Analysis of typical speed profiles from floating heavy vehicles

A. Pascale
Politecnico di Milano, Italy, 5th Faculty of Eng.
Dept. DEI,Telecommunication Eng.

F. Deflorio[*]
Politecnico di Torino, Italy, 1st Faculty of Eng.
Dept. DITIC, Transport Eng.

M. Pedroli
W.A.Y. s.r.l, Torino, Italy[†]

M. Nicoli
Politecnico di Milano, Italy, 5th Faculty of Eng.
Dept. DEI,Telecommunication Eng.

B. Dalla Chiara
Politecnico di Torino, Italy, 1st Faculty of Eng.
Dept. DITIC, Transport Eng.

Abstract

Topic of the paper is the reconstruction of the speed field and the recognition of typical daily patterns in floating truck data (FTD) collected by heavy vehicles operating on Italian motorways. We present a statistical analysis of velocity measurements from FTD and a procedure for the estimation of the typical speed profiles. Output of this procedure could be exploited for integration of missing data in real-time speed estimation and for incident detection procedures.
Keywords: *Floating vehicles, Speed profiles, Traffic monitoring*

Introduction

Traffic monitoring has great relevance in intelligent transport system (ITS), where accuracy and reliability of data are fundamental for the implementation of actions for traffic management. Recently, monitoring technologies based on probe vehicles have been emerging both as self-working systems and in cooperation with infrastructure-based systems [1, 2, 3]. Floating car data (FCD) systems are based on a number of probe vehicles, equipped with GPS positioning devices, that periodically send traffic data to a central processing unit through cellular radio terminals - usually general packet radio service (GPRS). Transmitted data include position, velocity and travel time.

FCD systems present a number of characteristics that make this technology very interesting for ITS. First of all they do not rely on a fixed monitoring infrastructure installation, this implies higher flexibility in area covering, lower installation and maintenance costs. On the other hand, the reliability of data collection, especially in **heavy vehicles** systems, is limited by the penetration rate; sampling of traffic parameters is nonuniform and time varying, with resolution depending on a combination of factors such as the number of probes, traffic demand patterns, traffic conditions and road features. In this scenario, data processing is crucial for the integration of sampled data. Nowadays, the wide use of positioning devices and the availability of advanced processing tools are making FCD systems able to overcome problems related to low penetration rate, providing an accurate and at certain conditions reliable way for monitoring traffic flows.

Applications of FCD range from traffic monitoring and forecasting, travel time estimation, con-

---

[*]Corresponding author - Address: C.so Duca degli Abruzzi, 24 - 10129 Torino - I; Tel. +39 011 5645601; e-mail: francesco.deflorio@polito.it.

[†]Tests have been carried out on a FTD system developed by the company WAY (Turin, I) supplying AVLS for heavy-duty vehicles; the system has been developed within a project supported by the research program POR-FESR, cofunded by the European Regional Development Fund, the Italian Government and the Piedmont Region.

Fig. 1: Map of the North Italy motorways and look on A4 highway Milan→Brescia. Circles denote the vehicle velocities matched to highway segments.

struction of typical traffic patterns to enable identification of anomalies or incident detection, fleet management, and advanced navigation based on real-time traffic conditions [3, 4, 5, 6]. In this paper we focus our attention on the analysis of daily speed patterns for the construction of a reliable and accurate historical database. We propose a procedure to estimate the velocity field from FCD and a method based on hierarchical clustering [7] to identify typical speed profiles. Experiments are conducted on data collected by a floating truck data (FTD) system managed by an Italian company named W.A.Y. (Turin).

Scenario of analysis

Data used for this analysis are provided by a fleet of more than 13000 probe vehicles that send measures of position and instantaneous speed via GPRS to a server of the W.A.Y. traffic center. Most vehicles are GPS-equipped trucks traveling on Italian motorways. Each motorway is mapped into a set of segments of length $100\text{m} \leq l_s \leq 500\text{m}$, numbered in the direction of traffic as $s = 1, 2, ..., N_S$. Temporal sampling ranges from 20 s to 3 min. Raw vehicle data are associated to the segments on a geographical map based on the computation of distance between the GPS vehicle positions and the segments. Velocities after map matching, v [km/h], are shown in Fig. 1 for motorway of North Italy. For the subsequent analysis we focus on the motorway A4 Brescia→Milan as highlighted in the box.

Objective of the paper is the construction of an accurate and easily accessible historical database of speed timeseries to be used as reference for algorithms of real-time velocity estimation and incident detection. These final goals require the computation of speed from raw FCD, a preliminary classification of probes data by vehicle's topology and a subsequent classification of segment speed profiles based on similarities in the daily trends, as discussed in the following sections.

Data aggregation

Probe vehicles send an array of data composed by: time stamp t_m [data,hh:mm:ss], vehicle ID [#], GPS position [lat,long], instantaneous velocity [km/h], incremental distance covered by vehicle [km]. Here we analyze a dataset collected by a fleet of 4463 trucks (a subset of vehicles selected based on consistency check) on 42 working days (from February and March 2011) on the motorway A4 Brescia→Milan (about 90 km). The penetration rate for the system has been estimated to be ~ 0.025 in capacity condition - segment 111-114 at 7am, in box in Fig. 2. We start from data of positions and velocities after the map-matching process. For speed computation we use only data collected by vehicles that have been active for a time window of at least 30 min, to exclude unreliable data from vehicles with either sporadic or not constant transmission rate. We compute the velocity of vehicle i as the ratio between the incremental distance covered from last measure $\triangle s_{i,m}$ and the time interval $\triangle t_{i,m} = t_m - t_{m-1}$ elapsed from the last measure; typically it is $20\text{sec} \leq \triangle t_{i,m} \leq 3\text{min}$. The velocity computed in this way is assigned to both the segments to which the measure is associated and to the intermediate ones that the vehicle passed in the time $\triangle t_{i,m}$. This step prevents that fast vehicles generate less samples per segment than

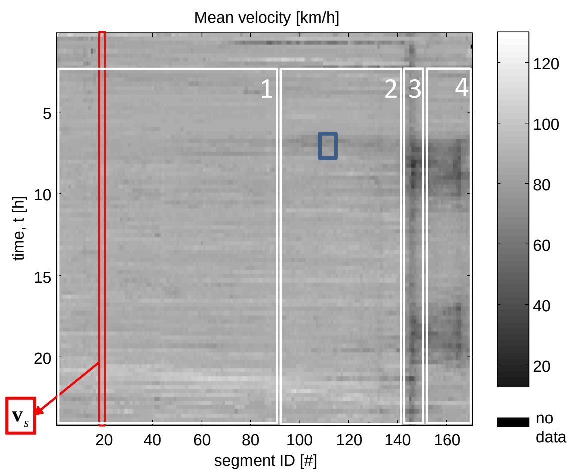

Fig. 2: Mean velocity map for the A4 motorway Milan–>Brescia. Black sections indicate lack of data, white boxes the 3 classes of segments.

slow vehicles. We do not use instantaneous velocity from GPS data as it is not averaged over $\triangle t_{i,m}$ and it is less reliable. Next steps are the average of associated velocities on each segment over a time interval $T = 15$min and, for each time sample of duration T in a daytime, the average over the 42 working days. Indeed objective of the paper is to classify segments based on their typical daily profiles during working days. The result of the aggregation is a daily speed profile for each segment s indicated as $\mathbf{v}_s = [v(s,t)]_{t=1}^{N_T}$, where $N_T = 96$ is the total number of time samples in 24h. The matrix collecting the profiles for all segments is $\mathbf{V} = [\mathbf{v}_1, ..., \mathbf{v}_S]$, where N_S is the total number of segments in the road. In case of poor data 2D linear interpolation is used to obtain a complete map.

Fig.2 shows the velocity map $\mathbf{V}$ for the A4 motorway Brescia→Milano computed on a total of $N_S = 170$ segments. Speed profiles present a number of features that are typical of the considered sources of data, i.e. trucks. The maximum velocity is the usual free-flow velocity of trucks on this motorway, about 90 km/h. In the early hours also higher speed can be observed, due to lower traffic volume. How we can observe from Fig. 2, data present high spatial correlation along the road. This spatial homogeneity of velocity suggests to associate in clusters segments with similar behavior over the day. We can recognize four typical profiles: on segments $1 \leq s \leq 90$, traffic is always in free flow; on segments $100 \leq s \leq 140$ a congestion in the morning rush hours in the can be observed

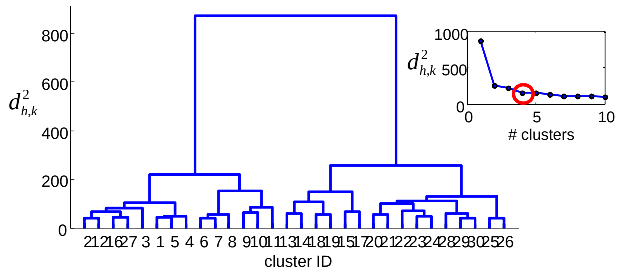

Fig. 3: Dendrogram representing the hierarchical construction of clusters using Ward's method. On the top the distance vs the number of clusters.

due to the increase of traffic when approaching the city; on $151 \leq s \leq 170$, congestions are observed in rush hours, this section corresponds to the north ring road of the city; on $142 \leq s \leq 150$ a lower mean speed is observed over the whole day, as these segments are located just before the toll station. In next section, we propose a procedure to automatically recognize these typical profiles. The classification starts from a preliminary analysis on the maximum velocities of vehicles to distinguish different vehicle typologies and avoid estimate bias.

Estimation of typical speed profiles

We classify vehicles based on their maximum detected speed, v_i^{max}, using ranges of 10km/h. For vehicle belonging to the same class (i.e., vehicles with maximum velocity within the same 10km/h range) we construct historical speed profiles. In this way profiles will not be polarized by faster or slower vehicles. We exclude from the analysis vehicles with maximum velocity $v_i^{max} \geq 120$[km/h] (cars), and $v_i^{max} \leq 70$[km/h] (slow vehicles). For identification of typical profiles we focus our attention on vehicles with $90\text{[km/h]} \leq v_i^{max} \leq 100\text{[km/h]}$ - the class with more vehicles, namely 1481 vehicles - and we consider the time window 2:30am-12pm on segments $1 \leq s \leq N_S$, where more data from probes are available.

Classification methods aim at separating the velocities profiles into N_h classes or clusters. Class h, for $h = 1, \ldots, N_h$, is associated with a typical speed profile $\mathbf{c}_h$ obtained as the average made on segment profiles $\mathbf{v}_s$ classified in that class. We select a hierarchical approach rather than a partitional one, as this approach does not require to choose in advance N_h; the number of classes follows directly from the classification procedure. We choose to employ the Ward's method [7] which has the advantage of providing homogeneous classes.

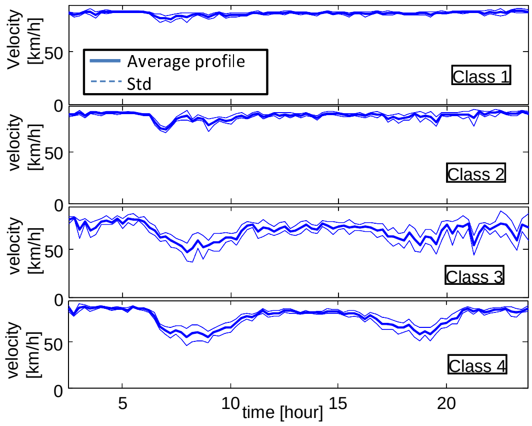

Fig. 4: Mean and std of velocity profiles.

Ward's method classifies speed profiles by constructing a tree in a "bottom up" approach. Each observation - segment profile $\mathbf{v}_s$ - starts its own cluster, then clusters are paired through the minimization of the normalized square distance metric:

$$d^2_{h,k} = n_h n_k \frac{\|\mathbf{c}_h - \mathbf{c}_k\|^2_2}{n_h + n_k},$$

where h, k are the cluster indexes, n_h and n_k are the number of components in clusters h and k. Pairing of clusters is repeated, until all observations are enclosed in one class. The procedure is illustrated in Fig. 3 by a dendrogram. Only the upper part of the total tree is shown because of lack of space. The number of classes is decided when a further increase of the number of clusters leads to a relative decrease of the normalized square distance below a given threshold (see top of Fig. 3).

The classification method applied to the available dataset for the test-case led to the identification of four clusters as shown in Fig. 4. The speed probability distributions for each cluster are in Fig. 5. The clustering into four classes is superimposed on the speed map in Fig. 2 by white boxes. We can observe that the method is able to automatically recognize the typical trends of the preliminary analysis made in the previous section.

Concluding remarks

In the paper we proposed a method to classify typical speed profiles on road segments starting from data collected by floating trucks. We proposed a procedure to construct the velocity field from traffic data provided by the probes. We made a preliminary analysis to recognize the vehicle types from the observed maximum velocity and to estimate the

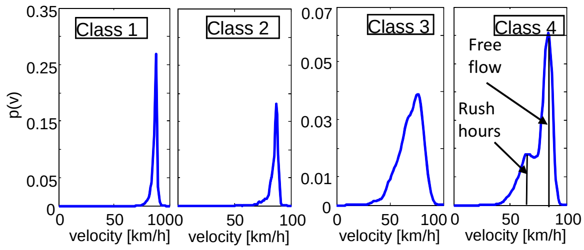

Fig. 5: Probability distributions of velocity.

typical speed profile. For classification of segments we employed a hierarchical technique based on the Ward's method that led to the separation of the dataset in homogeneous classes of speed trends.

Results of the case study presented in this paper showed that the proposed clustering procedure was able to identify the typical daily trends of velocity even from a dataset of moderate dimensions.Future development will be the extension of the analysis to a larger dataset in order to obtain a more accurate description of the traffic behavior over the segments and the different days. The proposed procedure is planned to be used for the construction of historical database for applications such as fleet management (e.g.,travel time estimation), navigation and incident detection.

References

[1] J. Yoon, B. Noble, M. Liu, "Surface Street Traffic Estimation", Proc. 5th MobiSys, San Juan, Puerto Rico, USA, June 11-14, 2007, pp.220-232.

[2] B.S. Kerner, C. Demir, R.G. Herrtwich, S.L. Klenov, H. Rehborn, M. Aleksic, A. Haug, "Traffic State Detection with Floating Car Data in Road Networks", Proc. 8^{th} IEEE ITSC, Vienna, Austria, Sept. 13-16 2005, pp.700-705.

[3] C. de Fabritiis, R. Ragona, G. Valenti, " Traffic Estimation and Prediction Based on Real Time Floating Car Data", Proc. 11th IEEE ITSC, Beijing, China, October 12-15 2008, pp. 197-203.

[4] J.S. Yang, "Travel Time Prediction Using the GPS Test Vehicle and Kalman Filtering Techniques", American Control Conference, June 8-10 2005, Portland, OR, USA, pp. 2128-2133.

[5] V. Sethi, N. Bhandari, F. S. Koppelman, J. L. Schofer, "Arterial incident detection using fixed detector and probe vehicle data", Transp. Res. Part C, Vol. 3, Issue 2, April 1995, pp. 99-112.

[6] Treiber, M., Kesting, A. and Wilson, R. E. , "Reconstructing the Traffic State by Fusion of Heterogeneous Data", Computer-Aided Civil and Infrastructure Engineering, Wiley, 2010.

[7] W. Weijermsrs, E. van Berkum, "Analyzing highway flow patterns using cluster analysis", Proc. 8th IEEE ITSC, Vienna, Austria, Sept. 13-16 2005, pp.831-836.