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 – 035

Analysis of the State-of-the-Art for Vehicular Traffic Prediction

D. Kehagias, T. Diamantopoulos

January 2013

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 – 035

Analysis of the State-of-the-Art for Vehicular Traffic Prediction

D. Kehagias, T. Diamantopoulos

January 2013

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

FP7 - Information and Communication Technologies

Grant Agreement No: 288094
Collaborative Project
Project start: 1 November 2011, Duration: 36 months

Analysis of the State-of-the-Art for Vehicular Traffic Prediction

Workpackage: WP2 - Algorithms for Vehicle Routing
Contributing Partners: CERTH

Nature: ☒ Report ☐ Prototype ☐ Demonstrator ☐ Other

Dissemination Level:
☒ PU: Public
☐ PP: Restricted to other programme participants (including the Commission Services)
☐ RE: Restricted to a group specified by the consortium (including the Commission Services)
☐ CO: Confidential, only for members of the consortium (including the Commission Services)

The eCOMPASS project (www.ecompass-project.eu) is funded by the European Commission, Information Society and Media Directorate General, Unit G4-ICT for Transport, under the FP7 Programme.

The eCOMPASS Consortium

Computer Technology Institute & Press 'Diophantus' (CTI) (coordinator), Greece

Centre for Research and Technology Hellas (CERTH), Greece

Eidgenössische Technische Hochschule Zürich (ETHZ), Switzerland

Karlsruher Institut fuer Technologie (KIT), Germany

TOMTOM INTERNATIONAL BV (TOMTOM), Netherlands

PTV PLANUNG TRANSPORT VERKEHR AG. (PTV), Germany

Contents

1 Traffic Prediction

Beyond any doubt, transportation plays a key role to the design and execution of the vast majority of the modern economic and social activities and as a result it significantly affects the perceived quality of the citizens' everyday life. Hence, the establishment of a reliable as well as time and cost efficient transportation system rises as a necessity of prodigious importance for the growth and prosperity of the contemporary societies.

In consequence, the research has focused on optimizing the allocation of the transportation network resources (roads' capacity) among the competing actuators, i.e. the vehicles that use the network or intend to use it in the imminent future. As an outcome, the notion of Intelligent Transportation Systems (ITS) [20] has been introduced as the integration of adequate technological infrastructure (data acquisition, storage, processing, transmission), Decision Support System (DSS) (modelling methodologies, routing and prediction algorithms) and interfaces (user equipment, traffic lights management, electronic information boards) that shall allow for the effective monitoring, assessment and management of the offered traffic load. In this context, the ultimate goal of ITS is to implement novel measures of both preventive as well as reactive nature against traffic congestion given a specific set of desirable performance criteria, such as travel time, cost and prioritization among different road segments.

However, despite the obvious merits from gathering the traffic data, having knowledge of the instantaneous or historical traffic conditions does not provide a solid ground for achieving optimal vehicle forwarding, since, especially in cases of intense traffic, the available data are bound to soon become obsolete. In contrast, for effectively performing the demanding task of traffic routing, it is required to obtain an as accurate as possible estimation of the network's forthcoming states. This information will allow the routing decisions to be based upon the network's status and requirements at the critical time that the managed vehicles are expected to reach the nodes/edges under investigation. As a result, the notion of *vehicular traffic prediction* has been soon developed as the process for forecasting the values of specific traffic intensity metrics at a given horizon in the future by taking into account the current and historical evolution of the traffic conditions in the transportation network of interest. It becomes apparent that efficient traffic prediction rises as a fundamental prerequisite for enhancing the performance of traffic management and succeeding in building fully exploitable ITSs with realistic applicability.

In this respect, within the overall ITS field, profound research interest has been concentrated in the specific area of traffic prediction, covering extensively all its various aspects and problematics. As a matter of fact, numerous related papers can be found in the bibliography [97, 103], while some of the introduced methodologies have also been applied and evaluated in real cases of transportation networks at least at an experimental level. A complete classification of the existing studies in traffic prediction on the basis of their fundamental characteristics is presented below, and an exhaustive list of the proposed techniques is also provided along with a short description of each contribution. Moreover, apart from the core traffic prediction algorithms, particular emphasis is laid upon the *map-matching* techniques, which form the basis for the adoption of the most modern traffic prediction methodologies that make virtue of the feedback from moving vehicles.

1.1 Taxonomy of traffic prediction techniques

There are several different criteria for categorizing the bulk of traffic prediction approaches. As a result, a multi-dimensional classification scheme is derived, depending on the selected set of the algorithms' corresponding predominant features. Traffic prediction technologies can be classified according to the following criteria:

- **Traffic Descriptors**. The term *traffic descriptor* refers to the metrics that are chosen to be implemented in order to quantify the performance of the transportation network in a manner that allows for

- Consistently monitoring the dynamics of the traffic evolution
- Evidently capturing the underlying causes of any abrupt variations or outliers of the traffic intensity
- Taking over the adequate measures for optimizing traffic forwarding and mitigating any traffic congestion phenomena

Furthermore, it should be noted that, although in the vast majority of the studies, the output of the traffic prediction algorithm is computed in the same metric as the input data, different traffic descriptors can be selected for the input and the output of the traffic prediction module. Further details are provided in Section 1.2.

- **Sources of Input Data**. This category regards the type and content of the information that is acquired about the conditions of the transportation network. Two separate types of data can be further isolated:
  - Quantitative real-time measurements of traffic descriptors. It refers to measurements of traffic intensity variables through the use of suitable technological means.
    * Loop detectors, aka induction loops. The vehicle detection loops are used to detect vehicles passing or arriving at a certain point, for instance approaching a traffic light. The presence of the vehicle's metal body causes the alteration of the inductance of an electrically conducting wire loop that is installed under the road surface in the pavement. The calculation of different metrics of the traffic intensity is facilitated with adequate data processing and transformation. Besides the straightforward traffic flow computation, speed can be also estimated [126, 88, 58].
    * Automatic vehicle identification systems, such as i) RFID and ii) Video cameras with image processing capabilities for license plate recognition. RFID detectors are being installed on many highways, especially for automated toll payment. As a RFID-enabled vehicle passes through two consecutive detection points, an estimate of the traffic density between the two points can be obtained by calculating the vehicles entering the section between the times of entrance and exit of the tagged vehicle, on the basis of its measured speed [27].
    * GPS-enabled devices, such as navigators and smartphones. These devices report on the location and the instantaneous speed of the vehicle, while enhanced versions also provide information of the vehicle's direction [2]. Depending on their operational purpose, the carrying vehicles are commonly referred to either as *prove cars* or as *floating cars*, while in some studies, these terms are also used interchangeably.
      · *Probe cars*, which are either appropriately equipped public-service vehicles (buses, taxis, police cars) that combine their normal functionality with providing traffic feedback or vehicles dedicated for traffic data collection (usually in short-term expenditures for experimental studies or pilot use cases)
      · *Floating cars*, which are private vehicles with commercial handheld user equipment whose owners have agreed to participate in traffic data collection either voluntarily or as part of an added value service [19].

    Particular emphasis must be laid on the fact that due to their profound proliferation, the GPS-enabled devices appear nowadays as the most promising source of traffic data, since they form a vast pool of real-time information with exhaustive temporal and spatial coverage. Moreover, considering the realistic implementation of such a data collection system of GPS-oriented traffic information, a key technical aspect regards the sampling methodology. In more detail, two sampling approaches are addressed:

    · Temporal. The GPS-devices transmit their feedback, consisting usually of timestamp, location coordinates and speed, periodically at specific time intervals [35]. Despite its simplicity from the equipment's point of view, this technique requires the deployment of efficient map-matching algorithms, in order to be able to accurately match the vehicle's location onto an actual road link and therefore to aggregate the total feedback from individual vehicles into statistics of specific links of the transportation network [82]. Given the favourable characteristics of GPS-oriented traffic data and the corresponding research and industrial trend towards their maximum exploitation, the optimization of the map-matching procedures arises as a rather complicated task of primary importance.
    · Spatial. The GPS-devices transmit their feedback when crossing predefined locations. To this aim, the devices are equipped with adequate software, capable of identifying the road link that the vehicle traverses and compare it with the list of predetermined observation points [39, 19]. Hence, the map-matching procedure is still executed but it is implemented locally at the user's side. Because the device can keep a short history of its trace without any privacy issues or significant storage overhead, the map-matching efficiency is augmented. Furthermore, the vehicle's location has only to be compared with a finitely limited set of road links or nodes and hence the processing overhead is dramatically reduced.
  – Semantic information. This data category entails information about qualitative attributes of the state of the transportation network that can substantially affect the traffic conditions at different spatiotemporal extent. The most common of these features are the weather conditions and the accidents occurrence. Castro-Neto et al. address the need for effectively incorporating semantic information into the traffic prediction process by diversifying their approach between typical and atypical conditions [9]. Furthermore, in [51] weather, holiday information and accident counts are utilized through a linear regression model for both traffic volume prediction and accident count prediction. Moreover, the authors in [22, 5] specifically study the exploitation of weather conditions in order to forecast significant traffic deviations.

- **Area of Implementation**. It concerns the type of the transportation network that the traffic prediction procedure is applied to. Two broad subcategories can be mainly identified:
  – Freeways/Highways. The majority of the existing studies refers to the case of freeways and highways, since the longer distances to be travelled and the large, yet slow, alterations in traffic intensity augment the precision of the forecasts, while at the same time make prediction more important for traffic management.
  – Urban roads. Traffic prediction for urban roads requires more complex approaches, due to the high density of the network, the interdependencies among neighboring road segments and the abrupt fluctuations in traffic conditions including great percentage of outliers.

- **Prediction Resolution**. It regards the time response of the prediction algorithm and it can be described by two parameters:
  – Prediction horizon. It is the extent of time ahead to which the forecast is referring. There exist methods that allow for prediction at multiple time windows ahead and/or at time windows of different width. Although the vast majority of the studies found in literature deal with the issue of short-term forecasting (horizon varying from 5 minutes to 1 hour), long-range trips in large areas, requiring long-term predictions in a large network are also considered [83].

  - Prediction step. It is the time interval upon which the forecasts are executed. It is closely connected to the time period of traffic data acquisition and aggregation.

- **Univariate vs. Multivariate**. These terms refer to the number of traffic descriptors that are utilized as the input (sources of data) of the prediction module or the number of traffic descriptors for which their forecasted values are provided as output. The term multivariate is also often used to describe prediction algorithms that incorporate into the prediction process measurements from multiple observation locations. Such methodologies are based upon the spatio-temporal analysis of the transportation network, i.e. in order to calculate the predicted values of a single traffic descriptor at a given location time series of data samples from different nodes/links of the network are taken into account.

- **Traffic Prediction Methodology**. It regards the core traffic prediction method and algorithm that is utilized. Due to its critical impact on the performance of the traffic prediction procedure as its core module, the traffic prediction methodology will be further scrutinized in Section 1.3.

1.2 Traffic Descriptors

According to the existing literature, the most broadly used traffic descriptors are:

- **Traffic Flow**. It is also referred to with the terms *traffic intensity* or *traffic volume* and it is defined as the number of vehicles passing through a point of observation per time unit. It is commonly measured in vehicles per hour.

- Traffic Density. It is defined as the number of vehicles per distance unit, i.e. the number of vehicles that are simultaneously travelling along a continuous road segment of known length. It is measured in vehicles per kilometer [60].

- **Occupancy**. One of the most common methods for acquiring traffic data is through the deployment of loop detectors, i.e. induction loops that are placed underneath the roads' surface at locations of particular interest so as to act as indicator of a vehicle's the presence. Occupancy is a measure of traffic stream concentration and is the percentage of time that the sensor is detecting vehicle presence, or, in other words, the percentage of time that the sensor is "on" [112].

- **Speed**. It is the average speed of all the vehicles that passed from a specific link/road during a certain time window. It is measured in kilometers per hour and it can be calculated either as the algebraic or the harmonic mean of the sample velocity values, depending on the exact implementation.

- **Travel Time**. It is defined as the necessary travel time between two fixed point along a highway, freeway or urban arterial [103]. Moreover, besides the aforementioned per link definition, travel time can be also defined on per aggregate route basis, i.e. total duration from start to destination. This later approach is mostly applicable in the case of Advanced Traveler's Information Systems (ATIS), such as navigators, where a travel time of reference exists, e.g. alternative routes to be evaluated, maximum acceptable travel time, historical information of the time performance of the route of interest. It must be underlined that, in comparison with all the other available traffic descriptors, travel time is the most easily and widely understood notion, since:

  - It is common for all transportation modes, e.g. car, bicycle, walking, bus, metro and thus it facilitates the integration of all the transportation means into a multi-modal routing framework.

- It expresses a basic physical quantity familiar to the human perception and intuition, i.e. how long will it take to move from one point to another or reach the final destination.
- Time is the fundamental variable behind well-established traffic analysis methods, such as the queuing theory [32]

Hence, it provides the most suitable interface for achieving the effective communication among all the transportation actuators, such as planners, engineers, administrator, users [95].

1.2.1 Selection of Traffic Descriptors

Based on the aforementioned definitions, it can be noted that there is a profound cross-correlation among the traffic variables. Nevertheless, despite their inherent inter-dependency, it is neither possible nor beneficial to consolidate the multiple traffic descriptors into one common metric. This desirable redundancy is dictated at first by the fact that the efficiency of each traffic descriptor in accommodating the monitoring and forecasting of the traffic conditions varies significantly as a function of the system's configuration, status and requirements. As a matter of fact, the several studies that dealt during the previous decades with the issue of assessing and comparing the performance of traffic prediction models based on different variables have reached contradictory results regarding the definition of the prevailing metric to be used in traffic prediction. A very early study shows traffic flow as the most stable traffic descriptor [57], while the authors in [63] argue in favour of occupancy. Finally, Dougherty and Cobbett showed an inefficiency of speed under conditions of rather high congestion [23]. Hence, the generic conclusion that can be deduced is that the selection of the optimum traffic descriptor heavily relies on the characteristics of the investigated scenario.

Moreover, the choice of the variable to be deployed for describing the traffic conditions is usually determined in a rather straightforward manner by the available sources of input data. The measured physical quantity is tightly connected to the type of technological means that are implemented for monitoring the traffic conditions. Hence, given a certain traffic monitoring infrastructure, the corresponding traffic descriptor is uniquely derived, at least for the input data of the traffic prediction algorithm.

In this respect, during the early years, research on traffic prediction was based predominantly on traffic flow analysis and secondarily on the exploitation of occupancy measurements, since this kind of information could be easily extracted from the technological means that were deployed for traffic monitoring at that period and which were merely restricted to induction loops. In contrast, time and speed calculations required the establishment of equipment for vehicle identification, such as cameras with image processing capabilities and RFID-enabled vehicles, in order to be able to track down the travel time and respectively the speed of a given sample-vehicle between two fixed points of reference. Therefore, given the technological status quo and the respective availability of test data, the bulk of the studies in traffic prediction, especially until the middle 00's, are primarily focused on traffic flow and occupancy [13, 117, 86, 90, 91, 112, 47, 102, 125], while substantial research work on the basis of traffic flow is still performed [29, 92, 9, 59, 10, 93, 11].

However, the public availability of GPS services for civilian applications and the respective broad proliferation of GPS-enabled handheld devices (smartphones, navigators) with high processing capabilities, along with the establishment of efficient mobile data communications, allowed for the collection of vast amounts of real-time speed traces. Additionally, the advent of Assisted-GPS allowed for the ubiquitous acquisition of location information of high accuracy even under conditions of absence of Line-of-Sight connection with the satellites, e.g. in urban territories of dense building environment. In consequence, along the most recent years research on traffic prediction strived towards speed-oriented analysis, in order to take advantage of the enhanced potentials provided by the new technological era in the field of traffic conditions monitoring and specifically in the field of collecting speed feedback from numerous vehicles with wide geographic distribution. For example *Mobile Century* is a traffic monitoring system based on the collection of speed data from GPS-enabled smartphones [37]. The extensive mobile coverage provided by the existing cellular network infrastructure is exploited for acquiring the position and velocity measurements of high accuracy that are provided by the modern GPS devices. Moreover, Li et al. study the optimization of the GPS data transmission interval, so as to minimize the communications overhead without decreasing the quality of the traffic estimation [62]. In this respect, extensive research activity has been concentrated on the development of traffic prediction algorithms specifically addressing the exploitation of the GPS data [14, 114, 19, 35, 82, 39, 68, 83, 8]. In parallel, travel time prediction has always been a hot topic in vehicular traffic forecasting due to its favourable features that have been described above [75, 81, 124, 41, 44].

Furthermore, the selection of the traffic descriptor is dependent upon the application type. In particular, ATIS operate optimally with physical quantities whose conceptual content is closer to the human perception, such as time and speed, since the measured values are targeted to be effectively communicated to the end-user. On the contrary, in the case of ATMS (Advanced Traffic Management System), where the ultimate goal is traffic control, traffic flow and occupancy can be proven to provide more exploitable information. Similarly, the area of the implementation (urban/suburban/highway/freeway) is a primary factor for traffic descriptor selection, since, for example, the time and speed measurements in urban environments pertain outliers and abrupt variations that result in high degree of uncertainty and call for more complex data pre-processing methods, while on the other hand most metropolitan areas are equipped with an extensive network of loop detectors that guarantee robust feedback of traffic flow and occupancy. Exactly the opposite holds for suburban or rural environments, where the speed data from GPS-enabled devices provide a cost-efficient solution for acquiring a reliable overview of the whole network.

1.2.2 Additional Traffic Descriptors

Besides the five basic traffic descriptors that have already been hereby described, additional metrics of traffic intensity have also been introduced aiming at providing a more explicit indication of the traffic conditions, according to the particular user/system requirements and specifications. One of the most prominent examples is queue length at signalized intersections, as formulating solid queue length estimations at signalized intersections are of prodigious interest for optimizing the dynamic control of the involved traffic signals. Comert and Cetin introduce an algorithm for calculating queue length by making virtue of the feedback (location and time) from the fleet of probe vehicles, while they also develop an analytical model for estimating the error in their results as a function of the percentage of probe vehicles in the traffic stream [18]. Moreover, Yingfeng et al. propose image processing methods for detecting the queue length by exploiting available video footage taken during the red cycle [118], while Liu et al. take also into consideration the phenomenon that the congested queue covers the whole length between two successive signalized junctions [65]. Furthermore, Kim and Park develop a model that allows for multi-time step queue prediction, based on discrete time point process [52].

In general, it should be pointed out that the concept of traffic estimation and forecasting at signalized intersections shares many common primitives with the area of queueing theory, which is regarded as the cornerstone of traffic (telecommunications, vehicular) analysis at servicing points, where bottlenecks are expected to arise. Therefore, apart from exploiting queue length as a metric of traffic intensity, the queuing theory principles (management of arrival/departure processes) have also been applied in order to model traffic behaviour [100] as one of the first approaches to the traffic flow problem [67]. Soh et al. implement Markov decision control methods for minimizing waiting time and queue length through optimal control of traffic signalling at junctions of interest with varying arrival rates of the vehicles.

Additionally, several studies have dealt with the issue of quantifying the qualitative concept of congestion, which merely refers to the human perception of the state of low speed drive and the corresponding delays induced to the travel time, due to traffic overload. Initially, the *congestion index* was introduced by Van Vuren and Leonard on the basis of travel time latencies [99]. Grant-Muller and Laird perform a rather thorough analysis of traffic congestion and attempt to define this notion from both the network's and the user's perspective, while they also provide multiple different quantitative approaches as functions of delay, journey time, speed and flow/capacity [31]. Moreover, Yunteng et al. claim that congestion quantification methods that are tailored for the freeway and which use operational characteristics (e.g., delay, speed and occupation etc.) are inconvenient to be used on signalized intersections. Hence, they combine traffic demand and traffic supply into a whole, to describe the congestion in signalized intersection based on loop detector data [121]. Marfia and Roccetti propose a novel definition of traffic congestion according to which a road is in a congested state only when the likelihood of finding it in the same congested state is high in the near future. Based on this new definition, an algorithm is devised that, exploiting probe vehicles, for any given road identifies if it is congested or not and provides the estimation that a current congested state will last for at least a given time interval [66].

Furthermore, special reference should be made to the detection and prediction of traffic accidents, which, although they are not of course considered as indicators of the traffic intensity, their occurrence still induces significant impact upon the evolution of traffic conditions. In this respect, Kamijo et al. developed an image processing algorithm that determines the state transit of each pixel along both the x-y image axes and the timeline and achieves event detection on the basis of the hidden Markov model (HMM) and the predefines behavioral patterns [48]. The authors in [109] follow a completely alternative approach to accident detection, through the use of the standard equipment of commercial smartphones, i.e. accelerometers and microphone. Huilin and Yucai take the step ahead towards accident prediction through the implementation of Neural Networks in order to enhance planning and traffic management [40].

1.2.3 Integration of Multiple Traffic Descriptors

As it has already been mentioned, each traffic descriptor facilitates a different perspective of analysis of the traffic conditions, while, at the same time, the availability of multiple data sources in the majority of real-life scenarios results in the co-existence of respectively multiple metrics of traffic intensity. Within this framework, several methods have been proposed for integrating disparate traffic information into a common traffic prediction model. Being more specific, the authors in [25] attempt to map the inter-relations among flow, speed and density, while their approach also aims at incorporating semantic information, such as weather and visibility conditions. The results include predictions of all the three input variables. Similarly, Chandra and Al-Deek are concerned with the forecasting of both traffic speeds and traffic flows at different locations in the investigated area [12]. Wei and Lee develop a functional relation between real-time traffic data as the input variables and real bus travel time as the output variable. Real-time traffic data are collected from bus GPS, loop detectors and incident databases and eventually the forecasted travel time is provided [107]. In the same way, Yuan et al. incorporate observation models for both Eulerian and Lagrangian sensor data that originate from loop detectors and vehicle trajectories respectively [119].

Alternatively, instead of independently modelling and predicting each traffic descriptor, Muraki and Kanoh implement the transformation of speed feedback into density values, so as to the new information format to better accommodate the dynamic control of traffic signals [70]. Moreover, in order to enhance the exploitation of the novel GPS-oriented speed data that are broadly available in the modern era, Herrera and Bayen incorporate speed measurements into existing traffic flow models [38]. Finally, by dividing the traffic volume with the road occupancy, Kamarianakis and Prastacos extract a new feature for traffic assessment that is called relative velocity [46].

1.3 Traffic Prediction Methodology

Traffic prediction algorithms aim at producing a solid forecast of maximum accuracy regarding the expected traffic conditions at given locations of the transportation network at the near future. To achieve this goal, the general approach that is followed is to make virtue of the most recent as well as historical traffic data, so as to deduce consistent conclusion about the network's upcoming states. As a result, the estimation of the system's forthcoming behaviour is deduced on the basis of the available information about its present and past status. In this respect, the research activity in traffic prediction has been focused on the development of the adequate processing scheme for the optimum exploitation of the available traffic data and the consolidation of this information towards the formulation of the requested traffic prediction. The performance of the implemented methodology is assessed according to its precision in reference to the values of the monitored traffic descriptors that are actually measured a posteriori, while the complexity of its deployment as well as the introduced processing overhead are also taken into account. Certain distance metrics are utilized for computing the divergence between the forecasted and the actually measured values of the analyzed traffic descriptors. The most widely accepted distance metrics for the evaluation of a prediction's error are the Mean Absolute Percentage Error (MAPE) and the Root Mean Square Error (RMSE) [86, 19] as well as the Mean Absolute Error (MAE), the Mean Relative Percent Error (MRPE) and the correlation coefficient between the actual and the predicted flow series [102]. Moreover, a key factor that characterizes the efficiency of any traffic prediction methodology is its robustness and reliability when applied to varying traffic environments.

According to the exact mathematical method that is implemented in order to formulate the traffic forecasting output on the basis of the monitored traffic conditions, a common taxonomy of the existing studies on traffic prediction is provided below [97] [103], including also an exhaustive presentation of the most interesting and promising solutions of each category.

1.3.1 Naive Methods (NM)

Although the term *naive* can have many interpretations, it is mainly used to refer to traffic prediction methods that are characterized by the absence of any advanced mathematical model or processing scheme for the exploitation of the traffic data. Hence, this category basically comprises of simplistic techniques, which should only be selected due to their minimum computational overhead and their ease of deployment, since they are characterized by significant drawbacks as far as the achieved accuracy is considered. Naive methods totally fail to capture any complex evolution of the traffic conditions. Moreover, they are completely inappropriate to follow the usually abrupt fluctuations of the traffic intensity and their efficiency is radically degraded by the existence of outliers, which is very common in traffic analysis, due to the inherently on-off nature of vehicular traffic, especially when congestion is faced. Therefore, they are outperformed by any other method for short-term prediction, while they can be proven tolerably trustworthy only in the case that longer prediction horizons are desirable.

Naive methods are basically limited to preliminary studies and implementations. The most common approach for performing forecasting in any type of discipline and for any type of quantitative metric is the computation of its *historical average*. Assuming the existence of a time series of traffic data for the network location of interest, the computation of historical average provides an estimation of the forthcoming state by averaging together the samples residing within the most recent time-window of a predefined width [72, 46]. Due to their minimum processing requirements, variations of the historical average model were applied to early deployments of ATISs, such as AUTOGUIDE [43] and LISB [50]. Historical average could also be considered as a simplified form of Moving Average [24] [92] and hence to be classified with the time series methodologies. Furthermore, historical average is often used as the scheme of reference for assessing the improvement achieved by novel prediction techniques [86]. In its simplified form, historical average of depth p is given by:

$$X_t = \frac{1}{p}\sum_{k=1}^{p} X_{t-k} \tag{1}$$

where X_{t-k} is the variable's value as measured at the k^{th} sample in the past. Moreover, a weighted scheme can be also implemented to apply different weight (w_{t-k}) at different historical samples:

$$X_t = \frac{1}{p}\sum_{k=1}^{p} w_{t-k}X_{t-k} \tag{2}$$

For the applied weights:

$$\sum_{k=1}^{p} w_{t-k} = 1 \tag{3}$$

From an utterly different perspective, the authors in [108] perform a draft, yet lightweight, forecasting of the traffic conditions, by clustering the weekdays into groups of common traffic attributes by means of Ward's hierarchical clustering procedure. Eventually, the currently measured conditions are matching with one of the resulting clusters. Similarly, Chrobok et al. define a set of days' classes according to their daily and seasonal attributes, by averaging the traffic flow data from all the monitoring loop detectors [16]. A new series of measurements is mapped to the classes derived from the training data, using the Mean Average Deviation and the Mean Relative Deviation.

1.3.2 Parametric Methods (PM)

The term *parametric* is used in the traffic prediction bibliography to imply that the corresponding techniques are based on specific models, whose general structure and primitives have been defined in advance and only the exact values of a given set of parameters needs to be determined through a learning procedure that is implemented heuristically on the basis of the available data that refer to the system's historical behaviour. The general methodology that is followed for the implementation of parametric methods, is that initially the most suitable model is selected and established according to its fundamental principles. The model is fed (input) with traffic data of the network locations for which the upcoming state for the desirable forecasting horizon is a priori known as historical information. The values of the model's open parameters are determined (training of the model) by iteratively calculating the prediction error for different sets of these parameters and choosing the set of parameter values that minimizes the prediction error for the majority of the scenarios.

The most common category of parametric traffic prediction algorithms, comprises 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. The rational behind the utilization and implementation of time series approaches instead of traffic theory methodologies, which would be mostly expected, is that traffic data do not correspond to stationary processes and hence a way of capturing their inherently seasonal behaviour is necessary.

Basic Approaches

One of the most elementary methods for prediction through time series analysis is exponential smoothing (ES), which can be also regarded as a specific simplified case of ARIMA models [111, 112]. According to the exponential smoothing technique, the present value of the variable can be estimated as a weighted sum of its previous measurement and the previously estimated value:

$$X_t = wX_{t-1} + (1-w)S_{t-1} \tag{4}$$

where w is the smoothing factor with $0 < w < 1$ and S_{t-1} is the estimated value at the previous step with $S_1 = X_0$.

Additionally, aiming also at combining minimum complexity with acceptably increased precision accuracy and versatility, Linear Regression (LR) has also been applied as a solution for vehicular traffic prediction [91]. In [81], Linear Regression is deployed for forecasting the travel time at freeway/highway segments. The selection of LR as the optimum prediction method is motivated by the empirical observation that the future travel time on a segment of a highway can be described by a linear model of the instantaneous and historical travel times on that segment. Actually, after extensive experimental measurements, it has been noticed that, although the slope and intercept of this linear relationships can alter heavily depending on the time of day and the interval until the departure, the linearity still persists. Hence, the prediction scheme is based on the implementation of Linear Regression with time-varying coefficients [124].

Univariate Auto-Regressive Integrated Moving Average

The Auto-Regressive Integrated Moving Average (ARIMA) family of models is the most widely deployed approach for vehicular traffic prediction and for time series forecasting in general. ARIMA models are also well known as Box-Jenkins models, since they were first introduced by Box and Jenkins in order to perform forecasting through determining the optimal matching of a given time series to a set of its past values [6]. ARIMA is a generalization of the Auto-Regressive Moving Average (ARMA) model, which is applied strictly to stationary time series, in order to capture the probable non-stationary nature of the investigated time series [7]. ARMA comprises of two components:

- Auto-Regressive (AR) part. It provides the current value of the process as the linear aggregate of a finite number of previous values of the process plus an error term. A time series X_t is expressed according to the AR model as:

$$X_t = \sum_{k=1}^{p} \phi_k X_{t-k} + \epsilon_t \tag{5}$$

  where p is a non-negative integer denoting the order of the AR part, i.e. the number of past lagged terms taken into account. ϵ_t is the error term, which is considered to follow a Gaussian distribution of type (0, σ_ϵ^2), i.e. White Noise (WN). The AR(p) model requires the estimation of $p+1$ parameters, i.e. the factors of the process's past values $\phi_1, \dots, \phi_p$ and the WN variation σ_ϵ^2, which are calculated by the data themselves.

- Moving Average (MA) part. It computes the current value of the process as the linear aggregation of a finite number of previous error terms. A time series X_t is expressed according to the AR model as:

$$X_t = \sum_{k=1}^{q} \theta_i \epsilon_{t-k} + \epsilon_t \tag{6}$$

  where q is a non-negative integer denoting the order of the MA part.

Hence, the total ARMA(p,q) model of a time series X_t is expressed as:

$$X_t = \sum_{k=1}^{p} \phi_k X_{t-k} + \sum_{k=1}^{q} \theta_k \epsilon_{t-k} + \epsilon_t \tag{7}$$

or equivalently, using the notation most commonly used in time series theory:

$$\big(1 - \sum_{k=1}^{p} \phi_k B^k\big) X_t = \big(1 + \sum_{k=1}^{q} \theta_k B^k\big) \epsilon_t \tag{8}$$

where B denotes the *backwards shift* or *lag* operator

$$B^k X_t = X_{t-k} \tag{9}$$

An even more compact notation commonly used is:

$$\phi\,(B)\, X_t = \theta\,(B)\, \epsilon_t \tag{10}$$

Moreover, let Y_t be the time series that is derived from the initial time series X_t by differencing X_t at the d^{th} degree, where d is a non-negative integer:

$$Y_t = \bigtriangledown^d X_t = (1-B)^d X_t \tag{11}$$

where $\bigtriangledown$ is the differencing operator. Differencing a time series results in creating a transformed time series that consists of the differences between lagged series observations. Hence, based on the assumption that the d^{th} difference of a non-stationary AR model in terms of lags can be described by a stationary ARMA(p,q) model, the ARIMA(p,d,q) model is defined by:

$$\big(1-\sum_{k=1}^{p}\phi_k B^k\big)(1-B)^d X_t = \big(1+\sum_{k=1}^{q}\theta_k B^k\big)\epsilon_t \tag{12}$$

or equivalently

$$\phi(B)(1-B)^d X_t = \theta(B)\epsilon_t \tag{13}$$

The first implementation of univariate ARIMA for traffic forecasting was performed more than three decades ago by Ahmed and Cook [4] and it immediately drew the massive attention of the research community [71] and gained broad acceptance as a well-established solution for vehicular traffic forecasting for both freeway [57] and urban [34] scenarios.

At this point, before proceeding further with the presentation of the time series methodologies, it should be once again mentioned in reference also to the time series terminology, that the term univariate refers to analysis of single-variable time series, i.e. prediction of a single traffic descriptor (e.g. speed, flow etc.) based on successive measurements of this traffic descriptor at a single location (node, link). On the contrary, the term multivariate, regards either the integration of time series from multiple traffic descriptors into the prediction of a single prediction output or, which is the most usual case, the integration of time series of the same forecasted traffic descriptor from multiple locations that are inter-correlated with the location of interest.

Nevertheless, ARIMA models are characterized by their inherent inclination to lay particular weight on the average values and highly disregard any outliers, i.e. ARIMA models are considered to follow the predominant pattern of the time series evolution in time. Hence, the performance of the ARIMA techniques can be proven to be rather limited in the particular case of vehicular traffic prediction, since they suffer from the severe drawback that they fail to promptly capture the frequent occurrence of transitions between stop-and-go situations and conditions of free flow, which is the most interesting phenomenon of transportation networks as far as the requirements for building a robust ATIS/ATMS are concerned. To overcome this deficient response of ARIMA models to significant flow alterations, several variations of the initial univariate ARIMA model have been so far proposed in the literature, aiming at providing an as solid as possible methodology for consistently as well as transparently predicting the fluctuations of traffic intensity in a timely manner.

In this context, Lee and Fabro suggested the use of subset ARIMA (SUBARIMA), so as to achieve higher adaptability, through the selective incorporation of time series components of specific lags [56]. The model for subset ARIMA is derived from 12 if the factors for all the other lags apart from the chosen ones are set to zero. It was found that subset ARIMA provided results of increased stability and accuracy under the tested scenarios. Moreover, given the fact that traffic data present profoundly periodic behaviour at different time-scales (weekly, daily, hourly), Seasonal ARIMA (SARIMA) soon caught the interest of the researchers as the most effective means for modelling these seasonal patterns. In SARIMA both the Auto-Regressive and the Moving Average component are multiplied by a seasonal factor (polynomial). Thus the Seasonal ARIMA(p,d,q)(P,D,Q) model of a time series X_t with period S is defined as:

$$\big(1-\sum_{k=1}^{p}\phi_k B^k\big)\big(1-\sum_{k=1}^{P}\Phi_k (B^S)^k\big)(1-B)^d\left(1-B^S\right)^D X_t = \big(1+\sum_{k=1}^{q}\theta_k B^k\big)\big(1+\sum_{k=1}^{Q}\theta_k (B^S)^k\big)\epsilon_t \tag{14}$$

or equivalently

$$\phi\left(B\right)\Phi\left(B\right)\left(1-B\right)^{d}\left(1-B^{S}\right)^{D}X_{t}=\theta\left(B\right)\Theta\left(B^{S}\right)\epsilon_{t} \tag{15}$$

where i) p, P are the orders of the non-seasonal and seasonal Auto-Regressive polynomials, ii) q, Q are the orders of the non-seasonal and seasonal Moving Average polynomials and iii) d, D are the order of non-seasonal and seasonal differencing. From the definition of SARIMA in 14 it can be drawn the conclusion that a time series X_t is a SARIMA$(p,d,q)(P,D,Q)$ process if the differenced series $Y_t=\left(1-B\right)^{d}\left(1-B^{S}\right)^{D}X_t$ can be described as a stationary ARMA model. Williams et al. were the first to propose the exploitation of Seasonal ARIMA for vehicular traffic prediction [111], while the authors in [112] provide an extensive theoretical background enriched with experimental results for the implementation of SARIMA in traffic forecasting. Additionally, Smith et al. compare the performance of Seasonal ARIMA, as the standard parametric method of traffic prediction, against non-parametric regression models [86]. It has been proved that SARIMA, when applicable, steadily outperforms the non-parametric regression techniques, drawing the conclusion that traffic conditions present stochastic rather than chaotic behaviour.

Moreover, besides the selection of the most suitable model for formulating the evolution of the traffic conditions, determining the optimum set of weighting parameters for the model of choice still remains a task of prodigious importance that dictates the accuracy of the prediction. Hence, the Ghosh et al. lay emphasis on developing an advanced methodology for inferring the parameters of the SARIMA model [28]. In more detail, in contrast to the traditional estimation that is based upon maximum likelihood and/or least-squares techniques, the Bayesian method is employed to estimate these parameters, since in Bayesian analysis the Markov chain Monte Carlo method is used to solve the posterior integration problem in high dimension. Each of the estimated parameters from the Bayesian method has a probability density function conditional to the observed traffic volumes and thus the forecasts from the Bayesian model are expected to better fit to a traffic behaviour of frequent outliers and abrupt alterations.

Multivariate Auto-Regressive Integrated Moving Average

In contrast to the aforementioned univariate ARIMA models, a major evolution to the ARIMA methodology concerns the development of space-time models, in order to succeed in incorporating measurements from multiple locations, i.e. multivariate analysis. In general, the spatiotemporal attributes of traffic data have been the target of substantial research activity [15], since it is evident that, due to the consistency of the traffic flow across the transportation network and its directional nature, a road segment cannot be studied in an isolated manner. Yue and Yeh introduce Pearson Coefficient as the most adequate distance metric for quantifying the dynamically altering cross-correlations among a given set of links, while they also propose an empirical rule for defining both the extent of each link's neighbourhood and their dependency weights according to the computed Pearson Coefficient value [120]. One of the preliminary attempts to integrate spatially disparate time series for traffic prediction resulted to the development of the ARIMAX model, which applies transfer functions with autoregressive integrated moving average errors [110]. Alternatively, in [29], a new Multivariate Structural Time-series (MST) model using the Seemingly Unrelated Time-series Equation (SUTSE) has been chosen to model the traffic flow time-series observations from multiple junctions within a congested urban transportation network.

Among all the multivariate space-time approaches, Space-Time ARIMA (STARIMA)and its variations rise as the most promising solution fro performing reliable traffic forecasting. Space-Time ARIMA was first introduced by Pfeifer and Deutsch back in the early eighties for studying the spread of diseases [78]. Thereafter, STARIMA has been applied to a wide range of disparate disciplines as a solution for exploiting the cross-correlation among multiple interdependent time series. Indicatively, the study of river flow [79] and spatial econometrics [30] can be noted as prominent cases of STARIMA implementation. In brief, STARIMA models capture the space-time dependency by expressing the requested variable at a given time t and location l as a weighted linear combination of all previous measurements performed within the monitor time window at all the correlated locations of l, introducing a lag both in space and time. To achieve this goal, neighbours of each location are hierarchically ordered through the definition of a corresponding sequence of weighting matrices, which quantify the physical properties of the spatial system being under investigation. Hence, if X_t is the $N \times 1$ vector of measurements at time instant t at N locations, the seasonal STARIMA model of order $(p_l, d, q_m)\ times(P_\Lambda, D, Q_M)_S$ can be computed from [47]:

$$\phi_{p,\lambda}(B)\,\Phi_{P,\Lambda}\left(B^S\right)(1-B)^d\left(1-B^S\right)^D \mathbf{X_t} = \theta_{q,m}(B)\,\Theta_{Q,M}\left(B^S\right)\epsilon_t \tag{16}$$

where

$$\phi_{p,\lambda}(B) = 1 - \sum_{k=1}^{p}\sum_{l=0}^{\lambda^k}\phi_{k,l}W_l B^k$$

$$\Phi_{P,\Lambda}\left(B^S\right) = 1 - \sum_{k=1}^{P}\sum_{l=0}^{\Lambda^k}\Phi_{k,l}W_l B^{kS}$$

$$\theta_{q,m}(B) = 1 - \sum_{k=1}^{q}\sum_{l=0}^{m^k}\theta_{k,l}W_l B^k$$

$$\Theta_{Q,M}\left(B^S\right) = 1 - \sum_{k=1}^{Q}\sum_{l=0}^{M^k}\Theta_{k,l}W_l B^{kS}$$

$\phi_{k,l}$ and $\Phi_{k,l}$ are the non-seasonal and seasonal autoregressive parameters at temporal lag k and spatial lag l, while $\theta_{k,l}$ and $\Theta_{k,l}$ are the respective moving average parameters. λ_k and Λ_k are the non-seasonal and seasonal moving average spatial orders for the k^{th} autoregressive term, while m_k and M_k are the non-seasonal and seasonal moving average spatial orders for the k^{th} moving average term. Finally, W_l is the $N \times N$ matrix for spatial lag l.

In case of no seasonality, the simple STARMA model is derived:

$$X_t = \sum_{k=1}^{p}\sum_{l=0}^{\lambda^k}\phi_{k,l}W_l X_{t-k} - \sum_{k=1}^{q}\sum_{l=0}^{m^k}\theta_{k,l}W_l\epsilon_{t-k} + \epsilon_t \tag{17}$$

The deployment of STARIMA models for the optimization of vehicular traffic prediction has been extensively studied by Kamarianakis and Prastacos who address the issue of STARIMA implementation as a special case of Vector Autoregressive Moving Average (VARMA) models [46, 47]. A set of N time series is described by using a set of $N \times N$ square matrices that include the autoregressive and moving average parameters among locations at different spatial order, so as to represent all autocorrelations and cross-correlations of the corresponding time series obtained at these locations. A STARIMA and VARMA model are separately applied and their performance is compared against the results acquired via the implementation of historical average and univariate ARIMA. One weight matrix is determined for each hop-level distance between road links, i.e. k^{th}-order matrix corresponds to the correlation between links that are connected through k hops (junctions), and the elements of the matrix (weights of k^{th}-order matrix) are computed as a function of the portion of traffic flow that each link inherits from its downstream links residing k hops away.

The effect of neighbouring downstream and upstream locations to the traffic conditions of a particular link are demonstrated in [12] and Vector Auto-Regressive models are applied for formulating these spatial interrelations of parallel time series. The authors in [21] calculate the weights as a function not only of the links' physical distance but also of their average speed, so as to take into consideration the time necessary for the traffic to propagate. In [69], Generalized STARIMA (GSTARIMA) is introduced to allow for the more flexible definition of the weighting parameters on per location basis. The AR and MA components can be adhusted dynamically to follow the

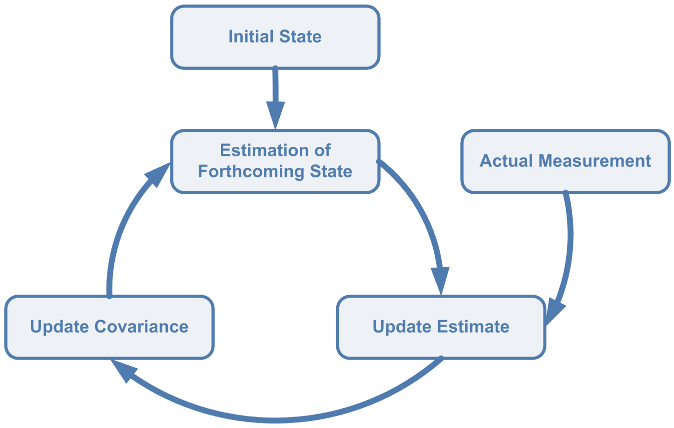

Figure 1: Block Diagram of Kalman Filter

respectively heterogeneous autocorrelations at different locations of the transport network. However, although the time parameters are space dependent, the space cross-correlations are still static according to a fixed predefined neighbourhood structure. A rather significant extension to the STARIMA model, as this has been established by Kamarianakis and Prastacos, is described by Min and Wynter who target at overcoming shortcomings of classical STARIMA, such as the supposed stationarity of the system and the constant relationship among the neighbouring links, which is depicted by the fixed space weighting matrices that are derived depending only on the links' distance, without considering the traffic status [68]. To this end, a Multivariate Spatial-Temporal Autoregressive (MSTAR) model is utilized.

Kalman Filter (KF)

Apart from the ARIMA-oriented methodologies that have been described above, Kalman Filters (KF) [45] was also one of the pivotal techniques that were proposed in the literature in the early eighties for the purposes of vehicular traffic prediction [73]. The main motivation behind the deployment of Kalman Filters lies within their ability of update the state variable continuously according to the new measurements [54].

According to the Kalman Filter theory, an estimation of the current state variable is performed, along with the covariance (uncertainty) of this estimation. Subsequently, upon the introduction of the next actual measurement, the estimates are recalculated as the weighted average of the components, laying additional weight to the estimates with higher certainty. This procedure is performed recursively, allowing for the prediction of the forthcoming state, by exploiting solely the current monitoring feedback and the immediately previously calculated state, without any additional past information being required. A draft description of the basic Kalman Filter functionality is presented in Figure 1 [94].

In this context, a common approach is to combine Kalman Filter with state-space modelling for multivariate analysis, where the Kalman Filter is deployed for the state estimation. In this respect, state space model has been implemented by Stathopoulos and Karlaftis for capturing the impact of upstream links along with a Kalman filter for calculating the sequential states [90]. Similar approach has been followed by the authors in [105, 106] who develop a stochastic macroscopic traffic flow model of freeway stretches, while some simple formulas are proposed to recreate real-time traffic measurements. This macroscopic traffic flow model along with the measurement model is organized in a compact state-space form, based on which a traffic state estimator is designed by use of the Extended-Kalman-filtering method, which is capable of capturing the non-linearity of the vehicular traffic data. Although Kalman Filter methodology has raised much lower research activity for traffic prediction than the ARIMA family of models, techniques for improving the Kalman Filter functionality for traffic prediction operation are still introduced. More concisely, Jula et al. exploit Kalman Filter as the tool of choice for forecasting travel times at freeway/highway segments, so as to eventually manage to calculate the anticipated arrival time at a given node [44]. Moreover, Hinsbergen at al. propose a method for reducing the processing overhead of Extended Kalman Filter, so as to overcome its inability to be implemented in real time in the case of large networks [98].

1.3.3 Non-Parametric Methods (NPM)

For the development of any of the *parametric* methods that have been described above, a model of basis is selected in advance and its parameters are determined through an optimization process for maximizing the prediction's accuracy either for a generalized approach or for the specific scenario under investigation. In juxtaposition, the term *non-parametric* is used, in order, not to indicate the absence of parameters, but to underline the fact that this category of traffic forecasting techniques does not presuppose a particular model structure. Hence, for the non-parametric methods, both the exact model structure and its parameters need to be specified along the processing of the traffic data, i.e. training of the model. Therefore, more extensive training procedure of the model is usually required as well as broader training dataset in juxtaposition with the parametric methods. The non-parametric prediction methods can be roughly classified into two broad categories:

- *Model-based.* These methods exploit the available historical data only during the training procedure, in order to build the model and define its parameters. Afterwards these data are discarded, since along the forecasting step no reference to the training data is necessary and only the current measurements are taken into consideration and are provided as input to the model. The most common example of model-based prediction techniques are the Artificial Neural Networks (ANN).
- *Memory-based.* On the contrary, these methods require to retain a database of historical samples, since, besides the model's training process, this information is also essential for formulating the estimation of the system's forthcoming states. The most prominent case of memory-based forecasting methods is non-parametric regression.

Non-Parametric Regression

Non-parametric regression is mainly identified with the k-Nearest Neighbour (kNN) methodology. According to the k-Nearest Neighbour techniques, let $\mathbf{S}$ be the set of available traffic observations of the $\mathbf{D}$ investigated traffic descriptors at the $\mathbf{L}$ locations of interest within a time window of depth w. Then, in order to acquire the forecasting for the system's forthcoming status, the maintained historical database is searched for the k clusters of traffic data (also referring to time windows of width w) that present the closest behavioural pattern in comparison with the reference set $\mathbf{S}$ of traffic data (pattern matching). After detecting these past states that show the highest similarity with the current status, the traffic intensity at an prediction horizon h is calculated as a parametric function of the traffic intensity measured at a step h ahead of the k-Nearest states that have been previously pinpointed in the historical database. As it becomes evident, the deployment of the k-Nearest Neighbour methodology for the purposes of vehicular traffic prediction exploits the inherently seasonal traffic attributes of the transportation networks, i.e. it is assumed that the evolution of the traffic intensity along the prediction horizon shall be similar with the evolution of the traffic intensity that has been so far monitored ahead of system states that resemble with the current one. As far as its implementation in traffic prediction schemes is concerned, the k-Nearest Neighbour technique was first introduced by Smith and Demetsky [85], who proved that it outperforms both the naive method of historical average and the parametric ARIMA model in terms of robustness against variable datasets. However, Williams et al. showed that the results obtained through Seasonal ARIMA modelling exceed the precision of k-Nearest Neighbour technique [111].

Further enhancing the study of non-parametric regression, the authors in [86], including both Smith and Williams, present a complete list of the open issues and challenges regarding the implementation of the k-Nearest Neighbour in real-life scenarios:

- State space. In the simplest deployment of kNN, the description of the system's current state refers to a set of w (width of monitoring window) sequential measurements (or a set of periodically aggregated measurements) of a single variable (speed, flow etc.) at a given node/link. Besides the issue of optimally defining w, even for a constant value of w there exists a practically infinite number of historical state spaces for different values of lag in reference to the current state.

- Distance metric. A distance metric needs to be defined for the pattern matching procedure, in order to quantify the proximity of each one of the available historical states with the investigated one that corresponds to the current traffic conditions of the transportation network.

- Forecast generation. As it has already been mentioned, the forecast is computed as the parametric function of the values of the resembling past states that fall at the prediction horizon. Such a parametric function can either be defined as the simple average of the considered values or a more complicated form can be selected. The most common approach is the calculation of the samples' weighted sum, where each weight is inversely proportional to the value of the distance metric between the current state and the respective historical state that the measurement refers to.

- Management of potential neighbor database. The performance of the k-Nearest Neighbour method is gravely dependent upon the extent and the quality of the historical data pool. Nevertheless, the larger the historical dataset becomes the higher the processing overhead introduced by the searching of the database is.

In this respect, the authors in [86] propose a method for heuristically improving the forecast generation by dynamically adjusting the weights of the forecast generation function. Moreover, an exhaustive assessment of this novel approach for non-parametric regression is performed against the Seasonal ARIMA model. The superiority of SARIMA is still deduced as a conclusion of the evaluation process, while, however, a combined model is suggested as an alternative when the requirements for the complete implementation of SARIMA are not satisfied.

Furthermore, k-Nearest Neighbour method has also been introduced for multivariate traffic analysis. Clark proposed a multivariate extension of non-parametric regression that exploits the three-dimensional nature of the traffic state in a multivariate manner that succeeds to incorporate traffic flow, occupancy and speed feedback [17]. Additionally, broadening the applicability of the kNN technique to multivariate prediction from multi-spatial point of view, a composite method has been suggested in [53] for utilizing the data from multiple loop detectors. The novelty lies within the fact that data from loop detectors in physical proximity with the loop detector where the forecasting is executed are also taken into account.

As a significant enhancement to the non-parametric regression, de Fabritiis et al. present a categorical k-Nearest Neighbour approach that is based upon the fundamental statement that, given the nature of traffic measurements, pattern matching for traffic prediction shall present higher efficiency when categorical data are processed [19]. The authors further support their choice of classifying traffic measurements into quantized levels of intensity, by putting forward the reasonable assumption that travellers have a better understanding of qualitative information due to the natural ordering and conceptual content. Within this framework, the time is discretized into 3-minutes frames and the speed observations for each monitored link are averaged on per timeframe basis. The spatiotemporal cross-correlation among neighbouring links is computed through the calculation of the Pearson Coefficient for different values of time lag for each pair of links and it is verified that every link is maximally interrelated with its upstream and downstream links. The speed value of each time frame is transformed into a categorical format that characterizes the link's traffic conditions: i) *free*, ii) *conditioned*, iii) *slow* and iv) *congested*. Moreover, if t is the index of the current timeframe, then the pattern of the current traffic state comprises of the categorical time series obtained from the target link and its immediate upstream and downstream links for different historical depths p, u, d respectively.

1. Target link: Speed at the $[t-p, t]$ timeframes.
2. Upstream link: Speed at the $[k-u, k]$ timeframes.
3. Downstream link: Speed at the $[k-d, k]$ timeframes.

For limiting state space, in order to minimize computational overhead, the pattern matching procedure is limited to searching the historical database of all previous days within only a specific time offset of $\pm f$ timeframes. Furthermore, two distance metrics are used for estimating the similarity between the current and each one of the past traffic patterns: Euclidean Distance (E_i) and the Spearman Coefficient. One separate threshold is defined for each of the distance metrics and only the past speed patterns that satisfy both criteria are taken into account for the calculation of the traffic prediction. Hence, in reference to the classical k-Nearest Neighbour methodology, where the value of k is defined in advanced, here the value of k is dynamically determined as the result of the forecasting procedure. Eventually, the forecast generation function is a weighted sum of the forthcoming speed values of all the k-Nearest Neighbours patterns, where the weight of each component is reversely proportional to the square root of the sum of the squares of the corresponding values of the two distance metrics.

The authors in [82] suggest an alternative method for deriving the traffic patterns to be compared, in order to overcome the phenomenon that, under heavy urban conditions, the GPS-oriented speed data that are obtained from the fleet of floating cars present rather abrupt fluctuations, including high percentage of zero values (stop-and-go conditions). According to this study's approach, the traffic pattern is formulated as the number of samples of zero speeds at the successive discrete timeframes. Furthermore, both a local and a global similarity measurement is calculated:

- Local similarity. The similarity between the current pattern at the link under investigation and previous observations of the same link. It is computed as the Euclidean distance between the two corresponding counts of zero speed values. Local similarity targets at capturing resemblance in the traffic conditions of the link itself.
- Global similarity. The similarity between the current and past patterns for the total transportation network as a whole. It is calculated as the Euclidean distance between the counts of zero speed values at the corresponding time windows throughout all the network's links. Global similarity aims at pinpointing timeframes where the traffic conditions of the aggregate network are similar.

Artificial Neural Networks (ANN)

Artificial Neural Networks (ANN), or just Neural Networks for simplicity, are by far the non-parametric models that have been mostly exploited in the field of vehicular traffic prediction. The ANNs, which are developed upon the fundamentals of artificial intelligence, can be roughly described as mathematical models that are being formulated in an automated manner by dynamically adjusting their structure and configuration according to the processed dataset, i.e. data-driven models. Thus, Artificial Neural Networks present outstanding performance in pattern classification and recognition [36]. The definition of the model's configuration is performed during the ANN's training procedure, which involves the execution of appropriate learning algorithms. In the same context, what is of prodigious importance as fas as their exploitation in real-life operations is concerned, ANNs present the highly desirable attribute of being capable to associate input and output patterns adaptively, without exhibiting any knowledge regarding the underlying physical relationships and processes that are actually responsible for the eventual result [74]. Hence, no a priori study or assumptions of the complicated and continuously altering network's characteristics is needed. Therefore, Artificial Neural Networks are regarded to prevail among the available traffic prediction methods in terms of their transparent applicability to the most wide range of scenarios. Due to this ability for reliable high-level analysis, which has established them as one of the most robust and accurate forecasting techniques, ANNs present the following advantages that are of particular interest in the special case of vehicular traffic predictions [103]:

- Efficient modelling of non-linear interrelations among the different functional entities of the transportation network as well as the variety of data sources and traffic descriptors.
- Efficient incorporation of the spatiotemporal correlations of the traffic conditions measurements.
- Traffic forecasting for multiple prediction horizons, without significant overhead.

The vast majority of the ANNs applications in the area of vehicular traffic prediction regard extensions of the Multi-Layered Perceptrons (MLPs) [55, 123, 41, 102]. The MLP is a distinctive subcategory of the Feed-Forward Neural Networks (FFNN) and the most commonly selected method for their training is the Error Back-Propagation (Back-Propagation Neural Networks - BPNN) learning rule. The basic architecture of such an ANN is depicted in Figure 2. FFNNs comprise of neurons (processing elements), which are arranged in multiple layers. More concisely, Feed-Forward Neural Networks are built upon the sequential interconnection of a single input layer, at least one hidden layer and the output layer, while each layer consists of different number of neurons. Additionally, relational functions are defined between the neurons residing at immediate neighbouring layers; on the contrary, neurons of the same layer respond completely independently from each other. These relationships between neurons of successive layers are determined by suitable weights that are assigned to the corresponding connection along the processing of data that is performed during the training stage and which is dictated by the Back-Propagation learning approach. According to the Back-Propagation training method, a two-phase iterative procedure is implemented:

- Forward. The sample data are fed as input to the hidden layers and an output is eventually produced, providing the estimated response of the transportation network, i.e. the prediction of the traffic conditions at the requested horizon when the current traffic intensity is equal to the initial training data that were given as input.
- Backward. The obtained state calculation is compared against the actually observed system response, i.e. the traffic forecasts are compared against the values that were eventually measured after the event described by the training data. The error is computed, using standard prediction error metrics, such as the MAPE, RMSE, MAE and MRPE, and the measured difference is propagated backwards into the ANN, so as the weights of the inter-layer connectors to be respectively adjusted.

This procedure is executed repeatedly until the error is stabilized, which means that no further optimization of the weights can be achieved [74].

As it has already been described above, the introduction of ANNs allows for the dynamical modelling of the system under investigation in a data-driven, completely unsupervised manner that guarantees the model's adaptability to the whole range of scenarios with acceptable performance. Nevertheless, it must be underlined that rather challenging issues still arise regarding the optimization of the ANN's implementation on per case basis, since significant aspects of the ANN's

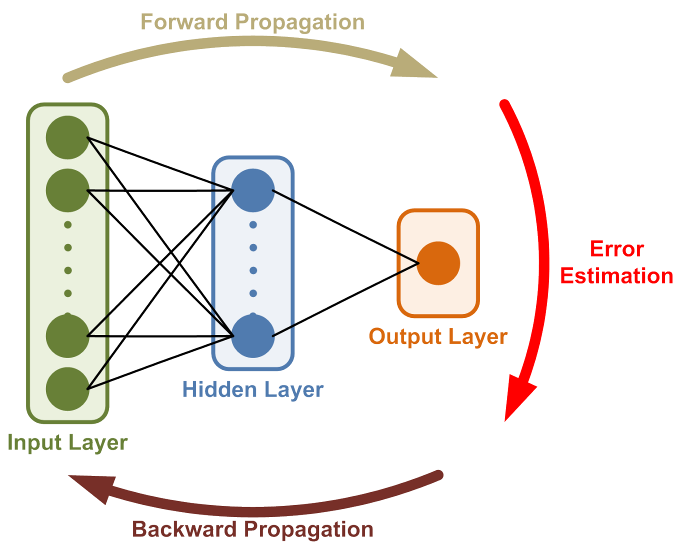

Figure 2: Feed Forward ANN with Back-Propagation

configuration remain to be defined by the analyst manually. More concisely, the maximization of an ANN's efficiency regards the addressing of three main issues:

- Preprocessing of the training data, so as to be offered to the ANN in a format that fits to the structure of the ANN and augments the effectiveness of the given training algorithm.
- The selection of the learning algorithm and its optimal parametrization. In brief, the learning algorithm refers to the methodology that is applied for pinpointing the local and global minima of the output error of so as to adjust the connections' weights during the back-propagation phase of the iterative training procedure. The most commonly used learning algorithms apply the Leverberg-Marquardt (aka Damped Least-Squares)[19] and the Gradient Descent [102] methods for minima calculation. Moreover, specific parameters of the learning algorithm need to be defined. For instant, the Gradient Descent implementation introduces two parameters for computing the step size along the adjustment of the weights' value:
  - Learning Rate. The current adjustment step is defined as a function of each weight's previous value. Learning Rate is the constant factor denoting the percentage of the previous weight value that is incorporated to the current weight calculation.
  - Momentum. For avoiding large oscillations, the current adjustment step is also estimated as a function of each weight's previous adjustment step. Momentum is the constant factor denoting the percentage of the previous weight adjustment step that is incorporated to the current weight computation.
- The definition of the ANN's structure, i.e. the number of hidden layers and the population of each one of them. The fundamental advantage of the ANNs lies within their ability to perform forecasting for system conditions for which the ANN has not been trained (no such data were fed during the training procedure). This feature of the ANNs, which is usually referred to as *generalization ability*, is affected by both the extent and content of the learning procedure and the structure of the ANN itself. In more detail, low structural complexity (few hidden layers and neurons per layer) results in decreased generalization ability, since the ANN's structure fails to support sufficient training. On the other hand, large number of connections is bound to cause the overfitting/overtraining of the ANN since it is incapable of differing from the

training patterns. Hence, although there exist several techniques for stopping the training procedure early enough to prevent from overfitting, the definition of the optimal structure is still considered as fundamental for building a high performance ANN.

Within the framework of the aforementioned optimization issues, numerous studies have focused on the development of ANN schemes particularly addressing the special requirements and intrinsic characteristics of vehicular traffic forecasting. Ishak and Alecsandru proposed the utilization of Principal Component Analysis (PCA) in order to project the input vector onto a smaller dimensional space and hence to improve the MLP performance by reducing the number of its inputs [42]. Moreover, the authors in [75] classify the historically available link travel times through an unsupervised clustering technique, so as an individual Artificial Neural Network (Modular ANN) to be defined and calibrated for each class and eventually each modular ANN to be used to perform the requested travel time forecasting. Furthermore, the use of a Self-Organizing Map (SOM, aka Kohonen map) has been proposed, in order to cluster the traffic data depending on the quality of the traffic conditions [13]. The output of the SOM-ANN are fed into the core MLP and the accuracy of the predictions is found to be superior of many parametric and naive methods. Moreover, given the time series nature of the traffic data, Lingras and Mountford utilized a Time Delay Neural Network (TDNN), which have been proven to be beneficial for time series analysis, since the neurons in a given layer can receive delayed input from other neurons in the same layer [64]. Additionally, a genetic algorithm is used by Vlahogianni et al. in order to develop the structure and the learning procedure of the ANNs, so as the ANN to become capable of capturing the spatial and temporal correlations of the fed source data [102]. Furthermore, the authors in [76] introduce an ANN for satisfying the specific requirements of speed prediction in modern transportation networks. The holistic model comprises of multiple NNs, in order to capture the dynamic nature of the traffic phenomena, through the incorporation of different prediction time intervals for each sensor location in the route.

Another method for traffic forecasting that originates in the area of artificial intelligence and machine learning, is Support Vector Regression (SVR) [87], which is a supervised learning method that is derived from the Support Vector Machine (SVM) theory. An application of SVR in vehicular traffic prediction has been introduced in [113], while the authors in [104] propose the implementation a variation that is based on the utilization of multiscale wavelet. Moreover, Castro-Neto et al. present Online Support Vector Regression machine for the prediction of short-term freeway traffic flow under both typical and atypical conditions [9]. Furthermore, SVR has also been exploited for the forecasting of accidents [61].

1.3.4 Hybrid Methods

As it has already been presented in the previous sections, numerous methods have been so far proposed for providing a reliable and robust solution to the complicated as well as rather important problem of vehicular traffic prediction. To this end, several different approaches have been implemented by exploiting a wide variety of mathematical theories and techniques that had already been successfully applied to other disciplines and which have been optimally adjusted by the transportation researchers for the field of traffic forecasting. In the same context, extensive studies have also been carried out regarding the assessment of the existing traffic prediction methodologies. In [84] a comparative study is presented regarding the efficiency of different non-parametric forecasting techniques. The performance of k-Nearest Neighbour models is evaluated against the accuracy of ANNs and it is shown that k-Nearest Neighbour rises as a rather promising technique for traffic forecasting. A comparison of non-parametric methods has also been performed in [101], where the precision and versatility of Support Vector Machines is benchmarked against the corresponding response of an ANN. Similarly, Time Delay and Recurrent ANNs are tested along with k-Nearest Neighbour models in [33], in order to evaluate their ability of providing acceptable results under both normal traffic conditions and in case of incidents' occurrence. The comparison of parametric and non-parametric models is also addressed in [116], where the forecasting efficiency of ARIMA, Support Vector Regression and Artificial Neural Networks is executed, through the exploitation of probe vehicle data for varying traffic scenarios. Actually, apart from the univariate ARIMA, a multivariate ARIMA version that incorporates an indicator variable for marking the unexpected conditions (e.g. incidents, weather deterioration) is evaluated as well. Moreover, a Generalized Auto-Regressive Conditional Heteroscedasticity (GARCH) model is deployed in [115] for assessing the effectiveness of the algorithms applied in the specific area of travel time predictions. It is pointed out that the forecasting errors are decreased for augmented demand of the reliability. Additionally, in [49] an exhaustive analysis of the characteristics of ANNs and statistical approaches is presented. From a broader perspective, the authors in [77] examine thoroughly the existence of specific factors that contribute to the deficiency of the traffic forecasting methods. Their findings that are based on multiple related projects carried out in the area of Minesota, USA, indicate that along the model development there is a substantial underestimation of significant factors such as the roads' type and function classification direction. Additionally, it is noted that semantic information as societal changes and trends are almost utterly neglected in the existing approaches.

The common conclusion that can be deduced from the analysis of all the evaluation studies that are described above as well as from the assessment process that is performed for each technique separately by the introducing researchers, is that no traffic prediction methodology can be decisively regarded to prevail among the bulk of available solutions. On the contrary, it can be safely taken for granted that the performance of the proposed forecasting algorithms is gravely dependent upon the configuration and structural characteristics of the transportation network, the user requirements and the sources and type of traffic data. Hence, different mathematical approaches and implementations can be optimally applied to respective transportation scenarios.

In this respect, during the recent years, the research activity has focused on the development of hybrid vehicular traffic prediction techniques, aiming at consolidating the favourable attributes of each constituent method into a highly efficient integrated solution. Hence, making virtue of the heterogeneity of its components, the aggregate modelling scheme is capable of serving under the most wide range of scenarios. One of the first hybrid approaches involved the utilization of Self-Organizing Maps as a data preprocessing step for an ARIMA-based predictor [96]. The technique, known as Kohonen-ARIMA (KARIMA), applies a Kohonen ANN for initially categorizing the input data into classes. This functional separation between the classification and forecasting tasks has been found to improve the prediction effectiveness in juxtaposition to both a univariate ARIMA model and a MLP ANN. A similar method for hybrid Kohonen-ARIMA modelling is also applied in [13] where, as it has already been mentioned, the Kohonen map has also been combined with an MLP forecasting module. In the same context, the authors in [10] proposes the deployment of Exponential Smoothing as the most adequate technique for preprocessing the source data, which will be thereafter fed to a Levenberg-Marquardt (LM) learning algorithm for training the ANN that plays the role of the core forecasting module. Such an approach is regarded to enhance the efficiency of the ANN implementation, since the new input presents more smoothed and coherent behaviour, without abrupt fluctuations. Furthermore, Zhang, proposes a methodology that amalgamates ARIMA and ANN processes [122]. In more detail, it is noticed that ARIMA models presuppose the existence of an underlying linear autocorrelation of the time series values and thus they fail to capture non-linear patterns, while, on the other hand, ANN are well-established due to their competence in following non-linear relationships. Therefore, initially an ARIMA model is applied, so as to analyse the system's linear constituents and subsequently an ANN is deployed to model the residuals from the ARIMA model, which are expected to contain all the information of the non-linearity of the traffic data. Eventually, the feedback from the ANN are provided as input for defining the error terms of the ARIMA model. Moreover, the authors in [26] utilize a Fuzzy Neural Network, which combines fuzzy logic and ANN for time series prediction, for developing a Non-linear Autoregressive Moving Average with exogenous inputs (NARMAX) model. To this end both feedforward and recurrent ANN structures are implemented. Similarly, a particular type of

Fuzzy Neural Network, which is known under the term Pseudo Outer-Product Fuzzy Neural Network using the Truth-Value-Restriction method (POPFNN-TVR), is deployed in [80] for performing short-term vehicular traffic forecasting.

The aforementioned hybrid methods have the common feature that they apply the heterogeneous components sequential, i.e. the output of the first element is provided as input to the second one. Alternatively, a hybrid method can also execute the comprising heterogeneous modules in an utterly separate manner and the individual results can be fused at a final integration step. Within such a framework, Stathopoulos and Dimitriou propose the exploitation of fuzzy logic principles for the fusion of predictions from multiple forecasting modules that operate completely individually from each other [89]. For achieving this goal, a Fuzzy Rule-Based System is implemented for integrating the separate forecasting outputs that result from an adaptive Kalman filter and an ANN model into a solid holistic prediction. A Fuzzy Rule-Based System is also implemented in [59] for fusing the traffic flow forecasts that are individually obtained from four independent modules: i) Exponential Smoothing, ii) ARIMA, iii) ANN and iv) Fuzzy Logic. Furthermore, in order to optimally treat the seasonal characteristics of the traffic data, the authors in [92] implement three different methods, the Moving Average, the Exponential Smoothing and the ARIMA model, for analyzing the time series that are derived from the weekly, daily and hourly cycle, respectively. Finally, an ANN is applied for incorporating the output from each one of the three components.

In Table 1 a summary of the most prominent techniques that have been proposed for traffic prediction in the area of transportation network is provided, laying particular emphasis on the most resent approaches.

1.4 Real-World Implementations

Among this bulk of available traffic prediction algorithms, additional emphasis should be laid on the approaches that have been designated to be implemented to the most modern applications of ATIS and/or ATMS, since the efficiency of these methodologies has been extensively proven under real-world scenarios and trials and therefore they can serve as a solid basis for further enhancements.

OCTOTelematics

As described in full detail in [19], *OCTOTelematics* has developed an ATIS by exploiting GPS data and implementing suitable traffic estimation and prediction methodologies. More concisely, *OCTOTelematics*, whose primary expertise lies in the field of telematics for insurance applications, has developed an On Board Unit (OBU) that is installed on the vehicles and is responsible for collecting statistics regarding the clients' driving behaviour, the encountered traffic conditions, any accidents' occurrence etc. The processing of this information shall eventually allow for the formulation of an as precise as possible insurance profiling on per client basis. Actually, in 2009 more than one million cars in Italy were equipped with the *OCTOTelematics* OBU (approximately 3% of the Italian cars), comprising a rather remarkable fleet of floating cars. Besides any other useful data, the OBU transmits the GPS measurements (coordinates, speed, direction) at specific spatial and temporal intervals (every 100 km or 12 minutes). Such a data gathering approach corresponds to the temporal sampling procedure and thus mapping of the traces onto the actual road links is appended. Finally, the obtained feedback is utilized for performing a categorical estimation of the traffic conditions, which is updated every 6 minutes and it is made freely available to the public through the company's website as shown in Figure 3. Specifically, Figure 3 includes a snapshot of the traffic conditions at the ring road and main arterials of Rome, where the different colouring corresponds to the respective classes of traffic intensity.

Additionally, *OCTOTelematics* deploys an heterogeneous scheme for forecasting traffic load at links of specific interest. The overall scheme comprises of a Pattern Matching and an ANN com-

Year	Proposed Technique	Traffic Descriptor	Univariate/ Multivariate	Area of Implementation	Methodology	
					PM	NPM
1999	[56]	F	UV	F/H	SUBARIMA	
2001	[110]	F	MV	F/H	ARIMAX	
2001	[13]	F, S, O	MV	F/H	ARIMA	SOM, ANN
2001	[64]	F	UV	F/H		TDNN
2002	[86]	F	UV	F/H	SARIMA	kNN
2003	[90]	F	MV	UR		KF
2003	[124]	T	UV	F/H	LR	
2003	[122]	-	UV	-	ARIMA	ANN
2003	[91]	O	UV	F/H	LR	
2003	[17]	F, S, O	MV	F/H		kNN
2003	[112]	F	UV	F/H	SARIMA	
2004	[81]	T	UV	F/H	LR	
2004	[113]	T, S	UV	F/H		SVR
2004	[42]	S	MV	F/H		ANN
2005	[47]	F	MV	UR	STARIMA	
2005	[26]	-	UV	-		Fuzzy Logic, ANN
2005	[41]	T	UV	F/H		ANN
2005	[102]	F	UV, MV	UR		ANN
2006	[80]	F, S	MV	F/H		Fuzzy Logic, ANN
2007	[53]	F, S, O	MV	F/H		kNN
2007	[28]	F	UV	F/H	SARIMA	
2008	[106]	F, S	UV	F/H		KF
2008	[19]	S	MV	F/H		kNN, ANN
2008	[44]	T	UV	F/H		KF
2008	[89]	F	MV	UR		Fuzzy Logic, KF, ANN
2009	[29]	F	MV	F/H	MST	
2009	[12]	F, S	MV	F/H	VARMA	
2009	[92]	F	UV	F/H	ES, MA, ARIMA	ANN
2009	[9]	F	UV	F/H		SVR
2010	[21]	F	MV	UR	STARIMA	
2010	[82]	S	UV	UR		kNN
2010	[69]	F	MV	UR	GSTARIMA	
2011	[68]	F, S	MV	UR	MSTAR	
2011	[76]	S	MV	F/H		ANN
2011	[59]	F	UV	UR	ES, ARIMA	Fuzzy Logic, ANN
2012	[10]	S	UV	F/H	ES	ANN

Table 1: Classification of existing traffic prediction techniques

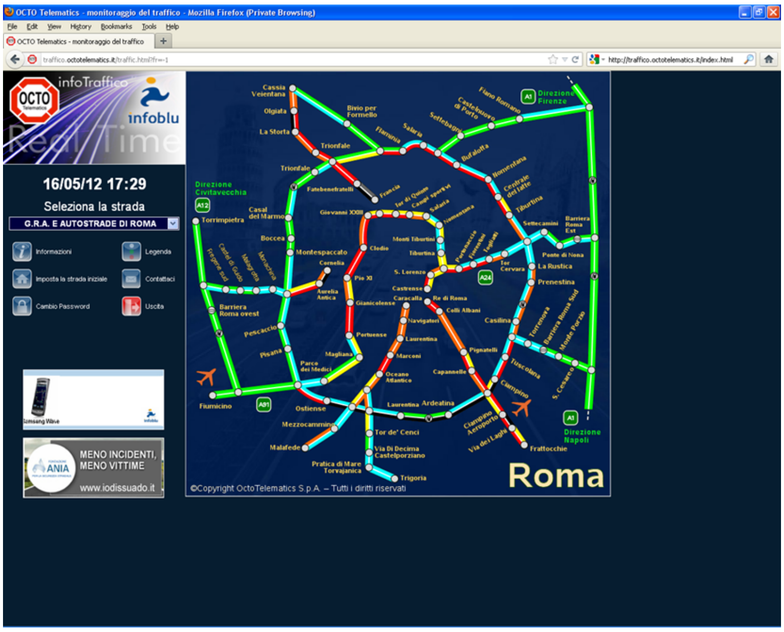

Figure 3: *OCTOTelematics* Traffic Monitoring Tool

ponent that operate utterly autonomously. The optimal forecasting module is selected depending on the format of the source data as well as the requested format of the prediction output. In particular, the Pattern Matching technique, which is a special variation of the k-Nearest Neighbour methodology and has been extensively described in Section 1.3.2, is carried out for identifying the network's forthcoming states when the travel speeds are categorically processed. On the other hand, a FFNN with a single hidden layer is utilized when the instantaneous speeds of the vehicles travelling at each link are provided as input and a respective speed forecast in terms of km/h is also required. Additionally, multiple FFNNs can be implemented for providing predictions at two different horizons of 5 and 10 minutes.

Mobile Century and Mobile Millennium Projects

The *Mobile Century* project as well as its extension, the *Mobile Millennium* project, are research projects that aim at designing, developing and deploying in full scale an holistic ATIS capable of performing in real-time:

- Exhaustive traffic data gathering
- Processing of the acquired information for estimation and short-term forecasting of the traffic conditions
- Distribution of the estimated and forecasted status of the transportation network back to the end-users (travelling citizens).

To this end, a public-private partnership has been established with the participation of UC Berkeley, Nokia Research Center and NAVTEQ, under the sponsorship of the California Department of Transportation, while the broader area of San Francisco, California, has been chosen for the case study [2].

The *Mobile Century/Millennium* system makes virtue of GPS data that are provided by GPS-enabled smartphones as well as by the GPS devices of the whole fleet of San Francisco taxis. The

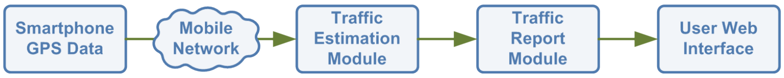

Figure 4: Block Diagram of *Mobile Century/Millennium* ATIS

data gathering architecture is based upon the innovative notion of Virtual Trip Lines (VTLs) [37], which is considered as a major advancements in the field of traffic monitoring infrastructures. Each VTL corresponds to a single location (point) of the transportation network (node/link) and hence it is completely identified by its coordinates. The rational behind the application of VTLs lies within the need for enhancing the data acquisition procedure by performing spatial sampling of the GPS data. In this respect, an adequate software, containing all the coordinates of all the VTLs, is developed and installed to the users' GPS devices (smartphones, navigators etc.), so as, whenever the vehicle crosses in proximity of a VTL, the GPS information to be transmitted to the data collection center through the mobile communications network. According to the VTL approach, a speed trace is not identified based on the device of origin but based on the index of the crossing VTL and hence the system refrains from maintaining historical information that could probably allow for the recreation of a vehicle's trajectory. Hence, besides the avoidance of the overhead that would be caused by the map-matching procedure, the introduction of VTLs holds the great advantage of guaranteeing the users' privacy, through achieving the data anonymization. A draft description of the *Mobile Century/Millennium* ATIS is provided by the Block Diagram in Figure 4.

Similarly to the case of the *OCTOTelematics* system, the researchers of the *Mobile Century/Millennium* project also implement both a discrete and a continuous traffic prediction methodology. In more detail, logistic regression is applied for the first time in vehicular traffic forecasting for performing the traffic estimation on the basis of clustering the traffic conditions onto a set of discrete congestion states. Additionally, STARMA is utilized for forecasting continuous travel times. The STARMA model is further exploited for predictions at a multi-step horizon, by using the single-step output of the model as input for forecasting at the subsequent step. Furthermore, in order to maximally exploit the potentials of the huge dataset of GPS observations, the authors in [38] address the issue of incorporating the GPS speed information into existing traffic flow models.

IBM Traffic Prediction Tool and Smarter Traveller Research Initiative

In cooperation with the Singapore Land Transport Authority, IBM carried out the implementation of its novel traffic prediction tool at the the pilot site of the Singapore's central business district [1]. The IBM tool is based on the deployment of a spatiotemporal analysis of the transportation network, according to the MSTAR model, which allows for incorporating the impact of neighbouring road segments into the calculations of a single location's conditions [68]. The system utilizes real-time data from loop detectors and is capable of performing traffic forecasting at multiple horizons of 10, 15, 30, 45 and 60 minutes.

Furthermore, upon the solid basis of the IBM traffic prediction tool and the *Mobile Century/Millennium* project, the *Smarter Traveller Research Initiative* is established in 2011 between IBM and two key partners of the *Mobile Century/Millennium* consortium, i.e. California Department of Transportation and UC Berkeley [3]. The goal of this common action is to develop a system that shall be capable of adaptively defining a commuters profile per citizen, so as to allow for determining the optimal as well as alternative routes in a manner of preventive traffic management. To this aim, data from GPS-enabled devices as well as feedback from classical traffic monitoring infrastructures is envisioned to be exploited.

1.5 Conclusions

Throughout the present Section, there has been provided an exhaustive presentation of the techniques that have been so far presented in the existing literature for carrying out the crucial as well as challenging task of vehicular traffic prediction. The primary axes for classifying the available traffic prediction approaches are:

- Traffic Descriptor. The quantitative metric that is selected for numerically representing the intensity of traffic load.
- Methodology. The exact algorithm that is implemented for calculating the forthcoming traffic conditions (values of the traffic descriptors) as a function of the present and past observations.

Moreover, a series of hybrid techniques has also been examined, so as to assess the merits from combining heterogeneous perspectives towards the ultimate goal of increasing the prediction's accuracy. Finally, a summary of the most modern real-world implementation has been provided, in order to acquire a concrete view of the issues and solutions that are addressed for the actual deployment of a theoretical approach into a fully functional system.

The common conclusion that is safely deduced from the analysis of all the proposed traffic prediction methods as well as from the available comparative studies is that there is no technique that can be safely regarded as the prevailing one in the general case or even in the majority of the possible scenarios. On the contrary, it becomes more than evident that different algorithms may present optimal performance under different traffic conditions and configuration of the transportation network.

Nevertheless, despite this grave dependency of each method's performance upon the overall system environment, there can be still identified approaches that impose significant enhancements to the evolution of the traffic forecasting. Specifically, particular emphasis must be laid upon the incorporation of spatial correlations into the calculation of a location's anticipated behaviour. The multivariate analysis of traffic conditions in terms of the spatial interrelations among neighbouring links is generally proven to offer advanced forecasting capabilities, since the impact from upstream/downstream links is mapped ahead of time. Therefore, a major issue that prominently rises within the research community is the development of suitable methods capable of promptly and accurately capturing the dynamic interdependencies among the network's segments as well as the efficient incorporation of these observations into the forecasting algorithm.

References

[1] IBM Traffic Prediction Tool. http://www-03.ibm.com/press/us/en/pressrelease/21971.wss.

[2] Mobile Century and Mobile Millennium Projects. http://traffic.berkeley.edu/.

[3] Smarter Traveller Research Initiative. http://www-03.ibm.com/press/us/en/pressrelease/34261.wss.

[4] M. S. Ahmed and A. R. Cook. Analysis of freeway traffic time-series data by using box-jenkins techniques. *Transportation Research Record: Journal of the Transportation Research Board*, (722):1–9, 1979.

[5] R. Alfelor, H. S. Mahmassani, and J. Dong. Incorporating weather impacts in traffic estimation and prediction systems. In *Proceedings of the ITE 2009 Annual Meeting and Exhibit*, October 2009.

[6] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel. *Time Series Analysis: Forecasting and Control*. Series in Probability and Statistics. Wiley, fourth edition, June 2008.

[7] P. J. Brockwell and R. A. Davis. *Introduction to Time Series and Forecasting*. Springer, second edition, March 2002.

[8] P. Castro, D. Zhang, and S. Li. Urban traffic modelling and prediction using large scale taxi gps traces. *Lecture Notes in Computer Science: Pervasive Computing*, 7319:57–72, 2012.

[9] M. Castro-Neto, Y. S. Jeong, M. K. Jeong, and L. D. Han. Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. *Expert Systems with Applications*, 36(3, Part 2):6164–6173, 2009.

[10] K. Y. Chan, T.S. Dillon, J. Singh, and E. Chang. Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg-Marquardt algorithm. *IEEE Transactions on Intelligent Transportation Systems*, 13(2):644 –654, June 2012.

[11] K. Y Chan, S. Khadem, T. S. Dillon, V. Palade, J. Singh, and E. Chang. Selection of significant on-road sensor data for short-term traffic flow forecasting using the taguchi method. *IEEE Transactions on Industrial Informatics*, 8(2):255 –266, may 2012.

[12] S. R. Chandra and H. Al-Deek. Predictions of freeway traffic speeds and volumes using vector autoregressive models. *Journal of Intelligent Transportation Systems*, 13(2):53–72, 2009.

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

[14] M. Chen and S. I. J. Chien. Dynamic freeway travel-time prediction with probe vehicle data: Link based versus path based. *Transportation Research Record: Journal of the Transportation Research Board*, 1768:157–161, 2001.

[15] T. Cheng, J. Haworth, and J. Wang. Spatio-temporal autocorrelation of road network data. *Journal of Geographical Systems*, pages 1–25, 2011.

[16] R. Chrobok, O. Kaumann, J. Wahle, and M. Schreckenberg. Different methods of traffic forecast based on real data. *European Journal of Operational Research*, 155(3):558–568, January 2004.

[17] S. Clark. Traffic prediction using multivariate nonparametric regression. *Journal of Transportation Engineering*, 129(2):161–168, April/May 2003.

[18] G. Comert and M. Cetin. Analytical evaluation of the error in queue length estimation at traffic signals from probe vehicle data. *IEEE Transactions on Intelligent Transportation Systems*, 12(2):563 –573, june 2011.

[19] C. de Fabritiis, R. Ragona, and G. Valenti. Traffic estimation and prediction based on real time floating car data. In *11th International IEEE Conference on Intelligent Transportation Systems, 2008 (ITSC '08)*, pages 197 –203, October 2008.

[20] G. Dimitrakopoulos and P. Demestichas. Intelligent transportation systems. *IEEE Vehicular Technology Magazine*, 5(1):77–84, March 2010.

[21] Q. Y. Ding, X. F. Wang, X. Y. Zhang, and Z. Q. Sun. Forecasting traffic volume with space-time ARIMA model. *Advanced Materials Research*, 156 - 157:979–983, October 2010.

[22] J. Dong, H. S. Mahmassani, and R. Alfelor. Weather responsive traffic management: Deployment of real-time traffic estimation and prediction systems. In *Proceedings of the 89th Annual Meeting of Transportation Research Board*, January 2010.

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

[24] K. Farokhi Sadabadi, M. Hamedi, and A. Haghani. Evaluating moving average techniques in short-term travel time prediction using an avi data set. In *Transportation Research Board 89th Annual Meeting*, January 2010.

[25] L. Florio and L. Mussone. Neural-network models for classification and forecasting of freeway traffic flow stability. *Control Engineering Practice*, 4(2):153–164, February 1996.

[26] Y. Gao and M. J. Er. Narmax time series model prediction: feedforward and recurrent fuzzy neural network approaches. *Fuzzy Sets and Systems*, 150(2):331–350, March 2005.

[27] D. Gazis and C. Liu. Kalman filtering estimation of traffic counts for two network links in tandem. *Transportation Research Part B: Methodological*, 37(8):737–745, September 2003.

[28] B. Ghosh, B. Basu, and M. O'Mahony. Bayesian time-series model for short-term traffic flow forecasting. *Journal of Transportation Engineering*, 133(3):180–189, 2007.

[29] B. Ghosh, B. Basu, and M. O'Mahony. Multivariate short-term traffic flow forecasting using time-series analysis. *IEEE Transactions on Intelligent Transportation Systems*, 10(2):246–254, June 2009.

[30] R. Giacomini and C. W. J. Granger. Aggregation of space time processes. *Journal of Econometrics*, 118(1-2):7–26, February 2004.

[31] S. Grant-Muller and J. Laird. The cost of congestion. In *Proceedings of the European Transport Conference 2007*, pages 441–459, October 2007.

[32] D. Gross, J. F. Shortle, J. M. Thompson, and C. M. Harris. *Fundamentals of Queueing Theory*. Series in Probability and Statistics. Wiley, fourth edition, May 2009.

[33] F. Guo, J. W. Polak, and R. Krishnan. Comparison of modelling approaches for short term traffic prediction under normal and abnormal conditions. In *13th International IEEE Conference on Intelligent Transportation Systems (ITSC '10)*, pages 1209 –1214, september 2010.

[34] M. M. Hamed, H. R. Al-Masaeid, and Z. M. Bani Said. Short-term prediction of traffic volume in urban arterials. *Journal of Transportation Engineering*, 12(3):249–254, May/June 1995.

[35] B. Hamner. Predicting travel times with context-dependent random forests by modeling local and aggregate traffic flow. In *2010 IEEE International Conference on Data Mining Workshops (ICDMW '10)*, pages 1357 –1359, december 2010.

[36] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, second edition, July 1998.

[37] J. Herrera, D. Work, R. Herring, X. Ban, Q. Jacobson, and A. Bayen. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C: Emerging Technologies*, 18(4):568–583, August 2010.

[38] J. C. Herrera and Bayen A. M. Incorporation of lagrangian measurements in freeway traffic state estimation. *Transportation Research Part B: Methodological*, 44(4):460–481, May 2010.

[39] R. Herring, A. Hofleitner, S. Amin, T. Abou Nasr, A. A. Khalek, P. Abbeel, and A. M. Bayen. Using mobile phones to forecast arterial traffic through statistical learning. In *89th Annual Meeting of the Transportation Research Board*, January 2010.

[40] F. Huilin and Z. Yucai. The traffic accident prediction based on neural network. In *Second International Conference on Digital Manufacturing and Automation, 2011 (ICDMA '11)*, pages 1349 –1350, August 2011.

[41] Satu Innamaa. Short-term prediction of travel time using neural networks on an interurban highway. *Transportation*, 32(6):649–669, 2005.

[42] S. Ishak and C. Alecsandru. Optimizing traffic prediction performance of neural networks under various topological, input, and traffic condition settings. *Journal of Transportation Engineering*, 130(4):452–465, July-August 2004.

[43] D. J. Jeffery, K. Russam, and D. I. Robertson. Electronic route guidance by autoguide: the research background. *Traffic Engineering & Control*, 28(10):525–529, 1987.

[44] H. Jula, M. Dessouky, and P.A. Ioannou. Real-time estimation of travel times along the arcs and arrival times at the nodes of dynamic stochastic networks. *IEEE Transactions on Intelligent Transportation Systems*, 9(1):97 –110, march 2008.

[45] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, (82):35–45, 1960.

[46] Y. Kamarianakis and P. Prastacos. 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.

[47] Y. Kamarianakis and P. Prastacos. Space-time modeling of traffic flow. *Computers & Geosciences*, 31(2):119–133, 2005.

[48] S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi. Traffic monitoring and accident detection at intersections. *IEEE Transactions on Intelligent Transportation Systems*, 1(2):108 –118, June 2000.

[49] M. G. Karlaftis and E. I. Vlahogianni. Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. *Transportation Research Part C: Emerging Technologies*, 19(3):387–399, June 2011.

[50] I. Kaysi, M. Ben-Akiva, and H. Koutsopoulos. An integrated approach to vehicle routing and congestion prediction for real-time driver guidance. *Transportation Research Board 72nd Annual Meeting*, (1408), 1993.

[51] K. Keay and I. Simmonds. The association of rainfall and other weather variables with road traffic volume in melbourne, australia. *Accident Analysis & Prevention*, 37(1):109 –124, January 2005.

[52] W. Kim and B. H. Park. Multi-time step traffic queue prediction model based on discrete time point process (dtpp). In *Second International Conference on Future Generation Communication and Networking Symposia, 2008 (FGCNS '08)*, volume 4, pages 60 –65, december 2008.

[53] M. D. Kindzerske and D. Ni. Composite nearest neighbor nonparametric regression to improve traffic prediction. *Transportation Research Record: Journal of the Transportation Research Board*, 1993:30–35, 2007.

[54] C. M. Kuchipudi and S. I. J. Chien. Development of a hybrid model for dynamic travel-time prediction. *Transportation Research Record: Journal of the Transportation Research Board*, 1855:22–31, 2003.

[55] E. Kwon and Y. J. Stephanedes. Comparative evaluation of adaptive and neural-network exit demand prediction for freeway control. *Transportation Research Record: Journal of the Transportation Research Board*, (1446):66–76, 1994.

[56] S. Lee and D. B. Fambro. Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting. *Transportation Research Record: Journal of the Transportation Research Board*, 1678:179–188, 1999.

[57] M. Levin and Y.-D. Tsao. On forecasting freeway occupancies and volumes. *Transportation Research Record: Journal of the Transportation Research Board*, 1980.

[58] B. Li. Bayesian inference for vehicle speed and vehicle length using dual-loop detector data. *Transportation Research Part B: Methodological*, 44(1):108 –119, January 2010.

[59] C. Li, S.G. Anavatti, and T. Ray. Short-term traffic flow prediction using different techniques. In *37th Annual Conference on IEEE Industrial Electronics Society (IECON '11)*, pages 2423 –2428, November 2011.

[60] X. Li, J. Han, J. G. Lee, and H. Gonzalez. Traffic density-based discovery of hot routes in road networks. In *Proceedings of the 10th international conference on Advances in spatial and temporal databases 2007 (SSTD '07)*, pages 441–459, July 2007.

[61] X. Li, D. Lord, Y. Zhang, and Y. Xie. Predicting motor vehicle crashes using support vector machine models. *Accident Analysis & Prevention*, 40(4):1611–1618, July 2008.

[62] X. Li, W. Shu, M Li, H. Y. Huang, P. E. Luo, and M. Y. Wu. Performance evaluation of vehicle-based mobile sensor networks for traffic monitoring. *IEEE Transactions on Vehicular Technology*, 58(4):1647 –1653, May 2009.

[63] W.-H. Lin, Q. Lu, and J. Dahlgren. A dynamic procedure for short-term prediction of traffic conditions. *Transportation Research Record: Journal of the Transportation Research Board*, 1783, 2002.

[64] P. Lingras and P. Mountford. Time delay neural networks designed using genetic algorithms for short term inter-city traffic forecasting. *Lecture Notes in Computer Science: Engineering of Intelligent Systems*, 2070:290–299, 2001.

[65] H. X. Liu, X. Wu, W. Ma, and H. Hu. Real-time queue length estimation for congested signalized intersections. *Transportation Research Part C: Emerging Technologies*, 17(4):412–427, August 2009.

[66] G. Marfia and M. Roccetti. Vehicular congestion detection and short-term forecasting: A new model with results. *IEEE Transactions on Vehicular Technology*, 60(7):2936–2948, september 2011.

[67] A. J. Miller. A queueing model for road traffic flow. *Journal of the Royal Statistical Society. Series B (Methodological)*, 23(1):64–90, 1961.

[68] W. Min and Wynter L. Real-time road traffic prediction with spatio-temporal correlations. *Transportation Research Part C: Emerging Technologies*, 19(4):606–616, August 2011.

[69] X. Min, J. Hu, and Z. Zhang. Urban traffic network modeling and short-term traffic flow forecasting based on GSTARIMA model. In *13th International IEEE Conference on Intelligent Transportation Systems, 2010 (ITSC '10)*, pages 1535 –1540, sept. 2010.

[70] Y. Muraki and H. Kanoh. Wide-area traffic signal control using predicted traffic based on real-time information. In *11th International IEEE Conference on Intelligent Transportation Systems, 2008 (ITSC '08)*, pages 1055–1060, october 2008.

[71] N. L. Nihan and K. O. Holmesland. Use of the box and jenkins time series technique in traffic forecasting. *Transportation*, 9(2):125–143, June 1980.

[72] 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, september 2005.

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

[74] A. K. Palit and D. Popovic. *Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications*. Springer, first edition, October 2005.

[75] D. Park and L. R. Rilett. Forecasting multiple-period freeway link travel times using modular neural networks. *Transportation Research Record: Journal of the Transportation Research Board*, 1617:163–170, 2008.

[76] J. Park, D. Li, Y. L. Murphey, J. Kristinsson, R. McGee, M. Kuang, and T. Phillips. Real time vehicle speed prediction using a neural network traffic model. In *The 2011 International Joint Conference on Neural Networks (IJCNN)*, pages 2991 –2996, August 2011.

[77] P. Parthasarathi and D. Levinson. Post-construction evaluation of traffic forecast accuracy. *Transport Policy*, 17(6):428–443, November 2010.

[78] P. E. Pfeifer and S. J. Deutsch. A three-stage iterative procedure for space-time modeling. *American Statistical Association: Technometrics*, 22(1):35–47, February 1980.

[79] P. E. Pfeifer and S. J. Deutsch. Variance of the sample-time autocorrelation function of contemporaneously correlated variables. *SIAM Journal of Applied Mathematics, Series A*, 40(1):133–136, February 1981.

[80] C. Quek, M. Pasquier, and B. B. S. Lim. Pop-traffic: a novel fuzzy neural approach to road traffic analysis and prediction. *IEEE Transactions on Intelligent Transportation Systems*, 7(2):133 – 146, june 2006.

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

[82] W. Shen, Y. Kamarianakis, L. Wynter, J. He, Q. He, R. Lawrence, and G. Swirszcz. Traffic velocity prediction using gps data: Ieee icdm contest task 3 report. In *Data Mining Workshops (ICDMW), 2010 IEEE International Conference on*, pages 1369–1371, december 2010.

[83] A. Simroth and H. Zahle. Travel time prediction using floating car data applied to logistics planning. *IEEE Transactions on Intelligent Transportation Systems*, 12(1):243 –253, march 2011.

[84] B. L. Smith and M. J. Demetsky. Short-term traffic flow prediction models-a comparison of neural network and nonparametric regression approaches. In *IEEE International Conference on Systems, Man and Cybernetics, 1994*, volume 2, pages 1706 –1709 vol.2, oct 1994.

[85] B. L. Smith and M. J. Demetsky. Multiple-interval freeway traffic flow forecasting. *Transportation Research Record: Journal of the Transportation Research Board*, 1554:136–141, 1996.

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

[87] A. J. Smola and Bernhard Scholkopf. A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222, 2004.

[88] F. Soriguera and F. Robust. Estimation of traffic stream space mean speed from time aggregations of double loop detector data. *Transportation Research Part C: Emerging Technologies*, 19(1):115–129, February 2011.

[89] A. Stathopoulos, L. Dimitriou, and T. Tsekeris. Fuzzy modeling approach for combined forecasting of urban traffic flow. *Computer-Aided Civil and Infrastructure Engineering*, 23(7):521–535, 2008.

[90] A. Stathopoulos and M. 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.

[91] H. Sun, H. X. Liu, H. Xiao, R. R. He, and B. Ran. Use of local linear regression model for short-term traffic forecasting. *Transportation Research Record: Journal of the Transportation Research Board*, 1836, 2003.

[92] M. C. Tan, S. C. Wong, J. M. Xu, Z. R. Guan, and P. Zhang. An aggregation approach to short-term traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, 10(1):60–69, March 2009.

[93] T.T. Tchrakian, B. Basu, and M. O'Mahony. Real-time traffic flow forecasting using spectral analysis. *IEEE Transactions on Intelligent Transportation Systems*, 13(2):519 –526, June 2012.

[94] N. A. Thacker and A. J. Lacey. Tutorial: The Kalman Filter. `http://www-flare.cs.ucl.ac.uk/staff/S.Prince/4C75/LaceyThackerTutorial.pdf`.

[95] S. M. Turner, W. L. Eisele, R. J. Benz, and D. J. Holdener. *Travel Time Data Collection Handbook*. Texas Transportation Institute, March 1998. FHWA-PL-98-035.

[96] M. Van Der Voort, M. Dougherty, and S. Watson. Combining kohonen maps with ARIMA time series models to forecast traffic flow. *Transportation Research Part C: Emerging Technologies*, 4(5):307–318, October 1996.

[97] C. P. Van Hinsbergen, J. W. Van Lint, and F. M. Sanders. Short term traffic prediction models. In *Proceedings of the 14th World Congress on Intelligent Transport Systems (ITS) 2007*, Beijing, China, October 2007.

[98] C. P. I. J. van Hinsbergen, T. Schreiter, F. S. Zuurbier, J. W. C. van Lint, and H. J. van Zuylen. Localized extended kalman filter for scalable real-time traffic state estimation. *IEEE Transactions on Intelligent Transportation Systems*, 13(1):385 –394, march 2012.

[99] T. Van Vuren and D. Leonard. Urban congestion caused by incidents. *Traffic Engineering and Control*, 35, August 1994.

[100] T. Van Woensel and N. Vandaele. Modeling traffic flows with queueing models: A review. *Asia-Pacific Journal of Operational Research (APJOR)*, 24(04):435–461, 2007.

[101] L. Vanajakshi and L. R. Rilett. A comparison of the performance of artificial neural networks and support vector machines for the prediction of traffic speed. In *2004 IEEE Intelligent Vehicles Symposium 2004*, pages 194 – 199, June 2004.

[102] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias. Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach. *Transportation Research Part C: Emerging Technologies*, 13(3):211–234, 2005.

[103] Golias J. C. Karlaftis M. G. Vlahogianni, E. I. Short-term traffic forecasting: Overview of objectives and methods. *Transport Reviews*, 24(5):533–557, September 2004.

[104] F. Wang, G. Tan, and Y. Fang. Multiscale wavelet support vector regression for traffic flow prediction. In *Third International Symposium on Intelligent Information Technology Application (IITA '09*, volume 3, pages 319 –322, november 2009.

[105] Y. Wang and M. Papageorgiou. Real-time freeway traffic state estimation based on extended kalman filter: a general approach. *Transportation Research Part B: Methodological*, 39(2):141–167, February 2005.

[106] Y. Wang, M. Papageorgiou, and A. Messmer. Real-time freeway traffic state estimation based on extended kalman filter: Adaptive capabilities and real data testing. *Transportation Research Part A: Policy and Practice*, 42(10):1340–1358, December 2008.

[107] C. H. Wei and Y. Lee. Development of freeway travel time forecasting models by integrating different sources of traffic data. *IEEE Transactions on Vehicular Technology*, 56(6):3682–3694, november 2007.

[108] W. Weijermars and E. van Berkum. Analyzing highway flow patterns using cluster analysis. In *Proceedings of the IEEE Intelligent Transportation Systems 2005*, pages 308 – 313, september 2005.

[109] J. White, C. Thompson, H. Turner, B. Dougherty, and D. C. Schmidt. Wreckwatch: Automatic traffic accident detection and notification with smartphones. *Mob. Netw. Appl.*, 16(3):285–303, June 2011.

[110] B. M. Williams. Multivariate vehicular traffic flow prediction: Evaluation of ARIMAX modeling. *Transportation Research Record: Journal of the Transportation Research Board*, 1776:194–200, 2001.

[111] B. M. Williams, P. K. Durvasula, and D. E. Brown. Urban freeway traffic flow prediction. application of seasonal autoregressive integrated moving average and exponential smoothing models. *Transportation Research Record: Journal of the Transportation Research Board*, (1644):132–141, 1998.

[112] B. M. Williams and L. 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, 2003.

[113] C. H. Wu, J. M. Ho, and D. T. Lee. Travel-time prediction with support vector regression. *IEEE Transactions on Intelligent Transportation Systems*, 5(4):276–281, december 2004.

[114] J. S. Yang. Travel time prediction using the gps test vehicle and kalman filtering techniques. In *Proceedings of the 2005 American Control Conference, 2005*, pages 2128–2133, june 2005.

[115] M. Yang, Y. Liu, and Z. You. The reliability of travel time forecasting. *IEEE Transactions on Intelligent Transportation Systems*, 11(1):162 –171, march 2010.

[116] M. Yildirimoglu and K. Ozbay. Comparative evaluation of probe-based travel time prediction techniques under varying traffic conditions. *91st Annual Meeting of the Transportation Research Board*, January 2012.

[117] H. Yin, S. C. Wong, J. 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.

[118] C. Yingfeng, Z. Weigong, and W. Hai. Measurement of vehicle queue length based on video processing in intelligent traffic signal control system. In *International Conference on Measuring Technology and Mechatronics Automation 2010 (ICMTMA '10)*, volume 2, pages 615 –618, march 2010.

[119] Y. Yuan, J. W. C. Van Lint, R. E. Wilson, F. Van Wageningen-Kessels, and S. P. Hoogendoorn. Real-time lagrangian traffic state estimator for freeways. *IEEE Transactions on Intelligent Transportation Systems*, 13(1):59 –70, march 2012.

[120] Y. Yue and A. G. O. Yeh. Spatiotemporal traffic-flow dependency and short-term traffic forecasting. *Environment and Planning B: Planning and Design*, 35(5):762–771, 2008.

[121] L. Yunteng, Y. Xiaoguang, and W. Zhen. Quantification of congestion in signalized intersection based on loop detector data. In *IEEE Intelligent Transportation Systems Conference, 2007 (ITSC '07)*, pages 904 –909, October 2007.

[122] G. P. Zhang. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 20:159–175, January 2003.

[123] H. M. Zhang. Recursive prediction of traffic conditions with neural network models. *Journal of Transportation Engineering*, 126(6):472–481, November-December 2000.

[124] X. Zhang and J. A. Rice. Short-term travel time prediction. *Transportation Research Part C: Emerging Technologies*, 11(3-4):187–210, June-August 2003.

[125] W. Zheng, D.-H. Lee, and Q. Shi. Short-term freeway traffic flow prediction: Bayesian combined neural network approach. *Journal of Transportation Engineering*, 132(2):114–121, 2006.

[126] L. Zhu and S. Jin. Speed estimation with single loop detector using typical effective vehicle length. In *2011 International Conference on Multimedia Technology (ICMT '11)*, pages 4096 –4099, july 2011.