Available online at www.sciencedirect.com
ScienceDirect
IFAC PapersOnLine 52-13 (2019) 957–962

A Genetic Algorithm for Assigning Train Arrival Dates at a Maintenance Centre

Hanyu Gu *, **Mark Joyce** **, **Hue Chi Lam** *,[1], **Mark Woods** **, **Yakov Zinder** *

* *University of Technology, Sydney, PO Box 123, Broadway, NSW, 2007, Australia (e-mail: hanyu.gu@uts.edu.au, hue.lam@student.uts.edu.au, yakov.zinder@uts.edu.au).*
** *UGL Unipart, Sydney, NSW, Australia (e-mail: mark.joyce@uglunipart.com, mark.woods@uglunipart.com).*

Abstract:
The paper is concerned with planning heavy maintenance of train-sets at a maintenance centre. The heavy maintenance process is complex and, for each train-set, the actual duration of maintenance is uncertain at the time of planning. The allocation of the dates when train-sets should arrive at the maintenance centre is crucial phase of the planning procedure. The objective function is a weighted sum of two components, the expected total penalty for not meeting the required number of train-sets in active service and the total cost for the deviation (earliness and tardiness) from the desired dates of arrival. A genetic algorithm is presented for the considered problem and its effectiveness is demonstrated by the computational experiments that used real-world data provided by a big maintenance centre.

Keywords: Planning of trains maintenance, Earliness and tardiness, Genetic algorithm

1. INTRODUCTION

The urban passenger trains have long been the backbone of the public transportation system and are one of the most popular modes of transport in many cities around the world. For example, in Sydney, Australia, the urban passenger rail network spans over 815 kilometres of track and delivered 340.7 million passenger journeys in the 2016 - 17 financial year (Sydney Trains, 2017). This paper presents an optimisation procedure for the key phase of planning the heavy maintenance of passenger trains in the urban transportation systems - the assignment of dates when the rolling stock should arrive at the maintenance centre.

The rolling stock is considered unsafe after a certain time period and the regular maintenance of the urban passenger trains must comply with the corresponding safety regulations. For this reason, each maintenance operation is assigned the desired date of its commencement with the possibility of some deviation from this due date. The increase of this deviation is undesirable since the performance of the maintenance too early would cause unnecessary loss of the rolling stock remaining mileage, whereas the performance of the maintenance too late would impact the network reliability. This leads to the introduction of an arrival time window within which the train can arrive at the maintenance centre. Any violation of this time window is undesirable.

The maintenance, in general, falls into two broad categories: light maintenance and heavy maintenance. The process of light maintenance, involving inspection and cleaning, has a short cycle time and the maintenance work can, therefore, be performed within a day or overnight at the rolling stock depot. On the other hand, the process of heavy maintenance, involving not only inspection and cleaning but also components change-out, has a long cycle time of at least one month and the maintenance work must be conducted in specialised maintenance centres. The duration of heavy maintenance depends on the fleet type and scope of work. Older fleet types typically require a longer duration of maintenance.

The rolling stock arrives at the maintenance centre in groups. Each group is typically comprised of four cars and is referred to as a set or a train-set (see for example Lai et al. (2015)). A train-set upon arriving at the maintenance centre is taken offline and shunted to the first operation line, where thorough inspections and replacement of some components and parts such as air-conditioning units are performed. After that, each train-set undergoes various maintenance operations such as bogies replacement, brakes testing, the public address system testing, etc. In order to undergo one type of service after another, a train-set has to be shunted between several lines. The duration of each operation depends on the condition of the train-set; the availability and composition of the workforce, which is comprised of technicians and engineers with a broad range of skills; the availability of spare parts; and many other factors, and is therefore uncertain at the time of planning. Furthermore, the necessity to use the special lines for each type of maintenance operations and the necessity to shunt train-sets between these lines impose the restriction on

[1] corresponding author

2405-8963 © 2019, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2019.11.318

the number of train-sets which can undergo maintenance simultaneously.

As previously mentioned, heavy maintenance has long and uncertain cycle time during which the train-sets must be completely withdrawn from service. This implies that the allocation of arrival dates at the maintenance centre has a direct impact on the number of trains in active service. If too many train-sets are taken out of service, there are not enough trains to meet the demand, especially during peak hours. Hence, for each type of train-sets there is a permissible number of train-sets that can be out of service simultaneously.

As a result of the complexity of the heavy maintenance process, the actual cycle time of each train-set at the maintenance centre is uncertain at the time of planning. This brings about challenges to the planner in assigning the train-sets' arrival dates that minimises a weighted sum of two components, the expected total penalty for violating the permissible number of out-of-service train-sets, and the total cost for the deviation (earliness and tardiness) from the desired dates of arrival.

The remainder of the paper is organised as follows. Section 2 presents a review of relevant publications. Section 3 presents the heavy maintenance planning problem of urban passenger trains with uncertain duration. Section 4 discusses the proposed Genetic Algorithm (GA). Section 5 presents the results of the computational experiments. The conclusion and directions of further research are given in Section 6.

2. RELATED WORKS

The literature on the planning heavy maintenance of trains is limited. As an early work, the problem is addressed by Sriskandarajah et al. (1998), in which a deterministic model with 27 trains and a planning horizon of one year is considered. Using GA, the author aims at finding a schedule that minimises the total cost of earliness and tardiness. The work of Doganay and Bohlin (2010) addresses the fleet level maintenance scheduling problem where they optimise the maintenance cost and the spare part cost. The authors present a deterministic mixed integer linear programming model with a planning horizon of two years for the studied problem.

The aforementioned publications and our paper are concerned with a long-term planning strategy for the planning heavy maintenance of trains. Furthermore, details of the maintenance activities are not considered. Instead, cycle time is used to represent the total duration of the stay of a train-set at the maintenance centre. The difference between our paper and the existing literature is that we treat the maintenance duration of train-sets as random variables while existing literature considers maintenance duration as a constant.

The problem considered in this study is closely related to the Resource Levelling Problem (RLP). In RLP, there is a set of activities which requires the consumption of various types of resource. The problem aims at finding a schedule which minimises the fluctuations in the consumption of resources under the restriction of precedence constraints and project deadline. The study in Li and Demeulemeester (2016) considers uncertain duration and aims at minimising the positive deviation from the desired resource utilisation. The approach is based on genetic algorithm. The study of Li and Demeulemeester (2016) considers activities with precedence constraints and project deadline constraint. In contrast, the study considered in our paper does not include precedence constraints. Furthermore, each train-set has a due date which leads to the introduction of the cost for earliness and tardiness with respect to the due date.

The problem considered in this study is also closely related to the Single Machine Scheduling Problem (SMSP) with the earliness and tardiness objective. In this problem, a set of jobs has to be processed on a single machine and preemptions are not permitted. Each job has a processing time and a due date by which the job should ideally be completed. The earliness and tardiness are the difference between the due date and the completion time of the job if the job is completed before or after the due date, respectively. It is well-known that the SMSP with the earliness and tardiness penalty is strongly NP-hard (Wan and Yuan, 2013). Apart from the linear earliness and tardiness penalty function which have been extensively studied in the literature, a recent trend has been to consider the quadratic penalty function (Vila and Pereira, 2013; Valente et al., 2011) because it is more practical to apply a heavy penalty for a large deviation from the desired due date.

3. PROBLEM DESCRIPTION

The heavy maintenance planning problem can be described as follows. Train-sets, constituting the set $N = \{1, \cdots, n\}$, are to undergo heavy maintenance at a maintenance centre during a planning period of T days. The planning horizon is discretised into calendar days which are indexed from 0 to $T-1$.

Each train-set $j \in N$ has an integer due date θ_j, which is the preferred day of the commencement of maintenance relative to the start of the planning period. There are m types of trains and the set of all train-sets is partitioned into m families. Each train-set belongs to a train family $F^k, k \in S = \{1, \cdots, m\}$. For each train family F^k and each day t, let C_{kt} be the permissible number of out-of-service train-sets. The value of C_{kt} is determined according to the predicted demand and is smaller if day t falls on public holidays. For each day t, let C_t be the permissible number of train-sets which can undergo maintenance simultaneously.

After the arrival of a train-set at the maintenance centre, inspection and several maintenance operations must be carried out without interruption on the first operation line. For each train family F^k, let p_k be the minimum duration on the first operation line. If a train-set of train family F^k arrives at the maintenance centre on day t, no other train-sets can arrive in the interval $[t, t + p_k - 1]$.

The cycle time of each train-set j is a random variable D_j, and $D = [D_1, \cdots, D_n]$. Train-sets in the same family F^k follow the same probability distribution. It is further assumed that all D_j are independent.

For each train-set $j \in N$, let nonnegative integer s_j be the arrival day of train-set j at the maintenance centre. The objective is to obtain an arrival plan $s = (s_1, \cdots, s_n)$ that minimises a weighted sum of two components. The first component of the objective function, denoted by G_1, is the sum of the expected penalties for violating the limits C_t and C_{kt}. More specifically, for any arrival plan s and any integers $1 \le k \le m$ and $0 \le t < T$, the number of trains of family k that reside at the maintenance centre on day t is

$$W^k_{s,t} = \sum_{j \in F^k} B(s_j, t), \tag{1}$$

where

$$B(s_j, t) = \begin{cases} 1 \text{ if } s_j \le t \text{ and } s_j + D_j \ge t+1 \\ 0 \text{ otherwise} \end{cases}. \tag{2}$$

Then, the total number of trains that reside at the maintenance centre on day t is

$$W_{s,t} = \sum_{k \in \{1,\ldots,m\}} W^k_{s,t}. \tag{3}$$

In what follows, the notation W_t and W^k_t can be used instead of $W_{s,t}$ and $W^k_{s,t}$ if it is clear what arrival plan is considered and therefore the omission of s does not cause any confusion or misunderstanding. Thus, G_1 is

$$G_1 = \sum_{t=0}^{T-1} \left(\delta \mathbb{E}\left[(W_t - C_t)^+\right] + \sum_{k=1}^{m} \delta_{kt} \mathbb{E}\left[(W^k_t - C_{kt})^+\right] \right), \tag{4}$$

where $(a)^+ := max(a, 0)$; $\mathbb{E}[\cdot]$ denotes the expectation operator; δ and δ_{kt} are the daily penalties for violating the respective limits.

The second component of the objective function, denoted by G_2, is the total cost of earliness and tardiness. The earliness and tardiness of train-set j are defined as follows:

$$E_j = max(0, \theta_j - \Delta - s_j), \tag{5}$$

$$T_j = max(0, s_j - \Delta - \theta_j), \tag{6}$$

where Δ defines the duration of the admissible earliness/tardiness time window. Hence, the earliness is the difference between the beginning of the arrival time window and the arrival day of a train-set, while tardiness is the difference between the end of the arrival time window and the arrival day of a train-set. As has been mentioned in Section 2, G_2 is a quadratic penalty function, i.e.

$$G_2 = \sum_{j \in N} (\lambda_1 E_j^2 + \lambda_2 T_j^2), \tag{7}$$

where λ_1 is the earliness cost factor and λ_2 is the tardiness cost factor.

We introduce a binary variable, y_{jk}, which is equal to 1 if train-set j arrives before train-set k and equal to 0 otherwise. The problem can be formulated as follows:

$$Z = \min\ (\alpha\ G_1 + \beta\ G_2) \tag{8}$$

subject to

$$0 \le s_j \le T-1, \ \ \forall j \in N \tag{9}$$

$$s_j + p_j \le s_k + M(1 - y_{jk}), \ \text{ for } j,k \in N, j<k \tag{10}$$

$$s_k + p_k \le s_j + M y_{jk}, \ \text{ for } j,k \in N, j<k \tag{11}$$

$$y_{jk} \in \{0,1\}, \ \ \forall j,k \in N, j<k \tag{12}$$

$$s_j \in \{0\} \cup \mathbb{Z}^+, \ \ \forall j \in N \tag{13}$$

where G_1 and G_2 are defined by (1)-(7); α and β are weights reflecting the relative importance of the two components of the objective function; s_j and y_{jk} are the decision variables; and M is a sufficiently large constant.

Constraint set (9) ensures that each train-set must arrive for maintenance on a particular day within the planning horizon. Constraint set (10) and (11) are disjunctive constraints which enforce that either train-set j arrives before train-set k or train-set k arrives before train-set j. Constraint set (12) and (13) are the integrality constraints.

4. GENETIC ALGORITHM

In this section, a Genetic Algorithm (GA) is presented. Then, the decoding procedure that is used to transform a chromosome into an arrival plan is described. Next, the method that is used to calculate the objective function is presented. Finally, the evolutionary strategy is also described.

4.1 The genetic algorithm

Since the pioneer publications of Holland (1975) and Goldberg et al. (1989), GA has been extensively applied to scheduling problems.

For the problem considered in this paper, a solution is an arrival plan $s = (s_1, \cdots, s_n)$ that specifies the arrival days for all train-sets. Each arrival plan is encoded as a chromosome whose size is equal to the number of train-sets. Each gene j in the chromosome is a random number generated according to the uniform distribution U(0,1) which determines the priority of a train-set $j \in N$ (see Figure 1 for an example). The benefit of using the random key to encode the solution is that we do not directly deal with the train-set indexes. Hence, all children chromosomes generated by crossover and mutation are guaranteed to be decoded into feasible solutions.

Train-set index	1	2	3	4	5
Random key	0.12	0.08	0.57	0.84	0.23
Sorted random key	0.08	0.12	0.23	0.57	0.84
Priority of train-set	2	1	5	3	4

Fig. 1. Random key encoding example

The GA starts with the generation of an initial population of POP chromosomes. Then, each chromosome of the initial population is decoded into a corresponding feasible arrival plan (see Section 4.2). Next, the fitness value of each arrival plan of the initial population is computed by evaluating the objective function according to (4) and (7) (see Section 4.3). In each iteration, the current population of chromosomes is evolved into the new population through 4 main steps, including parents selection, crossover, mutation, and offspring selection (see Section 4.4). This process is repeated until no more improvement is made for $stop\_iter$ iterations.

4.2 The decoding procedure

The decoding procedure transforms a chromosome into a feasible arrival plan. The decoding procedure is inspired

Algorithm 1 Genetic Algorithm

Input: Input data of all the train-sets
Output: Arrival plan $s = (s_1, \cdots, s_n)$
procedure
  Generate initial population $P_0[POP]$ randomly
  Decode $P_0[POP] \rightarrow s[POP]$
  Evaluate $s[POP] \rightarrow G[POP]$
  Set the current best objective $G^* = min(G[POP])$, and the current best plan s^*
  Set $t = 0$
  Set $i = 0$
  while $i < stop\_iter$ **do**
    Perform parents selection $\rightarrow Parents[POP]$
    Perform crossover $\rightarrow Children[POP/2]$
    Perform mutation $Children[POP/2]$
    Decode $Children[POP/2] \rightarrow s[POP/2]$
    Evaluate $s[POP/2] \rightarrow G[POP/2]$
    Select offspring $\rightarrow P_{t+1}[POP]$
    if (new best objective is found) **then**
      Update G^*, s^*
      Set $i = 0$
    else Set $i = i + 1$
    Set $t = t + 1$
  return s^*

by Li and Demeulemeester (2016) and is detailed in Algorithm 2. The mean cycle time is used in the decoding procedure for the following reasons: (i) the decoding procedure has to be computationally efficient because it is frequently invoked in the GA, and (ii) by using the mean cycle time of the train-sets, some levels of uncertainty has been taken into consideration.

The decoding procedure starts with an empty set of planned train-sets ($N^{'}$). For each day t, $W_t(N^{'})$ denotes the total number of train-sets, residing in the maintenance centre on day t, and for each family F^k, $W_t^k(N^{'})$ denotes the total number of train-sets of family F^k, residing in the maintenance centre on day t. For each train-set $j \in N$, $\bar{d}_j$ is the mean cycle time of train-set j. CB denotes the current best performance measure value. The value of the objective function $G = \alpha G_1 + \beta G_2$ is not used as the performance measure value because the violation of the limits is always equal to zero at the early stages of the decoding procedure. Instead, the performance measure PM is calculated as follow:

$$PM = \left(\delta \sum_{t=s_j}^{s_j+\bar{d}_j-1} \Big(W_t(N^{'}) + 1 \Big) + \sum_{k \in S} \sum_{t=s_j}^{s_j+\bar{d}_j-1} \delta_{kt} \left(W_t^k(N^{'}) + \frac{1}{|S|} \right) \right)^2 \tag{14}$$

In each iteration, a train-set $j \in N$, whose random key value is the smallest, is selected. From all feasible days within the planning horizon, the arrival day of train-set j is chosen such that it gives the smallest performance measure value as computed according to (14). The procedure is repeated until the arrival days of all the train-sets have been determined.

Algorithm 2 Decoding Procedure

Input: A chromosome
Output: An arrival plan $s = (s_1, \cdots, s_n)$
procedure
  Set $N^{'} = \emptyset$, $W_t(N^{'}) = 0$, $W_t^k(N^{'}) = 0$, $\forall t \in$ [0,T - 1]
  while $N \setminus N^{'} \neq \emptyset$ **do**
    $CB = \infty$
    Select a train-set j from $N \setminus N^{'}$ with minimum random key
    for $\theta \in [0, T-1]$ **do**
      if there is no s_k such that $s_k < \theta$ and $s_k + p_k > \theta$ and there is no s_k such that $s_k > \theta$ and $\theta + p_j > s_k$ **then**
        $PM = 0$
        for $t = \theta$ to $\theta + \bar{d}_j - 1$ **do**
          $PM = PM + \delta\Big(W_t(N^{'}) + 1\Big) +$
          $\sum_{k \in S} \delta_{kt} \Big(W_t^k(N^{'}) + \frac{1}{|S|} \Big)$
        $PM = PM^2 + \sum_{j \in N^{'} \cup j} (\lambda_1 E_j^2 + \lambda_2 T_j^2)$
        if $PM < CB$ **then**
          $CB = PM$, $s_j = \theta$
      else continue
    $N^{'} \leftarrow N^{'} \cup j$
  return s

4.3 Evaluation of the objective function

For each scheduled train-set and each day t, the arrival plan permits to calculate the probability that this train-set will be at the maintenance centre on day t. Hence, the number of train-sets on day t is a sum of several Bernoulli random variables and therefore is distributed according to the Poisson Binomial Distribution. This observation together with Biscarri et al. (2018) permit the efficient calculation of the distributions for all W_t and W_t^k. This, in turn, leads to a fast algorithm for evaluating the objective function.

4.4 Evolutionary strategy

The evolutionary strategy evolves the current generation towards better successive generation. It includes parent selection, resource-based crossover, two-point crossover, mutation, and offspring selection. The evolutionary strategy, inspired by Li and Demeulemeester (2016), is described in subsequent paragraphs.

The parent selection phase selects the top POP/2 best chromosomes in the current population as the father chromosomes. The remaining chromosomes form a pool of candidates from which two chromosomes are randomly chosen each time, and the one with a smaller fitness value is nominated as the mother chromosome. The process is repeated until POP/2 mother chromosomes are obtained. One father chromosome is paired with one mother chromosome to form a pair of parent chromosome.

The crossover phase operates on the POP/2 parent chromosomes to produce POP/2 children chromosomes. Following the idea of Li and Demeulemeester (2016), the crossover phase includes a resource-based crossover operator which is applied to the parent chromosomes with the top POP/4 best father chromosomes, and a two-point crossover operator which is applied to the remaining parent chromosomes. In the resource-based crossover operator, for a father chromosome and its corresponding arrival plan, a partial plan of length ε is randomly chosen in $[0.25T, 0.75T]$, a crossover point t is selected such that G_1 is minimised for the interval $[t, t+\varepsilon]$. Then, for the train-sets whose arrival days fall within the interval $[t, t+\varepsilon]$, the value of the corresponding gene is added by 5000 and given to the child chromosome. The values of other genes are obtained from the mother chromosome (see Figure 2 for an example).

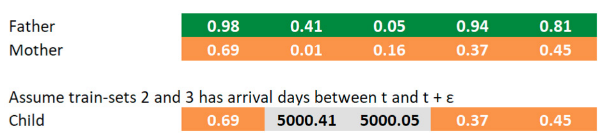

Fig. 2. Resource-based crossover example

In the two-point crossover operator, two crossover points t_1 and t_2 are randomly selected. The genes between t_1 and t_2 of the child chromosome are set equal to the corresponding genes in the father chromosome while the values of other genes are obtained from the mother chromosome (see Figure 3 for an example) .

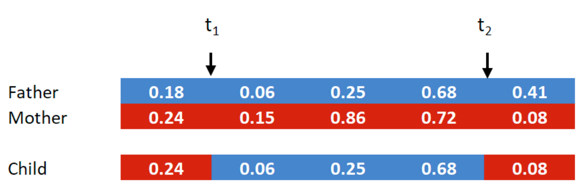

Fig. 3. Two-point crossover example

The mutation phase attempts to replace the values of some genes in the child chromosome. For each gene in the child chromosome, a random number is generated according to the uniform distribution $U(0, 1)$. If this random number is less than *mutation_prob*, the corresponding gene value is replaced by a new random key generated according to the uniform distribution $U(0, 1)$ (see Figure 4 for an example).

Fig. 4. Mutation example

Before the process of offspring selection, a migration phase is performed in which new chromosomes are randomly generated. The number of newly generated chromosomes is set equal to a proportion *mig_prob* of the population size. This step is added to the GA to prevent premature convergence (Valente et al., 2011). Finally, the top POP best chromosomes from the union set of parent chromosomes, children chromosomes, and the newly generated chromosomes are selected as the candidates of the new population.

5. EXPERIMENTAL RESULTS

All algorithms are implemented in Python 2.7. The computational experiment is conducted on a computer with Intel i5-6300U 2.4GHz processor and 8GB of RAM.

The proposed GA is tested on data provided by a big maintenance centre. The planning horizon is one year. There are three train families with 35 train-sets in total. The parameters of the train families are presented in Table 1. Limit normal and limit special indicate the permissible number of out-of-service train-sets on normal days and on special days, respectively.

Table 1. Parameters for the train families

Train Family	$\|F^k\|$	p_k	Limit normal	Limit special
1	25	4	3	1
2	5	5	2	1
3	5	5	1	1

Furthermore, the remaining parameters are given as follows: $C_t = 5$, $\Delta = 14$, $\delta_t = 1$, $\delta_{kt} = 1$ on normal days, $\delta_{kt} = 10$ on special days, $\lambda_1 = \lambda_2 = 1$.

Nine scenarios are generated with different weights for the two components, G_1 and G_2, of the objective function. Details of the assignment of weights α and β for all the scenarios are presented in Table 2.

Table 2. Assignment of α and β

Scenario	α	β	Scenario	α	β
1	1000	1	6	100	1
2	300	1	7	50	1
3	200	1	8	10	1
4	180	1	9	1	1
5	150	1			

The GA parameter values used in this study is specified as follows: $POP = 20$, $mig\_prob = 0.05$, $mutation\_prob = 0.05$, and $stop\_iter = 40$.

Table 3 provides the results of GA for all the scenarios. The best solution of the initial population is reported under the column titled "In Pop". The relative improvement in the objective function value over the best solution of the initial population is reported under the column titled "%Rel". For example, The relative improvement of GA for scenario 1 is calculated as $(417, 581 - 341, 962)/417, 581 \times 100\% = 18.11\%$.

On average, the relative improvement of the proposed GA over the best solution of the initial population is approximately 19.1%. The largest improvement of 31.69% is observed in scenario 7 while the smallest improvement of 11.42% is observed in scenario 3. Furthermore, it is noted that the relative improvement becomes worse as the relative weight of α increases.

Table 3. Results of GA

Scenario	In Pop	GA	%Rel
1	417,581	341,962	18.11
2	150,158	132,345	11.86
3	111,955	99,165	11.42
4	104,314	92,363	11.46
5	92,853	72,170	22.27
6	73,752	59,084	19.89
7	52,509	35,868	31.69
8	34,363	27,959	18.64
9	30,280	22,235	26.57
Average	118,640	98,128	19.10

The solution time (in minutes) of GA is reported in Table 4. The average solution time of the proposed GA is 20.56 minutes. The GA requires a significant amount of time for scenarios 5 and 6 as compared to other scenarios, at 41 and 40 minutes respectively.

Table 4. Solution time (in minutes)

Scenario	GA	Scenario	GA
1	16	6	40
2	10	7	23
3	10	8	10
4	20	9	15
5	41		
Average	20.56		

6. CONCLUSIONS

This paper considered the key phase of planning the heavy maintenance of passenger trains in which one has to specify for each train-set the date when this train should arrive at the maintenance centre. It is well-known that the deterministic resource levelling problem and the single machine scheduling problem with earliness and tardiness, which are closely related to the problem considered in this paper, are NP-hard. Hence, a genetic algorithm was proposed for the considered problem.

The proposed metaheuristic, inspired by Li and Demeulemeester (2016) and Valente et al. (2011), provided an efficient way to solve the planning heavy maintenance of train-sets. Indeed, computational results, with data provided by one of the leading maintenance centres in Australia, revealed that significantly improved arrival plan can be obtained for test cases consisting of 35 train-sets and a planning horizon of one year. At the best of the authors' knowledge, the considered problem has not been discussed in the literature previously. So, the results of the computational experiments can be viewed as a benchmark for further research.

Future work can further improve the proposed genetic algorithm with the application of a local search procedure as suggested by Valente et al. (2011). Another possibility for future research is to test other metaheuristics such as iterated local search or tabu search on the considered problem in order to improve the solution quality and reduce the solution time.

REFERENCES

Biscarri, W., Zhao, S.D., and Brunner, R.J. (2018). A Simple and Fast Method for Computing the Poisson Binomial Distribution Function. *Computational Statistics & Data Analysis*, 122, 92–100.

Doganay, K. and Bohlin, M. (2010). Maintenance Plan Optimization for a Train Fleet. *WIT Transactions on the Built Environment*, 114, 349–358.

Goldberg, D.E., Korb, B., and Deb, K. (1989). Messy Genetic Algorithms: Motivation, Analysis, and First Results. *Complex Systems*, 3, 493–530.

Holland, J.H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Michigan.

Lai, Y.C., Fan, D.C., and Huang, K.L. (2015). Optimizing Rolling Stock Assignment and Maintenance Plan for Passenger Railway Operations. *Computers & Industrial Engineering*, 85, 284–295.

Li, H. and Demeulemeester, E. (2016). A Genetic Algorithm for the Robust Resource Leveling Problem. *Journal of Scheduling*, 19(1), 43–60.

Sriskandarajah, C., Jardine, A., and Chan, C. (1998). Maintenance Scheduling of Rolling Stock Using a Genetic Algorithm. *The Journal of the Operational Research Society*, 49, 1130–1145.

Sydney Trains (2017). Sydney Trains Annual Report 2016-17. https://www.transport.nsw.gov.au Accessed 30 July 2018.

Valente, J.M.S., Moreira, M.R.A., Singh, A., and Alves, R.A.F.S. (2011). Genetic Algorithms for Single Machine Scheduling with Quadratic Earliness and Tardiness Costs. *The International Journal of Advanced Manufacturing Technology*, 54(1), 251–265.

Vila, M. and Pereira, J. (2013). Exact and Heuristic Procedures for Single Machine Scheduling with Quadratic Earliness and Tardiness Penalties. *Computers & Operations Research*, 40(7), 1819–1828.

Wan, L. and Yuan, J. (2013). Single-machine Scheduling to Minimize the Total Earliness and Tardiness is Strongly NP-hard. *Operations Research Letters*, 41(4), 363–365.