Centrally Coordinated Schedules and Routes of Airport Shuttles with LAX Terminals as Application Area

April 2023

A Research Report from the National Center for Sustainable Transportation

Petros Ioannou, University of Southern California

Pengfei Chen, University of Southern California

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NCST-USC-RR-23-14	N/A	N/A

4. Title and Subtitle Centrally Coordinated Schedules and Routes of Airport Shuttles with LAX Terminals as Application Area	**5. Report Date** April 2023 **6. Performing Organization Code** N/A
7. Author(s) Petros Ioannou, PhD, https://orcid.org/0000-0001-6981-0704 Pengfei Chen, PhD Student, https://orcid.org/0000-0003-4303-4676	**8. Performing Organization Report No.** N/A
9. Performing Organization Name and Address University of Southern California METRANS Transportation Consortium University Park Campus, VKC 367 MC:0626 Los Angeles, California 90089-0626	**10. Work Unit No.** N/A **11. Contract or Grant No.** Caltrans 65A0686 Task Order 070 USDOT Grant 69A3551747114
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology 1200 New Jersey Avenue, SE, Washington, DC 20590 California Department of Transportation Division of Research, Innovation and System Information, MS-83 1727 30th Street, Sacramento, CA 95816	**13. Type of Report and Period Covered** Final Report (November 2021 – December 2022) **14. Sponsoring Agency Code** USDOT OST-R Caltrans DRISI

15. Supplementary Notes

DOI: https://doi.org/10.7922/G2CC0Z03

16. Abstract

Today's airport terminals face a critical problem of traffic congestion in the terminal area partly caused by uncoordinated shuttle operations. The congestion near pick-up and drop-off points negatively affects passenger traffic leading to unnecessary idling, delays and congestion with negative impact on air quality and mobility. The need for an intelligent shuttle management system becomes more urgent with the development of information technologies, battery electric shuttles and autonomous vehicles. In this project, we developed a centrally coordinated shuttle scheduling and routing management system for mixed fleets of diesel and electric shuttles using a digital twin of LAX to LA downtown traffic road network by optimizing the total combined cost of energy consumption and travel time. A Co-Simulation Optimization method is used to solve the problem. The objective is to reduce congestion at the designated pick up and drop off points due to different shuttles showing up at these points during overlapping time windows which exceed the curb capacity. Another objective is to integrate into the system mixed fleet of shuttles that include diesel and battery operated. The proposed centrally coordinated shuttle scheduling and routing management system takes into account the characteristics of mixed shuttle fleets and is shown to reduce the operational cost such as energy consumption and delays. The results also suggest the deployment of electric shuttles in order to reduce emissions and improve air quality further.

17. Key Words	18. Distribution Statement
Shuttle Scheduling and Routing, Load Balancing System, Co-Simulation, Mixed Shuttle Fleet, Electric Shuttle, Autonomous Shuttle	No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	36	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

About the National Center for Sustainable Transportation

The National Center for Sustainable Transportation is a consortium of leading universities committed to advancing an environmentally sustainable transportation system through cutting-edge research, direct policy engagement, and education of our future leaders. Consortium members include: University of California, Davis; University of California, Riverside; University of Southern California; California State University, Long Beach; Georgia Institute of Technology; and University of Vermont. More information can be found at: ncst.ucdavis.edu.

Disclaimer

Acknowledgments

This study was funded, partially or entirely, by a grant from the National Center for Sustainable Transportation (NCST), supported by the U.S. Department of Transportation (USDOT) and the California Department of Transportation (Caltrans) through the University Transportation Centers program. The authors would like to thank the NCST, the USDOT, and Caltrans for their support of university-based research in transportation, and especially for the funding provided in support of this project. The authors also like to thank Justin Erbacci, David Reich and Michael Christensen from Los Angeles World Airports for their support in accessing data and making suggestions and describing the traffic issues and challenges the Los Angeles International airport has been facing.

NCST

Centrally Coordinated Schedules and Routes of Airport Shuttles with LAX Terminals as Application Area

A National Center for Sustainable Transportation Research Report

April 2023

Petros Ioannou, Ming Hsieh Department of Electrical Engineering, University of Southern California
Pengfei Chen, Ming Hsieh Department of Electrical Engineering, University of Southern California

NCST

TABLE OF CONTENTS

List of Tables

List of Figures

Centrally Coordinated Schedules and Routes of Airport Shuttles with LAX Terminals as Application Area

EXECUTIVE SUMMARY

A critical problem facing US airports as they respond to growth in services and operations is the limitation of curbside parking for shuttles to pick up and drop off passengers during peak hours. Today, shuttle companies and airport operations work independently without any schedule coordination, leading to frequent congestion near the pick-up and drop-off points that negatively affects passenger traffic leading to unnecessary idling, delays and congestion with negative impact on air quality. Accurate prediction of arrival times at the pick-up and drop off points depends on traffic conditions which are time varying as well as on the schedules of other shuttles sharing the same curbside spots. Without any form of central coordination, a single shuttle company cannot accurately develop a schedule that maintains a high quality of service at a reduced operational cost. This problem is exacerbated by existing and growing shuttle services provided by the airports themselves, centralized car rental facilities, and public transportation hubs. Furthermore, the transition of conventional shuttles to electric ones and the possibility of autonomous shuttles add additional complexities that necessitate the use of a centralized shuttle coordination system for optimum performance.

In this project, a CENtrally COordinated Shuttle system (CENCOS) is designed and evaluated by considering the recent development in theory, software and information technologies. The CENCOS system formulation is based on the combination of job assignment formulation and co-simulation optimization approach that combines real time traffic simulators with a route optimization algorithm in a feedback configuration. The system simulator captures the nonlinear impact of traffic loads on traffic flow which is then used in the optimization of the routes. Based on the more accurate predictions of traffic stated generated by the incorporated traffic simulator, the CENCOS system plans the schedules and routes of all shuttle vehicles in order to minimize curb congestion at the pick-up and drop-off points, reduce operational cost and improve quality of service. In the sight of emerging technology trends including burgeoning electrification and automation, the CENCOS system is designed to include multiple types of shuttles, such as electric shuttles and autonomous shuttles. The use of mixed fleet of shuttles powered by diesel and electricity, introduces additional constraints and cost criteria to be considered, since battery electric shuttles have a higher capital cost, shorter range, and longer refueling time than diesel shuttles. The benefits of the CENCOS co-simulation optimization system for a mixed shuttle routing system are compared with alternative shuttle management of schedules and routes without optimal centralized coordination. The results show that the proposed CENCOS system with the employed co-simulation optimization method provides significant reduction in total cost. Another result shows the total cost without including charging time and emissions decreases as the penetration of electric shuttles increases in the overall fleet. The CENCOS system is demonstrated in a real road network combining a ‘digital twin’ of the Los Angeles World Airport (LAWA) traffic network and a macroscopic network covering the road network from LAX to LA downtown. The CENCOS system can be applied to a

wide range of applications involving pick-up and drop-off points and shuttles in airports and other places.

The main outcomes of the project are listed as follows:

- The shuttles spend less waiting and traveling time in the airport area by participating in the proposed CENCOS system.
- The benefits generated by the CENCOS system depend on the background traffic and increase with heavy traffic.
- The CENCOS system shows savings in total cost of about 30% when compared with current shuttle management systems where each shuttle company operates individually without any coordination.
- The total cost of shuttle assignment from the CENCOS system decreases as the number of electric vehicles increases.
- The emissions go down drastically as the number of electric vehicles increases in the fleet.

We have to emphasize that the research performed is a preliminary step toward a coordinated shuttle scheduling and routing system using load balancing techniques and by no means captures the full complexity of shuttle transport. Some of the assumptions made need to be validated with experiments and some of the scenarios tested are rather simple when compared with the complexity of shuttle operations. This research however sets the foundations of the concept of centrally coordinated shuttle management system by solving some challenging problems whose solutions point to the direction for future research for eventual implementation.

1. Introduction

A critical problem facing US airports as they respond to growth in services and operations is the limitation of curbside parking for shuttles to pick up and drop off passengers. Today, shuttle companies and airport operations work independently without any schedule coordination, leading to frequent congestion near the pick-up and drop off points that negatively affects passenger traffic leading to unnecessary idling, delays and congestion with negative impact on air quality and quality of service to passengers. Accurate prediction of arrival times at the pick-up and drop off points depends on traffic conditions and the schedules of other shuttles sharing the same curbside spots. Without any form of central coordination, a single shuttle company cannot accurately develop a schedule that maintains a high quality of service and reduce operational cost. This problem is exacerbated by existing and growing shuttle services provided by the airports themselves to support “remote curb”, centralized car rental facilities, and public transportation hubs—particularly as more of these services are transitioned to autonomous vehicles. A CENtrally COordinated Shuttle system (CENCOS) is essential to improve the operations of the current system and support emerging technology trends including burgeoning electrification and automation and reduce operational costs. Research has shown that lowering operating costs and improving local air quality are the most compelling reasons for fleets to adopt electric vehicles [1], [2]. Further, fleets may also see reduced maintenance costs and a boost in public esteem [3]. Vehicle automation is a continuing trend in vehicle technologies in an effort to improve safety, passenger comfort and remove the randomness of human drivers leading to smoother traffic flows with expected energy savings. In the case of most airport shuttles, the routes are fixed and the challenging task of learning of the environment on new routes faced by automated vehicles does not apply. This simplifies the safety issues involved and makes automated shuttles quite feasible without costly infrastructure changes. An additional consideration for electrification is battery life and how it is affected by congestion as well as charging times and charging location. As some of the shuttles become automated operating on fixed routes the issue of safety comes up and issues such as safety gaps, collision avoidance, speeds, lane changes etc., add more constraints that the CENCOS needs to take into account. While some shuttles are in close vicinity of the airport and use the same daily routes, shuttles from longer distances may have the option of alternative routes which could be optimized.

Currently there is no existing system which can generate schedules and possibly routes for all shuttles in a way that it meets their operational objectives and service while minimizing congestion at the curbside. To centrally coordinate the shuttles in the traffic hubs such as LAX airport, it is important to notice that the traffic flow in the airport terminal area is not independent of the road network outside the airport area. To address the congestion of shuttles in the terminal area, we need to intelligently manage the schedule and route of each shuttle from their departure node, through stops and to the terminal nodes. The traditional bus scheduling and planning problem is divided into four aspects: timetabling, vehicle scheduling, crew scheduling and crew rostering [4]. The input to the schedule planning problem contains routes or lines to operate and how frequently. The average travel times between stops are also known parameters. The timetabling problem focuses on determining the decision of pairing routes or lines with a certain time, which is then called a trip. The vehicle scheduling problem then decides the assignment of shuttles or buses to each trip to make sure that every trip is assigned to exactly one vehicle. The solution of vehicle scheduling problem results in a set of vehicle blocks, where each vehicle block defines the trips paired to the vehicle. For each trip in a vehicle block, the crew scheduling problem then decides the assignment of one working period of one crew to a trip or a subsection of a trip. Crew rostering deals with the rosters based on crew duties. According to author's knowledge, the synchronization of bus timetabling problem that decides the departure times of trips of a whole day to maximize the synchronization of multiple events in order to smooth the passenger transfer process, is initially developed by Ceder et al. [5]. The recent trend of research on synchronization of bus timetabling can be found in [6]–[13]. Ibarra-Rojas and Rios-Solis developed dynamic bus timetabling techniques featured with oriented synchronization and evenly spaced departures [6]–[8]. Wu et al. studied a stochastic version of synchronization bus timetabling problem, where bus travel times are stochastic, and slack time is added into the timetable to mitigate the randomness of travel times so that the rate of transfer failures is reduced [9], [10]. Kang et al. developed a complex model to reduce the number of cases that passengers miss the connecting trains [11]. In [12], an event-driven model is developed for the train schedule to minimize the total travel time of all passengers and the energy consumption of trains.

Research on vehicle routing is very rich and many optimization tools have been developed over the years which are very useful in addressing some of the issues mentioned above. The Vehicle Routing Problem (VRP) formulation was first introduced by Dantzig and Ramser [14], as a generalization of the Traveling Salesman Problem (TSP) presented by Flood [15]. Since then, there is a significant amount of research on this topic which can be divided into 4 main categories. First, in static and deterministic problems, all inputs are known beforehand and vehicle routes do not change once they are in execution. This classical problem has been extensively studied in the literature, and we refer the interested reader to the recent reviews of exact and approximate methods by Baldacci et al. [16], Cordeau et al. [17], Laporte [18], [19], and Toth and Vigo [20]. Second, static and stochastic problems are characterized by inputs partially known as random variables, whose realizations are only revealed during the execution of the routes. Additionally, it is assumed that routes are selected a priori and only minor changes are allowed afterwards. Uncertainty may affect any of the input data like stochastic times where either service or travel times are modeled by random variables [21], [22]; and stochastic demands [23]–[27]. Third dynamic and deterministic problems have part or all of the inputs as unknown and appear dynamically during the design or execution of the routes. For these problems, vehicle routes are redefined in an ongoing fashion, requiring technological support for real-time communication between the vehicles and the decision maker (e.g., mobile phones and global positioning systems). Fourth, dynamic and stochastic problems have part or all of their inputs unknown and appear dynamically during the execution of the routes, but in contrast with the latter category, exploitable stochastic knowledge is available on the dynamically revealed information. As before, the vehicle routes can be redefined in an ongoing fashion with the help of technological support. For a comprehensive review of both the deterministic and the stochastic dynamic VRP, we refer the interested reader to [23]–[27]. Additional work on shortest route problems which cover the four categories mentioned can be found in [28]–[36] which also include work on multimodal routing and planning. With respect to electric vehicle routing, Ambrose and Jaller [37] examined the result of electric drayage trucks at the Port of Los Angeles and assessed emissions reductions with increased electrification of port truck operations. Nan et al. presented a mathematical programming model and solution method for path-constrained traffic assignment problem for electric vehicles in congested networks [38]. Bahrami et al. proposed a complementary equilibrium model for electric vehicles without violating driving range constraints [39]. Based on the assumption of large adoption of electric vehicles, Faridimehr et al. [40] proposed a two-stage stochastic programming model to determine the optimal network of charging stations for a community as well as the charging decision for each electric vehicle in this community. For a more detailed topic for electric vehicle traffic assignment, Yao et al. [41] compared electric vehicle's energy consumption rate on different road types from the floating car data collected from the road networks in Beijing.

Despite research in the area of scheduling and routing for shuttles and buses, there are still many issues that need to be addressed and new techniques need to be developed in order to make full use of the emerging electric vehicle technologies in a way that benefit the overall system and the environment. The complexity of the traffic network is immense due to the non-homogeneous dynamics of different vehicle classes at the vehicle level and nonlinear behavior at the traffic flow level. Mathematical models used by most TAP schemes cannot possibly capture the complexity of the real system in order to achieve the best possible outcome especially due to the added constraints of electric vehicles. The development of accurate mathematical models to describe traffic characteristics has always been a challenge and is becoming more of a challenge if electric vehicles are included in the model. The availability of fast computers and advanced software tools however, allows the development of traffic simulation models which can run in real time to provide the information and predicted states of the traffic network to choose routes that are more likely to be close to optimality than those based on simplified mathematical models. The challenge is how these simulation models can be integrated with optimization tools to generate more realistic outcomes. In past work [36], [42], we considered the use of real time traffic simulators as part of a centrally coordinated multimodal freight load balancing system and showed the significance of traffic simulators in planning freight routes to achieve a good balance of freight loads across the road and rail network. In [43], we explored the use of co-simulation optimization method to solve mixed freight truck routing problem. In this project we extended the work of [36], [42], [43] to incorporate a job assignment problem onto the load balancing layer in order to address the congestion around shuttle stop nodes. Considering that electric shuttles will be entering the market due to efforts to reduce emissions and most companies will be operating mixed fleets of shuttles, routing mixed fleets of shuttles in a coordinated manner that will have additional benefits to the environment and costs is an important research problem this project focused on.

The report is organized as follows. Section 2 deals with the main project content. Respectively Section 2.1 presents the digital twin of traffic network in an area that includes LAX all the way to downtown Los Angeles. Section 2.2 presents the main core of CENCOS system: its formulation and algorithm. Section 2.3 presents the key elements for the optimization algorithm as well as the emission model to evaluate vehicle emissions. Section 2.4 presents the simulation results that demonstrate the consistency of performance. Finally, conclusions are presented in Section 3 and appendices can be found in section 4.

2. Project Contents

2.1 LAX and digital twin

In this section, we configured and updated the LAX-LA downtown digital twin of traffic flow network that is developed in a previous project supported by LAX. The overall view of LAX-LA downtown road network is shown in Figure 1. The digital twin of traffic flow of this network is built using a microscopic traffic simulator, which details the interactions and modeling of dynamics of individual vehicles. However, during the test of the digital twin, we noticed that the computation time is exponentially large if the whole network is simulated using a microscopic traffic simulator. Considering the advantages of reducing computation time by aggregating traffic status from a macroscopic traffic simulator, we decided to combine the advantages of micro- and macroscopic traffic simulator and make a hybrid digital twin which uses a microscopic traffic simulator to model the traffic at the LAX area and a macroscopic traffic simulator to model the traffic at areas less important and complex than LAX. The microscopic LAX digital twin is shown in Figure 2 and the macroscopic complementary digital twin is shown in Figure 3.

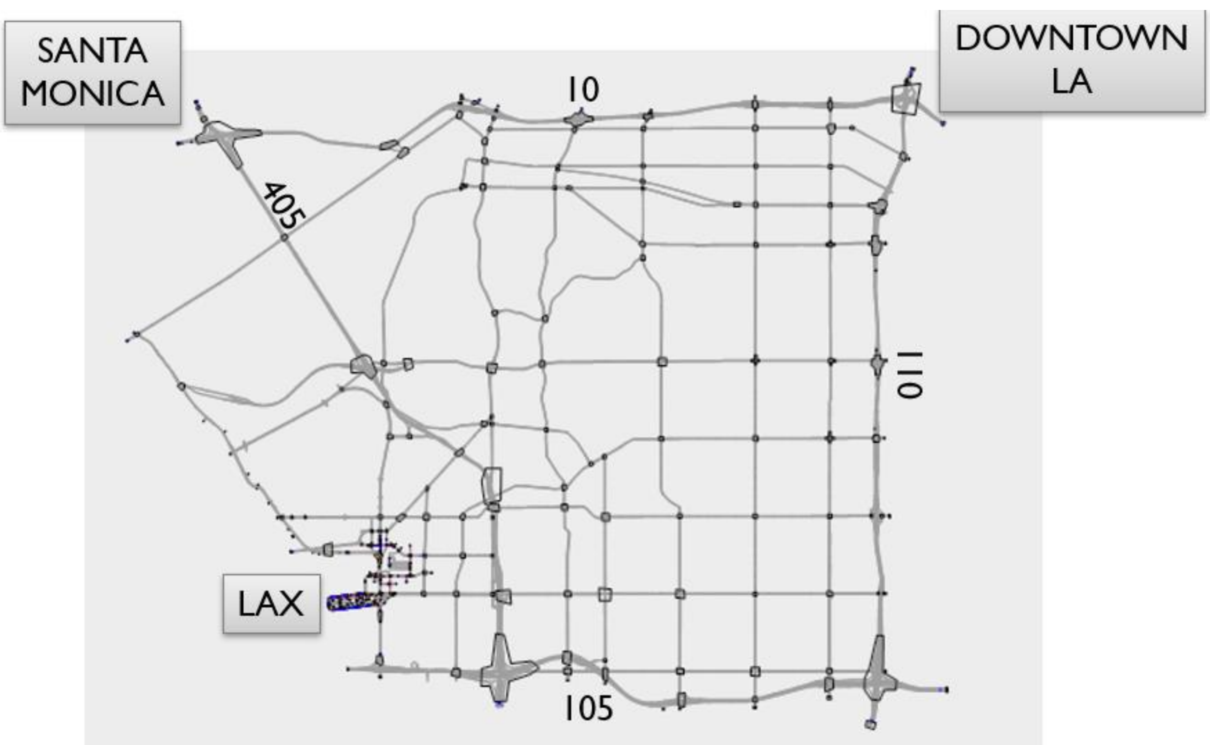

Figure 1. LAX-LA downtown digital twin of traffic network

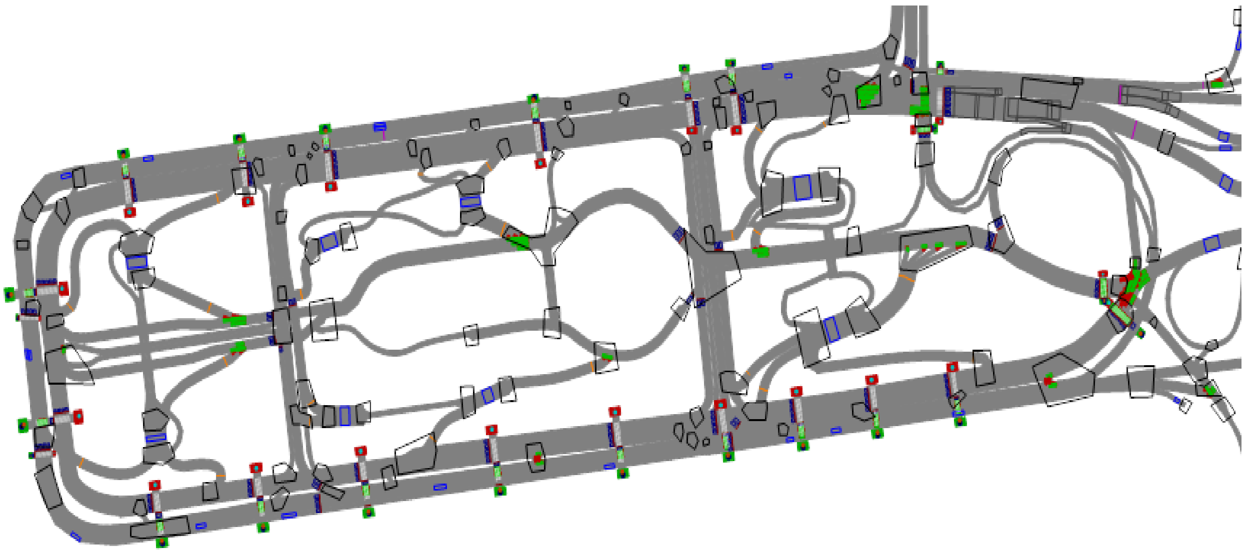

Figure 2. Microscopic digital twin of traffic flow network in the LAX area

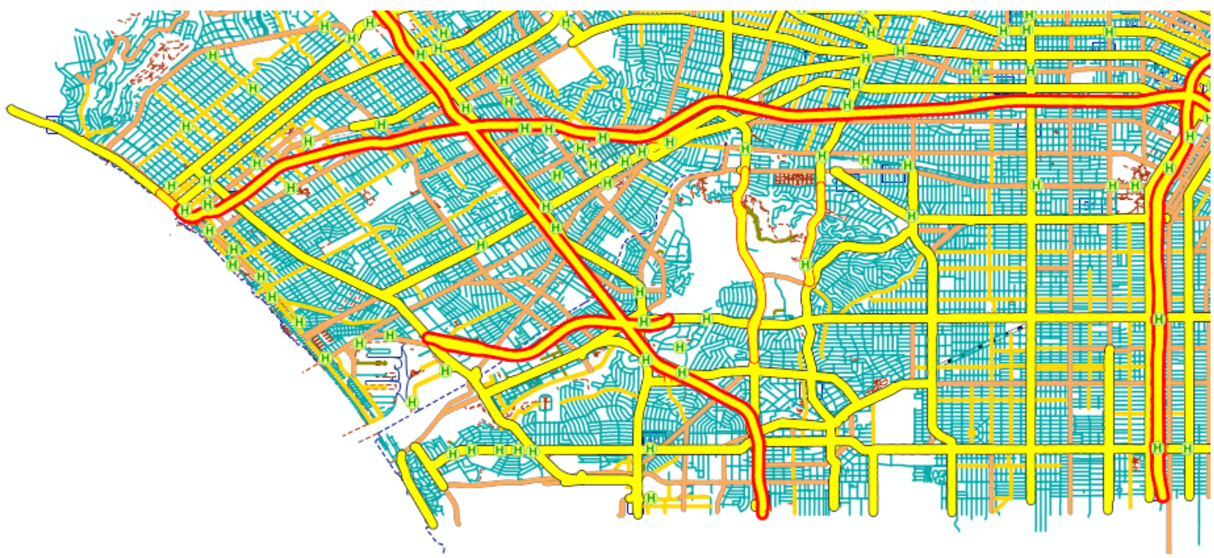

Figure 3. Macroscopic complementary digital twin of traffic flow network which connects with LAX traffic

For the microscopic LAX digital twin, the following components are configured: road geometry and characteristics, dynamic components such as traffic lights and signal control (.rbc file for signal control in the LAX terminal area), road sensors, background traffic flow. For the macroscopic complementary digital twin, we configured road geometry and used traffic flow data from sources such as Freeway Performance Measurement System (PeMS) of Caltrans, Los Angeles Department of Transportation (LADOT) and Southern California Association of Governments (SCAG). The shuttle demand data are collected from the shuttle company websites. Shuttle and bus services include FlyAway Bus Service, Scheduled buses (Antelope

Valley Airport Express, Central Coast Shuttle, Mickey's Space Ship Shuttle, etc.), buses to LAX City Bus Center and hotel shuttles.

The energy consumption is usually estimated using a set of static vehicle features, such as length, shape, et al. However, the shuttles are working under various conditions and their energy consumption characteristics vary under different working conditions. To achieve an accurate estimation of these characteristics, we used a method that combines mapping the driving speed to working mode and mapping the working mode to energy consumption rate developed by us in a previous project. In this method, the analytical model of typical diesel and electric engines are implemented and tested with driving cycles. Driving cycles are files that document the speed of a specific vehicle interval by interval under some driving mode. The analytic model [44] is used to describe the diesel engine and [45] to describe the electric engine of heavy-duty vehicles. More details can be found in [43].

2.2 CENtrally COordinated Shuttle system (CENCOS)

2.2.1 Optimization model

In this subsection, we developed a centrally coordinated airport shuttle system referred to as CENCOS that minimizes the overall cost by reducing congestion at airport curbside and improving quality of service using an innovative co-simulation optimization technique based on a digital twin of the traffic at the airport road network. Our approach involves a two-layer formulation. The upper layer referred to as the scheduling layer and the lower level layer referred to as the load balancing layer. Specifically, the upper layer is in charge of scheduling the order of shuttles arriving at the pick-up/drop-off stops around the network based on constraints such as traveling time, energy cost, charging time for electric shuttles, etc. This layer is formulated based on a job assignment problem with complex constraint specifications on charging behavior for each electric shuttle. Based on the scheduling of shuttles across the road network, a traffic load balancing assignment is used to assign the shuttle demand on the transportation system in order to minimize the travel time and energy consumption cost. This layer is formulated based on a Traffic Assignment Problem (TAP). The general framework of the optimization model can be described as follows: a central coordinator receives from individual users (shuttle companies) their shuttle origin/destination (O/D) demand with time windows and information about the mixed fleet of diesel and electric shuttles. Then the CENCOS system generates a scheduling plan by deciding the order of shuttles to be served at each stop based on the network link cost and desired service constraints. The network link cost consists of the travel time and energy consumption cost. Then the order of shuttles entering and leaving each stop is input into the load balancing assignment layer, which generates routes for each shuttle that minimize an overall system cost. The impact of the loads on each link is taken into account to achieve load balancing across the road network. The road network traffic states such as travel time and traffic flow on each road link are generated by a traffic simulator. The dynamic traffic states are then used in the overall co-simulation optimization procedure to intelligently calculate the direction and step size of the algorithm in each iteration. The traffic states of each segment are then fed back to the scheduling layer for the calculation of travel time and energy consumption. The travel time is also important in calculating battery life in the case of electric shuttles. The framework of CENCOS is summarized in Figure 4.

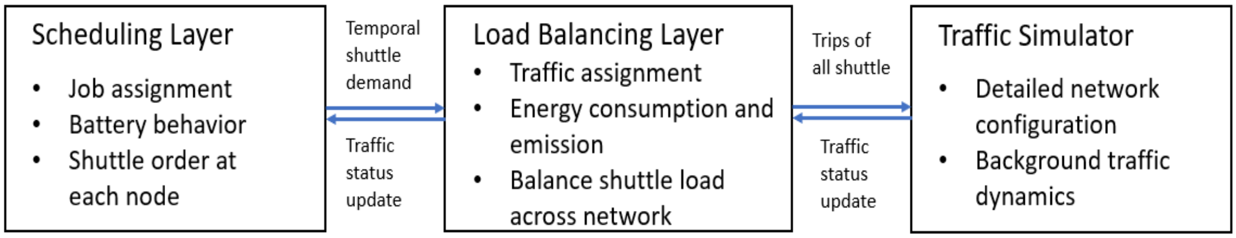

Figure 4. Framework of CENCOS formulation

2.2.1.1 Scheduling Formulation

We consider a road network represented by a directed graph $G = (V, E)$, where V is the node set and E is the link set. There are two sets of shuttle operation jobs that need to be fulfilled: short-distance job set J and long-distance job set U. The route of each operation job j is predetermined, while the routes for long-distance jobs need to be decided by the shuttle coordination system. The route can be represented by a sequence of ordered nodes $M^j = \{j(1), j(2), \dots, j(m^j)\}$, where m^j is the total number of stops in operation j. The shuttle coordination system also takes charge of scheduling, so that for each operation job $j \in J \cup U$, the shuttle arrives at the airport within a time window $[T_l^j, T_u^j]$, where T_l^j / T_u^j is the earliest/latest arrival time at the destination node $j(m^j)$. By taking into account the type of shuttles, the route that shuttles serving long-distance should take, the timing along the route as well as charging behavior if the shuttle is electric, the total cost which is a function of travel time and energy consumption can be minimized. In the following, we list the notation, variables and parameters used in the mathematical formulation of the problem:

- $t_{i,j}$: The arrival time of operation job j at node i;
- $h_j = 1$ if the operation job j is fulfilled by an electric shuttle; $h_j = 0$, otherwise;
- $x_{i,j,k} = 1$ if the operation job j precedes another operation job k at node i; $x_{i,j,k} = 0$, otherwise;
- $\hat{b}_j^i$: The battery status when the shuttle for operation job j arrives at node i in units of kWh;
- $\bar{b}_j^i$: The battery status when the shuttle for operation job j leaves node i in unit of kWh;
- η: value of time;
- $\varepsilon_{u,v}^e$: electric energy consumption from node u to node v in unit of \$;
- $\varepsilon_{u,v}^d$: diesel energy consumption from node u to node v in unit of \$;
- $p_{i,j}$: processing time at node i for operation job j;
- $d_{u,v}$: the travel time from node u to node v;
- ζ: electric energy charging rate in unit of kW;

- η: value of time;
- W, W_1, W_2: large numbers

Given the above notation we formulate the problem as follows:

$$Min_{t,\hat{b},\bar{b},x,h,M^j:j\in U} \sum_{j\in J}(\eta(t_{j(m^j),j} - t_{j(1),j}) + \sum_{r\in M^j\backslash\{m^j\}}(\varepsilon^e_{j(r),j(r+1)}h_j + \varepsilon^d_{j(r),j(r+1)}(1-h_j)))$$

Subject to,

$$W(1-h_j) + t_{j(r+1),j} \geq t_{j(r),j} + p_{j(r),j} + d_{j(r),j(r+1)} + \frac{\hat{b}_j^{j(r)} - \bar{b}_j^{j(r)}}{\zeta}, \forall j \in J, r \in M^j \quad (1)$$

$$W(1-h_j) + \bar{b}^{j(r+1)} \geq \hat{b}^{j(r)} + \varepsilon^e_{j(r),j(r+1)}, \ \forall j \in J, r \in M^j \quad (2)$$

$$W(2 - x_{i,j,k} - h_j) + t_{i,k} \geq t_{i,j} + p_{i,j} + \frac{\hat{b}_j^i - \bar{b}_j^i}{\zeta}, \ \forall j,k \in J, j \neq k, i \in V \quad (3)$$

$$W_1(1 - x_{i,j,k}) + W_2 h_j + t_{i,k} \geq t_{i,j} + p_{i,j}, \ \forall j,k \in J, j \neq k, i \in V \quad (4)$$

$$W_1 x_{i,j,k} + W_2(1-h_j) + t_{i,j} \geq t_{i,k} + p_{i,k} + \frac{\hat{b}_k^i - \bar{b}_k^i}{\zeta}, \ \forall j,k \in J, j \neq k, i \in V \quad (5)$$

$$W(x_{i,j,k} + h_j) + t_{i,j} \geq t_{i,k} + p_{i,k}, \ \ \forall j,k \in J, j \neq k, i \in V \quad (6)$$

$$T_l^j \leq t_{j(m^j),j} \leq T_u^j, \ \forall j \in J \quad (7)$$

$$t_{i,j}, \hat{b}_j^i, \bar{b}_j^i \geq 0, x_{i,j,k}, h_j \in \{0,1\} \quad \forall i \in V, j,k \in J \quad (8)$$

The objective function aims to minimize the sum of the travel time cost and energy consumption cost of all operation jobs. The total travel time for operation job j equals to the time duration from its first stop to the last stop. The energy cost equals to the energy consumed for each link along the route. Constraint (1) requires that for an operation job j, its arrival time at its $(r+1)$st node should be at least greater than the time after traveling along link $(j(r), j(r+1))$ and charging if it is fulfilled by an electric shuttle from its rth node. Constraint (2) states that if operation job j is allocated to an electric shuttle, then the battery energy after leaving a stop should be enough for it to cover the distance to its next stop. Constraints (3)-(6) describe the conflicting relations of different operation jobs. Constraint (3) makes the following statement: if operation job j precedes operation job k at some stop i, and operation job is fulfilled by an electric shuttle, then the shuttle of operation job k can arrive at stop i only if operation job j finishes all the processing and charging action. Constraint (4) describes the scenario when operation j fulfilled by a regular (non-electric) shuttle precedes operation k. Constraint (5) describes the situation that operation j does not precede operation k fulfilled by an electric shuttle at node i. Constraint (6) describes that operation j does not precede operation k fulfilled by a regular shuttle at node i. Constraint (7) declares the time window at destination node for each operation job. The solution for the scheduling problem provides a temporal assignment of shuttles across the whole road network detailing in the order of shuttles entering and leaving at each stop node as well as the energy being charged at each node. Note here, the travel time of each shuttle is updated by a traffic simulator, which is able to well capture the interactions between the shuttles and background traffic. The solution of the scheduling layer is then input into the load balancing assignment layer, where the solution is regarded as demand for the whole road network and needs optimal assignment arrangement spatially.

2.2.1.2 Load Balancing Formulation

Based on the spatially assignment of shuttle demand from the scheduling layer, the demand is then input into the load balancing layer to be assigned optimally spatially across the whole network. The formulation of load balancing layer can be described as follows: consider the road network to be a directed graph $G(E,V)$, where E is the set of all links and V is the set of all nodes. Among all the nodes, a subset of them are origin nodes, denoted as O, i.e., $O \subset V$. Another subset of nodes are destination nodes, denoted as D, i.e., $D \subset V$. For a certain pair of origin and destination nodes (i,j), $i \in O, j \in D$, the demand volume is $q_{i,j}$. All the shuttle types are included in a set U. To represent the distribution of shuttles, we use m_i^u as the number of total available shuttles of type u at node i. To cope with the temporal dimension, we discretize the time horizon into $|K|$ time intervals and use K as the set of all the time intervals. The following notation is used in the formulation to follow:

- $R_{i,j}^u$: The set of routes for shuttles of type u from i to j, $i \in O, j \in D$;
- $X_{i,j,r,k}^u$: The number of shuttles of type u from i to j, $i \in O, j \in D$, using route r in route set $R_{i,j}^u$ with a departure time k;
- $S_{i,j,r,k}^u(X)$: The average service cost per container fulfilled by a shuttle of type u from i to j, $i \in O, j \in D$, using route r in route set $R_{i,j}^u$ with a departure time k.

Given the above notation we formulate the problem as follows:

$$min_X \ \sum_{k\in K}\sum_{i\in O}\sum_{j\in D}\sum_{u\in U}\sum_{r\in R_{i,j}^u} S_{i,j,r,k}^u(X)X_{i,j,r,k}^u \tag{9}$$

$$\sum_{k\in K}\sum_{u\in U}\sum_{r\in R_{i,j}^u} X_{i,j,r,k}^u \ = \ q_{i,j}, \forall\, i \in O, j \in D \tag{10}$$

$$\sum_{k\in K}\sum_{j\in J}\sum_{r\in R_{i.j}^u} X_{i,j,r,k}^u \ \le \ m_i^u, \forall\, i \in I, u \in U \tag{11}$$

$$X_{i,j,r,k}^u \ge 0 \tag{12}$$

Equation (9) is the objective function, which aims to minimize the sum of the service cost of all the shuttle. $S_{i,j,r,k}^u(X)$ is the unit service cost of transporting a load with a shuttle of type u using route r from i to j at time k given X. The cost $S_{i,j,r,k}^u(X)$ is given by:

$$S_{i,j,r,k}^u(X) = \ C_{i,j,r,k}^u(X) + \eta\, T_{i,j,r,k}^u(X) \tag{13}$$

where $C_{i,j,r,k}^u(X)$ is the cost of the consumed energy, $T_{i,j,r,k}^u(X)$ is the travel time and η is the value of time used in the scheduling layer. The energy and travel time cost depend on the dynamics of the traffic network. The dynamics of the traffic network can be expressed as nonlinear dynamic functions of all decision variables, denoted as X, and is discussed in the following parts. In our case, the energy cost depends on the dynamics of the traffic network. More specifically, we formulate the energy cost coefficient of each shuttle type as a polynomial function of the speed of the road link, where the parameters of the function are estimated using regression over a set of testing data. Here we assume each shuttle is responsible for one demand, so that the total number of shuttles for an O/D pair is equal to the demand of the O/D pair, as shown in equation (10). Equation (11) represents the constraints on availability of a certain type of shuttle at each node. Equation (11) can also be used to formulate the distribution of available mixed shuttles over the road network at the beginning of the time horizon.

The dynamics of a traffic network are highly nonlinear and exhibit the following temporal-spatial relations: traffic flow dynamics in a link and between links. The dynamics in a link describe how the traffic flow moves from the upstream end of a link to the downstream end, while the dynamics between links describe how the traffic flow propagates across the traffic network. In most of the literature of vehicle routing, the complex dynamics of the traffic network are overly simplified and the dynamics between links are ignored. As a result, the calculated optimum routes may not be optimum in a real situation. In our approach, we introduce the following changes that make it more likely for a theoretical optimum to be closer to the one in practice:

- Instead of using a simplified mathematical model to account for the complex traffic dynamics, we use a traffic simulation model in a co-simulation optimization approach. The simulation model provides a far more accurate description of the traffic dynamical characteristics to be used by the optimum route generator.
- To efficiently apply the simulation model, we construct a service network layer as a connection between the optimizer and the simulation model.
- To speed up the iterative algorithm process, we propose a way to intelligently choose the direction and step size at each iteration based on the knowledge of the marginal cost.

In order to explain the proposed method, we first discuss the configuration of the service network and the changes it brings to the above formulation. To differentiate the notation between the service network and the road traffic network, we use the following terminologies:

- Road network link: edge in the road network
- Path: a sequence of concatenated road network links
- Service segment: edge in the service network
- Route: a sequence of concatenated service segments

A service network can be configured based on a traffic network in the following steps:

- Collect a subset of nodes in the traffic network including all O/D nodes as well as the nodes necessary for the routing of shuttles to form the service node set NS. These necessary nodes can be terminals, stops, charging stations and so on.
- Construct a set of segments L connecting nodes in NS.

The service network can be seen as an abstracted upper layer of the traffic network. An example of a road network with its associated service network is shown in Figure 5, where the bottom level is the road network and the upper level is the service network as an abstraction of the road network. The blue nodes in the service network are important nodes such as O/D nodes and intersections of freeways. Based on the service network nodes, the artificial segments are created to fully connect the service network.

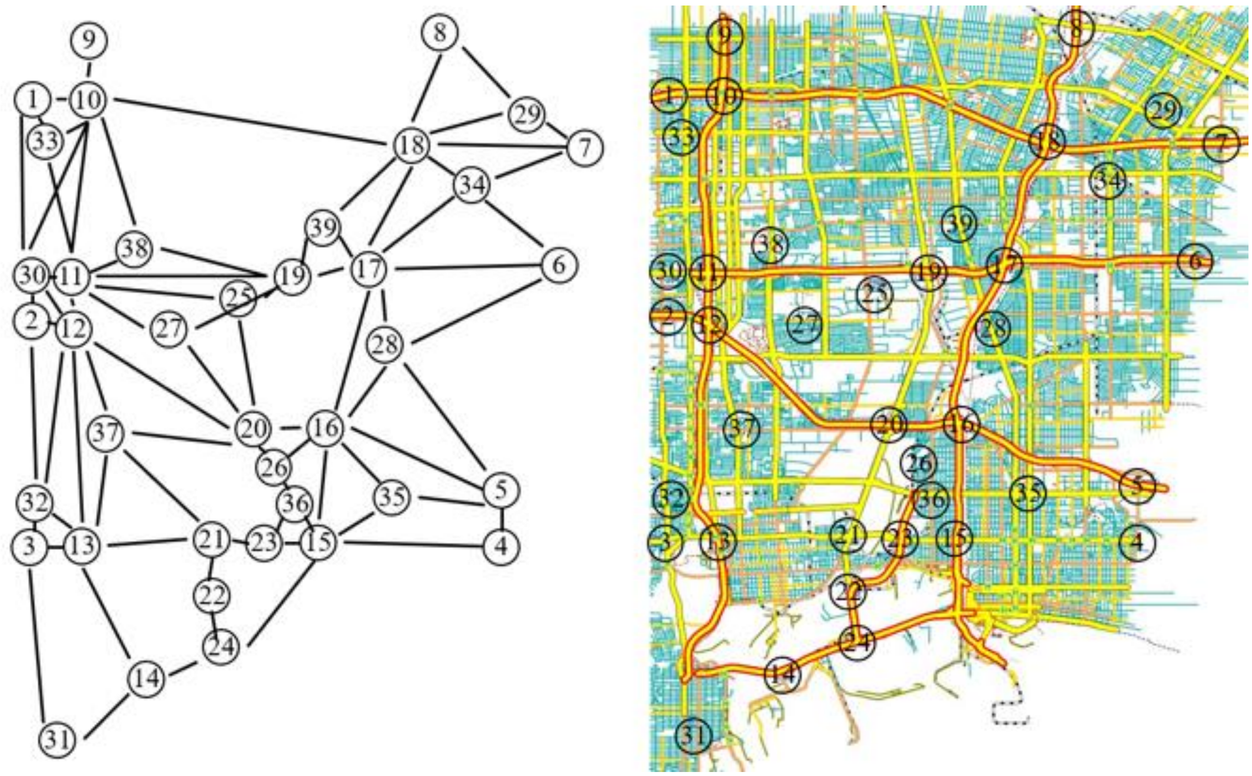

Figure 5. Service network example

With the inclusion of the service network, the relations between routes and links can be divided into two parts: relations between routes and service segments and relations between service segments and traffic network links. The relations between routes and service segments are specified as follows:

$$\sum_{i\in O}\sum_{j\in D}\sum_{u\in U}\sum_{r\in R_{i,j}^{u}}\sum_{\tau\leq k} X_{i,j,r,k}^{u}\delta_{l,r,\tau,k}^{u} = x_{l,k}^{u} \tag{14}$$

where $l \in L, k \in K$ and $\delta_{l,r,\tau,k}^{u} = 1$ when the shuttle of type u uses route r with departure time τ passing through segment l at time k, otherwise, $\delta_{l,r,\tau,k}^{u} = 0$. As for the relations between the service segment and traffic network links, we denote as $t_{l,k}^{p}$ the travel time on path p if a shuttle departs from the origin of segment l at time k. Assume links constituting path p to be $e_{p,1}, e_{p,2}, \ldots, e_{p,N_p}$, where N_p is the total number of links on path p. We define $\xi_{e,k}$ as the entering time at link e of a shuttle with a departure time k from the origin of that path. With $w_{e,k}$ to be the travel time of link e at time k, we now write the travel time of a path as follows:

$$t_{l,k}^{p} = \sum_{n_p=1}^{N_p} w_{e_{p,n_p}} \xi_{k,e_{p,n_p}} \tag{15}$$

$$\xi_{k,e_{p,1}} = 1 \tag{16}$$

$$\xi_{k,e_{p,n_p+1}} = \xi_{k,e_{p,n_p}} + w_{e_{p,n_p},\xi_{k,e_{p,n_p}}} \tag{17}$$

where $n_p = 1, \ldots, N_p - 1$. To make the notation simpler, we let $\widehat{w}_{p,n_p,k} \equiv w_{e_{p,n_p},\xi_{k,e_{p,n_p}}}$ to denote the travel time of link e_{p,n_p} on path p with the path departure time being $\xi_{k,e_{p,n_p}}$. Given the service segment volume $x_{l,k}^{u}$ and the path set of segment l, the vehicle dispatching problem in the traffic network can be expressed as follows:

$$\min_{y} TC = \sum_{k\in K} \sum_{l\in L} \sum_{p\in P_l} (c_{l,k}^{p,u} + \eta\, t_{l,k}^{p,u})\, y_{l,k}^{p,u} \tag{18}$$

where TC stands for the total cost of the assignment with mixed shuttles, which is a combined value of energy consumption and travel time cost; $c_{l,k}^{p,u}$ is the energy consumption coefficient for shuttles of type u passing through path p of segment l at time k; $t_{l,k}^{p,u}$ is the travel time of the path p in segment l that departs at time k; $y_{l,k}^{p,u}$ is the number of shuttles of type u assigned to pass through path p of segment l at time k and η is the value of time as mentioned before. The total cost is the sum of the energy consumption and travel time cost of all the segments with respect to time. The objective is to find an assignment for the mixed shuttles with minimum total cost. The constraints are defined by equations (14)-(17) generated from the service network as well as the complex dynamics from the simulated traffic network. In our method, the nonlinear dynamics of the traffic flow network are represented by the real time traffic flow simulation model that generates the states of the network to be used in the optimization problem. Aside from equations (14)-(17), the following equations are used to represent the relation between variables from the service network and the simulated traffic network:

$$\sum_{p\in P_l} y_{l,k}^{p,u} = x_{l,k}^{u}, \ \forall\, l \in L, k \in K \tag{19}$$

$$y_{l,k}^{p,u} \geq 0, \forall\, l \in L, p \in P_l, k \backslash in\, K \tag{20}$$

2.2.2 Optimization algorithm

In this subsection, we discuss the optimization algorithm used to solve the problem formulated above. Figure 6 gives a general overview of the method. The scheduling optimization and load balancing optimization parts are responsible for the scheduling and load balancing layer respectively. These two layers together play a central role in the optimization procedure; in practice, it can be a central coordinator whose aim is to manage shuttles scheduling and routing to fulfill demands at minimum system cost. The inputs to the system are shuttle demands, details of time window for each stop, the distribution of electric shuttles, emission model and other predetermined parameters. Shuttle demands represent the shuttles that transfer loads from origin to destination nodes. These demands are specified with time window on each stop as well as the type of shuttle fulfilling it, diesel or electric. The characteristics include the physical (weight, length, frontal area, et al.), dynamics (max speed, acceleration, et al.) and energy consumption (the amount of energy consumed based on the dynamic states). Based on the energy consumption characteristics of diesel/electric shuttles, the cost coefficients on each segment of both types of shuttles are calculated under different traffic conditions. An emission model from National Renewal Energy Laboratory (NREL) is used to calculate the emissions. A real-time traffic simulator is used to capture the dynamical characteristics of traffic and provide traffic status such as travel times along the links and routes as well as estimates of the energy cost of diesel and electric shuttles depending on the simulated traffic flow. The information from the simulator is used by the service graph optimization component to update the marginal cost of each service segment, which is used to update the route cost and the traffic status for the scheduling optimization part. Based on the updated traffic status, the Gurobi optimizer reoptimizes the order of shuttles entering and leaving each stop and updates the input into the load balancing layer. Based on the shuttle order decision and simulated route cost, the route collection for each O/D pair is updated as well. Then given the updated route collection, the assignment of diesel/electric shuttle for each O/D pair is updated by solving an integer combinatorial programming problem using a properly selected efficient step size. The new assignment is then generated and passed to the next iteration.

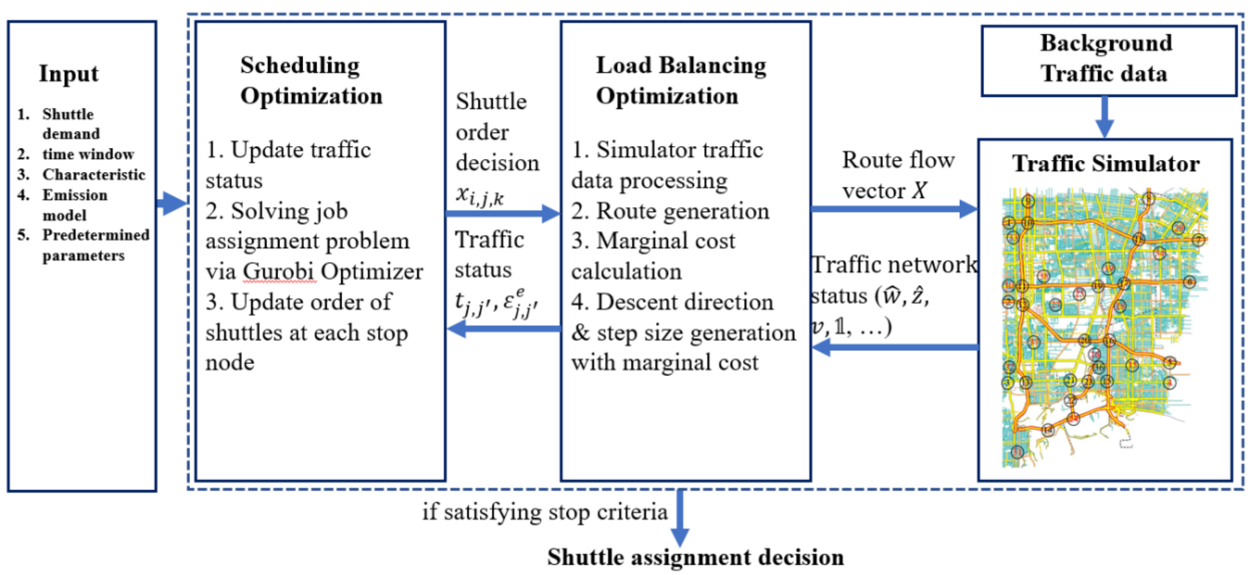

Figure 6. Framework of proposed method

The traffic simulator uses two types of inputs: background traffic flow and assignment traffic flow. The background traffic flow is obtained from various sources, such as PeMS [46] and Google Maps [47]. The assignment traffic flow is generated by the optimizer. The co-simulation optimization procedure iterates in a feedback loop that involves the traffic simulator, the scheduling and load balancing optimization. Through this procedure, the assignment and traffic flow states are sequentially updated until convergence is achieved. The difficulty in this procedure is calculating the marginal cost of each route, which is equal to the change in the total cost as a result of adding one unit of demand on that route. Since the total cost TC of equation (18) is complex, the marginal cost with respect to a route cannot be calculated directly. One way to calculate the marginal cost is to use Monte Carlo to simulate the impact of one unit of demand on each route at each time. However, it is impractical to enumerate all routes due to the fact that the number of possible routes grows exponentially with respect to the service network size. Our proposed approach bypasses this issue and works as follows:

1. Initialize cost coefficients based on the physical features such as speed limit for each segment l and iteration number $n = 0$. Configure the time window, diesel/ electric shuttle distribution for each demand in the job assignment problem. Initialize the travel time and energy consumption parameters from historic data. For the load balancing part, initialize the route collections for each O/D pair based on the segment cost calculated with the cost coefficients.
2. Check if the scheduling objective function value of the current iteration converges. If it converges, then stop the procedure and return with the order of shuttles at each stop node and corresponding routes; otherwise, continue to the next step.
3. Update the traffic status, including travel time and energy consumption for the scheduling optimization part.
4. Solve the job assignment problem using Gurobi Optimizer getting the order of shuttles at each stop node.
5. Based on the scheduling of each shuttle, establish the initial route flow vector $X^{(0)}$.
6. Check if the load balancing objective function value of the current iteration converges, i.e., $|TC(X^{(n)}) - TC(X^{(n-1)})| < \varepsilon$; ε is set to be a small number. If it converges, then go to step 2 and return with route flow vector; otherwise, continue to the next step.
7. Input the route flow vector $X^{(n)}$ into the traffic simulator and obtain the marginal cost of each segment.
8. Update the marginal cost of each segment as well as diesel/electric routes for each O/D pair and check whether there is a new minimal marginal cost route. If there is, then add it to the route collection.
9. Solve the following optimization problem for each origin node o to obtain a feasible route flow vector $\hat{X}^n$.

$$\min_{X} \sum_{u\in U} \sum_{k\in K} \sum_{j\in D} \sum_{r\in R^u_{o,j}} MC^u_{o,j,r,k}\ X^u_{o,j,r,k} \tag{21}$$

$$\sum_{u\in U} \sum_{k\in K} \sum_{r\in R^u_{o,j,k}} X^u_{o,j,r,k} = q_{o,j}, \forall\, j \in D \tag{22}$$

$$\sum_{k\in K} \sum_{j\in D} \sum_{r\in R^u_{o,j,k}} X^u_{o,j,r,k} \le m^u_o, \forall\, u \in U \tag{23}$$

where $MC^{u}_{o,j,r,k}$ is the marginal cost of route r from o to j with a shuttle of type u departing at time k. The marginal cost of a route is the sum of the marginal costs of the segments along the route.

10. Set the route flow vector for the next iteration as $X^{(n+1)} = X^{(n)} + \lambda^{(n)} \cdot (\hat{X}^{n} - X^{(n)})$, where $\lambda^{(n)}$ is the step size at the nth iteration, then go back to step 6. The step size $\lambda^{(n)}$ at the nth iteration is selected as in [36].

In the optimization algorithm, the marginal cost of each segment plays an important role, in pointing out the direction as well as the step size for the next iteration for the optimization algorithm. In the next subsection, we will present the calculation of marginal cost, which is used for the evaluation of the routes. In addition, we present the emission model used by the optimization procedure.

2.3 Evaluation of optimum routes

2.3.1 Marginal cost

The marginal cost represents the change in the total cost if one unit of demand/shuttle is changed on the path. It can be formulated as follows:

$$MCP^{p',u'}_{l',k'} = \frac{\partial TC}{\partial y^{p',u'}_{l',k'}} = \frac{\partial \sum_{k\in K}\sum_{l\in L}\sum_{p\in P_{;}}\left(c^{p,u}_{l,k} + \eta\, t^{p,u}_{l,k}\right) y^{p,u}_{l,k}}{\partial y^{p',u'}_{l',k'}}$$

$$= c^{p',u'}_{l',k'} + \eta\, t^{p',u'}_{l',k'} + \eta \sum_{k\in K}\sum_{l\in L}\sum_{p\in P_{;}} \frac{\partial t^{p,u}_{l,k}}{\partial y^{p',u'}_{l',k'}} y^{p,u}_{l,k} + \sum_{k\in K}\sum_{l\in L}\sum_{p\in P_{;}} \frac{\partial c^{p,u}_{l,k}}{\partial y^{p',u'}_{l',k'}} y^{p,u}_{l,k} \qquad (24)$$

where the first two terms are the cost of the path and the third term describes the travel time cost change due to the impact on the link travel time based on the dynamics of the traffic flow system. The fourth term accounts for the change of energy cost associated with the changes in link volume and can be calculated approximately using the traffic network states from the simulator. According to the derivative chain rule and equation (15):

$$\frac{\partial t^{p,u}_{l,k}}{\partial y^{p',u'}_{l',k'}} = \sum_{n_p=1}^{N_p} \frac{\partial \hat{w}_{p,n_p,k}}{\partial y^{p',u'}_{l',k'}} = \sum_{n_p=1}^{N_p} \frac{\partial \hat{w}_{p,n_p,k}}{\partial \hat{z}_{p,n_p,k}} \frac{\partial \hat{z}_{p,n_p,k}}{\partial y^{p',u'}_{l',k'}} \qquad (25)$$

where $\hat{z}_{p,n_p,k}$ is the traffic volume of the link e_{p,n_p} on path p with the path departure time being $\xi_{k,e_{p,n\_p}}$.The term $\frac{\partial \hat{w}_{p,n_p,k}}{\partial \hat{z}_{p,n_p,k}}$ represents the travel time change in link e_{p,n_p} at time $\xi_{k,e_{p,n\_p}}$. caused by changing the link volume by one unit. One of the most commonly used relationships between link volume and travel time is the Bureau of Public Roads (BPR) function [48].

$$w_e = t_f(1 + \alpha_e \left(\frac{z_e}{CAP_e}\right)^{\beta_e} \qquad (26)$$

where w_e is the link travel time, t_f is the link free-flow travel time, z_e is the vehicle volume on link e and CAP_e is the road link capacity. α_e and β_e are parameters for the model and can be estimated through historical traffic data. Then the link travel time derivative $\frac{\partial \widehat{w}_{p,n_p,k}}{\partial \hat{z}_{p,n_p,k}}$ based on equation (26) can be written as follows:

$$\frac{\partial \widehat{w}_{p,n_p,k}}{\partial \hat{z}_{p,n_p,k}} = \frac{\alpha_{e_{p,n_p}} \beta_{e_{p,n_p}} t_f \hat{z}_{p,n_p,k}^{\beta_{e_{p,n_p}}-1}}{CAP_{e_{p,n_p}}} \equiv B_{p,n_p,k} \hat{z}_{p,n_p,k}^{\beta_{e_{p,n_p}}-1} \tag{27}$$

After the derivation, the final form of marginal cost is:

$$\begin{aligned}
MCP_{l',k'}^{p',u'} &= c_{l',k'}^{p',u'} + \eta\, t_{l',k'}^{p',u'} \\
&\quad + \eta \sum_{k\in K} \sum_{l\in L} \sum_{p\in P_;} \sum_{n_p=1}^{N_p} \frac{\partial \widehat{w}_{p,n_p,k}}{\partial y_{l',k'}^{p',u'}} y_{l,k}^{p,u} \\
&\quad + \sum_{k\in K} \sum_{l\in L} \sum_{p\in P_;} \sum_{n_p=1}^{N_p} \frac{\partial h^u(v_{p,n_p,k})}{\partial \widehat{w}_{p,n_p,k}} \frac{\partial \widehat{w}_{p,n_p,k}}{\partial y_{l',k'}^{p',u'}} d_{p,n_p} y_{l,k}^{p,u} \\
&= c_{l',k'}^{p',u'} + \eta\, t_{l',k'}^{p',u'} + \sum_{k\in K} \sum_{l\in L} \sum_{p\in P_;} \sum_{n_p=1}^{N_p} \left(\eta + \frac{\partial h^u(v_{p,n_p,k})}{\partial \widehat{w}_{p,n_p,k}} d_{p,n_p}\right) \frac{\partial \widehat{w}_{p,n_p,k}}{\partial y_{l',k'}^{p',u'}} y_{l,k}^{p,u} \\
&\approx c_{l',k'}^{p',u'} + \eta\, t_{l',k'}^{p',u'} + \sum_{k\in K} \sum_{l\in L} \sum_{p\in P_;} \sum_{n_p=1}^{N_p} 1_{e'_{p,n_p},\xi_{k',e'_{p,n_p}}}(e_{p,n_p}, \xi_{k,e_{p,n_p}}) \\
&\qquad \cdot y_{l,k}^{p,u} \frac{1}{v_{p,n_p,k}\Delta t} \left(\eta + \frac{\partial h^u(v_{p,n_p,k})}{\partial \widehat{w}_{p,n_p,k}} d_{p,n_p}\right) B_{p,n_p,k} \hat{z}_{p,n_p,k}^{\beta_{e_{p,n_p}}-1}
\end{aligned} \tag{28}$$

Since the first and second terms are decomposable with respect to the links, the marginal costs of the paths belonging to the same segment will be placed in equilibrium by running a dynamic assignment algorithm. Then the marginal cost for a segment $MC_{l',k'}^{u'}$ is approximated by its marginal cost of path $MCP_{l',k'}^{p',u'}$. The calculation of the marginal cost of a segment requires the knowledge of the propagation of other segments $1_{e'_{p,n_p},\xi_{k',e'_{p,n_p}}}(e_{p,n_p}, \xi_{k,e_{p,n_p}})$, the basic traffic network status ($\widehat{w}_{p,n_p,k}$, $\hat{z}_{p,n_p,k}$, $v_{p,n_p,k}$, $h^u\left(v_{p,n_p,k}\right)$), as well as the aggregated segment-level information ($c_{l',k'}^{p',u'}$, $t_{l',k'}^{p',u'}$, $y_{l,k}^{p,u}$) from the simulator. With the marginal cost of each segment updated, route collections are updated by checking whether there are new lower marginal cost routes. Then the route flow vector X is updated to move along the descent direction with the step size described in the previous subsection based on the knowledge of the updated marginal cost. The algorithm stops when no more improvement on the total cost can be gained.

2.3.2 Emission models

The emissions generated as a result of assignment and routing procedure are estimated using the EPA model MOVES and include HC, CO, NOX, CO_2, $PM25$ [49]. MOVES is developed by EPA. It generates emission rates and emission inventories for both on-road motor vehicles and non-road equipment based on historical data of EPA. To use MOVES, the user specifies vehicle types, time periods, pollutants to observe, vehicle operating characteristics such as speed and acceleration, and road types as the inputs of the model. The model provides estimates of total emissions or emission rates per vehicle or unit of activity under different operation modes, such as operating, starting or idling.

2.4 Numerical results

This section presents the evaluation of the proposed CENCOS system using a hybrid digital twin network which covers LAX to LA downtown area built using the microscopic traffic simulator Vissim and a macroscopic traffic simulator Visum. The road network covered is shown in Figures 2 and 3. Lane characteristics such as length, capacity, speed limit et al. are incorporated in the Vissim and Visum network. Aside from the static road network characteristics, details of dynamic characteristics such as traffic light and signal control and sensor information are configured in the Vissim microscopic network. Due to the size and number of shuttle buses comparing with other road users in the LAX terminal, their management affects the overall transportation network. The background traffic is expressed as the number of trips between nodes that are origins and destinations. The historical freeway traffic flow data are obtained from PeMS [46] and Google Maps [47]. The raw traffic data are processed (formatted/truncated/aggregated) to fit the format of the traffic simulator. The background traffic conditions used in the numerical experiments is measured from the raw traffic data during medium traffic congestion hour (12pm to 4pm). The service network nodes used for load balancing layer are composed of demand nodes as well as intersections of freeways and major arterial ways. The demand of shuttles is provided with specifications on origin, destination and time window. To include the electric shuttles in the fleet, we assume the penetration of electric shuttles in the whole fleet varies from 0 to 100 percentage. The length of each interval is 30 minutes.

2.4.1 On CENCOS performance

To measure the performance of congestion alleviation of the proposed method, we compare the average shuttle operation time and background traffic speed in the airport between the current practice and the proposed system. The operation time in the airport is the difference between a shuttle entering and exiting time. It can be interpreted as the sum of travel time and waiting time in the airport area. The results of average operation time and background traffic speed under medium and heavy background traffic conditions are shown in Table 1 to Table 3, below.

Table 1. Performance on operation time and background traffic speed in airport under light traffic condition

	Operation time	Travel time	Waiting time	Background traffic speed
current practice	12.0 min	4.2 min	7.8 min	15.9 mile/h
CENCOS	11.8 min	4.1 min	7.7 min	16.1 mile/h

Table 2. Performance on operation time and background traffic speed in airport under medium traffic condition

	Operation time	Travel time	Waiting time	Background traffic speed
current practice	18.2 min	8.0 min	10.2 min	8.2 mile/h
CENCOS	13.6 min	5.5 min	8.1 min	12.1 mile/h

Table 3. Performance on operation time and background traffic speed in airport under heavy traffic condition

	Operation time	Travel time	Waiting time	Background traffic speed
current practice	23.2min	12.2 min	11.0 min	5.4 mile/h
CENCOS	15.5 min	6.9 min	8.6 min	9.5 mile/h

Under medium traffic condition, shuttles following the scheduling and routing management from CENCOS system spend 4.6 minutes per shuttle less than the current practice. The background traffic speed in the airport area increases from 8.2 mile/h to 12.1 mile/h. Under heavy traffic condition, shuttles following the scheduling and routing management from CENCOS system spend 7.7 minutes per shuttle less than the current practice. The background traffic speed in the airport area increases from 5.4 mile/h to 9.5 mile/h. The operation time of a shuttle reflects the sum of travel time on the road link and waiting time at each stop node spent in the airport area. From the results, under medium traffic condition, each shuttle saves average 2.5 minutes travel time and 2.1 minutes waiting time using CENCOS system. Under heavy traffic condition, the savings on travel time and waiting time are 5.3 minutes and 2.4 minutes.

Based on the results of the comparison of operation time and background traffic speed in the LAX airport area between current practice and the CENCOS system, we make the following observations:

- The shuttles spend less waiting time and traveling time in the airport area when following the management from CENCOS system.
- The benefits generated by the CENCOS system increase as the traffic becomes more congested and in fact it reduces the level of congestion.

We then perform comparison experiments between shuttle management systems with and without optimally centralized coordination under various penetration of electric shuttles in the fleet. To show the benefits of applying scheduling and load balancing co-simulation optimization assignment, we compared the proposed system against a mixed shuttle assignment system without optimally scheduling and routing. The non-coordinated system assumes that for each demand and the shuttle type associated with it, given the cost of each route between the origin and destination, the shuttle follows the individual minimal cost route according to the required time window. In the comparison we show that the lack of optimally centralized coordination of scheduling and routing of the shuttles and the lack of exchange of information among different shuttle providers may lead to many shuttles using overlapping routes and stop points which causes congestion. In the case of optimally CENCOS system, the scheduling layer first sorts the shuttles for each stop node to avoid congestion in each node and then in the load balancing layer, the changes of traffic flow characteristics on a certain route as well as the reactions of background traffic is reflected in the marginal cost so that the shuttles assigned on this route may be shifted to another route with lower marginal cost. In this way, the total cost of the assignment of overall shuttle cost can be reduced by reducing overlapping in routes and stop points . The comparison is shown in Figure 7. From the comparison results, we can see the average savings by applying CENCOS assignment versus system without CENCOS is 30.3%.

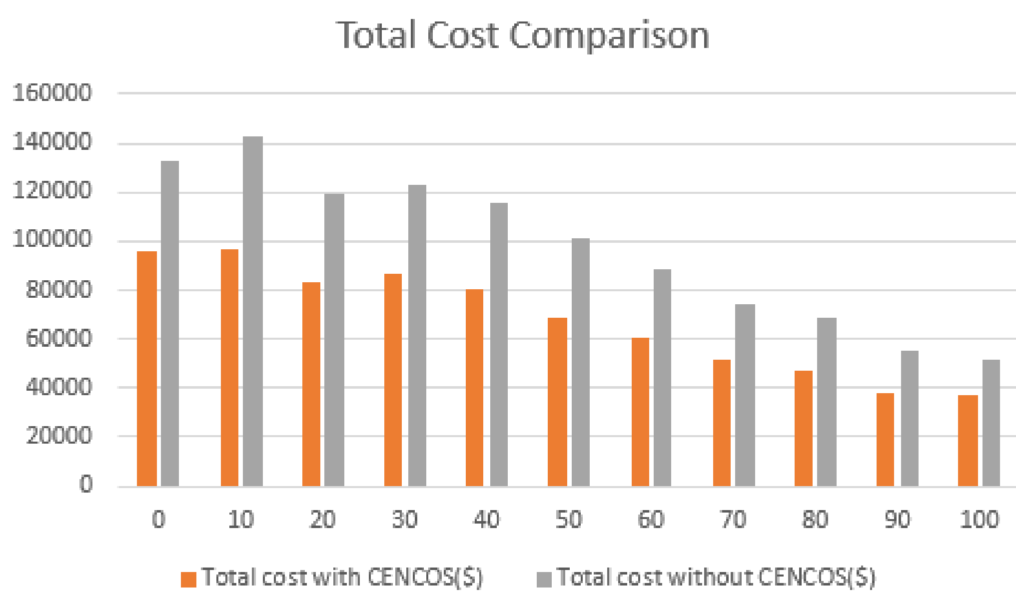

Figure 7. Comparison cases on system with and without CENCOS

We next test the system under various percentages of electric shuttle penetration. Under the predetermined background traffic condition, the percentage of electric vehicles in the whole shuttle fleet is varied from 0 % to 100 % in increments of 10 %. The results include total costs in US dollars of the assignment, the weight in unit of gram of several emissions (CO, NOX, CO2, PM25) as well as fuel consumed in unit of kg. The emissions are calculated by the modified EPA model MOVES [49] with speed as input and emissions in units of g/mile as output. The results are shown in Figure 8.

Figure 8. Results of total cost, fuel consumption and emissions

The above results lead to the following conclusions:

- The CENCOS system shows the savings on total cost of 30.3% versus traditional shuttle management system by optimally scheduling and routing shuttles with a co-simulation optimization method.
- The total cost of shuttle assignment from the CENCOS system decreases as the number of electric shuttles increases.
- The emissions go down drastically as the number of electric shuttles increases in the fleet.

2.4.2 On incorporated autonomous shuttles in the CENCOS system

The CENCOS system is compatible with the addition of autonomous shuttles. The autonomous shuttles can be incorporated as demand specified in origin, destination, time window and type of shuttle to serve. After the scheduling layer, the order of autonomous shuttles is temporally decided at each stop node. Then the autonomous shuttle demand is passed to load balancing layer together with the traffic load of non-autonomous shuttles. In the routing layer, since we assume that for safety autonomous shuttles follow a fixed route, the traffic load of autonomous shuttles is added in the simulator as another group with no capability of dynamic traffic assignment. In this way, the autonomous shuttles are fixed on running on a certain route and stops at stop nodes. Then the total cost of autonomous shuttles can be calculated through the traffic status provided by the traffic simulator.

3. Conclusions

In this project, we have proposed a centrally coordinated scheduling and routing system for shuttle management with LAX Terminals as application area. The dynamics and interactions with background traffic have been considered as well as inclusions of electric shuttles with their penetration varying from 0% to 100%. The electric shuttles have additional constraints that include limited range, longer refueling (charging) times and in addition the depletion rate of the battery life depends on traffic conditions. These characteristics introduce additional constraints that need to be taken into account in finding optimum schedules and routes that lead to shuttle load balance across the road network. The system is built on a two-layer framework, where the upper scheduling layer decides the order of shuttles entering and leaving each stop node and the bottom load balancing layer assign the shuttles across the road network aiming for minimum system cost. The integrated Gurobi Optimizer and co-simulation optimization method is used to solve the problem. A microscopic and macroscopic traffic simulation models are integrated in the CENCOS system to accurately predict the states of the transportation system as well as save computation time. A digital twin of the LAX to LA downtown traffic flow network is used to evaluate the system and the impact of electric shuttles in a mixed fleet. The CENCOS system shows the ability to alleviate traffic congestion in the airport area with respect to shuttle operation time and background traffic speed. The system shows 30.3% savings over one without the CENCOS system. Another result reveals that the use of electric shuttles can notably reduce the emissions and total cost without including the charging time. The application of CENCOS system provides a promising direction for the airport shuttle management that can be also utilized by other airports.

4. References

[1] H. Cai and M. Xu, "Greenhouse gas implications of fleet electrification based on big data-informed individual travel patterns," *Environ. Sci. Technol.*, vol. 47, no. 16, pp. 9035–9043, 2013.

[2] D. Y. Lee, A. Elgowainy, A. Kotz, R. Vijayagopal, and J. Marcinkoski, "Life-cycle implications of hydrogen fuel cell electric vehicle technology for medium- and heavy-duty trucks," *J. Power Sources*, vol. 393, pp. 217–229, 2018.

[3] M. Singer, "Consumer Views on Plug-in Electric Vehicles – National Benchmark Report," no. January, p. 38, 2016.

[4] A. Ceder, *Public transit planning and operation: Modeling, practice and behavior*. CRC press, 2016.

[5] A. Ceder, B. Golany, and O. Tal, "Creating bus timetables with maximal synchronization," *Transp. Res. Part A Policy Pract.*, vol. 35, no. 10, pp. 913–928, 2001.

[6] O. J. Ibarra-Rojas and Y. A. Rios-Solis, "Synchronization of bus timetabling," *Transp. Res. Part B Methodol.*, vol. 46, no. 5, pp. 599–614, 2012.

[7] O. J. Ibarra-Rojas, P. Fouilhoux, S. Kedad-Sidhoum, and Y. A. Rios-Solis, "Valid inequalities for the synchronization of bus timetabling," *Rep. Técnico PISIS-2012-02, Posgrado en Ing. Sist. Univ. Autónoma Nuevo León, México*, 2012.

[8] O. Ibarra-Rojas, F. López-Irarragorri, and Y. A. Rios-Solis, "Multiperiod synchronization bus timetabling. submited to Transportation Science,–," 2014.

[9] Y. Wu, J. Tang, Y. Yu, and Z. Pan, "A stochastic optimization model for transit network timetable design to mitigate the randomness of traveling time by adding slack time," *Transp. Res. Part C Emerg. Technol.*, vol. 52, pp. 15–31, 2015.

[10] Y. Wu, H. Yang, J. Tang, and Y. Yu, "Multi-objective re-synchronizing of bus timetable: Model, complexity and solution," *Transp. Res. Part C Emerg. Technol.*, vol. 67, pp. 149–168, 2016.

[11] L. Kang, J. Wu, H. Sun, X. Zhu, and Z. Gao, "A case study on the coordination of last trains for the Beijing subway network," *Transp. Res. Part B Methodol.*, vol. 72, pp. 112–127, 2015.

[12] Y. Wang, T. Tang, B. Ning, T. J. J. Van Den Boom, and B. De Schutter, "Passenger-demands-oriented train scheduling for an urban rail transit network," *Transp. Res. Part C Emerg. Technol.*, vol. 60, pp. 1–23, 2015.

[13] H.-Y. Shang, H.-J. Huang, and W.-X. Wu, "Bus timetabling considering passenger satisfaction: An empirical study in Beijing," *Comput. Ind. Eng.*, vol. 135, pp. 1155–1166, 2019.

[14] G. B. Dantzig and J. H. Ramser, "The truck dispatching problem," *Manage. Sci.*, vol. 6, no. 1, pp. 80–91, 1959.

[15] M. M. Flood, "The Traveling-Salesman Problem," *Oper. Res.*, vol. 4, no. 1, pp. 61–75, Jan. 1956.

[16] R. Baldacci, P. Toth, and D. Vigo, "Recent advances in vehicle routing exact algorithms," *4OR*, vol. 5, no. 4, pp. 269–298, 2007.

[17] C. Barnhart and G. Laporte, *Handbooks in operations research and management science: transportation*. Elsevier, 2006.

[18] G. Laporte, "What you should know about the vehicle routing problem," *Nav. Res. Logist.*, vol. 54, no. 8, pp. 811–819, 2007.

[19] G. Laporte, "Fifty years of vehicle routing," *Transp. Sci.*, vol. 43, no. 4, pp. 408–416, 2009.

[20] P. Toth and D. Vigo, *The vehicle routing problem*. SIAM, 2002.

[21] G. Laporte, F. Louveaux, and H. Mercure, "The vehicle routing problem with stochastic travel times," *Transp. Sci.*, vol. 26, no. 3, pp. 161–170, 1992.

[22] B. Verweij, S. Ahmed, A. J. Kleywegt, G. Nemhauser, and A. Shapiro, "The sample average approximation method applied to stochastic routing problems: a computational study," *Comput. Optim. Appl.*, vol. 24, no. 2–3, pp. 289–333, 2003.

[23] V. Pillac, M. Gendreau, C. Guéret, and A. L. Medaglia, "A review of dynamic vehicle routing problems," *Eur. J. Oper. Res.*, vol. 225, no. 1, pp. 1–11, 2013.

[24] P. Modesti and A. Sciomachen, "A utility measure for finding multiobjective shortest paths in urban multimodal transportation networks," *Eur. J. Oper. Res.*, vol. 111, no. 3, pp. 495–508, 1998.

[25] A. Lozano and G. Storchi, "Shortest viable path algorithm in multimodal networks," *Transp. Res. Part A Policy Pract.*, vol. 35, no. 3, pp. 225–241, 2001.

[26] K. F. S. Abdelghany, "Stochastic dynamic traffic assignment for intermodal transportation networks with consistent information supply strategies," 2001.

[27] M. Pickavet, P. Demeester, S. Demeyer, and P. Audenaert, "Dynamic and stochastic routing for multimodal transportation systems," *IET Intell. Transp. Syst.*, vol. 8, no. 2, pp. 112–123, 2014.

[28] T. Pajor, "Multi-modal route planning," *Univ. Karlsruhe*, 2009.

[29] L. Liu, "Data model and algorithms for multimodal route planning with transportation networks." Technische Universität München, 2011.

[30] D. Kirchler, "Efficient routing on multi-modal transportation networks." Ecole Polytechnique X, 2013.

[31] J. Guelat, M. Florian, and T. G. Crainic, "A multimode multiproduct network assignment model for strategic planning of freight flows," *Transp. Sci.*, vol. 24, no. 1, pp. 25–39, 1990.

[32] L. Castelli, R. Pesenti, and W. Ukovich, "Scheduling multimodal transportation systems," *Eur. J. Oper. Res.*, vol. 155, no. 3, pp. 603–615, 2004.

[33] H. Ham, T. J. Kim, and D. Boyce, "Implementation and estimation of a combined model of interregional, multimodal commodity shipments and transportation network flows," *Transp. Res. Part B Methodol.*, vol. 39, no. 1, pp. 65–79, 2005.

[34] K. G. Zografos and K. N. Androutsopoulos, "Algorithms for itinerary planning in multimodal transportation networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 1, pp. 175–184, 2008.

[35] L. Moccia, J. Cordeau, G. Laporte, S. Ropke, and M. P. Valentini, "Modeling and solving a multimodal transportation problem with flexible-time and scheduled services," *Networks*, vol. 57, no. 1, pp. 53–68, 2011.

[36] Y. Zhao, P. A. Ioannou, and M. M. Dessouky, "Dynamic Multimodal Freight Routing Using a Co-Simulation Optimization Approach," *IEEE Trans. Intell. Transp. Syst.*, 2018.

[37] H. Ambrose and M. Jaller, "ELECTRIFICATION OF DRAYAGE TRUCKS: ON TRACK FOR A SUSTAINABLE 1 FREIGHT PATH 2," 2016.

[38] N. Jiang, C. Xie, and S. Waller, "Path-constrained traffic assignment: model and algorithm," *Transp. Res. Rec. J. Transp. Res. Board*, no. 2283, pp. 25–33, 2012.

[39] S. Bahrami, H. Z. Aashtiani, M. Nourinejad, and M. J. Roorda, "A complementarity equilibrium model for electric vehicles with charging," *Int. J. Transp. Sci. Technol.*, vol. 6, no. 4, pp. 255–271, 2017.

[40] S. Faridimehr, S. Venkatachalam, and R. B. Chinnam, "A stochastic programming approach for electric vehicle charging network design," *IEEE Trans. Intell. Transp. Syst.*, no. 99, pp. 1–13, 2018.

[41] E. Yao, Z. Yang, Y. Song, and T. Zuo, "Comparison of electric vehicle's energy consumption factors for different road types," *Discret. Dyn. Nat. Soc.*, vol. 2013, 2013.

[42] A. Abadi, P. Ioannou, and M. M. Dessouky, "Multimodal Dynamic Freight Load Balancing," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 2, pp. 356–366, 2016.

[43] P. Chen, P. Ioannou, and M. Dessouky, "Mixed Freight Dynamic Routing Using a Co-Simulation Optimization Approach," *IEEE Trans. Intell. Transp. Syst.*, 2021.

[44] J. Wang and H. A. Rakha, "Fuel consumption model for heavy duty diesel trucks: Model development and testing," *Transp. Res. Part D Transp. Environ.*, vol. 55, pp. 127–141, 2017.

[45] K. N. Genikomsakis and G. Mitrentsis, "A computationally efficient simulation model for estimating energy consumption of electric vehicles in the context of route planning applications," *Transp. Res. Part D Transp. Environ.*, vol. 50, pp. 98–118, 2017.

[46] C. Chen, "Freeway performance measurement system (PeMS)," 2003.

[47] Google, "Google Maps." [Online]. Available: https://www.google.com/maps.

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

[49] Y. Zhang and P. A. Ioannou, “Environmental impact of combined variable speed limit and lane change control: a comparison of MOVES and CMEM model,” *IFAC-PapersOnLine*, vol. 49, no. 3, pp. 323–328, 2016.

Data Summary

Products of Research

The traffic flow data from Caltrans Performance Measurement System (PeMS) were collected for the study.

Data Format and Content

Data is in the format of zip file and includes following traffic information: timestamp, sensing station identifier, direction of travel, lane type, station length, total flow, average speed.

Data Access and Sharing

The general public can access the data through website https://pems.dot.ca.gov/.

Reuse and Redistribution

The data can be reused and redistributed by the general public through website https://pems.dot.ca.gov/.