Archive of SID

پیش‌بینی زمان ورود اتوبوس به ایستگاه با استفاده از داده‌های AVL: مطالعه موردی سیستم اتوبوس‌رانی شهر قزوین

حسین هاشمی، دانشجوی دکتری، دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه تربیت مدرس، تهران، ایران

امیر البدوی (مسئول مکاتبات)، استاد، دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه تربیت مدرس، تهران، ایران

E-mail: h.hashemi@modares.ac.ir

دریافت: ۱۳۹٤/۱۲/۱۸ پذیرش: ۱۳۹٦/۰۹/۲۵

چکیده

سیستم اتوبوس شهری در حال حاضر در دنیا به عنوان یکی از مهم‌ترین زیرساخت‌های سیستم حمل و نقل عمومی در جابجایی مردم شناخته می‌شود و به افزایش کیفیت خدمات این سیستم و برآورده کردن نیاز مردم منجر می‌شود تا نگرش مردم از حمل و نقل شخصی به حمل و نقل عمومی تغییر کند. در چند سال گذشته سامانه‌های موقعیت‌یاب خودکار خودرو جهت بهبود خدمات حمل و نقل عمومی در کشور راه‌اندازی شده است و داده‌های جمع‌آوری شده توسط این سامانه‌ها کمتر در جهت رضایت مشتریان (مسافران) استفاده شده‌اند. با توجه به اینکه اطلاع از زمان نسبتاً دقیق ورود اتوبوس به ایستگاه یکی از نیازهای مهم مسافران است در این مطالعه با استفاده از داده‌های موقعیت مکانی سیستم اتوبوس‌رانی شهر قزوین مبتنی بر زمان سفر هر اتوبوس و سرفاصله بین ایستگاه‌ها، بکارگیری داده های موقعیت‌یاب خودکار خودرو و بهره‌گیری از مدل‌های پایه در پیش‌بینی زمان ورود اتوبوس‌ها به ایستگاه نظیر شبکه عصبی مصنوعی و مدل‌های آماری نظیر مدل خودرگرسیو میانگین متحرک و ماشین‌های بردار پشتیبان، پیش‌بینی مدنظر، ارائه شده و مدل مطلوب توسعه یافته و بر اساس خروجی مدل‌های استفاده شده مدل شبکه عصبی عملکرد بهتری در پیش‌بینی زمان رسیدن اتوبوس‌ها به ایستگاه از خود نشان داده است. ازآنجایی‌که در ایران تاکنون مطالعه علمی در زمینه استفاده از داده‌های موقعیت‌یاب خودکار خودرو برای پیش‌بینی زمان ورود اتوبوس به ایستگاه انجام نشده است، این مقاله سعی کرده است با استفاده از تکنیک‌های مطرح و داده‌های دنیای واقعی ، نگاه علمی را در این حوزه کاربردی ارائه نماید

واژه‌های کلیدی: پیش‌بینی زمان ورود، اتوبوس‌رانی، داده‌های موقعیت مکانی، قابلیت اطمینان، شبکه عصبی مصنوعی

www.SID.ir

Archive of SID

۱. مقدمه

با توجه به افزایش روند شهرنشینی و افزایش تقاضای سفر در شهرها، توجه به سامانه‌های حمل‌ونقل عمومی افزایش پیدا کرده است. شرایط پیچیده ترافیکی، زمان‌بندی واقعی فعالیت اتوبوس‌ها را با شرایط غیرمطمئن مواجه ساخته است[Ghavami, Karimi and Mesgari, 2011]. بنابراین مسافران، نیازمند اطلاعات صحیح از زمان ورود هستند تا تصمیم خود را بگیرند و برنامه سفر خود را تنظیم نمایند. مضافاً، زمان‌سنجی دقیق یکی از عوامل بسیار مهم در ارزیابی کارایی عملیات شرکت‌های اتوبوس‌رانی و بخش مدیریت حمل‌ونقل عمومی است. محاسبه زمان ورود اتوبوس به ایستگاه یکی از مهم‌ترین مسائل سامانه‌های هوشمند حمل‌ونقل است [Zhou et al. 2016].

ایجاد سامانه‌های جدید در بسیاری از شهرها، امکان دسترسی به اطلاعات مکان و زمان اتوبوس‌ها را میسر کرده است. بنابراین سامانه‌ای قابل‌اطمینان که بتواند اطلاعات مناسب را به مسافران بدهد نقشی اساسی را در سیستم حمل و نقل شهری دارد [Lin at el. 2013].

ایده اصلی پیش‌بینی زمان سفر اتوبوس بر این حقیقت دلالت دارد که فرآیندهای رفتارهای ترافیک شهری هم‌زمان دارای هر دو خواص تا حدودی قطعی و تا حدودی نامنظم را دارند. نتایج این پیش‌بینی با استفاده از بازسازی بخش قطعی ترافیک و پیش‌بینی رفتارهای تصادفی که با عوامل پیش‌بینی نشده ایجاد می‌شوند، بدست می آید.

مقادیر داده‌های قبلی $f(t-1), f(t-2),\ldots, f(t-n)$ برای زمان های $t-1, t-2,\ldots, t-n$ داده شده است. مقادیر $f(t+1), \ldots,$ $f(t+2)$ با استفاده از تحلیل مجموع داده های قبلی قابل پیش‌بینی است. بنابراین مقادیر آینده از رابطه بین مجموعه متغیرهای وابسته به زمان و خروجی آن پیش‌بینی می‌شود [Zaki et al., 2013].

در حال حاضر بکارگیری داده‌های بسیار حجیم جمع‌آوری شده توسط سامانه‌های خودکار رهیاب خودرو در اکثر شهرهای ایران در جهت حل مسئله و نیازهای اساسی شهروندان بسیار ناچیز بوده و تحقیقات علمی در این حوزه می‌تواند منجر به افزایش رضایتمندی مردم شود. این مقاله با بکارگیری رویکرد توالی حرکت سه اتوبوس در دو ایستگاه نمونه در خط ۲ اتوبوس‌رانی قزوین به دنبال ارائه و توسعه مدل مناسب جهت پیش‌بینی زمان ورود اتوبوس به ایستگاه است. و این حوزه نوپا تاکنون در مقالات علمی داخلی مورد بررسی و تحقیق علمی قرار نگرفته است و این مقاله می تواندبه عنوان نقطه شروع در استفاده، تحلیل داده های بسیار حجیم موقعیت مکانی خودروها این خلا تحقیقاتی را در کشور برطرف نماید.

بنابراین دستاورد مهم این مقاله توجه به یکی از نیازهای اصلی شهروندان در دسترسی به اطلاعات مربوط به زمان رسیدن اتوبوس‌های شهری به ایستگاه مدنظر است و برای رسیدن به این هدف از مدل‌های آماری و داده‌کاوی موجود در ادبیات به منظور تحلیل داده‌های موردکاوی (که بیش از ۵ سال بدون تحلیل و دست‌نخورده در سیستم بایگانی شده بودند) استفاده شده است تا برای یک مشکل واقعی در حوزه اتوبوس‌رانی شهری که همان عدم داشتن اطلاعات از زمان رسیدن اتوبوس به ایستگاه است راه‌حلی علمی ارائه شود.

در بخش دوم ابتدا روش‌ها و مدل‌های پیش‌بینی زمان ورود ناوگان به ایستگاه در مقالات قبلی موردبررسی قرار می‌گیرد و در بخش سوم به تشریح نحوه جمع‌آوری داده و معرفی مورد کاوی، نحوه استفاده از داده‌ها در مدل و توسعه مدل مدنظر پرداخته می‌شود. در بخش چهارم نتیجه تحقیق و تحقیقات آتی بیان می‌شود.

۲. روش‌ها و مدلهای پیش‌بینی زمان ورود ناوگان

با توجه به بررسی مقالات مشابه قبلی و تحقیقات گذشته رویکردهای مختلفی برای پیش‌بینی زمان رسیدن اتوبوس به ایستگاه استفاده شده است که در ادامه به برخی از انواع دسته‌بندی رویکردها و تفاوت هر یک پرداخته می‌شود [Zaki et al. 2013].

رویکرد تاریخی: پیش‌بینی زمان سفر در یک زمان خاص به عنوان میانگین زمان سفر برای مدت مشابه در طول روزهای مختلف.

رویکرد لحظه‌ای: زمان سفر در فاصله زمانی بعدی را همانند فاصله زمانی حال حاضر پیش‌بینی می‌کند، این روش فرض می‌شود که روند نوسان زمان سفر اتوبوس بسیار محدود است که این فرض عملاً برای روند ترافیک واقعی، مانند حوادث، تراکم و دیگر شرایط ترافیک غیرقابل‌پیش‌بینی غیرممکن است.

www.SID.ir

Archive of SID

مدل‌های آماری: پیش‌بینی زمان ورود اتوبوس بر اساس یک تابع تشکیل شده توسط مجموعه‌ای از متغیرهای مستقل

رویکردهای بر اساس مدل: در ادبیات موضوع، برخی از محققان از الگوریتم فیلتر کالمن برای پیش‌بینی زمان ورود اتوبوس استفاده کرده‌اند که به نسبت از نظر دقت و صحت عملکرد قابل قبولی دارد. این مدل‌ها برای پیش‌بینی حداکثر یک یا دومرحله‌ای زمان(پیش‌بینی زمان لحظه‌ای و بسیار کوتاه) مناسب هستند [Jeong, 2005] و بر اساس مطالعات انجام شده مدل‌های شبکه عصبی مصنوعی کارایی بالاتری نسبت به فیلتر کالمن دارند [Lin et al. 2013] این الگوریتم با داشتن شرایط پویایی، خود را بر اساس داده‌های جدید به‌روزرسانی می‌کند. و این خاصیت در شرایط تغییر ویژگی‌های محیط حمل‌ونقل بسیار مهم است. به طوری که الگوریتم برای به‌روزرسانی متغیر حالت (سفر در زمان) مورد استفاده قرار گرفته و به طور مداوم مشاهدات جدید در دسترس قرار دارند.

روش‌های یادگیری ماشین: شبکه‌های عصبی مصنوعی (ANN) یکی از رایج‌ترین تکنیک‌های پیش‌بینی ترافیک است که عمدتاً به دلیل توانایی خود را برای حل روابط غیرخطی پیچیده استفاده می‌شود.

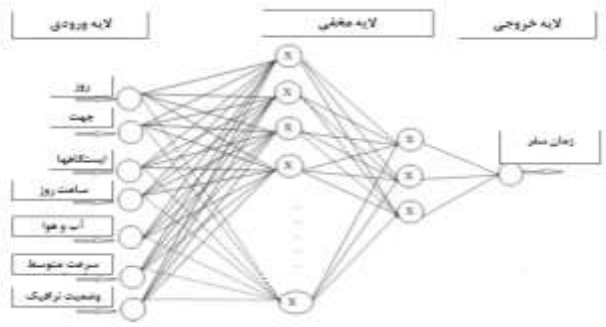
مظلومی و همکاران (۲۰۰۹) روش‌های ممکن برای پیش‌بینی زمان ورود ناوگان را به چهار دسته تقسیم می‌کند: (الف) مدل‌های رگرسیون (ب) مدل شبکه‌های عصبی مصنوعی (ANN) (ج) مدل‌های فیلتر کالمن و (د) روش‌های تحلیلی [Mazloumi et al. 2009].

البته تنوع دیگری از دسته‌بندی مقالات این حوزه وجود دارد به عنوان مثال لی، تیسنگ و تی‌سای (۲۰۰۹) متدولوژی‌های استفاده شده در تخمین زمان ورود اتوبوس‌ها در ٤ دسته‌بندی دیگر آورده است [Lee, Tseng and Tsai, 2009] : روش رگرسیون، روش سری زمانی ، روش ترکیبی داده‌ها یا مدل ترکیبی و روش‌های هوش مصنوعی مانند شبکه‌های عصبی.

۲-۱ ساختارکلی پیش‌بینی بر اساس شبکه‌های عصبی مصنوعی

شبکه‌های عصبی مدل‌های آماری در سامانه‌های واقعی جهان هستند که با استفاده از تنظیم مجموعه‌ای از داده‌ها ساخته می‌شود. این پارامترها به عنوان ورودی که با مقادیر مختلفی مرتبط هستند که همان خروجی هستند. فرآیند تنظیم اوزان به مقادیر صحیح- تنظیم- انجام می‌شود با استفاده از عبور از مجموعه‌ای از ورودی/ خروجی‌های مدل و تنظیم اوزان برای مینیمم کردن خطا بین جواب شبکه و نتیجه قابل‌تصور . هنگامی‌که اوزان تنظیم گردند، مدل قادر به تولید جواب برای داده‌های ورودی است که شامل داده‌های مورد امتحان نیست.

در شکل ۱ نمونه‌ای از شبکه عصبی شامل ٤ سطح : ورودی، دو سطح مخفی و خروجی است. ورودی دارای ۷ گره است و لایه مخفی اول دارای ۱۰ گره بوده و لایه دوم مخفی دارای ۳ گره است. درحالی‌که لایه خروجی دارای یک گره است.



شکل۱. ساختار کلی شبکه عصبی جهت مدل پیش‌بینی [Zaki et al. 2013]

۲-۲ مدل کلی سری زمانی مبتنی بر ARIMA

مدل خودرگرسیو میانگین متحرک ARIMA شامل سه بخش اتورگرسیو، میانگین متحرک و تفاضل است. به‌طورکلی این مدل را به صورت(p, d,q) ARIMA ارائه می‌کنند که (p) اولین رقم تعداد مراحل زمانی برای بهترین همبستگی به عقب بر می‌گردد. رقم دوم (d) بیانگر درجه تفاوت بوده و رقم سوم(q) بیانگر تعداد نقاط مورد استفاده برای محاسبه میانگین متحرک است. این محاسبات در رابطه (۱) ارائه می‌شوند:

$$y_t = \emptyset_1 y_{t-1} + \cdots + \emptyset_p y_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_p a_{t-p} \qquad (۱)$$

همبستگی بین نقاط داده‌ها در زمان t (y_t) و مراحل زمانی قبلی ($y_{t-1}, y_{t-2},\ldots,y_{t-n}$) یکی از مهم‌ترین ملاحظات این مدل است. به طور کلی هر مقدار y_t می‌تواند به عنوان ترکیبی خطی از مقادیر قبلی y_{t-1} و خطای تصادفی در زمان گذشته

www.SID.ir

Archive of SID

at-1 و خطای at از توزیع نرمال تبعیت می‌کند [Rashidi and Ranjitkar, 2013].

۲-۳ ساختار پیش‌بینی با استفاده از روش SVM

کاربرد مدل SVM در پیش‌بینی با استفاده از رویکرد طبقه‌بندی داده‌ها است. پیش‌بینی قیمت برق، مصرف انرژی، بازسازی سامانه‌های بی‌نظم و پیش‌بینی جریان ترافیک با استفاده از داده‌های ITS از کارکردهای این روش است [Wu, Ho and Lee, 2004].

ایده اصلی SVM نگاشت داده‌های ورودی به ابعاد بالاتر با استفاده از تابعی مانند Φ برای ایجاد ابر صفحه‌ای جدید است. امکان دسته‌بندی داده‌های در دو بعد امکان‌پذیری نبوده و این روش با استفاده از تابع $f(x) = (w.\Phi(x))$ می‌توان داده‌های ورودی را طبقه‌بندی کرد.

$$f(x) = \big(w.\Phi(x)\big) + b,\ \ w \subset R^n, b \subset R^n \qquad (۲)$$

هدف یافتن مقادیر w و b به نحوی است که ریسک رگرسیون کم شود:

$$R_{reg}(f) = C\sum_{i=0}^{l} \Gamma(f(x_i) - y_i) + \frac{1}{2}\|w\|^2 \qquad (۳)$$

۳. جمع‌آوری و تحلیل داده‌های تحقیق

۳-۱ تشریح سیستم مدیریت ناوگان شهر قزوین

سازمان اتوبوس‌رانی قزوین دارای سه بخش کلی سیستم اداری، توقفگاه و تعمیرگاه واقع و روزانه ۱۲۰ هزار نفر توسط ناوگان اتوبوس‌رانی در سطح شهر قزوین و حومه جابجا می‌شوند. همچنین در حال حاضر۲۳ خط اتوبوس تحت پوشش سازمان اتوبوس‌رانی قزوین مشغول به ارائه خدمات هستند که از این تعداد، ۱۷ خط درون‌شهری و ٦ خط نیز از مبدأ شهر قزوین به شهرهای مجاور رفت‌وآمد می‌کنند. سازمان اتوبوس‌رانی قزوین در بحث بهره‌وری و فنی، استفاده بهینه و کامل از اتوبوس‌ها و امکانات، برنامه‌ریزی و ارائه خدمات مطلوب به شهروندان در سال ۱۳۷۹ و سال ۱۳۸۷ مقام اول را در کشور توانسته کسب کند

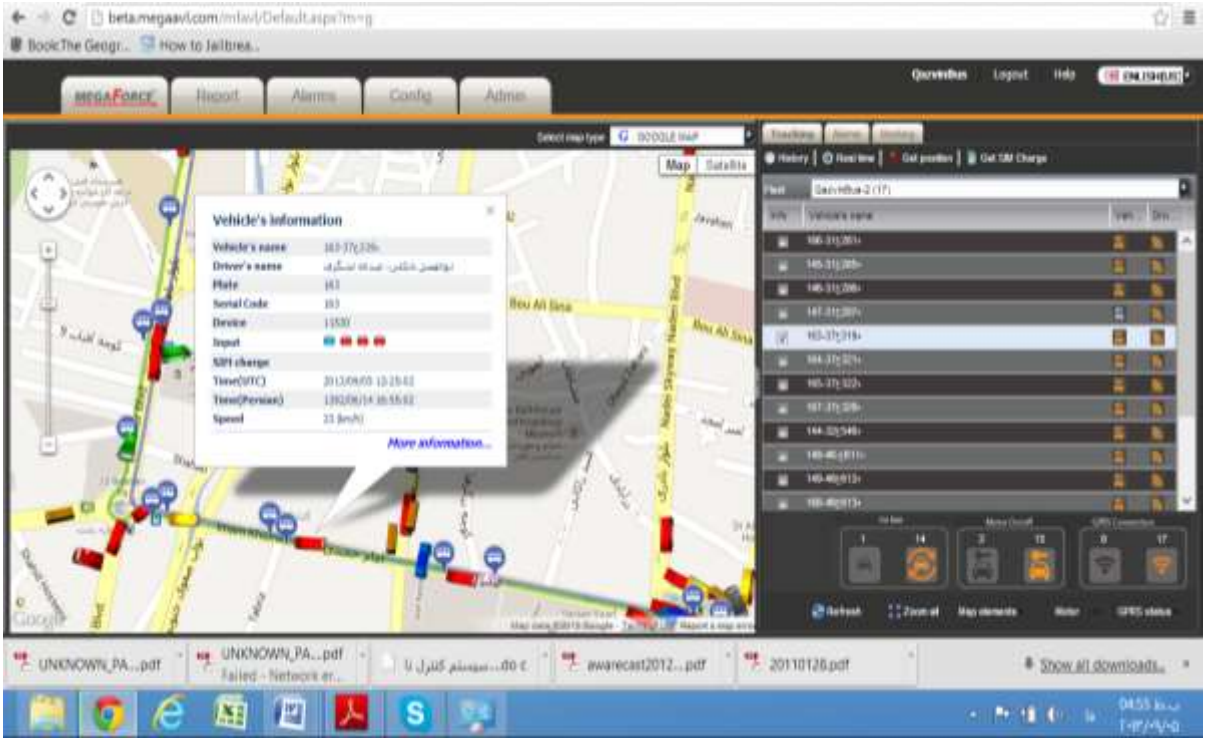
مدیریت حمل‌ونقل و ترافیک شهر قزوین یکی از سازمان‌های موفق در کشور است و در زمینه توسعه ITS در ناوگان حمل‌ونقل جزء اولین شهرها محسوب می‌شود. این سیستم توسط شرکت مگافورس آلمان در اتوبوس‌رانی و تاکسیرانی شهر قزوین پیاده‌سازی شده است. به‌طورکلی برای مدیریت بهینه ناوگان و مدیریت صحیح و نظارت بر عملکرد خودروها و همچنین ارائه خدمات بهتر به مشتریان این سیستم پیاده‌سازی شده است و تقریباً تمامی ۲۲۰ دستگاه اتوبوس و ۲۰۰۰دستگاه تاکسی شهری به سامانه AVL مجهز شده‌اند و موقعیت تمامی اتوبوس در طول شبانه‌روز به شکل برخط قابل‌دستیابی است. اتوبوس‌های شهری در ۲۳ مسیر مختلف طبقه‌بندی شده و هر مسیر شامل ایستگاه‌های مختلف در طول مسیر است. مهم‌ترین داده‌های مربوط به این سیستم شامل داده‌های موقعیت جغرافیایی شامل طول و عرض جغرافیایی سرعت (میزان سرعت، سکون/حرکت) ایستگاه‌های تعریف شده متناسب با GIS بوده و موقعیت هر اتوبوس بر اساس اعلام موقعیت خودکار دستگاه موقعیت‌یاب به شکل منظم با فاصله ۲۰ ثانیه یک‌بار به مرکز ارسال می‌گردد. با توجه به اطلاعات موجود در سیستم (لحظه‌ای و تاریخی) می‌توان تحلیل‌ها و گزارش‌های مختلف را متناسب با نیاز هر کدام از ذینفعان دریافت نموده و به شکل خروجی نرم‌افزار اکسل مشاهده کرد.

در شکل ۲ نمای کلی این سیستم و صفحه ابتدایی سیستم جهت مشاهده حرکت هر کدام از اتوبوس‌ها مشخص شده است. اجسام مشخص شده در روی نقشه اتوبوس‌های مسیر شماره ۲ به رنگ‌های واقعی خود می‌باشند که به عنوان مثال اتوبوس نشان داده در شکل ۲ به شماره ۳۱۹ع۳۷-۱۶۳ با سرعت ۲۳کیلومتر بر ساعت گزارش‌گیری۱٦:۵۵در حال حرکت است.

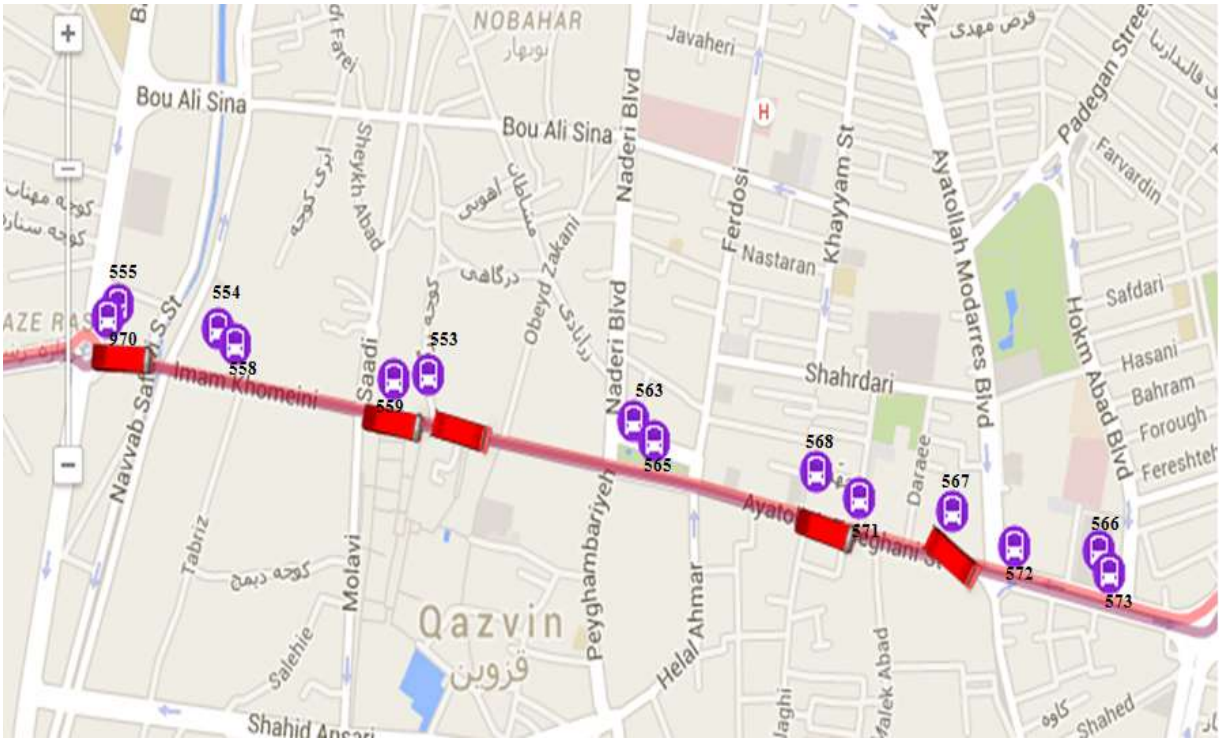
به منظور تحلیل سریع داده‌های در دسترس و ارزیابی نتایج بدست آمده در و تقلیل داده‌ها قسمت میانی خط ۲ قزوین که مهم‌ترین مسیر اتوبوس‌رانی این شهر است به تعداد ۲,۳۳ کیلومتر و به تعداد ١٤ ایستگاه رفت و برگشت در نظر گرفته شده است. مشخصات مسیر و تعداد ایستگاه‌ها در شکل ۳ اشاره شده است.

www.SID.ir

Archive of SID



شکل۲. نمای کلی سیستمAVL شهر قزوین



شکل ۳. قسمت میانی مسیر شماره ۲ اتوبوس‌رانی قزوین به طول ۲,۳۳ کیلومتر

www.SID.ir

Archive of SID

۳-۲ استخراج داده‌های موردنیاز برای تحلیل

برای رسیدن به نتایج این تحقیق یک ایستگاه به عنوان مبنای تحلیل داده و مدل‌سازی در نظر گرفته شده است. اطلاعات موردنیاز بر اساس زمان سفر هر اتوبوس از ایستگاه قبلی تا ایستگاه فعلی و همچنین سرفاصله‌های اتوبوس‌های پیشین در ایستگاه قبل استخراج شده است. برای پیش بینی زمان رسیدن هر اتوبوس k به ایستگاه j از ایستگاه قبلی i، اطلاعات جدول ۱ موردنیاز است.

جدول۱. متغیرهای موردنیاز برای پیش‌بینی زمان سفر

نام فیلد	توضیحات فیلد
VehicleName	کد kامین اتوبوس مورد بررسی
ObjectName	کد jامین ایستگاه اتوبوس مورد بررسی
EnterTime	زمان و تاریخ ورود به ایستگاه مورد بررسی
ExitTime	زمان و تاریخ خروج از ایستگاه مورد بررسی
t	ساعت ورود به ایستگاه
dw	روز هفته که اتوبوس k وارد ایستگاه j ام شده است
PreObj	کد ایستگاه قبل از ایستگاه j – یعنی i
PreObjEnterTime	زمان و تاریخ ورود به ایستگاه i ام
PreObjExitTime	زمان و تاریخ خروج از ایستگاه j ام
$tt_{i,j}(k)$	زمان سفر بین ایستگاه i و j برای اتوبوس k ام
$tt_{i,j}(k-1)$	زمان سفر بین ایستگاه i و j برای اتوبوس قبلی یعنی k-1 ام
$tt_{i,j}(k-2)$	زمان سفر بین ایستگاه i و j برای دو اتوبوس قبلی یعنی k-2 ام
$tt_{i,j}(k-3)$	زمان سفر بین ایستگاه i و j برای سه اتوبوس قبلی یعنی k-3 ام
$H_i(k)$	سرفاصله اتوبوس kام در ایستگاه iام
$H_i(k-1)$	سرفاصله اتوبوس k-1ام در ایستگاه iام
$H_i(k-2)$	سرفاصله اتوبوس k-2ام در ایستگاه iام
$H_i(k-3)$	سرفاصله اتوبوس k-3ام در ایستگاه iام

همانطور که در این جدول اشاره شده است، با توجه به ساختار ایجاد و توسعه مدل پیش‌بینی مبتنی بر در نظرگیری زمان سفر سه اتوبوس متوالی یا $tt_{i,j}(k)$ برای اتوبوس k-1,k-2,k-3 و سرفاصله در ایستگاه i ام $H_i(k)$ برای اتوبوس k-1,k-2,k-3 اثرات پارامترهای محیطی نظیر ترافیک و آب‌وهوا نیز در مدل دیده شده است. این بدان معنی است که ترافیک در یک محدوده زمانی مشخصی در طول روز اتفاق می‌افتد و اگر داده‌های مربوط به زمان ورود و خروج سه اتوبوس متوالی را در یک ایستگاه در نظر بگیریم می‌توان گفت در یک بازه زمانی مشخص اتوبوس‌های مدنظر از یک شرایط مشابه ترافیکی برخوردارند.

تمامی اطلاعات فوق برای مدت دو هفته از تاریخ ۱۱ خرداد تا ۲۵ خرداد ۱۳۹٤ جمع‌آوری شده است و پس از انجام مراحل آماده‌سازی و پاک‌سازی داده‌ها مورد استفاده تحلیل قرار گرفته است. جدول ۲ بخشی از داده‌های تحقیق را پس از انجام فرآیند آماده‌سازی نشان می‌دهد.

جدول۲. نمونه داده‌های مسئله مستخرج از سیستم AVL

کد وسیله نقلیه	زمان ورود	زمان خروج	$H_i(k)$	$H_i(k-1)$
١٤٦	٠٦/٠١/٢٠١٥ ٠٢:٤٩	٠٦/٠١/٢٠١٥ ٠٢:٥٠	٢٥:٣١	٠٦:٢٣
١٤٨	٠٦/٠١/٢٠١٥ ٠٣:٠٧	٠٦/٠١/٢٠١٥ ٠٣:٠٩	١٧:٣٦	٢٥:٣١
١٤٩	٠٦/٠١/٢٠١٥ ٠٣:١٧	٠٦/٠١/٢٠١٥ ٠٣:١٩	١٠:٤٧	١٧:٣٦
١٤٣	٠٦/٠١/٢٠١٥ ٠٣:٢١	٠٦/٠١/٢٠١٥ ٠٣:٢٤	٠٤:٤٣	١٠:٤٧
١٤٢	٠٦/٠١/٢٠١٥ ٠٣:٣٠	٠٦/٠١/٢٠١٥ ٠٣:٣٢	٠٩:١٤	٠٤:٤٣
١٦٦	٠٦/٠١/٢٠١٥ ٠٣:٣٩	٠٦/٠١/٢٠١٥ ٠٣:٤٢	٠٩:٠٠	٠٩:١٤
١٦٥	٠٦/٠١/٢٠١٥ ٠٣:٤٨	٠٦/٠١/٢٠١٥ ٠٣:٤٩	٠٧:٢٩	٠٩:٠٠
١٧٠	٠٦/٠١/٢٠١٥ ٠٣:٥٨	٠٦/٠١/٢٠١٥ ٠٤:٠١	١٠:٢٤	٠٧:٢٩

در این تحقیق، منظور از زمان سفر بین ایستگاه iو j برای اتوبوس k، مدت زمان خروج از ایستگاه قبلی (i) تا زمان رسیدن به ایستگاه فعلی (j) است که به دقیقه محاسبه شده است. همچنین سرفاصله به معنای مقدار زمانی است که اتوبوس k از اتوبوس k-1 در ایستگاه i جلوتر است. برای محاسبه زمان سفر و سرفاصله از روابط زیر استفاده شده است.

$$tt_{i,j}(k) = PreObjExitTime - EnterTime \quad (٤)$$

$$H_i(k) = PreObjEnterTime(k) - PreObjEnterTime(k-1) \quad (٥)$$

www.SID.ir

Archive of SID

در بررسی اولیه برای شناخت بیشتر داده ها و وضعیت حرکت ناوگان، گزارش ها و تحلیل های آماری گسترده‌ای صورت پذیرفته است. جدول ۳ اطلاعات آماری توصیفی مربوط به دو فیلد مهم tt(k) و H(k) را نشان می‌دهد.

جدول ۳. اطلاعات توصیفی مربوط به سرفاصله‌ها و زمان‌های سفر بین دو ایستگاه ۵۶۸ و ۵۶۳

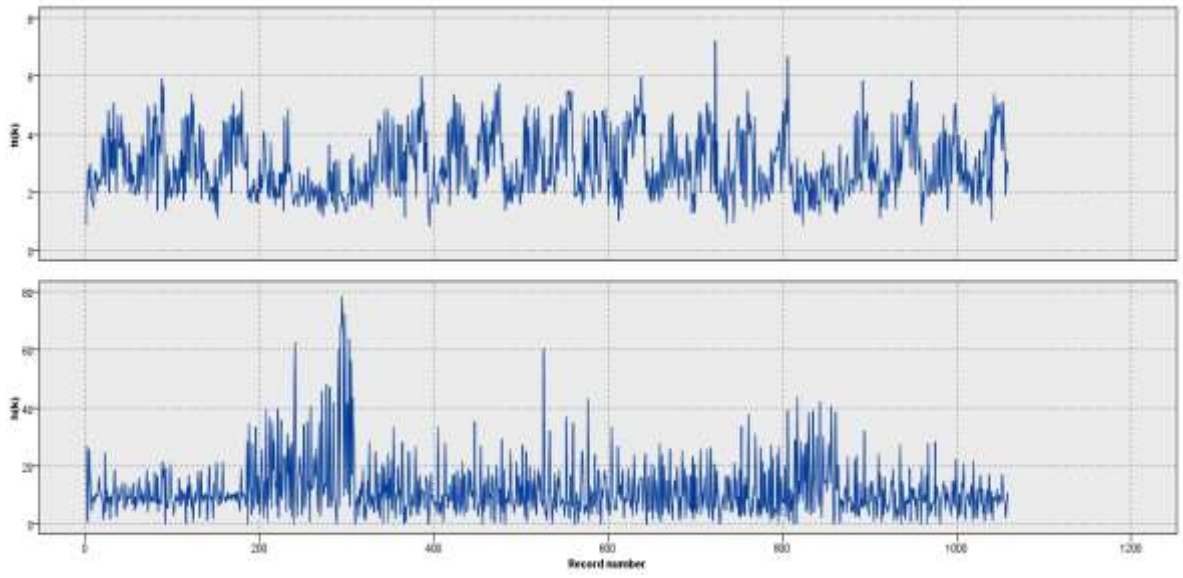
شاخص	اطلاعات توصیفی مربوط به سرفاصله‌ها	اطلاعات توصیفی مربوط به زمان‌های سفر
میانگین	۱۱,۸۰۶	۳,۰۰۹
حداقل	۰,۰۰	۰,۸۵۰
حداکثر	۷۸,۴۰	۸۷,۷۳
دامنه	۷۸,۴۰	۸۶,۸۸۳
واریانس	۹۶,۳۱۰	۷,۹۸۷
انحراف معیار	۹,۸۱۴	۲,۸۲۶
خطای استاندارد میانگین	۰,۳۰۲	۰,۸۰۷

همچنین نتایج تحلیل واریانس برای مقایسه میانگین سرفاصله‌ها و زمان های سفر در روزهای مختلف هفته صورت پذیرفته است (جدول ٤).

در نهایت شکل ٤ سری زمانی زمان سفر و سرفاصله‌ها را بین دو ایستگاه ۵۶۸ و ۵۶۳ خط دو اتوبوس‌رانی قزوین نشان می‌دهد.

جدول٤. مقایسه و تحلیل واریانس زمان سفر و سرفاصله به ازای روزهای مختلف هفته (خروجی نرم‌افزار SPSS)

فیلد	شنبه	یکشنبه	دوشنبه	سه شنبه	چهارشنبه	پنجشنبه	جمعه	ازمون F	اهمیت
tt(k)	۳,۴۹۹ ۶,۳۱۹ ۰,۴۶۵ ۱۸۵	۳,۰۷۲ ۱,۱۰۹ ۰,۰۸۱ ۱۸۸	۳,۱۲۵ ۱,۱۱۱ ۰,۰۸۵ ۱۷۱	۲,۶۰۳ ۱,۰۴۸ ۰,۰۹۰ ۱۳۵	۲,۷۵۳ ۰,۹۹۰ ۰,۰۸۵ ۱۳۶	۳,۱۷۰ ۱,۲۲۰ ۰,۰۸۴ ۱۷۷	۲,۰۹۳ ۰,۶۰۰ ۰,۰۷۳ ۶۷	۲,۹۴۳	۰,۹۹۳
h(k)	۹,۸۸۲ ۷,۱۲۳ ۰,۵۲۴ ۱۸۵	۹,۷۶۴ ۶,۳۹۷ ۰,۴۶۷ ۱۸۸	۱۰,۶۶۶ ۷,۳۸۴ ۰,۵۶۶ ۱۷۰	۱۳,۲۷۳ ۱۱,۷۲۶ ۱,۰۰۹ ۱۳۵	۱۲,۹۶۷ ۹,۴۶۱ ۰,۸۱۱ ۱۳۶	۱۰,۳۶۲ ۶,۲۶۹ ۰,۴۷۱ ۱۷۷	۲۴,۲۴۳ ۱۹,۵۷۰ ۲,۳۹۱ ۶۷	۲۵,۴۰۸	۱,۰۰۰



شکل٤. سری زمانی زمان سفر بین ایستگاه ۵۶۸ و ۵۶۳ (شکل بالایی) و سرفاصله (شکل پایینی) برای ایستگاه ۵۶۸

www.SID.ir

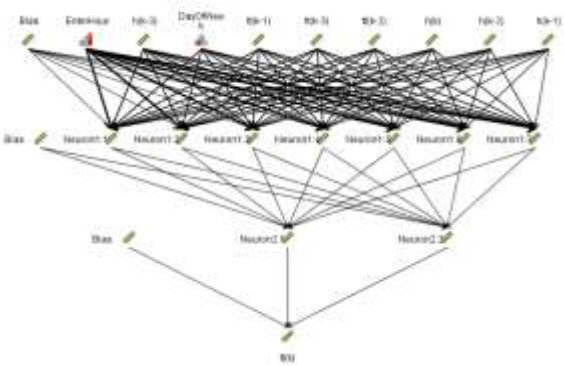
Archive of SID

٤. توسعه مدل مبتنی بر مدل‌های پیش‌بینی و داده‌های تحقیق

٤-١ پیش‌بینی مبتنی بر مدل ANN

یکی از ویژگی‌های مدل‌های ANN پیش‌بینی روابط غیرخطی بوده و در این مدل‌ها امکان استفاده از متغیرهای همبسته وجود دارد. بنابراین احتمال رسیدن به پیش‌بینی مطلوب متغیرها در صورت تنظیم درست پارامترها به حداکثر خود می‌رسد. از این‌رو در این تحقیق از مدل ANN با استفاده از بکارگیری و تحلیل داده‌های دو هفته اتوبوس‌های مسیر ۲ استفاده می‌شود.

در تحلیل ANN پیش رو، متغیر هدف tt(k) خواهد بود که بر اساس متغیرهای مستقل معرفی شده در جدول (۱) مورد پیش بینی قرار خواهد گرفت. مدل شبکه عصبی با الگوریتم انتشار رو به عقب آموزش داده شده است و وزن‌های آن تخمین زده شده است. بر اساس مدل های مختلف طراحی شده، مدل شبکه عصبی با دو لایه که لایه اول ۷ نرون و لایه دوم ۲ نرون دارا می باشد طراحی شده است. شکل ۵ ساختار شبکه عصبی طراحی شده و متغیرهای ورودی به تحقیق را نشان می‌دهد.



شکل۵. ساختار شبکه عصبی طراحی شده برای پیش‌بینی

مدل پایه برای پیش‌بینی زمان سفر با شبکه عصبی بر اساس روابط زیر تعیین شده است.

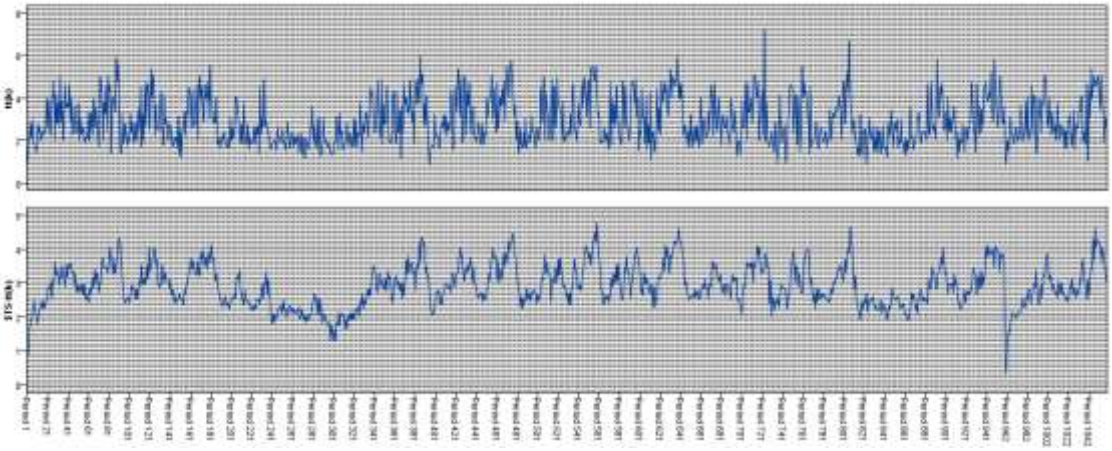
$$\widetilde{tt}_{i,j}(k) \cong f[t, dw, tt_{i,j}(k-b), H_i(k-b)] \quad (٦)$$

منظور از پارامتر b تعداد اتوبوس‌های پیشین است که در پیش‌بینی مورد استفاده قرار می‌گیرد و در این تحقیق b=3 است. بنابراین برای پیش‌بینی زمان رسیدن اتوبوس kام به ایستگاه jام، خواهیم داشت:

$$EnterTime_j(k) = PreObjEnterTime_i(k) + \widetilde{tt}_{i,j}(k) \quad (٧)$$

٤-٢ پیش‌بینی مبتنی بر ARIMA

با توجه به مدل‌های مختلف سری زمانی طراحی شده، مدل ARIMA(0,1,10) به عنوان مدل نهایی درنظر گرفته شده است. شایان ذکر است که انتخاب این پارامترها برای مدل ARIMA بر اساس منطق سعی و خطا صورت پذیرفته است و ممکن است برای داده‌های دیگری، پارامترهای دیگری به عنوان پارامتر بهینه درنظر گرفته شوند. شکل ٦ نتایج مقدار واقعی زمان سفر و مقدار پیش بینی شده توسط مدل ARIMA(0,1,10) را نشان می‌دهد.



شکل٦. مقدار واقعی (شکل بالایی) و مقدار پیش‌بینی شده (شکل پایینی) زمان سفر بر اساس مدلARIMA

www.SID.ir

Archive of SID

٤-٣ پیش‌بینی مبتنی بر SVM

مدل پیش‌بینی ماشین‌های بردار پشتیبان (SVM) نیز در این مقاله موردبررسی قرار گرفته است و داده‌های معرفی شده بر اساس این مدل مورد پیش‌بینی قرار گرفته است. تابع کرنل استفاده شده برای پیش‌بینی SVM تابع RBF است که با پارامترهای C=10 و اپسیلون ۰,۱ طراحی شده است. قابل ذکر است که انتخاب این پارامترها برای مدل SVM بر اساس منطق سعی و خطا صورت پذیرفته است و ممکن است برای داده های دیگری، پارامترهای دیگری به عنوان پارامتر بهینه درنظر گرفته شوند.

٤-٤ ارزیابی مدل‌ها

برای تحلیل نتایج و عملکرد مدل‌های پیش‌بینی، از منطق بسط دادن برای نگهداری ۳۰٪ داده ها به جهت تست کردن نتایج استفاده شده است. در نتیجه به عنوان مثال در مسیر بین دو ایستگاه ٥٦٨ و ٥٦٣ از بین ۱۰۵۸ سفر موجود در پایگاه داده، به تعداد ۷۲۷ سفر برای ایجاد مدل پیش بینی و ۳۳۱ رکورد باقیمانده در جهت ارزیابی نتایج استفاده شده است.

از جمله شاخص هایی که برای مقایسه نتایج مورد استفاده قرار خواهد گرفت، شاخص میانگین خطای پیش‌بینی، میانگین مطلق خطای پیش بینی، انحراف معیار خطای پیش بینی و وابستگی خطی بین مقدار واقعی و مقدار پیش‌بینی شده زمان سفر خواهد بود. همچنین شاخص متوسط خطای نسبی بر اساس رابطه زیر برای مدل ها محاسبه خواهد شد. این رابطه میزان خطا را با توجه به نسبت متوسط زمان سفر بین دو ایستگاه می‌سنجد.

$$RAE = \frac{Mean\ absolute\ error}{Average\ travel\ time} * 100\% \qquad (۸)$$

جدول ۵ مقایسه سه مدل پیشنهادی را بر اساس شاخص های کیفی عملکردی نشان می‌دهد.

جدول۵. مقایسه عملکرد سه مدل پیشنهادی بر اساس شاخص‌های کیفی متنوع

شاخص	میانگین خطا (دقیقه)		انحراف معیار خطا		میانگین مطلق خطا(دقیقه)		متوسط خطای نسبی	
مدل	آموزشی	تست	آموزشی	تست	آموزشی	تست	آموزشی	تست
ANN	۰,۰۰۶	۰,۰۳۰	۰,۸۰۵	۰,۹۶۹	۰,۶۲۴	۰,۷۴۵	۲۰,۸۰	۲۴,۸۰
ARIMA	۰,۰۰۵	۰,۰۲۷	۰,۹۱۶	۰,۹۴۸	۰,۷۳۳	۰,۷۵۷	۲۴,۴۰	۲۵,۲۰
SVM	۰,۰۶۰	۰,۰۶۳	۰,۷۷۲	۰,۹۹۲	۰,۷۱۲	۰,۷۶۲	۲۲,۰۰	۲۵,۴۰

بر اساس نتایج فوق، عملکرد شبکه عصبی از بقیه مدل‌ها بهتر است. این در حالی است که مدل شبکه عصبی به تعداد زیاد داده بسیار حساس است و در پروژه پیش‌رو اطلاعات مربوط به تنها دو هفته از عملکرد سیستم اتوبوس‌رانی مورد تحلیل قرار گرفته است. میانگین خطای مطلق برای شبکه عصبی برای داده‌های تست کمتر از سایر روش‌های دیگر بوده و در حدود ٤٤ ثانیه است که با توجه به شرایط داده‌های مدنظر مقدار قابل قبولی به حساب می‌آید. برای تضمین کاهش خطای پیش‌بینی به جز استفاده از یک مدل کارا عوامل دیگری نظیر کیفیت نمونه داده‌های با تناوب بیشتر است. به‌نحوی‌که به دلیل پاک‌سازی بخشی از داده‌های جمع‌آوری شده (به دلیل تنظیمات قبلی سیستم موقعیت یاب) ضمن اینکه ساختار و الگوی استفاده شده در این تحقیق اجازه دخیل نمودن همه عوامل مؤثر در زمان سفر(نظیر رفتار مسافران، تعداد ایستگاه‌ها، چراغ‌های راهنمایی و رانندگی، طول مسیر، رفتار رانندگان، عوامل مدیریتی و...) را نمی‌دهد و در واقع پیش‌بینی کوتاه‌مدت مبتنی بر زمان ورود و خروج سه اتوبوس در دو ایستگاه متوالی برای حجم داده‌های محدود همراه با حجم زیاد داده‌های گمشده انجام شده و این نتیجه منجر به کاهش دقت پیش‌بینی گردیده است. وجود حجم زیادی از داده‌های گمشده و تأثیر عوامل محیطی در این تحقیق و پیش‌بینی دقیق‌تر زمان ورود ناوگان به عنوان موضوع تحقیق آتی بوده و یکی از خروجی‌های این تحقیق در دنیای واقعی، تعریف پروژه بهبود برای سیستم ثبت داده‌های AVL مرکز اتوبوس‌رانی است. در تحقیقات گذشته با

www.SID.ir

Archive of SID

www.SID.ir

Archive of SID

استفاده از داده‌های لحظه‌ای موقعیت اتوبوس خطای پیش‌بینی کاهش داده شده است.

۵. نتیجه‌گیری

پیش‌بینی زمان ورود ناوگان یکی از اجزای مهم سیستم اطلاع‌رسانی به مسافران در ایستگاه بوده و ارائه اطلاعات دقیق می‌تواند منجر به کاهش زمان انتظار و انتخاب مناسب‌تر مد حمل‌ونقل برای مسافران را به دنبال داشته باشد. از انجاییکه در ایران داده‌های بدست آمده از سامانه‌های AVL در شبکه اتوبوس‌رانی مورد استفاده علمی و کاربردی قرار نگرفته است، این تحقیق هدف استفاده از داده‌های دنیای واقعی و ارائه مدل متناسب با داده‌ها برای پیش‌بینی زمان رسیدن ناوگان به هر ایستگاه اتوبوس‌رانی شهری پرداخته است.

پیش‌بینی صورت گرفته در این تحقیق مبتنی بر تحلیل هر ایستگاه و ایستگاه ماقبل آن و سرفاصله سه اتوبوس متوالی است بر این اساس متغیرهای مناسب برای مدل‌سازی بر اساس عملکرد اتوبوس‌های قبلی که در هر مسیر حرکت کرده‌اند تعریف شده است و سه مدل پیش‌بینی بر اساس مدل‌های شبکه عصبی، سری زمانی و ماشین بردار پشتیبان توسعه داده شده است. از مقایسه عملکرد سه مدل برای داده‌های موردکاوی سیستم اتوبوس رانی شهر قزوین، عملکرد مدل شبکه عصبی عملکرد بهتری نسبت به سایر مدل‌ها ارائه کرده است و متوسط خطای مطلق ٤٤ ثانیه برای پیش بینی زمان رسیدن اتوبوس به ایستگاه‌ها استخراج شده است.

پیشنهاد می‌شود تا برای تحقیقات آتی ضمن افزایش حجم داده‌ها برای زمان بیشتر، ترکیبی از ایستگاه‌های مختلف در یک مسیر مشخص در نظر گرفته شود چراکه در این تحقیق عملکرد هر ایستگاه خطوط اتوبوس‌رانی به صورت مستقل از یکدیگر درنظر گرفته شده است. همچنین درنظر گرفتن پارامترهای محیطی نظیر عملکرد راننده، سرعت، ترافیک، آب‌وهوا و ازدحام جمعیت مسافران به عنوان عوامل اثرگذار بر زمان رسیدن به اتوبوس می‌تواند مطرح شود.

برای ارائه تحقیقات علمی در این حوزه به دلیل ساختار جمع‌آوری این قبیل داده هنوز کاستی‌های زیادی وجود دارد. به عنوان نمونه بخش زیادی از داده‌های سیستم AVL به صورت از دست رفته در اختیار نویسندگان قرار گرفته است که بررسی علل از دست رفتن جهت بازیابی مجدد این داده‌ها و تعریف ساختار مناسب برای جمع‌آوری داده‌ها با دقت بالاتر خود نیازمند تحقیق دیگری است. در نتیجه ممکن است در برخی مدل‌ها دقت پیش‌بینی کمتر از استاندارد بین‌المللی باشد. در واقع نگاه محققان در این مقاله، ایجاد پایه اولیه برای تحلیل و بررسی عمیق‌تر حجم زیاد داده‌هایی است که در سیستم اتوبوس‌رانی کشور ثبت می‌گردد.

در تحقیق با هدف بررسی دلایل تأخیر ورود اتوبوس‌ها به ایستگاه به ارائه تحلیل علمی بر روی شاخص قابلیت اطمینان پرداخته می‌شود که این موضوع می‌تواند به بهبود عملکرد به موقع سیستم اتوبوس شهری کمک شایانی نماید و بر بهبود عملکرد پیش‌بینی تأثیر مثبت داشته باشد.

همچنین در تحقیق آتی از رویکردهای دیگر پیش‌بینی نظیر فیلتر مالمان و یا مدل‌های ترکیبی جهت ارائه مدل پیش‌بینی بهینه در شرایط مطلوب‌تر عملکردی قابلیت اطمینان و داده‌های جمع‌آوری شده در شرایط مؤثرتر استفاده خواهد شد.

برای توسعه تحقیقات آتی در این زمینه پیشنهاد می‌شود نحوه و شرایط پیش‌بینی پذیری مطلوب‌تر مدل با تغییر ساختار جمع‌آوری داده‌های موقعیت اتوبوس و یا بهبود قابلیت اطمینان سیستم موردبررسی و تحلیل علمی قرار گیرد.

٦. مراجع

-قوامی، سید مرسل؛ کریمی، علی و مسگری، محمد سعدی، "ارزیابی خطوط اتوبوس رانی با استفاده از سامانه اطلاعات مکانی و تحلیل پوششی داده ها، مطالعه موردی: خطوط اتوبوس رانی تهران"، فصلنامه مهندسی حمل و نقل، دوره ۲، شماره ۳، بهار ۱۳۹۰، صفحه ۲٦۱-۲۷۱.

-Jeong, R. H. (2005) "The prediction of bus arrival time using automatic vehicle location systems data" Doctoral dissertation, Texas A&M University.

-Lee, W. H., Tseng, S. S. and Tsai, S. H. (2009) "A knowledge based real-time travel time prediction system for urban network", Expert Systems with Applications, Vol. 36, No. 3, pp.4239-4247.

-Lin, Y., Yang, X., Zou, N. and Jia, L. (2013) "Real-time bus arrival time prediction: Case study for Jinan, China", Journal of

www.SID.ir

Archive of SID

Transportation Engineering, Vol. 139, No. 11, pp.1133-1140.

-Mazloumi, E., Currie, G., Rose, G. and Sarvi, M. (2009) " Using SCATS data to predict bus travel time", in 32nd Australian Transport Research Forum (ATRF), Auckland, New Zealand, 2009

-Rashidi, S. and Ranjitkar, P. (2013) "Approximation and short-term prediction of bus dwell time using AVL data", Journal of the Eastern Asia Society for Transportation Studies, Vol.10, pp.1281-1291.

-Wu, C. H., Ho, J. M. and Lee, D.T. (2004) "Travel-time prediction with support vector regression. IEEE Transactions On Intelligent Transportation Systems, Vol. 5, No. 4, pp.276-281.

-Zaki, M., Ashour, I., Zorkany, M. and Hesham, B. (2013) "Online bus arrival time prediction using hybrid neural network and Kalman filter techniques", International Journal of Modern Engineering Research, Vol. 3, No. 4, pp.2035-2041.

-Zhou, Y., Yao, L., Chen, Y., Gong, Y. and Lai, J. (2017) "Bus arrival time calculation model based on smart card data", Transportation Research Part C: Emerging Technologies, Vol. 74, pp.81-96.

www.SID.ir

Archive of SID

www.SID.ir

Archive of SID

حسین هاشمی، امیر البدوی

حسین هاشمی فارغ التحصیل دکتری مهندسی صنایع دانشکده مهندسی صنایع و سیستم‌ها از دانشگاه تربیت مدرس تهران است. ایشان مدارج کارشناسی و کارشناسی ارشد خود را در رشته ریاضی کاربردی و مهندسی صنایع از دانشگاه علم و صنعت و دانشگاه امام حسین(ع) به ترتیب در سال های ۱۳۸۲ و ۱۳۸۵ اخذ نموده است. زمینه های پژوهشی مورد علاقه وی شامل مسائل هوشمند سازی حمل و نقل، اقتصاد حمل و نقل، برنامه و ریزی حمل و نقل، پیش بینی زمان سفر و راهکارهای بهبود قابلیت اطمینان در سیستم های حمل و نقل است.

دکتر امیر البدوی، استاد دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه تربیت مدرس، است که درجه دکتری خود را از دانشگاه سیستم های اطلاعاتی از دانشگاه لندن (۱۹۹۷) دریافت کرده است. ایشان مدرک کارشناسی و کارشناسی ارشد خود در رشته مهندسی صنایع و سیستم را از دانشگاه صنعتی شریف اخذ نموده و زمینه های پژوهشی مورد علاقه وی شامل سیستم های اطلاعات مدیریت (MIS)، روش های تحقیق در کسب و کار، مدیریت استراتژیک، مدیریت تغییر استراتژیک، استراتژی تجارت الکترونیک (استراتژی الکترونیکی)، فرآیند کسب و کار، مهندسی الکترونیک، تجارت الکترونیک و بازاریابی، روش های توسعه نرم افزار در سیستم های اطلاعاتی (ADMIS) و مدیریت ارتباط با مشتری (CRM) است.

www.SID.ir