Data Article

A benchmark dataset for the multiple depot vehicle scheduling problem

Sarang Kulkarni [a,b,c,*], Mohan Krishnamoorthy [d,e], Abhiram Ranade [f], Andreas T. Ernst [c], Rahul Patil [b]

[a] IITB-Monash Research Academy, IIT Bombay, Powai, Mumbai 400076, India
[b] SJM School of Management, IIT Bombay, Powai, Mumbai 400076, India
[c] School of Mathematical Sciences, Monash University, Clayton, VIC 3800, Australia
[d] Department of Mechanical and Aerospace Engineering, Monash University, Clayton, VIC 3800, Australia
[e] School of Information Technology and Electrical Engineering, The University of Queensland, QLD 4072, Australia
[f] Department of Computer Science and Engineering, IIT Bombay, Powai, Mumbai 400076, India

ARTICLE INFO

ABSTRACT

This data article presents a description of a benchmark dataset for the multiple depot vehicle scheduling problem (MDVSP). The MDVSP is to assign vehicles from different depots to timetabled trips to minimize the total cost of empty travel and waiting. The dataset has been developed to evaluate the heuristics of the MDVSP that are presented in "A new formulation and a column generation-based heuristic for the multiple depot vehicle scheduling problem" (Kulkarni et al., 2018). The dataset contains 60 problem instances of varying size. Researchers can use the dataset to evaluate the future algorithms for the MDVSP and compare the performance with the existing algorithms. The dataset includes a program that can be used to generate new problem instances of the MDVSP.

© 2018 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Specifications table

Subject area	Operations research
More specific subject area	Vehicle scheduling
Type of data	Tables, text files
How data were acquired	Artificially generated by a C++ program on Intel® Xeon® CPU E5–2670 v2 with Linux operating system.
Data format	Raw
Experimental factors	Sixty randomly generated instances of the MDVSP with the number of depots in $(8, 12, 16)$ and the number of trips in $(1500, 2000, 2500, 3000)$
Experimental features	Randomly generated instances
Data source location	IITB-Monash Research Academy, IIT Bombay, Powai, Mumbai, India.
Data accessibility	Data can be downloaded from https://orlib.uqcloud.net/
Related research article	Kulkarni, S., Krishnamoorthy, M., Ranade, A., Ernst, A.T. and Patil, R., 2018. A new formulation and a column generation-based heuristic for the multiple depot vehicle scheduling problem. Transportation Research Part B: Methodological, 118, pp. 457–487 [3].

Value of the data

- The dataset contains 60 different problem instances of the MDVSP that can be used to evaluate the performance of the algorithms for the MDVSP.
- The data provide all the information that is required to model the MDVSP by using the existing mathematical formulations.
- All the problem instances are available for use without any restrictions.
- The benchmark solutions and solution time for the problem instances are presented in [3] and can be used for the comparison.
- The dataset includes a program that can generate similar problem instances of different sizes.

1. Data

The dataset contains 60 different problem instances of the multiple depot vehicle scheduling problem (MDVSP). Each problem instance is provided in a separate file. Each file is named as '*RN-m-n-k.dat*', where '*m*', '*n*', and '*k*' denote the number of depots, the number of trips, and the instance number for the size, '(m, n)', respectively. For example, the problem instance, '*RN-8–1500-01.dat*', is the first problem instance with 8 depots and 1500 trips. For the number of depots, m, we used three values, 8, 12, and 16. The four values for the number of trips, n, are 1500, 2000, 2500, and 3000. For each size, (m, n), five instances are provided. The dataset can be downloaded from https://orlib.uqcloud.net.

For each problem instance, the following information is provided:

The number of depots (m),

The number of trips (n),

The number of locations (l),

The number of vehicles at each depot,

For each trip $i \in 1, 2, \ldots, n$, a start time, t_i^s, an end time, t_i^e, a start location, l_i^s, and an end location, l_i^e, and

- The travel time, δ_{ij}, between any two locations $i, j \in 1, \ldots, l$.

All times are in minutes and integers. The planning duration is from 5 a.m. to around midnight. Each instance has two classes of trips, short trips and long trips, with 40% short trips and 60% long trips. The duration of a short trip is less than a total of 45 min and the travel time between the start

and end location of the trip. A long trip is about 3–5 h in duration and has the same start and end location. For all instances, $m \leq l$ and the locations $1, \ldots, m$ correspond to depots, while the remaining locations only appear as trip start and end locations.

A trip j can be covered after trip i by the same vehicle, if $t_j^s \geq t_i^e + \delta_{l_i^e l_j^s}$. If $l_i^e \neq l_j^s$, the vehicle must travel empty from l_i^e to l_j^s, otherwise, the vehicle may require waiting at l_i^e for the duration of $(t_j^s - t_i^e)$. A schedule is given by the sequence in which a vehicle can cover the trips. The MDVSP is to determine the minimum number of schedules to cover all trips that minimizes total time in waiting and empty travel. The following requirements must be satisfied:

1. Each schedule should start and end at the same depot.
2. Each trip should be covered by only one vehicle.
3. The number of schedules that start from a depot should not exceed the number of vehicles at the depot.

A sufficient number of vehicles are provided to maintain the feasibility of an instance.

For each instance size (m, n), Table 1 provides the average of the number of locations, the number of times, the number of vehicles, and the number of possible empty travels, over five instances. The number of locations includes m distinct locations for depots and the number of locations at which various trips start or end. The number of times includes the start and the end time of the planning horizon and the start/end times for the trips. The number of vehicles is the total number of vehicles from all the depots. The number of possible empty travels is the number of possible connections between trips that require a vehicle travelling empty between two consecutive trips in a schedule.

The description of the file for each problem instance is presented in Table 2. The first line in the file provides the number of depots (m), the number of trips, (n), and the number of locations (l), in the problem instance. The next n lines present the information for n trips. Each line corresponds to a trip, $i \in \{1, \ldots, n\}$, and provides the start location, the start time, the end location, and the end time of trip i. The next l lines present the travel times between any two locations, $i, j \in \{1, \ldots, l\}$.

The dataset also includes a program 'GenerateInstance.cpp' that can be used to generate new instances. The program takes three inputs, the number of depots (m), the number of trips (n), and the number of instances for each size (m, n).

Table 1
Average number of locations, times, vehicles and empty travels for each instance size.

Instance size (m, n)	Average number of			
	Locations	Times	Vehicles	Possible empty travels
(8, 1500)	568.40	975.20	652.20	668,279.40
(8, 2000)	672.80	1048.00	857.20	1,195,844.80
(8, 2500)	923.40	1078.00	1082.40	1,866,175.20
(8, 3000)	977.00	1113.20	1272.80	2,705,617.00
(12, 1500)	566.00	994.00	642.00	674,191.00
(12, 2000)	732.60	1040.60	861.20	1,199,659.80
(12, 2500)	875.00	1081.00	1096.00	1,878,745.20
(12, 3000)	1119.60	1107.40	1286.20	2,711,180.40
(16, 1500)	581.80	985.40	667.80	673,585.80
(16, 2000)	778.00	1040.60	872.40	1,200,560.80
(16, 2500)	879.00	1083.20	1076.40	1,879,387.00
(16, 3000)	1087.20	1101.60	1284.60	2,684,983.60

Table 2
Description of file format for each problem instance.

Number of lines	Number of columns in each line	Description
1	3	The number of depots, the number of trips, and the number of locations.
1	m	The number of vehicles r_d at each depot d.
n	4	One line for each trip, $i = 1, 2, \ldots, n$. Each line provides the start location l_i^s, the start time t_i^s, the end location l_i^e and the end time t_i^e for the corresponding trip.
l	l	Each element, δ_{ij}, where $i, j \in 1, 2, \ldots, l$, refers to the travel time between location i and location j.

2. Experimental design, materials, and methods

The procedure presented by Carpaneto et al. in [1] is used to generate the problem instances. The same procedure has been used by Pepin et al. in [4] to generate the benchmark dataset of the MDVSP. A detailed description of the procedure is presented in [3].

Our dataset provides start/end location and time of trips as well as the travel time between any two locations. The location and time information is required to model the MDVSP on a time-space network. The feasible connections and the cost of connections between the trips can be obtained as discussed in [3]. Thus, the dataset has all the information that is required to model the MDVSP on the time-space network (see [2]) as well as the connection-network (see [5]). The benchmark solutions for all the problem instances are presented in [3].

Transparency document. Supporting information

Transparency document associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2018.12.055.

References

[1] G. Carpaneto, M. Dell'Amico, M. Fischetti, P. Toth, A branch and bound algorithm for the multiple depot vehicle scheduling problem, Networks 19 (5) (1989) 531–548.
[2] N. Kliewer, T. Mellouli, L. Suhl, A time–space network based exact optimization model for multi-depot bus scheduling, Eur. J. Oper. Res. 175 (3) (2006) 1616–1627.
[3] S. Kulkarni, M. Krishnamoorthy, A. Ranade, A.T. Ernst, R. Patil, A new formulation and a column generation-based heuristic for the multiple depot vehicle scheduling problem, Transp. Res. Part B Methodol. 118 (2018) 457–487.
[4] A.S. Pepin, G. Desaulniers, A. Hertz, D. Huisman, A comparison of five heuristics for the multiple depot vehicle scheduling problem, J. Sched. 12 (1) (2009) 17.
[5] C.C. Ribeiro, F. Soumis, A column generation approach to the multiple-depot vehicle scheduling problem, Oper. Res. 42 (1) (1994) 41–52.