AMP: The future of rail traffic management is here

Optrail's Advanced Movement Planner (AMP) is an optimization-based planning tool for managing and optimizing train traffic in real-time. AMP produces scheduling and routing plans for all controlled trains several hours into the future, providing valuable suggestions to traffic controllers and automatically carrying out commands on the field in auto-mode. In November 2021, the rollout of AMP was completed on Union Pacific's network, one of the largest and most complex rail networks in the world. With over 50.000 kms covering 23 states in the USA dispatched using AMP, UP's is by far the largest rail network equipped with a fully automatic traffic management system. This remarkable feat is the fruit of over 15 years of research in mathematical optimization, graph theory and other fields of AI for automatic train scheduling.

Intro

Rail networks are complex systems that allow hundreds or even thousands of trains at any given time to carry passengers or goods between locations. The usage of rail resources (e.g. a track section) is heavily constrained, in the sense that they can generally be occupied by at most one train at a time. Since a common objective is for trains to traverse the network in as little time as possible, in some sense they "compete" for these constrained resources to reach their destination(s). For example, trains are often projected to occupy a track at the same time, like the simple case in Figure 1 in which two trains are "competing" for a track section *t*. Clearly, these trains cannot occupy *t* at the same time. One train must hold at the previous meeting location while the other train drives past.

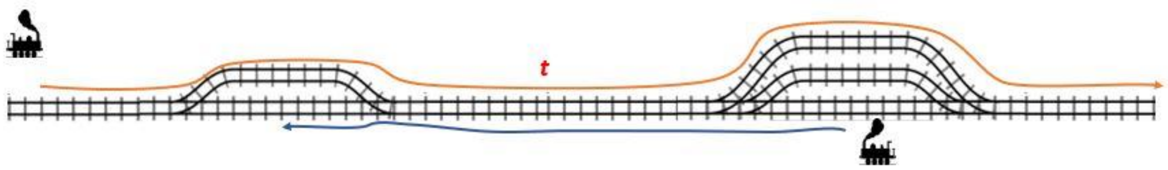


Figure 1: Trains "competing" for track section t

Zooming out to the network level, it becomes clear how questions such as "which train goes first?" or "is an alternative route available?" must be answered over and over by the controllers in charge of the network, the so-called *dispatchers*. These decisions are very challenging to take, because dispatchers have relatively little time to take them (sometimes only a few seconds) and because each decision can, potentially, have a cascading effect on the rest of the network. Factoring in the implications of each decision before taking it is extremely hard, even for experienced and (sometimes supernaturally) skilled dispatchers. What might appear the best decision locally can create unforeseen delays or issues down the line.

What is AMP

This is where Optrail's Advanced Movement Planner (AMP) comes into play. Thanks to its state-of-the-art algorithms, in a few seconds AMP finds the best possible decisions at a global level, factoring in the effect of each decision on the entire network for hours to come.

So to answer the question, in a nutshell AMP is a planning software that harnesses different optimization algorithms to produce real-time scheduling and routing plans for the entire controlled network, providing decision support to dispatchers or carrying out these decisions directly in automatic mode… all of this in real-time!

What are the main advantages of using AMP

By using AMP, rail operators can reduce operating expenses - OPEX - and capital expenditure - CAPEX - while increasing service quality. This is because AMP optimizes the efficiency of the network by improving the KPIs chosen by the operator, e.g. minimising delays or maximising train velocity (OPEX), and increasing line capacity without additional, and typically extremely expensive, infrastructure investments (CAPEX). To put things in perspective, every mile-per-hour increase in velocity can save a railroad hundreds of millions of dollars in capital and operational costs [9]. Similarly, increasing the number of trains on the network without laying additional track can greatly reduce capital expenditures, saving the cost of approximately 2 million dollars per mile of newly laid track. Improved dispatching also translates into greater schedule/timetable adherence, which in turn brings higher service quality both for passenger and freight traffic. Estimated times of arrival consequently become more precise when using AMP, which allows the operator to schedule crews more accurately and provide customers and/or passengers with reliable arrival times.

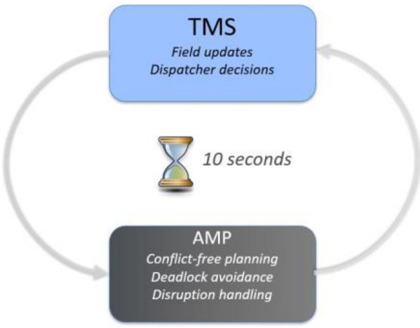
In terms of network and traffic management, AMP greatly aids dispatchers by removing much of their workload and stress. Thanks to AMP's ability to react extremely fast to changes in the field (only a few seconds), along with its "microscopic" modelling of the rail network, AMP plans can even be converted directly into commands that are executable "as-is" and can thus be sent directly to the field. In other words, AMP allows the network operator to fully automate the dispatching process if so desired. A dispatcher's time can thus be freed up for other tasks, while taking control of operations in special situations or should they deem it necessary to do so.

How does AMP work

The first step is for AMP to receive all the pieces of information required to make informed decisions. This is done thanks to a constant stream of up-to-date information regarding the current status of the network, which is achieved by integrating via a communication bus with the Dispatch System (also known as Traffic Management System - TMS), i.e. the command and control software interface to the field. In this way, AMP is provided with real-time updates on events such as train positions, or can factor into its plans any decisions taken by the dispatchers (such as aligning a particular switch or fixing a meet location for a pair of trains).

Once AMP receives this information, it starts its computation and, in as little as 10 seconds, seamlessly returns an updated plan to the TMS which then displays it for the dispatchers. In fully automatic mode, AMP's plans are carried out directly to the field.

A simplified sketch of this architecture is shown in Figure 2.



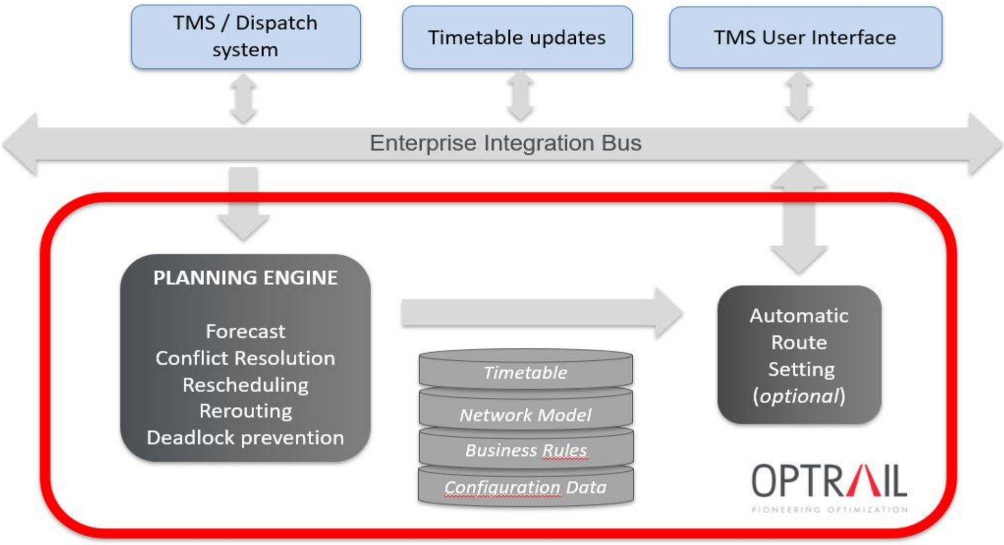


Figure 2: Overview of AMP architecture

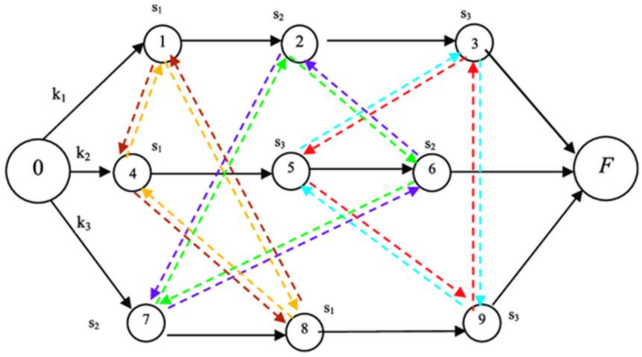
What are AMP's main functionalities

Routing and scheduling trains

AMP's first and foremost functionality is that of producing routing and scheduling plans for the controlled trains that are optimized at the global level. In practice, this means that every AMP plan contains a sequence of movements for each train and a (projected) time associated with each movement. These movements are defined and scheduled down to the granularity of track circuits, both at the main track and station level.[1] Naturally, the operator's goals and business rules are factored in during this process. Thanks to multiple projects and implementations in Europe and North America, AMP has a broad set of configurable, "off-the-shelf" business rules tailored both for passenger and freight train operations. Furthermore, AMP's mathematical (rather than rule-based) core model is highly flexible and allows easy customization for any rail operator's wishes and requirements. AMP's core planning module is based on a dedicated, proprietary routing and scheduling algorithm - the Continuous Rolling Horizon (CRH). In a nutshell, the CRH is a truncated

[1] AMP also supports moving block

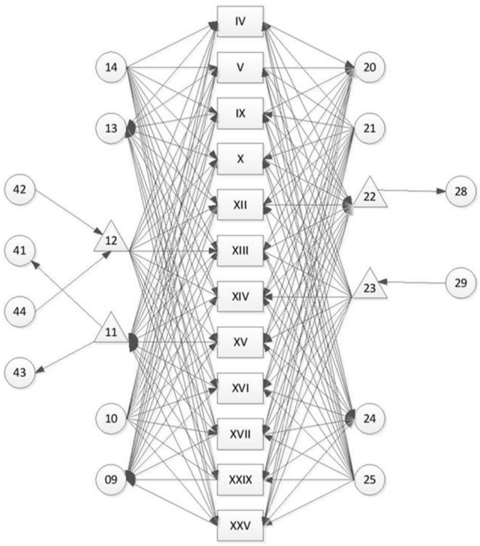
Branch & Bound algorithm in which branching is done on conflict resolution and routing options and bounds are computed very efficiently by exploiting the properties of an associated Directed Acyclic Graph (DAG) to perform dynamic longest path tree calculations. The entire mechanism is embedded in the spatial decomposition described in [6,7].



So far we have described AMP in terms of its main functionality, producing routing and scheduling plans. However it also provides other functionalities that can be equally important in specific contexts.

Station Platform assignment

AMP's dispatching algorithm is not limited to producing schedules/timetables at the main line level, but can also tackle stations of any level of complexity. This includes solving the so-called Platform Assignment problem, where trains must be assigned to platforms that can accommodate them in terms of length (train must fit), type (e.g. high speed, regional), passenger needs (e.g. ramps, elevators) etc. This includes determining the train's entire path (i.e. sequence of routes) from entrance to stopping point to exit even in complex stations. AMP supports all types of train activities, both for passenger trains (e.g. embarking, disembarking passengers) and freight trains (e.g. crew change, fueling, adding/removing cars) and allows users to specify multiple locations for performing an activity. Once this information is made available, AMP chooses the best amongst these locations, factoring current traffic conditions and any user-specified priorities.



Deadlock avoidance and handling

A group of trains are said to be *bound-to-deadlock* if every possible choice of movements ends up in a situation where no train can move because another train is blocking the next track along its route (*deadlock*). Releasing trains from a deadlock typically requires complex switching moves such as

pulling back trains, and, once a deadlock is in place, it may take many hours to resume regular traffic flow. Deadlocks are thus a highly expensive and operationally challenging event to recover from for an operator. For this reason, *deadlock avoidance* is arguably the primary responsibility of a dispatcher. This is particularly the case for dispatchers operating predominantly freight traffic, as is the case in North America. Indeed, in this context freight trains are typically "overlength", that is, trains whose footage exceeds that of most available sidings and station main tracks. This largely restricts the number of locations for trains to meet, because long stretches of the railroad are single-track and trains can meet and pass each other only when sufficiently long, parallel tracks are available.

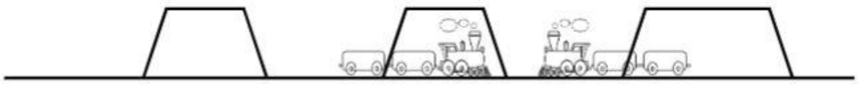


Figure 3: Stylised representation of a deadlock

AMP embeds dedicated algorithms to tackle this challenge. Firstly, a Deadlock Detection algorithm for identifying potential deadlocks in the given planning horizon, inspired by the mathematical optimization approach published in [3] and the work published in [2]. Additionally, AMP includes an algorithm that is dedicated solely to a special case: identifying deadlocks between pairs of (opposing) trains. Statistics show that these are the most common types of deadlocks, as stated in [8]. This “2 train” module embeds a combinatorial algorithm based primarily on Graph Theory, soon to be published [4]. This module is also used to speed-up the CRH’s branching procedure.

On top of this, AMP also includes a module dedicated exclusively to dispatching “overlength” trains, which, as mentioned, are a major deadlock hazard. The need presents itself specifically when the CRH’s “regular” planning horizon (generally set to a few hours) is insufficient to identify a deadlock risk coming from overlength trains that are very far apart. To tackle this issue, this “overlength train module” (the *Cupola*) looks beyond the regular planning horizon and helps the CRH to avoid deadlock-inducing decisions for this subset of critical trains.

Disruption handling and recovery

When a disruption occurs on a railway network, parts of the network may become unavailable for inbound trains, and decisions must be taken to mitigate the impact on overall traffic. This generally requires train movements to be re-planned in a short time, that is, to take appropriate rerouting and rescheduling decisions. Deadlocks are more prone to occur during disruptions, as the capacity of the network is often abruptly reduced and dispatchers face unfamiliar and critical circumstances. In order to prevent them and, in general, to mitigate the negative effects of a disruption, dispatchers are required to implement specific recourse actions, with some key objectives. Firstly, to leave, if possible, a viable path in the network with no obstructions, so as to allow possible work/maintenance trains to reach the disruption site(s). Secondly, to quickly restore the normal traffic regime once the disruption is over. Finally, to ensure continuity in the service offered to passengers, by taking actions such as minimizing cancellations, adding short-turn trains and substituting missing segments with alternative bus transport, etc.

AMP’s routing and scheduling core algorithm, the CRH, seamlessly factors in any track blocks or slowdowns in its planning process and accordingly produces optimized plans. Under more extensive

disruption, AMP’s *Safe Place* module comes into play. Firstly, it determines the subset of trains that are affected by the disruption and cannot be routed around it. Then, it identifies a location where the train can be “safely” parked until the disruption is over, i.e. a location that allows traffic to route around it.[2] The purpose is to avoid trains driving past the last location where they can be routed around, which could otherwise lead to further disruptions and deadlocks.

The algorithm powering the Safe Place module is based on the research published in [2], specialised in [1] and in another paper in preparation [5].

Boundary coordination

The overarching goal in movement planning is to manage and dispatch trains system-wide, from origin to destination. However, due to the enormous complexity of doing so, planners resort to the most natural, and generally successful, technique to tackle large scale problems: *decomposition*. The network is generally partitioned into different areas, each controlled by a regional operating center and/or by a group of dispatchers, where in turn each dispatcher is assigned a smaller portion of this region, which depending on average traffic conditions could be as granular as a line, a station or even more. This *divide et impera* approach allows dispatchers to tame the complexity of the task they face, at least to some extent. On the other hand, it may lead to suboptimal choices from a global perspective. As described in the introduction, each dispatcher is focussed primarily on dispatching her area of competence as well as possible. While there is of course some level of coordination between dispatchers (particularly those that sit in the same operating centre), it is virtually impossible for any human being to assess the impact of each decision on the entire network.

This is instead the kind of task AI-based algorithms tend to be very good at, which in turn motivates the implementation of movement planning technology like AMP. However, the nature of these problems is such that it becomes exponentially harder to solve them as their size grows. This means that, eventually, even the most advanced algorithm possible may struggle. For this reason, AMP follows a similar, decomposition approach in dispatching an arbitrarily large rail network, but does so in a way that ensures that decisions are taken with a global rather than local perspective.

How is this achieved? First of all, the network is divided into *planning regions*, each controlled by an AMP instance in charge of all, and only, the movements within the planning region. As described above this mimics the way dispatchers go by this process, albeit on a larger scale (each AMP instance can cover the territory of up to tens of dispatchers). The major difference lies in how the coordination between these instances occurs. A dedicated coordination module is assigned to each boundary to ensure smooth handover and make sure that decisions are optimized in a way that is beneficial *at the global level*. The partitioning of the network into planning regions can be established by the operator (for example based on existing administrative regions) or computed offline by a dedicated algorithm developed by Optrail, which aims to define regions that are balanced in terms of "planning complexity" while forming boundaries that have suitable properties for the coordination process to take place seamlessly.

[2] Note that the concept of "safe" location has nothing to do with the risk of accidents, which is responsibility of the underlying signaling system

Some key figures and facts

Below are some figures and facts related to the most recent deployment of Optrail's AMP. These numbers are dependent both on the network and traffic conditions and on specific requests made by the customer, however are helpful to give a broad idea.

- **Planning region size.** Each AMP instance on average operates a region between 40 and 100 trains for the given planning horizon[3]
- **Plans.** MP produces a conflict- and deadlock-free plan, including routing, scheduling decisions and timing for all train movements within such planning horizon
- **Computational time.** Timeout for returning a solution can be set as low as 10 seconds. In this way plans are always up-to-date with current field conditions and can be carried out automatically if configured to so by the operator
- **Business rules.** AMP is already compliant with a vast set of North American and European business rules, and new business rules can be easily configured thanks to the flexibility of the underlying routing and scheduling model
- **Objectives.** AMP supports a number of different objectives, which can be tuned based on the operator's business KPIs and preferences
- **Deadlock prevention.** AMP includes dedicated modules for deadlock prevention and handling:
    - Deadlock prevention within the planning horizon for all trains
    - Ad-hoc handling of "overlength" trains handling (extended planning horizon)
    - Assignment of "safe places" under disruptions
- **Boundary coordination.** MP coordinates the different planning regions to avoid inter-area deadlocks and improve the global quality of the plans

Contacts

This article was intended as an introduction to AMP, its purpose and functionalities. For any questions or for a further dive into its workings, please don't hesitate to reach out at leonardo.lamorgese@optrail.com

ABOUT OPTRAIL

Founded in 2014, OptRail is a digital tech company with long-standing experience in technology transfer from scientific research to the industry. Optrail's core competency is the application of quantitative decision sciences such as Operations Research and Machine Learning to real-world problems, with a particular focus on transport

[3] Commonly set to a few hours

and logistics. Its main application area has been the development of intelligent decision-support systems in the field of rail transport (urban and main line). www.optrail.com

References

[1] A. Croella, V. Dal Sasso, L. Lamorgese, C.Mannino, P. Ventura, "Disruption management in railway systems by safe place assignment", *Transportation Science,* to appear

[2] V. Dal Sasso, L. Lamorgese, C.Mannino, A. Onofri, P. Ventura, "The tick formulation for deadlock detection and avoidance in railway traffic control", *Journal of Transport Planning & Management, vol.18, 2021*

[3] L. Lamorgese, C. Mannino, "A non-compact formulation for job-shop scheduling problems in traffic management", *Operations Research*, vol. 67(6), 2019

[4] V. Dal Sasso, L. Lamorgese, C. Mannino, A. Tancredi, P. Ventura, "Easy cases of deadlock detection in train scheduling", submitted to *Operations Research* (2021).

[5] V. Dal Sasso, L. Lamorgese, C. Mannino, A. Onofri, P. Ventura, "The no-cycle formulation for deadlock detection with safe places", working paper (2021)

[6] L. Lamorgese, C. Mannino, M. Piacentini, "Optimal Train Dispatching by Benders'-like reformulation", *Transportation Science*, 50 (3), pp. 910-925, 2016.

[7] L. Lamorgese, C. Mannino, "An exact decomposition approach for the real-time Train Dispatching problem", *Operations Research*, 63 (1), pp. 48-64, 2015

[8] J. Pachl, "Deadlock avoidance in railroad operations simulations", *In 90th Annual Meeting of the Transportation Research Board,* 2011

[9] "The railroad with better profit margins than Google", Fortune 500, source: https://fortune.com/2015/06/04/union-pacific-railroad/