Traffic Flow Forecast Survey

Ella Bolshinsky
Computer Science Department
Technion
ellabo@cs.technion.ac.il

Roy Freidman
Computer Science Department
Technion
roy@cs.technion.ac.il

June 3, 2012

Abstract

Short-term traffic flow forecasting is an important aspect of the ITS as traffic predication can alleviate congestion, which causes drivers to incur a longer traveling time and economical loses. In addition, traffic congestion increases the pollution and the fuel usage. Thus, it is one of the severe problems in Metropolitan areas. Further, in tunnels the forecasting may help scheduling the ventilation fans. This way, the ventilation cost might be decreased while the air quality increased. Additional aspect of traffic prediction is that it may enable the drivers to plan their departure time and traveling path, as they posses the predictive information. In this paper, we survey the different techniques used for traffic forecasting, the input data for these techniques, the output provided by them, as well as some general insights.

1 Introduction

As already stated in the abstract, short-term traffic flow forecasting is an important aspect of the ITS [48]. In addition, traffic forecasting can be very beneficial to drivers (save time and money), to the environment (decrease pollution) and even to the authorities (decrease ventilation costs in tunnels) [26, 37].

Without traffic prediction, one can only deduce about the future traffic conditions according to the current traffic conditions, which is not an accurate deduction [7]. The prediction process may rely on historical data, current data, or both, to forecast the traffic in some interval or time period in the future.

As for the nature of the traffic flow, several works have claimed that traffic tends to exhibit chaotic behavior, especially in the developing world [34] and during congestion [44]. Yet, there are other evidences that show that traffic conditions data is characteristically stochastic, and not chaotic, suggesting that it can be modeled and predicted by stochastic predictors. Different approaches were proposed for traffic prediction including ones that are based on time series models [7, 33, 35, 46], Kalman filter theory [26, 39], Markov chain model [56], non-parametric methods [8, 36, 58], simulation models [14], local regression models [17, 43, 47], fuzzy-neural approach [55], sequential learning [12], ATHENA model [32], spectral analysis [37, 45], Bayesian networks [9, 48], and neural networks [19, 20, 22, 25, 27]. Some works combine several techniques and as the following survey [16] suggests, there is no technique that clearly outperforms the other techniques. Thus future works should consider combinations of the techniques or selecting a technique according to some circumstances. Many of the approaches proposed provide traffic forecasting that ranges from 5 minutes to 1 hour. However there are techniques and systems that can make longer prediction. For instance, the INRIX [4] predicts future traffic up to a year a head.

Different works consider different factors that might influence the traffic, such as weather conditions, holidays, special events, etc. In addition, sometimes missing data is considered. An additional factor noticed when considering the traffic on some link is the traffic on adjacent roads [48].

The sources of the data used for prediction include Traffic Management Bureaus, loop detectors, sensors and controllers placed on the roadways, video cameras, GPS-enabled vehicles, mobile devices and data originated by people. Section 2 presents the input data used for traffic prediction, while section 3 presents the sources of the data. Some works also addressed the architecture of the system and considered communication resources. This issue is addressed in section 5. We also provide a taxonomy of different techniques used for prediction in section 4 as well as few comparative works. In section 6 we outline the different forecasting horizons proposed by the techniques. This survey differs from the survey in [16] as it addresses also the input of the process as well as some general insights and system architectures. Another comprehensive survey is [51]. Our survey differs from this survey, as it covers also more recent works. In addition, we cover the sources of input data and some architecture issues not mentioned in [51].

Note that a large number of acronyms are used in this survey. A table of acronyms used is available in Table 1.

ANN	Artificial neural network
ARIMA	Autoregressive integrated moving average
CEM	Competitive expectation maximization
DGLM	Dynamic generalized linear model
EKF	Extended Kalman filter
EM	Expectation maximum
EN	Expert network
ES	Exponential smoothing
FAR	Full autoregressive
FL	Fuzzy logic
FNM	Fuzzy-neural model
GMM	Gaussian mixture model
GN	Gate network
ITS	Intelligent transportation system
K-NN	K nearest neighbor
LGP	Linear genetic programming
MA	Moving average
MLP	Multilayer perceptron
MPC	Model predictive control
MTL	Multi-task learning
NN	Neural network
SAR	subset autoregressive
SARIMA	Seasonal ARIMA
STL	single task learning
UTN	Urban traffic network
VAR	Vector Auto Regressive

Table 1: Acronyms table

2 Input Data

Some works use only current traffic conditions (for instance, Random Walk Forecast [7]), others take into account only historical data (for instance, [54] addresses the short-term traffic condition forecasts at a fixed location in the network, based only on previous observations at the forecast location), while [14, 40, 47] integrate historical and current data. A comparison between the use of only historical data and the use of both historical and real-time data for rule construction is made in [29]. The addition of real-time data gave some marginal performance improvement. However, some circumstances that might have caused this improvement to be marginal are noted.

There are works that are based on time series models, for instance [35] and ARIMA. In time-series modeling, one attempts to predict the value of a variable based on a series of past samples of the variable at regular intervals.

In [37] it is noted that 17 days are optimal for the prediction. [48] addresses the possibility of incomplete data and is designed to cope with it. In addition, [48] combines information from adjacent and current roads.

Some works address the seasonal, surrounding and time differences. For instance, [37, 42, 45, 46] consider only

Considered factor	Works	Remarks
Different week days	[14], INRIX [4]	
Week days versus week ends	[52]	
Calendar and time information	JamBayes project [21]	
Different time periods during the day	[45], [46]	The daily data in these works is divided into 6 separate time periods
Holidays	INRIX [4], [23], [52]	
Seasonal differences	[14], INRIX [4]	
Weather	JamBayes project [21], INRIX [4], [23]	
Special events	[14], INRIX [4]	For instance, INRIX [4] which uses some of the techniques and tools developed in the JamBayes project, considers special events, such as school schedules and concerts
Sporting events	JamBayes project [21]	
Road constructions and accidents	INRIX [4]	
Distributed sources like shopping centers, home communities and car parks	[49]	

Table 2: Considered factors when performing traffic prediction

weekdays as the authors mention that the traffic on the weekends is different, while, [22] considers only business days. Table 2 summarizes the various factors and the works that consider them.

3 Sources of the Input Data

The data is obtained from various sources that include Traffic Management Bureaus, loop detectors, sensors and controllers placed on the roadways, video cameras, GPS-enabled vehicles, mobile devices and data originated by people.

[48] and [49] use data recorded by the Traffic Management Bureau of Beijing, which records vehicles per hour. In [27], which also gets the data in units of vehicles per hour (veh/h), the data used for analysis are the vehicle flow rates of discrete time series that are recorded every 15 min.

Loop detectors is a common source of data, used in [6, 12, 13, 14, 19, 37, 42]. [37] uses inductive loop traffic detectors, while in [6] they are located in an urban area.

[31] obtains the data from sensors placed on the roadways. In [50], data is collected from a data-collection point that is located on National Highway 107, Guangzhou, Guangdong, China. The data in [46] was originated in controllers installed at 144 locations of the down-town urban network. UK Highways Agency collected the data for [44] by using an incident detection and automatic signalling (MIDAS) system. In [11] a video camera was used and a manual count of the flow observations was made. INRIX [4], which provides traffic forecasts, utilizes hundreds of sources including GPS-enabled vehicles, mobile devices and road sensors. Microsoft's Research Clearflow project, which followed the JamBayes project, [1] employs GPS data collected for over 5 years. This data was collected from volunteers, buses and paratransit vehicles. IBM Smarter Traveler traffic prediction tool [3] employs GPS monitoring and sensors located on the roads. Another IBM traffic prediction tool [2] uses real-time transportation data originated by people in cars, on trains and on buses that are moving in the city.

4 Forecasting Techniques

4.1 Bayesian networks

The conception and methodology of Bayesian networks is addressed in [48]. The advantages of the proposed model include an ability to create a Bayesian network to the traffic flow on a given road link in a given time, an ability to cope with incomplete data and good experimental results. For situations of incidents and accidents, it is mentioned in [48] that their approach might not be optimal and approaches for detecting abnormality might be applied. This work also uses GMM to depict the joint probability distribution between the data utilized for forecasting and the data to be forecasted. The CEM algorithm is used to estimate the parameters of GMM.

[9] estimates traffic flow, including origin-destination flow and link flow. It also addresses estimation updates upon receiving new information. The technique for traffic prediction presented in this work assumes that the origin-destination and link flows are capable of reproducing its real behavior and to construct a Gaussian Bayesian network employing the special characteristics of the variables of the traffic flow. This work states that Bayesian networks are powerful methods for the purpose of predicting traffic flows. A Bayesian network is also constructed in [21].

[49] addresses a spatio-temporal Bayesian network predictor. For the forecasting, the method employs all the spatial and temporal information from the traffic network. In order to select the most relevant input parameters the authors use the Pearson correlation coefficient. This work employs GMM and the CEM algorithm in order to approximate the joint probability distribution of the nodes in the Bayesian network (cause and effect nodes).

4.2 Simulation models

[14] uses on-line simulations into which a real-world data is incorporated. These simulations are based on a microscopic traffic flow model and can be done in real-time.

4.3 ATHENA model

[32] is based on statistical analysis of traffic volume profiles. This work presents a new method for traffic volume forecasting. A comparison of this model to a few other models will be described in section 4.6.

4.4 Neural networks

Some works address the use of NNs for forecasting. [19] discusses an object-oriented NN model that was developed for predicting short-term traffic conditions. This work, in contrast to some previous works, uses a dynamic NN architecture. In addition, due to the object-oriented approach, it is possible to model complex networks with a mixture of learning rules and processing element interactions. This modeling is difficult when using conventional NN paradigms. [25] depicts traffic volume forecasting with the help of decomposition and NN models. It also discusses updating the models in real time and the difference between multitvariate and univariate time series models and conclude that both multitvariate and univariate time series models reach the same accuracy level. The forecasting accuracy, by this work, is also not improved by increasing the number of input sources.

[22] describes the results of using a backpropagation NN algorithm in order to predict surface streets traffic flow in urban areas. NN methods that are based on a back-propagation algorithm have showed an ability to cope with complicated nonlinear predictions in several areas in previous studies. The model in this article uses a two-phase learning method (the first phase works on regular data, while the second addresses special cases and adapt to current situation).

[20] talks about reducing the size of NNs. This aspect is important as there is a large amount of possible input parameters and thus some NNs become impractical for implementation. The major difficulty is selecting the input parameters to the NN. This work states that a performance improvement might be achieved and that speed predictions are problematic. In addition, some phenomena like a single car traveling too fast or slow distort the training set (some ideas for solutions are suggested).

[12] aims to describe the application and performance of an alternative NN algorithm. This algorithm involves ”sequential or dynamic learning” of the traffic flow process. It investigates the potential of neural neutral networks

for forecasting motor-way traffic both in normal and in incident conditions. Three networks were constructed. One dynamic network was trained by using all the data, another dynamic network was trained with data obtained around incidents only and "simple dynamic model", that also used data around incidents only. "Simple dynamic model" differs from the first two in its hidden units. While in the first two models, the hidden units are able to move when new data is processed, in the "simple dynamic model", once the first 5 hidden units are fixed, all the other hidden units have to distribute around them upon processing new data. "Simple dynamic model" performed better than the first two on different data sets that they were tested on.

[27] uses multitask learning based NNs. The advantage of this method is that it can improve the generalization of the network, thus increasing forecasting accuracy. In this work the Levenberg-Marquardt algorithm is selected to train the network. MTL that was used in this work, proved to be more accurate than STL and by this work, traditional NN approaches for traffic flow forecasting are usually STL.

4.5 Time-series models and ARIMA

According to [43], statistical time series models, for instance, ARIMA, try to depict the past by utilizing mathematical models and then employ this model to forecasting. The ARIMA model is not suited for missing information and data filling technique might be problematic as the situation is complex. [7] presents a parametric model for traffic condition short-term forecasts at a predefined location in the network. The forecast is only based on previous observations at the location of the forecast. This model is the seasonal autoregressive moving average process. It is noticed that this type of prediction model is not the only one needed in the next generation ITS. In addition, it is stated that fitted SARIMA models provide equations that can be used in order to produce single and multiple interval forecasts. Moreover, SARIMA models allow the use of Kalman filtering techniques for the self-tuning forecast model. It is also claimed that there is no theoretical interest in investigating high level nonlinear mapping approaches (for instance NNs). This claim is supported by comparison to forecasting produced by NN and based on a common data set.

[42] evaluates the performance of a subset of ARIMA modeling. The structure of the evaluated model is identical to the structure of ARIMA with fewer parameters in it. Better results were achieved by a subset of ARIMA used in the experiments in comparison to the full ARIMA model. A comparison was also made to the ES method model, FAR and SAR models and its performance was superior.

The focus in [46] is on producing time-series state space models that are flexible and explicitly multivariate. This method enables jointly considering data from different detectors, and is able to model a wide variety of univariate models (for instance, ARIMA). According to the observed results, it seems that different specifications are appropriate for different time periods. In addition it seems that the use of multivariate state space models may be successfully applied to an urban roadway system. The results were compared to those achieved from the ARIMA model and were found to be superior (in some cases of the prediction the state space model "collapsed" into an ARIMA model).

[33] develops time-series models for predicting future traffic volume on urban arterials. In order to assess the time-series models, the Box-Jenkins approach is employed. Within the several ARIMA models evaluated, the ARIMA model (0, 1, 1) proved to be the most suitable.

4.6 Combination of NN and ARIMA

The technique introduced in [52] uses a Kohonen self-organizing map, which is a type of NN, as an initial classifier that makes it easier to define classes. In addition, there is an individually tuned ARIMA model for each class. The introduced method is called KARIMA. The task separation made in this technique improves the forecasting and it preforms better than both a single ARIMA model and a backpropagation NN. The authors note that the ATHENA model outperformed ARIMA and a backpropagation NN. However, it is very complicated. Thus the experiments were made with models that are similar to ATHENA but contain different sub-components. The numerical performance eventually achieved by this work is comparable to the performance of ATHENA. The disadvantages of the proposed model include not yet validated long term robustness. In addition, over a long period of time, as conditions alternate, its performance will deteriorate. On the other hand, Kohonen map can be automatically retrained.

[10] compared and combined two models: ANN and ARIMA. In ANN events from the past are analyzed and patterns are concluded. With these patterns, forecasts are made. This work states that traditional architectures of ARIMA or ANN models assume previous patterns will continue into the future, and thus cannot be expected to give

good results if this assumption is not valid. Therefore, this work introduced a judgemental adjustment technique that affects correcting future events that are seldom and irregular. Experiments justified that the judgemental adjustment technique helped to reduce the prediction error. Additionally, the combination of basic ANN model and ARIMA model, performed better then the models separately. In this work, in contrary to other works, the ARIMA model performed better than ANN.

4.7 Kalman filtering theory

In [39], two models for predicting short-term traffic volume are presented. These models utilize the Kalman filtering theory. By [39], Kalman filtering is applicable to short term stationary or nonstationary stochastic phenomena and it yields good traffic prediction accuracy. The models were compared with UTCS-2 and performed better. A dynamic state-space model is shown in [26]. It is noted that the state-space model has several features that make it suitable for traffic network applications. Dynamic models predict the traffic flow by utilizing on-line roadside measurements of traffic flow. Kalman filter is placed on the top of the state-space model.

[53] presents a dynamic state-space model. It also uses the Kalman filter, while the state-space model, a multivariate stochastic process, is underlying the filter. The state-space model has several advantages including being random, having an explicit measurement error and its simple, and flexible structure, which is very profitable. The predictor is compared to naive predictor and the results are depicted.

4.8 Markov Chain model

[56] models the traffic flow as a high order Markov Chain. The method employs current and recent values of the traffic flow and describes the future value. This future value is described by transition probability which is approximated by the GMM whose parameters are acquired by the EM algorithm. As it is assumed that the predicted state has a probability distribution and both current state and most recent states determine the next state, Markov Chain seems to be very suitable. Moreover, experiments showed that this technique performs well.

4.9 Historical average

Historical average, by [43], performs the forecasting based only on the past values. Thus its ability to respond to unanticipated events and incidents is low. However, it is easy for implementation and it is a fast working model.

4.10 Nonparametric regression

Nonparametric regression according to [43] tries to find past events that had input values identical to the current state of the system (when the prediction is made). In order to forecast short term traffic flows [58] uses a K-NN based nonparametric regression. It also performs analysis of the effect caused by key factors' settings in the model and makes suggestions about defining the state vector and choosing the forecasting method. Nonparametric regression, by this work, is an accurate method that has a strong transplant ability, while transcendental knowledge is not necessary in its data mining. In addition, its results are more accurate in unconventional road conditions. A comparison is made to BP NN and NPR is proven to be generally better.

[18] also suggests the k-NN method. For linear processes, this method is asymptotically as good as linear forecasters. However, it preforms worse in comparison to them when nonlinearities appear. The presented method performed comparably, but not clearly better than a simple univariate linear time-series forecasts in a comparison made in this study. This work suggests that larger databases may have a good impact on the method's accuracy and suspects that thin learning set used here made the model show no advantage in forecasting nonlinear transitions.

In [43], the following two models were discussed: back-propagation NN and nonparametric regression models, and compared to historical average and time-series models. For this purpose the four mentioned models were developed. Nonparametric regression gave the best performance. Moreover, its implementation was easy, and it is portable.

A nonparametric model based on using the kernel smoother for the autoregression function is presented in [36] for short-term traffic flow prediction. This model utilizes functional estimation techniques.

[47] proposes a local linear regression model for making short-term traffic prediction. Local linear regression can avoid boundary and cluster effects, and in the same time handle a wide range of data distributions. Local modeling methods are included in the nonparametric methods. In contrary to global models, local models are capable of learning locally. Simple local models can be used as building blocks to approximate a complex global function. The advantage of local models is that it can avoid negative interference exhibited by global models. Moreover, local models are real-time models. There is no need for reclassification of the data as these models can learn the functions from the infinite stream of training data. In addition, incorporation of exogenous variables to covariates is made easily with local models. Easy integration of traffic incidents, special events and other random factors can be made in the nonparametric prediction model. They can be integrated as exogenous variables if they are available. The method presented in [47] is compared to nonparametric approaches which are based on local constant regression (for instance the K-NN and kernel methods) and this work shows that the proposed method performs better.

4.11 Linear regression

[40] employs linear regression. This method is appropriate as the authors concluded that there is a linear relationship between any future travel time and the current status travel time. In addition this method is suitable for real time (the time consuming parts may be done offline).

4.12 Type-2 fuzzy logic

[29] proposes a method to forecast traffic that is base on type-2 FL. It produces prediction intervals instead of single forecast value. Moreover, as a by-product, this model yields time-dependent prediction intervals for forecasts in contrary to many other models that provide the user with a single forecast. The advantage of type-2 FL is its handling of uncertainties (including uncertainties in measurements and in the data used for parameters calibration). It is noted in the article that some simplifications were made and there is yet room for improvement and further research of this method.

4.13 Macroscopic model

[30] uses a macroscopic UTN model for traffic flow forecasting. In comparison to a microscopic model, the macroscopic model is less complex computationally. Thus, it is more suitable for short-term traffic forecast in real time. Additional advantage is that this model can cope well with a sudden huge change of the input traffic flow of the network. However, for a practical UTN a lot of parameters need to be fixed. Moreover, it is highly dependant on several factors, for example on detectors' precision and on future flow input. In the simulations, the real traffic is simulated by a microscopic model, while the forecasting is made by the macroscopic UTN model.

4.14 Hybrid methods

Some works combine or explore separately several techniques. [16] addresses different traffic prediction methods and states that there seems to be no "best technique". Thus, future work should consider combining different techniques. [50] creates 3 time series (daily similarity-s1, weekly similarity-s2 and hourly time series-s3), each time series is treated by another model (MA for s1, ES for s2 and SARIMA model for s3). Afterwards the data is aggregated based on NN and a prediction is made (based on the NN model). The prediction of the different models is also used as an input to the aggregation and NN phase. It is noted in this work that several works have shown that combining few models often results in a better prediction accuracy. MA and ES have the advantage of quick low cost updates and accurate short-term prediction. On the other hand, their disadvantage is not coping well with trend or seasonality. ARIMA's disadvantage is that this model is preassumed to be linear, while NN are able to cope well with nonlinear modeling. This work attempts to use the strength of each method. Data aggregation enables to combine information from different models. However, its disadvantage is that accidents and nonrecurring congestion affects its accuracy. The possible solution proposed in this work is subject to additional study.

[55] attempts to predict the traffic flows in an urban street network by developing and using a FNM. Two models construct the FNM: GN and EN. GN used the fuzzy approach while classifying the input data of similar characteristics

Method	Advantages	Disadvantages
MPL	Predicts well	The structure of the network must be defined a priori. In addition, it has only final synaptic weights for obtaining the outcome in parallel. Moreover, the resulting prediction equations are complex and have long expressions.
LGP	LGP-based equations are simple, straightforward and can easily be manipulated in practical circumstances. There is a high level of interactivity between the user and the methodology. Also, LGP can differentiate relevant input data and irrelevant input data. It gives similar prediction performance as MPL.	
FL	Has an explicit nature. The rules are written verbally.	A better understanding of the nature of the derived relationship between the interrelated input and output data is not provided, as there is no definite function to calculate the outcome using the input values. Concluding the fuzzy rules in this method is difficult.

Table 3: Advantages and disadvantages of the MPL, LPG and FL models by [57]

into clusters while the EN defines the input-output relationship in each cluster. Its advantage in comparison to the NN model is reduced training time. In addition, FNM was compared to a conventional NN model and performed better while converging much faster. This work also proposes an online rolling training procedure that is suitable for the proposed model due to its fast convergence. This training procedure improves the prediction quality.

[23] focuses on a quantificational method of dynamic factors from the perspective of nearness for traffic forecasting. It states that the dynamic forecasting model improvement should focus on the periodical changes in traffic flow while considering date types and characteristic times, as well as on the spatio-temporal law of traffic flow and on various external influences (for example, temperature, sunlight and visibility). This work also notices that ANNs can take into account some of the factors (e.g. visibility, sunlight and date type). The EKF-ANN algorithm is used here for traffic volume dynamic forecasting. This is because the EKF training algorithm is suitable for rapid calculation and to selecting dynamic samples.

[57] suggests the use of three techniques from the field of computational intelligence: LGP, MLP and FL. MLP is one of the most widely used ANN architectures. By utilizing these techniques, six different prediction models were developed. In the tests performed in the article for 5-minute and 30- minute intervals, the LGP and MLP models performed much better than the FL prediction models. The advantages and disadvantages of the proposed methods are summarized in Table 3. The disadvantages presented about MLP and FL limit their use by researchers. LGP behaves better and is clearly preferable in comparison with MLP and FL.

4.15 Other methods

[13] tests various VAR models and compares them to traditional models. In this work the authors notice that traffic data is correlated across space and time. One variable may be relevant to predicting another variable. It proves

that there were significant cross correlations among different traffic variables collected across very close locations at different time scales. Therefore, this work emphasizes the multivariate prediction of traffic variables. Some analysis are not possible in simple univariate formulations while they are possible in the VAR models. Another advantage of VAR estimation is its ability to estimate simultaneously several dependent variables (speeds, volumes, occupancies). In addition, it can provide a framework for estimating the influence of one variable on other variables. However, the accuracy of multivariate models at time scale larger than 15 minutes is questionable. VAR models in this work performed better than univariate ARIMA and SARIMA in 5 and 10 minutes prediction.

[11] explores the statistical nature of traffic flows aggregated at short time intervals and examines the potential of applying a DGLM to predicting such flows and providing prediction limit (it notes that other studies addressed only the prediction accuracy while ignoring the accuracy of the prediction limits). 3 methods are proposed for prediction bounds generating: Bayesian Predictive Distribution, F&S Asymmetric Error Distribution and Kalman Filter. It describes 3 prediction models based on: the Poisson distribution assumption, the negative binomial and the binomial distributions.

[41] mainly investigates MPC and more generally model-based control methods. Centralized MPC drawback is that its on-line computational complexity is significantly enlarged as the network becomes larger. Thus, increasing the prediction time and complicating the traffic model. [41] attempts to cope with this problem by utilizing a simplified traffic model. A simplified macroscopic model was established in this research. On the one hand, its advantages are reduced computing time, flexibility (it can handle different traffic scenarios and it has different cycle times). On the other hand, the accuracy is reduced (nevertheless, according to the simulation, the accuracy is rather sufficient for UTN control model). Additional methods that were not covered in this survey, are covered in [16]. In addition, this works divided the methods into three classes of Naive, Parametric and Non-Parametric methods.

4.16 Comparative analyses

Some works conduct a comparative analysis of several methods and depict the results.

[24] evaluates several short-term congestion prediction methods, namely: Multi-Linear Regression, Time Series Analysis (more specifically ARIMA), MLPs which are NNs described in this work, RBF networks, which are also NN that have different network construction in comparison to MLPs, Self-Organising Maps which are a type of NN and fuzzy logic. Eventually the Self-Organising Maps was not presented in the comparison due to the large error productions. The conclusion about the remaining 5 methods, was that ARIMA, MLP, RBF and fuzzy logic gave almost similar performance, which was better than the performance of the MLR. The authors note that if they had to select a winner, RBP would be selected.

[28] describes Multiple regression, ARIMA, Kalman filtering technique and NN and compares their results in predicting short term travel speeds. The authors note that Kalman filtering technique and NN performed best. It is also noticed that the Kalman filtering results are superior to those of the NN. Nevertheless, the transferability of the NN is expected to be better than the transferability of the Kalman filtering technique.

[38] compares the predictive ability of 5 methods: linear regression, NNs, regression trees, k-NN, and locally-weighted regression. The concluded result is that linear regression gave very accurate results. Moreover, its memory resource consumption and computational power make it suitable for real-world applications. It is important to note that this work considers a univariate short-term prediction of road travel times and the data was taken from a single road segment. Thus the authors note that a more complicated relationship may appear between few road segments causing to non-linear methods to perform better when making multivariate prediction.

[15] states that NNs, as opposed to conventional time series analysis techniques, do not require the model to be pre-specified. In addition, it is claimed that while the ARIMA models have a more explicit structure, the development and incorporation into local code of NN models can be easier. The statistical models checked in this work include linear regression, discounted linear regression, classical Box-Jenkins techniques (ARIMA) and transfer functions. Among these models, ARIMA generally yielded the best results.

4.17 Comparison of parametric and nonparametric models for traffic flow forecasting

By [44], SARIMA is a parametric approach that statistically seems to yield better results than non-parametric. ARIMA models are based on stochastic system theory. [44] claims that the disadvantage of SARIMA models is that incorporating this model into a production system is time consuming and requires a profound expertise. Nonparametric regression forecasting, on the other hand, is founded on chaotic system theory. Chaotic systems are defined by state transitions that are deterministic and non-linear. These state transitions are also ergodic and not periodic. It was previously reported that traffic flow shows chaotic behavior. Some of the implementation challenges of nonparametric regression include defining the distance metric for the nearness of historical events and current conditions, for a collection of nearest neighbors selecting a forecast generation method, etc. Theoretical foundation of nonparametric regression are a subject for examination in this research. Furthermore, it attempts to answer if predicting traffic flow of a single interval by using nonparametric regression (which is based on forecast generation methods improved by heuristics) approaches the performance of SARIMA models. An improvement was achieved in the performance of nonparametric regression by the heuristic forecast generation methods examined in this research. However, the achieved performance is yet not as good as the performance of SARIMA models. According to the authors, this strongly implies that traffic condition data is characteristically stochastic, and not chaotic. Nonparametric regression is applicable when ARIMA can not be implemented, as in this case they are preferable to naive forecast methods.

[8] also compares between the approaches. In this work, forecasts, comparable in their accuracy to SARIMA forecasts, are achieved by nonparametric regression based on a k-NN specification with adjusted and weighted output elements. [47] points that the basic difference between the two approaches is the assumption of the parametric approach that the model takes a form that can be represented by a finite number of parameters, while the nonparametric approach does not make such an assumption. Both [44] and [47] note that nonparametric methods are highly dependent on the size of the database. Larger databases enable these methods to have more patterns, thus making more accurate predictions.

5 Forecasting System Architecture

A distributed approach in which the central server performs various data mining tasks on historical data and sends the interesting patterns to the sensors is presented in [31]. The sensors in this work monitor and predict the coming traffic or raise alarms independently by comparing with the patterns observed in historical streams. These alarms about abnormal events are sent to other sensors in the relevant cluster. The amount of patterns stored on each sensor is rather small. The communication cost is also not high as there is a communication between the central server and a sensor only when an incorrect prediction is made.

[34], which does not deal with traffic prediction, presents TrafficSense for monitoring city road and traffic conditions (especially detecting bumps and potholes, braking, and honking). The work mainly focuses on sensing data performed by employing Smartphones and the sensors available on it. The data is afterwards sent to the server for aggregation. Smartphones enable monitoring without creating a specialized infrastructure and this work builds upon the growing ubiquity of mobile phones. The aspect of power consuming is also addressed. For instance, before sending data to the server, it is processed locally. In addition, there is an option to utilize the sensors in tendem. Monitoring is suitable also for roads in the developing world that has some additional constraints as road quality and chaotic traffic. In [21] there is a web service in which users can see the predictions made. It specifies when a relief is expected in congested roads, and on the other hand which roads are now free and will become congested. In addition, for mobile phone users it is noted that the system should be such that they can glance in quickly. The user can get alerts which he defines (for instance, when the route is likely to become congested in 30 minutes). This work also predicts situations that might surprise the user.

Mobile Millennium [5] is another research project that includes a pilot traffic-monitoring system that uses the GPS in cellular phones to gather traffic information, process it, and distribute it back to the phones in real time.

6 Output of the Process

Most works provide a prediction for up to an hour. [30, 48] provide a short term forecasting (determines the traffic in the next 5 minutes to half an hour). [19] predicts up to 15 minutes into the future (5 minutes into the future are predicted with high accuracy). [23] in its simulation checks short term prediction. In [57] 5-minute and 30-minute traffic flow predictions are made, while the performance of the 30-minutes predictions is better than the 5-minute prediction. In [22] on the other hand, the 5 minutes prediction is better than the 30 minutes prediction (as the 5 minutes prediction has a stronger correlation to the input data). The prediction ranges are 1,5,15 and 30 minutes ahead in [12] and eventually the 15 and 30 minutes forecasts were selected. [43] predicts 15 minutes into the future. The results were assessed also in case of incidents (and the model was trained with information prior to, during and after incidents). [52] forecasts half an hour and an hour. Some works provide a longer term prediction. In [27] the network is trained by using the first 22 days' data to forecast the traffic flow of later 3 days. INRIX [4] predicts the future traffic up to a year ahead and works with freeways, highways, and secondary roadways (for instance, arterials and side streets). [47] notes that the proposed method performance is decreased when the prediction horizon increases. The output provided by the model also differs. For instance, [2] provides volume and speed of traffic in the near future, while [40] predicts the time that travelling between two points will consume (when the departure time is in the future).

7 Insights about Traffic Prediction

[6] answers the question of the predictability of traffic in urban areas, and if it is predictable, at what levels. The answer provided by this work is that traffic does not vary during most months of the year and during weekdays and in this sense it is predictable. This work investigates the variation in different time resolutions (yearly, monthly, daily and time variations), in addition to spatial variation. It also checks the distributions in traffic flows, as this aspect is important in traffic simulations. It was generally noticed that urban area traffic flows are not normally distributed. However, there were some exceptions. [45] shows the need for parsimonious multivariate state-space models for short-term flow forecasting. It concludes that if considering data flows from different detectors together, much more information can be obtained, in comparison to separately considering each detector. For finding possible common cyclical components between two successive loop detectors, this work employed spectral analysis.

In [37] spectral analysis is also employed. It is used for time series analysis based on components of pre-computed eigenvector and current online data. This method is influenced strongly by the accuracy of the input data. The authors note that most of the errors are due to problems in data collection. The main drawback of the presented method is its lack of ability to account for fast unpredicted changes that are not represented in previous data (changes that are not seen in the covariance matrix). [7] notices that for time intervals exceeding some threshold (for example 15 minutes) univariate short-term predictions will be important and will give the most accurate forecast as when working with longer time intervals it will not be possible to theoretically establish and model stable correlation with other detection locations within the instrumented network.

8 Conclusions and Further Work

Traffic prediction, being an important aspect of ITS, was studies extensively and various techniques and improvements over them were proposed. However, there are yet some improvements and further work that can be done. [14] mentions that the driver reactions should be taken into account (this aspect has been addressed in few works, but these works are not about traffic forecasting). A large share of the works consider or test their work on freeways, while [23] states that the achieved prediction results can be used for controlling intersections, freeways, tunnels and parking lots. By [46] it seems that traffic flow in signalized urban arterials cannot be predicted, at least in the short-run, as accurately as flow in urban freeways. In addition, it is mentioned in this work that there is a necessity to develop new approaches that are able to capture boundary conditions of traffic behaviour. [12] notices that the technique they presented might be implemented using on-line data. Thus, the convergence time of different models should be considered. Another suggested improvement in [12] is to adjust the model to a potential end user instead of using a purely statistically base function. [13] states that additional research of VAR models in higher time scales (30-60 minutes) is needed.

Among the many different techniques proposed for traffic prediction, it is rather difficult to determine a single method that clearly outperforms the rest. First, the methods are tested with specific data and a comparison to only a small number of techniques is made. Second, there are some contradicting results about one's method superiority over the other (for instance, ARIMA model performed better than ANN in [10], while [10] states that other works concluded the opposite). [16] suggests that there is no technique that clearly outperforms the other techniques. Thus future works should consider combinations of the techniques or selecting a technique according to some circumstances. Indeed [50], which presents a method that combine several different techniques and is depicted more profoundly in 4.14, notes that several works have shown that combining few models often results in a better prediction accuracy. This work also employs the NN method which, according to [16], was addressed in tens of works and possess the ability to model dynamic and non-linear processes well.

References

[1] Clearflow project. http://research.microsoft.com/en-us/projects/clearflow/default.aspx.

[2] Ibm smarter traffic. http://www.ibm.com/smarterplanet/us/en/traffic_congestion/nextsteps/solution/N500945X17585D04.html.

[3] Ibm smarter traveler. http://www.engadget.com/2011/04/13/ibm-shows-off-smarter-traveler-traffic-prediction-tool/.

[4] Inrix. http://www.inrix.com/default.asp.

[5] Mobile millennium. http://traffic.berkeley.edu/.

[6] M. Karlaftis A. Stathopoulos. Temporal and spatial variations of real-time traffic data in urban areas. *Transportation Research Record: Journal of the Transportation Research Board*, pages 135–140, 2007.

[7] M.ASCE Billy M. Williams and F.ASCE Lester A. Hoel. Modeling and forecasting vehicular traffic flow as a seasonal arima process: Theoretical basis and empirical results. *Journal of Transportation Engineering*, 129(6):664–672, November 2003.

[8] R.K. Oswald B.L. Smith, B.M. Williams. Parametric and nonparametric traffic volume forecasting. Paper Presented at the 2000 Transportation Research Board Annual Meeting, Washington, DC, 2000.

[9] Enrique Castillo, Jos Mara Menndez, and Santos Snchez-Cambronero. Predicting traffic flow using bayesian networks. *Transportation Research Part B: Methodological*, 42(5):482 – 509, 2008.

[10] S.C. Chang, R.S. Kim, S.J. Kim, and M.H. Ahn. Traffic-flow forecasting using a 3-stage model. In *Intelligent Vehicles Symposium, 2000. IV 2000. Proceedings of the IEEE*, pages 451 –456, 2000.

[11] Shaw-Pin Miaou Chang-Jen Lan. Real-time prediction of traffic flows using dynamic generalized linear models. *Transportation Research Record: Journal of the Transportation Research Board*, pages 168–178, 2007.

[12] Haibo Chen and Susan Grant-Muller. Use of sequential learning for short-term traffic flow forecasting. *Transportation Research Part C: Emerging Technologies*, 9(5):319 – 336, 2001.

[13] Srinivasa Ravi Chandra Chilakamarri Venkata. *SPATIO-TEMPORAL ANALYSES FOR PREDICTION OF TRAFFIC FLOW, SPEED AND OCCUPANCY ON I-4*. PhD thesis, University of Central Florida, 2009.

[14] R. Chrobok, J. Wahle, and M. Schreckenberg. Traffic forecast using simulations of large scale networks. In *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE*, pages 434 –439, 2001.

[15] S D CLARK, M S Dougherty, and H R KIRBY. The use of neural networks and time series models for short term traffic forecasting: A comparative study. pages 151–62, 1993.

[16] F.M. Sander C.P.IJ. van Hinsbergen, J.W.C. van Lint. Short term traffic prediction models. ITS World Congress, Beijing, China, 2007.

[17] G. A. Davis, N. L. Nihan, M. M. Hamed, , and L. N. Jacobson. Adaptive forecasting of freeway traffic congestion. *Transportation Research Record*, 1287:29–33, 1990.

[18] Gary A. Davis and Nancy L. Nihan. Nonparametric regression and short-term freeway traffic forecasting. *Journal of Transportation Engineering*, 117(2):178–188, 1991.

[19] Hussein Dia. An object-oriented neural network approach to short-term traffic forecasting. *European Journal of Operational Research*, 131(2):253 – 261, 2001.

[20] Mark S. Dougherty and Mark R. Cobbett. Short-term inter-urban traffic forecasts using neural networks. *International Journal of Forecasting*, 13(1):21 – 31, 1997.

[21] R. Sarin E. Horvitz, J. Apacible and L. Liao. Prediction, expectation, and surprise: Methods, designs, and study of a deployed traffic forecasting service. In *Twenty-First Conference on Uncertainty in Artificial Intelligence, UAI-2005.*

[22] J. Gilmore and N. Abe. A neural networks system for traffic congestion forecasting. *IEEE Proceedings of International Joint Conference on Neural Networks*, pages 2025–2028, 1993.

[23] Yan hong Tang and Bao Xi. Dynamic forecasting of traffic volume based on quantificational dynamics: A nearness perspective. *Scientific Research and Essays*, 5(4):389–394, 2010.

[24] Giovanni Huisken and Martin van Maarseveen. Congestion prediction on motorways: A comparative analysis. *Department of Civil Engineering & Management, University of Twente*, 2000.

[25] Chen J. *Characterization and implementation of neural network time series models for traffic volume forecasting*. Dissertation for the doctor of philosophy degree in engineering science, The University of Toledo, 1997.

[26] K. Lindveld J. Whittaker, S. Garside. Tracking and predicting a network traffic process. *International Journal of Forecasting*, pages 51–61, 1997.

[27] Feng Jin and Shiliang Sun. Neural network multitask learning for traffic flow forecasting. In *Neural Networks, 2008. IJCNN 2008. (IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on*, pages 1897 –1901, june 2008.

[28] Seungjae Lee, Daehyon Kim, Juyoung Kim, and Bumchul Cho. Comparison of models for predictin. short-term travel speeds. 1998.

[29] L. Li, W.-H. Lin, and H. Liu. Type-2 fuzzy logic approach for short-term traffic forecasting. *Intelligent Transport Systems, IEE Proceedings*, 153(1):33 – 40, march 2006.

[30] Shu Lin, Yugeng Xi, and Yanfei Yang. Short-term traffic flow forecasting using macroscopic urban traffic network model. In *Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on*, pages 134 –138, oct. 2008.

[31] Ying Liu, Alok Choudhary, Jianhong Zhou, and Ashfaq Khokhar. A scalable distributed stream mining system for highway traffic data. In Johannes Frnkranz, Tobias Scheffer, and Myra Spiliopoulou, editors, *Knowledge Discovery in Databases: PKDD 2006*, volume 4213 of *Lecture Notes in Computer Science*, pages 309–321. Springer Berlin / Heidelberg, 2006.

[32] M. Aron M. Danech-Pajouh. Athena: a method for short-term inter-urban motorway traffic forecasting. *Recherche Transports Securite*, pages 11–6, 1991.

[33] Z.M.B. Said M.M. Hamed, H.R. Al-Masaeid. Short-term prediction of traffic volume in urban arterials. *Journal of Transportation Engineering*, pages 249–254, 1995.

[34] Prashanth Mohan, Venkata N. Padmanabhan, and Ramachandran Ramjee. Trafficsense: Rich monitoring of road and traffic conditions using mobile smartphones. Technical report, Microsoft Research India, Bangalore, April 2008.

[35] C. K. Moorthy and B. G. Ratcliffe. Short term traffic forecasting using time series methods. *Transportation Planning and Technology*, 12(1):45–56, 1988.

[36] ELFAOUZI N.-E. Nonparametric traffic flow prediction using kernel estimator. pages 41–54, 1996.

[37] H. Nicholson and C.D. Swann. The prediction of traffic flow volumes based on spectral analysis. *Transportation Research*, 8(6):533 – 538, 1974.

[38] D. Nikovski, N. Nishiuma, Y. Goto, and H. Kumazawa. Univariate short-term prediction of road travel times. In *Intelligent Transportation Systems, 2005. Proceedings. 2005 IEEE*, pages 1074 – 1079, sept. 2005.

[39] Iwao Okutani and Yorgos J. Stephanedes. Dynamic prediction of traffic volume through kalman filtering theory. *Transportation Research Part B: Methodological*, 18(1):1 – 11, 1984.

[40] J. Rice and E. van Zwet. A simple and effective method for predicting travel times on freeways. *Intelligent Transportation Systems, IEEE Transactions on*, 5(3):200 – 207, sept. 2004.

[41] Y. Xi S. Lin, B. De Schutter and J. Hellendoorn. A simplified macroscopic urban traffic network model for model-based predictive control. In *Proceedings of the 12th IFAC Symposium on Transportation Systems, Redondo Beach, California*, pages 286–291, September 2009.

[42] Daniel B. Fambro Sangsoo Lee. Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting. *Transportation Research Record: Journal of the Transportation Research Board*, pages 179–188, 2007.

[43] B. L. Smith and M. J. Demetsky. Traffic flow forecasting: Comparison of modeling approaches. *ASCE Journal of Transportation Engineering*, 123(4):261–266, 1997.

[44] Brian L Smith, Billy M Williams, and R Keith Oswald. Comparison of parametric and nonparametric models for traffic flow forecasting. *Transportation Research Part C: Emerging Technologies*, 10(4):303 – 321, 2002.

[45] A. Stathopoulos and M.G. Karlaftis. Spectral and cross-spectral analysis of urban traffic flows. In *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE*, pages 820 –825, 2001.

[46] Anthony Stathopoulos and Matthew G. Karlaftis. A multivariate state space approach for urban traffic flow modeling and prediction. *Transportation Research Part C: Emerging Technologies*, 11(2):121 – 135, 2003.

[47] Hongyu Sun, Henry X. Liu, Heng Xiao, Rachel R. He, and Bin Ran. Use of local linear regression model for short-term traffic forecasting. *Transportation Research Record: Journal of the Transportation Research Board Publisher*, 1836 / 2003:143–150, 2007.

[48] Shiliang Sun, Changshui Zhang, and Guoqiang Yu. A bayesian network approach to traffic flow forecasting. *Intelligent Transportation Systems, IEEE Transactions on*, 7(1):124 –132, march 2006.

[49] Shiliang Sun, Changshui Zhang, and Yi Zhang. Traffic flow forecasting using a spatio-temporal bayesian network predictor. In Wlodzislaw Duch, Janusz Kacprzyk, Erkki Oja, and Slawomir Zadrozny, editors, *Artificial Neural Networks: Formal Models and Their Applications ICANN 2005*, volume 3697 of *Lecture Notes in Computer Science*, pages 752–752. Springer Berlin / Heidelberg, 2005.

[50] Man-Chun Tan, S.C. Wong, Jian-Min Xu, Zhan-Rong Guan, and Peng Zhang. An aggregation approach to short-term traffic flow prediction. *Intelligent Transportation Systems, IEEE Transactions on*, 10(1):60 –69, march 2009.

[51] Eleni I. Vlahogianni, John C. Golias, and Matthew G. Karlaftis. Short-term traffic forecasting:overview of objectives and methods. *Transport Reviews*, 2004.

[52] Mascha Van Der Voort, Mark Dougherty, and Susan Watson. Combining kohonen maps with arima time series models to forecast traffic flow. *Transportation Research Part C: Emerging Technologies*, 4(5):307 – 318, 1996.

[53] Joe Whittaker, Simon Garside, and Karel Lindveld. Tracking and predicting a network traffic process. *International Journal of Forecasting*, 13(1):51 – 61, 1997.

[54] Billy M. Williams and Lester A. Hoel. Modeling and forecasting vehicular traffic flow as a seasonal arima process: Theoretical basis and empirical results. *Journal of Transportation Engineering*, 129(6):664–662, 2003.

[55] Hongbin Yin, S.C. Wong, Jianmin Xu, and C.K. Wong. Urban traffic flow prediction using a fuzzy-neural approach. *Transportation Research Part C: Emerging Technologies*, 10(2):85 – 98, 2002.

[56] Guoqiang Yu, Jianming Hu, Changshui Zhang, Like Zhuang, and Jingyan Song. Short-term traffic flow forecasting based on markov chain model. In *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*, pages 208 – 212, june 2003.

[57] Shahriar Afandizadeh Zargari, Salar Zabihi Siabil, Amir Hossein Alavi, and Amir Hossein Gandomi. A computational intelligence-based approach for short-term traffic flow prediction. *Expert Systems*, pages no–no, 2010.

[58] Tao Zhang, Lifang Hu, Zhixin Liu, and Yuejie Zhang. Nonparametric regression for the short-term traffic flow forecasting. In *Mechanic Automation and Control Engineering (MACE), 2010 International Conference on*, pages 2850 –2853, june 2010.