Project Number 288094

eCOMPASS

eCO-friendly urban Multi-modal route PlAnning Services for mobile uSers

STREP
Funded by EC, INFSO-G4(ICT for Transport) under FP7

eCOMPASS – TR – 048

Recent Research Results on Traffic Prediction (an overview)

Dionisis Kehagias

January 2014

Recent Research Results on Traffic Prediction

Dionisis Kehagias, Senior Researcher

Centre for Research and Technology Hellas – Information Technologies Institute (CERTH/ITI)

Introduction

This report outlines the most representative research results on Traffic Prediction at the Information Technologies Institute of CERTH that were developed within the scope of the eCOMPASS project (http://www.ecompass-project.eu/). The main objectives of traffic prediction techniques is to predict traffic, expressed either on travel time or average speed, within a time interval ahead in time using real time data from GPS devices (i.e., transmit timestamp, location, and speed). The main techniques for Traffic Prediction can be classified into two main categories: (a) *Parametric techniques*, whose purpose is to calculate the parameters of a model so that the model fits into an existing dataset in the best possible way and then use this model for making prediction on real data; and (b) *Non-parametric techniques* that typically use a machine learning approach for making predictions without a predefined model.
CERTH-ITI has recently demonstrated (2nd eCOMPASS review) its latest findings from the application of existing as well as novel traffic prediction techniques.

Parametric techniques

The most common category of parametric traffic prediction algorithms, consists of methods that are based on time series analysis. In brief, time series forecasting algorithms calculate the variable under study as a function of its previous values and an error-term. Among all multivariate space-time approaches, Space-Time Auto Regressive Integrated Moving Average (STARIMA) and its variations have been proposed as the most efficient models for performing reliable traffic forecasting, e.g. see [1]. The simple STARIMA model is derived as:

$$X_t = \sum_{k=1}^{p} \sum_{l=0}^{\lambda^k} \phi_{k,l} W_l X_{t-k} - \sum_{k=1}^{q} \sum_{l=0}^{m^k} \theta_{k,l} W_l \varepsilon_{t-k} + \varepsilon_t$$

The goal of the prediction technique is to calculate the ϕ, θ parameters that minimize the prediction error ε_t. A challenging task of the STARIMA design process is the selection of the most appropriate roads to participate in the model (expressed by the *W* matrices). A variation of the STARIMA model, called Lag-STARIMA has been proposed in eCOMPASS [2] that takes into account past snapshots of roads that are located at a long distance, (as opposed to neighboring ones).

Non-parametric techniques

The category of non-parametric techniques does not presuppose a particular model structure. Hence, for non-parametric methods, both the exact model structure and its parameters need to be specified along with traffic data processing, for model training. Therefore, a more extensive training procedure of the model is usually required as well as a broader training dataset in juxtaposition with the parametric methods. Some examples of popular non-parametric methods that are used for traffic prediction include:

- *K-Nearest Neighbour (kNN),* which performs forecasting by searching the historical database for the *k* clusters of traffic data that present the closest behavioral pattern with respect to the reference traffic data set (pattern matching).

- *Random Forest,* a learning method for classification (and regression) that operates by constructing a multitude of decision trees at training time and outputting the class that appears most often in the data by individual trees.
- *Clustering and classification techniques,* which try to create distinct groups of common traffic patterns by means of a clustering procedure with common characteristics.

In eCOMPASS we have developed a clustering-based technique which creates 200 road profiles and predicts traffic by classifying each road (based on its real-time traffic value) into the best matched profile.

Comparison of techniques

In order to assess the accuracy of each method we use the traffic prediction demo tool (Figure 1), which comprises a test-bed developed by CERTH/ITI for evaluating various traffic prediction algorithms.

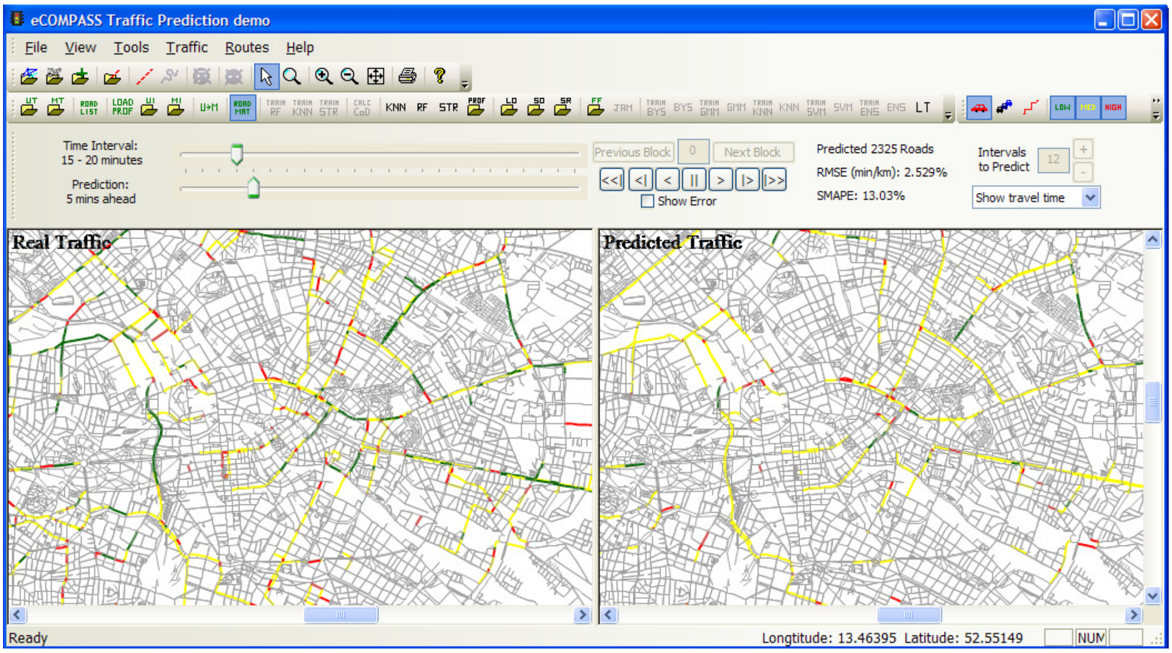

Figure 1: The traffic prediction demo tool

By using the traffic prediction tool we are able to evaluate the aforementioned approaches. In what follows we show the results of the comparison of the following techniques: (non-parametric) kNN, RF, Clustering-based, (parametric) STARIMA, Lag-STARIMA. The results are depicted in the following table for the prediction of traffic that occurs for the next 5, 10, 15, …, 60 minutes beyond current time. Traffic is expressed as average travel time (minutes/km).

Prediction Technique	*Prediction Accuracy for each 5-min interval ahead (RMSE in minutes)*											
	5min	**10min**	**15min**	**20min**	**25min**	**30min**	**35min**	**40min**	**45min**	**50min**	**55min**	**60min**
kNN	2.83	2.83	2.86	2.86	2.88	2.92	2.94	2.95	2.95	2.92	2.97	2.97
RF	2.77	2.79	2.82	2.81	2.84	2.86	2.83	2.85	2.83	2.81	2.85	2.82
Clustering	*2.437*	*2.435*	*2.418*	*2.404*	*2.411*	*2.408*	*2.433*	*2.459*	*2.483*	*2.516*	*2.501*	*2.494*
STARIMA	2.63	2.65	2.64	2.67	2.65	2.69	2.72	2.78	2.71	2.74	2.74	2.69
Lag STARIMA	2.53	2.52	2.52	2.56	2.53	2.54	2.56	2.65	2.56	2.56	2.56	2.51

From the table above it is shown that the clustering-based technique yields the best prediction accuracy for all time intervals compared to both parametric and non-parametric techniques.

References

[1] Kamarianakis Y. and Prastacos P. Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches. Transportation Research Record: Journal of the Transportation Research Board, 1857:74-84, 2003.

[2] Diamantopoulos T., Kehagias D., König F., Tzovaras D. Investigating the Effect of Global Metrics in Travel Time Forecasting. In Proceedings of the 16th International IEEE Annual Conference on Intelligent Transport Systems (ITSC 2013), The Hague, The Netherlands, October 6-9, pp.412-417, 2013.