UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Tecnica e Gestione dei Sistemi Industriali

Master's Thesis in Management Engineering:

"High-level Architecture and Compelling Technologies for an Advanced Web-based Vehicle Routing and Scheduling System for Urban Freight Transportation"

Supervisor:
Ch.mo Prof. ALESSANDRO PERSONA
Università degli Studi di Padova

Master Candidate:
LEONARDO CREATO
Matr. 1156580

Co-supervisors:
Ch.mo Prof. ILIAS TATSIOPOULOS
National Technical University of Athens - S.I.M.O.R.

Academic Year 2018 – 2019

UNIVERSITÀ DEGLI STUDI DI PADOVA

&

NATIONAL TECHNICAL UNIVERSITY OF ATHENS

Dipartimento di Tecnica e Gestione dei Sistemi Industriali

Master's Thesis in Management Engineering:

"High-level Architecture and Compelling Technologies for an Advanced Web-based Vehicle Routing and Scheduling System for Urban Freight Transportation"

Supervisor:

Ch.mo Prof. ALESSANDRO PERSONA

Università degli Studi di Padova

Master Candidate:

LEONARDO CREATO

Matr. 1156580

Co-supervisors:

Ch.mo Prof. ILIAS TATSIOPOULOS

National Technical University of Athens - S.I.M.O.R.

Academic Year 2018 – 2019

Acknowledgements

I would like to thank my family who have always believed in me and supported me unconditionally, my girlfriend for the immense love and patience that have given me the serenity to face and overcome every difficulty, my friends for the carefreeness, laughter and joys shared.

A special thank you to those who supported me in my Erasmus experience in Athens, in particular to Prof. Andrea Lazzaretto (DII) who allowed me to pursue the most beautiful adventure of my life, to Prof. Ettore Bolisani and all the university staff, especially Prof. Alessandro Persona for agreeing to be supervisor of this master thesis project abroad.

A heartfelt thank you to Euripides Kechagias, Giorgos Papadopoulos and Eleftheria Skordalaki who supported me in carrying out the thesis at N.T.U.A.

This work crowns a study and life path that I am proud of and which I will take with me forever. If I have managed to achieve this wonderful goal and to be the man I am today it is also thanks to all of you.

L.C.

Abstract

In an increasingly globalized and competitive market, companies continually seek to find efficient business solutions at a reasonable cost. This phenomenon of tight competition affects companies operating in logistics and transportation sector as well. In particular, when it comes to logistic distribution in urban areas (urban logistics), it is necessary to refer to the routing and scheduling process, which in its multiple variants is called VRSP (Vehicle Routing and Scheduling Problem).

The need to increase the efficiency and competitiveness of a company operating in this field implies, consequently, to seek for effective solutions to implement the distribution process, especially in a business framework where, in recent years, the service's level is increasingly crucial and costly to achieve for companies.
This, in practice, translates into the search for systems that can support managers, planners and other actors involved in the process, both from a decision-making and a strictly operational point of view. In particular, extremely important variables need to be taken into account in an urban routing environment. These variables at stake generate a mathematical complexity upstream and a massive flow of heterogeneous information, ultimately an intrinsic decision-making complexity. Either way, the complexity is not a mathematical-computational matter only, in fact, the growing complexity of data (Big data) and intra and inter-corporate relationships are another essential aspect not to be overlooked.

On the other hand, the technological progress of recent years allows us a certain freedom in design a system, which until a few years ago was inconceivable. Exploiting for example Cloud computing, wireless technologies, smart devices, latest generation traffic data collection systems and integration with high-performing software technologies. The search for a more efficient routing and scheduling, the improvement of service's level, the increasing complexity of real-world distributive contexts and articulated inter-and intra-company information, are contingent variables that generate the need for a system's architecture that may take into account all these aspects in a holistic, innovative, scalable, reliable manner. This solution aims to be capable to

capture this complexity in its entirety. Hence, new technologies and a lucid awareness of involved actors and infrastructures, provide the basis to create a more efficient routing and scheduling architecture for enterprises.

In a nutshell, the aim and focus of this thesis is to propose a high-level and web-based software architecture, in which are identified infrastructures and systems that can communicate and elaborate sensitive information and heterogeneous data, to the routing software; how they relate to each other and how new technologies can contribute to the system, enhancing its competitiveness. The architecture also aims to identify the actors who generate and request such information, the inputs and outputs of the process and how this informative flow develops within the system.

Sommario

In un mercato sempre più globalizzato e competitivo, le aziende cercano continuamente di trovare soluzioni di business efficienti a un costo ragionevole. Questo fenomeno di forte concorrenza colpisce anche le aziende operanti nel settore della logistica e dei trasporti. In particolare, quando si tratta di distribuzione logistica in aree urbane (urban logistics), è necessario fare riferimento al processo di routing e scheduling, che nelle sue molteplici varianti è chiamato VRSP (Vehicle Routing and Scheduling Problem).

La necessità di incrementare l'efficienza e la competitività di un'impresa operante in questo settore implica, di conseguenza, la ricerca di soluzioni efficaci per l'attuazione del processo di distribuzione, soprattutto in un contesto imprenditoriale in cui, negli ultimi anni, il servizio livello è sempre più cruciale e costoso da raggiungere per le imprese.

Questo, in pratica, si traduce nella ricerca di sistemi in grado di supportare manager, planner e altri attori coinvolti nel processo, sia dal punto di vista decisionale che da quello strettamente operativo. In particolare, è necessario tenere conto di variabili estremamente importanti in un ambiente di routing urbano. Queste variabili in gioco generano una complessità matematica a monte e un flusso massiccio di informazioni eterogenee, conseguentemente una complessità decisionale intrinseca. In entrambi i casi, la complessità non deve essere vista soltanto come una questione matematico-computazionale, infatti, la crescente complessità dei dati (Big data) e le relazioni intra e inter aziendali sono un altro aspetto essenziale da non trascurare.

D'altra parte, il progresso tecnologico degli ultimi anni ci permette una certa libertà nella progettazione di un sistema, che fino a pochi anni fa era inconcepibile. Sfruttando, ad esempio, il cloud computing, tecnologie wireless, smart devices, sistemi di raccolta dati sul traffico di ultima generazione e l'integrazione con tecnologie software ad alte prestazioni. La ricerca di routing e scheduling più efficienti, il miglioramento del livello del servizio, la crescente complessità dei contesti distributivi reali e articolate informazioni inter e intra aziendali, sono variabili contingenti che generano la necessità di un'architettura di un sistema che possa prendere in considerazione tutti questi aspetti

in modo olistico, innovativo, scalabile e affidabile. Questa soluzione mira ad essere in grado di catturare questa complessità nella sua interezza. In tal modo, le nuove tecnologie e la lucida consapevolezza degli attori e delle infrastrutture coinvolte, forniscono le basi per creare un'architettura di routing e scheduling più efficiente per le imprese.

In sostanza, l'obiettivo e il focus di questa tesi è quello di proporre un'architettura software di alto livello e web-based, in cui si vogliono individuare infrastrutture e sistemi che possano comunicare e elaborare dati variegati e sensibili ad un software di routing; come si relazionano tra loro e come le nuove tecnologie possono contribuire al sistema, potenziandone la competitività. L'architettura ha anche lo scopo di identificare gli attori che generano e richiedono tali informazioni, gli input e output del processo e come questo flusso informativo si sviluppa all'interno del sistema.

Table of contents

List of Figures

List of Tables

Chapter 1: Introduction

The proposed research work has been developed at the *National Technical University of Athens (N.T.U.A.)* through the *Erasmus+ Study programme* (figure 1.1) for Master Thesis finality, for a total duration of six months (September 2018 - March 2019). In particular, the thesis was proposed and written at the department *Sector of Industrial Management and Operations Research (SIMOR)*, which activates in educational and research activities regarding the design, planning and control of production systems i.e. production management.

Figure 1.1 – Erasmus + Programme Logo

This thesis aims to be a contribution to a broader and ambitious initiative co-financed by the *European Union* and *Greek national funds* through the *Operational Program Competitiveness, Entrepreneurship and Innovation*, named "*Research-Create-Innovate"* (project code: T1EDK-00527). The project, named with the acronym *"SMARTRANS"* is developed by the university National Technical University of Athens (NTUA) and in particular by the Operational Research Laboratory (ORL), Athens Information Technology (AIT) together with partner logistic, consulting and software companies. The target market is the Greek logistic distribution industry, but the beneficiaries are not limited to Greece, as the output is suitable wherever urban freight transportation plays a key role on economy, urban society and businesses, hence is globally valuable.

The *"SMARTRANS"* roadmap is articulated and structured in specific sections, namely 8 work packages (WP), whose focal points, with estimated duration, are briefly presented below and summarized in figure 1.2***Figure 1.2***:

- **Work Package 1: Project Management** (Duration: 36 months). This work package includes project management and risk management. It spans throughout the whole project;

- **Work Package 2: State-of-the-art Review** (Duration: 4 months). The state-of-the-art review will lead not only to a selection of algorithms which are well suited for the specific problem under study, but also a research on emerging, suitable and compelling technology that may enhance its performance and efficiency;
- **Work Package 3: Requirements Engineering** (Duration: 6 months). It aims to perform requirements engineering methodology comprising data, technical and process requirements including its integration with further focus on enterprise systems integration (ERP, WMS or GIS), pollution issues, drivers' regulations;
- **Work Package 4: Routing and Scheduling Algorithms** (Duration: 8 months). The specific algorithms selected in WP2 will be examined in-depth within practical problems scenarios analysed in WP 3;
- **Work Package 5: Travel Times Calculation Methods** (Duration: 8 months). This work package will conclude to methods and techniques for the determination of travel times in traffic congestion situations, through the design of forecasting algorithms, which may generate input for both static and dynamic purposes. Moreover, knowledge-based approaches will incorporate techniques for collecting, consolidating, elaborating and managing data from various online traffic data providers;
- **Work Package 6: System Development** (Duration: 20 months). The specialized algorithms of WP4 and travel times calculation methods of WP5, will be implemented in an integrated web-based software solution for static and dynamic routing suitable for enterprises and their systems.
- **Work Package 7: Demonstration and System Validation** (Duration: 6 months). The prototyped system will be demonstrated and tested in a real business distribution cases and environment. The feedback will be utilized to fine tune and improve the algorithms and the related software.
- **Work Package 8: Dissemination and Exploitation** (Duration: 30 months). It is the work package responsible for providing project results with a wide spectrum of visibility (through workshops, scientific journals, press, web site) and prepare the potential market for its exploitation.

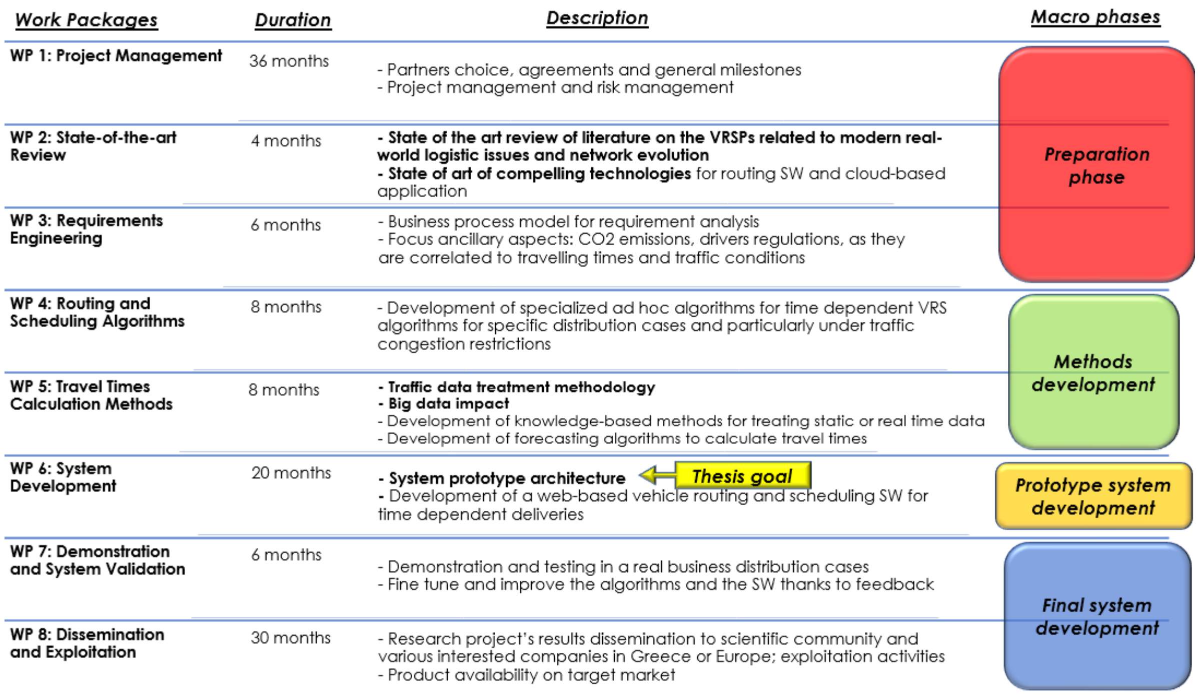

Figure 1.2 - "SMARTRANS" roadmap and some references of the thesis.

The project started in 2018, with the phase WP1, therefore it is still in its early stage, also, as shown in figure 1.2, its completion is projected in medium-long term horizon, since it is necessary the collaboration and the coordination of several complex entities and a specific Leader is assigned to each specific Work Package (WP). For this reason the student has been proposed a research thesis that has the dual purpose of, on the one hand, contributing to the aforementioned phase preparation, but at the same time allows the student to show his own and original idea of system's architecture in perspective, which anticipates the actual proposition of the prototype, hence the title "High-level Architecture and Compelling Technologies for an Advanced Web-based Vehicle Routing and Scheduling System for Urban Freight Transportation". In essence, the thesis aims to reproduce the project in a small scale and over a short period, accordingly with the limited Erasmus stay and the actual resources commitments. Moreover, given the current inapplicability of the project, the thesis on hand may be defined as a theoretical research and bibliographic (monographic) work. In a nutshell, contribution to project constitutes mostly the *preparation phase* of the project, but does not renounce to a first presentation of a possible *prototype system architecture*.

The research matrix never falls in the proper quantitative level since the actual development of the quantitative phases, such as Requirements Engineering (WP3) Vehicle Routing and Scheduling Algorithms (WP4) and Travel Time calculation methods (WP5), is subordinate to the early WPs and they have also been assigned to specific pool of specialized resources. Therefore, the approach of thesis research is modelling, that is high-level architecture (module-based).

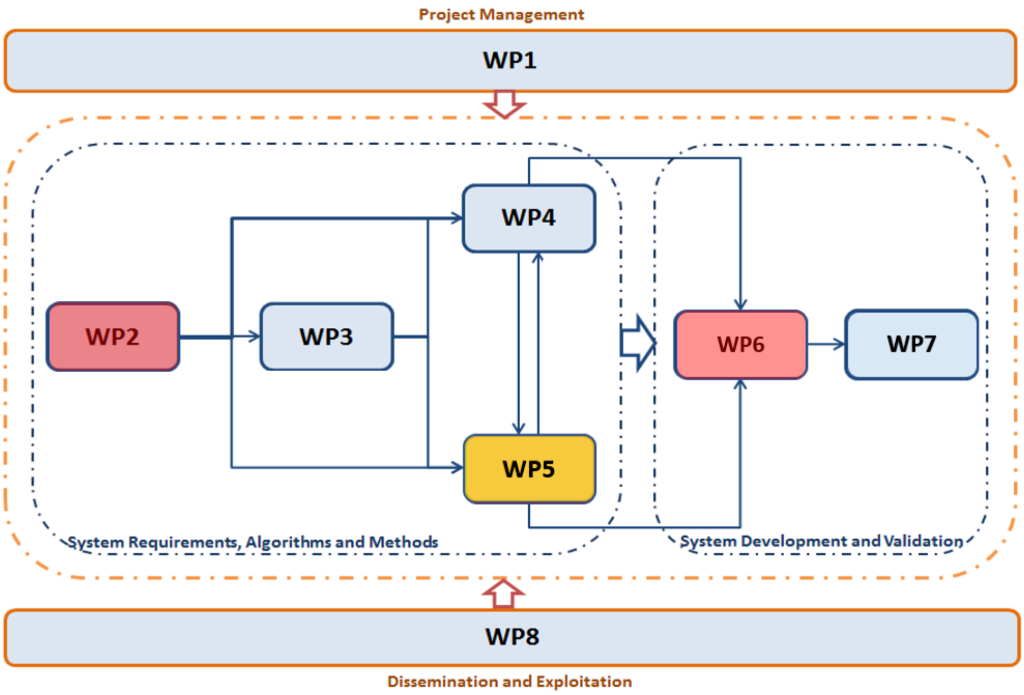

Figure 1.3 - Project pert chart with comparison between the actual (full, in blue) conceptual path and the thesis (partial, in orange and red) conceptual path.

Even if the thesis has not officially or formally been assigned to any of these work packages, it may be referred to the *"SMARTRANS"* project as the research work may help the initiative itself contributing to its scientific research. To understand more clearly it has been tried to refer the research work done in the thesis to the WPs of the project. The figure 1.3 shows the conceptual (shortened or partial) path followed by this thesis (red and orange WPs) compared to the actual (full) path (blue WPs). It is shown that the thesis focus could be identified in the WP2, WP5 and WP6. WP2 (discussed in the Chapter 2 and 4) includes state of art review, stood as state of art of vehicle routing problem, with a proper introduction of operational research concept, and the state of art of compelling technologies, which consists in the very core of the thesis. The WP5 (discussed in the

Chapter 3), is orange coloured to indicate that it has been reduced to the sole description of traffic (big) data treatment process instead of the elaboration of forecasting algorithms and knowledge-based approaches for traffic data, ad hoc for the SW, that, as said, will instead be studied in due time by a dedicated team. Finally, the WP6 (discussed in the Chapter 5), also reduced, take advantage of the considerations WP3 and WP5 to gather the basis to represent a prototype high-level system architecture.

Basically, attempts have been made to find the minimum conceptual path that allow to propose a model of architecture, which may be original, innovative and sustainable. This latter will achieve the true final software system structure only subsequently, upon completion of all stages presented.
The difference in objectives between the thesis and the full project is that the former proposes a *high-level system model architecture (module-based)*, the latter propose a *software system* holistically declined.

The main roles involved, who somehow contribute or are directly influenced, within this system's architecture might be:

- **Transportation planners (Logistic company)**: they are the first to receive the output of the system's architecture and, notably, from the software. Once analysed the proposed solution they might modify or confirm it; this constitutes the initial routing of vehicles and scheduling of deliveries;
- **Drivers**: they physically operate the assigned route. Specific devices are assigned, where they receive the information necessary to complete the route. Such devices must be user-friendly and allow manual modifications of the path.
- **System providers**: besides being the developers of the software functions (IT company), they are also responsible for the support, maintenance and updating of the routing software under discussion.
- **Traffic data providers**: they are responsible for the geo-localization data, of which the software, as well as the drivers, need. The role of these providers is all the more important the more dynamic the system wants to be.

1.1 Framework

The proposed thesis fits in a logistic and distributive business context, but, more widely, in supply chains framework that occur to compete in increasingly globalized and tightened business contexts, in which the technological process and alignment between different players of the supply chain are key determinants for maintaining a dominant role and competitiveness on the market. Natural consequence of these responsibilities towards the supply chain, and obviously of the production process, is that also the terminal part of this chain, which consists precisely in the distribution, must be more than ever efficient, convenient and effective. In particular, the focus of the thesis will be precisely on this last section of the business process: the distribution, which declines in the study of routing and scheduling problems, with all the directly and indirectly related business complexities arising.

Specifically, it will seek to design a web-based routing and scheduling software system architecture which may take advantage of the compelling and boosted technologies nowadays available to build an innovative and versatile platform delivered as SaaS. This will help managers, planners, drivers to operate in an efficient and holistic distribution management, with the ultimate purposes of improving enterprises performance and customer service, which is strongly conditioned by the complex urban environment. In particular logistic market trend requires more and more dynamic orders, higher demand, low batch shipments, variable productions, shorter lead times, rapid market changes, this translates in higher service standard to promise and to higher costs to bear.

This context, in fact, turns out to be a highly dynamic, variable and binding environment, which therefore requires considerable forward-looking, decisional and operational efforts.

1.2 Problem formulation

When it comes to innovative freight transportation paradigm, we are required to take into consideration the contribution that emerging and mature technologies, software, network and enterprise related, may provide to the routing software idea, as well as we should be aware of infrastructure technological advances, stood as road network, mobile network and ICT technologies. At the same time we should not forget that to achieve the output, the delivered routing service, we inevitably need to refer to the concept of VRSP (Vehicle routing problem), namely, the routing and scheduling of vehicles in an urban network context, which, with predefined routes, determined in advance (statically) or dynamically, must complete tasks, such as deliveries or pickups, at different drop points of the network, respecting constraints and optimizing a cost and a service. The generation and planning of these routes are, of course, a fundamental aspect as well as it is highly sensitive to a multiplicity of variables.

Since the aim is to seek for innovative technological solutions that can be inserted into a system architecture and that can improve the overall performance of the current routing systems, the identification of the variables that impact on it and the systems from which it draws cannot be ignored.

In particular, the variables at stake impact on the routing and scheduling system from two fronts: on one hand we have real variables deriving from the road network in which the company operates, on the other variables of a company nature, deriving from the company systems that provide the information of the customers' orders, warehouse, availability and resources of the company. Some examples are:

- The total daily time available;
- Customer importance and time windows
- Product constrains
- Departure time
- Loading and unloading times;
- Working hours Drivers;
- Fleet constrains (drivers shifts, capacity, availability, load, speed),
- Traffic congestions;
- Access restrictions (environmental constraints);
- Unforeseen events (incidents, weather issues...)

In recent years the interest has shifted in particular on the variants of routing that meet the recent necessities of huge amount of relevant and sensitive data, dynamism and real-time issues, as well as ancillary aspects. In terms of service level, for example, this translates into more freedom to the customer (or request of flexibility from the customer) which means generations or modifications of order in real-time, perhaps with a route already underway. It is quite clear that this leads to a dynamic issue, with equally dynamic information. All these considerations, inevitably force the system to be considered in all its entirety, considering both fronts and exploiting as much as possible systems capable of providing accurate, secure, global, real-time information. Only considering this wide spectrum of variables it is possible to create a truly innovative, original and solid system's architecture that can give consistency and continuity to the required goal.

Hence, this system will be an innovative architecture, applicable to any logistics and distribution company, that will eventually realize in a software solution which will absorb the burden usually handed by the company itself. Hence, deliveries planning and execution will be more efficient and effective and user companies are expected to improve their customer service level, which is a crucial success business factor in the current competitive business environment.

1.3 Thesis structure

The literature of the Virtual Routing Problem is extensive and well-documented, but the innovative and original nature of the project requires a new selection and revision of the approaches, technologies, models and algorithms that can represent a viable solution to the objective of the study and its unedited nature. Therefore the ideological path of the thesis develops, at first, through a revision of the key concepts of the study (algorithms, models and technologies) to comprehend the "catalogue" of choices that we dispose nowadays and we can draw from, then focus on specific aspects of concern from each field, which will be combined appropriately afterwards to develop the desired output.

The Thesis is divided in 2 main sections: a first descriptive part which contains literature and scientific review (chapter 2, chapter 3, chapter 4), and a second presentational part which contains results and discussion (chapter 5, chapter 6). Specifically, the logical thread of the thesis develops as follows.

The intent of **Chapter 2** is to revise the current state of the art of VRP, to capture the concepts and variety of approaches more suitable for the innovative nature of the project. Therefore, **Chapter 3**, aims to show the processes of traffic data treatment, namely the complexities of big data and data combination, to review the methods and sources of traffic data collection and to present the traffic data providers dynamics. In the **Chapter 4**, a catalogue of the new applicable and compelling technologies is presented and their features, advantages, drawbacks and possible contribution to the final output of the project is assessed. **Chapter 5** summarize, revise and elaborate the information gathered from the previous chapters, seeking to identify the data flow and processes, and unify the best scientific and technological solutions, deemed more valid, in a proposal for innovative and holistic software architecture, operating in the cloud environment, that aims to be highly competitive on the market. Finally, **Chapter 6** provide final conclusions, statement, issues and future developments concerning the obtained results.

Chapter 2: Theory of Vehicle Routing Problem and State of Art

This chapter provides an overview of State-of-the-art "Routing Problem": from static theories (Off line), up to the development of the Dynamic routing's (Real-time) problem with related theories, which adds the real-time constraint to the already numerous constraints that make the problem itself this complex.

2.1 Operational research approach for VRP

2.1.1 Graph theory and OR basics

Although the purpose of the thesis is not to deepen the purely mathematical and theoretical aspect of the VRP's theory, it is advisable to introduce at least some definitions necessary to characterize the nature of the problem and its variants, in order to allow a full understanding of the considerations that will be proposed in the following chapters. The decision-making problems related to VRP are strongly connected to the language of graph theory, both for their physical representation and for their mathematical formulation, some examples can be mass transport problems, organization of the timetables or service localization.

Besides being particularly interesting for the immediacy of its graphic form, the graph theory can also be described by a mathematical formulation: linear, nonlinear or mixed linear programming.

The linear programming is generally the milestone of the operational research as it is the most used programming model and its purpose is to find, starting from specific instances for each problem, a region of optimal solutions (figure 2.1) that respect the constraints imposed by the problem itself. However, it should be stressed that it is not always possible to get a solution, as the solution might be non-allowable, unlimited, or the set of solutions might be too vast.

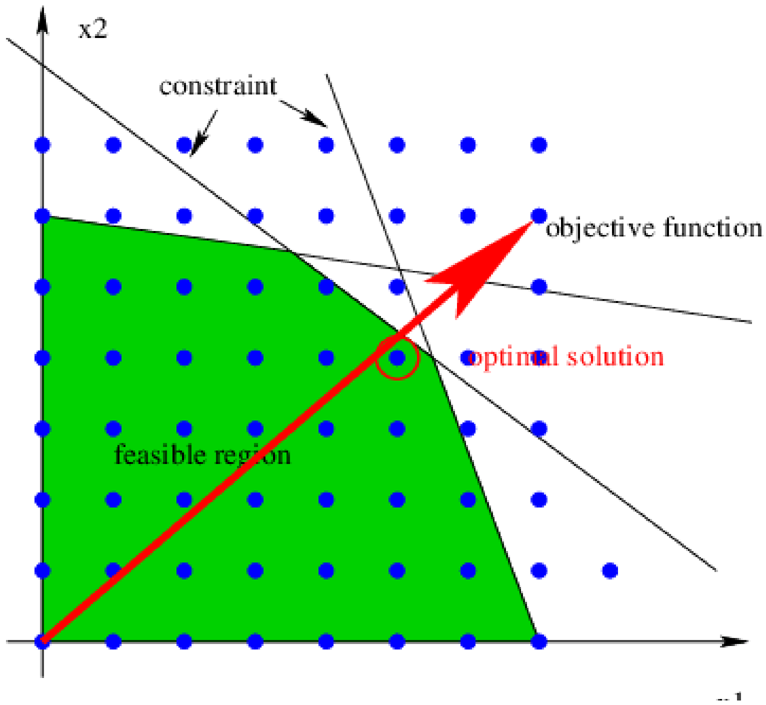

Figure 2.1 - Feasibility region of an ILP. Source: Helwig, S., Hüffner, F., Rössling, Weinard, M., Chapter 3. Selected Design Issues, Conference: Algorithm Engineering: Bridging the Gap between Algorithm Theory and Practice, 2010, p. 58-126

The integer linear programming (ILP), on the other hand, is a subcategory of the abovementioned LP (figure 2.1). The problems that are the object of the ILP are all those that have, among the constraints, the entirety of the variables and the objective function (although a relaxation of this constraint is admissible), and because of this peculiarity they are particularly suitable for representing real-world situations, related to Routing logistics. Most of the integer linear programming problems are demonstrated to be NP-hard; on the contrary, in the continuous an optimal solution is obtainable in polynomial time. For the problems of ILP there are several categories of resolution, including exact methods, heuristics and metaheuristics. The first ones, among which we cite the Branch and Bound method, provide excellent global, the latter ones a sub optimal solution. However, any further consideration concerning resolution method is so far suspended to be subsequently deepened.

In general, the interest of the operational research is to find, through the linear programming, a globally optimal solution.

In this particular study the theory of degrees is a crucial component for the design of effective and innovative algorithms that have to be implemented in the software, as it allows to achieve a mathematical formulation of the problem with the aim of finding an optimal, or at least an efficient solution.

Hence, it is thus possible to formulate operative research problems through the graph theory (figure 2.2 shows main types of edges found in a network), for example problems of maximum flow along the graph, in case the arches are associated with numerical values, or problems of minimum paths, for which the theory offer numerous algorithms.

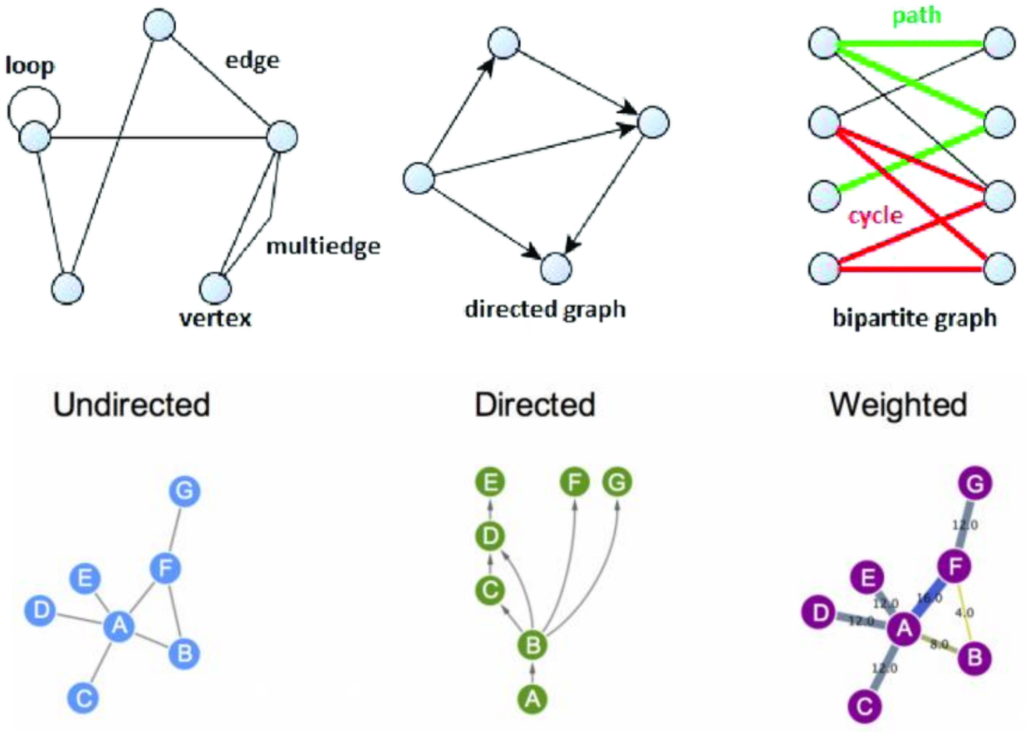

Figure 2.2 - Graphs examples and properties. Adapted from: https://www.ebi.ac.uk/

2.1.2 Resolution methods

Although the literature of the resolutive methods is wide, there is no uniquely shared classification. The proposal integrates several of the most authoritative and adopted classifications found in the literature and is enriched by the student's personal university experience during the related courses and it is shown in figure 2.3.

Two macro categories can be identified:

- Exact methods
- Approximation algorithms

While the first category has specific peculiarities strictly related to the case study on hand, the second category can be further decomposed into:

i. Heuristics
ii. Metaheuristics

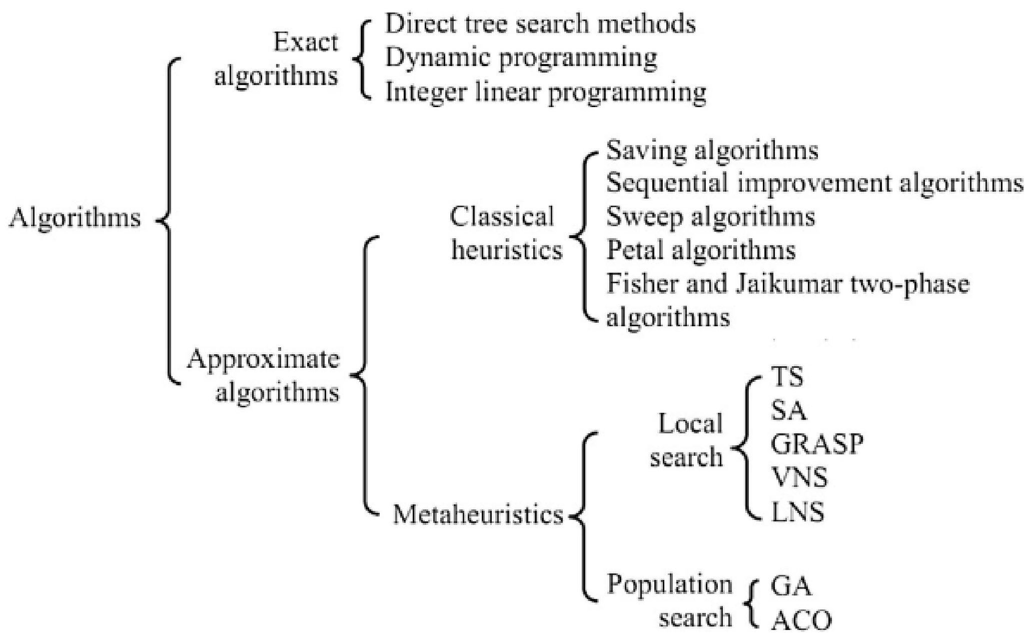

Figure 2.3 – Resolution methods for VRP and their relation. Source: Lin, C., Choy, K.L., Ho, G.T.S. et al., Survey of Green Vehicle Routing Problem: Past and future trends, 2014, p. 1118-1138

EXACT METHODS

In summary, the advantages of the exact methods, compared with the approximated, ones are remarkable and can be summarized as follows:

a) Valuable information on lower and upper limits of the optimal solution could be obtained (even if the algorithm crashes before the completion of the research procedure);
b) If an exact algorithm succeeds you will get a certainly optimal solution;
c) The exact methods allow to examine and prune the research parts in which no optimal solutions can be found.

However, in spite of the advantageous characteristics, the exact methods also have some considerable disadvantages:

a) Often the generation of such algorithms could indeed require huge development time by experts in Integer Linear Programming;
b) The memory needed for exact algorithms can also be very large and lead to the early abortion of a program;
c) The size of instances that are resolvable in practice is rather limited and the time needed for computing strongly increases with the size of the instance;
d) For many cases, the specific exact algorithms for a given problem are often hardly extendible, in case of modification of some details in the formulation of the problem itself.

Some examples of the most applied exact algorithms (for ILP) are: Branch-and-bound method; Cutting plane algorithms; Lagrangian relaxation methods; Dynamic programming methods.

APPROXIMATION ALGORITHMS

The approximations algorithms, namely heuristics and metaheuristics, have the following advantages:

a) In many cases they easily adaptable to slight variants of a problem, thus more flexible than exact methods;
b) They can examine a great amount of possible solutions in a time reasonably short computational time;
c) They are usually easier to implement and understand than exact methods;
d) They represent the algorithms that are most used for the resolution of a considerable number of problems.

However, there are also several disadvantages:

a) They have no well-defined ending criteria;
b) Sometimes optimality cannot be demonstrated;
c) Generally, they cannot reduce demonstrably the research's space.

HEURISTICS

An alternative way to the afore mentioned methods, and that may be the only viable whether:

- The optimal resolution of the problem can be too costly in terms of time or processing capacity;
- The optimum resolution of the problem may be impossible;

is to seek for a heuristic-type solution. That means these solutions are obtained by applying an algorithm conceived in such a way that, for certain size of heuristic's instances, it is able to derive a solution roughly close to the optimum, but without a priori guarantees on optimum's proximity.

Despite this property cannot be systematically verified beforehand, it is often a solution available in a reasonable time. Hence, the result is achieved by trying to balance the objectives of greater optimization, speed of execution, completeness and accuracy.

The main heuristic methods include: Greedy, Beam search (constructive), Local research.

METAHEURISTICS

Has to be noticed that classical heuristics, by its very nature, do not allow the intermediate solution to deteriorate during the process of finding a better solution (optimal, where possible). Consequently, they often get trapped in local optimum neighbourhood. Metaheuristics, by contrast, include mechanisms that avoid getting trapped in local optima and in recent years they have become increasingly important. The structure and the basic idea of each metaheuristics are substantially fixed, but the implementation of the various components of the algorithm depends on the individual problems. All the different metaheuristic approaches can actually be seen as homogeneous, as generalizations of a single fundamental approach, which is that of local research. Some of the most widely used meta-heuristics are Randomized multistart, Simulated annealing, Tabu Search, Genetic algorithms.

2.1.3 Computational complexity's problem

For computational complexity in operational research and in general optimization problems, we mean that magnitude that measures the amount of resources needed by an algorithm to solve a given problem.

This magnitude is paramount because it is generally associated with the processing time required, compared to the parameter of problem's size N: notably, it evaluates how much the calculation time increases as the instances dimension of the problem gets larger. It is noticeable that computational complexity is a crucial feature of an algorithm: since, as the size of the instance increases, if the calculation time required by the calculator to provide a solution raise excessively, this may cause some practical problems, especially when the algorithm, thereby its instances, is designed from real world situations, which are usually rather complex and broad. For this reason, in case of facing a problem, for which an algorithm able to provide a solution in polynomial time is not known, other methods (non-exact) are advisable, quicker in calculating an acceptable solution. In fact, when faced with an NP-hard problem, one frequently resorts to heuristic, metaheuristic or approximate procedures to obtain near-optimal solutions in lieu of seeking optimal solutions.

2.2 VRP definition and determinants

2.2.1 Definition of Vehicle Routing Problem

Before starting any consideration, it is essential to get clarity about what is meant by Vehicle Routing (a visual idea is shown in figure 2.4). It should be taken into account that although the literature is broad, there is no univocal definition to define this subject because of its breadth and variety. Nevertheless, a combination of some of the most authoritative and exhaustive definitions of VRP found in the literature is proposed.

The definition that clarifies the most Vehicle Routing's concept from a qualitative point of view is that one suggested by [1] who defines VRP as "*the problem of determining the shortest route from a vehicle that starts from one depot to multiple destinations to meet customer needs of n. Where the vehicle has a certain capacity, every vehicle starts from the depot and returns to the depot. Each customer can only be visited once. The purpose of VRP is to meet the needs of every customer with minimal cost. The cost is proportional to the total distance travelled by all vehicles so the VRP determines the shortest distance.*

The following are some of the constraints or limits that must be met in the VRP:

i. *The vehicle route starts from the depot and ends at the depot.*

ii. *Each customer must be visited once with one vehicle.*

iii. *Vehicles used are homogeneous with a certain capacity so that consumer demand on each route traversed should not exceed the capacity of the vehicle.*

iv. *If the vehicle capacity has reached the limit, then the next consumer will be served by the next shift.*

General purpose of VRP:

a) *Minimize the distance and costs associated with vehicle use.*

b) *Minimize the number of vehicles needed to service the demand of all consumers.*

c) *Balancing routes in terms of travel time and vehicle load.*

d) *Minimize services that are less satisfactory to consumers, such as delays in delivery."*

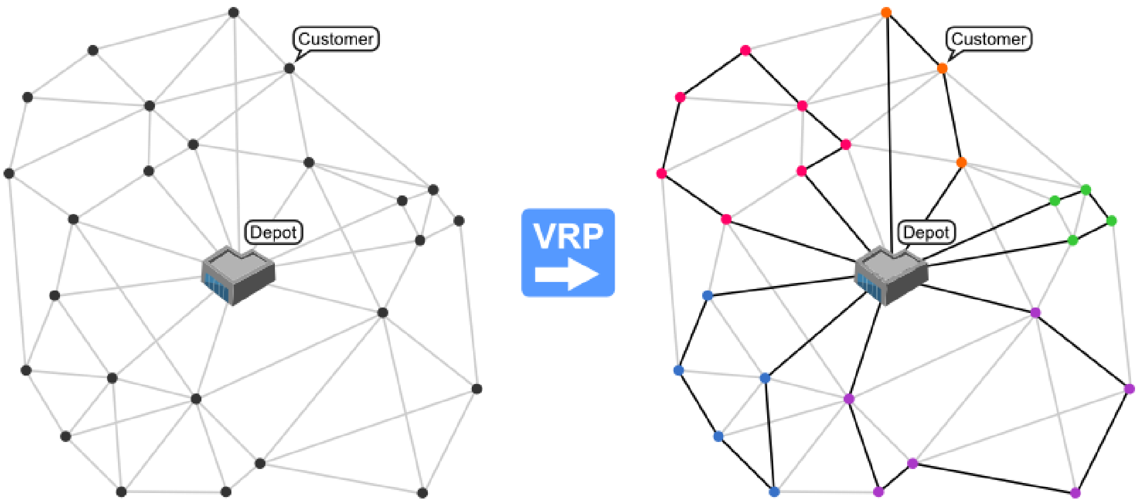

Figure 2.4 - Example of routing elaboration starting from several customers nodes. Source: http://neo.lcc.uma.es/vrp/vehicle-routing-problem/

For a formal definition, instead, reference is made to [2] who assert that *"The VRP can be represented as the following graph-theoretic problem. Let G = (V, A) be a complete graph where V= {0, 1, ..., n} is the vertex set and A is the arc set. Vertices j =1, ..., n correspond to the customers, each with a known non-negative demand, d_j, whereas vertex 0 corresponds to the depot. A non-negative cost, c_{ij}, is associated with each arc (i, j) ∈ A and represents the cost of traveling from vertex i to vertex j. If the cost values satisfy $c_{ij} = c_{ji}$ for all i, j ∈ V, then the problem is said to be a symmetric VRP; otherwise, it is called an asymmetric VRP. In several practical cases the cost matrix satisfies the triangle inequality, such that $c_{ik} + c_{kj}$ P cij for any i, j, k ∈ V (Toth & Vigo, 1998)".*

2.2.2 Conceptual schema

According to [3], Vehicle routing problem may be seen as *"Service of a set of customers by vehicles, located in one or more warehouses, which carry out their displacements using a suitable road network"* (*"servizio di un insieme di clienti da parte di veicoli, localizzati in uno o più depositi, che effettuano i loro spostamenti utilizzando una opportuna rete stradale")* and one possible classification of Routing problem main components could be the one proposed below: Road Network, Customers, Depots, Vehicles, Drivers, Operational constraints, Objectives, Other Features. The figure 2.5 below summarize the approach.

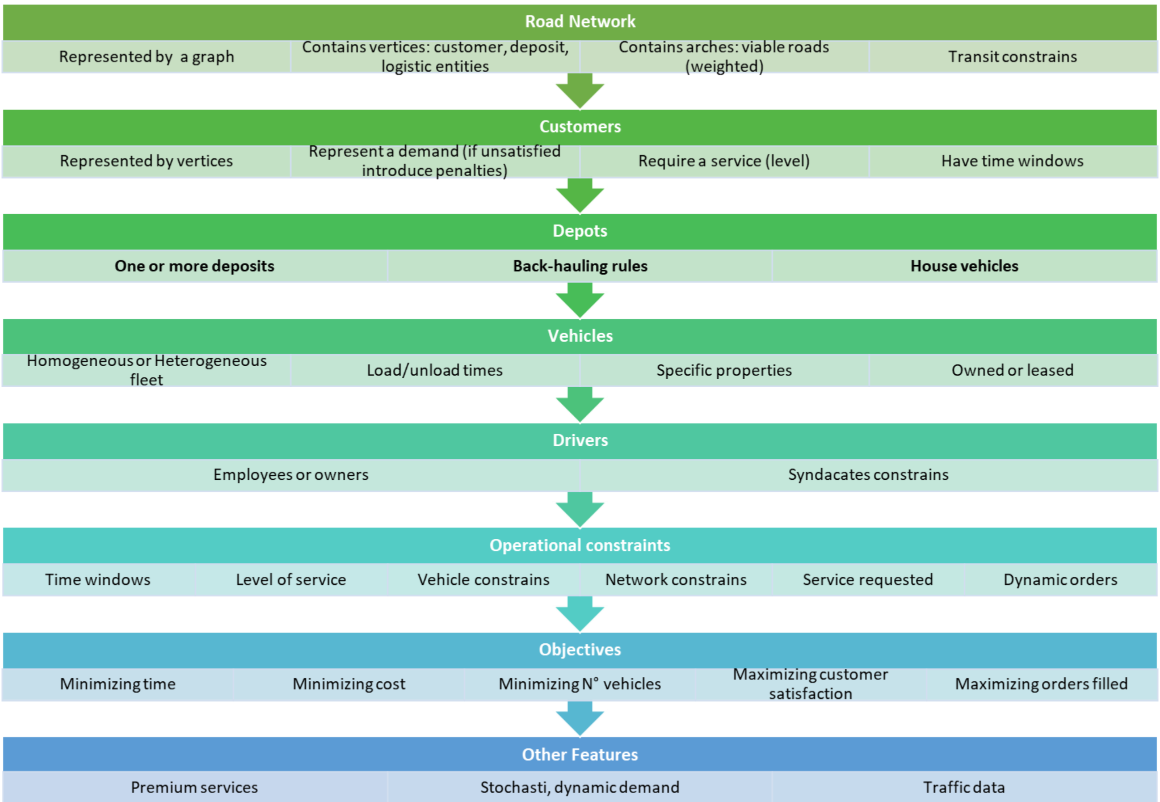

Figure 2.5 - Conceptual schema of factors involved in a Vehicle Routing Problem.

2.3 VRP variants of interest

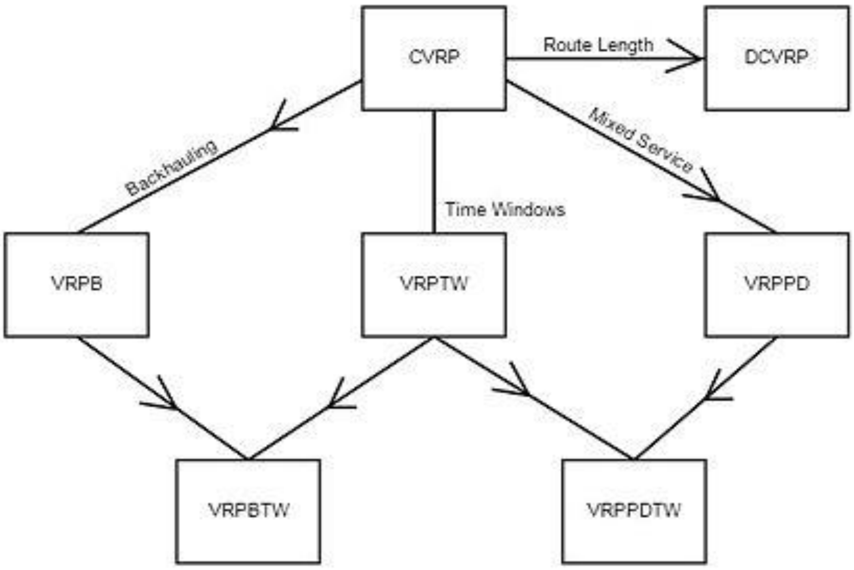

Figure 2.6 - Relations between different VRP variants. Source: https://en.wikipedia.org/wiki/Vehicle_routing_problem

The distribution of goods to customers is a crucial burden for private firms, hence the reliance on the fleet of vehicles performance, as well as the whole routing planning, has to be absolute. To have an effective management of vehicles fleet and routes generation means dealing with so-called "routing and scheduling problems". In particular, two concepts are identified in the aforementioned nomenclature, the concept of routing and the concept of scheduling that can be tackled simultaneously or not depending on the problem on hand.

Since the output of routing and scheduling systems is common, namely a minimum cost routing plan, the proper way to distinguish one from another is according to their specific characteristics or assumptions of the problem on hand.
One of the most common problems, could be, for example, delivering goods to different locations by a fleet housed at a deposit: without a priori restrictions on delivery times and whether goods can be delivered in a reasonable time, then temporal considerations might be ignored so that a pure vehicle routing problem is obtained.
Nonetheless, time constrains are often crucial since in real life situation customers may acquire the delivery only in a limited range of time. More concretely the increasingly

tight competition in the world of distributive logistics, means that the service's level required is becoming more and more a discriminating factor, which more often translates into the need for absolute flexibility of suppliers towards companies, mostly in time terms, in fact, time restrictions are no longer negligible, thus they guide the routing and scheduling activities.

In general, different problems frameworks, as well as specific characteristics and restrictions, lead to different categories of problems that require different modelling assumptions, hence led to different VRP variants (figure 2.6).

2.3.1 Static (Offline) Vehicle Routing Problem

VEHICLE ROUTING PROBLEM (VRP)

As stated by [4]*: "We are given a set of nodes and/or arcs to be serviced by a fleet of vehicles. There are no restrictions on when or the order in which these entities must be serviced. The problem is to construct a low-cost, feasible set of routes-one for each vehicle. A route is a sequence of locations that a vehicle must visit along with an indication of the service it provides. The routing of vehicles is primarily a spatial problem. It is assumed that no temporal or other restrictions impact the routing decision except (possibly) maximum route-length constraints."*

The basic vehicle routing problems, of interest for the thesis aim, that have been chosen to discuss are described below.

Particular attention is placed on the so-called **Travelling Salesman Problem** (TSP), figure 2.7) on one hand because of its historical importance, on the other hand because of its peculiarity of being a strongly topical problem so far. This problem, in fact, might be considered the simplest routing problem, as well as even the most studied among all. The same attention will be spent for the related Multiple travelling salesmen problem as well, due to its particular interest for the thesis purpose.

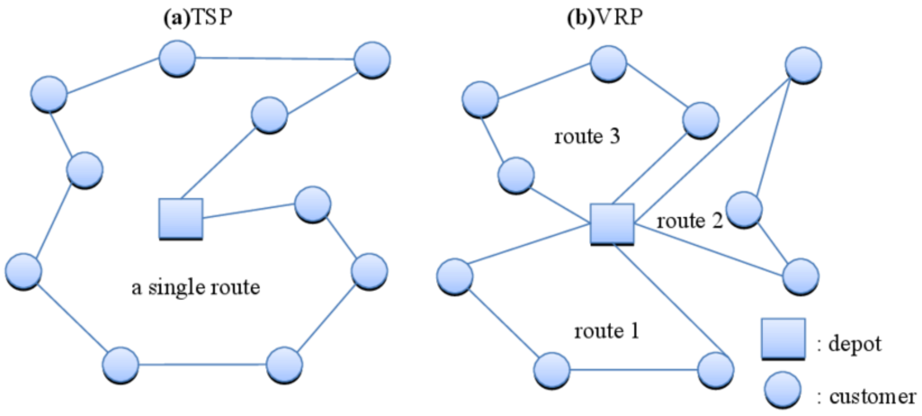

Figure 2.7 - Comparison between TSP and VRP. Source: Liu, W., Lin, C., Chiu, C. et al., Minimizing the Carbon Footprint for the Time-Dependent Heterogeneous-Fleet Vehicle Routing Problem with Alternative Paths

The celebrated traveling salesman problem is structured as follows: consider an undirected network where the demand for each node is supposed to be known (deterministic) and no route's time restrictions are settled. A single vehicle, with infinite capacity and housed in a single depot, must service each node to satisfy its demand. The salesman (who can be seen as a vehicle, especially in our case) must visit all the customers of the network, starting from the deposit and returning to the end of the route. The objective function consists in minimizing routing costs in the form of total distance travelled.

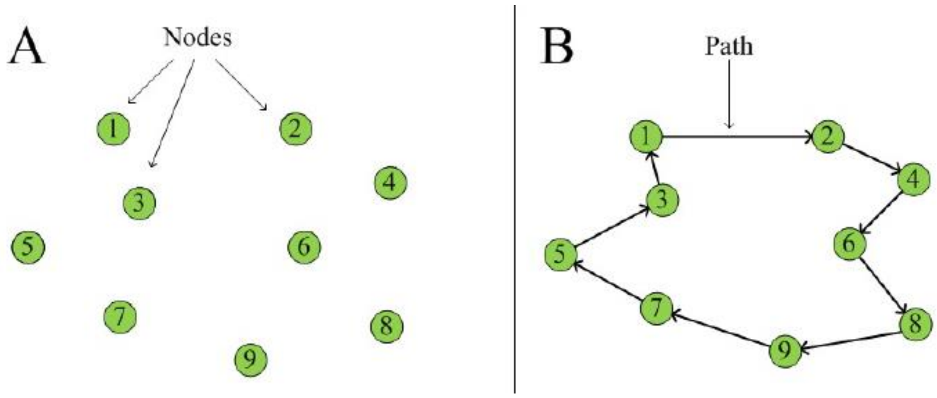

Figure 2.8 - Shortest path between a network of nodes. Source: Osaba, E., Carballedo, R., Diaz, F et al., Crossover versus Mutation: A Comparative Analysis of the Evolutionary Strategy of Genetic Algorithms Applied to Combinatorial Optimization Problems, 2014, p. 1-22

The TSP is a purely geographic problem (figure 2.8), in the sense that the problem's constraints and the objective function (which represents the magnitude to optimize) depend only on a geographic component.

Despite the descriptive simplicity of the TSP problem, finding the path that minimizes the total travelled distance or total costs and which visiting all graph nodes only once, it is proven to be a problem of difficult resolution, more formally an NP-hard problem.

Assuming to have at disposal a number m of salesmen, who can serve a certain set of customers, the **multiple traveling salesman problem** (MTSP) is obtained. In this problem these salesmen are to visit the nodes of a given n node network so that the total distance travelled by all them is a minimum. In other words, the MTSP problem is a generalization of the traveling salesman problem that aims to be closer to accommodating real-world issues where the need of more than one vehicle (salesman) arise more frequently. [4]

The **single depot, multiple vehicle, node routing** problem is considered the very classic vehicle routing problem, hereby it is particularly relevant since it represents the basis for the most complex real-life related models. The goal is to service all the customers demand, generating appropriate delivery routes in order to minimize the total travelled distance. In this problem customers' demands are to be known (deterministic demand) and vehicles, which have a certain limit capacity, departure and return at the same deposit as they are housed at a unique central depot. [4]

VEHICLE SCHEDULING PROBLEM (VSP)

Vehicle scheduling problems are issues that require the introduction of the time variable as an additional constraint to the routes. Since the transit of vehicles in the network compel a certain travel time, the route scheduling must consider both this time, necessary for the physical movement of the fleets, their sequencing and compliance with the crews' regulations, and of any possible additional time constraints imposed by customers. In the latter case in particular, it is clear that it is essential to make the time component explicit. Therefore, in a vehicle scheduling problem, a certain number of tasks within routes, that are associated with time, must be completed. The feasibility of a route depends on both space and temporal specifications, consequentially its scheduling. [4]

Some basic problems included in the VSP category, of particular interest for the thesis aim, are discussed in a row.
Generally, three fundamental constraints, which imply the complexity of a VSP, can be identified: (a) The return to depot constraint, after a period of service, for refuelling or end-of-day, to which a vehicle is subject; (b) The constraint on the vehicles compliance to operate certain specific tasks; (c) The constraint linked to the number of deposits to which the vehicles belong.

In the **single depot vehicle scheduling problem (VSP),** a series of paths consisting of nodes subsets, which belong to an acyclic network, are obtained. Since each path is assigned to one and one only vehicle, minimizing the paths necessarily entails a minimization of capital costs. An alternative objective function can be to minimize total arc weight, since, if we consider dead times as a measure to weigh the arcs of a network, it would correspond to minimize a total travel time, thus the operating costs.

In the **vehicle scheduling problem with multiple depots (VSPMD)** vehicles can depart from more than one depot to perform and complete tasks, it should be born in mind, though, that each vehicle must return to the depot it belongs to, same as the classic VRP. Furthermore, it should be taken into account that there is an upper limit (and lower) to the number of vehicles of a fleet that might be housed in each deposit.

VEHICLE ROUTING AND SCHEDULING PROBLEM (VRSP)

Real applications cope with both the abovementioned problems simultaneously, hence considering both the precedence and eventual time windows constrains imposed either by the network or customers. It is worth noting that the constraints implicitly considered in the classic VRP, including in particular the optimization of an objective function and the singularity of the visit to the customer for each courier, are to be taken into account still.
Two problems fall in particular in the sake of the thesis and are presented below.

Tractor-trailer Routing and scheduling with full loads refers to a relatively frequent problem especially between supplier, such as raw material suppliers, where the trailer is fully filled when a delivery occurs, which requires a certain quantity not exceeding the capacity of the trailer.
It is assumed that the trailer is both unique and that the destination of the entire load is sole. For scheduling purposes may be taken into account that the trailer can be filled either at the tractor's arrival pickup point or a full trailer already disposed for the pickup can be arranged; therefore, it heads directly to the final destination (it is trivial that the precedencies are to be considered), speeding up operations and eliminating load-unloading times.

Tractor trailer Routing and scheduling with partial loads is a problem clearly related to the previous one, with the difference that in this case the customers require quantities other than the trailer unit, but its fractions. The resolution of the problem is, therefore, to find the optimal route accomplished by the tractor by dividing the load to different destinations requesting different and specific quantities. An example of a scenario is offered by [4]: *"Leave domicile, pick up part of a load, Deadhead, pick up part of a load, Deadhead, deliver part of a load, pick up part of a load, Deadhead,..., deliver part of a load, return to domicile."* This problem in particular, in addition to the implicit precedence constraints, yields more complexity to the time windows problem. Indeed, it has to be considered in scheduling that temporally consecutive deliveries must nevertheless fit in time windows. Windows that can, moreover, be different and specific for each customer, both in terms of working hours, and of different service required (priority or not).

In conclusion it could be synthesized, as [4] observes, that: *"The basic output of all routing and scheduling systems is essentially the same: for each vehicle or driver, a route and a schedule is provided. Generally, the route specifies the sequence of locations to be visited and the schedule identifies the times at which the activities at these locations are to be carried out".*

2.3.2 DVRP: Dynamic Vehicle Routing Problems

Not only the technological advances we have witnessed in the last decade have strongly influenced the logistics and distribution sector, also the prospects and standards that society and networks require to the sector, specifically in Logistical management problems, have changed. In fact, they push more and more to consider the necessity of a more efficient, sustainable and long term, understood as the awareness of the inevitable increase of the population, therefore of the traffic and transports problems. On the other hand, the need to keep up with other companies in the industry, which exploit the adoption of the latest technologies, in order to stay or excel on the market, force to implement new more efficient and cost-effective technologies.

DYNAMISM IN VRP

The Dynamic Vehicle Routing Problem (which we will call DVRP from here onwards) can make a significant contribution in the field of Logistic management problems, and represents, in fact, a natural consequence of the abovementioned instances, as well as it is among the more studied VRP branches in recent years, considering its peculiarity to exploit online technologies and its potentialities from the point of view of saving, customer service, sustainability and increase in performance.
A comprehensive definition of DVRP is the one provided by [5]: *"DVRP consists in assigning routes at minimal costs, to a fleet of vehicles in order to serve a set of customers by also considering the information evolution (dynamism) and uncertainty (stochastic knowledge)."* Also, again [5], claims that *"a VRP is dynamic if information (input data) on the problem is made known to the decision maker, or is updated, concurrently with the determination of the routes",* for this reason DVRP is also called online routing.
These sources of dynamic variation of an original set of inputs, called **sources of dynamism**, can be diverse and constitute different inputs compared to those provided to the original optimization's problem, and are unknown until the moment of the event. They may correspond to:

- Variations in demand;
- Order cancellations;
- Network's traffic;

- Travel times;
- Vehicle conditions;
- Other sources of uncertainty that characterize the world of goods distribution in urban areas (e.g. accidents, deviations, time limits, bad-weather, etc.).

DIFFERENCIES WITH STATIC ROUTING

The DVRP incorporates some complexities that were not taken into account in the traditional Static routing, so-called offline routing, which although has been extensively studied and analysed, therefore it can be useful to make some considerations regarding these two types of vehicle routing.

From a mathematical point of view, it should be recalled that most of the static problems are NP-Hard ([4]), and, since the DVRP incorporates greater complexity and restrictions than the former, it appears natural that it also belongs to that category of problems not solvable in polynomial time. This implies greater resolution times, which explains why the most widely used solution methods are heuristics. A further consequence is that, usually, if the problem is particularly complex, it will be less likely that a new order will be dynamically inserted in an ongoing route ([6]).

Conversely, from the standpoint of objective functions, static and dynamic routing have substantial differences, as noted by [7]. Generally, the most common objective functions in static routing are minimization of the vehicles, total travel cost, total travel time, travel distance or fleet's operating cost. Dynamic routing, given its greater flexibility to the real world, widens its objectives to less purely economic, but more strategic variables, such as maximizing the throughput rate, service's level and customer satisfaction, or minimization of order's denials, delay or reaction time.

TRADE-OFF OF DVRP

In this regard, it is clear that in order to fully exploit DVRP potential, namely to harness an online platform to accept orders and possibly satisfy them usually "as soon as possible" by aggregating them to a route already in execution, it is necessary that both the time taken for the decision to accept an order and the time required for algorithm's construction and elaboration, must require a reasonable amount of time. Only by doing so, in order, decision, response and quick execution are allowed, with the additional scope to prevent a further variation of inputs. What we are looking for in the DVRP is therefore a trade-off between the proposed solution's quality and the desired responsiveness ([5]).

It is clear that the decision maker is faced with intrinsically complex and frenetic decisions, for this reason it is often supported by a DSS (decision Support Systems), of which the DVRP solvers are part. Moreover, the compromise quality-reactivity is one of the reasons why in this field heuristics are widely more used than the exact methods. In this direction, important researches are underway for the development of the so-called parallel resolution of algorithms, which can allow the search of solutions in parallel manner therefore faster. Of great interest is also the search for so-called decentralized solutions that may on the one hand prevent the central system from redesign an entire route for a request that may involve only one vehicle of the fleet, on the other allow vehicles to derive a new route by themselves thanks to dedicated devices.

EVOLUTION AND QUALITY OF INFORMATION

Another substantial peculiarity of the DVRP, as [7] claims "*In contrast to the classical definition of the vehicle routing problem, real-world applications often includes two important dimensions: evolution and quality of information*".

Specifically:

- The **evolution of information**: refers to the fact that dynamic routing considers that the instances established at the beginning of route's planning can undergo variations, more or less sudden, during the journeys, while
- The **quality of information**: refers to the uncertainty of some information (such as the demand of a customer), but it can be estimated from historical data rather than stochastic data.

It is interesting the consideration proposed by [7], which observes that these different routing scenarios might basically identify four routing categories by combining the characteristics of static, dynamism and the attributes of evolution and quality of information. A summary scheme of the characteristics of each category is offered below (figure 2.9).

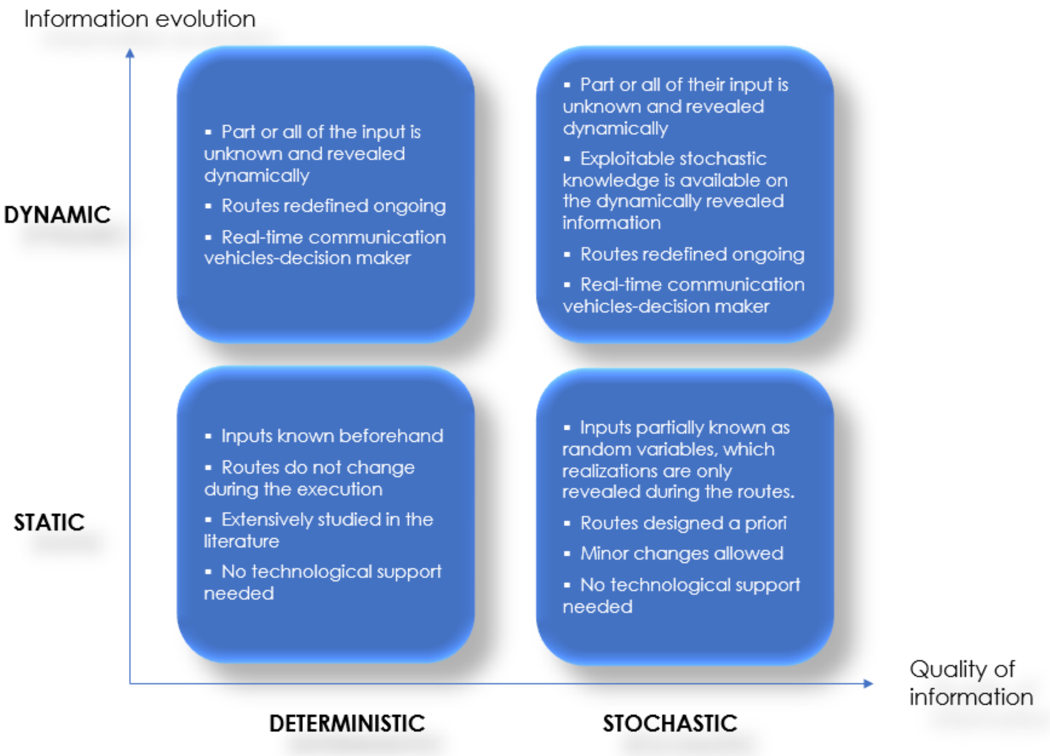

Figure 2.9 - Four cases of routing generated by different type of information

It is worth nothing that, even in DVRP contexts, static routing algorithms are often used, in fact they coexist. Therefore, it should be recalled that the advent of this new DVRP world, does not entail the abandon or disuse of the static routing, indeed, the first exploits the latter when possible, given its vast literature and the many applications still highly valid nowadays. Moreover, it is improper to make a blind use of the DVRP, as the amount of information and the effort of a whole system to sustain it is remarkable; in doing so, in fact, a scenario for example stable can become unnecessarily heavy, if treated in a dynamical manner. Indeed, inappropriate uses can be even counterproductive. Sometimes it is not necessary to resort to DVRP if not required, in fact, by doing so, time dynamic time-consuming procedures may be avoided, as a result a considerable amount time can be saved. Therefore, a highly innovative and original architecture, such as the one we want to propose, cannot be separated from taking into account both and the considerations all above.

TECNOLOGICAL SUPPORT

The efficiency of the DVRP lies in the fact that it wants to absorb and be more adherent to the complexity that characterizes the real world, understood as a network and its actors, consequently a continuous communication of information between the environment and the dispatcher is needed. Such information is necessary to update the routes, with real-time information, namely not only the latest information available from the network, but also the state of the vehicle and its location; for this reason, the exchange is twofold.
In this sense, over recent years in the field of real-time and online communication giant strides have been made and this has led to the development of a wide range of applications for routing, both from the point of view of global information infrastructures, such as the Global Positioning System (GPS) and Geographic Information Systems (GIS), and in terms devices available to drivers and decision makers (tablets, smartphones, navigators, in-vehicle systems, etc.). These technologies, that are particularly exploitable in the world of logistics, together with the massive development of Information and Communications Technology (ICT) tools, have gradually enabled the search for strategies that exploit the saving potential and performance awaited by dynamic routing.

A PRACTICAL EXAMPLE

To give an example in the thesis field of interest, let's think to the case of a courier destined to pick up at a customer in urban environment. If a variation of one or more of the above variables occurs (e.g. the arrival of a new order), it may first be accepted or discarded, both for reasons of impossibility or convenience, whether it is considered convenient, the dispatcher resolves online algorithms and updates the routes. At this point the necessary information is transmitted to the vehicles involved, all respecting the delivery constraints of the previous route and of the new instance (for example in terms of time windows). Obviously, the adjustment of a route, so the readdress of a vehicle to a new destination, requires prompt and precise communication of its location in the network, its current capacity and its general availability, and this would not be possible without a Dual and real-time data transmission between vehicles and dispatcher.

An example taken from [7] may be representative of the aforementioned case (figure 2.10). Consider a single vehicle that is initially assigned to a route that expects to serve, in sequence, the nodes A, B, C, D, E, where t_0 indicates that this route is derived from a first optimization with known input, performed at the beginning of the day. During the execution, two new orders X, Y are accepted by the system and a new route is recalculated and communicated to the vehicle. In fact, t_1 shows the change of the arcs to traverse, with the insertion of the new ones and the deletion of the previous ones, while t_f indicates the final transmitted route, which is (A, B, C, D, Y, E, X).

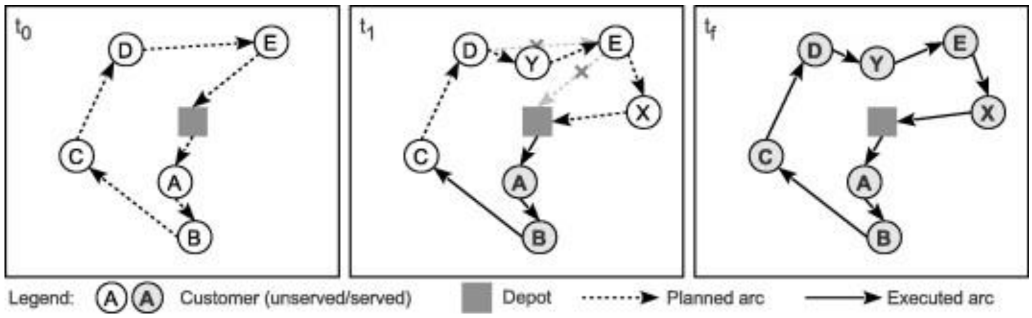

Figure 2.10 - Evolution of a route due to dynamic arrival of new order. Source: [7]

In terms of the dynamic communication, the underlying figure 2.11, proposed by [7], summarizes the process of transmitting and updating data between dispatcher and vehicle respectively, the latter operating in the so-called environment, which stands for the network.

Once the vehicle is ready (first dotted arrow), it is assigned the first destination, the customer (first double-headed arrow). Next, once the vehicle has communicated the dispatcher that it has started and completed the request to (respectively second and third dotted arrows), the dispatcher update data concerning the next request B (second double-headed arrow). The same process of notification and updating occurs in the case of new requests (fourth dotted arrow).

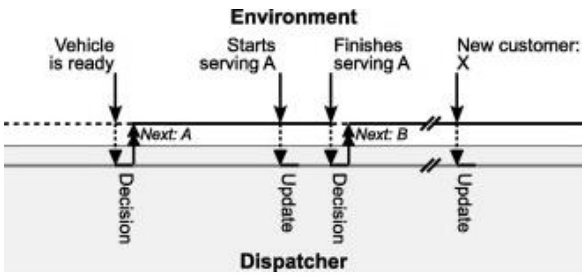

Figure 2.11 - Exchange of information in dynamic context. Source: [7]

PROBLEM'S DYNAMISM

Obviously not all problems are characterized by the same volatility of information, nor do all companies adopt identical decision-making strategies. For this reason it is good to introduce the concept of level of dynamism which, accordingly to [7]: "*Can be characterized according to two dimensions: The frequency of changes and the number of requests. The former is the rate at which new information becomes available, while the latter is the time gap between the disclosure of a new request and its expected service time.*". This leads us directly to the definition of *degree of dynamism* ([7]) understood as the relationship between the number of dynamic requests n_d and the total number of requests n_{tot}:

$$\delta = \frac{n_d}{n_{tot}}$$

Other more complex formulations were developed by Larsen [8] to meet the real world related problems. Among these is worth mentioning the *degree of effective dynamism*, that takes into account the time of communication of each new request in relation to the temporal horizon considered, still in relation to the total requests; And the *degree of dynamism for problems with time windows* in which the notions of urgency and reaction time are introduced, to indicate that the greater the timeliness is, the greater the flexibility of inserting a new request will be.

It must be noted that the multiplicity of sources of dynamism means that the previous considerations are not absolute and, indeed, they renege if other factors such as geographic distances and travel times greatly affect the level of dynamism of the problem under discussion.

APPLICATIONS FIELDS OF INTEREST

In the final section, some interesting applications of the DVRP in the field of services and freight transport are stated.

In the field of services, the application of the DVRP translates into the dynamic acquisition of instances characterized by a location, corresponding to the client's position within the network, and usually a time window that it requires to comply with. It also may require a service, that may concern pickup, delivery or both. Given these peculiarities, the famous TSP problem may be identified as basic problem, since here, like in the latter, capacity related constraints are not taken into account. An interesting example of services is the area of maintenance operations. Contracts with companies in this sector can be diverse and include both periodic maintenance or scheduled order. By nature of maintenance itself, however, even the best maintenance plan of a plant is subject to unexpected breakdowns or stops, this implies that, in the context of routes planned for scheduled maintenance, possible urgent requests may be added dynamically due to breakages or malfunction, requiring a "as soon as possible" service. Furthermore, as observed by [7], if then the concomitance of the workers skills, spare parts and maintenance tools is added to this scenario, is immediate to understand the complexity of the problem on hand. This explains why Decision Support Systems (DSS) are usually required in this field.

Even more interesting from the point of view of the thesis in question are the applications of urban freight transport. The definition of urban freight through DVRP recalls another category of applications called *City logistics* by the experts. By City Logistics we mean that category of problems that deal with managing, planning and coordinating the demand in urban environment, trying to offer a desired level of service and, whether possible, optimizing pollution, congestion and costs. [9] These problems are therefore closely linked and influenced by the nature of the urban environment itself, which is tackled in a holistic way, and which, in particular, is subject to high dynamism, in particular in terms of variability of the traveling times.
In these applications the company usually receives a pickup or delivery order from a customer. Together with the order the company, also acquires a client location, a time window and a desired level of service (at a certain price). Once this acceptance process has taken place, a courier is assigned to this order, which, in turn, can be placed in an ongoing or new route. Later on, then, travel times and traffic congestion are taken into account for further adjustments, thanks to dynamic real-time information.

In the field of newspapers and magazines delivery, for example, applications of DVRP have been proposed that offer the creation of online routes, with direct communication to the vehicles through smartphones, aiming to improve the level of service perceived.
Cases of grocery delivery ([7]), however, plan not only that products can be chosen by customers directly from the online platform of the store, but also that they can choose a certain arrival time. These products ordered online, once the purchase is confirmed, constitute a new dynamic request that is fed to the system.

FUTURE OF DVRP

In the last decade technological progress has been a huge opportunity for companies, because it has allowed for online and real-time routing management, therefore dynamically. Consequently, however, a growing complexity has led a need for support instruments.

As noted by [5] the future developments of the DVRP will cover:

- The implementation of parallel algorithms that can speed up the resolution of a problem, perhaps in a decentralized way;
- Reinforce the dual communication infrastructures between vehicles in the field and within the fleet itself;
- Greater consideration of the stochastic nature of the sources of dynamism;
- The implementation of electric vehicles.

2.3.3 TWVRP: Vehicle Routing Problem with Time Windows

A particular attention should be earmarked for an ever-rising variant of the VRP, the one that it includes time constrains, the so-called Vehicle Routing Problem with Time windows, often referred to simply by its initials, VRPTW. This problem extends the basic one through a further constraint: customers must be served within a certain chronological interval represented by these time window, which, in practice, curtail customer's readiness or willingness in accepting a delivery.

This problem responds to the ever more urgent and realistic need to consider the temporal facet of routing problems, which resolves to the need for companies not only to cut as much as possible logistics costs, but also to be competitive on the differentiation of the service. This has led to a growing and extraordinary development of research on Time-constrained Routing and scheduling problems. Hence, the Time dimension has been incorporated into VRPS in the form of time window constraints imposed on customers to reach and serve. More precisely, the VRPTW involves a fleet of vehicles which, starting from a deposit, must reach and serve a number of customers, located in different geographic locations, each with specific demands and time windows within which the service must take place (figure 2.12). As in the VRP, each customer must be served exactly by a vehicle, consequently, only once in the entire deployment process. [10] Furthermore, each vehicle must be assigned a route that contains deliveries not exceeding its capacity, and which must return to the depot from which it has departed at the end of the deliveries.

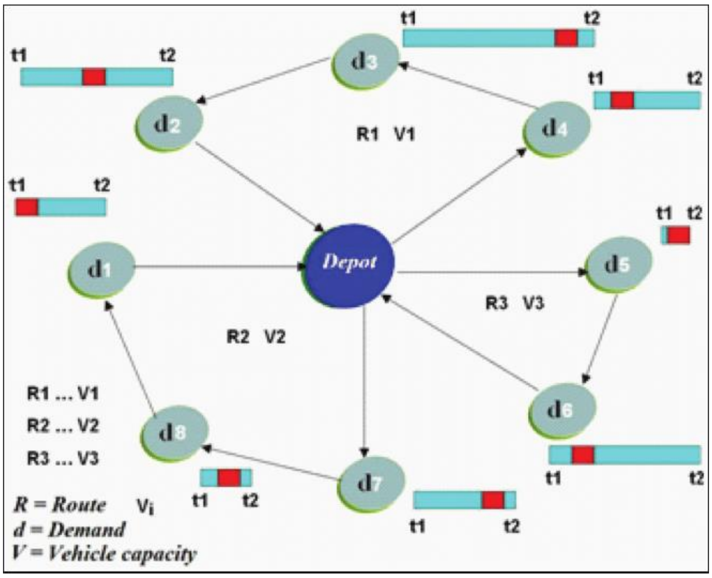

Figure 2.12 - VRP with Time Windows explicated and different for each customer. Source: https://www.slideserve.com/tyrone/a-hybrid-genetic-algorithm-for-the-periodic-vehicle-routing-problem-with-time-windows

In terms of the abovementioned temporal constrains, we may find to variants:

- **Hard TW**: in which the time windows are to be respected rigidly, as meaning that a vehicle may arrive in advance from a customer, but in that case it is forced to wait until the beginning of its time window and anyway it can never get from the customer after the end of the window;
- **Soft TW**: in which the time windows are soft, i.e. they can be violated at the cost however of the introduction of a penalty contribution on the objective function.

In general, researchers considered VRPTW problems with Hard Time Windows and homogeneous fleet of vehicles. However, the maturation of research in the field led to the developments of new sophisticated methods which allowed to treat also heterogeneous fleets. As regards the size of the fleet, could be fixed a priori or assumed free, that is, to be determined simultaneously with the best set of routes for vehicles, although many methods can currently optimize the fleet size during the process.
Despite the introduction of time windows in the VRP problem generates obviously more complexity, in terms of constructing and maintaining a viable set of routes, this

extension bears the specification of more realistic objective functions, rather than a simple minimization of the total distance travelled. In fact, in addition to the latter or the number of vehicles used, one can minimize the total trip duration or the overall cost involved in the routing and scheduling, which consists of fixed vehicle usage costs and in variable costs (e.g. time needed for travel, time waiting and loading and unloading, etc.). In spite of its intrinsic complexity, VRPTW has received considerable attention in recent years in the operational research community. First of all because it is still a NP-hard optimization problem and therefore represents a great challenge for each researcher; secondly, as earlier introduced, because of its strict connection with the achievable costs reduction within the logistic field. [10]
The VRPTW has been studied and tackled using many different techniques that include both exact, heuristic and metaheuristic methods. Given the arduous implementation of truly efficient methods (most of the times a prohibitive time is required to find a good solution to an instance of the problem), most of the research has focused on the development of different and efficient heuristic techniques able to provide in a reasonable time a good solution to the problem (acceptable, but certainly not excellent). [6], [10]

2.3.4 TDVRP: Time Dependent Vehicle Routing Problem

The rise of urban traffic in the last decade, in continuous ascent so far, has been dramatic, thus, today, it occurs to be a key factor strongly influencing the routes in terms of travel times, but, more generally, the entire operating network. The growing importance of customer service and service differentiation have, in turn, increased the awareness that today traffic conditions are no longer ignorable, and indeed have taken a pivotal role, as the service to Customer is clearly influenced by the delays arising. [6]
For this reason, special attention has been paid to a category of VRP called The Time Dependent Vehicle Routing and Scheduling Problem or by shortening TDVRSP.
This category addresses the problem of the dependence of route departure time on urban deliveries (or pickups) in terms of resulting travel time accordingly to congestion of network arcs involved at that given time (figure 2.13).

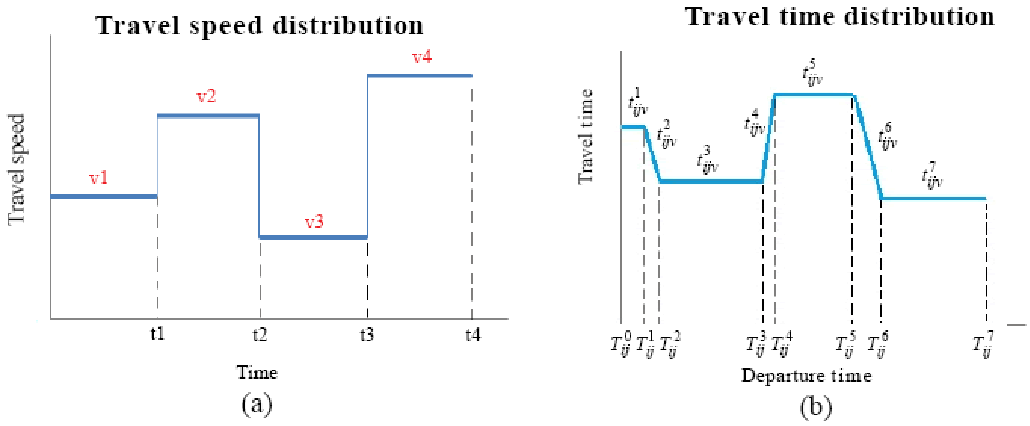

Figure 2.13 - Travel speed (a) and Time (b) distribution. Source: YenYi, H., Lee, Y., Two-stage Particle Swarm Optimization Algorithm for the Time Dependent Alternative Vehicle Routing Problem, 2014, p. 1-9

The objective function consequently aims at trip time minimization (with the hypothesis of a fleet with limited capacity), rather than travelled path minimization; accordingly, the output of the algorithm changes too, as network's congestion in critical time periods is introduced as a new instance. This means that it can return, as best route, a longer path, most likely not straight directed towards the node to be served. The best route selection approach of a TDVRSP, in this case, underpins the fact that longer routes, travelled at higher average speeds, allow to serve a customer in less time, because travelled far from the most congested areas, which certainly provide more direct ways, but equally characterized by even zero-speed tracts in case of clogging. This latter would have been the solution produced in case a common VRSP problem would have been adopted, since it does not consider the dependence on time. The consideration of time dependence allows also to meet other temporal requirements imposed, i.e. time windows. The benefit is therefore twofold in this sense, since the introduction of more stringent constraints enables more adherence to the real network's situation, in terms of travel times, hence in terms of arrival time within or extra time intervals imposed by consumers. Consequently, this may yield a solution that will probably be best fitted to the time windows.

The growing interest that TDVRSP is experiencing, is also due to its close correlation with the environmental impact (specifically considered in the so-called Pollution routing problem or Green routing). The increasingly stringent regulations on urban pollution, in some cases unsustainable, obligate to consider the emissions of couriers, which are less if travel's speeds are averagely high compared to continuous stop-and-go typical of traffic jams, as observed by [6]: *"Less pollution is created when vehicles are traveling less time and at environmentally friendly speeds."*

During the process of route's selection (with the shortest travel time), the algorithm needs two basic data, that is, temporal coordinates (date, time of start route) and realistic data on average speeds on involved arcs.

More generally one can say that the resolution of this problem involves a larger number of variables at stake, compared to the classic problems (where, for example, the distance between knots is enough), here it is to have access to the database of road travel times (in different time slices) regarding a given network.

It is also necessary that the travel time forecasts are as accurate as possible, in light of this, various techniques have been developed in order to accomplish this objective, such as the ones observed by [6] :*"[…] historical road travelling time profiles, time-series models, neural networks, non-parametric regression models, traffic simulation models, dynamic traffic assignment models or knowledge based real-time travel time prediction methods."*

As a result more and more authors agree that today it is highly necessary to have a wide range of scenarios and historical series, as well as real-time surveys where possible, that can serve as a basis for the travel times calculation concerning different hours, days, weeks, months, but also, more extensively, concerning seasonal phenomena or trends and special events (concerts, matches, parades, etc.); with the ultimate goal of having database of road travelling times regarding a given network.

This approach, based on the avoidance of being in congested areas at a certain time, may be expressed in precautions, such as those suggested by [6], including for example:

- **Change the order of customers to visit**: if, for example, a visit to a given customer would mean passing through a particularly critical area at that time;
- **Assign a delivery to a different route**: subsequent to a paramount traffic moment;
- **Real-time modification of a route**: in case of unexpected clogging due to an accident occurring in an ongoing route (very frequent occurrence);
- **A priori modification of a route**: if you have information about critical traffic peaks at the time of the scheduled departure of the route.

Chapter 3: Road traffic data collection

This chapter aims to presents a discussion on the road Traffic data collection, a field as growing as well as crucial for logistics companies and especially for an accurate fleet management. The fleet of vehicles operates, by its very nature, in a network inhabited by a myriad of actors, each with different needs and peculiarities, and external factors, such as emergencies and weather. Each of these factors inevitably influences the current state of the network itself by generating, for example, traffic, incidents, deviations and unforeseen events. The planners must take account of this during the daily activity, as this translates, from the point of view of operative business, in variations of the couriers travel times, of the viable paths, often to be managed in real-time or "as soon as possible" manner, therefore the feasibility of a delivery, this inevitably affect the customer service and, at longer term, the profitability and success of the company.

The discussion will also consider the fact that, more broadly, network's complexities are to be faced by entities not only private, but also country's agencies and departments that have the task of managing this complex reality, stood as traffic management and infrastructure transport, in terms of quality and safety for all the actors. This should not be considered a divagation from the main theme, but a way to realize how the need for real-time information is becoming a daily necessity.

The development of ICT and the growing interest on Intelligent Transportation Systems (ITS) in recent years have inevitably become an added value in this field, allowing not only new and more efficient data detection and transmission technologies, but also representing a cost-effective alternative to traditional methods, as well as offering new opportunities for combining heterogeneous traffic information. These advances, constitute a valuable source for providers because the market and, in particular the logistics companies, are increasingly willing to pay for a service or applications of dynamic traffic management.

From the sake of final thesis goal, it is clear the influence on the architecture of the system and the underpinning greed of that system to work in a context of reliable information and in real-time. Therefore, we will seek for a lucid solution that incorporates and integrates all this into the superstructure, so as to manage dynamic routing in a prudent and integrated way.

3.1 Traffic Data

By traffic data we mean that information circulating in a network in a certain temporal and spatial context and which in transport and mobility involve a multitude of stakeholders in different application scenarios: Citizen (Pedestrians and road Users), public sector (Governments and authorities) and the private sector. Such data can be of a heterogeneous nature and their load may vary over time. These dates are the basis for the network traffic measurement, network traffic control and simulation. [11]. Of course, first users of these traffic data are road agencies which use the data obtained from the network on one hand in the short term, for the daily monitoring of traffic flow, detention of accidents and data processing made available then to other Network users, on the other hand in the long term, to plan the future of network infrastructures. As part of the interest of the thesis, traffic data useful for logistics companies are, for example, network travel times, speeds, real-time updates and reports and origin-destination travel matrices; which exploit this information to improve their performance, for example to achieve reactivity of route change in case of accidents or queues, provision of optimal routes, improvement of customer service. [12]

The urban network traffic issue has seen an increase of its importance during the last decade because both from the point of view of the enterprises, and public-state, the rise of the new technologies has allowed to improve the quality, reliability, safety and forecasts that concern the network mobility, for the benefit of private users, public transport and freight transport.

Concretely speaking, the interest of the company has shifted gradually towards the ITS or "intelligent transport systems", so called because they propose to integrate competences from heterogeneous fields such as telecommunications, electronics and informatics, to exploit the peculiarity of new advances, with the ultimate aim of making the context of planning, design, forecasting and management of transport in a network, safer, more efficient and sustainable for all stakeholders. [11] The intelligence of these systems lies in the effective interconnection of these different branches thanks to the potential of new devices and ICT. Not by chance a possible variation of the acronym could be Interconnected Transport Systems. In short, these systems define the efforts and technologies to add information and communication technology to transport and vehicle infrastructures, such as dynamic information scoreboards on traffic, weather or accidents, car Navigator systems, dynamic driving directions, ANPR systems (license plate recognition), systems that compare live information with historical data or smart cities applications [12].

In practice, this has resulted into the advent of new devices and technologies exploiting ICT's new capabilities to manage more complex and wireless information and that have allowed a progressive evolution of the fabric of the transport system. This category leap has renewed and strengthened from multiple viewpoints: detection, collection, distribution, elaborations and data transmission. To date, moreover, we are facing a scenario in which all this information does not have a single direction, but everyone is producer, manager and receiver of useful traffic data, which appropriately treated, may become relevant to another actor. Even in this sense the transport context becomes interconnected (figure 3.1).

The evolution of traffic data world we are witnessing, will bring clear benefits to all the actors of the road network improving the accuracy, the coverage and the reliability of real-time information all at a reduced cost and with less inconvenience for the drivers. In view of the project on hand, this will impact positively on the network planners and manager offering the basis for a more competitive logistics business, not only in the short terms, but also in the medium-long terms, especially in terms of service offered to the customer.

The traffic data collection methods have changed over time, following the technological trends. The following paragraphs we will try to create a distinction between conventional and innovative methods. It has to be noticed that, despite considerations all above, the objective remained to consider a trade of between cost of data collection and the degree of coverage and accuracy obtained.

To begin, it is worth nothing the following consideration of [12], which proposes *"As a broad generalisation, most conventional traffic data sources are based on stationary, fixed point measurement devices distributed across the network. [...] By contrast, new data sources are becoming increasingly mobile, and provide scope for increased network coverage and more information about travel patterns, travel routes and travel speeds".*

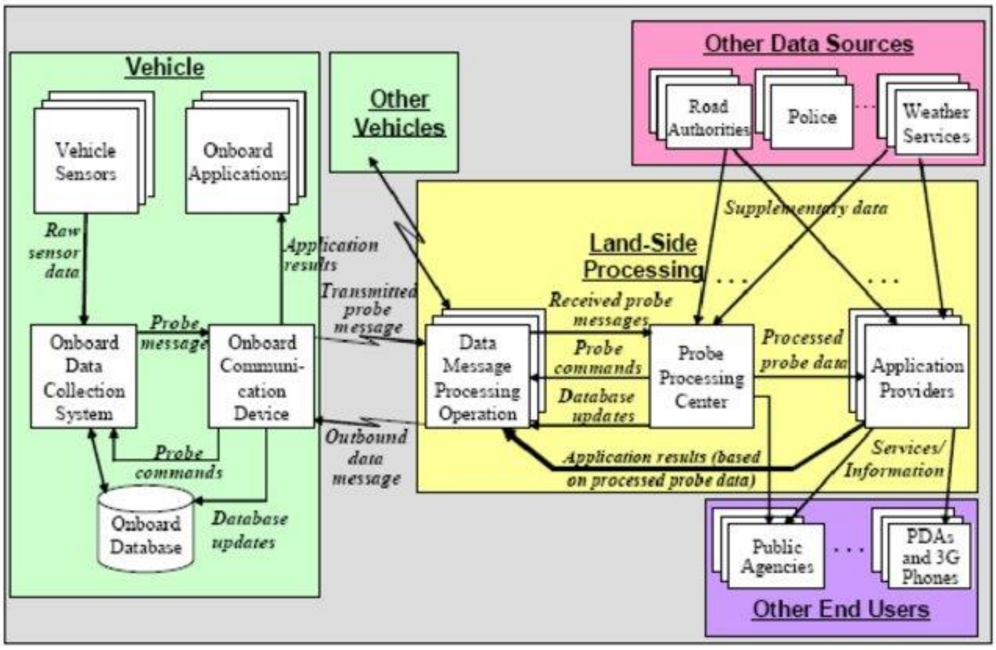

Figure 3.1- Traffic data processing flow. Source: [17]

3.2 Big data

This highly dynamic (massive, volatile and relevant information), complex (given the heterogeneity of the actors and their objectives) and competitive nature (given the close dependence between information from the network and the customer service) means that the stakeholders of the current context require more and more.

Not only the objectives of the stakeholders have varied, also from the point of view of the network two things may be observed: the first is that the technological coverage which is equipped with is considerably wider, more heterogeneous and performing compared to a few decades ago, the second is that today the network requires a better performance in terms of efficiency, connectivity and security.

The technological support available to both users and the network itself (ICT, mobile devices, sensors etc.), allows to point to these ambitious objectives, but, at the same time, this translates into an unprecedented frantic and massive circle of information. The concomitance of all this has generated the need for a so-called Big data paradigm, thus massive, highly variable and high-speed data that need to be generated, captured, collected, processed, transmitted and distributed. These characteristics constitute a breakthrough. These data incorporate an intrinsic complexity and, therefore, need to be treated with adequate countermeasures, methods and tools. However, at the same time they constitute an enormous opportunity for the world of transport, as such complexity translate into wealth of information and, therefore, is the bearer of enormous exploitable knowledge (figure 3.2). No wonder they are essential for ambitious and futuristic applications such as smart cities, IoT and ITS. [13]

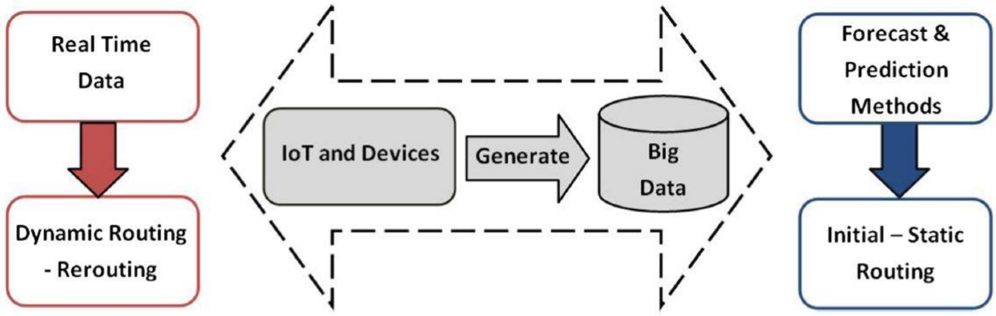

Figure 3.2 - How big data bring knowledge to the system

3.2.1 Contextualized definition

The rise of big data in the transport industry, in fact, does not revolutionize, those that have been for years well-known optimization functions in the field, rather it is proposed as an evolution, on one side to acquire and treat the datasets that these problems need, on the other side, to witness increasingly ambitious and conflicting objectives, so that big data techniques and analytics can be a solution. In essence it is about solving the same problem, but with data (due to new technologies involved) characterized by speed, variety and size all new, which amplify simultaneously both the complexity and the time space-time resolution. Namely, relevant information such as real-time traffic conditions, travel times and speeds, if detected in a big data context, are derived from an analysis of data captured from heterogeneous sources combined, with broad coverage and, characterized by different formats, time periods, hence, with greater accuracy, and consequently strongly representative of the mobility in the network. [13]

To give an exhaustive definition of big data, the combination of two of the most authoritative and recent definitions has been chosen, such as [14] and [13]. The former gives us a rather categorical definition stating that *"Big Data is high volume, high velocity and/or high variety information assets that require new forms of processing to enable enhanced decision-*

making, Insight Discovery and process optimization.", the second claims more holistically than the concept of Big data:*" Blends together the collection of large volumes of high-velocity, heterogeneous, evolving domain data and the use of advanced techniques and models to store, retrieve, manage, process and analyse the Captured information".*

Concretely speaking, both definitions refer to the fact that big data are a solution, consisting in techniques, tools and methods, to the aforementioned complexities, in order to improve the overall network efficiency, (understood as safer infrastructure and ITS) and, consequently, the effectiveness of the deriving data. Finally, at a deeper level, both private users (e.g. everyday travelling) and logistic businesses (e.g. freight Transportation) will benefit in terms of relevant information.

The technological advances (ICT, ITS and new detectors), ensure nowadays an adequate management of big data, and this allows to fulfil the other concomitant aspect, that is the new objectives of today's society meant as, an evolution of the urban context, towards an interconnected, dynamic and intelligent environment, as such flourishing of plethora of big data sources (figure 3.3).

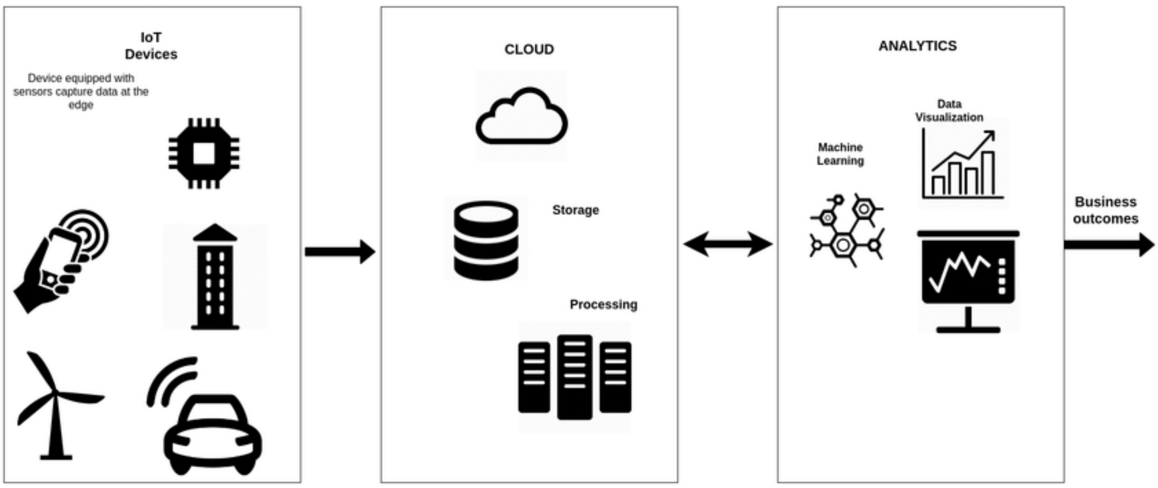

Figure 3.3 - Work-flow of IoT systems dealing with big data in the cloud. Source: Zobaed, S.M., Amini, M., Big Data in the Cloud, 2019, p. 1-*6*

3.2.2 Big data features

Despite the definition of Big data is widely agreed, in literature still persist some differences concerning the features. For the sake of this thesis a basic model but equally complete, has been chosen, i.e. the 3V model of [14] that represents Big Data According to three main features.

- **Volume and size**: refer to the distinctiveness of big data to be quantitatively significant to manage. As such they need ad hoc methods and tools (Big data techniques) to obtain solutions and to guarantee them in a reasonable time. Concretely speaking, a more broadly instrumented and covered network (e.g. sensored-vehicles, mobiles, on road detectors, wearables sensors), as well as a more accurate and reliable network's picture, are achievable only at the cost of a large amount of data to be processed. This proliferation of sensors is due to the extension of the coverage to rural areas and to the enhancement of particularly critical areas.
- **Velocity**: this peculiarity is understood as the speed of the data circulation through the different actors and in the life cycle (from generation to processing to distribution). In a nutshell, it can be said that the ambitious application of futuristic applications and new objectives has grown in step with the speed at which incoming and outgoing data, between sources and destinations, must be transferred. Speed in a context of big data almost necessarily requires a real-time management, especially if we think of applications such as smart cities and IoT. As [13] sometimes observes, *"Big data technology allows analysing the data while they are generated, without even storing them into databases"*. Another possible interpretation could be that of the speed (volatility) with which the data changes according to the environment changes.
- **Variety**: the need for more accurate information from the network has led to, as mentioned, the increase in coverage, but also the increase in density of sensoristics. This means that in the same road today might be found various types of sensors, each providing different and diverse traffic data (as explained in the section dedicated), but, above all, each characterized by specific and different formats and time periods. In practice this spread of conventional and new source of data has translated into the generation of an impressive variety of data (for example, think

about the profoundly different nature of FCD compared to social media source). It is clear that this heterogeneity that characterizes today's technological and digital domain requires appropriate big data tools for data integration and fusion functionalities [13]. Furthermore, this in order, on one hand, to exploit the wealth of carried information, on the other hand to connect those data to lay the foundations for ITS infrastructures, IoT and Smart Cities. The accomplishment would benefit a huge basin of stakeholders: citizens, authorities and, in particular in our case, enterprises.

3.2.3 New Big Data sources

The paradigm of big data not only allowed and forced to adopt a new approach to the management of data involving significantly different characteristics compared to those to which we were accustomed, but also expanded the pool of sources from which to draw. This openness certainly allows to widen the obtainable knowledge, but at the same time it has meant integrating data of a new nature and greatly different from the conventional sources. Here are 3 examples, proposed by [13] that obviously are not the only ones, but certainly the most significant.

- **Social media**: The massive and universal use of social media has made them an inexhaustible source and high-coverage of variegated information concerning the user's status and, as such, big data. The potentialities of the Big data techniques allow to analyse this data and exploit the social user as data provider since their activities are usually localized (mapped). In practice, by the usual use, they leave traces of themselves in the network, with which can be traced back. Some big data techniques can extrapolate relevant information even from string format information which may, for example, consist in tweets of the traffic situation or of a particular state of a user within the network. [13] considers in this typology also the so-called crowdsourcing, that is the growing interest of companies or individuals to request ideas, suggestions, opinions, to internet users for the realization of a problem, the solution of a problem or the improvement of a mobility service (e.g. taxi applications). From the considerations made is evident the deeply new nature of this data source.

- **New Sensors**: As already explained, in the last decade we have witnessed the explosion of the sensor equipment and devices, triggered by new objectives in the context of transport. In particular, not only the density, but also the integrated technology is enhanced accordingly to meet the new needs and potential that applications offer to make the future network a dynamic, flexible, connected and safe platform. Specifically, we are talking about the use of *"physical devices embedded with electronics, software, sensors and actuators"* ([13]) that are now reality in applications such as drones and connected cars.
- **Open data**: Refers to the new concept of making available data and relevant information to all those interested, therefore, in this context, anyone can be either provider or user. Obviously, this approach has assumed and will assume importance soon, meeting the need of the new transport industry to evolve, becoming more and more an interconnected environment, that might link and eliminate the barriers between private users, governments, and business users. Of course, it is not trivial to think about the complexities related to the privacy and data integrity that this approach incorporates.

3.3 Conventional data collection sources and method

By conventional data collection we mean those methods or devices that first enslaved the function of collating data. Nowadays they have not been not abandoned, indeed, they are still today an important share of the methods of detection. Usually by conventional we refer to "in-situ" sensors, placed along the road network and in particular under the road surface, the traffic cameras, the manual counts and the traditional surveys.

Although the technologies considered by several authors are fundamentally the same, there is no univocal classification between traditional and modern sources, also within the same category there may be further. The proposed classification (for both new and conventional sources) aims to propose an alternative classification that takes the cue mainly from those proposed by [15], [12], [16], [17] and it takes into account that, as noted by [16]: *" Each of these technologies has different technical characteristics and operating principles, which determine the types of data collected, accuracy of the measurements, levels of maturity, feasibility and cost, and network coverage "*

3.3.1 On-road sensors

On-road sensors are defined those sensors usually placed below, on the surface or somehow integrated in the road mantle [12]. They are sensors that exploit different technologies or physical principles, hence the technology involved inevitably depends on the nature of the information they can procure, for this reason they are often used in combination with each other. In particular, it is clear that for information such as speed or classification the combination of multiple sensors is necessary, while for information such as weight or counting, even a single sensor can suffice. In general, this category can provide information such as traffic counts, classified vehicle counts, vehicle speeds, real-time information, such as travel times, to traffic management systems, private users and providers, road agencies.

Usually this type of in-pavement sensors, are placed in roads with intense traffic (e.g. highways) and are also considered expensive both from the economic point of view and from the point of view of the installation. In fact, by their nature, need massive and intrusive works, consequently stopping the traffic with all the consequences that derive [12].

A list of these technologies is presented below.

- **Pneumatic road tubes**: the material's deflection is exploited to counting vehicles, which through their passage compress them, modifying the pressure inside then triggering the count;
- **Piezoelectric sensors**: the piezoelectric principle is underpinning: the weight of the vehicle causes a change in material's density, generating an electrical charge as a potential difference between the electrodes it is equipped with. [17] Then the system, starting from the frequency and intensity, goes back to the speed, the weight or the count (thanks to the proportionality ratio between deformation and the two magnitudes);
- **Magnetic/inductive loops**: the operative principle is based on electromagnetism, namely a current that generates an electro-magnetic field is induced, then the variation of inductance, generated by the passage of cars within this field is exploited as a measure to detect the speed or classification of vehicles [17];

- **Passive magnetic sensors**: another type of sensor placed under the road surface that can return relatively basic information with precision, such as counting and speed;
- **Ultrasonic Sensors**: sensors installed in the roadbed that emit sound waves that cross the roadway and find information about the vehicles that are hit by the particles, sending them back (figure 3.4). [17]

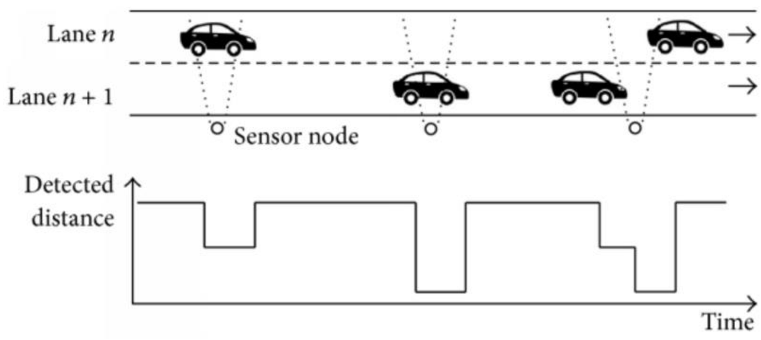

Figure 3.4 - Vehicle detection using ultrasonic sensors. Source: Jo, Y., Choi, J., Jung I., Traffic Information Acquisition System with Ultrasonic Sensors in Wireless Sensor Networks, 2014, p. 1-12

3.3.2 Traffic cameras

Method which nowadays, it can be considered conventional so far, è the traffic camera source (figure 3.5). These are devices placed along the roads, usually the most congested, which record the traffic of cars and provide crucial information such as traffic conditions (e.g. critical crossings, bridges or tunnels), support information for traffic lights management or speed control, incident management.
Ultimately, they are representing a key tool for ITS or Smart infrastructure solutions. [12], [16]

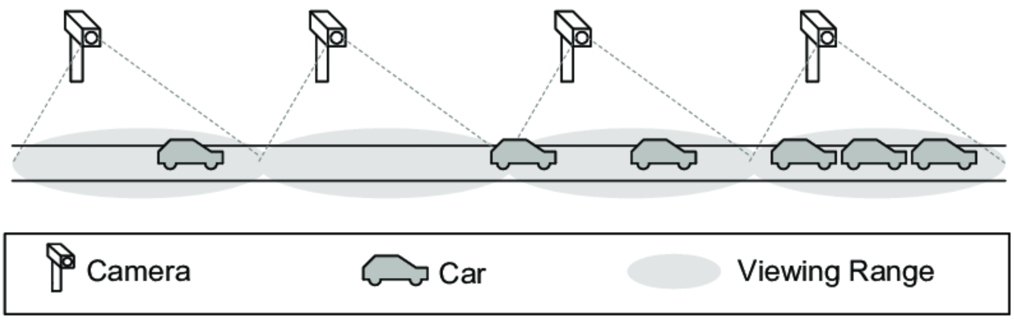

Figure 3.5 - An example of a highway with traffic-cameras. Source Weyns, D., Haesevoets, R., Helleboogh, A., The MACODO Organization Model for Context-driven Dynamic Agent Organizations, 2009, p. 1-25

3.3.3 Others conventional

In this section we will briefly introduce alternative conventional methods that do not fall within the previous categories and which, by their nature, are obviously not methods that provide information in real-time, which is why they are now sparsely used.

- **Manual counts**: by observing the traffic of vehicles through video, they retrieve complex information that (the earliest) software could not gather otherwise, such as classification, vehicle occupancy or information on roadside agents (e.g. pedestrians); [17]
- **Household Travel Surveys (HTS)**: these are questionnaires sent to a certain sample. They were used for long-term goals regarding the network, policies. [12] They are an expensive method of data collection (sometimes in fact the number of surveys sent is relatively low to save money and this generates untrustworthy and unprecise information) and now they have been definitively replaced by Floating car surveys (FCS) more efficient, economical and reliable.

A general resume is offered in table 3.1.

Table 3.1 - Type of data provided by the count technologies. Source: [17]

Detector Type			Volume/Count	Speed	Classification	Occupancy	Presence
I	Inductive Loop		✓	✓ (1)	✓ (2)	✓	✓
	Magnetic		✓	✓ (3)	✓ (3)	✓	✓
	Pneumatic Road Tube		✓	✓	✓	×	×
N	Active Infrared		✓	✓	✓	×	×
	Passive Infrared		✓	✓ (4)	✓	✓	✓
	Microwave Radar	Doppler	✓	✓	✓	✓	
		True Presence	✓	✓	✓	✓	✓
	Ultrasonic		✓	×	×	×	✓
	Passive Acoustic		✓	✓	✓	✓	✓
	Video Image Processing		✓	✓	✓	✓	✓

3.4 New data collection sources and methods

As mentioned in the conventional source of traffic data, in particular on-road sensors, they are still an important slice of the technologies used, and are thereby still necessary and fundamental so far. However, it would be dull not to exploit the opportunities that the current innovation offers, as it can constitute an important support and evolution, both from the point of view of the evolution of traffic data requirements (e.g. in terms of coverage and reliability), and in terms of costs (of implementation and maintenance).

As noted by [12], in fact: *"Over the last decade, advances in information and communications technology (ICT), particularly wireless communications technology, and burgeoning uptake in mobile telephone and GPS-enabled devices, has created interest in how These new systems may be used to provide traffic data."*

This section attempts to portray the traffic data collection methodologies that most have spread in recent years and which will constitute, or already constitute, new important sources of information that meet the necessity of more accurate, real-time and reliable information that today's context requires. It will be reviewed both sources involving road-side sensors (which get information from point-to-point surveys), and so-called probe-based sources, i.e. GPS-based, mobile-based and Bluetooth.

3.4.1 Side-road sensors

The development of wireless technology has led to the gradual introduction of devices that, given their peculiarity of being able to obtain information without direct contact with vehicles, can be placed at road side and possibly installed on pre-existing structures. For this reason are also called non-intrusive methods, such that they are based on observation which are far-off, which therefore do not cause an obstruction of traffic, neither heavy installation or maintenance works, since they are not to be installed in the roadbed. [16], [17]

- **Passive and active infrared**: these non-intrusive sensors are laid at the roadway's side and work by exploiting the emission of a high-energy (infrared) beam in the direction of the vehicle's flow. The passage of the vehicle constitutes an obstacle to the beam and causes the return of the particles emitted to the detector, which, according to their characteristic, returns information on the speed, counting or classification of the vehicles. This source is however susceptible to bad weather and the readings lose effectiveness during the night, also the road cover that offers is not optimal; [12], [17]

- **Microwave radar**: these sensors overcome the limits of the previous ones because, due to their operating principle based on the emission of high frequency radio waves (doppler radar), they are not affected by weather conditions, nor by the lack of light during the night-time. [12] They allow to detect the traffic flow and speed, in particular the high frequency allows to record precise vehicles profiles and, consequently, an accurate classification. The weaknesses are due to the fact that they suffer in the case of hidden vehicles and that require considerable efforts for the calibration;
- **LIDAR sensors**: non-intrusive sensors whose functioning does not differ from that of the previous microwave sensors, with the difference that here the waves cast belong to the ultraviolet spectrum (visible). They share all the strengths and weaknesses of the former; [12]
- **Passive acoustic sensors**: the passive acoustic sensors are usually installed on the road of highly trafficked roadways and whose operating principle is based on the detection of the sound (of the wheels) of the vehicles circulating. In this case the achievable information is counting, speed and classification. [17] As well as the operating principle they share the same weaknesses of infrared and ultrasonic sensors. They are an expensive technology both in terms of set up and maintenance, so the investment is justified only in the case of long-term installations; [12]
- **Video image detection**: this method consists of the use of so-called CCTV cameras placed along roadways to film the traffic of an affected artery. Afterwards, thanks to specific software (e.g. OCR-Optical Character Recognition), the images from the video are analysed (there are different video techniques [17]) to trace information such as the classification of vehicles and speed (figure 3.6). They are a key tool for incidents, traffic monitoring and license plate recognition. [12] They are a widely used source for the traffic data collection used extensively in European networks and among their strengths we may find the low maintenance costs, while the main weaknesses consist of the huge initial investment and low reliability during the night or with bad weather. [16]

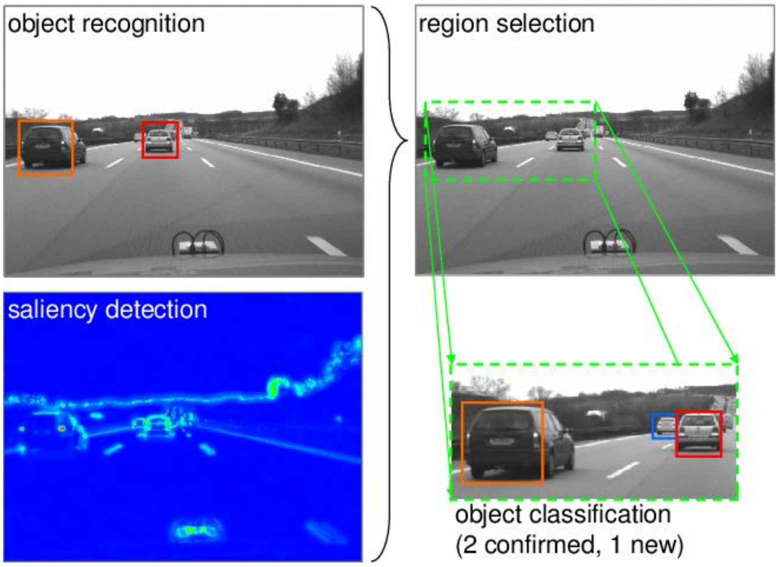

Figure 3.6 - Object recognition and detection through algorithm from video frames. Source: Hermann, A., Matzka, S., Desel, J., Using a proactive sensor-system in the Distributed Environment Model, 2008, p. 703-708

- **Weight in Motion systems (WIM)**: They are non-intrusive systems specifically dedicated to the continuous detection of data concerning the weight of vehicles, particularly used in the case of critical sections such as bridges or overpasses intensely trafficked and with limited capacity. [16]

3.4.2 Probe-based collection technologies

An in-vehicle and wireless based technology that surely has proven to have a tremendous potential and that today has reached full maturity today, is the so-called Probe-based technology, also called Floating Car Data (FCD). It constitutes the present and future of the data collection from the road network, as well as being crucial to the development of ITS systems.

As the name recalls, it bases its principle of operation exploiting the vehicles circulating in the network as probes, i.e. as real sensors, to acquire wirelessly information, transmitted by mobile devices (through GSM/3G/4G, GPS or Bluetooth) or specific GPS devices (e.g. navigators) inside the cockpit, allowing identification and tracking. [12], [17]

It is clear that, while previous sources used wireless technology for short distances, this data collection mode leverages a long-range wireless communication, taking advantage of the mobile network infrastructure as a means for finding information. For this

reason, it is also called mobile-based technology and the acronym FCD may be interpreted as floating Cellular Data.
To date, such data collection source is the most efficient, effective and convenient support solution (and sometimes substitutive solution) to expensive road based sensoristics and their limits, which, in any case, still constitute a key basin for traffic data collection up to now.
The probe-based approach is extremely convenient especially for the fact that it does not require:

- Investments in specific infrastructures situated along the network;
- Calibration, installation and maintenance.

Conversely, it also allows to obtain information:

- Reliable;
- Continuously;
- Real-time;
- High quality;
- Accurately (excellent from GPS, discrete from mobile);

regarding position, direction and speed (with sometimes a lack in precision), from which can be extracted, once elaborated an processed, useful information such as traffic volume, travel times and speeds, OD matrices, delays, incidents etc. [12], [17]
The floating car data method can be declined in 3 ways to collect raw traffic data.

GPS - BASED FCD

By GPS-based FCD we mean those raw traffic data that can be obtained from the presence of devices equipped with GPS (Global Positioning system) within the vehicle (figure 3.7). Until a few years ago the owners of these devices were pretty much professionals and companies, used for example by taxi drivers or couriers of logistics companies (e.g. fleet Management Service), while they were usually present less frequently in private vehicles. For this reason, the traffic gained by GPS, were transmitted to the processing centres in a massive and frequent way in the neuralgic or urban centres only, given the concentration of taxis and deliveries, while the measurements were less consistent in the rural environment. As a result, the processing of data from GPS devices generated useful information (precise and reliable) only for urban traffic and routes, and only for specific network users (professionals) classes, therefore not generally valid for all. However, while in the cities the main generators of info from GPS are taxis (who use it to work), on the other hand are mainly private individuals to have a crucial role in sending information on the traffic of rural areas or highways as well as, of course, the couriers.

The data service providers (essential supplier for logistics companies) use GPS based data as a primary source (by tracing GPS device paths), combined with other traffic data from a fixed location (e.g. offered by the Transport's department), to yield information on travel times, delays incidents and best routes, then sold to logistic enterprises for their fleet management. [17] Whenever providers exploit this mode they may be called wireless service providers because they systematically collect via wireless traffic data from probe vehicles to extract relevant information. [16]

Being devices specifically dedicated to geolocation, the accuracy of the data sent is remarkable (up to meter) and certainly greater than the localization via telephone operator. On the other hand, however, they are a relatively expensive technology and often do not constitute a sample subsequently representative to be reliable, as the vehicles that are equipped with it are not numerous. This could be considered true until a few years ago. In fact, this mode is more than ever exploitable so far, since with the rise of smartphones, worldwide spread, mobile phones are equipped with a multitude of functionalities (Bluetooth, GPS, Wi-Fi, Weather info etc.) and especially GPS, hence they provide and require continuous feedbacks and updates for geolocation. [12] Today many navigation apps are available for smartphones (easily downloadable) which make use GPS, with user-friendly interface, such as Google Maps and Waze.

The principle with which the data of GPS devices operate, may be resume by [12]: *"A device is located via GPS technology when that device bound signals directly to three or more satellites and the calculated position is sent either to that device's receiver or to another Receiver."*

In general terms, the GPS is a device that establishes a bidirectional communication with a traffic data provider and receives and collects the coordinates of its position from satellite, then this information is sent directly to a processing centre (or the centre itself capture it), as shown in figure 3.7, that collects, analyses and reprocesses appropriately transforming into useful information for network users, to which are later redistributed. [16], [17] Among the weaknesses that characterize the GPS probe-based data collection we can identify the possible network overload (issue perhaps solved by big data paradigm), the presence of obstacles that can degrade the transmission, the possible inadequacy of mobile devices to transmit or receive information (Insufficient power load) and privacy issues. [12]

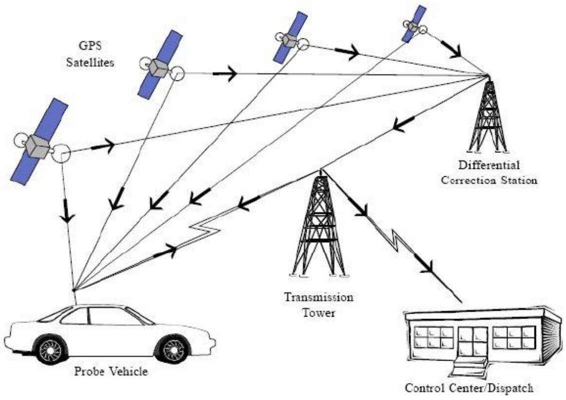

Figure 3.7 - Communication through GPS. Source: [17]

MOBILE BASED FCD

To introduce this methodology is worth considering [17], who states: *"Since nowadays most of the driving vehicles are equipped with at least one or several mobile phones, it may be worth using mobile phones as anonymous traffic Probes"*. Mobile based method is therefore a method of floating Car based on cellular phones (at least switched on) in cars, used as sensors, exploiting the cellular network data to detect raw information from the road network (figure 3.8). The signals, which encompass information, that the antennas and mobile phones transmit continuously each other (for calls and other data), are then exploited to intercept the devices, consequently the probe is localized in space and time by the use of telephone infrastructure (GSM/3G/4G/5G). [12]

By GSM we mean Global System for Mobile Communications, the standard (first generation) protocol, initially European then extended almost globally, at the base of telephone networks for mobile phones or small devices. With the technology progress, the evolution of the requirements and potential of mobile devices, the second generation, GPRS, was replaced first by the third, 3G (UMTS standard), and finally the fourth, 4G (LTE advanced standard) currently in force. With the evolution of the generations, the efficiency of the network has also increased and consequently the consistency (understood as precision, reliability and coverage) of the FCD mobile-based method.

And this basically consists of the "conventional" part of the method since the use of probe vehicle through mobile data dates back to the mid-year 2000. To date, as already mentioned, innovation lies in the fact that we are able to extract geolocation data from smartphones GPS system, of which they are nearly universally equipped, rather than through the network of telephone operators or special hardware devices.

Moreover, as noted by [17] , unlike the methodology based on GPS the cellular based method has on the one hand, additional advantages:

- Does not require the provision of in-vehicle hardware systems;
- Assures a huge degree of coverage;
- Covers all network users' classes;

On the other hand, new weaknesses:

- More algorithmic complexity to process the information (to be distributed then to users);
- Less precision (approx. 100 m; proportional to network accuracy);
- Privacy issues;
- Ambiguity of the number of mobiles in the vehicle (in-vehicle users vs non-vehicle users);

The geolocation of a phone is due to the so-called triangulation, i.e. exploiting the fact that even when the phone is actively used (but switched on) signals are exchanged with and through networks bidirectionally (figure 3.8); consequently, these signals vary their direction with the movement of the probe through several cell-sectors. By analysing this data information about traffic conditions can be gathered, e.g. information about traffic jams may be retrieved, and made available to users, whether more probes (therefore more devices) are detected in a given intersection.

The triangulation may be performed through different modalities and techniques. The passive monitoring and active monitoring techniques are worth mentioning.

- **Active monitoring**: The signal of the mobile device is picked up real-time, as a result is recorded the path that the probe accomplishes;
- **Passive monitoring**: It traces back to the routes traced by the car in an indirect way, consulting the data collected by the telephone operators.

Regardless of the technique used it is not exempt from inaccuracies, due to the degradation of transmissions, in the case of same antennas serving different roads, or the unperfect continuity of the signal (towers). It is straightforward that in urban areas the accuracy will inevitably be greater, given the greater concentration of antennas. [12], [17]

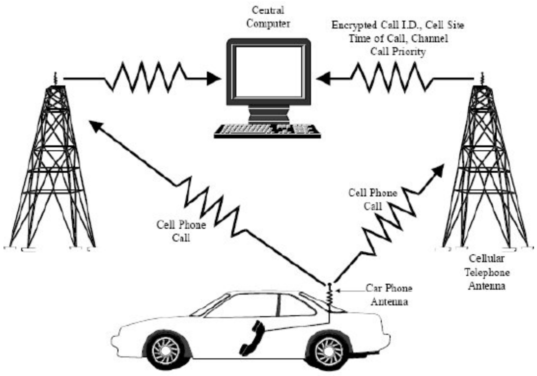

Figure 3.8 - Communication through cellular phones. Source: [17]

BLUETOOTH

Bluetooth technology is not a newly discovered technology, what was innovative was the idea of applying it to collect traffic data from the network. It turns out to be a hybrid between probe based (in vehicle) and on-road source: Because on the one hand it exploits the wireless nature of the Bluetooth signal, of short range, to trace the information of the probe (like average travel times or speeds), on the other it requires the Construction of Bluetooth detection infrastructures along the roads, which can instead provide information such as OD paths and traffic flow (typical on-road sensors purposes). [12] This field appears to be the more applicable and effective the more we think about the enormous expansion that this technology has had in all the industries and applications during the last decade; in fact, now all the devices are equipped with such technology, e.g. televisions, radios, navigators and, in particular, also cars and Mobile phones.

Consequently, this also explain the choice of inserting it into probe or mobile-based technologies category. If once the main weakness of this source was the insufficiency of vehicles equipped, nowadays the times are ripe to say that it is highly competitive. No wonder it will be a pillar of the traffic data collection methods in the coming years.

In essence, this source is based on a Vehicle re-identification, on the one hand because obviously it requires the presence of detectors that establish the identity (address), on the other because the passage of the probe is recorded a first time, then re-identified later (figure 3.9).

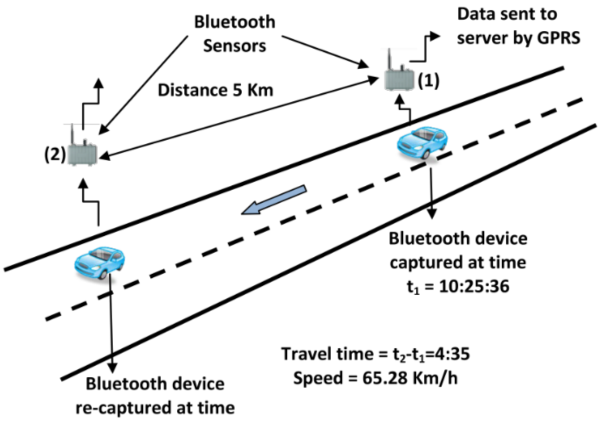

Figure 3.9 - Vehicle monitoring with Bluetooth sensors. Source: Barceló, J., Montero, L., A Kalman-Filter Approach For Dynamic OD Estimation In Corridors Based On Bluetooth And Wifi Data Collection, 2010, p. 1-29

RADIO FREQUENCY IDENTIFICATION (RFID)

This technology is based on Automatic identification and Data Capture (AIDC) technology. Namely RFID is used, in the logistic-distributive field, as System AVI (Automated Vehicle Identification). These in-vehicle systems have acquired consensus especially in the last period in the industrial-distributive field. The growing need for traceability of goods along the supply chain (especially in Lean production optics) has found solutions in these systems based on identification tags (serial numbers) placed on goods transported by couriers, and which communicate their location in the network through radio signals, unambiguously thanks to the uniqueness of the MAC address (with the same mechanism presented for Bluetooth). A future potential could be to exploit the RFID technology to go back to metric OD, travel times and traffic conditions but this would entail massive investments of RFID detectors along the roadways, to date not economically feasible. In fact, fundamental difference with Bluetooth technology is that it does not provide the same coverage, since it is not commonly found in private cell phones (as said requires electronic tags), therefore it would represent only a particular “Business” class and, consequently, would be useful only to them. Although not useful as a source for traffic information so far, constitutes instead a reality in the Supply chain and Logistics field. [12], [18]

AUTOMATED VEHICLE LOCATION (AVL)

A similar method consists in the so-called Automatic Vehicle Location (AVL), which differs from the RFID since instead of electronic tags, general transponders installed inside the cockpit are used, which transmit a signal to transmitters located along the roadways.

3.4.3 Other new sources

Some new sources that do not fit into the previous categories are presented below. The detailed description of them is beyond the scope of the thesis so they will only be mentioned.

- **Wi-Fi**: Recent studies are considering a solution (still at the experimental phase) to use Wi-Fi of telephones present in the couriers as an alternative to electronic tags. Wi-Fi technology, now universally spread, is based on a wireless communication and can characterize a network, which in this case consists of a real telecommunications network and might be seen as a small mobile network, whether connected to the Internet (through special access points). Hence, in this regard, it is evident how it can be exploited as a possible source of traffic data, by harnessing a mechanism similar to that seen for Bluetooth and RFID. It also allows the transmission of data via satellite as well as obviously via cable (ADSL). The idea is that a series of hot-spots (comparable as small towers with a signal's range of 100 m, possibly extendable), positioned across the network, can act as a detectors of Wi-Fi mobile devices in vehicles; furthermore the Wi-Fi addresses are fitted, like the previous two technologies, of MAC addresses consequently they are uniquely recognizable. In addition to investment and the design related problems of such infrastructures, this source incorporates problems involving privacy and the protection of transmitted data.
- **XFCD (extended Floating Car Data)**: The idea of the extended FCD is nothing more than an extension of the already described FCD method for Data collection approach. The extension lies in the fact that the technologies of which are equipped both the current cars and the mobile devices, allow to manage a heterogeneity of real-time and high-precision data. This information, may still be potentially useful to gather more significant information to enrich the knowledge about the current

network condition. In practice this means exploiting, not only the GPS position of the probe, but also different environment related information. Some examples of useful and potentially relevant information in this regard are offered by [17]: *"In-vehicle information can be used, for instance, to report immediately about traffic jams, detect weather conditions (e.g. data from the activation of windshield Wipers, temperature sensors and headlights), road surface state (e.g. the operation of ABS system can be used to detect slippery road conditions, risk of aquaplaning or black ice) and many others"*, as well as pollution related information, noise pollution, fuel consumption or other data from smartphone's sensors (accelerometers) to achieve further speed, road or density data. However, the extension of this data collection method is only conceptual, as it does not require the installation of additional hardware to operate, but software suffice. Through these latter it is also possible to filter the information to be communicated. In this sense, the quality and quantity of incoming data to the elaboration centres are enhanced, therefore, this leads to a reduction of the transmission costs, and, consequently, a greater affordability.

- **Floating car Surveys (FCS)**: This method, also called GPS based Household Travel surveys, has replaced the Household Travel Surveys (HTS), since as opposed to the latter, can find real-time data (concerning Travel times, delays and speeds) and is cheaper and easier to conduct. Another great advantage is the coverage of the information that GPS data can provide, which is considerably more extensive than any coverage offered by traditional questionnaires. [12] The only drawback may be the ambiguity of the number of active telephones in the vehicle (which can be obviously multiple).

3.5 Combining Data

The plethora of data sources seen so far, including big data, is evidently an essential resource for having a picture of the current situation and collecting raw data from the network. As seen in the previous sections, however, the sources differ, in terms of operating principle, time intervals, provided information and formats etc. For this reason, they are combined (data collected, processed and merged) to give life to more complete and consistent information, with the ultimate aim of offering as much as possible a personalized service and performance to decision makers. The following section proposes to present the problem of combining data, articulating it in two directions: to combine historical data and real-time data, and to combine the information gathered by different sources.

3.5.1 Combining Historical and Real-time data

Given the abundance and variety of data sources would be a waste not exploit this wealth of information, since then it would translate in more performing algorithms that are the basis for efficient and innovative routing (as presented in chapter 2). The Dynamic Traffic Assignment (DTA) is a *"framework for off-line and on-line calibration of DTA models"* [16] that offers not only the possibility to integrate this variety of sources, to extrapolate an improved information for the benefit of innovative model systems, but also to combine the data collected with historical data available, which are undoubtedly an essential asset. In this sense, the richness of information would inevitably result in more precise calibrations, therefore more adherent to the current situation. For the sake of this thesis this approach will be described, without deepen the details, also given the abundance of literature.
This framework combines the historical data with those of different sensoristics, through the use of network supply and demand simulators, to achieve two goals. [16] The first phase is called State estimation and consists in the exploitation of the sensors coverage in the network to gather fine-grained, real-time information and compare it directly with the available historical data. The ultimate objective of this phase is, starting from the above data, estimating (inferential statistics) the current state of areas not covered by sensors. The second phase consists in the elaboration of the collected data, then in the production and redistribution of relevant traffic information to the network users; for this reason, it is called Prediction-based information generation.

HISTORICAL DATA FOR OFF-LINE CALIBRATION

Off-line calibration draws from historical data (detected for a certain period of time: days, weeks, months up to year) to estimate variables, parameters and functions concerning mobility and network conditions. Deriving from archived data, so not real-time, these sets will reflect the reality of an "average historical" network and not a current state. Consequently, for the construction of an offline model the calibration will use conditions of expected flows, or matrices and traffic conditions, but probably not corresponding to the reality. Sometimes the stratification of historical data is applied to improve the quality and consistency of the estimates. [16] In essence, a characterization of the data is performed in the archive and, consequently, it is possible to take advantage of stratified sets. In doing so, despite historical estimations, adherence to a particular scenario of the network is reinforced. The importance of a non-real-time calibration is emphasized by the fact that, as we know, the current state of the road network is highly volatile, especially nowadays, so it is important to generate and have historical data sets and trends for wide scenarios range and situations. The more you have a rich and stratified database, the more the inputs will be reliable, and finally, the more accurate the estimates generated in situations of uncertainty will be. It, in fact, constitutes a fundamental baseline patrimony for new models, which, possibly, draw from on line calibrations to fill misalignments towards the current state.

DYNAMIC DATA FOR ON-LINE CALIBRATION

The traffic models can also be calibrated online, in the sense that you are directly in line with the network, thus the various sensoristics send real-time data to the model, which therefore receives dynamic inputs. In practice, a continuous comparison is made between the measurements in progress across the instrumented road network and the inputs (derived from the off-line calibration) initially fed to the model, in order to adjust the deviations. The resulting estimations not only happen to be more adherent to the real traffic conditions, as real-time data collection, but also more high-resolution. [16] Operating in this dynamic way the model ensures output, understood as traffic estimations, more accurate and reliable and, consequently, allows a more efficient and effective decision-making process. In conclusion we can say that the combination of historical data and real-time data assumes perhaps the clearest example in the calibration phase (see figure 3.10), phase as critical as essential for each traffic model. In conclusion, the considerations above show that, in the first place, although nowadays the need for real-time information is essential, the historical data continue to play a key role (especially in case of lack of direct information); secondly also these historical data can take advantage of the variety of traffic data sources that we now have.

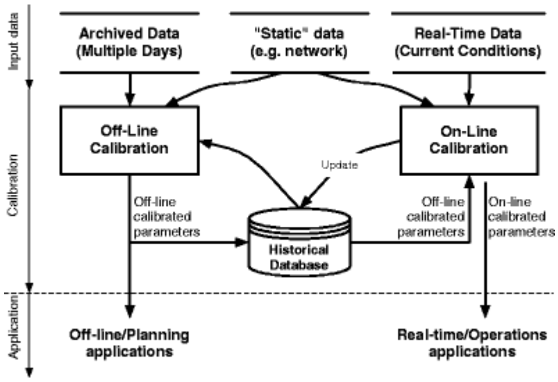

Figure 3.10 - The role of off-line and on-line calibration. Source: [16]

3.5.2 Combining different data sources and formats

It is out of the question that the broader the data sources basin from which it draws is, the greater the benefit in terms of coverage, precision and reliability of the data produced will be. As mentioned, in fact, not only a greater amount of information is collected, but we can combine different traffic data information from different sensors, allowing to acquire a holistic knowledge from the network. On the one hand because they are compensatory, on the other because, on equal terms of traffic data detected, some sensors ensure greater accuracy of others in function of different environments. For example, traffic data speeds and travel times are achievable accurately from Bluetooth sources (between two points along the road), traffic density, volume and flow data conversely are more accurate when obtained from point-based detection devices (on-road) such as magnetic loops or piezoelectric sensors, still particularly accurate route choices and OD flows when collected via GPS-probe sources, weather information are retrievable from traffic cameras as well as traffic jams length and incident monitoring; microwave radars allows to accurately detect the profiles of vehicles by helping their classification or still mobile based data are more reliable in urban environment (for the greater density of antennas), but they are not a source infallible in rural areas.

Moreover, the combination of traffic data sources, although it makes the network more controlled and structured, introduces a complexity related to the treatment of disparate data, requiring to be integrated and processed properly before extrapolating a usable and useful knowledge to the various network users and stakeholders (figure 3.11). Another complexity is related to the fact that not everyone needs the same latency that undoubtedly depends on the uses that the various users make, clearly incident related information require the almost immediacy of the data (real-time surveillance), while planners do not have such stringent needs (overall network occupancy and travel times). [12]

The problem of data fusing is not new in the context of transport, in fact there is no shortage of methods to treat such data, especially among the data providers that, through these, merge, for example, traffic data gathered by sensoristics and floating cars data, to extract knowledge and sell the service information. Mathematically speaking, as [12] observes: *"Typical data fusion methods are based on various statistical averaging techniques, such as simple and weighted convex combinations, Kalman Filters, ordered average weighting and artificial neural network Estimation methods "*. The mathematical complexity beyond these algorithms is due to the fact that to be effective and return useful outputs to individuals and businesses, they require a real-time resolution.

Figure 3.11, taken from [12], summarizes the overview.

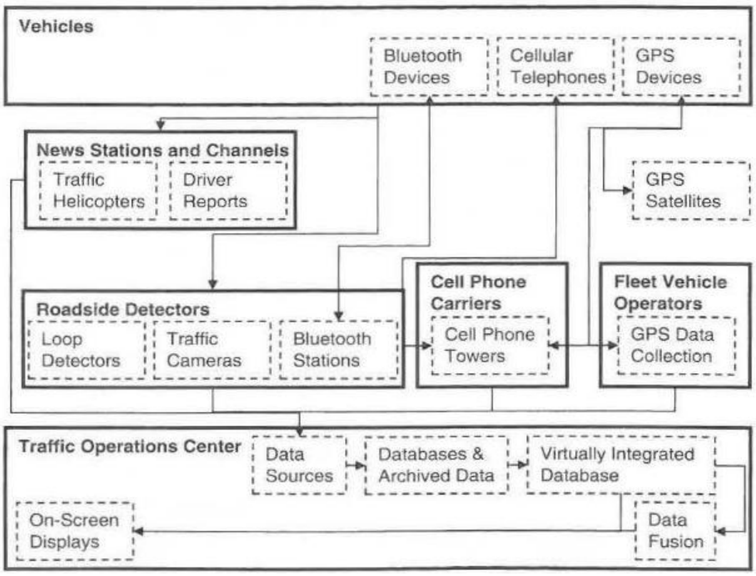

Figure 3.11 - Exchange of information between network and operation center. Source [12]

3.5.3 Combining sensors: a practical example

To give substance to what was said is aimed to present an application example developed by [15] .The authors propose an innovative traffic estimation model, integrating GPS data, Bluetooth and loop detectors, to estimate traffic in the urban environment, which is characterized by radically different dynamics from those that typify the roadways, and, consequently, which require a different approach (model) than those dedicated to them (more frequent in literature). For the sake of this thesis we will take advantage of this example by exploiting it to demonstrate and verify the actual benefit of combining different traffic data sources. This project proposes the integration because both the richness of data and the collection of heterogeneous data, are considered key important to ensure an improved consistency, resolution and robustness of the model, then, ultimately, the generated outputs. In particular, they adopt an incremental EKF algorithm incorporating a specific LWE model for urban environment, setting the hypothesis of independent detections with respect to the state of the system and its inputs [15].

The idea behind the model is to divide the urban corridor into Ds-size cells and Dt time intervals; by doing so a demand has been defined, understood as the amount of cars requiring you to enter a cell in a certain time, and a supply, understood as the amount of cars that can be accommodated in that cell. [15] In the mathematical processing, functions, parameters and variables to characterize the system are progressively introduced, such as density and flow functions and parameters such as free-flow speed, critical density and jam density, as well as independent variables such as space and time.

Note that sources have different characteristics in terms of traffic variables considered, latency and operational principle:

- **GPS Source**: detect time-stamped data concerning instantaneous speeds, with wide coverage, at variable order intervals depending on the device used;
- **Bluetooth Source**: detect average travel times and speeds by dual identification of the vehicle, at intervals from the order of seconds (available only when the scanned section is complete);
- **Loop Detectors**: dedicated to vehicles count, make data available with a latency of a few minutes.

The combination allows a holistic picture of the current traffic state, but at the same time the obvious differences constitute a complexity that has to be managed by the model, especially when you consider that the estimation latency request can differ from the intervals at which the data are available (which are equally variable, as just seen). At the end of the experiment considerations are made on how the different sensors influence the output estimates (see table 3.2). It turns out that the estimates derived from single source traffic data are particularly lacking using only Bluetooth, while the GPS alone ensures a discreet accuracy. The combination of both does not lead to sensitive improvements, conversely, the combination of GPS and magnetic loops, produces consistent increments of performance. However, the most important and expected result is that, instead, the merge of the three data collection methodology leads to the most significant percentage increase of accuracy (table 3.2).

Table 3.2 - Absolute mean similarity (AMS), relative mean similarity (RMS) and standard deviation similarity (SDS) similarity scores, penetration rate of 20%. Source: [15]

Data fused	AMS	RMS (%)	SDS
BT	0.27	45.32	0.07
GPS	0.36	59.36	0.10
BT + GPS	0.33	55.69	0.09
LD + BT	0.33	54.43	0.09
LD + GPS	0.46	77.29	0.06
LD + BT + GPS	0.46	76.81	0.06

3.6 Traffic Data Providers

Data providers are responsible for providing information to network users that require traffic data for different purposes. They consist of are research institutes, industry associations and, sometimes different consultancy companies, which are paid by different enterprises to supply data necessary to operate their business. Therefore, they are clearly pivotal partners for the logistic-distributive companies because they provide the data (dynamically or not) for the routing algorithms. Data providers who can provide reliable, accurate and cost-acceptable information are the key for ensuring a certain level of service offered to the customer. They combine and make use of all the methodologies, paradigms and technologies seen to capture, transmit, analyse and distribute the network data, in order to have as much as possible consistent data, for the sake of producing estimates as accurate as possible. In this section are portrayed some of the European main data providers, proposed by [12], [17] and are briefly presented their distinctive features.

3.6.1 Main European data providers

- **HERE**: it is among the most widespread data providers in Europe and worldwide, with tens of thousands of remote sources in almost 200 countries. [12] Provides driver services with maps, real-time current state updates, Maps services, best-route and collective transportation services. It works both in mobiles and cars framework, even though it is especially used by business customers given its accuracy and the extensive services.
- **TOMTOM**: company that manufactures Hardware and navigation software for cars and mobile. It boasts combined on-road and side-road sensors across the road network (mainly magnetic loops), FCD (both cellular-based and GPS) and transport agencies that offer reliable information on almost the entire motorway network and numerous other roads. [12], [17] The services offered vary depending on the package purchased, and range from Routing Navigator, to live weather conditions, to speed detectors, to traffic, travel and incidents updates. In Some cases, it is adopted by some cities to detect daily traffic data conditions.
- **INRIX**: data provider boasting to be a leader in EU and USA, with hundreds of business clients. It is a provider of apps, real-time information from the network (weather, incidents) and estimations concerning weather, ongoing works and traffic. [12] The information is collected through probe-based methodology, engaging a broad spectrum of networks users (taxi, trailers, couriers, shuttle, privates), and instrumented network (live traffic information). To make INRIX a complete service also contributes a consistent historical database that allows you to have time-dependent information (trends) and concerning special events.
- **ITIS Holding**: the most famous British data provider. It boasts an efficient data gathering and processing system that allows real-time picture of the network traffic situation. The majority of traffic data source is based on the FCD method, which therefore assures a large data coverage and to detect information concerning travel times, or OD matrices. In this context, probe-based network users and consumers contribute to the system by being both generators and users of network information.

- **Google Maps**: although not being, among the providers, the one that offers the best accuracy and services especially for business requirements (fleet management etc.), turns out to be a world-wide famous application and software, known for its immediacy and friendly user interface, thanks also to the compatibility with most mobile devices on the market. Google Maps gathers network traffic information basically via two methods: GPS data from consumers-users (enabling location sharing) and data from third parties, such as transport agencies, compensating lack of knowledge in case of portions of network with poor representation of smartphones. As a result, as in the case of other data providers, the customers are both provider and data receiver. [17] It offers navigation services, live updates of traffic jams, travel times, and best route choice according to criteria chosen by the user, such as, shortest travel time, shortest travelled distance, fuel economy, no-fees road; e not only by car, in fact, the same services are offered for public transport, by foot paths or taxis. [12] The Google Maps platform is constantly updated and many improvements have been operated over the years, for example, not only a point of interest on the map may be chosen as destination (by Category: Supermarket, gas station, train station, Church, ATMs etc), but even the proposed best route takes into account the arrival within the opening and closing times of the same (when this information is shared). In such case, algorithmically, a Vehicle Routing Problem with Time Windows is resolved.
- **Waze**: like the previous, it exploits probe-mobile of customers who utilize the corresponding app for gather information across the road Network. Once collected this information, they are then redistributed in the form of real-time services such as traffic flow, accidents, weather. The accuracy of this app is greater than in Google Maps, especially for the resolution to very minor roads (compared in figure 3.12). [12] The principle below this system is a community in which each driver contributes, through the use of the app, to compose a current state picture in the Waze system, sharing the events he witnesses across the network, like a dynamic puzzle. And this can happen either in a passive way (with the only activation and use in the background of the app on the smartphone), or active way (manually inserting the event encountered).

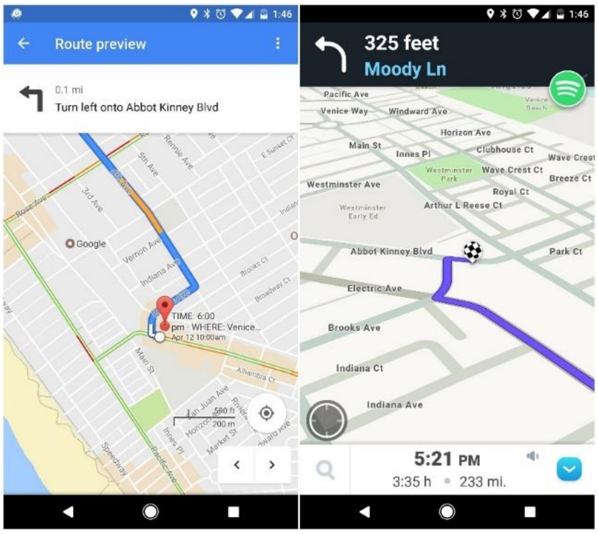

Figure 3.12 - Comparison between Waze and Google Maps digital visualization. http://www.globalsupportinitiative.com/get-directions-google-maps-by-car/waze-vs-google-maps-100717695-orig-get-directions-by-car-5/

- **Telstra**: specialized data provider for business customers offering customized routing, navigation and live network information services with a wide variety of applications including, in particular, fleet management.
- **Optus**: a company that collects network data through the signals that mobile phones in vehicles cockpits exchange with the telephone operator infrastructure. This data is then analysed and redistributed to the subscribed customers, as relevant information such as routing services and traffic updates [12]
- **Vodafone**: as one of the most widespread telephone companies in Europe and worldwide, it exploits its coverage to capture traffic data via Floating Cellular data method, using subscribers like probes, hence, like data provider. [12] Nevertheless born as telephone operator, nowadays it also plays an important role as a service provider by processing the data obtained and making available to customers smartphones network related information.

Chapter 4: Advanced and Compelling technologies

In this chapter are presented, in order of interest, modern technologies for Routing's management; furthermore, is proposed how each technology can contribute to the system architecture. Starting from a presentation and explanation of the broad world of Cloud and Big data, moving from the explanation of the computerized Geolocation world and Mapping, up to the radical advent of IoT (Internet of Things), to reach finally the main topic and goal of study, namely the development of a SaaS (Software as a Service) system.
The chapter concludes with a brief discussion about the emergent technologies, which are not mainly used yet in the practice, but, however, of possible scientific and practical interest over the coming years.

4.1 Software as a Service (SaaS)

Cloud Computing and Software as a Service (SaaS) are two closely related concepts that, on the one hand, meet the new business needs, on the other they arose thanks to ICT technological advances, which allowed the opening to Web-based and wireless applications. Below is given an overview of the context in which the new concept of Software as a service and, as a result, of cloud computing have taken hold.

CLOUD COMPUTING DEFINITION

By Cloud computing we mean that infrastructure which allows users to access remotely a platform of shared resources, whose source and actual location is unknown by the latter, which may consist of servers, networks, services, or applications, requested on-demand.
The tendency not only to require services (e.g. CRM or storage), but also to manage applications, on the Internet is nowadays an universally observed trend, and that is opposed to the previous "on-premises" services, especially as regards the management of maintenance, upgrades and development that in the new paradigm are made dynamically and scalable by the cloud service provider. To date, users prefer to disengage from local possession, preferring to request the service "as needed". [19] Consequentially, it can be observed that cloud services release significant initial

investments; are characterized by scalability, rapid delivery and development; the cost depends on the subscription; they are accessible by a multiplicity of users (called multi-tenant). In everyday life we can test the potential of such cloud environment in the usage of e-mailing sites or e-banking, instant messaging, web marketplaces or social media. In particular, the latter are called public cloud networks, as they have no access restriction and are commonly used by the internet community. Conversely, specialized providers may construct private clouds which are usually used for business purposes. A company that operates in this way works in a so-called context of the SaaS (Software as a Service), which will be discussed in the following sections. In particular, among the most famous cloud framework we find: Java Google Web Toolkit and Python Django (both Google App Engine), Microsoft.NET (Azura Service Platform).
In the panorama of routing, applications such as fleet management software can now be requested on-demand through the cloud platform (figure 4.1). [20]

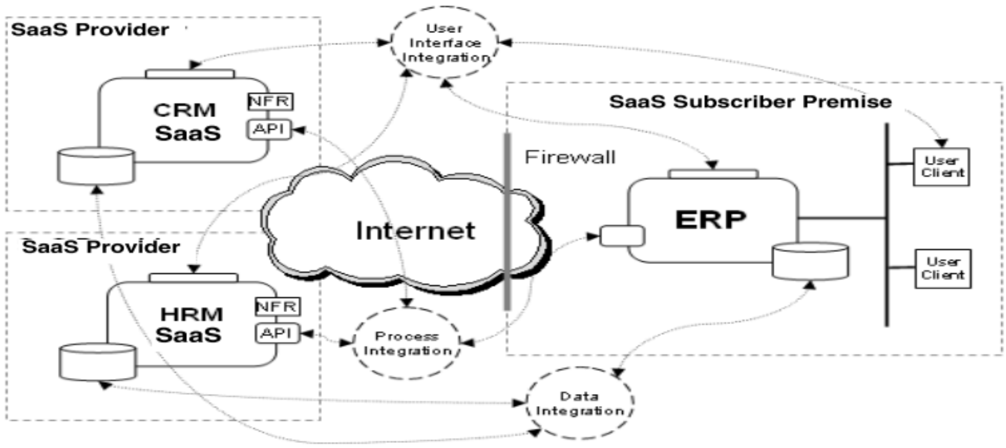

Figure 4.1 - SaaS Consumption Environment and Integration Requirements. Source: Sun, W., Zhang, K., Chen, S. et al., Software as a Service: An Integration Perspective. 2007, p. 558-569

SAAS DEFINITION

The gradual neglect of local applications is a reality and the advantages of the Cloud environment are becoming more obvious and clearer, especially in terms of sharing information in complex business realities. In general, an approach, Internet-as-a-service allows web users to take advantage of a software made available in a cloud network, in this way the provider's application is used and delivered as real service through the Internet, without the need to have it physically installed (in fact, whenever you want to access the software, you only need to access the cloud platform via the

Internet), because not only it is executed on the servers of the vendor, but also the support services are entrusted to it (figure 4.2). For this reason, a Software as a Service model is also called "Web-based Software", "on-demand software", or "hosted Software". [19]

By its very nature this context involves evidently the software industries, both in terms of software development and design, and sales and service delivery, as well as purchasing and usage terms have changed from end users' point of view. In our particular interest we are talking about routing solutions vendors, key partners for logistics companies.

Not by chance, in this field the desktop-based routing is leaving space to cloud-based routing, where a supplier software sells not only a product (routing software), but also a number of auxiliary services (performance improvements, maintenance, Availability, updates, testing, security, implementation, troubleshooting, maintenance, hosting, training, storage etc.), that allow the company to disengage from tasks not related to its competence and focus only on its own know-how. [19], [21]

In other words, a SaaS model compared with the classic desktop-based application, (i.e. traditional software model where the everything is managed by the company itself), brings the usage of an application on a new level, delegating back and front-office activities to the cloud provider, which manages a number of other customers unknown to each user. The latter, vice versa, only cares to utilize (or "rent") that service according to conditions customized and established in the subscription. The definition offered by [22] is particularly emblematic and summarizes this concept: *"SaaS focuses on separating the possession and ownership of software from its use. Delivering Software's functionality as a set of distributed services that can be configured and bound at delivery time can overcome many current limitations constraining software use, deployment, and evolution."*

We may define the cloud-based model as an advantageous and profitable experience for both parties, since even for vendors the management of services and customers through this multi-tenant architecture is more cost-efficient (compared to traditional), with the ultimate goal of allowing scalability, flexibility and support through cloud computing. [19]

Another possible scenario is represented by the combination of SaaS solutions with various sellers, in a single comprehensive solution that is offered to companies. This is possible, thanks to the fact that a service Oriented architecture (SOA), like the one in which SaaS operates, allows interaction between multiple hosted applications, so that each one of them can either provide its service or require one, being able to act both from provider and requester, integrating different software services, therefore features and functionalities. A figure that operates this service between SaaS suppliers and enterprises is called aggregator. [19]

In conclusion, we could summarize the new SaaS paradigm in two fundamental breakthrough: the first consists of the service-based mentality, i.e. not only a software product is sold, but of a complete set of service and features; the second in the success based revenue model, namely dependence of the success (i.e. continuation of the subscription) from customer's satisfaction. [23]

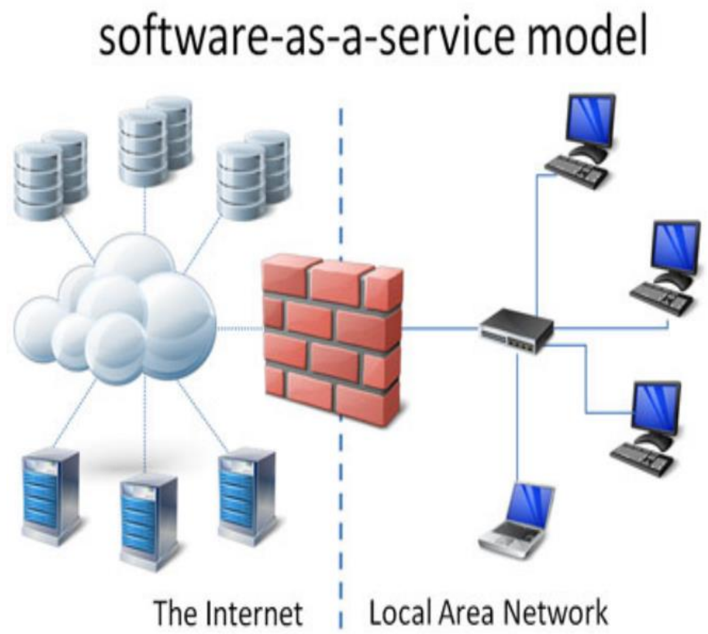

Figure 4.2 - Software as Service Model. Source: [19]

SAAS COSTS AND SUBSCRIPTION

It is worth to spend a section about how the way to sell/buy software in a SaaS model has changed. As mentioned among the major paradigm shifts compared to conventional hosting is the fact that only whether the customer is satisfied with the software will continue to request and pay the software service. Compared to a model conventional hosting where the company makes a substantial initial investment buying licenses, software, hardware and manages and customizes the application, the delivery of software as a service usually consists in a periodical fee comprising the entire set of support and auxiliary services operated by the software's vendor. [23]

This consequently allows to cut down on initial costs by reducing entry barriers even for small businesses, and allows you to treat software expenses more like *"[…] utility charges rather than an up-front capitalization"* [24]

Also, from the point of view of the software vendors, smaller and constant cashflow have replaced peak cashflows (due to the purchase of the software package on-premises). However, this requires an adaptation of the financial structure and costs far from being trivial.

Considering in detail the scenery of routing applications, we can say that the necessity of customization in the traditional model entailed to pay larger sums (up to several thousand euros) for the purchase of an ad-hoc routing software. Conversely, thanks to the SaaS model an investment in a routing service has been reduced to all-inclusive monthly payments around a hundred euros, to which costs for support services are usually to be added (e.g. integration of the service with pre-existing business systems).

Routing software services usually includes maintenance, support, hardware related costs, digital maps, training, and other costs, and should inevitably be considered if you want to switch to a SaaS model; as well as the cost variability according to the level of service and customization desired. [24]

From a more strictly economic point of view, the costs that we find in a traditional software hosting can be traced to the following:

- **Direct costs**: Servers, personnel, hardware, licenses;
- **Indirect costs**: e.g. storage, hardware management, upgrading, security and privacy management.
- **Other Costs**: Market offerings comparison, infrastructure evaluation, risks related to the physical "in-loco" presence of data, servers and books, depreciation, obsolescence etc.

Thanks to SaaS model all these costs become variable costs allowing less constraints and more ductility. [20]

SAAS FEATURES

The aim of this section is to present a structured list of SaaS features. Despite most of them are shared in the literature, they differ in some authors; the following are those that are considered the most significant, mostly taken from [19], [21], [23], to give a complete characterization of SaaS.

- **Configuration**: although the access to the application is carried out via Internet, by accessing a cloud platform of shared resources, each customer has its own personalization needs that require to be taken into account during the structuring phase (architecture) that relates provider and end users. We talk about configuration rather than customization, since the latter belongs to the conventional model of software hosting. In a SaaS model, in fact, you should not forget that you have to deal with shared resources. Despite this what ensures customization are the metadata that modifies the way the application run and serves several customers with different requirements. Of course, the customization process must be as agile as possible;
- **Requirement change**: a unique characteristic of hosted services is the ability to make recurrent changes to the service offered both for customers' needs and for system needs;
- **Integration**: most of the times it is not enough to buy a software service from a provider, in fact it must be integrated with existing business systems, this usually happens through protocols (HTTP, REST, SOAP and JSON);
- **Scalability**: key feature of SaaS and cloud-based applications, as enables to adapt the performance of the application according to the evolution of customer's requirements and needs;
- **Multi-tenancy**: software providers must be able, on the one hand, to efficiently construct and manage an architecture in which a multitude of clients (tenants) can access shared resources, in terms of peak usage, storage, etc., on the other, to ensure that each tenant has the requested personalization;
- **Online access**: the means by which you can access the application, but also other provider's services, is the Internet network, usually through of a Web browser. This free from software's and hardware's installation on site, and differs from conventional mode where Desktop-based clients used to access via WAN;

- **Centralization**: the resources have passed from being decentralized, desktop-based, to be centralized, as well as the front and back-office tasks (e.g. control, monitoring ad update) have passed from being managed by the company itself (IT department), which has purchased the software license, to be outsourced to the software vendor. Any centralized modification will take place in the cloud application, so any changes won't impact locally on the devices, but on the application to which they access;
- **Communication**: despite the main goal of SaaS model is to enable cloud computing, this structure also inherently allows for intra and inter high-efficiency communication;
- **Subscription pricing**: as explained in the dedicated section, instead of buying a product and assuming responsibility for its management, the service is rented and paid through monthly payments.

SAAS MODELS

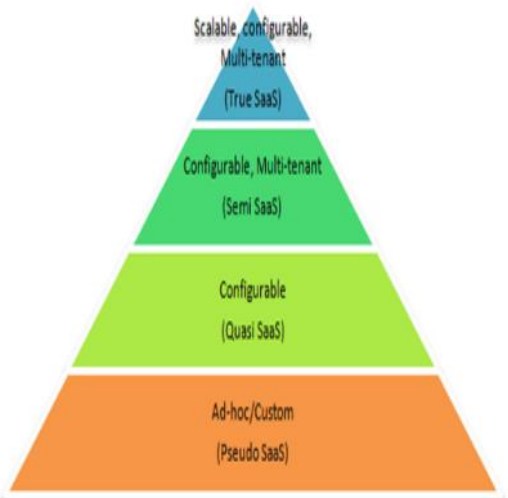

Figure 4.4 - Hierarchy of SaaS variants. Source: [19]

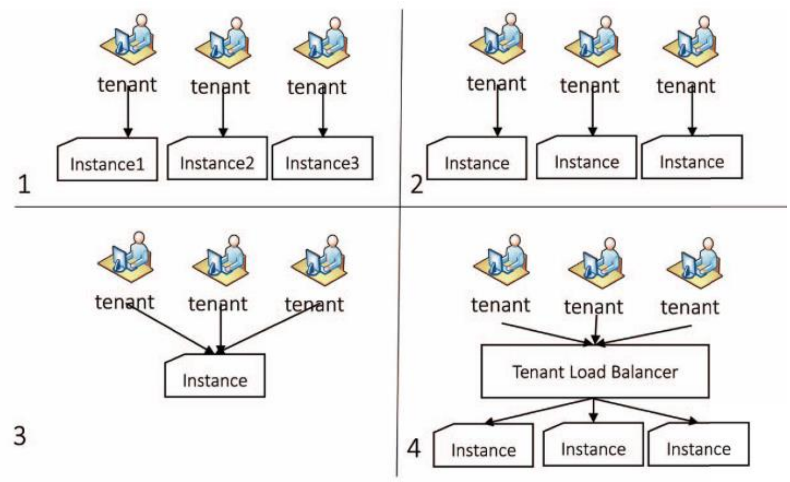

Figure 4.3 - Four-level SaaS maturity model. Source: [21]

The SaaS model can be implemented differently in different companies, in fact, it can be declined in four different ways (figures 4.4 and 4.3), depending on the degree of maturity to which it aspires, namely according to the so called Maturity Model, explained below in order of level of maturity. [19]

1. **Ad-hoc/custom**: this level of maturity provides an instance (understood as application code) dedicated, personalized and different for each tenant who uses it through the cloud servers;
2. **Configurable**: as in the previous level here too, each tenant has its own instance, except that, in this case, all instances share the same code. In spite of this, every instance works compartmentalized, in the sense that each one is isolated from the others. In fact, the instance (common) dedicated, is modified and configured subsequently according to the specific customer's needs (interface, behaviour etc.). This level and the previous one, because of their structure, require considerable storage and hardware capacity (especially for complex instances) due to the dedicated architecture to each tenant;
3. **Configurable, Multi-tenant-efficient**: the limits of previous levels are overcome in a model that reaches this level of maturity (although scalability problems take place), because instead of dedicating instances, here each customer draws from a single shared instance (then also code) in common to each tenant. The number of instances in this way is drastically reduced, therefore this allows for a more efficient (and even cost-effective) management of the cloud framework, as well as a more performing cloud computing. The metadata then allows you to apply the application configuration (instance) requested by the different users. Despite sharing, the sensitive data of each customer is isolated and protected.
4. **Scalable, configurable, multi-tenant-efficient**: scalability problems of the previous level are solved in this level thanks to a common load balancing layer (figure 4.5) between the instance (still unique and shared) and the end users. The configuration and issues are insured by the metadata and the data protection respectively as well. This level is not only an efficient solution to ensure scalability without further instance's changes, thanks to the middleware, but allows you to make changes or updates to all tenants as easily as it would be for a single one. In addition, this architecture allows you to provide a unique experience to each tenant, without the need to keep as many instances as there are clients.

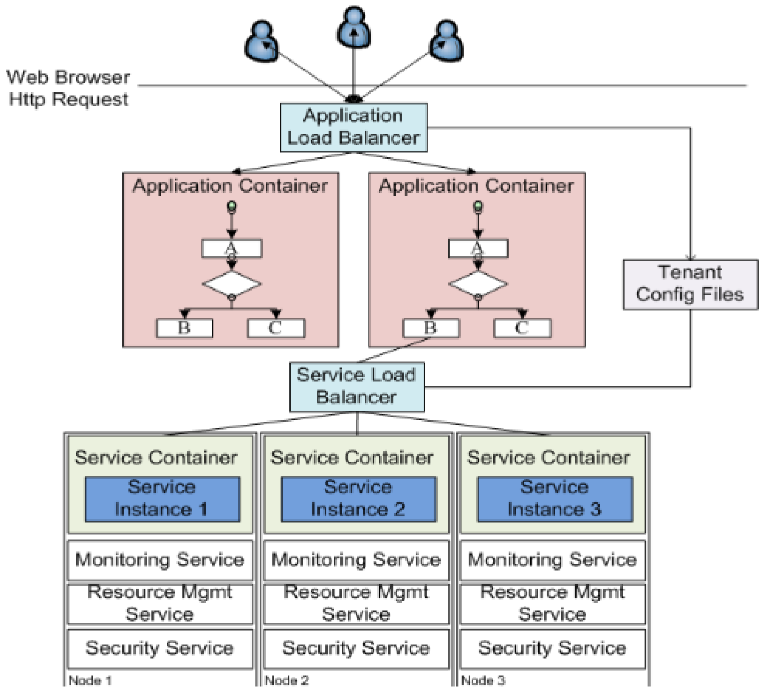

Figure 4.5 - Single Application Instance Multiple Service Instances. Source: Tsai, W., Sun, X., Balasooriya, J., Service-Oriented Cloud Computing Architecture, 2010, p. 684-689

COSTS-BENEFITS ANALISYS

In conclusion of this discussion on the cloud and SaaS, is appropriate to do a cost-benefit analysis (taking cue from the considerations of [19], [20], [23]) that can explain why the Software as a Service model has become so widespread. However, the purpose of this section is to present not only the reasons why it is desirable to adopt such a model, but to remember the presence of important drawbacks. The following are the main benefits of adopting a SaaS model (figure 4.6).

- ❖ **Cost-efficient implementation**: a traditional model implies both the purchase of costly licenses and physical infrastructure, and the costs of integration with company systems, training, personnel who manage maintenance, security and updates, which last over time. Moreover, even earlier the evaluation and testing phase of different solutions are to be taken into account. All above with the risk of dissatisfaction and dealing with sunk costs. These costs are not covered by a SaaS model.
- ❖ **Speed of implementation**: as no installation is required, all the time that would be spent on the evaluation of hardware and software offerings, the construction of

the infrastructure and the retrieval of other resources, is time saved. Once the subscription is decided, the implementation is carried out simply by accessing the Internet and the hosted service is "ready to use";

- ❖ **Low entry and exit barrier**: especially for small businesses this is a considerable advantage, since it allows the entry into the market of highly specialized solutions, which otherwise in a traditional on-premises model, would have suffered heavy initial investments. A further consequence of this is a broader offer on the market, in fact in this way, the business customers have the possibility to find on the market highly specialized solutions to choose from, at reasonable cost;
- ❖ **Centralized software**: those that were once necessary tasks (maintenance, updates, etc.), although extraneous to corporate know-how, are now outsourced and centralized to the software vendor, which deliver them easily from the cloud to the end user. This allows a total focus on value-added assets for the business enterprise. Also, the access to the service is centralized on a cloud platform, so one may access it from any device.
- ❖ **Low switching costs**: the relatively unfair numbers (if compared with those of the investments of "in-house" software solutions) are periodically paid, guarantee the company a certain freedom of choice of eventual alternative providers, in the sense that in the case of dissatisfaction the sunk costs and change are reasonable and non-prohibitive;
- ❖ **Zero infrastructure**: cloud computing allows the run the application on provider's servers, as well as the vendor is entrusted with the ancillary management; therefore, all the efforts concerning the infrastructures are avoided;
- ❖ **Easy and cost-effective infinite scalability**: the subscription-based model allows clients to easily adapt the software structure to varying needs (e.g. with adjustments of the monthly fee). The provider does nothing, but adapt the instance according to customer's new needs, on-demand;
- ❖ **Accessibility**: the Internet as a service facility allows you to easily access relevant information from anywhere in the world (as these are present in the cloud), as well as you can access and run the hosted service at any time, even without having it installed, from the device you are connecting on. Moreover (usually) accessibility is guaranteed independently of the platform and the operating system from which it is accessed;

- **Competitive offering**: from the provider's point of view as said, deliver software as a service constitutes a significant benefit in terms of openness to a new market, revenue, easier management of services and customers. Inevitably the lower costs arising from the adoption of this model translate in more competitive and quality offer, which further broaden the market by allowing further benefits, thus creating a virtuous circle.
- **Budgeting**: the expense management of a SaaS is much easier given the rent-based structure, since there are no hidden or unexpected costs, but they are known a priori, therefore the calculation of the budget is trivial.

Figure 4.6 - Advantage of SaaS in Software Development. Source: [19]

Despite the clear advantages of adopting a SaaS architecture we find also some crucial disadvantages, among which:

- **Security and data sovereignty**: in a SaaS model not only the ownership of the software is centralized, but also relevant data are uploaded in the cloud and handled by the provider. In other words, if you want to adopt a SaaS model you must be aware that you are giving up control over considerable information. Despite it is of course vital for companies to ensure the security, privacy and integrity of customer data, it is clear that the risk of theft or tampering with information on the web is exponential. If you think about the number of information circulating in a cloud composed of thousands of tenants is even more obvious the degree of risk and complexity, as well as the amount of data that the provider has to deal with. In the traditional model it was possible to have control of the information, as they were physically stored in the servers or corporate archives. It is no wonder that companies with particularly sensitive data still prefer to rely on desktop-based applications instead of SaaS models. The challenges in this field are still all open;

- **Lower customization**: although this may amaze, in a way we could say that the total customization that in the traditional model was possible, here is less, given the relative standardization that the multi-tenant structure requires.
- **Internet Connections requirements**: internet dependence, although it is one of the main advantages of the SaaS structure, does not allow (usually) to work offline or in the absence of a connection, and this constitutes an important drawback. In addition, a weak connection can compromise important service features (upload or download), especially for particularly heavy services.
- **Load Balancing Feature lack**: the aforementioned last level of the maturity model is also called "True SaaS" and is more complete SaaS model. At the same time, it is also the most complex to implement, particularly with regard to its focal point, i.e. the load balancing middleware that allows scalability, avoids system overloading and ensures stability. Since the latter are key features, load balancing issues are an important lack that need to be filled.

4.2 Decision Support Systems-Software (DSS)

In the last few decades, even in the company field, we are witnessing the development of informatics towards activities requiring an "intelligent" intervention. In this context, the Decision Support Systems (DSS) are placed, aimed at solving problems of higher level than those treated by traditional information systems.

As emphasized in [5], integrating dynamic vehicle routing problems into a DSS (Decision support System), *"[...] Together with modules for acquiring real-time traffic data and forecasting travel times on the base of real-time and historical data."* Can greatly help managers in decision-making processes.

DSS DEFINITION

A Decision Support System is a software system, or a computer architecture, that provides the decision-maker:

- information necessary to understand the problem;

- a series of models and data analysis features that give the possibility to explore data from different points of view according to the user's needs;
- the possibility to evaluate the scenario's choices consequences.

What is required of a DSS is the ability to consolidate information, to produce reports or forecasts, to allow simulations. [25]

The main purpose of a DSS is therefore to allow to extract information from a large amount of data, to suggest the most appropriate actions to different business needs, either strategic (long term) or tactical/operative (short period). All of this in an interactive, flexible and simple way, in order to increase the efficiency and effectiveness of the decision-making process. The latter can be classified into four categories: strategic planning (establishing policies, setting objectives, selecting resources), management control (ensuring effectiveness in the acquisition and use of resources), operational control (ensure effectiveness in the performance of operations), operational process (decisions taken during operations). Likewise, the decision-making processes, even the types of decisions may be highly heterogeneous (they can range from optimization problems to sorting of alternatives), consequently also the underpinning data, this implies that a DSS envisages distinct elaboration modalities.

To conclude, the elements included in a decision-making support system must be fully integrated with the other company information systems. Only in this way, in fact, the DSS will be able to operate on real data and may return reports which can be trustworthy to the operative reality. That is why, it is often useful to associate a GIS with the DSS, as it allows to manage the Database containing georeferenced data in a more intuitive and immediate way.

KNOWLEDGE MANAGEMENT

A fair classification of how the knowledge is structured in DSSs, is given by [25]:

- **Datawarehouse**: by the term data warehouse we mean a technological architecture aimed at supporting the decision-making processes of the company's business users. It is the heart of the DSS, a large container that collects data from different data bases and makes them: organized, consistent, updated, non-volatile and related. In practice makes them suitable to be worked to extract relevant information;

- **Data mart**: a data mart contains a subset of the data contained in the company Datawarehouse. It is therefore a departmental data warehouse, a sort of smaller-sized and more specific Datawarehouse;
- **Data mining**: indicates the process of navigating and analysing a set of data to find any regularity, extract knowledge, and derive meaningful recurring rules. It is the technique whit which to find information in a data base: Clustering (homogeneous subgroups), association (between the various data), sequencing (analysis of historical series), forecasting (forecast data);
- **Query**: interrogate the data base through specific instructions. Query support the decision-making process, speed up the elaboration of information and provide evidence or denial of the hypotheses formulated by decision makers;

DSS COMPONENTS

Here are the three main components (figures 4.7) that characterize a DSS's architecture, in order to give a schematic idea of how it is structured (still according to [25]).

- **Knowledge system (KS)**: consists of all the information related to the problem on hand and consists of three parts: DATA BASE (textual, numerical, structured geographic data, etc.), MODEL BASE (statistical, mathematicians, simulation, linear programming models, etc.) and KNOWLEDGE BASE (logical rules and relationships).
- **Problem processing system (PPS)**: it is the component that, once the problem has been defined through the interface, generates useful answers to the decision-making process thanks to the manipulation of information contained in the KS. It also consists of three parts: data-base management (Database structure definition, acquisition, modifications, access to data, etc.), model-base management system (management, creation, model updating, sensitivity analysis, etc.) and inferential engine (carries out the Artificial intelligence functions to provide a clear, consistent and logical explanation of the relationship that binds input to output, often difficult to perceive in a complex context).

- **User-System interface (USI)**: the main interaction mechanism between user and system. The interface must transform the commands expressed in natural language by the user in understandable terms for PPS and KS. Conversely, the output must be presented in a clear and unambiguous manner.

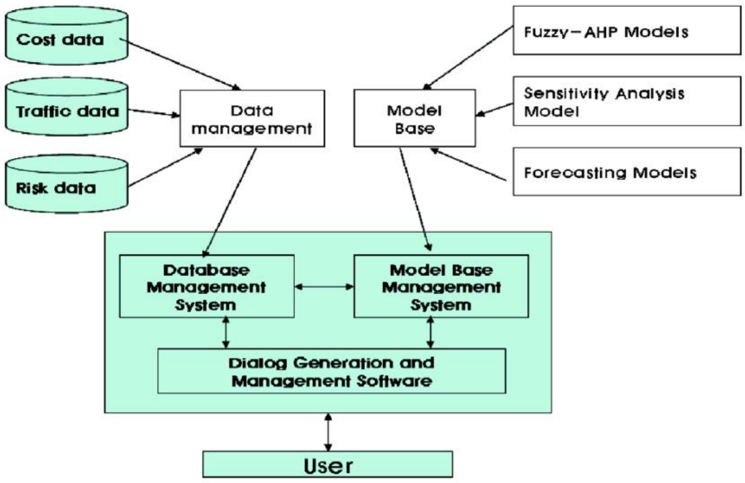

Figure 4.7 - Architecture of decision support system for multimodal transport network. Source: Ko, H., A DSS approach with Fuzzy AHP to facilitate international multimodal transportation network, 2009, p. 1-20

4.3 Geographical Information Systems (GIS)

The GIS (Geographical Information system) is a system composed of software tools or instrument that allows to analyse, represent, interrogate entities or events occurring on the territory. Nowadays geographic software may integrate common operations that can be carried out on the data base, such as researches and statistical analyses. For all problems that have a geographic component, the GIS allows to create maps, locate objects, integrate information, study the evolution of a network, view scenarios (even three-dimensionally), plan strategies, design infrastructure, solving complicated mobility problems and developing the most effective solutions. GIS is not a sole software, but it is rather a system consisting of the combination of hardware, specific software for GIS applications, alphanumeric and geospatial data, and finally, but not least, by the human factor, thus the people that manipulates these data. One of the leading companies of such systems, ESRI, defines the GIS as a *"[...] A computer software that links geographic information (where things are) with descriptive information (what things are). [...] Can present many layers of different information.* " [26] A GIS system enable to acquire digital maps, the software then allows you to georeference the maps with respect to a specific reference system and to connect to these maps descriptive alphanumeric data, which can be originally registered on a database (e.g. ERP databases) or on spreadsheet tables, referring to events, conditions or objects represented on the map (e.g. their geometry and relationship). In our case data referring to a courier can be the position, speed, state of the load, served orders or data relating to a road network, such as critical intersection, restrictions, speed limits and so on; all of these can all be associated with a graph representing the network. [27]

GIS allows you to edit maps by inserting symbols, objects referenced both in terms of attributes and spatial coordinates, on which you can perform specific analyses and processes at the user's will, as all the information is collected in a manageable set of data. Once this smart map has been built, interactive representation of georeferenced data, the system will be able to overlay it perfectly to the original digital map as a kind of specific information layer (figure 4.8). From the organizational point of view, the GIS works through stackable and integrable layers according to geographical location attributes, but this is only possible when you have precise coordinates for each layer; GIS-oriented data archives are also typically divided into separate thematic layers. In general, this categorization of georeferenced elements allows to execute "query by attribute", namely it is possible to ask, obtain and visualize not only which, but also where are the elements that have certain characteristics or attributes. In our case, for example, which and where are the customers to be served within a specific area.

GIS systems are characterised by user-friendly interfaces, with masks, drop-down menus or guided wizards that make them usable even by generic operators, which in this way can focus on the real goal rather than interpreting the software.

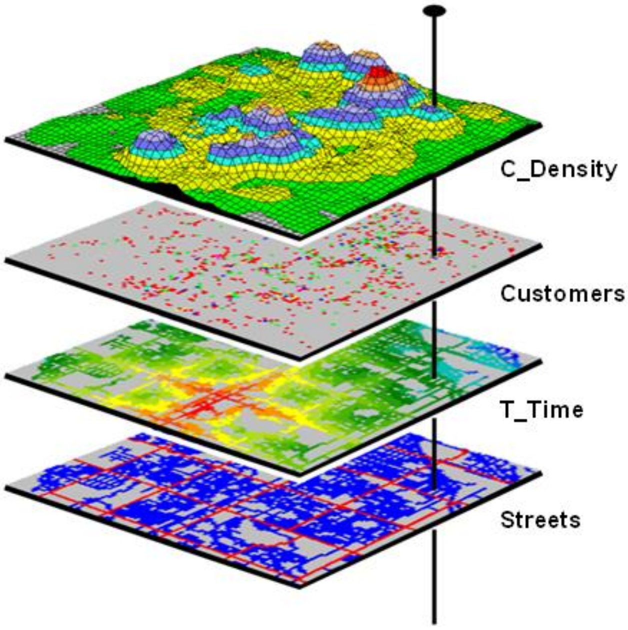

Figure 4.8 - GIS layered structure. Source:
https://gis.stackexchange.com/questions/239309/displaying-multiple-stacked-layers-in-arcgis-desktop

LOGISTIC APPLICATION

In a distribution company, the daily activity expects to keep track of a considerable variety of information: keeping accurate and updated information about resources availability, defining orders and daily shipments, providing info, keeping and analyse historical data about traffic and real-time, shipments monitoring, plan trips. All this is necessarily related to spatially traceable information.

As a result, it is clear that the application of GIS technology to a routing context is widely used. In particular, it is proposed an example referring to the management of the fleets of mobile vehicles through radiolocation and spatial representation on a cartographic basis of each vehicle in service in real-time.

The application is constituted by the management system of the operative station, the on-board's system and the communication system that allows the first two components to communicate. In particular, the management system of the operation centre is designed to evaluate the acceptance of new orders to ongoing routes, generate and manage the archive of customer practices, manage cartography, manage the fleet, assign the order, calculate the most suitable, to monitor the state and the position of the moving vehicles and visualization of each medium on a cartographic basis (figure 4.9).

The on-board system, in the most extensive configuration, includes a navigation system for the calculation of the routes, a GPS satellite tracking device, a communication and data exchange system with the operations centre, which allow the automatically send the position to the control unit, receiving a destination (or text message) from the control unit, sending coded messages to the control unit. [27]

Finally, another function that GIS can play concerns the possibility of developing complex spatial models as "what if" scenarios for transport planning and responding to circumscribed spatial queries (e.g. premium customers in Radius of 5 km).

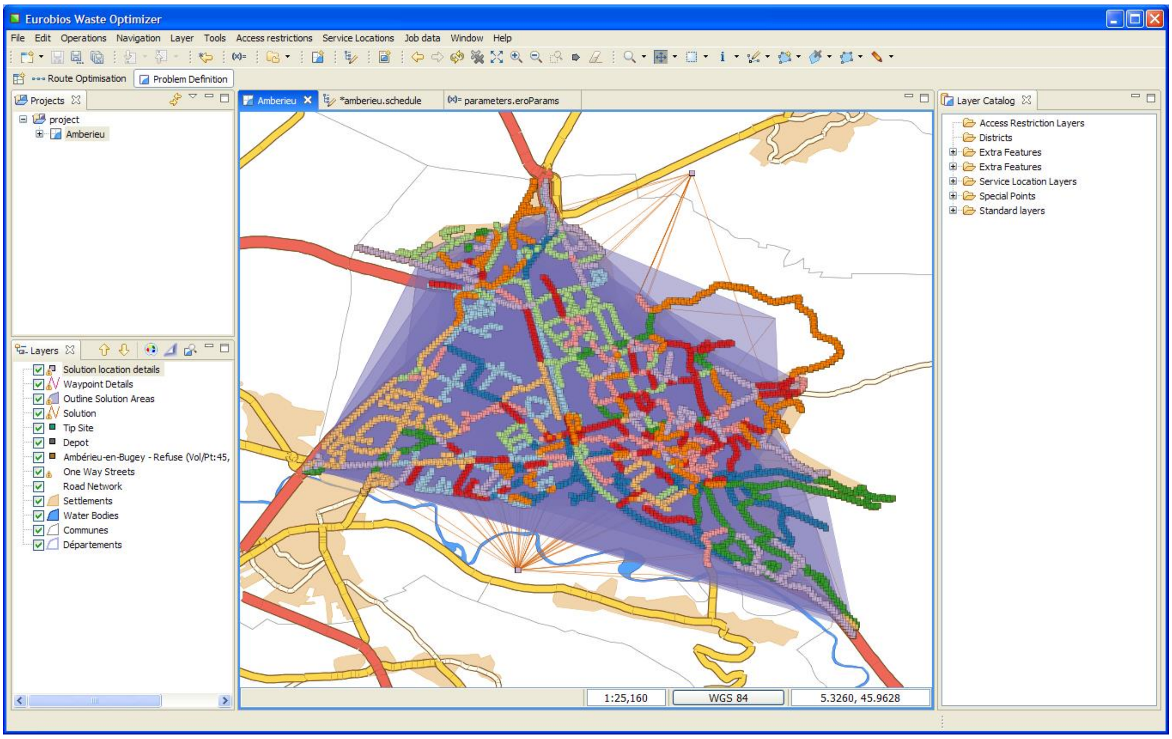

Figure 4.9 - Window of GIS application for optimization of waste collection routing. Source: http://udig.refractions.net/gallery/eurobios/

4.4 Digital Twins

The paradigm of the Digital Twins (also shortened DT), meets the recent need to lay the foundations for the so-called industry 4.0, which is supported by the development of the IoT. It should be recalled that the latter is stood as the search to yield "things" intelligent (e.g. devices, but not solely), thanks to integration with network, detectors, electronics, hardware and software, with the ultimate goal of creating a Web-based data environment for achieve an improved cost-efficiency, performance and system management. [28]

Although the concept of Digital Twins has been known for a few decades, starting from the first germ born in NASA, only in recent years its effective applicability in other fields is becoming a reality, due to new opportunities offered in the field of information technology that now enable to interconnect physically dispersed entities in a wireless manner.

DIGITAL TWINS FRAMEWORK

To date, we might recognize two distinct, but closely related worlds: physical one, whereby a physical asset is physically in a system, and the virtual one, for which the state and the information of the latter are traced from afar. In recent years this context has given birth to a new scientific field, that is the Cyber-Physical Systems (CPS) of which the Digital Twins are part. [28] Thanks to this approach, the digital twins allow to decline the virtual control in a physical control of devices (in our case for example vehicles) acting or sensing in the real network. This is done by generating a "virtual twin", a sort of digital replica, corresponding to a physical asset including processes, people, places, systems and devices. By doing so, a continuous flow of real-time data between the two levels allows the simultaneously existence of both the virtual and the physical asset. Detectors and software underpin and ensure this layer connection providing dynamic data. This digitalization satisfies IoT devices requirements during all their operating phases. In other words, the virtual counterpart disposes of information collected (from the real world) by the corresponding physical twin by means of technologies (sensors, software etc.) which either the network or the vehicle it is equipped of. The advantage of having a virtual counterpart enable to operate simulations of the evolution of the system, then estimates of the performance expected in a given dynamic scenario, as well as other purposes (e.g. different fleet management strategies). [28]

It is clear that these peculiarities meet the well-known and discussed needs that the modern context of routing (and IoT applications) requires, as well as real-time requirements, synchronization of sparse entities (intra and extra corporate) and the cloud-based framework, thanks to the physical-virtual connection (figure 4.10).

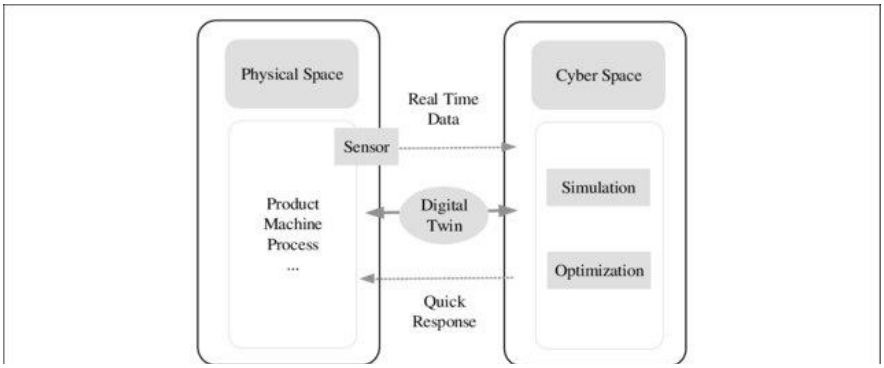

Figure 4.10 - The relation between cyber and physical space Source: Sakari, M., Enabling Digitalization with Plant Engineering Software COMOS, 2018, p. 1-45

TECHNICAL DEFINITION

The definition of digital twins, although maintaining its fundamental core concept, has evolved over time in function of openness to new application's fields.

Initially, digital twins in the aerospace field were used to monitor the overall status of aerospace equipment through representative models of the system and the environment in which the vehicle (interacting with the outside world) operated, which took into account of data collected, historical and dynamic environment.

In order to define this technology as comprehensively as possible, also in the perspective of the thesis, we chose to refer to the following definitions. Some authors define digital twins as a system or model as [29]. *"An integrated multi-physics, multi-scale, probabilistic simulation of an as-built system, enabled by Digital Thread, that uses the best available models, sensor information, and input data to mirror and predict activities/performance over the life of its Corresponding physical twin",* or [30]: *"Very realistic models of the process current state and its behaviour in interaction with the environment in the real world"*; Others, like [31] , define the DT paradigm by focusing on the concept of digital counterpart (figure 4.11): *"Virtual* *substitutes of real world objects consisting of Virtual representations and communication capabilities making up smart objects acting as intelligent nodes inside the Internet of things and services.".*

There is no real definition in the literature for Digital twins applied to transport contexts. However, we can understand how, according to the considerations above, digital twins may find a significant applicability in a context of Vehicle Routing, as it is essential to keep track of information physically detected by the network and present simultaneously in the cloud database, in order to ensure real-time feedback and to apply Big data analytics which may generated relevant information to drivers and decision makers. This is perfectly summed up in the definition of [32] ,which considers digital twins as a: *"Coupled model of the real machine that operates in the cloud platform and simulates the health condition with an integrated knowledge from both data driven analytical Algorithms as well as other available physical knowledge"*

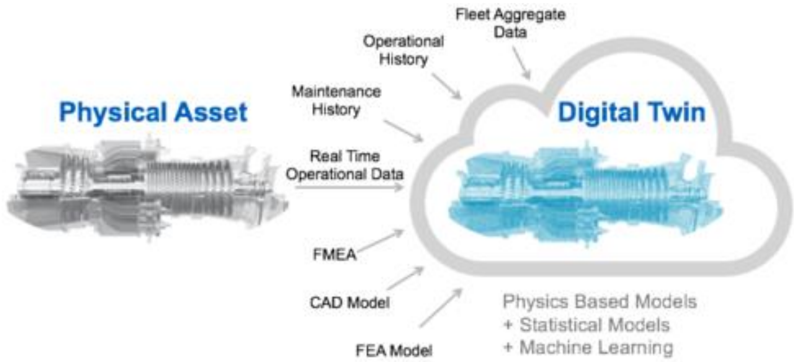

Figure 4.11 - Physical asset and its digital twin. Source: https://www.vizexperts.com/blog/digital-twin-and-its-impact-on-industry-4-0

DT PUROPOSES

Digital twin purposes have evolved through years, depending on the extent to different fields of application. The main purposes are presented in timeline below [28].

- **Current health status**: For control purposes, maintenance management, physical state of physical counterpart and model adherence to the real system;
- **Physical Twin's life digitalization**: to study and forecast system's evolution, hence performance, ensure seamless of data gathering, provide a digitalization of the asset, IoT entities management;
- **Decision-making support**: Thanks to multi-disciplines synergy and big data analysis; Estimates and performance optimization's purposes thanks to historical and real-time data.

GENERAL CONCLUSION

Essentially on the one hand, thanks to detectors in the physical layer, it is possible to obtain information, and thanks to the virtual level remote tracing and control are provided; on the other hand, the simulation, also fed by data gathered from the field, allows planning or forecasting purposes. [28]

For this reason, digital twins are not only particularly useful to keep control of complex systems affected by a multitude of dynamic factors, relationships and limits, but also, they are particularly suitable for cloud-based systems, given the digitalization of physical data in the virtual layer.

In conclusion, we could summarize the Digital Twins as a means to virtually represent a complex system, capable of simulating its evolution thanks to the synchronization between a virtual and physical level performed by an instrumented network, involving intelligent "things" communication, complex formal models, data management and processing. All above to monitor and optimize the performance of the entire system according to the evolution of the environment, either in real-time, thanks to the data detected on the field, or statically thanks to the prior knowledge of future scenarios, derived from estimates. [28]

In conclusion, this permit an enhanced decision-making processes and overall system's maintenance. The relevance of this technology is evidenced by the fact that some software solutions to implement Digital Twins are already available on the market, which is expected to widen in the coming years.

4.5 Blockchains

The approach of blockchain technology lends itself to the increasingly digital and web-based nature of the modern context, understood either in terms of cloud or in terms of IoT. This technology over the last decade (2008 with the advent of Bitcoin cryptocurrency) has revolutionized the way of conceiving the exchange of assets between involved parties in a transaction and its application's field is recently spreading in various applications.

THE NATURE OF BLOCKCHAINS

The underpinning concept of blockchain is to eliminate the presence of a third party (usually bank or provider) in transactions that makes the process centralized and controlled by a party not directly involved, in order to yield a decentralized process (controlled directly by the protagonists of the transaction), more fluid, safe, robust and economically advantageous for both parties. [33] The blockchain platform consists of a public and shared distributed database (although there are also Private Blockchains and Federated Blockchains), consisting of a public ledger, which collects all the information about all the transactions operated by the users (called nodes in such a framework) that are part of it and that are responsible directly for the confirmation and verification of the latter, creating a kind of network consensus (peer-to-peer network). [34], [35] The public and shared nature of this ledger allows transparency and accessibility of information, but at the same time the respect of privacy is ensured thanks to anonymity (users are represented by avatars); this peculiarity is particularly suitable in IoT applications and frameworks. In addition, distributed accountability and cross-checking allow you to easily detect possible frauds, allowing system's self-control. [35] The assumptions of information integrity and cyber-security, pivotal features to ensure stability of the system, are still open challenge, as well as the technical requirements to support such a system.

BLOCKCHAINS MECHANISM AND FEATURES

At the most detailed level, we can define the underlying mechanism as a public key infrastructure (PKI), whereby a public key identifies the user's wallet and a private key allows user's certification. The two fundamental entities in the Blockchains framework are transactions and the blocks. A valid recognized transaction is encoded and contained in a block, which in turn is connected to a preceding block of other previous and verified transactions, forming a block chain called blockchain precisely. By doing so, the integrity is verified by "block-by-block" and stored along the chain. [35] Users, called nodes, are responsible for storage, mining, and transactions verification because only new blocks verified by them are to be hooked to the chain, who are therefore guarantors of database integrity (figure 4.12).

In summary we could identify among the main features of a blockchain framework [35] :

- **Security of operations**: thanks to consensus-based architecture, the encrypted codes and authentication;
- **Decentralization of storing/computing**: due to the public ownership of all the data, avoiding third part involvement;
- **Data integrity and Transaction immutability**: due to the verified block structure of the chain and the traceability and verified validity of all the operations in the Ledger, as no modification or elimination after nodes approval is allowed.

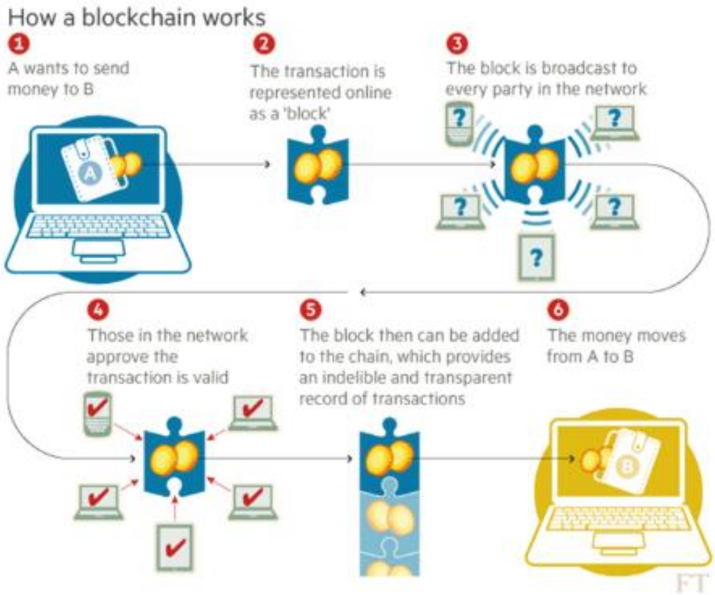

Figure 4.12 - Operation of a Blockchain transaction. Source: [35]

CONCLUSIONS

This technology is taken into account since cloud-based software implies the paying of monthly fees digitally, and, therefore, have all the advantages in seeking a more convenient and reliable solution to structure payments and subscriptions. In this sense digital contracts, such as Smart contract, are a great opportunity for software companies offering cloud-based services. In particular nowadays [36] refers to: *"[…] Blockchain-oriented software, that is a software that uses the implementation of Blockchain in its components."*

4.6 Internet of Vehicles (IoV)

More and more the current technological landscape aims to bring intelligence in the applications, in the sense that we do not seek a mere increase in performance, but rather a better interaction between technologies to bring benefits in a holistic way to a multitude of users. In everyday life we are all accustomed to passing in a completely natural way from one network to another (connection to the Hi-fi system, network of Smart TV, connected vehicles, wireless headphones, social media, etc.). This search for interconnection between networks can be achieved by means of wireless technologies and thanks to the progress achieved so far in the field of ICT. The solid basis offered by them, together with the awareness of this multi-network environment has given birth to the so-called Internet of Things paradigm, where for things we mean the proliferation (in heterogeneous environments such as offices, home and in the transport) of Cars, mobile phones, TV, home appliances, audio systems, home automation etc. with the common denominator of being internet connected.

Internet of Vehicles is nothing more than the declination of the internet of things in the world of transport, with the aim of making the network an intelligent context, making users experience more enjoyable, efficient and secure, thanks to pivotal changes in the way which man, devices and vehicles interact with one another.

The purpose of this section is to present the model at the base of the IoV, the architecture and the potentials with the aim of potentially considering it a means for the improvement of the innovative routing architecture which will be proposed later, since the future of vehicular communication across the network is inevitably significant.

IOV MODEL, FEATURES AND LAYERED ARCHITECTURE

The proposed models of IOV are different in the literature, but although each emphasizes or points of interest on some aspect at least the three components turn out to be a common denominator.

- **Humans**: They are both generators of information and service/data users. They also include all the actors in the network (from drivers to passengers to public transport users, to pedestrians) and contribute more or less actively to Request or provide information.

- **Vehicles**: They include all vehicles involved in the road network that provides useful information via the IOV infrastructure and at the same time provide this information to drivers to offer an improved and safer driving experience.
- **Things**: This is all that is around the two previous components, are sources of relevant information scattered in the network that consists of the environment. [37]

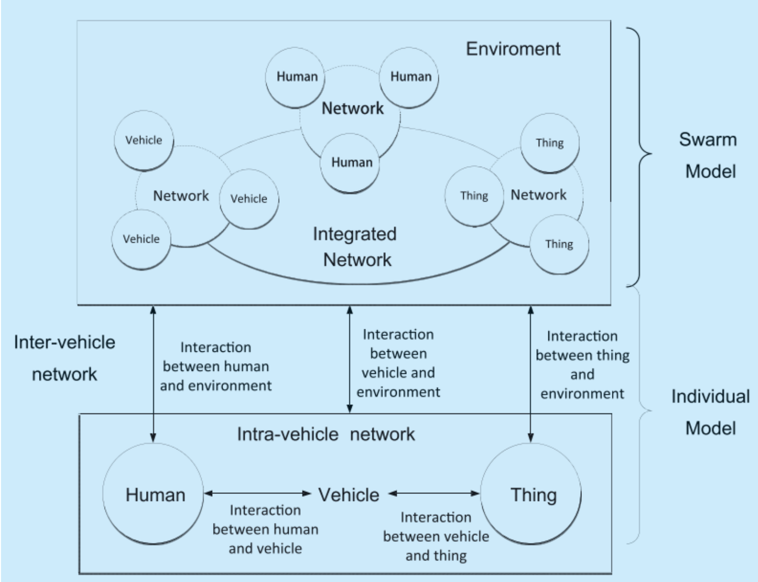

Figure 4.13 - Network model of IoV. Source: [37]

The image represents the model proposed by [37] and summarizes comprehensively how the three components operate and interact with each other. In particular, three macro divisions are identified: an individual model consisting of the representation of a single human-vehicle-thing network; of a swarm model, which represents how different networks are linked to one another in a broader integrated heterogeneous multinetwork cloud based; end in an intermediate layer that retaliates as individual model and swarm model are tied. It is through the swarm model that the dynamic information circulating inter and intra networks is worked through cloud computing, crowd sensing and crowd sourcing to become services. (Figure 4.13)

In an Internet of Vehicles environment, you can find three main features which will be discussed below. (Figure 4.14)

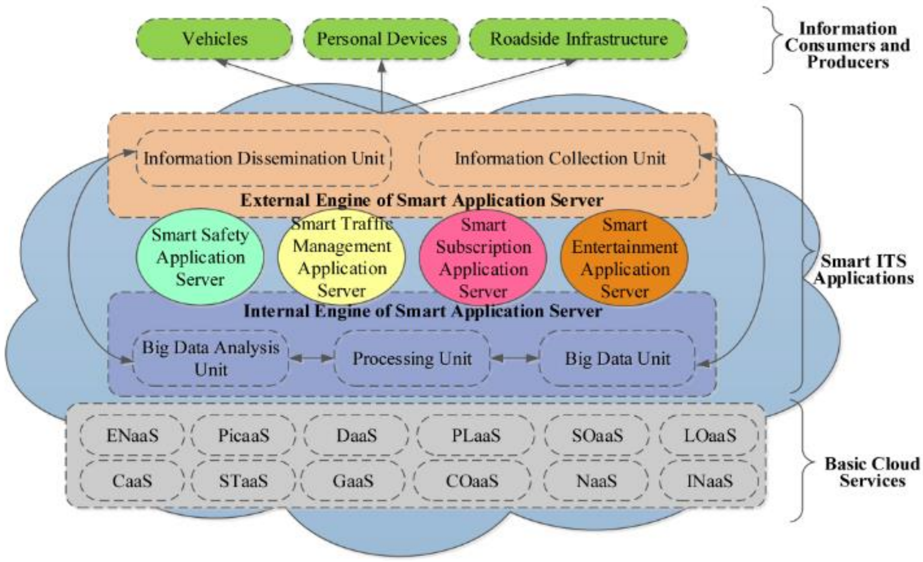

Figure 4.14 - The role of cloud computing as three operation levels. Source: [38]

- **The cloud**: This infrastructure has the main role in the IOV context, as it absorbs and treats in a single cloud-based platform all the information found across the network, and this consists of the so-called primary cloud services, or XaaS, or Anything as a Service. Secondarily by the big data management of the servers it can deploy different categories of smart applications, disseminating intelligent data to the end users (clients), stood as vehicles or personal devices. Still, as stated by [38]: *"The primary responsibility of the cloud servers is to process and apply artificial intelligence in real-time big traffic data to make intelligent decisions for smart client applications."*
- **The connection**: Core component of the IoV since it allows the whole context to work sustainably. It enables the access to IoV framework and keeps a reliable connection between the different entities and heterogeneous networks.
- **The Clients**: they are the beneficiaries of the work of the artificial intelligence operated on the servers, as well as they are the end-users of smart services and applications discussed in the preceding paragraphs.

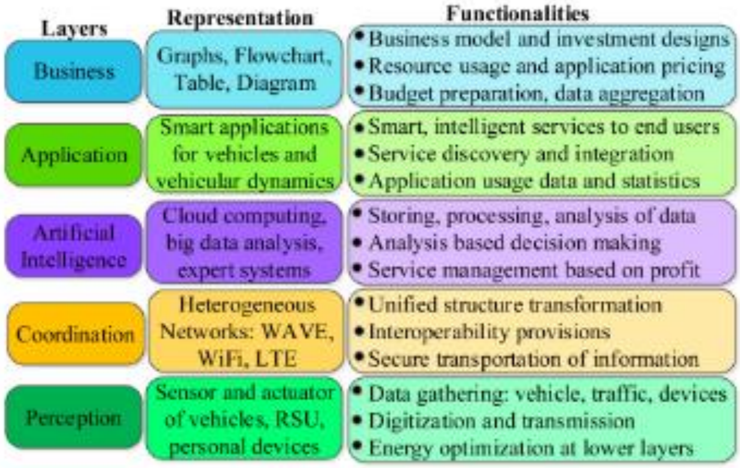

Figure 4.15 - The five layered architecture of IoV. Source: [38]

Usually the structure of IOV models is seen as a layered structure in which each layer is made up of heterogeneous elements used for specific objectives. The proposed structure is taken from [38] and has to be stood as a bottom-up and internet based frame where each layer is a step and depends on the previous one. In particular, the heart of the IOV resides in artificial intelligence, as it is the innovative element that allows smart and business applications, for this reason we may define it a Service Oriented Architecture (SOA), XaaS stood as Anything as a service. (Figure 4.15)

IOV STEPS

It is clear that in this panorama of relations and interconnections between the network, the WAT (Wireless Access Technologies) is an essential aspect, especially given the heterogeneity of the networks, because they must ensure continuity of service, security, Quality of information. In particular it is shared by many authors that the best wireless transmission technology depends on the device and the service offered. [37] asserts that probably the best technology to date can be used to implement an IOV network either 4G or LTE, given the extensive coverage and compatibility with almost all devices. In IoV Optics, WLANs, WiMAX, Cellular Wireless, and satellite communications are also used; Although in particular it is to be said that the latter does not guarantee sufficient precision and is also more precarious and costly. In particular, in an IOV context, 5 types of vehicular communication: vehicle-to-vehicle (V2V), vehicle-to-Infrastructure (V2I) of mobile networks, vehicle-to-Personal devices (V2P), vehicle-to-Sensors (V2S) can be established; Each one characterized by different best WAT, more or less efficient depending on the case on hand. (Figure 4.16)

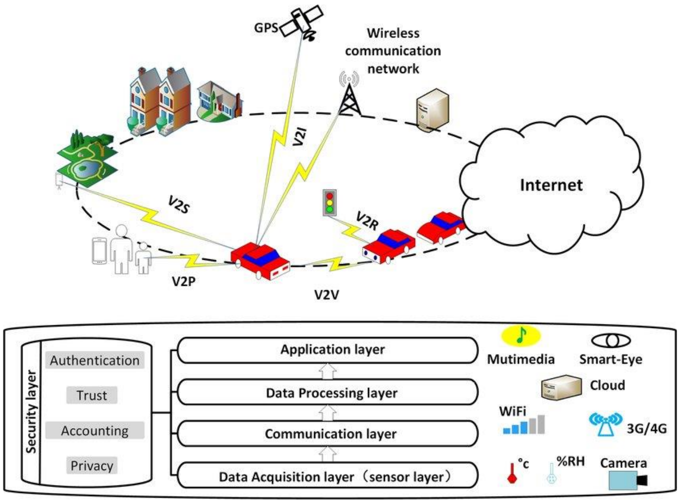

Figure 4.16 -VTX communication in IoV. Source: Shen, S., Wei, Z., L. Sun et al., The Shared Bicycle and Its Network: Internet of Shared Bicycle (IoSB): A Review and Survey, 2018, p. 1-24

IOV APPLICATIONS

The world of transport in some way is gearing for the advent of the IoV, in some cases at the level of the individual model is already in front of a mature framework. Indeed, for some years, chips have been installed for unique identification of any connected entity (smartphones, trains, Metro, Tv, washing machines, conditioners, gates, etc.) that allows to automatically provide information necessary to take advantage of a service in the Intelligent IoV framework.

More generally we can identify two categories of applications: safety applications and user applications.

Safety applications, also called safety and management-oriented applications, using big data traffic in the servers aim to improve the driving experience and enhance the safety of not only drivers, but more generally all network users, Through intelligent countermeasures. Here are some examples that may be of interest to the purpose of the thesis. [37]

- **Collision avoidance**: To date is an order-qualifier for any manufacturer and consists in the management and intelligent control of the motion of the vehicle (speed, steering, brakes etc.) according to what is happening in the network, thanks to communication with other vehicles, this system can share its functionality, thanks to IoV, warning other networks cars of an accident. Fulfils both the collision warning, concerning anticipated contact notices with nearby cars, accidents, queues and road conditions with sound and on-display notifications, and driver assistance, intended as a real temporary assistance in case of emergency Detected by the sensors of which the vehicle is equipped. It is both intended as a semi-automatic guide (e.g. lane change), as well as other automatic operations, e.g. emergency call.

- **Intelligent intersection**: Futuristic application that if implemented successfully allows to liquify and make safer and trafficked and crucial crossings. The basic principle is a real-time communication between the vehicles involved (through WAT, GPS and in-vehicle processing) of information regarding speed, trajectory, destination;
- **Navigation**: Navigation services are achievable through GPS systems integrated into the vehicle as well as other related ITS services. Users, but in general any entity in the IOV framework, can obtain information about the state of live traffic through the interaction of different sensors and other networks of things, vehicles or humans and the real-time information that they share in the cloud.
- **Remote diagnostic**: Thanks to the holistic and extensive nature of the cloud and to the fact that each connected entity is traced and identified, it is possible to establish a continuous connection with the state of the vehicle, as several sensors of which the cars today are equipped, they allow real-time communication concerning the operation of the organs of the car. Therefore, a vehicle black-box is always available in the case of emergency or preventive communications. For example, near a service centre, all vehicle data, as well as information on the nature of the failure, is communicated in advance to the garage. This is called cloud-based data management, and can generate obvious benefits in terms of maintenance costs, breakdown of service, therefore, consequently, reliability of customer service. A related application is the so-called self-repair, which consists of step-by-step instructions, delivered through the vehicle display, with the possibility of live-video assistance, to allow the repair of any simple malfunctions.

Other applications in this category that are not closely related can be intelligent parking Helper, multi modal transportation, remote telematics.

User and business-oriented applications are IoV's value-added services for the user, not in terms of safety but in terms of quality of travel experience. They consist of surplus not necessarily real-time and usually interactive and user centred, such as smartphone connection, social media access, audio and video streaming and downloading, entertainment applications and games, points of interest information such as ratings, reviews and timetable. From a more strictly business standpoint, however, we can talk about global Internet services, which includes insurance services, diagrams and graphs display, Routing and Fleet Management, generally software based. Here are some examples taken from [38] and [37].

- **Insurance**: This is a new innovative way to establish insurance parameters. As explained in the diagnostic application, thanks to the nature of the environment IoV it is possible to create ad-hoc insurances according to the performance of the driver, as specific statistics of the vehicle are calculated and shared in this case, we talk about Insurance on driver statistics;
- **Connected driving**: Widespread application with great potential in an IOV framework, is revolutionizing the concept of driving experience, offering the user a wide range of distraction-free applications. In fact, many manufacturers offer smart driving services by enabling the indirect use of devices, such as smartphone (messaging, maps, music, social, voices support) while driving, i.e. Apple CarPlay or Google Android with user-friendly interfaces. [38] Even remote devices may be connected directly, namely office, home, studio devices through the car's screen, for work or entertainment purposes.
- **Cloud services**: they relate to the possibility of establishing while driving network of vehicles, or at the same time, connecting to the Multinetwork cloud. In this way the platform draws information from the connected vehicles, freeing itself in this way from data storage issues, and by operating a big data pooling then to be computed (cloud computing), at the same time is ready to deliver intelligent cloud applications to Users themselves.

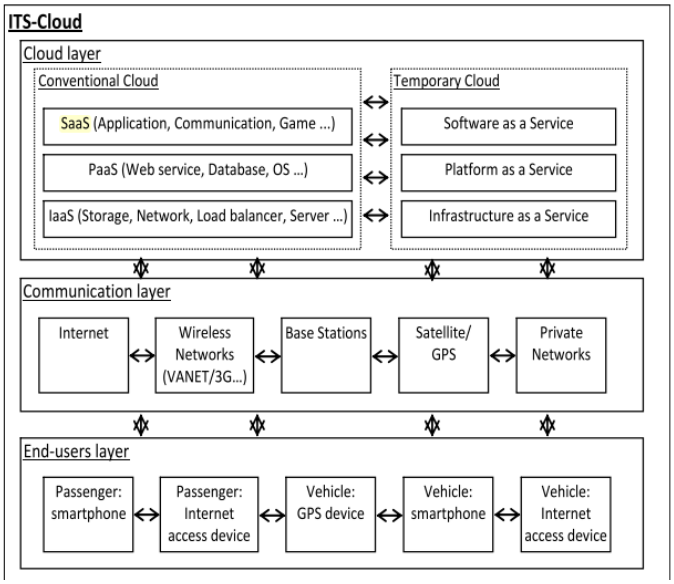

Figure 4.17 - ITS-Cloud architecture. Source: Bitam, S., Mellouk, A., ITS-cloud: Cloud computing for Intelligent transportation system, 2012, p. 2054-2059

These applications are expected to become more and more effective in the next years as the market interest, the potential for safety and driving experience are becoming a standard. In addition, the growing processing and decision-making capability of vehicles [38] will generate a ripple effect, since more and more connected cars (thanks to the dissemination of new standards intelligent in-vehicle systems, e.g. Apple CarPlay) will become part of this heterogeneous multinetwork by expanding the IoV basin, generating more interaction between network entities, therefore increasing big data circulation, computation and reliability, finally transport efficiency both in terms of driving experience and in terms of Routing-related issues. This comes together with a necessary development of WAT infrastructure in order to accomplish the abovementioned steps. (Figure 4.17)

INFOTAINMENT SYSTEMS

The word infotainment is a neologism oftentimes used in the automotive sector nowadays, it comes from the combination of two terms, “information” and “entertainment”. It stands for an information and entertainment system of hardware and software (specified by company) of which most of the cars are now equipped.
The infotainment systems can be controlled from steering-wheel controls (for safety reasons) and they may include: audio speakers, main and auxiliary displays, satellite navigation systems, video player, USB and Bluetooth connectivity, wi-fi router

(internet in the car), on-board diagnostic computer, voice controls, external cameras. Given the smartphones spread, Google and Apple have developed systems (Android Auto and Car Play) that can integrate its apps to the on-board infotainment software, reproducing what is displayed on the smartphone.

FUTURE ASPECTS

According to what has been discussed, the infrastructure scenario is expected to be profoundly changed since the advent of the IoV. Some new expectations, suggested by [38] are given below.

- **Online vehicles**: each entity will be connected to the Internet to be present in the IoV cloud framework, in order to exploit the advantages of the online service availability of information in a cloud based multinetwork.
- **Global Internet ID**: as already stressed, in the future, each entity will be traceable and identifiable in the context of IoV, so all the relevant information is available, safe and reliable for a further smart use by a network stakeholder.
- **Combination of RFID and GPS**: these two technologies will be widely exploited to ensure the indispensable protocol performance in terms of accuracy, security and dissemination, requested by IOV functioning.
- **On road Internet**: following the IoV's deployment, we will generate a kind of on-the-road Internet created by the interaction of the circulating entities across the road network.
- **Big data explosion**: internet of things will be led to an exponential growth of big data generated by entity entities in the heterogeneous multinetwork, and consequently that supposed be treated properly in a cloud environment; this constitutes an opportunity and an intrinsic complexity at the same time.

Chapter 5: System's Architecture proposal

This chapter presents the output that gives substances to the considerations made in the previous chapters and which consists in the final project's objective, that is a Routing and Scheduling system's architecture. At first, this architecture will therefore be represented and commented upon as a diagram, afterwards it will be explained how each element and compelling technology is interconnected and contributes to the final output. Hence is explained the overall significance of the proposed architecture and the original contribution of the thesis as compared to the state of the art.

5.1 Overall Idea

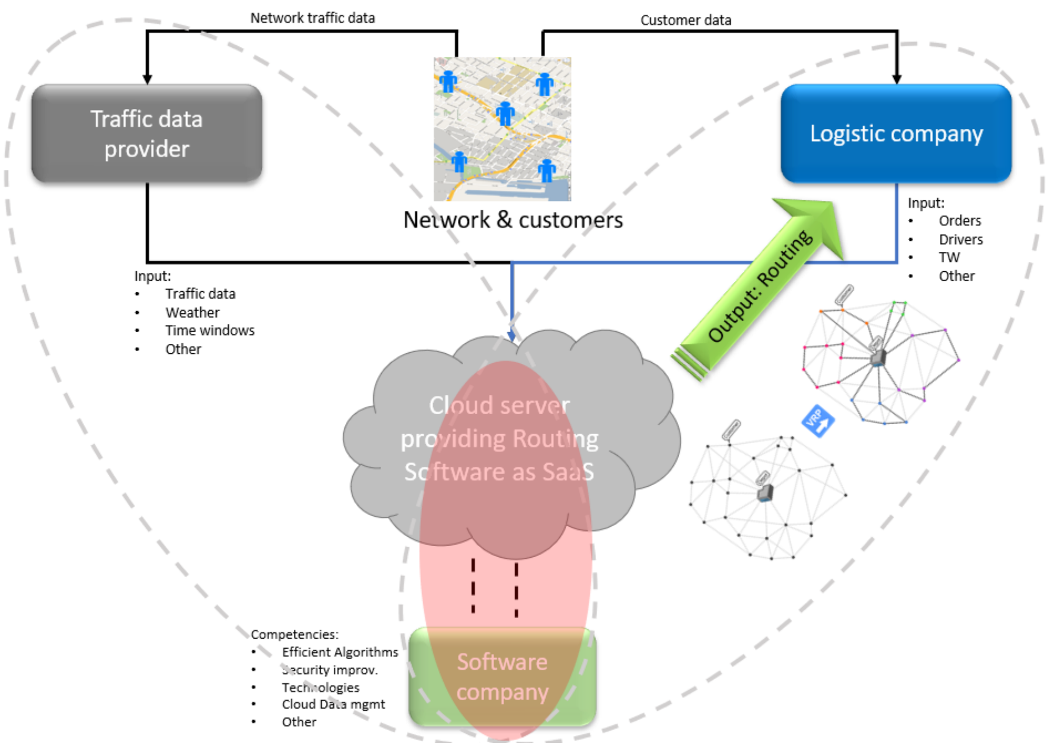

Figure 5.1 - Overall system relations and idea

The overall architecture's idea (figure 5.1) may be resumed in three main concurrent entities at stake:

- Logistic company
- Software company
- Traffic data provider

5.1.1 Logistic company

Logistics businesses are the beneficiaries of the service since they outsource the entire operational routing and scheduling management to a routing software provider, since the specialization and the infrastructure of the latter allow to perform much better achievement than the company itself would do and at a lower cost, thanks to economies of scale and specialization. It is necessary that the logistic companies submit necessary information to the architecture of the software system such as customer related information, fleet related and others useful for system's architecture modules. Fleet operating vehicles are not only the real end users of the routing and scheduling service, but they are an active part of the system because communication with the cloud software is twofold: downstream the application is delivered, upstream position, status data and drivers' messages are broadcasted.

5.1.2 Traffic data provider

The traffic data provider (or providers) plays a key role in meeting the dynamic needs of the system, feeding the architecture's components at regular intervals that need real-time traffic data information for both simulative purposes and rerouting. The traffic data provider assumes the processing burden of data detected across the network through sensoristics, FCD method or other, producing as output relevant data (live travel times, speeds, OD matrix, incidents) then used by software. Unlike the logistics company, the traffic data provider is a third party, a supplier of the software company. The system envisages the traffic data communication to the software architecture to take place straight to the cloud, so that the software, stood as SaaS, has the necessary current network data without interruption or intermediaries.

5.1.3 Software provider

The core of the system is not the on-cloud software company itself, but rather the SaaS it offers. As shown in the figure, in the cloud environment, in fact, we realize the synthesis of the service that incorporates logistics data, real-time traffic data, routing and scheduling algorithms, decision-making process and other accessory components. It all materializes with the delivery of the output to fleet's drivers after logistic company's user confirmation or modification of the delivery plan. as an application, that runs in cloud and offers a continuous service of static and dynamic routing directly on the on-board devices connected to the web platform, which in turn act as retroactive feedback sources. The SaaS framework means that any update, modification or variation of cloud platform components is automatically declined on all the clients. Despite the shared resources in cloud, it should be remembered that the system structure is proposed as configurable multi-Tenant Scalable SaaS ("true SaaS"), so that the different companies needs are met.

5.2 Back-end

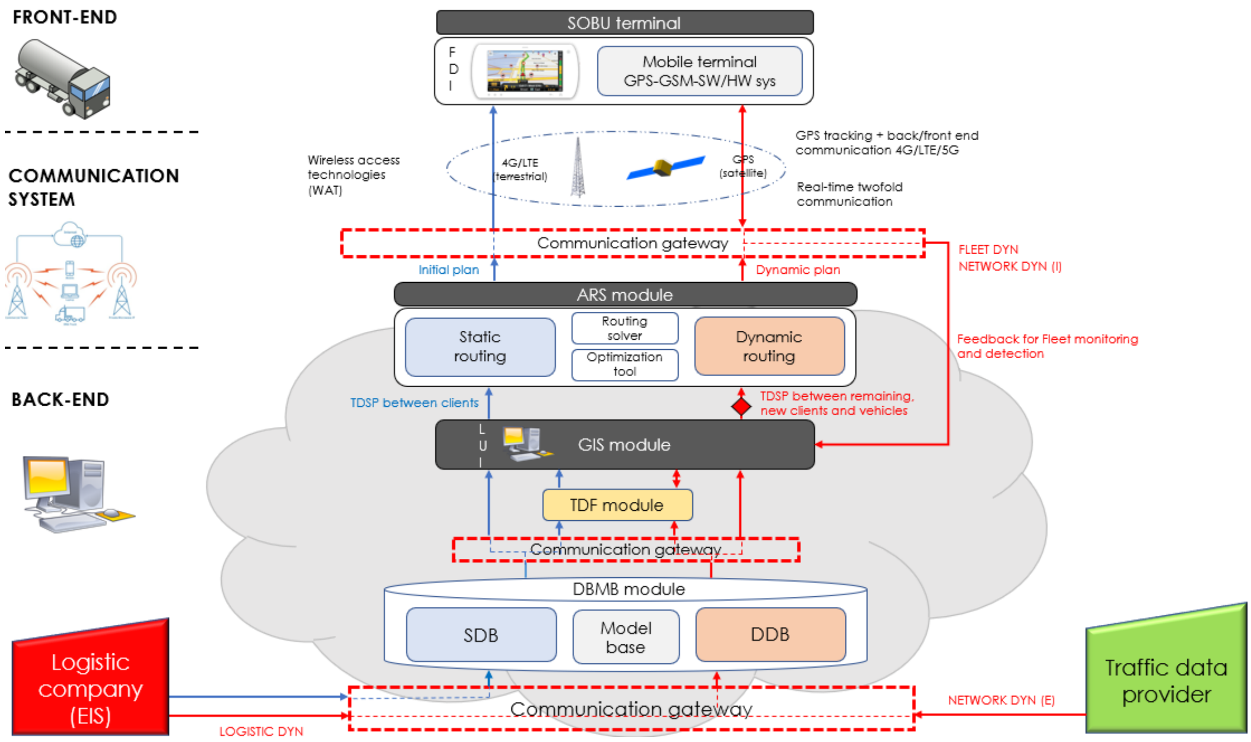

Figure 5.2 - System architecture (module-based)

The back-end of the proposed software architecture (figure 5.2) consists in a **Decision Support system (DSS)** structured as follows:

- Fleet Drivers Interface (FDI);

- Logistic User (LUI) Interface;
- Data and Model Base module (DBMB);
- GIS module;
- Advanced Routing and Scheduling module (ARS);
- Time Dependent Forecasting module (TDF);
- Smart on-board unit module (SOBU).

The input data providers are the logistics company and the traffic data provider. In the following section is explained in detail how these operate, communicate and interact with each other.

The back-end of the proposed System aims to be an advanced fleet management system (AFSM) tool that supports initial routing plans assignment decisions, but also helps to monitor the delivery plan and possibly be able to react to changes system dynamics by updating routing plans in an automated, efficient, and sustainable manner. The DSS is in this sense seen as a VRSS (vehicle routing and scheduling system) since it is designed to address the complexity of big data management deriving from the SaaS nature of system, in fact, one has to think about the amount of data (dynamic and not) even only one single company faces to operate the scheduling and the routing. Now one has to evaluate the amount of data that the system has to sustain and treat in a multi-tenant structure (each customer of the service is a client).

In substance this back end system should allow to integrate, incorporate and perform the three main steps which involves the shipments: the routing, the scheduling, the dispatching.

- **The routing (transportation planning)**: fleet planner's job, whose purpose is to minimize the travelled distance and vehicle's number. It determines how and where the work is to be done (spatial problem), what type of vehicles are to be used and it prescribes the path and sequence of clients (nodes) to be followed. It also includes items pickup list and pickup-delivery locations list.
- **The scheduling**: fleet scheduler's job, whose purpose is to minimize operating costs. The objective of the dispatcher is route's assignment, namely to decide who (drivers) is to do the transportation task, with what type of vehicle and when the work is to be done, thus time sequence schedules (the required time windows for arrival).

- **The dispatching**: it is defined as the physical release of work authorization to an operating facility (vehicle) in accordance with a previously established plan developed by routing and scheduling functions.

Hence, one single user logistic user is enabled to perform route planner, scheduler and the dispatcher's job in a holistic manner, though one single interface.

LOGISTIC COMPANY (LC)

The *logistic company* is the party involved in the software routing service (pre-back end), so as a first step it is necessary to first establish the terms of purchase of the product service, which we remember, is delivered as a SaaS, requires the monthly payment of a subscription fee, whose amount depends on the service features requested and whose continuation depend on the client's satisfaction, who in case can unsubscribe (mechanism explained in figure 5.3). (For further information please refer to the appropriate section 4.1). The different requirement of each company, as well as the different goals, translates into an identical, but dedicated and configured instances in architectural terms ("True SaaS"), in the formulation of VRP with different constrains, inputs, policies and rules in algorithmic terms.

From this point onwards the company must establish a close cooperation relationship with the software company, providing all the data that the system architecture needs in order to operate in a continuative and integrated way, both in terms of historical data (already present in the company databases: customer locations, depots, warehouses, fleet data, historical travel times etc.) and dynamic data (new orders, customers time windows, availability of vehicles, unforeseen events, modifications involving the shipments etc.), as well as others useful for system's architecture modules. Especially in the context of dynamic customer request, the timeliness of communication is essential. This is the reason why the proposed system aims to be able to manage a connectivity with Enterprise Information Systems (namely ERP, WMS as data sources) through a web Application Program Interface (API), i.e. adaptors for standardizing inputting information from various EISs and devices.

The proposed architecture envisages that this information is shared and stored directly in the system's cloud platform, which compartmentalize and isolate it from other logistic companies' information, specifically in the DBMB. The sorting of the specific information is accomplished by the Communication gateway (CG), which will be explained further on. Even dynamic modifications or requests, whether occur, enter directly into the cloud. That is why the Logistic company is responsible what we will call "logistic dynamism". Once the integration has been stabilized, the service becomes operational in a short period of time, given its SaaS nature.

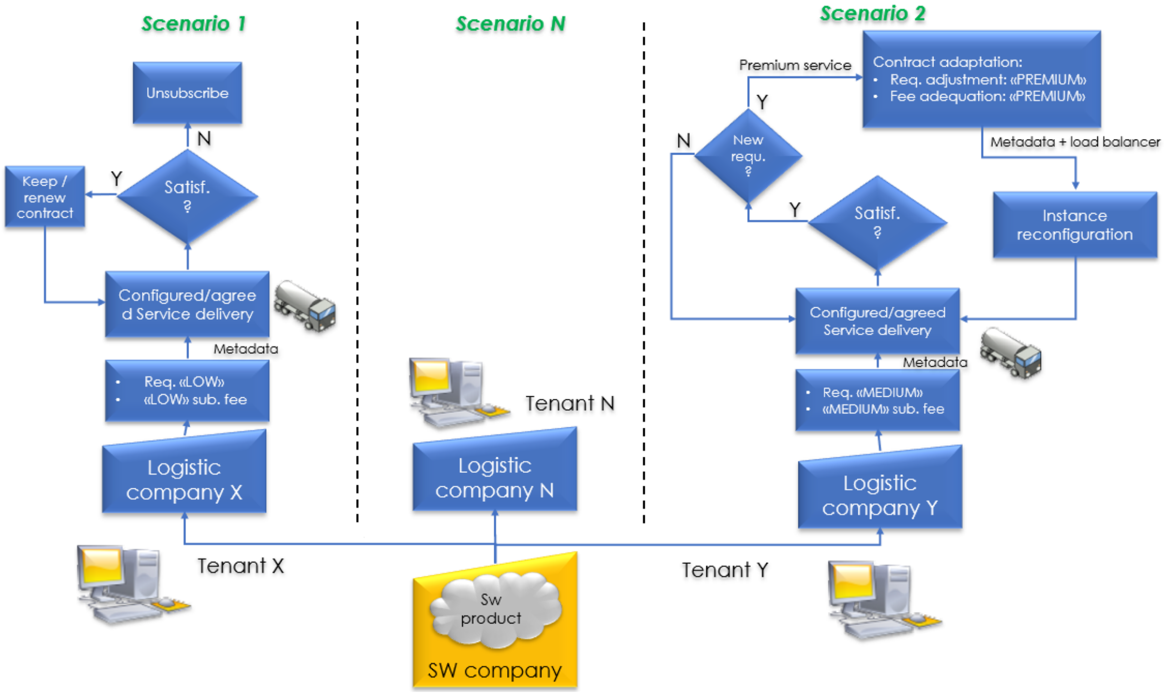

Figure 5.3 - Distribution of Software product and subscription contracts

TRAFFIC DATA PROVIDER (TDP)

The *Traffic data provider*, or providers, (TDP) is the third part involved in the system and is essential to guarantee the dynamic features of the system. As explained in the chapter 3, it assumes the burden of traffic data treatment, understood as collection, validation, conversion, combination, processing, model calibration (static and dynamic), visualization and data distribution; detected by a plethora of road sensors scattered across the network. The different data collection technologies described provide software company's servers, via wireless, a wide range and heterogeneous data that combined and worked allow the extraction of a network information which is relevant to the network users and commercial application (figure 5.4). Some of them may offer travel times, some others raw data. It has to be reminded that the current trend towards smart applications and high ICT performance and the current technological state of sensing devices, make them assume the form of big data considering their peculiarities of volume and size, velocity and variety. Thus, the traffic data provider must cope with big data complexities, therefore disposing big data analytics' tool. Most of the traffic data providers mentioned are used as a possible choice, however it is proposed as provider a company that ensures wide European coverage of the service (which points strongly on FCD method via worldwide smartphones base), fair accuracy (cost-efficient accuracy) and high upgrade frequency. The real-time traffic data are sent directly into the cloud and in particular in the dynamic DB sub module, which subsequently will be incorporated in the GIS module and in the Forecasting module, allowing to feed the system with real-time data from the network, fundamental for the functions of fleet detection, monitoring and rerouting that will be explained in detail afterwards. That is why the Traffic data provider is responsible for what we will call "(external) network dynamism". It is worth remembering that other data from the field are detected by the company through GPS fleet tracking, which will be deepened subsequently and that will be responsible for what we will call "(internal) network dynamism".

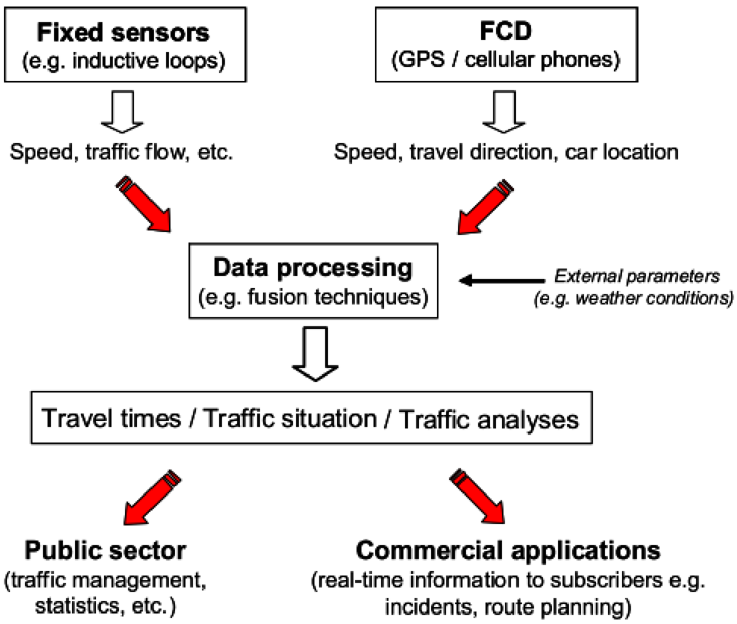

Figure 5.4 - Traffic data sources and data treatment output.
Source: [20]

LOGISTIC USER INTERFACE (LUI)

The *logistic user* interacts with the system in cloud thanks to the DSS interface (GUI). The framework in which it has been choose the Route Planner to work in is the GIS cartographic environment, thanks to its intuitive interface, so that it has access to submitted data and manages them directly into the system. Subsequently, dynamic events such as new orders, cancellations, changes in the service or any need of variation either requested by customers or due to business needs (logistic dynamism), are transmitted straight into the cloud and are made human readable and manageable. The interface is essential and key to the system as it must allow the user to have a tool that allows him to: enter data, view information about routes, network, customers and fleet, operate set-up analysis (es. Scenarios), simulations, visualize problem's results, in such a way to help him in the decision-making process. In essence, DSS system's interface allows you to manage in a holistic, integrated and smart way all data concerning customer's shipments, in order to make optimal decisions in the shortest time possible.

COMMUNICATION GATEWAY (CG)

The *Communication Gateway (CG)*, which has been inspired by the architecture proposed by [44], has been conceived as a mandatory passage for any information circulating within the architecture. Its role is fundamental given the criticality and the amount of information at stake. Specifically, it is designed as a sorter and filter for every data communication and has the role of ensuring access and data validation. In fact, given the heterogeneity of information that crosses the CG, it has the key role of sorting data to the specific module of the architecture, transforming its format to make it tractable by the module at hand. For example, data from the TDP are sorted first to the DB's dynamic submodule, then to TDF or to GIS depending on the needs, as well as logistic, qualitative and quantitative information from the various companies (directly from their ERPs) might be sorted to the DB 'static or dynamic submodule. In this latter case the CG can access EIS's various data through the API.
It is clear that the CG operates transversally from the back-end to the front-end, constituting a component that surrounds and extends to the whole architecture (as shown in figure 5.2). For this reason, the idea is to assign to this module also basic operational pre-processing tasks. For example, acting on the data that impacts on routing by aggregating customer orders according to load type, node localization in the

network, in some way streamlining and absorbing part of the GIS and ARS modules solver's work, therefore facilitating the computation. Furthermore, the CG provides readable information to the GIS providing the right format for visualization of the routes through the interface, facilitating the decision maker's job. The information sent to the SOBU systems passes from here as well, therefore it ensures the operation of the aforementioned real time twofold communication that must exist between the cloud system and the drivers, acting like an intermediary.

DATA-BASE AND MODEL-BASE MODULES (DB / MB)

The *Data Base and Model base module (DBMB)* incorporates Data Base module's and Model Base module's functions and represent the knowledge container of the system. It consists not only of numerical, textual, spatial data, (utilized by the GIS) but also includes mathematical models, algorithms, network characterization, rules, linear programming models that will be used by the system in the ARS module. It includes all the data that constitutes both the system inputs and all the logic rules, priorities, policies and other relationships that characterize the system. The data contained could be subdivided into logistic data, network data and VRP data, and can be further subdivided into **static [S]** and **dynamic [D] data**.

- **Logistic data** we find: customer data (geographic location [S]), depots or warehouses locations [S], product data [S] (capacity, constrains, weight), soft and hard time windows [S/D], premium or standard service [S], dynamic orders [D], orders modifications [D], known demands [S], fleet data (size [S], capacity [D], general technical [S], depot constrains [S/D], GPS tracking data [D] etc.) and schedule data (historical service times [S], driver's shifts [S/D], drivers regulations [S], assignment rules [S], load acceptance policy [S]);
- **Network data**, (shared with the GIS module) we find: Network graph [S], Historical travel times [S], Network Restrictions [S/D], periodical or special events [D] (accidents, emergencies) and the real time traffic data from the TDPs [D];
- **VRP data**: saved VRP models and methods [S/D], VRP rules [S/D], policies [S], weights [S/D]).

Given the twofold nature of this data, two submodules have been created, namely the *Dynamic DB submodule (DDB)* and the *static DB submodule (SDB)*.

The specific logistical data of customers and the fleet have been submitted directly from the logistic company at stake, therefore it consists of an aggregate DB or knowledge deposit formed by routing related companies' data. In order to standardize accessing from different EISs adaptors will be used. However, the nature of shared resources that characterizes a cloud framework allows the company to implement a pooling of information from which to extract useful knowledge in general. This is a further advantage of the deliver the service as a SaaS and in general of cloud platforms. Finally, this module includes both the Database and Model base management systems, whose role is to organize, acquire, modified, update models and regulate the access to the knowledge. To conclude we can say that this module provides the system all the dynamic and static data it needs to operate routing and scheduling optimization (the only "internal" data are the ones from the fleet).

TIME DEPENDENT FORECASTING MODULE (TDF)

The function of the *Time Dependent Forecasting (TDF)* module is to obtain time dependent travel times, through prediction models and forecasting methods. These are applied either to historical data in the case of static routing, or to real time data (raw data from TDP) for dynamic routing. Basically, it retrieves on two different databases depending on the purpose (figure 5.5).

- **Initial plan forecasts (static routing)**: the TDF is therefore fundamental in the preliminary phase of the initial routing plan's creation, the time dependent travel times are calculated by applying forecasting and prediction methods to the stored historical data in the SDB, from which it is possible to extract historical trends and traffic patterns, seizing the variability of traffic conditions over shorter or longer time period. Efficient and effective forecasting methods produce robust and reliable routing and scheduling plans which consequently reduce the likelihood to resort to dynamic rerouting, which in any case requires a significant computational effort (especially in a scalable SaaS architecture context, where several companies may request the recalculation at the same time).
- **Dynamic forecasts (rerouting)**: the system envisages to have one or more TDPs, some of them may procure raw data (traffic conditions, traffic analysis), some other

travel times, whatever type of data they are providing they constitute a real time traffic data DB (see figure 5.5) from which the TDF can draw to acquire knowledge on the actual network's condition and evolution, then to calculate dynamic travel times (due to traffic congestions), enriching the available architecture's knowledge in terms of *network dynamism*. In this sense the TDF module aims to extend its functionality to the dynamic facet of the system, allowing the system to have a vision of the future, in two ways: processing live raw data coming from the TDP ("external" network dynamism) and the live raw data collected by the operating fleet ("internal" network dynamism, through CG and GIS), thereby retrieving forecasts for possible rerouting, limiting the risk of traffic congestion, which means delays or missing orders, therefore enhancing the service (this consists in a further support, since the GPS already carries out travel times).

This addition to the work of the TDP during dynamic routing increases the accuracy of real time estimates, enhancing the robustness of the DSS even in the face of important developments or particularly volatile networks, as well as obviously enriching the knowledge on traffic patterns. The TDF is therefore the module that allows both network dynamism and logistic dynamism to be accessed, both provided by the TDP and Logistic company in the DDB. The first through the processing of raw data from the urban network, has been explained above.

The *logistic dynamism* is a crucial driver since a new customer pending request (ongoing route) require the routing feasibility to be verified among the criteria that determine the acceptance of the order, and this happens through travel times evaluation. For this reason, the TDF module continuously communicate with the GIS module: in fact the latter first of all updates the map with the coordinates of the new customer coordinates, vehicles, fleet status and traffic information, therefore calculating the distances, afterwards the TDF operates the travel times estimation, to send them again to the GIS, so that the feasibility and the delays are verified and eventually the rerouting is triggered. It may happen that an incoming order is particularly important (premium client or service), in this case the rerouting is triggered even if the feasibility check is not met, therefore some vehicles will be redirected and maybe less important customers will remain unsatisfied. This is based on the weight of the customers and the criteria established by the company, which translates into modification of the weights in the function that the routing algorithm has to optimize. This process is automated and made human readable to enhance the decision-making effectiveness. For what concern the internal network dynamism (from the fleet) it is not necessary to switch from the TDF module, the detection suffice, because the travel times can be directly obtained from the tracking process.

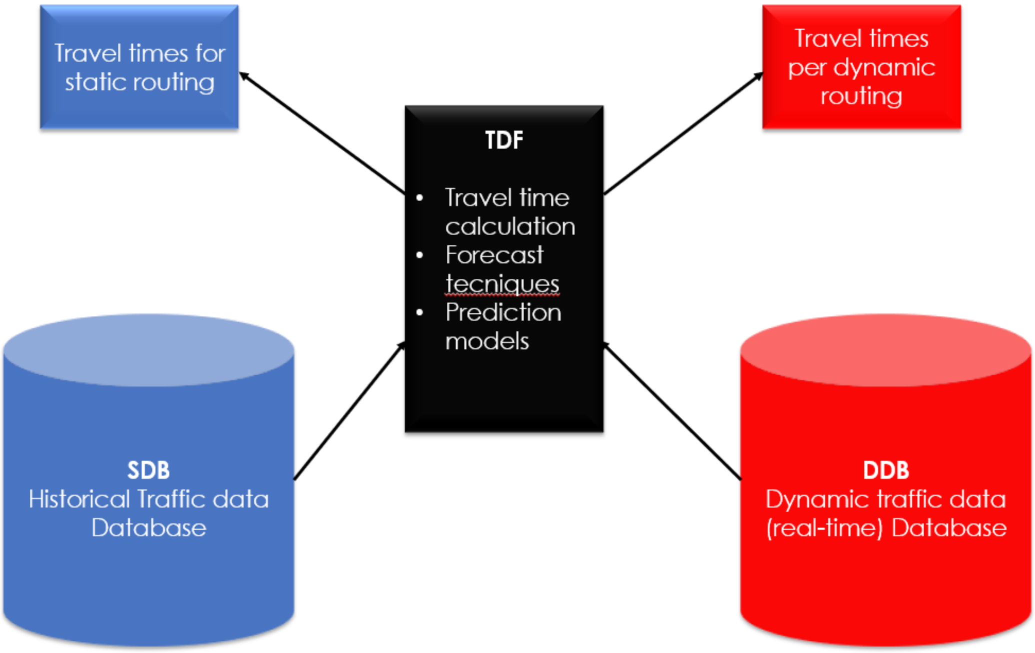

Figure 5.5 – Forecasting methods applied on different databases

The TDF module in this architecture wants to be an efficient tool for static routing and further support for the DSS in terms of estimations in dynamic cases, as it should be remembered that the TDP absorbs the burden of big data analytics and treatment. This is not only in an expensive data treatment job, but also in a great opportunity, in fact: the great variety (heterogeneity) of data constitutes a source of rich knowledge; the great speed allows to short interval data communication, making the system more adherent and reactive to reality, to the evolution of the state of the network; the big amount of data allows more solid statistical analysis.

From the operational point of view it is a matter of solving forecast algorithms, consequently the module draws on models, rules, etc. from the SDB as it also has a forecasting solver to fed. Whose final goal is to assign to each arc of the graph (fed to the ARS) a certain expected time dependent traveling time. In conclusion, the pivotal role of this module is clear, since an effective travel time estimation ensures more accurate routing solutions, namely closer to the ideal optimum. This is especially important when time windows constrain are to be met.

GIS MODULE

Given the matrix of the problem and the need for decision support is proposed a 3D graphics GIS environment as a working interface (LUI) which may work closely with the system's operative modules. In essence the role of this module is to handle the geographic data pursuing geospatial tasks whose output is required by the ARS module of the architecture.

This allows to dynamically manage and handle the information on hand, making the process smarter and under control. The purpose of this module is to provide the user with a comprehensible working interface and to facilitate the static and dynamic planning of routing, scheduling and monitoring, supporting the user thanks to the multiple features described in the appropriate section (see 4.3).

A GIS platform allows to make the GPS tracking visual on the map and to report a series of information on the real time status of the fleet and the delivery plan that constitute the input for the ARS. Thanks to the GPS and GPRS connection the GIS module is responsible for monitoring, detection and eventual adjustments, which consists in the key function of the system, namely static e dynamic purposes the architecture aims to achieve.

The GIS module works in close contact with the DB and MB, TDF and with the ARS modules, acting like an intermediary, in fact the close collaboration between ARS and GIS consists in the very pivotal core of the system.

Once the digital urban map is available, an editor allows to stratify database's data and geo-reference interested entities such as depots, customer locations, trucks, warehouses, and reporting their respective peculiarities as known demand and load. All this information is caught and imported from the SDB and DDB submodules.

Thanks to this information and the forecasted travel times provided by TDF (and as one of GIS software features is the provision of shortest path algorithms, which can be harnessed as a part of a routing system), the GIS module can create a time dependent shortest path between clients or nodes (which have already been organized in neighbourhoods by the CG), not yet optimized (the output values form an OD matrix that will serve the routing solver). This, it sent to the ARS which optimizes it by the routing solver. Subsequently the near optimal routing solutions are evaluated, through the LUI, by the logistic user that, through GIS functionalities, is able to visualize, analyse, accept or modify them manually according to his experience. At the end of this decision-making process the optimized plan is sent to fleet vehicles, which can start the first daily routing plan.

Once the work shift has been initiated and the path to be followed is communicated to each vehicle, the GIS's module key functions are performed: namely it deals with the management of the dynamic facet of the routing, which it could be divided into passive and active.

- The former follows the system in a *dynamic passive management*, in the sense that it is to detect feedback from the network, without elaborating changes.
  - It consists of the **fleet service monitoring**, which verifies the status of completion of the service and the initial delivery plan, keeping track of each route thanks to the tracking GPS mechanism (devices are equipped with GPS systems, of which we will speak later, which enable the knowledge of the geographic coordinates) and the Internet connection. Operationally the module keeps track of the current routes, not only in term of geographical displacement of the vehicles, but also in terms of parameter's communication, such as instantaneous residual capacity, served customers, cost, (through GSM/GPRS), speed, delay, time, distance (through GPS) etc. as well as it allows human readable monitoring to the route planner.
  - The fleet monitoring is accompanied by the **fleet detection**, which is the simultaneous calculation of deviations from the initial (initial/static route plan) plane (or forecasts). In practice the system compares initially estimated travel times (from historical data e.g. customer to customer or customer to depot matrices) with the actual ones detected from the field, which are provided dynamically by the traffic data provider (or the TDF module depending on TDP) and by the operating fleet which acts as a probe. In this sense another key function of the module is the integration of current traffic data. Hence, fleet

detection relies not only on the historical archive, but also on live updates from the network in such a way as to allow an analysis of the dynamic and continuous deviations compared to network's evolution (e.g. in terms of delays generation that may impact on the time windows, consequently the expected initial sequence of customers). However, the rerouting is kept under control (verified rerouting) in order to regulate system's nervousness (frequent rerouting). This depends also from companies' specifications (delay tolerance): the higher the required service level is, the more the system will be nervous, hence there is an optimal trade-off between system's nervousness and service level.

➢ The *active dynamic management* is **triggered by an event** belonging to one of the following sources of dynamism: logistic dynamism (new dynamic request, demand or time windows variation, order modifications), network dynamism (accumulation of delay such that time windows constrains are violated, other unforeseen events such as accidents, deviations, time limits, bad-weather, etc.), fleet dynamism (vehicle breakdown, load, capacity or working hour constrains exceeding). Fleet or network's dynamism triggers are detected either through retroactive feedback mechanism from the system, ("internal" whether they are data collected from the field by the fleet and communicated wirelessly to the online system), or "externally by the TDPs. The system admits also a proactive feedback mechanism, stood as a recalculation not due to delays, but due to the finding of a faster way (new best solution) operated by the optimization tool. Logistic dynamism is an "external" trigger as well, since it came from Logistic companies which communicate eventual new events. In this sense the decision-making process involves also this facet, in fact, operationally before starting with any rerouting operation the system evaluates the acceptance of an order from a routing and scheduling feasibility point of view, this is possible since it draws incoming new orders information from the DDB submodule.

Regardless of the nature of the dynamism the GIS manages this evolution of the state of the network by updating the cartographic environment with new travel times, new customers to serve and fleet status, creating a new time dependent shortest path between remaining clients, new clients and vehicles, which will also be fed to the ARS module for optimization (of heuristic algorithm).

Another GIS function is the translation of the referenced network map, that is an urban network with the entities inserted (road representation), in a weighted directed asymmetric graph (arches representation), so as to feed the vehicle routing solver (presented in the next module) the problem declined in algorithmic terms.

Finally, on a practical algorithmic level the incorporation of updated travel times provided by the TDF module may cause either the modification of the weights assigned to the viable arcs of the road network, or possible violations of time windows constrains, in both cases the provided shortest path/graph to the ARS module's routing solver may change, causing a possible new best route solution (dynamic routing plan, see figure 5.6).

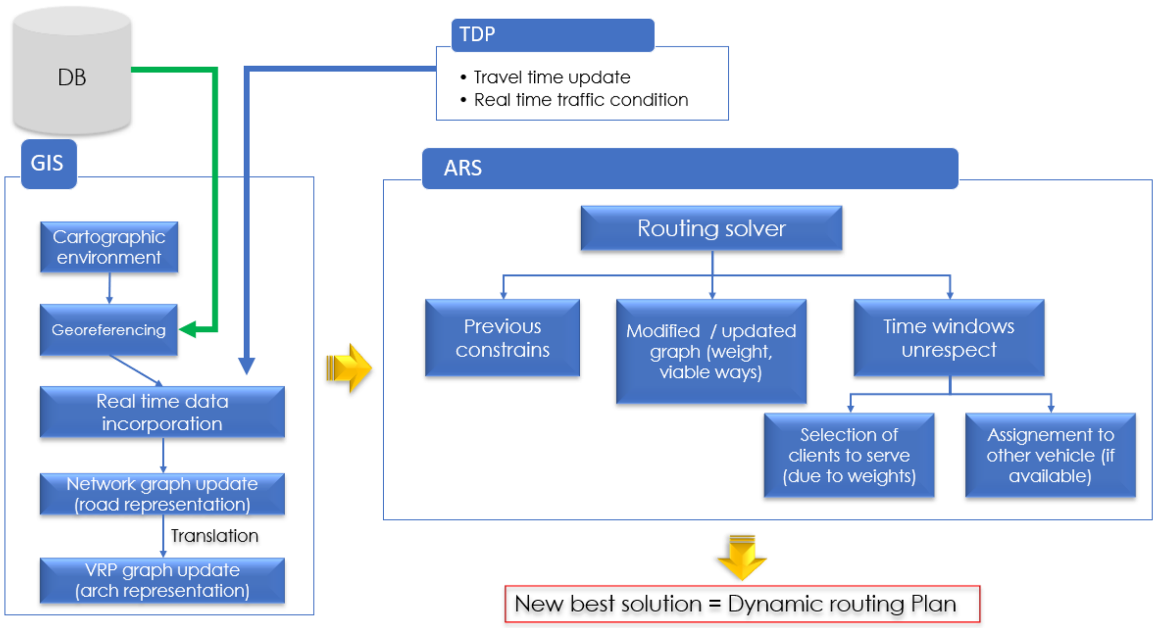

Figure 5.6 – GIS graph translation and consequences on Routing solver

ADVANCED ROUTING & SCHEDULING MODULE (ARS)

The *advanced routing and scheduling (ARS) module* together with the GIS module represent the thinking brain of architecture, as they incorporate the operating functions of the system. Once the problem has been defined and deployed through the GIS module's interface, useful answers have emerged and are available to the decision-making process.

The problem processing module incorporates a **Routing Solver (RS)**, a mixed-linear programming component that operates in the software company's cloud servers. Here we find the mathematical resolution of the Vehicle routing problem. In particular the system aims to be solid solver of *Time dependent VRSP of pickup and delivery with time windows* (for detailed explanation refer to Chapter 2). The routing solver has an initial library of routing mathematical models (which reside in the DBMB module), but the service specification, the routing targets, the criteria and weights, etc. and in general the VRP features can vary depending on the different logistic companies (and their customers) requirements. The proposed architecture meets these assumptions. The configurability features of the routing service offered as SaaS, are achievable thanks to the scalable, configurable, multi-tenant-efficient structure. In practice the architecture is proposed as a "true SaaS" in the maturity scale. This SaaS structure has been chosen because it is computationally efficient in cloud, since one common instance is shared, it allows configuration, thanks to metadata and allows scalability, thanks to the load balancing layer.

As the sources of dynamism can be contemporary and overlapping and the system must be able to manage them simultaneously. This new set of modified inputs, is fed, through the GIS, to the routing solver in cloud that solves the updated VRP with real-time data. It is important to emphasize that the previous constraints are still to be respected in the new solution. Obviously, the system focuses on a resolution of the problem in a reasonable time in order to be adherent to the real scenario. For this reason, heuristic or metaheuristic approach is proposed instead of an exact method, with suboptimal, but more reactive solution (as suggested by most of the authors in the literature).

This module is responsible for generating both start-of-day routing plans, with initially known (or partially known) customer demands, and dynamic plans, with live adaptation, whether there are significant deviations from the expected initial routing plan or new dynamic data, in order to achieve the established service objectives even in response to system's changes. From both a static and a dynamic point of view it optimizes the initial and the rerouted solution elaborated by the GIS as shortest paths respecting the constrains.

- **Optimized initial routing plan (static routing)**: it takes into consideration only data already known about a delivery and demand. the goal is to assign an initial route to each vehicle in the fleet, in order to create an optimal sequence of customers (inspired from the TSP) that minimize a function. The aim is to satisfy the demand (assumed at the beginning known a priori) and requested service by the customers, respecting the constraints of capacity, constraints imposed by the customers (time windows imposed and priorities based on the assigned weight) and constraints imposed by the network. The choice of an initial path is assessed taking into account time dependent forecasted travel times, which means that depend on the departure time. This is done either by the TDF, exploiting the knowledge of historical traffic data and distances, or they are provided by the TDP. On the algorithmic level this translates into giving a cost, or a weight, to network's arches to which also contribute the length of the tract, tolls etc. Furthermore, the network loses its symmetry property, because the weight of the same trait is the function of the travelling direction, determined by the travel time or other constraints (unique senses, weight limits, restrictions, events etc.). The routing solver thereby yields to optimize a weighted function, whose objective depends on the company's objectives, such as, the maximization of served clients, minimization of vehicles, maximization of load saturation, maximization of premium customers, even if usually the target is the minimization of transport (operational) costs. The DSS supports the logistic user (in a GIS environment) in this initial delivery plan decision-making process by presenting data, reports, analyses that can help you take a final decision. In addition, the architecture is made to allow the routing planner to be able to vary parameters or initial inputs in such a way that it can compare scenarios and alternative routing solution.

The system is designed to react accordingly and create a new plan optimized to serve the customers still dissatisfied, but more generally general it focuses on optimal modification of existing assigned routes according with new data.

- **Optimized dynamic plan in case of Logistic dynamism (re-routing):** whether the source of dynamism is the customer the online communication of new dynamic request, demand variation or order cancellation to the software company is necessary. Operationally first of all the system evaluates the acceptance of an order in feasibility terms and due to its additional marginal cost, which is function of the extra distance to be travelled. Then, if so, the updated algorithm is resolved online and the updated routes (with eventually the new sequence of customers), or a new route with new trucks in case of unavailability (for time windows or capacity constrains), are communicated to the drivers' devices.

- **Optimized dynamic plan in case of Network dynamism (re-routing):** on operational level in this system real-time updates from the network are provided either by the traffic data provider ("external"), or through tracking ("internal"). On a practical algorithmic level these travel times update may provoke two things: modification of the weights assigned to the viable arcs of the network, possible violations of time windows constrain, in both cases this may result in possible new best route solution (dynamic routing plan).

- **Optimized dynamic plan in case of Fleet dynamism (re-routing):** thanks to the tracking and GSM system real-time updates from the point of view of the state of the vehicle are gathered, intended both as loading conditions and as delivery plan's progress. The case of vehicle breakdown, goes beyond the scope of the thesis, however, the underpinning idea it is to verify the customers still to be served that were foreseen in the route of the broken vehicle, and try to reassign them to other vehicles or to create eventually a new route. The complexity is in the joint evaluation of the availability of load, time of service and respect of the time windows and weight of dissatisfied customers, taking into account that in any case should be assured the deliveries planned in the route plan of the other fleet vehicles.

Whatever the trigger is, the rerouting may imply a vehicle dynamic re-sequencing (rerouting) and re-assigning, or the generation of a new route (new vehicle).
The proposed system offers also a proactive feedback mechanism (figure 5.7) thanks to the **optimization tool**, which allows the continuous search for a best solution with the latest input set. Obviously, the current actual best solution is retained waiting for a new best one and a minimum improvement threshold is to be set and whether significant improvements to the solution are obtained, they subjected through LUI to be visualized and eventually adopted as a new plan in case of confirmation. The support of DSS to the routing and scheduling manager in the process of dynamic routing, is the same as that happens in the initial plan (here of course it is real-time

data processing), however in this case is even more fundamental, as the timeliness of decision is essential to be adherent to the volatility of the system. For detailed explanations of how the dynamic and static routing process works, please refer to Chapter 2, the thorough description has been dodged to avoid repetitions.

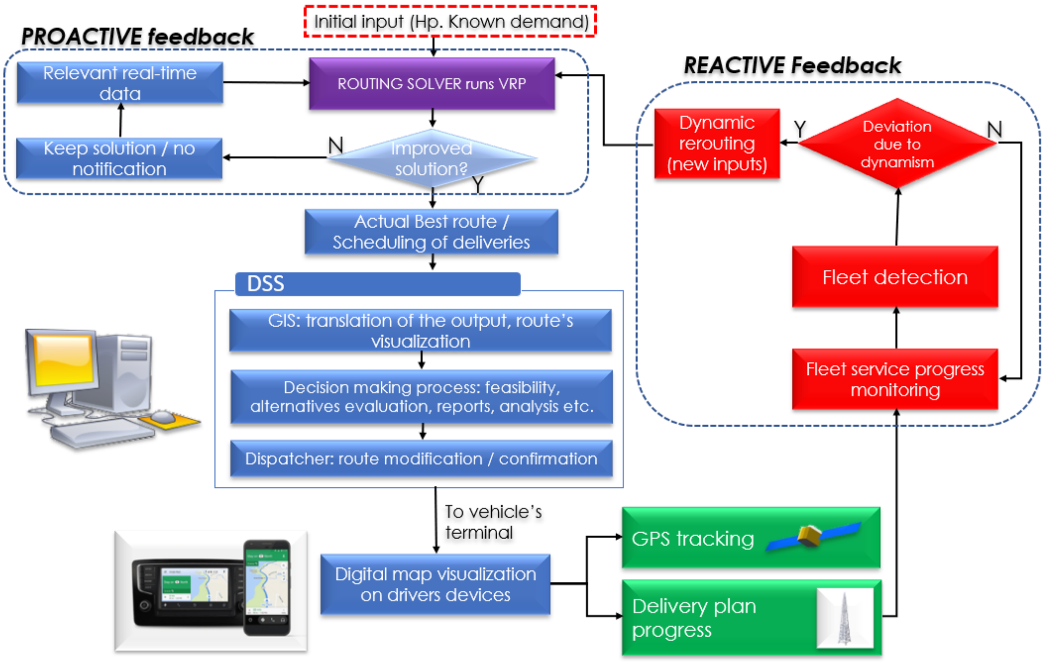

Figure 5.7 – Proactive and Retroactive (reactive) feedback

5.3 Communication layer

The middle layer, communication system, provide the connection as well as it allows informative and operative communication between the back-end and front-end of the system, so it must be a dual and real-time data transmission between vehicle and dispatcher. The communication infrastructure is obviously the means by which the whole system can operate and it happens through Wireless Access Technologies (WAT). The architecture envisages the use of two systems: the GPS communication systems and the terrestrial network.

✓ **Global Positioning System**: it consists in the use of satellites that localize the coordinates of the fleet vehicles; therefore, it is fundamental for the tracking function. Allows on-board devices or in-vehicle systems to communicate their position to the satellite which in turn communicates it to a GPS receiver, which finally sends the coordinates to the systems requiring them, at regular intervals. In our case these data enter the CG that subsequently will sort them to the GIS module for further detection, monitoring and, eventually, rerouting. The latter module allows the spatial representation on a cartographic basis of each vehicle in service in real-time. Thanks to the GPS infrastructure it is also possible to go back to the live travel times (additional to the one provided by the TDP or TDF), that will be used by the system to address the retroactive feedback and implement the fleet detection with the expected service times to possibly plan a rerouting.

- ✓ **Terrestrial mobile communication network**: through this the operational transmissions between the on-board devices/in-vehicle systems and the system in the cloud take place. These are global standard communication protocols for mobile devices and terminal. Given the dynamic purposes of the system is inevitable that key parameter is the bandwidth. The timeliness of communication and, therefore, execution can be ensured only if the technologies at the base ensure stable and rapid connection of a large amount of data in both directions, upstream and downstream. An adequate timeliness for the proposed architecture translates data transmission at very short intervals considering real-time needs. The choice of the communication protocol, is beyond the objectives of the thesis, however in the dedicated chapter, the considerations led to consider adequate to the system the Protocol 4G/LTE Advanced (5G in future perspective). The reasons for the proposal of this technology are both of a performance nature (it ensures average download speeds of up to 400 Mbps) and because it is expected to become widely spread in medium-short term, moreover it ensures more security and integrity. Obviously, the twofold correspondence between the cloud system and the on-board devices, already fundamental for the monitoring of the (initial) delivery plan or the fleet detection, in the case of dynamic purposes assumes even more importance. It is necessary, in fact, on the one hand that the system knows the state of the vehicles and the vehicles know the routes, thanks to the communications via mobile network, and on the other hand geographic coordinates, through GPS infrastructure, to understand possibly how to redirect the vehicles. The fleet communication system allows the automatic dispatch of the position to the central system, to receive a destination (or text message) from the control unit, sending messages to the control unit. The typical process of dynamic communication consists of a continuous and twofold transmission of notification and updating of updated real-time data between vehicles and the central system.

From the point of view of the TDP, the communication system is the basis for the data collection from the field through FCD method, nowadays no doubt the most common method (see Chapter 3 for complete considerations). To date the fabric of the road network is sprinkled with WAT that allow to detect data and communicate them at the centre for further elaboration, as explained in the chapter the heterogeneity of data, the coverage, the accuracy recent technology are capable of, allow the use of this information for intelligent transportation systems, what our system is meant to be. At the system's level these real-time data upcoming from the network, pass through the CG that not only sort them to the modules mentioned, but also aims to send possible consistent data to the TDF module for further analysis and combination with already archived historical data, hence retrieving reliable forecasts for future rerouting.

5.4 Front-end

The front-end is the part of the system that physically operates in the network and runs the shipments, which is why we could define the fleet as the **mobile terminal**. Physically it consists of all the fleet's onboard software and hardware that may ensure a user-friendly user interface and allow back-end communication (through GSM protocols): collection and communication of fleet data, as well as the delivery plans progress and the execution of routing planes sent by the central system.

SMART ON-BOARD UNIT MODULE (SOBU)

Basically, the *smart on-board system*, in the most extensive configuration, includes a range of wireless communication sensors, software and hardware that retrieve trip data subsequently forwarded, through GSM infrastructure, to the central system for the fleet monitoring and detection. The devices usually include: navigation system for the route guidance; driver's display showing basic trip and load parameters and information; a GPS satellite tracking device; a terrestrial network communication and data exchange system which provide internet connection with the operation centre, in order to automatically send route data to the control unit, receiving a destination (or text message) from the control unit, sending coded messages to the control unit (figure 5.8).

An example of what can happen in the case of new (accepted) order occurs is that at the beginning of the day the delivery plan is communicated to the vehicle and whenever the vehicle completes the service to a customer it communicates it to the system. If a new customer request occurs (and the system evaluates its convenience) the GIS and ARS module resolves online algorithms, then, after the confirmation of the logistic user, the communication gateway CG transmit the updated routing plan (rerouting) and to the terminals of vehicles involved.

In the proposed architecture fleet's vehicles will be equipped with an Infotainment system (consisting of hardware and software that work in integration with vehicles sensors), which is assumed to be soon available in most of the manufactured vehicles, as they provide a complete routing assistance to drivers and enclose all the above mentioned devices functions. The main ones we are interested in include: LCD high resolution touchscreen, satellite navigation systems, Bluetooth connectivity, wi-fi router (internet in the car, then cloud application suitability), on-board diagnostic computer, voice and messaging controls, smartphone connection, messaging, external cameras. A further motivation for the adoption of the such systems, which is also a further novelty of the system, lies in the fact that, unlike the normal fleet management systems, it is not to have specialized on-board devices with data potential processing because the computational burden is borne by the ARS module in cloud, so in essence, the only requirement that on board equipment must meet are high resolution digital maps, user-friendly interface and high-performing dual remote communication (figure 5.8) and hosting the developed cloud software application. The telematic equipment on board as a whole allows execution in the order of: reception task, travel route guidance, items and load checking.

The fleet vehicles in addition to acting as executors of service, in the proposed architecture, also act as probes to gather traffic data for the central system because, thanks to the vehicle's equipped sensors (varying on the nature of the loaded products), provide real-time update about the actual traffic conditions, therefore travel time variations, hence they broadcast a retroactive feedback to the GIS module that eventually may plan a rerouting. Of course, this "internal" feedback (gathered from fleet vehicles only) does not replace the role of the traffic data provider (that gather information from the overall road network), which consist in a more reliable and accurate source of relevant traffic information, but are rather combined. In fact, feeding the VRP algorithm solver with more realistic information implies that even the solutions (routes) will be closer to the optimum.

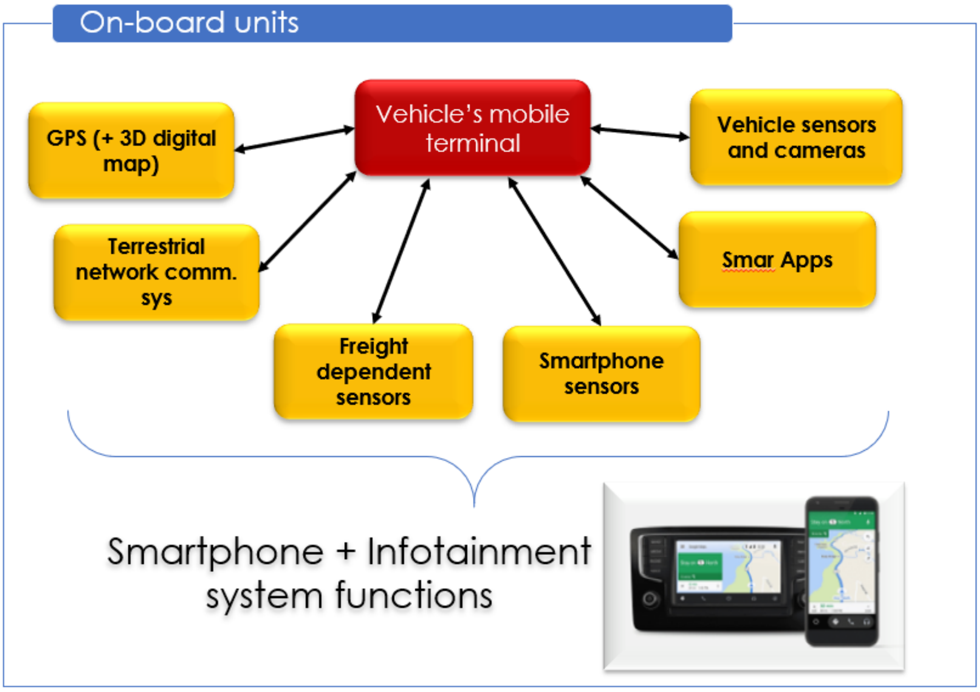

Figure 5.8 – SOBU devices

FLEET'S DRIVER INTERFACE (FDI)

The architecture is structured so that on-board telematics receive delivery plan indications and send back a continuous feedback to the system passing through the CG, then submitted to the GIS module in the cloud DSS.

We believe that Infotainment systems provide the best *fleet's drivers interface* solution, and since they are connected, they allow to exploiting smart business application, enabled by IoV framework, e.g. route guidance applications together with connected driving.

The routing sequence resulting from the resolution of the *time dependent VRP with Time windows* is sent from the cloud to the vehicle's infotainment system (smartphone-navigator integration), and specifically it is displayed on high resolution LCD touchscreen in a digital 3D map environment (for a better orientation especially in big cities). In this context, SOBU's interface assist the driver, while performing the trip, with traffic alarms, routing navigation indications, notifications, possibility of sending/receiving tasks, delivery status, messages or notifications to/from back-end, showing vehicle status information overview.

During the execution of the route there is a continuous transmission between smartphone and cloud system through terrestrial network (4G/LTE advanced), so that the user/route planner follows up the movements of the fleet and the progress of the delivery plan through the LUI (incorporated to GIS module). Thanks to the bandwidth of 4G/LTE advanced, served customer updates, vehicle's failure, delay or unforeseen events, take place at nearly real-time intervals.

Eventually, the driver also has the possibility within certain parameters and limits, to vary the assigned plan to follow a preferred one (simply through the touchscreen), the system still keeps track of these modifications and controls them by operating the usual monitoring and detection.

5.5 Role (where) and contribution (what) of the proposed technologies within the architecture

5.5.1 Software as a Service (SaaS) role and contribution

The role of *Software as a Service* (SaaS) technology contributes to the architecture in the **pre back-end** phase, i.e. **service purchasing** (monthly fee based), in the back-end of the system, as the entire application and architecture operate web-based (Cloud computing), and in the front-end phase, through the communication layer, from the cloud directly to the fleet's on-board devices connected to the Web platform (Service delivery to drivers).

Overall SaaS contribution is achieved in a synthesis of a service that incorporates and involves logistics data, real-time traffic data, routing algorithms, decision-making process and driver's execution, by offering a continuous service of static and dynamic routing.

Specifically, company requirements', decided by contract and modifiable during the use of the service, and that lie in the **SDB** translate into an identical, but dedicated and configured instances (through metadata) in architectural terms ("True SaaS"), and in the formulation of VRP with different target, weight, criteria constrains, inputs, policies and rules (of the function that the routing algorithm has to optimize) in algorithmic terms. In this sense, SaaS acts in **ARS** and GIS because they draw from the above-mentioned specifications to create the input sets to be fed to the routing solver.

In operational terms this is useful also to the **TDF** for the calculation of travel times, as forecasting methods are applied on a rich data archive shared by the subscribed tenants from which the system can derive reliable traffic patterns and trends, consequently, effective solutions performed by the ARS and available to the fleets.

The cloud computing implies that both of these modules solve the algorithms directly in the cloud, in fact even dynamic modifications or requests, whether occur, enter directly into the cloud.

The **GIS** can also take advantage of the shared context to characterize and georeference its map with information of all the companies subscribed by creating a very rich context of relevant information available to the decision maker, which translates in an even more effective decision-making process. The figure 5.9 provide some highlights of what explained.

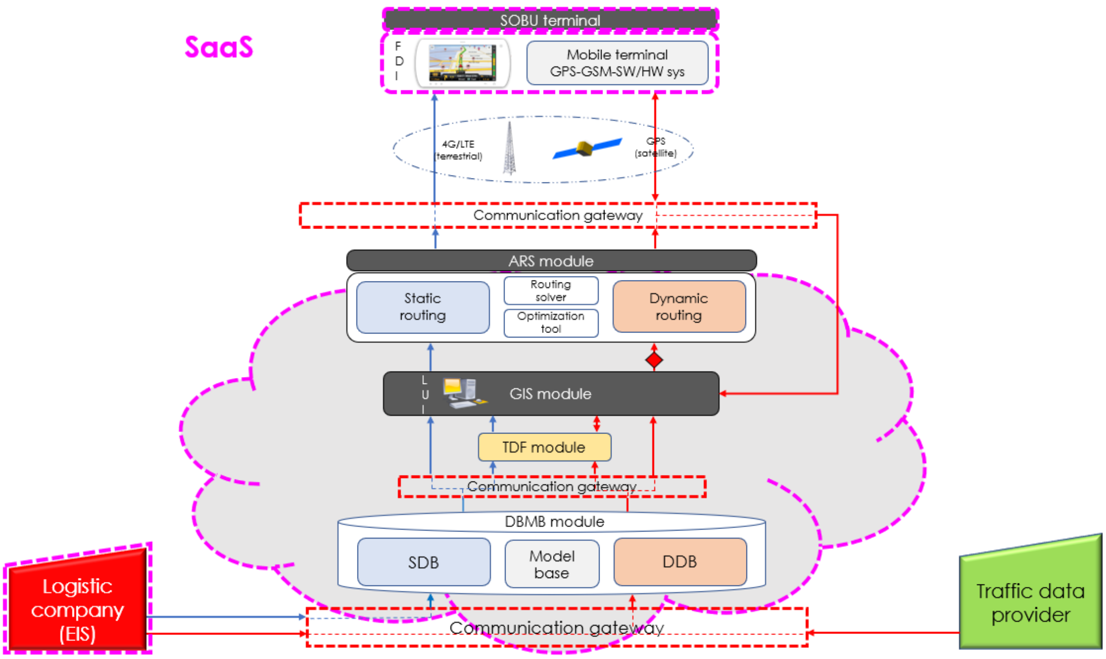

Figure 5.9 – SaaS's role and contribution in the system's architecture

5.5.2 Digital Twins role and contribution

Digital twins (DT) technology allows the system to meet and **cope with IoT** (therefore IoV, whose contribution is further explained) and Industry 4.0 requirements, since the need of dynamism, timeliness, real time data, system synchronization and intelligent management are nowadays essential in the modern routing context, and are to be addressed in order to develop an innovative routing solution.

Digital Twins enable the architecture to exploit the new opportunities offered in the field of ICT that today allows to integrate, in a wireless manner, network, detectors, electronics, hardware and software, physically dispersed, creating a web-based data environment to achieve an improved cost-efficiency, performance and system's management.

This technology inevitably involves the front-end of the system, as it exploits the devices (sensors of which the vehicle is equipped and smartphone) present in the **SOBU** as sensing units, and detectors in the network (physical layer), where the physical assets (counterparts) correspond to operating fleet vehicles.

In the back-end of the system, in the cloud, is recreated simultaneously the virtual twin of the vehicle that retrieve the information and status of its physical counterpart, thanks real time synchronization twofold communication enabled by WATs (communication layer), therefore bringing dynamic feedback to the system, ultimately creating smart entities in the Internet of things and services.

Operatively speaking the virtual counterpart provide the **GIS** module the input data to mirror/trace, predict activities, try simulations of the evolution of the system (allowing planning or forecasting purposes, therefore exercise also of TDF module), to finally estimates of the expected performance in a given dynamic urban scenario. In this sense could be a useful to assist the GIS module with Microsimulator appliances (preferable to macrosimulators, due to their fine-graned simulation suitable for volatile contexts), which may absorb some of its work, operating some activities involving virtual twin. The consequence of all above translates into a substantial and improved support for the logistic user.

In this sense, the DT does not replace the work of the existing twofold communication via GPS and GSM's tracking and detection, as the DT characterization consists in a more holistic knowledge, but rather constitute a further support in terms of predictive foresight of the system, therefore more flexibility and robustness.

Summarizing, Digital Twins main purposes in the system are listed below.

- **Current health status**: for control purposes, maintenance management, physical state of physical counterpart, model adherence to the real system, overall system's maintenance
- **Physical twin's digitalization**: to study and forecast system's evolution, hence performance, ensure seamless data gathering, provide a synchronized digitalization of the asset, IoT entities management; decision making support: thanks to multi-disciplines synergy and big data analysis; estimates and performance optimization purposes thanks to historical and real time data.

In a nutshell, DT prove to be a powerful and pivotal tool (for an innovative and competitive application) to keep control and represent (synchronize) complex systems affected by a multitude of dynamic factors and relationships, but, above all, they are particularly suitable for cloud-based systems, given the digitalization of physical data in the virtual layer (intelligent "things"). Finally, they can monitor and optimize the performance of the entire system according to the evolution of the environment, either in real-time, thanks to the data detected on the field, or statically, thanks to the prior knowledge of future scenarios, derived from estimate. The figure 5.10 provide some highlights of what explained.

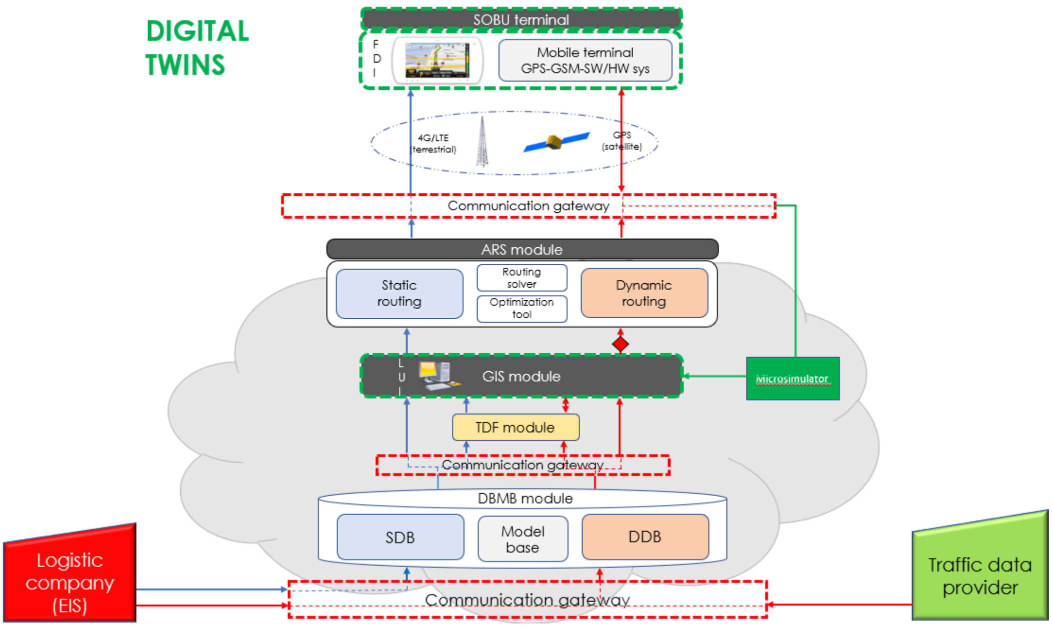

Figure 5.10 - Digital Twins' role and contribution in the system's architecture

5.5.3 Blockchains role and contribution

Blockchains technology allows the system to meet the increasingly digital and web-based nature of the modern context, which hides threads and complexities together with opportunities, understood either in terms of cloud or in terms of IoT. In essence the goal is to eliminate the presence of a third party in transactions, making the process controlled directly by the actors without intermediaries, with benefits in terms of efficiency, security and costs. We want to create a service delivery routing context that operates in a block chain platform, namely a public shared ledger collects all the information about all the transactions (data in our case) operated and the control role of the traditional third party is replaced by a Network consensus, thus public network of users. Basically, a valid recognized transaction is encoded and contained in a block, which in turn is connected to a preceding block of other previous and verified transactions, forming a block chain called block chain precisely. By doing so, the integrity is verified by "block-by-block" and stored along the chain. Platform users, are directly responsible for storage, mining, and transactions verification and confirmation because only new blocks verified by them are to be hooked to the chain, who are therefore guarantors of database integrity.

The idea behind the integration of this technology of the high-level system's architecture is to apply the same concept applied to the transaction (described above) to the data circulating in the system, forming a block chain of verified data where encoded data are contained into the blocks instead of transactions.

Hence, the Blockchain technology will be applied to the **CG** since, as explained in section 5.2, it is mandatory passage for any information circulating within the architecture, acting like a sorter (routing it towards the specific module) and a filter for every data communication. Moreover, since the CG operates transversally across the whole system, this concept will be extended and will flow in the entire architecture. Although the CG already has the function of ensuring access and data validation, this innovative concept can further **enhance system's reliability**.

The Blockchain technology is suitable, however, as it stands, in the system's architecture in the pre back-end phase, or in service purchasing. As a cloud-based software implies the paying of monthly fees digitally, the advantages of seeking a more convenient and reliable solution to structure payments and subscriptions are clear.

Hence, digital contracts, such as Smart contracts, are a great opportunity that this technology offers to software companies offering cloud-based services. In this sense the proposed architecture aims to host a Blockchain-oriented software, that is software that uses the implementation of Blockchain in its components.

The public and shared nature allows cross-checking, system self-control, transparency and accessibility of information, but at the same time the respect of privacy is ensured thanks to anonymity, which is an essential peculiarity in IoT applications and frameworks which the proposed architecture aims to be. This should provide:

- **System cyber-security and privacy**: thanks to consensus-based architecture, the encrypted codes and authentication; data integrity, traceability, sovereignty
- **Transaction immutability**: due to the verified block structure of the chain, the traceability and verified validity of all the operations in the ledger.

Finally, the Blockchain technology proves to be a fundamental technology given the multi-tenancy and the mole, relevance, rapidity and heterogeneity of data (Big data) circulating in the system and therefore have to be protected. The figure 5.11 provide some highlights of what explained.

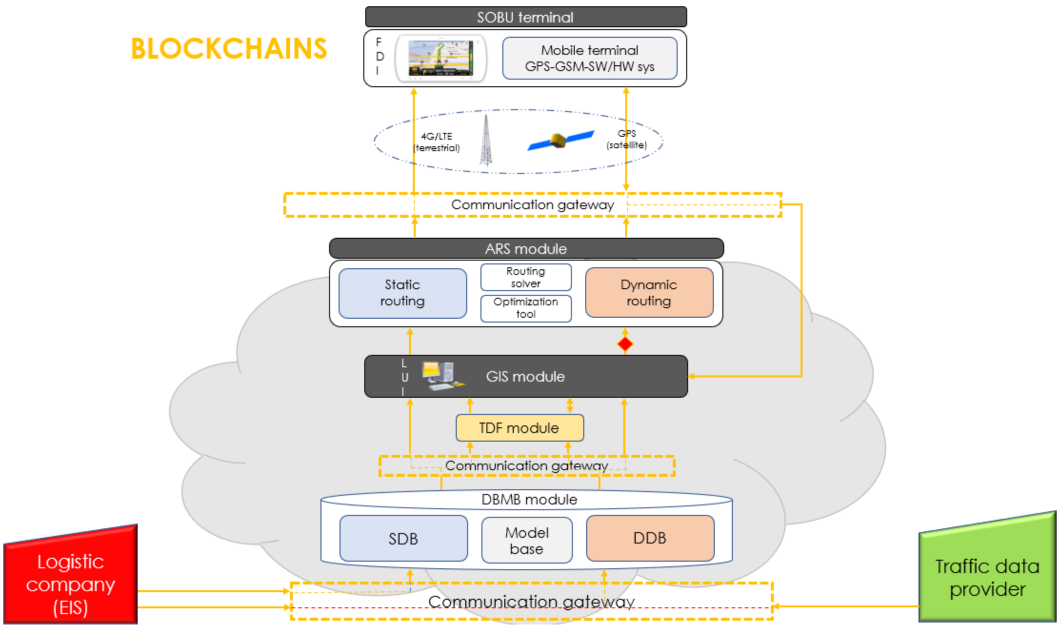

Figure 5.11 – Blockchains' role and contribution in the system's architecture

5.5.4 Internet of Vehicles (IoV) role and contribution

Another paradigm that architecture wants to exploit is the *Internet of Vehicles* (IoV), which is nothing more than the declination of the **Internet of Things in the world of transport**, where for things we mean the proliferation cars, buses, trains, mobile phones, TV, home Appliances, audio systems, home automation etc. with the common denominator of being internet connected entities that allows to automatically provide information necessary to take advantage of a service in the Intelligent IoV framework. The IoV follows the trend of extending the concept of intelligence to applications for road users, understood in terms of interaction between networks of technologies, through the connection of WAT (wireless technologies applications) and ICT, to bring benefits in a holistic way to a multitude of users.

It consists of the three actors who inhabit the network, namely people (including drivers, passengers, pedestrians, public transports users), vehicles and things (stood as everything that surround two previous components and are source of relevant information scattered in the network). All of them interact with each other and act both as data provider (for the IoV cloud) and as data users (end users), as well as all forming a network within and between themselves, then creating a cloud based integrated heterogeneous multinetwork, which is the IoV. It is through this broad cloud that the dynamic information circulating across networks is absorbed and worked through cloud computing (artificial intelligence applied to real time big traffic data) to become services, thus smart applications and useful data to the end users (clients).

The idea is that, by circulating, fleet's vehicle establishes vehicle-to-vehicle (V2V), vehicle-to-Infrastructure (V2I) of mobile networks, vehicle-to-Personal devices (V2P), vehicle-to-Sensors (V2S) connection. All this lets you see that the application of this in the system is at the **front-end** level, and in particular involving the **SOBU** and the **FDI**. As a result, logistic companies' fleets will be able to benefit specifically from safety applications, to improve the driving experience and enhance the trip safety, and business application, to exploit IoV's value-added services in terms of quality of travel experience (since both are becoming a standards).

Safety application that will be applied are listed below.

- **Collision avoidance**, which it includes collision warning, accidents, queues and road conditions alerts with sound and on-display notifications thanks to communication with other vehicles (e.g. this system can share its functionality, thanks to IoV, warning other networks cars of an accident occurred further on)
- **Driver assistance**, in case of emergency detected by the sensors of which the vehicle is equipped; semi-automatic guide and emergency call;
- **Navigation**: through GPS systems integrated into the vehicle, the driver can obtain all the data that vehicles, human or devices share on IoV's cloud (e.g. live traffic information);
- **Remote diagnostic**: as vehicles are nowadays equipped with several sensors, enables real time communication to centres concerning the functioning of car's motion organs (therefore remote health status control), also self-repairs (step-by-step) tool delivered through the vehicle display. This benefits in terms of maintenance costs, breakdown of service, therefore, consequently, reliability of customer service.

Business applications that will be applied are listed below.

- **Insurance on driver statistics**: allows ad-hoc insurances according to the performance of the driver;
- **Cloud service**: means the opportunity of establishing (temporary) connection while driving to cloud multinetwork;
- **Connected driving**: means distraction-free applications, thanks to smart driving services such as, indirect use of smartphones (messaging, maps, voices support) while driving, i.e. apple CarPlay or Google Android, with user-friendly interfaces through fleet screen (Driver interface). Hence, smartphone, cloud system's connection and point of interest information.

The key idea is to equip fleet vehicles with **infotainment systems** (smartphone-navigator integration), with which the driver can interact, which we believe to provide the best interface solution, since they can connect with IoV cloud framework to make available smart business and safety application, especially navigation and connected driving.

The customer sequence sent from the cloud to the vehicle's infotainment system is displayed on high resolution LCD touchscreen in a 3D digital map environment (for a better orientation especially in cities). Basically, we want to leverage the IoV framework via infotainment systems to bring complete routing assistance to drivers while driving through SOBU's interface. Thus, providing especially traffic alarms, external cameras view, routing navigation indications, voice and messaging controls, possibility of sending/receiving tasks, delivery status, messages or notifications to/from back-end, showing vehicle diagnostics and Other information overview, smartphone connection, Wi-Fi router (Internet in the car) to IoV cloud so access business and safety applications.

In addition, this carries foresight to the System since, given the trend of tremendous expansion IoV's basin (every car will be connected, traceable in an "on-road Internet"), equipping intelligent vehicle systems will become an absolute standard. The system will cope and be ready for this scenario and will cope with these certain necessities.

We may concludes saying that the system's architecture aims to introduces intelligent in-vehicle systems to establish a symbiotic relation between driver and vehicle by exploiting connected driving. The figure 5.12 provide some highlights of what explained.

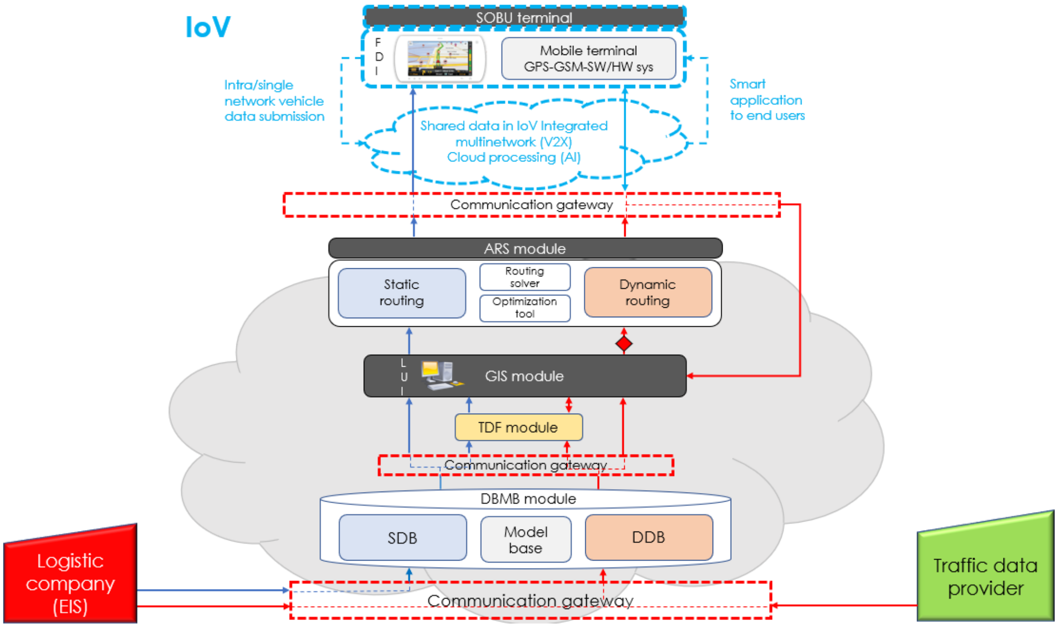

Figure 5.12 – IoV' role and contribution in the system's architecture

5.5.5 Decision Support System (DSS) role and contribution

The literature proves *Decision Support Systems* (DSS) systems to be a compelling solution to assist decision making in high complexity contexts that require and seek for an **“intelligent” intervention** in order to be innovative and performing. Moreover, they demonstrate to be even more useful whether coupled with GIS systems and traffic data processing modules, especially for dynamic transportation related contexts.

In the specific one has chosen to resort to a DSS as it is a particularly suitable tool from the operational and infrastructural point of view, as it consists of a knowledge systems (KS), a problem processing system (PPS) and a user-system interface (USI) structure which perfectly fits the system’s architecture needs.

The KS is covered in architecture by **DB-MB** module, which also incorporate, the MB and DB management systems role, namely to organize, acquire, modified, update models, access knowledge. It is worth mentioning that DSS also structure and enable to manage the knowledge, through data warehouse (organizes data, in order to make them suitable to be worked), data mart (subsets data), data mining (to analyse data to extract valuable knowledge) and queries (to interrogate the data base and verify hypotheses formulated by the decision maker).

The PPS is represented by the combination of **GIS**, **TDF** and, more indirectly, the **ARS** (algorithmic engine), in that (in particular the GIS) they manipulate the information stored and shared by the tenants in the cloud DB (KS), allowing analysis to generates useful output and answers to the decision-making process.

The DSS introduces into the system, through the PPS’s inferential engine, Artificial intelligence functions to provide a clear, consistent and logical explanation of the relationship that binds input to output. Finally, the USI is clearly represented by the **LUI** in the GIS environment, as transform the commands expressed in natural language by the user in understandable terms for PPS and KS, and, conversely, allows to present the output in a clear and unambiguous manner.

All these necessary tasks reach their most extensive realization within the architecture, and this explains why DSS constitutes almost the whole back-end of the system.

In conclusion DSS provides the user a series of information to fully understand the problem; models and data analysis that give the possibility to inference data and compare solutions or scenarios; reports or forecasts, to allow simulations. In general, it offers some basis to suggest the most appropriate action enhancing the efficiency and effectiveness of the decision-making process, in particular concerning optimization problems to choose between alternative solutions.

Finally, it also provides allow long term strategic actions (strategic planning, e.g. setting objectives) or short period tactical actions (operational). The latter include management control (ensuring effectiveness in the acquisition and use of resources), operational control (ensure effectiveness in the performance of operations, i.e. monitoring and detection), operational process (decisions taken during operations, i.e. rerouting). The figure 5.13 provide some highlights of what explained.

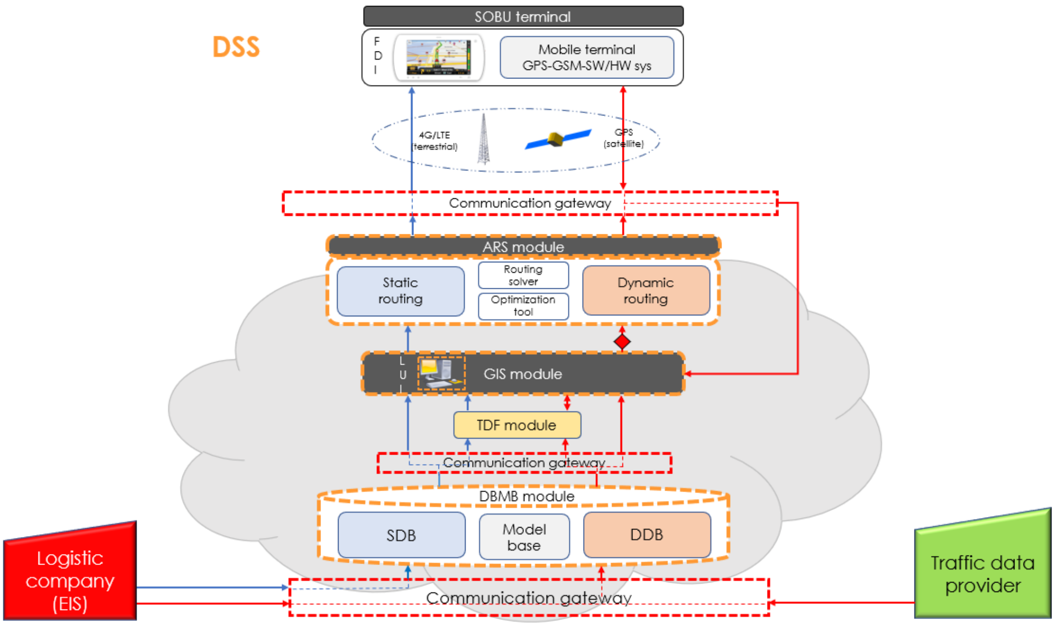

Figure 5.13 – DSS's role and contribution in the system's architecture

5.5.6 GIS role and contribution

The role and contribution of Geographical Information Systems (GIS) in the system are evidently represented by the **GIS** module in architecture, whose function is extensively expressed in the description of the system back end in Section 5.2. Therefore, to avoid repetitions is reported below only some highlights concerning the work of the module.

The choice to resort to GIS system that can work in close contact with the **DB and MB, TDF** and with the **ARS** modules as mentioned is due to the fact that the framing of its functions in a DSS context can enhance systems performance. First of all, it offers a cartographical environment interface which suits perfectly with the matrix of the problem and the need for decision support in an urban scenario. Moreover, it **handles the geographic data pursuing geospatial tasks** whose output is required by the ARS and it can geo-reference interested entities imported from DB. It faces the static facet of the problem, in fact, thanks to travel times provided by TDF, the GIS module can create a time dependent shortest path between clients or nodes then sent to ARS module to be optimized by the routing solver. Later near optimal solutions are evaluated by the route planner.

It is responsible of the **dynamic passive management** (feedback from the network, without elaborating changes) operating two of the key functions of the system: fleet service monitoring, which verifies and report overviews about the service's status of competition and the initial delivery plan, keeping track of each route thanks to the GPS tracking mechanism made visual on the map through the interface; and fleet detection, stood as dynamic continuous calculation of deviations from the initial (initial/static route plan) plane (or forecasts) compared to network's evolution, involving also integration of current traffic data from the network. It also performs **active dynamic management** thus triggered by an event: fleet dynamism, network dynamism and logistic dynamism. The first two are detected either through retroactive ("internal") or proactive ("external") feedback. The module manages this evolution of the state of the network by updating the cartographic environment and triggering the rerouting, initially a shortest path between remaining clients, new clients and vehicles, again to be fed to the routing solver. All the graphs forwarded to ARS have been translated from referenced urban map to weighted directed asymmetric graph (arches representation). The figure 5.14 provide some highlights of what explained.

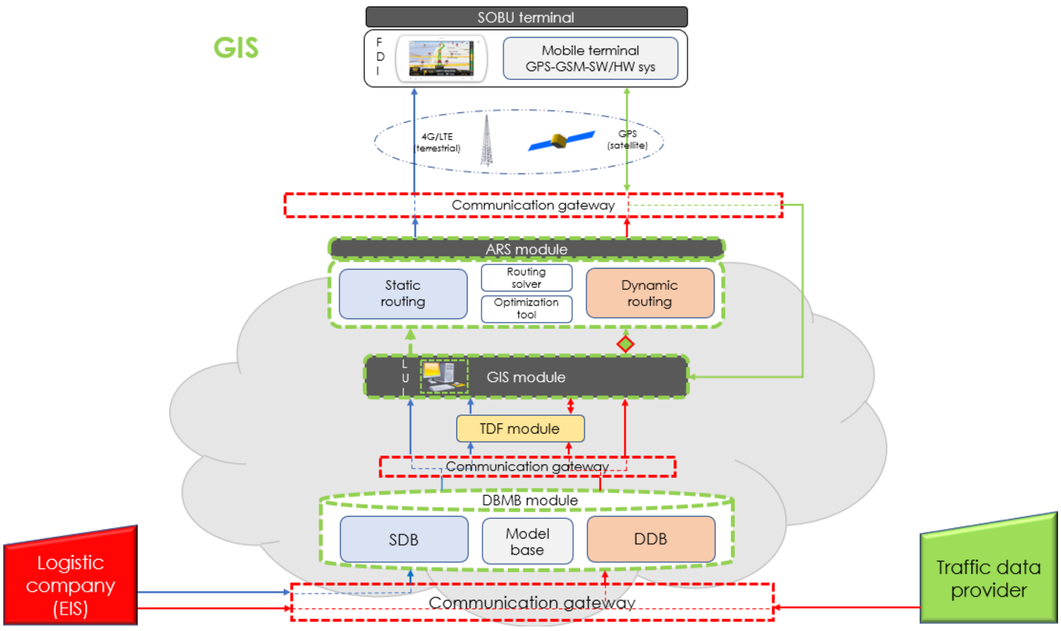

Figure 5.14 – GIS's role and contribution in the system's architecture

Chapter 6: Conclusions

6.1 General conclusions and statement

The proposed high-level (module based) system's architecture proves to be highly benefitting and suitable for the dynamism and unpredictability that nowadays characterize the routing context, both in terms of orders (e.g. dynamic orders, higher demand, low batch shipments, variable productions, shorter lead times etc.), fleet and in terms of road network (e.g. traffic congestion, travel times), that imply frequent rerouting of fleet in a timely fashion, in order to succeed in serving customers within increasingly stringent constrains (asking for reduced lead times for dispatching).

Compared to other software solution already available in the market, it aims to face and manage more complexity in less time, addressing the need to control environment's volatility, in order to plan, route and schedule and, anticipating and reacting appropriately, without compromising and even enhancing customer service, which is an increasingly key determinant for the competitive advantage in the actual distribution industry.

The value added for logistics companies will be a software service developed ad-hoc for urban freight transportation contexts, which provide holistic, responsive and effective time dependent routing and scheduling solutions that will benefit both in terms of cost, saving travelling kilometres, and in terms of resource allocation and utilization. These solutions are ready to implement, and are visualized in an user-friendly interface of a DSS, in such a way that this way one single user logistic user is enabled to perform route planner, scheduler and the dispatcher's job in a holistic manner, though one single interface, offering decision analytic support and assistance to take the right decision in the shortest time possible.

This become a real Vehicle and Scheduling Routing system (VRSS), which integrates dynamic and static VRSPs elaboration, real-time traffic data integration and forecast elaboration support, becoming a Vehicle routing system. This already solid framework is enhanced by the adoption of compelling and modern technologies (Blockchains, Digital Twins, IoV) that render the system not only, even more original, innovative, sustainable, integrated, holistic and forward-looking, but also opening the perspectives to future scenarios and forthcoming requirements of urban logistics. All of this is offered as a web-based flexible SaaS therefore entailing cost-efficiency (so attractive also for small logistic businesses), configurability, scalability and cloud computing, offering short-term as well as long term benefits and perfectly integrating with EIS. Even fleet drivers will benefit since their work will become less stressful, more satisfying and easier, with less mistakes, idle and overtimes as the travel experience will be enriched, supported and assisted by infotainment systems connected with the cloud. Finally, considerable savings are achievable thanks to lower transportation cost (minimization of total travelled distance), improved resources utilization, reduced number of owned and hired vehicles and personnel reduction.

6.2 Expected benefits

Despite the benefits offered by the system are not limited to those indicated, the most important and obvious ones are listed below.

❖ **Customer and premium service**: by its nature, SaaS is a customer-tailored on-demand service, either in terms of configuration, depending on its requirements, and in terms of scalability, with economical and application performance adjustments of the monthly fee. Moreover, the subscription fee model, allows the freedom of contract's interruption since based on customer satisfaction ("Low switching costs"), this pushes the software vendor not only to offer a top service in terms of Software, but also ancillary services (performance improvements, maintenance, updates, testing, security, training, storage etc.) Still, start-up times are immediate, the application is ready to use and the results are immediately verifiable and tangible. Finally, the TDF produces reliable forecasts, to limit the risk of traffic congestion, which means delays or missing orders, therefore enhancing the service. Also, the delivery time is expected to drop.

❖ **Know how concentration**: the logistic company outsources and centralizes the software-related front and back office activities to the software company. This allows a total focus on value-added assets for the business enterprise, disengaging from tasks extraneous to corporate know-how.

- **Cloud computing suitability (exploitation):** as the proposed system aims to be able to manage a connectivity with Enterprise Information Systems, cloud computing creates enormous opportunities for the system and facilitate software adoption for companies. The cloud environment not only ensure the input of data (from TDP, LC) straight into the platform, but also allows you to adopt on-board devices solutions without any particular computational requirements since the computational burden is absorbed by the cloud processing modules. Moreover, the information pooling obtainable ensure rich knowledge at system's disposal. Finally, web cloud also proves to be fertile ground for advanced and futuristic applications.

- **Cost efficiency**: cost efficiency that emerges from system's architecture derives from three fronts.
  1) *Software expenses*: direct, indirect and various software related costs since all these activities are now supported by the SaaS provider to which an all-inclusive fee is payed (basically a charge, whose amount depends on the service features requested). Finally, low switching costs guarantees freedom of choice of alternative software providers, since in case of dissatisfaction the sunk costs are reasonable.
  2) *Market*: a market with lower barriers (for logistics companies), therefore more extensive and competitive, imply competitive SW prices, given the high demand. At the same time, market's profitability, lead to the proliferation of new software providers, involving both competitiveness, contributing to further decrease the price (price differentiation), or to increasingly specialized companies (product differentiation), thus creating a virtuous circle.
  3) *Algorithms efficiency*: as effective routing solution that evidence to be effective in highly dynamic urban environment imply less total travelled miles, (therefore lower transportation cost) and effective scheduling solutions, imply higher resources utilization (utilization of fixed cost), less owned and hired vehicle and better allocation. Finally, the adoption of the software expects to reduce the need of personnel dedicated to routing and scheduling related tasks.

- **Enhanced decision-making process**: although the system delivers the user one or more near optimal solutions, which can be accepted or not, the system provides the user also with a series of tools that enable, primarily to fully understand the problem, secondly to edit and explore the solution in a cartographic GIS interface, which ensure to have visual impact of the outputs on a digitalized urban network. Hence, the architecture, thanks to DSS Framework, supports user offering decision analytic support and assistance in complex environment affected by several factors and actors, and that thereby requires an "intelligent intervention". Moreover, through Digital Twins, the digitalization of the virtual counterpart enables additional data to exploit. In this way the support becomes even more solid thanks to multi-disciplines synergy. Also, the goal of the decision making may vary: minimization of initial plan's deviations, or compliances, risk etc.

- **Complex system control suitability (enhanced monitoring)**: to keep the complex urban system under control is a priority of the system, in fact the architectures, is made in such a way to retrieve information from any actor either involved or which affect the system, through robust feedbacks, in particular through twofold real-time Communication between back end and drivers enabled by WAT (communication layer). Surely the central role is played by the GIS that, through Fleet's retroactive feedback, assures both real-time monitoring (data acquisition of position, speed, delay, distance travelled, via GPS, and vehicle status instantaneous residual capacity, served customers, via terrestrial GSM network, and in general constrains violations), and detection (deviations from schedule or time windows unrespect due to network's evolution), which controls two of the three triggers of the rerouting (network and fleet dynamism). These tasks integrate current network data provided by TDP and TDF that help strengthen network knowledge to be controlled. Hence the system is kept monitored in its entirety thanks to the synergy of different expedients. For what concern innovative technologies, digital twins allow further system's (fleet) control thanks to digitalization and the IoV allows remote diagnostic (fleet health status) via connections with car's organs and sensors. Both bring benefits in terms of maintenance (costs and preventiveness), breakdown of service, therefore, consequently, reliability of customer service. Finally, the architecture is made in such a way as to ensure a control of the triggers: Network, fleet and logistic dynamism.

- **Static and Dynamic routing optimization**: represents the core competence of the proposed system, and is enclosed in the ARS module, as it aims to cope with Static and dynamic routing problems (specifically TDVRPTW), fed with complex inputs, stringent constrains and in a timely fashion, as the routing Solver runs a heuristic algorithm. The specific VRP algorithm will be optimized taking into account:
  a) customers still to be served (dynamic),
  b) weights of the customers (due to premium services, amount, product cost),
  c) soft and hard time windows (of new and planned nodes to be served),
  d) fleet constrains (drivers shifts, availability, load),
  e) product constrains,
  f) route precedencies due to load sequences (especially when accepting a new order),
  g) departure time.
  h) loading unloading times;
  i) forecasts

  Taking into account what listed above, the system will emit:
  i. an initial plan (static routing),
  ii. process a reassignment, (dynamic rerouting),
  iii. resequencing (dynamic rerouting),
  iv. generate a new route (dynamic rerouting),

  by considering the position of customers, vehicles and schedule's progress. Moreover, thanks to the optimization tool, the actual best solution is retained until significant improvements to the solution are obtained (minimum threshold to limit nervousness). Finally, optimized solutions provided owe their effectiveness to the system's predictive effort (explained below) through static and dynamic forecasts. It has to be reminded that the final decision of optimized routing plans is still up to logistic user.

- **Predictive foresight**: another key system's feature is the orientation to anticipate future events and needs. The foresight in static terms is given by time dependent travel time estimates elaborated by the TDF. Dynamic foresight is achieved through data submitted by the TDP and GIS through retroactive feedback (network's data detection), providing dynamic service times therefore forecasting for possible deviations from the schedule. Another support is given by digital twins that enable to access prior knowledge of future scenarios thanks to digitalization. As stated, provisional effort is not redundant but rather in parallel, to make the forecasts more reliable.

- **Future oriented and forward looking**: the idea is to exploit advanced technology not only that may support and enhance already solid system's functions, but also can anticipate future requirements that will become soon absolute standards in order to be competitive in future transportation contexts (e.g. smart cities). Some contributions are: SaaS model, adoption of blockchains to enhance the security, especially in web framework, Infotainment systems, for smart assistance, Digital twins to enable digital simulation, and IoV to exploit IoT paradigm, nowadays mature for real life applications, providing smart applications for drivers.

- **Drivers work conditions:** the benefits extend to drivers as well. In operational terms, efficiency in routing and scheduling plans' optimization, guarantee less mistakes, idle and overtimes and traffic congestions, and especially ongoing dynamic order management is expected to be less frustrating and more controlled. Furthermore, the travel experience become enriched as infotainment systems provide user-friendly interface and connected assistance (to back-end) through business (e.g. navigation to better guide through the network) and safety applications (e.g. collision avoidance, remote diagnostic). Eventually, the driver will have the possibility within certain parameters and limits, to manipulate the assigned plan to follow a preferred one (simply through the touchscreen). This are expected to result in a less stressful, more satisfying and easier routing performance thereby enhancing customer service. Indirectly, drivers' job conception will change as IT predisposition is required, therefore becoming attractive for new technologically skilled and qualified staff.

- ❖ **System reliability and data security**: the Cloud-SaaS context requires special attention in terms of cyber-security, especially in multi-tenant contexts, with relevant and massive data exchange. The decision to adopt blockchains deals with this necessity that are expecting to become increasingly important, as they boost privacy, security and integrity, namely data transparency, traceability, accessibility, sovereignty, integrity, system self-control, transaction immutability (safer and more convenient "smart" contracts). The software provider must strive to ensure these requirements, otherwise subscription's cancellation, reputation loss or company shutdown may occur. On the other hand, also at the infrastructural level new terrestrial communication protocols are expected to cope with this issue.

6.3 Issues and future developments

Before listing possible issues, it may be useful to ask what a company should evaluate before subscribing to the service:

- Does the urban logistics environment in which the company operates generate a complex problem requiring support?
- Is network's volatility such as to make elaborated routing plans (in progress) unsuitable or inadequate?
- Is there a need for more responsiveness in serving customers (competitiveness), therefore in generating routing solutions?
- Does the company want to invest in the long term with a system that will help your business in future prospects?

The main issues that the system is expected to face are presented, but not limited, by the list below:

- **Unknown demand or demand modification**: to evaluate stochasticity issues in the system was not the purpose of the treatment, however the importance of better capturing uncertainty is remarkable given the clear consequences on the vehicle's routing. In the literature VRP with stochastic demand is well known and has been deeply analysed, thus is integrable in the proposed architecture.

- **Stable and high-performing connection**: the entire architecture functioning depends on internet connection as does not allow to work offline. Moreover, weak connection can compromise important service features affecting the whole system. In this sense infrastructures (and WAT) play a key role, but this depends on external factors (provider, governments etc.)

- **Lower customization**: in the traditional model (software hosting) a total personalization was possible, here we talk about configuration (through) rather than customization, given the relative standardization that the multi-tenant structure requires.

- **Multi-tenancy:** it is not trivial to efficiently construct and manage an architecture in which a multitude of clients (tenants) can access shared resources, therefore computational capacity (for peak usage), instances configuration (to ensure personalization) and storage issues are to be addressed and evaluated.

- **Security and data sovereignty:** despite the Blockchains countermeasure, it is necessary to be aware of the risks of data ownership's outsourcing, especially in web platforms and where the amount and data sensitivity are crucial. Hence, is vital for software companies to ensure the security, privacy and integrity of customer (tenants) data. This means that any security solution or improvement must be adopted as challenges in this field are still all open.

- **Batch data recording through RFID:** this functionality can facilitate and speed driver's task related to the load (recording, check, load/unload) as thanks to RFID technology the freight is recorded automatically and traceable in the system. This can reduce bureaucracy as well as human mistakes, increasing accuracy hence customer service.

- **Complexity reduction:** the future momentum for VRP software application, will probably be linked to the reduction of complexity in order to boost system's reactivity, therefore competitiveness. A possible solution can be the hierarchic approach of Multi Agent Systems (MAS) whose function is to decompose the macro problem in subproblems independently solvable by agents interacting with each other to find a global optimum, allowing less onerous and paralleled resolutions, thus exponentially more efficient.

- **Drivers' resistance to implementation**: although the system brings clear benefits in terms of travel experience and performance, the abandon of work habits must be thought as gradual. It should therefore be offered initial support and a period of adaptation should be expected, in order to get familiar with all the system's ancillary technologies.

- **Green VRP:** even though less traffic congestion and less travelled kilometres consequently lead to less CO2 emissions (which is a core objective of the complete project *"SMARTRANS"*), real expedients that bind the algorithms to consider the environmental aspect, which is nowadays extremely crucial, have not been proposed. In this sense a branch of problems that specifically address these issues, namely the Green or Pollution VRP, have been deeply studied in the last years in literature and are integrable in the system's architecture.

Glossary

- VRSP: vehicle routing and scheduling problem
- VRP: vehicle routing problem
- VSP: vehicle scheduling problem
- TSP: travelling salesman problem
- TDVRP: time dependent vehicle routing problem
- VRPTW: vehicle routing problem with time windows
- DVRP: dynamic vehicle routing problem
- OR: operational research
- ILP: integer linear programming
- GPS: global positioning system
- GIS: geographic information systems
- ITS: intelligent transportation systems
- ICT: information and communications technology
- WAT: wireless access technologies
- FCD: floating car data
- GSM: Global System for Mobile Communications
- LTE: Long Term Evolution
- OD: origin destination
- RFID: Radio-Frequency Identification
- SAAS: software as a service
- DT: digital twins
- IOT: internet of things
- IOV: internet of vehicles
- DSS: decision support system
- EIS: enterprise information systems
- ERP: enterprise resource planning
- AFMS: advanced fleet management system
- VRS: vehicle routing system
- USI: user system interface
- PPM: problem processing module
- LC: logistic company
- SP: software provider
- TDP: traffic provider
- DBMB: data and model base module
- DDB: dynamic data base
- SDB: static data base
- LUI: logistic user interface
- CG: communication gateway
- TDF: time dependent forecasting module
- ARS: Advanced routing and scheduling module
- SOBU: smart on-board unit
- FDI: fleet's driver interface

Bibliography

[1] C. D. Purnamasari and A. Santoso, "Vehicle Routing Problem (VRP) for courier service: A review," *MATEC Web Conf.*, vol. 204, p. 07007, Sep. 2018.

[2] B. Eksioglu, A. V. Vural, and A. Reisman, "The vehicle routing problem: A taxonomic review," *Comput. Ind. Eng.*, vol. 57, no. 4, pp. 1472–1483, Nov. 2009.

[3] D. Vigo, "Il Vehicle Routing Problem: Modelli ed Algoritmi," *Operations Research Group Bologna*. [Online]. Available: http://www.or.deis.unibo.it/didatt_pages/mmsd_ce/L_VRP97.pdf. [Accessed: 15-Nov-2018].

[4] S. Raff, "Routing and scheduling of vehicles and crews: The state of the art," *Comput. Oper. Res.*, vol. 10, no. 2, pp. 63–211, Jan. 1983.

[5] L. Abbatecola, M. P. Fanti, and W. Ukovich, "A review of new approaches for Dynamic Vehicle Routing Problem," in *IEEE International Conference on Automation Science and Engineering*, 2016.

[6] S. Gayialis, Sotiris; Konstantakopoulos, Grigorios; Papadopoulos, Georgios; Kechagias, Evripidis; Ponis, "Developing an Advanced Cloud-Based Vehicle Routing and Scheduling System for Urban Freight Transportation," *IFIP Int. Conf. Adv. Prod. Manag. Syst.*, vol. 536, pp. 190–197, 2018.

[7] V. Pillac, M. Gendreau, C. Guéret, and A. L. Medaglia, "A review of dynamic vehicle routing problems," *Eur. J. Oper. Res.*, vol. 225, no. 1, pp. 1–11, Feb. 2013.

[8] A. ; Larsen and O. B. G. Madsen, "General rights The dynamic vehicle routing problem," 2018.

[9] "CITY LOGISTICS: come procedere?," 2015. [Online]. Available: http://www.logisticamente.it/Articoli/5894/CITY_LOGISTICS_come_procedere/. [Accessed: 11-Dec-2018].

[10] C. Malatesta, Marina; Filippi, "Metodi Esatti di Risoluzione per il Problema di Vehicle Routing con Finestre Temporali .," 2004.

[11] "What is Network Traffic? - Definition from Techopedia." [Online]. Available: https://www.techopedia.com/definition/29917/network-traffic. [Accessed: 29-Dec-2018].

[12] B. April, "New traffic data sources – An overview 1," no. April, 2014.

[13] A. I. Torre-Bastida, J. Del Ser, I. Laña, M. Ilardia, M. N. Bilbao, and S. Campos-Cordobés, “Big Data for transportation and mobility: recent advances, trends and challenges,” *IET Intell. Transp. Syst.*, 2018.

[14] Mark Beyer; Douglas Laney, “The Importance of ‘Big Data’: A Definition.” [Online]. Available: https://www.gartner.com/doc/2057415/importance-big-data-definition. [Accessed: 21-Jan-2019].

[15] A. Nantes, D. Ngoduy, A. Bhaskar, M. Miska, and E. Chung, “Real-time traffic state estimation in urban corridors from heterogeneous data,” *Transp. Res. Part C*, vol. 66, pp. 99–118, 2016.

[16] C. Antoniou and R. Balakrishna, “A Synthesis of emerging data collection technologies and their impact on traffic management applications,” pp. 139–148, 2011.

[17] G. Leduc, “Road Traffic Data : Collection Methods and Applications Road Traffic Data : Collection Methods and Applications,” no. May, 2015.

[18] A. Persona, “Lezioni di Logistica Industriale,” Padova, 2018.

[19] G. Kulkarni, J. Gambhir, and R. Palwe, “Cloud Computing-Software as Service,” *Int. J. Cloud Comput. Serv. Sci. (IJ-CLOSER*, vol. 1, no. 1, pp. 11–16, 2012.

[20] C. Thoma, “Fleet management on the basis of cloud software.”

[21] H. Chen, “Architecture strategies and data models of Software as a Service: A review,” *2016 3rd Int. Conf. Inf. Cybern. Comput. Soc. Syst. ICCSS 2016*, pp. 382–385, 2016.

[22] P. Version, “into a Service,” *Computer (Long. Beach. Calif).*, vol. 44, no. October, pp. 38–44, 2008.

[23] J. Ju, W. Ya, J. Fu, J. Wu, and Z. Lin, “Research on key technology in SaaS,” *Proc. - 2010 Int. Conf. Intell. Comput. Cogn. Informatics, ICICCI 2010*, pp. 384–387, 2010.

[24] J. Partyka and R. Hall, “On the Road to Connectivity,” 2010.

[25] G. Cosentino, “DSS Decision Support System,” 2008.

[26] Esri, “What is GIS?,” no. July, 2012.

[27] D. F. Venuda, “Il GIS (Geographic Information System) in biblioteca.”

[28] E. Negri, L. Fumagalli, and M. Macchi, “A Review of the Roles of Digital Twin in CPS-based Production Systems,” *Procedia Manuf.*, vol. 11, pp. 939–948, Jan. 2017.

[29] E. M. Kraft, “The Air Force Digital Thread/Digital Twin - Life Cycle Integration and Use of Computational and Experimental Knowledge,” in *54th AIAA Aerospace Sciences*

Meeting, 2016.

[30] R. Rosen, G. von Wichert, G. Lo, and K. D. Bettenhausen, "About The Importance of Autonomy and Digital Twins for the Future of Manufacturing," *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 567–572, Jan. 2015.

[31] M. Schluse and J. Rossmann, "From simulation to experimentable digital twins: Simulation-based development and operation of complex technical systems," in *2016 IEEE International Symposium on Systems Engineering (ISSE)*, 2016, pp. 1–6.

[32] J. Lee, E. Lapira, B. Bagheri, and H. Kao, "Recent advances and trends in predictive manufacturing systems in big data environment," *Manuf. Lett.*, vol. 1, no. 1, pp. 38–41, Oct. 2013.

[33] F. Casino, T. K. Dasaklis, and C. Patsakis, "A systematic literature review of blockchain-based applications: Current status, classification and open issues," *Telemat. Informatics*, vol. 36, pp. 55–81, Mar. 2019.

[34] J. Yli-Huumo, D. Ko, S. Choi, S. Park, and K. Smolander, "Where Is Current Research on Blockchain Technology?—A Systematic Review," *PLoS One*, vol. 11, no. 10, p. e0163477, Oct. 2016.

[35] S. Almeida, A. Albuquerque, and A. Silva, "An Approach to Develop Software that Uses Blockchain," Springer, Cham, 2019, pp. 346–355.

[36] S. Porru, A. Pinna, M. Marchesi, and R. Tonelli, "Blockchain-Oriented Software Engineering: Challenges and New Directions," in *2017 IEEE/ACM 39th International Conference on Software Engineering Companion (ICSE-C)*, 2017, pp. 169–171.

[37] F. Yang, S. Wang, J. Li, Z. Liu, and Q. Sun, "An overview of Internet of Vehicles," *China Commun.*, vol. 11, no. 10, pp. 1–15, 2014.

[38] O. Kaiwartya *et al.*, "Internet of Vehicles: Motivation, Layered Architecture, Network Model, Challenges, and Future Aspects," *IEEE Access*, vol. 4, pp. 5356–5373, 2016.