Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 22–41

AN INTEGRATED ASSIGNMENT, ROUTING, AND SPEED MODEL FOR ROADWAY MOBILITY AND TRANSPORTATION WITH ENVIRONMENTAL, EFFICIENCY, AND SERVICE GOALS

T. GIOVANNELLI AND L. N. VICENTE

Abstract: Managing all the mobility and transportation services with autonomous vehicles for users of a smart city requires determining the assignment of the vehicles to the users and their routing in conjunction with their speed. Such decisions must ensure low emission, efficiency, and high service quality by also considering the impact on traffic congestion caused by other vehicles in the transportation network.

In this paper, we first propose an abstract trilevel multi-objective formulation architecture to model all vehicle routing problems with assignment, routing, and speed decision variables and conflicting objective functions. Such an architecture guides the development of subproblems, relaxations, and solution methods. We also propose a way of integrating the various urban transportation services by introducing a constraint on the speed variables that takes into account the traffic volume caused across the different services. Based on the formulation architecture, we introduce a (bilevel) problem where assignment and routing are at the upper level and speed is at the lower level. To address the challenge of dealing with routing problems on urban road networks, we develop an algorithm that alternates between the assignment-routing problem on an auxiliary complete graph and the speed optimization problem on the original non-complete graph. The computational experiments show the effectiveness of the proposed approach in determining approximate Pareto fronts among the conflicting objectives.

Keywords: Assignment-routing-speed optimization; Urban road networks; Environmental, efficiency, and service goals; Integrated transportation system.
Math. Subject Classification (2010): 90C11, 90C29, 90C90.

1. Introduction

In this paper, we propose an innovative integrated transportation model for the management of possibly all vehicles traveling on the streets and roads of a city, which are assumed to have different levels of autonomy (with or without a driver). Our main goal is to provide an optimization model that can be used to effectively manage mobility and transportation within a city by adopting a green logistics perspective and pursuing efficiency and service

Received December 06, 2022.
Support for LNV was partially provided by the Centre for Mathematics of the University of Coimbra under grant FCT/MCTES UIDB/MAT/00324/2020.

quality. The integrated model introduced in this work will qualify the cities implementing the proposed transportation network as smart.

The achievement of our goal will be enabled by the recent (and future) advances in information and location-sensitive technologies, which facilitate acquiring the necessary high-quality (granular) data for supporting the decision-making process for vehicles traveling on the streets and roads of a smart city. Several definitions of smart city have been proposed in the literature (see, e.g., [46], [11], and [62]) and a universal characterization cannot be provided since smart cities involve different types of features that vary based on the specific context considered [33]. In general, the phrase *smart city* refers to cities provided with technological infrastructure based on advanced data processing that pursue several goals, such as a more efficient city governance, a better life quality for citizens, increasing economic success for businesses, and a more sustainable environment [64].

In our paper, we focus on smart cities with intelligent transportation systems (see [63] for a survey). In particular, all the smart cities provided with technologies for real-time transportation data acquisition fit within the scope of our work. However, we point out that the process leading to the acquisition of such technologies, which requires a multitude of social, political, and economic factors involving public-private partnerships and local authorities, is omitted from this paper. The importance of the role played by mobility and transportation in smart cities is highlighted in many works [64]. For instance, in [58], four different groups of smart cities are identified based on a cluster analysis of cities around the world, and one of them gathers cities adopting smart transportation systems. The goal of such systems is to control traffic congestion by public transportation, car sharing, and self-driving cars. In [2], the urban congestion is seen as a challenge arising from the persistent need of citizens to use their private cars. A solution proposed to alleviate such a problem requires smart control of the traffic in the existing road infrastructure to ensure a sustainable transportation, which is exactly the goal of our paper.

A crucial tool for managing transportation in smart cities is the concept of urban artificial intelligence (AI) [17]. Among the examples of urban AI, self-driving cars and city brains represent two important categories. In particular, the author points out that the number of cities where autonomous cars are allowed to drive is increasing (see also [1]), and these also include cars with the highest level of autonomy, when no human input or supervision

[58] Z. Tang, K. Jayakar, X. Feng, H. Zhang, and R. X. Peng. Identifying smart city archetypes from the bottom up: A content analysis of municipal plans. *Telecommunications Policy*, 43(10), 2019.

[59] T. Vidal, T.G. Crainic, M. Gendreau, and C. Prins. A unified solution framework for multi-attribute vehicle routing problems. *European Journal of Operational Research*, 234:658 – 673, 2014.

[60] T. Vidal, G. Laporte, and P. Matl. A concise guide to existing and emerging vehicle routing problem variants. *European Journal of Operational Research*, 286:401 – 416, 2020.

[61] A. Wächter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.*, 106:25–57, mar 2006.

[62] R. Wenge, X. Zhang, C. Dave, L. Chao, and S. Hao. Smart city architecture: A technology guide for implementation and design challenges. *China Communications*, 11:56–69, 2014.

[63] Z. Xiong, H. Sheng, W. Rong, and D. Cooper. Intelligent transportation systems for smart cities: A progress review. *Science China Information Sciences*, 55, 12 2012.

[64] C. Yin, Z. Xiong, H. Chen, J. Wang, D. Cooper, and B. David. A literature survey on smart cities. *Science China Information Sciences*, 58:1–18, 2015.

[65] K. G. Zografos and K. N. Androutsopoulos. A decision support system for integrated hazardous materials routing and emergency response decisions. *Transportation Research Part C: Emerging Technologies*, 16:684–703, 2008.

T. Giovannelli
Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA 18015-1582, USA (tog220@lehigh.edu).

L. N. Vicente
Department of Industrial and Systems Engineering, Lehigh University, 200 West Packer Avenue, Bethlehem, PA 18015-1582, USA and Centre for Mathematics of the University of Coimbra (CMUC), Portugal (lnv@mat.uc.pt).

[39] B. Li. Optimization of people and freight transportation : pickup and delivery problem variants. *Ph.D. Thesis, TU/e Eindhoven University of Technology, Eindhoven, The Netherlands*, 2016.
[40] Y. Ma and J. Xu. Vehicle routing problem with multiple decision-makers for construction material transportation in a fuzzy random environment. *International Journal of Civil Engineering*, 12:332–346, 06 2014.
[41] Y. Marinakis and M. Marinaki. A bilevel genetic algorithm for a real life location routing problem. *International Journal of Logistics: Research and Applications*, 11:49–65, 02 2008.
[42] Y. Marinakis, A. Migdalas, and P. Pardalos. A new bilevel formulation for the vehicle routing problem and a solution method using a genetic algorithm. *Journal of Global Optimization*, 38:555–580, 08 2007.
[43] V. Markevicius, D. Navikas, A. Idzkowski, A. Valinevicius, M. Zilys, and D. Andriukaitis. Vehicle speed and length estimation using data from two anisotropic magneto-resistive (amr) sensors. *Sensors*, 17:1–13, 08 2017.
[44] K. Miettinen. *Nonlinear Multiobjective Optimization*, volume 12. Springer Science & Business Media, New York, 2012.
[45] A. V. Mugayskikh, V. V. Zakharov, and T. Tuovinen. Time- dependent multiple depot vehicle routing problem on megapolis network under wardrop's traffic flow assignment. *2018 22nd Conference of Open Innovations Association (FRUCT)*, pages 173–178, 2018.
[46] T. Nam and T. A. Pardo. Conceptualizing smart city with dimensions of technology, people, and institutions. In *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*, page 282–291, New York, NY, USA, 2011. Association for Computing Machinery.
[47] M. Nasri, T. Bektaş, and G. Laporte. Route and speed optimization for autonomous trucks. *Computers & Operations Research*, 100, 07 2018.
[48] J. Oyola, H. Arntzen, and D. Woodruff. The stochastic vehicle routing problem, a literature review, part i: models. *EURO Journal on Transportation and Logistics*, 7, 10 2016.
[49] J. Oyola, H. Arntzen, and D. Woodruff. The stochastic vehicle routing problem, a literature review, part ii: solution methods. *EURO Journal on Transportation and Logistics*, 6, 10 2016.
[50] J. Paszkowski, M. Herrmann, M. Richter, and A. Szarata. Modelling the effects of traffic-calming introduction to volume–delay functions and traffic assignment. *Energies*, 14, 2021.
[51] L. Perron and V. Furnon. Or-tools.
[52] V. Pillac, M. Gendreau, C. Guéret, and A.L. Medaglia. A review of dynamic vehicle routing problems. *European Journal of Operational Research*, 225:1 – 11, 2013.
[53] J. Qian and R. Eglese. Finding least fuel emission paths in a network with time-varying speeds. *Networks*, 63:96–106, 2014.
[54] R. Raeesi and K. G. Zografos. The multi-objective steiner pollution-routing problem on congested urban road networks. *Transportation Research Part B: Methodological*, 122:457–485, 2019.
[55] U. Ritzinger, J. Puchinger, and R.F. Hartl. A survey on dynamic and stochastic vehicle routing problems. *International Journal of Production Research*, 54:215–231, 2016.
[56] J. Rodríguez-Pereira, E. Fernández, G. Laporte, E. Benavent, and A. Martínez-Sykora. The steiner traveling salesman problem and its extensions. *European Journal of Operational Research*, 278:615–628, 2019.
[57] I. Sung and P. Nielsen. Speed optimization algorithm with routing to minimize fuel consumption under time-dependent travel conditions. *Production & Manufacturing Research*, 8:1–19, 2020.

are required. A city brain is a digital platform applied to the management of a city, including urban transportation, where the goal is to control traffic lights and flows of vehicles by using advanced data collected throughout the city. In this paper, we assume that our integrated model is used by a city brain for the urban transportation management of autonomous vehicles. However, the development of the features of such a platform is left for future work.

The presence of autonomous vehicles in a transportation network allows for the determination of the assignment of vehicles to users and their routing in conjunction with their speed, thus increasing the complexity of the decision-making process, which is naturally formulated as a hierarchy of three levels (assignment, routing, and speed). Transportation and mobility solutions should be determined with low environmental impact but, at the same time, with high efficiency for service providers and high service quality for users, leading to conflicting goals which need to be optimized simultaneously. Moreover, decisions unilaterally made in one component or part of the network can have a strong impact on the overall system. Consequently, the optimal solutions of the single components considered separately are different from the solutions obtained when such interaction is taken into account, thus requiring a proper integration of the network components. Finally, the current modeling techniques for assignment-routing problems are based on assumptions that are not satisfied on urban road networks and, therefore, require innovative solution methodologies. We aim at developing a future-oriented *integrated assignment, routing, and speed system for roadway mobility and transportation with environmental, efficiency, and service goals*, and this paper represents an instrumental building block towards this main goal and where we make the following three main contributions.

A trilevel multi-objective formulation architecture for vehicle routing problems with speed optimization. We propose an innovative formulation architecture to decompose every *Vehicle Routing Problem* (VRP) with speed optimization into three levels, each addressing one of the following vehicle-related decisions: assignment to users, routing to accommodate all the requests, and speed along each segment of the route*. In particular, our

*Note that although controlling the speed in roadway transportation is considered an impractical task (see, e.g., [60]), considering smart and autonomous vehicles allows one to take into account speed decisions.

paper extends the bilevel formulation developed by [42] to propose a trilevel multi-objective formulation for a VRP with speed optimization (referred to as a VRP/speed problem). Such an architecture allows a comprehensive understanding of the overall problem complexity, guides the development of subproblems, relaxations, and solution methods, and provides new insights into VRP/speed problems. In this paper, we use the trilevel multi-objective architecture to formulate a (bilevel) VRP/speed problem (where upper level is assignment-routing and lower level is speed), develop a corresponding optimization method, and identify and report the trade-offs among the three considered goals.

Integration among all the different transportation problem components in a smart city. The transportation services arising in an urban transportation network can be modeled through a proper VRP variant. All the frameworks, models, and methods proposed in the literature address the different transportation services arising in a city independently, without any attempt to consider such services as parts of the same system. Several studies in the literature considered modeling a general framework accounting for as many transportation services as possible. For example, [59] proposed a unified model capable of separately describing different transportation *problem components* (later referred to as *components*). However, the goal of [59] was to propose a general-purpose optimization algorithm that can quickly provide efficient solutions to different problems, each related to a transportation service. In contrast, our paper aims to develop a new general model, where different transportation components (such as personal trips, freight transportation, ride-sharing, car-pooling, dial-a-ride, and vehicle sharing) can be integrated into a comprehensive optimization framework. In this regard, we propose a VRP/speed formulation for a specific problem component, which is sufficient to fully address the integration among all the components. In particular, in such a formulation, an innovative constraint on the speed variables is used to model the impact of the traffic congestion caused by routing decisions (made in other components) on the component under consideration.

Use of non-complete graphs for vehicle routing problems with speed optimization. Using non-complete graphs is essential to address a VRP/speed problem on urban road networks. However, to find the value

[20] E. W. Dijkstra. A note on two problems in connexion with graphs. *Numer. Math.*, 1:269–271, dec 1959.
[21] B. Donovan. Link level traffic estimates. 2015.
[22] M. Ehrgott. *Multicriteria Optimization*, volume 491. Springer Science & Business Media, Berlin, 2005.
[23] R. Eshtehadi, M. Fathian, and E. Demir. Robust solutions to the pollution-routing problem with demand and travel time uncertainty. *Transportation Research Part D: Transport and Environment*, 51:351 – 363, 2017.
[24] K. Fagerholt, G Laporte, and I. Norstad. Reducing fuel emissions by optimizing speed on shipping routes. *Journal of the Operational Research Society*, 61:523–529, 03 2010.
[25] B. Fleischmann. A cutting plane procedure for the travelling salesman problem on road networks. *European Journal of Operational Research*, 21(3):307–317, 1985.
[26] R. Fukasawa, Q. He, F. Santos, and Y. Song. A joint vehicle routing and speed optimization problem. *INFORMS Journal on Computing*, 30:694–709, 2018.
[27] T. Garaix, C. Artigues, D. Feillet, and D. Josselin. Vehicle routing problems with alternative paths: An application to on-demand transportation. *European Journal of Operational Research*, 204:62–75, 2010.
[28] A. García Nájera and J. Bullinaria. Bi-objective optimization for the vehicle routing problem with time windows: Using route similarity to enhance performance. volume 5467, pages 275–289, 04 2009.
[29] V. Ghilas, E. Demir, and T. van Woensel. Integrating passenger and freight transportation: model formulation and insights. BETA publicatie : working papers. Technische Universiteit Eindhoven, 2013.
[30] K. Ghoseiri and S. Ghannadpour. Multi-objective vehicle routing problem with time windows using goal programming and genetic algorithm. *Applied Soft Computing*, 10:1096–1107, 09 2010.
[31] A. Gupta, Y. Ong, A. Zhang, P. Tan, H. Handa, H. Ishibuchi, and K.C. Tan. A bi-level evolutionary algorithm for multi-objective vehicle routing problems with time window constraints. *Proceedings of the 18th Asia Pacific symposium on intelligent and evolutionary systems—Volume 2. Proceedings in adaptation, learning and optimization*, 2:27–38, 01 2015.
[32] B. Huang, L. Yao, and K. Raguraman. Bi-level GA and GIS for multi-objective TSP route planning. *Transportation Planning and Technology*, 29:105–124, 2006.
[33] A. Karvonen, F. Cugurullo, and F. Caprotti. *Inside Smart Cities: Place, Politics and Urban Innovation*. Routledge, 2018.
[34] G. Kim, Y. Ong, C. Heng, P. Tan, and A. Zhang. City vehicle routing problem (city VRP): A review. *IEEE Transactions on Intelligent Transportation Systems*, 16:1–13, 08 2015.
[35] R. Kramer, A. Subramanian, T. Vidal, and L. Cabral. A matheuristic approach for the pollution-routing problem. *European Journal of Operational Research*, 243, 04 2014.
[36] R. Kucharski and A. Drabicki. Estimating macroscopic volume delay functions with the traffic density derived from measured speeds and flows. *Journal of Advanced Transportation*, 2017, 2017.
[37] R.S. Kumar, K. Kondapaneni, V. Dixit, A. Goswami, L.S. Thakur, and M.K. Tiwari. Multi-objective modeling of production and pollution routing problem with time window: A self-learning particle swarm optimization approach. *Computers & Industrial Engineering*, 99:29 – 40, 2016.
[38] A. Letchford, S. Nasiri, and A. Oukil. Pricing routines for vehicle routing with time windows on road networks. *Computers & Operations Research*, 51:331–337, 11 2014.

decisions based on private objectives. The acquired data would then be used to feed the integrated optimization model proposed in this paper.

References

[1] R. A. Acheampong and F. Cugurullo. Capturing the behavioural determinants behind the adoption of autonomous vehicles: Conceptual frameworks and measurement models to predict public transport, sharing and ownership trends of self-driving cars. *Transportation Research Part F: Traffic Psychology and Behaviour*, 62:349–375, 2019.

[2] A. Arroub, B. Zahi, E. Sabir, and M. Sadik. A literature review on smart cities: Paradigms, opportunities and open problems. In *2016 International Conference on Wireless Networks and Mobile Communications (WINCOM)*, pages 180–186, 2016.

[3] B. Beirigo, F. Schulte, and R.R. Negenborn. Integrating people and freight transportation using shared autonomous vehicles with compartments. *IFAC-PapersOnLine*, 51:392–397, 01 2018.

[4] T. Bektas and G. Laporte. The pollution-routing problem. *Transportation Research Part B: Methodological*, 45:1232–1250, 09 2011.

[5] H. Ben Ticha, N. Absi, D. Feillet, and A. Quilliot. Empirical analysis for the vrptw with a multigraph representation for the road network. *Comput. Oper. Res.*, 88:103–116, dec 2017.

[6] H. Ben Ticha, N. Absi, D. Feillet, and A. Quilliot. Vehicle routing problems with road-network information: State of the art. *Networks*, 72:393–406, 2018.

[7] H. Ben Ticha, N. Absi, D. Feillet, and A. Quilliot. The steiner bi-objective shortest path problem. *EURO Journal on Computational Optimization*, 9:100004, 2021.

[8] H. Ben Ticha, N. Absi, D. Feillet, A. Quilliot, and T. Woensel. The Time-Dependent Vehicle Routing Problem with Time Windows and Road-Network Information. *SN Operations Research Forum*, 2:1–25, March 2021.

[9] G. Berbeglia, J.F. Cordeau, and G. Laporte. Dynamic pickup and delivery problems. *European Journal of Operational Research*, 202:8 – 15, 2010.

[10] B. Boyacı, T. Dang, and A. Letchford. Vehicle routing on road networks: How good is euclidean approximation? *Computers & Operations Research*, 129, 05 2021.

[11] Z. Bronstein. Industry and the smart city. *Dissent*, 56:27–34, 06 2009.

[12] Bureau of Public Roads. *Traffic Assignment Manual for Application with a Large, High Speed Computer*. U. S. Department of Commerce, Bureau of Public Roads, Office of Planning, Urban Planning Division, 1964.

[13] D. Cattaruzza, N. Absi, D. Feillet, and J. Gonzalez-Feliu. Vehicle routing problems for city logistics. *EURO Journal on Transportation and Logistics*, 6:51–79, 04 2017.

[14] W. Chen, M. Mes, and M. Schutten. Multi-hop driver-parcel matching problem with time windows. *Flexible services and manufacturing journal*, 30:517–553, 9 2018. Springer deal.

[15] A. Corberán and G. Laporte. *Arc Routing: Problems, Methods, and Applications*. Society for Industrial and Applied Mathematics, 2015.

[16] G. Cornuéjols, J. Fonlupt, and D. Naddef. The traveling salesman problem on a graph and some related integer polyhedra. *Math. Program.*, 33:1–27, sep 1985.

[17] F. Cugurullo. Urban artificial intelligence: From automation to autonomy in the smart city. *Frontiers in Sustainable Cities*, 2, 2020.

[18] E. Demir, T. Bektaş, and G. Laporte. The bi-objective pollution-routing problem. *European Journal of Operational Research*, 232:464 – 478, 2014.

[19] E. Demir, T. Bektaş, and G. Laporte. A review of recent research on green road freight transportation. *Eur. J. Oper. Res.*, 237:775–793, 2014.

of the speed variables for each edge in the network, we cannot directly resort to commonly used approaches for non-complete graphs (which build a complete graph by combining the edges in the shortest path between two customer nodes into a single edge). Therefore, we propose an approach that computes an approximate solution to the assignment-routing (upper level) problem on the road network by first solving such a problem on an auxiliary complete graph. This mechanism allows one to use existing VRP methods for complete graphs, without the need to formulate the assignment-routing problem on the non-complete graph, where some of the classical VRP constraints are most likely infeasible. Then, considering the approximate assignment-routing solution as a parameter, we develop a formulation for the speed optimization (lower level) problem on a non-complete graph. It is important to remark that VRP with green-oriented objectives and speed optimization is well-known in the literature [4]. However, considering such a problem on a road network represents, to the best of our knowledge, a novel contribution, thus requiring a new solution methodology.

1.1. Organization of this paper. This paper is organized as follows. Section 2 provides a general overview of the main optimization models adopted to solve transportation problems on complete graphs and road networks. The trilevel multi-objective formulation architecture is described in Section 3. Section 4 presents an innovative constraint to model the integration among all the different transportation components. In Section 5, the trilevel multi-objective architecture is used to formulate a (bilevel) VRP/speed problem on a non-complete graph and an optimization method is developed to solve such a problem. The computational experiments are described in Section 6. Finally, in Section 7 we draw some concluding remarks and we outline several ideas for future work.

2. Literature review

2.1. Optimization models for transportation on complete graphs. Transportation services across a city have been widely studied within the field of Operations Research (see, e.g., [34] and [13]). Optimization problems related to such services are variants of the basic VRP, where a fleet of vehicles is used to deliver goods to a set of customer nodes. Two important decisions are considered: assigning groups of customers to each vehicle and defining the corresponding route.

The Pickup and Delivery Problem (PDP) is a VRP variant where people or objects need to be transported from an origin to a destination (examples of PDPs are ride-sharing, carpool problem, dial-a-ride problem, and vehicle-sharing). Classical VRPs and PDPs can be combined in the so-called *people and freight integrating transportation* problems, which deal with the integration of passenger and freight transportation. Their objective is to increase the occupancy rate by letting the spare seats in the vehicles be used to transport goods (see, e.g., [3], [14], [39], and [29]). In this way, each vehicle can carry passengers, goods, or both of them.

The *pollution-routing problem*, introduced by [4], extends the classical VRP by taking into account not only the traveled distance between origin and destination, but also the fuel consumed and the emissions generated by the vehicles. The aim is to find a proper route and speed for each vehicle, allowing customers' requests to be met, by minimizing the overall operational and environmental cost, while respecting time windows and capacity constraints (see, e.g., [35], [53], [26], [47], and [57]). More challenging problems arise when the stochastic ([55], [48], [49], [23]), dynamic ([52], [9]), and multi-objective ([28], [30], [18], [37]) versions are taken into account. When route is considered fixed and the only variable is the vehicle speed, the problem is called *the speed optimization problem* (see [24]). The problem considered in our paper can be included in the class of pollution-routing problems because the environmental-impact objective is considered together with speed optimization.

In the literature, few works address a VRP (on a complete graph) by adopting a multi-level formulation, and all of them propose a bilevel problem (see, e.g., [31], [42], [41], and [40]). In particular, in [31] and [42], the authors use an assignment-routing formulation to solve a classical VRP. Similarly, [41] propose two nested optimization levels to deal with a VRP integrated with a facility location problem but, in contrast with [31] and [42], the objective functions involved in each level are conflicting with each other. In our paper, we take advantage of the hierarchical structure at stake by proposing an optimization method that alternates between the assignment-routing problem and the speed optimization problem until a satisfactory solution is returned. Differently from the bilevel approach of [31], which extends the work of [42] to the bi-objective case by considering efficiency and service quality in each level, we also account for the environmental impact as a third objective.

allows one to model the impact of the traffic congestion caused by routing decisions made in other components on the component considered (which is a VRP-type one in our paper). Based on the formulation architecture, we developed formulations for the assignment-routing problem and the speed optimization one for the chosen freight transportation component. The approach proposed to solve the VRP/speed problem considered computes an approximate solution to the assignment-routing problem on a non-complete graph (i.e., problem (5.10)) by first solving such a problem on an auxiliary complete graph (i.e., problem (5.22)). The resulting algorithm (i.e., Algorithm (1)) alternates between the auxiliary assignment-routing problem (5.22) and the speed optimization problem (5.35) on the non-complete graph.

The computational experiments show that the proposed approach is able to determine a set of Pareto optimal solutions among the conflicting terms of the objective functions considered. Moreover, the experimental results show the importance of considering the impact of different components due to traffic congestion. Therefore, the integrated model can be used as a decision support tool to help find the best trade-off among competing criteria.

The development of concrete formulations for PDP and PFITP components and the analysis of potential conflicts among the objective functions when such components are considered is left for future work. Moreover, further research is needed to improve the proposed concrete formulation for road networks by removing the assumption that the vehicle speed used on an edge traversed more than once is the same as the first time the edge was traversed. Finally, an additional avenue of research will be focused on the speed optimization problem, which has been so far solved by using an algorithm that provides solutions with local convergence guarantees but, at the same time, high computational cost. Although the use of an exact algorithm provides us with a good solution in terms of vehicle speed given the assignment and routing decisions, in practice the use of heuristic algorithms for speed optimization is needed to ensure a real-time response.

To manage the transportation network and grant users access to the smart decision-making system, one should develop an application for smart devices. Such an app would acquire data from the users by asking them for the type of service they need (use their own vehicles, car sharing, ride sharing, carpooling, etc.) and other relevant information (such as origin and destination). Moreover, users will need to specify whether to concede the vehicle-related decisions on assignment, routing, and speed to the app or make their own

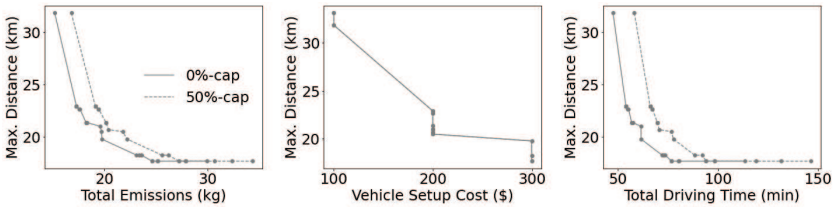

FIGURE 6. Approximate Pareto fronts for the large instance.

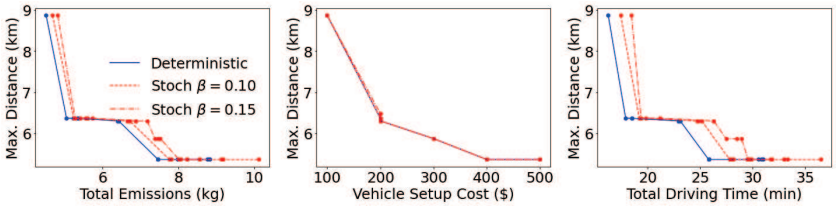

FIGURE 7. Approximate Pareto fronts for the deterministic case and the stochastic one.

case, both the CO_2 emissions and total driving time significantly increase as compared to the deterministic case due to the decrease in speed caused by traffic congestion, according to formula (4.3) and constraint (4.4). Reducing the number of vehicles from 5 to 1 leads to 53.77% ($\beta = 10\%$) and 43.87% ($\beta = 15\%$) savings in CO_2 emissions and 52.29% ($\beta = 10\%$) and 42.08% ($\beta = 15\%$) savings in total driving time. The maximum driving distance increases from 5.37 km to 8.87 km. Once a set of objective function weights is given, the average CPU time required by Algorithm 1 for $\beta = 10\%$ is 422.54 s and for $\beta = 15\%$ is 637.11 s.

7. Concluding remarks and future work

In this paper, we laid the groundwork for the development of a green-oriented integrated system aimed at managing mobility and transportation services within a smart city. In particular, we proposed a new formulation architecture consisting of three nested optimization levels associated with assignment, routing, and speed decisions. To address the integration among all the different transportation problem components, we developed an innovative constraint on the speed variables (i.e., formula (4.3) and constraint (4.4)) that

2.2. Optimization models for transportation on road networks. When considering road networks and, consequently, non-complete graphs (i.e., graphs that do not contain an edge for every pair of nodes), routing problems face additional challenges because key assumptions are not satisfied (see, e.g., [25], [16], [6], [8]). In particular, only a subset of nodes are customers since most of the nodes are associated with cross-roads, which do not have a demand to meet. Therefore, only some nodes need to be visited by the vehicles involved in the problem, which is in contrast with the traditional VRP formulations considering each node in the graph as a customer. Moreover, it may not be possible to find a route that visits nodes and edges only once, thus leading to problems with an empty feasible set. Recently, exact approaches to solve routing problems on a subset of nodes have been proposed in [54] and [10] for VRP, [56] for the traveling salesman problem (TSP), and [7] for the shortest path problem. Such papers belong to the stream of works focusing on the so-called Steiner TSP, which was addressed for the first time by [25] and [16].

On road networks, arcs may be associated with several attributes (distance, cost, time, etc.), and this implies that the shortest path does not coincide with the cheapest or quickest path. Therefore, an additional approach consists in representing the related non-complete graph by using a multi-graph [27], which allows for multiple arcs between each pair of nodes to account for the best paths associated with each attribute. The computational efficiency of multi-graphs has been questioned in [38] and reaffirmed in [5]. Instead of resorting to a multi-graph, [65] and [32] define their VRPs on a road-network, and then propose heuristic procedures that aggregate the attributes associated with each arc. They transform the original graph into a complete one by shortest paths, which is also the approach adopted in our work.

Finally, it is important to mention the class of problems known as *arc routing problems* [15], which are based on road networks and, unlike VRPs, the demand is located along the edges of the network. As pointed out in [6], the main difference between arc routing problems and VRPs on road networks is that in the former the demand is associated with a large subset of arcs, while in the latter the demand is only on a small subset of nodes.

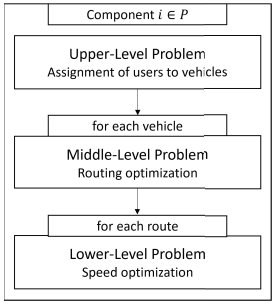

FIGURE 1. Diagram illustrating the trilevel optimization problem for each transportation component.

3. A new trilevel multi-objective formulation architecture for vehicle routing problems with speed optimization

The management of transportation services in smart cities is naturally formulated as an integrated trilevel multi-objective VRP/speed problem. Figure 1 presents the main features of the resulting trilevel formulation, which is particularized in Problem (3.2) below. The three hierarchical levels involved in such a problem are repeated for all the P possible transportation services. In particular, for each component $i \in$ P, given a set of vehicles, the Upper-Level (UL) problem determines the optimal assignment of users to such vehicles, represented by a vector a^i of binary variables. Solving the UL problem requires computing the optimal route associated with each feasible assignment a^i. In fact, given the number of available vehicles and the users assigned to them (this means that a^i is a parameter), the Middle-Level (ML) problem finds an optimal route for each vehicle, represented by a vector r^i of binary variables. Solving the ML problem requires, in turn, computing the speed (represented by a vector v^i of real variables) of the given vehicle on each segment of the given route. To this aim, for each possible route (r^i is now handled as a parameter as well), the vehicle speed v^i is determined by solving the corresponding Lower-Level (LL) problem.

Problem (3.2) depends on several data parameters, among which we highlight r^{-i} and w. In particular, since the routing decisions r^i made in each component affect the overall traffic congestion along the streets (and v^i as a

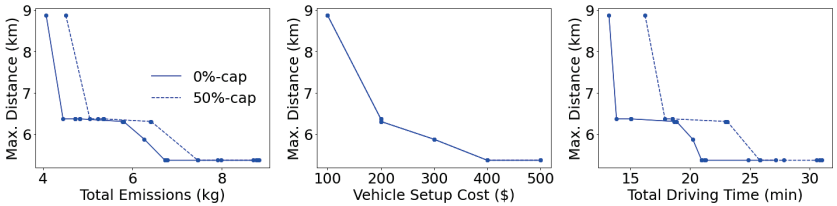

FIGURE 5. Approximate Pareto fronts for the small instance.

From the plots related to the small instance, one can observe that reducing the number of vehicles from 5 to 1 (which corresponds to a decrease in the vehicle setup cost from \$500 to \$100) leads to an increase of the maximum driving distance from 5.37 km to 8.87 km, which allows achieving 53.27% (in the 0%-cap case) and 48.98% (in the 50%-cap case) savings in CO_2 emissions and 51.29% (0%-cap) and 47.50% (50%-cap) savings in total driving time. Moreover, the first and third plots show that without increasing the maximum driving distance, it is possible to reduce the total emissions from 8.71 kg to 6.73 kg (0%-cap) and from 8.84 kg to 7.47 kg (50%-cap) and the total driving time from 27.16 min to 20.97 min (0%-cap) and from 31.01 min to 25.82 min (50%-cap). Note that using 5 vehicles does not lead to any significant improvement in terms of maximum driving distance compared to the Pareto solution with 4 vehicles.

On the large instance, only solutions with a number of vehicles between \$100 and \$300 are Pareto optimal. Reducing such a number from 3 to 1 (which corresponds to a decrease in the vehicle setup cost from \$300 to \$100) leads to an increase of the maximum driving distance from 17.68 km to 31.85 km, and the savings that can be obtained are 49.33% (0%-cap) and 51.17% (50%-cap) for CO_2 emissions and 58.17% (0%-cap) and 60.27 (50%-cap) for total driving time.

Stochastic case. In this subsection, we limit the analysis to the small instance and we consider the 50%-cap case. The random variable $\omega_{n_1 n_2}$ in (4.3) has a Poisson distribution with parameter $\lambda_{n_1 n_2}$ for all $(n_1, n_2) \in N_{\text{sub}}$. Different values of the parameter have been considered based on the capacity of each edge, i.e., $\lambda_{n_1 n_2} = \beta\, k^{\max}_{n_1 n_2}$ with $\beta \in \{0.10, 0.15\}$.

Figure 7 shows the comparison of the approximate Pareto fronts obtained in the deterministic and stochastic cases. In particular, in the stochastic

Speed (NC):
$$L = \alpha_1 \left(\frac{\text{TGE}}{w_1}\right) + \alpha_3 \left(\frac{\text{TDT}}{w_3}\right) + \alpha_5 \left(\frac{\text{LAC}}{w_5}\right) + \alpha_6 \left(\frac{\text{MAT}}{w_6}\right). \tag{6.3}$$

Constraint function ℓ is the same as problem (5.35).

Among the solutions returned by the algorithm, we eliminated the dominated points to obtain an approximation of the Pareto front. We point out that although all the terms are considered in the experiments, in the plots we only report the comparison of the approximate Pareto fronts among the conflicting terms in the environmental impact and efficiency objective functions. The terms in the service quality objective function are omitted from the figures since they have been observed not to be conflicting with the other ones. In particular, the maximum driving distance (MDD), which is referred to as *Max. Distance* in the figures, has been compared against the total generated CO_2 emissions (TGE) (referred to as *Total Emissions*), the total setup cost (TSC), and total driving time (TDT).

All tests were run on a Linux server with 32GB of RAM and an AMD Opteron 6128 processor running at 2.00 GHz. The assignment-routing problem (Step 2.1 in Algorithm 1) has been implemented in Python 3.7 and solved by using OR-Tools 9.1 [51] with default options. To solve the speed optimization problem (Step 2.3), the interior point method implemented in the solver IPOPT has been used [61] with the parameter *tol* set to 10^{-2}.

Deterministic case. In the deterministic case, the parameter $\omega_{n_1 n_2}$ in (4.3) has been set to 0 for all $(n_1, n_2) \in N_{\text{sub}}$. Figures 5–6 show the comparison between the approximate Pareto fronts obtained for 0%-cap and 50%-cap on the small and large instances. Such figures confirm that the integration among the transportation components has a significant impact on the objectives considered, as illustrated in the plots. Once a set of objective function weights is given, the average CPU time required by Algorithm 1 on the small instance is 408.58 s and on the large instance is 2995.91 s. The majority of the time is due to the computation of an approximate solution for the speed optimization problem. The development of a heuristic for speed optimization to reduce the computational cost is left for future work.

consequence), we introduce the following parameters

$$r^{-i} := (r^j \mid j \in \{1, \dots, \text{P}\} \text{ and } j \neq i) \tag{3.1}$$

to address the interplay among the different P components, i.e., how routing decisions made in one problem affect the settings in others. The uncertain parameter w accounts for uncertain factors like the weather or unforeseen events. Furthermore, w can account for the randomness in traffic congestion caused by users that decide not to share their GPS data for privacy concerns (therefore, they are not included in any component).

$$\begin{aligned} \min_{\{a^i, r^i, v^i \mid i \in \{1,\dots,\text{P}\}\}} \quad & F^i(a^i, r^i, v^i; r^{-i}, w) \\ \text{s.t.} \quad & f^i(a^i, r^i, v^i; r^{-i}, w) \leq 0 \\ & r^i \in \operatorname*{argmin}_{r^i, v^i} \; G^i(a^i, r^i, v^i; r^{-i}, w) \\ & \qquad \text{s.t.} \; g^i(a^i, r^i, v^i; r^{-i}, w) \leq 0 \\ & \qquad\quad v^i \in \operatorname*{argmin}_{v^i} \; H^i(a^i, r^i, v^i; r^{-i}, w) \\ & \qquad\qquad \text{s.t.} \; h^i(a^i, r^i, v^i; r^{-i}, w) \leq 0. \end{aligned} \tag{3.2}$$

Since the task of the integrated model is to accommodate competing goals to achieve societal benefits, all of the three problems have a multi-objective nature. Accordingly, each objective function is a vector-valued function having three scalar-valued functions as components, each related to one of the three categories of objectives: (i) generating a low impact on the environment; (ii) increasing efficiency in providing the transportation services; (iii) meeting requests of users by providing high quality. Functions f^i, g^i, and h^i represent the constraints.

The advantage of considering first such a complex trilevel formulation is to comprehensively understand the overall problem complexity and guide the development of subproblems or relaxations and of efficient solution methods. The three levels can be considered separately in a hierarchical order, choosing the most suitable optimization method for each one. Alternatively, the UL and ML problems can be combined into a single one, leading to the following bilevel problem

$$\begin{aligned} \min_{\{a^i, r^i, v^i \mid i \in \{1,\dots,\text{P}\}\}} \quad & U^i(a^i, r^i, v^i; r^{-i}, w) \\ \text{s.t.} \quad & u^i(a^i, r^i, v^i; r^{-i}, w) \leq 0 \end{aligned} \tag{3.3}$$

$$v^i \in \underset{v^i}{\text{argmin}} \; L^i(a^i, r^i, v^i; r^{-i}, w) \quad \text{s.t.} \; \ell^i(a^i, r^i, v^i; r^{-i}, w) \leq 0, \tag{3.4}$$

where the (upper level) problem (3.3) is an assignment-routing problem and the (lower level) problem (3.4) is a speed optimization problem.

4. Integration among different transportation components on a road network

We now describe the approach proposed to model the integration among the different transportation components. In the urban transportation literature, *volume-delay functions* (VDFs) are commonly used to describe the fundamental relationships between average speed (km/hour) or travel time (hour) on the one hand and traffic flow (vehicles/hour) or density (vehicles/km) on the other hand [36, 50]. One of the most widely adopted VDFs is the BPR function proposed by the [12]. To the best of our knowledge, the work by [45] is the only one using the BPR function in a VRP problem (without speed optimization). While [45] has used such a formula to compute the travel time as a function of the traffic flow, [36] have derived an expression to compute the approximate average speed v on a street as a function of the traffic density. In particular, such an expression is given by

$$v \; = \; \frac{v_0}{1 + \gamma \; (k/k^{\max})^{\eta}}, \tag{4.1}$$

where v_0 is the free-flow speed (when there is no traffic congestion), k is the traffic density (number of vehicles per unit distance), $k^{\max}$ is the traffic density when the street is at full capacity, and γ and η are positive parameters. To keep notation simple, we will use the same k and $k^{\max}$ to denote numbers of vehicles instead of densities.

We now want to propose an approximate reformulation of (4.1) to relate the average speed on each edge of a non-complete graph to the routing decisions made in other problem components. The non-complete graph associated with a road network is denoted by $\mathrm{G} = (\mathrm{N}, \mathrm{E})$, where N is the set of nodes representing all the locations relevant for making decisions (which includes not only customer nodes, but also road intersections and connections), while E is the set of directed edges connecting pairs of nodes in N. For the sake of simplicity, we will omit the argument i associated with the component under consideration, and we will use the argument j to denote

time). Recalling the definition of $k_{n_1 n_2}$ in (4.2), one has

$$k_{n_1 n_2} = \begin{cases} 0 & \text{(Case 0\%-cap)}, \\ 0.5 \, k^{\max}_{n_1 n_2} & \text{(Case 50\%-cap)}, \end{cases}$$

for all $(n_1, n_2) \in \mathrm{E}$.

We recall that in multi-objective optimization one is interested in obtaining a set of points that cannot improve one objective without worsening the values of the other ones. Points with this property are called Pareto optimal solutions (or non-dominated points). To obtain the approximate Pareto fronts, the weighted-sum method with normalization has been applied to all the terms in the objective functions U_1, U_2, and U_3 of the original assignment-routing problem (5.10), all the terms in U_1^{C}, U_2^{C}, and U_3^{C} for the auxiliary assignment-routing problem (5.22), and all the terms in L_1, L_2, and L_3 for the speed optimization problem (5.35). In particular, given non-negative weights α_i, with $i \in \{1, \ldots, 6\}$, the problems considered in the experiments are (6.1)–(6.3) below. Each term in the objective functions has been normalized using additional weights w_i or $\bar{w}_i$, with $i \in \{1, \ldots, 6\}$, which represent the value of each term at the initial solution, i.e., $w_1 = \mathrm{TGE}(a_{0,0}, r_{0,0}, v_{0,0})$, $w_2 = \mathrm{TSC}(a_{0,0}, r_{0,0}, v_{0,0})$, and similarly for the remaining terms and for the weights $\bar{w}_i$.

Assignment-Routing (NC):

$$U = \alpha_1 \left(\frac{\mathrm{TGE}}{w_1}\right) + \alpha_2 \left(\frac{\mathrm{TSC}}{w_2}\right) + \alpha_3 \left(\frac{\mathrm{TDT}}{w_3}\right) + \alpha_4 \left(\frac{\mathrm{MDD}}{w_4}\right) + \alpha_5 \left(\frac{\mathrm{LAC}}{w_5}\right) + \alpha_6 \left(\frac{\mathrm{MAT}}{w_6}\right). \tag{6.1}$$

Constraint function u is the same as problem (5.10).

Assignment-Routing (C):

$$U^{\mathrm{C}} = \alpha_1 \left(\frac{\overline{\mathrm{TGE}}}{\bar{w}_1}\right) + \alpha_2 \left(\frac{\overline{\mathrm{TSC}}}{\bar{w}_2}\right) + \alpha_3 \left(\frac{\overline{\mathrm{TDT}}}{\bar{w}_3}\right) + \alpha_4 \left(\frac{\overline{\mathrm{MDD}}}{\bar{w}_4}\right) + \alpha_5 \left(\frac{\overline{\mathrm{LAC}}}{\bar{w}_5}\right) + \alpha_6 \left(\frac{\overline{\mathrm{MAT}}}{\bar{w}_6}\right). \tag{6.2}$$

Constraint function u^{C} is the same as problem (5.22).

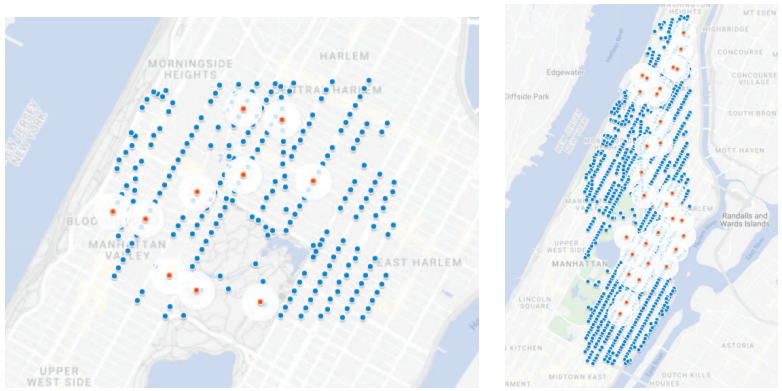

FIGURE 4. Small (left) and large (right) instances. Customers are represented by the big circles.

where $v_{q\,n_1\,n_2}$ is in km/h.

The integration among the different transportation components is modeled through formula (4.3) and constraint (4.4), which is analogous to constraint (5.17) used in the original assignment-routing problem (5.10). We recall that to ensure that the approximate solution obtained by applying (5.33) is feasible with respect to constraint (5.17), we adopt the penalty term (5.34), which requires solving the auxiliary assignment-routing problem (5.22) with such a penalty term in the objective function every time such a constraint is violated on the original non-complete graph. According to [36], $\gamma = 1$ and $\eta = 2$ are reasonable values for the parameters used in formula (4.3) when considering urban road networks. Denoting by i the current component, we handle the variables $r^j_{q\,n_1\,n_2}$ in (4.3) and (4.4) as parameters, with $j \in \{1, \ldots, \mathrm{P}\}$ and $j \neq i$. In particular, two cases are taken into account to represent the total number of vehicles controlled by other components that traverse an edge (n_1, n_2). In the first case, which we denote as 0%-cap, no vehicles controlled by other components traverse the edge (n_1, n_2), while in the second one, denoted as 50%-cap, the number of such vehicles is equal to 50% of the capacity of the edges (the word *capacity* is used to refer to $k^{\max}_{n_1\,n_2}$, which is the maximum number of vehicles that can traverse the edge at the same

the parameters that represent the routing decisions made in the other problem components, according to the notation used in (3.1). Let $r_{q\,n_1\,n_2}$ be a binary variable equal to 1 if vehicle $q \in \mathrm{V}$ traverses $(n_1, n_2) \in \mathrm{E}$. For each component $j \in \{1, \ldots, \mathrm{P}\}$, with $j \neq i$, we denote the corresponding routing parameters as $r^j_{q\,n_1\,n_2}$. Let $k_{n_1\,n_2}$ be the number of vehicles traversing edge $(n_1\,n_2) \in \mathrm{E}$ that are not controlled by component i, i.e.,

$$k_{n_1\,n_2} \;=\; \sum_{j=1,\,j\neq i}^{P} \sum_{q\in \mathrm{V}^j} r^j_{q\,n_1\,n_2} + \omega_{n_1\,n_2}. \tag{4.2}$$

In (4.2), the first term is given by the sum of the routing parameters associated with the other problem components. The second term, represented by $\omega_{n_1\,n_2}$, is a non-negative integer random parameter representing the number of vehicles traversing the edge (n_1, n_2) that are associated with users not sharing their GPS data for privacy concerns (therefore, they are not included in any component). Moreover, we define $k^{\max}_{n_1\,n_2}$ as the maximum number of vehicles that can traverse edge (n_1, n_2) at the same time, which can be estimated dividing $d_{n_1\,n_2}$ by the average length of the vehicles in the smart city streets.

Denoting the varying speed bounds on edge (n_1, n_2) as $v^{\min}_{n_1\,n_2}$ and $v^{\max}_{n_1\,n_2}$, introduced for the sake of clarity, the speed upper bound is affected by the traffic congestion according to the relation

$$v^{\max}_{n_1\,n_2} \;=\; \max\left\{ v^{\mathrm{maxlim}}_{n_1\,n_2} \left(1 + \gamma \left(\frac{k_{n_1\,n_2}}{k^{\max}_{n_1\,n_2}}\right)^{\eta}\right)^{-1},\; v^{\mathrm{minlim}}_{n_1\,n_2} \right\}, \tag{4.3}$$

which is inspired by (4.1). The speed lower bound $v^{\min}_{n_1\,n_2}$ is set equal to $v^{\mathrm{minlim}}_{n_1\,n_2}$ for all $(n_1, n_2) \in \mathrm{E}$. To ensure $k_{n_1\,n_2} \leq k^{\max}_{n_1\,n_2}$, one should also include the following constraint

$$\sum_{q\in \mathrm{V}} r_{q\,n_1\,n_2} + k_{n_1\,n_2} \;\leq\; k^{\max}_{n_1\,n_2}, \quad \forall (n_1, n_2) \in \mathrm{E}. \tag{4.4}$$

In (4.3), the maximum speed limit $v^{\mathrm{maxlim}}_{n_1\,n_2}$ plays the same role as the free-flow speed in (4.1). In particular, such a speed limit is decreased based on the ratio between $k_{n_1\,n_2}$ and $k^{\max}_{n_1\,n_2}$. Therefore, the more vehicles traverse an edge, the slower the speed on that edge is. The max function is used to ensure that the speed upper bound $v^{\max}_{n_1\,n_2}$ is greater than the minimum speed limit $v^{\mathrm{minlim}}_{n_1\,n_2}$. Since in practice vehicles may traverse a given edge at different times, the

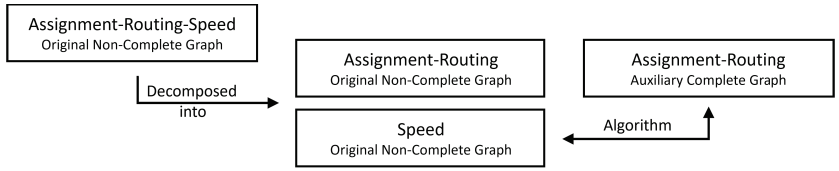

FIGURE 2. Diagram of the approach proposed to solve a VRP/speed problem on a road network (non-complete graph).

maximum speed resulting from formula (4.3) gives a pessimistic upper bound on the actual speed of vehicles.

5. Formulating and solving an integrated VRP/speed model for freight transportation on a road network

The transportation services (also called problem components) arising in a smart city can be categorized into VRPs, PDPs, or a combination of the two (*People and Freight Integrating Transportation Problems* (PFITP)). Although PDPs can be considered variants of the classical VRP, we denote as VRPs all the problem components where a set of customers wait at fixed locations for the delivery of orders, while with PDPs we refer to the components characterized by people or objects that need to be picked up at their origins and dropped off at their destinations. In our paper, we provide an integrated formulation for a VRP/speed problem concerning a specific VRP-type component, i.e., freight transportation, which is sufficient to illustrate all our contributions.

After introducing the notation used to formulate an integrated (bilevel) VRP/speed model for freight transportation on a road network (Subsection 5.1), we describe the environmental, efficiency, and service objective functions (Subsection 5.2). Then, we present the assignment-routing problem (Subsection 5.3) and the speed optimization problem (Subsection 5.4) that are obtained by decomposing a VRP/speed problem on a non-complete graph according to upper and lower level problems in (3.3)–(3.4). Finally, to solve the VRP/speed problem considered, we propose an optimization algorithm that alternates between the assignment-routing problem on an auxiliary complete graph and the speed optimization problem on the original non-complete graph (Subsection 5.5). The resulting approach is illustrated in Figure 2.

Algorithm 1

Input: Initial number of vehicles $\mathrm{VEH}_0 = 1$, initial speed $v_{0,0}$, N_{sub}, $k = 0$, $k_{\max}$, $j_{\max}$,
$\hat{\hat{U}} := +\infty$.
While $k \le k_{\max}$ **and** $\mathrm{VEH}_k \le |N_{\mathrm{sub}}|$ **do**
Step 1. $\hat{U} := +\infty$, $j = 1$.
While $j \le j_{\max}$ **and** $(a_{k,j},\, r_{k,j},\, v_{k,j}) \neq (a_{k,j-1},\, r_{k,j-1},\, v_{k,j-1})$ **do**
Step 2.1. Obtain $(\bar{a}^{\mathrm{C}}_{k,j}, \bar{r}^{\mathrm{C}}_{k,j})$ by solving the assignment-routing problem (5.22) on the auxiliary complete graph with fixed VEH_k and $v_{k,j}$.
Step 2.2. Obtain an approximate solution $(\tilde{a}_{k,j}, \tilde{r}_{k,j})$ to the assignment-routing problem (5.10) by applying (5.33).
Step 2.3. Obtain $\tilde{v}_{k,j}$ by solving the speed optimization problem (5.35) with fixed $\tilde{a}_{k,j}$ and $\tilde{r}_{k,j}$. Set $(a_{k,j},\, r_{k,j},\, v_{k,j}) = (\tilde{a}_{k,j}, \tilde{r}_{k,j}, \tilde{v}_{k,j})$.
Step 2.4. If $U(a_{k,j},\, r_{k,j},\, v_{k,j}) < \hat{U}$.
$\hat{U} := U(a_{k,j},\, r_{k,j},\, v_{k,j})$.
$(\hat{a}_k,\, \hat{r}_k,\, \hat{v}_k) := (a_{k,j},\, r_{k,j},\, v_{k,j})$.
Step 2.5. $j = j + 1$.
End do
Step 3. If $\hat{U} < \hat{\hat{U}}$.
$\hat{\hat{U}} := \hat{U}$.
$(\hat{\hat{a}}_k,\, \hat{\hat{r}}_k,\, \hat{\hat{v}}_k) := (\hat{a}_k,\, \hat{r}_k,\, \hat{v}_k)$.
Step 4. $\mathrm{VEH}_{k+1} = \mathrm{VEH}_k + 1$ and $k = k + 1$.
End do

The demand parameter DEM_n has been set to 0 for all $n \in \mathrm{N}_{\mathrm{sub}}$ to omit the capacity constraints from these experiments. Both the vehicle setup cost SC_q and the cost for late arrivals p_n have been arbitrarily set to \$100 for all $q \in \mathrm{V}$ and $n \in \mathrm{N}_{sub}$. The capacity $\mathrm{cap}_{n_1 n_2}$ of each edge (n_1, n_2) has been estimated dividing $d_{n_1 n_2}$ by the average length of vehicles in a city, set to 4.5 m (which is a reasonable assumption, according to [43]). Finally, to model the speed-dependent function f_q computing the CO_2 emissions per unit of distance (gram/km) generated by a vehicle $q \in \mathrm{V}$ traversing an edge (n_1, n_2), we have used the following formula, reported in [19]:

$$f_q(v_{q\,n_1 n_2}) = 871 - 16 v_{q\,n_1 n_2} + 0.143 v^2_{q\,n_1 n_2} + 32031/v^2_{q\,n_1 n_2},$$

speed variables. At each outer iteration, an assignment-routing on the non-complete graph is obtained transforming the original network into an auxiliary complete graph and determining the optimal assignment-routing on such a graph by using standard VRP methods. The resulting assignment-routing solution will be passed to the speed optimization problem as a parameter in order to obtain a speed value for each edge traversed by the vehicle on the road network. More specifically, given the current number of available vehicles VEH_k and the current speed vector $v_{k,j}$, Step 2.1 determines $\bar{a}^{\text{C}}_{k,j}$ and $\bar{r}^{\text{C}}_{k,j}$ by solving the assignment-routing problem on the auxiliary complete graph (5.22). Step 2.2 determines an approximate solution $(\tilde{a}_{k,\,j}, \tilde{r}_{k,\,j})$ to the assignment-routing problem (5.10) by applying (5.33). Step 2.3 determines $\tilde{v}_{k,j}$ by solving the speed optimization problem (5.35) given the assignment and routing determined at the previous step. The outer loop proceeds until either the maximum number of inner iterations is reached or the current solution is equal to the previous one, meaning that no further improvement is possible. The final solution resulting from the application of the algorithm is denoted as $(\hat{\hat{a}}_k,\, \hat{\hat{r}}_k,\, \hat{\hat{v}}_k)$.

6. Computational experiments

Two instances have been randomly generated from a dataset gathering hourly traffic information in the New York City streets from 2010 to 2013 (with more than 95,500 nodes and 260,850 arcs), which has been created by [21]. In particular, we consider a first instance (referred to as *small* in the remainder of this section) consisting of a graph with 269 nodes, 656 edges, and 10 customer nodes, and a second instance (indicated as *large*) based on a graph with 1130 nodes, 2703 edges, and 30 customer nodes. The small and large instances have been obtained by including all the nodes with longitude in $[-73.970, -73.941]$ and latitude in $[40.791, 40.810]$ and all the nodes with longitude in $[-73.970, -73.938]$ and latitude in $[40.760, 40.840]$, respectively.

In all the experiments, the customer nodes in N_{sub} have been randomly generated from the graphs. The time window upper bound b_n for each node $n \in N_{\text{sub}}$ has been set to $2/3 * |N_{\text{sub}}| * \text{TS}(\text{dep}, n)$, where $\text{TS}(\text{dep}, n)$ denotes the time required to reach n from the depot by traveling along the shortest path at a speed of 10 miles/h (about 16.09 km/h), which has resulted in tight windows. Since we are considering an urban setting, the maximum speed limit $v^{\text{maxlim}}_{n_1 n_2}$ has been set to 25 miles/h (about 40.23 km/h), while the minimum speed limit $v^{\text{minlim}}_{n_1 n_2}$ has been set to 5 miles/h (about 8.05 km/h).

5.1. Notation for an integrated VRP/speed model for freight transportation. Tables 1–2 introduce the sets, parameters, functions, and optimization variables required to formulate an integrated VRP/speed model for freight transportation. Throughout this section, the following notation is used to denote the vectors of assignment, routing, and speed variables:

$$\begin{aligned} a &:= (a_{q\,n} \mid q \in \text{V},\, n \in \text{N}_{\text{sub}}), \\ r &:= \left((r_{q\,n_1\,n_2} \mid q \in \text{V},\, (n_1, n_2) \in \text{E}),\, (r^{\text{rep}}_{q\,n_1\,n_2} \mid q \in \text{V},\, (n_1, n_2) \in \text{E})\right), \\ v &:= \left((v_{q\,n_1\,n_2} \mid q \in \text{V},\, (n_1, n_2) \in \text{E}),\, (\text{arr}^{i}_{q\,n} \mid q \in \text{V},\, n \in \text{N})\right), \end{aligned} \tag{5.1}$$

where $(\mathbf{u}, \mathbf{v})$ is adopted to denote the concatenation of two vectors $\mathbf{u}$ and $\mathbf{v}$. We will use the superscript C when the vectors defined in (5.1) refer to a problem defined on a complete graph.

We point out that each customer node in the set $\text{N}_{\text{sub}} \subseteq \text{N}$ is associated with a demand DEM_n for a product (for simplicity, we assume single commodity), with $n \in \text{N}_{\text{sub}}$. Only the nodes in N_{sub} need to be visited by the vehicles involved in the problem, as opposed to the traditional VRP formulations considering each node in the graph as a customer node (see Subsection 2.2).

5.2. Environmental, efficiency, and service objective functions. Objective function (5.2) below represents the environmental impact, which is given by the total generated emissions (TGE). Objective function (5.3) below represents the efficiency, which is given by the total setup cost (TSC), total driving time (TDT), and maximum driving distance (MDD). To compute TSC, we have included a copy of the depot in the set N (denoted as cop) and the edge (dep, cop) in the set E, and we have replaced all the edges (dep, n) in E with (cop, n) to ensure that a vehicle visits the copy after leaving the depot and before visiting other nodes. This ensures that the setup cost is only paid once for each vehicle regardless of the number of times the copy of the depot is visited. To compute MDD, we consider the total number of times each edge (n_1, n_2) in E is traversed, which is given by the sum of $r_{q\,n_1\,n_2}$ and $r^{\text{rep}}_{q\,n_1\,n_2}$. Objective function (5.4) below represents the service quality, which is given by the late arrival cost (LAC) and maximum arrival time (MAT), where the non-negative real variable $\text{arr}^{1}_{q\,n}$ represents the arrival time of vehicle $q \in \text{V}$ at node $n \in \text{N}$ when it is visited for the first time. We point out that in the computational experiments the terms in objective functions (5.3)–(5.4) are normalized to avoid numerical issues due

N	Set of nodes representing all the locations relevant for making decisions.
$\mathrm{N_{sub}}$	Subset of N composed of customer nodes ($\mathrm{N_{sub}} \subseteq \mathrm{N}$).
E	Set of directed edges connecting pairs of nodes in N.
$\mathrm{E_{sub}}$	Set of directed edges connecting each pair of customer nodes in $\mathrm{N_{sub}}$, i.e., $\{(n_1, n_2) \in \mathrm{N_{sub}} \times \mathrm{N_{sub}} \mid n_1 \neq n_2\}$.
V	Set of available vehicles, i.e., $\{1, \ldots, \mathrm{VEH}\}$, with VEH maximum number of vehicles.
$\mathrm{N}^q(r)$	Set of nodes visited by vehicle $q \in \mathrm{V}$ based on the routing vector r, i.e., $\{n \in \mathrm{N} \mid r_{q\,n\,\bar{n}} = 1 \text{ for some } (n, \bar{n}) \in \mathrm{E}\}$.
$\mathrm{N}^q_{\mathrm{sub}}(a)$	Set of customer nodes visited by vehicle $q \in \mathrm{V}$ based on the assignment vector a, i.e., $\{n \in \mathrm{N_{sub}} \mid a_{q\,n} = 1\}$.
$\mathrm{E}^q(r)$	Set of edges traversed by vehicle $q \in \mathrm{V}$ based on the routing vector r, i.e., $\{(n_1, n_2) \in \mathrm{E} \mid r_{q\,n_1\,n_2} = 1\}$.
$\mathrm{E}^q_{\mathrm{ord}}(r)$	Ordered set of the edges in E (possibly repeated) traversed by vehicle $q \in \mathrm{V}$ based on the routing vector r.
$\mathrm{E}^q_{\mathrm{sub}}(r^{\mathrm{C}})$	Set of edges in $\mathrm{E_{sub}}$ traversed by vehicle $q \in \mathrm{V}$ based on a routing vector r^{C} defined on a complete graph, i.e., $\{(n_1, n_2) \in \mathrm{E_{sub}} \mid r^{\mathrm{C}}_{q\,n_1\,n_2} = 1\}$.
$\mathrm{SP}(\bar{n}_1, \bar{n}_2)$	Sequence of edges in E corresponding to the shortest path on the non-complete graph between the customer nodes $\bar{n}_1 \in \mathrm{N_{sub}}$ and $\bar{n}_2 \in \mathrm{N_{sub}}$.
P	Number of transportation services (problem components).
$d_{n_1\,n_2}$	Length of the edge $(n_1, n_2) \in \mathrm{E}$.
$v^{\mathrm{maxlim}}_{n_1\,n_2}$	Maximum speed limit on the edge $(n_1, n_2) \in \mathrm{E}$.
$v^{\mathrm{minlim}}_{n_1\,n_2}$	Minimum speed limit on the edge $(n_1, n_2) \in \mathrm{E}$.
dep	Node in N corresponding to the depot.
cop	Node in N corresponding to a copy of the depot (it can be visited more than once).
$f_q(v)$	Speed-dependent function computing the emissions (per unit distance) generated by vehicle $q \in \mathrm{V}$ when traveling at speed v.
$r^j_{q\,n_1\,n_2}$	Binary parameter equal to 1 if vehicle q from component j traverses edge $(n_1, n_2) \in \mathrm{E}$, with $j \in \{1, \ldots, \mathrm{P}\}$ and $j \neq i$.
$\omega_{n_1\,n_2}$	Positive integer random parameter representing the number of vehicles traversing the edge $(n_1, n_2) \in \mathrm{E}$ that are not included in any component.
$k_{n_1\,n_2}$	Number of vehicles traversing the edge $(n_1, n_2) \in \mathrm{E}$ not controlled by component i.
$k^{\max}_{n_1\,n_2}$	Maximum number of vehicles that can traverse edge $(n_1, n_2) \in \mathrm{E}$ at the same time.
$v^{\max}_{n_1\,n_2}$	Speed upper bound on the edge $(n_1, n_2) \in \mathrm{E}$.
$v^{\min}_{n_1\,n_2}$	Speed lower bound on the edge $(n_1, n_2) \in \mathrm{E}$.
$d^{\mathrm{SP}}_{n_1\,n_2}$	Length of the shortest path $\mathrm{SP}(\bar{n}_1, \bar{n}_2)$ for all $(\bar{n}_1, \bar{n}_2) \in \mathrm{E_{sub}}$.
C_q	Capacity of vehicle $q \in \mathrm{V}$.
SC_q	Setup cost of vehicle $q \in \mathrm{V}$.
DEM_n	Demand of customer $n \in \mathrm{N_{sub}}$.
$[a_n, b_n]$	Time window for delivery at customer $n \in \mathrm{N_{sub}}$.
p_n	Penalty cost incurred when customer $n \in \mathrm{N_{sub}}$ is visited out of the time window.
$c^q_1(p, r)$	Function that returns the p-th element of $\mathrm{E}^q_{\mathrm{ord}}(r)$, with $p \in \{1, \ldots, \lvert\mathrm{E}^q_{\mathrm{ord}}(r)\rvert\}$.
$c^q_2(n, p, r)$	Function that reads $\mathrm{E}^q_{\mathrm{ord}}(r)$ to count the number of times node $n \in \mathrm{N}$ has been visited by vehicle $q \in \mathrm{V}$ until the edge associated with the p-th element is traversed.
$c^q_3(n_1, n_2, \bar{n}_1, \bar{n}_2)$	Function that reads the shortest path $\mathrm{SP}(\bar{n}_1, \bar{n}_2)$ to count the number of times edge $(n_1, n_2) \in \mathrm{E}$ is visited by vehicle $q \in \mathrm{V}$ in $\mathrm{SP}(\bar{n}_1, \bar{n}_2)$.

TABLE 1. List of sets, parameters, and functions used in the integrated VRP/speed model for freight transportation.

all $(n_1, n_2) \in \mathrm{E}^q(r)$.

$$t_{q\,n_1\,n_2} = (r_{q\,n_1\,n_2} + r^{\mathrm{rep}}_{q\,n_1\,n_2})\, d_{n_1\,n_2} / v_{q\,n_1\,n_2}, \quad \forall q \in \mathrm{V},\ (n_1, n_2) \in \mathrm{E}^q(r) \tag{5.36}$$

$$\mathrm{emis}_{q\,n_1\,n_2} = (r_{q\,n_1\,n_2} + r^{\mathrm{rep}}_{q\,n_1\,n_2})\, d_{n_1\,n_2}\, f_q(v_{q\,n_1\,n_2}), \quad \forall q \in \mathrm{V},\ (n_1, n_2) \in \mathrm{E}^q(r) \tag{5.37}$$

$$\mathrm{arr}^i_{q\,n_1} - \mathrm{arr}^j_{q\,n_2} + t_{q\,n_1\,n_2} \leq M(1 - r_{q\,n_1\,n_2}),$$
$$\text{with } (n_1, n_2) = c^q_1(p, r),\ i = c^q_2(n_1, p, r),\ j = c^q_2(n_2, p, r)$$
$$\forall q \in \mathrm{V} \text{ and } p \in \{1, \ldots, \lvert\mathrm{E}^q_{\mathrm{ord}}(r)\rvert\} \tag{5.38}$$

$$v_{q\,n_1\,n_2} \leq v^{\max}_{n_1\,n_2}, \quad \forall q \in \mathrm{V},\ (n_1, n_2) \in \mathrm{E}^q(r) \tag{5.39}$$

$$v_{q\,n_1\,n_2} \geq v^{\min}_{n_1\,n_2}, \quad \forall q \in \mathrm{V},\ (n_1, n_2) \in \mathrm{E}^q(r) \tag{5.40}$$

$$v_{q\,n_1\,n_2} \geq 0,\ t_{q\,n_1\,n_2} \in \mathbb{R},\ \mathrm{emis}_{q\,n_1\,n_2} \in \mathbb{R}, \forall q \in \mathrm{V},\ (n_1, n_2) \in \mathrm{E}, n \in \mathrm{N}^q(r)$$
$$\text{and } \mathrm{arr}^i_{q\,n_1},\ \mathrm{arr}^i_{q\,n_2} \geq 0 \text{ in (5.38) } \forall q \in \mathrm{V},\ p \in \{1, \ldots, \lvert\mathrm{E}^q_{\mathrm{ord}}(r)\rvert\} \tag{5.41}$$

Note that all the optimization variables in the speed-optimization problem (5.35) are real-valued. Given that constraints (5.36)–(5.37) are nonlinear, such a speed optimization problem is a nonlinear constrained optimization problem.

5.5. An algorithm for assignment-routing-speed optimization. To solve the VRP/speed problem on road networks, we propose an optimization method that alternates between the auxiliary assignment-routing problem (5.22) and the speed optimization one (5.35). To handle the multiple objective functions in such problems, we use the weighted-sum method [22], [44]. Therefore, given non-negative weights α_i, with $i \in \{1, 2, 3\}$, such objective functions are weighted into single-objective scalar functions, i.e., $U(a, r, v) = \sum_{i=1}^{3} \alpha_i\, U_i(a, r, v)$ and $L(a, r, v) = \sum_{i=1}^{3} \alpha_i\, L_i(a, r, v)$.

The schema of the proposed optimization method is reported in Algorithm 1, where we use the notation introduced in (5.1). To keep track of the values of the variables across the iterations in the two *while loops*, we use two subscripts, i.e., $a_{k,j}$, $r_{k,j}$, $v_{k,j}$, where k refers to the outer loop and j to the inner loop. Moreover, the number of vehicles available at iteration k is denoted by VEH_k.

Given an initial number of available vehicles VEH_0 and an initial speed vector $v_{0,0}$, the algorithm performs an exhaustive enumeration until either the maximum number of outer iterations $k_{\max}$ or the maximum number of customer nodes $|N_{sub}|$ is reached. For each number of vehicles, the algorithm runs an outer cycle until no change occurs in the assignment, routing, and

edges $(\hat{n}_1, \hat{n}_2)$ will not be included in the vehicle's route, i.e.,

$$\Phi(r^{\mathrm{C}}_{\hat{q}\,\bar{n}_1\,\bar{n}_2}) \;=\; \begin{cases} +\infty & \text{if } r^{\mathrm{C}}_{\hat{q}\,\bar{n}_1\,\bar{n}_2} = 1, \\ 0 & \text{if } r^{\mathrm{C}}_{\hat{q}\,\bar{n}_1\,\bar{n}_2} = 0, \end{cases} \tag{5.34}$$

where $(\bar{n}_1, \bar{n}_2) \in \mathrm{E_{sub}}$ is such that $(\hat{n}_1, \hat{n}_2) \in \mathrm{SP}(\bar{n}_1, \bar{n}_2)$.

5.4. The speed optimization problem. When solving the speed optimization problem, the assignment and routing variables on the non-complete graph, represented by the vectors a and r, respectively, are considered parameters. In particular, according to the notation introduced in Table 1, let $\mathrm{N}^q(r) \subseteq \mathrm{N}$ be the set of nodes visited by vehicle q, with $q \in \mathrm{V}$, based on the routing vector r, i.e.,

$$\mathrm{N}^q(r) \;=\; \{n \in \mathrm{N} \mid r_{q\,n\,\bar{n}} = 1 \text{ for some } (n, \bar{n}) \in \mathrm{E}\}.$$

Similarly, let $\mathrm{N}^q_{\mathrm{sub}}(a) \subseteq \mathrm{N}$ be the set of customer nodes visited by vehicle q, with $q \in \mathrm{V}$, based on the assignment vector a, i.e.,

$$\mathrm{N}^q_{\mathrm{sub}}(a) \;=\; \{n \in \mathrm{N_{sub}} \mid a_{q\,n} = 1\}.$$

Finally, let $\mathrm{E}^q(r) \subseteq \mathrm{E}$ be the set of edges traversed by vehicle q, with $q \in \mathrm{V}$, based on the routing vector r, i.e.,

$$\mathrm{E}^q(r) \;=\; \{(n_1, n_2) \in \mathrm{E} \mid r_{q\,n_1\,n_2} = 1\}.$$

We can now write the speed optimization problem on the non-complete graph $\mathrm{G(N, E)}$ as a particular case of problem (3.4) for our choice of freight transportation component

$$\begin{aligned} &\text{Speed (NC):} \qquad && L_1 = \mathrm{TGE},\ L_2 = \mathrm{TDT}, \text{ and } L_3 = \mathrm{LAC} + \mathrm{MAT}. \\ &&& \ell \text{ is given by (5.36)–(5.41) below.} \end{aligned} \tag{5.35}$$

The terms in the objective functions of problem (5.35) are defined in (5.2)–(5.4) (note that E can be replaced by $\mathrm{E}^q(r)$ and $\mathrm{N_{sub}}$ can be replaced by $\mathrm{N}^q_{\mathrm{sub}}(a)$). Constraints (5.36)–(5.38) below are analogous to (5.18)–(5.20). Constraints (5.39) and (5.40) below restrict the value of the vehicle speed on each edge, where $v^{\max}_{n_1\,n_2}$ is given by (4.3), while $v^{\min}_{n_1\,n_2}$ is set equal to $v^{\mathrm{minlim}}_{n_1\,n_2}$ for

$a_{q\,n}$	Binary variable equal to 1 if vehicle $q \in \mathrm{V}$ is assigned to customer $n \in \mathrm{N_{sub}}$.
$r_{q\,n_1\,n_2}$	Binary variable equal to 1 if vehicle $q \in \mathrm{V}$ traverses $(n_1, n_2) \in \mathrm{E}$.
$r^{\mathrm{rep}}_{q\,n_1\,n_2}$	Integer variable representing the number of times that vehicle $q \in \mathrm{V}$ traverses $(n_1, n_2) \in \mathrm{E}$ after the first time.
$v_{q\,n_1\,n_2}$	Speed of vehicle $q \in \mathrm{V}$ on the edge $(n_1, n_2) \in \mathrm{E}$.
$t_{q\,n_1\,n_2}$	Time spent by vehicle $q \in \mathrm{V}$ to traverse $(n_1, n_2) \in \mathrm{E}$.
$\mathrm{emis}_{q\,n_1\,n_2}$	Emissions generated by vehicle $q \in \mathrm{V}$ to traverse $(n_1, n_2) \in \mathrm{E}$.
$\mathrm{arr}^i_{q\,n}$	Non-negative real variable representing the arrival time of vehicle $q \in \mathrm{V}$ at node $n \in \mathrm{N}$ when it is visited for the i-th time, with $i = c^q_2(n, p, r)$, $\forall q \in \mathrm{V}$, $n \in \mathrm{N}^q(r)$, and $p \in \{1, \ldots, \lvert\mathrm{E}^q_{\mathrm{ord}}(r)\rvert\}$.

TABLE 2. List of optimization variables used in the integrated VRP/speed model for freight transportation.

to their different units of measurement.

$$\begin{aligned} &\text{Environmental impact:} && \mathrm{TGE} && (5.2) \\ &\text{Efficiency:} && \mathrm{TSC} + \mathrm{TDT} + \mathrm{MDD} && (5.3) \\ &\text{Service quality:} && \mathrm{LAC} + \mathrm{MAT} && (5.4) \end{aligned}$$

$$\mathrm{TGE} = \sum_{q \in \mathrm{V}} \sum_{(n_1, n_2) \in \mathrm{E}} \mathrm{emis}_{q\,n_1\,n_2} \qquad \mathrm{TSC} = \sum_{q \in \mathrm{V}} \mathrm{SC}_q\, r_{q\,\mathrm{dep\,cop}}$$

$$\mathrm{TDT} = \sum_{q \in \mathrm{V}} \sum_{(n_1, n_2) \in \mathrm{E}} t_{q\,n_1\,n_2} \qquad \mathrm{MDD} = \max_{q \in V} \left\{ \sum_{(n_1, n_2) \in \mathrm{E}} (r_{q\,n_1\,n_2} + r^{\mathrm{rep}}_{q\,n_1\,n_2})\, d_{n_1\,n_2} \right\}$$

$$\mathrm{LAC} = \sum_{q \in \mathrm{V}} \sum_{n \in \mathrm{N_{sub}}} \max\{\mathrm{arr}^1_{q\,n} - b_n, 0\}\, p_n \qquad \mathrm{MAT} = \max_{\substack{n \in N_{\mathrm{sub}} \\ q \in V}} \{\mathrm{arr}^1_{q\,n}\}$$

Note that the max functions used in the objective functions (5.3)–(5.4) can be linearized using additional non-negative variables (i.e., z, u, w), rewriting MDD, LAC, and MAT as

$$\mathrm{MDD}_{\ell in} = w, \qquad \mathrm{LAC}_{\ell in} = \sum_{q \in \mathrm{V}} \sum_{n \in \mathrm{N_{sub}}} z_{q\,n}\, p_n, \qquad \mathrm{MAT}_{\ell in} = u, \tag{5.5}$$

and considering the following constraints.

$$w \;\geq\; \sum_{(n_1, n_2) \in \mathrm{E}} (r_{q\,n_1\,n_2} + r^{\mathrm{rep}}_{q\,n_1\,n_2})\, d_{q\,n_1\,n_2}, \quad \forall q \in \mathrm{V} \tag{5.6}$$

$$z_{q\,n} \geq \mathrm{arr}^1_{q\,n} - b_n, \qquad \forall q \in \mathrm{V},\ n \in N_{\mathrm{sub}} \tag{5.7}$$

$$u \;\geq \mathrm{arr}^1_{q\,n}, \qquad \forall q \in \mathrm{V},\ n \in N_{\mathrm{sub}} \tag{5.8}$$

For the rest of the paper, in similar situations where max functions appear, one can linearize them as it was done in (5.5)–(5.8).

5.3. The assignment-routing problem. When considering road networks, which are represented by non-complete graphs, the standard VRP/speed formulations based on complete graphs can easily lead to infeasible problems since feasible assignment-routing-speed solutions may not exist. In particular, each customer node may be visited by more than one vehicle and/or more than once by the same vehicle. Moreover, a vehicle may need to both visit some nodes (including the depot) and traverse some edges more than once.

In the assignment-routing problem, we minimize objective functions (5.2)–(5.4) to determine an assignment-routing solution (a, r) on a non-complete graph $\mathrm{G(N, E)}$. The resulting vector of routing variables r is associated with a sequence of edges in E that starts and ends at the depot and possibly visits nodes and edges more than once. According to the notation introduced in Table 1, let $\mathrm{E}^q_{\mathrm{ord}}(r) \subseteq \mathrm{E}$ be an ordered set of the edges (possibly repeated) traversed by vehicle q, where edges are sorted based on the order in which they are traversed along the route given by r. We can now introduce two functions based on $\mathrm{E}^q_{\mathrm{ord}}(r)$. Function $c^q_1(p, r)$ returns the p-th element of $\mathrm{E}^q_{\mathrm{ord}}(r)$, with $p \in \{1, \ldots, |\mathrm{E}^q_{\mathrm{ord}}(r)|\}$, while function $c^q_2(n, p, r)$ reads $\mathrm{E}^q_{\mathrm{ord}}(r)$ to count the number of times node n has been visited by vehicle q until the edge associated with the p-th element is traversed. Moreover, we define the optimization variable $\mathrm{arr}^i_{q\,n}$ as the arrival time of vehicle q at node n when it is visited for the i-th time, i.e.,

$$\mathrm{arr}^i_{q\,n} \geq 0, \text{ with } i = c^q_2(n, p, r), \quad \forall q \in \mathrm{V},\ n \in \mathrm{N}^q(r),\ p \in \{1, \ldots, |\mathrm{E}^q_{\mathrm{ord}}(r)|\}. \tag{5.9}$$

By considering the speed and arrival time variables as parameters, we can write the assignment-routing problem on the non-complete graph $\mathrm{G(N, E)}$ as a particular case of problem (3.3) for our choice of freight transportation component (we use "(NC)" to denote that we are here considering a non-complete graph):

$$\begin{aligned}\text{Assignment-Routing (NC):}\quad & U_1 = \mathrm{TGE},\ U_2 = \mathrm{TSC} + \mathrm{TDT} + \mathrm{MDD}, \text{ and } U_3 = \mathrm{LAC} + \mathrm{MAT}.\\ & \text{Constraint function } u \text{ is given by (5.11)–(5.21) below.}\end{aligned} \tag{5.10}$$

shortest paths associated with each edge in the optimal routing on the auxiliary complete graph. In the example illustrated in Figure 3, assume that an optimal route for the VRP defined on the auxiliary complete graph $\mathrm{G(N_{sub}, E_{sub})}$ is 0-8-3-1-0. Then, a route for the original non-complete graph $\mathrm{G(N, E)}$ can be obtained by considering the shortest paths between 0-8, 8-3, 3-1, and 1-0, which leads to 0-6-7-8-6-0-5-3-5-2-1-2-3-5-0.

More formally, according to Table 1, let us define the shortest path on the non-complete graph between two customer nodes as a sequence of edges in E denoted by $\mathrm{SP}(\bar{n}_1, \bar{n}_2)$, with $(\bar{n}_1, \bar{n}_2) \in \mathrm{E_{sub}}$. The function $c^q_3(n_1, n_2, \bar{n}_1, \bar{n}_2)$ reads $\mathrm{SP}(\bar{n}_1, \bar{n}_2)$ to count the number of times edge $(n_1, n_2) \in \mathrm{E}$ is visited by vehicle $q \in \mathrm{V}$. Moreover, given an optimal solution $(\bar{a}^{\mathrm{C}}, \bar{r}^{\mathrm{C}})$ to the assignment-routing problem (5.22), let $\mathrm{E}^q_{\mathrm{sub}}(\bar{r}^{\mathrm{C}}) \subseteq \mathrm{E_{sub}}$ be the set of edges on the auxiliary complete graph traversed by vehicle q, with $q \in \mathrm{V}$, based on the routing vector $\bar{r}^{\mathrm{C}}$, i.e.,

$$\mathrm{E}^q_{\mathrm{sub}}(\bar{r}^{\mathrm{C}}) \ = \ \{(n_1, n_2) \in \mathrm{E_{sub}} \mid \bar{r}^{\mathrm{C}}_{q\,n_1\,n_2} = 1\}.$$

For all $q \in \mathrm{V}$, $n \in \mathrm{N_{sub}}$, and $(n_1, n_2) \in \mathrm{E}$, we can set

$$\begin{aligned}
\tilde{a}_{q\,n} \ &= \ \bar{a}^{\mathrm{C}}_{q\,n},\\
\tilde{r}_{q\,n_1\,n_2} \ &= \ \begin{cases} 1 & \text{if } (n_1, n_2) \in \mathrm{SP}(\bar{n}_1, \bar{n}_2) \text{ for some } (\bar{n}_1, \bar{n}_2) \in \mathrm{E}^q_{\mathrm{sub}}(\bar{r}^{\mathrm{C}}),\\ 0 & \text{otherwise,}\end{cases}\\
\tilde{r}^{\mathrm{rep}}_{q\,n_1\,n_2} \ &= \ \left(\sum_{(\bar{n}_1, \bar{n}_2) \in \mathrm{E}^q_{\mathrm{sub}}(\bar{r}^{\mathrm{C}})} c^q_3(n_1, n_2, \bar{n}_1, \bar{n}_2) \right) - 1.
\end{aligned} \tag{5.33}$$

Note that constraints (5.23)–(5.28) in the auxiliary assignment-routing problem (5.22) ensure that constraints (5.11)–(5.16) in the original assignment-routing problem (5.10) are satisfied at the point $(\tilde{a}, \tilde{r})$ resulting from (5.33). In the formulation of problem (5.22), we do not include a constraint enforcing (5.17) because it would not allow solving problem (5.22) by using existing VRP methods, which are not designed to handle such a constraint. Therefore, to ensure that constraint (5.17) evaluated at $(\tilde{a}, \tilde{r})$ is satisfied for all $q \in \mathrm{V}$ and $(n_1, n_2) \in \mathrm{E}$, we use a penalization approach. In particular, when such a constraint is violated for some vehicles $\hat{q} \in \mathrm{V}$ and edges $(\hat{n}_1, \hat{n}_2) \in \mathrm{E}$, we solve the auxiliary assignment-routing (5.22) again by introducing penalty terms in the objective function to ensure that the

$$\overline{\text{TDT}} = \sum_{q \in \text{V}} \sum_{(n_1,n_2) \in \text{E}_{\text{sub}}} t^{\text{C}}_{q\, n_1\, n_2} \qquad \overline{\text{MDD}} = \max_{q \in V} \left\{ \sum_{(n_1,n_2) \in \text{E}_{\text{sub}}} r^{\text{C}}_{q\, n_1\, n_2}\, d^{\text{SP}}_{n_1\, n_2} \right\}$$

$$\overline{\text{LAC}} = \sum_{q \in \text{V}} \sum_{n \in \text{N}_{\text{sub}}} \max\{\text{arr}^{\text{C}}_{q\, n} - b_n, 0\}\, p_n \qquad \overline{\text{MAT}} = \max_{\substack{n \in N_{\text{sub}} \\ q \in V}} \{\text{arr}^{\text{C}}_{q\, n}\}$$

Constraints (5.23)-(5.28) below are standard in the VRP literature and represent the user assignment (each customer is served by one vehicle), capacity constraint (demand cannot exceed vehicle capacity), route assignment (each customer is included in a route), starting flow (each vehicle departs from the depot), and flow conservation (incoming flow is equal to outgoing flow). Constraints (5.29) and (5.30) below compute the travel time and the generated emissions on each edge. In (5.31) below, M is a sufficiently large positive constant, whose value can be set equal to the sum of the largest values of $\text{arr}^{\text{C}}_{q\, n_1}$ and $t^{\text{C}}_{q\, n_1\, n_2}$.

$$\sum_{q \in \text{V}} a^{\text{C}}_{q\, n} = 1, \quad \forall n \in \text{N}_{\text{sub}} \tag{5.23}$$

$$\sum_{n \in \text{N}_{\text{sub}}} a^{\text{C}}_{q\, n} \text{DEM}_n \leq \text{C}_q, \quad \forall q \in \text{V} \tag{5.24}$$

$$\sum_{(n_1,n_2) \in \text{E}_{\text{sub}}} r^{\text{C}}_{q\, n_1\, n_2} = a^{\text{C}}_{q\, n_1}, \quad \forall q \in \text{V},\, n_1 \in \text{N}_{\text{sub}} \tag{5.25}$$

$$\sum_{(n_1,n_2) \in \text{E}_{\text{sub}}} r^{\text{C}}_{q\, n_1\, n_2} = a^{\text{C}}_{q\, n_2}, \quad \forall q \in \text{V},\, n_2 \in \text{N}_{\text{sub}} \tag{5.26}$$

$$\sum_{(dep,n) \in \text{E}_{\text{sub}}} r^{\text{C}}_{q\, dep\, n} = 1, \quad \forall q \in \text{V} \tag{5.27}$$

$$\sum_{(n_1,n_2) \in \text{E}_{\text{sub}}} r^{\text{C}}_{q\, n_1\, n_2} - \sum_{(n_2,n_1) \in \text{E}_{\text{sub}}} r^{\text{C}}_{q\, n_2\, n_1} = 0, \quad \forall q \in \text{V},\, n_2 \in \text{N}_{\text{sub}} \tag{5.28}$$

$$t^{\text{C}}_{q\, n_1\, n_2} = (r^{\text{C}}_{q\, n_1\, n_2}\, d^{\text{SP}}_{n_1\, n_2})/v^{\text{C}}_{q\, n_1\, n_2}, \quad \forall q \in \text{V},\, (n_1, n_2) \in \text{E}_{\text{sub}} \tag{5.29}$$

$$\text{emis}^{\text{C}}_{q\, n_1\, n_2} = r^{\text{C}}_{q\, n_1\, n_2}\, d^{\text{SP}}_{n_1\, n_2}\, f_q(v^{\text{C}}_{q\, n_1\, n_2}), \quad \forall q \in \text{V},\, (n_1, n_2) \in \text{E}_{\text{sub}} \tag{5.30}$$

$$\text{arr}^{\text{C}}_{q\, n_1} - \text{arr}^{\text{C}}_{q\, n_2} + t^{\text{C}}_{q\, n_1\, n_2} \leq M(1 - r^{\text{C}}_{q\, n_1\, n_2}), \quad \forall q \in \text{V},\, (n_1, n_2) \in \text{E}_{\text{sub}},\, n_2 \neq \text{dep} \tag{5.31}$$

$$a^{\text{C}}_{q\, n} \in \{0, 1\},\, r^{\text{C}}_{q\, n_1\, n_2} \in \{0, 1\},\, \text{arr}^{\text{C}}_{q\, n} \geq 0,\, t^{\text{C}}_{q\, n_1\, n_2} \in \mathbb{R},\, \text{emis}_{q\, n_1\, n_2} \in \mathbb{R},$$
$$\forall q \in \text{V},\, n \in \text{N}_{\text{sub}},\, (n_1, n_2) \in \text{E}_{\text{sub}} \tag{5.32}$$

Note that the assignment-routing problem (5.22) is solely composed of binary variables (the variables $t^{\text{C}}_{q\, n_1\, n_2}$ and $\text{emis}^{\text{C}}_{q\, n_1\, n_2}$ are only introduced for the sake of clarity in the presentation of the formulation).

After solving the auxiliary problem (5.22), an approximate solution $(\tilde{a}, \tilde{r})$ for the original assignment-routing problem (5.10) can be derived from the

Constraint (5.11) below ensures that each customer node is assigned to at least one vehicle (not necessarily just one). Constraint (5.12) below ensures that the demand assigned to a vehicle does not exceed the vehicle capacity. Constraints (5.13) and (5.14) below ensure that each customer node is visited at least once by the assigned vehicle. Constraint (5.15) below ensures that all the vehicles in V are used and their routes start at the depot. Constraint (5.16) below requires flow conservation at each node. Constraint (5.17) below is analogous to (4.4) and accounts for integration with other components. Constraints (5.18) and (5.19) below compute the travel time and generated emissions for each edge, considering the total number of times the edge is traversed. In particular, $d_{n_1 n_2}/v_{q\, n_1\, n_2}$ represents the time required to travel along the edge (n_1, n_2), while $f_q(v_{q\, n_1\, n_2})\, d_{n_1\, n_2}$ computes the emissions generated along the same edge. Constraint (5.20) below extends the classical precedence constraint to the case with nodes visited more than once.

$$\sum_{q \in \text{V}} a_{q\, n} \geq 1, \quad \forall n \in \text{N}_{\text{sub}} \tag{5.11}$$

$$\sum_{n \in \text{N}_{\text{sub}}} a_{q\, n} \text{DEM}_n \leq \text{C}_q, \quad \forall q \in \text{V} \tag{5.12}$$

$$\sum_{(n_1,n_2) \in \text{E}} (r_{q\, n_1\, n_2} + r^{\text{rep}}_{q\, n_1\, n_2}) \geq a_{q\, n_1}, \quad \forall q \in \text{V}, n_1 \in \text{N}_{\text{sub}} \tag{5.13}$$

$$\sum_{(n_1,n_2) \in \text{E}} (r_{q\, n_1\, n_2} + r^{\text{rep}}_{q\, n_1\, n_2}) \geq a_{q\, n_2}, \quad \forall q \in \text{V}, n_2 \in \text{N}_{\text{sub}} \tag{5.14}$$

$$r_{q\ \text{dep cop}} = 1, \quad \forall q \in \text{V} \tag{5.15}$$

$$\sum_{(n_1,n_2) \in \text{E}_{\text{sub}}} (r_{q\, n_1\, n_2} + r^{\text{rep}}_{q\, n_1\, n_2}) - \sum_{(n_2,n_1) \in \text{E}_{\text{sub}}} (r_{q\, n_2\, n_1} + r^{\text{rep}}_{q\, n_2\, n_1}) = 0, \quad \forall q \in \text{V},\, n_2 \in \text{N}_{\text{sub}} \tag{5.16}$$

$$\sum_{q \in \text{V}} r_{q\, n_1\, n_2} + k_{n_1\, n_2} \leq k^{\max}_{n_1\, n_2}, \quad \forall (n_1, n_2) \in \text{E} \tag{5.17}$$

$$t_{q\, n_1\, n_2} = (r_{q\, n_1\, n_2} + r^{\text{rep}}_{q\, n_1\, n_2})\, d_{n_1\, n_2}/v_{q\, n_1\, n_2}, \quad \forall q \in \text{V},\, (n_1, n_2) \in \text{E} \tag{5.18}$$

$$\text{emis}_{q\, n_1\, n_2} = (r_{q\, n_1\, n_2} + r^{\text{rep}}_{q\, n_1\, n_2})\, d_{n_1\, n_2}\, f_q(v_{q\, n_1\, n_2}), \quad \forall q \in \text{V},\, (n_1, n_2) \in \text{E} \tag{5.19}$$

$$\text{arr}^{i}_{q\, n_1} - \text{arr}^{j}_{q\, n_2} + t_{q\, n_1\, n_2} \leq M(1 - r_{q\, n_1\, n_2}),$$
$$\text{with} \quad (n_1, n_2) = c^q_1(p, r),\ i = c^q_2(n_1, p, r),\ j = c^q_2(n_2, p, r)$$
$$\forall q \in \text{V} \ \text{ and } p \in \{1, \ldots, |\text{E}^q_{\text{ord}}(r)|\} \tag{5.20}$$

$$a_{q\, \bar{n}} \in \{0, 1\},\, r_{q\, n_1\, n_2} \in \{0, 1\},\, r^{\text{rep}}_{q\, n_1\, n_2} \in \mathbb{Z}^+,\, t_{q\, n_1\, n_2} \in \mathbb{R},\, \text{emis}_{q\, n_1\, n_2} \in \mathbb{R}, \forall q \in \text{V},$$
$$\bar{n} \in \text{N}_{\text{sub}},\, (n_1, n_2) \in \text{E}, \text{ and } \text{arr}^{i}_{q\, n_1},\, \text{arr}^{j}_{q\, n_2} \geq 0 \text{ in (5.20) } \forall q \in \text{V},\, p \in \{1, \ldots, |\text{E}^q_{\text{ord}}(r)|\} \tag{5.21}$$

In our approach, to obtain an approximate assignment-routing solution for problem (5.10), we first solve the assignment-routing problem on an auxiliary

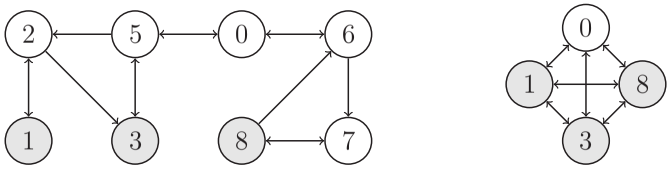

FIGURE 3. Example showing the approach used to obtain an auxiliary complete graph from a non-complete graph. The non-complete graph $\mathrm{G}(\mathrm{N}, \mathrm{E})$ reported in (a) is transformed into the complete graph $\mathrm{G}(\mathrm{N}_{\mathrm{sub}}, \mathrm{E}_{\mathrm{sub}})$ in (b) solely based on the customer nodes.

complete graph $\mathrm{G}(\mathrm{N}_{\mathrm{sub}}, \mathrm{E}_{\mathrm{sub}})$ with set of nodes given by $\mathrm{N}_{\mathrm{sub}}$ and set of edges given by $\mathrm{E}_{\mathrm{sub}} = \{(n_1, n_2) \in \mathrm{N}_{\mathrm{sub}} \times \mathrm{N}_{\mathrm{sub}} \mid n_1 \neq n_2\}$. In particular, the original road network is transformed into a complete graph by considering the lengths of the shortest paths between each pair of customer nodes as the lengths of the new edges (such lengths can be computed efficiently by using, for example, the Dijkstra algorithm proposed in [20]). The approach proposed is illustrated in Figure 3. The non-complete graph $\mathrm{G}(\mathrm{N}, \mathrm{E})$ in (a) represents a small road network with single vehicle where the depot is labeled as 0, the customer nodes are denoted as 1, 3, and 8, and all the edges are assumed to have a unitary length except $(2, 3)$, whose weight is equal to 4. To transform such a graph into the complete graph $\mathrm{G}(\mathrm{N}_{\mathrm{sub}}, \mathrm{E}_{\mathrm{sub}})$ in (b), we need to compute the shortest paths between the depot and each customer node and between each pair of customer nodes, which are reported in Table 3 along with their lengths. Note that the distances between the start and end nodes of the shortest paths are used as lengths for the edges in the complete graph. Although this approach can be computationally expensive when the number of nodes in the original graph is large as compared to the number of edges, as is the case for urban road networks, we pursue this strategy since in our case only a relatively small subset of nodes are customers and, therefore, the resulting complete graph can be computed relatively quickly.

After building the auxiliary complete graph $\mathrm{G}(\mathrm{N}_{\mathrm{sub}}, \mathrm{E}_{\mathrm{sub}})$, standard VRP techniques can be used to obtain an optimal assignment-routing on such a graph by solving problem (5.22) below ("(C)" denotes that we are considering a complete graph), where the optimization variables are the ones listed in

$\bar{n}_1$, $\bar{n}_2$	$\mathrm{SP}(\bar{n}_1, \bar{n}_2)$	$d^{\mathrm{SP}}_{n_1 n_2}$
0, 8	$\{(0,6),(6,7),(7,8)\}$	3
8, 0	$\{(8,6),(6,0)\}$	2
0, 3	$\{(0,5),(5,3)\}$	2
3, 0	$\{(3,5),(5,0)\}$	2
0, 1	$\{(0,5),(5,2),(2,1)\}$	3
1, 0	$\{(1,2),(2,3),(3,5),(5,0)\}$	7
8, 3	$\{(8,6),(6,0),(0,5),(5,3)\}$	4
3, 8	$\{(3,5),(5,0),(0,6),(6,7),(7,8)\}$	5
3, 1	$\{(3,5),(5,2),(2,1)\}$	3
1, 3	$\{(1,2),(2,3)\}$	5
1, 8	$\{(1,2),(2,3),(3,5),(5,0),(0,6),(6,7),(7,8)\}$	10
8, 1	$\{(8,6),(6,0),(0,5),(5,2),(2,1)\}$	5

TABLE 3. Shortest path and corresponding length between the depot and each customer node and between each pair of customer nodes in the non-complete graph $\mathrm{G}(\mathrm{N}, \mathrm{E})$ considered in Figure 3.

a^{C}_{qn}	Binary variable equal to 1 if vehicle $q \in \mathrm{V}$ is assigned to customer $n \in \mathrm{N}_{\mathrm{sub}}$.
$r^{\mathrm{C}}_{q n_1 n_2}$	Binary variable equal to 1 if vehicle $q \in \mathrm{V}$ traverses edge $(n_1, n_2) \in \mathrm{E}_{\mathrm{sub}}$.
$v^{\mathrm{C}}_{q n_1 n_2}$	Speed of vehicle $q \in \mathrm{V}$ on the edge $(n_1, n_2) \in \mathrm{E}_{\mathrm{sub}}$.
$\mathrm{arr}^{\mathrm{C}}_{qn}$	Arrival time of vehicle $q \in \mathrm{V}$ at node $n \in \mathrm{N}_{\mathrm{sub}}$.
$t^{\mathrm{C}}_{q n_1 n_2}$	Time spent by vehicle $q \in \mathrm{V}$ to traverse $(n_1, n_2) \in \mathrm{E}_{\mathrm{sub}}$.
$\mathrm{emis}^{\mathrm{C}}_{q n_1 n_2}$	Emissions generated by vehicle $q \in \mathrm{V}$ to traverse $(n_1, n_2) \in \mathrm{E}_{\mathrm{sub}}$.

TABLE 4. List of the optimization variables used in the formulation of the assignment-routing problem on the auxiliary complete graph.

Table 4. In such a problem, we assume that each vehicle can visit the nodes (including the depot) and edges at most once.

$$\text{Assignment-Routing (C):} \quad U^{\mathrm{C}}_1 = \overline{\mathrm{TGE}},\ U^{\mathrm{C}}_2 = \overline{\mathrm{TSC}} + \overline{\mathrm{TDT}} + \overline{\mathrm{MDD}},\ \text{and}\ U^{\mathrm{C}}_3 = \overline{\mathrm{LAC}} + \overline{\mathrm{MAT}}.$$
$$\text{Constraint function } u^{\mathrm{C}} \text{ is given by (5.23)–(5.32) below.} \tag{5.22}$$

$$\overline{\mathrm{TGE}} = \sum_{q \in \mathrm{V}} \sum_{(n_1, n_2) \in \mathrm{E}_{\mathrm{sub}}} \mathrm{emis}^{\mathrm{C}}_{q n_1 n_2} \qquad \overline{\mathrm{TSC}} = \sum_{q \in \mathrm{V}} \mathrm{SC}_q \sum_{(dep, n) \in \mathrm{E}_{\mathrm{sub}}} r^{\mathrm{C}}_{q\, dep\, n}$$