می‌توانیم با استفاده از الگوریتم kmeans هر داده را ازنظر زمانی به یک ایستگاه مرتبط کنیم.

الگوریتم k-means نیز یک الگوریتم خوشه‌بندی است که تفاوت عمده آن با DBSCAN این است که از تعداد خوشه‌ها (تحت عنوان k) بهره می‌برد.

یک پایگاه داده معتبر برای به دست آوردن تعداد ایستگاه‌های خطوط اتوبوس‌رانی اصفهان اپلیکیشن اصرو است که داده‌های آن را با پایش وب استخراج کردیم.

در این مرحله ابتدا حدس زدیم به کمک شمردن ایستگاه‌های خطوط اتوبوس در نقشه ارائه‌شده و اعتبار سنجی آن با تعداد ایستگاه‌های به‌دست‌آمده از اپلیکیشن اصرو می‌توان k مناسب را به دست آورد. اما با مقایسه آن‌ها متوجه شدیم تعداد ایستگاه‌ها در این دو منبع اختلاف قابل‌توجهی دارد. به کمک grid search برای الگوریتم k-means در یک بازه پیرامون اعداد به‌دست‌آمده از اصرو و نقشه جست‌وجو انجام دادیم و با معیار silhouette ارزیابی کردیم. در اینجا به دلیل اعتبار اپلیکیشن اصرو برای silhouette های قابل‌قبول k را برابر با تعداد ایستگاه‌های اصرو قراردادیم. و با این مقدار k در الگوریتم k-means به‌صورت مجزا برای هر گروه (که هر گروه خط اتوبوس ، پلاک اتوبوس ، تاریخ و way مشابه دارند) فاصله‌ی عددی ایستگاه سوارشدن نسبت به ایستگاه اولیه‌ی اتوبوس را پیدا کردیم و برای خطوط دیگر k بهینه را در یک بازه اطراف تعداد ایستگاه‌های اصرو به دست آوردیم.

در ادامه برای اختصاص یک ایستگاه سوارشدن به هر مسافر کافی است بفهمیم هر پلاک اتوبوس از مسیر رفت یا برگشت روز خود را آغاز کرده است که به این منظور امتیاز silhouette را برای هر دو حالت محاسبه کردیم و بهترین حالت را انتخاب کردیم.

با توجه به تعداد زیاد داده و زمان اجرای بالا، کد دو الگوریتم خوشه‌بندی به‌صورت multiproccess توسعه یافت که سرعت اجرا را به‌صورت قابل‌توجهی بالا برد. هر proccess به یک گروه از داده‌ها که خوشه‌بندی بر روی آن‌ها انجام می‌گرفت اختصاص یافت.

۳-۳- به دست آوردن ایستگاه‌های پیاده شدن و تشکیل ماتریس OD

از چندین داده برای ساخت این ماتریس از تحلیل نهایی استفاده می‌شود. اولین داده فایلی است که با استفاده از نتایج به‌دست‌آمده از مراحل پیشین، داده‌های کارت‌های اتوبوس‌رانی را به تفکیک روز و در هرروز به تفکیک هر شناسه کارت اتوبوس درمی‌آورد. هدف این است که در این داده جدید تمام موارد استفاده‌شده از هر کارت در ایستگاه‌های متمایز بررسی شود. لازم به یادآوری است که نتایج مراحل پیشین مشخص‌کننده‌ی این ایستگاه‌ها برای هر استفاده از کارت اتوبوس هستند.

داده‌ی موردنیاز دیگر، ایستگاه‌های نزدیک به یکدیگر در خطوط اتوبوس متفاوت است. این داده با استفاده از اطلاعات جغرافیایی ایستگاه‌ها به‌دست‌آمده است و تمام ایستگاه‌های خطوط متفاوت که فاصله آن‌ها از یکدیگر کمتر از ۵۰۰ متر است را مشخص می‌کند.

در ادامه پروژه و با استفاده از این داده‌های تولیدشده، به دست آوردن ماتریس مبدأ مقصد میسر می‌شود. ماتریس‌ها اول برای هرروز به‌صورت جدا تهیه

می‌شوند تا اطلاعات هرروز را در نسخه وب بتوان جداگانه نشان داد. سپس تمام نتایج در یک ماتریس جامع تجمیع می‌شود.

برای تولید این ماتریس‌ها دو فرض در نظر گرفته می‌شود که به ترتیب در دو مرحله از آن‌ها برای به دست آوردن ماتریس استفاده می‌شود. اولین فرض این است که اگر کسی در یک روز از یک خط اتوبوس در دو مسیر متفاوت استفاده کند احتمالاً در مسیر برگشت به مبدأ اولیه است. این فرض به این معناست که مقصد سفر اول، مبدأ سفر بعدی و همچنین مقصد سفر آخر، مبدأ سفر اول بوده است. برای جلوگیری از محاسبه چندباره در این مرحله استفاده‌های تکراری در یک مسیر شناسایی و کنار گذاشته می‌شود.

فرض دوم و درنتیجه آن مرحله دوم به این صورت تعریف می‌شود که بسیاری از سفرهای درون‌شهری درواقع زنجیره‌ای از سفرهاست که مقصدهای میانی را می‌توان از فایل ایستگاه‌های نزدیک به هم و همچنین گذاشتن شرط زمانی به دست آورد. همچنین اگر مسافر در سفر آخر از خطی که اولین بار در اولین سفر استفاده‌شده بود استفاده کند، درواقع زنجیره سفر یک زنجیره بسته است و مقصد نهایی نیز همان مبدأ اولیه است که البته الگوریتم استفاده‌شده در مرحله اول چنین شرایطی را نیز لحاظ می‌کند.

بعد از به دست آوردن اطلاعات در این دو مرحله؛ اطلاعات با یکدیگر تجمیع می‌شود تا اطلاعات مبدأ مقصد برای هرروز به دست بیاید که در نسخه وب قابل‌مشاهده است.

در مرحله آخر از تجمیع اطلاعات تمام‌روزها، اطلاعات تجمعی مربوط به ایستگاه‌های پیاده شدن و همچنین ماتریس نهایی مبدأ مقصد به دست می‌آید.

۴- نتیجه‌گیری

خروجی سیستم به دو صورت موجود است. ۱- ماتریس مبدا-مقصد ۲- سامانه تحت وب .

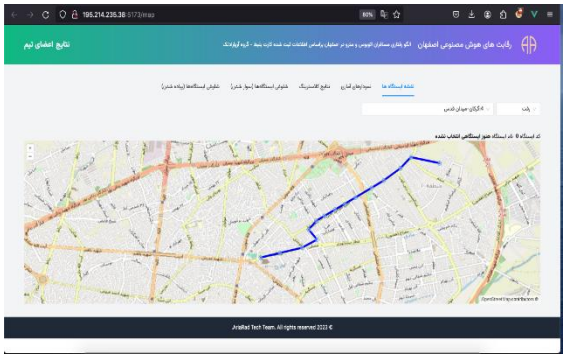
ماتریس مبدأ – مقصد

ماتریس مبدأ مقصد نهایی با فرمت CSV به‌عنوان خروجی اصلی به‌دست‌آمده است. این ماتریس یک ماتریس مربعی است که شامل ۱۴۳۱ در ۱۴۳۱ ایستگاه اتوبوس است که با استفاده از تجمیع تمام ایستگاه‌هایی که بین آن‌ها سوارشدن و پیاده شدنی اتفاق افتاده به‌دست‌آمده است. در این ماتریس سطرها نشان‌دهنده‌ی ایستگاه‌های سوارشدن و ستون‌ها نشان‌دهنده ایستگاه‌های پیاده شدن است که با شناسه ایستگاه مشخص‌شده است. همچنین یک پَروَنجا راهنما در کنار این ماتریس وجود دارد که برای هر شناسه ایستگاه نام ایستگاه و مکان آن را مشخص می‌کند.

سامانه تحت وب

برای این چالش یک سیستم نرم‌افزار تحت وب تهیه گردید. این نرم‌افزار از دو پروژه مجزا در سمت بک‌اند و فرانت‌اند با معماری سرویس گرا و با زبان‌های پایتون و جاوا اسکریپت توسعه‌یافته است. این نرم‌افزار از قسمت‌های زیر تشکیل گردیده است . لازم به ذکر است که این نمودارها و نقشه‌ها در سایت در محیط کاملاً تعاملی قرار دارد و با استفاده از گزینه‌های متفاوت می‌توان جزییات بیشتری را مشاهده نمود.

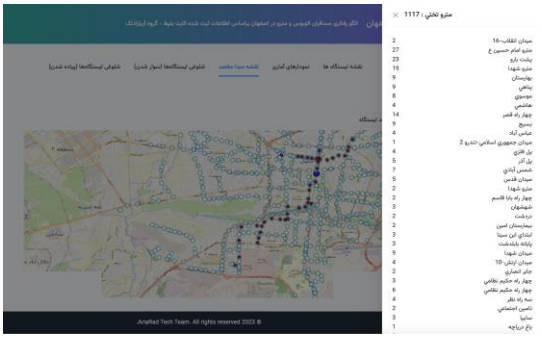
قسمت اول: ایستگاه‌های به‌دست‌آمده برای هر خط به‌وسیله‌ی نقشه اتوبوس‌های اصفهان و همین‌طور داده‌های به‌دست‌آمده از برنامه کاربردی اصرو را بر روی نقشه نشان می‌دهد. یک نمونه از این قسمت در شکل ۴-۱ که نمایش‌دهنده جهت رفت خط ۴ است. این قسمت قابلیت انتخاب خط و انتخاب جهت رفت ، برگشت و یا هردو جهت را دارد.



شکل ۴-۱نقشه خطوط در سامانه توسعه‌یافته

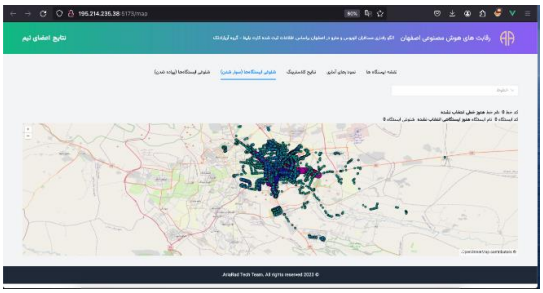
قسمت دوم : نمودارهای آماری به‌دست‌آمده از روی‌داده‌ها که شامل شلوغ‌ترین ساعات روز ، شلوغ‌ترین روزها ، شلوغی ایستگاه‌های مترو، BRT و اتوبوس است.

قسمت سوم: نقشه خروجی ماتریکس به‌دست‌آمده در این نقشه که بعد از انتخاب تاریخ و بعد از انتخاب ایستگاه سوارشدن، تمام ایستگاه‌هایی که با استفاده از این ایستگاه به آن‌ها رسیده است نمایش داده می‌شود. شکل ۴-۲ نمایش‌دهنده این قسمت است. همان‌طور که دیده می‌شود نمایش خروجی بر اساس روز تقسیم‌بندی شده است. با کلیک بر روی هر ایستگاه می‌توان اطلاعات جریان مسافرانی که این ایستگاه سوار شده‌اند را مشاهده نمود. به این معنی که تعداد تخمینی پیاده شدن در هر زنجیره سفر مشخص می‌گردد.



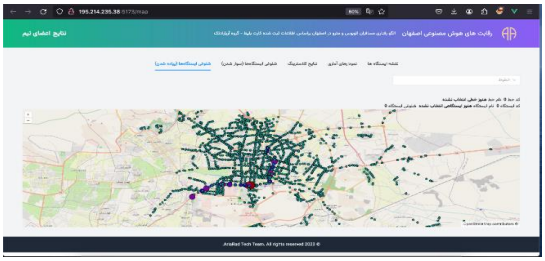
شکل ۴-۲نتایج ماتریس مبدأ مقصد

قسمت چهارم : نقشه شلوغی ایستگاه‌ها بر اساس تعداد افراد بر روی ایستگاه‌های موقع سوارشدن . شکل ۴-۳ نمایش‌دهنده شلوغی ایستگاه‌ها هنگام سوارشدن است .در این نقشه در ابتدا تمام ایستگاه‌های سوارشدن با توجه به میزان شلوغی آن‌ها که با بزرگ و کوچک بودن آن ایستگاه نمایش داده‌شده است قابل‌مشاهده است. بعد از انتخاب خط مربوطه می‌توان به شکل جزئی‌تری در مورد اطلاعات شلوغی هر ایستگاه را به دست آورد.



شکل ۴-۳ شلوغی ایستگاه‌ها برای سوارشدن

قسمت پنجم : نقشه شلوغی ایستگاه‌ها بر اساس تعداد افراد بر روی ایستگاه‌ها موقع پیاده شدن از اتوبوس است که در شکل ۴-۴ نمایش داده‌شده است و توضیحات آن مانند مورد قبلی است با این تفاوت که این اطلاعات از ماتریس مبدأ مقصد به‌دست‌آمده است و مقدار پیاده شدن در هر ایستگاه را نمایش می‌دهد. طبق نتایج بدست آمده ایستگاه میدان آزادی شلوغ ترین ایستگاه است.



شکل ۴-۴ شلوغی ایستگاه‌ها هنگام پیاده شدن.

منابع

[۱]

[1] M. Utsunomiya, J. Attanucci, and N. Wilson, “Potential Uses of Transit Smart Card Registration and Transaction Data to Improve Transit Planning,” *Transportation Research Record*, vol. 1971, no. 1, pp. 118–126, Jan. 2006, doi: 10.1177/0361198106197100114.

[2] L. Sun, A. Tirachini, K. W. Axhausen, A. Erath, and D.-H. Lee, “Models of bus boarding and alighting dynamics,” *Transportation Research Part A: Policy and Practice*, vol. 69, pp. 447–460, Nov. 2014, doi: 10.1016/j.tra.2014.09.007.

[3] Y. Ji, J. Zhao, Z. Zhang, and Y. Du, “Estimating Bus Loads and OD Flows Using Location-Stamped Farebox and Wi-Fi Signal Data,” *Journal of Advanced Transportation*, vol. 2017, pp. 1–10, 2017, doi: 10.1155/2017/6374858.

[4] X. Ma, Y. Wang, F. Chen, and J. Liu, “Transit smart card data mining for passenger origin information extraction,” *J. Zhejiang Univ. - Sci. C*, vol. 13, no. 10, pp. 750–760, Oct. 2012, doi: 10.1631/jzus.C12a0049.

[5] Z. Dafu, Z. Xinming, and W. Jian, “Commuter Travel Identification Based on Bus IC Data,” *Procedia - Social and Behavioral Sciences*, vol. 96, pp. 1547–1555, Nov. 2013, doi: 10.1016/j.sbspro.2013.08.176.

[6] J. M. Farzin, “Constructing an Automated Bus Origin–Destination Matrix Using Farecard and Global Positioning System Data in São Paulo, Brazil,” *Transportation Research Record*, vol. 2072, no. 1, pp. 30–37, Jan. 2008, doi: 10.3141/2072-04.

[7] J. Zhao, A. Rahbee, and N. H. M. Wilson, “Estimating a Rail Passenger Trip Origin-Destination Matrix Using Automatic Data Collection Systems,” *Computer aided Civil Eng*, vol. 22, no. 5, pp. 376–387, Jul. 2007, doi: 10.1111/j.1467-8667.2007.00494.x.

[8] M. Jafari Kang, S. Ataeian, and S. M. M. Amiripour, “A procedure for public transit OD matrix generation using smart card transaction data,” *Public Transp*, vol. 13, no. 1, pp. 81–100, Mar. 2021, doi: 10.1007/s12469-020-00257-7.

[۲]