ABSTRACT
This paper presents an event-driven simulation-based optimization method for solving the train timetabling problem to minimize the total trains’ traveling times in the hybrid single and double track railway networks. The simulation approach is well applied for solving the train timetabling problems. In present simulation model, the stations and block sections of the railway network are respectively considered as the nodes and edges of the network model. The developed software named SIMARAIL has the capability of scheduling trains in the large scale networks regards to capacity constraints and infrastructure characteristics. This simulation model for railway timetabling is based on a detailed microscopic infrastructure model, which includes the most detailed infrastructure information. This research is based on integration of a discrete event simulation and GA meta-heuristic algorithm to generate near optimal train timetable. In other words, the simulation model is used to construct a feasible solution for train timetabling problems.

Keywords:
Railway simulation, discrete-event simulation, genetic algorithm, simulation based optimization, timetable generation.

1 Introduction
The passengers demand on railway transportation is expected to significantly increase in future. Hence, scheduling of the trains in order to have the minimal traveling time becomes the challenging optimization problem. Train scheduling is one of the most interesting problems in transportation planning systems. Finding the best solution regards to all constrains increases the complexity of such problems to be solved. In this paper, the scheduling of Iranian National Railway Network is discussed and the approach to solve the problem is presented. The remaining of the paper is as follow. In section 2 a literature review of train timetabling problems and related issues discussed. In section 3, the problem is described in detail. In section 4 the simulation model is presented and the optimization part using GA algorithm is addressed. Section 5 clarifies the model by an example. Finally, the discussion and conclusion are stated in section 6.
2 Literature Review:

Train timetabling is one of the most interesting problems in transportation planning systems. The goal of train timetabling models is to determine the departure time of trains from stations while satisfying a set of operational constraints. A review of analytical railway optimization models is provided by Assad (1980). Caprara et al. (2006) also provided an excellent state-of-the-art review of railway optimization problems.

Several train timetabling problems are known to be NP-hard with respect to the number of conflicts in the schedule (Cai et al., 1998; Caprara et al., 2002). It is difficult to determine optimum solutions to industry sized problems in a reasonable time and this raises the need for efficient heuristic and meta-heuristic algorithms. In the literature, several complex search procedures are introduces such as look-ahead search (Sahin, 1999), backtracking search and meta-heuristics algorithms (Higgins et al., 1996; Liu and Kozan, 2009). In Higgins et al. (1997), an enhanced local search heuristic (LSH), genetic algorithms (GA), tabu search (TS), and two hybrid algorithms (HA1 and HA2) are applied to the train scheduling problem. Babar Khan et al. (2006) present an intelligent search technique to train scheduling problem based on genetic algorithm to minimize delays at the intermediate and final train stations. There has been some earlier work in simulation modeling for train timetabling problem. Petersen and Taylor (1982) present a simulation model for railway scheduling. They partition the railway line into track sections representing the stretches of track between adjacent stations and develop arithmetical formulation to represent the model logic. Dessouky and Leachman (1995) developed a simulation modeling approach for single-track or double-track rail networks by considering single-speed limit without deceleration rates. A hybrid methodology of the event-driven simulation and the network-based simulation methods is proposed in Cheng (1998) for resolving resource conflicts in train rescheduling. Li et al. (2008) presented an advance simulation method based on the global information of the train for solving the train scheduling problem to reduce the total travel time on the single-track railway.

3 Problem Clarification

Generated timetable by the custom software can only pictures the timetable by the initial user decisions and the software could not design an optimal timetable by itself. On the other hand, it was found that because of not customizing for local requirements, the software has so many unsatisfying factors which are unbearable for the National Railway Authority anymore. The major problems were:

- Praying Time: Based on regulations, trains have to stop for praying in time range based on the sun position in day. So there are different praying zones across horizon which a train could stop.
- Track Maintenance: Because of the aged railway tracks, there is need for maintenance daily interval. Hence a train is not allowed to be on the rail during maintenance period.
- Decision Support System: Deciding system to calculate near the best departure time of trains in stations. New system has to deliver a full travelling schedule from the origin to destination and recognize trains stop's stations for each reason such as praying or fuelling etc.

3.1 Model Assumptions

The problem is concerned with a railway network with a set of linked corridors. Railway network consists of single and double tracked routes. A train service is defined as a trip of a train that travels from its origin station to the destination station. Each train is assumed to have a pre-specified traveling route in the network. Furthermore, the traveling times of trains at segments are assumed to be constant. The minimal headway of trains is also defined as a constant. Train services are determined by a set of scheduled stops and unscheduled. There are three category of stopping for trains:

1. Scheduled stops in predetermined stations with predefined stop time
2. Scheduled stops in undetermined stations with predefined stop time
3. Unscheduled stops in undetermined stations with undefined stop time
A successor block section is considered available if the following conditions are held:
1. The block section contains no train
2. The block section is available according to maintenance schedule
3. The dispatching of a train to this successor block section should not create a deadlock situation.
4. There is at least one free line for passing train or one free platform for stopping train in the following station.

4 Architecture of the system

Designed system consists of four major parts that could solve the mentioned problem and also generate an accurate timetable of the network. SIMARAIL is a simulation-based optimization environment that consist of an integrated set of modules that all work together. The main core of SIMARAIL is a discrete-event simulation system. SIMARAIL simulates trains movements on the railway network. The simulation module helps to handle effectively different types of capacity constraints and the entire rules of railway systems. Figure 1 shows the SIMARAIL component architecture.

4.1 Data requirements and outputs

Some Infrastructure of the network and also Characteristics of the trains are listed below:
- The infrastructure characteristics:
  - number of stations and block sections
  - number of line and platforms which are existing in the stations
  - track maintenance periods (section unavailability times)
  - length of the block sections
- The train characteristics:
  - maximum speed
  - priority
  - dispatching frequency
  - route in the network
  - initial departure interval with tolerance
  - section free running time and scheduled stop times
  - safety rules
  - headway times between trains

4.2 Infrastructure details in the simulation model

In the train scheduling literature, the station capacity is typically related to the number of tracks and platforms. There are usually two types of stations according to their configuration (Fig. 2). In one-lane sta-
tions, there is no restriction for assignment of trains into tracks (Fig. 2a). On the other hand, in two-lane stations, inbound trains traverse in the inbound lane and outbound trains on outbound lane respectively (Fig. 2b). Each platform is connected with at least one track (platform2 in Fig. 2b) and at most two tracks (platform1 in Fig. 4a).

![Fig. 2. Two general station configurations and train movement constraints in the station yard](image1.png)

For entering a train to a station yard, a free track should be reserved. If the train needs a platform in a station, then a free track which is connected with a platform should be reserved. For microscopic modeling of train operations, these constraints need to be taken into account in the train timetabling problem. So the station configuration details about the number of platforms are considered in station attributes of the simulation model. Train movement on double-track is similar to car movement on a 2-lane freeway. Trains traveling in the same direction use the same track.

### 4.3 Platform

The designed model are built on the Enterprise Dynamics Platform and all the programming needed are done under 4DScript (the ED programming language).

Enterprise Dynamics® is a leading simulation platform to design and implement simulation solutions. It allows a problem solver to model virtually any problem and, by experimentation, look for a solution for a given problem or an answer to a specific question. Most of the problems or questions for which simulation is used are:

- Capacity investigations
- Investment evaluations
- Time-to-Market vs. Costs evaluations

Enterprise Dynamics has been chosen for solving this problem because of:

- Capability of developing required objects
- Powerful simulation engine
- Ability to implement customized optimization algorithms

### 4.4 Simulation model

The simulation approaches can be categorized as synchronous simulation and asynchronous simulation. Synchronous simulation models process the movements of all trains at the same time; but asynchronous models rank trains according to their priorities and insert trains into simulation procedure sequentially based on the ranks. Synchronous simulation model is the most applicable simulation model and many popular software tools (such as RailSys and OpenTrack) are designed based on this model. The synchronous simulation approach was applied for the scheduling trains. The simulation model parameters such as platforms capacity, praying time period, availability of infrastructure for fuel/water charging and finally train travel time for each block/class has been considered deterministic.

#### 4.4.1 Event-driven simulation
Discrete-event simulation systems are extensively used to model the complex systems within a discrete time frame usually through a sequential sequence of events. Event-driven simulation is also a very popular technique for synchronous simulation. Clearly an event can be identified as any change in the state of the variables in the simulation model (e.g. departure time of the trains from stations) or any updating of the trains’ conditions or railway infrastructure information.

An event signifies the happening of a possible change in the status of system in a specific time. In this event-driven simulation model, an event list is used to store system events in the order of the occurring time. The event list updates at the end of the current iteration. This process continues to execute the next processing step until a terminating condition of the simulation model is satisfied or the event list becomes empty. Modeling train operations is a key issue of simulation systems. The entities of this simulation model consist of the trains, stations, block sections, waiting queues and headway justifiers. Figure (3) shows the part of the simulation model entities and their connections.

The list of events and attributes of the entities are summarized in table (1).

<table>
<thead>
<tr>
<th>Entity \Event</th>
<th>On event</th>
<th>On entered</th>
<th>On exited</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>-</td>
<td>- Update counter attribute</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- calculate the total delay</td>
<td></td>
</tr>
<tr>
<td>Station</td>
<td>Move involved train into next block section</td>
<td>Calculation of the maximum scheduled stop time</td>
<td>-</td>
</tr>
<tr>
<td>Block section</td>
<td>- if the content of the current block section is at least 1 then move the last train in the block section into the headway justifier object - update content attribute</td>
<td>- destroy events of the involved train - update Occupancy attribute - Create exit event for entered train</td>
<td>- check the existence of the low priority train in the block section - Create update event for block section</td>
</tr>
<tr>
<td>Waiting queue</td>
<td>- check the destination station of the involved train - check the condition of releasing the involved train</td>
<td>- logging the entering time of the involved train</td>
<td>- update capacity attributed of the previous and next stations</td>
</tr>
<tr>
<td>Headway justifier</td>
<td>- Move involved train into next station</td>
<td>- check the headway times and select the train for delay</td>
<td>- update the occupancy attributed of the current headway justifier</td>
</tr>
</tbody>
</table>
4.5 Optimization Model

The role of the optimization part is to minimize traveling time of all trains in the rail network and also finding the best station to stop for refueling and praying. The procedure should be done subject to some constrains which is listed below:

- maximum traveling time
- departure time interval
- running time on segments
- minimum station stop time
- overtaking and crossing conflict
- station capacity

We considered religious constraints in the optimization model that appear in Iran railway network. Train movements on the railway network in Iran are regulated by the daily praying. All trains should stop within the praying time window for a period of 20-25 minutes. Daily praying consists of three religious services (The sunrise, afternoon and night). For each of these services there is a specified time window for praying. Each train should select the best station for stopping for every prayer service that is active. After recognizing whether a prayer service is active for a train or not, the optimization model determines the best station to stop for praying.

4.5.1 Genetic algorithm:

Genetic Algorithms (GAs) implement optimization strategies based on simulation of the natural law of evolution of a species by natural selection. The basic GA operators are:

- Encoding
- Recombination
- Crossover
- Mutation

GA has been applied to various optimization problems. It has been shown to be highly effective in searching a large, poorly defined search space even in the presence of difficulties such as high-dimensionality, multi-modality, discontinuity and noise. In this problem, the genetic parameters are listed in table 2.

<table>
<thead>
<tr>
<th>Table 2- GA parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Real Life</strong></td>
</tr>
<tr>
<td>Chromosome</td>
</tr>
<tr>
<td>Gene</td>
</tr>
<tr>
<td>Genotype</td>
</tr>
<tr>
<td>Environment</td>
</tr>
<tr>
<td>Breeding Procedure</td>
</tr>
</tbody>
</table>

The typical genetic algorithm has the main following stages:

<table>
<thead>
<tr>
<th>Table 3-GA procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>T=0</td>
</tr>
<tr>
<td>Init population P(0)</td>
</tr>
<tr>
<td>evaluate P(0)</td>
</tr>
<tr>
<td>while (not termination condition) do</td>
</tr>
<tr>
<td>begin</td>
</tr>
<tr>
<td>T = T+ 1</td>
</tr>
<tr>
<td>select P(i) from P(T-1)</td>
</tr>
<tr>
<td>recombine P(T)</td>
</tr>
<tr>
<td>mutate P(T)</td>
</tr>
<tr>
<td>end</td>
</tr>
</tbody>
</table>

| end |

| end |
Each train has a departure time interval which could be set by user. The default length of initial interval time is 30 minutes. It means that if the user set the assumed time of departure to 10, system can select any time between 9:30 to 10:30 for departure. The first population was created randomly in the area that are allowed to be. Sizes of chromosomes are proportional to number of the trains defined in the model. After generating the first population, evaluation of each chromosome is started by simulation. For calculating the fitness of each chromosome, a simulation model runs, and total delay of the trains recorded as the value of fitness. Afterwards, the process of creating new generation would be started.

GA functions in this research are designed basis on the Rank Selection method. After sorting the solutions (all the population) by their fitness (total traveling time), choosing them for crossing over would be done. The chromosome with higher rank is more likely to be chosen. In each generation, there are some chromosomes that some of their genes had been changed by mutation. The mutation changes the chromosomes in the way that they remain feasible solution already.

Chromosomes consist train departure times from origin station, during the 24 hour period. For example a chromosome which starts with the following genes, shows that the first train will depart at 8:00, second in 8:20, etc. Each chromosome includes more than 150 genes, equals to the total number of trains in the whole network, ready to depart. Each gene must be set in pre-defined period and in 5 minute steps (8:00, 8:05, 8:10 ...).

Origin of each train is defined in the model and it’s not part of the GA parameters. The praying, fuel charging, water charging (for steam locomotives) and other service providing stations are assign to each train though simulation run, taking into account the line, platform or equipment capacity of the stations & praying allowed period.

Optimization will force simulation to run each chromosome and gather the fitness of them. In some answers (chromosomes), so many delays occur because of limit of infrastructure capacity in stations or other constraints. But some answers, leveled the delays between trains or lowered them. After calculating fitness of all chromosomes (total network delay), GA sort the chromosomes ascending (better chromosome placed upper). Selection opportunity of each chromosome is increased by its rank thorough GA operations in terms of cross over and mutation. The new population consists of three following types of chromosome:
- Children created by cross over operation
- Children created after mutation of previous generation
- Best of the previous generation

Share of each type is configured by user by GUI form.

The operation tends to find the minimal delay for each train in the network. Some trains are traveled in two way lines thorough their journey and most delays of them are caused by the capacity of stations and difference in speed of trains. The finish rule of the GA is defined when the sum of such delays get equal
to zero. Experiments showed that around 400 generations have to be passed to achieve the mentioned result.
All limitations are managed in the simulation model and optimization takes the model as a black box in which a chromosome fitness are assigned. If a chromosome is infeasible, (for example a train couldn’t find the station for praying until the time is passed away) the simulations model turns back fitness and a penalty equals to 10 hours of delay. After some generations, high percentage of feasible solution could be gained.

5 Results

Extended railway network of Iran is covering large area with the main line about 9,036 kilometers. Because of old structure of rails in many parts, more than 7500 kilometers of the mentioned network is single-track section and only less than 1500 kilometers is double-track section. According to this fact, the scheduling of the trains and the estimation of the traveling time of them, turns to be a very complicated and hard job to do. Also Iran has the second line of railway network about 2,457 Kilometers. All lines together connect more than 400 City and Town Stations together with an increasing rate of usage by the people. In the table bellow, the increasing in traveling distance during recent years is noticeable.

<table>
<thead>
<tr>
<th>Year</th>
<th>Distance traveled by Locomotives</th>
<th>Person*kilometer In Million</th>
<th>No of travelers In Million</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>67,924,754</td>
<td>12,549</td>
<td>21</td>
</tr>
<tr>
<td>2009</td>
<td>64,358,772</td>
<td>13,900</td>
<td>24</td>
</tr>
<tr>
<td>2010</td>
<td>67,915,918</td>
<td>15312</td>
<td>26</td>
</tr>
</tbody>
</table>

In this section, we apply SIMARAIL to a large-scale railway network. The network used in this paper has 152 stations and 151 sections. In the designed simulations, 148 trains traverse the network. Figure (4) shows the result of the optimization phase.
The information & analyzes which could be gained in figure 6 are:

- The horizontal axis is the time.
- The vertical axis shows station name.
- Each colored line shows the train from its origin to destination. The number in parentheses beside each line is the train number. The number beside each line between two stations shows the travel time for that train. The number under or above the line in the station shows the minute section of the arrival/departure time to/from that particular station. The horizon lines indicate the stop time of trains in that station. In this example the praying time is between 4 to 5 o’clock, hence the trains are stopped for praying purpose.
- In this example the block of JahanAbad – Abrisham is single track due maintenance, so the trains has to be stopped for the opponent path get free.
- In this example, the green train which has been arrived 4:28 at Bastam station. First it has stopped for praying 20 minutes. But after praying there was no free capacity at next station. So it has stopped for 40 minutes. The GA tries to minimize such delay by departing the train later from origin of its boundaries allows. It could be seen it is possible to depart train 20 minutes later to avoid such delays.

6 Conclusion

In this paper, we presented SIMARAIL, in which train traveling in the Iranian railway network is modeled by discrete-event simulation approach. SIMARAIL is introduced, which could properly develop the optimal timetable, subject to all constrains. For achieving the goal, we enhanced the simulation model with a optimization module. The simulation model is used for finding the fitness of each chromosome, which is generated during the GA optimization. In other
words, simulation-based optimization has been used to develop proper timetables. Using simulation helps op-
timization procedures by making modeling of constraints easier in comparison to mathematical modeling.

References


ACKNOWLEDGMENTS

The authors would like to acknowledge the support & effort of Islamic Republic of Iran Railways, especially, Mr. Hosseini, head of IT department, to do the project successfully.
AUTHOR BIOGRAPHIES

Arman Sajedinejad is a Ph.D. candidate of Industrial Engineering at the Industrial Engineering Department, Faculty of Engineering, Tarbiat Modares University, in Tehran, Iran. His research background includes: Supply Chain Management, JIT Production Systems, Simulation based optimization and also Performance Evaluation of Organizations. His email address is sajedinejad@modares.ac.ir

Soheil Mardani is a founder of Simaron Pardaz Co., dedicated company to develop simulation solution for different sector of industries. He is specialist in applying discrete-event simulation for complex manufacturing, logistics & transportation systems and for modeling business systems as well. Besides to his professional life, he also teaches simulation corresponding courses at universities for BSc & MSc student of Industrial & Transportation Engineering. His email address is s.mardani@simaron.com

ERFAN HASSAN NAYEBI is M. Sc. Student of industrial engineering at Sharif university of Technology. He received his B. Sc. for industrial engineering at Sharif university of Technology in 2009. During his career at SIMARON he has contributed in a wide variety of simulation-based optimization studies for railway companies. His research interests are in the area of simulation and meta-heuristic algorithm. He can be reached at e_hasannayebi@ie.sharif.edu

S. Ahmad Reza Mir Mohammadi K. received his bachelor’s degree in Industrial Engineering – Systems Planning and Analysis in 2010 from Amirkabir University of Technology, Iran. His research interests are simulation modeling, evolutionary algorithms, scheduling and expert systems. He had been working in Simaron Pardaz Co. from April 2010 to July 2011. His e-mail address is a.r.mirmohammadi@gmail.com