Schedule Optimization at SNCF: From Conception to Day of Departure

Nejib Ben-Khedher
SABRE Technology Solutions
11/13 Avenue de Friedland
75008 Paris, France

Josephine Kintanar
SABRE Technology Solutions
116 ter rue de Saussure
75017 Paris, France

Cecile Queille
Societé Nationale des Chemins de Fer Français
116 ter rue de Saussure
75017 Paris, France

William Stripling
SABRE Technology Solutions
4255 Amon Carter Boulevard
Fort Worth, Texas 76155, USA

Because of the upcoming deregulation and the advent of high-speed rail networks, European passenger railroads are battling for customers among themselves and with other means of transportation. To maintain a competitive advantage, they rely on the scheduling process as a key factor in winning market share. Societé Nationale des Chemins de Fer Français (SNCF), in partnership with SABRE Technology Solutions, has developed RailPlus, a strategic schedule planning system, and RailCap, a tactical capacity-adjustment system, to meet the new challenges of the rail business. The heart of these systems is a set of advanced operations research models that enable analysts to take a global approach to decision making. The two systems combined provide 110 million francs in incremental revenue per year and substantially reduce operating costs. In addition, the two systems have become the foundation for the reorganization of SNCF/TGV.

The passenger rail industry in Europe will be facing growing challenges in the next decade. A European Union directive requests that the infrastructure be separated from the operations, thus opening the way for rail operators to compete across Europe. The same directive invites governments to tighten their subsidies to national operators, forcing them to improve their control of costs and to use

0092-2102/98/2801/0006/$5.00

TRANSPORTATION—SCHEDULING
NETWORKS/GRAPHICS—MULTICOMMODITY

their resources more efficiently. In the meantime, airline deregulation is driving travel fares down and customer expectations up. The rail industry, especially in its growing high-speed component, is directly affected by all the changes in the travel environment.

To meet these challenges, the European rail operators need competitive products. The schedule or timetable is their product, and it must be well designed, distributed, and managed to attract passengers, enhance revenue, and reduce costs.

The Scheduling Function

Like airlines, rail operators today are looking for advanced decision-support tools in the areas of pricing, yield management, schedule planning, and control. These needs reflect the growing importance of marketing functions over the traditionally more influential transport functions in the railroad organizations.

Scheduling, with its product-positioning component, is the first step in the marketing planning process. Analysts identify needs of the customer, determine the route structure of the railroad, and design the operating schedule that represents the services the rail operator offers. Scheduling is therefore a fundamental component in maximizing overall profit.

Today scheduling has become more complex. This complexity stems from the operators' need to build schedules to fit a changing demand, to meet both constraint-driven and market-driven criteria, and to allow adjustments. The scheduling process has also become more quantitative. More data on revenue, passengers, and costs are available to support in-depth analysis of schedules. In particular, the reservations and inventory-management systems, with their extensive capabilities, offer valuable information on passenger behavior and provide revenue and demand forecasts as the day of departure approaches. It was in response to these challenges that Societé Nationale des Chemins de Fer Français (SNCF), the national railroad of France, contracted with SABRE Technology Solutions (SABRE), a division of the SABRE Group, to develop a comprehensive set of scheduling tools that would allow SNCF to maintain a competitive advantage over other European rail and air operators.

The SNCF Example

In addition to managing conventional passenger trains, freight trains, and commuter lines, SNCF operates the world's most extensive high-speed network. SNCF's high-speed electric-powered *Trains à Grande Vitesse* (TGV) carry over 50 million passengers per year between 140 cities in France and Europe. SNCF deploys over 300 train sets, ranging in size from a 368-seat unit to a 516-seat unit (double-deck). There are approximately 3,000 departures scheduled per week. Total revenue from the TGV operation is 10 billion francs per year. The TGV activity represents more than 53 percent of the total 40 billion intercity passenger-kilometers SNCF realizes each year. SNCF will continue to expand its high-speed network into the next decade, with new service to the Mediterranean coast, the east region of France, and enhanced international service to Belgium, The Netherlands, Germany, Italy, and Spain.

For nearly a decade, SNCF and SABRE as partners have worked to design, develop, and implement a sophisticated rail-reservation-and-distribution system and a comprehensive suite of peripheral decision-support systems, including yield-management (RailRev), schedule-planning (RailPlus), and capacity-management (RailCap) systems.

The RailPlus and RailCap systems have been developed with the objective of meeting the challenges of all the schedule-construction phases. RailPlus is an integrated planning and analysis system designed to support the base schedule development for the high-speed train service at SNCF. RailCap is the capacity-adjustment system supporting tactical changes to the schedule based on up-to-date forecasts provided on a daily basis by RailRev, the SNCF revenue-management system.

SABRE introduced the scheduling tools at SNCF in phases to ensure that potential users reached consensus on their functionality. We used iterative prototyping as a means to help analysts describe their functional specifications. Moreover we conducted a proof-of-concept study to prove the ability of the mathematical models to capture operational constraints and the solution method to process millions of scenarios in a timely manner. Basically SABRE had to prove to the SNCF commercial analysts that a global-network approach—virtually impossible to master manually—would result in a profit increase. SABRE also had to convince the transport analysts that the solution provided was indeed operationally feasible.

As a token of its commitment to both RailCap and RailPlus, SNCF has made organizational changes. It formed the Center of Operations of TGV (COTGV) in the spring of 1994 mainly to perform the yield-management and capacity-adjustment activity. More than 30 commercial analysts proactively control sales, manage fares, and modify schedules, relying heavily on the RailRev and RailCap decision-support tools. RailPlus has also become the centerpiece system in the reorganization of the *Grandes Lignes* (intercity passenger division) into geographical business units (GBUs) in the summer of 1996. The GBUs use RailPlus to construct their schedules. Each GBU typically has three to four marketing analysts and one transport analyst, jointly using the tool to construct the most profitable, yet feasible, schedule for a given season (for example, peak summer). In addition to the four GBUs (South-East, Atlantic, North-East, and Inter-Section), the future-market-project units use RailPlus to perform prospective studies. A coordinator who reconciles the GBUs in case of resource conflicts also uses RailPlus.

Terminology

In this paper, we use the term *train* to refer to a particular service between connecting points along an itinerary called a *train path;* it is identified by a numerical *train symbol.* Each train has an origin station and a terminal station and may have several intermediate stops in between. The train path between any two adjacent stations is called a *train segment*, or *train leg.* A train covers one or more origin/destination (OD) markets on its itinerary.

Capacity is the number of seats available on a particular train. As illustrated in Figure 1, a unit of capacity on the TGV railroad is the *train set*, which is a set of six to

10 articulated passenger cars with a power unit on each end. Train sets with the same capacity in first and second class comprise a *fleet* (as an example, all train sets with 195 first-class seats and 346 second-class seats make up the 2N fleet in Figure 1).

A train consists of one or two *train units* that operate over all or part of its train path. A train unit always contains a train set and can be thought of in terms of the role it plays over the train itinerary (first train unit or second train unit). Train service on the TGV network consists of the following types:

—A *single-unit train* consists of one train unit containing a train set of any fleet type and is open to reservations.

—A *double-unit train* consists of two train units that contain train sets of operationally compatible fleet types. The first train unit must be open to reservations and travels the entire itinerary. The second train unit may be closed to reservations (the train set is being positioned for future use) and may be part of the train for only part of the train path. A *coupling* operation is performed to join two train sets together. The reverse operation is called *decoupling*. Coupling and decoupling operations are performed only at identified stations.

—An *optional train path* is in the schedule but has no regularly assigned capacity. It is like a single-unit train with no train set assigned. During peak periods, optional trains can be assigned a train set and opened for reservations.

—A *shuttle* is an empty train consisting of one or two train sets of any fleet type.

RailCap at SNCF

SNCF operates two published base schedules per year for their TGV service. These schedules include adjustments for peak holiday periods, and they require approximately 85 percent of the available TGV train sets. The remaining 15 percent are used to augment regularly scheduled service during periods of peak passenger activity. The major responsibilities of RailCap are to monitor the reservation activity for all trains and proactively to add capacity (train sets), called *forcements,* to the schedule before trains are closed to reservations.

At SNCF, RailCap receives the latest forecasts produced by the yield-management system RailRev. The forecasts, based on long-term trends and short-term booking activity, are constantly updated. RailCap also draws on the current-schedules database to get the most accurate snapshot of the schedule in operations. Using this input information, RailCap may suggest the following changes to train capacity:

—Add a second train unit to single-unit trains;

—Drop empty second train units or open them to reservations on double-unit trains;

—Open an optional train to reservations and assign it an itinerary-compatible fleet type.

The heart of RailCap is a capacity-adjustment model that seeks to maximize the expected incremental profit while ensuring that a set of operational constraints are met (Figure 2). Capacity adjustments can be suggested from 15 to three days before the train departure. RailCap incorporates the latest information and runs overnight to provide analysts with suggestions in the morning.

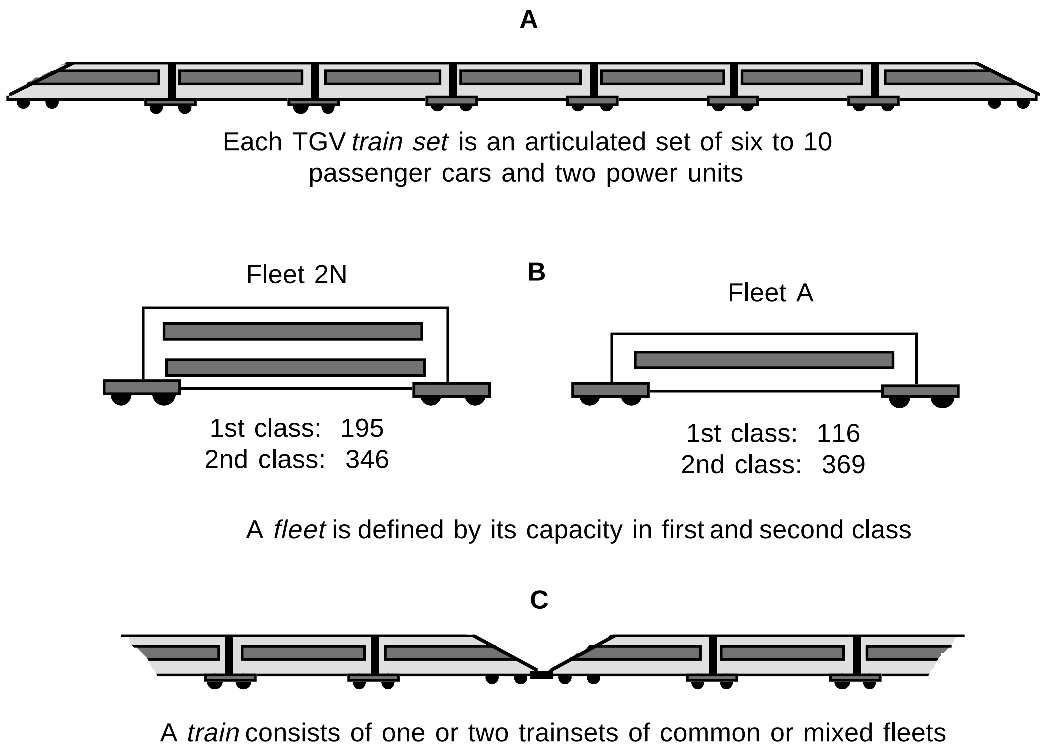

Figure 1: The SNCF TGV train set is an articulated set of two power units and six to 10 passenger cars (A), having one of several capacity configurations (B), and traveling singly or coupled with another train set to make up a multiunit train (C).

The Capacity-Adjustment Model of RailCap

The capacity-adjustment model is an integer multicommodity network flow problem with operational side constraints. The model's network is a mathematical representation of the flow of train sets over the TGV rail network. The model allocates additional train sets to the TGV daily schedule so as to maximize the additional operating profit over a designated time window. The mathematical structure of the capacity-adjustment model is a set of fleet-specific circularized networks composed of nodes and arcs. Network nodes are fleet specific and represent events that occur over time at specific locations. Network arcs represent the possible movements of train sets within a specific fleet between node pairs. The fleet-specific networks are connected by a set of multicommodity constraints. The flow of train sets over arcs is represented mathematically as a set of variables in an integer-programming formulation. (See Figure 3.)

Network nodes can represent events at stations that occur over time. These events include trains originating or terminating at a station and trains making intermediate stops at a station to couple or decouple train sets where permitted. Network nodes can also represent maintenance bases on a given day, and fleet supply and sink amounts at the beginning and end of the

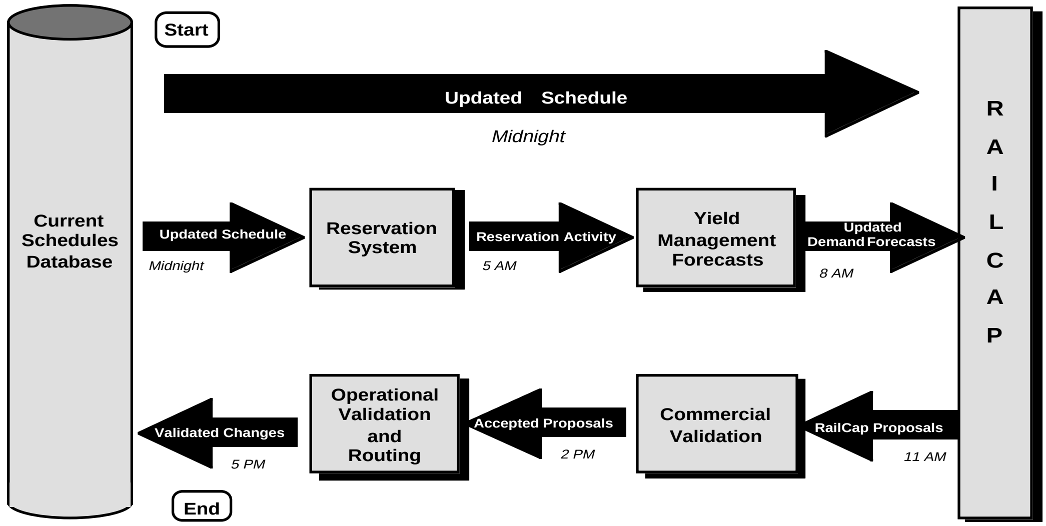

Figure 2: The daily capacity-adjustment process starts at midnight when the current schedule is downloaded to RailCap. At 8:00 AM, the yield-management demand forecasts are downloaded, and RailCap processing begins. When the run is completed, RailCap sends the capacity-adjustment recommendations to the commercial schedule analysts for validation of the expected traffic. The solution is then passed on to the operations analysts who validate its feasibility. The validated solution is then sent to the current-schedules database to update the current schedule.

network. Maintenance nodes represent the daily pool of train sets for each fleet at the four TGV maintenance bases.

Network assignment arcs represent the flow of train sets over the TGV network. Ground arcs carry the flow of train sets between two adjacent station nodes during the same day and have no associated costs. Overnight arcs carry the train-set flow between the last station node of one day and the first station node of the next day. The objective function coefficients of these arcs represent the cost of storing a train set overnight and preparing for service the next day. Maintenance-hub arcs carry train-set flow between consecutive daily maintenance base nodes. These arcs carry a lower bound equal to the number of train sets required to be in maintenance. Shuttle arcs are used at the end of the day for shuttling train sets from stations with limited storage capacity toward stations with greater or unlimited storage capacity.

There are three major types of side constraints in the capacity-adjustment model. The limitation-of-assignment constraint is a multicommodity constraint that is applied to every train departure in the run period, and it requires that exactly one capacity pattern is selected for the train departure. The station storage constraint is a multicommodity constraint that is applied to every station that can store only a limited number of train sets overnight. It requires that the total flow through overnight arcs of all fleets going into the station not exceed the maximum storage limit. The train-set-inventory constraint is applied to every fleet and requires that the available train-set inventory in each fleet not be exceeded.

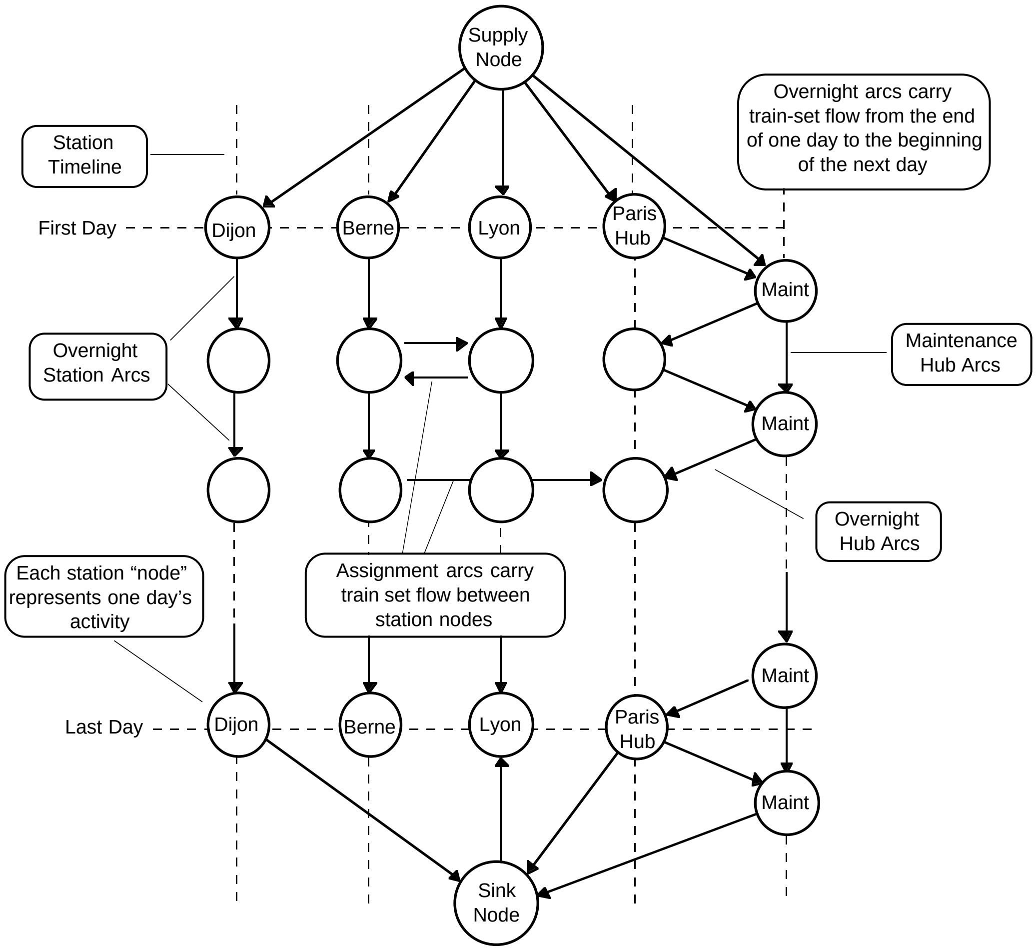

Figure 3: In this simplified representation of the network structure of RailCap, each node represents a day's activity at that site.

Generating the Capacity Patterns

A *capacity pattern* is a particular capacity configuration of a train and represents the number of first- and second-class seats on each leg. Each capacity pattern becomes a decision variable in the capacity-adjustment model. Generating the decision variables requires two steps. First we generate all feasible patterns, and then we assign each pattern an objective function value. One of the classic problems in capacity scheduling for airlines and passenger trains has been reconciling origin/destination (OD)-based demand with leg-based capacity.

Because SNCF forecasts the demand over the TGV network by passenger origin and destination, we had to mathematically represent this OD flow in the decision variables and assignment arcs. The traffic and revenue effect of assigning extra capacity to a train leg depends on the capacity assignments on the other legs of the train because an OD covers several legs and legs are shared by several OD markets. If the model is to evaluate forcements accurately, the assignment variable should relate not only to a leg on the train but to a forcement pattern over all the train legs.

Each decision variable is a set of assignment arcs that is a mathematical representation of a particular capacity-assignment decision made on a train. Each variable represents a scenario, or pattern, of a certain capacity configuration for the train under consideration, and each assignment arc of that variable represents movement of a train set in that pattern. The capacity-adjustment model uses this scenario-based method of OD train-set flow instead of the method of assigning capacity by train leg.

The model tests each possible capacity configuration for feasibility before considering it in the optimization. The feasibility constraints include allowance for coupling or decoupling, operational compatibility between the fleet types in the first and second train units, and operational compatibility of fleet type with stations on the train path (length of platform, electrical equipment).

After RailCap generates a pattern and tests for feasibility, it adds corresponding assignment arcs to the network. The flow over the arcs corresponds to the decision variable, equal to one if the pattern is selected and zero otherwise. As an example, assume that a single-unit train from Paris to Marseille has two intermediate stops (Lyon and Avignon) where it is possible to couple or decouple train sets. Assuming that the second-unit fleet type has to be the same as the first-unit fleet type, there are seven possible forcement patterns (Figure 4).

Scenario 1 adds a second train unit from Paris to Marseille, while scenario 2 adds a second train unit from Paris to Lyon, decouples it in Lyon and continues with one train unit to Avignon and Marseille. The remaining scenarios can be interpreted in a similar manner. Each decision variable is represented in the RailCap network by one or more arcs. For example, there are two arcs in the network for scenario 2: One arc connects Paris and Marseille and represents the first train unit; the second arc connects Paris and Lyon and represents the second train unit.

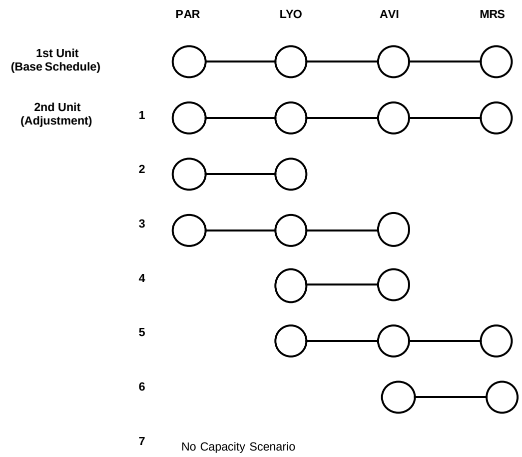

Figure 4: There are seven possible capacity scenarios for a Paris to Marseille train that passes through two intermediate stops (Lyon and Avignon) where it is possible to couple or decouple train sets.

Evaluating the Capacity Patterns

We defined the value of a forcement as the incremental profit produced by the forcement decision. We compute this profit as the incremental revenue minus the incremental cost incurred by the forcement.

In RailCap, the incremental revenue is equal to the expected spilled revenue from the original capacity that is recaptured by the additional capacity in the pattern.

Spill results when a passenger is denied boarding because of insufficient capacity. Given a demand distribution (usually assumed normal) with a certain mean and standard deviation and given a capacity, we can calculate the expected number of spilled passengers as the (conditional) expected demand minus capacity, given that demand is greater than capacity. We developed a closed-form polynomial approximation for the normal distribution [SABRE 1987]. This analytical formula expresses spill as a function of capacity, demand mean, and demand standard deviation or coefficient of variation.

To compute spill at the OD level, we had to convert leg capacity to OD capacity. The spill valuation is further complicated by the fact that the capacity-adjustment model must consider the effects of the SNCF yield-management system (RailRev) on spilled revenue. Long-haul ODs compete with short-haul ODs for the leg capacity. The capacity-adjustment model determines spill at the OD level by allocating the train capacity on each leg to all the ODs over those train legs such that it minimizes total spilled revenue. This allocating capacity at the OD level is called the *capacity-distribution problem* [SABRE 1991], and RailCap solves this problem for every capacity pattern.

The capacity-distribution problem is a nonlinear optimization problem because the expected spilled revenue is a nonlinear function of the OD demand distribution and the capacity allocated (decision variable) on each leg to each OD that a passenger might travel. We used a gradient search method to solve the problem. This optimization approach uses a Lagrangian relaxation [Fisher 1985] of the problem with a subgradient algorithm [Held, Wolfe, and Crowder 1974] and can be very fast if appropriately tuned. The routine is executed more than 20,000 times in a typical run.

Once we have calculated the incremental traffic and revenue for each scenario, we must take into account the effects of recapture. *Recapture* is that percentage of passengers spilled from a train that will be accommodated on other trains. Thus, the net value of the capacity added to a train by a forcement is affected by the potential of accommodating on adjacent trains the passengers that would be spilled with a no-forcement decision. An adjacent train is a train that shares at least one common OD with the reference train and the departure times of the common ODs are within a specified time window. If adjacent trains exist and are expected to have enough capacity to recapture all or some of the first-unit spill, the value of the forcement will be reduced accordingly.

To determine the value of a forcement, RailCap computes the incremental revenue induced by the forcement minus the incremental operating cost, which consists of

—Fixed costs, such as the origination cost;

—Schedule-related costs, such as driver cost, power cost, coupling/decoupling cost, maintenance cost, and toll cost, which are based on the train itinerary and capacity configuration; and

—Passenger-related costs, such as service cost, reservation cost, and controller cost, which are a function of the projected traffic on the train.

RailPlus at SNCF

RailPlus is an integrated schedule-planning-and-analysis system designed to support the development of base schedules for the TGV system at SNCF. To develop schedules, it uses a time-phased approach that begins with SNCF's long-term TGV plans to service profitable markets and to acquire new fleets. These long-term decisions are usually market driven and based on projected profitability. As the actual period of operation approaches, SNCF refines the schedule so that it meets the various operational requirements, such as fleet sizes and maintenance constraints.

SNCF uses RailPlus to develop a base schedule at least twice a year (corresponding to SNCF's schedule-planning periods). In developing the base schedule it considers both profitability and feasibility. RailPlus has a modularized design, which allows the schedule analyst to move easily between market-driven decision making and constraint-driven decision making at any stage in developing the base schedule. The main modules in RailPlus are the work-set manager, profitability module, feasibility module, capacity-allocation module, and routing module (Figure 5).

The RailPlus Work-Set Manager

In the RailPlus system, we created the work-set concept as a mechanism for organizing and managing groups of data that analysts use to develop their schedules. A *work set* consists of a version (or scenario) of a train schedule, a set of model parameters and output data generated by executing the model. Each analyst owns an independent set of work sets. Analysts manage work sets so that they can develop schedules concurrently for different periods. Or for a given period, analysts can evaluate different scenarios at the same time.

The work-set manager module in Rail-

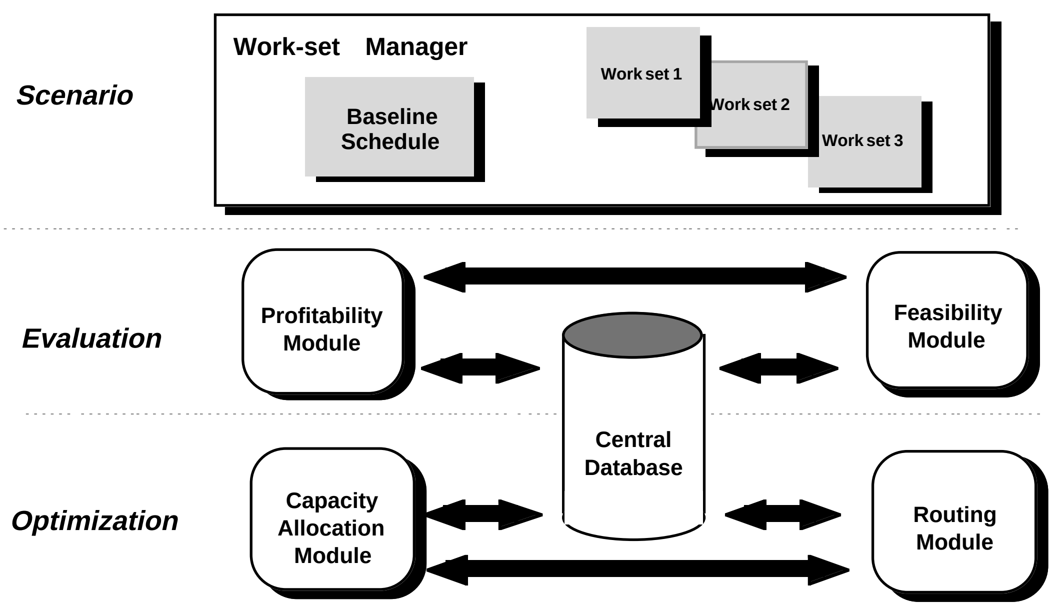

Figure 5: This is a functional diagram of the RailPlus process. The work-set manager tracks the changes made by the profitability and feasibility modules during the evaluation phase. The work set is then passed to the capacity-allocation module, which allocates fleet sizes to each train. Finally, the schedule is routed to conform with maintenance constraints.

Plus fully automates the scenario-creation process for analysts. It reduces the time required to prepare scenarios, allowing analysts to evaluate more scenarios than they could manually. It supports the operations typically found in file-management systems.

The RailPlus Profitability Module

The profitability module evaluates the demand, traffic, revenue, and operating costs associated with a schedule. We designed it to be used interactively and iteratively by the analyst to develop profitable schedules. The analyst can perform what-if profitability analyses by making schedule changes (via the work-set manager). The profitability module incrementally evaluates the estimated train-by-train changes in profits resulting from such schedule changes. It computes the demand, traffic, and revenue data at the finest level of detail possible (for OD, train, date, and class of service). It computes cost data at the train level. To estimate, it goes through the following steps:

(1) It evaluates market size.
(2) It allocates daily aggregate demand to trains (market share).
(3) It projects on-board train traffic through a spill and recapture analysis.
(4) It estimates total train-operating costs.

It estimates OD traffic on each train as OD demand less OD spill plus OD recapture. It computes OD revenue on each train as OD traffic multiplied by average revenue per passenger. It calculates train demand, spill, recapture, and operating costs for each day of the week.

SNCF forecasts daily aggregate demand for every OD served in the TGV system. Using estimates of elasticity for price, travel time, and frequency (the days of the week during which a train number operates), it can adjust daily aggregate demand if edits to the work-set schedule cause average price, travel time, or frequency to change. (For example, a general increase in OD fares would tend to decrease OD demand.)

Next, for each OD, RailPlus allocates aggregate daily demand among the trains that serve the OD using estimates of demand share by train. It estimates demand share based on passenger-preference theory. Specifically, we used a multinomial-logit choice model [Ben-Akiva and Lerman 1985] to estimate the following parameters:

—Passenger preference for a given departure time,
—The sensitivity of passengers to displacement time (the difference between preferred and actual train departure time),
—Price sensitivity of passengers, and
—Travel-time sensitivity of passengers.

The model predicts the mean demand share of a train in a given OD, given the departure time, price, and travel time of all trains serving the OD. We calibrated the parameters of the model using nonlinear regression, basing them on untruncated historical traffic data. The model groups the parameters by type of OD market (short vs. long haul) and day of week.

Spill is the number of passengers that are denied passage on a train because of a lack of seats. RailPlus estimates spill by OD for each train in the schedule using the method described earlier.

Recapture is the portion of demand spilled from one train that is accommodated on another train that has sufficient capacity. RailPlus estimates recapture by OD for each train in the schedule. It computes expected recapture by applying the same multinomial-logit choice model used to estimate demand share on a limited set of trains, in particular, on only those trains with remaining capacity.

RailPlus calculates the total operating cost for each train, including the same list of train-operating costs as RailCap.

The RailPlus Feasibility Module

The RailPlus feasibility module performs a series of checks on a work-set schedule and flags any violations of schedule imbalance, train set availability, and station storage constraints. Analysts use the feasibility module early in schedule development to analyze the operational feasibility of the schedules they have constructed. We designed the feasibility module to be used in an iterative manner. Analysts can perform feasibility checks, edit the work-set schedule based on the results, and rerun the checks to see how the schedule changes affect feasibility. In addition, for schedules that are already operationally feasible, the feasibility module provides detailed reports on schedule activity at the fleet and station levels. Analysts can use graphical displays showing station and track to identify opportunities to increase fleet utilization and to resolve violations.

The RailPlus Capacity-Allocation Module

The RailPlus capacity-allocation module integrates the profitability and feasibility aspects of constructing the schedule. Given a list of train frequencies in a schedule, it selects the set of train-frequency capacity patterns that maximizes the total operating profit of the schedule.

Analysts use a range of parameters to control the operational and marketing constraints considered in an execution of the capacity-allocation module. Because analysts can tailor the optimization problem, the capacity-allocation module can support many types of short-term and long-term strategic studies.

We represented the capacity-allocation problem as an integer multicommodity network flow problem with operational side constraints. The network flow problem is very similar to that used in RailCap (Table 1).

The RailPlus Routing Module

The path that an actual train set follows as it moves from one train-unit assignment to another over the course of a schedule cycle (one week) is called a *routing.* A routing can be described as a sequential set of *turns.* A turn occurs when a train set has completed its assignment to one train unit and is assigned to a new train unit, and the amount of time a train set spends between assignments is called the *turn time.* It follows that the number of routings for any given week is equivalent to the number of train sets required to cover the schedule. The routing module builds routings so that

—The number of routings is minimized,
—Each train set passes through a station with a maintenance center when routine maintenance checks are required,
—Directional connection rules are followed at stations, and
—The set of routings covers all the trains in the schedule.

Constructing train-set routings is one of

Function	RailCap	RailPlus
Schedule input	Uses a dated schedule. Station time-lines run from the first date to the last date in the run window, typically covering departure in the next three to 15 days.	Uses a base schedule by day of week. Station time-lines run from Monday to Sunday, representing a typical weekly schedule.
Capacity pattern generation	Because the problem is capacity adjustment, the first train unit is constrained to cover the entire train itinerary and retain its original fleet assignment.	The RailCap restrictions on the first train unit are relaxed, and the module generates more capacity patterns than in RailCap.
Capacity pattern valuation	Short-term train-OD demand forecasts come from RailRev. Analysts provide recapture rates.	RailPlus uses a passenger preference model to generate train OD demand forecasts, ignoring recapture effects.
Network structure	The structure is not cyclical. The network represents train-set flow over a rolling 12-day horizon.	The structure is cyclical. The network represents train-set flow that is repeatable from one week to another.
Side constraints		Additional constraints in RailPlus: —Set a minimum train frequency on specific ODs. —Set user-defined fleet types on specific trains.
User parameters to control operational and marketing constraints		Additional parameters in RailPlus: —Allow trains to be dropped from the schedule or constrain all trains to have some capacity assigned. —Allow legs to be dropped from the train itinerary. —Provide fleet-availability data at the work-set-level. —Provide exceptional-turns data at the work-set level. Turn on or off any of the operational or marketing constraints.

Table 1: This is a summary of the main differences between the capacity-adjustment problem in RailCap and the capacity-allocation problem in RailPlus.

the final checks analysts perform in developing the schedule. Maintenance checks and the directional connection rules are the critical constraints when determining the number of train sets needed to cover a schedule. Within these constraints, the number of routings needed to cover a schedule is directly tied to the train-set turn times. It is therefore critical to construct train-set routings that minimize turn time due to maintenance checks and station maneuvering.

The maintenance checks for the TGV train sets are both distance based and time based with respect to when maintenance is due. They also vary with respect to the downtime a train set spends in maintenance. However, it is the directional-connection conditions that bring additional complexity to the problem and constitute the fundamental difference between the rail and the airline environment.

The directional-connection rules reflect the conditions that exist in handling multiunit trains. These rules establish where a particular train set can be located (front or back) in a multiunit train and what couplings are possible at hub stations in assigning inbound train sets to outbound trains. These rules depend on several factors.

—Stations have different configurations. After a train enters a station, it can depart one of two ways. If it continues in the same direction to the next stop, the station configuration is called a *through.* If it reverses direction upon departure (as when the tracks dead end into the station), the station configuration is called a *reverse.*

—The position of the maintenance center with respect to the train-set order of an inbound multi-unit train will affect its handling. Can the train set designated for maintenance be decoupled from the train and shuttled to the maintenance base without having to maneuver the other train set out of the way?

—The order (front or back) in which train sets in a multiunit train arrive at a station will affect their handling. If the station configuration is a reverse, only the back train set can be coupled to an outbound train. If the station configuration is a through, only the front train set can be coupled to an outbound train.

In all, we recorded more than 60 connection rules, which must be checked to ensure feasibility.

The routing problem can be stated as "find the minimal set of feasible routings that covers once and only once every train unit in the weekly base schedule." We formulated the problem as a set-partitioning problem [Vance et al. 1997]. The solution method uses the concept of column generation to implicitly evaluate all possible routings and selects the optimal set that covers the schedule. The column-generation process iterates between a master problem and a subproblem. The master problem is a set-partitioning problem with multiple side constraints that enforce the conditional connection rules. We formulated the subproblem as a shortest-path problem with the time-based and distance-based maintenance constraints [Ahuja, Magnanti, and Orlin 1993]. We constructed a train-unit time-space network, where nodes represent train units and arcs represent connections. The arc costs are derived from the master-problem dual variables. The iterative movement between the subproblem and the master problem ends when all dual variables are negative. We then find the optimal solution to the LP relaxation of the problem. In general, such a solution is not integer.

The iterative procedure to generate columns is embedded in a branch-and-bound algorithm. The branching depends on whether a connection (train unit T1 to train unit T2) out of the set of fractional routings is to be selected. We then define a new master problem and iterate between the subproblem and the master problem.

The typical-size problem solved covers

around 1,000 train units. Representing the problem as a network requires an equivalent number of nodes and 50 times more arcs. The column-generation technique reduces the number of routings the module investigates to about 5,000. The solution time ranges in minutes—a significant reduction from the work weeks it took the analyst to manually construct the routings. The automation of routing construction offers unparalleled what-if capabilities that permit planners to evaluate changes to schedules and to evaluate the economic effects of maintenance rules, locations, and turn times. The capacity of the mathematical model and its related solution to incorporate complex connection and maintenance rules brought a lot of credibility to the automation process, which SNCF routing analysts had initially viewed with skepticism.

Benefits of RailCap

The COTGV considers RailCap's capacity-allocation optimizer to be responsible for an increase of 10 million francs in revenue per year. This corresponds to a two percent increase in revenue due to the optimization, based on the 500 million francs annual revenue associated with the capacity-adjustment activity at SNCF. RailCap also decreased costs by three percent. Adjusting capacity prior to RailCap was a reactive and local process. As demand is usually unidirectional on a peak day, SNCF needs empty shuttles from or to a hub station to position train sets at stations where added capacity is to originate or terminate. RailCap's proactive and global approach resulted in a dramatic decrease in cost and better utilization of train sets, allowing SNCF to make more adjustments and gain higher profits.

RailCap has also increased customer satisfaction, since one of its primary objectives is to better fit capacity to demand. With RailCap, SNCF can accommodate three percent more passengers in their first choice. In the long term, this will help the company to retain its market share.

Benefits of RailPlus

In 1995, SNCF conducted a study on the effects of the RailPlus capacity-allocation module on schedule profitability. It determined that the capacity-allocation process alone increases revenue by one percent and reduces operating costs by two percent relative to a manually prepared base schedule [SNCF 1995]. The one percent increase in total TGV revenues translates to 100 million francs per year.

RailPlus results in better quality service for travelers. It helps SNCF to design its schedule product (frequency, departure time, and itinerary) to best match customers' preferences. It allocates capacity to best fit demand and minimize loss of passengers. SNCF estimates that by using RailPlus it has increased total traffic by one percent and first-choice passenger accommodation by about three percent.

RailPlus has dramatically increased analyst productivity by automating scheduling functions and, in particular, applying optimization techniques in areas where the volume of options to evaluate is combinatorial.

The integration of feasibility and market-driven modules within one tool, RailPlus, has truly facilitated communication between marketers and operational analysts, improving their coordination. SNCF's adoption of RailPlus reduced the

Problem	Modules	Formulation	Solution Techniques
Adjusting and allocating capacity	Capacity allocation module in RailPlus RailCap	Multicommodity network flow with side constraints	Branch-and-bound algorithm plus fine tuning of algorithm parameters
Estimating traffic and spill	Profitability and capacity allocation modules in RailPlus RailCap	Nonlinear stochastic optimization problem	Lagrangian relaxation and subgradient algorithm
Estimating market size and market share	Profitability and capacity allocation modules in RailPlus	Demand elasticity model Multinomial-logit choice model	Definition of attributes and calibration using regression analysis
Building routings for a base schedule	Routing module in RailPlus	Set partitioning	Column generation (shortest path with resource constraints) imbedded in a branch-and-bound algorithm
		Bipartite assignment	Assignment algorithm (CPLEX Library)

Table 2: Operations research contributed greatly to the development of RailCap and RailPlus.

number of iterations in schedule production. In the past, the large number of required iterations was time consuming and forced analysts to limit the number of changes to the schedule from one year to the next.

RailPlus helps scheduling analysts to make consistent and quantifiable assessments of schedule scenarios in terms of revenues and costs and of resource utilization.

Marketers can use RailPlus as a powerful negotiation tool to demonstrate when certain operational procedures might harm profits. The capacity-allocation module permits analysts to change the parameters of the constraints very easily to analyze their impact.

Problem	Modules	Solution Techniques	Average Size	Average Run Time
Allocating capacity	Capacity allocation module in RailPlus	Branch-and-bound algorithm (CPLEX) plus fine tuning of algorithm parameters	15,000 rows 35,000 columns	60 minutes
Adjusting capacity	RailCap	Branch-and-bound algorithm (OSL) plus fine tuning of algorithm parameters	30,000 rows 50,000 columns	80 minutes
Building routings for a base schedule	Routing module in RailPlus	Column generation (shortest path with resource constraints) imbedded in a branch-and-bound algorithm Assignment algorithm	1,000 nodes 50,000 arcs	10 minutes

Table 3: The optimization models in RailCap and RailPlus solved very large problems quickly.

The Use of Operations Research

The importance of operations research techniques in the RailPlus and RailCap projects can be measured by estimating the amount of OR-related effort. The tasks related to the OR modules (including design, development, testing, and implementation) represented 65 percent of the total work hours. A team of experienced OR analysts with backgrounds in optimization and statistics participated in this effort. They used a number of OR models and algorithms in formulating the business problems (Table 2).

The problems varied in size and corresponding run times (Table 3). We made a special effort to tune the algorithms, and in some cases the parameters of the optimization software, to ensure reasonable computational time. This is particularly true for the daily run of RailCap, which is considered time-critical.

References

Ahuja, R. K.; Magnanti, T. L.; and Orlin, J. B. 1993, *Network Flows: Theory, Algorithms, and Applications*, Prentice-Hall, Englewood Cliffs, New Jersey.

Ben-Akiva, M. and Lerman, S. R. 1985, *Discrete Choice Analysis: Theory and Applications to Travel Demand*, MIT Press, Cambridge, Massachusetts.

Fisher, M. L. 1985, "An applications oriented guide to Lagrangian relaxation," *Interfaces*, Vol. 15, No. 2, pp. 10–21.

Held, M. H.; Wolfe, P.; and Crowder, H. D. 1974, "Validation of subgradient optimization," *Mathematical Programming*, Vol. 6, No. 1, pp. 62–88.

SABRE Technology Solutions 1987, "Spill analysis—Phase I final report," Internal document.

SABRE Technology Solutions 1989, "Recapture model development final report," Internal document.

SABRE Technology Solutions 1991, "The hub-and-spoke traffic estimation model," Internal document.

Societé Nationale des Chemins de Fer Français 1995, "Rapport d'activité," Internal document.

Vance, P. H.; Barnhart, C.; Johnson, E. L.; and Nemhauser, G. L. 1997, "Airline crew scheduling: A new formulation and decomposition algorithm," *Operations Research*, Vol. 49, No. 2, pp. 188–200.

Thierry Mignauw, Senior Vice President, SNCF Grandes Lignes, Paris, France, writes: "SNCF is facing growing challenges that require us to enhance our competitiveness, reduce our costs and increase our revenues. We are seeking to move from a constraint-driven environment to a market-driven one. Our main objective is to capture and retain market share by providing the most adequate service to our customers. We would like to attain this goal while ensuring the best use of our resources.

"The RailPlus schedule planning system allows us to build the right schedule for every season. It enables our schedulers, through quantitative analyses, to generate profitable schedules. RailPlus translates very well our effort to bring more synergy between the technical and commercial analysts. It integrates the feasibility and profitability aspects in one system. RailPlus also assists us in our attempt to rationalize our decisions to open a new service, add frequency, alter fares or assign capacity. It helps us make these decisions while measuring the impact on a global rather than local scale.

"As an example, the RailPlus system has been used in 1996 to design our new schedule for the TGV Méditerannée (planned for year 1999–2000). It has also helped evaluate both the technical and profit viability of a shuttle service on our busiest markets, such as Paris-Lyon and Paris-Lille. It is currently assisting analysts from all business units to identify low-demand trains for which no compulsory reservation/yield management is needed. The tool is also helping analysts from the North Business Unit to redefine their fare structure. The powerful what-if capabilities of the tool have considerably increased our ability to perform scenario analyses.

"The RailCap capacity adjustment system, in combination with the Yield Management system, are the core applications at our Centre d'Opérations des TGV (COTGV). The COTGV was created four years ago around these tools and *groups* operational and marketing tacticians. The analysts are in charge of adapting the schedule and optimizing revenues in view of daily updated demand and revenue forecasts. The volume of revenue optimized by RailCap is in excess of 100 million francs a year. RailCap's capacity allocation optimizer is credited with nearly five percent of the total revenues. As a matter of fact, we are planning to extend RailCap's utilization in the future to perform the scheduling of the peak periods.

"Upon Mr. Poinssot's arrival to Grandes Lignes a couple of years ago, the senior management carefully reviewed the pertinence of all the applications that were under development or in the process of being implemented. We retained the ones that were proven to be strategic for our long-term business goals. RailPlus and RailCap were among the leading applications. We expect these applications, in conjunction with Yield Management, to provide us a steady three to five percent increase in profit for the next five years. These tens of millions of francs are to be added to the gain in productivity and the improvement in the quality of service brought by the use of these decision support systems.

"Both the end users and myself are very excited about the opportunity that is offered to the SNCF scheduling applications to be recognized by the Operations Research/Management Science community. We feel that the recognition will go beyond the SDT recipients, to reward a rail operator, i.e., SNCF which is seeking to play a visionary role in one of the most conventional industries."