Bus speed estimation by neural networks to improve the automatic fleet management

G. Salvo [1], G. Amato [1], P. Zito [1*]

[1] Università degli Studi di Palermo,
Dipartimento di Ingegneria dei Trasporti

Abstract

In the urban areas, public transport service interacts with the private mobility. Moreover, on each link of the urban public transport network, the bus speed is affected by a high variability over time. It depends on the congestion level and the presence of bus way or no. The scheduling reliability of the public transport service is crucial to increase attractiveness against private car use. A comparison between a Radial Basis Function network (RBF) and Multi layer Perceptron (MLP) was carried out to estimate the average speed, analysing the dynamic bus location data achieved by an AVMS (Automatic Vehicle Monitoring System). Collected data concern bus location, geometrical parameters and traffic conditions. Public Transport Company of Palermo provided these data.

Keywords: Radial Basis Neural Network; Public Transport Performances; AVM system.

1. Introduction

Liberalization and privatisation of the public transport impose the adoption of appropriate strategies to improve the efficiency and the competition. The rationalization of the resources and the competition among different sectors (costs reduction, waste elimination, profits increase, improvements of the perceived service quality) have required the use of telematics and control systems for the automatic fleet management. Furthermore, an improved service quality of the public transport can attract new users. Under this aspect: by the company's view point, it produces an increase of revenue; whereas by the citizens' view point, it implies a reduction of the travel time and delay; whose effects can be led up to an increase the life quality, de-congestioning urban areas, and reduction of noise and air pollution levels.

The recent advances in information and communication technologies, which support phases of collection, storage, processing and dissemination of information, allow to have spatially referenced data. Intelligent Transport Systems (ITS) defined as: *the*

---

[*] Corresponding author: Pietro Zito (zito@ditra.unipa.it)

application of communications and information technologies in the sector of transport, allows:

- – to improve mobility of vehicles fleet (by i.e. Automated Vehicle Monitoring, bus priority system), increasing transport safety, maximizing the performance and the comfort, and reducing conflict points with private cars;
- – to keep travellers/drivers more informed (with pre and on trip information by i.e. internet kiosks, variable message sign at the bus stop, mobile phone, on board speaker and so on);
- – to improve the accessibility transport infrastructures increasing their capacity, optimizing the traffic flows and keeping the existing transport/traffic infrastructures;
- – to improve traffic management and increase the average travel speeds;

The main outcome of ITS applications, for the public transport has been a new approach for the management of fleet as such as AVMS (Automatic Vehicle Monitoring System). An AVMS is generally composed: by GPS (Global Position System) that updates the fleet vehicles' positions frequently, by remote control system (with an equipment of sensors on-board) to check the safety and security of vehicles and by a communication system that broadcasts information to the Control Centre (CC), whose task is essentially to coordinate the vehicles fleet. The CC can also receive real-time updates on the state of the road network (i.e. an incident on a link) and uses together with position data to determine a new scheduling of the affected bus. The AVMS allows to communicate in real time to the Control Centre the exact vehicle's position and various other parameters of each vehicle as load condition on board, vehicle diagnostic conditions (check of the engine and closure of doors).

The main task of Control Centre is to check the positions on the road network of the transit fleet, respect to scheduled ones at the bus stops. Whenever delays due to traffic congestion or difficulties in loading and unloading passengers occur on public transport network, the CC can provide alternative solutions. Another specialised application of AVMS is the emergency service, to improve security of the driver and passengers (in case of crimes on-board and/or an incident). Whether an incident occurs involving the bus, the control system will broadcast automatically its position directly to security services (police, ambulance, fire fighters, etc.). Moreover, this service is particularly useful in rural areas or in case of night incident. Thus, the AVMS allows of having a whole control of bus fleet in real-time, improving quality service under the following aspects:

- - to check any non recurrent events (incidents, dangerous situations, etc.), taking the right decision to keep a high service level by on-line fleet management data;
- - to plan a demand-oriented service (routes, frequencies, etc) by off-line fleet management data;
- - to up-date information provided for customers, as such as forecasted arrival/departure time at the bus stop, by using various communication devices (mobile phone, web site, variable message sign at bus stops, etc).

In the time scheduling of the public transport service, the bus speed is assumed constant on every road link, nevertheless the speed varies over time. It depends essentially on traffic and road conditions that in particular cases of traffic congestion can cause service reliability issues and delays on routes.

In literature, the automatic vehicle location (AVL) systems appeared since 1990, they used data coming from AVL for traveller information (Balogh 1993, and Daley 1994).

Furthermore, Daley (2003) forecasted transit arrival/departure time using Kalman Filter and automatic vehicle location data. Data coming from the AVMS allows the estimation of the average speed of buses on any road link of public transport network, planning scheduled times on each route in a more reliable way. This work aims to point out a novel approach to estimate the average speed of buses using data coming from an AVMS and elaborated by a neural network. The knowledge of the bus speed allows to plan and to keep the maximum levels of service improving its quality (efficiency, frequency, and regularity of the courses). In literature, many works were carried out to estimate the arc cost (running time) for private cars by the knowledge of the road characteristics and traffic conditions (slope, width of lane, tortuosity, number of intersection per km, flow and capacity (for further details see High Capacity Manual, 1990, and Cascetta, 1998). Nevertheless, for the public transport and in particular for bus service, similar relationships have been used by a car equivalency factor. Thus, these relationships can be improved by using tools (as neural models) which better fit to the real traffic conditions. The main task of the analysis is to model the bus average speed in terms of the road characteristics and traffic conditions, by using neural network.

The neural networks for the transport sector were mostly applied to forecast short-term traffic flows Mussone (1996), Dougherty (1996, and 1997), Chen (2001) and Yin et al. (2002), using a feed forward neural network and the back propagation learning algorithm or hybrid approaches (neurofuzzy and Kohonen maps). Neural networks were also implemented to estimate the average journey time (Tafti, 2001 and La Franca et al. 2004). Amin et al. (1998) used a radial basis function network to predict the flow of traffic. In 1997, Costa and Markelos appraised the public transport efficiency by neural networks. More recently Celikoglu and Cigizoglu (2007) developed a radial basis function to forecast daily trip flows of public transport.

The neural networks can approximate highly non-linear functions and do not require a prior knowledge on the nature of these relationships. In this work, a comparison was carried out between different neural paradigms namely: Multi Layer Perceptron (MLP) and Radial Basis Function network (RBF). Two neural networks were used to estimate the average speed analysing the dynamic bus location data achieved by AVMS. The Radial Basis Function network (RBF) is an effective method to solve complex mapping problems between input and output variables. Furthermore, few works based on RBF networks are present in literature among which it is notable Celikoglu's one.

In this paper, the collected data concern (used as input variables in our models): bus locations, geometrical parameters and traffic conditions on each link of public transport network (number and width lanes, rate flow/capacity, reserved lane, number of intersections with or without traffic lights per kilometre, number of bus stops per kilometre, legal or illegal parking on the way, etc.). Public Transport Company of Palermo provided them. Moreover, the calibrated neural networks provide the estimated speed for two traffic conditions (rush hour or off peak conditions).

The paper has the following structure: next section describes the RBF network and its advantages; description of input data is included in Section 3, Section 4 presents the methodology, whereas Section 5 shows results of the model proposed and finally Section 6 discusses conclusions and further research direction.

2. Multi Layer Perceptron and Radial Basis Function Networks

The multi layer perceptron (MLP) or feed forward neural network consists of a set of simply interconnected neurons. The output of each neuron (except those in the input layer) is scaled by the connecting weights, modified by a transfer or activation function which can be either linear or non-linear, and fed forward to be an input to the neurons in the next layer of the network.

The multi-layer feed-forward neural networks (MLP) have the well-known ability to learn through training. The back propagation algorithm (BP) is usually used to train MLP networks. The learning is based on some nonlinear optimization techniques, which avoid during the learning period of falling into local minima. During training, the neural network is repeatedly presented with the training data and the weights in the network are adjusted until the output vector, produced by the network, does not match the target vector within a certain error. The training process uses this error to adjust the weights of the network according to the gradient descent learning algorithm to minimise the error.

An RBF network is regarded as a feed-forward neural network with a single layer of hidden units, whose responses are the outputs of radial basis function. Broomhead and Lowel (1988) developed the concept of Radial Basis Function (RBF) networks and applied it in the neural network modelling. The radial basis neural networks are another class of neural networks mainly used to approximate functions and for cluster classification. An RBF network not only has good performance of generalization, but it avoids the over-elaborated, lengthy computing like Back Propagation (BP) algorithm. Its learning speed is $10^3 - 10^4$ time faster than the BP algorithm, and this makes the RBF network applied widely.
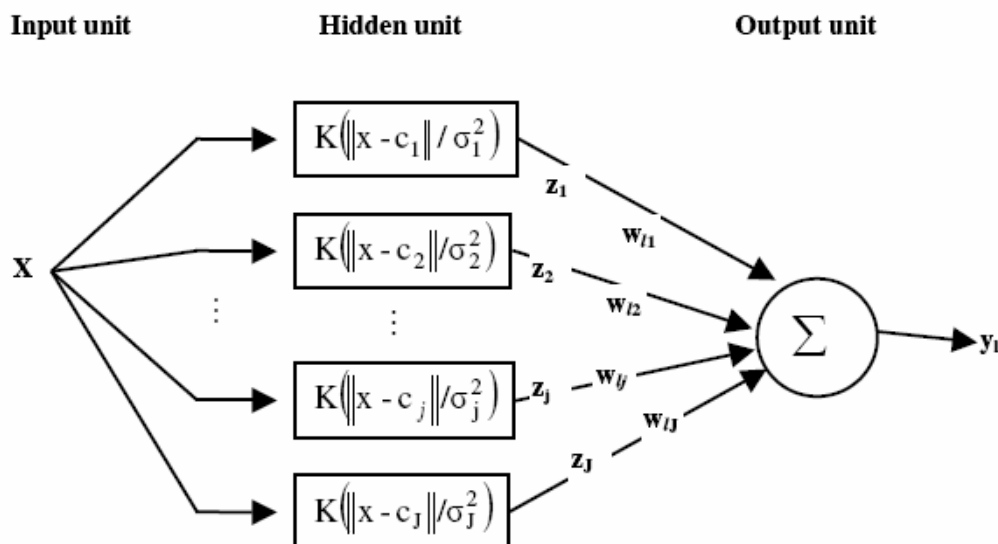

Figure 1: The structure of an RBFNN.

The adjustable parameters of such networks are the centres (the location of basis functions), the width of the receptive fields (the spread), the shape of the receptive field and the linear output weights. The Figure 1 shows the structure of an RBFNN. Radial

basis networks have their origin in techniques for performing exact interpolation of a data set. The input of each radial basis function of a RBF neural network is the distance between the input vector (activation) and its centre (location). The radial basis function approach introduces a set of N basis functions, which take the form ϕ ($\parallel$ x - x_n $\parallel$), where ϕ is a non-linear function. Thus the nth function depends on the Euclidean distance $\parallel$ x - x_n $\parallel$ between x and x_n. A basis function commonly used is the Gaussian:

$$\phi(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{1}$$

where σ is a parameter whose value controls the smoothness of the interpolating function.

The radial basis function networks are similar to the feed forward neural networks and assume the following form:

$$y_k(\mathbf{x}) = \sum_{j=1}^{M} w_{k,j}\phi_j(\mathbf{x}) + w_{k0} \tag{2}$$

$$\phi_j(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x}-\boldsymbol{\mu}_j\|^2}{2\sigma_j^2}\right) \tag{3}$$

where $\mathbf{x}$ is the n-dimensional input vector with elements x_i, and μ_j is the vector determining the centre of basis function ϕ_j and has elements $\mu_{j,i}$. By matrix notation, we can write:

$$\mathbf{y}(\mathbf{x}) = \mathbf{W}\Phi \tag{4}$$

where $\mathbf{W} = (w_{k,j})$ and $\Phi = (\phi_j)$. The weights can be achieved by minimization of a sum of square error function given by:

$$E = \frac{1}{2}\sum_{n}\sum_{k}\left\{y_k(\mathbf{x}^n) - t_k^n\right\}^2 \tag{5}$$

where: t_k^n is the target value for output unit k when the network is presented with input vector $\mathbf{x}^n$. An important aspect of radial basis network is the distinction between the roles of the first and second layer of weights. The first layer weights contain the parameters governing the basis function (σ_j and μ_j), which can be determined through unsupervised training. The second phase consists to find the weights $w_{k,j}$ of second layer, keeping fixed the parameters of radial basis functions, by a supervised training. Therefore, an unsupervised procedure allows the optimisation of the basis function parameters, which depend only on the input data from training set, and which ignores any target information. The basis function centres μ_j can be considered as prototypes of the input vectors. The determination of the weights $w_{k,j}$ of second layer, knowing the target vector, is achieved by a supervised procedure, resolving a linear matrix equation (Bishop, 1995):

$$\mathbf{W^T} = \mathbf{\Phi^\dagger T} \tag{6}$$

where: $\mathbf{T} = (t_k^n)$, $\mathbf{W^T}$ is the transpose matrix of weights, and $\mathbf{\Phi^\dagger}$ is the pseudo inverse of $\mathbf{\Phi}$. Girosi and Poggio (1990) showed that radial basis function networks possess the property of best approximation. The optimum number of nodes required in the hidden layer is related to the complexity of the input and output mapping, the amount of noise in the data and the amount of training data available. During training process, the RBF network will automatically add nodes into the hidden layer when needed until the actual error is below the given one.

3. Data collection and analysis

The main data source was the Public Transport Company of Palermo (AMAT). Dynamic bus location data along some bus routes were achieved by AVMS implemented by AMAT for fleet management. Collected data concern bus positions, geometrical parameters and traffic conditions.

Let **F/C** be the rate flow/capacity; **RL** the reserved lane (dummy variable 1 if present 0 otherwise); **NJ** the number of intersections with or without traffic lights per kilometre; **NB** the number of bus stops per kilometre; **IP** the illegal parking on the way (dummy variable); **FP** the free parking (dummy variable); **NI** the number of inflows per kilometre; **NO** the outflows per kilometre; **PC** the number of pedestrian crossings; **CA** with or without commercial activities (dummy variable). The neural networks provide the estimated speed for various traffic conditions (rush hour or off peak conditions) on each link of public transport network.

A cleaned dataset of 112 data points, each representing 40-seconds location of the line number 102 along its route, was used for the analysis (see figure 2).

Figura 2: The route of the bus line number 102.

After a statistical analysis of position data was carried out in order to delete outlier points. For each time interval between two positions of the bus in succession by GPS, the travel time was estimated. Thus, the average speed of the bus was computed for every time interval. The average bus speed is the target vector used during training process. Training data set took into account both rush hour and off peak traffic conditions.

The statistical analysis was carried out to determine whether any obvious correlations were present. The linear regression analysis carried out by SPSS, highlighted low coefficients of determination equal to 0.28 (see table 1).

Table 1: Results of the statistical model.

Multiple R	0.537						
R^2	0.288						
Adjusted error	0.268						
Standard error	0.2121						
ANOVA		DF	Sum of square			Mean Square	
Regression		3	1.967			0.565	
Residual		108	4.859			0.045	
F = 14.572		Sign. F = 0.0000					
Variable	B	SE B	Beta	Toler.	VIF	T	Sign. T
Const	.456	.044				10.242	.000
F/C	-.389	.070	-.465	.949	1.053	-5.585	.000
RL	.117	.044	.218	.977	1.023	2.651	.009
IP	.0958	.041	.194	.961	1.041	2.337	.021

4. Methodology

All data collected was normalised in a range (0,1) ensuring that values of different input variables are in the same range, in order to avoid overflows due to very large or very small weights.

Let $\hat{z}$ be the normalised values; z_{min} and z_{max} the minimum and maximum values of z respectively.

$$\hat{z} = \frac{z_i - z_{min}}{z_{max} - z_{min}} \tag{7}$$

The performance index used during the training process was the mean sum of squared errors between the estimated t_i and the actual values a_i, N is the number of observations, according the following equation:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \tag{8}$$

After the initial training of the MLP and RBF network, some statistics indexes (the Root Mean Squared Error RMSE, the Mean Absolute Percentage Error MAPE, the correlation and determination coefficient R and R^2) were determined in order to compare different models and to appraise the goodness of fit.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2} \tag{9}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|t_i - a_i|}{t_i} \times 100 \tag{10}$$

The structure of MLP was found by a pruning approach whereas the RBF during training process, the learning algorithm adds automatically nodes in the hidden layer until when the error function is below a given threshold.

Finally, a sensitivity analysis was carried out for input variables to see which of them is the most important to the estimation of the bus speed along its route. The sensitivity analysis investigated the behaviour of the RBF network to various increments of input variables. The conclusion from the sensitivity analysis was that if an increase in an input variable causes significant change (either positively or negatively) to the output variable, this input variable is regarded as an important input and should be retained in the model. A task of this work is to study a generic approach, which allows the neural networks developed for the selected route to be transferable to other bus routes. The goodness of results allows to extend the use of the calibrated neural network on new links and routes.

5. Results

The outcomes achieved after the training process of the RBF and MLP network are showed in table 2. Such values highlight the goodness of fit between real and estimated bus speed. Also, for sake of note RBF outperforms MLP in any studied case.

Table 2: Training and testing performance of RBF and MLP.

	RBF Network		*MLP Network*	
	On Training Set Bus line 102	**On Test Set Unseen data**	**On Training Set Bus line 102**	**On Test Set Unseen data**
RMSE	1.675	1.985	3.438	8.701
MAPE	6.010	8.507	22.429	33.789
R	0.950	0.932	0.786	0.405
R²	0.903	0.865	0.618	0.164

Figure 3 and 4 show the comparison between real and estimated bus speed in rush hour and off peak traffic conditions respectively for RBF. It should be noted that the estimated bus speed is very close to real one. Nevertheless, in the range between 18 and 28 (numbers of bus location along route corresponding to Maqueda Street) the RBF network provides underestimated values of speed in off peak traffic conditions whereas it provides overestimated values of speed in rush hour traffic conditions. Probably, since this street is characterised by traffic conditions and commercial activities much heavy that cause low performances in speed estimations.
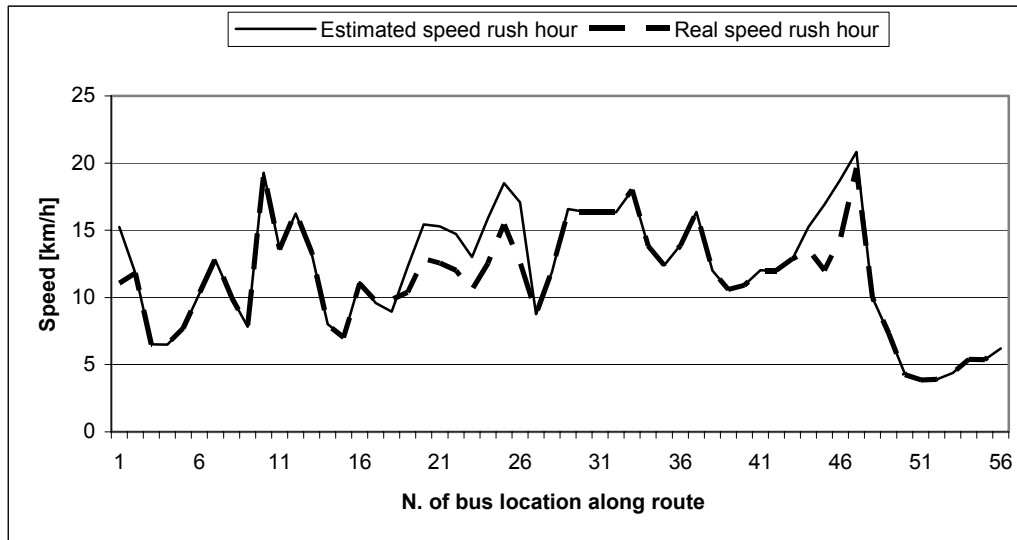

Figure 3: The comparison between real and estimated bus speed in rush hour traffic condition and RBFNN.
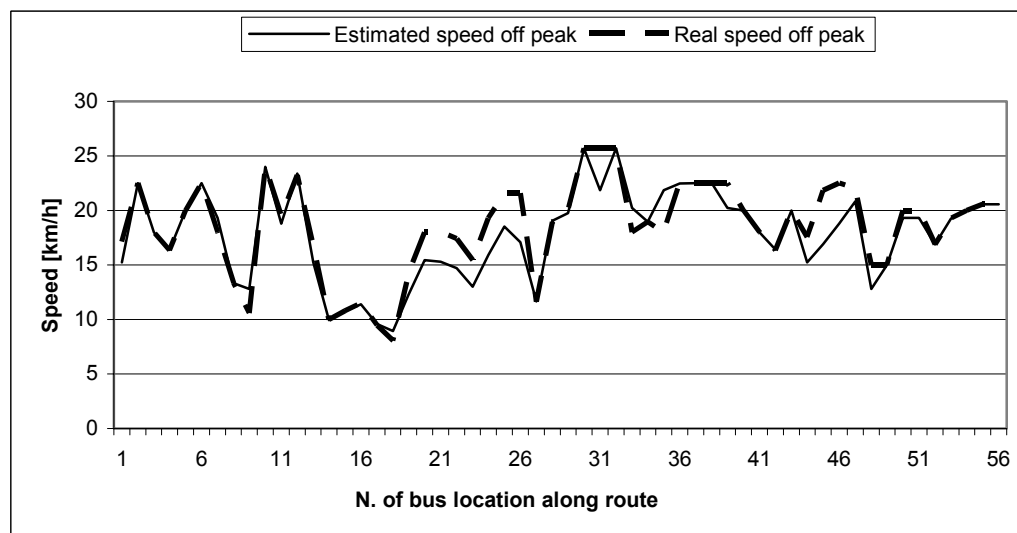

Figure 4: The comparison between real and estimated bus speed in off peak traffic condition and RBFNN.

Finally, the figure 5 shows sensitivity analysis discussed before. It should be noted that major important input variables are: the rate flow/capacity F/C, the reserved lane RL, commercial activities CA and illegal parking IP.
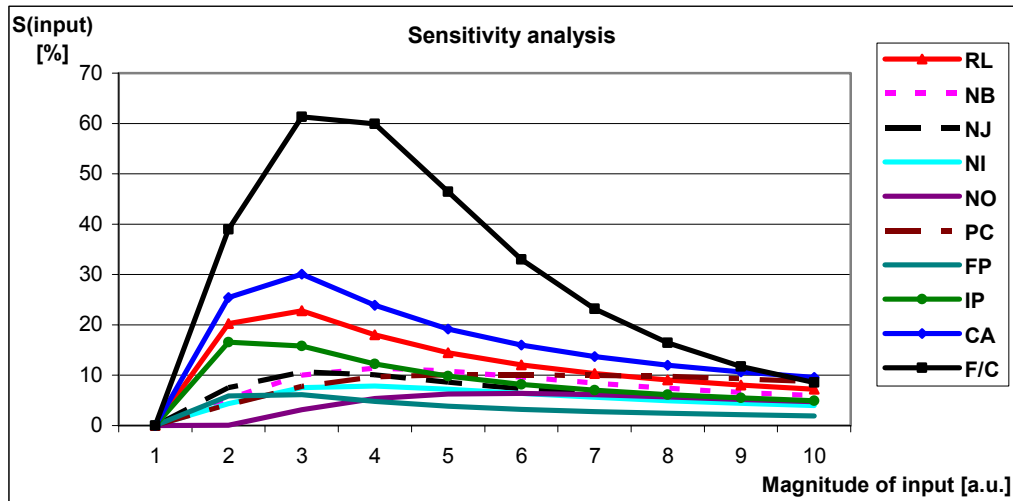

Figure 5: The sensitivity analysis.

The maximum error between real and estimated bus speed is equal to 4.5 km/h in Maqueda Street. After trained, the neural network was tested with unseen data on a new bus route along Strasburgo Road in Palermo. The test carried out gave a correlation coefficient between real and predicted values of 0.93, showing that the approach used still produces good performance on transferability (see table 2).

6. Conclusions

The effect on the average speed of bus produced by various parameters was investigated by using data coming from an AVMS and a neural network. The AVMS allows the communication in real time of the exact vehicle location and various other parameters of each vehicle as load condition on board, vehicle diagnostic conditions. Collected data concern bus location, geometrical parameters and traffic rules.

A comparison was carried out between Radial Basis Function network (RBF) and Multi Layer Perceptron (MLP). These neural netoworks were used to estimate the average speed analysing the dynamic bus location data achieved by AVMS. Input parameters are divided into geometrical parameters and traffic conditions (number and width lane, rate flow/capacity, reserved lane, number of intersections with or without traffic lights per kilometre, number of bus stops per kilometre, legal or illegal parking on the way). It was showed that RBF outperformed MLP in any studied case. The neural networks provide the estimated speed for various traffic conditions (rush hour or off peak conditions) on some link of public transport network. The RBF neural network (as well as MLP) was tested on different links of the public transport network with

different geometric and traffic characteristics from ones used for training, producing good results compared with the real measured average speed and highlighting a good generalization capabilities without over fitting issues. The calibrated neural model, on some links of public transport network, is able to estimate the bus average speed producing good results compared with the real measured average speed by knowledge of geometric and traffic characteristics.

The use of trained neural network to estimate the running average speed along all arcs of road public transport network, for various traffic condition (rush hour and off peak conditions) allows to assess the round-trip time achieved by sum of estimated running times for every arc along the bus route and to compare it with scheduled one.

The estimation of running time in terms of traffic conditions implies advantages both the Public Transport Company (optimising crew and means scheduling) and the costumer (reliability of the trip time and the scheduled timetable of transit at the bus stop).

The results presented in the paper are a specific case but encouraging for further applications of the proposed methodology on an enhanced database. The goodness of results allows to extend the use of the calibrated neural network on new links and any routes.

References

AA.VV. (1985) "Highway Capacity Manual" Transportation Research Board, National Research Board.
Amin, S. M., Rodin, E. Y. and Garcia-Ortiz, A. (1998) "Traffic Prediction and Management via RBF Neural Nets and Semantic Control", *Computer Aided Civil and Infrastructure Engineering*, 13(5), pp. 315-327.
Balogh, S. and Smith, R. (1993) "London transport's countdown system – a leader in the bus transport revolution", *IEE Colloquium* on "Public Transport Information and Management Systems. IEE London UK.
Bishop, C. M. (1995) *Neural Network for Pattern Recognition*, Oxford University Press, Oxford.
Broomhead, D. and Lowe, D. (1988) "Multi variable functional interpolation and adaptive networks", *Complex Systems* 2, pp. 321–355.
Cascetta, E. (1998) *Teoria e metodi dell'ingegneria dei sistemi di trasporto*, Utet.
Celikoglu, H. B. and Cigizoglu, H. K. (2007) "Modelling public transport trips by radial basis function neural networks", *Mathematical and computer modelling*, pp. 480-489.
Chen, H. and Grant Muller, S. (2001) "Use of sequential learning for short term traffic flow forecasting" *Transportation Research Part* C, 9, pp. 319-336.
Costa, A. and Markelos, R. N. (1997) "Evaluating public transport efficiency with neural network models", *Transportation Research Part C*, 5, pp. 301-312
Crainic, T. G. and Laporte, G. (1998) *Fleet Management and Logistics*, Kluwer Academic Publishers.
Dailey, D. J. and Cathey, F. W. (2003) "A prescription for transit arrival/departure prediction using automatic vehicle location data" *Transportation Research Part C*, 11, pp. 241-264.
Daley, D. and Haselkorn, M. (1994) "Demonstration of an advanced public transportation system in the context of an ivhs regional architectur". *First World Congress on Application of Transport Telematics and Intelligent Highway System*, pp 3024-3031.
Dougherty, M S (1998) "Exploring traffic systems by elasticity analysis of neural networks". In: Himanen, V., Nijkamp, P. and Reggiani, A. (eds), *Neural networks in transport applications*. (Ashgate)
Dougherty, M. S. and Cobbett, M. R. (1997) "Short term inter urban traffic forecasts using neural networks" *International Journal of Forecasting* 13, pp.21-31.
Dougherty, M., Van Der Voort, M. and Watson, S. (1996) "Combining Kohonen maps with ARIMA time series models to forecast traffic flow" *Transportation Research Part C*, 4, pp. 307-318.
Greenfeld, O. and Koncz, N. (1995) "The development of a gis-based transit advanced traveller information system". ARISA Proceedings. *Papers fron the Annual Conference of the urban and*

Regional Information Systems Association. Urban & Regional Inf. Syst. Assoc., Washington, DC, USA, pp. 695-709.

Hornik, K., Stinchcombe, M. and White, H. (1989) "Multilayer feed forward networks are universal approximators", *Neural Networks*, 2(5), pp. 359-366.

La Franca, L., Migliore, M., Salvo, G. and Carollo, F. (2004) "The automatic vehicle monitoring to improve the urban public transport management", *XI convegno CODATU Towards a more attractive Urban Transportation,* Bucharest, 22-24 aprile 2004.

Mussone, L. and Florio, L. (1996) "Neural network models for classification and forecasting of freeway traffic flow stability" *Control Engineering Practice*, 4, pp. 153-164.

Poggio, T. and Girosi, F. (1990) "Network for approximation and learning". *IEEE Proceeding* 78, pp. 1481-1497.

Tafti, M. F. (2001) "The application of artificial neural networks to anticipate the average journey time of traffic in the vicinity of merges" *Knowledge-Based Systems*, 14, pp. 203-211.

Yin, H., Wong, S. C., Xu, J. and Wong, C. K. (2002) "Urban traffic flow prediction using a fuzzy – neural approach" *Transportation Research Part C*, 10, pp. 85-98.