Lancaster University Management School
Working Paper
2010/035

Railway Revenue Management: Overview and Models

Alexander Armstrong and Joern Meissner

The Department of Management Science
Lancaster University Management School
Lancaster LA1 4YX
UK

The LUMS Working Papers series can be accessed at http://www.lums.lancs.ac.uk/publications/
LUMS home page: http://www.lums.lancs.ac.uk/

Railway Revenue Management: Overview and Models

Alexander Armstrong and Joern Meissner

Department of Management Science, Lancaster University Management School, United Kingdom.
a.r.armstrong@lancs.ac.uk, joe@meiss.com

July 26, 2010

Abstract

The railway industry offers similar revenue management opportunities to those found in the airline industry. The railway industry caters for the delivery and management of cargo as well as the transport of passengers. Unlike the airline industry, the railway industry has seen relatively little attention to revenue management problems.

We provide an overview of the published literature for both passenger and freight rail revenue management. We include a summary of the some the available models and include some possible extensions. From the existing literature and talks with industry, it is clear that that there is room to exploit revenue management techniques in the railway industry, an industry that has revenues of $60 billion in the US and promises huge growth in Europe in the forthcoming years.

Keywords: railway, revenue management, pricing, passenger, freight, models

1 Introduction

During the past 40 years there have been numerous advances in the field of revenue management. Though the field is now well-established, academic researchers have neglected to investigate certain industries. While the airline and hotel industries have received their fair share of attention, the passenger rail and freight rail industries have been overlooked. There is still little published research on these industries.

The lack of attention on the part of academic researchers is inexplicable. Consider the following. The rail industry boasts large revenues: \$60billion for freight rail in the US (Association of American Railroads 2009), Amtrak saw \$2billion in 2009 (Amtrak 2009) and combined revenues for all passenger rail operators in the UK for 2009 were in excess of £6billion (Office of Rail Regulation 2009).

In addition, rail is a *green* alternative to other modes of transport. A number of countries, including the UK, are making a big effort to reduce carbon emissions. Passenger and freight rail traffic will almost certainly increase massively over the next decade. The number of kilometres travelled by passenger rail has been increasing year-on-year in the UK (Office of Rail Regulation 2009). In the US, where the rail network is less developed than Europe and Asia, \$8billion has been made available under the American Recovery and Reinvestment Act of 2009 for the purpose of rebuilding high speed rail links throughout the country.

In this article we provide an overview of the literature on railroad passenger revenue management (RPRM) and railroad freight revenue management (RFRM). The objective of this paper is to aggregate the available work and to present a list of the available models along with the problems faced by the rail industry. The paper does not provide a general overview of revenue management – the interested reader should refer to McGill and van Ryzin (1999) and Chiang et al. (2007). Similarly, there is some overlap between the application of revenue management and operations research to rail freight, detailed overviews can be found in Macharis and Bontekoning (2004) and Bontekoning et al. (2004). Rail freight car scheduling is the most closely related application, a survey of such methods can be found in Cordeau et al. (1998).

This paper is structured in the following way. We first discuss the similarities and differences between the passenger and freight rail industries. We then present a summary of all the published work related to RFRM and RPRM. The literature is then discussed §3 followed some minor extensions to passenger rail models in §4. Finally, in §5 we present future research opportunities and in §6 we conclude the document.

2 Passenger and Freight Rail

Passenger and freight services vary country to country. Obviously, they all serve to transport either cargo or passengers from one place to another along some route. For clarity, let's define a service as a train travelling from an origin to a destination at a specific time. Additionally, a service may have intermediate terminals (or railway stations) where *something* can be loaded or unloaded. The service travels along some route, where a route is made up of at least one leg, where a leg is defined as two adjacent stations.

In both industries, the goal of revenue management goal is to find the optimal max of passengers or cargo travelling along each leg in order to maximise the overall revenue. For passenger services, this can be achieved by the pricing of tickets a specific way or by the limiting of the availability of tickets to passengers. For freight services, the problem is more complex in organisation. The remainder of this section highlights why.

The (obvious) key difference between the two industries is that they carry very different items. This is significant because freight has the additional operational burden of loading and unloading the cargo. Additionally, the cargo needs to be managed around these situations. This leads to a variety of complex issues pertaining to the way each carriage (or truck) is loaded so as to reduce the pickup and setoff times at intermediate terminals. The normal strategy is to form a *block* of carriages that share a common origin and destination thus simplifying the operational procedures at the terminals.

In freight rail, it is possible to annul and/or consolidate services should it suit the service operator. Typically, this is not possible in passenger rail: time-tabled services are only cancelled under exceptional circumstances. It is also possible to add or remove carriages, which provide rail freight operators with variable capacity and more freedom to match demand. Conversely, in the passenger rail industry, the number of carriages is only modified in exceptional circumstances. On some routes, passenger services will travel with a very low load factor.

The two industries deal with very different types of customers. Freight rail deals with a small number of customers who can be dealt with at a fairly personal level. There may also be contractual agreements between the operator and customer. On the other hand, passenger rail deals with a large number of customers. Moreover, rail fares can be purchased in a variety of ways. In some countries, passenger services allow customers to travel on any service on the same open ticket without their having to seek authorisation. Naturally, this, coupled with the lack of check-in procedures can make demand estimation a difficult process.

Finally, the booking horizons are significantly different. Passenger rail services can see demand occurring from as far out as three months. For freight rail it can be as little as 0–24 hours (Campbell 1996). Both of these booking profiles provide different challenges for the operator.

In summary, it is clear that passenger rail is more closely related to the airline industry than freight rail. There are to be sure some differences: the services are more frequent, there are more walk-up customers and there are multiple legs that share the same resource. In general, though, the problems are quite similar. The freight industry shares some similarities with the airline industry, but has additional constraints and factors that make the problem of revenue management difficult to solve. The revenue management problems of freight rail are very closely related to those of the car scheduling business. In any case, the common goal for both industries can be defined as which orders to accept/reject and where appropriate, how to handle accepted orders in a way that is beneficial to revenue.

3 Literature

There is little literature on either passenger rail or freight rail. This is probably due to the limited number of such services in the USA, where air travel is the more common form of transport. Ciancimino et al. (1999) came to similar conclusions. Similarly, the use of rail for freight has been on the decline in the USA for years (Association of American Railroads 2010). Despite this, both types of rail services are very common in the UK and mainland Europe. In countries like the UK, the use of rail freight is on the increase (Institution of Mechanical Engineers 2009) and will almost certainly continue to grow as more and more companies strive to reduce their carbon emissions and governments seek to reduce road congestion via modal switch.

Previous RM surveys have identified a few pieces of work related to rail, the common set is typically: Kimes (1989), Strasser (1996) and Ciancimino et al. (1999). Kimes (1989) presents a general approach to revenue management and identifies areas where RM can be applied – it does not explicitly deal with rail. Strasser (1996) is often categorised as rail passenger revenue management, this is incorrect as it deals with rail freight revenue management. Ciancimino et al. (1999) appears to be the first published piece of work to deal with revenue management for passenger rail services and Campbell and Morlok (1994) and Campbell (1996) to be amongst the first that deals directly with revenue management for freight services. Table 1 provides a chronological listing of literature that deals with revenue management for rail.

The following subsections provide an overview of the literature for both fields.

Sector	Count	Reference
Freight	12	Allman (1972), Strasser (1992), Nozick (1992), Kwon et al. (1993), Campbell and Morlok (1994), Kwon (1994), Strasser (1996), Campbell (1996), Kraft (1998), Kraft et al. (2000), van Slyke and Young (2000), Kraft (2002)
Passenger	6	Ciancimino et al. (1999), Kraft et al. (2000), Hood (2000), Sibdari et al. (2008), Bharill and Rangaraj (2008), You (2008)

Table 1: Summary of Railway Revenue Management Literature

3.1 Rail Freight Revenue Management

Rail freight can be categorised into three types: intermodal, general carload and unit (or bulk flow) train. Intermodal freight transport is the utilisation of a variety of modes of transport (e.g. truck, rail or ship) to move a container without handling the cargo within. General carload freight can be considered as anything that is not transported by container, but in some form of railcar, for example: a hopper, tank car or box car. Intermodal and carload freight have many differences, but can be simplistically seen as cars carrying shipments from many origins to many destinations. Unit trains are generally made up of the same type of car and travel from one point of origin to a single destination. Campbell and Morlok (1994) cites that revenue management techniques primarily apply to intermodal and carload freight (with emphasis on intermodal) and that it is less applicable to unit trains due to the nature of the business, that is, high volume customers who have service and capacity fixed under a contract.

The generalisation between intermodal and carload freight allows models to be discussed as a single topic – generally, terminals can handle cars and containers with little trouble. Campbell's thesis and Campbell and Morlok (1994) investigate the applicability of revenue management techniques to intermodal freight. He looks at the techniques used in airline revenue management and derives a series of changes in the models to reflect the different goals of intermodal freight. The applicability of RM to rail freight is tested under the perishable asset revenue management (PARM) model (Weatherford and Bodily 1992). The variable capacity nature of freight is noted, but passes the PARM model under the inclusion of additional logistical complexities. A general model coined the "Periodic Train Capacity Allocation (independent periods)" model is developed in Campbell and can be characterised by:

- T available train departure times denoted $t = 1, ..., T$.

- A network of m legs where each leg is constrained by the maximum serviceable capacity, o_k, for $k = 1, .., m$.
- The capacity for a train departing at t on leg k, $c_k t$ for $k = 1, ..., m$.
- The set of all allowed blocks is denoted $\mathbb{F}$.
- Each block is denoted by the origin-destination pair (ij).
- The set $I_k = \{(ij) \in \mathbb{F} \mid ij$ includes leg/link k$\}$

We denote $x_{ijt} \in \mathbb{N}$ as the variable used to control the available capacity on block (ij) for departure time t. The variable b_{ijt} denotes any previous bookings for the block (ij). $P_{ijt}(x_{ijt})$ is some function that represents the expected profit for block (ij) for the departure at t.

We then aim to maximise the objective

$$\sum_{(ij)\in\mathbb{F}} \sum_{t} P_{ijt}(x_{ijt}) \tag{1}$$

subject to the constraints

$$\begin{aligned} \sum_{(ij)\in I_k} x_{ijt} &\leq \min\{o_k, c_{kt}\} \ \forall k, t & (2)\\ x_{ijt} &\geq b_{ijt} \ \forall t = 1, ...T, (ij) \in \mathbb{F}. & (3) \end{aligned}$$

Those familiar with Campbell's model will notice that we have dropped the multiple containers per car constraint. This can be trivially reintroduced by constraining taking $x_{ijt} \in s\mathbb{N}$, where s is the number of slots per car. The model can be reduced to a steady state formulation by assuming that demand is stationary, dropping the time subscripts and thus fixing capacity for each leg over the entire planning period.

While Campbell and Morlok were perhaps the first to explicitly talk about yield management, prior work had exploited revenue management techniques in the form of service differentiation and allocation strategies. Allman (1972) developed a linear programming formulation to maximise profit given some configuration of freight cars. The significance of the work is that it demonstrates that different car configurations can have substantial impact on profit.

Strasser (1992) looks at the effects of scheduling decisions on performance and revenue via simulation. Her work looks into the suggestions of two unnamed railroads in the US and evaluates these suggestions via a simulation model. She concludes that her model provides an economical way for yards to investigate scheduling decisions without the need for historical data. It is not clear of the applicability of the model to the *real world* due to difficulty in estimating the required parameters, but for those lacking historical data it provides a starting point to investigate new scheduling/allocation policies. Based on the results of the simulation, Strasser suggests a new scheduling policy for the two railroads that can increase revenue at no further cost to the shipper.

Nozick (1992) develops a framework for analysing strategies available to the railroads to increase the profitability of intermodal services. The aim of Nozick's thesis was to develop a model, a solution technique and a general improved understanding of intermodal services. Her work develops a model of similar structure to a multicommodity network flow model that allows railroads to determine optimal fleet size and the cost to provide different price-service combinations.

Kwon et al. (1993) and Kwon (1994) looked at applying service differentiation to the freight market. Their work cited that shippers are highly sensitive to service reliability in mode and carrier selection and have noticed the inability of the railroad to achieve the same standard of reliability as the trucking industry. Kwon (1994) suggests that the freight market can be divided into two markets defined by service quality and willingness to pay. In their work, they differentiate products by low, medium and high priorities and design three heuristics to test how well the differentiation works. Using simulation they show that trip time and reliability are improved for all three of their heuristics and that there is a clear trade-off between service and cost. They conclude that highly reliable services are not required for all customers and that service differentiation enables the railroad to better cater for their market. Kwon et al. note that further work needs to be performed on a more realistic network before these results are generalised. Kwon (1994)'s thesis provides a more thorough analysis of the work and discusses service differentiation further. It also includes a multi-commodity network flow model formulation.

Strasser (1996) looks at the potential of service differentiation to intermodal services and the problems faced with adopting revenue management practices. In particular, the problems that occur at management level and more generally, at what level RM should be applied; should it be managed locally, making small gains with risk of loss at other yards; or globally, requiring co-operation at yard level. Strasser also includes a literature review, which in summary, suggests that adoption of service differentiation and variable capacity models would help with revenue management with the rail freight industry.

Kraft (1998) concentrated on rail service reliability. The aim of the work was to find a way of increasing the reliability of rail in order to compete with trucks. He develops a novel approach to managing the day-to-day railroad network operations. Kraft develops two models: the 'dynamic car scheduling' model, a deterministic model to handle observed demand; and the 'train segment pricing' model, a stochastic model that implements a bid-price approach in order to determine the acceptance of future orders. Both these models are based on multi-commodity network flow models. These new models can be reduced to a series of small sub-problems. This allows the approach to be scaled to large networks with the aid of parallel (or now, high performance computing). The model is run through a simulation and the results suggest that should the model be fully implemented, an expected improvement of more than 10 points could be achieved in operation ratio. Similar work is also available in Kraft (2002). The train segment pricing model can be characterised in the following manner.

Assume we have a set of K customers, where for each $k \in K$ we have some expected volume of shipping required, denoted d_k. Suppose that we can define our rail network in the following way:

- The set of nodes, N, that defines the entire network in terms of yards at some point in time.
- The subset of nodes, $T_k \subset N$, that represents destination nodes for each customer $k \in K$.
- A single fictitious node, Ω, that represents the common destination for all traffic.
- The set of all train schedules over space and time, $\mathbb{F}$, where $(ij) \in \mathbb{F}$ denotes an origin-destination pair where $i, j \in \mathbb{F}$.
- The set $\mathbb{F}$ also includes pairs of the form $(j, \Omega), j \in T_k \ \forall k \in K$ and define a link.
- The set of leg/segment, S, where a leg $s \in S$ is defined as two adjacent yards and has physical capacity as an integer number of cars, denoted as c_s.

We also define the following notation:

- The subsets, $S_{ij} \subset S$, include all the legs that make up the schedule (ij) for all $(ij) \in \mathbb{F}$.
- The subsets, $I_s \subset \mathbb{F}$, includes all links that utilise the leg s for all $s \in S$.
- The revenue for a customer k shipping over (ij) is denoted r^k_{ij}.

We further define:

- The flow volume over the schedule (ij) for customer $k \in K$ as $x^k_{ij} \geq 0$.
- The probability that customer k accepts a delivery slot that terminates at node $j \in T_k$ as P^k_j.

- An optional user-defined booking limit, b_s for all $s \in S$, that can be greater than or less than c_s and can be used instead to account for demand uncertainty. In either case, we denote the choice of booking/capacity limit as C_s.

Kraft (2002) define P_j^k as the logit function in the form of

$$P_j^k = \frac{\exp\left(\alpha - \frac{\Delta\alpha}{\rho_k}\right)}{1 + \exp\left(\alpha - \frac{\Delta\alpha}{\rho_k}\right)} \tag{4}$$

Where α is a calibration parameter, the Δ parameter corresponds to the difference between delivery slot j and the desired level of service for the customer k and ρ_k should be set to the number of delivery slots beyond the base transit time that provides a 50% acceptance level.

The model itself can then be formulated such that we seek to maximise

$$\sum_{k \in K} \sum_{(ij) \in \mathbb{F}} r_{ij}^k \cdot x_{ij}^k \tag{5}$$

subject to the following constraints

$$\sum_{j \in T_k} \frac{1}{P_j^k} \cdot x_{j\Omega}^k \leq d_k \ \forall k \in K, \tag{6}$$

$$\sum_{(iQ) \in \mathbb{F}} x_{iQ}^k - \sum_{(Qj) \in \mathbb{F}} x_{Qj}^k = 0, \ Q \neq \Omega \text{ and } Q \neq \text{ shipping origin}, \tag{7}$$

$$\sum_{k \in K} \sum_{(ij) \in I_s} x_{ij}^k \leq C_s \ \forall s \in S. \tag{8}$$

van Slyke and Young (2000) formulate the accept/reject booking problem as a time-dependent, finite horizon stochastic knapsack model. They propose that their model extends to freight yield management. However, this is only true for cases where customers cannot cancel their order and the capacity of the vehicle is fixed. In cases where the previous assumptions hold, the model provides a way to allocate capacity to variable sized orders over multiple-legs.

Table 2 provides a summary of what each piece of literature covers. Table 3 provides an overview of what each of the optimisation models offer. Optimisation models that allocate a number of carriages to a specific trip are flagged as capacity allocation models. Models that attempt to optimise over the entire booking horizon are denoted with the booking horizon flag. If a model can handle different levels of service, it is marked with the service differentiation flag. Models are formulated as time-space problems are also flagged.

Reference	PhD Thesis	Summary	Opt Model	Sim Model
Allman (1972)	×	×	√	×
Strasser (1992)	×	×	×	√
Nozick (1992)	√	√	√	×
Kwon et al. (1993)	×	×	×	√
Campbell and Morlok (1994)	×	√	×	×
Kwon (1994)	√	√	√	√
Strasser (1996)	×	√	×	×
Campbell (1996)	√	√	√	×
Kraft (1998)	√	×	√	×
Kraft et al. (2000)	×	√	×	×
van Slyke and Young (2000)	×	×	√	×
Kraft (2002)	×	×	√	×

Summary: provides a discussion about the rail freight problem, previous/related literature and future work, **Opt Model:** includes a revenue optimisation model, **Sim Model:** provides a model to analyse different pricing or differentiation policies

Table 2: Summary of Railway Freight Revenue Management Literature

Reference	CA	SD	BH	Det	Stoch	TS
Allman (1972)	√	×	×	√	×	×
Nozick (1992)	×	√	×	√	×	√
Kwon (1994)	×	√	×	×	√	√
Campbell (1996)	√	√	×	√	√	×
Kraft (1998)	√	×	×	×	√	√
van Slyke and Young (2000)	×	√	√	×	√	×
Kraft (2002)	√	×	×	×	√	√

CA: capacity allocation, **SD:** service differentiation, **BH:** booking horizon, **Det:** deterministic, **Stoch:** stochastic or probabilistic, **TS:** time-space network formulation

Table 3: Summary of Railway Freight Revenue Management Models

3.2 Rail Passenger Revenue Management

Passenger rail services tend to vary country to country and in some cases, between operators within the same country. For the purpose of simplicity, we assume that passenger services all share a common theme. That is, they all transport passengers from an origin to destination along a common set of legs and there is at least one class of travel. Passengers are able to purchase tickets in advance for a finite amount of time and we further assume that the the train operating company (TOC) can select or set different prices for tickets during the booking horizon.

Before we dive into the literature, we briefly touch on the fundamental differences between the airlines and passenger rail.

- There is no check-in procedure on passenger rail services
- Open tickets generally allow passengers to travel on any rail service (without check-in or authorisation to travel)
- Walk up tickets are very common; a large number of passengers purchase their tickets on the day from the station
- A large number of services run at a load factor of less than 100%, thus the overbooking paradigm does not need to be considered
- Passengers are often allowed to stand during train journeys hence increasing capacity beyond the number of seats
- Legs cannot be considered independently as the majority of journeys are composed of multiple adjacent legs

The first three points can make it very difficult to estimate how many passengers actually travel on specific services. In the UK, TOCs perform manual counts in order to determine how many passengers are travelling on the service, but this can be error prone and inaccurate, especially on busy services. The fourth and fifth point allows us to simplify the problem with respect to overbooking and cancellations. That is, in the majority of passenger rail problems we do not need to consider either.

Finally, whilst nested fare classes are somewhat common in the airline industry, their use within passenger rail is mostly dependent on the TOC and the services offered. Most of the existing literature does not consider nested fare classes. Ciancimino et al. (1999) explicitly stated that there is no interest in considering nested fare classes in their model. You (2008) extended the model to incorporate a *bumpable* second fare. Other pieces have considered single-fares over resource differentiable products. Hence, whilst the existing literature does not really consider nested fares, they are actively used within the industry.

In summary, the passenger rail and airline problems are quite closely related, however, the large amount of dependence on the legs within the network along with the differences in ticket structure and travel regulations differentiate the problem. The remainder of this section provides a summary of the existing literature.

The first piece of work directly concentrating on RPRM is Ciancimino et al. (1999). Ciancimino et al. developed a model for a single-fare, multi-leg capacity allocation problem. Here the goal is to allocate a specific quantity of seats to each of the origin-destination pairs in order the maximise the revenue for entire journey. They develop a deterministic and probabilistic approach to solving the problem. The deterministic model is formulated as a linear programming problem, whereas the the probabilistic formulation utilises a truncated normal distribution to model service demand with parameters to be provided by the end user. Ciancimino et al. ran numerical experiments based off a real-world data-set provided by the Italian railway company Ferrovie dello Stato and contrasted the results to an existing first-come-first-served (FCFS) booking policy. They report results for a variety of different cases. For the single step optimisation case where optimisation is performed only at the start of the booking horizon, the deterministic approach yielded approximately a 3% decrease in revenue against the FCFS policy whereas the probabilistic approach gained around about a 1% increase. When both models are optimised repeatedly over a 60 day booking period (5 times, 15 times and daily). They observe that optimising daily on their data set does not result in a significant improvement. Their results clearly show that the greater the number of legs, the larger the potential for revenue improvement. When optimising 5 times at day 60, 45, 10, 4 and 1, the deterministic model saw revenue increases ranging from 0.4% to 6.2% as the number of legs increased. The probabilistic model increased from 0.7% to 6.9%. Similar results are seen when optimising the model 15 times over the booking period with increases from 0.9% to 6.83% for the deterministic problem and 1.16% to 7.46% for the probabilistic model. The Single-Fare, Multi-Leg model developed in their paper can be characterised by:

- A unique fare class.
- A route of m legs, with each leg constrained by a service capacity c_k for $k = 1, ..., m$.
- n origin-destination pairs (ij), where $n \leq \binom{m+1}{2} = \frac{m(m+1)}{2}$.

For each pair (ij), we know

- The rate (cost), t_{ij}.
- The mean value for service demand, μ_{ij}.
- The standard deviation of service demand, σ_{ij}.

We denote x_{ij} as the variable used to control the available capacity for each leg (ij) and

$$\mathbb{F} = \{(ij) \mid (ij) \text{ is a valid O-D pair}\}$$

to denote the set of feasible origin-destination pairs on the route. Let l_{ij} and u_{ij} denote the minimum and maximum capacity for a each pair (ij). Then the following constraints should be satisfied

$$0 \leq l_{ij} \leq x_{ij} \leq u_{ij} \quad \forall (ij) \in \mathbb{F} \tag{9}$$

and

$$\sum_{(ij) \in I_k} x_{ij} \leq c_k \tag{10}$$

where

$$I_k = \{(ij) \in \mathbb{F} \mid ij \text{ includes leg k}\}.$$

Ciancimino et al. note that equations 9 and 10 can be rewritten as

$$l \leq x \leq u \tag{11}$$

and

$$Bx \leq c \tag{12}$$

where $B \in \{0,1\}^{m \times n}$, $c = \{c_k\} \in \mathbb{R}^m$ and $x, l, u \in \mathbb{R}^n$ denote vectors for their respective (ij) components. We denote the revenue for (ij) as $t_{ij} \cdot \min\{x_{ij}, d_{ij}\}$ where d_{ij} is the service demand for (ij). Naturally, it is unlikely that we know d_{ij} so we assume it is a continuous random variable. Hence, let $p(d_{ij})$ denote the probability density function of d_{ij}, then the expected revenue over $\mathbb{F}$ is

$$\sum_{(ij) \in \mathbb{F}} t_{ij} \left(\int_{l_{ij}}^{u_{ij}} y p(y)\, dy + x_{ij} \int_{x_{ij}}^{\infty} p(y)\, dy \right), \tag{13}$$

where $p(d_{ij}) = 0$ for $d_{ij} < l_{ij}$. This problem can be transformed into a deterministic integer programming problem by setting $p(d_{ij}) = \mu_{ij} \;\; \forall (ij) \in \mathbb{F}$ and by adding the additional constraint

$$x_{ij} \leq \mu_{ij} \tag{14}$$

that adjusts the upper-bound in (9) to $\min\{\mu_{ij}, u_{ij}\}$, we would like to maximise our revenue function,

$$\sum_{(ij)\in\mathbb{F}} t_{ij}x_{ij}. \tag{15}$$

A probabilistic formulation can developed by taking d_{ij} to have truncated normal distribution with mean μ_{ij} and standard deviation σ_{ij}.

Kraft et al. (2000) provide a discussion about both RFRM and RPRM. Kraft et al. discuss some of the advantages of applying a bid-price methodology to both fields and the problems that EMSR leg-based approaches face with railroad RM problems. Their work also focuses on the network aspect of railroad RM and talks about problems faced with traffic-mix optimisation, that is, determining optimal combination of origin-destination fares across the route. They discuss some of the successful implementations of RPRM such as the system implemented for SNCF – Soceiété Nationale des Chemins de fer français (Ben-Khedher et al. 1998), but do not mention some of the initial teething problem faced when the system first went active (Mitev 1996). Another *horror* story regarding the implementation of an RM system can be found in Link (2004). Link discusses the issues that Deutsche Bahn AG faced.

Though not strictly revenue management, Hood (2000) developed a choice model to aid with demand estimation in order to improve time tabling and pricing decisions. Hood's *MERLIN: Model to Evaluate Revenue and Loadings for InterCity* work draws upon a number of factors that passengers face when deciding which train to travel on. The model takes the various factors and translates them into a generalised cost which is then fed into a logit model. They claim their results to produce demand estimates similar to the observed demand, however they also cite problems with computation time.

Bharill and Rangaraj (2008) consider a premium segment of Indian Railways, the Rajdhani Express and how revenue management strategies can be applied in order to increase average revenue. The Rajdhani Express is a high speed service that offers connections between the capital (New Delhi) to state capitals and other prominent Indian cities. The service offers three resource differentiable products and a single fare for each of these products. Their work looks into how they can estimate cross-price elasticity of demand for three products. Using these values, they develop a model that allows them to estimate demand subject to changes in the price of the fares and additional costs such as booking cancellation fees. They use their models to analyse existing pricing strategies and to suggest changes to IR. Bharill and Rangaraj suggests that analysis similar to their own on similar train routes over differing demand periods (peak/off-peak etc.) is feasible and can be expected to yield positive results.

Sibdari et al. (2008) develop a series of pricing policies for a multi-product revenue management problem for the Amtrak Auto Train. Auto Train is a service that allows passengers to bring their own vehicles onboard and then ride the train. Passengers ride the train in one of the available types of accommodation ranging from seats to sleepers. The passengers vehicle type and accommodation type are bundled together and sold in a two-stage process: passenger selects vehicle type, then selects accommodation type. A booking is only complete when both parts of the process have taken place. The bundle price is made up by the vehicle type and accommodation type, the price of each product is selected from a choice of four fares and can be changed daily upon manager authorisation. Sibdari et al. take a reduced subset of the available products (capturing the majority of the product offering) and perform a variety of analyses to learn about demand. The analysis of their historical booking data suggests that you cannot learn about future demand based on the past sales. They then develop a model and optimise over it using a variety of methods. Their results show that problem formulated as dynamic program yields the greatest improvement in revenue, 17%-31% over their test case at different points in the booking horizon. A myopic policy and static price heuristic (fixed price for the duration) yields a loss in revenue during the later stages of the booking period whilst generating a 4-5% increase at the start of the booking period.

You (2008) extend the single-fare, multi-leg model presented by Ciancimino et al. (1999) to a two-fare, multi-leg model. Their model has an underlying assumption that passengers on discount fares can be *bumped*. You developed a hybrid optimisation algorithm to solve the model. Their hybrid algorithm first solves a relaxation of the problem using linear programming to locate the so-called *solution generating point*. It then uses particle swarm optimisation to locate feasible and high-quality solutions. As with Ciancimino et al. (1999), the model does not directly deal with multi-stage aspect of the booking process, but again, it is suggested to overcome this by running the algorithm sequentially at different points in the booking horizon. They test their optimisation approach on a series of 60 theoretical test cases of differing complexity and show that their approach outperforms existing optimisation tools. You's two-fare, multi-leg model can be characterised similarly to the single-fare case and we again assume there is m legs, with $m+1$ stations, with $n = \frac{m(m+1)}{2}$ origin-destination pairs. This implies that a passenger can now travel from any station to any future station on the route. We now let $\alpha = 1, \alpha = 2$ denote the standard and discount fares and $t_{\alpha ij}$ denote the rate for fare α between i and j. We assume that discount passengers can be bumped to another train at a cost of h_k for leg $k = 1, ..., m$ with total bumping cost for the origin-destination pair (ij) equal to $\sum_{k=i}^{j} h_k$. We also assume that sales of the full fare for all legs k are constrained by c_k, that is, we cannot serve excess demand for full fare passengers.

Similarly, demand is assumed demand to be independently normally distributed for each fare and origin-destination pair with mean $\mu_{\alpha ij}$, standard deviation $\sigma_{\alpha ij}$ and $p_{\alpha ij}(\cdot)$ denote the density function. Once again, let $x_{\alpha ij}$ denote the capacity decision variable for fare α on the pair (ij) and let

$$R_{\alpha ij}(x_{\alpha ij}) = t_{\alpha ij}\left(\int_0^{x_{\alpha ij}} y p_{\alpha ij}(y)\, dy + x_{\alpha ij}\int_{x_{\alpha ij}}^{\infty} p_{\alpha ij}(y)\, dy\right) \tag{16}$$

denote the potential revenue for fare α over the pair (ij) when the booking limit is $x_{\alpha ij}$. Let $X \in \mathbb{R}^{m(m+1)}$ with $x_{\alpha ij}$ at the $0.5(\alpha - 1)m(m + 1) + \sum_{k=1}^{i-1}(m - k + 1) + (j - 1)$th entry. Denote $R(X)$ to be the total revenue when the booking limits are set as X, then

$$R(X) = \sum_{\alpha=1}^{2}\sum_{i=1}^{m}\sum_{j=i+1}^{m+1} R_{\alpha ij}(x_{\alpha ij}). \tag{17}$$

Let $S_{\alpha ij}(x_{\alpha ij})$ be the expected sales for fare α over (ij) with booking limit $x_{\alpha ij}$ given by

$$S_{\alpha ij}(x_{\alpha ij}) = \int_0^{x_{\alpha ij}} y p_{\alpha ij}(y)\, dy + x_{\alpha ij}\int_{x_{\alpha ij}}^{\infty} p_{\alpha ij}(y)\, dy$$

Observing that seats on leg $k = 1, ..., m$ are consumed by itineraries of form $\{(ij) \mid 1 \leq i \leq k, k + 1 \leq j \leq m + 1\}$ we have sales of leg k for fare α defined as

$$S_{\alpha k}(X) = \sum_{i=1}^{k}\sum_{j=k+1}^{m+1} S_{\alpha ij}(x_{\alpha ij})$$

and the total expected sales for train leg k as $\sum_{\alpha=1}^{2} S_{\alpha k}$. Returning to the assumption that discount passengers can be bumped, we define the expected excess demand for leg k, O_k as

$$O_k = \max\left\{\sum_{\alpha=1}^{2} S_{\alpha k}(X) - c_k, 0\right\} \forall k$$

with total bumping cost given by

$$H(X) = \sum_{k=1}^{L} O_k h_k.$$

Finally, to enforce an upper bound on demand for full fare seats, the constraint

$$S_{1k}(X) \le c_k, \ \forall k$$

should be satisfied whilst maximising the objective

$$R(X) - H(X)$$

with $x_i \ge 0, \ \forall x_i \in X$.

In summary, there is very little literature that attacks RPRM problems. The approach by Ciancimino et al. (1999) makes the assumption that there is a single fare that is resource differentiable, You (2008) makes an assumption about bumping costs – in the UK, rail passengers with pre-booked tickets are only bumped under exceptional circumstances. The approach by Sibdari et al. (2008) could potentially be adapted to cater for the multi-fare nature in the UK, but estimating the distributions for each fare could prove difficult and the model is designed for a single-leg.

Table 4 summarises what each of the models presented in the literature offers. A model is said to be multi-fare if can offer more than one price for the same set of resources at the same time, for example: a standard fare and a discount fare. In the **MF** and **ML** columns, integers indicate the number of fares and legs available to be used in the model, n indicates the model has been generalised. We make the differentiation between fare and products by noting multiple fares are offered for the same shared resource, whereas multiple products utilise different resources, for example: standard and first class. We make this differentiation due to the nature of railway where there may exist a set of n fares over m products that utilise m distinct vehicle resources, all travelling over the same network. Naturally, each fare for each vehicle on each leg can be treated as a single product if necessary, that is, a multi-product, multi-resource problem. Multi-leg denotes that the model handles the sequential network nature that is seen on the railway. Dynamic pricing implies that the price of products is changed during the booking horizon via the optimisation step. We note that capacity allocation can be used to achieve a pseudo dynamic pricing policy: in a shared resource case we can limit the capacity of all but one fare to zero and change the available fare at different time-steps to achieve the desired effect. To satisfy the revenue optimisation criteria, models had to present a methodology to solve the model. The demand estimation criteria implies that the model takes a set of parameters and produces an estimate of the passenger count.

Reference	MF	MP	ML	DP	CA	Opt	DE
Ciancimino et al. (1999)	1	$\times$	n	$\times$	$\surd$	$\surd$	$\times$
Hood (2000)	1	$\surd$	$\times$	$\times$	$\times$	$\times$	$\surd$
Sibdari et al. (2008)	1	$\surd$	$\times$	$\surd$	$\times$	$\surd$	$\surd$
Bharill and Rangaraj (2008)	1	$\surd$	$\times$	$\times$	$\times$	$\surd$	$\surd$
You (2008)	2	$\times$	n	$\times$	$\surd$	$\surd$	$\times$

MF: Multi-Fare, **MP:** Multi-Product, **ML:** Multi-Leg, **DP:** Dynamic Pricing, **CA:** Capacity Allocation, **Opt:** Revenue Optimisation, **DE:** Demand Estimation

Table 4: Summary of Railway Passenger Revenue Management Models

4 Rail Passenger Revenue Management Model

The following subsections show some small generalisations to the models to cater for different cases.

4.1 Multi-Fare, Multi-Leg Model

The previous models can be easily extended to the q fare case. First, we remove the assumption that passengers can be bumped. This assumption is unrealistic for the UK railway where passengers are not routinely bumped. It can be reintroduced by assuming that passengers who pay more to travel receive a larger bumping cost should it occur.

Suppose we have a set of q fares $\{1, \dots, q\}$ and again, let α denote the fare. Then, noting that $I_k = \{(ij) \mid 1 \le i \le k, k+1 \le j \le m+1\}$, then the capacity constraint for each leg can be adjusted to

$$\sum_{\alpha=1}^{q} \sum_{(ij)\in I_k} x_{\alpha ij} \le c_k$$

with $0 \le l_{\alpha ij} \le x_{\alpha ij} \le u_{\alpha ij}$ enforcing any capacity requirements for the fare α on itinerary (ij). Similarly, we can calculate the total revenue generated when the booking limit for each itinerary (ij) at fare α is set to $x_{\alpha ij}$ as

$$\sum_{\alpha=1}^{q} \sum_{(ij)\in\mathbb{F}} t_{\alpha ij} \left(\int_0^{x_{\alpha ij}} y p_{\alpha ij}(y)\, dy + x_{\alpha ij} \int_{x_{\alpha ij}}^{\infty} p_{\alpha ij}(y)\, dy \right)$$

and maximise this quantity in order to maximise our revenue over all the legs and fares in the journey.

The difficulty of this model is obtaining an accurate estimation of $p_{\alpha ij}(\cdot)$ and the underlying assumption that all the fares are independently distributed–something that is unlikely to be true when groups of fares are differentiable only by price.

4.2 Single-Leg, Dynamic Pricing Model

It is possible to reduce the problem by Sibdari et al. (2008) to cater for multiple-fares on a single-leg. For the case over the booking horizon $t = T, ..., 0$, denote $F = \{f_1, ..., f_q\}$ to represent the q available fares, let C represent the capacity of the train.

Over the booking period it is possible to offer a single price for a period of time. Demand for each fare is dependant on the price and time in the booking horizon and denoted $D(f_i, t)$. Assume $D(f_i, t)$ has a Poisson distribution with mean $\lambda(f_i, t)$ where $f_i \in F$.

At time t suppose we have n seats remaining with the maximum expected revenue from day t to 0 denoted $R_t(n)$. On day $t-1$, say we have m seats remaining. We can then calculate $R_t(n)$ as follows:

$$R_t(n) = \max_{f_i \in F} \sum_{m=0}^{n} (f_i(n-m) + R_{t-1}(m)) \times P_t(f_i, m, n) \tag{18}$$

where $R_0(n) = 0$ for all $n > 0$ and $P_t(f_i, m, n)$ is the probability of selling $n - m$ seats at price f_i. The transition probability, P_t is calculated as:

$$P_t(f_i, m, n) = \begin{cases} Pr(D(f_i, t) = n - m) & \text{if } n \geq m \\ 0 & \text{if } n < m. \end{cases} \tag{19}$$

5 Future Work

Given the differences between the two types of rail, we believe there are opportunities for future work in two subsections.

5.1 Freight Rail

Freight rail would yield great benefits from a pricing optimisation system that can learn about demand and adjust its parameters and existing allocations accordingly. The very short period in which demand occurs can make it difficult to plan schedules and as such can lead missed opportunities in revenue. Switching existing processes to an on-line learning model would benefit both car scheduling and revenue decisions. Alternately, a model (similar to that presented in Kraft (2002)) that can create a loop between car scheduling and pricing decisions could also yield potential improvements for revenue, whilst still considering service reliability.

5.2 Passenger Rail

Our suggestions for future work for passenger rail stem from conversations with train operating companies in the UK. Through our talks we discovered that a clear that a better understanding of passenger purchase behaviour is required. Work should be undertaken that seeks to understand:

1. How passengers react to price changes for a single journey,
2. How passengers react to price changes for a group of adjacent journeys, and
3. How to stimulate demand, that is, at what price customers switch to no-purchase or alternate mode of transport.

Or, more generally, efforts should be focused on deriving accurate choice models for passenger rail. It would be hugely advantageous to know the substitution effect for journeys with respect to price and time.

While there is some literature available for passenger rail pricing models, we observed that TOCs favour a more systematic approach to their pricing and favour allocation decisions based on analyst prior knowledge and sets of rules. A more scientific approach to pricing and allocation decisions could yield significant gains in revenue. Thus, we suggest that work is performed to bring passenger rail pricing to the same level that is currently seen in more mature areas of revenue management.

6 Conclusion

In this paper we have listed the available literature and models for both freight and passenger rail revenue management. It is clear that as compared to other areas of revenue management such as airlines and hotels there is comparatively little literature.

Kraft et al. (2000) list a number of cases of revenue management systems having been implemented into the rail passenger industry. However, the lack of literature would imply either that the technology is proprietary or that the systems are based on airline revenue management technology. In agreement with Ciancimino et al. (1999), we also hypothesise that the lack of literature in passenger rail is also down to the low usage of rail in the United States. This hypothesis may also hold true for freight rail, where the industry has been in decline for over a decade.

In closing, despite the lack of literature, it is clear that revenue management techniques can be made use of by the rail industry to solve the complex issues involving pricing.

References

Allman, William P. 1972. An optimization approach to freight car allocation under time-mileage per diem rental rates. *Management Science* **18**(10) 567–574.

Amtrak. 2009. Audited consolidated financial statements - fiscal year 2009. `http://www.amtrak.com/servlet/ContentServer/Page/1241245669222/1237608345018`.

Association of American Railroads. 2009. Railroad statistics (2009). `http://www.aar.org/~/media/AAR/Industry%20Info/Statistics%2020090910.ashx`.

Association of American Railroads. 2010. January 2010 rail time indicator. `http://www.aar.org/NewsAndEvents/RailTimeIndicators.aspx`.

Ben-Khedher, N, J Kintanar, C Queille, W Stripling. 1998. Schedule Optimization at SNCF: From Conception to Day of Departure. *Interfaces* **28**(1) 6–23.

Bharill, R, N Rangaraj. 2008. Revenue Management in Railway Operations: A Study of the Rajdhani Express, Indian Railways Rajdhani Express, Indian Railways. *Transportation Research Part A: Policy and Practice* **42**(9) 1195–1207.

Bontekoning, Y.M, C Macharis, J.J Trip. 2004. Is a new applied transportation research field emerging?-a review of intermodal rail-truck freight transport literature. *Transportation Research Part A* **38** 1–34. doi:10.1016/j.tra.2003.06.001.

Campbell, Keith. 1996. Booking and revenue management for rail intermodal services. Ph.D. thesis, University of Pennsylvania.

Campbell, Keith, Edward K Morlok. 1994. Rail freight service flexibility and yield management. *Transportation Research Forum Proceedings* **2** 529–548.

Chiang, WC, JCH Chen, X Xu. 2007. An Overview of Research on Revenue Management: Current Issues and Future Research. *International Journal of Revenue Management* **1**(1) 97–128.

Ciancimino, A, G Inzerillo, S Lucidi, L Palagi. 1999. A Mathematical Programming Approach for the Solution of the Railway Yield Management Problem. *Transportation Science* **33**(2) 168.

Cordeau, JF, P Toth, D Vigo. 1998. A survey of optimization models for train routing and scheduling. *Transportation Science* **32**(4) 380–402.

Hood, ISA. 2000. *MERLIN: Model to Evaluate Revenue and Loadings for Intercity*. 2nd ed. Yield Management: Strategies for the Service Industries, Thomson London, 98–110.

Institution of Mechanical Engineers. 2009. Rail freight: Getting britain on the right track. `http://www.imeche.org/NR/rdonlyres/3B84627A-8FC9-42FB-83E1-32CCFBDEC664/0/Rail_Freight_Report.PDF`.

Kimes, Sheryl E. 1989. Yield Management: A Tool for Capacity-Contrained Service Firms. *Journal of Operations Management* **8**(4) 349–363.

Kraft, E. 2002. Scheduling railway freight delivery appointments using a bid price approach. *Transportation Research Part A* **36** 145–165.

Kraft, E. R., B.N. Srikar, R.L. Phillips. 2000. Revenue Management in Railroad Applications. *Transportation Quarterly* **54**(1) 157–176.

Kraft, ER. 1998. A reservations-based railway network operations management system. Ph.D. thesis, University of Pennysylvania.

Kwon, Oh Kyoung. 1994. Managing hetrogeneous traffic on rail freight networks incorporating the logistics needs of market segments. Ph.D. thesis, Massachusetts Institute of Technology.

Kwon, Oh Kyoung, Joseph M Sussman, Carl D Martland. 1993. Developing insights on effects of service differentiation in rail freight transportation systems. *Transportation Research Forum Proceeding* **14**(1) 549–573.

Link, Heike. 2004. PEP – A Yield-Management Scheme for Rail Passenger Fares in Germany. *Japan Railway & Transport Review* **38** 1–6.

Macharis, C, Y.M Bontekoning. 2004. Opportunities for or in intermodal freight transport research: A review. *European Journal of Operational Research* **153** 400–416. doi:10.1016/S0377-2217(03)00161-9.

McGill, J, G van Ryzin. 1999. Revenue Management: Research Overview and Prospects. *Transportation Science* **33**(2) 233–256.

Mitev, NN. 1996. More than a failure? The computerized reservation systems at French Railways. *Information Technology & People* **9**(4) 8–19.

Nozick, Linda Karen. 1992. A model of intermodal rail-truck service for operations management, investment planning, and costing. Ph.D. thesis, University of Pennsylvania.

Office of Rail Regulation. 2009. National rail trends. `http://www.rail-reg.gov.uk/server/show/nav.1862`.

Sibdari, S, K Lin, S Chellappan, K Lin. 2008. Multiproduct Revenue Management: An Empirical Study of Auto Train at Amtrak. *Journal of Revenue and Pricing Management* **7**(2) 172—184.

Strasser, Sandra. 1992. The effect of railroad scheduling on shipper modal selection: A simulation. *Journal of Business Logistics* **13**(2) 175–196.

Strasser, Sandra. 1996. The Effect of Yield Management on Railroads. *Transportation Quarterly* **50** 47–56.

van Slyke, Richard, Yi Young. 2000. Finite horizon stochastic knapsacks with applications to yield management. *Operations Research* **48**(1) 155–172.

Weatherford, L, S Bodily. 1992. A taxonomy and research overview of perishable-asset revenue management: Yield management, overbooking, and pricing. *Operations Research* **40**(5) 831–844.

You, P. 2008. An efficient computational approach for railway booking problems. *European Journal of Operational Research* **185**(5) 811–824.