GÜLTEKIN GÜNDÜZ, M.Sc.[1]
E-mail: ggunduz@sabanciuniv.edu
TANKUT ACARMAN, Ph.D. [1]
(Corresponding author)
E-mail: tacarman@gsu.edu.tr
[1] Computer Engineering Department
  Galatasaray University
  Ciragan Cad., 36, Ortakoy, Istanbul 34349, Turkey

VEHICLE TRAVEL TIME ESTIMATION USING SEQUENCE PREDICTION

ABSTRACT

This paper proposes a region-based travel time and traffic speed prediction method using sequence prediction. Floating Car Data collected from 8,317 vehicles during 34 days are used for evaluation purposes. Twelve districts are chosen and the spatio-temporal non-linear relations are learned with Recurrent Neural Networks. Time estimation of the total trip is solved by travel time estimation of the divided sub-trips, which are constituted between two consecutive GNSS measurement data. The travel time and final speed of sub-trips are learned with Long Short-term Memory cells using sequence prediction. A sequence is defined by including the day of the week meta-information, dynamic information about vehicle route start and end positions, and average travel speed of the road segment that has been traversed by the vehicle. The final travel time is estimated for this sequence. The sequence-based prediction shows promising results, outperforms function mapping and non-parametric linear velocity change based methods in terms of root-mean-square error and mean absolute error metrics.

KEY WORDS

travel time estimation; recurrent neural networks; sequence prediction; intelligent transportation systems;

1. INTRODUCTION

Short-to-medium-term travel time and traffic speed prediction has been one of the crucial components of Intelligent Transportation System (ITS) applications. Travel time and traffic speed prediction methods are separated based on the underlying data acquisition mechanisms and prediction approaches categorized either parametrically or non-parametrically. Data acquisition is largely performed by stationary sensors such as remote traffic microwave sensors [1] or loop detectors [2, 3]. As an alternative solution to remote sensors, travel time is predicted using mobile phone GNSS data or Floating Car Data (FCD) [4-6]. Parametric approaches include prediction using regression trees [7], time-series analysis [8] while non-parametric approaches adopt pattern searching and artificial intelligence, like estimating network speed by processing traffic images taken from a traffic camera with Convolutional Neural Networks (CNNs) [9].

The proposal includes travel time and traffic speed estimation method based on sequence prediction that is extracted from a collection of large spatio-temporal GNSS data. Districts are smaller regions selected from the overall data collection region subject to their local speed distribution and noise variance. Non-linear traffic dynamics related to the district and meta information are captured using Recurrent Neural Networks (RNNs) that are formed of Long Short-term Memory (LSTM) cells. For evaluation and test purposes, data were collected in a metropolitan city of Istanbul, Turkey, over a period of 34 days, from a fleet of 8,317 passenger cars.

District specific travel time and traffic speed estimation are learned by breaking down the entire trip of vehicles into smaller sub-trips. A sub-trip is defined as the trip between two consecutive GNSS measurement-based records with at most an elapsed time of 15 seconds. In *Figure 1*, for cluster centres with possible links, a link starting and ending speed is established by relying on the available data. Afterwards, for each sub-trip belonging to the links, elapsed travel time Δt is predicted with $RNN_{\Delta t}$ according to the meta-information of the day of the week, which is fine-tuned using district coordinates and link start and end velocities. Using Δt, velocity v_{t+1} is predicted by $RNN_{v_{t+1}}$ at the next time instance as well.
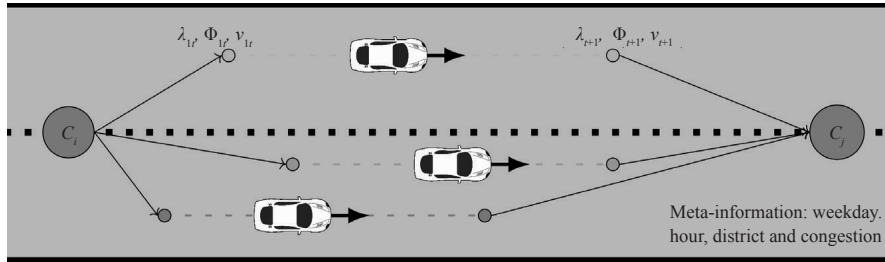
Figure 1 – Illustrative representation of prediction instances

Then, sub-trips are processed by district specific RNNs and predicted travel time estimations are stitched together to re-create the actual trip. The link travel speed is constituted by dynamic link speed of the roads traversed by vehicles. The starting and ending average link speed is calculated for the interval of 5 minutes by using the map-matched vehicle sub-trip speed data, respectively. Furthermore, this method provides continuous arterial speed information due to the analysis of sub-trips.

2. RELATED WORK

Short-term freeway travel time prediction based on a dynamic linear model using Bayesian inference was studied in [10]. The data were acquired from a loop detector and on a sequential basis, the state of a priori knowledge has been revised by newly available information. In [11], time lags related to elapsed time in segments were conducted using seasonal auto-regressive integrated moving average and Kalman filter. The gradient boosting method with time-lags related to previous elapsed times was presented in [7], the data were provided both by GNSS receivers mounted on vehicles and stationary sensors. An online freeway corridor travel time prediction was studied using Takagi-Sugeno-Kang Fuzzy Neural Network in [12]. Non-parametric estimation using Euclidean distances of FCD was studied in [13]. Travel time estimation based on GNSS measurement data has been performed in [14] using a suffix-tree-based method. In [15], the uncertainties in travel time predictions were regulated by the proposal of probabilistic prediction intervals, which are constructed using Bayesian techniques with coverage above the nominal confidence level of 90%.

Traffic speed estimation performed using the fusion of loop-detectors and probe vehicles was evaluated with GNSS data in [16]. The use of stationary sensors such as loop detectors and interpolating speed between measurements has been evaluated as inaccurate in congested and transition conditions [17]. Probe vehicle data sampling was investigated in [18] and concluded that sampling protocol regarding time or space affects the accuracy of estimation.

3. FLOATING CAR DATA COLLECTION

For the time period between January and February 2018, data from 8,317 vehicle on-board tracking systems were collected in Istanbul, Turkey. These tracking systems involve a GNSS receiver, an odometer measuring the vehicle wheel angular velocity and a cellular wireless communication modem to transmit measurement data to the back-end server. From each vehicle alongside the spatio-temporal information of latitude-longitude coordinates and time provided by a GNSS receiver, and vehicle velocity information measured at a higher frequency by an odometer on the vehicle was transmitted to the back-end server. For travel time and speed estimation purposes, twelve districts have been selected, subject to different characteristics constituted by either high and low-speed roads, curvatures, and different levels in GNSS accuracy due to the non-line-of-sight effects of urban canyons. *Figure 2* shows data collected in these selected districts, where the first district D_1 is shown in the top-left, the second district D_2 is shown in the top-centre and the last district D_{12} is plotted in the bottom right. The reported FCD coordinates are plotted with different colours denoting various vehicle speed ranges in order to give an insight into the types of roads, areas that get congested throughout the time of the day and to illustrate the characteristics of each individual district. Each district has different characteristics, with varying legal speed limit roads, road curvatures and GNSS noise relating to urban canyons. The data are coloured according to the vehicle velocity between 0-20 km/h, 20-50 km/h, 50-90 km/h and the velocity higher than 90 km/h coloured in blue, green, orange, and red, respectively.

Table 1 gives north-west and south-east coordinates of the district bounds in the forms of latitude and longitude, denoted by P_{NW} and P_{SE}, respectively.
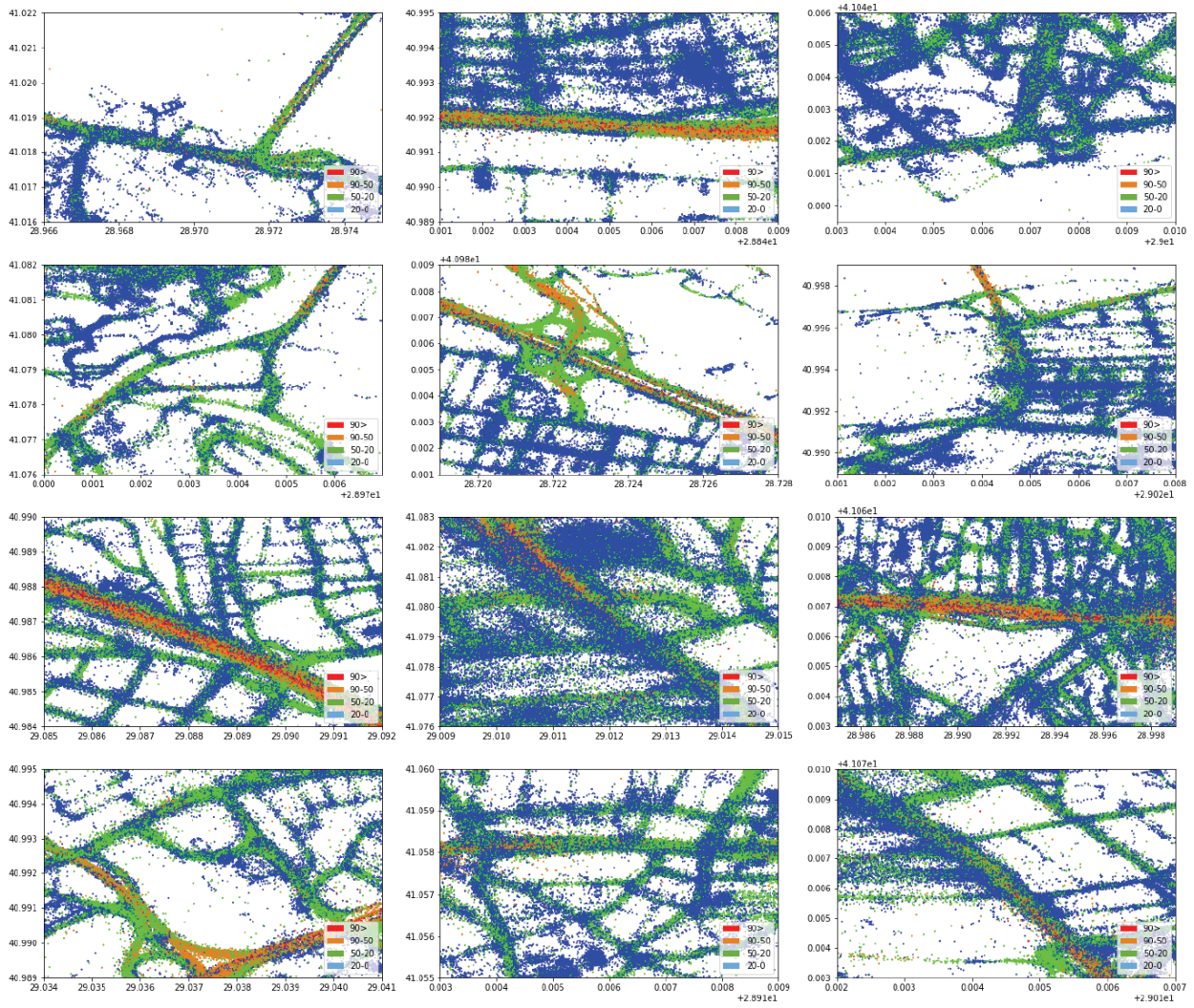
Figure 2 – Representation of GNSS data collected from the selected districts

Table 1 – Number of unique vehicles data has been collected, north-west/south-east bounds and area dimensions of districts

	ID	P_{NW}	P_{SE}	Width [m]	Height [m]
D_1	1876	(41.022000, 28.966000)	(41.016000, 28.975000)	755.8	667.9
D_2	3354	(40.995000, 28.841000)	(40.989000, 28.849000)	672.1	667.9
D_3	2132	(41.046000, 29.003000)	(41.039490, 29.010000)	587.6	724.9
D_4	1539	(41.082000, 28.970000)	(41.076000, 28.976976)	585.3	667.9
D_5	2654	(40.989000, 28.719000)	(40.981000, 28.728000)	756.2	890.5
D_6	1248	(40.999000, 29.021000)	(40.989000, 29.028000)	588.1	1,113.1
D_7	3584	(40.990000, 29.085000)	(40.984000, 29.092000)	588.1	667.9
D_8	2925	(41.083000, 29.009000)	(41.076000, 29.015000)	503.4	779.2
D_9	4000	(41.070000, 28.985000)	(41.063000, 28.999000)	1,174.9	779.2
D_{10}	2461	(40.995000, 29.034000)	(40.989000, 29.041000)	588.1	667.9
D_{11}	958	(41.060000, 28.913000)	(41.055000, 28.919000)	503.6	556.5
D_{12}	3073	(41.080000, 29.012000)	(41.073000, 29.017000)	419.5	779.2

Table 2 – Statistical analysis of collected vehicle speed and sub-trip distance

District	Collected vehicle speed				Sub-trip distance			
	Mean [km/h]	Standard deviation	Skew	Kurtosis	Mean [m]	Standard deviation	Skew	Kurtosis
D_1	18.67	16.95	0.7765	0.1874	24.60	22.07	2.3196	9.2757
D_2	16.76	25.42	1.8807	2.9224	22.00	22.93	1.5826	2.6389
D_3	11.99	12.73	1.2090	1.5411	15.17	17.71	2.3737	7.1933
D_4	18.08	13.12	1.0954	1.6214	18.15	20.18	2.3834	7.2759
D_5	23.55	21.92	1.0413	1.0449	27.53	25.53	1.9070	4.1603
D_6	10.42	14.81	1.6927	3.0640	32.51	27.92	1.4865	3.0003
D_7	25.37	26.76	1.4888	1.8157	21.13	21.20	2.1015	5.6636
D_8	8.02	13.97	2.3915	6.8270	16.23	20.88	2.8880	10.3628
D_9	19.54	23.22	1.9410	3.7079	17.92	24.08	3.2460	13.8090
D_{10}	24.69	19.39	0.6294	0.0578	35.52	27.57	1.3212	1.9767
D_{11}	14.70	15.16	1.1494	1.0894	14.82	12.85	2.0523	6.6082
D_{12}	17.60	16.78	1.3092	1.9920	17.76	18.32	2.1042	5.3458

Column ID denotes the unique number of vehicles from the whole dataset that has reported valid trips from the corresponding district, and the dimension of the area is denoted by its width and height.

Table 2 gives the mean, standard deviation, skew and kurtosis properties of the reported vehicle speeds; D_8 has the lowest and D_{10} has the highest mean speed. Also, *Table 2* presents information about the sub-trips. The mean distance travelled, standard deviation and related statistics of the sub-trip travel distance distribution are given.

Particularly, the number of unique vehicles traversing D_1 for different days of the week and hour is plotted in *Figure 3* to show how a particular day of the week and hour affects the district traffic congestion. Traffic density is reached at its maximum peak on Friday afternoon. The number of unique vehicles from our dataset is used as a function that represents overall district traffic congestion.

Figure 3 – Number of unique vehicles present in D1 for different day of the week and hour

4. SEQUENCE PREDICTION

Deep learning refers to an Artificial Neural Network (ANN) and is a sub-field of machine learning. The deep term refers to the number of hidden layers in the neural network. ANN is biologically inspired by the architecture and function of the human brain that learns from large-scale observational data. The usefulness and applicability of neural networks is enhanced by the introduction 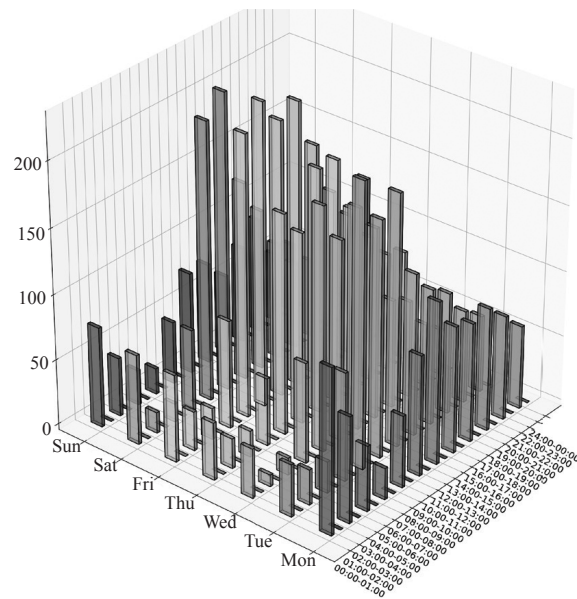of CNN and RNN architectures. These architectures are de-signed to manipulate and to encode certain hidden properties of the input data, which is generally a multi-dimensional array. The dimension of the input arrays can be adapted such as 1D for sequences of text, 2D for grey scale images, and 3D for coloured images or videos. Three main layers are namely convolution, pooling, and fully-connected. Convolutional layer applies the convolution operation, i.e. performs dot products between given filters and local regions of the input data and creates the convolved features. The fully-connected layer is a former neural network layer and it has connections

to all units in the pooling layer. The pooling layer corresponds to a gradual reduction of the spatial size of the convolved features and reduces the computation cost and the size of parameters to be tuned. The current trend is seeking the promising and practical solution at discovering complicated patterns in high-dimensional and large-scale data [19].

LSTM units proposed to overcome the vanishing gradient problem of RNNs have been successful at improving the level in prediction accuracy [20], and they are widely used in time series [21] and sequence prediction [22, 23] applications. For each district, an RNN consisting of an LSTM layer and a densely connected layer is trained. The district specific RNNs are deployed to learn non-linear spatio-temporal and district geometric characteristics like GNSS noise errors, traffic congestion related to a day of the week, hour and position. From the FCD, sub-trips are extracted subject to the constraint that a vehicle should not be stationary, i.e. a travelling speed has to be greater than zero, and at most 15 seconds should elapse between two consecutive GNSS reports such that a sub-trip can be established between them. For two latitude and longitude coordinates denoted by λ_s, φ_s and λ_e, φ_e, and vehicle speed v_s, v_e at these given coordinates, where s and e denote starting and ending, respectively, a sequence is defined related to $|N|$, which is the number of unique vehicles from our dataset in the district, Δd denotes the distance between the coordinates for a week of the day and minute bin, the following network input and output is defined:

- $F_{RNN_{\Delta t}} = \{\text{day}, \lambda_s, \lambda_e, \varphi_s, \varphi_e, v_s, v_e, \Delta d\}$ is a sequence fed to $RNN_{\Delta t}$ to estimate the travelling time;
- $\hat{Y}_{RNN_{\Delta t}} = \{\Delta t\}$ denotes travelling time estimation as an output of $RNN_{\Delta t}$;
- $F_{RNN_{vt+1}} = \{\text{day}, \lambda_s, \lambda_e, \varphi_s, \varphi_e, v_s, \Delta t, \Delta d\}$ is a sequence fed to the RNN_{vt+1} to estimate the ending speed of the link at the next time instance;
- $\hat{Y}_{RNN_{vt+1}} = \{v_e\}$ denotes the ending speed of the link estimation as an output of RNN_{vt+1}.

To summarize, $RNN_{\Delta t}$ generates the estimation of time to elapse for a sub-trip given in a sequence constituted by the week of the day, time of the day, historical district traffic congestion, position and velocity at given coordinates. For instance, *Figure 1* illustrates the travel time prediction of Δt. RNN_{vt+1} generates the final speed output when the vehicles complete their sub-trip.

5. LINK EXTRACTION

For each district contributing to the collected FCD dataset, links are created in the first stage. A link is a directed edge between roads of the district, whose dynamic traffic data are collected. An unsupervised link extraction method is adopted, which is compatible with the ad-hoc FCD and scalability of the district number or size is ensured. In order to extract links, the FCDs collected from each district are divided into clusters. Since the latitude and longitude distributions of the data are not uniform, i.e. they are reported from roads with different speeds, traffic congestions, road size and traffic flows, the 'bottom-up' hierarchical clustering with connectivity measures, agglomerative clustering is preferred. Agglomerative clustering establishes clusters with uneven sizes, which also solves the problem of uneven data distribution. A cluster satisfying the connectivity constraints assures the inclusion of a segment from a single road in the district. This connectivity procedure excludes data that do not belong to that particular road. The following clusters and their extracted centres, valid links are checked for each cluster centre and the closest neighbours. Between pairwise cluster centre coordinates, an edge is created and divided into 50 evenly spaced points. If an FCD exists within a range of 5 metres in each interval between these evenly spaced points then the edge is accepted as a valid link.

Figure 4 shows the resulting clusters for each district where FCD data are plotted in different colours denoting its membership in an individual cluster. Also, cluster centres are marked with blue circles and the accepted links are plotted with a black line. In order to extract the link speeds for each sub-trip in the FCD, the following procedure is applied: the starting position and the ending position are matched with the closest cluster centre, and the closest starting cluster centre is removed from the set of ending cluster centres. The vector constituted by a change in latitude-longitude coordinates is guaranteed to be in compliance with the heading of the link with a tolerance of 45º.

In *Figure 4*, if the sub-trip is assigned to an accepted link, it is added in the set to establish the starting and the ending speed of the link. For intervals of five minutes, the set belonging to the link is averaged where the average of the starting velocities of the sub-trips is used to create the starting velocity of the directed link, and the average of the ending velocities of the sub-trips is used to create the
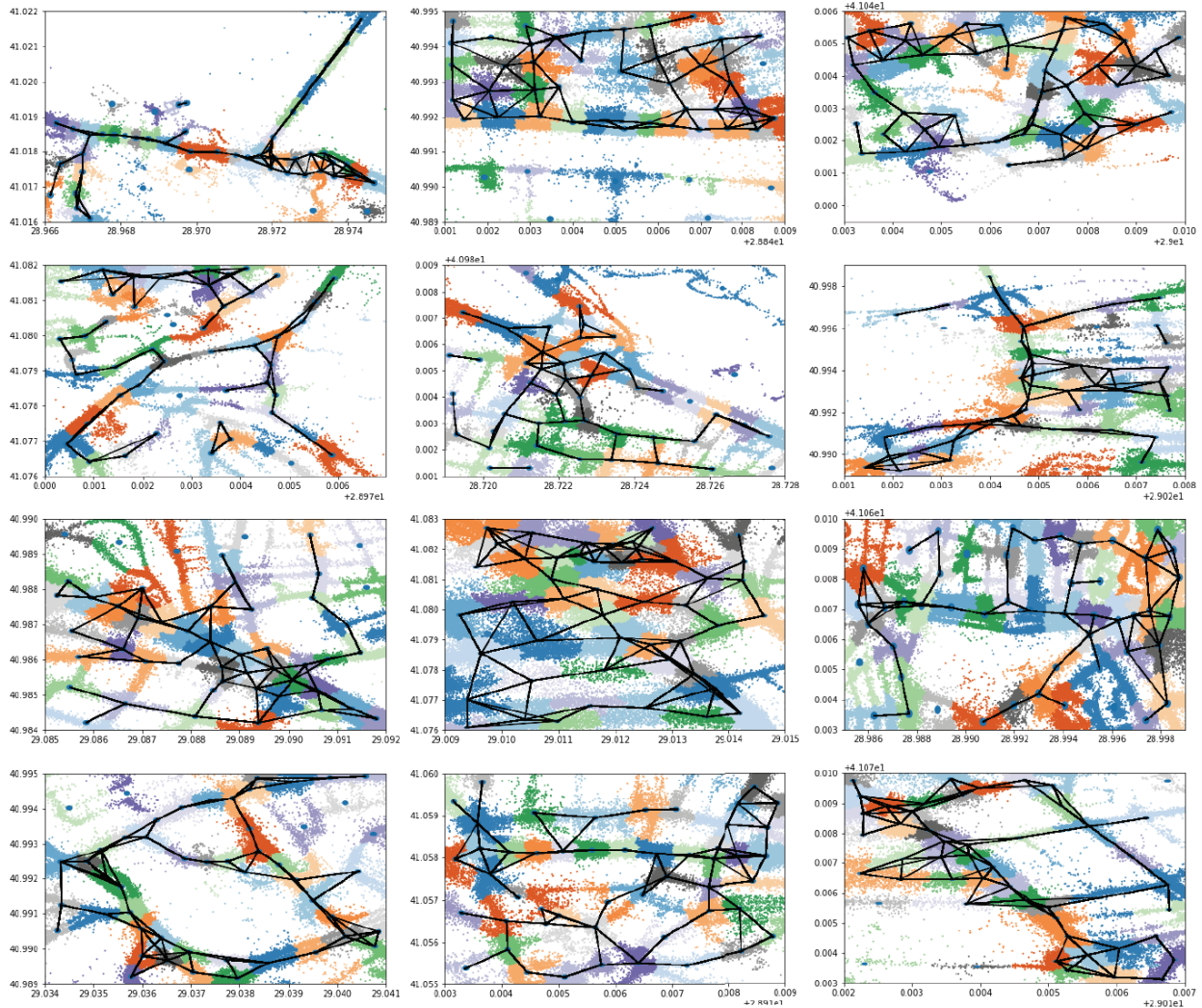
Figure 4 – Representation of unsupervised link extraction

ending velocity of the link. After the link speeds are extracted, all FCD velocities are replaced with the speeds of the link.

6. EXPERIMENTAL EVALUATION

For each district, the collected FCD is sorted according to the day of the week and time in descending order, and every other element is divided into training and testing data in order to provide a balanced number of observations for each day of the week and minute bin. Twenty-five percent of the dataset is used for training, the other 25% for validation and the remaining 50% of the dataset is used for testing. Since a large number of trip data is available, the training ratio is kept low for the purpose of increasing the generalization performance of the RNNs. The training batch size is chosen as 32 observations and it is done for 50 epochs. *Figure 5* shows the Cumulative Distribution Function (CDF)

of elapsed time between consecutive FCD data. *Figure 6* shows the CDF of travelling speed of vehicles and *Figure 7* shows the CDF of the travelled sub-trip distance for the selected districts. Districts D_2, D_6, and D_{10} have a higher distribution of longer elapsed times between data reports, and districts D_1 and D_3 have a higher distribution of shorter elapsed times. Also, districts D_5, D_{10}, and D_{11} have a higher distribution of high-speed trips. Lastly, districts D_6 and D_{10} show a higher distribution of sub-trips with higher travelled distances.

The total number of sub-trips for each district is given under the column entitled |*Data*| in *Table 3*; 50% of these data are used for testing purposes. For comparison purposes with respect to the proposed RNN with LSTM cell layer, a Multilayer Perceptron (MLP) with two hidden layers is also trained to test the function mapping approach versus the sequence-based prediction system. Since acceleration values between the consecutive coordinates of
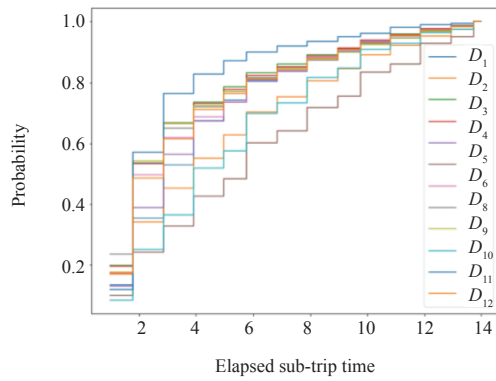
Figure 5 – Cumulative Distribution Function regarding elapsed time between consecutive GNSS sampling data for different districts
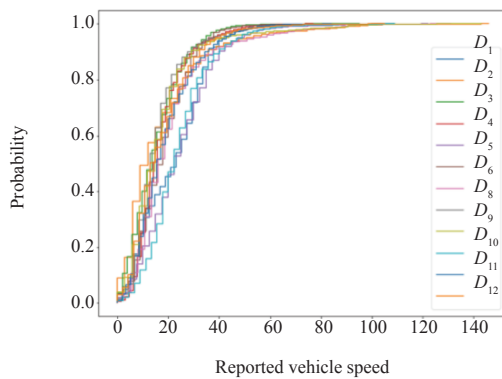

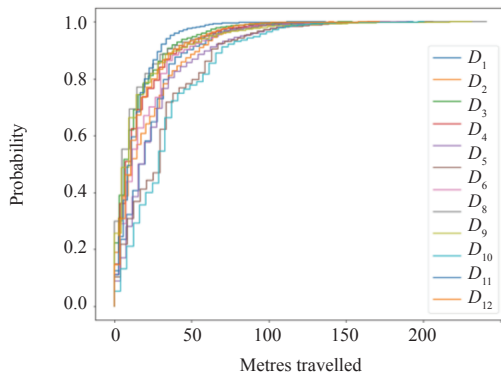*Figure 6 – Cumulative Distribution Function regarding vehicle speed in each district*

Figure 7 – Cumulative Distribution Function regarding distance travelled in each district

the sub-trip are not measured, the travel time and the travel speed are also predicted by using a linear velocity change model, which is denoted by symbol *L* in *Table 3*. The smallest error is highlighted for easier comparison. For testing and comparison purposes, the root-mean-square error (RMSE) and the mean absolute error (MAE) are adopted. RMSE is useful in capturing undesirable large errors, while the number of observations can affect RMSE. MAE weights differences equally [24]. Since each district

has a varying number of sub-trips, the estimation approaches are tested and evaluated using both metrics.

Figure 8 illustrates the estimation responses of trained RNNs compared to the ground-truth values. The top row of *Figure 8* shows the travel time estimations of $RNN_{\Delta t}$, sub-trip travel time in seconds is plotted with respect to the sub-trip index. Ground-truth values are sorted according to travel time and plotted with red step line; the corresponding $RNN_{\Delta t}$ output of the sub-trip sequences are shown with blue points. Similarly, the bottom row of *Figure 8* shows the estimation responses of ending travelling speed with respect to the sub-trip index. For all twelve districts, the travel time estimations are compared to the ground-truth in *Figure 9*, and ending travel speed estimations are plotted in *Figure 10*. In *Figure 9*, the first figure in the first row compares travel time versus prediction for district D_1 and the last figure compares the responses for D_{12}. In *Figure 10*, the first figure in the first row compares the ending travel speed versus prediction for district D_1 and the last figure compares the responses for D_{12}. District numbering is increased in each row from left to right. The travel time prediction belonging to some districts includes higher prediction error variance but the overall step-like change is captured.

Table 3 gives RMSE and MAE of predicted sub-trip times compared to the actual sub-trip time relating to districts and prediction method, whereas *Table 4* presents the final speed prediction responses
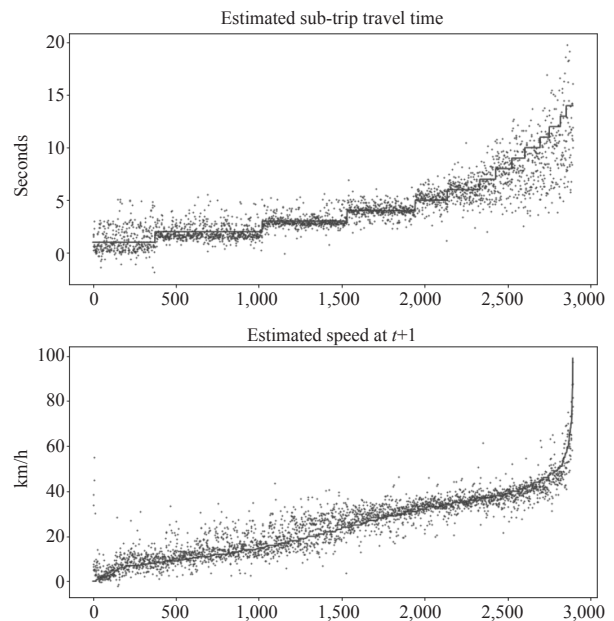

Figure 8 – Predicted travel time and final speed of sub-trips compared with respect to the ground-truth for district D2

Table 3 – RMSE and MAE related to districts and predicted sub-trip time

	|Sub-trip|	$RNN_{\Delta t}$ [s]	$MLP_{\Delta t}$ [s]	$L_{\Delta t}$ [s]
D_{1RMSE} D_{1MAE}	6,250	1.384 0.683	2.076 1.193	1.436 0.671
D_{2RMSE} D_{2MAE}	6,833	1.704 1.021	2.775 1.949	1.854 1.037
D_{3RMSE} D_{3MAE}	17,429	1.333 0.785	1.968 1.165	1.440 0.750
D_{4RMSE} D_{4MAE}	13,135	1.280 0.654	2.038 1.223	1.362 0.687
D_{5RMSE} D_{5MAE}	15,996	1.208 0.648	2.012 1.283	1.255 0.627
D_{6RMSE} D_{6MAE}	18,222	1.691 1.104	2.303 1.579	1.943 1.165
D_{7RMSE} D_{7MAE}	14,055	1.357 0.730	2.141 1.210	1.495 0.768
D_{8RMSE} D_{8MAE}	32,679	1.452 0.754	2.242 1.315	1.571 0.837
D_{9RMSE} D_{9MAE}	42,006	1.314 0.765	2.514 1.448	1.547 0.836
D_{10RMSE} D_{10MAE}	22,424	1.574 0.942	2.049 1.312	1.710 0.893
D_{11RMSE} D_{11MAE}	16,823	1.190 0.574	1.688 0.875	1.231 0.612
D_{12RMSE} D_{12MAE}	13,190	1.433 0.745	2.164 1.289	1.612 0.831

Table 4 – RMSE and MAE related to districts and ending travel speed estimation

|	$RNN_{v_{t+1}}$ [km/h]	$MLP_{v_{t+1}}$ [km/h]	$L_{v_{t+1}}$ [km/h]
D_{1RMSE} D_{1MAE}	5.205 3.608	7.633 5.365	5.910 3.906
D_{2RMSE} D_{2MAE}	4.506 3.283	7.143 5.171	6.371 4.344
D_{3RMSE} D_{3MAE}	4.045 2.810	6.105 4.287	6.737 4.632
D_{4RMSE} D_{4MAE}	4.048 2.887	5.690 4.211	6.614 4.668
D_{5RMSE} D_{5MAE}	4.185 3.046	5.510 3.945	6.513 4.367
D_{6RMSE} D_{6MAE}	4.946 3.631	7.011 5.252	7.971 5.750
D_{7RMSE} D_{7MAE}	5.372 3.904	7.003 4.983	8.171 5.407
D_{8RMSE} D_{8MAE}	3.937 2.752	5.894 3.958	7.729 5.543
D_{9RMSE} D_{9MAE}	4.103 2.725	6.146 4.045	8.418 5.806
D_{10RMSE} D_{10MAE}	5.771 4.209	8.216 6.346	8.649 5.916
D_{11RMSE} D_{11MAE}	4.013 2.940	5.725 4.186	6.657 4.594
D_{12RMSE} D_{12MAE}	4.157 2.977	6.659 4.871	6.812 4.778
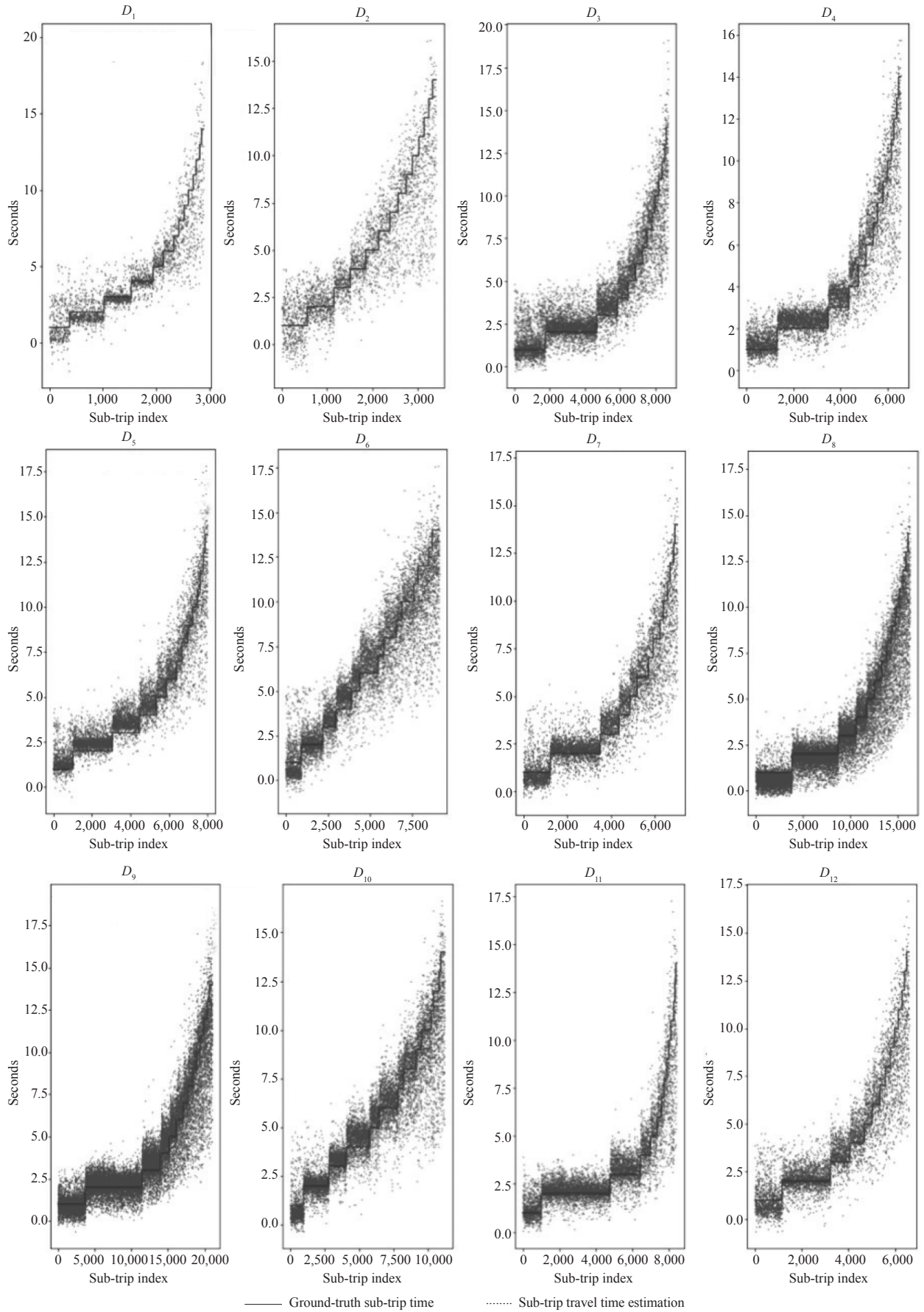
Figure 9 – *Predicted travel time of sub-trips compared with respect to the ground-truth for all twelve districts*

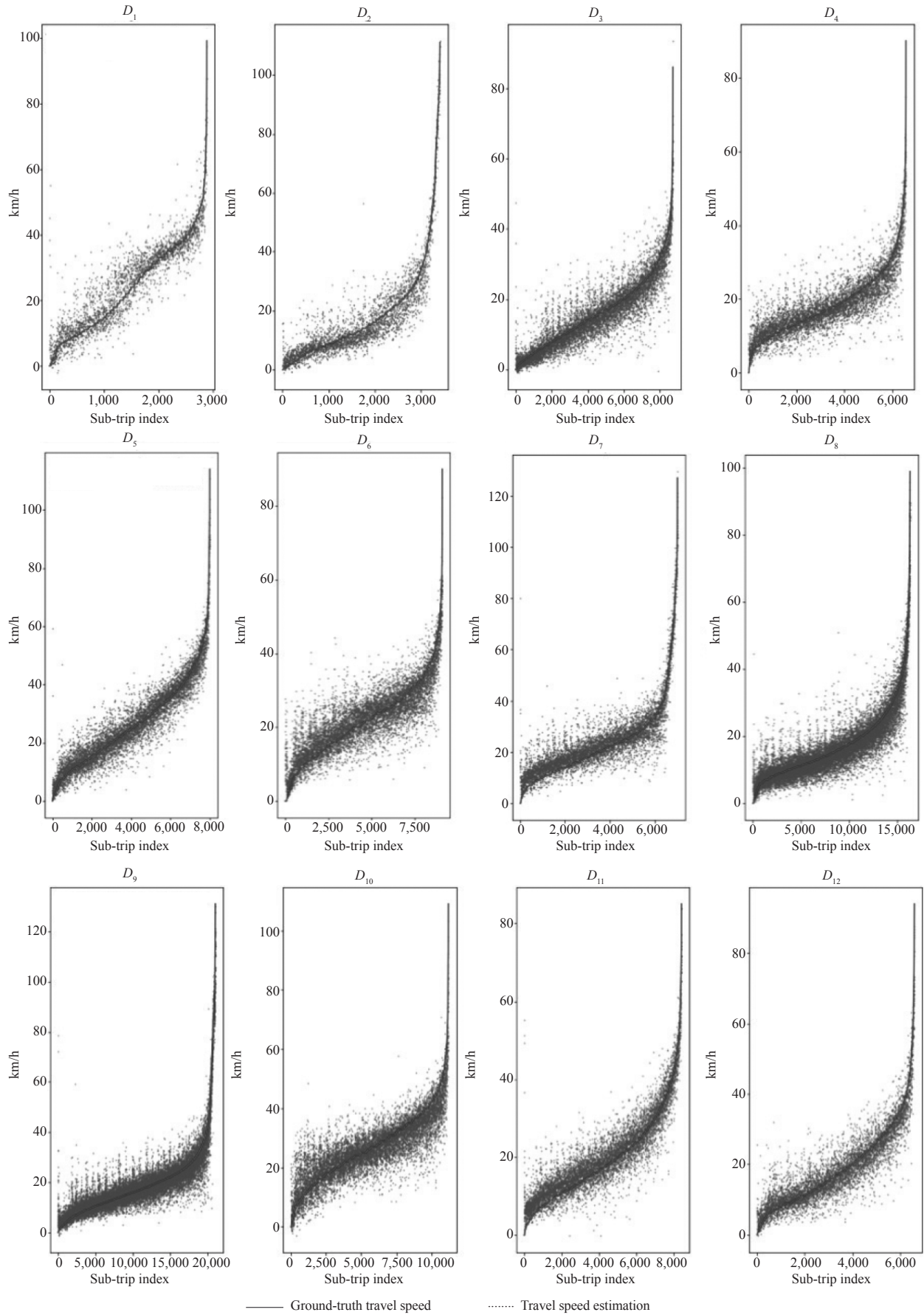Ground-truth travel speed .... Travel speed estimation

Figure 10 – RMSE and MAE related to each district and the prediction responses of ending speed with respect to the sub-trip index

concerning the same districts. Sequence prediction assured by the RNN method outperforms both predictions by linear speed change and MLP function mapping in terms of RMSE and performs better than the linear speed change in 8 out of 12 districts in terms of MAE. Linear speed change approach performs better than MLP in both error metrics. For ten out of the twelve districts, RNN performs sub-second MAE subject to the ground truth elapsed times in the interval of 1 to 15 seconds. The lowest RMSE value is achieved for D_{11}, which has the smallest mean sub-trip travel distances. Following comparisons plotted in *Figure 5*, district D_5, which has a higher portion of longer sub-trip times than D_3, lower RMSE and MAE values are achieved. This observation can underlie that prediction time interval length may not be the most significant feature to improve the estimation accuracy. Sub-trip final speed estimation is also performed most accurately by the RNN, where the linear change prediction and MLP approaches exchange rankings according to the district and evaluation metrics.

7. CONCLUSION

The region-based travel time estimation by dividing the total trip into sub-trips between the observed GNSS coordinates and by introducing RNNs using sequence prediction has been proposed. During experimental evaluation studies, using a low training split such as in our study 25% training and 25% validation, the generalization performance of RNNs outperforms both MLP and linear velocity change prediction results in root-mean-square error. This performance is assured in 8 out of 12 district results with lower mean absolute error values. For the twelve inspected districts, the sub-trip travel times for ten out of twelve districts are predicted with sub-second MAE value with respect to ground-truth observations that are reported in the interval of 1 to 15 seconds. With sequence-based approach performing the best travel time prediction using linear speed change, denoted by L ranks second compared to a prediction by function mapping MLP. However, for the given the elapsed time of the sub-trip, the final link speed prediction by the RNN approach reaches at the highest level in prediction performance metrics. Furthermore, the prediction performance of the Multilayer Perceptron method and the linear velocity change model varies for depending on the districts and the evaluation metrics.

GÜLTEKIN GÜNDÜZ[1]
E-mail: ggunduz@sabanciuniv.edu
TANKUT ACARMAN[1]
E-mail: tacarman@gsu.edu.tr
[1] Bilgisayar Mühendisliği Bölümü
Galatasaray Üniversitesi
Çırağan Cad. No:36, 34349 Ortaköy/İstanbul, Türkiye

ARAÇ SEYAHAT SÜRESININ DIZI TAHMINI KULLANILARAK KESTIRILMESI

ÖZET

Bu makalede, dizi tahmini kullanan bölge esaslı seyahat süresi ve trafik hızı tahmin yöntemi önerilmektedir. 34 gün boyunca 8,317 araçtan toplanan hareketli araba verisi değerlendirme amaçlı olarak kullanılmaktadır. On iki ilçe seçilerek uzay-zamansal doğrusal olmayan ilişkiler Tekrarlayan Sinir Ağları kullanılarak öğrenilmektedir. Toplam seyahatin zaman tahmini, ardışık iki GNSS ölçüm verisi arasında oluşan bölünmüş alt yolcuların seyahat süresi tahmini ile çözülmektedir. Uzun Kısa-Süreli Bellek hücreleri ile yolculuk süresi ve alt yolcukların son hızı dizi tahmini kullanılarak öğrenilir. Dizi, hafta içerisinde gün meta-bilgileri, aracın rota başlangıç ve bitiş konumlarıyla ilgili dinamik bilgiler ve araç tarafından katedilen yol bölümünün ortalama seyahat hızı dahil edilerek tanımlanır. Nihai seyahat süresi tanımlanan dizi için tahmin edilmektedir. Dizi esaslı tahminleme ümit vaat eden sonuçlar vermektedir, kare kök ortalama hata ve ortalama mutlak hata metrikleri karşılaştırıldığında fonksiyon haritalama ve parametrik olmayan doğrusal hız değişimine dayalı yöntemlerden daha yüksek performans sağlamaktadır.

ANAHTAR KELİMELER

seyahat süresi tahmini; tekrarlayan sinir ağları; dizi kestirimi; akıllı ulaşım sistemleri;

REFERENCES

[1] Ma X, Tao Z, Wang Y, Yu H, Wang Y. Long Short-Term Memory Neural Network for Traffic Speed Prediction Using Remote Microwave Sensor Data. *Transportation Research Part C: Emerging Technologies.* 2015;54: 187-197.

[2] Kim S, Coifman B. Comparing Inrix Speed Data Against Concurrent Loop Detector Stations Over Several Months. *Transportation Research Part C: Emerging Technologies.* 2014;49: 59-72.

[3] Coifman B. Empirical Flow-Density and Speed-Spacing Relationships: Evidence of Vehicle Length Dependency. *Transportation Research Part B: Methodological.* 2015;78: 54-65.

[4] Yuan Y, Van Lint H, Van Wageningen-Kessels F, Hoogen-doorn S. Network-Wide Traffic State Estimation Using Loop Detector and Floating Car Data. *Journal of Intelligent Transportation Systems.* 2014;18(1): 41-50.

[5] Yao B, Chen C, Cao Q, Jin L, Zhang M, Zhu H, Yu B. Short-Term Traffic Speed Prediction For An Urban Corridor. *Computer-Aided Civil and Infrastructure Engineering.* 2017;32(2): 154-169.

[6] Woodard D, Nogin G, Koch P, Racz D, Goldszmidt M, Horvitz E. Predicting Travel Time Reliability Using Mobile Phone GPS Data. *Transportation Research Part C: Emerging Technologies.* 2017;75: 30-44.

[7] Zhang Y, Haghani A. A Gradient Boosting Method to Improve Travel Time Prediction. *Transportation Research Part C: Emerging Technologies.* 2015;58: 308-324.

[8] Min W, Wynter L. Real-Time Road Traffic Prediction With Spatio-Temporal Correlations. *Transportation Research Part C: Emerging Technologies.* 2011;19: 606-616.

[9] Ma X, Dai Z, He Z, Ma J, Wang Y, Wang Y. Learning Traffic As Images: A deep convolutional neural network for large-scale transportation network speed prediction. *Sensors.* 2017;17(4): 818. Available from: doi:10.3390/s17040818 [Accessed 18th September 2019].

[10] Fei X, Lu C-C, Liu K. A Bayesian Dynamic Linear Model Approach for Real-Time Short-Term Freeway Travel Time Prediction. *Transportation Research Part C: Emerging Technologies.* 2011;19(6): 1306-1318.

[11] Xia J, Chen M, Huang W. A Multistep Corridor Travel-Time Prediction Method Using Presence-Type Vehicle Detector Data. *Journal of Intelligent Transportation Systems*. 2011;15(2): 104-113.

[12] Zhang Y, Ge H. Freeway Travel Time Prediction Using Takagi–Sugeno–Kang Fuzzy Neural Network. *Computer-Aided Civil and Infrastructure Engineering.* 2013;28(8): 594-603.

[13] Rahmani M, Jenelius E, Koutsopoulos HN. Non-Parametric Estimation of Route Travel Time Distributions From Low-Frequency Floating Car Data. *Transportation Research Part C: Emerging Technologies.* 2015;58: 343-362.

[14] Wang Y, Zheng Y, Xue Y. Travel Time Estimation of A Path Using Sparse Trajectories. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 24-27 August 2014, New York, NY, USA*. New York, NY, USA: ACM; 2014. p. 25-34.

[15] Khosravi A, Mazloumi E, Nahavandi S, Creighton D, Van Lint J. Prediction Intervals To Account for Uncertainties in Travel Time Prediction. *IEEE Transactions on Intelligent Transportation Systems.* 2011;12(2): 537-547.

[16] Bachmann C, Abdulhai B, Roorda MJ, Moshiri B. A Comparative Assessment of Multi-Sensor Data Fusion Techniques for Freeway Traffic Speed Estimation Using Microsimulation Modeling. *Transportation Research Part C: Emerging Technologies.* 2013;26: 33-48.

[17] Soriguera F, Robuste F. Requiem for Freeway Travel Time Estimation Methods Based on Blind Speed Interpolations Between Point Measurements. *IEEE Transactions on Intelligent Transportation Systems.* 2011;12(1): 291-297.

[18] Jenelius E, Koutsopoulos HN. Probe Vehicle Data Sampled by Time or Space: Consistent Travel Time Allocation and Estimation. *Transportation Research Part B: Methodological.* 2015;71: 120-137.

[19] LeCun Y, Bengio Y, Hinton G. Deep Learning. *Nature.* 2015;521(7553): 436-444.

[20] Hochreiter S, Schmidhuber J. Long Short-Term Memory. *Neural Computation.* 1997;9(8): 1735-1780.

[21] Malhotra P, Vig L, Shroff G, Agarwal P. Long Short Term Memory Networks for Anomaly Detection in Time Series. In: *Proceedings of the 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 22-24 April 2015, Bruges, Belgium*. Louvain-la-Neuve, Belgium: Presses universitaires de Louvain; 2015. p. 89-94.

[22] Alahi A, Goel K, Ramanathan V, Robicquet A, Fei-Fei L, Savarese S. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 21-23 June 1994, Seattle, WA, USA*. New York, NY, USA: IEEE; 2016. p. 961-971.

[23] Sutskever I, Vinyals O, Le Q. V. Sequence to sequence learning with neural networks. In: Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ. (eds.) *Advances in Neural Information Processing Systems 27*. Red Hook, NY USA: Curran Associates, Inc.; 2014. p. 3104-3112.

[24] Willmott CJ, Matsuura K. Advantages of the Mean Absolute Error (MAE) Over the Root Mean Square Error (RMSE) in Assessing Average Model Performance. *Climate Research.* 2005;30(1): 79-82.