PhD Thesis

Data mining, optimization and simulation tools for the design of Intelligent Transportation Systems

Universidad de Castilla-La Mancha
June 2015

Author
María Teresa Sánchez Rico

Advisor
Ricardo García Ródenas, PhD
A mis padres y
hermanas
Abstract

Over the past years, the problems caused by traffic congestion in urban areas are increasingly severe and visible. Accidents, traffic jams and environmental damage are only some of the negative externalities caused by traffic congestion that today concern society.

The so-called Intelligent Transportation Systems (ITS) are multidisciplinary tools which involve the application of advanced technologies and analytical approaches with the purpose of solving some of the referred transport problems. In the last decades different initiatives and projects have emerged in this field. Nevertheless, there is currently no proposal about a generic framework of ITS to be adapted in different traffic contexts.

In this thesis, different mathematical and optimization models as well as simulation and data mining approaches are provided with the aim of contributing to the ITS field. The initial motivation for the development of this thesis comes from considering that nowadays any ITS must be capable of fulfilling four main premises: scalability, technological independence, robustness and facing the use of big data.

The main contributions of this thesis are:

• **Literature review.** A literature review on the current advances in the field of ITS and its classification is done. Furthermore, a proposal for an ITS prototype for traffic control and management is carried out.

• **An approach to face the continuous dynamic network loading problem.** A continuous DNL model based on flow discretizations, instead of time discretizations is presented with the main goal of achieving a trade-off between precision and computational cost.

• **A proposal for the enhancement of traffic signal optimization.** A bilevel optimization model based on Time-Of-Day (TOD) intervals for traffic signal timing is proposed to address simultaneously the segmentation problems and the traffic control problems over these time intervals. The model has been solved by the use of efficient metaheuristic algorithms.

• **An approach for daily traffic patterns identification.** A prototype of urban traffic control system based on a prediction-after classification approach is presented. In an off-line phase, a repository of traffic control strategies for a set of (dynamic) traffic patterns is constructed through dynamic cluster techniques. In an on-line phase, the current daily traffic pattern is predicted within the repository and its associated control strategy is implemented in the traffic network.

• **A proposal for the design of ITS.** A proposal for an ITS for traffic control and management is presented by using TOD intervals. This approach replaces the short-term traffic predictions by a finite sequence of stationary states within a set of given traffic patterns.
Resumen

En los últimos años, los problemas que causa la congestión del tráfico en las ciudades se muestran con mayor intensidad. Accidentes, atascos y daño al medio ambiente son solo algunos ejemplos de los aspectos negativos que origina el constante incremento del tráfico urbano.

Los denominados Sistemas Inteligentes de Transporte (ITS) son herramientas multidisciplinarias que integran avances tecnológicos y métodos analíticos, con la finalidad de resolver algunos de los problemas descritos anteriormente. En la última década han surgido diversas iniciativas y proyectos en el campo de los ITS, a pesar de esto, hoy en día no existe alguna propuesta que aborde un marco genérico para el desarrollo de un ITS el cual pueda ser adaptado a distintos usos.

En esta tesis doctoral se han desarrollado diversos modelos matemáticos y de optimización, así como métodos de minería de datos y simulación con el objetivo de aportar herramientas analíticas al campo de investigación del diseño de ITS. La motivación inicial ha sido considerar que hoy en día cualquier ITS debería cumplir cuatro premisas básicas: escalabilidad, independencia tecnológica, robustez y hacer frente al uso de grandes cantidades de datos.

Las contribuciones de esta tesis se sintetizan en:

• **Revisión de la literatura.** Se ha llevado a cabo una revisión de la literatura sobre la clasificación de los ITS y de los avances actuales. Además se ha propuesto un prototipo de ITS para el control y la gestión del tráfico.

• **Un modelo de simulación para el problema de carga dinámica de tráfico.** Se ha propuesto un esquema de simulación para el problema de carga dinámica de tráfico basado en la discretización de flujos, en lugar del tiempo. Este aspecto es novedoso en la literatura. El objetivo principal ha sido conseguir un equilibrio entre la precisión del modelo y el coste computacional de su resolución.

• **Metodología para el diseño de planes prefijados de control de tráfico.** Se ha desarrollado un modelo de optimización binível basado en los denominados *Time-Of-Day* intervalos para la definición de planes de control de tráfico (por ejemplo control semafórico) con el objetivo de abordar de forma simultánea la segmentación temporal del tráfico y su control sobre estos intervalos de tiempo. El modelo ha sido resuelto eficientemente con varios algoritmos metaheurísticos.

• **Metodología para la identificación de los patrones diarios de tráfico.** Se ha desarrollado un prototipo de sistema de control de tráfico urbano basado en un enfoque de clasificación/predicción. Este sistema consta de una fase *off-line*, en la que se construye un repositorio con estrategias de control para un conjunto de patrones dinámicos de tráfico. Para
Este fin se han desarrollado técnicas cluster dinámicas basadas en distancias espectrales. En una fase on-line, el patrón diario actual es detectado dentro del repositorio y su estrategia correspondiente de control es implementada en la red de tráfico.

- **Un marco para el desarrollo de ITS.** Se ha desarrollado una propuesta de un ITS para control y gestión del tráfico mediante el uso del concepto de intervalos TOD. Esta propuesta reemplaza las predicciones de tráfico a corto plazo por un TOD dentro de un conjunto de patrones. Esto permite considerar el comportamiento dinámico del tráfico como una secuencia finita de estados estacionarios. Se ha focalizado el sistema al problema de recomendaciones de ruta a los usuarios.
Agradecimientos

Hay muchas personas a las que quisiera agradecer la realización de esta tesis doctoral, ya que sin ellas, incluso estando lejos, esto no hubiese sido posible.

Antes que nada, quiero agradecer al gran responsable de este trabajo, mi director de tesis Dr. Ricardo García Ródenas por todas sus enseñanzas, por su motivación y apoyo de forma desinteresada a lo largo de este camino. No es fácil resumir la gratitud y admiración que siento.

A mis compañeros del grupo MAT. Use, José, Mari Carmen y Mari Luz. Gracias por su paciencia, por sus conocimientos, por sus consejos y sobre todo por su amistad. Han sido mi ejemplo de perseverancia cada día.

A todos los que he conocido en mi paso por este laboratorio, desde mi llegada hasta el día de hoy, en especial a Alfredo, Víctor, Abril, Mateus y Christian. Qué gran experiencia haber coincidido, gracias por haberme echado una mano siempre que lo he necesitado.

A todos esos profesores del Tecnológico de Morelia y de la Universidad Michoacana de San Nicolás de Hidalgo, que despertaron en mí la inquietud por la investigación y que me han apoyado para que esto fuese posible. Gracias en especial al Mtro. Juan Carlos Olivares y al Dr. Juan Flores por las facilidades durante mi estancia de investigación en el año 2013.

A mis padres, Enrique y Azucena, y a mis hermanas, Isa y Rocío, por soportar la distancia y sobre todo gracias por creer en mí y por estar cada día y a cualquier hora disponibles para apoyarme y escucharme siempre. Mami, nada de esto sería posible sin ti.

A mis amigos de Morelia, porque desde la distancia también han sido un gran apoyo. Quiero agradecer en especial a Mike, Jorge, Fanny, Elsa e Iris, gracias por tanto cariño, por las videollamadas, los chats, los mensajes y en general por haber vivido conmigo, aunque sea en la distancia, cada etapa de esta andadura.

A mis amigos de Ciudad Real por tantas experiencias, los de aquí y los que también vinieron de fuera. Gracias por involucrarse y por estar pendientes de mí y de mis avances en la tesis. Gema e Iru, gracias por su gran amistad y por integrarme en su familia como una más.

A Tomi, por ayudarme tanto, por ser mi ejemplo y motivación, por escucharme durante horas en los momentos más difíciles y darme calma. Definitivamente, todo esto ha sido mucho más fácil por ti.

Por último, a la Universidad de Castilla-La Mancha y a la Escuela Superior de Informática por los medios y las facilidades durante estos años.

Esta tesis se ha podido desarrollar gracias Consejo Nacional de Ciencia
y Tecnología (CONACYT) del gobierno mexicano en conjunto con el Consejo Estatal de Ciencia, Tecnología e Innovación del estado de Michoacán (CECTI) dentro del programa de becas "Formación de Recursos Humanos de alto nivel en el Extranjero". Además el proyecto PEIC-2014-003-P de la Junta de Comunidades de Castilla-La Mancha ha financiado parte de la investigación realizada en este trabajo.
Contents

1 Introduction 3
  1.1 Motivation and thesis objectives ....................... 5
  1.2 Summary of contributions and thesis outline ............. 7

2 Intelligent transportation systems 15
  2.1 Taxonomies of intelligent transportation systems .......... 16
    2.1.1 A classic taxonomy ................................ 16
    2.1.2 A taxonomy from a point of view of the market areas ... 18
    2.1.3 A taxonomy according on their strategic objectives .... 20
    2.1.4 Current examples of ITS projects ...................... 21
  2.2 A prototype for the design of ITS for traffic control and management ...................................................... 23

3 Traffic modelling 29
  3.1 Introduction to the traffic flow theory and its modelling .. 29
    3.1.1 Basic concepts of traffic flow ......................... 31
  3.2 Dynamic traffic assignment .................................. 35
    3.2.1 The structure of the classic four-step model .......... 36
    3.2.2 Challenges for dynamic traffic assignment ............. 37
    3.2.3 Characteristics of traffic demand ....................... 40
      3.2.3.1 The demand intensity ............................. 40
    3.2.4 Characteristics of traffic supply ....................... 41
      3.2.4.1 Traffic network representation ..................... 41
      3.2.4.2 Requirements for continuous dynamic network loading problem ............................................ 42
      3.2.4.3 Analytical macroscopic link models ............... 44
    3.2.5 Reconciliation mechanisms between demand and supply 48

4 A Monte Carlo approach to simulate the stochastic demand in a CDNLP 53
  4.1 Introduction ................................................. 53
  4.2 A continuous dynamic network load model .................... 56
    4.2.1 Nomenclature ............................................ 56
    4.2.2 Formulation of the basic problem ....................... 58
    4.2.3 Adopted link dynamic model: a generalized point queue model ............................................... 59
  4.2.4 Stochastic demand .......................................... 64
  4.3 Proposed solution procedure ................................. 65
    4.3.1 Discretization of link dynamic model .................... 65
      4.3.1.1 Dynamic of centroid connectors .................. 65
      4.3.1.2 Dynamic of road links ............................ 67
    4.3.2 Discrete event algorithm for moving packets within traffic network ........................................... 68
    4.3.3 Monte Carlo simulation .................................. 70
  4.4 Numerical experiments ....................................... 70
4.4.1 Computational burden ........................................ 72
4.4.2 Example of deterministic demand ............................ 72
4.4.3 Example of stochastic demand ................................. 73
4.4.4 Distributed computing ......................................... 75

5 Time series data mining to enhance traffic signal optimization 81
5.1 Introduction .......................................................... 81
5.2 Mathematical formulation ......................................... 83
  5.2.1 Model to identify time-of-day breakpoints .................. 83
  5.2.2 A combined approach of TOD determination and optimal traffic signal timing ........................................... 87
5.3 Hybrid meta-heuristic algorithm for the TOD problem ....... 88
  5.3.1 A simulated annealing (SA*) for TOD problem ........... 88
5.4 Determination of number of TODs ............................... 90
  5.4.1 Bayesian Information Criterion (BIC) ......................... 90
  5.4.2 ΔBIC .............................................................. 91
  5.4.3 A modified PETE algorithm .................................. 92
  5.4.4 An approach oriented to the TOD problem ................ 92
5.5 Computational experiments ......................................... 92

6 An approach to dynamical classification of daily traffic patterns 105
6.1 Introduction .......................................................... 105
6.2 Methodology .......................................................... 109
  6.2.1 A prototype of urban traffic control system ............... 109
  6.2.2 A clustering method to identify daily traffic demand patterns ................................................................. 111
    6.2.2.1 A k−means clustering algorithm ......................... 113
    6.2.2.2 Determining the number of clusters ................... 117
  6.2.3 Dynamic classification of daily traffic pattern ............. 118
6.3 Implementation ........................................................ 120
  6.3.1 Experiment 1: A simulation study for assessing the proposed methodology ........................................... 120
    6.3.1.1 Data description ........................................... 120
    6.3.1.2 Clustering of daily traffic demand patterns in synthetic data ...................................................... 122
    6.3.1.3 Dynamic classification for synthetic data ............... 126
  6.3.2 Experiment 2: Implementation with real data ............. 128
    6.3.2.1 Data description ........................................... 128
    6.3.2.2 Clustering of daily traffic demand patterns in real data ...................................................... 129
    6.3.2.3 Dynamic classification for real data ................... 132

7 A proposal for the design of an ITS for traffic control and management 139
7.1 Introduction .......................................................... 139
7.2 Related literature .................................................... 140
7.3 A proposal of an intelligent traffic control system ............. 142
7.4 Numerical experiments .............................................. 144
  7.4.1 Route optimization .............................................. 145
## Contents

7.4.2 Experiment 1: Route recommendations .......................... 147  
7.5 Experiment 2: ITCS evaluation .................................. 150

### 8 Conclusions and further work

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1 Conclusions</td>
<td>157</td>
</tr>
<tr>
<td>8.2 Suggestions for further work</td>
<td>160</td>
</tr>
<tr>
<td>8.3 Conclusiones</td>
<td>163</td>
</tr>
<tr>
<td>8.4 Futuras líneas de investigación</td>
<td>166</td>
</tr>
</tbody>
</table>

### Bibliography

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Dynamic capacity given by a square wave</td>
<td>185</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Statistics on passenger vehicles per 1000 persons. Credited: 
Statistical PocketBook (2014), The International Council on Clean Transportatio
European commission. .................................................. 4
1.2 The vicious circle of mobility ...................................... 5
1.3 Budget allocation for the research theme “Intelligent Trans-portation Systems” in Europe. Source: Rafiq et al. (2013) .... 8
1.4 Thesis outline ........................................................... 9

2.1 A taxonomy from the point of view of the market areas and its roughly corresponding classic area ....................... 19
2.2 Prototype for the design of ITS for traffic control and manage-
ment ................................................................. 23

3.1 Mathematical modelling ............................................ 30
3.2 Methodology based on models .................................... 31
3.3 Trajectories and measurement intervals S₁ and S₂. Credited: 
Immers and Logghe (2003) ........................................... 32
3.5 Traditional four-steps schema for traffic modelling .......... 38
3.6 Example of road network in Ciudad Real (Spain) ........... 38
3.7 Traffic network representation .................................... 42
3.8 Road discretization used in the cell transmission model ..... 45
3.9 Example of trapezoidal function for Q(k) used in cell transmis-
sion model ............................................................. 47

4.1 Diagram of our proposed link divided into two sections ...... 61
4.2 The O-D demand intensity functions for the Nguyen-Dupuis 
network .............................................................. 71
4.3 Nguyen-Dupuis network: travel times on the different paths of 
the illustrative example .............................................. 73
4.4 Nguyen-Dupuis network: Dynamically varying flows over time in two links ................................................. 74
4.5 Nguyen-Dupuis network: arrival times for different paths versus departure time ............................................. 74
4.6 Results of Monte Carlo simulation for the flow on link a = 3 . 75
4.7 Empirical distribution of the flow on link a = 3 for 4 different instants ............................................................... 76
4.8 Results of Monte Carlo simulation for travel times on path p = 1 76
4.9 Empirical distribution of the travel times on path p = 1 for 4 different instants ..................................................... 77
4.10 Performance analysis of parallel computing using from 2 to 12 cores ................................................................. 77
# LIST OF FIGURES

5.1 Modelling of a traffic network ........................................ 84
5.2 An illustrative example of variable \( s \) ........................... 85
5.3 Nguyen-Dupuis network ................................................. 93
5.4 Synthetic data set based on flows and densities ..................... 94
5.5 Average performance of the metaheuristic algorithms in density-based tests ............................................... 95
5.6 Average performance of the metaheuristic algorithms in flow-based tests ...................................................... 95
5.7 A four-leg intersection .................................................. 96
5.8 Solution Test 3 .............................................................. 99
5.9 Real data set. Average daily occupancy and flow profiles .......... 100
5.10 Solution of the TOD problem using the \( \Delta \text{BIC} \) rule .................. 101
5.11 Solution of the TOD problems using the BIC rule .................. 101
5.12 PETE algorithm behaviour .............................................. 102
5.13 OP algorithm behaviour ................................................. 102
6.1 A prototype of urban traffic control system ......................... 110
6.2 Intensity of O-D demand for pattern \( P_1 \) ............................. 121
6.3 Histograms of the concordance coefficient \( \kappa \) for identification of patterns based on flows ............................ 124
6.4 Histograms of the concordance coefficient \( \kappa \) for identification of patterns based on densities ......................... 124
6.5 CTDP solution based on flow (centroids) ............................ 125
6.6 CTDP solution based on densities (centroids) ....................... 125
6.7 Architecture of the neural network for traffic classification .... 127
6.8 Dynamic classification of traffic patterns based on Euclidean distance .......................................................... 127
6.9 Dynamic classification based on discriminant analysis ............. 128
6.10 Dynamic classification of traffic patterns based on Neural Network .......................................................... 128
6.11 Location of the selected detectors for testing the methodology through OpenStreetMaps contributors ........................ 129
6.12 Number of clusters selected ............................................. 130
6.13 Average daily occupancy profiles of the resultant CTDP .......... 132
6.14 Average daily flow profiles of the resultant CTDP ................... 132
6.15 Traffic classification for a real problem ............................. 134
6.16 Dynamic non-nervous classification based on occupancy ......... 135
6.17 Dynamic non-nervous classification based on average speed .... 136
7.1 Agent-based architecture for the implementation of the ITCS .... 143
7.2 A generic ITCS prototype ............................................... 144
7.3 Data flows in the ITCS for routes recommendations ................. 145
7.4 Intensity of O-D demand and TOD breakpoints ..................... 146
7.5 Improvement of recommendations for each pattern ................. 150
7.6 Travel times with and without recommendations for pair \( \omega_2 \) .... 151
7.7 Links behaviour with and without recommendations ............... 152
7.8 Detection of the current demand pattern without the recommenda- tions ................................................................. 153
List of Figures

7.9 Detection of the current demand pattern with the recommendations ........................................ 153

A.1 Link capacity function $\beta_a(t)$ given by a square wave. ........................................ 186
# List of Tables

2.1 Summary of recents ITS projects ........................................... 22  
2.2 Main reviews of traffic modelling approaches .......................... 25  

4.1 Discrete event algorithm .................................................... 69  
4.2 Parameters of the Nguyen-Dupuis network example ................... 70  
4.3 Parameters of the O-D demand intensity functions ..................... 71  
4.4 Test of computational burden for simulation of Nguyen-Dupuis  
and Sioux Falls network examples ........................................... 72  

5.1 Hybrid meta-heuristic algorithm for the TOD problem ................. 89  
5.2 SA* algorithm for TOD problem ........................................... 90  
5.3 Comparison between the sequential and simultaneous methodologies  
................................................................. 98  
5.4 Optimal number of clusters by selection criterion and the  
problem ................................................................. 100  

6.1 Literature for CTDP ......................................................... 108  
6.2 The k-means algorithm for the CTDP .................................. 117  
6.3 Parameters for the intensity of demand functions ....................... 121  
6.4 Average and variance-covariance matrix of parameter δ for  
traffic patterns ............................................................ 122  
6.5 Descriptive statistics associated with the Kappa concordance  
coefficient ................................................................. 123  
6.6 BIC and R values for each clustering model based on occupancy  
and on speed ............................................................... 130  
6.7 Number of days of resulting clusters .................................... 131  

7.1 Prior route choice probabilities (user point of view) ................... 148  
7.2 Probabilities for route recommendations .................................. 149
Introduction
1

Introduction

Traffic arises because throughout history the movement of people and goods has been a need for society and currently traffic management and control is a critical component for maintaining a city’s functionality. The steady growth of cities over the years and the increase in their population have lead to new urban designs, more efficiency and sustainability in order to improve the quality of life in large cities. In the last decades traffic has become a social problem in some cities due to its negative externalities: accidents, congestion and environmental damage. This occurs because the number of vehicles has increased massively. In accordance with the International Council on Clean Transportation (ICCT), total sales of new vehicles were around 75 million in 2013, this is a 33% increase in vehicle sales in just four years, and about 12 million were sold in Europe. The European market remains centred on a handful of countries, around 75% of all new car registrations occur in the five largest markets (Germany, France, United Kingdom, Italy, and Spain). On the other hand China, the US, Mexico, Indonesia, Argentina, and Turkey are countries that currently represent the major vector of growth in vehicle sales. Figure 1.1 shows the number of vehicles per capita in some key countries at the present moment.

Nowadays, as a consequence of this increase in the number of vehicles, governments are committed to improving the infrastructures for urban road users. This brings us to a vicious circle in which every year there is more demand for private mobility with all the negative factors that this entails as shown in Figure 1.2. This tendency increases every day, thereby causing congestion growth.

Important technical developments in transport modelling have taken place since the mid-1970s, in particular at major research centres; despite this, we still encounter many of the same transport problems of the past: congestion, pollution, accidents and financial deficits. We are increasingly becoming money rich and time poor.
Traffic engineering is the application of the principles of engineering, planning, analysis and design to the disciplines comprising transportation: vehicles, physical infrastructure, safety, environmental impacts and energy usage among others. From the point of view of strategic planning the study of traffic phenomena involves the activities of traffic network design, geometric route design or vehicle design. The operational level focuses on the traffic control and the analysis of the impact of road network management on human behaviour and welfare economics among other social factors. The topic of this thesis is focused on the field of traffic control and management.

Bearing in mind the paradox described in Figure 1.2 in which the enlargement of an existing road network may lead to higher levels of congestion (Braess’s paradox), the main objective of this doctoral thesis is to improve the road network capacity through management and control actions.

Nowadays, the problems concerned with traffic and the developing of new technologies have motivated the research community to pay attention in the area of so-called Intelligent Transportation Systems (ITS). This research field studies technologies and analytical techniques with the purpose of

---

**Figure 1.1:** Statistics on passenger vehicles per 1000 persons. Credited: Statical PocketBook (2014), The International Council on Clean Transportation, European commission.
1.1. Motivation and thesis objectives

The main motivation after analysing the literature on this topic (see Chapter 2) is that any ITS prototype must be capable of fulfilling four main premises, scalability, technological independence, robustness and the use of Big Data. These characteristics are described below:

- The scalability for a traffic control tool can be described as the capacity of adaptation to the different sizes of networks without losing performance and reliability. It is mandatory that the proposed ITS will be able to achieve this characteristic in order to be applied in real networks, particularly in huge cities (for example in Mexico city.}

**Figure 1.2:** The vicious circle of mobility

devolving new systems capable of solving some of the referred transport problems. ITS are established on the basis of research activities spread over many different areas such as electronics, control, communications, sensing, robotics, signal processing and information systems.

The main motivation after analysing the literature on this topic (see Chapter 2) is that any ITS prototype must be capable of fulfilling four main premises, scalability, technological independence, robustness and the use of Big Data. These characteristics are described below:

- The scalability for a traffic control tool can be described as the capacity of adaptation to the different sizes of networks without losing performance and reliability. It is mandatory that the proposed ITS will be able to achieve this characteristic in order to be applied in real networks, particularly in huge cities (for example in Mexico city.}
Thus, to ensure that the proposed ITS possesses this feature a low computational burden is necessary and it may be processed in distributed computing.

- The technological independence ensures stability and persistence over time due to the fact that the technology is changing rapidly and thus the way in which the traffic data and available information are collected may be different in a couple of years. To ensure that the proposed ITS has the capability of projection to the future it should be designed based only on the acquired data independently of the devices used to collect the data. For this survival reason and to avoid a premature obsolescence it is necessary to ensure its technological independence.

- Another essential characteristic is robustness. The possible responses of the system must be assessed to avoid chaotic situations in the traffic network. It is necessary to ensure that the ITS will have a proper response to each situation even for unusual states of the system such as incidents. ITS should avoid over-reactions on the responses, maintaining a certain parsimonious behaviour.

- The use of Big-Data system architecture is an essential characteristic to achieve the scalability in ITS. Nowadays, there are different approaches that address this problem but depending on the focus of interest or study there are better technical solutions (see Philip Chen and Zhang (2014) and Araghi et al. (2015)). ITS must be able to work with historical data obtained from the traffic network and learn from them for better predictions in real time.

During the last years in the developed and developing countries a significant portion of the research budget has been allocated to ITS research. According to Rafiq et al. (2013) in Europe the budget in this research field has increased as is shown in Figure 1.3. Despite the increasing investment currently it is not possible to find a general methodology for the development of an ITS capable of covering all the traffic problems, rather, more recent projects are focusing their efforts on generating proposals for a specific traffic problem.

The starting point of this thesis is founded on the following research questions:

- Does a conceptual framework exist capable of describing in a unified way recent research and technological advances in the ITS field?
- Is there an instance of this conceptual framework in which it is able to satisfy basic premises proposed previously? In particular, satisfying the following guidelines:
  - Applicability to huge cities.
  - Technological independence to be integrated in smart cities and able to consider a collaborative users environment.
1.2 Summary of contributions and thesis outline

- Ability to handling, manipulate and retrieve big traffic data for identification of traffic patterns or other goals.
- Capacity of suitability to generic traffic control actions.

With the objective of finding answers to the previously stated questions, first it is necessary to review the most recent advances in the ITS field, going more deeply into the systems whose domain is urban traffic. In addition, to compare the techniques used, their scalability, their efficiency and their applicability in real contexts. This issue will be addressed in Chapter 2.

This doctoral thesis proposes a set of analytical tools and approaches for the design of ITS for traffic control and management based on the above four premises. One of our greatest challenges is the capacity of the approaches to be incorporated in real time (this implies a fast computational response) to increase the traffic network capacity. Currently there are technologies that allow the interaction between vehicles and the infrastructure in real time in order to give information about the state of the system (see Hounsell et al. (2009)), and, because of this, it is possible to develop traffic tools that make recommendations to the users (best routes, travel times, etc.). The objective of the approaches presented in this thesis is their integration into the development of ITS for traffic control and management which is able to reduce congestion and travel times based on route recommendations to users.

The scalability of a generic framework of ITS is based on two phases (off-line and on-line). In the off-line phase the heavy computational burden, such as the optimization of the system or retrieval information, is performed, and in the on-line phase real-time control is carried out. This thesis also tackles the Dynamic Traffic Assignment Models (DTA) as one of the main components for the off-line phase.

1.2 Summary of contributions and thesis outline

The process followed to attain the proposed objectives is described below:

In Chapter 2 a novel prototype for ITS is proposed. A crucial element in any ITS is the availability of adequate tools for dynamic traffic assignment, in particular for the dynamic traffic network loading model because it consumes around 80% of the CPU time of the process, for this reason this key issue is reviewed in Chapter 3 and Chapter 4 addresses the development of a fast network loading model. This is also relevant for ensuring the scalability of any system. At the present moment, the cheapest and the most used way for controlling traffic is based in the so-called Time-of-Day intervals (TODs) which apply up-dated-coordinated TOD plans independently of the current network state. Chapter 5 improves this control approach by integrating the TOD identification phase and the strategic control optimization in a single
model. The off-line TOD operations are easily applicable for any kind of city but nowadays there are other technological developments that improve this control. For this reason, in Chapter 6 an on-line system is proposed to exploit the traffic data and to determine the current state of the traffic network in real time. Chapter 7 by including all the procedures described, gives a proposal for the design of ITS for traffic control and management, compatible with the basic premises. Finally in Chapter 8 the conclusions of the thesis are presented.

Figure 1.4 depicts the thesis outline which is organized according on the maturation process of the tools proposed for the design of ITS for traffic control and management, experimented with throughout the development of this doctoral thesis.

The contributions of the thesis are summarised as follows:
1.2. Summary of contributions and thesis outline

Figure 1.4: Thesis outline
A Monte Carlo approach to simulate the stochastic demand in a continuous dynamic traffic network loading problem

Authored by Sánchez-Rico María Teresa and García-Ródenas Ricardo and Espinosa-Aranda José Luis


Dynamic Traffic Assignment models are mathematical tools used for traffic management and control. These require a Dynamic Network Load (DNL) model, a route choice model and a mechanism to ensure the relationship between the submodels. The DNL problem aims to find, in a congested network, dynamic traffic volumes and travel times for a given time period. The DNL problem involves a high computational cost, so the model becomes intractable in real-time and often on off-line applications. This paper proposes a continuous DNL model based on flow discretizations, instead of time discretizations, like the classic cell transmission model. These discretizations create homogeneous traffic packets according to their route. The algorithm propagates the packets synchronously across the links. The dynamic mechanism used in the network links is based on a generalization of the Whole-link Travel Time (WTT) model, which divides the links in the running section and the vertical queue section. The first is associated with the travel time and the second with the link capacity. A generalization of the point-queue model is introduced to tackle dynamic link capacities such as signalled intersections. Under weak assumptions, the resulting model satisfies the FIFO rule and it is used to obtain a computationally tractable model. The resulting model leads to a discrete events simulation approach. Numerical experiments indicate that the proposed algorithm is promising.

Time series data mining to enhance traffic signal optimization

Authored by García-Ródenas Ricardo and López-García María Luz and Sánchez-Rico María Teresa

Annals of Operations Research, Under Review

The segmentation of time series has been studied in a wide range of applications. This study investigates a challenging segmentation problem in traffic engineering, namely, identification of time-of-day breakpoints for traffic signal timing plans. A large number of urban centres have traffic control strategies based on the segmentation of time series. We propose a bilevel optimization model to address simultaneously the segmentation problems and the traffic control problems over these time intervals. Efficient metaheuristic algorithms have been developed for the bilevel model based on the hybridization of the simulated annealing and Nelder-Mead methods.
1.2. Summary of contributions and thesis outline

Numerically the effectiveness of the algorithm using real and synthetic data sets is demonstrated. We address the problem of automatically estimating the number of time-of-day segments that can be reliably discovered. We adapt the Bayesian Information Criterion, the PETE algorithm and an oriented-problem approach. The experiments show that this last method gives interpretable results about the number of reliably necessary segments from the traffic-engineering perspective. The experimental results show that the proposed methodology provides an automatic method to determine the time-of-day segments and timing plans simultaneously.

An approach to dynamical classification of daily traffic patterns

Authored by García-Ródenas Ricardo and López-García María Luz and Sanchez-Rico Maria Teresa

Computer-Aided Civil and Infrastructure Engineering, Under Review

This paper proposes a prototype of urban traffic control system based on a prediction-after-classification approach. In an off-line phase, a repository of traffic control strategies for a set of (dynamic) traffic patterns is constructed. The core of this stage is the $k$-means algorithm for daily traffic pattern identification. The clustering method uses the input attributes flow, speed and occupancy and it transforms the dynamic traffic data at network-level in a pseudo-covariance matrix which collects the dynamic correlations between the road links. An optimum number of traffic patterns is provided by Bayesian Information Criterion and the ratio of change in dispersion measurements. In an on-line phase, the current daily traffic pattern is predicted within the repository and its associated control strategy is implemented in the traffic network. The dynamic prediction scheme is constructed on the basis of an existing static prediction method by accumulating the trials on sets of patterns in the repository. This proposal has been assessed in synthetic and real networks testing its effectiveness as a data mining tool for the analysis of traffic patterns. The approach promises to effectively detect the current daily traffic pattern and is open to being used in intelligent traffic management systems.

A proposal for the design of Intelligent Transportation Systems for traffic control and management

Authored by García-Ródenas Ricardo and Sánchez-Rico María Teresa

In this chapter a proposal of an ITS for traffic management and control is presented and is based in the TOD concept. The TOD concept, within a given demand pattern, replaces the short-term traffic predictions. This made it possible to consider the dynamic state of the traffic as a finite sequence of
stationary states. This finiteness allows the control strategies to be optimized within an off-line context. The main objective of the proposal is the route recommendations and to this end different numerical experiments have been conducted to illustrate the approach.

The results show that the proposed ITS is able to identify the network state in real time but after route recommendations, the system may lose effectiveness in identifying the network state.
Intelligent transportation systems
Intelligent Transportation Systems (ITS) are tools that combine advanced technologies in communication and information in order to solve transportation problems such as, traffic congestion, safety and environmental conservation.

ITS appeared around 30s and they have been slowly creeping into our lives. The earliest developments on ITS were made in Europe, U.S and Japan. The evolution in the ITS can be classified in three phases through the years.

The preparation phase, between 1930 and 1980, includes the development of Electronic Route Guidance Systems (ERGS), which used a two-way road vehicle communications to provide route guidance. Other relevant advances in this phase are the Comprehensive Automobile Traffic Control System (CACS) and the Autofahrer Leit and Information System (ALI) in Japan and Germany respectively, which are dynamic route guidance systems based on real traffic conditions.

The feasibility study phase, between 1980 and 1995, is characterised by the development of many programs which can be considered as the evolution of the basic technologies carried out in the previous phase. The most important developments in Europe during this phase were part of the Program for European Traffic with Efficiency and Unprecedented Safety (PROMETHEUS) project which was formed for the purpose of reducing road accidents and to improve the traffic efficiency. Other important projects are the Dedicated Road Infrastructure for Vehicle Safety in Europe (DRIVE) for the development and test of the communication systems, for drive assistance and traffic management. The European Road Transport Telematics Implementation Coordination Organization (ERTICO) was set up to provide support for perfecting and implementing the Europe Transport Telematics Project. In the United States the major project was the formation of the Intelligent Vehicle Highway Systems (IVHS America) which is a public-private consortium for consolidating national ITS interests and promoting
international cooperation in ITS. In 1994 the United States Department of Transportation (USDOT) changed the name from IVHS to Intelligent Transportation Society of America (ITS America). Several projects were developed at more than eighty places across the US. Two important projects have been carried out in Japan in order to alleviate congestion, improve safety, and reduce the environmental consequences of road traffic, Road Automobile Communication System (RACS) by the Ministry of Construction and Advanced Mobile Traffic Information and Communication System (AMTICS) by the National Police Agency intended for the promotion of research and development of vehicle safety technologies.

The product development phase, from 1995 until the present, is characterised by the development of the transport telematics and ITS applications. The previous phase focused on creating a technical foundation, with high-level functions for ITS, and this purpose was achieved successfully. This lead to the present phase, dealing with the creation of feasible products, geared to specific needs and problems.

2.1 Taxonomies of intelligent transportation systems

ITS integrate Information and Communication Technology (ICT) with transport infrastructures, vehicles, and users. This implies exploiting the existing technologies and infrastructures, such as computers, communication networks, sensors, positioning systems, and automation technologies for collecting the relevant data. It can be viewed as a cooperative network, consisting of a large number of users and components, which either transmit their own data to the central node. The gathered data are then efficiently used to facilitate novel services that have the potential to relieve traffic congestion, to ensure safety, and to protect the environment. In this section three classifications of these systems are reviewed.

The classification based in the market areas is useful from the point of view of transfer of technology to society. This classification identify new business opportunities and new products. From a social point of view a classification based in the strategics objectives pursued by the ITS is more appropriate because these define the priorities and objectives of the public policies in the field of transport and allow to redirect the research efforts.

2.1.1 A classic taxonomy

ITS have been clustered under different perspectives. The initial taxonomy is based on the problem to be addressed.
2.1. Taxonomies of intelligent transportation systems

Advanced Traffic Management Systems (ATMS). These systems operate with a series of video and roadway loop detectors, variable message signs, network signal and ramp meter timing schedules including roadway incident control strategies from one central location to respond to traffic conditions in real time. ATMS are usually installed by operators of freeways in order to monitor, analyse and control traffic.

Ramp metering and freeway-to-freeway control regulate the flow on ramps by traffic signals. The aim is to avoid traffic congestion in the mainstream flow.

Variable Message Signs (VMS). Are used to set speed limits or to display information to the road user. In different countries the amount of information displayed on these devices vary a lot: in Germany, most VMS do not present recommendations, e.g., travel time, but only road signs. In other European countries, like the Netherlands, Great Britain or France, travel times and recommendations are given (see Emmerink et al. (1996), Wardman et al. (1997)). Additionally to VMS, lane control signals are employed which allow for a traffic responsive use of lanes.

Advanced Travellers Information Systems (ATIS). Have the goal of supply real time traffic information to the travellers. (see Adler and Blue (1998), Zito et al. (2011)). The information about the traffic conditions influences drivers so that they make a better use of the transport system, allowing the reduction of congestion, optimising the traffic flow and reducing pollution. With ATIS systems, travellers, from home, on work, or in stopping-place can decide which is the most advantageous road to reach its destination, the most favourable transportation service and the most appropriate schedule to adopt.

The provision of real-time information to travellers will lead to more efficient distribution of travellers to routes and modes. For the individual, using ATIS can lead to more efficient travel choices and help reduce anxiety and stress associated with travel planning.

Commercial Vehicle Operations (CVO). Use different ITS technologies to the special needs of commercial vehicles and fleets to increase safety and efficiency. CVO systems became useful for large and medium companies because they allow the management of all the vehicles, while controlling speed and stopping-place times, besides fulfilling the destination.

Intelligent Public Transportation Systems (IPTS). These systems are a branch of ITS, which aim to control public transportation networks, to maintain their performance, and to provide users (passengers) with up-to-date information about trips and network operating conditions. To reach these aims, IPTS rely on several technologies that can be embedded within different control architectures. In IPTS are also included the automatic payment systems, through the use of multiple usage smart cards which provide functions such as stored credit or automatic capture of passenger information and journey profile. This issue is addressed in Elkosantini and

*Advanced Vehicles Control systems (AVCS).* Joint sensors, computers and control systems to assist and alert drivers or to take part of vehicles driving (see Shladover (1995) and Baskar et al. (2011)).

*Advanced Rural transport Systems (ARTS).* Include systems that apply ITS technology to the special needs of rural areas, since these are characterised by a number of attributes, which are different from urban areas, e.g., blind corners, fewer passing lanes, longer distance travel or less supporting infrastructure.

### 2.1.2 A taxonomy from a point of view of the market areas

A current way of looking the ITS classification is attending to the market areas in accordance with the buyers and users of ITS. Companies supplying ITS equipments and services might typically divide the entire field into such market areas when defining customers, developing products, and carrying out marketing and sales strategies. Taking on this view, ITS can be classified by recognizing the way in which transportation is currently operated and organized and observing practical, institutional and organizational boundaries rather than theoretical technological ones (see McQueen and McQueen (1999)). A classification of the ITS market areas is illustrated in Figure 2.1, moreover the classification described in Section 2.1.1 is further superimposed. This classification appears in the preparation phase whereas the classification from a point of view of the market areas appears during the product development phase.

The main categories are:

*Traffic control and management.* In this market area, the customers are primarily local and central government agencies. These are the people responsible for using tax money to improve the transportation process, making it safer and more efficient. Their main motivation is to manage the entire road network as well as possible on behalf of the general public. This is achieved through collecting traffic data and influencing and managing traffic flowing on the network.

*Emergency management.* The responsible for road safety are in this market area. This group responds to incidents and emergencies and has primary responsibility for identifying problems, deciding on appropriate responses and resources, and then managing the incident.

*Transportation planning.* Customers in this market area try to match transportation supply with demand both now and in the future. These are the people who try to determine current transportation use patterns and make predictions about future use patterns. They include local, state, and federal transportation agencies as well as individual, traffic planners, and
2.1. Taxonomies of intelligent transportation systems

Figure 2.1: A taxonomy from the point of view of the market areas and its roughly corresponding classic area

transit planners.

Traveler information services. In this market area, the customers for ITS are the providers and users of travel information. Although central, state, and local transportation agencies can play a role in this area. It is likely that the significant users will be private sector information service providers and subscribers.

Commercial vehicles management. There are two primary groups of customers in this area: the trucking industry and the local agencies that regulate them. The trucking industry is comprised of potential customers for in-vehicle devices providing travel information and fleet management services. The regulatory authorities procure infrastructure to support electronic data interchange and automated licensing and inspection procedures.

Transit management. This area is populated by those who plan and operate our transit systems in both urban and rural areas. They are concerned with increasing the operational efficiency of all transit modes as well as achieving a tangible improvement in the attractiveness of these service offerings.

Smart cities. The vehicle manufacturers, transportation infrastructure companies, automotive electronics manufacturers, and truck and transit vehicle manufacturers all belong in this category. They are all concerned
with enhancing the capability of road vehicles through the use of electronics, sensors, communications, and control actuator technologies. Vehicle drivers also belong in this group, because they will be the eventual consumer or end user of the various ITS applications.

Vehicles and infrastructure safety. This area, again, consists of the central, state, and local government officials who are charged with improving the efficiency of the roadway network. The people responsible for incident and emergency management tend to have different needs and objectives and belong to different organizations than the traffic managers; accordingly, they have their own market area. Note that the market areas are not mutually exclusive. There is substantial overlap between many of the market areas because of the integrated nature of transportation demand. The market areas should be considered as a way of focusing on the needs of a particular group of people, rather than a means of rigid distinction.

Payment systems. This area encompasses all the people that try to take money from you in return for providing a service. It includes toll road operators, transit agencies, car parking operators, and all the customers and users of ITS involved in the process of fee payment for transportation services. The payment systems area is also expanding to include people responsible for general purpose payment services beyond the transportation field. This includes people involved in the development of applications such as Pay-per-view TV, telephone cards, and general purpose debit cards. There is a growing overlap between transportation and non transportation needs because of the potential for synergy in the payment systems area. For example, the same smart card could be used to pay for bus fares and make telephone calls within a region.

2.1.3 A taxonomy according on their strategic objectives

In this section a classification of ITS based on their strategic objectives is addressed (see Rafiq et al. (2013)) with the purpose of identifying directions of actual and future research in this field.

CooperativeMobility. In recent years, the development of cooperative systems based on vehicle communications has attained a considerable growth. The commercial realization of such systems requires an intelligent organization of system resources by integrating potential cooperative applications running on an in-vehicle devices. This requires an open in-vehicle platform to guarantee interoperability with respect to the communication infrastructure, transportation facilities, various vehicle types, etc.

EcoMobility. In a modern society, mobility and the corresponding modes of transportation are pivotal elements which have direct influence on daily life activities. However, with passing time a growing number of transport vehicles are negatively impacting our environment due to growing greenhouse gas emissions. Therefore, it is indispensable to intelligently
explore transportation resources such that they pose the least possible threat to our environment. EcoMobility pursues this objective by employing ITS to make the transport greener. This is constantly gaining recognition since eco-friendly ITS applications have the potential to reduce energy consumption, improve energy efficiency, cut greenhouse gas emissions, and limit our dependence on fossil fuels.

**SafeMobility.** A major challenge faced by the transportation authorities in Europe is to make the modes of transportation safer. One of the reasons is the fact that more than 40,000 people die on the road each year, which is a significant loss to the society. Moreover, road accidents cost the European economy around 200 billion every year. Research and initial deployments have shown a considerable improvement in the road safety via the use of ITS, incorporating intelligent advanced driver assistance systems (ADAS). Such systems can not only detect hazards on the road ahead but also warn the driver about them even before they are visible. Hence, such systems are very useful in keeping the vehicles at a safer distance from one another. Additionally, the drivers remain well aware of the local traffic-driving conditions.

**Infomobility.** Reliable, personalized, and anytime anywhere based Real-time Travel and Traffic Information (RTTI) is a key element of intelligent mobility services envisioned for the future. Such information is vital for the design and development of safer, smarter, and more efficient transport systems. Here, personalized RTTI means context aware services for a particular user, taking into account the needs, preferences, miscellaneous attributes (such as accumulated travel bonuses of individuals), environmental signature, travel habits, etc.

### 2.1.4 Current examples of ITS projects

There are currently ITS initiatives/projects in Europe, and they cover nearly all of the main segments of the transportation sector. The purpose is to appraise the milestones that have been reached so far in the area of ITS. Table 2.1 summarises these projects according with the taxonomies previously reviewed.
Table 2.1: Summary of recent ITS projects

<table>
<thead>
<tr>
<th>Control and management actions</th>
<th>Cooperative Mobility</th>
<th>Mobility and EcoMobility</th>
<th>SafeMobility</th>
<th>InfoMobility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Road traffic management</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smart cities</td>
<td></td>
<td>COSMO (2013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Charges</td>
<td></td>
<td></td>
<td></td>
<td>GSC (2009)</td>
</tr>
<tr>
<td>Incentives and discussion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raise awareness on save-driving</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic management</td>
<td>CITYLOG (2010)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident/emergency management</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other initiatives</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.2 A prototype for the design of ITS for traffic control and management

![Diagram of ITS prototype](image)

**Figure 2.2:** Prototype for the design of ITS for traffic control and management

In a cooperative scenario, ITS extract and exploit the useful information from the environment using smart travel-planning services available over the Internet together with the technology platforms installed in the vehicles. From this point of view, the term *data acquisition* (or collection) plays a primary role. Data acquisition answers the questions regarding the retrieval of relevant information in a particular environment enriched with sensors/smart devices. One can create a rich set of applications by connecting hundreds of smart devices/sensors in an environment to a...
cooperative network. In this context, **Crowd-sourcing** plays the first crucial role. Crowd-sourcing makes it possible to connect millions of users to a network creating a huge amount of information every second. The development of such a cooperative network of road users provides valuable information for the transportation authorities in real-time. This huge amount of information retrieved using the interconnected network of users and sensors has to be stored and mined by using an **information retrieval module** in order to produce traffic patterns and forecasts. This is the second key element, an e-learning module that allows to infer the current state of the system. Therefore, the third important factor involves the development of a **smart traffic management system**. This module is a conceptual framework and is responsible for designing the appropriate strategies to achieve the ITS objective. For example ITS may optimize the travel times of the users or to minimize the harmful CO\(_2\) emissions by assisting the drivers to choose the best possible route to their destinations. ITS services rely heavily on real-time information generated by sensors, devices, and ICT-network users. Thus, efficient exploitation and integration of the aforementioned three factors are vital to fulfil the envisaged goals for ITS in the near future.

It is possible to observe in Figure 2.2 the structure of a prototype for the design of ITS for traffic control and management, each box corresponds to a particular process that will be addressed within this thesis in its corresponding chapter, including details of its own literature review. Table 2.2 shows seminal reviews in the field of transportation planning according with the classification of ITS based in strategic objectives. The works presented in the table have been selected as key reviews with the aim of showing a general and introductory idea of the traffic problems that have been intensively studied due to the increase in traffic during the last two decades. Many of these models might represent the core of the smart traffic management systems and would define the ITS objectives without change the structure showed in Figure 2.2.
### Tabla 2.2: Main reviews of traffic modelling approaches

<table>
<thead>
<tr>
<th>Strategic objective</th>
<th>Reference</th>
<th>Key features</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mobility</strong></td>
<td><strong>Papageorgiou et al. (2003a)</strong></td>
<td>Analyse control strategies for road networks, freeways and route guidance</td>
</tr>
<tr>
<td></td>
<td><strong>Hounsell et al. (2009)</strong></td>
<td>Present a state-of-the-art reviews of Urban Traffic Management (UTM) in vehicles and the benefits of new UTM systems.</td>
</tr>
<tr>
<td></td>
<td><strong>Hardjono (2011)</strong></td>
<td>Categorize traffic jams reduction approaches and identify its strengths.</td>
</tr>
<tr>
<td></td>
<td><strong>Chen and Cheng (2010)</strong></td>
<td>Examine an agent-based approach and its application in different modes of transportation, including roadway, railway, and air transportation.</td>
</tr>
<tr>
<td></td>
<td><strong>Idris et al. (2009)</strong></td>
<td>Focuses on the car park management system introduced and reviews the evolution of vehicle detection systems developed over the years.</td>
</tr>
<tr>
<td></td>
<td><strong>Marsden (2006)</strong></td>
<td>Presents a review of the evidence base upon which commuter, leisure and shopping and residential parking policies are based.</td>
</tr>
<tr>
<td></td>
<td><strong>Cortés et al. (2010)</strong></td>
<td>Provide guidelines on how to incorporate the necessary entities and components for a proper simulation of public transport systems in a microsimulation environment.</td>
</tr>
<tr>
<td><strong>SafeMobility</strong></td>
<td><strong>Ozanne-Smith (2005)</strong></td>
<td>Analyse studies focused particularly on risk factors identified including widespread risky driving behaviour.</td>
</tr>
<tr>
<td></td>
<td><strong>Pirrera et al. (2010)</strong></td>
<td>Focuses on noise processing in general and in relation to sleep as a consequence of road traffic.</td>
</tr>
<tr>
<td></td>
<td><strong>Garg and Maji (2014)</strong></td>
<td>Presents an exhaustive comparison of principal traffic noise models adopted in recent years in developed nations.</td>
</tr>
<tr>
<td><strong>CooperativeMobility</strong></td>
<td><strong>Giannopoulos (2004)</strong></td>
<td>Examines the applications of information and communication technologies in the field of transport during the last decade.</td>
</tr>
<tr>
<td></td>
<td><strong>Papadimitratos et al. (2009)</strong></td>
<td>Surveys the state-of-the-art approaches, solutions, and technologies across a broad range of projects for vehicular communication systems.</td>
</tr>
<tr>
<td><strong>EcoMobility</strong></td>
<td><strong>Elkafoury et al. (2014)</strong></td>
<td>Discusses the limitations and potentials of different methods used for road traffic emission monitoring, models.</td>
</tr>
<tr>
<td></td>
<td><strong>Goel and Kumar (2014)</strong></td>
<td>A review of fundamental drivers governing the emissions, dispersion and exposure to vehicle-emitted nanoparticles at signalised traffic intersections.</td>
</tr>
<tr>
<td><strong>InfoMobility</strong></td>
<td><strong>Chorus et al. (2006)</strong></td>
<td>Presents such a review of both the empirical and the conceptual literature concerning the use and effects of travel information systems.</td>
</tr>
</tbody>
</table>
Traffic modelling
Traffic modelling

An essential element of the ITS, is a traffic mathematical model that allows to answer questions of the type What if...?, thus, it be able to evaluate the actions to be taken in the traffic network. For this reason one of the most important tools, the so-called Dynamic Network Loading Models are reviewed in this chapter. The second purpose of this chapter is to advocate that the ITS should be technologically independents focusing on what kind of information should be register instead of the used technology to obtain it. For this reason in this chapter we introduce the essential magnitudes that characterise the traffic flows, which should be as the basis of the ITS actions.

3.1 Introduction to the traffic flow theory and its modelling

A model is a simplified representation of a part of the real world (the system of interest) which focuses on certain elements considered important from a particular point of view.

A broad definition allows us to incorporate both physical and abstract models. In the first category we find, for example, those used in architecture or in fluid mechanics which are basically aimed at design. In the latter, the range spans from the mental models all of us use in our daily interactions with the world, to formal and abstract (typically analytical) representations of some theory about the system of interest and how it works. Mental models play an important role in understanding and interpreting the real world. Mathematical modelling is an Art, as such, there are many valid options which could be suitable to face one problem (see Figure 3.1). Basically there are two premises that a model should be hold:

Premise 1. The model must reproduce the available system observations.
Figure 3.1: Mathematical modelling

Premise 2. The model should be capable of answering the interest questions about the systems of study.

Therefore, a model is not a synthesis of reality and it is not expected that the mechanisms collected by the model are the true ones operating in reality.

Figure 3.2 shows a methodology based on models to face scientific and engineering problems. At the modelling stage it is proposed a model capable of answering the interest questions. At the calibration stage the capacity to reproduce the available data is evaluated. At the analysis and resolution stage if the accuracy is not enough, another model would be proposed (more complex, following a principle of parsimony). At the use stage the answers are obtained and applied on the system.

Traffic models can be classified as microscopic, macroscopic or mesoscopic. Microscopic traffic simulations assume that the behaviour of an individual vehicle is a function of the traffic conditions in its environment. Although microscopic simulations usually keep track of each vehicle from its origin to its destination. Moreover their assumptions are difficult to validate in practice because humans’ behaviour in real traffic is difficult to observe and measure.

Macroscopic models assume an aggregate behaviour of vehicles, these
models treat traffic in an aggregate manner, such as a uniform or homogeneous flow, without considering each individual vehicle. We assume that the ITS is able to collect the macroscopic magnitudes of flow, speed and density as the essential parameters which characterize the traffic state. The new technological advances as localization through mobile phones, vehicle registration plates, etc. allow to estimate the described magnitudes in more precisely manner, as well as other magnitudes such flow fractions following particular routes (see Castillo et al. (2008a), Castillo et al. (2010a) Castillo et al. (2010b)). It is necessary to focus in what is obtained instead in how is obtained, although it remains essential for its practice viability. The advantage of these models with respect to the microscopic ones is a more tractable computational burden. Appealing to our premise of scalability we focused on these latter models.

3.1.1 Basic concepts of traffic flow

In this section a theoretical framework for traffic flow modelling is described at macroscopic level. The objective of this section is presenting the basic concepts of traffic flows for its modelling based on the work presented by Immers and Logghe (2003).
We begin with a basic definitions:

*Trajectory* is the position of a vehicle through time. Figure 3.3 shows some trajectories on the $x - t$ space.

A *measurement interval* $S$ is defined as an area in the $t - x$ space and it is always done for a certain measurement interval. Figure 3.3 shows two measurement intervals:

- **S1.** This rectangular measurement interval covers a road section of length $\Delta X$ during an infinitely small time interval $dt$. This coincides approximately with a location interval $\Delta X$ at a specific moment $t_1$. We assume that $n$ vehicles move through this interval.

- **S2.** This rectangular measurement interval represents an infinitely small road length $dx$ during a time interval of $\Delta T$. This coincides approximately with a time interval $\Delta T$ at a location $x_2$.

*Density* is a typical variable from physics that was adopted by traffic science. Density $k$ reflects the number of vehicles per kilometre of road. That is:

$$k = \frac{n}{\Delta X}$$  \hspace{1cm} (3.1)
where the variable $n$ indicates the number of vehicles at instant $t_1$ on the location interval $\Delta_X$.

In general, the Density $k$ depends on the location $x_n$, time $t_n$ and the measurement interval $S_n$, thus:

$$k(x_n, t_n, S_n) = \frac{n}{\Delta_X} \quad (3.2)$$

The flow rate represents the number of vehicles that passes a certain cross-section per time unit and it is calculated as:

$$q = \frac{m}{\Delta_T} \quad (3.3)$$

where the variable $m$ indicates the number of vehicles passing some designated point in a time interval $\Delta_T$.

For a time interval $\Delta_T$ at any location $x_n$, time $t_n$ within a measurement interval $S_n$, the flow rate is calculated as follow

$$q(x_n, t_n, S_n) = \frac{m}{\Delta_T} \quad (3.4)$$

The flow rate is usually expressed in vehicles per hour. We call the maximum possible flow rate of any road its capacity. Depending on vehicle composition, the capacity of a highway lies between 1800 and 2400 vehicles per hour and per traffic lane.

The mean speed $u$ is defined as the quotient of the flow rate and the density. The mean speed is also a function of the location, the time and the measurement interval.

$$u(x_n, t_n, S_n) = \frac{q(x_n, t_n, S_n)}{k(x_n, t_n, S_n)} = \frac{Total \ distance \ covered \ by \ vehicles \ in \ S_n}{Total \ time \ spent \ by \ vehicles \ in \ S_n} \quad (3.5)$$

This definition of the mean speed is also called the fundamental relation of traffic flow theory and it is expressed as:

$$q = k \cdot u \quad (3.6)$$

The time-mean speed is the arithmetic mean of the speeds observed at some designated point along a highway and it is computed as:

$$\bar{u} = \frac{1}{n} \sum_{i=1}^{n} u_i \quad (3.7)$$
where $u_i$ is the spot speed (i.e., the speed of the vehicle at the fixed point on the highway) of the $i$th vehicle.

The space-mean speed $u$ is the more useful definition of average traffic speed in the context of traffic analysis and is determined on the basis of the time needed by a vehicle $i$ to traverse some known length of highway, $l_i$, and is

$$u = \frac{1}{n} \sum_{i=1}^{n} l_i;$$

(3.8)

where $l_i$ is the length of highway used for the speed measurement of vehicle $i$ and,

$$t = \frac{1}{n} \left[ t_1(l_1) + t_2(l_2) + \cdots + t_n(l_n) \right];$$

(3.9)

where $t_i(l_i)$ is the time necessary for vehicle $i$ to traverse a section of highway of length $l_i$.

The relative occupancy $b$ in time interval $S_n$ is given by

$$b(x_n, t_n, S_n) = L \cdot k(x_n, t_n, S_n)$$

(3.10)

where $L$ is the physical length of the vehicles if we assume that all vehicles have the same length.

Road traffic state is characterised by the flow rate, the density and the mean speed. We combine all the possible homogeneous and stationary traffic states in an equilibrium function that can be described graphically by three diagrams. The equilibrium relations presented in this way are better known under the name of fundamental diagrams.

Figure 3.4 shows the relation between two of the three variables. The third variable can always be recovered by means of the relationship $q = k \cdot u$. The third variable in the $q - u$ and the $k - q$ diagram is an angle. The flow rate in the $k - u$ diagram is represented by an area. A fundamental diagram applies to a specific road and is drawn up on the basis of observations. Thus stationary and homogeneous traffic is always in a state that is located on the bold black line. Some special state points require extra attention:

**Completely free flowing traffic:** when vehicles are not impeded by other ones, and they travel at a maximum speed of $u_f$ (free speed). This speed is dependent, amongst other things, on the design of a road, the speed restrictions in operation at any particular time and the weather. At free speed, flow rate and density will be close to zero.
3.2 Dynamic traffic assignment

In section 3.1 the magnitudes to describe the observed data has been reviewed (see Figure 3.1). In this section the mechanisms to develop mathematical models will be analysed.

In the practice, during the traffic assignment model elaboration, the so-called four-step scheme is followed, which allows the set of necessary data for its formulation. In this section we review this scheme in order to define the input data are required for the model and the data are obtained by the model output.

Saturated traffic: on saturated roads, flow rate and speed are down to zero. The vehicles are queuing and there is a maximum density of \( k_j \) (jam density).

Capacity traffic: the capacity of a road is equal to the maximum flow rate \( q_c \). The maximum flow rate of \( q_c \) has an associated capacity speed of \( u_c \) and a capacity density of \( k_c \). The diagram shows that the capacity speed \( u_c \) lies below the maximum speed \( u_f \).

3.2 Dynamic traffic assignment

![Fundamental diagram: relationship among flow, density and speed for vehicle movement. Credited: (Immers and Logghe, 2003)](image-url)
3.2.1 The structure of the classic four-step model

Years of practice and development have resulted in a general structure which has been called the classic four-step model. This structure is, in effect, a result from practice since 1960s but has remained more or less unaltered despite major improvements in modelling techniques since then.

- **Trip generation phase.** The approach starts by zoning a study area and the collection and coding of planning, calibration and validation data. These data would include base-year levels for population of different types in each zone of the study area as well as levels of economic activity including employment, shopping space, educational and recreational facilities.

  After this phase is obtained a modelling of the transport network by a graph \( G = (N, A) \) where \( N \) and \( A \) are the set of nodes and links (directed) respectively. The meaning of the links depends on whether the network is traffic or public transport. In the first case the links are associated with the streets and nodes with intersections and dummy nodes called centroids which represent the obtained zoning. In the second case, each node is associated to a stop and each link represents the possible movements between stops that a user can perform; in addition, other links are associated with travel time in public vehicles, moving or waiting.

- **Distribution phase.** The aim of this phase is to obtain the distribution in the space of the trip, that is, get the number of trips that are made from one area to another, providing the origin-destination (O-D) trip matrix. Trip generation problem involves answering the question of how many trips will begin or end in each analysed zone. At this stage we obtain a set of ordered pairs of \( N \times N \) and travel demand (initially considered fixed).

  An O-D matrix is essential for efficient traffic control and management. In the past two decades, the Matrix Estimation (ME) problem has been intensively researched, (see Yang and Zhou (1998), Cascetta and Nguyen (1988), Castillo et al. (2008a), Ortúzar and Willumsen (2001) ) This problem tries to estimate O-D trip matrices based one some observed link flows. Unfortunately, this is an under-specified problem, due to the fact that the number of O-D pairs is normally much larger than the number of links, and there is an infinitely set of solutions satisfying the conservation constraints. According with Jiménez Gómez (2009) the matrix estimation problems can be classified in three groups:

  *Direct sample estimation.* These methods use a large amount of data, for example family socio-economic data, number of vehicles per family, the reason of the trip, the modal split etc. It is an expensive process because it collects too much information demanding a lot of time and money, besides it requires time analysing it.
3.2. Dynamic traffic assignment

Demand model based methods. These methods are based on the trip generation and distribution phases of the classical model.

Traffic counts based methods. These methods combine, in an efficient way, the information obtained from link counts with other information (prior or target matrix, socio-economic data, etc.) in order to estimate or update the trip matrix. This is the most used and studied methods during the last years because traffic counts is the type of information that can be automatically collected on a subset of the links of the network and therefore, it is available in urban areas at low cost.

• Modal split phase. Explains the choice of mode, that is, how users choose the mode of transport which satisfies their travel needs. With these models the O-D matrix is obtained for each transportation mode present in the study.

• Assignment phase. The traffic assignment phase describes the behaviour of users on a traffic network. The models for this stage describe how users choose their route, and perhaps destination and mode departure time, in terms of (generalized) costs in each of these alternatives. The use of mathematical models has been useful in identifying solutions or policies to achieve the proposed objectives for the system, being a great help for planning and decision-making and obtaining macroscopic descriptions of traffic flow from the behaviour of drivers or users.

Predicting traffic flows is one of the main aims of traffic models. This includes link, OD pair and route flows. Though these flows are closely related, traditionally the prediction or estimation of link and OD flows have been treated separately, in the so-called traffic assignment and the trip matrix estimation problems, respectively. Nevertheless, they can also be combined into one problem, in which the trip O-D matrix flow estimation problem and the traffic assignment problem become a single one, which is solved using bi-level approaches.

3.2.2 Challenges for dynamic traffic assignment

Dynamic traffic assignment (DTA), though still in an unfinished state of flux, has evolved substantially since the seminal work of Merchant and Nemhauser (1978). There is currently heightened interest in DTA, particularly in the development of approaches that can be deployed for large-scale in real-time or off-line planning contexts applications. In addition, researchers have become increasingly aware that the theory of DTA is still relatively undeveloped, which necessitates approaches that account for challenges from the application domains as well as for the fundamental questions related to tractability and realism. Agencies and practitioners are also increasingly realizing the potential of DTA to address long-standing problems with the unrealistic assumptions of existing static
CHAPTER 3. TRAFFIC MODELLING

Trip generation phase

- Zoning of the studying area
- Modellization of the traffic network
- Data collection for the study

Distribution phase

- O-D matrix generation
- Travel demand for each O-D pair

Modal split phase

- An O-D matrix is obtained for each mode of transport

Assignment phase

- Each O-D matrix is allocated to a set of paths in the traffic network

Figure 3.5: Traditional four-steps schema for traffic modelling

Figure 3.6: Example of road network in Ciudad Real (Spain)

planning methods, as well as the potential of DTA to both evaluate ITS technologies, as well as be the main operational engine for deployment. DTA refers to a broad spectrum of problems, each corresponding to
3.2. Dynamic traffic assignment

different sets of decision variables and underlying behavioural and system assumptions, and possessing varying data requirements and capabilities in terms of representing the traffic system or control actions. One common feature of these models is that they depart from standard static assignment assumptions to deal with time-varying flow. Another feature shared by these models is that none presently provides a universal solution for general networks. Perhaps the one aspect that fosters unanimity among researchers is that the general DTA problem is inherently characterized by ill-behaved system properties that are imposed by the need to adequately represent traffic realism and human behaviour. This is further exacerbated by the time-dependency and randomness in system inputs. A fundamental consequence of this reality is that a theoretical guarantee of properties such as existence, uniqueness, and stability can be tenable only through compromises in depicting traffic theoretic phenomena and potentially restrictive assumptions on driver behaviour. Viewed from the complementary perspective, an ability to adequately capture traffic dynamics and driver behavioural tendencies precludes the guarantee of the standard mathematical properties. This inherent complexity of DTA has spawned a clear division of approaches that range from the analytical to the simulation-based. DTA researchers aim to develop deployable solution procedures that seek close-to-optimal solutions with a clear understanding that claims of uniqueness or global optimality are neither essential nor particularly meaningful in the real-world. This has manifested as the development of mostly heuristic implementation procedures that seek effectiveness, robustness, and deployment efficiency. Another outcome is the notion of commensurably of features among different approaches so that trade-offs among desirable features allow different degrees of responsiveness to different DTA problems given the broad scope of objectives and functional needs addressed under the general umbrella of DTA. A reassuring practical aspect of the general DTA problem is that its mathematical intractability is not an all-encompassing barrier to the real-world utility of the associated solution approaches. Substantial research over the past decade suggests that effective and efficient solutions can be obtained for several realistic scenarios characterized by inconsequential mild violations and infrequent physical manifestations of the mathematically intractable aspects. There are several actual situations where ill-behaved problem characteristics do not arise, and even if they do, it is with minimal practical, temporal and spatial consequences vis-à-vis the modelling assumptions. Hence, different approaches may address different functional needs with different degrees of robustness, precluding the notion of sweeping generalizations on mathematical intractability by focusing on pathological scenarios, specially if they are rare in practice. The consensus is that a deployable DTA approach should adequately represent traffic realism in the context of the problem objective.

DTA appeals to a broad audience of researchers and practitioners, and promises the potential for a wide range of applications. Recent trends suggest that it can impact a broad range of problems in transportation operations, such as the ITS design addressed in this thesis.
At the present moment, State-of-Art in dynamic equilibrium models is not yet completed and requires improvements. However, a significant number of models developed in literature exist. In order to be able to elaborate a unified theory of dynamic traffic assignment modelling (as it occurs in the static case) it is necessary to compare the existing models and to establish the relationships between them. Daganzo (1995a) considers that a basic theory of macroscopic traffic models for networks in the context of dynamic assignment have to include at least realistic models of:

- **Demand.** Path choice and demand intensity when the time-varying link times are known.

- **Supply.** Traffic behaviour when the vehicle paths are known (so that link travel times can be predicted). This submodel is the continuous dynamic network-loading problem (CDNLP). The continuous DNLP is fully defined by: (a) a continuous time link model, (b) intersection modelling and (c) partial flow modelling. The main modelling approaches for (a) are: Whole-link travel time models (WTT) and the Kinematic wave theory.

- **Equilibrium mechanisms.** To reconcile the predictions of items 1 and 2.

A specific dynamic assignment model can be viewed as a combination of the previous three submodels. Below, each of these elements is analysed.

### 3.2.3 Characteristics of traffic demand

The demand for transport is derived, it is not an end in itself. With the possible exception of sightseeing, people travel in order to satisfy a need (work, leisure, health) undertaking an activity at particular locations. This is equally significant for goods movements. In order to understand the demand for transport, we must understand the way in which these activities are distributed over space, in both urban and regional contexts. A good transport system widens the opportunities to satisfy these needs; a heavily congested or poorly connected system restricts options and limits economic and social development. The demand for transport services is highly qualitative and differentiated. There is a whole range of specific demands for transport which are differentiated by time of day, day of week, journey purpose, type of cargo, importance of speed and frequency, and so on. A transport service without the attributes matching this differentiated demand may well be useless.

#### 3.2.3.1 The demand intensity

For each defined O-D pair $i \in I$ we consider that there is a time-dependent function $d_i(t)$ which defines the flow intensity in the $i$ pair in function of
time. The number of users travelling during the interval \([a, b]\) for the \(i\) pair is calculated by the integral \(\int_a^b d_i(t) \cdot dt\). These functions define the dynamic O-D flows on the road network.

### 3.2.4 Characteristics of traffic supply

The first distinctive characteristic of transport supply is that it is a service and not a good. Therefore, it is not possible to stock it, for example, to use it in times of higher demand. A transport service must be consumed when and where it is produced, otherwise its benefit is lost. For this reason it is very important to estimate demand with as much accuracy as possible in order to save resources by tailoring the supply of transport services to it. Many of the characteristics of transport systems derive from their nature as a service. In very broad terms a transport system requires a number of fixed assets, the infrastructure, and a number of mobile units, the vehicles. It is the combination of these, together with a set of rules for their operation, that makes possible the movement of people and goods.

#### 3.2.4.1 Traffic network representation

Transport demand takes place over space. This seems a trivial statement but it is the distribution of activities over space which makes for transport demand. There are a few transport problems that may be treated, albeit at a very aggregate level, without explicitly considering space. However, in the vast majority of cases, the explicit treatment of space is unavoidable and highly desirable. The most common approach to treat space is to divide study areas into zones and to code them, together with transport networks, in a form suitable for processing with the aid of computer programs.

The representation of the transportation network is a key part of the supply modelling effort, i.e. what the transport system offers to satisfy the movement needs of trip makers in the study area. The aim of the network representation is to represent the real network of a city with its physical structure into a model.

The transport network may be modelled at different levels of aggregation. At one extreme one has models with no specific links at all; they are based on continuous representations of transport supply (see Smeed (1968)). These models may provide, for example, a continuous equation of the average traffic capacity per unit of area instead of discrete elements or links.

Normal practice, however, is to model the road network as a directed graph \(G = (N, A)\), i.e. a set of nodes \((N)\) and a set of links \((A)\) joining them (see Larson and Odoni (1981)), where most nodes are taken to represent intersections and the links represent roadways or streets. Links are characterised by several attributes such as length, speed, number of
lanes and so on, and are normally unidirectional. A subset of the nodes is associated with points of trip generation called centroids, and a subset of the links to centroid connectors. All trips are assumed that start or end in the centroids. An example of this modelling is given in Figure 3.7 and represent the Nguyen and Dupuis network topology (see Nguyen and Dupuis (1984)) which will be widely used throughout this thesis. The network shows the connectors as discontinued lines and the links as solid lines, the centroid nodes are represented as triangles and the intersections as circles.

Figure 3.7 shows a Nguyen-Dupuis network (see Nguyen and Dupuis (1984)) including 17 nodes connected by 23 links. Each link in the network is associated with a direction of flow. It is possible to distinguish two subsets of nodes \{c_1, c_4\} origin nodes and \{c_2, c_3\} destination nodes.

Currently, the principal source of network data would be one of the many digital maps available for most cities. The modelling of an area is a fundamental step in order to find a trade-off of accuracy and computational complexity for study area.

**Figure 3.7:** Traffic network representation

### 3.2.4.2 Requirements for continuous dynamic network loading problem

Traffic performance models for dynamic traffic assignment consist of the estimation of link travel times using the different sections of the road network and considering variability over time, thus forming an extension of the conventional static assignment problem. Another function of traffic performance models for DTA is to describe the propagation of traffic flows
3.2. Dynamic traffic assignment

on time-varying networks. This is a major difference from static traffic assignment. Flow propagation is a unique feature of DTA and is described through traffic performance models.

The importance of traffic performance models for DTA lies in the fact that the results of flow propagation and travel time estimation affect the traffic flows to be assigned to each route. Without an appropriate description of flow propagation on the network and a reasonable estimation of travel time, realistic time-varying route choice cannot be accomplished.

Deterministic dynamic traffic assignment models have been developed mainly under three approaches: i) simulation, ii) mathematical programming and iii) optimal control over networks. These models consider that users minimize their travel time, continuously updating their chosen routes according to traffic conditions. The problem is especially important for ITS in special for advanced traveller information systems (ATIS), which require information to recommend routes according to the future travel time in the routes. It aims to improve the behaviour of traffic, reducing congestion by providing uniform traffic conditions. In these systems there is a central controller (we call it smart traffic management system) that recommends routes to the users in real time.

There are requirements that for the appropriate traffic performance models should meet for their application to dynamic transportation networks. Most of them have been identified from the literature: flow conservation and flow propagation are the most important but others requirements as the first in/first out (FIFO) principle, causality, and reasonable behaviour of outflow are also relevant.

• Flow propagation. In static assignment, flow propagation is achieved implicitly because flow is assumed to be constant over each entire route from its origin to its destination. In dynamic assignment flow should propagate through a link in a manner consistent with the speed of vehicles. The minimum time taken for a vehicle to traverse the link should not be shorter than the free flow travel time. The flow propagation equation for a link $a$ can be expressed as follows:

$$E_a(t) = V_a[\tau_a(t)] \tag{3.11}$$

where $\tau_a(t)$ is the link exit time for a vehicle that entered at time $t$ at link $a$; and $E_a(t)$ and $V_a[\tau_a(t)]$ are accumulated inflow and accumulated outflow at times $t$ and $\tau_a(t)$ for the link $a$ respectively.

• Flow conservation. The conservation of traffic flow is the most important requirement that traffic performance models should meet. One cannot imagine that a traveller who has entered the network will vanish before reaching his/her destination, or that at any time the total outflow exceeds the total inflow to a link. The flow conservation relationship
can be expressed as follows:

\[ E_a(t) = V_a(t) + x_a(t) \text{ for all } t \geq t_0 \]  (3.12)

This equation states that traffic on the link \( a \) at time \( t \) is equal to the difference between total inflow and total outflow up to time \( t \) from the earliest time, \( t_0 \). For this relationship to hold it should be assumed that the link is empty at time \( t_0 \). By adopting the flow conservation equation, any traffic performance model is guaranteed to respect the flow conservation requirement provided the outflow is calculated through the flow propagation equation.

### 3.2.4.3 Analytical macroscopic link models

The link traversal time can be computed using the following analytical macroscopic link models:

1. **Kinematic model.** This model is founded on representing the phenomenon of traffic as a group of differential equations in partial derivatives. When traffic is stationary and homogeneous, we know that the values for these variables will remain constant along the entire road and for some extended periods. However, real traffic is neither homogeneous nor stationary and it is necessary to describe the evolution of traffic over time in order to describe the dynamic relation between \( q(x, t) \), \( u(x, t) \) and \( k(x, t) \). It is assumed that we are dealing with point variables which are singularly defined at any moment \( t \) and an every location \( x \).

The following traffic conservation laws are used to describe the changes in time and location of the macroscopic variables along a road. The fundamental relation in a link \( a \):

\[ q_a(x, t) = k_a(x, t)u_a(x, t) \]  (3.13)

The next following partial differential equation represents the conservation law of traffic:

\[ \frac{\partial k_a(x, t)}{\partial t} + \frac{\partial q_a(x, t)}{\partial x} = z_a(x, t) \]  (3.14)

The expression \( z_a(x, t) \) represents the volume of traffic that enters the road at time \( t \) and location \( x \).

We add another assumption to this conservation law: all possible dynamic traffic state comply with the stationary fundamental diagrams. This mean that although traffic state on roads can change over time, they still comply with the fundamental diagrams at each moment and at every location. Therefore the successive traffic states ‘move’ as it were across the bold black lines in the fundamental diagrams.
This assumption allows us to write the flow rate in function of density as follow:

\[ q_a(x, t) = Q_a(k_a(x, t)) \]  \hspace{1cm} (3.15)

Using the fundamental diagram in the traffic conservation law led to the first dynamic Lighthill traffic model in 1950s. This model was named the LWR-model, (see Lighthill and Whitham (1955)). Several schemes were developed to numerically solve this equation with the help of a computer in order to obtain a traffic model that could be applied to practical situations.

In order to describe the changes in a discretized time-processing of the macroscopic variables along a road, Daganzo proposed the so-called Cell-Transmission Model (see (Daganzo, 1995a)) as a method which automatically generates appropriate changes in density at locations where the hydrodynamic theory would call for shockwave.

\[ t \]

\[ t + \Delta T \]

**Figure 3.8:** Road discretization used in the cell transmission model

To this end the road is divided into homogeneous sections called cells (see Figure 3.8), numbered consecutively, starting with the upstream end of the road, from \( i = 1 \) to \( I \). The lengths of the sections are not chosen arbitrarily; they are set equal to the distances travelled in free traffic conditions by a typical vehicle in one clock tick denoted as \( T \).
Under light traffic then, all the vehicles in a cell can be assumed to advance to the next with each tick; it is unnecessary to know where within the cell they are located. Thus, the evolution of the system obeys:

\[ x_{i+1}(t + \Delta T) = x_i(t) \]  \hspace{1cm} (3.16)

where \( x_i(t) \) is the number of vehicles in cell \( i \) at time \( t \). Figure 3.8 shows the model in a graphic way, the expected inflow is given by \( q_{i-1}(t) \) and the outflow is given by \( q_i(t) \). It is assumed that the above recursion holds for all flows, unless traffic is slowed down by queuing from a downstream bottleneck.

To incorporate queuing, the following two variables are introduced:

- \( X_i(t) - x_i(t) \), the amount of empty space in cell \( i \) at time \( t \).
- \( x_{i-1}(t) \), the number of vehicles in cell \( i - 1 \) at time \( t \).

The last quantity will ensure that the vehicular density on every section of the road remains below jam density. The proposed simulation is based on a recursion where the cell occupancy at time \( t + \Delta T \) equals its occupancy at time \( t \), plus the inflow and minus the outflow; i.e.,

\[ x_i(t + \Delta T) = x_i(t) + (q_{i-1}(t) - q_i(t))\Delta T \]  \hspace{1cm} (3.17)

where the flows are related to the current conditions at time \( t \) as indicated below:

\[ q_i(t) = \min \left\{ \frac{x_i(t)}{\Delta T}, Q \left( \frac{x_i(t)}{T} \right) , \frac{X_{i+1}(t) - x_{i+1}(t)}{\Delta T} \right\} \]  \hspace{1cm} (3.18)

The simulation would step through time, updating the cell occupancies (for all \( i \)) with each tick of the clock.

The cell transmission model is a discrete approximation to the Lightill, Whitham, Richards hydrodynamic model with a density-flow \( (k - q) \) relationship in the shape of an isosceles trapezoid, as in Figure 3.9.

This relationship can be expressed as:

\[ q = Q(k) = \min\{u_c,k, u_c(k_j - k)\}, \text{for } 0 \leq k \leq k_j \]  \hspace{1cm} (3.19)

where \( u_c \) is the free-flow speed (and the speed of all backward moving waves), \( k_j \) is the jam density, and \( u_c \leq k_j v/2 \) is the maximum flow.

2. Whole Travel Time model (WTT).

This model assumes the existence of instantaneous changes in the propagation of density within the links of a network and reflects the behaviour of the vehicles in macroscopic form and the link i.e., in global form. A generic macroscopic way of this model is:
3.2. Dynamic traffic assignment

Flow conservation.

\[ \frac{dx_a(t)}{dt} = u_a(t) - v_a(t) \]  

(3.20)

Where \( u_a(t) \) and \( v_a(t) \) are the inflow and outflow for the link \( a \) at the instant \( t \).

Flow behaviour (one of these.)

\[ \tau_a(t) = f(\Gamma, \Omega(t)) \]  

(3.21)

or

\[ v_a(t) = g(\Gamma, \Omega(t)) \]  

(3.22)

Flow propagation.

\[ \int_t^{t+\tau_a(t)} v_a(s)ds = x_a(t) \]  

(3.23)

where \( \tau_a(t) \) is the link traversal time (the time it takes for the vehicle entering at time \( t \) to traverse the link \( a \)); \( x_a(t) \) is the number of vehicles on a link \( a \) at time \( t \) and \( \Gamma \) is a vector of parameters reflecting the physical characteristics of a specific link such as the free flow travel time and bottleneck capacity; and \( \Omega(t) \) is a vector of the current link state variables such as \((x_a(t), u_a(t), v_a(t))\).
In the above modelling framework, the two ways of specifying flow propagation speed lead to two types of link models: the delay-function model and the exit-flow function model. Examples of the former include the linear delay-function model studied by many researchers: Friesz et al. (1993), Astarita (1996), Wu et al. (1998), Xu et al. (1999), and examples of the latter include the seminal paper M-N model proposed by Merchant and Nemhauser (1978a, 1978b), Janson (1991), Ran et al. (1996), and Chen and Hsueh (1998) and the widely-used point queue (PQ) model (Smith (1984)).

In transportation, various queues arise when the demands of service are higher than the supplies, and the PQ model has been proposed to study such queueing systems. The basic PQ derives its name from the following assumptions: (1) the physical length of any vehicle is zero; (2) vehicles move along the link at free flow speed before they arrive at the exit node; (3) a point queue (no physical length) forms at the exit node if the traffic arriving at the exit is more than the link capacity at the exit. The model has been implemented in several DTA studies.

3.2.5 Reconciliation mechanisms between demand and supply

As every physical model, or situation in life, when a change occurs in the system, it searches for the equilibrium state. The same occurs in a transportation system, when the system is in disequilibrium (changes in infrastructure, control policies, etc.), there are forces that tend to direct the system toward the equilibrium state. Nowadays, the models used and developed in transportation planning are based in the balance between supply and demand in the system.

In an operational context, the objective of DTA models is to represent traffic evolution on road networks when conditions change. They seek to describe the assignment of demand on different paths connecting every OD pair in an equilibrium state.

*User equilibrium* is achieved when no driver can unilaterally reduce his travel costs by moving to another route (see Wardrop (1952)). Wardrop’s user equilibrium principle of route choice can be formulated for the dynamic problem as follows:

*Under equilibrium conditions in networks where congestion varies over time traffic arranges itself so that at each instant the costs incurred by drivers on those routes that are used are equal and no greater than those on any unused route.*

If travellers choose not only route but also departure time, Wardrop’s equilibrium expression can be further extended:
3.2. Dynamic traffic assignment

Under equilibrium conditions in networks where congestion varies over time and travellers can choose their time of travel, traffic arranges itself so that the total cost associated with travel on those route that are used by travellers at the time when they are used, are equal and no greater than those on any route at a time when it is not used.

Stochastic User Equilibrium will be reached when no traveller believes that his travel time can be improved by unilaterally changing route. The stochastic user equilibrium is a generalization of the user equilibrium definition. If the perceived travel times are assumed to be entirely accurate, all drivers would perceive the same travel time and the stochastic user equilibrium will be identical to the (deterministic) user equilibrium. Daganzo and Sheffi (1977) were the first authors to define the concept of SUE for static case, and they used several versions as the simple multinomial logit (MNL) model or the more complex multinomial probit (MNP) model.
A Monte Carlo approach to simulate the stochastic demand in a CDNLP
4

A Monte Carlo approach to simulate the stochastic demand in a continuous dynamic traffic network loading problem

4.1 Introduction

The complexity of current transportation systems implies the need for analytical tools capable of providing an adequate information system and make predictions in real time. The traffic assignment models describe the behaviour of users on a traffic network. The static assignment models assume that over a time period demand conditions remain the same, i.e. these models assume steady-state flows. The evolution of static models to dynamic ones is due to several reasons, including the inability to explain flow changes in short time periods and the underestimation of total travel times.

Dynamic traffic assignment (DTA) models consider that users try to minimize their travel time, continuously updating their chosen routes according to traffic conditions. The DTA has evolved substantially since the seminal work of Merchant and Nemhauser (1978).

In order to be able to create a unified theory of dynamic traffic assignment modelling (as occurs in the static case), it is necessary to compare and to establish the relationships between the existing models. In Daganzo (1995a) it is considered that a basic theory of DTA models should include at least realistic models of: 1) Traffic propagation when the paths used are known. This sub-model represents the continuous dynamic network loading (CDNL) problem. 2) Path choice model when the time-varying link times are known.
3) Equilibrium to reconcile the predictions of items 1) and 2). A specific DTA model can be viewed as a combination of the previous three sub-models.

Two types of loading procedures can be found in the DTA literature:

- **Simulation-based methods.** Generally speaking, loading procedures based on simulation work with individual entities (vehicle or vehicle packet). At each simulation step the location of each entity within the network is recorded. Such loading procedures provide more flexibility and can, accordingly, deal with more complex traffic phenomena obtaining more realistic loading results. Nevertheless, these procedures share certain drawbacks as well. For instance, as much more detailed information about both the network and vehicles is required, the resulting loading procedures usually prove to be computationally intensive (often slow and memory-consuming). The simulation is called microscopic when the movements of the vehicles are traced individually. DynaMIT Ben-Akiva et al. (2006) represents a simulation-based real time system that estimates and predicts the state of a transportation network adopting such an approach. For reviews on microscopic simulation-based models see Hoogendoorn and Bovy (2001), Panwai and Dia (2005), and Brockfeld and Wagner (2006).

The mesoscopic simulation assumes that the vehicles are clustered into packets and the link performances are expressed in an aggregate way (see Celikoglu (2013), Celikoglu et al. (2009), Celikoglu and Dell’Orco (2008), Celikoglu and Dell’Orco (2007), Dell’Orco (2006)). Since its emergence, the point-queue (PQ) model (see Vickrey (1969)) was widely used as link model due to its good predictions at low computational cost (see Nie and Zhang (2005)). Nowadays more realistic approaches of queue (see Adamo et al. (1999)) have been developed to give a more detailed representation of congestion effects.

Currently, mesoscopic CDNL model has been extended and applied to multimodal flow and pedestrian flow, in ordinary and emergency conditions (see Di Gangi (2011) Gangi and Velonà (2009) Di Gangi et al. (2008)).

- **Macroscopic approaches** are loading procedures based on analytical models. These approaches work with macroscopic traffic flow variables that describe the average behaviour of vehicles within the network. The variables are time-dependent and describe states such as inflow/out-flow rate, volumes and travel times for each link in the network. Thus, these procedures are comparatively simple to implement and efficient to compute, even for large road networks, but are often believed to be less accurate in capturing complex traffic phenomena than simulation based loading procedures. A review of analytical CDNL models is presented in Friesz et al. (2011).

All loading models mentioned above are deterministic because they capture average traffic states but they do not deal with the stochastic nature
4.1. Introduction

of the demand. The main existing types of stochastic dynamic traffic models are based on time series models (see Lee and Fambro (1999), Tan et al. (2009)) or Bayesian networks (see Zhang et al. (2004), Sun et al. (2006), Castillo et al. (2008b), Castillo et al. (2012b)).

These papers are focused on estimating travel times using sensors. The problem of travel time prediction is different from estimation. These methods are able to obtain the travel times by tracking the driving routes on real networks or on traffic models. In the context of travel time predictions the references for stochastic load models are scarce.

The parameter calibration of traffic models has a bilevel structure (see García-Ródenas and Verastegui-Ray (2013), Montanino et al. (2012), Antoniou et al. (2011), Ciuffo et al. (2008), Balakrishna et al. (1999), Wu et al. (2003)) in which the lower level is given by the traffic model and the upper level is focused on the estimation problem. Calibration and prediction under uncertainty requires massive evaluation of these models which makes the computational aspect a major barrier. Moreover, recent studies show that CDNL problem consumes the larger part of the running time in the solution algorithm (see Nie et al. (2004)). As the size of a real road network could be large, DTA solution algorithms require efficient loading procedures.

Discrete event algorithms are widely used in transportation research, for example in railway research for detecting the conflicts produced in a given schedule Espinosa-Aranda and García-Ródenas (2012) or for solving the vehicle re-scheduling problem in cases of emergency Almodóvar and García-Ródenas (2013).

This chapter discusses the efficient resolution of the mesoscopic CDNL model based on a discrete event simulation. The adopted mesoscopic CDNL model is a broader version of the whole-link model of Adamo et al. (1999). This link model aims to find a trade-off between computational tractability and congestion representation.

The main contributions of this chapter are:

- A discrete event algorithm proposed to solve the mesoscopic CDNL model. The main characteristic is high computational performance. In this scheme, unlike with the traditional time discretization models, at each time step only one packet is updated (event).
- Distributed computing approaches can be adopted to address the computational burden for solving the stochastic CDNL model (see Liu and Meng (2013) Huang et al. (2010)). In this chapter parallel computing techniques are used to address the demand uncertainty by implementing a Monte Carlo simulation approach.
- The adopted whole-link travel time model allows time-dependent reduction of the network capacity to be tackled. The resulting CDNL model presents a FIFO behaviour which is used to accelerate the
procedure.

The proposed discrete event algorithm requires reduced computational costs which allows its implementation in real time. This enables the development of efficient tools capable of evaluating travel times in real time. This challenge is central to route recommendation problems in Intelligent Transportation Systems (ITS). From an off-line point of view, this computational advantage enables the consideration of stochastic demands which allows operational planning of traffic network incorporating demand uncertainty. The CDNL model with stochastic demand obtains the probability distribution of the travel times. The reliability-based traffic assignment models (Shao et al. (2006), Lo et al. (2006)) require these inputs within the route choice model.

The chapter is structured as follows. In Section 4.2 the basic notation and the formulation of the model is described. In Section 4.3 the proposed solution procedure and its computing components for the network simulation are introduced. Numerical experiments are carried out in Section 4.4 and finally some conclusions are drawn in Section ??.

### 4.2 A continuous dynamic network load model

A CDNL problem consists of, given the path flows and link performance functions, determining time-dependent network flow conditions such as link travel times, link counts, link inflows and link outflows. An accurate way of utilizing a DNL approach is to formulate the DNL problem as a continuous-time system of non-linear equations expressing link dynamics, flow conservation, flow propagation and boundary conditions. In this section, we describe the proposed CDNL model.

#### 4.2.1 Nomenclature

The general notations of the variables and indices used in this chapter are given below:

- **Sets and index**
  - \( \mathcal{N} \) Set of nodes.
  - \( \mathcal{P} \) Set of paths.
  - \( \mathcal{O} \) Set of origins.
  - \( \mathcal{A} \) Set of links of the traffic network.
  - \( \mathcal{W} \) Set of origin-destination pairs.
  - \( i \) Packet.
4.2. A continuous dynamic network load model

- **a** Link.
- **p** Path.
- **w** Origin-destination pair.
- **t** Time.
- \(A_p\) Set of links of path \(p\).
- \(\mathcal{P}_\omega\) Set of paths for the origin-destination pair \(\omega\).
- \(\theta\) Random variable that defines the parameters for the demand.
- \((0, T]\) Simulation period.
- \((0, T_\infty]\) Time interval from the instant when flows enter the network to the last instant when all flows leave the network.

**Functions**

- \(h_p(t)\) Inflow rate at the origin of path \(p\) at time \(t\).
- \(u^a_p(t)\) Entrance flow (inflow) to link \(a\) from path \(p\) at time \(t\).
- \(v^a_p(t)\) Exit flow (outflow) to link \(a\) from path \(p\) at time \(t\).
- \(V^a_p(t)\) Cumulative exit flow (outflow) to link \(a\) from path \(p\) at time \(t\).
- \(x^a_p(t)\) Number of vehicles on link \(a\) from path \(p\) at time \(t\).
- \(\tau^a(t)\) Exit time from link \(a\) associated with entering at time \(t\).
- \(u^a_r(t), v^a_r(t), x^a_r(t), \tau^a_r(t)\) As previous definitions given for paths but for aggregate flows.
- \(u^a_q(t), v^a_q(t), x^a_q(t), \tau^a_q(t)\) As previous definitions given by paths but for the vertical queue section.
- \(d_\omega(t, \theta)\) Intensity of demand at time \(t\).
- \(\alpha_a\) Travel time in the running section of the link \(a\).
- \(\beta_a(t)\) Capacity of the link \(a\) at time \(t\).
- \(\mathbb{P}(p, t)\) Choice probability associated to path \(p\) at time \(t\).

**Algorithm**

- \(T\) Simulation clock.
- \(Q_a\) Set of packets that are currently in the running section of the link \(a\).
- \(\phi_a\) Last instant in which a packet has left the link \(a\).
- \(\hat{a}_i\) Current link for the packet \(i\).
- \(\tilde{a}_i\) Next link for the packet \(i\).
- \(\tau_i\) Instant when the packet \(i\) leaves its current link.
- \(n_i\) Ordinal number of the current link within the followed path \(p\).
CHAPTER 4. A MONTE CARLO APPROACH TO SIMULATE THE STOCHASTIC DEMAND IN A CDNLP

4.2.2 Formulation of the basic problem

Consider the Nguyen-Dupuis network topology \( G = (N, A) \), where \( N \) and \( A \) denote the sets of nodes and directed links, respectively. The set \( N \) contains two types of nodes which represent the intersections and the so-called centroids, which are dummy nodes representing generation and attracting zones for the traffic demand, i.e. access/egress nodes of trips. The links are classified into centroid connectors, which link the centroids with the intersections, and the normal links, which represent roads. An example of this modelling is given in Figure 3.7. The centroid connectors are represented as discontinuous lines and the road links as solid lines. The centroids are represented as triangles and the nodes as circles.

CDNL model can be formulated with the following system of equations:

- **Flow conservation** equations

  \[
  x_{ap}^p(t) = x_{ap}^p(0) + \int_0^t [u_{ap}^p(\xi) - v_{ap}^p(\xi)] d\xi \\
  \forall p \in P, \forall a \in A_p. \tag{4.1}
  \]

  \[
  u_{ap}^p(t) = h_p(t) \quad \forall p \in P\tag{4.2}
  \]

  where \( a' \) is the first link (centroid connector) of path \( p \).

  The flow conservation equation between two consecutive links can be expressed as

  \[
  u_{ap}^p(t) = u_{ap}^{-p}(t) \quad \forall p \in P, \forall a \in A_p, \tag{4.3}
  \]

  where \( a \) and \( a^{-} \) are two consecutive links of path \( p \).

- **Flow propagation** equations describe the flow progression over time as

  \[
  V_{ap}^p(t) = \int_{I_a(t)} u_{ap}^p(\xi) d\xi \quad \forall p \in P, \forall a \in A_p \tag{4.4}
  \]

  where

  \[
  I_a(t) = \{\xi : \tau_a(\xi) \leq t\}. \tag{4.5}
  \]

  and \( \tau_a(t) \) is the exit time from link \( a \) for a vehicle entering at time \( t \).

- **Boundary conditions.** For simplicity, they are assumed to be empty at \( t = 0 \), i.e.

  \[
  v_{ap}^p(0) = 0, \quad x_{ap}^p(0) = 0, \quad u_{ap}^p(0) = 0, \quad \forall p \in P, \forall a \in A_p. \tag{4.6}
  \]
4.2. A continuous dynamic network load model

Equations (4.1)–(4.3) model the conservation principles regarding hydrodynamic analogy. Equations (4.4)–(4.5) describe the flow progression over time. The flow entering link $a$ at time $t$ exits the link at $\tau_a(t)$ and the set $I_a(t)$ describes the elapsed time when the flow leaves the link $a$ by time $t$. Therefore, the relationships between aggregated and disaggregated variables are calculated as follows:

$$x_a(t) = \sum_p \delta_{ap} x_a^p(t)$$  \hfill (4.7)

$$v_a(t) = \sum_p \delta_{ap} v_a^p(t)$$  \hfill (4.8)

$$u_a(t) = \sum_p \delta_{ap} u_a^p(t)$$  \hfill (4.9)

where $(\delta_{ap})$ is the link-path incidence matrix, with $\delta_{ap} = 1$ if link $a$ is on path $p$ and $\delta_{ap} = 0$ otherwise. In the above expressions, $x_a(t)$ is the link volume, i.e. the number of vehicles on a link at time $t$.

4.2.3 Adopted link dynamic model: a generalized point queue model

The CDNL model described above is completely defined by specifying how to calculate travel times on the links.

The link traversal time can be computed by using Whole Travel Time, (WTT) models. As it was seen in Chapter 3, these models assume the existence of instantaneous changes in the propagation of density within the links of a network and reflect the behaviour of the vehicles in a macroscopic way. A generic macroscopic whole link model for a given link $a$ is formed by the following equations:

- Flow conservation equation

$$x_a(t) = x_a(0) + \int_0^t [u_a(\xi) - v_a(\xi)] \, d\xi$$  \hfill (4.10)

- Flow behaviour (one of:)

$$\tau_a(t) = f(\Gamma, \Omega(t)); \text{ delay-function model}$$  \hfill (4.11)
or

\[ v_a(t) = g(\Gamma, \Omega(t)); \text{ exit-flow function model} \quad (4.12) \]

- Flow propagation

\[
\int_{t}^{\tau_a(t)} v_a(s) ds = x_a(t) \quad (4.13)
\]

where \( \Gamma \) is a vector of parameters reflecting the physical characteristics of a specific link such as the free flow travel time and bottleneck capacity and \( \Omega(t) \) is a vector of the current link state variables such as \((x_a(t), u_a(t), v_a(t))\).

Daganzo (1995b) indicates that only two special cases of WTT models depict transportation-like phenomena. Each model denotes a link with no spatial dimension containing a point queue or a link with constant travel time and no queueing. Road segments exhibiting both phenomena must be represented by two links in series. This is how the basic point queue (PQ) model operates. The PQ model has been implemented in several DTA studies Drissi-Kaletoui and Hameda-Benchekroun (1992), Gangi et al. (1996), Kuwahara and Akamatsu (1997), Li et al. (2000).

In this work a generalized PQ model is proposed and can be described as a concatenation of two link models: the so-called \textit{running section} that models the travel time in the link and the \textit{vertical queue} that models the link capacity restrictions. A delay-function model is used in the running section while in the vertical queue section a exit-flow function model is employed.

Other methods Celikoglu et al. (2009) require the resolution of fixed-point problems and the maximizing of the problem of flow optimization in each node as a part of its analytical rules or interpolation method. Castillo et al. (2012a), Wu et al. (1998), Xu et al. (1999), Rubio-Ardanaz et al. (2003). This could indicate a priori that the proposed approach is less time-consuming than those developed in the literature and could be applicable to large scale networks.

We assume that a link \( a \in A \) is modelled by a single segment with two sections as shown in Figure 4.1. The following relationships describe the variables of each section with the whole link variables.

\[
\begin{align*}
x_a(t) &= x^r_a(t) + x^q_a(t) \quad (4.14) \\
\tau_a(t) &= \tau^r_a(\tau^q_a(t)) \quad (4.15) \\
u_a(t) &= u^r_a(t) \quad (4.16) \\
v_a(t) &= v^q_a(t) \quad (4.17) \\
v^r_a(t) &= v^d_a(t) \quad (4.18)
\end{align*}
\]
4.2. A continuous dynamic network load model

Figure 4.1: Diagram of our proposed link divided into two sections

Where the superscripts relate the meaning of the variable to the appropriate section.

\[ x_a(t) = x_r^a(t) + x_q^a(t) \] (4.19)
\[ \tau_a(t) = \tau^q_a(\tau_r^a(t)) \] (4.20)
\[ u_a(t) = u_r^a(t) \] (4.21)
\[ v_a(t) = v_q^a(t) \] (4.22)
\[ v_r^a(t) = u_q^a(t) \] (4.23)

Where the superscripts relate the meaning of the variable to the appropriate section.

1. Running section. This represents the traversal time of a vehicle in the whole link. In this chapter we assume a delay-function model. The most notable examples (static assignment traffic problem) are the so-called Bureau of Public Roads functions (BPR) (see TRB (1965)).

\[ f_a(x_r^a) = \alpha_a + \gamma_a \left( \frac{x_r^a}{\beta_a} \right)^{n_a} \] (4.24)

where \( x_r^a \) is the link volume in running section, \( \alpha_a \) is the travel time in the free-flow link, \( \beta_a \) is the link capacity, \( \gamma_a \) and \( n_a \) are parameters that should be calibrated (see García-Ródenas and Verastegui-Rayo (2013) Russo and Vitetta (2005)).

This function \( f_a \) defines the exit time as

\[ \tau_a^r(t) = t + f_a(x_r^a(t)) \text{ for all } t \in [0, T_\infty] \] (4.25)
where \( T_\infty \) denotes the latest instant when all flow leaves the network and \( f_a(x_a^r(t)) \) is the traversal time of link \( a \) at instant \( t \).

2. Vertical queue section. This section of the link models the capacity restrictions related to the flow that should be reached for each link \( a \in A \) and for each instant \( t \).

The first constraint of the model is that the outflow is lower than the link capacity for all instants.

\[
v_a^q(t) \leq \beta_a(t) \quad \text{for all } t \in [0, T_\infty]
\]

where \( \beta_a(t) \) is a function that dynamically defines link capacity.

The users in the queue satisfy the FIFO rule. This implies that a user is served at the first instant after all earlier users in the queue have been served. The exit time from the queue of a user who arrives at instant \( t \), is denoted by \( \tau_a(t) \) and satisfies the following equation:

\[
x_a^q(t) = \int_t^{\tau_a(t)} v_a^q(\xi)d\xi
\]

It is worth noting that the proposed model generalises both the delay-function model and the exit-flow function model. For example, the PQ model is obtained when \( \gamma_a = 0 \) and \( \beta_a(t) = \beta_a \) are used. Moreover, if \( \beta_a(t) = \infty \) it leads to the delay-function model.

**Theorem 4.2.1 (The Proposed Model is FIFO-Rule Consistent).** The model implied by expressions (4.19)–(4.27) together with the FIFO assumption of the running sections for all link \( a \), guarantees satisfaction of the FIFO rule.

**Proof.**

If all links in a path satisfy the FIFO condition, then the users that are on the same path also satisfy the FIFO condition. Note that for a generic link \( a \) the FIFO condition holds.

Let \( t_1 \leq t_2 \) where \( t_1, t_2 \in [0, T_\infty] \). If and only if \( \tau_a(t_1) \leq \tau_a(t_2) \). Using Eq. (4.20), the above relationship leads to:

\[
\tau_a^q(\tau_a^r(t_1)) \leq \tau_a^q(\tau_a^r(t_2)).
\]

It is assumed that the FIFO condition holds in the Running Section, thus

\[
\tau_a^r(t_1) \leq \tau_a^r(t_2)
\]
4.2. A continuous dynamic network load model

Let \( t^\prime_1 = \tau^\prime_q(t_1) \) and \( t^\prime_2 = \tau^\prime_q(t_2) \). Using the relationship (4.10) and the non-negativity of the inflow rate function \( u^a_q(t) \), we obtain:

\[
x^a_q(t_2^\prime) = x^a_q(t_1^\prime) + \int_{t_1^\prime}^{t_2^\prime} [u^a_q(\xi) - v^a_q(\xi)] d\xi \geq x^a_q(t_1^\prime) - \int_{t_1^\prime}^{t_2^\prime} v^a_q(\xi)d\xi \tag{4.28}
\]

Using Eq. (4.27) and the linearity of the integral, we obtain:

\[
x^a_q(t_1^\prime) = \int_{t_1^\prime}^{\tau^\prime_q(t_1^\prime)} v^a_q(\xi)d\xi = \int_{t_1^\prime}^{t_2^\prime} v^a_q(\xi)d\xi + \int_{\tau^\prime_q(t_1^\prime)}^{\tau^\prime_q(t_2^\prime)} v^a_q(\xi)d\xi \tag{4.29}
\]

Eq. (4.29) can be rewritten as:

\[
\int_{\tau^\prime_q(t_1^\prime)}^{\tau^\prime_q(t_2^\prime)} v^a_q(\xi)d\xi = x^a_q(t_1^\prime) - x^a_q(t_2^\prime) - \int_{t_1^\prime}^{t_2^\prime} v^a_q(\xi)d\xi \tag{4.30}
\]

Eq. (4.28) shows that the right side of Eq. (4.30) is negative. As \( v^a_q(\xi) \geq 0 \), we obtain \( \tau^\prime_q(t_2^\prime) \geq \tau^\prime_q(t_1^\prime) \).

Nevertheless, finding proper delay functions is challenging. For example, two decades after the work of Friesz et al. (1993), the linear delay-function is still the only delay-function of the form \( f_a(x^a_q(t)) \) that satisfies the FIFO condition under all inflow profiles. Additionally, recent studies indicate that probably there is no FIFO-consistent delay-function other than the linear one. In Garcia-Ródenas et al. (2006) it is shown that for BPR type the linear functions only satisfy the FIFO condition, and moreover that piecewise linear functions do not satisfy the FIFO condition. Taking this into account the linear delay-function model (i.e, \( n_a = 1 \) in BPR function) is assumed. This approach is presented in Astarita (1996).

One way of improving the predictions of the proposed model increasing the computational cost would be that, instead of applying the proposed travel-time model to the whole link, divide the link into segments, applying the model (suitably adjusted) sequentially to these segments. Carey and Ge (2005) proves that, as the discretization is refined, the solution converges to the solution of Lighthill-Whitham-Richards (LWR) model. Also the queue section can make close approximations of inhomogeneous links by treating the links as homogeneous, using the mean value as a capacity parameter.
4.2.4 Stochastic demand

The traffic network, barring any incidents, is accurately known while demand has a strong stochastic component to be considered. For this reason, we assume that the traffic demand intensity $d_w(t, \theta)$ for the origin-destination pair $w$ at the instant $t$ depends on a random variable $\theta$ which reflects the change in demand from day to day.

The flow intensity in the path $p$ is given by:

$$h_p(t, \theta) = \mathbb{P}(p, t)d_w(t, \theta) \text{ with } p \in \mathcal{P}_\omega$$  \hspace{1cm} (4.31)

where $\mathbb{P}(p, t)$ is the probability that a user of pair $\omega$ chooses path $p \in \mathcal{P}_\omega$ at instant $t$. Note that $\sum_{p \in \mathcal{P}_\omega} \mathbb{P}(p, t) = 1$. In this chapter, we assume that the function $\Pi(t) := (\cdots, \mathbb{P}(p, t), \cdots)$ and $D(t, \theta) := (\cdots, d_w(t, \theta), \cdots)$ are exogenous to the CDNL model and are known. The fact that $D(t, \theta)$ depends on a random variable means that flows and travel times on the paths of the network, denoted by $C(\Pi(t), \theta)$, are also considered random variables defined as the solution of the CDNL model. The aim in this chapter is to compute efficiently the vector of random travel times $C(\Pi(t), \theta)$. It does not imply that the route choice model $\Pi(t)$ is independent of the level of congestion but is rather considered outside of the CDNL model. This matter is analysed below.

DTA models integrate a traffic demand model, a supply model (the CDNL model) and a demand/supply interaction model. The demand model represents the decisions made by users on the traffic network, such as the choice of whether the journey is made, the destination, the starting time or the path selected. Discrete choice demand models McFadden (1974), Ben-Akiva and Lerman (1995), Cascetta (2001) can be used to explicitly define $\Pi(t)$ as a function of the vector of random travel times $C$ on the paths of the network.

$$\Pi(t) = G(t, C)$$  \hspace{1cm} (4.32)

The whole DTA model is defined by finding a fixed point $\Pi(t)$ such that:

$$\Pi(t) = G(t, C(\Pi(t), \theta))$$  \hspace{1cm} (4.33)

Eq. (4.33) shows that the route choice model $\Pi(t)$ depends on the level of congestion $C$. The demand models will allow the extension of the CDNL model to the context of multimodal networks, for example, García and Marín (2005) uses a nested logit distribution as a demand model which explicitly takes into account the choice of mode of transport by users and transfer node among modal networks.
4.3 Proposed solution procedure

The procedure proposed for solving the CDNL model is a hybrid of an analytical approach and a simulation-based approach. A discretization scheme is applied to analytical formulation of CDNL model and the analytical rules considering flow conservation, flow propagation, capacity and boundary conditions lead to a discrete-packet approach. Mesosimulation is employed in Leonard et al. (1989), Jayakrishnan et al. (1994), Celikoglu and Dell’Orco (2007), Dell’Orco (2006) and Celikoglu et al. (2009). The main difference between the proposed approach and the previous works is that instead of employing temporal discretization \([t, t + \Delta t]\), it is the flow \(\Delta x\) which is discretized.

The simulation algorithm discretizes the demand and obtains the so-called traffic packets travelling through the network. A packet is the set of vehicles leaving a centroid connector in a time interval (variable for each packet) and following the same path \(p\). For simplicity we assume that all packets have the same number of vehicles \(\Delta x\) (size of the packet).

In subsection 4.3.1 the rules which describe the link dynamic are summarized. In subsection 4.3.2 the mechanisms to synchronize all packets are described.

4.3.1 Discretization of link dynamic model

The basis of the mesosimulation procedure is the packet; the state of each packet \(i\) is defined by the pair of features \((\tau_i, n_i)\) where \(\tau_i\) is the departure time for the current link and \(n_i\) is the ordinal number of the current link in its path. In the beginning of the simulation all packets are in the first link, this is \(n_i = 1\) which represents the centroid connectors. Now the calculation of the first value of \(\tau_i\) is described.

Consider two types of link on the network: the centroid connectors and the road links. The former are used to make the demand load, and the latter are responsible for demand propagation within the network.

4.3.1.1 Dynamic of centroid connectors

We have two types of mesoscopic load computing components in centroid connectors: deterministic load, which assumes that the number of people who want to travel at every moment and the chosen path are known; and stochastic load, in which the existence of uncertainty in both choices is assumed.
* Deterministic load in centroid connectors. Suppose as starting point the set of flow paths \( \{ h_p(t) \}_{p \in P} \). In this case \( \theta \) is a known parameter. The associated packets are analysed for each path \( p \).

The path-link flow conservation equation (4.2) allows an inductive relationship to be obtained which determines the load of the packet \( i \) in the traffic network as a function of the previous packet \( i - 1 \). We assume that the \( i - 1 \)th packet has been loaded at the instant \( \tau_{i-1} \) of entry, and that \( \tau_{i-1} \) is less than \( T \) (the last time in the simulation). The next entry time for the packet \( i \) satisfies the equation:

\[
\Delta x = \int_{\tau_{i-1}}^{\tau_i} h_p(\xi) d\xi
\] (4.34)

If the solution holds \( \tau_i > T \), the packet would leave the temporal simulation horizon and would be considered null, proceeding to load demand in following path \( p = p + 1 \) by solving:

\[
\Delta x = \int_{0}^{\tau_i} h_p(\xi) d\xi
\] (4.35)

* Stochastic load in centroid connectors. The stochastic load is analogous to the deterministic case since the path associated with the packet \( i \) is chosen randomly according to the probabilities \( P(p, \tau_i) \) with \( p \in P_w \).

In order to define more formally the numerical scheme employed, suppose that a realization of the random variable \( \theta \) has been performed, and to do this, the demand \( \omega \) has been analysed and the packet \( i \) has been generated. The following equation is solved:

\[
\Delta x = \int_{\tau_{i-1}}^{\tau_i} d_\omega(\xi, \theta) d\xi
\] (4.36)

if \( \tau_i > T \) a null packet is generated and the following demand \( \omega = \omega + 1 \) is analysed by solving:

\[
\Delta x = \int_{0}^{\tau_i} d_\omega(\xi, \theta) d\xi
\] (4.37)
4.3. Proposed solution procedure

Then, the path \( p \) followed by the packet \( i \) should be randomly generated through the Monte Carlo simulation using the probabilities \( P(p, \tau_i) \) with \( p \in P_w \).

4.3.1.2 Dynamic of road links

Assume that packet \( i \) arrives at instant \( T \) at link \( a \). The new value \( \tau_i \), in which the packet leaves the link is computed as follows. It is necessary to consider the time to traverse the running section.

\[
\tau_i' = \tau_a'(T) = T + \alpha_a + \gamma_a \cdot x_a'(T)
\]  

(4.38)

As the model satisfies the FIFO condition, the instant in which the packet before \( i \) enters the vertical queue should be prior to \( T \) and therefore it is assumed that the entry time is known and denoted by \( \phi_a \).

Equation (4.26) implies:

\[
\int_{t_1}^{t_2} v_q^a(\xi) \, d\xi \leq \int_{t_1}^{t_2} \beta_a(\xi) \, d\xi \quad \text{for all } t_1, t_2 \in [0, T_\infty]
\]  

(4.39)

The last equation can be discretized. Let \( \phi_a \) be the last instant in which a packet leaves the link \( a \), then \( \tau_i \) is the next instant in which the packet \( i \) leaves the vertical queue according to the flow capacity restrictions. The equation (4.41) leads to:

\[
\Delta x = \int_{\max\{\tau_i', \phi_a\}}^{\tau_i} v_q^a(\xi) \, d\xi \leq \int_{\max\{\tau_i', \phi_a\}}^{\tau_i} \beta_a(\xi) \, d\xi
\]  

(4.40)

The value \( \tau_i \) is the least instant for which the previous relationship holds and so should satisfy:

\[
\Delta x = \int_{\phi_a'}^{\tau_i} \beta_a(\xi) \, d\xi
\]  

(4.41)

where

\[
\phi_a' = \max\{\tau_i', \phi_a\}.
\]  

(4.42)
It is worth noting that in equation (4.40) the value of $\tau''_i$ is previously computed using equation (4.38).

In the case of constant capacity $\beta_a(t) = \beta_a$ along the simulation period, equation (4.41) leads to

$$\tau_i = \phi'_a + \frac{\Delta x}{\beta_a}$$  

(4.43)

In appendix A the case in which $\beta_a(t)$ is a square wave has been solved. This type of capacity function allows for modelling signal-controlled intersections.

### 4.3.2 Discrete event algorithm for moving packets within traffic network

The discretized CDNL model can be interpreted as a queueing network in which each link is an individual queue. The customers (packets) follow an itinerary defined by the path initially chosen. Each packet is served at each link using the dynamic mechanism previously described. When a packet is served at a link, it can join another link and queue to be served or leave the traffic network.

We consider the following function

$$a := a(i, n)$$  

(4.44)

which returns the link $a$ that would be visited by the packet $i$th in its $n$th queue. This function defines the customer routing in the queueing network.

We will denote an event as the situation in which a packet leaves its current link in a given time. The proposed algorithm processes all events in time order. Let $T$ be the current instant of the simulation clock; the algorithm calculates the next event to be processed as the next packet that leaves the set of individual queues. This event is associated with a packet $i'$ and through the function:

$$\hat{a} = a(i', n_{i'})$$  

(4.45)

Where $n_{i'}$ is the number of links visited by the packet $i'$, and $\hat{a}$ is the current link. The link queue $\hat{a}$ is defined by the set $Q_{\hat{a}}$ of all packets that are currently
4.3. Proposed solution procedure

on the link \( \hat{a} \). This means that, when \( i' \) leaves the queue it should be updated. Thus, we can calculate the next link to visit:

\[
\tilde{a} = a(i', 1 + n_{i'})
\]  
(4.46)

and the queue \( Q_{\tilde{a}} \) is updated.

Table 4.1 summarizes the discrete event simulation algorithm.

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0. Initialization.</td>
<td>Let ([0, T]) be the simulation period and ( \Delta x ) be the packet size. Initialize the simulation clock ( T = 0 ) for all links ( a \in A ).</td>
</tr>
<tr>
<td>1. (Packet generation).</td>
<td>Load the packets deterministically or stochastically in the centroid connectors during the simulation period. Locate the packet ( i ) in the first link of its path (centroid connector) ( n_i = 1 ). By using equations (4.34)–(4.37) compute the exit times ( \tau_i ). Update all queues ( Q_a ) where ( a ) is a centroid connector. Otherwise ( Q_a = { } ) and ( \Gamma_a := +\infty ). Compute the next event for the queue ( a ) by: ( \Gamma_a := \min_{i \in Q_a} { \tau_i } ).</td>
</tr>
</tbody>
</table>
| 2. (Processing of events). | While \( T \leq T \).
| 2.1 (Update clock). | Let \( \hat{a} := \arg \min_{a \in A} \{ \Gamma_a \} \) be the link where the next event is produced (exit of packet from link \( \hat{a} \)) for analysis. Update clock \( T = \Gamma_{\hat{a}} \). Let \( i' \) be the packet associated with the event being analysed in the link \( \hat{a} \), that is, \( i' := \arg \min_{i \in Q_{\hat{a}}} \{ \tau_i \} \). and go to Step 2.3. If false
| 2.2 (Simulation of queue \( Q_{\hat{a}} \)). | If the packet \( i' \) has finished its trip calculate the exit time \( t_{i'} = T \). Take \( \phi_{\hat{a}} = T \).
Update queue \( Q_{\hat{a}} = Q_{\hat{a}} - \{ i' \} \).
Update its next event.
If \( Q_{\hat{a}} \neq \{ \} \) then compute \( \tau_i' := \min_{j \in Q_{\hat{a}}} \{ \tau_j \} \).
Using \( \phi_{\hat{a}} \), \( \tau_i' \) and equations (4.41)–(4.42) compute \( \tau_i \). Let \( \Gamma_{\hat{a}} = \tau_i \). Otherwise if \( Q_{\hat{a}} = \{ \} \) set \( \Gamma_{\hat{a}} = +\infty \).
| 2.2.2 Otherwise | let \( n_{i'} = n_{i'} + 1 \) and \( \bar{a} = a(i', n_{i'}) \). Compute \( x_{\bar{a}} = \{ j \in Q_{\bar{a}} : \tau_j \leq T \} \) \( \cdot \Delta x \) where \( \cdot \) is the cardinal of a set. Compute \( \tau_{i'} = \alpha_{\bar{a}} + \tau_{i'} \cdot x_{\bar{a}} \).
If \( Q_{\hat{a}} = \{ \} \) then compute \( \tau_i' \) by using \( \phi_{\bar{a}}, \tau_{i'} \) and equations (4.41)–(4.42); and let \( \Gamma_{\hat{a}} = \tau_i' \). \( Q_{\hat{a}} = Q_{\hat{a}} \cup \{ i' \} \). |

End
4.3.3 Monte Carlo simulation

The Monte Carlo method is very flexible and allows any random variables $\theta$ to be addressed. Assume that $d_\omega(t,\theta)$ depends on a random variable $\theta$ and the variables of interest are obtained by using the discrete event algorithm (travel times, flows, etc.) for each realization of the random variable. This process allows a random sample for any size of the variables of interest to be generated.

The main drawback of the Monte-Carlo approach is that it is a very time-consuming task. Nevertheless, the possibility of using distributed computing strategies makes it a useful alternative.

4.4 Numerical experiments

In this section, the proposed method is illustrated and evaluated. The proposed discrete event algorithm has been implemented in MATLAB.

The Nguyen-Dupuis Nguyen and Dupuis (1984) and the Sioux Falls networks have been used to test the model.

The Nguyen-Dupuis network consists of thirteen nodes and four O-D pairs as shown in Figure 3.7. The running section is defined by the linear cost functions (4.38) with parameters $\gamma_a = 0$ and $\alpha_a$ included in Table 4.2. In the queue section (4.41)–(4.42) the capacity is considered constant $\beta_a(t) = \beta_a$ over time. These parameters are also shown in Table 4.2.

**Tabla 4.2:** Parameters of the Nguyen-Dupuis network example

<table>
<thead>
<tr>
<th>Link $a$</th>
<th>$\alpha_a$ (hours)</th>
<th>$\beta_a$ (veh./hour)</th>
<th>Link $a$</th>
<th>$\alpha_a$ (hours)</th>
<th>$\beta_a$ (veh./hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1, 5)</td>
<td>7/60</td>
<td>2500</td>
<td>(1, 12)</td>
<td>9/60</td>
<td>2500</td>
</tr>
<tr>
<td>(4, 5)</td>
<td>9/60</td>
<td>2500</td>
<td>(4, 9)</td>
<td>12/60</td>
<td>2500</td>
</tr>
<tr>
<td>(5, 6)</td>
<td>3/60</td>
<td>2500</td>
<td>(5, 9)</td>
<td>9/60</td>
<td>2500</td>
</tr>
<tr>
<td>(6, 7)</td>
<td>5/60</td>
<td>2500</td>
<td>(6, 10)</td>
<td>13/60</td>
<td>2500</td>
</tr>
<tr>
<td>(7, 8)</td>
<td>5/60</td>
<td>2500</td>
<td>(7, 11)</td>
<td>9/60</td>
<td>2500</td>
</tr>
<tr>
<td>(8, 2)</td>
<td>9/60</td>
<td>2500</td>
<td>(9, 10)</td>
<td>10/60</td>
<td>2500</td>
</tr>
<tr>
<td>(9, 13)</td>
<td>9/60</td>
<td>2500</td>
<td>(10, 11)</td>
<td>6/60</td>
<td>2500</td>
</tr>
<tr>
<td>(10, 12)</td>
<td>9/60</td>
<td>2500</td>
<td>(11, 3)</td>
<td>8/60</td>
<td>2500</td>
</tr>
<tr>
<td>(12, 6)</td>
<td>7/60</td>
<td>2500</td>
<td>(12, 8)</td>
<td>14/60</td>
<td>2500</td>
</tr>
<tr>
<td>(13, 3)</td>
<td>11/60</td>
<td>2500</td>
<td>(13, 1)</td>
<td>0</td>
<td>50000</td>
</tr>
<tr>
<td>(16, 4)</td>
<td>0</td>
<td>50000</td>
<td>(2, 14)</td>
<td>0</td>
<td>50000</td>
</tr>
<tr>
<td>(3, 15)</td>
<td>0</td>
<td>50000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.4. Numerical experiments

The demand intensity is given by

\[ d_\omega(t, \theta) := \sum_{i=1}^{2} \frac{\theta_\omega}{\sigma_i} e^{-\frac{1}{2} \left( \frac{t - \mu_i}{\sigma_i} \right)^2} \]  (4.47)

where the values of the parameters are given in Table 4.3.

In the experiments the daily demand intensity presents two peak hours (around 8:00 and 15:00). This behaviour is modelled as the sum of two Gaussian functions. The parameter \( \mu_i \) is associated with the peak hour \( i \) and the parameter \( \sigma_i \) defines its amplitude. The amount of users for the pair \( \omega \) is proportional to \( \theta_\omega \).

**Tabla 4.3:** Parameters of the O-D demand intensity functions

<table>
<thead>
<tr>
<th>Pair</th>
<th>( \theta_\omega )</th>
<th>( \mu_1 )</th>
<th>( \mu_2 )</th>
<th>( \sigma_1 )</th>
<th>( \sigma_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w_1 )</td>
<td>4000</td>
<td>8</td>
<td>15</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>( w_2 )</td>
<td>8000</td>
<td>9.5</td>
<td>15.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>( w_3 )</td>
<td>6000</td>
<td>7.5</td>
<td>14.5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>( w_4 )</td>
<td>2000</td>
<td>8.5</td>
<td>16</td>
<td>1.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Figure 4.2 shows the demand intensity for the four O-D pairs considered in the Nguyen-Dupuis network.

The Sioux Falls network consists of seventy-six nodes, six hundred paths and five hundred and twenty-eight O-D pairs; in this chapter we describe the Nguyen-Dupuis network example and in order to save space, the Sioux Falls network data used are available at http://bit.ly/18KVDfM.

**Figure 4.2:** The O-D demand intensity functions for the Nguyen-Dupuis network
4.4.1 Computational burden

The test problem has been solved on an Intel core 2 Quad Q9550 (2.83GHz) CPU with 4 GB of RAM memory with three different sizes of packet for both networks over a simulation period ranging from 00:00 to 23:00 hours. Simulation results are shown in Table 6.3.

Table 4.4: Test of computational burden for simulation of Nguyen-Dupuis and Sioux Falls network examples

<table>
<thead>
<tr>
<th></th>
<th>Nguyen-Dupuis</th>
<th>Sioux Falls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Packet size ($\Delta x$)</td>
<td>Number of packets</td>
<td>Simulation time (seconds)</td>
</tr>
<tr>
<td>10</td>
<td>6807</td>
<td>2.9</td>
</tr>
<tr>
<td>5</td>
<td>13617</td>
<td>6.15</td>
</tr>
<tr>
<td>1</td>
<td>68093</td>
<td>68.64</td>
</tr>
</tbody>
</table>

The computational cost depends on the number of packets analysed on the network. Taking into account that the reduction of the packet size causes an increase of the simulation time, a non-linear relationship between the simulation time and the packet size is evident. The size of the network influences computational cost linearly. The results show a high performance for the algorithm mainly due to the FIFO behaviour of the model, which means not having to reorder queues once the packet has changed from one queue to another.

It is worth noting that despite thousands of packets being analysed, the computational cost is in the order of seconds.

4.4.2 Example of deterministic demand

This section shows the simulation results with the Nguyen-Dupuis network. The goal of this section is to test the internal consistency of the model.

Figure 4.3 shows the travel time for all the paths in the network. Each colour corresponds to a specific path. It is possible to see that travel times in each path increase at different moments; it shows also an increase in congestion between 09:30 - 10:00 and 15:00 - 16:00 therefore greater travel times are obtained during peak hours. Note that the maximum intensity of
the demand is reached at times 9:00 and 15:30 but the maximum congestion is delayed half hour.

It is important to note that the network behaviour is similar to the PQ model applied to each link separately. The reason is that travel times are affected by variations in demand only if the network capacity is exceeded, i.e. the network experiences congestion.

![Graph showing travel times on different paths](image)

**Figure 4.3:** Nguyen-Dupuis network: travel times on the different paths of the illustrative example

A second output of the model shows the link loads as a function of the time. Figure 4.4 shows the link flows and how they vary dynamically over time, resulting in a zig-zag effect as a consequence of the discretization of the problem (packet size). Note that the number of packets in the connector link 21 decreases during the simulation.

Figure 4.5 shows the arrival times for the different paths, depending on the instant at which the trip is initiated. We can observe a monotonically increasing behaviour in all paths, which ensures the FIFO condition, and is one of the main contributions of this model.

### 4.4.3 Example of stochastic demand

In this section we have considered that \( \theta \), which represent the total demand, is a multivariate normal variable \( N(\bar{\theta}, \Sigma) \), i.e., with the following probability density function:

\[
f(\theta) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp \left\{ -1/2(\theta - \bar{\theta})^T \Sigma^{-1}(\theta - \bar{\theta}) \right\}
\]

where \( \bar{\theta} \) is the \( n = 4 \) dimensional vector of means given in Table 4.3. \( \Sigma \) is the...
CHAPTER 4. A MONTE CARLO APPROACH TO SIMULATE THE STOCHASTIC DEMAND IN A CDNLP

Figure 4.4: Nguyen-Dupuis network: Dynamically varying flows over time in two links

Figure 4.5: Nguyen-Dupuis network: arrival times for different paths versus departure time

following variance-covariance matrix

\[
\Sigma := \begin{pmatrix}
10000 & 2500 & 0 & 0 \\
2500 & 10000 & 0 & 0 \\
0 & 0 & 10000 & 2500 \\
0 & 0 & 2500 & 10000
\end{pmatrix}
\]

|\Sigma| is the determinant of \(\Sigma\), and \(\theta^T\) denotes the transpose of \(\theta\).

In order to illustrate the stochastic demand model, we have employed the Monte Carlo simulation method to obtain a random sample of incoming traffic demand in the Nguyen-Dupuis network. This sample of size 100 has
been randomly generated and simulated by the discrete event algorithm given in table 4.1. Figure 4.6 shows the dynamic evolution of the flows in the link $a = 3$. Grey lines are individual simulations and black lines are the mean plus $\pm 1.96$ times the sample standard deviation. Note that the output of the model provides not only mean values (blue line) of predictions but the corresponding variabilities as well. In fact, they provide empirical probability density functions. To illustrate this fact 4 temporal instants are considered and are depicted by the red vertical lines on Figure 4.6. Also a histogram of the data and an adjustment of a normal probability density function is made. The results are shown in Figure 4.7. It is seen to be a good fit for the normal distribution with different means and variances in each instant.

Figure 4.8 shows the travel times in the path 1 corresponding to links (20, 2, 18, 11, 22, 30) for 100 simulations, the blue line represents mean values and the vertical lines indicate temporal instants as in the previous example. The figure shows that travel times are bounded below by the free-flow travel time. Grey lines are individual simulations and the black line is the mean plus 1.96 times the sample standard deviation.

Figure 4.9 shows the variation of the travel times on the time instants (red lines) through the 100 simulations. In this case it is possible to observe a poor fit to a normal distribution. Note that this variability in travel times can be used as input for route and departure time choice models which represent risk-averse behaviour.

![Figure 4.6: Results of Monte Carlo simulation for the flow on link $a = 3$](image)

### 4.4.4 Distributed computing

Finally, a case study focused on the assessment of the distributed computing of the Monte Carlo simulation has been carried out. In this experiment a set of 100 random simulations were executed in a parallel architecture with 12 cores (AMD Dual Opteron 6 Core 4226 2.7 Ghz), comparing the performance...
CHAPTER 4. A MONTE CARLO APPROACH TO SIMULATE THE 
STOCHASTIC DEMAND IN A CDNLP

Figure 4.7: Empirical distribution of the flow on link $a = 3$ for 4 different instants

Figure 4.8: Results of Monte Carlo simulation for travel times on path $p = 1$

of the non-parallel version with respect to the executions using from 2 to 12 cores. Figure 4.10 depicts the speedup ratio (cpu time with 1 core / cpu time with $n$ cores) of each simulation for each network. It can be seen that the two graphs have similar growth until they use 6 or more cores. Therefore the maximum speedup ratio reached is approximately 7 for the Sioux Falls network and 5 for the Nguyen-Dupuis network. The difference between the behaviour of the two graphs is related to the amount of access to memory required for each network. This example highlights the fact that the memory is the bottleneck of the procedure stating that for some networks other distributed computing techniques such as grid computing could be more convenient.
Figure 4.9: Empirical distribution of the travel times on path $p = 1$ for 4 different instants

Figure 4.10: Performance analysis of parallel computing using from 2 to 12 cores
Time-series data mining to enhance traffic signal optimization
5

Time series data mining to enhance traffic signal optimization

5.1 Introduction

Applied optimization and data mining are powerful techniques that allow a number of challenges raised in many scientific fields to be tackled (Fu (2011), Chaovalitwongse et al. (2011), Chaovalitwongse et al. (2012), Wang et al. (2011)). This work addresses how these techniques allow congestion in large cities to be reduced.

Real-world travel demands are intrinsically dynamic which makes for significant variability at the congestion level even for the same time of day and day of the week. These dynamic variations in traffic demand show a recurrent behaviour which imposes a pattern over the day although other factors can be considered completely random and are produced by unpredictable phenomena such as accidents, bad weather, etc.

The study of the dynamic aspects is essential for a proper modelling of the traffic congestion phenomena. As was seen in Chapter 2 Intelligent Transportation Systems (ITS) address this level of uncertainty through advanced monitoring systems of the traffic network in real time, which makes it possible to determine the system state and to respond to unexpected situations. The implementation of these systems (traffic-responsive) requires significant economic investment and complex maintenance processes, which means that at the moment, the number of urban areas equipped with these systems remains small. Many studies and methods for traffic planning and control in urban networks are based on the assumption of recurrent traffic which determines dynamic congestion patterns. The planner establishes time-of-day intervals (TODs) in which it is possible to consider stationary traffic and then, off-line, to plan the best strategy for the current TOD. Typically 3 – 5 plans are run in a given day.
With adequate design and with regular updates these pre-timed systems obtain acceptable results (see Koonce et al. (2008)) when compared with Intelligent Transportation Systems. For this reason it is crucial to have specialized tools capable of automating the planning process and making planning changes when new mobility patterns are detected. In addition robust optimization techniques allow robust pre-timed systems to be obtained that are less sensitive to fluctuations in traffic flow (see Yin (2008)).

Even in areas with traffic control and management centres where a huge amount of dynamic data about the network status is generated, it is imperative to have data-mining tools to determine optimal steady-state freeway conditions for planning and controlling the traffic network under normal conditions.

Traffic signal timing plans are an essential control strategy for traffic networks, particularly in major cities (see Papageorgiou et al. (2003b)). Signal control can be categorized as pre-timed plans, actuated signal plans (off-line approaches) and adaptive methods (on-line approaches). The first group, which deals with traffic demand changes during the day, considers TOD intervals which assume a stationary traffic demand, and a timing plan for each TOD is selected. In the last group signal times are determined by using real-time traffic data (see Angulo et al. (2011)). Off-line techniques involve two broad steps: a) identification of the appropriate TOD intervals in which traffic flows remain constant and b) an optimization approach to carry out phasing-timing plans for each TOD interval.

In practice TOD determination is a task that has been performed visually by traffic engineers making it challenging and subjective. The literature contains methods focused on the identification of changes in traffic over time. Cluster analyses have been used to determine TOD intervals automatically. These methods, such as the $k$-means algorithm (Wang et al. (2005), Ratrout (2011)) or hierarchical methods (Smith et al. (2001)) determine the breakpoints by minimizing within-cluster distances and maximizing between-cluster distances. Thus, the clusters do not directly reflect the performance of the timing plans. Some studies have attempted to solve this problem. Wong and Woon (2008) employ an interactive approach for traffic signal optimization. The first step uses the $k$-means algorithm to identify the TOD intervals and thus optimize the signal timing plans and subsequently, by applying the new traffic controls, the new flows are recalculated to iterate the procedure. Park et al. (2004) propose a bilevel approach: the upper level determines the TOD breakpoints and lower level problem optimizes the timing plans of the corresponding time intervals. A GA-based algorithm obtains breakpoints in a relatively small number of iterations. Lee et al. (2011) address the determination of the TOD problem as a coordinated actuated traffic signal control system. This method explicitly considers transition costs by including a microscopic simulation to assess the transition costs and genetic algorithms.

The studies of Park et al. (2004), Park and Lee (2008) and Lee et al. (2011)
show the bilevel structure for the problem (see Migdalas et al. (1998)). In this chapter the problem is formalized through a bilevel model based on López-Garcia et al. (2014). The model is formed to simultaneously determine the TOD’s and the traffic control strategy. In the upper level the TOD breakpoint is determined by optimization and the lower level problem is represented by a traffic responsive control problem. The integration of both levels allows local optimal in the TOD breakpoints to be avoided. This chapter focuses on a control strategy based on timing plans for intersections but can be easily applied to methods using the study of the whole traffic network.

This model cannot be addressed with the generic segmentation schemes developed in the literature (see Fu (2011), Warren Liao (2005)) because they are designed for one-level schemes. For this reason a metaheuristic methodology has been applied to the bilevel model. In this chapter the hybridization of particle swarm optimization and the Nelder-Mead method, and the hybridization of simulated annealing with Nelder-Mead have been analysed in comparison with a genetic approach.

An aspect which has not been previously studied in the literature as far as we know is automatic determination of the number of TOD’s. This has been considered in this chapter through the Bayesian Information Criterion (BIC), the PETE algorithm and an oriented-problem approach.

The chapter is organized as follows. In section 5.2 the bilevel model is formulated. In section 5.3 the metaheuristic methodology for resolution is described. The automatic determination of the number of TOD’s is carried out in section 5.4. In section 5.5 the numerical experiments over real and synthetic data are studied and finally the conclusions obtained are analysed.

5.2 Mathematical formulation

5.2.1 Model to identify time-of-day breakpoints

Urban traffic networks are mathematically modelled by a directed graph $G = (N,A)$ in which the set of nodes $N$ represents the intersections and the so called centroids. The centroids are dummy nodes which model city areas with generation/attraction of trips. The set of links $A$ represent urban roads and the so-called connectors, which are dummy links joining the centroid nodes with the intersection nodes.

Figure 5.1 shows a representation of a traffic network. This network consists of 4 centroids, 4 connectors and 11 links representing the streets of the urban area modelled. In addition to the network (supply) in these systems the origin-destination matrix is considered representing the demand between different centroids. In this example we consider four origin-destination pairs $A \rightarrow C, A \rightarrow D, B \rightarrow C, B \rightarrow D$. The index $\omega$ denotes one
of these pairs and \( W \) the set of all origin-destination pairs. To model the variation of demand over time the intensity of demand \( q_\omega(t) \) is introduced. These functions are not directly observable in the network.

![Traffic Network Diagram](image)

**Figure 5.1:** Modelling of a traffic network

New technologies allow the monitoring of traffic networks in real-time. These traffic control systems are located in a subset of links of the network, denoted by \( \hat{A} \subset A \) and we will call this the set of sensors. The traffic parameters in the link \( a \in \hat{A} \) are the traffic flow \( q_a \) (veh./hour), the speed \( v_a \) (km./hour) and the density (or alternatively occupancy) \( k_a \) (veh./km.), which is linked to the two above by the equation:

\[
q_a = v_a k_a, \quad a \in A.
\]  

The main objective of this section is to state the optimization model for TOD’s based on the time-domain-constrained data-clustering problem. The goal is to define time intervals in which traffic demand is approximately stationary and therefore the dynamic component within each interval can be considered negligible. This is a resolution strategy for addressing the management and the control of traffic for non-stationary demand (dynamics).

Suppose we have a data set, \( \{(t_j, q_j)\}_{j=1}^N \) where \( q_j \in \mathbb{R}^d \) is a vector of link flows. Indeed, data \( q_j \) are labelled with the time instant \( t_j \) for which the data has been obtained. We assume the data are time ordered, this is, \( t_j < t_{j+1} \) where \( j = 1, \ldots, N - 1 \). We also assume that data pairs are drawn from the following regression model

\[
q_j = F(t_j; \Theta_k) + e_j; \quad j = 1, 2, \cdots, N
\]  

where \( e_j \) represents the error term modelled through a random variable with mean zero and \( \Theta_k \) denotes the vector of (unknown) parameters of the model.
5.2. Mathematical formulation

that works for a specific period of time \( k \). In this chapter we address the TOD problem via cluster analysis taking into account the time instant at which the different observations have been taken. The goal is to find \( K \) disjoint time segments (clusters) so the goodness-of-fit of regression model (5.2) to the observations in each segment could be the best one.

Suppose \( T = [a, b] \) is the time horizon to be partitioned. Consider the boundary points \( a = s_0 < s_1 < s_2 < \cdots < s_{K-1} < s_K = b \) (see Figure 5.2) as decision variables. Then we define the set \( C_k(s) = \{ j \in \{1, \cdots, N \} / s_{k-1} \leq t_j < s_k \} ; \ k = 1, \cdots, K \)

![Figure 5.2: An illustrative example of variable \( s \)](image)

The objective function of the TOD problem can be stated in terms of \( s \) as

\[
\tilde{J}(s) := \text{Minimize} \sum_k \sum_{j \in C_k(s)} L(q_j, t_j; \Theta_k) = \sum_k \tilde{J}_k(s) \tag{5.3}
\]

where \( \tilde{J}_k(s) \) is given by the following statistical estimation problem of the unknown parameter \( \Theta_k \)

\[
\tilde{J}_k(s) := \text{Minimize} \sum_{j \in C_k(s)} L(q_j, t_j; \Theta_k) \tag{5.4}
\]

where \( L \) is a loss function, \( \Theta = (\Theta_1, \Theta_2, \cdots, \Theta_K) \in \mathbb{R}^{Q \times K} \) is the matrix formed by all the parameter vectors of the regression models. The usual loss function \( L \) is the sum of squared errors which is given by

\[
L(q_j, t_j; \Theta_k) = \|q_j - F(t_j; \Theta_k)\|^2 \tag{5.5}
\]

where \( \| \cdot \| \) is the Euclidean norm. Observe that \( C_k(s) \) can be the empty set or have an insufficient number of observations to make the estimation problem
(5.4) well posed. In these cases we consider that $\tilde{J}(s) = +\infty$. The problem (5.4) depends on the application that is being addressed.

Finally, the TOD problem can be written as:

$$\begin{align*}
\text{minimize} & \quad \tilde{J}(s) \\
\text{subject to} & \quad s_{k-1} < s_k; \quad k = 1, \cdots, K
\end{align*}$$

(5.6)

A bilevel optimization problem (5.6) involves two optimization levels (upper and lower level). Equation (5.4) defines the lower level problem. At the upper level the clustering is decided and the lower level adjusts all statistical models for the upper level variable $s$. The vector $s = (s_1, \cdots, s_{K-1}) \in \mathbb{R}^{K-1}$ is the upper level variables of the problem, since when it is fixed, the optimization problem (5.6) is separable in $K$ independent parameter setting problems (the lower level problem).

The objective is to reformulate the previous model as an unrestricted optimization model. Now if we define

$$\hat{J}(s) := \begin{cases} 
\hat{J}(\text{sort}(s)) & s \in [a, b]^{K-1} \\
+\infty & s \notin [a, b]^{K-1}
\end{cases}$$

(5.7)

where \text{sort}(s) orders the components of $s \in \mathbb{R}^{K-1}$ from low to high and $[a, b]^{K-1}$ is $[a, b] \times \cdots \times [a, b]$ hypercube in $\mathbb{R}^{K-1}$, then (5.6) can be reformulated as an unconstrained optimization problem as follows:

$$\begin{align*}
\text{minimize} & \quad \hat{J}(s) \\
\text{subject to} & \quad s \in \mathbb{R}^{K-1}
\end{align*}$$

(5.8)

We will refer to problem (5.8) as the TOD problem.

The basic problem which will be covered is to assume time intervals where the traffic has a stationary behaviour, this is, constant flows within each time interval, and thus the regression function to be considered is

$$F(t_j, \Theta_k) = \Theta_k, \text{ with } j \in C_k$$

(5.9)

In this case $\Theta_k$ is the flow vector in the $k$ period. In this case the lower level problems can be solved explicitly by fixing the parameter vector $\Theta_k$ as the mean flow in the time interval $k$, i.e.:

$$q_k(s) = \frac{1}{|C_k(s)|} \sum_{j \in C_k(s')} q_j$$

(5.10)
5.2. Mathematical formulation

where $|\cdot|$ is the cardinal of a set and $s' = \text{sort}(s)$. The optimization model is stated as:

$$\text{Minimize} \sum_{k=1}^{K} \sum_{j \in C_k} \sum_{i} \| q_{ij} - q_k(s') \|^2$$  \hspace{1cm} (5.11)

5.2.2 A combined approach of TOD determination and optimal traffic signal timing

Note that the formulation is general enough in order to tackle the situation in which $\Theta_k$ could represent the optimal timing plan for the TOD $k$, and the loss function $L(q_j, t_j; \Theta_k)$ for period $j$ is the travel time in the network in that period $j$. In that case the lower level problem cannot be explicitly solved and the problem must have a bilevel structure. This type of problem is difficult to solve and requires metaheuristic optimization algorithms (see Wang et al. (2014)). Without loss of generality, we consider an isolated fixed-time signalled intersection at TOD $k$ (see Figure 5.7). In this problem the control variables are transformed into $\Theta_k := (C_k, g_k)$ where $C_k$ is the cycle length (seconds) and $g_k = (\cdots, g_{ki}, \cdots)$ denotes the vector of effective green time for each line group $i$ at TOD $k$.

We use the delay equation in the Highway Capacity Manual HCM (2010) to estimate the delay per vehicle for period $j \in C_k$

$$d(q_j, t_j; C_k, g_k) := \sum_{i=1}^{n} \left[ \frac{c_i^k (1 - \lambda_i^k)^2 q_{ji}}{x_i^k s_i} + 900 T_j q_{ij} \left( x_i^k - 1 + \sqrt{(x_i^k - 1)^2 + \frac{4x_i^k}{c_i^k T_j}} \right) \right]$$  \hspace{1cm} (5.12)

where

- $n$: is the number of lane groups.
- $\lambda_i^k$: is the effective green split per lane group $i$ at TOD $k$, i.e. $\lambda_i^k = \frac{q_{ij}}{x_i^k}$. 
- $s_i$: is the saturation flow for lane group $i$ (veh/h).
- $x_i^k$: represents the degree of saturation in line $i$ at TOD $k$, i.e. $x_i^k = \frac{q_{ij}}{\lambda_i^k s_i}$. 
- $T_j$: is the duration of the analysis period, i.e $T_j = t_j - t_{j-1}$ and $t_0$ the starting time.
- $c_i^k$: is the capacity for lane group $i$ (veh/h), i.e $c_i^k = \lambda_i^k s_i$.

The total delay time at period $j$ can be computed as

$$L(q_j, t_j; C_k, g_k) := d(q_j, t_j; C_k, g_k) T_j \left[ \sum_{i=1}^{n} q_{ji} \right]$$  \hspace{1cm} (5.13)
Therefore, the optimization model of signal timing can be written as:

$$\hat{J}_k(s) := \text{Minimize} \sum_{j \in C_k(s)} L(q_j, t_j; c^k, g_k)$$

subject to linear constraints on $c^k$ and $g_k$

5.3 Hybrid meta-heuristic algorithm for the TOD problem

Espinosa-Aranda et al. (2013) propose a meta-heuristic framework to incorporate a local search in promising regions into a population-based algorithm.

These methods are probabilistic and maintain a population of candidates. The population-based algorithm avoids being trapped at a local optimum. Moreover, to achieve faster convergence, the hybrid method applies a local optimization method starting from the best point in the population set. This hybrid algorithm provides a trade-off between accuracy and computational cost. The hybrid meta-heuristic algorithm is shown in Table 5.1. This method has two essential parameters $n_c$ and $n_r$. The parameter $n_c$ (exploration phase) guarantees that on the number of successful iterations achieved by the global optimization algorithm is obtained, a local exploitation method (exploitation phase) is conducted starting from the best found solution. The parameter $n_c$ controls the escape from a given neighbourhood.

The method employed in the exploitation phase is the Nelder-Mead simplex (NM). This algorithm was introduced in Nelder and Mead (1965) applied to unconstrained optimization problems. The main advantage of this method is that it frequently gives significant improvements in the first few iterations and quickly produces quite satisfactory results.


5.3.1 A simulated annealing (SA) for TOD problem

Simulated annealing is a popular global search meta-heuristic. The key feature of simulated annealing is that it provides a means to escape local optima by allowing hill-climbing moves in the hope of finding a global optimum.
5.3. Hybrid meta-heuristic algorithm for the TOD problem

<table>
<thead>
<tr>
<th>Tabla 5.1: Hybrid meta-heuristic algorithm for the TOD problem</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1.</strong> <em>(Initialization).</em> Initialize the number of iterations <em>(N)</em>, the global optimizer parameters and randomly generate an initial population of solutions. Initialize the number of iterations <em>(n_c, n_r)</em> associated with the global optimizer and NM respectively. Set the counters <em>(t = 1)</em> and <em>(n = 0)</em> and let <em>(J_aux)</em> = (+\infty).</td>
</tr>
<tr>
<td><strong>Step 2.</strong> <em>(Exploration stage).</em> Apply one iteration of global optimization algorithm to the current population. Let <em>(s)</em> be the current solution and <em>(J)</em> its objective value. If <em>(J)</em> (&gt; J_aux), then let <em>n = n + 1</em> and <em>(J_aux) = J</em>.</td>
</tr>
<tr>
<td><strong>Step 3</strong> <em>(Exploitation stage).</em> If <em>n = n_c</em> apply <em>n_r</em> iterations of NM algorithm by initializing the method from <em>(s)</em>. Let <em>(s_NM)</em> be the solution found, then replace the best solution of the population by <em>(s) = s_NM</em> and take <em>(n = 0)</em>.</td>
</tr>
</tbody>
</table>
| **Step 4** *(Stopping criterion).* If the current iteration is *
* \(= N\), Stop; otherwise set *t = t + 1* and go to Step 2. |
| Output: The best segmentation obtained of *(a, b)* defined by *(s_1^Y, s_2^Y, ..., s_{K-1}^Y)* and \(J^*\). |

SA starts with an initial solution *(s)*. At each iteration of a SA algorithm the current solution *(s)* and a newly selected solution *(s')* are compared. The new solution is generated (either randomly or using some pre-specified rule) in a neighbourhood *(N(s))* of the current solution *(s)*. |

A key problem-specific choice concerns the neighbourhood function definition. The efficiency of simulated annealing is highly influenced by the neighbourhood function used. In this work the following problem-specific neighbourhood is proposed,

\[
N(s) := \{ s' \in [a,b]^{K-1} / s_k' = s_k; \text{ for all } k \neq k' \text{ with } k' \in \{1, \cdots, K - 1\} \}\tag{5.15}
\]

The generation probability function has been chosen as a uniform distribution with probabilities proportional to the size of the neighbourhood *(N(s))* of the current solution *(s)*. To generate the new candidate *(s' \in N(s))* we carried out three steps:

\[
s' = s
\]

Choose a random number \(k' \in \{1, \cdots, K - 1\}\)

\[
s_{k'}' = a + Rand() \cdot (b - a).
\]

Where \(Rand()\) is a random uniform number on \([0, 1]\). |

The candidate solution *(s')* is accepted based on the rule

\[
P(\text{Accept } s' \text{ as next solution}) = \begin{cases} 
\exp\left[\frac{-(\hat{J}(s') - \hat{J}(s))}{T_n}\right] & \text{If } \hat{J}(s') - \hat{J}(s) > 0 \\
1 & \text{If } \hat{J}(s') - \hat{J}(s) \leq 0
\end{cases}
\]

Improving solutions are always accepted, while a fraction of non-improving solutions are accepted in the hope of escaping local optima.
in search of global optima. The probability of accepting non-improving solutions depends on a temperature parameter $T_n$. The resulting SA is called SA$^*$ and it is outlined in Table 5.2.

### Tabla 5.2: SA$^*$ algorithm for TOD problem

1. **(Initialization).** Initialize the number of iterations ($N$). Select an initial solution $s$, a temperature cooling schedule ($T_n$), and an initial temperature $T = T_0$. Select a repetition schedule, ($M_n$), that defines the number of iterations executed at each temperature. Set the temperature change counter $n = 0$ and repetition counter $m = 0$.

2. **Step 1.** Generate a solution $s' \in N(s)$ using equation (5.16)

3. **Step 2.** Calculate $\Delta = \tilde{J}(s') - \tilde{J}(s)$. If $\Delta \leq 0$ then $s = s'$, $\tilde{J}^* = \tilde{J}(s')$ and $s^* = s'$. Otherwise $\Delta > 0$, set $s = s'$ with probability $\exp[-\Delta/T_n]$. Take $m = m + 1$.

4. **Step 3.** If $m = M_n$, then $n = n + 1$ and $m = 0$.

5. **Step 4.** (Stopping criterion). If the current number of iterations is $t = N$, Stop; otherwise let $t = t + 1$ and go to Step 1.

Output: The best segmentation of $[a, b]$ defined by $s^*$ and its objective value $\tilde{J}^*$.

### 5.4 Determination of number of TODs

The clustering algorithm assumes that the optimum number of clusters $K$ is known. This chapter looks at four methods for determining the optimal number of clusters. The first two are based on the widely used Bayesian Information Criterion and employ all the registered time series, while the third algorithm is based on Vasko and Toivonen (2002) and the fourth is oriented to the problem. Those last two methods operate with the mean observed values.

#### 5.4.1 Bayesian Information Criterion (BIC)

The BIC is a likelihood criterion for model comparison that penalizes models with additional parameters. The BIC is defined mathematically as:

$$ BIC(K) = -2 \log L_K(q_1, \ldots, q_n) + \lambda m_K \log(n) $$

where $\{q_1, \ldots, q_n\}$ is the complete data to be modelled. The first term term represents the maximum log likelihood of the data under the model with $m_K$ parameters. The second term $\lambda m_K \log(n)$ is responsible for penalizing the candidate models according to their number of parameters $m_K$ and $\lambda$ is the penalty weight ($\lambda = 1$ according to the BIC theory). The optimum model corresponds to the one for which the value of BIC, given by equation (5.17) is the least. If we assume that the observations in a TOD are drawn from
5.4. Determination of number of TODs

A full-covariance Gaussians \( \{q_j\}_{j \in C_k} \sim N(\mu_k, \Sigma_k) \), the BIC for the \( K \)-TODs solution is defined

\[
BIC(K) = \sum_{k=1}^{K} n_k \log(|\Sigma_k|) + K \left( d + \frac{d(d+1)}{2} \right) \log(n)
\]  (5.18)

in which \( n_k \) denotes the number of articles in cluster \( C_k \) and \( d \) is the dimension of the flow vector space.

A two-stage process was presented by Chiu et al. (2001) for determining the optimum number of clusters. The procedure is described below, following the work of Xia and Chen (2007).

The two-stage process first examines the BIC for all potential clustering solutions. The goal is to find the smallest number of clusters that have the lowest BIC, because the BIC decreases first and then increases as the number of clusters increases, the BIC for each \( K \) (clustering solution) is calculated. Beginning from \( K = 1 \), the first \( \hat{K} \) value that satisfies \( BIC(\hat{K}) < BIC(\hat{K}+1) \) is chosen as a coarse estimate of the number of clusters. In the second stage, the ratio of changes in dispersion measurement is used to determine the optimum number of clusters based on the coarse estimate obtained in the first stage. The ratio of changes in dispersion measurement is defined as \( R(K) = s_{K-1} / s_K \) for \( K = 2, \ldots, \hat{K} \), in which \( s_{K-1} \) denotes the change in dispersion measurement if \( K \) clusters are merged into \( K-1 \) clusters.

The parameter \( s_K \) can be computed as \( s_K = l_K - l_{K+1} \), in which \( l_K = \sum_{k=1}^{K} n_k \log(|\Sigma_k|) \). This second stage is based on the understanding that a significant increase in \( R(K) \) will be observed when two clusters that should not be merged are merged. The \( R(K) \) value for each \( K(=2, \ldots, \hat{K}) \) is calculated, and the two largest \( R(K) \) values are identified as \( K = K_1 \) (the largest) and \( K = K_2 \) (the second largest). Xia and Chen (2007) uses an empirical threshold value of \( R(K_1)/R(K_2) = 1.15 \); that is, if \( R(K_1)/R(K_2) > 1.15 \), \( K \) is set to \( K_1 \); otherwise, \( K \) is set to \( \max(K_1, K_2) \).

5.4.2 \( \Delta BIC \)

In segmenting an audio stream the BIC has widely used (Chen and Gopalakrishnan (1998), Wang et al. (2008)). It can be shown Chen and Gopalakrishnan (1998) that if the expression

\[
\Delta BIC(s_k(K)) = (n_k + n_{k+1}) \log(|\Sigma_k \cup \Sigma_{k+1}|) - n_k \log(|\Sigma_k|) - n_{k+1} \log(|\Sigma_{k+1}|) - \left( d + \frac{d(d+1)}{2} \right) \log(n_k + n_{k+1}) \text{ with } k = 1, \ldots, K-1.
\]  (5.19)
is positive, the time $s_k$ is a good candidate for a segment boundary. Note that $\Sigma_k \cup \Sigma_{k+1}$ represents the variance-covariance matrix of the observations $(q_j)_{j \in C_k(s) \cup C_{k+1}(s)}$. It is possible to apply this criterion to the $K$-solution to determine if its border points $s_k$ are significant. The criterion used is to choose the solution with significant points of maximum cardinality $K$.

### 5.4.3 A modified PETE algorithm

Vasko and Toivonen (2002) present the so-called PETE algorithm to determine the number of time segments. This method generates a $p$-value for each increase in the number of segments. In this chapter a modification of this method is adapted in order to reduce its computational cost. Let $e(s^K)$ be the segmentation error of the solution $s^K$. By using a Monte Carlo simulation the random error of adding a new segment in the partition is calculated as follows: A random segment $k \in \{1, \ldots, K\}$ is selected, then the observations $C_k(s^K)$ are randomly ordered and the segment is randomly partitioned. The new error denoted by $e_j(s^K)$ is calculated. Drawing from the random sample $\{e_j(s^K)\}$ the $p$-quantile $e_p$ of the random error is calculated, if it satisfies

$$e(s^{K+1}) < e_p \tag{5.20}$$

then the $K+1$-solution is selected and the procedure is repeated.

### 5.4.4 An approach oriented to the TOD problem

The desired objective is a suitable model of the dynamic mechanisms in traffic parameters. In this context the selection of a high number of TOD’s has as a consequence a greater analytic cost but gives more satisfactory results. For this reason a natural method is to require a number of clusters $K$ so that the average relative mean error $\bar{e}$ in describing the time series as a set of stationary segments does not exceed a value $\tilde{e}$ by the planner, that is

$$K^* = \text{Arg Minimize}_K \left\{ K : \bar{e}(s^K) < \tilde{e} \right\}. \tag{5.21}$$

This value $\bar{e}$ has a physical interpretation and thus allows a priori determination.

### 5.5 Computational experiments

The objectives of these computational tests are:
5.5. Computational experiments

1. To analyse the performance of the proposed metaheuristic algorithms applied to the TOD problem. This goal has been analysed in Experiment 1.

2. To evaluate the proposed methodology in a real case. The purpose is to compare a sequential methodology in which, primarily the number of TOD’s is determined and later the signal control for each TOD is planned, with the intention of addressing both problems simultaneously. The numerical results are collected in Experiment 2.

3. The objective in Experiment 3 is to analyse the above four indices to determine the optimal number of TOD’s over real data.

**Experiment 1: A comparison of the performance of different metaheuristics**

The first data set was generated through a simulation experiment, using the dynamic traffic load model developed in Sánchez-Rico et al. (2014); the Nguyen-Dupuis network shown in Figure 5.3 has been used. The network shows the sensor locations to obtain the data.

![Nguyen-Dupuis network](image)

**Figure 5.3:** Nguyen-Dupuis network

The results are shown in Figure 5.4. These results consist of 4 daily traffic patterns, considering traffic flow or density over a set of three sensors located in the network.

The combined model given in Subsection 5.2.2 has a bilevel structure, in which the evaluation of the objective function requires solving $K$
optimization problems. This has motivated the need for efficient resolution algorithms. In this experiment the algorithms standard PSO, SA, SA’ and their hybridizations with the NM have been tested. The selected parameters for the hybridization are $n_c = 1,5$ and $n_r = 100$. Moreover in references Park et al. (2004) and Lee et al. (2011) GA algorithms are introduced as a possibility for solving this problem. GA has been introduced as a baseline for this work.

Eight synthetic data sets (see Figure 5.4) are considered and $K^* = 5$ TOD’s are defined. As the algorithms have a probabilistic nature each instance was run 10 times over the eight test problems. In order to visualize the results obtained, the average change in the 10 runs and for all the test problems is considered, and in addition the value of the objective function has been standardized as

$$Z^* = \frac{\widehat{J}(s) - J_{\text{MIN}}}{J_{\text{MAX}}}$$  \hspace{1cm} (5.22)

where $J_{\text{MAX}}$ and $J_{\text{MIN}}$ are respectively the maximum and minimum found for all the algorithms in a given problem. This standardization means that the 8 test problems are equally important.

The results are shown in Figures 5.5 and 5.6. It is observed that the hybridization speeds up the original algorithm. In addition it can be seen that the $SA^* + NM$ algorithm with $n_c = 1$ outperforms the other algorithms.
5.5. Computational experiments

Figure 5.5: Average performance of the metaheuristic algorithms in density-based tests

Figure 5.6: Average performance of the metaheuristic algorithms in flow-based tests

Experiment 2: A real case study

This example is taken from Yin (2008) and consists of a real-world intersection between 164th Street SW and Alderwood Mall Parkway in the City of Lynwood, Washington. The flows were recorded in March-April 2005, 36 observed flow patterns were retrieved for the PM peaks (4:30-6:30 p.m.) between Tuesday and Thursday. Based on these data two test problems were designed, in the first the 36 patterns are considered as consecutive over...
a day (named Test 1). The second (named Test 2) is obtained by modifying the original data to introduce a dependence on the current time of day, in particular the flows of the \( j \)-th-period have been multiplied by the weighting factor 
\[
f_j = 0.85 + 0.65 \exp(-0.15 \times (j - 12)^2) + 0.65 \exp(-0.15 \times (j - 32)^2)
\]
with \( j = 1, \cdots, 36 \).

The intersection is shown in Figure 5.7 and the saturation flow rates \( s_i \) for groups \( i = 1, \cdots, 8 \) are 1900, 3800, 3800, 1900, 1900, 3800, 3800, 1900. A specific lead-lag phasing sequence is used in the example and the resulting constraints for the traffic-signal timing problem are:

\[
\begin{align*}
  g_1 + g_2 + g_3 + g_4 + L &= c \
  g_1 &= g_6, g_2 = g_5, g_3 = g_8, g_4 = g_7 \
  g_1 &\geq g_{\text{min}}, i = 1, 2, \cdots, 8 \\
  c_{\text{min}} &\leq c \leq c_{\text{max}}
\end{align*}
\]

where \( L \) is the total lost time per cycle, 14s used in the example; \( g_{\text{min}} \) is the minimum green time, 8s used, and \( c_{\text{min}} \) and \( c_{\text{max}} \) are the minimum and maximum cycle length, specified as 50s and 140s, respectively. When using Eq. (5.13) to calculate the total delay, the duration of each time period \( j = 1, \cdots, 36 \) is set as \( T_j = 0.25h \).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure5_7.png}
\caption{A four-leg intersection}
\end{figure}

Over the two test problems a sequential and a simultaneous methodology is applied. The sequential methodology is the one followed in practice, first the TOD’s are determined and later the optimal traffic signal timing is calculated for each TOD. The simultaneous methodology is the one described in this chapter and it directly minimizes total time in the intersection. Both methods require an algorithm to solve the problems (5.14). He and Hou (2012) use ant colony and a genetic algorithm to solve this problem. These algorithms present two disadvantages, the first is its high computational cost due to the large number of solutions and the second is, because the simulated annealing algorithm is used for the bilevel model, conflicts may
appear between the accuracy of resolution of the lower level problem and the objective function value as shown in García and Marín (2002).

This numerical experiment employs an interior-point algorithm (implemented in the MATLAB function \texttt{fmincon}). In our numerical tests, there is evidence that this option is faster and more efficient than the GA algorithm implemented in the MATLAB function \texttt{ga}.

The results obtained are shown in Table 5.3. The first column shows the number of TOD’s considered, the second column the methodology used and the third the time required to obtain the best solution. The third column shows $\tilde{J}_{\text{delay}} = \sum_k \tilde{J}_k(s)$ where $J_k(s)$ is calculated by using Eq. (5.13) and the fourth shows the value $\tilde{J}_{\text{tod}}$ where $J_k(s)$ is calculated by Eq. (5.11). The results obtained agree with expectation. Each methodology achieves better results for the objective function it is trying to minimize. The simultaneous methodology minimizes the total waiting time while the sequential gives TOD’s with lower variability. It is seen, however, that the sequential approach is very efficient and is capable of obtaining similar solutions to the simultaneous method. Both methodologies allow the systems to be recalibrated automatically. One conclusion is this, considering that traffic regulation in different time intervals is advantageous. The reduction in the waiting time by considering $K = 4$ instead $K = 2$ is about 3.1%. To go into this question more deeply, Figure 5.8 shows the solutions obtained by employing the simultaneous methodology in Test 2. It depicts the TOD’s obtained for several values of $K$ and how the optimal traffic signal timing is strongly dependent on the number of TOD’s. It shown that the cycle amplitude $C$ is dependent on the level of congestion of the TOD.

**Experiment 3: Determination of the optimal number of TOD’s**

In this section the algorithms described in Section 5.4 to identify the optimal number of TOD’S are tested. In order to assess the methods, real traffic data collected by the California Freeway Performance Measurement System (PeMS) is used. The PeMS collects the traffic data in real time from over 25,000 individual detectors across all major metropolitan areas of the State of California. In order to test the methodology proposed in this article, observations over 100 days in 2013 have been collected from two dual loop detector stations in the Bay Area of California. The data have been previously classified into different traffic patterns on different days. The traffic profiles based on speed and occupancy are shown in Figure 5.9. The data used are available at http://bit.ly/1hsTEjO.

For each of the 8 traffic patterns the TOD’s have been identified for $K = 1, \cdots, 20$, and by using the BIC, $\Delta$BIC, modified PETE and the oriented-problem (OP) methods, the optimal number of TOD’s have been selected. The $p$ value employed in the modified PETE method was $p = 0.05$ and the mean
CHAPTER 5. TIME SERIES DATA MINING TO ENHANCE TRAFFIC SIGNAL OPTIMIZATION

<table>
<thead>
<tr>
<th>Test</th>
<th>K</th>
<th>Approach</th>
<th>CPU time (s)</th>
<th>$\bar{D}_{\text{delay}} (s)$</th>
<th>$\bar{D}_{\text{todd}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>Sequential</td>
<td>0.1</td>
<td>2051660.5</td>
<td>6036.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simultaneous</td>
<td>212.1</td>
<td>2021479.3</td>
<td>6123.2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Sequential</td>
<td>1.7</td>
<td>1992496.1</td>
<td>5500.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simultaneous</td>
<td>170.7</td>
<td>1989842.6</td>
<td>5509.4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Sequential</td>
<td>1.3</td>
<td>1979965.9</td>
<td>5242.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simultaneous</td>
<td>169.0</td>
<td>1977440.8</td>
<td>5451.0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Sequential</td>
<td>0.2</td>
<td>2750230.8</td>
<td>10340.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simultaneous</td>
<td>145.9</td>
<td>2737846.5</td>
<td>10759.0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Sequential</td>
<td>1.4</td>
<td>2708806.7</td>
<td>9540.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simultaneous</td>
<td>152.4</td>
<td>2703892.0</td>
<td>10038.9</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Sequential</td>
<td>1.2</td>
<td>2655422.0</td>
<td>7326.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simultaneous</td>
<td>192.9</td>
<td>2653730.4</td>
<td>7340.1</td>
</tr>
</tbody>
</table>

The observed results indicate that the BIC is the most parsimonious method since it determines a reduced number of TOD’s, the $\Delta$BIC method is highly sensitive and establishes a higher number of TOD’s. On the other hand it seems that the modified PETE algorithm finds an acceptable number of TOD’s. The OP method employs the same threshold for both types of data (occupancy or flow) but the variability level is different for each, which means that for the occupancy a high number of TOD’s is established but this is not the case for the speed.

To gain additional insight into the results the different procedures are displayed. The clustering results using the BIC and the $\Delta$BIC criteria are shown in Figures 5.10 and 5.11. The first column is associated with the occupancy and the second is based on the speed; finally the vertical lines separate the different TOD’s according to their algorithm. As can be observed in the Figures the $\Delta$BIC criterion produces a non-significant number of TOD’s which may not be helpful for traffic control; in addition the BIC algorithm gives an insufficient number of clusters as a result. These methods tackle the existing variability between days.

Unlike the two previous methods, the modified PETE and the OP algorithms work with a mean traffic profile. Figure 5.12 illustrates the way the modified PETE algorithm works. The first graph corresponds to pattern 1 (occupancy) and the optimal number of time segments for this particular
5.5. Computational experiments

Figure 5.8: Solution Test 3

case is 12. It is possible to observe in the graph that the stopping criterion of the algorithm is when the \( p \)-value of the mean error of adding a new random segment is less than the error obtained by increasing one TOD in the current solution.

Figure 5.13 shows the results obtained from the OP method; the first graph shows the occupancy-based tests and the second the speed-based tests. In both graphs the red line represents the required threshold for the mean relative error value that must not be exceeded when selecting the number of clusters. This method has resulted in practice in more consistent and comprehensive partitions, for example the optimal number of partitions for the occupancy problem varies between 10 and 18 while for the speed problem it varies between 2 and 4.

The criterion of fixing a maximum mean relative error is easily interpreted by planners. This makes the method more easy to apply than the previous.
Figure 5.9: Real data set. Average daily occupancy and flow profiles

Table 5.4: Optimal number of clusters by selection criterion and the problem

<table>
<thead>
<tr>
<th>Criterion</th>
<th>BIC</th>
<th>ΔBIC</th>
<th>PETE</th>
<th>OP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pattern 1</td>
<td>5</td>
<td>20</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>Pattern 2</td>
<td>3</td>
<td>20</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>Pattern 3</td>
<td>2</td>
<td>20</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Problem</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pattern 1</td>
<td>8</td>
<td>17</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Pattern 2</td>
<td>5</td>
<td>15</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Pattern 3</td>
<td>8</td>
<td>20</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>Pattern 4</td>
<td>3</td>
<td>9</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Pattern 5</td>
<td>3</td>
<td>16</td>
<td>16</td>
<td>2</td>
</tr>
</tbody>
</table>
5.5. Computational experiments

Figure 5.10: Solution of the TOD problem using the $\Delta$BIC rule

Figure 5.11: Solution of the TOD problems using the BIC rule
CHAPTER 5. TIME SERIES DATA MINING TO ENHANCE TRAFFIC SIGNAL OPTIMIZATION

Figure 5.12: PETE algorithm behaviour

Figure 5.13: OP algorithm behaviour
An approach to dynamical classification of daily traffic patterns
An approach to dynamical classification of daily traffic patterns

6.1 Introduction

The main goals of the so-called Traffic Management and Control Centers (TMCC) are automatic incident detection, real time traffic information provision and the coordination of traffic control and surveillance systems (see Chapter 2). The theoretical basis of most of these Intelligent Transport Systems (ITS) is short-term traffic forecasting. The quantitative aspect of short-term traffic flow predictions has been intensively researched for more than two decades. This problem can be addressed through different methods, the most popular in the past ten years are Bayesian inference, time series models and neural networks but it can be studied also with traffic assignment models whose aim is to predict link flows and travel times on a traffic network as can be seen in Sánchez-Rico et al. (2014) and Codina et al. (2014), these models involve considerable work in order to provide consistency and accuracy to the traffic models (see Castillo et al. (2014)). Critical reviews on the topic of short-term traffic flow predictions are given in Vlahogianni et al. (2014), Adeli (2001)) but also, in the existing literature, not much attention is paid to qualitative predictions. This aspect is relevant to add robustness to the implemented systems and to facilitate the functioning of TMCC, which can serve as a basis for the development of TMCC for both real-time traffic monitoring and off-line traffic-system performance evaluation for decision-making.

Traffic Pattern Recognition (TPR) is a process to identify qualitative congestion levels from real-time traffic information in conjunction with historical traffic records. Traffic presents conditions of variations within a single and day-to-day variations. The work described in the literature for TPR is focused on one of the two types of variations. In the first case the
traffic state can be classified into different traffic operating conditions [such as the Level of Service (LOS)] and in the second case the days are classified into different traffic patterns. These groups are reviewed below:

- **Identification of the traffic network operating conditions.**

  These studies are focused on TPR for limited periods of time and the most widely used techniques are clustering algorithms. Xia and Chen (2004) uses a $k$-means algorithm to categorize freeway operating conditions using continuous traffic data obtained through ITS detectors.

  In this context other methods have been proposed in order to improve the cluster techniques for traffic identification, for example Park (1998) and Park (2002) categorize freeway traffic flow patterns in the short term by using a neural network model and the $k$-means algorithm.

  The work of Xia and Chen (2007) propose a nested cluster technique for analysing freeway operating conditions. This method has the advantage of being able to identify the appropriate number of clusters, a parameter necessary for performing the $k$-means algorithm. Qu et al. (2010) address the network-level traffic status using traffic assignment ratio matrices and self-organizing maps, the key point of this chapter is the use of network-traffic state measurements, rather than a particular highway or corridor. This type of approach is essential when designing vehicle guidance systems. Yang et al. (2012) present a $k$-means fuzzy algorithm with reinforcement learning to estimate the network congestion.

  Yu et al. (2013) classify traffic-condition patterns via a support vector machine using the three-dimensional space defined by traffic volume, speed and occupancy.

  Montazeri-Gh and Fotouhi (2011) study the application of driving-condition recognition to intelligent control of hybrid electric vehicles.

  The on-line applicability of these methods entails a high computational cost due mainly to the training phase. This challenge is addressed in some of the studies in the literature.

  Innamaa (2009) develops a method for classifying traffic flow status in real time that has self-learning capability based on self-organising maps (SOM).

  Yanguo et al. (2011) present a real-time traffic condition identification based on Fuzzy $c$-means clustering and the traffic parameters flow, speed and occupancy. The work of Xia et al. (2012) applies the approach given in Xia and Chen (2007) to an on-line context based on the traffic variables flow, speed and occupancy and in a second step, apply the Bayesian Information Criterion for classifying current state of the traffic network to the resulting clusters.

  Celikoglu (2013) combines neural network theory with a macroscopic
6.1. Introduction

traffic model to figure out freeway traffic flow patterns as a function of density measures obtained by multiple sensors. The work of Zhang and Ge (2013) develops a neural network approach to predict freeway corridor travel times with an on-line algorithm that not only has the capabilities of neural networks but also takes into account the human-like thinking and reasoning of fuzzy logic systems.

- **Classification/prediction of Daily Traffic Demand Patterns (CTDP)**

These studies have addressed the construction of classification and prediction systems for daily traffic demand. The main aspect of these proposals is that they classify a full day rather than short time periods in traffic states (or clusters). They assume the existence of recurrent congestion and use a two stage approach. In the first phase the typical daily traffic profile is obtained and in the second phase a short-term traffic prediction is performed based on clusters which define the baseline prediction (see Thomas et al. (2010) and Chiou (2012)).

Most of the studies implement a rough classification by weekdays and weekends but more sophisticated studies have emerged. There are studies addressing this topic from different perspectives. A coarse taxonomy of these studies can be performed by analysing the traffic network modelling, the used methodology and the followed goals. The traffic network is defined by the kind of area (freeway, arterial or traffic network); the traffic parameters collected [volume, speed, and density (or occupancy)] and the number of their locations. The studies seek to anticipate the traffic conditions in the long term and for this purpose “prediction-after-classification” approaches are developed. Some studies are focused on clustering traffic patterns and others also address the prediction phase. Another essential feature is the focus on the level of prediction (prediction link by link, prediction on specific paths or prediction of the global state of the network). Table 6.1 shows the relevant literature on CTDP.
### Table 6.1: Literature for CTDP

<table>
<thead>
<tr>
<th>Reference</th>
<th>Area&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Traffic param.</th>
<th>Data collection</th>
<th>Approach</th>
<th>Goal&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stutz and Runkler (2002)</td>
<td>UA</td>
<td>V</td>
<td>SD</td>
<td>Partially supervised clustering with cluster merging based on the Fuzzy c-means algorithm</td>
<td>C+P</td>
</tr>
<tr>
<td>Weijermars and van Berkum (2005)</td>
<td>UM</td>
<td>V</td>
<td>MD</td>
<td>Hierarchical clustering (Ward's method)</td>
<td>C</td>
</tr>
<tr>
<td>Thomas et al. (2010)</td>
<td>UM</td>
<td>V</td>
<td>MD</td>
<td>Baseline prediction+correlation with consecutive days</td>
<td>C+P</td>
</tr>
<tr>
<td>Chiou (2012)</td>
<td>UM</td>
<td>V</td>
<td>SD</td>
<td>Functional mixture prediction+functional classification</td>
<td>C+P</td>
</tr>
<tr>
<td>Zhang et al. (2012)</td>
<td>UM</td>
<td>S</td>
<td>MD</td>
<td>Dynamical clustering of spatial correlations+on-line traffic prediction based on neural network for each cluster 7 daily profiles for each day +regime-switching modelling framework</td>
<td>C+P</td>
</tr>
<tr>
<td>Kamarianakis et al. (2012)</td>
<td>UM</td>
<td>A</td>
<td>MD</td>
<td>Daily traffic profiles with functional data analysis+multivariate control chart</td>
<td>C+P</td>
</tr>
<tr>
<td>Guardiola et al. (2014)</td>
<td>UA</td>
<td>V</td>
<td>MD</td>
<td>Parameter estimation for day to day traffic models by Bayesian approaches+Markov Monte Carlo</td>
<td>C+P</td>
</tr>
<tr>
<td>Parry and Hazelton (2013)</td>
<td>UM</td>
<td>V</td>
<td>MD</td>
<td></td>
<td>C+P</td>
</tr>
</tbody>
</table>

<sup>a</sup> Urban motorway (UM); Urban arterial (UA)

<sup>b</sup> Volume (V); Speed (S); Density (D); Occupancy; (A) a general traffic parameter

<sup>c</sup> Multiple detectors (MD); Single detector (SD)

<sup>d</sup> Long-term traffic classification and prediction system (C+P); Daily traffic classification (C)
6.2. Methodology

The main contributions of this chapter are:

- This chapter describes a prototype of urban traffic control system based on clustering/classification techniques. The goal of this system is to control and manage the recurrent traffic profiles (without incidents). The system consists of an off-line and an on-line stage. The off-line phase allows knowledge of what typical daily network-traffic demand patterns can be distinguished by using historical data and cluster analysis. A repository of control and management strategies can be computed on the basis of the resultant clusters. The on-line approach to traffic control can predict the current traffic-network status and, based on the repository obtained in the off-line phase, implement the proper control strategy.

- As a by-product of the proposed methodology the CTDP has been used to identify daily traffic patterns based on density (occupancy), flows and speed presented for one day and their outputs are a basis for defining traffic management scenarios which are the input for macroscopic traffic models (García-Ródenas and Verastegui-Ray (2013)).

- To our knowledge, a dynamic multivariate $k$-means algorithm based on spectral distances has not yet been analysed in the literature. This scheme takes into account the multivariate nature of multiple counts on the network. The studies presented in the literature are applicable to univariate data (see Jacques and Preda (2013)).

This chapter is organized as follows: Section 6.2 sets out the methodology, in Section 6.3 two tests were conducted in order to evaluate the consistency of the proposed methodology, the first with synthetic data obtained by the simulation of a dynamic traffic loading model (see Sánchez-Rico et al. (2014)) and the second with real data collected on highways in the United States.

6.2 Methodology

6.2.1 A prototype of urban traffic control system

Figure 6.1 shows a prototype of urban traffic control system which exploits the traffic data collected by Traffic Management and Control Centers. The essential ideas of this approach were suggested in Angulo et al. (2011). This scheme has an off-line phase and an on-line phase. In the off-line stage the most demanding computational calculations are addressed. In this stage it is assumed that some daily traffic patterns are recurrently presented in the network and for this reason the set of patterns can be identified from the historical data and subsequently the traffic control strategy for each pattern is optimized (e.g signal timings or/and optimal route
recommends) taking into account the time-dependent characteristics of both traffic (demand) and the infrastructure (supply).

FIGURE 6.1: A prototype of urban traffic control system

In the on-line stage, the traffic control strategies obtained are implemented in the real traffic network. This system requires a module capable of detecting incidents. If the current state is an incident then an answer oriented to the disruption is given, in another case the appropriate pattern from the repository is estimated and its associated optimal control strategy is chosen for the current time of day. The module of the incident detection system has already been set out in the literature, for example Ghosh-Dastidar and Adeli (2003) proposes a freeway incident-detection model using a combination of wavelet de-noising, statistical cluster analysis and neural network pattern recognition, also Vlahogianni and Karlaftis (2013) presents an approach for predicting incident durations by using a fuzzy entropy methodology and neural network models.

The advantages of this architecture are two: i) the system is both fast and scalable to full urban networks because the most intensive computational calculations (such as system optimization) are addressed in the off-line phase and ii) robustness. since the system only implements (automatically) a finite set of control strategies (defined in the repository), these strategies are previously evaluated by the planner.

This chapter is focused in the second stage of the off-line and on-line phases. This is, in the recognition of the daily traffic patterns and its prediction. The third stage of the off-line phase is beyond the scope of this chapter but some available methods can be seen in the reviews of
6.2. Methodology

Papageorgiou et al. (2003b), Cascetta et al. (2006) among others.

6.2.2 A clustering method to identify daily traffic demand patterns

The purpose of this section is to determine the patterns of demand that prevail across days in a traffic network and the available data to consider this problem.

Let us continue considering the traffic network shown in Figure 5.1 and previously described in Chapter 5, which consists of 4 centroids, 4 connectors and 11 links representing the streets of the urban area modelled. To model the variation of demand over time the intensity of demand $q_\omega(t)$ is introduced. This function indicates the number of vehicles per hour that want to make the trip $\omega$ at instant $t$.

A daily demand for the road network can be represented by the function

$$Q : [0, 24] \subseteq \mathbb{R} \rightarrow \mathbb{R}^{|W|}$$

$$t \mapsto Q(t) = (\cdots, q_\omega(t), \cdots)$$

It is assumed that the traffic network operates under a recurrent congestion which makes the number of observable situations part of a finite set $\{1, \cdots, K\}$ of states.

The purpose of the CTDP is to discover these $K$-demand patterns which can operate in the traffic network analysed.

New technologies allow the monitoring of traffic networks in real-time. The key difficulty is that the functions $q_\omega(t)$ are not directly observable. The most commonly used data collection scheme is based on sensors (such as loop-detectors). We assume that these point sensors are located in a subset of links of the network, denoted by $\tilde{A} \subset A$ (see Figure 5.1). They are used to collect traffic data averaged over time intervals across all the lines in one direction.

The traffic measurements in the link $a \in \tilde{A}$ are the traffic flow $q_a$ measured in number of vehicles per hour per lane (vphpl), the traffic speed $v_a$ measured in miles per hour (mph) or kilometres per hour (kph) and the density $k_a$ measured in number of vehicles per mile per lane (vpmpl) or per kilometre per lane (vpkpl). In practice, the measure of occupancy can be employed (usually as a proportion of the design maximum value) instead of the density.

The above three traffic parameters satisfy:

$$q_a = v_a k_a, \quad a \in A.$$
The sensors located in the set $\hat{A}$ provide information about the link status (traffic flow, occupancy and speed) at regular interval times $t_1, \ldots, t_m$. The information collected during a given day represented by three matrices of $m \times n$ dimension, where $m$ is the number of rows and coincides with the number of sampled instants and the value $n$ represents the cardinality of the set $\hat{A}$. This information is denoted as

$$X = (X^q, X^k, X^v) \in M^3_{m \times n}(\mathbb{R})$$

where $M^\alpha_{m \times n}(\mathbb{R}) = M_{m \times n}(\mathbb{R}) \times M_{m \times n}(\mathbb{R}) \times M_{m \times n}(\mathbb{R})$ is the space of $m \times n$-matrices of real numbers.

The network topology imposes an underlying structure on the observed data due to the flow conservation restrictions and the satisfaction of demand. Let us formally analyse this question, to simplify the analysis, assume a stationary situation on the system in which the dynamic component has been removed. The satisfaction of demand requires:

$$\sum_{p \in P_\omega} h_k = q_\omega$$

(6.4)

where $h_k$ is the flow in the path $k$ and $q_\omega$ is the origin-destination pair demand intensity $\omega$ and $P_\omega$ is the set of paths that can be employed to meet the travel demand in pair $\omega$.

On the other hand, let $q_a$ be the link flow circulating in a link $a$, then the following must be satisfied

$$\sum_{\omega \in W} \sum_{p \in P_\omega} \delta_{ak} h_k = q_a,$$

(6.5)

where the value of $\delta_{ak}$ is 1 if the path $p$ uses the link $a$ and 0 otherwise. This restriction requires that the flow circulating through a link is the sum of all path flows that use the specific link.

Equation (6.5) shows that there is a correlation between the link flows sharing paths (with delays for the dynamic case). Because of this, to characterize the demand pattern should be appropriate to consider this correlation structure. This has led us to transform the original data as follows:

**Definición 6.2.1** (Operator $\circ$). Let $X = (X^q, X^k, X^v) \in M^3_{m \times n}(\mathbb{R})$ be daily observation data, we consider as input

$$X' \circ X'' = (X'^q X'', X'^k X'', X'^v X'') \in M^3_{n \times n}(\mathbb{R})$$

(6.6)
6.2. Methodology

where

\[ X^\xi X^\xi = \sum_{t=1}^{m} (X_{t*}^\xi)(X_{t*}^\xi) \quad \text{with} \quad \xi \in \{q,k,v\} \]  

(6.7)

with \( X_{t*}^\xi \) is the \( t \)-th row of the matrix \( X^\xi \) and \( \cdot \) denotes the transpose of a matrix or vector. Note that this transformation eliminates the time component \( t \) of the data and preserves the correlations between links.

6.2.2.1 A k-means clustering algorithm

The use of distances derived from the Euclidean norm may not be suitable for matrices since they are not invariant under isometries. This problem has been solved working with the spectrum of matrices because it is invariant with respect to common transformations.

We start the discussion by introducing the concept of matrix pencil.

**Definición 6.2.2** (Matrix pencil). Let \( A \) and \( B \) be two matrices \( n \times n \), the function \( L(\mu) = A - \mu B \) is called the matrix pencil and is represented by the pair \( (A,B) \). We say that this matrix pencil is a definite pencil if the matrices \( A \) and \( B \) are symmetric positive definite.

In the following definition, the concept of generalized eigenvalues in a matrix pencil is introduced.

**Definición 6.2.3** (Generalized eigenvalues). Given two square matrices \( A \) and \( B \) of dimension \( n \times n \), the generalized eigenvalues are defined as the roots of the polynomial \( |A - \mu B| = 0 \). These values will be denoted by \( \mu_s(A,B) \) with \( s = 1, \ldots, n \).

The following theorem characterizes the generalized eigenvalues for a definite pencil.

**Theorem 6.2.1** (Parlett (1997)). Let \( (A,B) \) be a definite pencil, then there exists a non-singular matrix \( V \) so that it can simultaneously diagonalize the matrices \( A \) and \( B \), which means:

\[
\begin{align*}
V^TAV &= \Lambda \\
V^TBV &= \Sigma
\end{align*}
\]

being \( \Lambda = diag(\lambda_1, \ldots, \lambda_n) \) and \( \Sigma = diag(\sigma_1, \ldots, \sigma_n) \). The generalized eigenvalues are finite, real and are given by the expression:

\[
\mu_s(A,B) = \frac{\lambda_s}{\sigma_s}
\]

(6.8)
Below, various measures are defined based on the spectrum of a matrix.

**Definición 6.2.4** (Dissimilarity measures for square matrices). We define the following dissimilarity measures between square matrices:


\[
 d_{PD}(A, B) = \sum_{s=1}^{n} \left( \mu_s^* - \frac{2}{\sqrt{2}} \right)^2 
\]

(6.9)

where

\[
 \mu_s^* = \frac{1 + \mu_s(A, B)}{\sqrt{1 + \mu_s(A, B)^2}}
\]

- **Geodesic distance (G)**, Lang (1999). The geodesic distance between two positive-definite matrices \(A\) and \(B\) is given by

\[
 d_G(A, B) = \left( \sum_{s=1}^{n} \log^2 (\mu_s(A, B)) \right)^{\frac{1}{2}}.
\]

(6.10)

- **Derived distance from the spectral norm (E)**, Golub and Loan (1997).

\[
 d_E(A, B) = \sqrt{\mu_{\text{max}}}
\]

(6.11)

where \(\mu_{\text{max}}\) represents the largest eigenvalue of the matrix \(A - B\).

In calculating these distances some computational problems can arise derived from the singularity of one of the matrices. González-Hernández (2010) provides a discussion about the calculation of the generalized eigenvalues in this case.

Also note that for \(d_G\) the matrices must be positive definite. This property holds for our data. Be \(v \in \mathbb{R}^n\) be \(\xi \in \{q, k, v\}\), we have

\[
v' X^\xi v = \sum_{t=1}^{m} v'(X^\xi_{t*})' (X^\xi_{t*}) v \\
\sum_{t=1}^{m} [(X^\xi_{t*}) v]' (X^\xi_{t*}) v = \sum_{t=1}^{m} (w_t)^2 \geq 0
\]

(6.12)

where \(w_t = (X^\xi_{t*}) v\) and therefore is a positive-semi-definite matrix. In addition if the vector \(\sum_{t=1}^{m} (X^\xi_{t*})' > 0\) component by component the matrix is positive definite.
6.2. Methodology

**Proposición 6.2.1** (Invariance of dissimilarity measures with respect to scale of data). Let $A, B \in M_{n \times n}(\mathbb{R})$ and let $\alpha \in \mathbb{R} - \{0\}$ and $\beta > 0$, thus

$$d_\xi(\alpha A, \alpha B) = d_\xi(A, B) \quad \text{with} \quad \xi \in \{\text{PD, } G\} \quad (6.13)$$

$$d_E(\beta A, \beta B) = \sqrt{\beta} d_E(A, B) \quad (6.14)$$

**Proof.** Equation (6.13) is proved by showing that the generalized eigenvalues of $(A, B)$ and $(\alpha A, \alpha B)$ are the same. The relationship

$$|\alpha A - \mu \alpha B| = |\alpha (A - \mu B)| = \alpha^n |A - \mu B| \quad (6.15)$$

proves that $\mu$ is a generalized eigenvalue of $(\alpha A, \alpha B)$ if and only if it is a generalized eigenvalue of $(A, B)$.

In order to prove the equation (6.14), we consider the following relationship

$$|\beta A - \beta B - \hat{\mu} I_n| = \left|\beta (A - B - \frac{\hat{\mu}}{\beta} I_n)\right|$$

$$= \beta^n \left|A - B - \frac{\hat{\mu}}{\beta} I_n\right| \quad (6.16)$$

Equation (6.16) shows that $\hat{\mu}$ is an eigenvalue of $(\beta A - \beta B)$ if and only if $\frac{\hat{\mu}}{\beta}$ is an eigenvalue of $(A - B)$. By assumption $\beta > 0$,

$$d_E(\beta A, \beta B) = \sqrt{\max_s \{\hat{\mu}_s\}}$$

$$= \sqrt{\beta \max_s \left\{\frac{\hat{\mu}_s}{\beta}\right\}} = \sqrt{\beta} d_E(A, B) \quad (6.17)$$

is satisfied.

Below, the composed dissimilarity measure used in our cluster analysis is introduced.

**Definición 6.2.5** (Composed dissimilarity measure $D_\xi(\cdot, \cdot)$). Let $A, B \in M_{n \times n}(\mathbb{R})$ with $A = (A^q, A^k, A^v)$ and $B = (B^q, B^k, B^v)$ the composed
dissimilarity measure between $A$ and $B$ is defined as:

$$
D_\xi(A, B) = \eta^q d_\xi(A^q, B^q) + \eta^k d_\xi(A^k, B^k) + \eta^v d_\xi(A^v, B^v)
$$

(6.18)

where $\xi \in \{P, D, G, E\}$, $\eta^q, \eta^k, \eta^v \geq 0$ and $\eta^q + \eta^k + \eta^v = 1$.

Proposition 6.2.1 guarantees that if $D_\xi(\cdot, \cdot)$ is used with $\xi \in \{P, D, G\}$ the units used for gauging the occupancy, the flow or the speed are irrelevant and these measures are comparable. This allows the $\eta^q, \eta^k$ and $\eta^v$ coefficients to be interpreted as confidence levels in the data or in its discriminant capacity and its weighting will depend also on the purpose of the traffic study.

One of the most widely used methods of cluster analysis is the $k$–means algorithm. The essential difference in our algorithm with other proposed methods is that we use the composed dissimilarity measures for square matrices (Definition 6.2.5).

The algorithm begins with a sample collected from the traffic network on a set of possible scenarios (working days, weekend days, holidays, days within vacation periods, Sundays, etc.) We denote by $\{X_i\}_{i=1}^N$ the sample obtained and by $X_i$ the matrix of observations collected on day $i$. We assume that the observations taken on different days are collected at the same time instants and in the same set of $n$ links. This implies that $X_i$ with $i = 1, \cdots, N$ are matrices $m \times n$ for each sampled day.

The first task is to transform the data set into observation matrices:

$$
A_i = (X_i)' \circ (X_i) \in M_n^3(\mathbb{R}), \ i = 1, \cdots, N.
$$

(6.19)

The $k$–means algorithm shall be applied on the set of matrices $\{A_i\}_{i=1}^N$ by using the composed dissimilarity measure given in definition 6.2.5. To describe the clustering algorithm, let $D_\xi$ be a generic dissimilarity measure and the following notation is introduced.

**Definición 6.2.6** (Voronoi regions). We will call $C = (C_1, \cdots, C_K)$ the set of centroids in the space of matrices $M_n^3(\mathbb{R})$. The so-called Voronoi regions are defined for each centroid $C_j$ as the set of matrices of $M_n^3(\mathbb{R})$ whose nearest centroid relative to the metric $D_\xi(\cdot, \cdot)$ is $C_j$. This is,

$$
R^\xi_j = \{A \in M_n^3(\mathbb{R}) \mid j = \arg \min_s D_\xi(C_s, A)\}.
$$

(6.20)

**Definición 6.2.7** (Voronoi set). The Voronoi set is defined $\Pi_j^\xi$ by:

$$
\Pi^\xi_j = \{i \in \{1, \cdots, N\} \mid j = \arg \min_s D_\xi(C_s, A_i)\}.
$$

(6.21)
Applying definitions (6.20) and (6.21), the $k$–means algorithm is described in Table 6.2.

**Tabla 6.2: The $k$-means algorithm for the CTDP**

**Step 0.** *(Data transformation).* For the set of observations $\{X_i\}_{i=1}^N$ on the traffic network calculate:

$$A_i = (X_i)' \circ X_i; \quad i = 1, \ldots, N.$$ 

Randomly initialize the set of centroids $C = \{C_1, \ldots, C_K\}$ with $\{C_j\}_{j=1}^K \subseteq \{A_i\}_{i=1}^N$. Choose a dissimilarity measure $\xi \in \{P \circ G, G\}$.

**Step 1.** *(Cluster update).* Calculate for each $C_j$ the set $\Pi_j^\xi$.

**Step 2.** *(Centroids update).* Calculate the new centroids $C_j$

$$C_j = \frac{1}{|\Pi_j^\xi|} \sum_{i \in \Pi_j^\xi} A_i; \quad j = 1, \ldots, K. \quad (6.22)$$

Where $|\cdot|$ is the cardinal of a set.

**Step 3.** *(Stopping criterion).* If the centroids have changed their coordinates then go to step 1. Otherwise, the algorithm has obtained the $K$–best clusters.

**Step 4.** *(Transfer the solution to the original space).* Calculate the $K$ dynamic traffic patterns $P_j$:

$$P_j = \frac{1}{|\Pi_j^\xi|} \sum_{i \in \Pi_j^\xi} X_i; \quad j = 1, \ldots, K. \quad (6.23)$$

### 6.2.2.2 Determining the number of clusters

The $k$–means algorithm assumes that the optimum number of clusters $K$ is known. This chapter extends the widely used Bayesian Information Criterion (BIC) and a two-stage process presented by Chiu et al. (2001) for determining the optimum number of clusters (see Section 5.4.1). The essential difference is that it is originally designed for cluster algorithms based on the Euclidean distance between vectors and in this chapter this is extended to the distance based on spectral methods for matrices. Below, the procedure is described by following the work of Xia and Chen (2007).

BIC is a likelihood criterion for model comparison that penalizes models with additional complexity. The BIC for the $J$–cluster solution is defined as:

$$BIC(J) = -2 \sum_{j=1}^J \xi(j) + 2JA \log(N), \quad (6.24)$$

in which $N$ denotes the total number of days in the dataset and the parameter $\xi(j)$ is a measurement of the estimated variance of the attribute variables for
cluster \( j \) defined as:

\[
\xi(j) = -n_j \left( \sum_{a=1}^{A} \frac{1}{2} \log (\hat{\sigma}_a^2 + \hat{\sigma}_{ja}^2) \right)
\]

in which \( n_j \) denotes the number of articles in cluster \( j \), \( \hat{\sigma}_a^2 \) denotes the estimated variance of the the \( a \)-th attribute for all articles in the dataset, and \( \hat{\sigma}_{ja}^2 \) denotes the estimated variance of the attribute \( a \)-th for those elements in cluster \( j \).

In this chapter the articles are matrices and the \( a \)-th attribute is not clearly defined. To overcome this difficulty we have considered that only one attribute \( A = 1 \) is measured which consists of the dissimilarity measure \( D_\xi \) and for this reason the sub-index \( a \) is deleted. Theoretically, the smaller BIC is associated with the better fit of the clustering model to the dataset.

The values \( \hat{\sigma}_j^2 \to 0^+ \) are interpreted as spherical clusters with respect to the dissimilarity measure \( D_\xi \) in which the articles are uniformly distributed in the sphere. Selecting the appropriate number of clusters for the kind of system presented in this work seeks to obtain homogeneous groups which at the same time contain a significant number of articles.

The two-stage process first examines the BIC for all potential clustering solutions. The goal is to find the smallest number of clusters that has the lowest BIC, because the BIC decreases first and then increases as the number of clusters increases, the BIC for each \( J \) (clustering solution) is calculated. Beginning from \( J = 1 \), the first \( J = J^* \) value that satisfies \( BIC(\hat{J}) < BIC(\hat{J}+1) \) is chosen as a coarse estimate of the number of clusters. In the second stage, the ratio of changes in dispersion measurement is used to determine the optimum number of clusters based on the coarse estimate obtained in the first stage. The ratio of changes in dispersion measurement is defined as \( R(J) = s_{J-1}/s_J \) for \( J = 2, ..., \hat{J} \), in which \( s_{J-1} \) denotes the change in dispersion measurement if \( J \) clusters are merged into \( J-1 \) clusters.

The parameter \( s_J \) can be computed as \( s_J = l_J - l_{J+1} \), in which \( l_J = \sum_{j=1}^{J} \xi(j) \). This second stage is based on the understanding that a significant increase in \( R(J) \) will be observed when two clusters that should not be merged are merged. The \( R(J) \) value for each \( J = 2, ..., \hat{J} \) is calculated, and the two largest \( R(J) \) values are identified as \( J = J_1 \) (the largest) and \( J = J_2 \) (the second largest). Xia and Chen (2007) uses an empirical threshold value of \( R(J_1)/R(J_2) = 1.15 \); that is, if \( R(J_1)/R(J_2) > 1.15 \), \( K \) is set to \( J_1 \); otherwise, \( K \) is set to \( \max(J_1, J_2) \).

### 6.2.3 Dynamic classification of daily traffic pattern

The prototype of urban traffic control system requires a dynamic classification of the current traffic pattern in real time. The system implements the
appropriate rules of traffic control for the prediction of the current traffic state. An essential characteristic that this approach must have is parsimony. This is, the prediction of the traffic state should be changed only if it is strongly necessary. Each prediction change involves a policy change in the traffic control (for example light controls) and the couplings between policies can be chaotic.

This section discusses how a parsimonious system of prediction should be built (that we will call non-nervous) from a prediction system based only on the current state \( t \) (that in contrast we will call nervous system).

Cluster analysis allows a matrix \( Y \in M_{N \times K}(\mathbb{R}) \) to be obtained defined by

\[
Y_{ij} := \begin{cases} 
1 & \text{ if } i \in \Pi_j \\
0 & \text{ Otherwise} 
\end{cases}
\]

which classifies the \( N \) observed days into \( K \) patterns.

A non-nervous prediction system can be conceptually formulated in the following form. Let us assume that \( Y \) and the set of observations in the traffic network are known. \( \{(X_i^{qi}, X_i^{ki}, X_i^{vi})\}_{i=1}^N \) within period \( t \). The so-called nervous prediction system is used as a basis for predicting the current traffic pattern which corresponds to the current observations \( z^t \). Suppose that this method is able to calculate

\[
y^t = F \left( z^t, \left\{ \left( X_i^{qi}, X_i^{ki}, X_i^{vi} \right) \right\}_{i=1}^N, Y \right); \tag{6.26}
\]

and the next current traffic profile is selected

\[
j^* = \arg \max_{j \in \{1, \ldots, K\}} \left\{ \sum_{s=1}^t y^s_j \right\} \tag{6.27}
\]

For example, discriminant analysis adjusts a multivariate normal in each cluster and with the equation (6.26) the probability of belonging to the group is calculated. In this case the optimization criteria is maximisation and by (6.27) the method assigns the more likely cluster.

From this nervous system (which only considers the current state) a non-nervous system is defined which considers the events of the entire day. The proposed classification criterion is:

\[
j^* = \arg \max_{j \in \{1, \ldots, K\}} \left\{ \sum_{s=1}^t y^s_j \right\} \tag{6.28}
\]

This criterion is interpreted as a collection of evidence from the preceding periods.
For the purpose of providing equal importance to all $t$ periods, the values of $y_t^j$ must be in the same range for all $t$. This occurs in discriminant analysis and in the Neural Networks where $y_t^j \in [0,1]$. There are prediction methods in which $y_t^j$ does not belong to a predefined range and it could occur that the predictions of the whole day might be conditioned by an outlier (incident) in a specific period. For example if the function (6.26) is employed as the Euclidean distance between the observations and the cluster centroids. For these cases it is appropriate to standardise the measures dividing by the maximum score of the period.

### 6.3 Implementation

#### 6.3.1 Experiment 1: A simulation study for assessing the proposed methodology

The purposes of these computational tests are:

1. Analysing the performance of the $k$–means algorithm applied to the CTDP problem with respect to the dissimilarity measure used in the solution obtained.

2. Assessing the computational efficiency of various classifying approaches on the basis of non-nervous systems.

#### 6.3.1.1 Data description

The data for this experiment have been generated through a simulation experiment, using the dynamic traffic load model developed in Sánchez-Rico et al. (2014). The Nguyen-Dupuis network shown in Figure 5.3 was used.

The network has 4 origin-destination pairs: $\omega_1 = (C_1, C_2)$, $\omega_2 = (C_1, C_3)$, $\omega_3 = (C_4, C_2)$ and $\omega_4 = (C_4, C_3)$. The demand intensity in pair $\omega$ is given by the following expression

$$q_\omega(t, \delta_\omega) := \sum_{i=1}^{2} \frac{\delta_\omega}{\sigma_i^w \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{t - \mu_i^w}{\sigma_i^w} \right)^2}$$

(6.29)

whose parameters are included in Table 6.3.
6.3. Implementation

<table>
<thead>
<tr>
<th>Pair</th>
<th>δω1</th>
<th>μω1</th>
<th>σω1</th>
<th>δω2</th>
<th>μω2</th>
<th>σω2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ω1</td>
<td>4000</td>
<td>8.0</td>
<td>15.0</td>
<td>2.0</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>ω2</td>
<td>8000</td>
<td>9.5</td>
<td>15.5</td>
<td>1.5</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>ω3</td>
<td>6000</td>
<td>7.5</td>
<td>14.5</td>
<td>2.0</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>ω4</td>
<td>2000</td>
<td>8.5</td>
<td>16.0</td>
<td>1.5</td>
<td>1.5</td>
<td></td>
</tr>
</tbody>
</table>

In this experiment, four demand patterns have been considered (see Figure 6.2) and are defined by the vector δ. Pattern P1 is considered the basis demand, pattern P2 is 75% of the basis demand, pattern P3 employs the basis demand in pairs with origin C1 and the 75% of the basis demand in pairs with origin C4 and pattern P4 is identical to P3 but exchanging the centroid percentages.

![Figure 6.2](image.png)

**Figure 6.2:** Intensity of O-D demand for pattern P1

For each of these demand patterns a random sample of size 25 was generated taking into account that the parameter vector δ = (δω1, δω2, δω3, δω4) ~ N(μδ, Σδ) has a normal distribution with an average of μδ and variance-covariance Σδ. These parameters are shown in Table 6.4. For each of the 100 daily traffic demands generated (defined by the parameter δ), through a traffic simulation model, the density and flow in links a5, a6, a7 are obtained every 5 minutes during the study period. By sampling these links it is possible to detect demand variation in either centroid C1 and C4.
CHAPTER 6. AN APPROACH TO DYNAMICAL CLASSIFICATION OF DAILY TRAFFIC PATTERNS

Table 6.4: Average and variance-covariance matrix of parameter \( \delta \) for traffic patterns

<table>
<thead>
<tr>
<th>Pattern</th>
<th>( \omega_1 )</th>
<th>( \omega_2 )</th>
<th>( \omega_3 )</th>
<th>( \omega_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>4000</td>
<td>8000</td>
<td>6000</td>
<td>2000</td>
</tr>
<tr>
<td>P2</td>
<td>3000</td>
<td>6000</td>
<td>4500</td>
<td>1500</td>
</tr>
<tr>
<td>P3</td>
<td>4000</td>
<td>8000</td>
<td>4500</td>
<td>1500</td>
</tr>
<tr>
<td>P4</td>
<td>3000</td>
<td>6000</td>
<td>6000</td>
<td>2000</td>
</tr>
</tbody>
</table>

\[ \Sigma_\delta \text{ for all pattern} \]

<table>
<thead>
<tr>
<th>Pair</th>
<th>( \omega_1 )</th>
<th>( \omega_2 )</th>
<th>( \omega_3 )</th>
<th>( \omega_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \omega_1 )</td>
<td>10000</td>
<td>2500</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \omega_2 )</td>
<td>2500</td>
<td>10000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \omega_3 )</td>
<td>0</td>
<td>0</td>
<td>10000</td>
<td>2500</td>
</tr>
<tr>
<td>( \omega_4 )</td>
<td>0</td>
<td>0</td>
<td>2500</td>
<td>10000</td>
</tr>
</tbody>
</table>

6.3.1.2 Clustering of daily traffic demand patterns in synthetic data

This experiment uses the \( k \)-means algorithm (Table 6.2) with the dissimilarity measures \( d_{PD} \), \( d_{E} \) and \( d_{G} \). The \( k \)-means algorithm reaches local optima, and so to overcome this drawback the \( k \)-means algorithm has been replicated 100 times starting from random initializations on centroids. To measure the quality of the clustering obtained by the algorithm, the Kappa concordance coefficient has been applied.

The Kappa concordance coefficient \( \kappa \) (see Carletta (1996)) is a statistic that emerged to measure the concordance between two experts. This magnitude is also applicable in the context of cluster analysis as is suggested in López-García et al. (2014). Each object of the sample is classified twice, once using the true categories system \( c_1, \ldots, c_K \) and again using the category system employed by the cluster analysis algorithm. It is assumed that a mapping \( \text{map}(c_k) \) is performed between the two category systems. The objective is to evaluate if the observed concordance between the classifiers is higher than expected. This statistic is defined by:

\[
\kappa = \frac{P_o - P_e}{1 - P_e}, \tag{6.30}
\]

where \( P_o = \sum_{k=1}^{K} p(c_k, \text{map}(c_k)) \) is the observed proportion of objects in the sample that have been classified in the same category and the parameter
6.3. Implementation

\[ P_e = \sum_{k=1}^{K} \hat{p}(c_k) \hat{p}(\text{map}(c_k)) \]

is the expected concordance proportion between groupings where \( \hat{p}(c_k) \) is the original probability of choosing an object from category \( c_k \) and \( \hat{p}(\text{map}(c_k)) \) is the probability that the cluster algorithm might assign an object to category \( \text{map}(c_k) \). To calculate the Kappa value, a mapping \( \text{map}(\cdot) \) which maximizes the value of \( \kappa \) is computed. The numerator of \( \kappa \) is the observed proportion which is higher than expected and its denominator is the maximum value that the numerator could take. The \( \kappa \) value takes values between \([-1, 1]\). A concordance level \( \kappa = 1 \) is produced only when there exists concordance in 100% of the observations. The result obtained by the \( k \)-means algorithm depends on the centroid initializations. In order to control this effect, the \( k \)-means algorithm is repeated one hundred times with random initializations. For each of these the corresponding kappa value is determined and the procedure is performed for two types of data, occupancy and flow.

The results obtained are shown in Table 6.5. In this table it is possible to observe the mean value, the standard deviation and the minimum and maximum value obtained from the sample; we see that the maximum value of \( \kappa \) for the dissimilarities is \( \max(\kappa_E) = 0.3467 \), \( \max(\kappa_{PD}) = \max(\kappa_c) = 1 \) which indicates that the distances \( d_G \) and \( d_{PD} \) allow an exact reconstruction of the original demand patterns but also in no case for the spectral norm \( E \).

**Tabla 6.5:** Descriptive statistics associated with the Kappa concordance coefficient

<table>
<thead>
<tr>
<th></th>
<th>Based on flows</th>
<th></th>
<th>Based on densities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \bar{\kappa} )</td>
<td>( \sigma_\kappa )</td>
<td>( \min(\kappa) )</td>
</tr>
<tr>
<td>( d_G )</td>
<td>0.817</td>
<td>0.201</td>
<td>0.333</td>
</tr>
<tr>
<td>( d_{PD} )</td>
<td>0.852</td>
<td>0.202</td>
<td>0.507</td>
</tr>
<tr>
<td>( d_E )</td>
<td>0.124</td>
<td>0.157</td>
<td>0.000</td>
</tr>
</tbody>
</table>

The question is which of the two distances \( d_{PD} \) or \( d_G \) that are able to reconstruct the original clusters gives better computational performance. This is analysed by evaluating which requires the lowest number of \( k \)-means algorithm repetitions to ensure the optimal global solution (reconstruction of the original clusters).
Figures 6.3 y 6.4 show respectively the sampling distribution of $\kappa$ based on flow and density observations. In the flow case both distances achieve the global optimum in more than 50% of cases. In the density case the distance $d_G$ obtains 20% success but also the $d_{PD}$ only 2%.

Figure 6.3: Histograms of the concordance coefficient $kappa$ for identification of patterns based on flows

Figure 6.4: Histograms of the concordance coefficient $kappa$ for Identification of patterns based on densities

A conclusion derived from table 6.5 and the histograms of Figures 6.3-6.4 is that the use of the metric $d_G$ is recommended.

The solution obtained by the CTDP are shown in Figure 6.5 y 6.6. It is possible to verify in these graphs the difficulty of performing this task visually as different demand patterns can share the same pattern in some subset of
6.3. Implementation

Since this is a simulation experiment, the number of clusters is controlled and known by the user but this situation is not available with real data.
6.3.1.3 Dynamic classification for synthetic data

In this section 100 new days have been generated through simulation, 25 for each of the four existing patterns. With this new data we evaluate the predictive capacity of a non-nervous classifier.

Three basic classifiers have been considered:

**Euclidean Distance.**

The most basic classifier is to assign the \( j \) pattern in the instant \( t \) whose centroid is the nearest to the observed data \( z^t \) in the current period \( t \). We consider the relative distance measure based on

\[
y^t_j = \frac{\|z^t - P^t_j\|^2}{\max_{j' \in \{1, \ldots, K\}} \|z^t - P^t_{j'}\|^2}
\]

where \( \| \cdot \| \) is the Euclidean norm.

**Discriminant Analysis.** Discriminant analysis proposes estimating the class-membership probability \( P_{j/z} \) adjusting multivariate normal densities with covariance estimates stratified by group. Following the discriminant analysis, the best cluster membership \( j^* \) given \( z^t \) is determined by maximizing the posterior probability \( P_{j/z} \). The resulting non-nervous classifier is given by

\[
j^* = \arg \max_{j \in \{1, \ldots, K\}} \sum_{s=1}^{t} P_{j/z} (j/z^s)
\]

**Neural Network.** The third classifier is based on neural networks. The architecture of neural network for daily traffic pattern classification is shown in Figure 6.7. It consists of an input layer with one node for each traffic parameter observed. The number of nodes of the hidden layer is chose to be twenty. The network with a single hidden layer failed to produce satisfactory results because of the complexity of the pattern-recognition problem. However, as the aim of this experiment is to show how the corresponding non-nervous classifier improves the basis classifier, this configuration has been maintained.

The training of neural networks is initiated with random values for the weights of the links. The weights of the neural network are optimized using the Levenberg-Marquardt back propagation algorithm until system error is reduced to a given tolerance (a value of \( 1e-06 \) is used in this research) or a number of given iterations is achieved (a value of 1000 for this research). The neural network is trained for each period and more than 200 time periods are considered. The computational time for each period is around 20 seconds. Over a number of periods the limit of iterations is reached but to limit the total time, this parameter has been maintained. However the training is done in the off-line phase, and thus is applicable in a real context.
6.3. Implementation

The results are shown in Figures 6.8, 6.9 y 6.10. The most important conclusion is that a non-nervous classifier improves the accuracy of the basis classifier. This is because the non-nervous classifier incorporates in its forecast historical data instead of only using the information of the current time interval. In addition it is possible to observe that from a period the true patterns are correctly selected by the classifier in the 100 days. The classifier with the best performance is the discriminant analysis. This could be due to the fact that for data generation a normal multivariate variable has been employed to generate the demands in the O-D pairs and possibly produces flows and densities observed in the measured links which could also be distributed as normal multivariates. This is the assumption of the discriminant analysis.

Figure 6.8: Dynamic classification of traffic patterns based on Euclidean distance
CHAPTER 6. AN APPROACH TO DYNAMICAL CLASSIFICATION OF DAILY TRAFFIC PATTERNS

6.3.2 Experiment 2: Implementation with real data

6.3.2.1 Data description

The dataset was obtained through the California Freeway Performance Measurement System (PeMS) which collects the traffic data in real-time from over 25,000 individual detectors across all major metropolitan areas of the State of California. In order to test the methodology described in this article, observations over 100 days in 2013 have been collected from two dual loop detector stations in the Bay Area of California as shown in Figure 6.11 and traffic measurements include flow rate, speed, and occupancy. The data used are available at http://bit.ly/1hsTEjO.

For the experiment the data for the months of May, June, July and August from 2013 have been used and only the days with missing data have been deleted.
6.3. Implementation

Figure 6.11: Location of the selected detectors for testing the methodology through OpenStreetMaps contributors

6.3.2.2 Clustering of daily traffic demand patterns in real data

As a consequence of the numerical experiment in the previous section, in this test the dissimilarity measure $D_G$ is used as it is the most appropriate, giving improved results in our methodology. The experiment was repeated twice in order to reach a better interpretation of the data, in the test trial only the occupancy $(\eta^q, \eta^k, \eta^v) = (0, 1, 0)$ was considered and in the second test only the speed $(\eta^q, \eta^k, \eta^v) = (0, 0, 1)$.

The first task to be addressed is the optimum number of clusters. Table 6.6 lists the BIC and the ratio of changes in dispersion measurement for each number of clusters ranging from $J = 1$ to $J = 10$ based on occupancy. It is possible to observe that the BIC decreases from $J = 1$ to $J = 5$, therefore an estimation of the number of cluster should be $J = 5$, on the other hand the largest ratios of changes in dispersion measurement for $J$ are $R(J_1) = 3.99$ when $J_1 = 3$ and $R(J_2) = 0.89$ when $J_2 = 2$, respectively, then the optimum number of cluster ($K$) is set to $K = J_1 = 3$. The same procedure has been conducted for the speed. It is observed that the BIC declines from $J = 1$ to $J = 5$, the largest ratios of changes in dispersion measurement for $J$ are $R(J_1) = 24.30$ when $J_1 = 5$ and $R(J_2) = 1.35$ when $J_2 = 4$, respectively and thus for the speed the optimum number of clusters is set to $K = J_1 = 5$. Figure 6.12 depicts BIC versus number of clusters. The BIC shows a decreasing behaviour and at some point starts to settle around a value. The applied method tries to determine the optimum number of clusters where a change in the behaviour is seen.
Table 6.6: BIC and R values for each clustering model based on occupancy and on speed

<table>
<thead>
<tr>
<th>Number of clusters</th>
<th>Occupancy</th>
<th>Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>(J)</td>
<td>BIC</td>
<td>R(J)</td>
</tr>
<tr>
<td>1</td>
<td>76.01</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>-20.71</td>
<td>0.89</td>
</tr>
<tr>
<td>3</td>
<td>-130.57</td>
<td>3.99</td>
</tr>
<tr>
<td>4</td>
<td>-151.22</td>
<td>2.88</td>
</tr>
<tr>
<td>5</td>
<td>-152.38</td>
<td>-120.39</td>
</tr>
<tr>
<td>6</td>
<td>-143.08</td>
<td>0.13</td>
</tr>
<tr>
<td>7</td>
<td>-136.68</td>
<td>0.60</td>
</tr>
<tr>
<td>8</td>
<td>-149.88</td>
<td>-8.52</td>
</tr>
<tr>
<td>9</td>
<td>-178.06</td>
<td>0.13</td>
</tr>
<tr>
<td>10</td>
<td>-164.46</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Figure 6.12: Number of clusters selected

After selecting the optimum number of clusters \( K \), the \( k \)-means algorithm is applied to the dataset. The number of days in each cluster is shown in Table 6.7. It can be seen that the size of the clusters is homogeneous and is high enough to establish that each cluster represents a daily traffic profile. The optimum number of clusters \( K \) represents a trade-off between the accuracy of the clusters and the necessary work to optimize the network for a high number of clusters.
6.3. Implementation

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Number of days (Occupancy)</th>
<th>Number of days (Speed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>43</td>
<td>19</td>
</tr>
<tr>
<td>P2</td>
<td>34</td>
<td>13</td>
</tr>
<tr>
<td>P3</td>
<td>23</td>
<td>33</td>
</tr>
<tr>
<td>P4</td>
<td>-</td>
<td>17</td>
</tr>
<tr>
<td>P5</td>
<td>-</td>
<td>18</td>
</tr>
</tbody>
</table>

The clustering algorithm is first implemented on the dataset for the occupancy. Figure 6.13 shows the average daily traffic occupancy profile of the resulting clusters.

The LOS measures for multi-lane freeways in the HCM (Transportation Research Board (TRB) (2010)) are used to describe the patterns. We transform the LOS measures A through F (Transportation Research Board (TRB) (2010)) for static flow pattern representation in terms of vehicle per kilometer-lane, as given by the equation (6.33)

\[
FP = \begin{cases} 
A, & \text{if } k \leq 7 \\
B, & \text{if } 7 < k \leq 11 \\
C, & \text{if } 11 < k \leq 17 \\
D, & \text{if } 17 < k \leq 22 \\
E, & \text{if } 22 < k \leq 28 \\
F, & \text{otherwise} 
\end{cases}
\]  

(6.33)

To transform the LOS measure thresholds into occupancy thresholds for two traffic lanes, we use the relationship

\[
b = 1 - (1 - kL)^2
\]  

(6.34)

where \( k \) is the density (vehicles per kilometer), \( L \) length of vehicles and \( b \) is the occupancy in a loop detector for a two traffic-lane location. For descriptive purposes, we have used \( L = 0.005 \) kilometers of length for the vehicles.

In Figure 6.13 it is possible to observe that the occupancy in the third pattern corresponds to a type of day without congestion, which could be for example a Sunday or a holiday, and the corresponding percentage for this pattern is around 23% of our sample. The number of Sundays in the sample is around 17%.

In patterns \( P1 \) and \( P2 \) the occupancies correspond to a recurrent congested day. Traffic pattern \( P2 \) has a higher congestion level than traffic pattern \( P1 \), reaching the level \( E \) during the peak-hour in the \( P2 \) pattern.
CHAPTER 6. AN APPROACH TO DYNAMICAL CLASSIFICATION OF DAILY TRAFFIC PATTERNS

Figure 6.14 shows the patterns obtained based on speed. In this figure traffic pattern $P_4$ can be highlighted because it represents uncongested behaviour where all users have a constant speed. In contrast traffic pattern $P_2$ shows a high level of congestion, where jams on the road are observed ($v = 0$). Traffic patterns $P_1, P_3$ and $P_5$ represent different degrees between the two states.

![Traffic pattern 1](image1)

![Traffic pattern 2](image2)

![Traffic pattern 3](image3)

Figure 6.13: Average daily occupancy profiles of the resultant CTDP

![Traffic pattern 4](image4)

![Traffic pattern 5](image5)

Figure 6.14: Average daily flow profiles of the resultant CTDP

6.3.2.3 Dynamic classification for real data

In this section fifty new days have been considered, different from the data employed before, to carry out the cluster analysis. This data has been
collected during the months of January and February 2014. The 100 days employed in the cluster analysis have also been used as a training sample for the classification procedure. In contrast to what happened in the simulation experiment, in this procedure the pattern to which each of the 50 days belongs is unknown and because of this it is not possible to determine the accuracy of the classification. However, it is possible to assume that if one procedure is reliable and capable of detecting the correct pattern for the current day, this forecast should not change throughout the day. Note that this characteristic system is a necessary but not sufficient condition of the procedure. In this section, the percent of change in the forecast from one period to the next is studied. The classification methods are the same as used for the synthetic data.

Figure 6.15 shows the evolution in the percentage of prediction not changed for all methods and for all occupancy and speed data. The NN method presents the greatest changes. This is due to the need to improve the network specification and/or the training phase.

For a better understanding of the procedures, the number of days is dynamically shown in each pattern and the classification changes every hour. Figures 6.16 and 6.17 show this information. The procedures based on the Euclidean distance and cluster analysis work in a similar way while the NM has a slightly different behaviour in the speed case. The patterns $P_1$ and $P_3$ can be considered a graduation of the $P_3$ pattern. The NM did not make the distinctions considered by other methods. Nevertheless, for the following objective, which is to impose a control system on the current traffic pattern, this confusion is not serious because it means that similar patterns should correspond to similar control strategies.
CHAPTER 6. AN APPROACH TO DYNAMICAL CLASSIFICATION OF DAILY TRAFFIC PATTERNS

Figure 6.15: Traffic classification for a real problem
Figure 6.16: Dynamic non-nervous classification based on occupancy
CHAPTER 6. AN APPROACH TO DYNAMICAL CLASSIFICATION OF DAILY TRAFFIC PATTERNS

Figure 6.17: Dynamic non-nervous classification based on average speed
A proposal for the design of Intelligent Transportation Systems for traffic control and management
A proposal for the design of Intelligent Transportation Systems for traffic control and management

7.1 Introduction

At the present moment the so-called ITSs concentrate a large amount of researches in the transportation field. This is due to nowadays half of the population lives in cities and this tendency is increasing every day, becoming necessary to develop traffic management tools for facing this continued growth of cities through reducing travel times and emission of pollutants by means of making effective traffic management such as route recommendations, making accurate traffic light controls among others.

The evolution of the development of ITS has progressed through different stages since the 70s, during the first decade called preparation phase was introduced the microprocessor, the beginning of GPS development and some dynamic route guidance systems based on real traffic conditions were developed. Despite that, the first stage was characterised by the construction of new roads due to the fact that the technologies had not yet matured enough. The feasible study phase from 1980 to 1995 is characterised by an explosion of development programs by governments and industry mainly in the European Union, Japan and United States. The underlying concepts and basic technologies of these programs were established during this phase. The most relevant programs that appears in this stage were the PROMETHEUS, DRIVE and ERTICO projects in Europe, the ITS America in the US and the RACS and AMTICS projects in Japan (see Figueiredo et al. (2001)). The current stage is called product development phase and it is characterised by the creation of feasible products focused on large-scale networks. There is increasing interest of governments and research groups in the development
CHAPTER 7. A PROPOSAL FOR THE DESIGN OF AN ITS FOR TRAFFIC
CONTROL AND MANAGEMENT

of efficient ITS. The objectives are to improve traffic conditions and to palliate
the problems derived from traffic congestion. As show Rafiq et al. (2013)
the investment for research and for developing ITS projects has increased
significantly in the last years. The current tendency is the developments
d of products covering specific sectors, and these products are associated
to specific companies. Many methodologies have been emerged in the
development of ITS, however there is not yet a general prototype of ITS that
might be capable of integrating new products of different companies or new
functionalities and technologies.

Any ITS standard should possess four essential characteristics indepen-
dently of their final purpose. The first one, the scalability which will enable
to be applied both in small cities and in megalopolis like Mexico city, Tokyo,
etc. The second characteristic is the robustness, this is that the system is
capable to give an appropriate response for any possible situation, including
incidents in the traffic network. The third characteristic is that the system
 should be technologically independent, this is, focusing on the data flows and
not in the technological devices used in the data acquisition. This last char-
acteristic is critical in order to avoid the system obsolescence, therefore capa-
ble of integrating new technologies in the future. Finally the system should
be able to manage Big Data in the system, this is, the system must learn from
the registered experience by different data sources (sensors, users, etc.). In
the last years novel technologies and paradigms have emerged with the pur-
pose of developing more suitable solutions of ITS in real contexts, this is the
 case of the agent computing paradigm, which is rapidly emerging as one of
the powerful technologies for the development of ITS due to its benefits for
deal with the uncertainty in a dynamic environment and in real-time context.

This chapter has focused on proposing an ITS for route recommendations,
specifically an Intelligent Traffic Control System (ITCS) based on the
integration of the work developed in previous chapters. The main objective is
the adaptation of the general framework described in Chapter 2, Figure 2.2,
for the development of ITS which also fulfill the characteristic of scalability,
technological independence, use of big data and robustness.

7.2 Related literature

No recurrent and recurrent congestion in a traffic network is a problem
that can be attenuated by re-routing the traffic flows in a large part of
the network. Usually the traffic control centres implement different control
actions in order to accomplish the enhancement of the traffic network.
To achieve this, it is necessary to predict the most likelihood evolution of
the state of the network and select the most appropriate control action.
The traffic control actions such as route recommendations, require expert
knowledge and much experience. ITCS can be used as a tool to face this
task through integrating short-term predictions with traffic monitoring and
control software. Several works have been developed in order to propose tools for supporting traffic control actions in control centres when non-recurrent congestion exists. Dahal et al. (2013) present an ITCS for supporting decision makers in traffic management centres to identify coordinated control actions from a global view. This proposal uses fuzzy logic systems and neural networks, first the fuzzy techniques are used to build the structure of the system and the neural network updates the fuzzy rule-system and the structure of the system when new data become available. Hegyi et al. (2001) and Liu et al. (2011) are important works developed under this aspect as well. Hounsell et al. (2009) present a review of the the current State-of-the-Art of Urban Traffic Management (UTM) technologies and discuss the potential impacts of new vehicle technologies on UTM and opportunities they provide.

The management and control of a traffic network can be tackle in two main directions which have been widely studied in the last years from different approaches:

1. Traffic light control and coordinations. Araghi et al. (2015) present a review of intelligent methods that have been applied for controlling traffic signals and a comparison between the three key intelligent methods that have been used in this area, Q-learning, neural network, and fuzzy logic systems. Angulo et al. (2011) present a methodology in two main phases which address control problems of traffic lights in real time by using soft-computing techniques, clustering techniques and optimization methods.

2. Route recommendations. Wunderlich et al. (2000), propose a prediction technique for decentralized route guidance architectures to identify time-dependent link travel times which communicated to drivers, leads to time dependent fastest paths consistent with this forecast. Jula et al. (2008) investigate methods to predict travel times along the links and estimate arrival times at the nodes of a stochastic and dynamic network in real time.

Multiagent-Systems (MAS) can successfully be applied in many domains under the following three conditions: i) the problem domain is geographically distributed; ii) the subsystems exist in a dynamic environment; and iii) the subsystems need to interact with each other in a flexible manner. On urban traffic networks these conditions hold, and an agent-based approach is well suited. Chen and Cheng (2010) present a review of agent applications classified into five categories: 1) agent-based traffic control and management system architecture and platforms; 2) agent-based systems for roadway transportation; 3) agent-based systems for air-traffic control and management; 4) agent-based systems for railway transportation; and 5) multiagent traffic modelling and simulation, the proposal presented in this chapter is focused within the first classification. Claes et al. (2011) propose a multiagent system for anticipatory vehicle routing useful in large-scale dynamic environments. This proposal consist in a decentralized approach
CHAPTER 7. A PROPOSAL FOR THE DESIGN OF AN ITS FOR TRAFFIC
CONTROL AND MANAGEMENT

7.3 A proposal of an intelligent traffic control system

In this section a conceptual prototype of an ITCS is presented, which, according on the classic taxonomy of ITS corresponds to the field of Advanced Traffic Management Systems (see Section 2.1.1). The proposal is focused on a specific functionality, however the ITCS structure allows its adequacy to general functionalities. The main goal of the ITCS is the anticipatory route recommendation and this problem can be defined as the computation of paths from a set of predicted future link travel times, which, when disseminated to drivers, cause the same set of predicted link travel times to be realized by vehicles in the network. The travel-time prediction allows guided vehicles judiciously to avoid the most serious delays. The desired objective is reduced travel times for the guided vehicles as well as reduced delay for unguided vehicles.

Starting from a realistic proposal, the ITCS developed in this chapter is based on the assumption that it is possible to get vehicular information on a portion of the traffic network.

This proposal combines the tools developed in the previous chapters with the aim of fulfilling the four basic premises for an ITS described in 1.1. The characteristic of scalability will be achieved by using an agent-based framework (see Figure 7.1). Our approach was motivated by the MAS-based model for anticipatory route recommendation problems presented in Claes et al. (2011) in which there are three basic types of entities: 1) vehicle agents; 2) infrastructure agents; and 3) virtual environment. We consider that a part of vehicles are represented by a situated vehicle agent, deployed on (a smart device within) the vehicle. A vehicle agent provides information to the infrastructure agents about its destination, location and current route. Moreover, vehicle agent guides the driver by providing information on the best route in real time to reach their destination. The core elements of the road infrastructure (such as roads and crossroads) are represented and managed by infrastructure agents. These agents are deployed on computation and communication devices in the road infrastructure. Infrastructure agents exchange information about the current traffic measures with the virtual environment, which sends the optimal route recommendation according on the current prediction of the network state. The infrastructure agents recommend to the vehicular agents the best route according with the destination of the trip. The virtual environment is a centralised software representation of the environment which estimates the current traffic state in real time, and

based in the coordination between agents that are embedded in the real-world environment (vehicles and road infrastructure) and a decentralized software which conceptually hosts the infrastructure and vehicle agents.
7.3. A proposal of an intelligent traffic control system

delivery recommendations to infrastructure and vehicle agents. The route selection is based on the point of view of the optimum system, which consists of the overall travel time minimization. Figure 7.1 shows the described entities and its exchange of information for the proposed agent-based framework.

**Figure 7.1**: Agent-based architecture for the implementation of the ITCS

Figure 7.2 shows the tasks to be carried out in the virtual environment. Since the main objective of our proposal is the recommendation to users about the best route to reach their destination according on the current system status. The approach is based on two phases within the virtual environment: an 1) off-line phase, in which is built a recommended routes repository for different demand patterns. In this stage the most time-consuming computations are carried out and 2) on-line phase to estimate the current state and to make recommendations on real-time. We have adopted the proposal of Angulo et al. (2011) for the anticipatory vehicle routing problem. Figure 7.2 specifies the computing process that should be performed for each phase described above. The stages of traffic pattern recognition and temporal segmentation analysis corresponding to the off-line phase in the global proposal have been addressed in Chapter 5 and Chapter 6. The stages of traffic pattern identification, current instant identification and the implementation of the optimal policy corresponding to the on-line phase in the global proposal have been addressed in Chapter 6. In this chapter we deal with the traffic optimization approaches for all TODs.
The characteristic of technological independence will be achieved by projecting our proposal on the acquisition of data as speed, densities and flows, independently of the device used as a data source. Figure 7.3. shows the information flows for the proposed ITCS.

7.4 Numerical experiments

The ITCS objective is the route recommendations, for this reason, the system must be capable of computing within an off-line context the set of optimal recommendations for each traffic pattern presented in the network. The
7.4. Numerical experiments

equilibrium traffic assignment models adopt the user equilibrium or the system optimal with the aim of modelling the route selection. In this chapter the system optimal criterion is adopted, under which, the recommendations are made in such a way that the total travel time in the system is minimized, although for some particular user, the system gives a recommendation that might not be his/her best option. As in previous chapters, for this numerical experiment the dynamic traffic load model developed in Sánchez-Rico et al. (2014) has been used. The test is conducted on the Nguyen-Dupuis network (see Figure 5.3).

7.4.1 Route optimization

A method to tackle the non-stationary demand patterns is the so-called Time-Of-Day intervals. This method divides the study period [a,b] in time
subintervals in which the demand is approximately stationary. In this section the followed procedure for TOD selection explained in Chapter 5 is considered.

![Image](image.png)

**Figure 7.4:** Intensity of O-D demand and TOD breakpoints

It is assumed that the study period $[a, b]$ has been partitioned in $K$ TODs (see Figure 7.4), this is $[a, b] = \bigcup_{k=1}^{K} TOD_{k}$ and $TOD_{k} \cap TOD_{k'} = \emptyset$ where $k \neq k'$ and within each time interval the routes are recommended with constant probabilities, this is:

$$\mathbb{P}(p, t) = x^k_p, \text{ with } t \in TOD_k, \text{ and } K = 1, \ldots, K$$ (7.1)

By interpreting the variables $x^k_p$ as the probabilities of choosing a route, this variables must hold

$$\sum_{p \in P_w} x^k_p = 1, \text{ for all } w \in W$$ (7.2)

$$0 \leq x^k_p \leq 1, \text{ for all } p \in P_w \text{ and for all } w \in W$$ (7.3)

Equation (7.2) imposes the users in pair $w$ must choose a path to make its route and equation (7.3) indicates that are probabilities.

It is denoted by $x^k = (\cdots, x^k_p, \cdots)$ the recommendation probabilities within the TOD$_k$. The set of restrictions (7.2)-(7.3) it is denoted by:

$$x^k \in X, k = 1, \ldots, K$$ (7.4)
7.4. Numerical experiments

and it defines the recommendations during the $TOD_k$.

The total time spent by the demand in the traffic network in function of the recommendations provided can be conceptually expressed as:

$$Z(x^1, x^2, \cdots, x^K)$$

(7.5)

In this work the function $Z(\cdot)$ is computed by using the loading network model of the Chapter 4. The problem that is addressed in this chapter can be defined as the following optimization problem:

$$\begin{align*}
\mathbf{P} \quad & \text{minimize} & & Z(x^1, x^2, \cdots, x^K), \\
\quad & \text{subject to} & & x^k \in X, k = 1, \cdots, K
\end{align*}$$

(7.6)

The problem $\mathbf{P}$ consists of optimizing a simulation model with a high computational cost. In this numerical experiment, different algorithms have been considered with the main goal of obtaining the optimal recommendations for users in the traffic network.

7.4.2 Experiment 1: Route recommendations

In the experiments it is considered that the users choose his/her route with a prior probability distribution (see Table 7.1). We have employed the proportion of trips in each path obtained from the static traffic assignment problem as the prior probability. With the challenge of route recommendations, it is intended to improve these probabilities through their suitability with the demand intensity observed in each TOD. Basically, the dynamic TAP is being solved.

The numerical test has been addressed through different generic optimization methods, such as genetic algorithms, the Nelder Mead method, and a greedy algorithm. This last method was designed for recommending only one route for each pair and TOD. All of these proposals, for different reasons, have not been succeeded in improving the prior choices made by the users.

Finally, a projection algorithm has been used, in which within each TOD works as the projection algorithm for the static TAP (see Patriksson (1994)).

This method looks for a user equilibrium, and therefore there is no guarantee of improvement in the system optimal criterion. However, we thought that, if the static equilibrium had been adapted to the dynamic equilibrium, this process could enhance also the system optimal criterion.

The solutions obtained with the projection algorithm are shown in Table 7.2. These results correspond to the system recommendations. Figure 7.6
depicts the travel times with and without the system recommendations. Despite the fact that in general there is no significant improvement in the travel times (see Figure 7.5), the results are more consistent. For example, in Patterns 1 and 4 the path 6 is less recommended and therefore the travel times are more close, on the other hand in Patterns 2 and 3 the path number 6 is not recommended by the system since generates the greatest travel times. Another example of how the route recommendations affect the system is shown in Figure 7.7. It is possible to observe that in the same link the distribution of packets change when the route recommendations have been carried out.

### Tabla 7.1: Prior route choice probabilities (user point of view)

<table>
<thead>
<tr>
<th>Pair</th>
<th>Path</th>
<th>TOD₁</th>
<th>TOD₂</th>
<th>TOD₃</th>
<th>TOD₄</th>
<th>TOD₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>ω₁</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.469</td>
<td>0.469</td>
<td>0.469</td>
<td>0.469</td>
<td>0.469</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
</tr>
<tr>
<td>ω₂</td>
<td>4</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.031</td>
<td>0.031</td>
<td>0.031</td>
<td>0.031</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
<td>0.125</td>
</tr>
<tr>
<td>ω₃</td>
<td>7</td>
<td>0.333</td>
<td>0.333</td>
<td>0.333</td>
<td>0.333</td>
<td>0.333</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.667</td>
<td>0.667</td>
<td>0.667</td>
<td>0.667</td>
<td>0.667</td>
</tr>
<tr>
<td>ω₄</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
### 7.4. Numerical experiments

**Tabla 7.2:** Probabilities for route recommendations

#### pattern 1

<table>
<thead>
<tr>
<th>Pair</th>
<th>Path</th>
<th>TOD₁</th>
<th>TOD₂</th>
<th>TOD₃</th>
<th>TOD₄</th>
<th>TOD₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td>0.469</td>
<td>0.492</td>
<td>0.475</td>
<td>0.487</td>
<td>0.479</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.125</td>
<td>0.208</td>
<td>0.189</td>
<td>0.204</td>
<td>0.19</td>
</tr>
<tr>
<td>ω₂</td>
<td>4</td>
<td>0.250</td>
<td>0.250</td>
<td>0.235</td>
<td>0.251</td>
<td>0.238</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.031</td>
<td>0.021</td>
<td>0.044</td>
<td>0.021</td>
<td>0.038</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0.125</td>
<td>0.028</td>
<td>0.057</td>
<td>0.037</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.379</td>
<td>0.379</td>
<td>0.381</td>
<td>0.379</td>
<td>0.377</td>
</tr>
<tr>
<td>ω₃</td>
<td>8</td>
<td>0.621</td>
<td>0.621</td>
<td>0.619</td>
<td>0.621</td>
<td>0.623</td>
</tr>
</tbody>
</table>

#### pattern 2

<table>
<thead>
<tr>
<th>Pair</th>
<th>path</th>
<th>TOD₁</th>
<th>TOD₂</th>
<th>TOD₃</th>
<th>TOD₄</th>
<th>TOD₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td>0.469</td>
<td>0.498</td>
<td>0.496</td>
<td>0.496</td>
<td>0.496</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.125</td>
<td>0.240</td>
<td>0.238</td>
<td>0.238</td>
<td>0.238</td>
</tr>
<tr>
<td>ω₂</td>
<td>4</td>
<td>0.250</td>
<td>0.253</td>
<td>0.253</td>
<td>0.253</td>
<td>0.252</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.031</td>
<td>0.009</td>
<td>0.013</td>
<td>0.013</td>
<td>0.014</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0.125</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.438</td>
<td>0.438</td>
<td>0.437</td>
<td>0.438</td>
<td>0.438</td>
</tr>
<tr>
<td>ω₃</td>
<td>8</td>
<td>0.562</td>
<td>0.562</td>
<td>0.563</td>
<td>0.563</td>
<td>0.563</td>
</tr>
</tbody>
</table>

#### pattern 3

<table>
<thead>
<tr>
<th>Pair</th>
<th>Path</th>
<th>TOD₁</th>
<th>TOD₂</th>
<th>TOD₃</th>
<th>TOD₄</th>
<th>TOD₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td>0.469</td>
<td>0.496</td>
<td>0.496</td>
<td>0.496</td>
<td>0.497</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.125</td>
<td>0.232</td>
<td>0.238</td>
<td>0.238</td>
<td>0.238</td>
</tr>
<tr>
<td>ω₂</td>
<td>4</td>
<td>0.250</td>
<td>0.254</td>
<td>0.252</td>
<td>0.251</td>
<td>0.251</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.031</td>
<td>0.017</td>
<td>0.014</td>
<td>0.015</td>
<td>0.015</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0.125</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.437</td>
<td>0.437</td>
<td>0.438</td>
<td>0.437</td>
<td>0.438</td>
</tr>
<tr>
<td>ω₃</td>
<td>8</td>
<td>0.563</td>
<td>0.563</td>
<td>0.563</td>
<td>0.563</td>
<td>0.562</td>
</tr>
</tbody>
</table>

#### pattern 4

<table>
<thead>
<tr>
<th>Pair</th>
<th>Path</th>
<th>TOD₁</th>
<th>TOD₂</th>
<th>TOD₃</th>
<th>TOD₄</th>
<th>TOD₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td>0.469</td>
<td>0.478</td>
<td>0.472</td>
<td>0.478</td>
<td>0.474</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.125</td>
<td>0.161</td>
<td>0.156</td>
<td>0.161</td>
<td>0.157</td>
</tr>
<tr>
<td>ω₂</td>
<td>4</td>
<td>0.250</td>
<td>0.251</td>
<td>0.245</td>
<td>0.251</td>
<td>0.246</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.031</td>
<td>0.025</td>
<td>0.034</td>
<td>0.026</td>
<td>0.032</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0.125</td>
<td>0.085</td>
<td>0.093</td>
<td>0.085</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.354</td>
<td>0.354</td>
<td>0.354</td>
<td>0.354</td>
<td>0.354</td>
</tr>
<tr>
<td>ω₃</td>
<td>8</td>
<td>0.646</td>
<td>0.646</td>
<td>0.646</td>
<td>0.646</td>
<td>0.646</td>
</tr>
</tbody>
</table>
CHAPTER 7. A PROPOSAL FOR THE DESIGN OF AN ITS FOR TRAFFIC CONTROL AND MANAGEMENT

7.5 Experiment 2: ITCS evaluation

The ITCS assessment has not been based on the reductions of the travel times because, a key factor is the employed optimization method for route recommendations. This is that the merit of the ITCS strongly depends on the quality of the route optimizer. For this reason the objective in this experiment is to assess the ITCS capacity on recognizing the current traffic pattern in real time.

During the experimentation two random samples have been generated. One of them employed in the training phase and other for its assessment. Each sample consists of twenty five days for each of the four patterns. Once more, the Nguyen Dupuis network described in previous chapters has been used with the same characteristics (see section 4.4).

The first experiment carried out consists of evaluating whether the ITCS is able to identify the current system status. One hundred days have been generated and the results are shown in Figure 7.8. It is possible to observe that the system is capable to detect the demand pattern in the 100% of cases.

In the next experiment the previous procedure was repeated but executing the ITCS in the route recommendation. In this example we have considered that the system recommends the route to all users. This fact has a strong
7.5. Experiment 2: ITCS evaluation

Figure 7.6: Travel times with and without recommendations for pair $\omega_2$

impact over the pattern of mobility in the network.

Figure 7.9 shows the obtained results. It is observed that the system is
capable of correctly detecting the status of the network for patterns 1 and 4. However, for patterns 2 and 3 the recommendations affect significantly the observations and the ITCS is not capable of identifying the patterns. These results illustrate a key factor: the ITCS is trained without intervention but during its working the observations are affected by its own intervention. This result shows the need of the simulation models which may allow to detect how the perturbed observations should be after the intervention in the system. In addition, the data must be incorporated to the module that predicts the system state.

Figure 7.7: Links behaviour with and without recommendations
**Figure 7.8:** Detection of the current demand pattern without the recommendations

**Figure 7.9:** Detection of the current demand pattern with the recommendations
Conclusions and suggestions for further work
Conclusions and suggestions for further work

8.1 Conclusions

Human challenges grow in complexity and multiply their relationships. To deal with this, new methodologies are required from different technological and scientific fields. During the development of this thesis it has been necessary to work in different areas, such as traffic engineering, computer science and optimization techniques. From a general point of view, the work carried out during the realization of this thesis has contributed to traffic simulation models, dynamic cluster techniques, dynamic methods for traffic pattern identification and different optimization techniques.

This thesis develops analytical tools and approaches for the design of Intelligent Transportation Systems (ITS) for traffic management and control capable of satisfying the four premises scalability, technological independence, robustness and use of Big Data. The approaches presented in this thesis are focused on traffic control and management strategies in order to increase the network capacity in the cities through the advantage that modern technologies now offer.

The state-of-the-art in the field of ITS was analysed in Chapter 2 with the object of knowing its current situation, reviewing taxonomies on this topic and assessing the applications as well as possible gaps in order to locate our proposal within the field of traffic control and management tools in an online context. Chapter 3 describes the basic concepts of traffic flow theory, its essential magnitudes, its dynamic behaviour and its modelling.

This doctoral thesis has been organized according to the different stages required under the proposed scheme for the design of ITS for traffic control and management shown in Figure 2.2, each stage has been described in
its corresponding chapter. The conclusions obtained from each research chapter are described below.

* Chapter 4. A Monte Carlo approach to simulate the stochastic demand in a continuous dynamic traffic network loading problem.

A discrete event algorithm is proposed for solving a mesoscopic CDNL model. Compared with the proposals presented in the literature, we have opted for discretizations of the packet size $\Delta x$ instead of time discretizations. The numerical tests show a high performance for the algorithm.

The Monte Carlo simulation approach and the distributed computational strategies allow the stochastic CDNL model for general random variables to be tackled. This kind of approach is necessary for risk-averse traffic modelling.

The goal of the developed approach for the CDNL model is a trade-off between precision and computational cost. Our approach is based on a macroscopic whole link model which presents the following characteristics: i) it generalizes both the PQ model and the delay-function model, ii) it is compatible with the FIFO condition in the case of linear delay-function models and iii) it explicitly reflects time-dependent link capacities.

This last property is key in the evaluation of interventions in real time systems. The FIFO condition is used to speed up the discrete event simulation procedure.

The simulation code is very versatile, it permits on-line observations of the state of traffic to be incorporated easily and can be used for working with discontinuous intensities and capacities. From previous experience, the proposed approach is easily integrable as a component of flow propagation in DTA models with a wide range of applications. The application considered in this thesis was its use in advanced traveller information systems (ATIs). Dynamic signal optimization and ramp metering are other possible issues that a general model can study, in the topics of capacity management and speed regulation. In Chapter 7 a preliminary study using this model has been conducted for route recommendations.

* Chapter 5. Time series data mining to enhance traffic signal optimization.

In this chapter a methodology for automatically updating pre-timed traffic control systems is set out. For this purpose a bilevel model has been formulated, which simultaneously includes the problem of determining the time-of-day (TODs) breakpoints and the traffic control problem for each time interval. The proposed model has been solved by the use of a novel class of
metaheuristic algorithms. The effectiveness of the hybrid algorithm SA∗+NM has been demonstrated on a collection of synthetic and real problems and it outperforms GA, SA+NM and PSO+NM methods. Furthermore, this feature allows us to apply the SA∗+NM algorithm to bilevel problems.

The sequential and the simultaneous methodologies have been illustrated on a real problem. As might be expected the simultaneous method achieves better results than the sequential but the computational cost is higher. It can be seen that the sequential methodology, which is used in practice, achieves satisfactory results.

Finally, the automatic determination of the optimal number of TODs has been studied by employing the most promising methods, based on BIC, ΔBIC, a modified PETE algorithm and a problem-oriented approach. It is observed numerically that the criterion based on the BIC is conservative and produces a reduced number of TODs in comparison with the other methods. The index ΔBIC produces a large number of TODs. The problem-oriented method has an interpretation which allows its easy application and achieves fair values. The modified PETE method also achieves solutions in accordance with what is expected.

* Chapter 6. An approach to dynamical classification of daily traffic patterns.

In this chapter, a two-stage urban traffic control system is analysed. In the off-line stage, a dynamic cluster technique is developed to classify the daily traffic profile in a road network, based on the \( k \)-means algorithm and on a definition of the composed spectral dissimilarity measures that are applicable to the data transformation \( \circ \) on traffic parameters: flow, speed and density (occupancy).

The optimum number of clusters is determined by applying the BIC criterion and the ratio of changes in dispersion measurement. In the on-line stage the current traffic pattern is dynamically predicted. A conceptual approach is constructed on the basis of a method of prediction that focuses on the current period.

The numerical tests have been conducted on two datasets. The first set of data is acquired by simulations, using the results obtained in Chapter 4, the second set is real data obtained from U.S. freeways through the California PeMS (Freeway Performance Measurement Systems). The numerical tests on synthetic data show that the dissimilarity measure \( d_\circ \) offers the best performance, and is able to reconstruct the original clusters with an accuracy of 100%. Moreover, the numerical tests illustrate that the non-nervous prediction outperforms the baseline prediction system.

Over real data it has been shown that the procedure described determines clusters with a significant number of patterns. In addition the prediction of the non-nervous systems shows parsimonious behaviour.
This study addresses the essential phases of the approach. It would be possible to add a preprocessing of noisy traffic data and a pre-classification into working days and non-working days (Weijermars and van Berkum (2005)), and to analyse the performance of the entire system.

* Chapter 7. A proposal for the design of Intelligent Transportation Systems for traffic control and management.

Traffic control systems based on TODs (as analysed in Chapter 5) are applied in urban areas without a proper technological infrastructure for real-time traffic monitoring. In this chapter we have introduced the use of these systems for urban areas equipped with traffic control centres. To this end, the ITS developed in Chapter 6 has been enriched by the TOD concept. The TOD concept, within a given demand pattern, replaces the short-term traffic predictions. A TOD-based approach considers the dynamic nature of the traffic as a finite sequence of stationary states. The finiteness of TODs allows the control strategies to be optimized in an off-line context.

In this chapter the proposed ITS focuses on route recommendations and a numerical experiment has been conducted to illustrate this approach.

The computational experiment carried out shows the difficulty of the problem of performing route recommendations dynamically; this problem is intrinsically related to the dynamic TAP. Some generic optimization procedures such as GA, the Nelder-Mead method and a greedy algorithm have been unsuccessfully applied. We believe that to tackle successfully this problem in real networks, it is required to apply problem-oriented methods, in which all the characteristics are exploited.

In a test problem, it was stated that the ITS is able to identify the network state in real time. The negative aspect detected is that the intervention of the ITS itself can strongly influence the observations and the predictions derived from them. This aspect must be analysed in future research.

### 8.2 Suggestions for further work

Working on this thesis has been a learning process in research; it is not the end in itself but a midpoint from which to continue. This is the reason why the further lines of research are focused on desired goals, as distant points on the horizon, instead of improvements on the current work.

1. **Dynamic clustering.** In this thesis, techniques for discovering dynamic patterns have been analysed. The essential characteristic is the monitoring of some attributes over time (several sensors). In the literature this problem has been deeply studied for univariate
8.2. Suggestions for further work

time series. We assume that the application of this method can be interesting in different fields, not necessarily in the transportation domain. Nowadays the so-called Analytics in a Big Data World is gaining importance in the research community and these techniques which incorporate the dynamic component are certainly relevant.

2. **Public transport.** The proposed ITS has a conceptual design. This characteristic makes it flexible to be adopted in public transport networks. These systems have specific characteristics that must be fulfilled, different from those of traffic networks. For example, the sensors register passengers instead of vehicles; new data sources exist for this purpose, such as mobile phones, ticket sales etc.; there are incidents on fleets of vehicles; lines are recommended instead of routes; etc. Despite these differences, we assume that the three following essential bases, on which the proposed ITS is supported, are still valid in a public transport context: i) The existence of a set of mobility patterns which are repeated over time; ii) A set of optimization techniques that allow the best system response to each of the patterns to be obtained; iii) A set of data mining tools which allows the current state of the system in real time to be inferred.

3. **Dynamic Traffic assignment (DTA).** A natural extension of Chapter 4 is to embed the CDNL within a route choice model, that is, to address the dynamic traffic assignment problem. The literature shows its extreme complexity and undoubtedly it is a big challenge. Nowadays effective computational resources exist and it is desirable to use them. The decomposition algorithms have been applied successfully to the static TAP. In particular the disaggregate simplicial decomposition presented in Larsson and Patriksson (1992). This method properly exploits the decomposition by O-D pairs. Another class of decomposition approaches based on O-D pairs are the Jacobi-type schemes which drive parallel computing approaches. The exploration of these ideas will be attempted in a dynamic context. The proposed simulation model is based on the concept of packet, which is characterised by its O-D pair, its route and its departure time. The key idea that must be tackled, in order to parallelize the DTA solution, is how to consider the effect (link queues) on a given O-D pair of a set of O-D pairs which are not being analysed.

4. **Agent-based approaches.** The approaches and techniques presented in this thesis and mainly the general proposal for the design of an ITS for traffic control and management have been developed with the intention of being applicable in real contexts. We consider that agent-based proposals are well placed to achieve this goal due to their benefits in the development of large-scale distributed systems. The main reason for the growing success of agent technology is the decomposition of the system into multiple agents that interact with each other to achieve a desired global goal. Currently, in the traffic and transportation area the Multiagent-Systems (MAS) have been widely applied in modelling, simulation, dynamic routing, congestion management and intelligent
traffic control, also MAS are in line with so-called "smart-cities" which seek to maintain the city's functionality through the use of information technologies such as mobile phones, GPS devices, video cameras, etc.
Conclusiones y futuras líneas de investigación

8.3 Conclusiones

Los retos humanos crecen en complejidad y multiplican sus interrelaciones. Afrontarlos requiere nuevas técnicas provenientes de campos diversos. En esta tesis se ha trabajado en temas de ingeniería del tráfico, ciencias de la computación y técnicas de optimización. Se han desarrollando modelos de simulación de tráfico, se han adaptado y testado técnicas cluster dinámicas, se han mejorado métodos de predicción de patrones dinámicos y se han usado técnicas de optimización.

Las propuestas presentadas en esta tesis persiguen incrementar la capacidad de la red de tráfico de las ciudades mediante el aprovechamiento de los avances tecnológicos que existen hoy en día. El objetivo principal de esta tesis es el desarrollo de un conjunto de herramientas para el desarrollo de Sistemas Inteligentes de Transporte (ITS) para la gestión y control de tráfico. Se ha perseguido dotar a estos sistemas de la capacidad de escalabilidad, independencia tecnológica, robustez y uso de grandes cantidades de datos.

El estado del arte en el campo de los ITS ha sido analizado en el capítulo 2, con los objetivos de conocer su situación actual, estudiar sus taxonomías y detectar posibles carencias para así ser capaces de situar nuestras propuestas dentro del campo del desarrollo de ITS para gestión y control de tráfico dentro de contextos en tiempo real. En el capítulo 3 se han descrito los conceptos básicos de la teoría de tráfico, sus magnitudes esenciales, su comportamiento dinámico y su modelización.

Esta tesis doctoral se ha organizado de acuerdo a las diferentes etapas que requeriría el esquema de la figura 2.2 para el desarrollo de un ITS para control y gestión de tráfico, cada etapa del prototipo ha sido llevada a cabo en su capítulo correspondiente. Las conclusiones obtenidas dentro de cada capítulo se describen a continuación.

* Capítulo 4. Una esquema de simulación Monte Carlo para abordar la demanda estocástica en el problema de carga dinámica de la red (CDNL).

En este capítulo se propone un algoritmo de eventos discretos para resolver un modelo mesoscópico para el problema CDNL. En contraste con los trabajos encontrados en la literatura, en esta propuesta se hace
una discretización basada en flujos (tamaño de paquete $\Delta x$) en lugar de una discretización basada en tiempos. Las pruebas computacionales han demostrado un alto rendimiento del algoritmo.

El esquema de simulación Monte Carlo y las estrategias de computación paralela empleadas han permitido que el modelo estocástico sea abordado para variables aleatorias genéricas. Este tipo de enfoques son necesarios para la modelización del llamado problema de asignación de tráfico con adversión al riesgo.

El objetivo del modelo propuesto ha sido encontrar un equilibrio entre la precisión y el coste computacional de su resolución. Nuestra propuesta se basa en una generalización del modelo macroscópico de arco Whole-link Travel Time model (WTT) el cual tiene las siguientes características: i) generaliza dos modelos existentes, el denominado Point-queue model y el delay-function model, ii) es compatible con la condición FIFO en el caso de los modelos lineales delay-function y iii) recoge la dependencia dinámica de la capacidad de los arcos.

Esta última propiedad es clave para la evaluación de intervenciones en tiempo real en el sistema. La condición FIFO es usada para acelerar el proceso de simulación de eventos discretos.

El código de simulación es muy versátil, éste permitiría la incorporación de observaciones en tiempo real para conocer el estado del tráfico de forma sencilla. Además puede ser usado en un contexto de capacidades e intensidades discontinuas. El enfoque propuesto es fácilmente integrable como un componente para la propagación de flujo en los modelos de asignación dinámica de tráfico (DTA) para un amplio rango de aplicaciones. La aplicación considerada en esta tesis ha sido los sistemas avanzados de información al viajero (ATIs). Otras posibles aplicaciones pueden ser la optimización dinámica de semáforos y el control de acceso a vías principales, específicamente en los temas de gestión de capacidad y regulación de velocidad. En el capítulo 7 se presenta un estudio preliminar para la recomendación de rutas basado en este modelo.

* Capítulo 5. Minería de datos de series temporales para mejorar la optimización de los planes prefijados de control tráfico.

En este capítulo se presenta una metodología para actualizar de forma automática los llamados pre-timed traffic control systems que funcionan de acuerdo a estrategias establecidas previamente para las distintas hora del día. Para este propósito, se ha formulado un modelo binivel el cual incorpora de manera simultánea dos problemas. Por un lado determina puntos de cambio en la intensidad comportamiento de la demanda, a través de diversos intervalos temporales denominados Time-Of-Day (TODs). Por otro lado, el modelo binivel también aborda el problema de control de tráfico en cada intervalo de tiempo. Esta propuesta ha sido resuelta mediante el uso de
algoritmos metaheurísticos. La eficacia del algoritmo híbrido SA∗+NM ha sido demostrada en una colección de datos sintéticos y en una colección de datos reales, además se han probado los métodos GA, SA+NM and PSO+NM. El algoritmo permite la aplicabilidad del SA∗+NM en el modelo binivel.

Una metodología secuencial y otra simultánea se han ilustrado sobre un problema real. Tal como era de esperar el método simultáneo logra mejores resultados que el método secuencial pero el coste computacional ha sido mayor. Se ha podido comprobar que la metodología secuencial, usada en la práctica, alcanza resultados satisfactorios.

Finalmente, la determinación automática del número óptimo de TODs ha sido estudiada mediante el uso de los métodos BIC, ΔBIC, un algoritmo PETE modificado y un enfoque orientado al problema. Se ha podido observar numéricamente que el criterio basado en el método BIC es conservativo y produce un número de TODs reducido en comparación con los otros métodos. El método ΔBIC produce un gran número de TODs. El método orientado al problema es sencillo de aplicar y alcanza valores esperados en el número de TODs. El algoritmo PETE modificado también obtiene una solución de acuerdo a lo previsible.

* Capítulo 6. **Una metodología para la clasificación dinámica de patrones de tráfico diarios.**

En este capítulo se analiza un sistema de control de tráfico urbano basado en dos etapas. En una etapa off-line, se aplica una técnica dinámica de análisis clúster para identificar los perfiles de tráfico diarios observados en la red. Esta técnica se basan en el algoritmo de las $k$-medias y en medidas de proximidad espectral, las cuales son aplicables a la transformación sobre datos basados en flujo, velocidad y densidad (ocupación).

El número óptimo de clusters se determina mediante la aplicación del criterio BIC y la razón de cambio de las medidas de dispersión. En la fase on-line el patrón de tráfico actual es determinado de forma dinámica. Un método de predicción de patrones es desarrollado sobre la base de un método de predicción enfocado en el periodo actual. Este método se ha denominado método de predicción no nerviosa.

Los experimentos numéricos se han llevado a cabo en dos conjuntos de datos. El primer conjunto se ha generado mediante simulación, usando los resultados obtenidos en el capítulo 4, el segundo es un conjunto de datos reales obtenidos a través de sistema PeMS (Freeway Performance Measurement Systems) el cual proporciona datos reales de las autovías del estado de California en EEUU. La experiencia computacional con datos sintéticos muestra que la distancia $d_C$ ofrece el mejor rendimiento, siendo capaz de reconstruir los clusters originales con una eficacia del 100%. Además, la experiencia computacional ilustra que la predicción no nerviosa mejora el método de predicción de referencia sobre el que se construye el sistema.
Los resultados sobre datos reales muestran que el procedimiento descrito determina clusters con un número de patrones significativo. Además la predicción no nerviosa del sistema muestra un comportamiento parsimonioso.

Este estudio aborda las fases esenciales del sistema. En futuras investigaciones sería interesante añadir un procedimiento para incorporar un conjunto de estados previamente clasificados, como pueden ser, días laborales, días de vacaciones, días no laborales, y así analizar el rendimiento completo del sistema.

Capítulo 7. Una propuesta para el diseño de Sistemas Inteligentes de Transporte para la gestión y el control del tráfico.

En este capítulo hemos introducido el uso de los TODs descrito en capítulos anteriores, para áreas urbanas dotadas de centros de control. Para ello hemos enriquecido el ITS desarrollado en el capítulo 6 con la gestión y control a través de TODs. Esto permite considerar el estado dinámico del tráfico como una secuencia finita de estados estacionarios. Su finitud permite, teóricamente, optimizar las estrategias de control en un contexto off-line. El ITS se ha enfocado hacia la recomendación de rutas y se han llevado a cabo distintas pruebas computacionales para ilustrar el rendimiento.

La experiencia computacional desarrollada muestra que el problema de efectuar dinámicamente recomendaciones de rutas es un problema difícil, intrínsecamente relacionado con el problema TAP dinámico. Se han aplicado infructuosamente propuestas genéricas de optimización. Creemos que para afrontar exitosamente este problema en redes reales se requiere de métodos orientados al problema en los que se explote todas sus características.

Hemos constatado en un problema test que el ITS tiene una capacidad en tiempo real de reconocer el estado del sistema. El aspecto negativo que hemos detectado es que la propia intervención del ITS puede incidir fuertemente en las observaciones y en las predicciones derivadas de ellas. Este aspecto debe ser analizado en futuras investigaciones.

8.4 Futuras líneas de investigación

La realización de esta tesis doctoral es un ejercicio de aprendizaje, de aprender a investigar. No se trata en sí de un fin, de una llegada, sino de un punto intermedio desde donde partir. Es por esto que las futuras líneas de investigación no se focalizan como mejoras de las imperfecciones que tiene la tesis sino como puntos lejanos en el horizonte donde se desea llegar.

1. Cluster dinámico. En esta tesis se han analizado técnicas para
8.4. Futuras líneas de investigación

el descubrimiento de patrones dinámicos. La característica esencial es que el método monitoriza varios atributos a lo largo del tiempo (varios sensores). En la literatura se ha estudiado profusamente este problema para series temporales univariantes. Pensamos que puede ser interesante la aplicación de estos métodos a otros campos, no necesariamente en el dominio del transporte. Hoy por hoy el denominado Analytics in a Big Data World está adquiriendo una gran importancia en la comunidad investigadora y estas técnicas que incorporan la componente dinámica son sin duda relevantes.

2. **Transporte público.** El ITS propuesto en el capítulo 2 tiene un diseño conceptual. Esta característica le dota de gran flexibilidad y permite ser adaptado a problemas de transporte público. Estos sistemas tienen características propias, diferentes a las redes de tráfico, y son las que se deben incluir. A modo de ejemplo, los sensores no registran vehículos sino pasajeros; existen nueva fuentes de información como los teléfonos móviles, ventas de tickets, etc; existen incidentes en la flota de vehículos; no se recomiendan rutas sino líneas (hiperrutas); etc. Pese a estas diferencias consideramos que los siguientes tres pilares esenciales sobre los que se soporta el prototipo de ITS presentado siguen siendo válidos en un contexto de transporte público: i) Un conjunto de patrones de movilidad que se repiten reiteradamente en el tiempo, ii) Un conjunto de técnicas de optimización que permiten obtener la mejor respuesta del sistema a cada uno de dichos patrones y iii) Un conjunto de técnicas de minería de datos que permiten inferir en tiempo real el estado actual del sistema.

3. **Asignación dinámica de tráfico** Una extensión natural del capítulo 4 es embeber el modelo CDNL dentro un modelo de elección de rutas, esto es, abordar el llamado *dynamic traffic assignment problem*. La literatura muestra su extremada complejidad y sin duda es un gran reto. Hoy por hoy se han desarrollado grandes facilidades computacionales y son las que se desean usar. Los algoritmos de descomposición se han aplicado exitosamente al TAP estático. En concreto el denominado *disaggregate simplicial decomposition algorithm* presentado por Larsson and Patriksson (1992) explota exitosamente la descomposición por pares origen-destino. Los esquemas tipo Jacobi para el TAP estático explotan las técnicas de computación paralela. Estas ideas se intentarán explorar en un contexto dinámico. El modelo de simulación se basa en el concepto de paquete que está caracterizado por su par origen-destino, la ruta seguida y el instante de salida. La idea clave que se debe afrontar para poder paralelizar la resolución del DTA, es cómo poder considerar el efecto (colas en los arcos) de un conjunto de pares origen-destino no analizados en un procesador determinado sobre el conjunto de pares origen-destino analizados en dicho procesador.

4. **Computación basada en agentes.** Los métodos presentados en esta tesis doctoral para el desarrollo de ITS para la gestión y control del tráfico han sido desarrollados con la intención de adaptarse en contextos reales. Consideramos que los sistemas basados en agentes
son un enfoque capaz de ser adaptado a nuestro trabajo, gracias a su posibilidad de desarrollo de sistemas distribuidos de gran escala. El motivo principal del aumento en los enfoques basados en agentes se deben a su capacidad de descomponer un sistema, en múltiples agentes que interactuando unos con otros son capaces de alcanzar objetivos globales. Actualmente los denominados Multiagent-Systems (MAS) han sido aplicados en problemas de tráfico, principalmente en los temas de modelado, simulación, control de la congestión y control inteligente de tráfico. Además estas propuestas están alineadas con los objetivos perseguidos en las denominadas smart-cities, que hoy en día representan gran interés para la comunidad científica debido a su búsqueda por mejorar la funcionalidad de las grandes ciudades a través del uso de tecnologías de la información.
Bibliography


Bibliography


McQueen, B. and McQueen, J. (1999). *Intelligent transportation systems architectures*. Artech House Publishers, Norwood, MA, USA.


A

Dynamic capacity given by a square wave

In this section we will analyse a link capacity \( \beta_a(t) \) given by a square wave as shown in Figure A.1. We introduce the following notation.

- \( w_a \) Wave amplitude associated with the square wave \( \beta_a(t) \).
- \( f_a \) Period of time in which the link capacity has the maximum value.
- \( T_0 \) Initial instant of the simulation.
- \( t_0^a \) Instant lesser than the initial instant \( T_0 \) in which the square wave begins its maximum capacity, (see Figure A.1).

The solution of equation (4.41) yields the value of \( \tau_i \) and coincides with the moment in which the shaded area below the wave, is exactly \( \Delta x \). This value is given by the expression:

\[
\tau_i := \begin{cases} 
\phi_a' + \frac{\Delta x}{\beta_a}, & \text{if } \Delta x \leq (z_a - \phi_a') \beta_a; \\
(N_a + M_a + 1)w_a + s_a f_a - t_0^a, & \text{otherwise},
\end{cases}
\]  

where

\[
M_a := \left[ \frac{\phi_a' - T_0 + t_0^a}{w_a} \right] 
\]
\[
z_a := T_0 - t_0^a + M_a w_a + f_a 
\]
\[
N_a := \left[ \frac{\Delta x - (z_a - \phi_a') \beta_a}{\beta_a f_a} \right] 
\]
\[
s_a := \frac{\Delta x - (z_a - \phi_a') \beta_a}{\beta_a f_a} - N_a 
\]
and $[\cdot]$ is the integer part of a number. Note that in the paper it is assumed that the initial instant of the simulation is $T_0 = 0$.

Figure A.1: Link capacity function $\beta_a(t)$ given by a square wave.
MAT research group
Universidad de Castilla-La Mancha

Ciudad Real, June 2015