A SMART DATA APPROACH TO ANALYZE VEHICLE FLOWS

M. Mazzei*
CNR, Italian National Research Council - Institute of Systems Analysis and Computer Science – LabGeoInf - 00185 Rome, Italy - mauro.mazzei@iasi.cnr.it

Commission IV, WG IV/9

KEY WORDS: Smart Data, Smart Cities, Machine Learning, Transportation.

ABSTRACT:

In the logic of Smart Cities it is of fundamental importance to analyze the traffic situation through dedicated sensors and networks. According to this approach and through the potential of smart data is based this study. Improve prediction of traffic patterns by analyzing and counting vehicles in a virtualized scene in real time. In the past, the technique of hardware inductive coils was used that were dropped in the asphalt to exploit the principle of magnetic induction in order to verify the transit of vehicles. This technique is not able to classify vehicles or estimate their speed, unless using multiple inductive coils. The proposed system provides for the virtualization of an area of interest which requires a selection and mapping of the areas where the control areas are to be included. The “image detection” techniques allow us to classify the vehicles in transit. With the techniques of “machine learning” can to able to verify the flow, count the vehicles present in the scene and classify them by vehicle type in real time. The vehicle counting and classification data available in the cloud platform allow to model and update the main nodes of the network in order to improve the prediction and estimates of the best routes of the road network according to the degree of saturation of the flows and the length of the line of the graph. The model can also indicate additional information of an environmental nature in an ITS system present in the cloud.

1. INTRODUCTION

This work aims to explore some object detection technologies and to produce a model capable of identifying, in real time, cars and motor vehicles on the road, starting from the processing of images acquired by appropriately positioned cameras.

The identified vehicles are tracked and counted, opening up to the possibility of using the data for statistical purposes, traffic control and surveillance and to proposals of alternative routes. The scenario I hypothesized that involves monitoring vehicle flows through fixed cameras that frame an area of road, the complexity of a traffic monitoring system varies in relation to the objectives of the monitoring itself, the technologies adopted, the methods used, the transfer and information processing. Generally they are divided into three main components for an ITS system: quantities to be detected (traffic variables), data collection techniques, functions that allow data analysis. In my scenario I hypothesized that to use automatic detection techniques through video images based on the use of fixed cameras, which allow to continuously detect the traffic scenes that take place on a road section. The cameras offer a spatial-temporal representation of the vehicular outflow, therefore it is necessary to process the video sequences through methodologies that allow an interpretation of the content of each image (spatial analysis) and their correlation in the temporal sequences (temporal analysis).

To design an automatic detection system with video images it is necessary to use specialized hardware and software, especially if consider to automatically analyze the video sequences in real time, important aspects concern the transmission and reception of data, number of cameras per control system, identification of areas for monitoring and finally the processing and analysis of “image detection” in real time. The results of a good classification of objects are obtained by processing images or video streaming with algorithms based on Artificial Intelligence.

1.1 Vehicle counting and virtualized scenes

Artificial Intelligence techniques based on algorithms able to train neural network models allow to identify and classify objects framed by low-cost fixed cameras, the overlap of virtualized scenes on real scenes allows to monitor changes with respect to an area of check. By virtualized scene, mean an overlap of polygons in a digital environment that identify an area on the real scenes acquired by the fixed cameras, these virtualized scenes allow the counting of the vehicles present in the area of interest, in a certain way a bit like it happens. With the hardware inductive coils that are drowned in the asphalt to exploit the principle of magnetic induction in order to verify the transit of vehicles. The classic inductive coils are not able to classify vehicles or estimate their speed, unless multiple ones are used, moreover, these systems are overrnn and very expensive.

The meet system of virtualized areas and real data acquired by fixed cameras that frame the selected area, turns out to be an innovative and economically more advantageous approach than inductive coils techniques. Telecommunications systems and in particular 5G technology, into the low latency on the communication network, allow a wide possibility in the exploitation of digital systems to support of monitoring vehicle flows in real time. This type of systems requires strong integration in the implementation and development phase,
especially in real-time environments. As regards the data detection phase, “image detection” techniques are used to classify the vehicles in transit as they pass the analyzed area (see Fig. 1).

**Figure 1.** Example of a virtualized area at point A at time t1 taken by a fixed camera.

### 1.2 Theoretical aspect

In theory, some useful information is provided to have a better knowledge of the method used. The fundamental relationship of vehicular traffic represents the join existing between the macroscopic variables of the outflow, i.e. density, velocity and flow (respectively k, v and q), expressed by the identity:

\[
q = k \times v
\]  

(1)

If consider a selected area of unit length, potential mismatches between the variables emerge. The flow is, in fact, substantially referred to the transits that cross the area in a given time interval, and as such they can be measured in a section of the same, while the density considers the number of vehicles that are on the area in an instant of time, and therefore it is measurable along the selected area. The speed of the flow was introduced as the average speed of the traffic units that compose it, but it is not defined how this average should be considered, i.e. whether as an average of the vehicle speeds in the section or as an average of the vehicle speeds in the area.

To solve this problem, Edie (1963) proposed the following definitions for the flow q, the density k and the mean velocity v, valid for a region A in the space (x) - time (t) plane:

\[
q = \frac{\sum_{i=1}^{N} x_i}{t_i} \\
\]

\[
k = \frac{\sum_{i=1}^{N} f_i}{|A|} \\
\]

\[
v = \frac{\sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} f_i}
\]

(2)

where \(x_i\) is the distance traveled by the i-th vehicle between the \(N\) present in region A, \(t_i\) is the time spent by the i-th vehicle between the \(N\) present in region A and \(|A|\) is the (Euclidean) area of region A, having the dimension [length] [time], with the length usually corresponding to the physical length of the segment (in km) regardless of the number of lanes. The average velocity calculated according to the previous equation is commonly called the average spatial velocity (SMS - space mean speed) and denoted by \(v_s\), which satisfies the fundamental relation of traffic:

\[
v_s = \frac{\sum_{i=1}^{N} x_i \cdot v_i \cdot dt}{\sum_{i=1}^{N} f_i \cdot dt} = \frac{1}{N_s} \sum_{i=1}^{N} v_i
\]

(3)

In these terms, the number of traffic units that cross a cross section in the unit of time (flow rate \(q\)) is equal to the product of the number of units present in the unit of length (density \(k\)) by the distance covered by them, in the unit of time (speed \(v\)). If consider the vehicles transited in the area selected at point A in time \(t1\) and the transit speed, I can calculate the average travel speed from point A to a hypothetical point B in time \(t2\). In this way I can deduce the vehicular outflow of the analyzed line of road as long as there is no condition for any other alternative routes.

Having said this with the techniques of “machine learning” are able to count all the vehicles present in the scene and classify them by vehicle type, so that upon passing point B at time \(t2\), I can also determine the classification of the vehicles in circulation.

### 2. RELATED WORK

Most of the ITS systems allow you to estimate the arrival time of the journey, possibly suggesting an alternative route. Many works propose solutions of robust graph models with refined processing techniques and algorithms that analyze the data and estimate the best predictions, but do not take into account the effects due to the monitoring of critical nodes, since they are not based on real data processing time.

Many researchers have extensively studied network traffic prediction models. The traditional Markov model is a typical short-range dependency model that fails to meet the demands of long-range dependency network prediction. With the development of machine learning, prediction models such as the posterior propagation neural network and the supporting vector machine have emerged. Although the neural network can very well represent the long-range dependency and the short-range dependence of network traffic, it suffers from difficult parameter selection, low convergence rate, and susceptibility to premature convergence, making it unsuitable for real-time forecast of network traffic. The SVM model can be a solution only for machine learning with small samples and is difficult to implement for large-scale training samples. The most widely applied predictions models are time series models, which are divided into moving average model, stationary time series model, and exponential smoothing model. The moving average model can effectively eliminate random fluctuations in the forecast, but it can only use the average value of the last actual dataset for the predictions, which does not consider subsequent trends, making it suitable for instant predictions where the data does not they neither increase rapidly nor decrease rapidly. The search for network traffic characteristics is based on the theory of time series analysis, which is aimed at models of traffic prediction. The time series are an objective record of the...
historical behaviours of the system under study, which consists in analyzing these historical records, discovering the statistical dependence between the data, grasping the structural characteristics and operating rules of the system and predicting the future behavior of the system. Research shows that the network traffic model has statistically self-similarity, which can accurately describe network traffic, correctly establish functions and dependencies on existing data, analyze intrinsic data models, and control and predict future data.

Google Maps estimates the duration of the journey and therefore the arrival time based on traffic conditions in real time. To predict future traffic conditions, including delays that can increase the duration of the journey, Google also takes into account historical data (for example, it is known that traffic congestions occur at certain times of the day on certain roads). Traffic prediction is done through a machine architecture learning called “Graph Neural Network”.

In this way the forecasts are very influenced by the historical results and by the saturation of the flows that is meets along the way, but they do not take into account that the historical flows can adapt in real time and this information is not recorded in the necessary time.

My proposal is to model the graphs starting from the analysis of O/D matrices that allow accurate analysis of daily and peak traffic conditions as a historical model, however, I believe it is of fundamental importance to fit this model with the analysis of crucial nodes with sensors that can verify the flow of vehicles in real time (through cameras and fixed sensors) and in turn modify the forecast model with constantly updated data.

With this work I propose to integration of two systems, the flow monitoring detection system (deterministic system) from the evaluation, and system of the traffic model (stochastic system) based on historical data, and use the results in a targeted way. The real-time updating of monitoring data from sensors will facilitate the updating of the traffic forecasting model, especially in the evaluation of the most significant nodes in the monitored graph through the installed technology.

3. PROJECT

3.1 Development with AI

The development of the project (SUMMa: Smart Urban Mobility Management - Support program for emerging technologies - Ministry of Economic Development - Italy) was to create a low-cost software model in an open source environment able to receiving as input a video of a traffic scene on the road and identifying and classification between visible objects, categories of cars, trucks and motorcycles, etc. To obtain these results, it was decided to use a system artificial intelligence and deep techniques learning, be able of obtaining much better results. The choice of the classification system has been oriented on the Yolo system on Darknet architecture, all in an open source environment, this system also from the literature in this regard is particularly suitable for object detection in real time.

The activity conducted is derived from the need to evaluate the issues of the mobility system in the experimentation area of the SUMMa project based on the study of specific models and algorithms, both in terms of current scenarios and in the perspective of planned mobility growth in future years. The availability of current information technologies in terms of traffic assignment and planning makes it possible to create computerized models of the transport network for be able of estimating the effects resulting from planned interventions and evaluate its effectiveness and impact. Through the use of simulation models, in fact, it is possible to assessment the systemic effectiveness of the proposed interventions and the localized performance provided by each single element of the road network (arcs and nodes).

Traffic and transport assignment and simulation models now play a central and irreplaceable role in transport planning and represent a fundamental element in addressing logistical and settlement choices, as well as providing valid support in the field of communication. The profound evolution of modelling applications, together with the computational potentials now achieved by computer systems, I can now to improve the number of variables considered and the quality of the results that can be obtained. Traffic simulation models are the only tools are able of representing the states of circulation (and of the induced environmental conditions) in a continuous procedural form, and therefore are able of allowing a degree of surveillance on the evolution of the systems while this evolution is taking place.

Today the computerized models of the road network have been refined in the technical level and in the use strategies, it is possible to evaluate the effects of planned interventions, not only in terms of transport, but also environmental and urban planning.

Within a perspective of sustainability, the modelling applied to transport allow to evaluate the loads of sound and polluting emissions connected to vehicular traffic, the variation in behaviour in terms of transport costs between two or more settlement or infrastructural systems. alternatives, and depending on the transformations in the network and in the nodes, the probabilities of saturation of the load capacity of a given territorial area.

A fundamental element of the quality of a traffic simulation model is constituted, in general, by the reliability of the basic information support with the current state of demand of mobility. These databases, from which all subsequent analyzes and processing are derived, are available and of good quality (ie corresponding to the observed reality), the more precise the results and estimates from the traffic simulation model will be. commonly a “system” is understood as a set of connected parts that can interact to form a functional unit. I believe it is useful to give a more precise definition which will be useful in the following: I consider “system” a set of parts that communicate with each other, whose activity is aimed at achieving a result. The communication between the parts of the “system” is aimed at the implementation of the necessary coordination of the activities that allow to achieve the expected result. Communication between the parts of the system involves different types of interaction such as the transmission (transport or spatial transformation) of material goods, people, messages. And just as the activities of the parts of a system are always connected to a place, the same thing happens for communications. The first step I want to define a system is to recognize the activities connected by communications. A system can be represented graphically with a connected graph, whose nodes are representative of the parts of the system and the connection arcs are representative of the communication.
channels existing between the parts of the system. Such canals can represent constructed connections, such as roads, trails, railways, canals, pipelines, telephone cables, or natural connections such as rivers, air corridors, mountain, bottoms of valleys.

The activities of the parts of a system, in order to get the expected result, must be subjected to a controlled regulation of the error. This regulation (feedback) is implemented through a control mechanism provided with information on the actual state of the system, compared with the desired state. In my case study, the zones considered correspond to the system i wish to control; the desired states are those possibly present in a plan, while can evaluate the actual state of the system through various forms of surveys, including census surveys on population, employment, services, housing. Cities can be influenced by adding, removing or altering parts of the system or connections between components, thereby affecting land uses and their communications.

This can be achieved in two ways. Firstly, through public intervention relating to better planning of places such as hospitals, schools, residences, services, roads, parking, etc. Secondly, indirectly, by better regulating the processes of control and regulation of development. It is clear that must have means that allow us to predict the effects of my actions, since by the time they occur, the system may have gone beyond the pre-established limits and corrective action may be too late. The means that allow us to predict the effects of my actions are the simulation models that increase and expand the planner's experience, indicating the need for any corrective actions and allowing the experimentation of different solutions. Since the temporal variable can be introduced in the simulation models, it is also possible to verify the different aspects of the actions in the short, medium and long term. I must consider that the communication of multiple systems (deterministic and stochastic) can improve both the data acquisition phases and the actual needs between offer and demand, especially in the context of transport solutions, and the modelling and predictions phases of a more accurate solution to urban transport problems.

3.2 The transport models

The mathematical models of transport supply systems use on the one the theory of graphs and networks to represent the topological and functional structure of the system and on the other the results of different engineering disciplines to describe the “performances” and interactions of the elements that compose it.

This models have a dual function. The first is to allow simulating the performance of transport services for users and the impacts on the external environment. The performances offered by a transport system in a certain area, i.e. the service level attributes, such as times, costs, etc., are used for the simulation of the demand for mobility while the external impacts are used in the design and evaluation phase interventions. The second function of the models is to participate, in the simulation of the flows that the various elements of the offer system in the reference period. Both of these functions require very efficient computational algorithms given the size that networks can expand to in applications.

3.3 CNNs: convolutional neural networks

Convolutional neural networks (CNN), are a specialization of artificial neural networks, with characteristics that make them the best solution for image classification applications. The fundamental aspect that characterizes them is the use of the convolution operation that gives the network properties that type of applications can successfully.

The neural networks are composed of layers mainly of the Dense type (the classic layers completely interconnected with each other) or of the Dropout type (layers that contrast the overfitting phenomenon). CNNs will introduce new types of layers including those of Convolution, ZeroPadding, MaxPooling. The technique of convolutional neural networks came into global focus in 2014, in the full phase of expansion of the deep learning, when K. Simonyan and A. Zisserman made a decisive contribution to the object recognition technique with their “Very Deep Convolutional Networks for Large-Scale Image Recognition”. Another decisive step occurred in 2017 when G.Huang , Z.Liu and L.van der Maaten proposed their “Densely Connected Convolutional Networks”, in which it was stated that deep learning and deep neural networks could not fully benefit from the effects of backpropagation, and the first layers lost their ability to learn simple low-level patterns.

The cause was to be attributed to the concatenated structure of the layers in the network, therefore a new approach was proposed that added to the traditional concatenated path the possibility for each layer to reach all the others. The convolution operation that makes this possible is a mathematical calculation that is applied to a tensor. The matrix object of the convolution is “scanned” by a second smaller matrix, called kernel or filter. The matrix product at each step generates the final result.

3.4 Deep learning on Cloud

The algorithms of deep learning requires a great deal of hardware resources. The possibility of using cloud services, in some cases free, offers the opportunity to work on artificial intelligence even to those who are not equipped with up-to-date hardware. Each of the main services currently available on the cloud has its specific peculiarities, but all of them guarantee concrete advantages, including the ability to work on very powerful machines from anywhere, maximum flexibility, already optimized configurations. A network was designed and trained with a dataset downloaded from Open Images v6.

Through the cloud service it is possible to request the availability of GPUs or TPUs, trying not to abuse when not necessary because an algorithm limits its use to distribute it equally among users. The developed application takes care of processing the video file and operating the object detection, selecting between three different classes of objects: cars, trucks and motorbikes. It provides for the counting of objects and offers a timer controlled by the user to define the counting window. I chose to use OpenCV DNN with the Yolov4 model pre-trained with the Darknet55 framework. The analysis of a video is nothing more than the analysis of many frames one after the other at the maximum possible speed for the machine.
3.5 The SORT algorithm

To be able to track the objects, I used the SORT algorithm, which uses the Kalman filter. Once an object has been identified, it analyzes speed and acceleration to be able to predict with a certain degree of confidence where it might be at the next instant. Used within an object code tracking offers great potential. The algorithm is initialized at the beginning of the program, then it is constantly updated in the program loop. Once the object is uniquely identified and tracked, with a unique id assigned, the program can count the objects that transit within a certain area. The area is defined as a polygon which is then not shown on the screen. When the object coordinates are inside the area, the program increments the counter and updates all the data on the screen. I have implemented a user-controlled timer, which can start and reset it. The data shown on the screen include the total of each class of objects in both directions of circulation, and the average transit of objects per second.

Shown at the top left in Fig. 2 are the statistical data of vehicle count and classification and the average vehicular flow. Another relevant aspect with this approach is that for a section of road, only one camera is needed to count and classify vehicles in both directions of the road, this implementation simplifies the detection system and reduces its cost considerably on the devices to be installed.

The sensed data is automatically available on the cloud (deterministic system), and as I explained above, this data is also available to improve the analysis of prediction model (stochastic system) that loaded this data. The implemented system has shown to have good object detection capabilities. Figure 2 shows a screenshot of the running system.

4. ANALYSIS AND MODELLING

4.1 Evaluation of traffic flows

In this paper the development of traffic models of the road network is described, below is the conceptual model developed. The data present on the cloud allow in real time to identify the main nodes (see figure 3 as an example) for sorting the traffic which is then model within the graph. These remarks foreplay must be validated come on counts of the traffic in order to better calibrate the traffic pattern.

Figure 3. Graph of an urban centre where the nodes of the virtualized monitoring areas are highlighted.

4.2 Automatic network optimization procedures

The procedure implemented searches, starting from an initial network, the optimal configuration (from the traffic point of view) of the road network, generating a better configuration than the previous one at each iteration. As is evident, it leads to the determination of a local optimal point. The objective function to be minimized is the demand-weighted travel time between OD pairs per OD pair:

\[
F(A) = \min_{A} \sum_{od} d_{od} \cdot t_{od}(A)
\]

where: 
A = indicates a network configuration according to the direction of travel 
\(d\) = demand flow between zone o and zone d 
\(t_{od}\) = average travel time at equilibrium to go from o to d in correspondence with the network configuration A, function, among other things, of the flows.

The average travel times between pairs or d are calculated with an incremental assignment of the origin-destination demand to the configuration itself.
4.3 Excellent Local Search Methodology

Once the traffic flows have been assigned to the network, the degree of saturation weighted on the flows and on the length of the branches \( g_{Sp} \) is calculated for each arc of the aforementioned set, given by:

\[
g_{Sp} = \frac{f}{C} \cdot f \cdot l
\]

where:
- \( g_{Sp} \) = degree of saturation weighted on the flows and on the length of the branches
- \( C \) = bow capacity
- \( f \) = flow over the arc
- \( l \) = length of the arc

A list of arcs is created in descending order of \( g_{Sp} \). The arcs for which the \( f / C \) ratio exceeds a predetermined limit value are identified. The first branch of the list is extracted, the extracted arc is "preserved" and the reverse is eliminated, thus choosing the direction of travel for each street. Each time an arc is eliminated, the connection is checked; if the check is negative, the opposite direction of travel is chosen. The procedure diagram is shown in the following figure 4.

![Flow chart algorithm](image)

4.4 Logical flow of the optimization procedure

When starting the optimization procedure, it is possible to fix the value of the \( f / C \) ratio. Consider that, for the trend that the cost functions have, the default value is assumed to be \( f / C = 0.75 \). For each configuration obtained following the intervention, the value of the objective function and its difference with respect to the initial value are calculated; if among these configurations there are one or more that have improved the value of the objective function with respect to the previous one, the one corresponding to the maximum improvement is assumed as a new configuration and a new iteration is carried out. Continue until you reach a network configuration for which any variation made no longer causes an improvement in the scheme. The convergence of the algorithm is ensured by its structure and by the fact that the number of solutions is finite; in fact, since the number of solutions is finite and a different solution is generated in each step, which corresponds to an improvement in the objective function, in the worst case scenario all the solutions will be explored, but in any case a finite number of iterations will be carried out. Finally, the calibrated model can provide a traffic light system managed by ITS always active in the cloud with the correct saturation information of the arcs to adapt in real time with respect to the state of the traffic.

5. CONCLUSION

In conclusion, this work aims to demonstrate how the integration of two systems communicating through the cloud, the first that monitors the flow and the second system receives the data from the first to best estimate the optimal solutions. This approach can be a valid solution for all ITS techniques aimed at analyzing vehicle flows to improve circulation, especially in urban centres where there is more density.

The integration of aspects related to sensors, artificial intelligence, prediction model and communication offer solution to improve a logic of smart data with the aim of making smart cities. It must be recognized that it is of fundamental importance to have knowledge of several specialized sectors and the integration of these areas with smart data offers enormous possibilities for intelligent solutions in the smart cities, in the not too distant future these systems will also be used to improve the lifestyles and well-being of people in cities, for monitoring environmental pollution by adding ad hoc sensors, this last activity will be a starting point for the next work.

REFERENCES


European Commission: Sustainable and Smart Mobility Strategy - putting European transport on track for the future. COM (2020) 789 final (2020)


Maldonado Silveira Alonso Munhoz, PA, et al.: Smart mobility: the main drivers for increasing the intelligence of