




PHILINE SCHIEWE¹
MARC GOERIGK²
NIELS LINDNER³

Introducing TimPassLib – A library for integrated periodic timetabling and passenger routing

¹  0000-0002-4223-3246

²  0000-0002-2516-0129

³  0000-0002-8337-4387

Zuse Institute Berlin
Takustr. 7
14195 Berlin
Germany

Telephone: +49 30 84185-0
Telefax: +49 30 84185-125

E-mail: bibliothek@zib.de
URL: <http://www.zib.de>

ZIB-Report (Print) ISSN 1438-0064
ZIB-Report (Internet) ISSN 2192-7782

Introducing TimPassLib – A library for integrated periodic timetabling and passenger routing

Philine Schiewe*¹, Marc Goerigk², and Niels Lindner³

¹Department of Mathematics and Systems Analysis, Aalto University, Espoo, Finland

²Network and Data Science Management, University of Siegen, Siegen, Germany

³Institut für Mathematik, Freie Universität Berlin, Berlin, Germany and
Department Network Optimization, Zuse Institute Berlin, Berlin, Germany

Abstract

Classic models to derive a timetable for public transport often face a chicken-and-egg situation: A good timetable should offer passengers routes with small travel times, but the route choice of passengers depends on the timetable. While models that fix passenger routes were frequently considered in the literature, integrated models that simultaneously optimize timetables and passenger routes have seen increasing attention lately. This creates a growing need for a set of instances that allows to test and compare new algorithmic developments for the integrated problem. Our paper addresses this requirement by presenting TimPassLib, a new benchmark library of instances for integrated periodic timetabling and passenger routing.

Keywords: periodic timetabling; optimization in public transport; data sets; benchmarking

1 Introduction

Improvements in public transport systems have the potential to better the life of millions. In the sustainable development goals of the United Nations Development Programme, goal 11 (“Sustainable Cities and Communities”) states the need to “provide access to safe, affordable, accessible and sustainable transport systems for all, improving road safety, notably by expanding public transport” [UN15]. Operations research can help achieve these goals by making better use of the existing infrastructure.

One part of the process is to design timetables that allow passengers short travel times. Since the introduction of the periodic event scheduling problem (PESP) in the late 80s (see [SU89]), periodic timetables have been a major challenge to optimize. In this model, we assume that services are offered periodically (e.g., there is a train departure of a specific line from a station at 20 minutes past the hour throughout the day). Periodic timetabling models are surprisingly flexible in the aspects that can be included [LM08], and have been driving algorithmic development through the theoretical (see, e.g. [BHKL20, LR22]) and computational challenge that they pose (see, e.g., [GS13, GL17, BLR20, LL20, LL22, BLM22]). Algorithmic development requires instances that allow the research community to compare performance results. The PESPLib benchmark set [Goe12] has been offering such an opportunity in the case of pure periodic timetabling problems.

*Corresponding author. Email: philine.schiewe@aalto.fi

At an increasing rate, recent research has gone beyond the possibilities of PESP by looking at more integrated models for public transport [Sch17, Sch20, SS22]. This approach is particularly natural in the case of periodic timetabling. To find a good timetable, we need to know the routes that the passengers wish to use. But passenger routes depend on the timetable. In classic timetabling problems, we simply fix a passenger routing that is calculated in a heuristic way. This can potentially lead to solutions that are far from optimal, see [BHK17, SS20], even if we follow an iterative timetabling and passenger routing strategy, see for example the discussion in [SG13]. For this reason, integrated models need to be considered, that optimize both the timetable and the passenger routes simultaneously. In many recent publications, solution methods for integrated models have been discussed, see, e.g., [SS15, GGNS16, BHK18, LLB20, SS20].

This creates a demand for new benchmark instances that allow researchers to compare the results that solution methods can achieve. The purpose of this paper is to address this demand by introducing a new, freely available benchmark set of instances for the integrated periodic timetabling and passenger routing: the TimPassLib [SGL23].

The remainder of this paper is structured as follows. In Section 2, we give a formal definition for the integrated problem that we study. The problem instances that constitute the TimPassLib are described in Section 3. We conclude our work in Section 4 and described the data format in Appendix A.

2 Problem Definition

We first introduce the classic periodic timetabling problem where passenger routes are fixed, which is based on the periodic event scheduling problem [SU89]. We consider a set of events \mathcal{E} . For each event $i \in \mathcal{E}$ we would like to schedule a time π_i in the discrete interval $\{0, \dots, T-1\}$, where T denotes the period length. The schedule is supposed to be repeated every T time units. In timetabling, typical events are the arrival or the departure of a train at a station. Additionally, there is a set of activities $\mathcal{A} \subseteq \mathcal{E} \times \mathcal{E}$ that connect events with each other. Such an activity may model a train driving from one station to the next, a train waiting at a station for passengers to embark or disembark, or a group of passengers changing from one train arrival to the departure of another train. Another class of activities, which is not related to passengers, could model headway between trains or synchronization of lines running more frequently than once every T time units. Events and activities together result in the event-activity network $\mathcal{N} = (\mathcal{E}, \mathcal{A})$. Each activity $a \in \mathcal{A}$ has a lower bound $\ell_a \in \mathbb{N}$ and upper bound $u_a \in \mathbb{N}$ that reflect requirements on minimum and maximum durations of activities. We denote by y_a the slack of an activity, i.e., difference of an activity duration to its lower bound. Finally, we assume that some weight w_a is known for each activity, which represents the number of passengers who wish to use this activity.

The periodic timetabling problem can then be formulated as the following optimization problem:

$$\begin{aligned}
\min \quad & \sum_{a \in \mathcal{A}} w_a (y_a + \ell_a) \\
\text{s.t.} \quad & y_a = [\pi_j - \pi_i - \ell_a]_T & \forall a = (i, j) \in \mathcal{A}, \\
& 0 \leq y_a \leq u_a - \ell_a & \forall a \in \mathcal{A}, \\
& \pi_i \in \{0, \dots, T-1\} & \forall i \in \mathcal{E}.
\end{aligned}$$

Solution approaches based on mixed-integer programming typically focus on one set of variables; either the timetable variables π_i (node-based formulations) or the slack variables y_a (cycle-based formulations) [Lie06]. The symbol $[\cdot]_T$ denotes the modulo bracket $[x]_T = \min\{x + zT : x + zT \geq 0, z \in \mathbb{Z}\}$ and is usually modeled by introducing additional integer variables.

In the *integrated* periodic timetabling and passenger routing problem, we choose to model the activity weights w_a in more detail. We assume that an origin-destination (OD) matrix is given, where every entry d_{st} denotes the number of passengers who wish to travel from origin station $s \in V$ to destination station $t \in V$, where V denotes the set of all stations. Note that multiple events in the event-activity network can be assigned to the same station. Let P_{st} denote all simple paths in \mathcal{N} that connect an event corresponding to station s with an event corresponding to station t , and use exclusively drive, wait or change activities as described above. For each path $p \in P_{st}$, let f_p denote the fraction of passengers that travel from s to t along p . This means that for each OD-pair (s, t) with $d_{st} > 0$, we require

$$\sum_{p \in P_{st}} f_p = 1.$$

The passenger weights w_a for each activity $a \in \mathcal{A}$ are then determined as

$$w_a = \sum_{(s,t) \in V \times V : d_{st} > 0} \sum_{p \in P_{st} : a \in p} d_{st} f_p \quad \forall a \in \mathcal{A}.$$

From a passenger's perspective, there are more criteria which connection to choose, apart from just the travel time. Another frequently considered criterion is the number of transfers along a journey [BDG⁺16]. This can be included in this model by using an additional penalty if passengers need to use a change activity, which represents the discomfort in comparison to a direct connection. Let c denote this penalty value and let $\mathcal{A}_{\text{change}} \subseteq \mathcal{A}$ denote the set of change activities in the event-activity network. The objective function is extended by the term

$$\sum_{a \in \mathcal{A}_{\text{change}}} c w_a$$

to include change penalties.

The integrated periodic timetabling and passenger routing problem is now to find both the timetable (π, y) as well as passenger routes and weights (f, w) simultaneously. The resulting model is thus non-linear because of the bilinear term in the objective. Observe that we can assume that variables f_p are binary, as there is always an optimal solution where passengers are routed along a single shortest path, since the model does not contain capacity constraints. Hence, the overall model for our problem can be summarized as follows [Lü09, SG13, BHK17, LLB20]:

$$\begin{aligned} \min \quad & \sum_{a \in \mathcal{A}} w_a (y_a + \ell_a) + \sum_{a \in \mathcal{A}_{\text{change}}} c w_a \\ \text{s.t.} \quad & y_a = [\pi_j - \pi_i - \ell_a] T & \forall a = (i, j) \in \mathcal{A}, \\ & \sum_{p \in P_{st}} f_p = 1 & \forall (s, t) \in V \times V : d_{st} > 0, \\ & w_a = \sum_{\substack{(s,t) \in V \times V \\ d_{st} > 0}} \sum_{p \in P_{st} : a \in p} d_{st} f_p & \forall a \in \mathcal{A}, \\ & 0 \leq y_a \leq u_a - \ell_a & \forall a \in \mathcal{A}, \\ & \pi_i \in \{0, \dots, T-1\} & \forall i \in \mathcal{E}, \\ & f_p \in \{0, 1\} & \forall (s, t) \in V \times V : d_{st} > 0, p \in P_{st}. \end{aligned} \quad (\star)$$

3 Data sets

All benchmark data sets presented in this paper are available at <https://timpasslib.net> [SGL23]. The file format is described in Appendix A. The website collects the best known solutions and bounds for each data set. A potential new solution can be checked by means of a verification script written in Python, which also computes an optimal routing for a given timetable.

In the remainder of this section, we briefly describe the TimPassLib data sets. Table 1 presents an overview of some key features of the instances. The columns contain the following information:

instance: The name of the instance. More information can be found in the following subsections.

stations: The number of stations.

lines: The number of operated lines. Note that a line can have a frequency higher than one.

OD pairs: The number of OD pairs with $d_{st} > 0$.

OD total: The total number of passengers.

events: The number of events in the event-activity network.

activities fixed: The total number of activities in the event-activity network.

activities fixed: The number of activities in the event-activity network with $\ell_a = u_a$.

activities free: The number of activities in the event-activity network with $u_a - \ell_a = T - 1$.

activities restricted (restr.): The number of activities in the event-activity network with $\ell_a < u_a < \ell_a + T - 1$.

reference objective (ref. obj.): The objective value w.r.t. (\star) of the best timetable computed by the concurrent PESP solver by [BLR20] within a wall time limit of one hour. The computations have been executed on an Intel Xeon CPU E3-1270 v6 running at 3.80 GHz with 32 GB RAM using Gurobi 9.5.2 [Gur23] as MIP solver. The weights for the periodic timetabling problem have been obtained from a passenger routing according to lower bounds ℓ_a and change penalty c . The aim of this reference objective is to provide initial solutions obtained with uniform computational power.

lower bound A lower bound on the objective value of (\star) obtained by routing all passengers on shortest paths w.r.t. lower activity bounds ℓ_a and change penalty c .

gap The gap per passenger between reference objective and lower bound:

$$\frac{\text{reference objective} - \text{lower bound}}{\text{OD total}}.$$

This gap can be used to estimate the gains of solving the integrated problem.

3.1 hamburg

The instance *hamburg* models the suburban commuter rail network of S-Bahn Hamburg. The infrastructure is mostly independent from other railways and has only few single-track sections. The network is operated with six lines, where one of the lines has two branches. The period length is 10, as all lines run every 10 minutes in the rush hour. Bounds for travel and dwell times are derived from the annual timetable for 2023. The passenger demand is based on data from trains with sensors for automatic passenger counts, which is publicly available on the open data portal of Deutsche Bahn AG [Deu17].

3.2 schweiz_fernverkehr

The instance *schweiz_fernverkehr* is an excerpt of the ICE, TGV, InterCity and InterRegio trains within Switzerland and contains 80 unidirectional trips. The period length is 120 minutes. The data is based on GTFS timetable data for 2023 and on station passenger counts, which are both publicly available on the open data platform of the Swiss federal office of transport [SBB23].

3.3 toy and toy_2

The instances *toy* and *toy_2* are based on a small artificial data set in the software library LinTim [SAS⁺22]. The line concepts are generated algorithmically and consist of two and six lines, respectively. The period length is 60. Note that the same instances are used in [SS20].

3.4 regional

The instance *regional* is based on the regional train network in Lower Saxony, Germany, and is available as part of the software library LinTim [SAS⁺22]. The line concept is generated algorithmically and consists of 8 lines. The period length is 60. Note that the same instance is used in [SS20].

3.5 grid

The instance *grid* is a benchmark data set, originally introduced in [FHSS17] and available as part of the open source data set [FOR18] as well as the software library LinTim [SAS⁺22]. The line concept is generated algorithmically and consists of 8 lines. The period length is 60. Note that the same instance is used in [SS20].

3.6 long-distance

The instance *long-distance* is inspired by the long-distance train network in Germany and part of the developer version of software library LinTim [SAS⁺22]. The line concept is generated algorithmically and consists of 42 lines. The period length is 60. Note that the same instance is used in [SS20].

3.7 metro

The instance *metro* is based on the metro system in Athens, Greece, and is available as part of the software library LinTim [SAS⁺22]. The line concept consists of 4 lines. The period length is 150 and the timetable is planned in increments of 6 seconds. Note that the same instance is used in [SS20].

3.8 Erding_NDP_S020 and Erding_NDP_S021

The instances *Erding_NDP_S020* and *Erding_NDP_S021* are based on the transport supply in Erding, Germany, and are available as part of an open source data set [FOR18]. The corresponding line concepts consist of 21 lines and the period length is 60.

3.9 stuttgart

The instance *stuttgart* is based on the transport supply in Stuttgart, Germany, and is available as part of an open source data set [FOR18]. The corresponding line concept consists of 156 lines and the period length is 3600, i.e., the timetable is planned in seconds.

3.10 RxLy

This group consists of 16 instances, numbered from *R1L1* to *R4L4*. They are an extension of the 16 core instances of the PESPlib [Goe12], which in turn was based on long-distance train network data in Germany from the software library LinTim [SAS⁺22]. The numbers of lines are in the range from 54 up to 133, with a period length of 60.

4 Outlook

Integrating passenger routing into periodic timetabling is a natural step to resolve the mutual interdependence of both problems. For more than ten years, the PESPlib library [Goe12] has stimulated a variety of research on periodic timetabling with fixed routing. As of now, recent improvements are only marginal, disregard passenger flow, and therefore do not necessarily improve passenger comfort. With the initiation of TimPassLib, we hope to foster the investigation of periodic timetabling with passenger routing, to extend the scope of the PESPlib library to a solid base of benchmark instances for the integrated problem, and to contribute to the attractiveness of public transport with mathematical optimization methods. We are therefore very much looking forward to submissions of new solutions or lower bounds as well as new benchmark instances.

Acknowledgements

We want to thank Alexander Schiewe for his help in setting up the website <https://timpasslib.net>.

Declarations

Ethical Approval Not applicable.

Competing Interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors' Contribution Philine Schiewe: Conceptualization, Methodology, Software, Investigation, Validation, Formal Analysis, Data Curation, Writing – Original Draft, Writing – Review & Editing; Marc Goerigk: Conceptualization, Methodology, Investigation, Data Curation,

instance	OD				activities				ref. obj.	lower bound	gap	
	stations	lines	pairs	total	total	events	total	fixed				free
hamburg	68	7	2030	9694166	508	823	449	284	90	141629305	139892927	0.18
schweiz_fernverkehr	140	80	12082	1347686	2234	18467	1563	14787	2117	63273307	60084289	2.37
toy	8	2	46	2622	32	44	2	16	26	21714	20194	0.58
toy_2	8	6	46	2622	156	1088	112	868	108	19316	19114	0.08
regional	34	8	330	325968	412	1520	176	958	386	2068934	1804642	0.81
grid	25	8	567	2546	392	2382	176	1842	364	50692	47824	1.13
long-distance	250	42	6106	385868	2800	14958	2716	12242	0	67697819	64906980	7.23
metro	51	4	2385	63323	964	3240	708	2050	482	12021720	11978129	0.69
Erding_NDP_S020	51	21	675	558164	1132	5300	320	3944	1036	12291729	12206083	0.15
Erding_NDP_S021	51	21	675	558164	1132	5396	320	3944	1132	12307765	12206083	0.18
stuttgart	560	156	132439	39517136	21328	297094	22856	268606	5632	49009914897	45072189100	99.65
R1L1	250	55	48842	3147382	3664	6381	642	2827	2912	542908145	522575407	6.46
R1L2	250	54	48842	3147382	3668	6543	632	2983	2928	542381697	522212362	6.41
R1L3	250	65	48842	3147382	4184	7025	752	2971	3302	543067240	522199838	6.63
R1L4	250	71	48842	3147382	4760	8528	830	3910	3788	537879494	520799059	5.43
R2L1	280	66	61110	3666720	4156	7356	814	3332	3210	681061389	650575045	8.31
R2L2	280	66	61110	3666720	4204	7561	820	3489	3252	676836085	650293220	7.24
R2L3	280	80	61110	3666720	5048	8285	970	3397	3918	677126860	649767761	7.46
R2L4	280	116	61110	3666720	7660	13168	1496	5740	5932	667537183	647184195	5.55
R3L1	296	73	68284	3878392	4516	9140	794	4770	3576	694086648	665804283	7.29
R3L2	296	70	68284	3878392	4452	9249	774	4937	3538	694334373	665719574	7.38
R3L3	296	95	68284	3878392	5724	11165	1038	5631	4496	691688857	665595680	6.73
R3L4	296	120	68284	3878392	8180	15655	1478	7715	6462	682090300	662251432	5.12
R4L1	319	86	77878	4183088	4932	10262	996	5502	3764	755138437	723276168	7.62
R4L2	319	89	77878	4183088	5048	10733	984	5863	3886	754453547	724254447	7.22
R4L3	319	115	77878	4183088	6368	13236	1240	7098	4898	751351849	722434044	6.91
R4L4	319	133	77878	4183088	8384	17753	1572	9635	6546	738792466	720103154	4.47

Table 1: Overview of key features for the data sets.

Writing – Original Draft; Niels Lindner: Conceptualization, Methodology, Software, Validation, Investigation, Data Curation, Writing – Review & Editing

Funding Not applicable.

Availability of data and materials The TimPassLib data set is publicly available at <https://timpasslib.net>.

References

- [BDG⁺16] Hannah Bast, Daniel Delling, Andrew Goldberg, Matthias Müller-Hannemann, Thomas Pajor, Peter Sanders, Dorothea Wagner, and Renato F Werneck. Route planning in transportation networks. *Algorithm engineering: Selected results and surveys*, pages 19–80, 2016.
- [BHK17] Ralf Borndörfer, Heide Hoppmann, and Marika Karbstein. Passenger routing for periodic timetable optimization. *Public Transport*, 9:115–135, 2017.
- [BHKL18] Ralf Borndörfer, Heide Hoppmann, Marika Karbstein, and Fabian Löbel. The modulo network simplex with integrated passenger routing. In *Operations Research Proceedings 2016: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR), Helmut Schmidt University Hamburg, Germany, August 30-September 2, 2016*, pages 637–644. Springer, 2018.
- [BHKL20] Ralf Borndörfer, Heide Hoppmann, Marika Karbstein, and Niels Lindner. Separation of cycle inequalities in periodic timetabling. *Discrete Optimization*, 35:100552, February 2020.
- [BLM22] Enrico Bortoletto, Niels Lindner, and Berenike Masing. Tropical Neighbourhood Search: A New Heuristic for Periodic Timetabling. In Mattia D’Emidio and Niels Lindner, editors, *22nd Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2022)*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.
- [BLR20] Ralf Borndörfer, Niels Lindner, and Sarah Roth. A concurrent approach to the periodic event scheduling problem. *Journal of Rail Transport Planning & Management*, 15:100175, 2020.
- [Deu17] Deutsche Bahn AG. Passagierzählung S-Bahn Hamburg. <https://data.deutschebahn.com/dataset/passagierzahlung-s-bahn-hamburg.html>, 2017. Accessed: 2023-01-23.
- [FHSS17] Markus Friedrich, Maximilian Hartl, Alexander Schiewe, and Anita Schöbel. Angebotsplanung im öffentlichen Verkehr - planerische und algorithmische Lösungen. In *Heureka’17*, 2017.
- [FOR18] Collection of open source public transport networks by DFG Research Unit “FOR 2083: Integrated Planning For Public Transportation”. <https://github.com/FOR2083/PublicTransportNetworks>, 2018.

- [GGNS16] Philine Gattermann, Peter Großmann, Karl Nachtigall, and Anita Schöbel. Integrating passengers’ routes in periodic timetabling: a sat approach. In *16th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2016)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.
- [GL17] Marc Goerigk and Christian Liebchen. An Improved Algorithm for the Periodic Timetabling Problem. In Gianlorenzo D’Angelo and Twan Dollevoet, editors, *17th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2017)*, volume 59 of *OpenAccess Series in Informatics (OASICs)*, pages 12:1–12:14, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. ISSN: 2190-6807.
- [Goe12] Marc Goerigk. PESplib – a benchmark library for periodic event scheduling. <http://num.math.uni-goettingen.de/~m.goerigk/pesplib/>, 2012. Accessed: 2023-01-26.
- [GS13] Marc Goerigk and Anita Schöbel. Improving the modulo simplex algorithm for large-scale periodic timetabling. *Computers & Operations Research*, 40(5):1363–1370, 2013.
- [Gur23] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2023.
- [Lie06] Christian Liebchen. *Periodic timetable optimization in public transport*. PhD thesis, Berlin, 2006.
- [LL20] Niels Lindner and Christian Liebchen. Determining All Integer Vertices of the PESP Polytope by Flipping Arcs. In Dennis Huisman and Christos D. Zaroliagis, editors, *20th Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2020)*, volume 85 of *OpenAccess Series in Informatics (OASICs)*, pages 5:1–5:18, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. ISSN: 2190-6807.
- [LL22] Niels Lindner and Christian Liebchen. Timetable merging for the periodic event scheduling problem. *EURO Journal on Transportation and Logistics*, 11:100081, 2022.
- [LLB20] Fabian Löbel, Niels Lindner, and Ralf Borndörfer. The restricted modulo network simplex method for integrated periodic timetabling and passenger routing. In *Operations Research Proceedings 2019: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR), Dresden, Germany, September 4-6, 2019*, pages 757–763. Springer, 2020.
- [LM08] Christian Liebchen and Rolf H Möhring. The modeling power of the periodic event scheduling problem: railway timetables—and beyond. In *Computer-aided Systems in Public Transport*, pages 117–150. Springer, 2008.
- [LR22] Niels Lindner and Julian Reisch. An analysis of the parameterized complexity of periodic timetabling. *Journal of Scheduling*, 25(2):157–176, 2022.
- [Lü09] Jan Lübbecke. *Passagierrouting und Taktfahrplanoptimierung*, 2009. Diploma thesis, Technische Universität Berlin.
- [SAS⁺22] Alexander Schiewe, Sebastian Albert, Philine Schiewe, Anita Schöbel, Felix Spühler, and Moritz Stinzenhöfer. Documentation for LinTim 2022.08. 2022.

- [SBB23] SBB AG. Open data platform mobility Switzerland. <https://opentransportdata.swiss/en>, 2023. Accessed: 2023-01-23.
- [Sch17] Anita Schöbel. An eigenmodel for iterative line planning, timetabling and vehicle scheduling in public transportation. *Transportation Research Part C: Emerging Technologies*, 74:348–365, 2017.
- [Sch20] P. Schiewe. *Integrated Optimization in Public Transport Planning*, volume 160 of *Optimization and Its Applications*. Springer, 2020.
- [SG13] Michael Siebert and Marc Goerigk. An experimental comparison of periodic timetabling models. *Computers & Operations Research*, 40(10):2251–2259, October 2013.
- [SGL23] Philine Schiewe, Marc Goerigk, and Niels Lindner. TimPassLib – A library for integrated periodic timetabling and passenger routing. <https://timpasslib.net>, 2023.
- [SS15] Marie Schmidt and Anita Schöbel. Timetabling with passenger routing. *OR spectrum*, 37(1):75–97, 2015.
- [SS20] Philine Schiewe and Anita Schöbel. Periodic timetabling with integrated routing: Toward applicable approaches. *Transportation Science*, 54(6):1714–1731, 2020.
- [SS22] Philine Schiewe and Anita Schöbel. Integrated optimization of sequential processes: General analysis and application to public transport. *EURO Journal on Transportation and Logistics*, 11:100073, 2022.
- [SU89] Paolo Serafini and Walter Ukovich. A mathematical model for periodic scheduling problems. *SIAM Journal on Discrete Mathematics*, 2(4):550–581, 1989.
- [UN15] Sustainable Development Goals — United Nations Development Programme. <https://www.undp.org/sustainable-development-goals#sustainable-cities-and-communities>, 2015. Accessed: 2023-01-30.

A Data Format

In- and output data for the benchmark instances is given in `csv` files that are designed to be easily readable. Each file starts with a header that describes the content of the columns.

A.1 Input File `Config.csv`

The file `Config.csv` contains additional information needed for the data set.

```
# config_key; value
ptn_name; toy
period_length; 60
ean_change_penalty; 5
```

The file contains a `value` for each following `config_key`:

`ptn_name`: name of the data set (string)
`period_length`: period length T (integer)
`ean_change_penalty`: change penalty c (integer)

A.2 Input File `OD.csv`

The file `OD.csv` contains information about the OD pairs.

```
# origin; destination; customers
1; 2; 10
1; 3; 20
1; 4; 10
1; 5; 60
...
```

Here, `origin` gives the id of the start station s of the OD pair, `destination` the corresponding destination t and `customers` the number of passengers d_{st} who wish to travel from s to t . Note that all values are integers.

A.3 Input File `Events.csv`

The file `Events.csv` contains information about the events \mathcal{E} as events represent arrival and departures of trips at stations.

```
# event_id; type; stop_id; line_id; line_direction; line_freq_repetition
1; "departure"; 2; 2; >; 1
2; "arrival"; 3; 2; >; 1
3; "departure"; 3; 2; >; 1
4; "arrival"; 6; 2; >; 1
...
```

`event_id`: id of the event (integer)

`type`: whether the event represents the arrival ("`arrival`") or departure ("`departure`") of the corresponding trip

`stop_id`: id of the corresponding station (integer)

`line_id`: id of the corresponding line (integer)
`line_direction`: whether the trip belongs to the line in forward ($>$) or backward ($<$) direction
`line_freq_repetition`: repetition of the line in the planning period to which the trip belongs (integer)

A.4 Input File `Activities.csv`

The file `Activities.csv` contains information about the activities.

```
# activity_index; type; from_event; to_event; lower_bound; upper_bound
1; "drive"; 1; 2; 3; 4
2; "wait"; 2; 3; 1; 3
3; "drive"; 3; 4; 1; 3
4; "wait"; 4; 5; 1; 3
...
```

`activity_index`: id of the activity (integer)

`type`: type of the activity. Possible values are:

- "drive" represents a vehicle driving from one stop of a line to the next (can be used for routing passengers)
- "wait" represents a vehicle waiting at a stop (can be used for routing passengers)
- "change" represents a transfer between two lines at the same stop (can be used for routing passengers)
- "sync" used for synchronizing trips of the same line (cannot be used for routing passengers)
- "headway" represents headway constraint between vehicles using the same infrastructure (cannot be used for routing passengers)

`from_event`: id of event i for activity (i, j) (integer)

`to_event`: id of event j for activity (i, j) (integer)

`lower_bound`: lower bound ℓ_a of activity a (integer)

`upper_bound`: upper bound u_a of activity a (integer)

A.5 Solution File `Timetable.csv`

The solution file `Timetable.csv` should be given in the following format.

```
# event_id; time
1; 0
2; 3
3; 4
4; 5
...
```

`event_id`: id of the event (integer)

`time`: time of the event ($\{0, \dots, T - 1\}$)

Since an optimal routing for a given timetable can be computed efficiently by solving a series of shortest path problems, passenger routing data is not required when submitting a solution.