This is the author version published as:


Copyright 2009 IEEE
A Study of Heuristic Approach on Station Track Allocation in Mainline Railways

W.Z. Jia\textsuperscript{1,2}, T.K. Ho\textsuperscript{2}, B.H. Mao\textsuperscript{1}
\textsuperscript{1}Integrated Transport Research Center of China, Beijing Jiaotong University
\textsuperscript{2}Department of Electrical Engineering, Hong Kong Polytechnic University

Abstract

Station track allocation is the critical component in the overall railway timetabling. Because of its intrinsic complexity and lack of modeling on station track layouts and train movement within station, analytical approach to attain optimal solution is not feasible. This study investigates the possibilities of applying a heuristic approach and identifies possible difficulties in practice. It is the first and important step to resolve one of the burning issues in the mainline railway operation in China.

1. Introduction

For high-speed mainline railway operation, service punctuality is the collective indication of the system reliability and service quality. With the ever increasing expectation from the customers nowadays, the railway operators are committed to very demanding performance pledges. For example, delays of over 5 minutes may trigger a full refund in Spain while loss compensation on delayed train services will be mandatory within the EC members from 2010 \cite{1}. Major stations are the hubs for various service lines but they are the primary sources of delays, particularly under the heavy traffic demands. The allocation of tracks in the vicinity of the main stations is thus crucial for quality of service provision and capacity utilization.

There are three main sources of service delays in railway services, unexpected incidents, such as adverse weather; trackside infrastructure and rolling stocks; and service timetables. While the unexpected incidents are out of the hands of the operators and the well-being of the infrastructure is a long-term factor, the timetables and their robustness are essential in providing buffers for certain service disturbances. A timetable usually specifies the arrival and departure times of trains at stations and its demand on the station tracks is implicit. Further scheduling is required to ensure the fulfillment of the arrival and departure times under the given track resources at the stations. As a result, track allocation becomes one of the critical components of the timetabling process for mainline railways.

This paper describes the characteristics of the track allocation problem, followed by the discussions of the possible applications of the heuristic approach to attain the optimal solution to satisfy the timetable requirements and to provide high tolerance of service disturbances.

2. Timetabling and track allocation

2.1. Mainline railway timetabling

Timetabling is a time-consuming and complicated process as it attempts to satisfy the requirements of a vast amount of train service provisions. The overall process consists of 3 levels, demand identification, scheduling and operations, as illustrated in Fig. 1.

The first level investigates the service requirements, in terms of timings, frequencies, routes, destinations and service types, under the given information on track and train availability. It establishes the demands and constraints of the overall scheduling problem. The Scheduling level usually provides a feasible train service timetable which satisfies most, if not all, of the requirements and constraints. The timetable also ensures that adequate train fleet and crews are available to realize the timetable such that reasonable time for turn-over is included. In addition, the timetable inevitably imposes tracks and platforms utilization at stations and the station resources have to be allocated accordingly in order to cope with the demands.

On the Operation level, certain services may suffer from delays and the subsequent services have to be re-scheduled in order to recover from the disturbance. The time buffers embedded in the timetable are then used to provide such flexibility. However, when the
delay becomes excessive, significant disturbances in the services are still unavoidable. The re-adjustment of the timetable may not be as such a large-scale problem as the timetabling, but it requires on-line computation and real-time implementation.

Train-movement conflict identification and resolution is the core issue in the station track allocation problem. It is the common problem for the scheduling of inter-lacing parallel processes. The discussions on the possible introduction of buffer times between two parallel processes in close chorological proximity are not very common. The complication of the track layouts and the traffic volume at the stations only adds to the complexity of the problem. Additional constraints of fitting particular train types to certain tracks and uncertainties of driving behavior further extend the dimensionality of the problem.

Most of the previous works focuses on the stability of the station operation as timetabling is often regarded as a periodic event scheduling problem. The track allocation problem has been dealt with as a 2-stage routing problem [3]. A feasible solution is obtained on an independent set modeling through a fixed-point iteration method. The initial solution is then amended by applying a local search optimization in order to increase the time slot of a chosen route, and hence the stability of the timetable.

Flexibility is introduced in the timetable through proper distribution of time buffers within the timetable [4] so that operational conflict may be resolved when service disturbance arises. Another study looks into the knock-on delay effects through an inter-lacing track layout due to a primary delay [5]; and hence compares the stability of the timetable.

Randomness in operation has recently been considered in the studies of timetabling. In a study of service reliability [6], a feasible timetable is first attained and it is subject to random disturbances on services. The timetable is then evaluated by predefined stability criteria under these disturbed conditions. The timetable is enhanced or compensated (by introducing buffers at identified bottlenecks) accordingly. Another approach is to adopt the timetable stability criteria directly in the formulation of timetables [7]. A random re-start local-search heuristic method has been employed to attain the timetable.

2.4. Discussions

From the previous studies on the station track allocation problem, the objective functions usually include the balance of track utilization, maximization of train assignments to the tracks available, adherence to the timetabled arrival and departure times. From train operation viewpoint, the objectives ensure a safe and conflict-free route for each train to enter and leave
the station. On the other hand, they allow balanced utilization of the station facilities and acceptable workloads on the station crews. However, the actual operation is always subject to unexpected disturbances. It is desirable to incorporate certain flexibility in the station track allocation in order to accommodate the unexpected disturbance within a defined extent and hence keep its impact to a minimum. Such flexibility is referred as the robustness of the allocation.

Indeed, robustness has been implicitly considered in the objective functions in the previous studies, but it is not defined and expressed analytically. While train movement in stations is often represented as a deterministic dynamic process, the stochastic nature of the service disturbance does not necessarily fit well in the formulation of robustness with the traditional models. It should also be noted that a comprehensive mechanism to evaluate station track allocation schemes with appropriate indices or features, such as delay severity and propagation rate upon a single disturbance; and operation bottleneck identification and comparison, is not yet available.

Because of the above difficulties, the models and algorithms in the previous studies only provide a feasible, but not the optimal, track allocation scheme. In other words, given the same station track layout and required arrival and departure times, different feasible track allocation schemes are attained from different approaches of modeling and computation. As they are not optimal with respect to any specific evaluation index, comparison among schemes prior to implementation may not be straightforward and thus further improvement is rather limited. For railway operators, an optimal station track allocation scheme with well-defined evaluation indices is of utmost importance for planning and operation.

3. Methodologies

From the above discussions, the station track allocation problem contains the following concerns.

a) There is a lack of a well-defined and analytical set of criteria for the evaluation of allocation schemes
b) Formulation of the optimal allocation scheme has to satisfy the criteria of both deterministic and stochastic natures
As station track allocation is an off-line scheduling problem, the requirement on computation time is not the most critical consideration. It is possible to seek for the optimal, instead of a feasible, solution by allowing more computation time. An allocation scheme with better quality (with respect to the specified evaluation indices) will certainly pay off, particularly if on-line re-adjustment is required when an unexpected disturbance occurs during operation.

The possible algorithm to attain the optimal track allocation scheme should be able to handle the mixture of criteria for evaluation and their non-analytical relationships. The optimality of the solution can be a trade-off with the computation time available and a near-optimal is accepted if time available is limited. Heuristic approach and evolutionary computing techniques are the potential methodologies to meet the requirements.

The solution algorithm consists of two steps, as shown in Fig. 2. The first step employs heuristics derived from experience to establish an initial feasible solution. From a pre-defined set of evaluation indices (e.g. robustness and facilities utilization etc.), the solution is then improved gradually with evolutionary computing techniques.

Fig. 2 Basic algorithm for track allocation

3.1. Heuristic approach

For engineering applications, the heuristic approach is to provide rules of thumb from experience in order to enable quick feasible solution. The rules may not come with any strict scientific proof but they usually offer a reasonable starting point for the solution. It has been widely used in constrained resource allocation problems and found successful applications [8-10].

From the experience on train movement and station operation, a number of simple rules on track allocation are given below as examples.

a) Direction consistency – trains are running up-track and down-track through the station. It is common to keep one track for one direction only.
b) Initial conflict elimination – there are two types of train movement conflicts, trains across tracks and trains on the same track. Elimination is realized by operating the trains as far apart as possible but the conflict may still remain.

c) Balance workload on tracks – each track is assigned one train service in turn whenever possible in order to maintain a roughly equal workload.

These rules may be gathered from experienced timetable planner and station operators; and they may not be exhaustive.

3.2. Evolutionary computing

Evolutionary Computing (EC) is one of the topical areas of artificial intelligence. It mimics the behavior (or evolution) of humans or biological processes in the search of a solution in a multi-dimensional solution space. When an intermediate solution is attained, it is evaluated and certain parts of its characteristics are then retained for the next round of evolution. The EC techniques have been successfully applied in a wide range of engineering problems, such as pattern recognition, scheduling and robots.

In this track allocation problem, EC techniques are the suitable tools to search for the optimal solution when given an initial feasible solution, through a complex solution space. The evolution allows gradual improvement of the solution optimality.

4. Implementation

Genetic Algorithm (GA), Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO) are the popular techniques in EC as their theoretical development has been very mature and practical applications scatter in every discipline of engineering and management.

Discussions are given here on the suitability of GA and PSO for the implementation of this station track allocation study. It is the definitive step to establish the objectives of the further work and to identify possible difficulties.

4.1 GA and PSO

GA is one the early adopted techniques in EC. It emulates the process of biological evolution by reproducing through generations [11]. PSO has emerged in recent years as a popular technique of EC and it is based on the random food-searching behavior of a flock of birds or animals [12].

In GA, a possible solution is represented by a chromosome, in which the genes are subject to selection, mutation and cross-over during evolutions in order to look for better solution. With PSO, the position of a particle denotes a solution and it looks for a new direction for a better solution through communications with other particles who collectively search for the optimal solution over the solution space. The process is similar to mutation and cross-over operations in GA. In general, the advantages of PSO are its computation simplicity and easy implementation. They have been widely employed in scheduling, parameter optimization, system control and even system training [13-15].

With the similarities between GA and PSO, the discussions here take PSO as an example to illustrate the application of EC techniques in the station track allocation problem.

Each particle in PSO is traveling at a certain speed within the solution space. The particle position is the current solution attained by the particle. The particle re-adjusts its speed and direction according to evaluation of the whole flock during each evolution. The process repeats until a pre-defined condition is met.

In station track allocation, each train is given a track (and its platform) for stopping and a route for entrance and departure. The solution space is represented in terms of the allocation in space on the same track, as well as time. An example of a solution (with 4 tracks and 15 trains), i.e. a particle location, is denoted schematically as in Fig. 3. The track space and time are given in the vertical and horizontal directions respectively. The entrance and departure times are also included as the ‘black triangles’. The exact timings are calculated through train movement simulation [16]. In order words, the particle location is defined by a multi-dimensional vector quantity which encapsulates the time and space information on track allocation. The scale of the vector quantity inevitably depends on the numbers of tracks at the station and trains to be allocated.
Fig. 3 Elements of a possible solution
Each solution has to be evaluated by a fitness function. Robustness, reflecting the capability to absorb service disturbance, is introduced in this study as the necessary performance indicator of a track allocation scheme. The formulation of the robustness index has been discussed in a recent study [3].

4.2 Difficulties

As the entrance and departure routes of the trains usually cover more than one track, the space conflict across tracks is common but it is not yet modeled in the solution space. When a particle updates its location, there are additional constraints on the track allocation. The inter-track space relationship has to be modeled properly and incorporated in the solution space. Graph theory offers the framework to identify and represent the physical inter-relationship of tracks.

5. Conclusions

Station track allocation is a combinatorial optimization problem and it is an urgent real-life problem in the midst of the rapid development and increasing demands in the mainline railways in Mainland China. While traditional analytical methods may not be applicable due to the lack of proven deterministic models on train movement and track layouts, this study investigates the feasibility of a heuristic approach.

An initial solution is first established by a set of rules derived from experience. Evolution computation technique is then applied to improve the solution gradually. Robustness is included as the solution-evaluation index.

It is only the first step of the study. Investigations on the implementation and proper modeling of the track layout are required. As far as EC techniques are concerned, it is also intended to examine the performance of GA and PSO, such as chromosome or particle representation, impact of the quality of the initial generations, computational demands and convergence rate to optimality.

Acknowledgements

The authors would like to express their gratitude for the National Scientific Foundation of China (60634010) and the Hong Kong Polytechnic University for their financial support on the project.

References