Trends in Models and Algorithms for Fleet Management

Maurizio Bielli\textsuperscript{a,*}, Alessandro Bielli\textsuperscript{b} and Riccardo Rossi\textsuperscript{c}

\textsuperscript{a}Institute of Systems Analysis and Informatics, National Research Council, Viale Manzoni 30, 00185 Rome, Italy
\textsuperscript{b} KIG Group International, Rome, Italy
\textsuperscript{c} Department of Structural and Transportation Engineering, University of Padova, Italy

Abstract

Fleet management represents a relevant activity at tactical and operational levels to be faced by private companies and public agencies devoted to passengers and freight transportation services. Mathematical models and computation techniques have been developed for optimizing and simulating the operation of transport fleets in order to serve the customers demand with the objective of cost efficiency. Many fleet management problems correspond to combinatorial optimization problems, such as vehicle routing and scheduling, that are notoriously difficult to solve, even in a static context. Moreover, dynamic fleet management concern real-time requirements of the provided service and the control of unforeseen events that could affect the performances of the transport operations. The aim of this paper is to identify the most relevant problems in fleet management relatively to different transport modes, each one with specific characteristics, and to present an overview of recent contributions both in the development of mathematical models and in the design of computation algorithms.

Keywords: Fleet management; Vehicle routing and scheduling; Transport modes; Optimization; Models; Algorithms

1. Introduction

The management of fleets of different transport modes presents a relevant activity at tactical and operational levels to be faced by private companies and public agencies devoted to passengers and freight transportation services. In the framework of Operations Research and Management Science disciplines many mathematical models and quantitative methods have been developed for optimising and simulating the operation of transport fleets in order to serve the customers demand. These tools with the relative implemented software packages provide a valuable support to decision makers and industrial managers. However, several models are idealized or simplified with assumptions non realistic for practical applications, because real-world situations almost always present aspects which evade mathematical structuring. Therefore, in order to meeting the need of more powerful and flexible tools, the research has been focused on more complex and general models and advanced computation techniques, able to
deal with large scale practical problems. For example, with reference to the classical Vehicle Routing Problem (VRP), representing the most relevant problem in freight distribution in road transport, the research community has turned to variants of the VRP which before were considered too difficult to handle. The development of “Rich VRP” denotes in the literature models able to include many aspects that are essential to the routing of vehicles in real-life, by considering fleet composition, soft time windows, fixed and variables costs, constraints on crew, and so on (Hartl et al. 2006).

Indeed, in a typical large or medium-sized city, many private firms and public organizations operate fleets of vehicles of different types to cater to various needs of the population: emergency vehicles (fire trucks, ambulances, etc.), police cars, commercial delivery vehicles, taxis, courier fleets, etc. Some of these fleets have to perform tasks that may be known well in advance or that are sometimes repetitive (e.g., vehicles making regular deliveries to food and retail stores). Many of them, however, operate essentially in a demand-responsive mode: the demands for services are not known beforehand and the fleet has to be deployed and managed (re-routed) in real-time to handle them as effectively as possible. The same description applies to “local” pick up and delivery operations performed within a relatively short time period (e.g., a day) in the surrounding area of major intermodal terminals such as ports and major rail yards (Crainic et al. 2009).

Moreover, scheduled transportation networks (e.g. transit and rail) give rise to very complex and large-scale network optimization problems, which require innovative solution techniques and ideas from mathematical optimization and theoretical computer science. Applicable tools and concepts include those from graph and network algorithms, combinatorial optimization, approximation and online algorithms, stochastic and robust optimization. Therefore, new models and algorithms can improve the productivity of resources, efficiency and network capacity. Today, senior executives realize that they need to develop knowledge-based decision-support systems where modelling, algorithmic intelligence and optimization systems constitute an integral part of decision making processes (Ahuja and Liebchen, 2011).

From a modelling standpoint, many fleet management problems correspond to combinatorial optimisation problems (e.g., vehicle routing, scheduling, network design) that are notoriously difficult to solve, even in a static context. This, coupled with real-time requirements, explains to a large extent the reliance up to now on human dispatchers. Fortunately, recent developments in the area of algorithms, in particular the emergence of powerful meta-heuristics, and advances in computing technology, in particular distributed and parallel computing and artificial intelligence, now make it possible to contemplate tackling in real-time large combinatorial problems in a reasonably effective way. In fact, currently, the main obstacle in most advanced fleet management systems applications is the need to handle dynamic (stochastic) data (Crainic et al. 2009).

The aim of this paper is to identify the most relevant problems in fleet management relatively to different transport modes, each one with specific characteristics, and to present an overview of recent contributions both in the development of mathematical models and in the design of computation algorithms.

2. Vehicle routing and scheduling

The Vehicle Routing Problem (VRP) is a general class of problems in which a fleet of vehicles with limited capacity based at one or several depots has to be routed serving a certain number of customers to minimize the number of routes and total travelling time as well as distance of all vehicles. VRP with time windows is one generalization of the VRP with additional restrictions where each customer has to be supplied in a specific time interval. More realistic problems must consider that most companies have heterogeneous vehicles and 24-hour driving/working per day is infeasible, because total daily driving and working time is limited, with total working time including total driving time, waiting time, and loading/unloading time. Nowadays, due to higher fuel prices, minimizing delivery costs has become increasingly important. Therefore, VRP receives considerable amount of attention from industries and arises as a central problem in the field of transportation, distribution and logistics.

Mathematical formulation of these problems with relative solution algorithms are presented by Cordeau et al. (2007). Moreover, Golden et al. (2008) collect in their book significant contributions on methodological advances and new approaches for solving existing VRP, and on novel problems that have arisen in the vehicle routing domain by highlighting new challenges for this field. In particular, overview and surveys, new directions in modelling and algorithms, practical applications are illustrated. Most relevant problems are discussed and examined, such as
routing of a heterogeneous vehicle fleet, capacitated arc routing, dynamic VRP, multi-objective VRP, parallel solution methods and metaheuristics.

The VRP has already been proven as NP-Hard problem. Currently, application of exact methods is still limited to about 100 customers, therefore, it is desirable to obtain approximate solutions instead of optimal ones so that they could be reached quickly and accurate enough to be acceptable, compared to optimal solutions. In this case, heuristic approaches have been introduced to solve large instances. In fact, the last fifteen years an incremental amount of metaheuristic algorithms have been designed. Simulated annealing, genetic algorithms, artificial neural networks, tabu search, ant colony optimization, Greedy Randomized Adaptive Search Procedure, Guided Local Search, Variable Neighborhood Search together with a number of hybrid techniques are the main categories of the metaheuristic procedures. In particular, these algorithms have been proposed for the solution of the Vehicle Routing Problems and they have the ability to find their way out of local optima (Marinakis, 2009).

Local search is the most frequently used heuristic technique for solving combinatorial optimization problems. It is also the basis for modern metaheuristics, like, e.g., Tabu Search, and Variable Neighborhood Search. The paper by Irnich et al. (2006) introduces sequential search as a generic technique for the efficient exploration of local-search neighborhoods. One of its key concepts is the systematic decomposition of moves, which allows pruning within local search based on associated partial gains. The application of theoretical concepts to several well-known neighborhoods of the vehicle-routing problem is also demonstrated. Moreover, Khebbache-Hadji et al. (2011) present heuristics and memetic algorithms for two-dimensional loading capacitated vehicle routing problem with time windows.

2.1. Fleet composition and routing

In this class of problems the fleet is heterogeneous and a mix of vehicle types differing in capacity and costs is given. Then, the fleet composition must be determined, as regards the type, the number of vehicles of each type, as well as the order in which to serve the customers with each vehicle, with the objective of minimizing the total costs.

Hoff et al. (2010) present a complete classification of these problems with reference to maritime and road-based transportation. Moreover, a wide literature survey is described and categorized.

In particular, from a computational point of view, the design and implementation of a genetic algorithm based heuristic has been developed by Liu et al. (2009), while a heuristic algorithm based on tabu search has been proposed by Brandao (2009).

Repoussis and Tarantilis (2010) present a novel Adaptive Memory Programming (AMP) solution approach, where the target is to design a set of depot returning vehicle routes able to service a set of customers with known demands. The proposed method utilizes the basic concept of an AMP solution framework equipped with a probabilistic semi-parallel construction heuristic, a novel solution re-construction mechanism, an innovative Iterated Tabu Search algorithm tuned for intensification local search and frequency-based long term memory structures. Computational experiments on well-known benchmark data sets illustrate the efficiency and effectiveness of the proposed method, which improves the best reported cumulative and mean results over most problem instances with reasonable computational requirements.

Prins (2009) presents two memetic algorithms (genetic algorithms hybridized with a local search). They are based on chromosomes encoded as giant tours, without trip delimiters, and on an optimal evaluation procedure which splits these tours into feasible trips and assigns vehicles to them. The second algorithm uses a distance measure in solution space to diversify the search. Numerical tests on standard instances show that the two methods, especially the one with distance measure, compete with published metaheuristics and improve several best-known solutions.

The contribution of Tütüncü (2010) concerns a visual interactive approach based on a new greedy randomised adaptive memory programming search (GRAMPS) algorithm is proposed to solve the heterogeneous fixed fleet vehicle routing problem (HFFVRP) and a new extension, which is called heterogeneous fixed fleet vehicle routing problem with backhauls. This problem involves two different sets of customers. Backhaul customers are pickup points and linehaul customers are delivery points that are to be serviced from a single depot by a heterogeneous fixed fleet of vehicles, each of which is restricted in the capacity it can carry, with different variable travelling costs.
The proposed approach is implemented within a visual decision support system, which was developed to allow users to produce and judge alternative decisions by using their knowledge and experience about the requirements of the HFFVRP.

Bettinelli et al. (2011) present a branch-and-cut-and-price algorithm for the exact solution of a variation of the vehicle routing problem with time windows. Different pricing and cutting techniques are illustrated and an experimental evaluation of their combinations presented. Computational results are reported on the use of the algorithm both for exact optimization and as a heuristic method.

Moreover, the solution of routing problems with soft time windows has valuable practical applications. Soft time window solutions are needed when: (a) the number of routes needed for hard time windows exceeds the number of available vehicles, (b) a study of cost-service tradeoffs is required, or (c) the dispatcher has qualitative information regarding the relative importance of hard time-window constraints across customers. Figliozzi (2010) proposes a new iterative route construction and improvement algorithm to solve vehicle routing problems with soft time windows. Due to its modular and hierarchical design, the solution algorithm is intuitive and able to accommodate general cost and penalty functions. Experimental results indicate that the average run time performance is of order O(n2). The solution quality and computational time of the new algorithm has been compared against existing results on benchmark problems and improved solutions demonstrated.

3. Dynamic fleet management

Classical fleet management problems address the development of routing and scheduling plans for the transport means with the objective of cost efficiency and customer satisfaction. However, unforeseen events could affect the performances of the transport operations, such as customer orders arriving in real-time, travel times delays, breakdowns and so on. Dynamic fleet management exploits information provided by communication technologies (mobile and satellite) in order to improve the real-time use of transport resources. Moreover, the deployment of Intelligent Transportation Systems technologies, in particular accurate positioning devices and in-vehicle computing and communication equipments, opens up the possibility of enhanced customer service and increased productivity by re-routing vehicles in real-time to serve new requests. One only needs the appropriate methodology to transform these data into accurate and timely decisions. It is thus normal that a significant line of research addresses the issues of real-time dispatching, routing, and re-routing of vehicles in response to changes in demand, travel time or other conditions of travel. This information can be conveyed via Advanced Traveller Information Systems, as well as wireless or on-board communication devices. Zeimpekis et al. (2007) in their book collect important advances in dynamic vehicle routing and fleet dispatching from methodological and computational point of view. Moreover, several case studies relative to different transport modes, such as air transport, intermodal transport of hazardous materials, courier fleet and city logistics, are illustrated. Cheung et al. (2008) develop and analyse a mathematical model for dynamic fleet management that captures the characteristics of modern vehicle operations. The model takes into consideration dynamic data such as vehicle locations, travel time, and incoming customer orders. The solution method includes an effective procedure for solving the static problem and an efficient re-optimization procedure for updating the route plan as dynamic information arrives. Computational experiments show that the re-optimization procedure can generate near-optimal solutions.

In this framework, the research focused on models and algorithms for the real-time vehicle routing problem. There are well-known integer programming approaches that give an almost exact solution to such a model, but they may require much longer computational time for large-size problem instances. For the dynamic-update module, the quick heuristic procedure can generate good updated solutions in seconds, and the effectiveness of the heuristic is also demonstrated computationally. In fact, a critical issue in real-time settings is that of response time. In situations such as emergency vehicle management, or when a customer is waiting for a decision, there is no time to compute an “optimal” response when a call is received. This does not preclude, however, the use of deliberate decision-making to optimise the response: one simply has to find ways of anticipating future events in an effective fashion. Thus, for example, one may combine data processing and forecasting methods, optimisation-simulation models, and decision heuristics into comprehensive decision-support systems. The optimisation-simulation models continuously generate and evaluate future conditions and deployment scenarios, while rapid, simpler heuristics respond in real-time to customer requests or changing conditions (congestion, accidents, and so on). Note, however, that this may
result in significant computational requirements, since one has to prepare for many potential outcomes. Parallel computing may help address this issue as well as provide more robust solutions (Crainic, 2008).

4. City Logistics

Another field of increasing interest is the urban freight transportation and the development of new organisational models for the management of freight movements within the city. As for any complex system, city logistics transportation systems require planning at strategic, tactical and operational levels (Benjelloun and Crainic, 2008) and the development of formal mathematical models (Taniguchi and Thompson, 2002). Moreover, the development of Intelligent Transportation Systems allows to exploit the availability of data and information provided by advanced communication technologies for better use of the transportation system, infrastructure and services (Crainic et al. 2009) and for intelligent vehicle routing and scheduling (Thompson 2004). In fact, in wide area road networks, vehicle routing is usually based on distances. However, vehicle routing in city logistics networks demands for time-dependent travel time estimates for every route section. Vehicle routing based on actual travel time estimates requires empirical traffic data as a key input, but, recently, such data arises from modern vehicular communication networks. In order to benefit from telematics based data collection, time-dependent travel time estimates have to be integrated into time-dependent vehicle routing approaches. Whereas static approaches are well studied, time-dependent vehicle routing forms a field of potential research. Especially the preparation and the integration of time-dependent data into vehicle routing approaches is rarely focused.

Van Woensel et al. (2008) consider queuing theory to provide time-dependent travel time estimates. They use a tabu search approach to solve the time-dependent capacitated vehicle routing problem. Donati (2008) focuses on ant colonies to heuristically solve time-dependent vehicle routing problems. Both publications are more focused on large area networks. A holistic approach to the design and evaluation of City Logistics applications has been proposed by Barcelo et al. (2007) within an integrated framework of a decision support systems, where vehicle routing and fleet management represent the core models together with models able to include the dynamic aspects of traffic on the underlying road network., such as the AIMSUN simulation model. This is able to track individually the fleet vehicles, emulating their monitoring in a real time fleet management system and gathering dynamic data. This information allows a “Dynamic Router and Scheduler” to determine which vehicle will be assigned to the new service and its new route.

5. Urban public transport

The global problem faced by the urban public transport agencies consists of determining how to offer a good-quality service to the passengers while maintaining reasonable asset and operating costs. The public-transport (transit) operation planning process commonly includes four basic activities, usually performed in sequence: (1) network route design, (2) frequency setting and timetable development, (3) vehicle scheduling, and (4) crew scheduling and rostering.

Desaulniers and Hickman (2007) presents a wide and complete review state-of-the-art models and approaches for solving all the public transit problems at strategic, tactical and operational levels. Guihaire and Hao (2008) propose a classification of approaches dealing with the design, frequencies setting, timetabling of transit lines and their combinations. Hickman et al. (2008) collect several contributions on vehicle routing and crew scheduling, timetabling optimisation and service management for different modal public transport.

With reference to the first two aforementioned activities, Zhao and Zeng (2008), present a metaheuristic method for optimizing transit networks, including route network design, vehicle headway, and timetable assignment. Given information on transit demand, the street network of the transit service area, and total fleet size, the goal is to identify a transit network that minimizes a passenger cost function. Transit network optimization is a complex combinatorial problem due to huge search spaces of route network, vehicle headways, and timetables. The methodology described includes a representation of transit network variable search spaces ; a user cost function based on passenger random arrival times, route network, vehicle headways, and timetables; and a metaheuristic search scheme that combines simulated annealing, tabu and greedy search methods. This methodology has been tested with problems reported in the existing literature, and applied to a large-scale realistic network optimization
problem. The results show that the methodology is capable of producing improved solutions to large-scale transit network design problems in reasonable amounts of time and computing resources.

Bielli et al. (2002) proposed a heuristic approach to solve transportation bus network optimisation problems. The method involves genetic algorithms and a number of additional ingredients which allows to compute fitness function values by means of a multicriteria analysis executed on the performance indicators.

Instead, Ceder (2011) addresses the vehicle scheduling problem, while taking into account the association between the characteristics of each trip (urban, peripheral, inter-city, etc.) and the vehicle type required for the particular trip. The problem is based on given sets of trips and vehicle types, where the categories are arranged in decreasing order of vehicle cost. Therefore, each trip can be carried out by its vehicle type, or by other types listed in prior order. This problem can be formulated as a cost-flow network problem with an NP-hard complexity level. Thus, a heuristic algorithm is developed, based on the Deficit Function theory. Two examples are used as an expository device to illustrate the procedures developed, along with a real-life example of a bus company.

Moreover, the vehicle scheduling problem, arising in public transport bus companies, must address the task of assigning buses to cover a given set of timetabled trips with consideration of practical requirements, such as multiple depots and vehicle types as well as depot capacities. An optimal schedule is characterized by minimal fleet size and minimal operational costs including costs for unloaded trips and waiting time. Kliever et al. (2006) discuss the multi-depot, multi-vehicle-type bus scheduling problem for timetabled trips. They use time–space-based instead of connection-based networks and this leads to a crucial size reduction of the corresponding mathematical models compared to well-known connection-based network flow or set partitioning models. The proposed modeling approach enables to solve real-world problem instances with thousands of scheduled trips by direct application of standard optimization software. The largest problems solved to optimality could not be solved by any existing exact approach.

A different approach proposed by Belmonte et al. (2008) is based on the design and implementation of a multi-agent decision support system for the bus fleet management domain. In particular, they show a complete description of the proposed multi-agent architecture and focus mainly on knowledge and software engineering features.

Another class of public transport problems faced by bus companies concerns the school transport service. The paper by Park and Kim (2010) aims to provide a comprehensive review of the school bus routing problem. It seeks to plan an efficient schedule for a fleet of school buses where each bus picks up students from various bus stops and delivers them to their designated schools, while satisfying various constraints, such as the maximum capacity of a bus, the maximum riding time of a student in a bus, and the time window of a school. This class of problem consists of different sub-problems involving data preparation, bus stop selection, bus route generation, school bell time adjustment, and bus scheduling. The various assumptions, constraints, and solution methods used in the literature are summarized and a list of issues requiring further research is also presented.

Therefore, public transit has provided interesting and challenging problems that have been successfully and efficiently solved by the development of innovative solution approaches (models and computation techniques). The research trends concern the implementation of software tools able to solve larger instances and additional complexities, but also to represent the core of a decision support system.

6. Dial-a-ride transport

The realization of innovative passengers transport services requires more and more often a greater flexibility and inexpensiveness of the service. To answer this request in many cases the solution is to realize demand responsive transportation systems (DRTS), that require the planning of travel paths (routing) and customer pick-up and drop-off times (scheduling) on the basis of received requests. In particular, it has to tackle the problem of multiple vehicles with limited capacity and temporal constraints (time windows). The problem of working out optimal service paths and times is called a Dial-a-Ride Problem (DaRP), which derives from the well-known Vehicle Routing Problem with the addition of precedence constraints between pick-up and drop-off locations as well as an upper bound on the riding time. The complete mathematical model formulation of this problem is presented by Cordeau and Laporte, (2007).

Its computational complexity makes DaRP as a well known NP-Hard combinatorial optimization problem, so attempts to develop optimal solutions have been limited to simple and small-size problems. It can be argued that
heuristic procedures are more suitable for realistic networks and demand, because they allow to obtain good solutions in a limited amount of time.

Furthermore, in this context, customers often formulate two requests per day, specifying an outbound request from the pick-up point to a destination and an inbound request for the return trip. A DRTS may operate according to a static or to a dynamic mode. In the static setting, all the customer requests are known beforehand, and the DRTS produces, by solving a DaRP instance, the tour each vehicle has to make, respecting the pick up and delivery time windows while minimizing the solution cost. In the dynamic mode, the customer requests arrive over time to a control station and, consequently, the solution may also change over time. Then, the computation must operate dynamically without interrupting the optimization cycle until the end of the service, providing the best solution found each time the system is modified. It must to point out that the DaRP has to notify customers as soon as possible the acceptance or denial of their requests. Moreover, according to the dynamic mode, the DaRP has to be capable of easily inserting new customer requests into one of the already initiated tours without violating any previously accepted customer requests. The goal is not only to minimize the operating costs incurred to satisfy all requests but also to maximize the quality of the service, expressed by indicators such as the average passenger waiting time and the on-board (ride) time.

Gomes and de Sousa (2009) develop a dynamic vehicle routing model able to optimality solve very small instances, but, nonetheless, it should be noted that the resulting formulation is quite flexible in accommodating new constraints and problem variations. Using greedy insertion heuristics, one can quickly construct feasible solutions for the problem. In fact, from a computational point of view, several algorithms have been developed and tested, such as tabu search heuristics, dynamic programming, branch and cut, a heuristic two phase solution method, genetic algorithms, variable neighborhood search.

Moreover, in many real world applications, different types of users must be explicitly taken into account. For example, in the field of patient and disabled people transportation, up to four different transportation modes can be distinguished. In the paper by Parragh (2011), staff seats, patient seats, stretchers and wheelchair places are considered. Furthermore, most companies involved in the transportation of the disabled or ill dispose of different types of vehicles. Then, both aspects are introduced into state-of-the-art formulations and branch-and-cut algorithms for the standard dial-a-ride problem are developed. Also a recent metaheuristic method is adapted to this new problem. In addition, a further service quality related issue is analysed: vehicle waiting time with passengers aboard.

Finally, it must be remarked that few studies have been conducted on the quality of service provided by the organizations responsible for the operation of dial-a-ride services for people with reduced mobility. In order to study quality in such contexts, Paquette et al. (2009) first provide several definitions of quality in the service sector. Then, some models based on the definitions of quality in marketing are quickly surveyed, and various measurement scales presented. After a brief review of the tools and practices used in the service sector to improve quality, the notion of quality in dial-a-ride services is discussed more in depth. In particular, the dimensions and attributes of various scales of measurement used by researchers in dial-a-ride studies are reviewed. Then, the impact on quality of various elements, like the size and type of organization and the operational rules used, are discussed.

7. Air transport

Another relevant class of fleet management problems concern the air transport and the activities carried out by airline companies. Yu and Thengvall (2009) collect several contributions on Airline Optimisation and, in particular, on network design and schedule construction, fleet assignment, aircraft routing, crew scheduling, revenue management, irregular operations, air traffic control and ground delay programs, gate assignment, fuel management, short term fleet assignment swapping. Operations Research techniques have been applied to solve these problems. The majority of applications utilize network-based models and their solutions range from traditional mathematical programming approaches (integer programming and stochastic linear programming) to a variety of novel heuristic approaches.

The airline scheduling process is usually carried out sequentially, so that flight, aircraft and crew schedules are created one after another over several months prior to the day of operations. A detailed flight schedule, consisting of given departure and arrival times for all flights, is based on marketing decisions and provides the basis for operational scheduling tasks. The first step is typically the assignment of an aircraft fleet type to each flight, considering the demand forecasts, the capacity and the availability of the aircraft. After fleet assignment a specific
aircraft is assigned to each flight with respect to maintenance constraints (aircraft routing). Crew scheduling is typically decomposed into two processes. In the first crew pairing phase anonymous crew itineraries, called pairings, are generated regarding general regulations, such as maximal allowed working time or flying time per duty. In the second phase (crew rostering), individual crew members are assigned to the itineraries. The main goal of this scheduling process is cost optimality.

Fleet routing and flight scheduling are essential to a carrier’s profitability, its level of service and its competitive capability in the market. Normally, network flow techniques are adopted for modelling and solving such complex mathematical problems. It is believed that the full optimization problem is computationally intractable, and hence the constituent subproblems are optimized sequentially so that the output of one is the input of the next. The sequential approach, however, provides an overall suboptimal solution and can also fail to satisfy the maintenance constraints of an otherwise feasible full problem. Papadkos (2009) proposes several integrated models for the optimization of airline scheduling, solved by applying an enhanced Benders decomposition method combined with accelerated column generation. As a result, the integrated approach significantly reduces airline costs. Finally, a comparison of alternative formulations has shown that keeping the crew scheduling problem alone in the Benders subproblem is much more efficient than keeping the aircraft routing problem.

The fleet assignment problem (FAP) deals with assigning aircraft types, each having a different capacity, to the scheduled flights, based on equipment capabilities and availability, operational costs, and potential revenues. However, due to the large number of flights scheduled each day, and the dependency of the FAP on other airline processes, solving the FAP has always been a challenging task for the airlines. Sherali et al. (2006) present a tutorial on the basic and enhanced models and approaches that have been developed for the FAP, including: (1) integrating the FAP with other airline decision processes such as schedule design, aircraft maintenance routing, and crew scheduling; (2) proposing solution techniques that include additional considerations into the traditional fleeting models, such as considering itinerary-based demand forecasts and the recapture effect, as well as investigating the effectiveness of alternative approaches such as randomized search procedures; and (3) studying dynamic fleeting mechanisms that update the initial fleeting solution as departures approach and more information on demand patterns is gathered, thus providing a more effective way to match the airline’s supply with demand. Bélang er et al. (2006) propose a model for the periodic fleet assignment problem with time windows in which departure times are also determined. Anticipated profits depend on the schedule and the selection of aircraft types. In addition, short spacing between consecutive flights which serve the same origin–destination pair of airports are penalized. A non-linear integer multi-commodity network flow is formulated and new branch-and-bound strategies developed, which are embedded in the branch-and-price solution strategy.

In airline traffic disruptions occur frequently and cannot be totally avoided. They may lead to infeasible aircraft and crew schedules during the day of operations, due to absence of resources or violation of crew rules. The process of finding new schedules in such cases is called recovery or disruption management. The short-term recovery actions usually imply additional costs meaning that the total operational costs of a crew schedule can be significantly higher than the original planned costs. It is generally desirable to construct the schedule already in the planning phase in such a way that not just the planned costs, but the total operational costs are minimized. The goal is thus to construct schedules which remain feasible or can be recovered without high costs in cases of disturbances. This approach is generally called robust scheduling.

Mercier et al. (2005) propose an integrated aircraft routing and crew scheduling problem incorporating a robustness measure. For the robustness measure the available ground time for crews during an aircraft change is considered. Aircraft changes with ground time below a threshold are called restricted. Schedules with less restricted aircraft changes are assumed to be more robust, therefore a penalty value for such aircraft changes is used. Two Benders decomposition methods combined with branch and price are proposed to solve the integrated model. Both methods lead to very high solution times. The proposed robustness measure does not consider any recovery actions as well as no information about delay probabilities. Thus, only a local evaluation of the aircraft changes is carried out. Weide et al. (2007) propose to solve the integrated aircraft routing and crew scheduling model heuristically and iteratively using a shared objective function. A non-robustness measure is used to penalize restricted aircraft changes according to the amount of slack time during such an aircraft change.

Dück et al. (2009) aim to improve stability of aircraft routing and crew scheduling and formulate a stochastic integrated model, which involves propagation of delays through aircraft as well as crew through an integrated recourse function. The formulated model is hard to solve, therefore, several changes leading to a more tractable
model for robust aircraft routing and crew scheduling were proposed. The new model can be solved by a combination of the iterative approach, proposed by Weide et al. (2007), and classical column generation methods, thus providing a good starting point for considering robustness with real-life problem instances. Thus, the iterative approach yields a set of different solutions regarding the trade-off between costs and robustness whereas an integrated approach returns mostly one near-optimal solution for a given robustness penalty. The selection of the “right” robustness penalty is difficult. Therefore, it is believed the iterative approach is more advantageous for a decision maker, who can review several solutions with different trade-offs and select the best one.

Recently, as a means of forming global networks and improving operation efficiency, major air carriers have increasingly entered into alliances with other carriers. Fleet routing and flight scheduling are not only important in individual airline operations, but also affect the alliances. The setting of a good flight schedule can not only enhance allied airline operating performance, but can also be a useful reference for alliance decision-making. Yan and Chen (2007) develop several coordinated scheduling models, which will help the allied airlines solve for the most satisfactory fleet routes and timetables under the alliance. The models are formulated as multiple commodity network flow problems which can be solved using a mathematical programming solver.

8. Maritime transport

The term maritime transportation refers to seaborne transportation, but also to other water-borne transportation, namely inland waterways. The field is characterized by ship and port characteristics, logistics, containers management and interconnections among vessels with port and terminal operations.

Christiansen et al. (2007) discuss various aspects of maritime transportation operations and presents associated decision making problems, models and solution approaches for the three modes of operations in maritime transportation, namely liner, industrial, and tramp. In particular, tactical problems include: adjustments to fleet size and mix, fleet deployment (assignment of specific vessels to trade routes), ship routing and scheduling. The issue of robustness in maritime transportation planning is also addressed, because there are many uncertain factors in the ocean shipping industry resulting in delays and lack of timely fulfilment of plans. Therefore, it may often be important to consider robustness in optimization models used for planning. Several methodologies have been applied for problem solutions, such as simulation, dynamic programming, re-optimization for different scenarios or input parameters, deterministic models that incorporate penalties, stochastic optimization models.

In particular, in the framework of the fleet composition and mix routing problem, Hoff et al. (2010) presented a wide survey of the different specific problems and the state of art of the main contributions in the literature. Moreover, the paper by Fagerholt et al. (2010) presents a decision support methodology for strategic planning in tramp and industrial shipping. The proposed methodology combines simulation and optimization, where a Monte Carlo simulation framework is built around an optimization-based decision support system for short-term routing and scheduling. The simulation proceeds by considering a series of short-term routing and scheduling problems using a rolling horizon principle where information is revealed as time goes by. The approach is flexible in the sense that it can easily be configured to provide decision support for a wide range of strategic planning problems, such as fleet size and mix problems, analysis of long-term contracts and contract terms. The methodology is tested on a real case for a major Norwegian shipping company, by providing valuable decision support on important strategic planning problems.

As regards the ferry scheduling problem, it deals with finding routes and schedules for a fleet of ferries in order to serve transportation requests between a set of ports. The most common objectives in the literature and in practice include minimization of ferry costs and maximization of customer satisfaction. The paper by Mitrovic Minic and Punnen (2009) deals with a real-world problem of a ferry company that serves the demand for transportation of cars, trucks, and passengers using a heterogeneous fleet of re-configurable ferries. The scheduling model aims at supporting the company whose management wanted to re-evaluate their current schedules and consider possible improvements. The goal is to use the available resources (ferries and crews) efficiently and to reliably meet the existing demand for the transportation among the residents and tourists between the ports. Considered improvements of the ferry service include: decreasing the operational costs, reducing excess travel time and waiting time for passengers. Wang et al. (2008) formulate a mixed-fleet ferry routing and scheduling model while considering passengers’ choices for differential services. Ferry services with different operation characteristics and passengers with different preferred arrival time-windows are considered in the model. The logit model is applied to determine
passengers’ service choices. Then, the formulation determines the best mixed-fleet operating strategy, including interlining schemes, so as to minimize the objective function that combines both the operator and passengers’ performance measures. Mathematically, this mixed-fleet routing and scheduling problem is formulated as a mixed integer non-linear program and an iterative heuristic algorithm to solve this problem is developed.

9. Rail and intermodal transport

The intermodal transport represents a complex system composed by different transport networks, infrastructures, different transport means and operators, such as drayage operators, terminal operators, network operators and intermodal operators, who take care of the route selection for a shipment through the whole intermodal network. These operators face operational problems with different time horizons: strategic, tactical and operational problems can be distinguished. Intermodal transport, by definition, involves many decision makers who all need to work in collaboration in order for the system to run smoothly. If intermodal transport is to be developed it will require more decision-making support tools to assist the many actors and stakeholders involved in the operation.

In particular, the operational level involves the day-to-day management decisions about the load order of trains and barges, redistribution of railcars or push barges, and load units (fleet management). A typical management problem in intermodal rail/road transport is the assignment of a set of trailers and containers to the available flatcars that can move this equipment. In fact, decisions made by intermodal operators deal with route and service choices in existing intermodal networks. This type of decision, by its nature, is an operational one, because it concerns the assignment of shipments to routes and carriers. Intermodal routing is rather more complex than the routing problems of road haulage. In road haulage, the minimum cost path algorithm was most commonly used in order to find the route most suited to meeting the objective of the company, for example, the least costly or less time consuming route. A large variety of combinations of transportation modes is possible. In this case the routing decision is a more modal choice problem for specific trajectories between beginning and end points, taking into account specific freight volumes and, possibly, specific time constraints. Operations Research has been used for various strategic, tactical and operational problems faced by all the operators (Macharis and Bontekoning, 2004).

Moreover, container transportation is a major component of intermodal transportation carried out by a combination of truck, rail, and ocean shipping, dedicated rail services to move massive quantities of containers and trailers over long distances. Fleet management covers the whole range of planning and management issues from procurement of power units and vehicles to vehicle dispatch and scheduling of crews and maintenance operations. Often, however, the term designates a somewhat restricted set of activities: allocation of vehicles to customer requests and repositioning of empty vehicles (Crainic and Kim, 2007).

Whereas, the rail transport is characterized by different kind of trains travelling on the network and the subdivision of resources between passenger and freight trains. Main problems concern train scheduling and routing taking into account given timetables.

D’Ariano et al. (2007) study a train scheduling problem faced by railway infrastructure managers during real-time traffic control. When train operations are perturbed, a new conflict-free timetable of feasible arrival and departure times needs to be re-computed, such that the deviation from the original one is minimized. The problem can be viewed as a huge job shop scheduling problem with no-store constraints. The scheduling problem is modelled with an alternative graph formulation and a branch and bound algorithm is developed, which includes implication rules enabling to speed up the computation.

Mao et al. (2007) describe the key models and algorithms applied in train scheduling, locomotive and crew work plan, yard operational plan and experimental analysis of train timetables. Authors advance an integrated framework which deals with its major functions, key theoretical and technological progresses and several applications of the platform, including ones in improvements of railway station design, signal layout optimization, train scheduling evaluation at microscopic level and possible other applications in the construction of high-speed railways of China.

Mu and Dessouky (2011) present planning and scheduling tools needed to effectively manage the scarce resources, in order to cope with the rapidly increasing demand for railway transportation. This research develops optimization-based approaches for scheduling of freight trains. Two mathematical formulations of the scheduling problem are introduced and several heuristics based on mixtures of the two formulations are proposed. The proposed algorithms are able to outperform two existing heuristics, namely a simple look-ahead greedy heuristic and a global
neighborhood search algorithm, in terms of railway total train delay. For large networks, two algorithms based on the idea of decomposition are developed and are shown to significantly outperform two existing algorithms.

Burdett and Kozan (2009) consider techniques for scheduling additional train services integrated into current timetables and involving general time window constraints, fixed operations, maintenance activities and periods of section unavailability. The problem is characterised as a hybrid job shop scheduling problem with time window constraints. To solve this problem constructive algorithm and meta-heuristic scheduling techniques that operate upon a disjunctive graph model of train operations are utilised.

Hong et al. (2009) present a two-phased train-set routing algorithm to cover a weekly train timetable with minimal working days of a minimal number of train-sets. First, relax maintenance requirements and obtain minimum cost routes by solving the polynomial relaxation. Then, maintenance-feasible routes are generated from the crossovers of the minimum cost routes. This pragmatic approach seems particularly effective for the high-speed railway systems, where the railway topology is relatively simple with few end stations while the trains are frequent. An optimal feasible routing is found, which is an 8.8% improvement over the current routing generated by a set partitioning approach based on a path generation scheme.

Corman et al. (2010) address the problem of train conflict detection and resolution, which is dealt every day by traffic controllers to adapt the timetable to delays and other unpredictable events occurring in real-time. A number of algorithmic improvements are described, implemented in the real-time traffic management system ROMA (Railway traffic Optimization by Means of Alternative graphs), achieved by incorporating effective rescheduling algorithms and local rerouting strategies in a tabu search scheme. A fast heuristic and a truncated branch and bound algorithm are alternate for computing train schedules within a short computation time, and investigate the effectiveness of using different neighborhood structures for train rerouting. The computational experiments are based on practical size instances from a dispatching area of the Dutch railway network and include complex disturbances with multiple late trains and blocked tracks. Several small instances are solved to optimality in order to compare the heuristic solutions with the optimum. For small instances, the new tabu search algorithms find optimal solutions. For large instances, the solutions generated by the new algorithms after 20 sec. of computation are up to more than 15% better than those achieved within 180 sec. by the previous version of ROMA.

Kuo et al. (2010) develop a train slot selection model based on multicommodity network flow concepts for determining freight train timetables for scheduling rail services along multiple interconnected routes. The model seeks to minimize operating costs incurred by carriers and delays incurred by shippers while ensuring that the schedules and demand levels are mutually consistent. A column generation-based methodology is proposed for train slot selection to meet frequency requirements. This methodology is embedded in a simulation-based iterative framework, where demand for rail services is re-computed in accordance with the train schedule obtained by solving the freight train scheduling problem.

Cacchiani et al. (2010) study the problem of freight transportation in railway networks, where both passenger and freight trains are run. While the passenger trains have a prescribed timetable that cannot be changed, freight train operators send the infrastructure manager requests to insert new freight trains. For each freight train, the associated train operator specifies a preferred ideal timetable, which can be modified by the infrastructure manager in order to respect safeness operational constraints. In particular, this modification may correspond to routing the train along a path which is different with respect to the one in the ideal timetable. Roughly speaking, the objective is to introduce as many new freight trains as possible by assigning them timetables that are as close as possible to the ideal ones. For this timetabling problem on a generic railway network, an integer linear programming formulation is presented, that generalizes some formulations already presented for the case of a single railway line, and a Lagrangian heuristic based on this formulation.

Zhou and Zhong (2005) address a double-track train scheduling problem for planning applications with multiple objectives. Focusing on a high-speed passenger rail line in an existing network, the problem is to minimize both the expected waiting times for high-speed trains and the total travel times of high-speed and medium-speed trains. By applying two practical priority rules, the problem with the second criterion is decomposed and formulated as a series of multi-mode resource constrained project scheduling problems in order to explicitly model acceleration and deceleration times. A branch-and-bound algorithm with effective dominance rules is developed to generate Pareto solutions for the bicriteria scheduling problem, and a beam search algorithm with utility evaluation rules is used to construct a representative set of non-dominated solutions. A case study based on Beijing–Shanghai high-speed railroad in China illustrates the methodology and compares the performance of the proposed algorithms.
Flamini and Pacciarelli (2008) address a scheduling problem arising in the real time management of a metro rail terminus. It mainly consists in routing incoming trains through the station and scheduling their departures with the objective of optimizing punctuality and regularity of train service. The purpose is to develop an automated train traffic control system, able to directly implement most traffic control actions, without the authorization of the local area manager. The scheduling problem is modeled as a bicriteria job shop scheduling problem with additional constraints. The two objective functions, in lexicographical order, are the minimization of tardiness/earliness and the headway optimization. The problem is solved in two steps. At first a heuristic builds a feasible solution by considering the first objective function. Then the regularity is optimized without deteriorating the first objective function. Computational results show that the system is able to manage the terminus very efficiently.

The determination of fleet size of locomotives and of a policy to deadhead them are tactical issues that influence the level of customer service in a rail network. Godwin et al. (2008) consider a railroad system in which a priori freight train schedule does not exist. A simulation-based approach is proposed for tactical locomotive fleet sizing. Their study shows that the throughput increases with the number of locomotives up to a certain level; after that the congestion caused by the movements of large number of locomotives in the capacity-constrained rail network offsets the potential benefit of a large fleet.

As regards the intermodal transport, a formulation and solution procedure for optimizing the fleet size and freight car allocation under uncertainty demands is proposed by Sayarshad and Tavakkoli-Moghaddam (2010). There are important interactions between decisions on sizing a rail–car fleet and utilizing that fleet. Consequently, the optimum use of empty rail–cars for demands response in the length of the time periods is one of advantages the proposed model. The model also provides rail network information such as yard capacity, unmet demands, and number of loaded and empty rail–car at any given time and location. Consequently, the model helping managers or decision makers of any train company for planning and management activities. A two-stage solution procedure for solve rail–car fleet sizing problem is proposed.

An improvement on drayage operations is necessary for intermodal freight transport to become competitive. When drayage takes place in cities or urban centres transit times are usually random, as a consequence finding the optimal fleet schedule is very difficult, and this schedule can even change during the day. Escudero et al. (2010) present a dynamic optimisation model which uses real-time knowledge of the fleet's position, permanently enabling the planner to reallocate tasks as the problem conditions change. Stochastic trip times are considered, both in the completion of each task and between tasks. Tasks can also be flexible or well-defined. The algorithm is described in detail for a test problem and then applied to a set of random drayage problems of different size and characteristics, obtaining significant cost reductions with respect to initial estimates.

Andersen et al. (2009) present a new optimization model for the tactical design of scheduled service networks for transportation systems, where several entities provide service and internal exchanges and coordination with neighboring systems is critical. Internal exchanges represent border crossings necessitating changes of vehicles, while the coordination with neighboring systems represents intermodal operations. For a given demand, the model determines departure times of the services such that throughput time of the demand in the system is minimized. The model is an extension of the design-balanced capacitated multicommodity network design model, denoted service network design with asset management and multiple fleet coordination to emphasize the explicit modeling of different vehicle fleets. Data from a real-world problem addressing the planning of new rail freight services across borders are used to illustrate the capabilities of the formulation. Authors analyse how synchronization with collaborating services and removal of border-crossing operations impact the throughput time for the freight, moreover, identify a significant potential for system performance enhancement from synchronization among collaborating services for the problem studied.

10. Conclusions

Advanced Fleet Management Systems based on Fleet Telematics (Goel, 2008) represent the new opportunities to process dynamic information to be integrated into the transportation plans, in order to achieve a more timely operation, efficient allocation and utilization of the fleet, and satisfaction of customer requests (Crainic et al. 2009).

As a matter of fact, many analytical approaches from Operations Research and Management Science are well suited to fleet management problems. These include modelling and simulation, multi-objective programming,
optimisation techniques, heuristic algorithms, decision support systems, data mining, performance evaluation, customer satisfaction evaluation, benchmark analysis, statistical analysis and data envelopment analysis.

This paper aimed to show how classical and new methodologies and algorithms could improve knowledge in fleet planning and management with reference to all the transport modes. Moreover, new transport and logistics services represent a real challenge for the efficient and sustainable mobility of passengers and goods, and as a consequence for the management of available resources. The multi-disciplinary nature of transport and logistics services requires the support of different scientific disciplines and the cultural background of applicable projects carried out in different countries (Bielli and Filippi, 2009).

The research trends consist in developing new effective models for fleet management and to combine data mining and forecasting methods, optimisation and simulation models, decision heuristics, and to embed them into comprehensive decision support systems. The computational efficiency of solution methods for fleet management problems may be significantly enhanced through parallel and distributed computing. The integration of exact algorithms and meta-heuristics into co-operative search methods, and the development of co-operation mechanisms based on mathematical programming principles, decomposition methods in particular, integration of operations research methods and artificial intelligence techniques are promising research directions (Crainic et al. 2009).

References


Mitrovic Minic S. and Punnen A.P., 2009, “Routing and scheduling of a heterogeneous fleet of re-configurable ferries: a model, a heuristic, and a case study”, Proc. 13th Meeting of the EURO Working Group Transportation, Padova, Italy


