NORTHWESTERN UNIVERSITY

Integrated Transit System Design with Autonomous Vehicle Fleet Services: Mathematical Formulation, Solution Approach and Large-Scale Application

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Civil and Environmental Engineering

By

Helen Karla Ramalho de Farias Pinto

EVANSTON, ILLINOIS

March 2021

ABSTRACT

Integrated Transit System Design with Autonomous Vehicle Fleet Services:

Mathematical Formulation, Solution Approach and Large-Scale Application

Helen Pinto

Providing quality transit service to travelers is a constant challenge for transit agencies. The advent of fully-autonomous vehicles (AVs) and their inclusion in mobility service fleets may allow transit agencies to offer better service or reduce their own capital and operational costs. This study focuses on the problem of allocating resources between transit patterns and operating shared-use AV mobility services (SAMSs) in a large metropolitan area. To address this question, a joint transit network redesign and SAMS fleet size determination problem (JTNR-SFSDP) is introduced, and a bi-level mathematical programming formulation and heuristic solution approach are presented.

The problem definition, modeling and solution of the JTNR-SFSDP are presented to recommend frequencies for existing transit lines and SAMS fleet size (level of service supplied), allowing the complete removal of existing lines. It demonstrates robustness of the proposed model with a sensitivity analysis at the design level, and shares takeaways from a large-scale application. It also proposes an agent-based model to capture travelers' mode choices and the system-level performance of the integrated transit-SAMS system in terms of travelers' simulated wait and travel times, transfers, shared rides in SAMS and denied transit boardings. This part is referred to as the dynamic combined mode choice and traveler assignment problem (DCMC-TAP). Results indicate the modeling framework can improve the travel experience of transit users in terms of average user waiting time as well as generalized travel costs, while under comparable operating subsidy constraints by reducing inefficient services and reallocating resources between transit and SAMS.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 Introduction

1.1 Motivation and problem description

Transit is often subsidized by local and federal government because it is essential for a healthy and happy urban population (Vuchic 1999). It plays a crucial role in social equity by providing affordable mobility to those who cannot afford other modes and those who cannot drive or use active modes (children, elderly, disabled and people without a driver's license). It also reduces congestion on the road network and associated gas emissions by allowing many riders inside a single vehicle. Congestion and emissions have direct impacts on stress levels and respiratory health of a whole population, as well as on the economy due to the wasted time on the road.

It is well known that transit services face disadvantages compared to car modes due to less comfort, privacy and security, and accessibility barriers to fixed transit stations (first/last-mile access problem). Transit agencies and urban planners must find ways to remain sustainable and competitive with the emergence of autonomous vehicle (AV) technology, and leverage the potential synergies from an integrated multimodal system composed of traditional fixed-route transit and on-demand autonomous mobility services. This will have to be done on the strategic and tactical planning side (transit network design and route frequency setting). However, the transit planning process is very complex and new planning tools will be needed to predict the supply-demand dynamics in such an integrated system.

Car manufacturers, ridesourcing and information technology companies are devoting substantial resources to the development of fully autonomous vehicles (AVs). These vehicles promise to transform the dynamics of the transportation sector in aspects ranging from operating cost structure and traveler convenience to environmental impacts and social welfare (Fagnant, Kockelman, and Bansal 2015). The innovative aspect of this technology has motivated research seeking to quantify these impacts and recommend operational models to guide efforts for its successful implementation at scale.

Similarly to how ridesourcing services (Uber, Lyft, Via, etc) have conveniently enabled flexible and affordable door-to-door rides at the push of a smartphone button, mobility services operated by AVs are expected to significantly impact the demand for transit in urban and suburban areas (Correia et al. 2018). They may pose competition to transit services depending on their fares, travel time and other attributes that reflect the user experience. Studies assessing impacts of ridesharing found the potential for both substituting and complementing transit on a case-by-case basis (Hall, Price, and Palsson 2017; Martin and Shaheen 2014).

Beyond competing with public transportation and sometimes even replacing it, AVs can also play a big role in making transit options more efficient and economical. Two examples may help in visualizing practical solutions that AVs bring to society. The first is elimination or reduction in service of bus lines that are not frequently used around the urban area, especially in lower-density residential areas. Replacing these bus lines with judiciously designed and efficiently provided AVs could result in benefits in terms of cost savings for agencies and better mobility service levels for individuals commuting in those areas. The second widespread potential benefit of AVs resides in the multimodal service that these vehicles offer to society in general, especially if planned as part of one seamless public service; hence, individuals now have the option of choosing between solely using AVs or, alternatively, combining the use of AVs with other modalities of transportation. This multimodal transportation service concept will be discussed in section 2.7.

Shared-use autonomous mobility service (SAMS) fleets may offer transit agencies a potential solution to the challenge of designing public transit networks to efficiently serve heterogeneous travelers subject to various constraints (e.g. budgetary, historical or political background, service policies, equity issues). In a cost-based analysis of future transport systems, Bösch et al. (2018) indicate that there is an opportunity for SAMSs to replace bus service in low-density areas and allow transit agencies to focus their resources on mass transit in dense urban areas. Agencies could replace inefficient transit routes/patterns operating in certain regions, and during certain times of the day, and reallocate those resources to operate (or subsidize) SAMSs. Given the considerable operational and capital cost advantages of SAMS over fixed-route transit services and driver-operated mobility services (e.g. flex-transit, ridesourcing, and taxi service), it is conceivable that reallocating resources from less cost-efficient transit patterns may produce better service for travelers and/or reduce overall transit agency costs.

At the time when this doctoral study started, research exploring the integration of transit and AVs was in its early stages and no previous studies had proposed an approach to redesign transit in an integrated system with AV-enabled mobility services, except for (Shen, Zhang, and Zhao 2018), which served as the initial reference for this work. Shen et al performed a supply-side simulation of first-mile shared AV services integrated with Singapore's transit system, identifying synergy opportunities between transit and SAMS. Similarly to Shen et al, I use an agent-based modeling approach that captures the system-level impacts of such integrated system of transit and

SAMS. However, my contributions go a step further by also providing an optimization-based approach to redesigning transit networks and operating (or subsidizing) SAMSs considering traveler behavior, the interactions between traveler agents in the system, and operational and budgetary constraints. The agent-based model used in this dissertation has three combined components: a mode choice model, a transit traveler assignment and simulation model, and a SAMS fleet assignment and simulation model. The mode choice model analyzes and predicts the choice of travel mode for all travelers' trips (private car, SAMS, transit, walking, etc), the assignment predicts their path choices and the simulation evaluates their experience as they move through their chosen paths.

To help achieve the Pareto-improving outcome of better service for travelers and lower costs for agencies, I present a modeling framework to optimize the joint design of transit networks and SAMS fleets. Specifically, the objective is to formulate and solve the joint transit network redesign and SAMS fleet size determination problem (JTNR-SFSDP) subject to user-equilibrium constraints at the mode and route choice levels. To illustrate the effectiveness of the modeling framework, I use traveler demand from Chicago's metropolitan region along with the region's existing multimodal transit network composed of urban and suburban buses, and urban and commuter rail.

Features of the modeling framework presented in this dissertation include (i) formulating a bi-level optimization framework; (ii) capturing congestion effects in the transit network; (ii) capturing four travel modes: walk, transit-only (bus and/or rail), SAMS-only and SAMS+Transit; (iii) considering spatial and temporal heterogeneity of demand in lower level, while upper-level takes a fixed demand supplied by the lower-level; and (iv) measuring mode choice response from changes in bus service. Additionally, the concept of route patterns used in the Transit Network Frequency Setting problem (TNFSP) by Verbas and Mahmassani (2013) is adopted to solve the frequency setting problem, modified to allow the removal of certain bus patterns. Route patterns are subsets of ordered stops for a certain route and dispatch time.

1.2 Objectives

The general goal is to develop a modeling framework to support the planning, design and analysis of integrated transit and shared-use AV-enabled mobility services (SAMS).

This will be accomplished with the following specific objectives:

(1) Propose a flexible model that captures the transportation supply and demand impacts of an on-demand SAMS by combining agent-based mode choice, and dynamic assignment-simulation models;

(2) Evaluate the potential impacts of SAMS on (a) the demand for transit and personal vehicle, and on (b) the transportation level of service provided to travelers;

(3) Develop an optimization model to jointly design transit networks and SAMS fleets considering the user response to design decisions as well as budgetary and operational constraints.

1.3 General Approach

The framework to support the joint design of integrated SAMS and transit systems is formulated as a two-level problem (called upper and lower levels) where the design part determines transit pattern frequencies and SAMS fleet size in the upper level, and the lower level evaluates the supply decision from the upper level through simulation of traveler experience and mode choice response. This evaluation is performed using (i) a multimodal agent-based time-dependent assignment-simulation tool called NU-Trans, (ii) a SAMS assignment-simulation tool, and (iii) a multinomial logit mode choice model. The traveler assignment-simulation and mode choice steps loop through an iterative process until user equilibrium state is reached in the lower level, meaning all travelers have chosen their best paths and modes from what is available. Another iterative process between upper and lower levels also reach an equilibrium point between supply and demand response.

1.4 Structure of this Dissertation

This chapter has discussed the potential impacts that autonomous vehicles are expected to have in passenger transportation and the importance of redesigning public transit systems so that they can better accommodate and leverage this emerging technology, improve the provided level of mobility service and continue to play their essential role in social equity. Additionally, the purpose of the study along with specific objectives and the general approach were presented.

Chapter 2 provides a literature review discussing earlier studies on transit network design, frequency setting and assignment, as well as research on the intersection of autonomous vehicles and transit systems. Later the bi-level optimization model and the concept of transit patterns used in this dissertation are introduced. Finally, I provide a description of the transit system in the Chicago metropolitan region (which is used as a testbed to illustrate application of the developed methodology) and discussion of the characteristics that motivated the theme of this dissertation.

Chapter 3 describes the proposed joint design of multimodal transit networks and shared autonomous mobility fleets model. The mathematical formulation of the problem considering budgetary and operational constraints as well as user equilibrium at the mode and route choice levels is described. A solution approach that builds upon the framework from the upcoming chapter is also presented. Results from a large-scale application in Chicago's metropolitan region illustrate robustness of the model.

Chapter 4 presents the solution framework for the dynamic combined mode choice and transit assignment-simulation problem to capture the impacts of a fleet of shared autonomous vehicles (SAMS). The mathematical formulation for the fixed-point problem of finding near-optimal user equilibrium modal flows is presented. To solve the problem heuristically, an agent-based modeling framework with three components is proposed: a mode choice model, a transit assignment-simulation model and a SAMS assignment-simulation model. Results from a large-scale application show that the solution approach provides a satisfactory solution that can be used to predict users' response to the supply of new mobility services as well as their overall experience in the urban transportation system.

Chapter 5 summarizes the presented work, discusses applications, limitations, future work and contributions.

2 Background Review

This chapter aims to discuss the scientific foundation upon which this dissertation has been built. I will cover concepts taken from transit network design and frequency setting, transit assignment, stochastic user equilibrium and combined transportation models. The stochastic user equilibrium is a concept that applies to the combined mode choice and traveler assignment in the lower level used to evaluate the joint design of SAMS fleet size and transit route frequencies.

An introduction to bi-level modeling will be given to better understand the solution approach used in the methodology. Later I discuss several previous works involving integration of emerging mobility services, including SAMS, and transit systems. Finally, a review of Chicago's transit system and its characteristics that motivated the development of this work is provided.

2.1 Transit Network Design and Frequency Setting Problems

Ceder and Wilson (1986) decompose the transit planning process in five parts: network design, frequency setting, timetabling, vehicle scheduling, and crew scheduling/rostering. All these subproblems are complex nonconvex NP-hard problems, hence very difficult to solve analytically. Ideally, they would be planned simultaneously but this is intractable in practice, so they are usually solved sequentially or as a combination of subproblems. The disadvantage of solving subproblems separately is that finding a global optimal solution is not guaranteed.

Ceder and Wilson (1986) present the transit network design problem (TNDP) and the transit network frequency setting problem (TNFSP) as interdependent public transit service subproblems. The TNDP is inherently a long-term strategic planning problem, wherein the network designer chooses transit lines and stops (i.e. transit routes) in a certain area. A significant volume of research exists on the TNDP (seminal work: Baaj and Mahmassani 1995; Ceder and Wilson 1986; Clarens and Hurdle 1975).

The TNDP instances are usually large-scale with multiple objectives, where the main user cost objective function is nonconvex with respect to the decision variables. Due to the nonconvex, multiobjective and combinatorial aspects of the problem, the literature adopts simplifying assumptions and proposes solutions reached through heuristic procedures. Objective functions include optimizing customer experience (i.e. area coverage, accessibility, travel time, trip directness, number of transfers, demand satisfaction, wait time, closeness to shortest path, etc), operator costs (number of routes, total route length, fleet size, hours of operation, etc), or both metrics simultaneously (Ceder and Israeli 1998; Guan, Yang, and Wirasinghe 2006). The input data typically include origin-destination (OD) demand, supply data, route performance indicators and area's topology. The output can be route changes, new routes, and operating strategies. Constraints to this problem involve policies based on local historical or political background.

The TNFSP is a medium-term tactical planning problem. The objectives are similar to those for transit network design aiming to optimize customer experience and/or operating costs. The output will be service frequencies. Its input data typically include available subsidy, service policies, current patronage, transit route network, OD demand, available fleet size and capacity. Constraints to this problem include demand satisfaction (avoiding overcrowding and large headways), headway bounds from regulating authorities, historical route runs, and number of route runs (Guihaire and Hao 2008).

As summarized by Iliopoulou, Kepaptsoglou, and Vlahogianni (2019), the complexity of the TNDP and TNFSP is explained in parts by different authors: (i) there is no simple solution procedure short of direct comparison of various local optima (Newel 1979); (ii) the discrete nature of the decision variables (Baaj and Mahmassani 1991, Chakroborty 2003); (iii) nonlinearity and the existence of logical conditions (Chakroborty 2003); (iv) NP-hard (Baaj and Mahmassani 1991, Kechagiopoulos and Beligiannis 2014); (v) different criteria for evaluation of solution quality based on contradicting targets (due to multiple objectives); (vi) cumbersome data collection due to the dynamic and time-dependent demand.

Due to TNDP convexity issues and problem size (see: Ceder and Wilson 1986), the current study modifies the TNFSP formulation of Verbas and Mahmassani (2015) to model the JTNR-SFSDP – further explained in section 3.3. The tactical aspect of the TNFSP also makes it more suitable to our application, where the focus is to change aspects of an already existing transit network to adapt it to an innovative environment of SAMS. Unlike most TNFSP formulations, the modified TNFSP in this study allows transit frequencies to be set to near-zero, allowing the removal of transit lines and improving the service of other lines with higher demand. For a detailed review of the TNDP, see reviews by Ibarra-Rojas et al. (2015) as well as Guihaire and Hao (2008).

Seminal work on the TNFSP includes analytical approaches that aimed to determine route attributes like spacing and length rather than actual routes (Newell 1971; Salzborn 1972) and heuristic approaches (Furth and Wilson 1981; Marguier and Ceder 1984; Schéele 1980). Early heuristics set the grounds for the development of metaheuristics but they could not solve large networks and could not provide accurate representations (Iliopoulou, Kepaptsoglou, and

Vlahogianni 2019). Metaheuristics came to solve hard combinatorial optimization problems. These are classified in two types based on whether it improves a single candidate solution (single-solution based) or selects from a pool of candidate solutions (population-based) (Gendreau and Potvin 2005). The problem solved in this dissertation improves an initial single solution (before scenario).

Recent research employs improved solution methods including problem decomposition, simulation, genetic algorithms, and optimization solvers such as KNITRO and CPLEX (Ibarra-Rojas et al., 2015). Chapter 2.6 outlines the common bi-level mathematical programming modeling framework typically employed to model the TNDP and the TNFSP (e.g. Constantin and Florian 1995; Fan and Machemehl 2011; Gao, Sun, and Shan 2004; Yoo, Kim, and Chon 2010; Yu, Yang, and Yao 2010).

In a more recent review of applications of metaheuristics for the TNDP/TNFSP, Iliopoulou, Kepaptsoglou, and Vlahogianni (2019) affirm that metaheuristics have been widely used to solve the TNDP/TNFSP since the 1990s. What makes them attractive for this purpose is the adaptability to different problem structures, the capability to represent complex problems efficiently and the computational performance. The authors identify an implementation framework containing common algorithmic components and different solution representations and methods across the literature.

2.2 Transit Service Patterns

A transit service pattern refers to a subset of ordered stops along a directed transit route. An example is the "short-turn" pattern, described by Furth and Day as a service pattern entirely overlapped by a longer service pattern on the same route, where the shorter pattern is a "short-turn" variation of the full-length pattern (Furth and Day 1985). Figure 2.1, from Verbas and Mahmassani (2013), depicts an example of a bus route with three different service patterns running on the same time interval. The dots represent bus stops served. The black dots are all the stops that belong to the bus route. The red pattern is the full-length pattern, and the blue and green patterns are "short-turn" variations. Short-turn patterns are used in the operation of transit services to better address variations in demand at certain times of the day and certain points of the route.

Given that transit service patterns run at various times of the day and can overlap other patterns, the transit unit of analysis used in this study is the pairing of transit service pattern and time interval of the day. Across this study, a "pattern" $p \in P$ refers to the transit service pattern associated with a specific dispatch time interval (d_p). This notation allows the solution algorithm to set the frequency of each transit service pattern at each time interval of the day (or study horizon). In this study, the time-intervals are 30 minutes. For example, if a transit service pattern runs from 6am to 7am, it is represented by two patterns: one for each 30-min interval. This way, we can set the headway for each interval independently (h_1, h_2). The variable q represents the frequency of each pattern, which is the interval length divided by the pattern's headway (h_p). For a traveler waiting at a certain stop, her perceived frequency of a bus route is the sum of the frequencies of the patterns that run at her departure time and that serve the stop she is waiting at and the stop she plans to get off at.

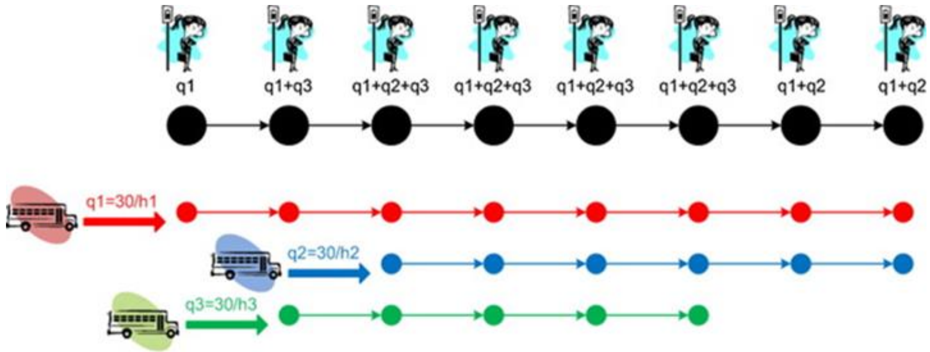

Figure 2.1: Illustration of transit service patterns along a transit route

2.3 Transit Assignment Problem

The formulation of the assignment of transit travelers to paths in the transit network has long been developed in the literature, starting with the straightforward search of a minimum cost transit path (Dial 1967) and later the adoption of the hyperpath concept formally defined by Nguyen and Pallottino (Nguyen and Pallottino 1988). Hyperpaths enabled a complete representation of the complexity of travelers' journeys by defining links for each segment of the path in the transit network graph: not only the transit lines but also segments of walking, transferring, waiting, etc, for all possible paths that connect the traveler's origin to their destination through the transportation network. This concept has been widely adopted until today in transit and traffic assignment models and requires modifying the original network representation into a graph of higher complexity to include the hyperpath links (Oliker and Bekhor 2018).

Spiess and Florian attempted to find more realistic optimal solutions in polynomial time and conceptually extended the model to incorporate some nonlinear congestion effects (Spiess and Florian 1989). Congestion becomes explicitly addressed in the works by de Cea and Fernandez and Wu et al. (de Cea and Fernandez 1993; J. H. Wu, Florian, and Marcotte 1994), where the user wait time at the transit station is penalized according to vehicle occupancy or passenger flow in the transit system. Cominetti and Correa model the passenger assignment in congested networks including the possibility of walking between transit stations through walking links (Cominetti and Correa 2001). Transit congestion effects are later modeled as hard capacity constraints (Hamdouch and Lawphongpanich 2008) and the availability of seated versus standing space (Hamdouch et al. 2011) in each individual vehicle, which can impede the user from boarding in a crowded vehicle. For Schmöcker et al., seat capacity is understood as a factor influencing the traveler route choice, although standing capacity is disregarded (Schmöcker et al. 2011).

Transit congestion effects must be included in realistic[1] models because they represent an additional burden to the traveler by either adding to the perceived travel time and/or by adding discomfort to the travel experience, hence affecting their route choice. Links associated with waiting, walking, transfers and standing time can have their inconvenience converted into additional perceived travel time through a multiplier, often referred as the value of waiting, walking, etc, typically up to 3 times the perceived travel time when passenger is seated. Because the final route choice depends on this perceived cost in the network, transit assignment models that

[1] Realism in the context of dynamic transportation assignment models refers to the need to adequately represent the complex human behavior and traffic dynamics, which can be further aggravated by time-dependency and randomness in system inputs (Peeta and Ziliaskopoulos 2001).

consider congestion effects require solving a user equilibrium model on the hyperpath space, often done iteratively.

The behavioral foundation of the user equilibrium is that each traveler seeks to minimize their travel costs. Cepeda et al characterizes the equilibrium in large-scale congested transit networks through the formulation of an optimization problem that seeks to minimize a gap function that represents the user equilibrium when it reaches its minimum value (Cepeda, Cominetti, and Florian 2006). The authors use the method of successive averages (MSA) to solve the problem. Travelers are assumed to make route choices based on least cost hyperpaths.

Transit assignment models can be frequency-based or schedule-based. The frequency-based models use aggregated representation of transit service for strategic planning problems and they are more suitable to handle large-scale networks; schedule-based models are helpful in the operations planning, given that it is based on the trajectories of the vehicles, allowing for a disaggregate understanding of passenger flows (Liu and Ceder 2017). Verbas and Mahmassani (2015) propose a time-dependent and frequency-based least-cost hyperpath algorithm that considers congestion effects to be used in transit assignment-simulation. In their study, the experience of travelers and movement of people and vehicles is captured through a multiagent particle simulation, where transfers, vehicle capacity through hard and soft constraints, seating and standing space availability and boarding rejections are reflected in the travel cost. This framework, later integrated with a mode choice model (O. Verbas et al. 2016), is adapted in this dissertation to include a new multinomial logit mode choice model and the presence of SAMSs.

Verbas et al. (2016) have shown, advancing on work by Sbayti, Lu, and Mahmassani (2007) and Lu, Mahmassani, and Zhou (2009), that gap-based dynamic transit assignment simulations are able to reach convergence much faster than the common method of successive averages (MSA). This is because the gap-based approach shifts travelers from their current best path to another path based on the cost gap between the two paths, whereas the MSA treats paths equally. This convergence method is applied in a dynamic traveler assignment and simulation framework for multimodal large-scale transit networks (including bus, rail, walking and biking), considering different service patterns of transit routes. This gap-based approach is used in my work.

2.4 Stochastic User Equilibrium

The stochastic user equilibrium (SUE) proposed by Daganzo and Sheffi (1977) and Fisk (1980) became a widely used behavior model for traveler path assignment. Sheffi (1985) defines the SUE as the state in which no traveler believes that their perceived travel cost can be improved unilaterally by changing paths. This comes from the understanding that travelers do not have full knowledge of the network, so they choose their best path alternative based on their perceived travel costs rather than the true travel cost. In addition to the variability and incompleteness of the available information to the user, in contrast to deterministic models, stochastic models allow users to choose paths according to their different preferences or perceptions of the information, effectively overcoming the assumption of homogeneous users in Wardrop’s equilibrium. Wardrop’s first principle states that, at equilibrium, the cost of all used routes is equal and not higher than those of unused routes.

Besides the assumption of heterogeneous users, SUE also includes the random effect of the stochastic assignment problem on a congested network (Lam et al. 1999). It reflects the variability of travel times along links, represented through a distribution of perceived costs (by travelers); various forms have been assumed for this distribution in the literature. A common form is the normal distribution with mean equal to the (measured) average travel cost and variance proportional to the mean. This yields a probabilistic equilibrated assignment.

Gentile (2018) provides a recent discussion of SUE models, warning that "trying to match real flow data with a deterministic assignment model can lead to fictitious increases of demand and/or relevant distortions in the model, so as to generate congestion on some links while letting other links become comparably attractive and then used".

2.5 Combined Mode Choice and Assignment

Combined models found in the transportation literature are those that propose formulations that aim to solve a combination of steps of the transportation planning process. The latter is traditionally divided in four steps: trip generation, trip distribution, modal split and assignment. Wong et al. (2004) provides a review of early studies proposing combined models involving mode choice and assignment. These early models adapt Wardrop's first principle to the mode choice where the equilibrium is reached when no traveler can improve their travel cost by switching modes. As Verbas et al. (2016) properly summarizes, major shortcomings of these early models were their deterministic assignment of passenger to modes, their static (aka independent of time, not dynamic) nature, and the lack of behavioral realism in the transit assignment (user heterogeneity, incomplete information, congestion effects, overlapping links, etc). This is why my work, based on Verbas et al. (2016), seeks to overcome such shortcomings by using stochastic assignment of passengers to modes through a multinomial logit mode choice model, and by using a time-dependent simulation-based assignment tool that captures congestion effects.

Wong et al. (2004) propose an optimization model that not only combines a hierarchical mode choice and assignment model but also solves the trip distribution stage. Abdelghany and Mahmassani (2001) present a dynamic trip assignment (DTA) model for urban intermodal transportation networks. It captures the dynamic interactions between mode choice and traffic assignment and estimates the effect of this interaction on overall network performance. The model implements a multiobjective dynamic trip assignment procedure in which travelers choose their mode route based a multicriteria least cost path formulation. Their experiments illustrate the significance of including mode choice dimension in the DTA framework and show the importance of a multiobjective assignment procedure incorporated in the model. Zhou, Mahmassani, and Zhang (2008) perform an agent-based simulation that captures traveler choices of route, departure time and mode, where departure time and mode are determined through a logit model. Zhang, Mahmassani, and Vovsha (2011) integrated a nested logit mode choice model and a DTA model.

In a more recent study done after the publication of my work, Kamel et al. (2019) propose an integrated model for travel mode, departure time and route choice for large multimodal transportation networks with five modes. Departure time and mode choices are modeled through a nested logit structure, and route assignment through a simulation-based dynamic traffic and transit assignment tool. The modeling framework is similar to the one used in this doctoral research in that it uses an iterative process to capture the effects of the supply-demand decisions over one another; the main loop ends with the convergence of the estimated demand. Within each iteration the route assignment estimates the level of service attributes based on the demand and other fixed inputs on one level, whereas the nested logit model estimates the demand based on the simulated level of service.

2.6 Bi-Level Modeling Framework

This study models the JTNR-SFSDP, described in section 3.3, as a bi-level mathematical program. The generic formulation of a bi-level mathematical program is presented in Eqns. (2.1)-(2.2).

Upper Problem: $F[x, y]\ ;\ G[x, y] \leq 0$ (2.1)

Lower Problem: $f[x, y]\ ;\ g[x, y] \leq 0$ (2.2)

F[·]: objective function of the upper-level decision maker(s)

x: decision vector for the upper-level decision maker(s)

G[·]: constraint set of the upper-level decision vector

f[·]: objective function of the lower-level decision makers

y: decision vector for the lower-level decision makers

g[·]: Constraint set of the lower-level decision vector

$y = y(x)$ is typically referred to as the reaction or response function.

If a response function can be found, the variable y in the upper-level problem can be replaced with the relationship between y and x in the response function (Sun, Gao, and Wu 2008). In this study, an analytical relationship between the upper-level decision variables and the lower-level decision variables does not exist because of the complex, non-convex, nonlinear, and discrete nature of the lower-level decision problem. Hence, this study employs a heuristic approach to solve the bi-level problem, rather than an exact analytical solution method. The conceptual framework of this heuristic approach is depicted in section 3.1.

2.7 Integrating Shared AV Mobility Services and Public Transit Systems

It has been argued that autonomous vehicles will promote sustainability and transit efficiency through fleet fuel economy, marginal cost pricing, raising vehicle occupancies through dynamic ridesharing, lower energy consumption and reduced carbon footprint (Fagnant and Kockelman 2014). Several studies have investigated the applicability and demand of integrating transit services with SAMS fleets serving as first/last mile feeders.

Although this integration may be the best option from the standpoint of optimization and traffic flow improvement, the demand for it may not realize as expected. SAMS may compete with transit rather than complement it, as they offer convenient user experience with competitive rates. This will depend on the target population group. Certain population groups, such as the elderly and disabled, may not like to transfer between modes, whereas captive transit riders, price-sensitive and technology-oriented population groups are expected to welcome and experience the greatest benefits of an integrated system. Nevertheless, with proper policy incentives and optimal designs that are demand-responsive, the integrated SAMS and transit service may prove to be feasible, as discussed in the following studies.

Greenblatt and Shaheen provide a review on the history and future trends of AVS and mobility on demand, such as ridesourcing (Greenblatt and Shaheen 2015). They argue that ridesourcing users are replacing taxi and bus trips. A notable portion of them have the origin or destination of the trip at a public transit station, which suggests ridesourcing and SAMSs as a potential transit feeder mode. In comparison to taxis, ridesourcing presents advantages in “hail and dispatch” wait time besides often having greater passenger loads due to demand pair-matching technology. According to Yan et al., ridesourcing services provide an effective solution to last-mile problem by filling service gaps of large-volume transit lines and extending the catchment area of transit (Yan, Levine, and Zhao 2019).

Other recent studies explore the integration of emerging shared and on-demand mobility services (e.g. ridesharing, bikesharing, SAVs, etc.) within transit networks. At the operational-level, Stiglic et al. investigate the potential benefits of integrating ridesharing and public transit systems using an optimization model (Stiglic et al. 2018). The authors considered several ride matching options, such as ridesharing only, ridesharing and dropping off passengers at transit stops, and a park-and-ride option (in which the driver uses the transit line after dropping off passengers). The optimization algorithm maximizes the number of matched riders, and minimizes the total increase in driving distance for all drivers. To test the benefits of the integrated service, the authors conducted an extensive computational study on a simulated network that has a stylized transit network similar to that of the Bay Area Rapid Transit. The optimized integrated service directed more drivers to transit and provided enhanced mobility and a sustainable system, by reducing the negative externalities of private automobile travel. Liang et al. model and optimize a first- and last-mile electric automated taxi feeder mode to transit (Liang, Correia, and van Arem 2016). Using a continuous approximation model, Wu et al. model the optimal design of a network with bikesharing stations and transit lines (L. Wu, Gu, and Fan 2018). Previous research examines the integration of demand-responsive and traditional fixed-route transit services (Hickman and Blume 2001).

According to Meyer et al. (2017), AVs will substantially increase accessibility and compete with transit. Based on a simulation in Swiss municipalities, they affirm that, from a capacity standpoint, AV fleets will generally be able to serve the full transport demand, including car and transit. They continue explaining that transit will only still be required in the centers of large agglomerations, where highest transport demands meet limited road capacities (Meyer et al. 2017). Shen et al. (2018) employ an agent-based supply-side simulation model to test different bus network and SAV integration scenarios. The results indicate significant potential benefits of replacing low-utilization bus routes with SAVs. The authors define a priori the bus routes that will be replaced by SAMSs as well as the modal share for both transit and SAMSs (Shen, Zhang, and Zhao 2018).

Yan et al. use a joint revealed and stated preferences model to evaluate response to an integrated system of ridesourcing services and public transit (Yan, Levine, and Zhao 2019). Revealed preference choices included driving, walking, transit, or biking. Stated preference choices presented new alternative system designs that reflected an integrated service including variables for number of transfers and wait times. They predict that using existing bus stops as drop-off locations and substituting ridesourcing services for low-demand bus routes would slightly boost transit use while reducing operational costs. The increase in transit use is largely due to reduced wait times and in-vehicle travel times. Feigon and Murphy (2016) showed that individuals who already use shared-ride services are more likely to use transit, as well as that shared vehicles have a great potential to complement public transit (Transportation Research Board and National Academies of Sciences 2016).

Lu et al. analyzed social preferences and demand for autonomous vehicles and transit-oriented developments (TOD) based on a survey conducted in Atlanta (Z. Lu et al. 2017). The survey focused on future technologies that affects housing, living, and transportation decisions of individuals. Using sentiment and latent-class analysis, the authors found that most Atlanta residents prefer to live in TODs rather than in an automobile-dependent environment. They also found public support for future innovations in Georgia, as well as positive attitudes towards the prospect of improving traffic congestion with the use of AVs. From these results, the authors predict a future with integrated AV and transit services, but also argue that this may be the case because Atlanta currently needs a stronger public transportation system. Hence, they suggest that policy makers need to consider local conditions before the introduction of integrated services. Here, I would add that this does not mean that the benefits from an integrated SAMS and transit service would only be significantly felt in urbans areas with weak public transportation systems. In Chicago, where there is a much stronger transit system, there are still several accessibility gaps that can be filled by AVs (refer to section 2.8 for more details).

The modeling framework presented in this study for the JTNR-SFSDP involves an integration of transportation supply and demand models. Chapter 4 (Pinto et al. 2018) introduces the DCMC-

TAP modeling framework that is employed in this study to model the lower-level problem in the bi-level JTNR-SFSDP. Similarly, Verbas et al. (2016) and Zhang et al. (2011) introduce dynamic combined mode choice-transit assignment and dynamic combined mode choice-traffic assignment models, respectively. Verbas et al. (2015) introduce the dynamic transit assignment-simulation model employed in this dissertation to obtain performance metrics for the transit mode based on transit demand and transit route/pattern frequency inputs. Hyland and Mahmassani (2018) present the operational strategies and simulation model for the on-demand shared-ride SAMS (Zhang, Mahmassani, and Vovsha 2011; O. Verbas et al. 2016; İ. Ö. Verbas, Mahmassani, and Hyland 2015; Hyland and Mahmassani 2018).

2.8 Chicago's Transit System

This review section addresses the quality of transit services in Chicago metro region and indicates that there is room for improvement. This may be facilitated with the emergence of AVs to fill accessibility gaps and/or to allow reallocation of public resources to areas that most need them. As a particular example, there is an opportunity for AVs to supply targeted services that could fill reverse commute needs and reduce the spatial mismatch that exists between the unemployed labor force and their job opportunities, effectively helping develop the local economy as well. Focus is given to the metropolitan region of Chicago because it is also the transit network chosen for our case study and large-scale application of our framework.

2.8.1 Chicago Transit and its Accessibility

Chicago is known for its variety and wide coverage of transit services. In the metropolitan area, the Regional Transportation Authority has divided responsibilities for administration and operation of transit in the region between the Chicago Transit Authority (CTA) and suburban operators. CTA is in charge of rail and bus services in the city of Chicago as well as some connections to suburban Cook County, whereas the suburban operators are Metra, in charge of commuter rail, and Pace, in charge of fixed-route bus and paratransit services. Despite its extensive transit network, several areas lack adequate accessibility to public transportation.

Studies show that last-mile access barriers are most prevalent in the suburbs (low density and scarce transit coverage) but also correlated to socio-demographic factors like poverty, unemployment, perception of crime and walkability of sidewalks. In Greater Chicago, most residential areas with job access within 30 min are concentrated on the north side (Owen and Levinson 2014), and areas most affected by last-mile access problems are suburbs and certain south and west neighborhoods (Tilahun and Li 2015). While Chicago ranks third in total employment among US cities, it is fifth in job accessibility.

2.8.2 Spatial Mismatch and Reverse Commute in Chicago

An analysis of the transit quality of service and employment accessibility (Minocha et al. 2008) examined Chicago's region by comparing supply and demand indicators of transit: transit availability index (TAI) and transit employment accessibility index (TEAI). The TAI was based on transit frequency, hours of service and coverage, and TEAI represented the potential of residents in certain traffic analysis zones (TAZ) to access jobs in other TAZs in the Metropolitan area. Their results indicate most south Chicago neighborhoods to have high TAI and the lowest TEAI values from the studied areas, implying job remoteness in these areas and a mismatch between where residents and their potential jobs are located.

Besides the Loop (CBD), four out of the top five largest employment centers in the region are located along the I-90 corridor (172,000 jobs) and in Lombard (32,000 jobs), Naperville (35,000 jobs), and Oak Brook (33,000 jobs) (Urban and Smith, 2012), in the social and economically diverse collar counties. Such spatial mismatch continually worsens with the lack of efficient transit within and between Chicago and these key suburban job centers (Keil and Addie 2015).

Job availability in the suburbs (Cook and collar counties) in industries like retail and manufacturing as well as suburban institutions and corporate campuses, opposed to the disappearance of local low-wage manual labor jobs, implies the need for reverse commute (from the city to the suburbs) for many residents. These people would certainly benefit from investments on ways to improve the user experience for reverse commutes and potentially offering of direct transportation services to far west areas like Naperville (in DuPage County), reducing travel times to work that currently take more than 90 minutes (with connections at the Loop to Metra's BNSF Railway).

Early surveys by Pace of two reverse commute runs from south Chicago to job centers in DuPage County revealed that the services influenced the decision of over 60 percent of surveyed passengers to take and retain the jobs (Cervero, 1994), indicating that providing affordable direct transportation to areas outside the CBD is a successful way of enhancing the potential for job growth. With this behavioral finding and given the current demand for labor in suburban areas around Chicago, this is expected to remain true nowadays. This means that such an affordable direct transportation can either be provided through a fixed-route service, such as the PACE suburban bus, or it may be a flexible on-demand service performed by SAMS. The answer to which of these alternatives would be more beneficial depends on the existing demand, the latent demand as well as the operational characteristics of the supplied service. The methodology to answer such question is one of the contributions of this doctoral research.

2.9 Summary of the Background Review

This review explained that the transit network design and frequency setting problems can be both formulated separately or together. The formulation used in this dissertation is a modification of the frequency setting problem that allows the removal of transit patterns. Hence, it is a combined transit design and frequency setting formulation. The use of the concept of transit patterns (subsets of route stops) gives flexibility to the model because it allows targeting specific stops with high/low demand where the level of service needs to be adjusted, instead of adjusting all stops in a single route equally.

Later the traveler assignment problem is described as the process that predicts travelers’ path choices in the network, where each traveler chooses their respective minimum-cost hyperpath, considering congestion costs in the transit network. Costs considered in the assignment are frequency-based, time-dependent and simulation-based. Furthermore, the transit network is seen separately from the SAMS network. The simulation-based traveler assignment problem is solved by minimizing a gap-based function that represents the difference between the best estimated user travel cost and the experienced user travel cost after congestion effects. A multinomial logit mode choice model predicts travelers' stochastic choices between the transit and SAMS networks, or a combination of both. A discussion of previous works that combine mode choice and assignment models is given. Later the concept of stochastic user equilibrium is explained for better understanding of the transit assignment solution, which considers users' heterogeneity and randomness. The importance of stochastic and time-dependent models is highlighted to obtain realistic solutions.

The generic formulation of a bi-level mathematical program is presented to represent the joint design of SAMS fleet size and transit frequencies subject to the users' mode choice and path assignment equilibrium. I explain that in our case there is no analytical relationship between the decision variables in the upper and lower levels. Recognizing this, I developed a heuristic approach to solve it, which is a contribution explained in Chapter 5.2.

Previous studies integrating on-demand mobility services and transit systems are discussed. I highlight potential benefits that autonomous vehicles are expected to provide, especially as a shared service and judiciously integrated with public transit. Previous surveys also predict higher user acceptance of AV technology by specific population groups. A gap in this literature that is filled by my work is the scarcity of supply-demand interaction models applied to the design of transit systems with AVs. Further details of the contributions can be found in the chapter.

Finally, as the chosen location for case study, Chicago's transit system is shown to have accessibility gaps that could be potentially solved through targeted implementation of shared AV-enabled mobility services.

3 Joint Design of Multimodal Transit Networks and Shared Autonomous Mobility Fleets

This chapter presents the joint transit network redesign and SAMS fleet size determination problem (JTNR-SFSDP) subject to user equilibrium at the mode and route choice level.

The mathematical formulation of the JTNR-SFSDP is based on a transit network frequency setting problem formulation (TNFSP), in which the set of transit service patterns is fixed. Two major changes to the base TNFSP formulation in the JTNR-SFSDP formulation are the inclusion of a SAMS fleet size decision variable as well as the removal of maximum headway (i.e. minimum frequency) constraints for bus services. Removing the maximum headway constraint allows transit patterns to be effectively removed.

3.1 Conceptual Framework

The general framework and heuristic solution approach to the bilevel model is depicted in Figure 3.1 to facilitate the understanding of the mathematical model for the JTNR-SFSDP.

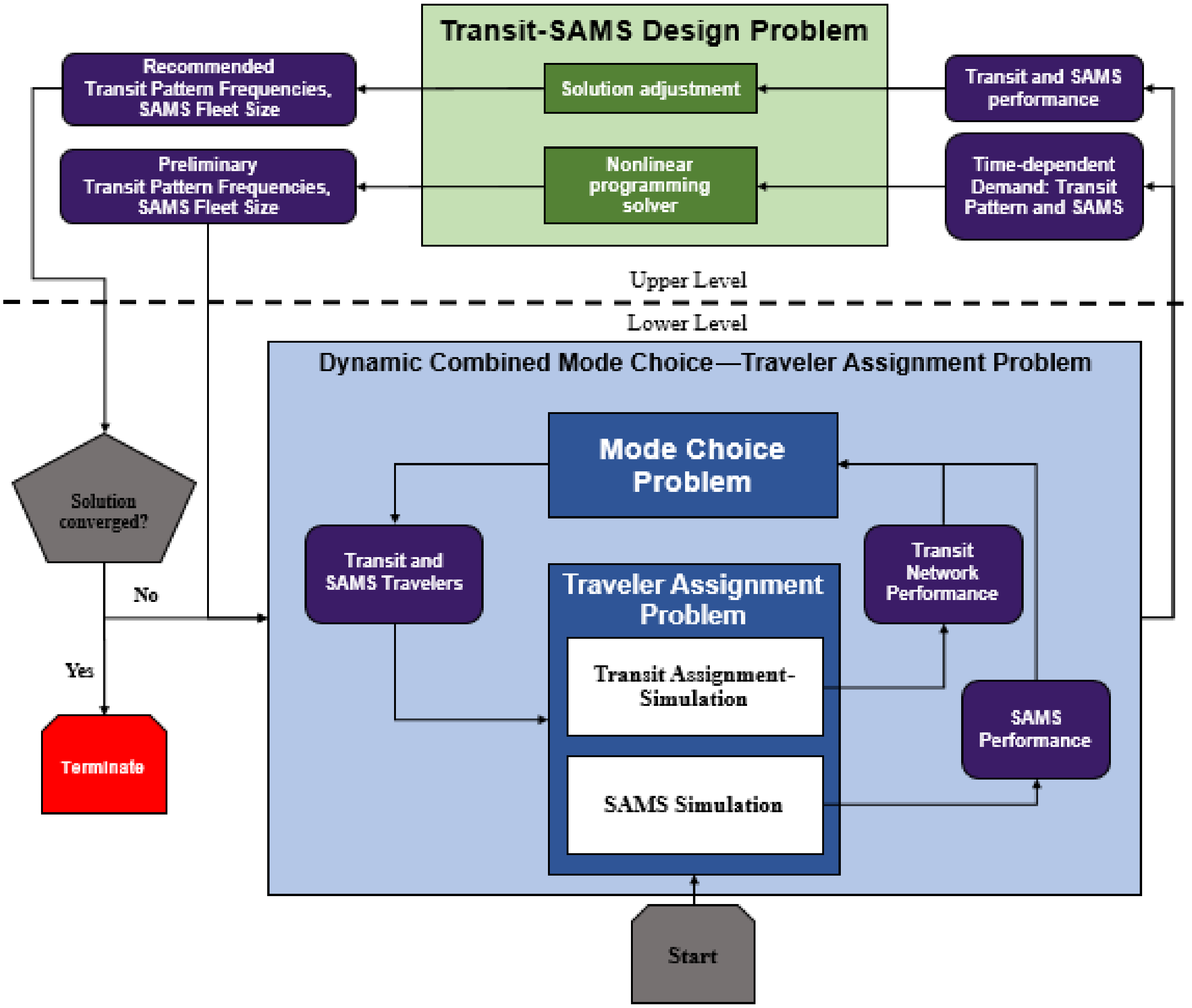

Figure 3.1: Framework to solve the joint transit network redesign and SAMS fleet size determination problem

3.2 Nomenclature

Sets and Indices

P set of transit patterns; indexed by $p \in P$
P_b subset of bus transit patterns; $P_b \subseteq P$
P_r subset of rail transit patterns; $P_r \subseteq P$
T set of time intervals; indexed by $t \in T$
J set of transit vehicle types; indexed by $j \in J$
M set of modes, indexed by m∈M
M^t set of transit modes; indexed by $m \in M^t = \{bus, rail\}$
Z set of transit vehicle trips; indexed by $z \in Z$
Z_p subset of transit vehicle trips belonging to pattern p, $Z_p \subset Z$
O set of origin microanalysis zones (MAZs), indexed by $o \in O$
D set of destination MAZs, indexed by $d \in D$
T^a set of assignment time intervals, indexed by $t^a \in T^a$
Q set of paths, indexed by $q \in Q$

Parameters

l_p trip duration of a transit vehicle on pattern $p \in P$
d_p dispatch time interval of pattern $p \in P$; $d_p \in T$
θ_p vehicle type used on pattern $p \in P$
B_p capacity of vehicle used on pattern $p \in P$
F_p transit fare
F^{SAMS} SAMS fare
τ'_p trip duration of a transit vehicle on pattern $p \in P$, bounded above by time interval length ($\tau_p = min\left(l_p, 30\right)$
Γ operating budget/subsidy
c_1 transit operating cost per hour
c_2 SAMS average operating cost per vehicle in the planning horizon
w_o minimum average traveler wait time for SAMS
Δ_1, Δ_2 slopes in piecewise linear function ($\Delta_2 > \Delta_1$)
a_1, a_2 cut-off points in the piecewise linear equation
r^{AV} average service rate of AVs in the SAMS fleet (passengers/vehicle-hour)
V_j^o original number of transit vehicles of type $j \in J$
S^o original number of AVs in the SAMS fleet
k_j equivalent number of AVs associated with a transit vehicle of type $j \in J$
h_m^- minimum headway for transit mode $m \in M^t$
h_m^+ maximum headway for transit mode $m \in M^t$
γ_p^L coefficient converting transit overcrowding [persons] into units of time
γ_t^R coefficient converting SAV overcrowding [persons] into units of time

μ coefficient that allows change in maximum vehicle resources (measured in equivalent number of AVs)

Input from Lower-Level Problem

e_t^{SAMS} demand for SAMS during time interval $t \in T$

e_p demand for pattern $p \in P$

g_p maximum flow on pattern $p \in P$

Upper-Level Problem Decision Variables

h_p headway of pattern $p \in P$

S SAMS fleet size

v_j number of transit vehicles of type $j \in J$

Upper-Level Problem Auxiliary Variables

u_t estimated average SAMS traveler wait time during time interval $t \in T$

ρ_t estimated utilization rate of SAMSs during time interval $t \in T$

L_p boarding rejections due to transit vehicle crowdedness (transit rejection penalty term)

R_t^{SAMS} boarding rejections due to congestion in SAMS system (SAMS rejection penalty term)

Input from Upper-Level

$R_{o,d}^{t^a}$ list of travelers with origin $o \in O$, destination $d \in D$, and departure time $t^a \in T^a$

Lower-Level Problem Decision Variables

d_p dispatch time interval of pattern $p \in P; d_p \in T$

$R_{o,d}^{t^a,m,q}$ list of travelers with origin $o \in O$, destination $d \in D$, and departure time $t^a \in T^a$ on mode $m \in M$, assigned to path $q \in Q$

Lower-Level Problem Endogenous Variables

$\Psi_{o,d}^{t^a,m}$ probability of a traveler with origin $o \in O$, destination $d \in D$, and departure time $t^a \in T^a$ choosing mode $m \in M$

$L_{o,d}^{t^a,m}$ least cost paths between origin $o \in O$ and destination $d \in D$ on mode $m \in M$ with departure time $t^a \in T^a$

3.3 Mathematical Formulation

The mathematical program for the JTNR-SFSDP is defined as follows:

$$min \sum_{p\in P} \frac{e_p h_p}{2} + \sum_{\substack{t\in T \\ p\in P|d_p=t}} (e_t^{SAMS}+L_p)\, u_t + \sum_{p\in P} \gamma_p^L L_p + \sum_{t\in T} \gamma_t^R\, R_t^{SAMS} \tag{3.1}$$

$$\rho_t = \frac{e_t^{SAMS}}{r^{AV}S} \qquad \forall t \in T \tag{3.2}$$

$$u_t \geq w_o \qquad \forall t \in T \tag{3.3}$$

$$u_t \geq w_o + \Delta_1(\rho_t - a_1) \qquad \forall t \in T \tag{3.4}$$

$$u_t \geq w_o + \Delta_1(a_2 - a_1) + \Delta_2(\rho_t - a_2) \qquad \forall t \in T \tag{3.5}$$

$$c_1 \sum_p \frac{l_p}{h_p} + c_2 S - \sum_p (F_p e_p + F^{SAMS} e_t^{SAMS}) \leq \Gamma \tag{3.6}$$

$$\sum_{p|d_p=t,\theta_p=j} \frac{\tau_p'}{h_p} \leq v_j \qquad \forall\, j \in J, t \in T \tag{3.7}$$

$$\sum_{j\in J} k_j v_j + S \leq \mu \left(\sum_{j\in J} k_j V_j^o + S^o \right) \tag{3.8}$$

$$h_m^- \leq h_p \leq h_m^+ \qquad \forall p \in P_i, \quad m \in M^t \tag{3.9}$$

$$L_p \geq g_p h_p - B_p \qquad \forall\, p \in P \tag{3.10}$$

$$R_t^{SAMS} \geq\ e_t^{SAMS} - r^{AV} S \qquad \forall\, t \in T \tag{3.11}$$

$$\{e_t^{SAMS}, e_p, g_p\} = \varphi\left[DCMCTAP\left(h_p, S\right)\right] \tag{3.12}$$

$$S \geq 0 \tag{3.13}$$

$$v_j, L_p, R_t^{SAMS} \geq 0 \qquad \forall\, j \in J, p \in P, t \in T \tag{3.14}$$

The objective function in Eqn. (3.1) aims to minimize the disutility of travelers, particularly their wait time and boarding rejections. The first term represents the cumulative wait time of travelers assigned to transit. Assuming random arrival from uniform distribution, the average traveler wait time on pattern $p \in P$ is $\frac{h_p}{2}$. Multiplying $\frac{h_p}{2}$ by e_p, the demand for pattern $p \in P$, gives the cumulative wait time of travelers using pattern $p \in P$.

The second term in the objective function gives the cumulative wait time of SAMS travelers. The auxiliary decision variable u_t estimates the average SAMS traveler wait time during time interval $t \in T$ and it is directly related to the main SAMS decision variable – SAMS fleet size (S) – via the SAMS fleet utilization rate (ρ_t). The constraints in Eqn. (3.2)-(3.5) define the relationship between S and the auxiliary decision variables (ρ_t, μ_t). Multiplying the estimated wait time u_t by the SAMS demand gives the cumulative wait time of SAMS travelers with departure time t. Here it is assumed that transit travelers who are denied boarding (L_p) will become SAMS demand ($e_t^{SAMS} + L_p$) in the pattern time interval ($d_p = t$).

The third term in the objective function captures transit vehicle crowding. Eqn. (3.10) defines the transit overcrowding auxiliary variable (L_p) as a linear function of the headway (h_p) and the maximum flow of the pattern (g_p). L_p is the number of passengers exceeding the vehicle capacity B_p of each pattern, summed over all patterns. The parameter γ_p^L converts units of transit travelers with denied boarding into units of time. Similarly, the fourth term captures the disutility from the SAMS system crowding. In Eqn. (3.11), R_t^{SAMS} reflects the number of travelers exceeding the capacity of the SAMS system as a function of the SAMS fleet size (S). The parameter γ_t^R converts units of SAMS travelers with denied service in a 30-minute time interval into units of time. This objective function captures the pertinent planning-level performance metrics most impacted by the decision variables (S and h_p) for a joint transit-SAMS system.

Eqn. (3.2) defines the SAMS fleet utilization rate (ρ_t) as the ratio of the SAMS demand rate (e_t^{SAMS}) and SAMS fleet service rate ($r^{AV}S$), analogous to a queueing system model. The SAMS service rate is a product of the SAMS fleet size (S) and the average service rate of an AV in the SAMS fleet (r^{AV}). r^{AV} is based on simulations of a shared-ride SAMS fleet. Several factors impact the service rate, such as trip length, total demand, and maximum detour distance for shared-ride pickups; the service rate in this study was calibrated to the specific shared-ride SAMSs (with a maximum in-vehicle traveler detour distance/time of 30%) proposed for a specific service region using Chicago synthetic demand and taxi data. Although there are economies of density with shared-ride SAMS, these are beyond the scope of a planning design-level transit-SAMS design model.

Eqn. (3.3)-(3.5) define the piecewise linear relationship between u_t and the estimated ρ_t displayed in Figure 3.2. In a queuing system, the relationship between utilization rate and average wait time is highly nonlinear. In this model, as utilization ρ_t increases from 0 to a_1, wait time (u_t) remains flat and relatively low at the minimum average traveler wait time (w_o), as there are always empty AVs to serve new requests. When $\rho_t > a_1$, the possibility of all AVs in the fleet being occupied exists and hence average wait time u_t increases slightly (Δ_1) with ρ_t between a_1 and a_2. Finally, when $\rho_t > a_2$, the average utilization rate ρ_t approaches full utilization, at which point average wait time u_t increases rapidly (Δ_2) with ρ_t. This piecewise linear relation acts to prevent the transit-SAMS designer from setting the SAMS fleet size parameter (S) too low, such that the SAMS fleet cannot serve the SAMS demand (e_t^{SAMS}).

Eqn. (3.6) represents an operating budget constraint. The first term is the transit operational cost of pattern $p \in P$ with trip duration l_p and frequency $\frac{1}{h_p}$, where c_1 is cost per unit time and $\sum_p l_p \frac{1}{h_p}$ is the total operating time of all patterns. The second term represents the operational cost of the SAMS, or the cost to subsidize its operation. The third term is the farebox revenue (including SAMS fares). The total operational cost of the transit lines and SAMS fleet, after subtracting farebox revenue, must not exceed the available operational budget (Γ).

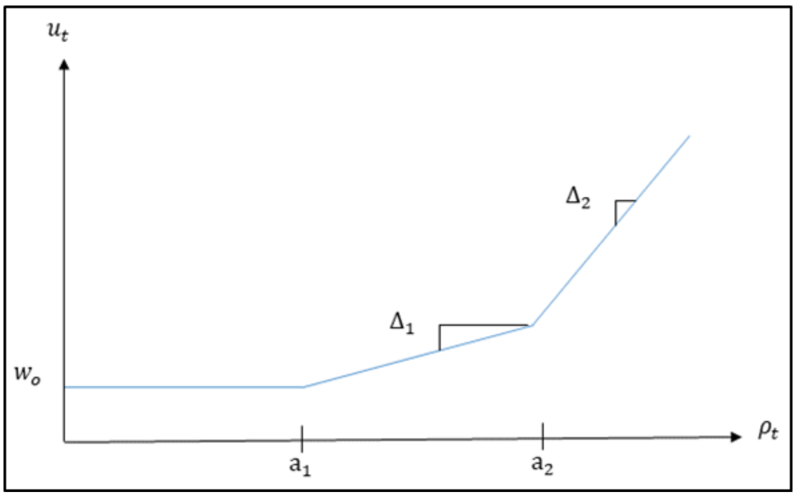

Figure 3.2: Piecewise linear relationship between average SAMS traveler wait time (u_t) and estimated SAMS fleet utilization rate (ρ_t)

Eqn. (3.7) ensures enough buses of type $j \in J$ (v_j) exist to complete all the necessary transit trips made by vehicles of type j, during time interval $t \in T$. τ'_p is the pattern operating time bounded from above by the time interval length (30 minutes). The bound is necessary because a pattern p, by definition in Chapter 2.5, is only active in a time interval t (i.e. pattern headways cannot be greater than the time interval it belongs to). $\frac{\tau'_p}{h_p}$ gives the required number of transit vehicles for pattern p. Summing the required fleet of all patterns with dispatch interval $d_p = t$ and vehicle type $\theta_p = j$, gives the minimum fleet size of transit vehicle type θ_p required during time interval d_p. For simplification, patterns are considered to remain with the same vehicle type; hence, the decision problem does not aim to find the vehicle type for pattern p, but rather to optimize pattern headway given the available number of transit vehicles of type $j \in J$.

The constraint in Eqn. (3.8) represents a capital cost budget measured in terms of AV-equivalents. V_j^o is the initial number of transit vehicles of type $j \in J$; S^o is the initial SAMS fleet size; and k_j is the AV-equivalents associated with transit vehicle type j. For example, if buses of type j cost $10x$ and AVs cost $2x$, then k_j would be 5. The summation of the chosen number of AVs (S) and the chosen number of AV-equivalents ($\sum_{j \in J} k_j v_j$) must be no greater than the summation of the initial AVs (S^o) and initial AV-equivalents ($\sum_{j \in J} k_j V_j^o$). This upper bound can be adjusted by a factor μ. The purpose is to capture the capital cost budget constraint facing transit agencies. Purchasing AVs to operate an SAMS will limit the agency's ability to purchase buses. Many transit agencies treat capital and operational cost budgets differently. Hence, both Eqn. (3.6) and Eqn. (3.8) are needed to properly model transit agency budgets.

The constraint in Eqn. (3.9) represents the pattern headway policy for each transit mode $m \in M^t = \{bus, rail\}$, for all patterns in the pattern set of mode m, $p \in P_m$. Rail transit patterns are required to remain and therefore have upper and lower bounds. Bus patterns can be removed; hence, their pattern headway is only bounded from below.

Eqn. (3.12) represents the user equilibrium constraint at the mode choice and route choice levels, as well as the relationship between the upper-level decision variables (h_p, S) and the dynamic combined mode choice—traveler assignment problem (DCMC-TAP). This is the only constraint in (3.1)-(3.14) that refers to the lower-level problem.

The constraints in Eqn. (3.13)-(3.14) require the SAMS fleet size, the overcrowding auxiliary variables (L_p, R_t^{SAMS}) and fleet size of transit vehicle type $j \in J$ (v_j) to be nonnegative.

3.4 Solution Approach

3.4.1 Overview

To solve the bi-level problem in Eqn. (3.1)-(3.14), this study employs a heuristic implementation of the approach outlined in Figure 3.2. The upper-level problem is the JTNR-SFSDP, except that the lower-level decision variables (mode choice and route choice) are fixed. The outputs of the upper-level model are transit pattern headways and SAMS fleet size. This information is passed to the lower-level model, which is a dynamic combined mode choice-traveler assignment agent-based model. The lower-level model returns transit pattern demand and time-dependent SAMS demand to the upper-level module.

The solution procedure is as follows:

Step 0: Set parameters and initialize

Set upper-level model parameters $k_j, V_j^o, S^o, r^{AV}, w_o, a_1, a_2, \Delta_1, \Delta_2, c_1, c_2, \Gamma, h_m^-, h_m^+, \gamma_p^L, \gamma_t^R, \mu$.

Set transit network parameters: dispatch time interval d_p, trip length l_p, vehicle type θ_p and capacity B_p.

Initialize SAMS fleet size S and transit pattern headways h_p.

Initialize ODT-dependent modal demand $R_{o,d}^{t^a,m}$.

Step 1: Create a transit timetable

Convert transit pattern headways h_p obtained from upper level (or initial configuration) into transit pattern trips Z_p. The first trip of pattern p is set to start at time $t_{1,p}$ based on the start time t_o of the pattern dispatch time interval d_p; t_e is the end time of the dispatch time interval. For a trip $z \in Z_p$, the trip start time $t_{z,p}$ is computed as follows:

$$t_{z,p} = \{\frac{t_o + h_p}{2}, \quad if\ z = 1\ t_{z-1,p} + h_p, \quad if\ t_{z-1,p} + h_p \le t_e\ t_{z-1,p} + \frac{h_p}{2} + \frac{h_{p'}}{2}, \quad if\ t_{z-1,p} + h_p > t_e \quad \forall\, z \in Z_p, p \in P, p' \in P: d_{p'} = d_p + 1$$

Step 2: Run simulation models

SAMS simulation: Given fleet size S from the upper-level, and ODT demand for SAMS ($R_{o,d}^{t^a,m} | m = SAMS$), simulate the SAMS fleet serving SAMS demand to obtain ODT-dependent performance metrics.

Transit assignment-simulation: Given the transit pattern trips Z_p obtained in Step 1, and ODT demand for transit ($R_{o,d}^{t^a,m} | m = transit$), solve the dynamic transit traveler assignment problem using an iterative assignment-simulation approach to obtain the ODT-dependent performance of the transit system. Go to Step 3.

Step 3: Run time-dependent mode choice model

Using the ODT-aggregated SAMS fleet and transit network performance output from Step 2, update traveler mode choice probabilities $\Psi_{o,d}^{t^a,m}$ and modal flows $R_{o,d}^{t^a,m}$. If mode choice probabilities converge, go to Step 4; if not, go to Step 2 with updated modal flows $R_{o,d}^{t^a,m}$.

Step 4: Run transit-SAMS design model (A) and adjust solution to account for lower level response (B)

A: Using time-dependent SAMS demand (e_t^{SAMS}) and pattern-level transit demand (e_p), aggregated from Step 3, employ a nonlinear programming solver designed to find local optimal solutions to solve the mathematical program described in Section 3.1. Let X be the solution vector containing SAMS fleet size S and pattern headways h_p. The obtained solution for current iteration k (X_k^{dir}) is used as a recommended direction to be further evaluated.

B: Repeat Step 2 and Step 3 to evaluate lower level response for X_k^{dir}. Compare lower level performance of solution X_k^{dir} with that of previous iteration (X_{k-1}, or initial configuration if $k = 1$), and move from solution X_{k-1} in the direction of X_k^{dir} proportionally to the improvement of the objective function, denoted as W. The adjusted solution X_k is found in Eqn. (3.15).

$$X_k = \left[1 + \frac{W_{k-1} - W_k^{dir}}{W_{k-1}} * \frac{X_k^{dir} - X_{k-1}}{X_{k-1}}\right] * X_{k-1} \tag{3.15}$$

In the computation of the adjusted solution X_k, the first term in the product represents the improvement in the objective function value, W; hence it is positive if W decreases. The second term in the product is the step size relative to the previous solution. If the objective value decreases, the adjusted solution will be between X_{k-1} and X_k^{dir}. Otherwise, the adjusted solution moves away from X_k^{dir}.

The objective value W is computed as the sum of cumulative transit wait time ($\sum_p e_p w_p$), SAMS wait time ($\sum_t e^{SAMS} w_t$) and a penalty for travelers who could not be served (γe_u), as seen in Eqn. (3.16), where $\gamma = 30min$, e_u are unserved travelers, and w_p and w_t are *experienced* traveler wait times in the transit and SAMS systems (as evaluated in the lower level).

$$W = \sum_p e_p w_p + \sum_t e^{SAMS} w_t + \gamma e_u \tag{3.16}$$

If the solution converges, then terminate. If not, then go to Step 1 with new fleet size S and pattern headways h_p. The solution converges when the upper-level decision variables (transit headways and SAMS fleet size) converge (i.e. no further improvement in the objective function is possible), and the lower level time-dependent SAMS demand (e_t^{SAMS}) and pattern-level transit demand (e_p) converge (i.e. the demands are internally consistent with the design variables).

In Step 2, the SAMS simulator obtains the respective experience for individual travelers via running a simulation and dynamically operating a SAMS fleet, using assignment, routing, and scheduling algorithms. Simultaneously, the transit assignment-simulation model solves a congested multi-modal time-dependent assignment problem via iteratively (1) determining least-cost transit hyperpaths on a time-dependent network; (2) assigning transit travelers to a transit hyperpath; and (3) simulating the performance of transit travelers and vehicles in a congested urban transit network. The simulation captures crowding on transit vehicles and at transit stops. The transit assignment-simulation model returns the performance of the transit network and the experience of individual travelers. See Verbas et al. (2015) for details of the transit assignment-simulation solution approach with the NU-TRANS tool.

In Step 3, the respective performance of the SAMS and the transit network at the ODMT-level provide the attribute values input to the mode choice model, which assigns (or reassigns) individual travelers to walking, transit, SAMS or SAMS+Transit, based on the ODT performance of each mode. System performance is captured in terms of fare, in-vehicle travel time, wait time, walk time, in-vehicle standing time (when unable to find a seat), number of transit transfers, and SAMS probability of sharing a ride. The mode choice model then feeds this demand into the traveler assignment-simulation module (Step 2). This process repeats until the modal flows converge. The DCMC-TAP is challenging because it is a fixed-point problem with interdependencies between the mode choice probabilities and the transit and SAMS system performance. Hence, several iterations of the mode choice model may be required. See Chapter 3 and Verbas et al. (2016) for details of the modal assignment procedure.

In Step 4, the upper-level transit-SAMS design module receives converged mode and route flows from the lower-level dynamic combined mode choice-traveler assignment-simulation model (i.e. Step 3 and Step 2). From the lower-level (Step 3), Step 4 also receives an evaluation measure of the user experiences that incorporates user wait times and crowding in transit and SAMS vehicles. This measure is what the upper-level objective function tries to approximate at an aggregate level. Given the converged mode and route flows, the non-linear programming solver obtains values for the upper-level decision variables – transit pattern headways and SAMS fleet size – that minimize cumulative traveler wait time. The values of the two upper-level decision variables are then fed back into the lower-level model in order to evaluate the user response to the new proposed transit pattern headways and SAMS fleet size. The lower-level model then evaluates the user response and returns an evaluation measure of user experience. The algorithm then compares the new evaluation measure of user experience (based on the new proposed upper-level decision variables) with the previously obtained evaluation measure from Step 3. The gap between these two measures is used to drive the upper-level decision variables in the direction of the local design optima.

The evaluation measure, W, is analogous to the upper-level objective function except that the attribute values, instead of being estimates, reflect the user experiences simulated in the lower level. Additionally, the unserved travelers accounted for in W are only partially accounted for in X_k^{dir} because some of them are not directly associated to a travel mode and therefore do not affect the direction of the decision variables. In other words, while the two last terms in Eq. (3) penalize the rejection of travelers who were assigned to either transit or SAMS (in case of insufficient transit frequency or SAMS fleet), there can still be other travelers who had not been assigned to either mode (called unserved). This can happen if the prior decision variable configuration did not provide any viable mode for a certain traveler, such that the modal demand only represents a part of the population. Hence, the adjustment of the direction decision variables completes the solution approach by including the response of travelers.

This step-by-step process repeats until the transit-SAMS design problem decision variables and the modal flows and traveler experience in the lower level stabilize at a local optimum, and are mutually consistent between the upper-level and lower-level. It is important to note that this is a simulation-based heuristic procedure and the solution is not guaranteed to reach global optimum.

In the application of this modeling framework, it is recommended to test various starting points in order to possibly obtain several local optimal solutions.

3.4.2 Upper-Level Model Formulation

The upper-level model is displayed in Eqn. (3.1)-(3.14) with one important change. The time-dependent SAMS demand (e_t^{SAMS}), the transit pattern demand (e_p), and the transit pattern maximum flow (g_p) are fixed values (not dependent on the upper-level decisions). The values for these three variables are taken from the previous iteration of the lower-level model.

3.4.3 Lower-Level Model Formulation

The formulation of the lower-level DCMC-TAP is introduced in Chapter 3 and originally proposed by Verbas et al. (2016). This section presents the mathematical formulation of the DCMC-TAP.

DCMC-TAP Formulation

This section presents the mathematical formulation of the time-dependent mode choice model in the lower-level DCMC-TAP. The equilibrium-based formulation follows the logic presented in Zhang et al. (2011) who formulate an integrated mode choice—traffic assignment problem.

Equation (3.17) displays the mathematical relationship between modal probabilities $\Psi_{o,d}^{t^a,m}$ and the list of travelers assigned to each mode $R_{o,d}^{t^a,m}$.

$$\Psi_{o,d}^{t^a,m}\left(\left|R_{o,d}^{t^a,m'}\right|_{m'\in M}\right) = \frac{\left|R_{o,d}^{t^a,m}\right|}{\left|R_{o,d}^{t^a}\right|} \qquad \forall o \in O, d \in D, t^a \in T^a, m \in M \tag{3.17}$$

Since the probability $\Psi_{o,d}^{t^a,m}$ of choosing mode $m \in M$ is itself a function of the modal flows of all modes $\left|R_{o,d}^{t^a,m'}\right|_{m'\in M}$, the time-dependent mode choice problem can be defined as a fixed-point problem. The objective of this fixed-point problem is to find the optimal modal flows $\left|R_{o,d}^{t^a,m'}\right|_{m'\in M}$, satisfying the condition in (3.18).

$$\left|R_{o,d}^{t^a,m}\right|^* = \left|R_{o,d}^{t^a}\right| \times \Psi_{o,d}^{t^a,m}\left(\left|R_{o,d}^{t^a,m'}\right|_{m'\in M}\right) \qquad \forall o \in O, d \in D, t^a \in T^a, m \in M \tag{3.18}$$

The fixed-point problem can be re-formulated as a gap-based nonlinear program (Zhang et al., 2011), as in Eqn. (3.19)-(3.21):

$$GAP_M = \frac{1}{2}\sum_{q\in Q}\sum_{r\in R}\sum_{t\in T}\sum_{m\in M}\left(\left|R_{o,d}^{t^a,m}\right| - \left|R_{o,d}^{t^a}\right| \times \Psi_{o,d}^{t^a,m}\left(\left|R_{o,d}^{t^a,m'}\right|_{m'\in M}\right)\right) \tag{3.19}$$

such that:

$$\sum_{m\in M}\left|R_{o,d}^{t^a,m}\right| = \left|R_{o,d}^{t^a}\right| \qquad \forall o \in O, d \in D, t^a \in T^a \tag{3.20}$$

$$\left|R_{o,d}^{t^a,m}\right| \geq 0 \qquad \forall o \in O, d \in D, t^a \in T^a, m \in M \tag{3.21}$$

The objective displayed in Eqn. (3.19) minimizes the discrepancy between the assigned modal flow $\left|R_{o,d}^{t^a,m}\right|$ and the expected modal flow $\left|R_{o,d}^{t^a}\right| \times \Psi_{o,d}^{t^a,m}\left(\left|R_{o,d}^{t^a,m'}\right|_{m'\in M}\right)$ summed over all origins $o \in O$, destinations $d \in D$, departure time intervals $t^a \in T^a$ and modes $m \in M$. The convergence of GAP_M to zero satisfies the fixed-point problem in Eqn. (3.18). Equation (3.20) is the flow conservation constraint, and Eqn. (3.21) satisfies the nonnegativity of modal flows.

Despite the relatively simple mathematical formulation in Eqn. (3.18)-(3.21), the amount of information contained in it and the interdependencies between terms $\Psi_{o,d}^{t^a,m}\left(\left|R_{o,d}^{t^a,m'}\right|_{m' \in M}\right)$ and $R_{o,d}^{t^a,m}$ make the problem analytically intractable; i.e., it cannot be solved through an explicit analytical relation. For this reason, the problem is solved with an agent-based simulation approach. The next two subsections present the dynamic transit assignment simulation model and the SAMS fleet simulation model that feed into the mode choice model to solve the DCMC-TAP.

Dynamic Transit Assignment-Simulation Model

Determining least cost ODT paths for each mode $L_{o,d}^{t^a,m}$ and path-dependent ODT modal flows $\left|R_{o,d}^{t^a,m,q}\right|$ involves solving a shortest hyperpath problem and the dynamic traveler assignment problem (DTAP), respectively. The agent-based simulation approach used to solve this part of the problem is shown in Figure 3.3. As described in Verbas (2015), characteristics of the least cost hyperpath calculation algorithm include (i) a multimodal formulation that considers transit modes such as bus, rail, and commuter rail as well as walking and biking (option disabled in this study); (ii) it is time-dependent because it considers frequency and availability of transit service patterns in different time intervals; (iii) generalized cost is movement- and approach- dependent; (v) link and node costs are frequency-based; (vi) it enables penalization of transfers; and (vii) it accounts for probability of standing or being denied boarding due to overcrowding in transit vehicles.

$L_{o,d}^{t^a,m}$ and $\left|R_{o,d}^{t^a,m,q}\right|$ also directly depend on the transit pattern headways (h_p) and SAMS fleet size (S) determined in the upper-level problem. Smith (1993) shows that the DTAP is not necessarily convex and that multiple solutions may exist. The decision variables in the lower-level problem – the list of ODT travelers assigned to each path $R_{o,d}^{t^a,m,q}$, and the list of ODT travelers assigned to each mode $R_{o,d}^{t^a,m}$ – are easily converted to the upper-level input parameters – pattern-level transit demand e_p and time-dependent SAMS demand e_t^{SAMS}. Hence, the lower-level problem captures traveler responses to changes in the SAMS fleet size (S) and transit pattern headways (h_p).

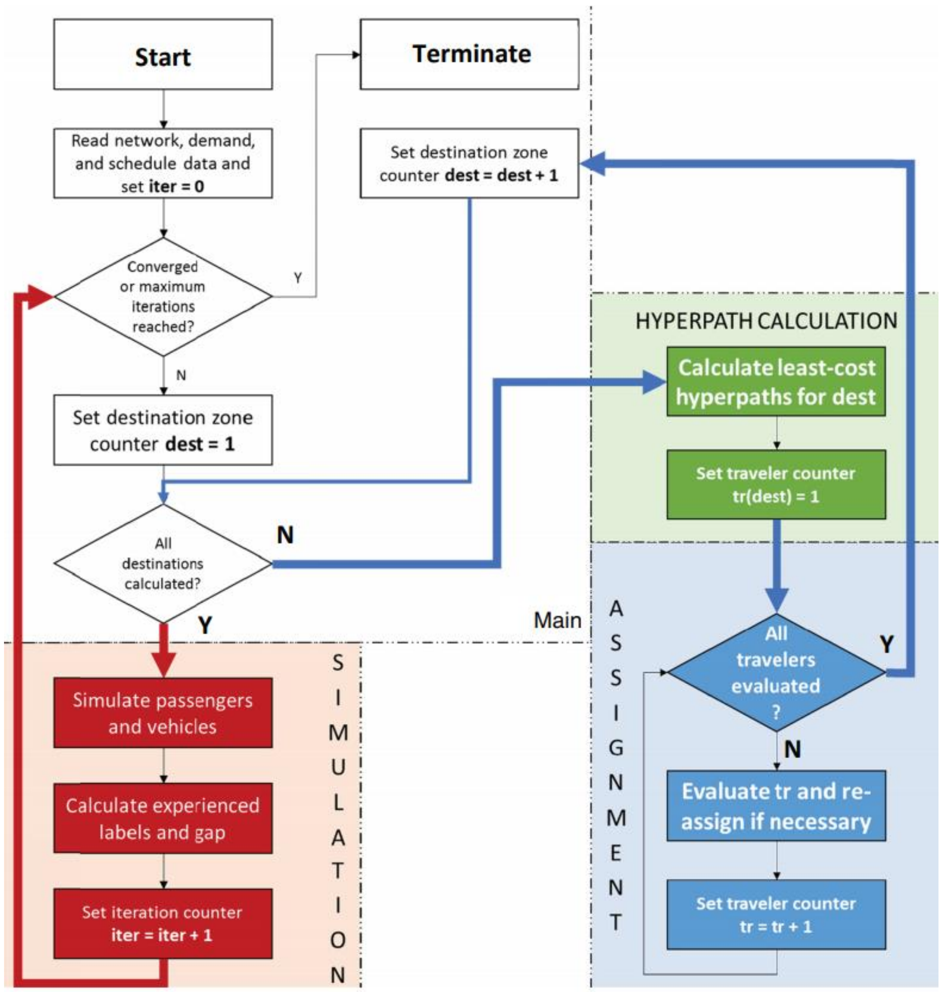

Figure 3.3: Flowchart of transit assignment-simulation algorithm in lower level (Verbas et al., 2015)

SAMS Simulation

This section presents the simulation model of an on-demand shared-ride SAMS fleet. This model is adopted from work by Hyland and Mahmassani (2018) and it has the following characteristics:

- Travelers request rides dynamically via a mobile application;

- A request includes a pickup location and a drop-off location, both of which must be within a pre-defined geographical service region;
- Travelers want to be served (i.e. picked up) immediately;
- Travelers will always be served, assuming they are willing to wait;
- A single AV picks up and drops off a traveler request i but the same AV may pick up and/or drop off other traveler requests while traveler request i is in the AV;
- The AVs in the fleet are functionally homogeneous;
- The AVs can only have two traveler requests inside at one time (similar to Lyft Line);
- The AV fleet operator has complete control over each AV.

The operational problem associated with an on-demand shared-ride SAMS is a stochastic dynamic control problem. The problem is dynamic as travelers make requests while the SAMS fleet is in operation, and the travelers want to be served immediately. The problem is stochastic because these user requests are random (drawn from a spatial-temporal demand distribution) from the perspective of the SAMS fleet operator.

This study employs an optimization-based solution approach to assign idle/empty and en-route drop-off AVs to open user requests (meaning, they have not been assigned to an AV yet). The solution approach involves solving an optimization problem every $\Delta\tau$, the inter-decision time.

Let C^o and C^{IV} denote the set of open user requests and in-vehicle user requests. If τ is the current time and t_i^r is the request time of user $i \in C^o$, then user i's elapsed wait time (w_i) is $w_i = \tau - t_i^r$. Similarly, let V^I, V^P, and V^D be the set of idle, en-route pickup, and en-route drop-off AVs respectively; $V = \{V^I, V^P, V^D\}$. Moreover, let V' denote the subset of AVs that are available to be assigned to user requests. V' only include idle and en-route drop-off AVs, not en-route pickup AVs; $V' = \{V^I, V^D\}$.

All idle AVs V^I can be assigned to all open user requests C^o; however, some en-route drop-off AVs V^D are not eligible to be assigned to any open user requests C^o. Let d_i and d_i^{max} denote the cumulative detour distance of user i and the maximum detour distance of user i, respectively. Then if $d_i \geq d_i^{max}$, the en-route drop-off AV $j(i) \in V^D$ carrying user $i \in C^{IV}$ is not allowed to be assigned to another user; these AVs are not considered in the assignment problem. Similarly, if assigning en-route drop-off AV $j \in V^D$ to an open user request $i \in C^o$ would increase the detour distance of either the in-vehicle user inside the AV $d_{i(j)}$ or the open user request d_i above their respective maximum detour distances $d_{i(j)}^{max}$, d_i^{max}, then the AV-user assignment is not feasible. Let f_{ij} equal one if there is a feasible match between en-route drop-off AV $j \in V^D$ and open user request $i \in C^o$, and zero otherwise.

At every decision epoch, the SAMS fleet operator solves the mathematical programming problem defined below. The time between epochs is the inter-decision time $\Delta\tau = 15\ s$. The math program utilizes the assignment (bi-partite matching) problem structure. The formulation of the myopic AV-user shared-ride assignment problem is given in Eqns. (3.22)-(3.26):

$$\min_{x_{ij}} \sum_{i \in C^o} \sum_{j \in V'} x_{ij} \{ c^{VOT} (t_{ij}^t + t_{ij}^d - w_i) + c^{EDCR} (d_{ij}) - r^{asgn} \} + c^{share} \sum_{i \in C^o} \sum_{j \in V^{IV}} x_{ij} \tag{3.22}$$

s.t.

$$\sum_i x_{ij} \leq 1 \qquad \forall j \tag{3.23}$$

$$\sum_j x_{ij} \leq 1 \qquad \forall i \tag{3.24}$$

$$x_{ij}\left(1-f_{ij}\right)=0 \qquad \forall i,j \in V^D \qquad (3.25)$$

$$x_{ij} \in \{0,1\} \qquad \forall i,j \qquad (3.26)$$

where x_{ij} equals 1 if AV $j \in V'$ is assigned to user $i \in C^o$, and zero otherwise.

The objective function includes penalty terms for remaining empty pickup time t_{ij}^t, added user detour time t_{ij}^t, and empty distance d_{ij} to pick up a user. The parameters $c^{VOT} = \frac{\$23}{hour}$; $c^{EDCR} = \frac{\$0.50}{mile}$; and $c^{share} = \$0.0$ denote the value of time, empty distance cost rate, and the penalty for assigning an open user request to an en-route pickup AV. The objective also includes a reward for assigning an AV to a user ($r^{asgn} = \$10.0$) and a reward that increases as a function of the elapsed wait time of user i.

Eqn. (3.23) ensures that each AV j is assigned to at most one open user request. Eqn. (3.24) ensures that no more than one AV is assigned to a single open user request. Eqn. (3.25) ensures only feasible AV-user assignments are made. Eqn. (3.26) is the integrality of the decision variable.

Figure 3.4 displays the agent-based SAMS simulation model for the on-demand shared-ride SAMS that incorporates the mathematical program in Eqn. (3.22)-(3.26). The model assumes it takes 15 seconds to drop-off a traveler and 45 seconds to pick up a traveler. The vehicles travel at 35 miles per hour, a relatively high speed. The model also assumes the maximum detour distance for traveler i (d_i^{max}) is a 30% increase in distance relative to the traveler's shortest path distance from her origin to destination without sharing a ride.

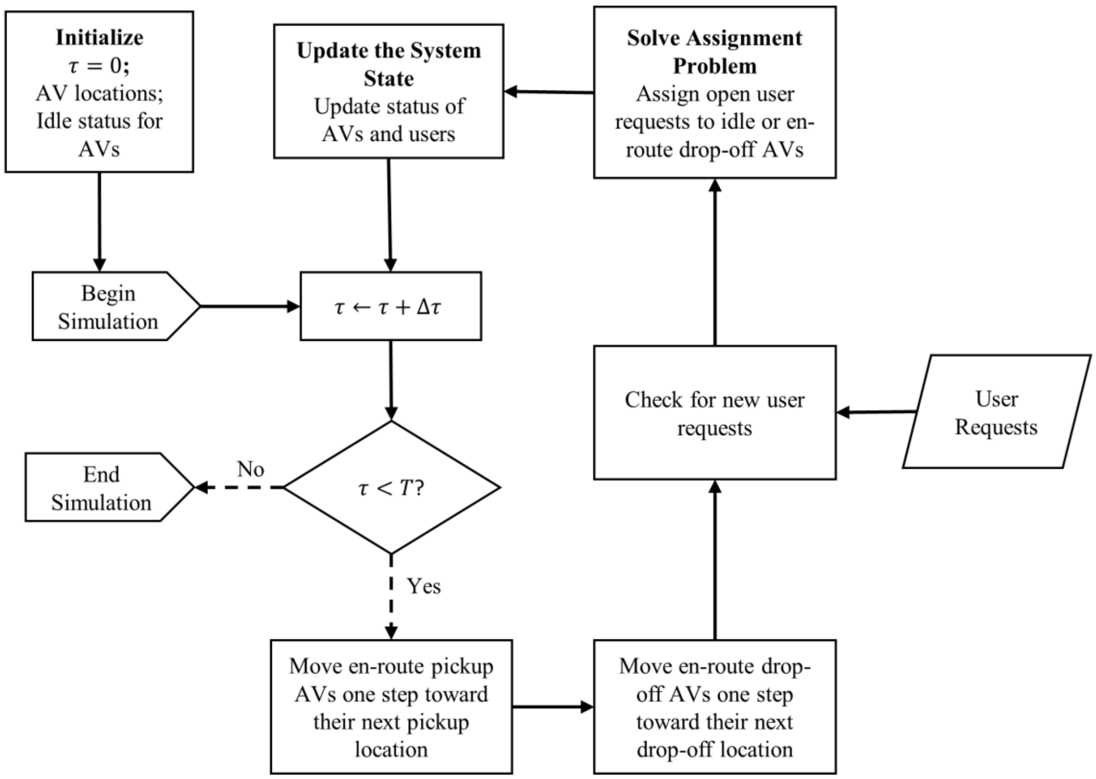

Figure 3.4: Flowchart of the SAMS simulation algorithm in lower level

3.5 Experimental Design

The modelling framework is demonstrated at scale using the actual network of the Greater Chicago metropolitan area (USA). The application is performed in the large-scale urban network for the transit system, with SAMS fleet coverage tested in a suburban area. Transit data is taken from the General Transit Feed Specification for services provided by the Chicago Transit Authority and Metra (operators of urban bus and heavy rail, and commuter rail respectively).

The SAMS fleet coverage test area includes the city of Evanston and an 8 km (5 mi) buffer area that surrounds it. Evanston is a north suburb of Chicago with a population over 74000. A significant flux of travelers who live in Evanston commute to work in Chicago daily. Similarly, an influx of travelers commute to Evanston from Chicago. Given the seamless integration of the transit lines in Evanston and Chicago, we model the entire Chicago transit network to have a realistic picture of the congestion in the transit vehicles.

The transit network is composed of 14259 transit stops, 64083 links and 146 transit routes. Moreover, the transit routes are distributed across 1081 transit service patterns (subsets of routes) wherein 660 are in the simulated period (morning peak). The demand data includes a list of approximately 630000 riders in the greater study area, from which 88000 riders are located in the suburban test area. Travelers listed in the demand data are expected to be transit travelers based on a household travel survey. Other general parameters and properties assumed for a typical weekday service during the morning peak period (6 to 10 a.m.) are described in Table 1.

The simulations were performed in a compute node running the Red Hat Enterprise Linux 6 operating system, with the following characteristics: Intel Xeon E5-2680, v4 14C 2.4 GHz, Intel® QPI, 2500 MHz, 28 cores and 128GB RAM. The memory requirement is approximately 50GB.

The dynamic transit assignment and simulation portion of the lower level (step 2 in Section 4.4.1) is solved with the NU-TRANS[2] tool and parallelized using OpenMP threads. Computational time is approximately 10 min for every iteration in the transit assignment-simulation. Still in the

[2] Northwestern University dynamic transit assignment-simulation tool, described in (Ö. Verbas et al., 2016).

lower level, SAMS simulation time varied according to SAMS fleet size and assigned demand from each iteration, taking 30 minutes to 1 hour to run with Python and Gurobi optimizer.

Table 1: Input parameters and characteristics of bi-level model

Sets		**Variable**	
P	$\lvert P \rvert = 2680$	F_p	U$2.25
T	$\lvert T \rvert = \lvert T^a \rvert = 48$ (time intervals in a day)	F^{SAMS}	Upper level: U$2.25
J	2 bus types (small and large); 15 train types		Lower level: U$2.25 (SAMS+Transit)
M	{walk, transit, SAMS, SAMS+Transit}		$min(\$4.85, \$3.64 + \$0.20/minute + \$0.81/mile)$ (SAMS)
M^t	{bus, rail}	Γ	U$ 311,000 (estimated based on initial configuration with some extra space)
Z	$\lvert Z_o \rvert = 20920$ (initial transit vehicle trips)	c_1	U$2.26/minute (bus) or U$2.40/minute (rail)
O, D	16819 MAZs (wherein 1348 are transit zones)	c_2	U$250/day
		w_o	3.0 minutes
		Δ_1, Δ_2	$\Delta_1 = 20$; $\Delta_2 = 50$
Simulation characteristics		a_1, a_2	$a_1 = 0.50$; $a_2 = 0.80$
Relative weight of waiting for transit:	2.0	r^{AV}	2 passengers/vehicle-hour
Relative weight of waiting for SAMS:	1.0	S^o	0 AVs
Relative weight of walking:	1.5	k_j	6 AVs for a small bus; 9 AVs for a large bus
Relative weight of standing:	2.0	h_m^-	2.0 minutes
Penalty for transfers (minutes):	5.0	h_m^+	30.0 minutes
Walking speed (meters/second):	1.4	γ_p^L	30.0 minutes per denied boarding
Simulation interval length (seconds):	1.0	γ_t^R	30.0 minutes per denied boarding
Transit assignment and least-cost hyperpath calculation interval length (minutes):	30.0	μ	6.0

The upper level takes up to 10 min to calculate pattern demands and the SAMS fleet size (direction solution, step 4A). The solution adjustment (step 4B) requires repetition of steps 2 and

3, which takes about 4 hours to complete 28 inner iterations and adjust the upper level solution based on lower level response. For every iteration of the upper level, the lower level performed 28 inner iterations. The mode choice model and SAMS simulation are run in the first and every fourth inner iteration (lower level). Altogether, the computational time of a full iteration of the bi-level approach takes approximately 8 hours.

The nonlinear programming solver used in the upper level transit-SAMS design model is KNITRO. The tool is configured to use a feature called *parallel multistart*, which uses parallelization to solve the problem simultaneously from different feasible and randomly selected starting points. The returned solution is the local optimum with the best objective function value.

3.6 Experimental Results

The output is shown for lower level and upper level iterations. The lower level output represents the result of simulated experiences after 28 inner iterations of the dynamic transit traveler assignment-simulation. The upper level results show features associated with the recommended design (SAMS fleet size and pattern frequencies) across 30 upper iterations. An additional “iteration 0” is added to some of the charts to show conditions before any changes in the transit design and before an SAMS fleet is added. Finally a “No SAMS” scenario is shown as a benchmark to scenarios with an SAMS fleet.

The “iteration 0” is a lower level run where transit and walking are the only modes available for the users, and the transit system has its original network and schedule characteristics. It serves as a reference representing the user-equilibrium state of the unchanged system (*status quo*). Results are presented for multiple starting points of the SAMS fleet size to address possible convergence issues (500, 1000 and 2000 initial SAMS fleets). The initial SAMS fleet size is used to obtain an initial time-dependent SAMS demand, which is leveraged by the upper level to provide design recommendations starting from the first upper iteration.

Table 2: Mode Splits across Upper Iterations

Iteration	Unserved (%)			Transit (%)			Walk (%)			SAMS (%)			SAMS+Transit (%)		
Initial SAMS fleet	500	1000	2000	500	1000	2000	500	1000	2000	500	1000	2000	500	1000	2000
0 (base)		0.7			69.5			29.8			-			-	
5	0.2	0.2	0.2	59.4	59.3	59.3	29.2	29.2	29.2	5.7	5.9	5.8	5.5	5.5	5.5
10	0.2	0.2	0.2	59.7	59.9	59.7	29.1	29.0	29.1	5.6	5.7	5.7	5.4	5.3	5.3
15	0.2	0.2	0.2	60.2	60.1	60.1	28.8	28.9	28.9	5.6	5.6	5.6	5.2	5.3	5.3
20	0.2	0.2	0.2	60.3	60.4	60.1	28.8	28.7	28.8	5.4	5.4	5.7	5.2	5.2	5.2
25	0.2	0.2	0.2	60.3	60.4	60.3	28.7	28.7	28.7	5.5	5.6	5.5	5.3	5.2	5.3
30	0.2	0.2	0.2	60.4	60.5	60.5	28.7	28.6	28.7	5.5	5.4	5.5	5.2	5.3	5.2

Table 2 presents the observed modal splits at the end of each upper level iteration for transit, walk, SAMS and SAMS+Transit. The transit and walk modes are part of the transit path assignment in the NU-TRANS tool. The walk mode share is obtained by filtering the traveler trajectories that only include walking links. The transit ridership is shown to decrease by 13% from the base case to the last iteration; SAMS and SAMS+Transit ridership increase significantly from zero to approximately 5% each, the number of unserved travelers decreases by 70% and, to a smaller extent, the walking mode decreases by 4%. Unserved travelers represent travelers that either do not have a feasible travel mode from the origin to the destination in the departure time of interest, or travelers that were assigned to transit but were unable to reach their destination due to missed transfers or boarding rejections from overcrowded vehicles found in the simulation.

The similarity of the mode shares for different initial SAMS fleet sizes shown in Table 2 reinforces the robustness of the solution vis-a-vis different initial fleet size start points.

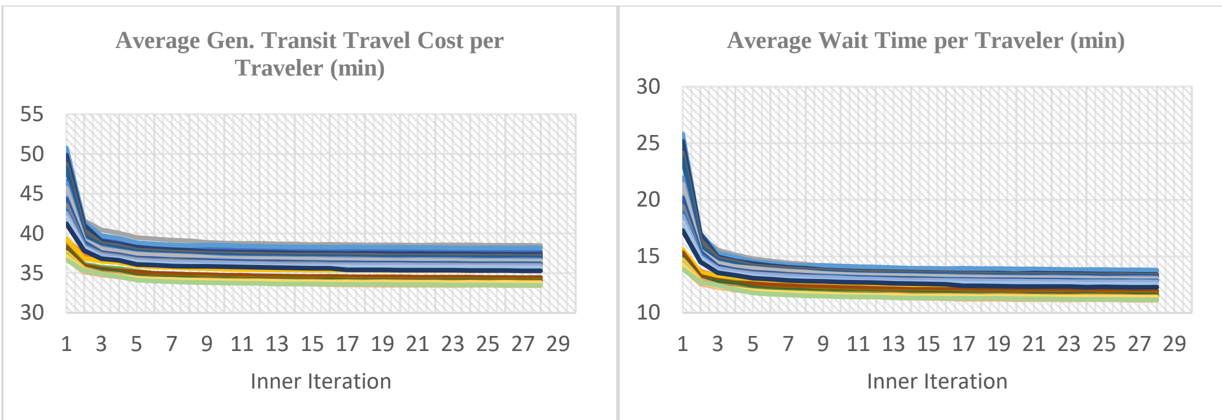

Figure 3.5: Average generalized transit travel cost and wait time per transit/walk traveler for different upper level iterations (lower level output)

Figure 3.5 shows that the lower level moves towards convergence of the average generalized transit travel cost and average wait time per traveler with 28 inner iterations; in this figure, each curve is an upper level iteration. Similarly, the mode choice model indicates that the travelers' mode decisions converge, as they switch less and less between modes across inner iterations (Figure 3.6).

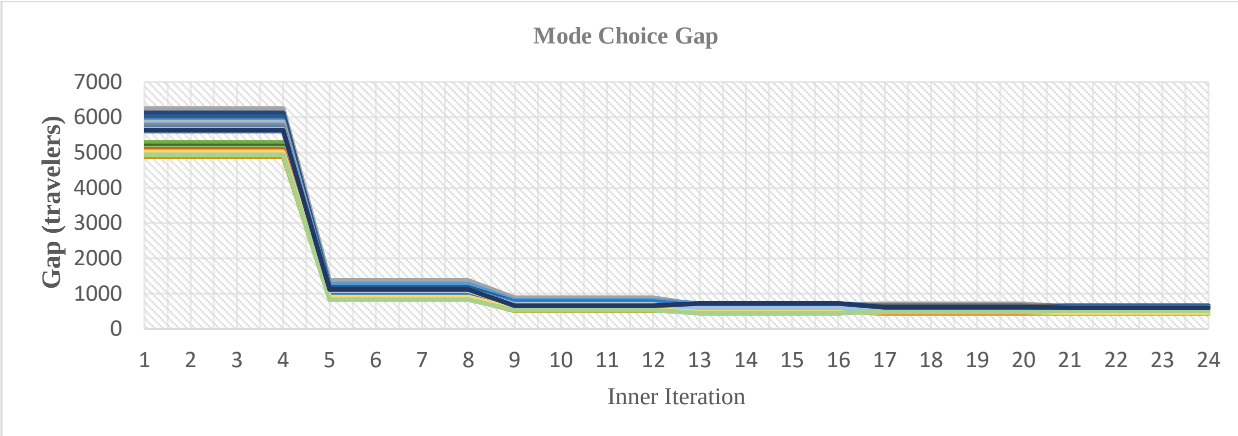

Figure 3.6: Convergence of mode choice gap for different upper level iterations (lower level output)

This is also reflected in the wait times experienced by riders (Figure 3.7). There is a decreasing trend for the wait time in the transit mode approaching 6 minutes in the "No SAMS" case, and it narrows down to approximately 5 minutes for SAMS and transit users in the SAMS scenarios in the Evanston case area. The objective value described in Step 4 of the solution approach is depicted in the bottom chart of the figure (showing wait times and penalties for rejections and unserved travelers). The curves account for cumulative wait times and boarding rejection penalties for all modes. It is noticeable that the SAMS scenarios present improved (lower) objective value than the "No SAMS" case in Evanston.

Figure 3.8 shows improving average transit wait times and experienced objective value for the general population. The "No SAMS" case however does not perform particularly different from the SAMS cases given that the SAMS population size is relatively small compared to the general population size. An additional insight taken from the lower level response is the effect on transit boardings, which increase with upper level decisions but boarding denials decrease (Figure 3.9).

Performance metrics of the upper level can be seen in Figure 3.10. The different initial SAMS fleet sizes do not seem to significantly affect the performance of the solution approach. The progress of the SAMS fleet size determination over different iterations lead to a common local optimal SAMS fleet size (approximately 1000 vehicles). The number of transit vehicle trips is shown to increase across upper level iterations in the same manner for all initial SAMS fleets. These vehicle trips are an immediate response to the changes in pattern frequencies. They increase more drastically for the “No SAMS” case given the same budget availability and they get much higher than initial conditions because we assume a generous initial operating budget.

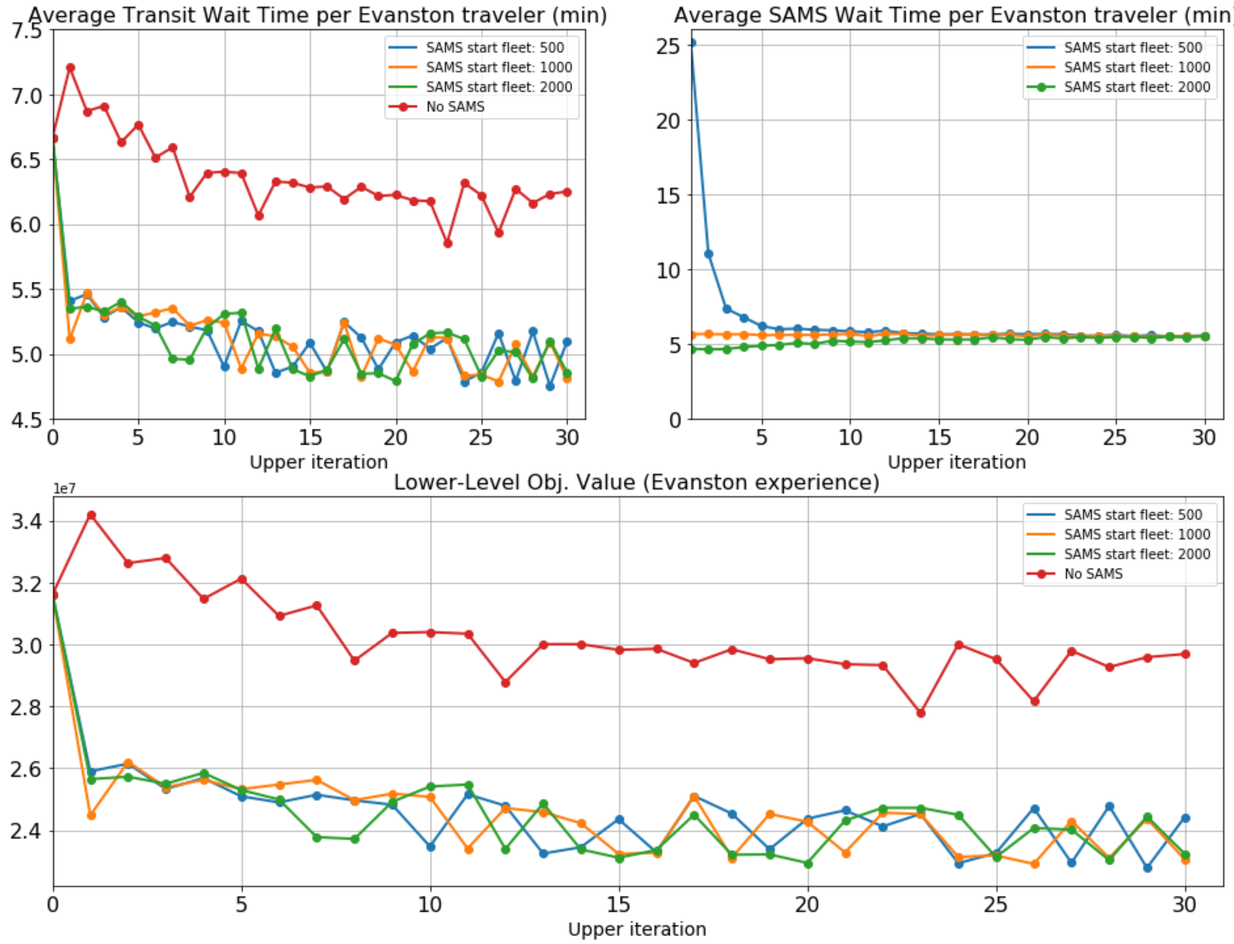

Figure 3.7: Experienced wait times on transit and SAMS and Objective Value for Evanston population

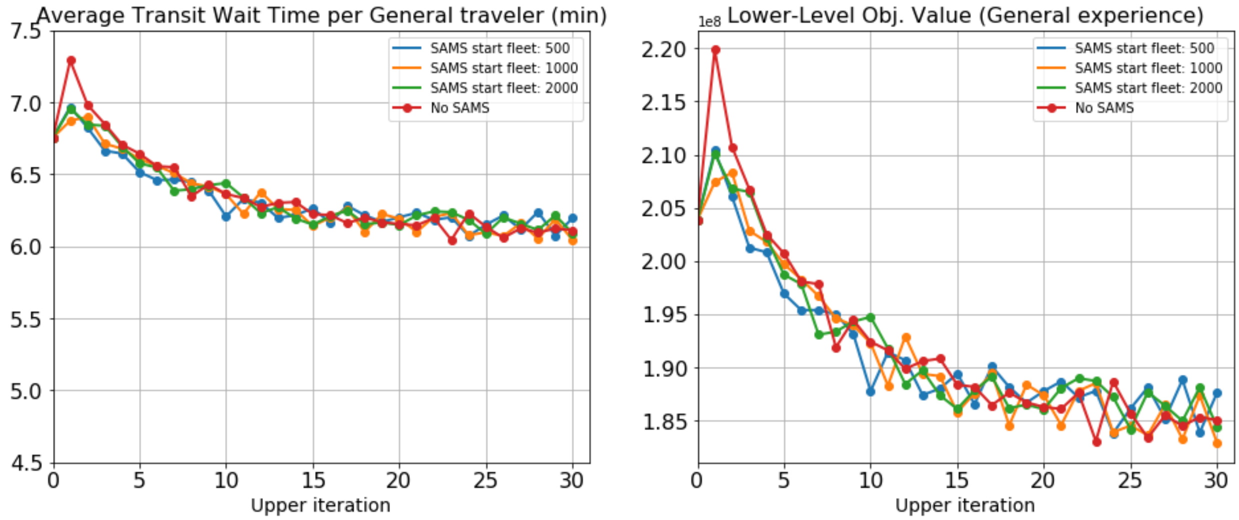

Figure 3.8: Experienced wait times on transit and SAMS for general population

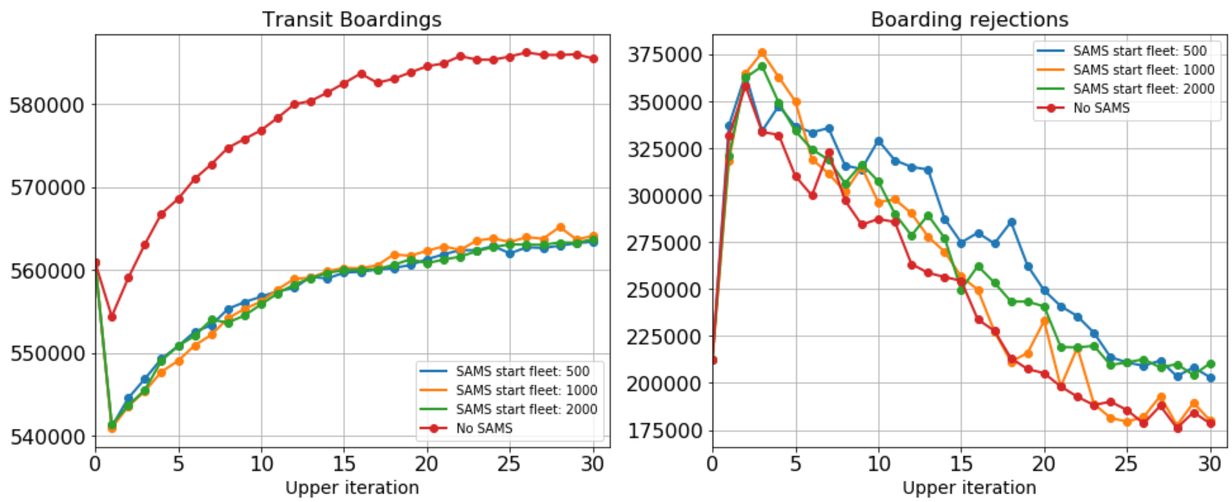

Figure 3.9: Experienced transit pattern boardings and transit rejections

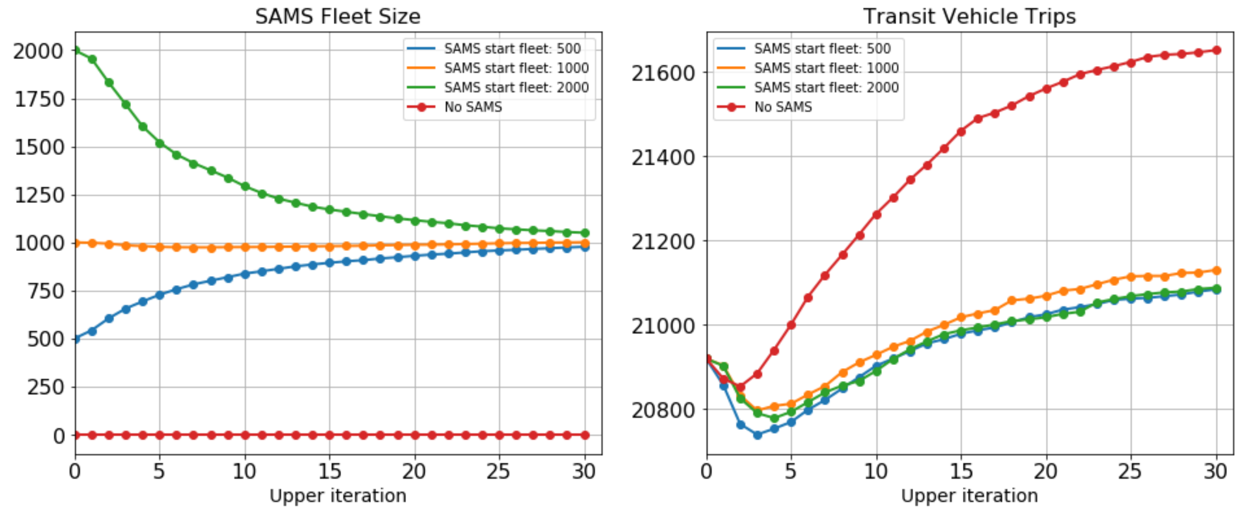

Figure 3.10: Performance metrics from the upper level

Table 3 displays a sample transit route and its patterns. The results suggest the removal of pattern 144, which is a “limited-stop” (“short-turn” or not full-length) pattern. Conversely, the results indicate that the solution procedure decreases the headway of most of the other patterns for this particular route, except for patterns 151-152. The headways estimated above 30 minutes are interpreted as suggestions of pattern removal but are not actually removed in the simulations.

Taken together, the traveler experience results and the results in Table 2 indicate that the modeling framework and solution procedure may effectively reallocate resources between transit pattern headways and SAMS fleet size in order to improve traveler experience. The nature of the solution clearly illustrates a trade-off between SAMS users and regular transit users. While overall wait times decrease, not all transit users benefit, as reduced frequencies on certain lines result in much reduced service levels. This suggests that agencies may wish to explore the impact of additional design constraints in the formulation.

Table 3: Upper level output sample of recommended pattern headways

Pattern	Time Interval	Cycle time (min)	Initial headway (min)	Solution headway (min)	Initial cost ($)	Solution cost ($)	Initial fleet size	Solution fleet size	Route	Direction	Stops	Pattern type
35	7	24.5	15	15.3	110.9	108.6	1.9	1.6	112	West	51	All stops
136	8	24.5	15	14.7	110.9	113.0	1.8	1.6	112	West	51	All stops
137	9	24.5	15	14.6	110.9	114.1	1.9	1.6	112	West	51	All stops
138	10	24.5	15	14.4	110.9	115.2	2	1.6	112	West	51	All stops
139	11	24.5	15	9.7	110.9	171.5	2	2.5	112	West	51	All stops
140	12	24.5	15	7.4	110.9	224.9	2	3.3	112	West	51	All stops
141	13	24.5	15	7.5	110.9	220.8	3	3.3	112	West	51	All stops
142	14	24.5	15	14.5	110.9	114.6	2	1.6	112	West	51	All stops
143	7	11	30	16.5	24.9	45.4	1.6	0.7	112	West	25	Limited stop
144	14	14	30	300.0	31.7	0.0	2	0.0	112	East	24	Limited stop
145	7	28	15	9.3	126.7	205.1	2	2.8	112	East	50	All stops
146	8	28	10	6.2	190.0	308.4	1.7	4.7	112	East	50	All stops
147	9	28	15	7.5	126.7	254.1	2	3.7	112	East	50	All stops
148	10	28	15	9.7	126.7	195.4	0.4	2.8	112	East	50	All stops
149	11	28	30	14.2	63.3	133.6	2	1.9	112	East	50	All stops
150	12	28	15	9.8	126.7	194.3	0.5	2.8	112	East	50	All stops
151	13	28	15	28.4	126.7	67.0	1	0.9	112	East	50	All stops
152	14	28	15	28.6	126.7	66.4	1.9	0.9	112	East	50	All stops

3.7 Sensitivity Analysis

Several parameters are provided as input to the bi-level solution approach in order to obtain results. Assumptions are made for initial operating subsidy, operating costs of transit vehicles, fares, initial SAMS fleet, etc. In this section, I perform an analysis to demonstrate how sensitive the framework is to changes in the available operating subsidy.

3.7.1 Impact of Subsidies on Transit Performance

Operating subsidy is a critical component for urban transit operation as the profit associated with this service is minimal, typically covering 33% of the total cost (Reeven and Karamychev 2016). To increase ridership, lower average fares are needed which can only be realized with subsidies. Although the general belief has been that subsidies increase ridership, research shows that the impact of subsidies in terms of user costs, transit performance, and labor productivity is not always positive.

Cervero and Pucher showed that higher operating costs per hour and lower productivity have been associated with higher operating subsidies (Cervero 1984; Pucher 1982). Reeven and Karamychev (2016) analyzed ridership data from 1991 to 2012 and found that subsidies have had insignificant effects on ridership. The same result was found by Malalgoda and Lim who suggest that policy makers should not rely on boosting transit funding to increase ridership, but rather consider incorporating innovative solutions in their operations, such as ride-hailing services (Malalgoda and Lim 2019). The above studies point to the need to study optimal allocation of funds in a world of transit integrated with SAMS.

3.7.2 Sensitivity to Operating Subsidy

The solution approach for the joint transit network design and SAMS fleet size determination problem was tested using five different available operating subsidy constraints. The available subsidy is provided only for the simulated morning peak period (6AM to 10AM). The different maximum available operating subsidies and their corresponding outcomes from running nine upper-level iterations of the framework are presented in **Table 4**. The value of $311,000 serves as reference because it was used to illustrate the large-scale application presented in the previous section.

All transit network characteristics and parameters used in the model remain the same as those used to provide the results from the previous section, except for the multiplier in the upper bound of the fleet capital budget. This time a value of 1.5 times the original fleet capital budget was used instead of 6 from the previous section. However, it was noticed that this did not change the results because the operating budget is a binding constraint that does not allow full use of the fleet capital budget in case of multipliers higher than 2.

Table 4: Ridership levels by mode in Evanston buffer zone by operating budget

Budget (US$)	SAV	TRW	Walk	SAV + Transit	Unserved	Evanston Total
220,000	470	69515	8387	9393	63	87828
265,000	487	69494	8452	9336	59	87828
311,000	433	69869	8416	9063	47	87828
356,000	434	69959	8421	8963	51	87828
401,000	500	69783	8424	9014	107	87828

The table shows ridership levels for each mode in Evanston and its 5-mi buffer zone. The possible modes are SAMS, Transit-only and SAMS + Transit. Travelers who cannot complete their trip with either of these available mode options are called "unserved".

It can be seen that with an increase in the available operating subsidy (budget), there is an increase in transit ridership, which supports the idea that there is an improvement in the transit level of service and therefore its attractiveness to riders. The number of riders who use SAMS as a feeder mode to transit also seem to decrease with an increased operating budget. That being said, all remaining modes (SAMS-only, walk, and unserved) seem unaffected by the changes in the operating budget.

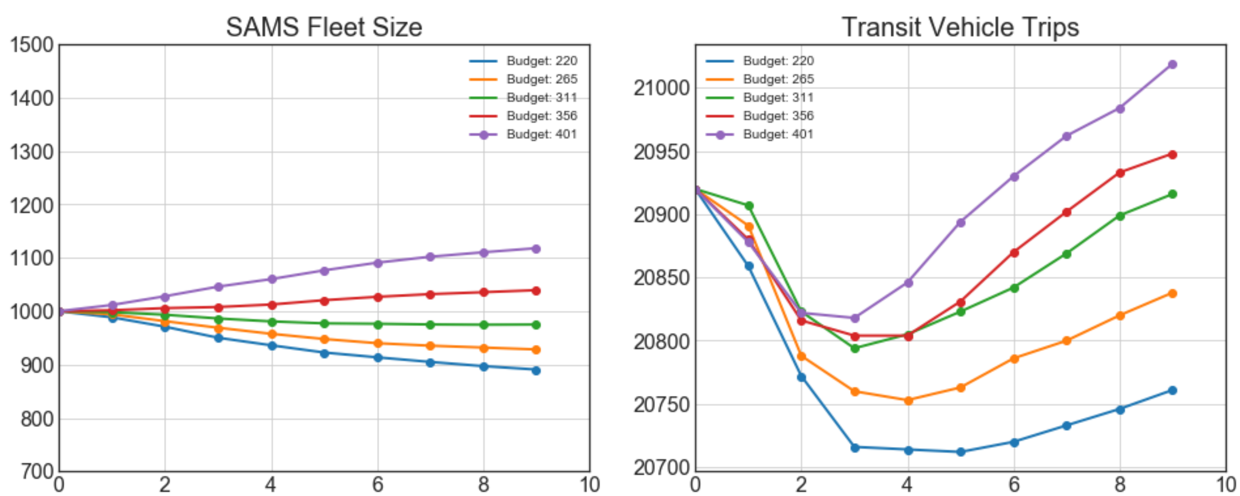

Figure 3.11: Sensitivity of fleet size and transit supply in the upper level

Figure 3.11 shows that the point of convergence of the final SAMS fleet size depends on the budget. Based on the reference, it increases for higher budgets (than U$311,000) and decreases for lower budgets. As for the supplied transit service levels, number of supplied transit trips decrease and then keep increasing, which seems to indicate that higher budgets are allocated preferably to improve transit levels of service rather than SAMS.

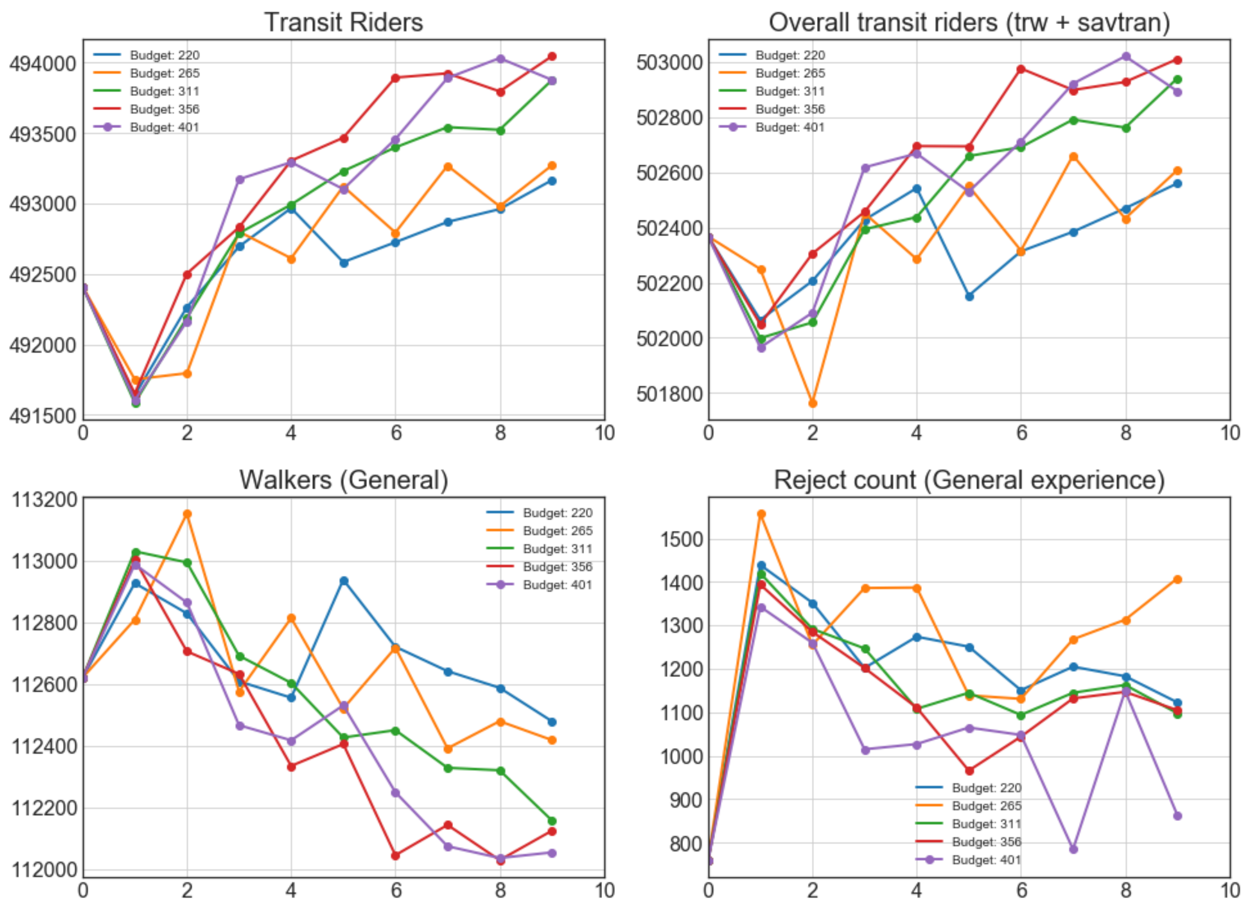

Figure 3.12: Sensitivity of demand in entire Chicago network

Figure 3.12 shows that transit ridership increases across upper iterations, while walkers and rejections decrease. Comparing different budgets, more budget means more transit ridership. This is due to the improved transit level of service.

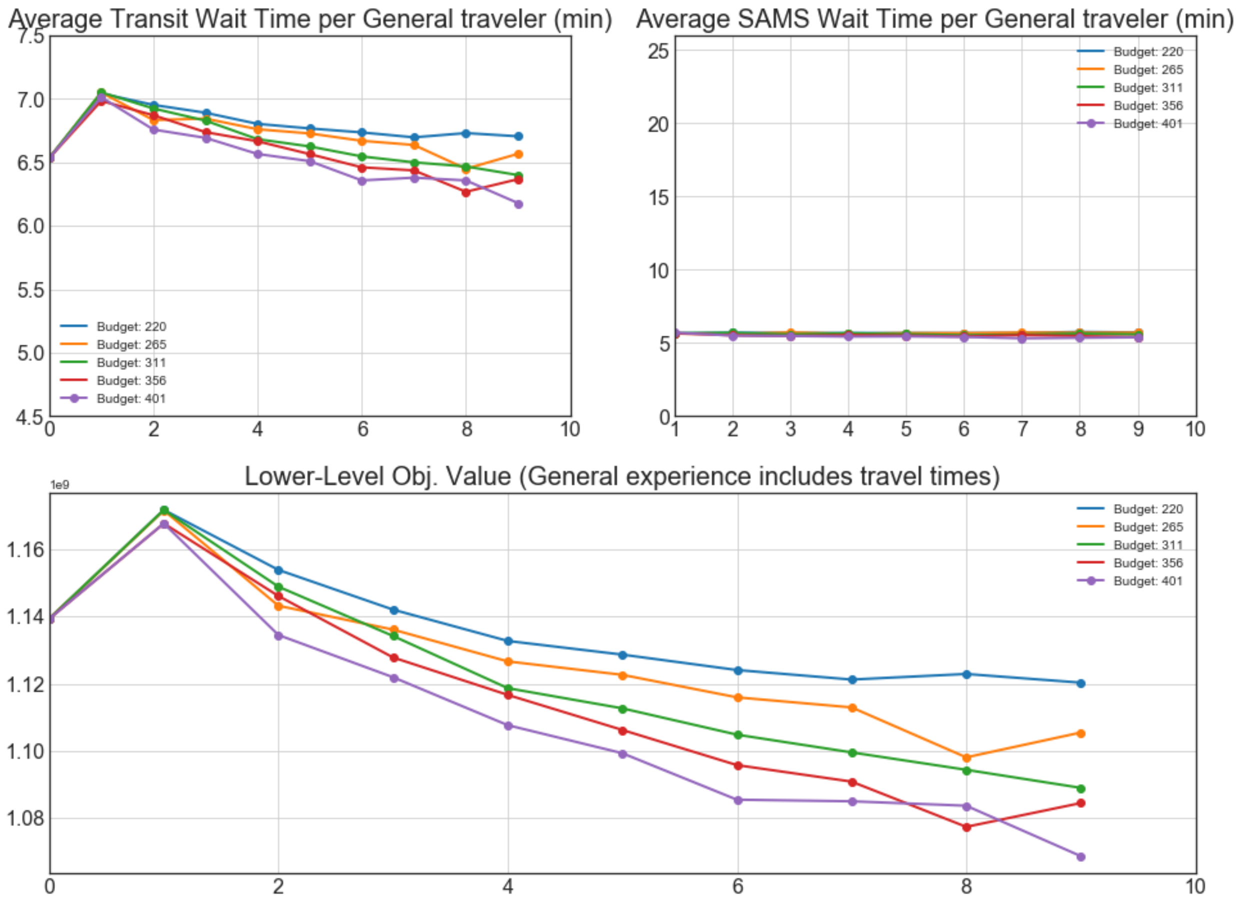

Figure 3.13: Sensitivity of wait times and generalized cost in entire Chicago network

Figure 3.13 shows that SAMS wait times are similar across different budgets. We can also see that, for transit wait times, they all decrease with upper iterations and higher budgets provide lower transit wait times. The bottom plot shows that the generalized cost (including wait times, rejection penalties, in-vehicle travel times and number of transfers) decreases across upper iterations, which shows that the framework not only improves travelers wait times but also improves the overall transportation experience of travelers.

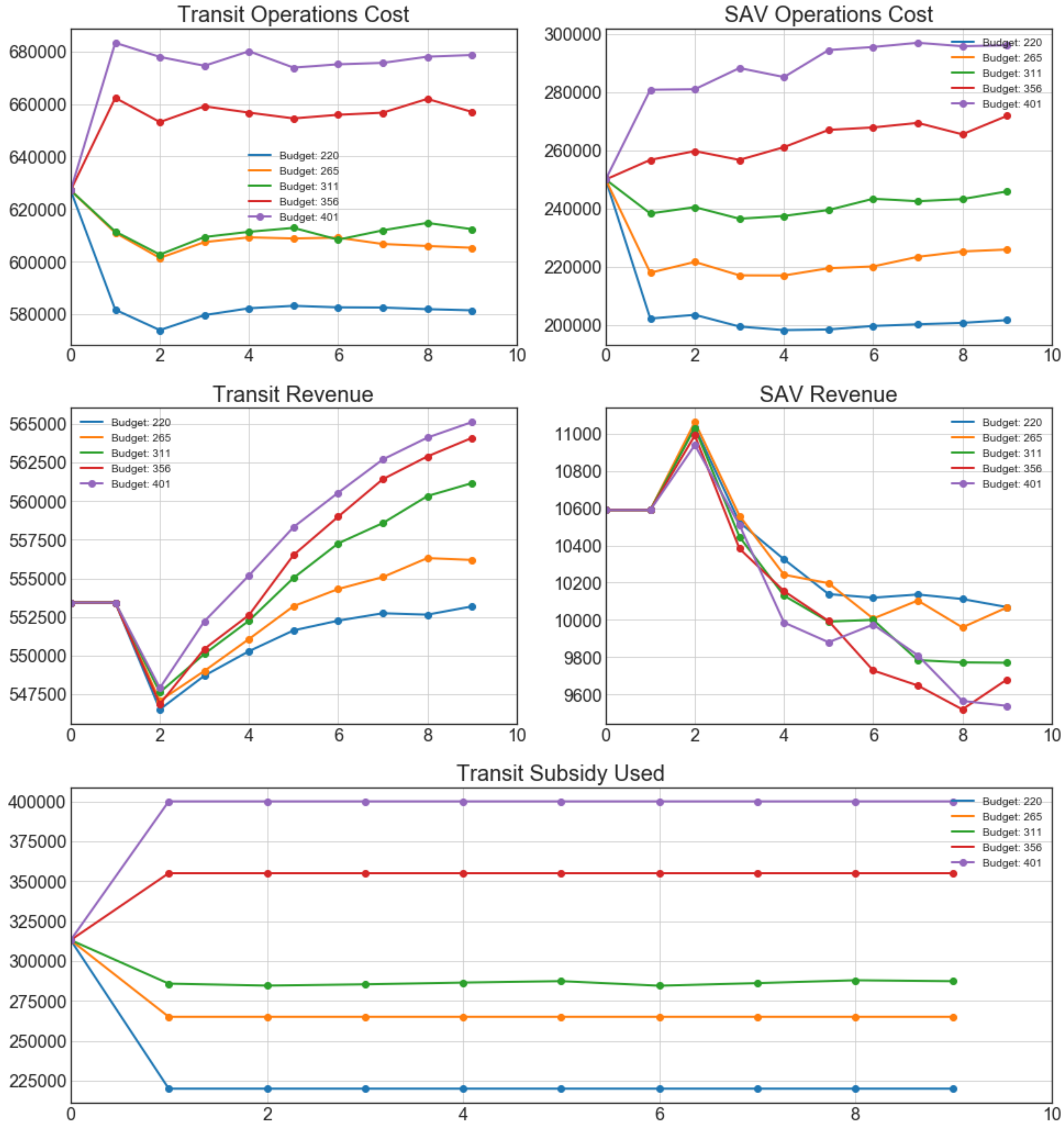

Figure 3.14: Estimated costs, revenues and used subsidies in upper level

Figure 3.14 shows the estimated costs, revenues and used subsidies in each of the budget scenarios across upper-level iterations.

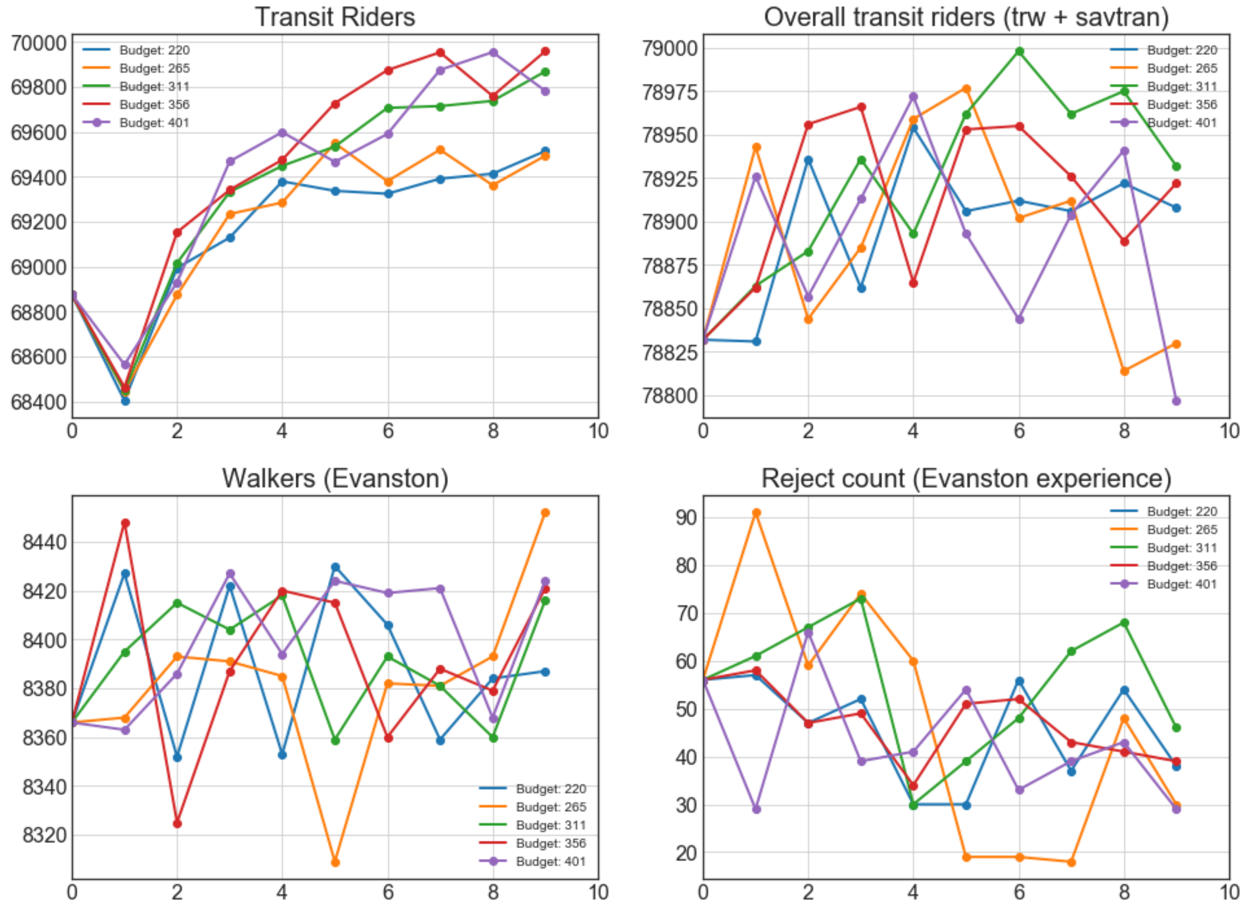

Figure 3.15: Sensitivity of demand in Evanston buffer network

Figure 3.15 shows that transit ridership in Evanston buffer area also increased after the joint design, and we see higher ridership for higher budgets levels. Number of walkers and rejections do not seem to vary significantly. It must also be noted that the results presented for Evanston do not seem as clearly delineated as the general Chicago scenarios because they would require more iterations to reach sharper results in such a small area compared to the overall network. Many changes happen across the entire Chicago network in each of these upper-level iterations.

3.8 Summary

This chapter provided a methodology and modeling framework to support the joint re-design of multimodal transit networks and SAMS fleets to explore and plan mobility scenarios with on-demand, shared and fully-autonomous vehicles. It introduces the joint transit network redesign and SAMS fleet size determination problem (JTNR-SFSDP), along with a heuristic bilevel solution approach demonstrated on an actual large-scale network. The development and effectiveness of the lower-level of this framework are described and illustrated in a stand-alone application in Chapter 4.

4 Assessing the Impact of Shared-use AV Mobility Services on Transit Demand: A Combined Dynamic Mode Choice-Traveler Assignment Modeling Approach

This is a separate application of the lower level alone from the previous chapter. This section introduces an integrated mode choice and dynamic traveler assignment-simulation modeling framework that explicitly models the dynamics of, and congestion in, the transit network and SAMS system. First, we present a mathematical formulation of the dynamic combined mode choice—traveler assignment problem (DCMC-TAP). The problem is analytically intractable because it is a fixed-point problem with interdependencies between the mode choice probabilities and the transit and SAMS system performance, including stochasticity and vehicle capacity constraints; therefore, we present a simulation-based, iterative heuristic solution approach. In the iterative modeling framework, the outer level assigns travelers to one of five modes: car, park-and-ride, transit, SAMS, or transit with SAMS feeder. The inner level, both (1) iteratively determines minimum cost transit hyperpaths, assigns travelers to hyperpaths, and simulates their experiences, and (2) simulates a SAMS fleet providing service to travelers. Time-dependent network performance data is then fed to the mode choice model. This process repeats until the mode choice probabilities converge.

4.1 Mathematical Formulation

This section presents the mathematical formulation of the dynamic combined mode choice—traveler assignment problem (DCMC-TAP). This mathematical equilibrium-based formulation presents a modeling framework to assess the impact of SAMSs on transit and car demand. The mathematical formulation follows the logic presented in (O. Verbas et al., 2016; Zhang et al., 2011).

Let Q, R, M, T, and P represent the sets of origins, destinations, modes, assignment time intervals, and network paths respectively. These sets are indexed by origin $q \in Q$, destination $r \in R$, mode $m \in M$, assignment time interval $t \in T$ and path $p \in P$. Let TR be the set of all travelers and $TR_{q,r}^{t}$ be the set of travelers with origin $q \in Q$, destination $r \in R$, and departure time interval $t \in T$. $\left|TR_{q,r}^{t}\right|$ represents the demand flow originating at $q \in Q$, at departure time interval $t \in T$ and terminating at destination $r \in R$. Let $TR_{q,r}^{t,m}$ be the subset of travelers with origin $q \in Q$, destination $r \in R$, and departure time interval $t \in T$ assigned to mode $m \in M$. Similarly, let $Pr_{q,r}^{t,m}$ be the probability of a traveler with origin $q \in Q$, destination $r \in R$, and departure time interval $t \in T$ choosing mode $m \in M$.

Equation (4.1) displays the mathematical relationship between $Pr_{q,r}^{t,m}$ and $TR_{q,r}^{t,m}$.

$$Pr_{q,r}^{t,m}\left(\left|TR_{q,r}^{t,m'}\right|_{m' \in M}\right) = \frac{\left|TR_{q,r}^{t,m}\right|}{\left|TR_{q,r}^{t}\right|} \qquad \forall q \in Q, r \in R, t \in T, m \in M \tag{4.1}$$

Since the probability $Pr_{q,r}^{t,m}$ of choosing mode $m \in M$ is itself a function of the modal flows $\left|TR_{q,r}^{t,m\prime}\right|_{m\prime \in M}$, the dynamic mode choice problem (DMCP) can be defined as a fixed-point problem with the objective of finding the optimal modal flow $\left|TR_{q,r}^{t,m\prime}\right|_{m\prime \in M}^{*}$ that satisfies the condition in Eqn. (4.2).

$$\left|TR_{q,r}^{t,m}\right|^{*} = \left|TR_{q,r}^{t}\right| \times Pr_{q,r}^{t,m}\left(\left|TR_{q,r}^{t,m'}\right|_{m' \in M}^{*}\right) \qquad \forall q \in Q, r \in R, t \in T, m \in M \tag{4.2}$$

The fixed-point DMCP can be re-formulated as a gap-based nonlinear program (Zhang et al., 2011) in Eqn. (4.3)-(4.5):

$$\min_{TR_{q,r}^{t,m}} GAP_M = \frac{1}{2}\sum_{q \in Q}\sum_{r \in R}\sum_{t \in T}\sum_{m \in M}\left(\left|TR_{q,r}^{t,m}\right| - \left|TR_{q,r}^{t}\right| \times Pr_{q,r}^{t,m}\right) \tag{4.3}$$

such that:

$$\sum_{m \in M}\left|TR_{q,r}^{t,m}\right| = \left|TR_{q,r}^{t}\right| \qquad \forall q \in Q, r \in R, t \in T \tag{4.4}$$

$$\left|TR_{q,r}^{t,m}\right| \geq 0 \qquad \forall q \in Q, r \in R, t \in T, m \in M \tag{4.5}$$

The DMCP objective displayed in Eqn. (4.3) minimizes the discrepancy between the assigned modal flow $\left|TR_{q,r}^{t,m}\right|$ and the expected modal flow $\left|TR_{q,r}^{t}\right| \times Pr_{q,r}^{t,m}$ summed over all origins $q \in Q$, destinations $r \in R$, assignment time intervals $t \in T$, and modes $m \in M$. The convergence of GAP_M to zero satisfies the fixed-point DMCP in Eqn. (4.2). Equation (4.4) is the flow conservation constraint, whereas Eqn. (4.5) satisfies the non-negativity of mode flows.

Solving the simple fixed-point DMCP formulation in Eqn. (4.3)-(4.5) requires knowledge of the origin-destination time-dependent (ODT) modal flows ($\left|TR_{q,r}^{t,m'}\right|_{m' \in M}$) and subsequently the function $Pr_{q,r}^{t,m}\left(\left|TR_{q,r}^{t,m'}\right|_{m' \in M}^{*}\right)$. However, the ODT modal flows are unknown. They depend on the route choices of travelers in the road and transit networks (dynamic traveler assignment problem – DTAP). Without an explicit analytical relation, an agent-based simulation model is needed to obtain the solution. The proposed iterative simulation-based solution approach is described in the next section.

4.2 Solution Approach

This section details the solution approach employed to solve the DCMC-TAP. The next several subsections detail the individual mode choice model, dynamic transit assignment-simulation model, and the SAMS simulation model.

4.2.1 Overview

As mentioned, the DCMC-TAP is analytically intractable due to the interdependencies between the mode choice problem and the DTAP, as well as the underlying complexity of the DTAP as the problem size increases and vehicle capacity constraints are included in the problem formulation. To address the interdependencies, this study employs an iterative solution approach wherein the outer level problem is the mode choice problem, and the inner level problem is the DTAP. To address the complexity and intractability of the DTAP, this study employs a simulation-based solution approach. The solution method includes a SAMS simulation model, and a dynamic transit assignment-simulation model.

To report on the underlying road traffic conditions, a fixed ODT matrix is used with estimated travel metrics for the car mode based on previous dynamic traffic assignment (DTA) simulations. Explicitly integrating a dynamic road traffic assignment model in the solution procedure is out of the scope of this dissertation. However, this should be considered in future work to depict the effect of the SAMS fleet on road congestion.

In the iterative heuristic solution procedure, the outer-level mode choice model determines ODT modal shares given performance metrics for the considered modes. The performance metrics come directly from the SAMS simulation model, the transit assignment-simulation model, and the road network ODT matrix. The ODT modal shares from the outer-level mode choice model are informed to the two inner-level simulation models. The SAMS simulation model solves a dynamic stochastic pickup and delivery problem with immediate demand requests and outputs the ODT performance for the SAMS mode. Concurrently, the inner-level transit assignment-simulation model solves the DTAP and determines the ODT performance for the transit mode.

Based on the prevailing road, transit, and SAMS performance characteristics, the mode choice model solves for new mode choice probabilities $\left(Pr_{q,r}^{t,m}\right)^k$, where k is the iteration number. Then, the number of travelers assigned to mode m at iteration k $\left|TR_{q,r}^{t,m}\right|^k$ is updated according to Eqn. (4.6).

$$\left|TR_{q,r}^{t,m}\right|^k = \left|TR_{q,r}^{t,m}\right|^{k-1} + \alpha_M^k \left(\left|TR_{q,r}^{t}\right| \times \left(Pr_{q,r}^{t,m}\right)^k - \left|TR_{q,r}^{t,m}\right|^{k-1}\right)$$

$$\forall q \in Q, r \in R, t \in T, m \in M \qquad (4.6)$$

where α_M^k is the step size and $\left(\left|TR_{q,r}^{t}\right| \times \left(Pr_{q,r}^{t,m}\right)^k - \left|TR_{q,r}^{t,m}\right|^{k-1}\right)$ is the descent direction.

With the new modal flows $\left|TR_{q,r}^{t,m}\right|^k$, the dynamic assignment in the transit and SAMS networks need to be re-equilibrated – technically the road network too but we assume this is fixed.

After equilibration, the network performance attributes are fed into the outer-level mode choice model.

Figure 4.1 displays a flowchart of the solution approach. At the top level, travelers are assigned to one of five transport modes, based on the performance of the road, transit, and SAMS systems. System performance comes in the form of mode-specific origin-destination-departure time (ODT) performance metrics that include:

- out-of-pocket cost
- in-vehicle travel time
- wait time
- walk time
- in-vehicle standing time (when unable to find a seat)
- number of transit transfers
- SAMS sharing probability

The travelers assigned to transit, park-and-ride, or feeder SAMSs, are moved to the dynamic transit assignment-simulation tool where only the transit portion of the trips are modeled. In parallel, travelers that are assigned to the general SAMS or feeder SAMS modes are moved to the SAMS simulator, where the SAMS portion of their trips is simulated. For consistency, the speed of the SAMS vehicles in the simulation is set to be the same as the passenger car mode, which is an average from all trajectories for each ODT triplet provided by the ODT matrix. The road network performance, and therefore, the ODT metrics for the car mode, and road portion of the park-and-ride mode are fixed.

To obtain the transit ODT performance metrics, we iteratively calculate traveler least weighted generalized cost hyperpaths, assign travelers to hyperpaths, and simulate the movement of vehicles and travelers. This approach is similar to that presented by Verbas (Verbas et al., 2015; Ö. Verbas et al., 2016). Updated costs at the link and node level from the simulation are used to find updated least cost hyperpaths and reassign travelers to their least cost routes on the following iteration. This process continues until the dynamic transit assignment converges or reaches an acceptable number of iterations l.

To obtain the SAMS ODT performance, we use another agent-based simulation tool to model a SAMS fleet operator serving travelers with a fleet of SAMSs. The updated transit and SAMS ODT performance metrics are fed back into the mode choice model. The mode choice model then recalculates mode choice probabilities and re-assigns travelers to one of the five travel modes. This process continues until the mode choice probabilities converge; i.e. minimal change from one iteration to the next, or reach a satisfactory number of iterations.

A mode choice iteration k is run after $l = 4$ inner iterations of the dynamic transit assignment-simulation loop. In every mode choice iteration, using a gap-based approach, a traveler is given a chance to switch modes if there is a decrease in the probability of choosing the same selected mode from one iteration to the next. In iteration k, the probability that traveler $tr \in TR$, who is going from $q \in Q$ to $r \in R$ at departure time interval t, will take the chance to switch from her mode choice n from iteration $(k - 1)$ to another mode is described in Eqn. (4.7).

$$\left(SProb_{tr,n}^{M}\right)^{k} = max\left\{0, \frac{\left(Pr_{q,r}^{t,n}\right)^{k-1} - \left(Pr_{q,r}^{t,n}\right)^{k}}{\left(Pr_{q,r}^{t,n}\right)^{k-1}}\right\} \quad \forall\, tr \in TR, m \in M, n \in M \tag{4.7}$$

The probability that the traveler will switch from mode n to mode $m \in M$ in iteration k is described in Eqn. (4.8).

$$\left(Prob_{tr,n}^{m}\right)^{k} = \left(SProb_{tr,n}^{M}\right)^{k}\left(Pr_{q,r}^{t,m}\right)^{k} \quad \forall\, tr \in TR, m \in M, n \in M \tag{4.8}$$

$\left(Pr_{q,r}^{t,m}\right)^{k}$ is the probability of choosing $m \in M$ computed in the multinomial logit form by the mode choice model based on the utility function of the modes.

The following subsections detail the mode choice model, the dynamic transit assignment-simulation model, and the SAMS simulation model. The following section discusses integration challenges.

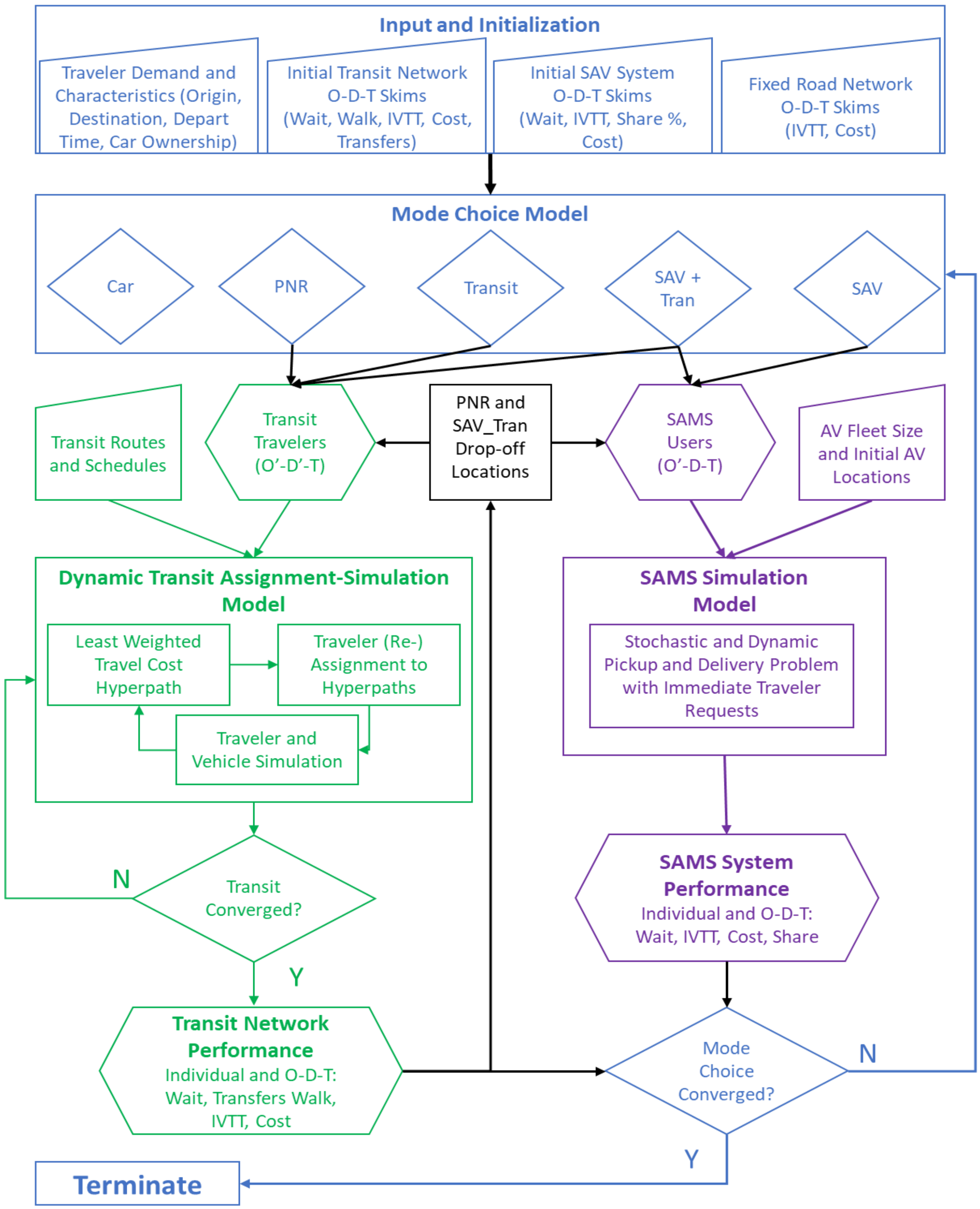

Figure 4.1: Solution Approach for the dynamic combined mode choice and traveler assignment simulation

4.2.2 Mode Choice

Mode Choice Input

The input to the mode choice model includes (1) a list of all traveler-trips (travelers for short) and their respective characteristics, and (2) ODT performance metrics for all modes. Traveler characteristics include origin, destination, departure time and car ownership.

Mode Choice Model

This study employs a multinomial logit (MNL) discrete choice model. The modeling framework can handle more complex choice models including nested logit and cross-nested logit. However, reliable data was not available to estimate such models because the SAMS modes had not been implemented in real life yet at the time of this study.

The mode choice model assigns travelers to one of the modes. The functional form of the utility function is given in Eqn. (4.9), where b is the parameter index and m is the mode. Wait time is the only modal specific coefficient in the model. The coefficient estimates listed below were obtained from a variety of sources in the literature including a stated-preference survey on flexible demand-adaptive transit (Frei et al., 2017), as well as SAMSs (Krueger et al., 2016). The ASC values for car, park-and-ride, transit, SAMS, and transit with SAMS are 0, -1.5, -1.7, -0.5, and -1.5, respectively. The performance variables $IVTT_m$, $Stand_m$, $Wait_m$, and $Walk_m$ should be in minutes. The fare variable ($Fare_m$) should be in US dollars; the $Share \in [0,1]$ variable is the probability of sharing the SAMS ride with another traveler, the $Transfers$ variable counts number of transfers along the journey.

$$V_m = ASC_m + \beta_{b,m} X_{b,m} \tag{4.9}$$

$$\begin{aligned} V_m = ASC_m &+ \beta_{IVTT}(IVTT_m - Stand_m) + \beta_{wait,SAV} Wait_{SAV} + \ \beta_{wait,tran} Wait_{Tran} \\ &+ \ \beta_{walk} Walk_m + \beta_{fare} Fare_m + \beta_{Share} Share + \beta_{Transfer} Transfers \\ &+ \beta_{Stand} Stand \end{aligned}$$

$$\begin{aligned} V_m = ASC_m &+ (-0.12)(IVTT_m - Stand_m) + (-0.12) Wait_{SAV} + \ (-0.24) Wait_{Tran} \\ &+ (-0.18) Walk_m + (-0.6) Fare_m + (-0.5) Share + (-0.5) Transfers \\ &+ \ (-0.24) Stand + (-0.24) Stand \end{aligned}$$

In multinomial logit models, the probability of a traveler with origin $q \in Q$, destination $r \in R$, and departure time interval $t \in T$ being assigned to mode $m' \in M$ is given in Eqn. (4.10).

$$Pr_{q,r}^{t,m'} = \frac{exp\ (V_{m'})}{\sum_m exp\ (V_m)} \tag{4.10}$$

At the mode choice level, the mode characteristics are evaluated. Road network performance (travel time and distance for passenger car mode and driving portion of the Park-and-Ride mode) is fixed. It is provided as an average for every OD and departure time interval based on previous DTA simulations with DYNASMART-P. Walking and transit experiences are obtained from the least-cost hyperpath calculation, assignment and simulation at the lower level.

At the mode choice level, we learn each traveler's ODT, estimate their travel characteristics for every one of the five potential modes, and build their mode choice set according to their personal or travel characteristics. The passenger car mode is feasible for those whose household own a car and whose estimated car travel time is no more than 24h (1440 min). The feeder SAMS and Park-and-Ride modes are feasible for travelers whose origin is within a 3-mile radius of a transit service that reaches the traveler's destination MAZ in no more than 6 hours. However, the feeder SAMS is only allowed for those whose origin is in Evanston, whereas Park-and-Ride is only feasible for those whose household own a car and whose origin is not in Evanston. If a traveler's origin and destination are both located in Evanston, she may take the general SAMS mode. Lastly, if the traveler's origin is in a MAZ served by transit, and the estimated transit travel time is no more than 6 hours, then transit is in their mode choice set.

4.2.3 Dynamic Transit Assignment-Simulation

A dynamic transit assignment-simulation tool applicable to large-scale multimodal networks, NU-TRANS (O. Verbas et al., 2016), is adapted and used to assess the experience of travelers at user equilibrium. The multinomial logit mode choice model described in the previous section was added to NU-TRANS. NU-TRANS comprises the following features:

- Multimodal least-weighted travel time hyperpath calculation
- Multimodal time-dependent assignment
- Multi-agent particle simulation
- Mode choice model

The least-cost hyperpath calculation and assignment use a gap-based formulation applied to a multimodal transit network including bus, rail and walk modes. The provided hyperpaths are time-dependent on a frequency-based network. They capture different service patterns of a transit route and consider seating and vehicle capacity constraints. The multi-agent particle simulation moves vehicles and travelers second-by-second, capturing the experience at the agent level, including transfers or missed connections, seating or standing, boardings and rejections.

A procedure to update the traveler's drop-off transit MAZ was added to the tool. After every $l = 4$ iterations in the lower level, the travelers' drop-off MAZs are updated with the MAZ within a 3-mile radius of their origin that provides the least generalized cost by transit, based on the most recent hyperpath calculation. The updated drop-off MAZ becomes the origin of the transit portion of their trip if the Park-and-Ride or feeder SAMS modes are selected.

NUTRANS Input

The necessary input for NUTRANS includes (1) time-dependent agent-based trip demand (ODT matrix); (2) traveler characteristics (value of time and car ownership); and (3) transit network (i.e. routes) and schedule of transit vehicles, which are obtained from the General Transit Feed Specification data.

4.2.4 SAMS Fleet Simulator

SAMS Simulation Input

The input to the SAMS simulation model includes (1) the SAMS fleet size, (2) the initial location of all SAMSs in the coverage area and (3) the origin $q \in Q$, destination $r \in R$, and departure time $t \in T$ of every traveler $i \in TR^{SAV}$ assigned to general SAMS or feeder SAMS.

SAMS Simulation Model

The SAMS simulation model is an agent-based micro-simulation tool that models the movements of travelers and SAMSs, and the operational decisions of a SAMS fleet operator. This study assumes the SAMS fleet operator has complete control over all the SAMSs in the fleet. However, the fleet operator has no a priori information about the location and time of traveler requests. Hence, the SAMS fleet operator must dynamically assign SAMSs to traveler requests. The SAMS fleet operator aims to minimize traveler wait times. This study assumes that no more than two traveler requests can be inside a SAMS at one time.

The underlying problem for the SAMS fleet operator involves using a fleet of SAMSs $V = \{1,2, \dots j, \dots |V|\}$ to serve travelers $TR^{SAMS} = \{1,2, \dots i, \dots |TR^{SAMS}|\}$ who make immediate/on-demand requests over the finite horizon $T = [0, |T|]$. The requests arrive randomly according to a Poisson process with a spatial-temporal demand rate that is unknown to the SAMS fleet operator. Each traveler's origin ($q_i \in Q$) and destination ($r_i \in R$) can be located anywhere within a pre-defined service region. Each traveler's request time ($t_i \in T$) also denotes her earliest pickup time, her desired pickup time, and the time the SAMS fleet operator becomes aware of the request.

Each SAMS traveler $i \in TR^{SAMS}$ can be in one of four mutually exclusive states: unassigned, assigned, in-vehicle, or served. Unassigned travelers TR_U^{SAMS} have made a request but have not yet been assigned to an AV. Assigned customers TR_A^{SAMS} have been assigned to an AV but have not yet been picked up. In-vehicle customers TR_{IV}^{SAMS}have been picked up and are en-route to their destination. Finally, served customers TR_S^{SAMS}have been dropped off at their destination. At every time instant $\tau > t_i$, unassigned and assigned customers have an elapsed wait time w_i, wherein $w_i = \tau - t_i$. Additionally, let $posTR_i$ denote the current position of traveler i.

At any point in time $\tau \geq 0$, each AV j has a physical location $posV_j$. AV $j \in V$ can be in one of three states: idle, en-route to pick up a traveler, and en-route to drop off a traveler, respectively. These three states correspond to three mutually exclusive and collectively exhaustive subsets of AVs; idle AVs V_I, en-route pickup AVs V_P, and en-route drop-off AVs V_D.

Let ϱ_{ij} equal 1 if AV $j \in V_D$ can feasibly pickup traveler $i \in C_U$, and 0 otherwise. For any $i \in C_U$, $j \in V_D$ pair, ϱ_{ij} equals 1, if $j \in V_D$ is 20% closer than all idle AVs V_I, and the assignment does not cause traveler $i \in C_U$, or the passenger in SAMS j (i'), to ever backtrack (i.e. move in a direction opposite of their destinations). Let $d(pt_1, pt_2)$ denotes the distance between points pt_1 and pt_2, and let $0 < \alpha < 1$. Using the notation in Figure 4.2, the back-tracking conditions are formalized as follows:

- for all $i \in C_U$:
    - for all $j \in V_D$:
        - $\varrho_{ij} = 0$
        - If $d(o_{i'}, D_i) < d(posTR_i, D_i)$ and $d(posTR_i, o_{i'}) + d(o_{i'}, D_i) < (1 + \alpha) \times d(posTR_i, D_i)$:
            - if $d(o_{i'}, D_i) < d(o_{i'}, D_{i'})$:
                - if $d(D_i, D_{i'}) < d(o_{i'}, D_{i'})$ and $d(o_{i'}, D_i) + d(D_i, D_{i'}) < (1 + \alpha) \times d(o_{i'}, D_{i'})$:
                    - $\varrho_{ij} = 1$
            - else:
                - if $d(D_i, D_{i'}) < d(o_{i'}, D_i)$ and $d(o_{i'}, D_{i'}) + d(D_i, D_{i'}) < (1 + \alpha) \times d(o_{i'}, D_i)$:
                    - $\varrho_{ij} = 1$

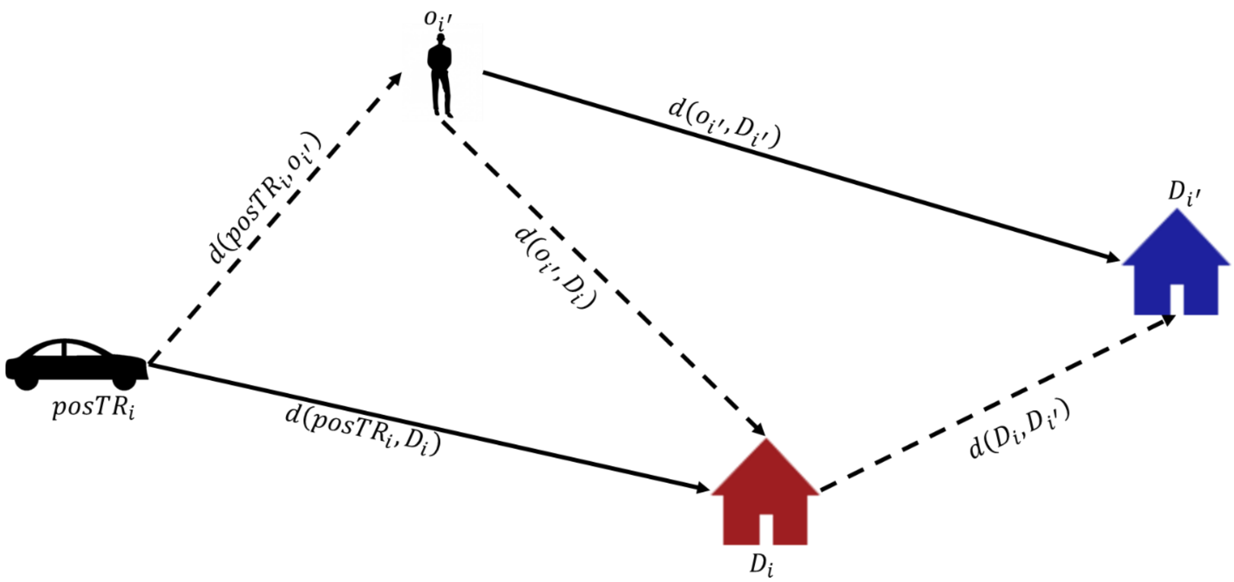

Figure 4.2: Diagram for determining feasible AV-traveler assignments

The mathematical programming formulation of the AV-traveler assignment problem with shared rides, for the case where the number of unassigned travelers is greater than the number of idle AVs $|C_U| > |V_I|$) is presented in Eqn. (4.11)-(4.16). The parameter γ converts the units of time, to units of distance. The parameter φ denotes the penalty for assigning a traveler to an en-route drop-off SAMS.

$$\min_{x_{ij}} \sum_{i \in C_U} \sum_{j \in V_I \cup V_D} \left(d_{ij}x_{ij} - \gamma w_i x_{ij}\right) + \sum_{i \in C_U} \sum_{j \in V_D} \left(\varphi x_{ij}\right) \tag{4.11}$$

subject to:

$$x_{ij} - \varrho_{ij} \leq 0 \qquad \forall i \in C_U, \forall j \in V_D \tag{4.12}$$

$$\sum_{i \in C_U} \sum_{j \in V_I \cup V_D} x_{ij} \geq \sum_{j \in V_I} 1 \qquad \forall q \in Q, r \in R, t \in T, m \in M \tag{4.13}$$

$$\sum_{j \in V_I \cup V_D} x_{ij} \leq 1 \qquad \forall i \in C_U \tag{4.14}$$

$$\sum_{i \in C_U} x_{ij} \leq 1 \qquad \forall j \in V_I \cup V_D \tag{4.15}$$

$$x_{ij} \in \{0,1\} \qquad \forall i \in C_U, \forall j \in V_I \cup V_D \tag{4.16}$$

The first term in the objective function in Eqn. (4.11) determines the distance between each assigned AV-customer pair. The second term in the objective function, rewards the SAMS fleet operator for assigning SAMSs to travelers that have a large elapsed wait time. The third term in the objective function slightly penalizes the SAMS fleet operator for assigning en-route drop-off AVs to travelers.

The constraint in Eqn. (4.12) ensures that an inefficient assignment of an en-route drop-off AV ($j \in V_D$) to a traveler request does not occur. The constraint in Eqn. (4.13) requires the number of assigned AVs to be greater than or equal to the number of idle AVs in the system at the current time. The two constraints in Eqn. (4.14) and Eqn. (4.15) require travelers to be assigned to at most one AV, and AVs to be assigned to at most one customer, respectively. The last constraint in Eqn. (4.16) ensures that the decision variable is binary. Fortunately, because the constraint matrix is unimodular, the LP-relaxation of the IP in Eqn. (4.11)-(4.16) returns integer solutions. Hence, this formulation is computationally very efficient.

For the case where the number of unassigned customers is less than the number of idle AVs ($|C_U| \leq |V_I|$), the inequality constraint in Eqn. (4.14) becomes an equality constraint; i.e. all travelers must be assigned to an AV. Additionally, the constraint in Eqn. (4.13) is removed.

SAMS Simulation Model Output

The output of the SAMS simulation model includes average wait time, in-vehicle travel time, travel distance, and probability of sharing an SAMS for every ODT. This information is then immediately fed back into the mode choice model. There is nothing to equilibrate within the SAMS simulation model. However, it is important to note that the SAMS simulation model captures congestion in the SAMS system. Given that the SAMS fleet size is fixed, if the mode choice model assigns more and more travelers to either SAMS mode, traveler wait times will increase significantly. Hence, the next mode choice iteration will assign fewer travelers to the SAMS modes.

4.3 Integration challenges and limitations

A challenge found in the process of integrating the models lies on selecting the feeder SAMS and park-and-ride drop-off points. A user who has an origin in a microanalysis zone (MAZ) that does not have transit service may decide to request a SAMS (feeder) to connect to transit; in such case, she will need to pick a drop-off point that is convenient for the transfer, not necessarily the closest transit station, depending on the transit schedule and her destination of interest. Here the origin of the transit portion of the trip for park-and-ride and feeder SAMS travelers is selected individually based on the user's ODT. The user's drop-off point is set to be the MAZ served by transit and within a 3-mile radius of the origin, that provides the “shortest path” to the destination MAZ at the traveler’s time of interest. This “shortest path” is the least cost hyperpath obtained from the most recent transit assignment-simulation iteration in NU-TRANS, where the cost is the perceived generalized cost, including in-vehicle travel and wait times, boarding rejections and seat availability.

As previously noted, the integration involves three separate programs for the dynamic transit assignment-simulation (NU-TRANS), the SAMS fleet simulation and the DTA simulation (DYNASMART-P) to provide performance metrics of the transit, SAMS and passenger car modes. However, the DTA output currently serves only to give a one-time fixed input of the passenger car performance to the framework’s main loop. Furthermore, even though buses, SAMSs and passenger cars share the same road network, our integration does not yet reflect the traffic conditions caused by their intermodal interaction in lanes with mixed use. This can be done in future work by assuring that the DTA simulation (i) belongs to a parallel feedback loop such that the road network performance gets updated with the most recent passenger car mode shares and (ii) includes the buses and assigned SAMSs to capture their interaction and effect on the passenger car performance. To improve consistency between the SAMS system and the passenger car mode in terms of performance in the road network, we provide the SAMS fleet simulator with a priori location- and time-dependent (ODT) speeds from the DTA simulation with passenger cars.

On the transit assignment and simulation platform, walking is allowed on transit links so that travelers can have flexibility to take transit from a different start point or to make transfers that may require walking. It is therefore possible that the traveler completes the trip by only walking even though her assigned mode is “transit”. This is a weakness of the framework because at the mode choice level the traveler is considered to be charged a transit fare despite the fact that she may eventually be assigned to a walking-only path.

Origins and destinations are modeled in terms of microanalysis zones (MAZs) in NU-TRANS and SAMS fleet simulator. This improves the accuracy of transit access and egress modeling in the traveler path choices due to the fine level of spatial resolution. Trips are assumed to originate or end at the MAZ centroid. Since the road DTA performance is obtained from a mesoscopic analysis, it is provided in a traffic analysis zone (TAZ) to superzone level. Hence the traffic conditions do not reflect the same fine resolution.

4.4 Application to a large-scale network

This section describes a large-scale implementation of the integrated modeling framework presented in the previous section. The modeling framework was tested on the region served by the Chicago Transit Authority (CTA). The study area comprises 16,819 MAZs; 1348 of which are served by transit. The transit network represents services provided by two agencies: Chicago Transit Authority (buses and heavy rail operator) and Metra (suburban commuter rail operator). The network includes 14,259 nodes, 64,083 links and 1081 transit service patterns[3].

To examine the impact of a suburban SAMS transit feeder service (composed of feeder SAMSs and general SAMSs), the simulation is performed for the entire metro area for different network scenarios, then observing the user experience in the municipality of Evanston, north suburb of

[3] A service pattern is a subset of stops served from the entire stop set of a route (Verbas and Mahmassani, 2013).

Chicago. Evanston has four CTA bus lines, six heavy rail stations, and three commuter rail stations; hence, transit accessibility, especially to the urban core, is quite high. In the mode choice framework, only trips originating in Evanston can use the feeder SAMSs; moreover, only trips with an origin and destination in Evanston can use the general SAMS mode. These assumptions ensure that the SAMSs mostly stay in the Evanston area.

The authors obtained a 25% sample of traveler trips generated in the Chicago metro area on a typical weekday. Each trip was replicated four times to represent 100% of all trips in the Chicago metropolitan area. For the current study, trips with destinations outside of the CTA operating region and trips within the same zone were removed. A total of 2,136,876 trips are simulated, of which 35,756 have the start point in Evanston. The demand data is synthetized based on the 2007 Household Travel Survey conducted by the Chicago Metropolitan Agency for Planning (CMAP) and results from previous work with an activity-based and dynamic traffic assignment (ABM-DTA) integration model for the Chicago metropolitan region (Halat et al., 2017; Xu et al., 2017; Zockaie et al., 2015). The observed fields for each traveler trip are person ID, start time, car availability, value of time, origin and destination in terms of MAZs, and travel mode.

Passenger car travel characteristics (toll, time and distance) are obtained from a one-time DYNASMART simulation. Park-and-ride facilities are assumed to be available in all transit MAZs. Transit travel characteristics such as vehicle schedules and station locations are obtained from the General Transit Feed Specification (GTFS). Additionally, Geographic Information System (GIS) shapefiles obtained from the City of Chicago Data website are used to identify MAZs located within the area of interest.

Trips are observed during the morning period between 6 AM and 10 AM in three transit network scenarios: (1) the current transit network in the Chicago Metropolitan area, comprising services operated by the Chicago Transit Authority and Metra; (2) the addition of a SAMS fleet to the current transit network; and the replacement of the existing CTA bus service in Evanston for a SAMS fleet. Table 5 describes the simulation scenarios.

Table 5: Characteristics of the scenarios

Scenario	Description	Modes	Fare type
1	Current network	Private automobile, Transit (bus, rail and walking)*	Transit: flat fee Private car: toll + distance-based Park-and-Ride: toll + distance-based + parking ticket
2	Current network + SAMS	Private automobile, Transit (bus, rail and walking), general SAMS, feeder SAMS *	Transit: flat fee Private car: toll + distance-based Park-and-Ride: toll + distance-based + parking ticket General SAMS: flat fee + distance-based + duration-based, or minimum fare Feeder SAMS: 1.5 x transit fare
3	Current network + SAMS, without buses	Private automobile, Transit (rail and walking), general SAMS, feeder SAMS *	Transit: flat fee Private car: toll + distance-based General SAMS: flat fee + distance-based + duration-based, or minimum fare Feeder SAMS: 1.5 x transit fare

*The simulation included the Park-and-Ride mode only for the Chicago area outside Evanston so that it would not overlap with SAMS.

4.4.1 Pricing

In this section, the fare types briefly introduced in Table 5 (user out-of-pocket costs) are explained in more details. The fare is one of the many parameters used by the mode choice model in the calculation of the travel mode's disutility.

The transit mode, including both bus, rail and walking, has a fixed fare (\$1.40 or \$2.25). We do not differentiate transit pricing in terms of transit service type (local bus, express bus, different commuter rail destination zones, etc). Transit riders are charged a fixed amount regardless of how many transfers or different service types existed in their trip.

If a traveler is assigned to the general SAMS mode, a minimum fare ($SAV_{min} = \$1.70$) is applied based on current prevailing transportation network companies' (TNC) pricing practices. The SAMS fare is then estimated from time-based, mileage-based and fixed service fees as shown in Eqns. (4.17)-(4.18).

$$Fare_{SAV} = (SAV_{min}, 0.7 * Trip\ Cost_{SAV}) \tag{4.17}$$

$$Trip\ Cost_{SAV} = minute\ rate * IVTT_{SAV} + mileage\ rate * Dist_{SAV} + service\ fee \tag{4.18}$$

The price of the general SAMS ride is a function of the in-vehicle travel time ($IVTT_{SAV}$) and travel distance ($Dist_{SAV}$). Travelers in this mode are not given the option of deciding whether their ride will be shared. We assume all passengers are requesting a shareable service, so that they will be paying 70% of a full ride cost even if their ride is not matched with another passenger. The parameters used in this calculation are average values for the passengers' ODT given by the SAMS fleet simulator.

The passenger car mode and driving portion of the Park-and-Ride mode have their out-of-pocket costs estimated based on user-specific data (car ownership and value of time[4]) and fixed road network performance metrics. At a time-dependent OD level, each trip is associated with a toll, travel time and distance based on its value of time (DYNASMART DTA output). The DTA tool detects value of time breakpoints for each ODT to assign travelers to certain paths. The toll associated to the trip refers to the existing tolls, if applicable, in the assigned path. The cost associated with the passenger car and Park-And-Ride (PNR) mode is calculated in Eqns. (4.19)-(4.20). The park ticket is \$1.00 and the mileage fee is \$0.95 per mile.

$$Cost_{PNR} = Toll_{CAR} + mileage_{fee} * Dist_{PNR} + Fare_{Transit} + Park\ Ticket \tag{4.19}$$

$$Cost_{CAR} = Toll_{CAR} + mileage_{fee} * Dist_{CAR} \tag{4.20}$$

The feeder SAMS mode is set to allow a similar pricing structure to the trip cost of the general SAMS ($Trip\ Cost_{SAV}$) with time and mileage-based as well as a constant fee. However, in this case study, we only considered a fixed fare for this mode. SAMS (feeder) riders who are dropped off at a transit station are charged a fixed fare equivalent to 1.5 times the prevalent transit fare, which covers both the SAMS and transit portions of the trip. If the traveler opts to walk to the destination after the SAMS drop-off at the transit station, they are still charged the same. This

[4] The user's value of time was assigned by the ABM model in a previous ABM-DTA integration work (Zockaie et al., 2015) based on household income and assumed the same for all household members. There a continuous VOT distribution was estimated for each of four different user groups based on income level (0-30k, 30-60, 60-100k, >100k). Each traveler trip was then assigned a different value of time following the distribution of the user group it belongs to.

pricing strategy reflects a potential policy to encourage transit use and increase chances of matching travelers to shared rides.

4.4.2 Replacing a Bus Fleet with a SAMS fleet

Fleet size is probably the most important factor impacting the service quality of a SAMS feeder mode. Estimating how many SAMSs will be in the Evanston suburban feeder system is a hard task. However, to get some estimate of the SAMS fleet size, we assume that the bus purchasing costs can be used to purchase SAMSs, and that the operating life of a SAMS will be the same as a current transit bus.

The Chicago Transit Authority operates four bus routes in the city of Evanston. Assuming there are six buses running in each direction during the morning peak for each route, the number of buses on all four routes is 48. According to Czerwinski et al., (2016), the cost of a bus is between $400,000 and $700,000; we assume $500,000. Future AV purchasing costs are unknown, but after a few years on the market they should not cost more than $100,000 for a regular-sized sedan. Considering $80,000 for a vehicle, the SAMS fleet size is assumed to be 300 AVs.

4.5 Results

This section displays the computational results for the simulation-based, iterative heuristic solution approach to the DCMC-TAP. The results illustrate the gap convergence, mode shares and traveler experience for the three scenarios described in Table 5.

4.5.1 GAP Convergence

Figure 4.3 and **Figure 4.4** illustrate the convergence of the integrated framework in all scenarios for the Evanston area.

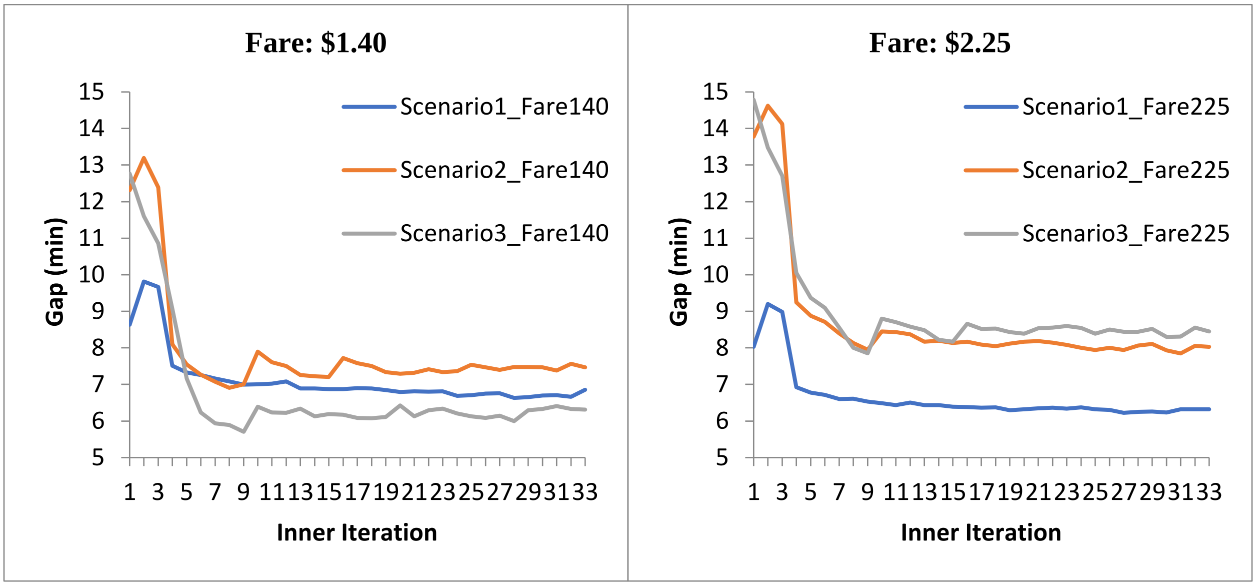

Figure 4.3: Lower level - Average transit assignment gap per traveler

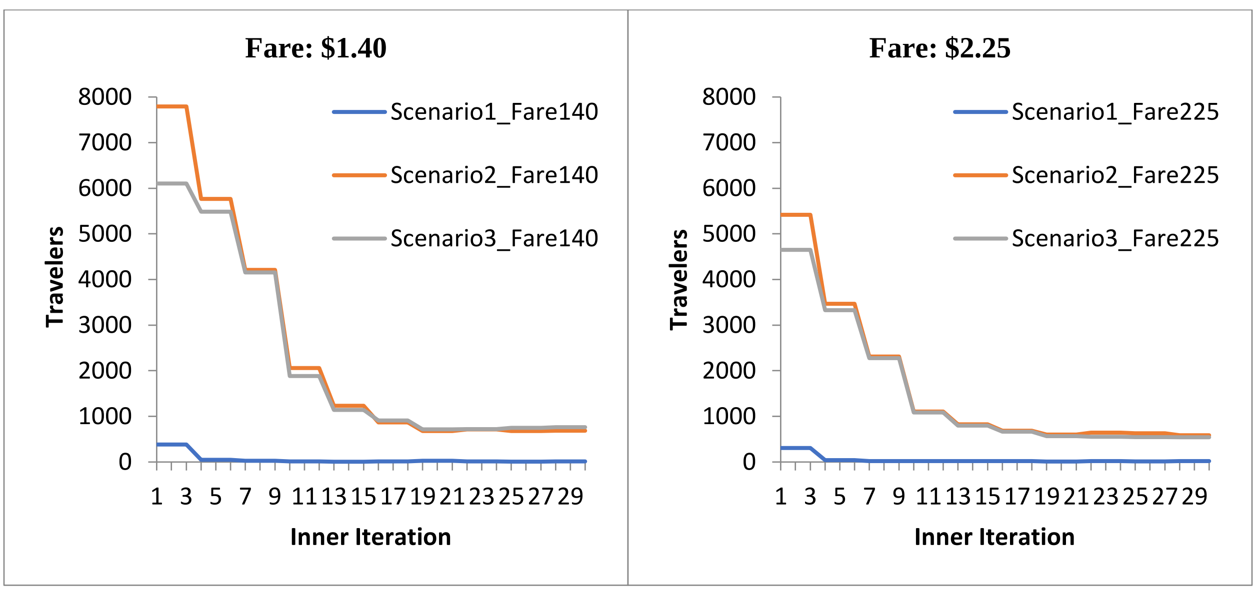

Figure 4.4: Upper level - Mode choice gap

The outer gap calculation in the mode choice level is based on the sum of differences of mode choice flows between outer iterations. Figure 4.4 shows that the mode choice gap has an overall reduction with the iterations for all scenarios.

4.5.2 Evanston Mode Share

The modal share obtained in the last iteration of the scenarios in Evanston can be seen in Table 6. For a more detailed specification of the mode shares, we segment the trips with feeder SAMS mode in two groups based on the traveler's final itinerary. If the traveler decided to walk after the SAMS drop-off, the trip mode is called "SAMS + Walk'. Otherwise, the trip mode is "SAMS + Transit".

As a representation of the current situation in Evanston (scenario 1), our simulation predicts 69.9% (70.9%) mode share for private car, considering a transit fare of $1.40 ($2.25 for values in parentheses). This number reduces by 13.3% (13.0%) after the addition of SAMSs (scenario 2). After removing the local bus lines (scenario 3), the private car share increases slightly compared to scenario 2, but it is 11.2% (11.2%) lower than the base case (scenario 1). As expected, the transit share gets impacted, with a reduction of 6.6% (7.2) with the SAMSs (scenario 2). The removal of buses does not have a significant impact on the total transit share between the scenarios with SAMSs (scenarios 2 and 3). Another notable change is in the walking mode share, which reduces by 8.1 (9.8%) after the addition of SAMSs. Finally, general SAMS and feeder SAMS modes seem much more desirable alternatives than transit alone, even if the traveler decides to walk to their destination after the SAMS drop-off at a transit station.

According to a Chicago Metropolitan Agency for Planning analysis of the 2008 household travel survey data on travel to work for Cook county residents, 73% of work trips were done with a private car (driver or passenger), 21% were completed with transit and 5.6% by walking or biking. Five-year estimates based on the 2006-2010 American Community Survey of mode travel to work with Evanston commuters presented 64.6% of the trips with passenger car (driver or passenger), 22.2% with transit and 13.2% walk.

Table 6: Mode Share in Evanston (%)

	Fare	**1.40**			**2.25**			**2010 American Community Survey**
	Scenario	**1**	**2**	**3**	**1**	**2**	**3**	
Mode	Incomplete Trips	0.1	0.0	0.0	0.1	0.0	0.0	
	Car	69.9	56.6	58.7	72.0	57.9	59.7	64.6*
	Transit	18.0	5.1	3.6	16.8	4.9	3.3	22.2*
	Walking	11.9	3.8	1.8	11.1	2.0	1.1	13.2*
	General SAMS	-	22.2	24.2	-	26.4	26.5	-
	SAMS + Transit	-	6.3	6.8	-	5.1	5.9	-
	SAMS + Walk	-	6.1	5.0	-	3.7	3.4	-
	Transit Total	**18.0**	**11.4**	**10.4**	**16.8**	**10.0**	**9.2**	-
	Transit Change		**-6.6**	**-7.6**	**-**	**-7.2**	**-8.0**	-

* values adjusted after removing work from home responses and other minor modes

4.5.3 Transit Traveler Experience

The average generalized travel cost of transit riders, including those who arrive in the transit station by car or feeder SAMS, is shown to converge in all scenarios in Figure 4.5. Depending on the prevailing transit fare, the scenarios behave differently. The addition of SAMSs to the current scenario did not induce a meaningful change in the average generalized travel cost with the transit fare as $1.40 (40 minutes for all three scenarios). However, with a $2.25 fare, this variable increased from 39 (scenario 1) to 46 minutes (scenarios 2 and 3).

These results are better understood when matched with the mode shares previously shown in Table 6. The cost increase from after the addition of SAMSs with the higher transit fare can be explained by more people in the private car mode (greater loss in car mode compared to lower fare), as well as lower loss in walking mode.

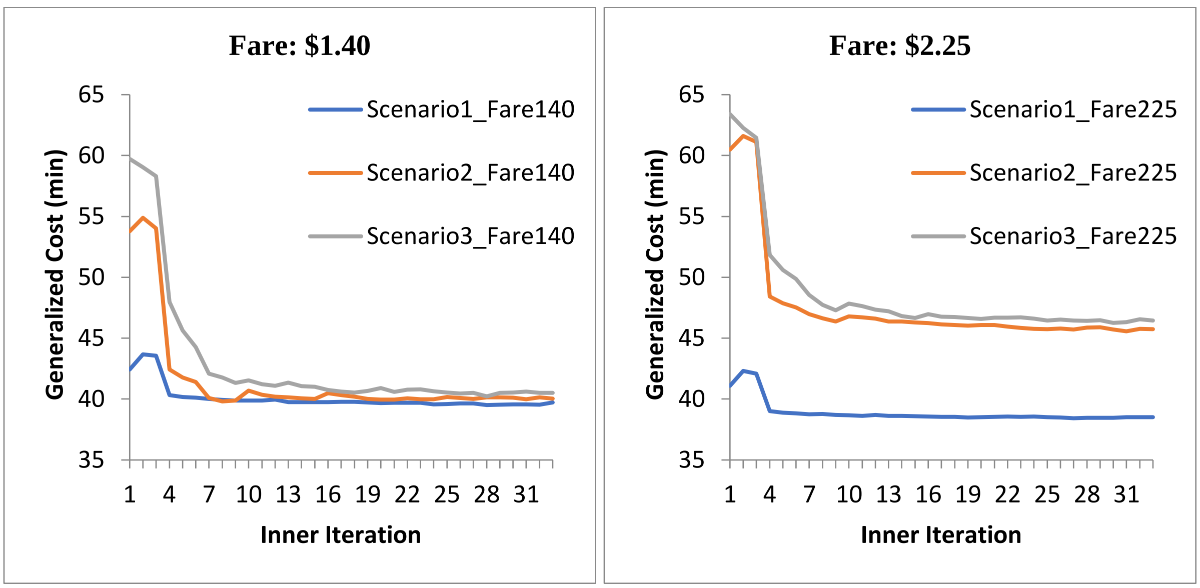

Figure 4.5: Average generalized transit travel cost per traveler

4.6 Summary

AVs and SAMSs promise to disrupt existing urban and suburban transportation systems, especially, transit systems. The impacts of SAMSs are highly uncertain; yet transportation planners still need to plan for SAMSs. To support transportation planners and modelers, this chapter presents a flexible modeling framework for dynamic transit assignment and simulation that explicitly incorporates a SAMS. The service is a first-mile suburban transit feeder SAMS.

The flexible modeling framework integrates a mode choice model with a dynamic transit assignment-simulation model and a SAMS fleet simulation model. The integrated model endogenously determines traveler mode choice as well as the performance of the transit network and the SAMSs at equilibrium. This study presents a mathematical formulation of the DCMC-TAP. The problem is analytically intractable; therefore, we present a simulation-based, iterative heuristic solution approach. In the iterative modeling framework, the upper level assigns travelers to one of five modes: car, park-and-ride, transit, SAMS, or transit with SAMS feeder. The lower level: (1) iteratively determines minimum cost transit hyperpaths, assigns travelers to the hyperpaths, and simulates their experiences, and (2) simulates a SAMS fleet providing service to suburban travelers. Individual traveler experiences and time-dependent network performance skims are then fed to the mode choice model. This process repeats until the mode choice probabilities converge. This integrated modeling framework, which endogenously determines traveler mode choice and transit and SAMS system performance, provides transportation planners and modelers a powerful tool to test various scenarios related to SAMSs.

To illustrate the capabilities of the integrated modeling framework, this study presents in application in the city of Chicago. The dynamic transit assignment-simulation tool models the entire Chicago Transit Authority network; moreover, the SAMS simulation tool models Evanston, IL, a suburb of Chicago. The model results illustrate that the iterative bi-level solution approach effectively solves the DCMC-TAP. Moreover, the results indicate that the integrated modeling framework can be used to assess transit and SAMS modal share as well as the performance of the transit network and the SAMS system.

This research makes several unique contributions to the literature. This is the first study, as far as the authors are aware, that explicitly integrates a SAMS simulation model with a dynamic transit assignment-simulation tool. Both simulation models are high-resolution and are able to capture the dynamics of a transit system and a SAMS system, such as the non-linear impacts of congestion. Second, this research develops the first integrated modeling framework that endogenously determines the modal shares for passenger car, park-and-ride, transit, general SAMS, and feeder SAMS. Moreover, these modal shares depend on and are consistent with the transit network and SAMS system performances. Verbas et al. present an integrated mode choice and dynamic transit assignment model (O. Verbas et al., 2016) but do not incorporate the SAMS modes. The integrated modeling approach provides reliable forecasts of transit and SAMS demand that explicitly consider the impacts of SAMS and transit system performance on mode choice, and the impact of traveler mode choices on SAMS and transit system performance.

4.7 Future Work

The results presented in this chapter show that the mode choice probabilities converge. However, more validation needs to be completed to ensure that the results are reliable. Several assumptions were made for parameters used in the modeling framework, hence a sensitivity analysis of the impacts of these parameters from the mode choice, assignment and simulation models should be performed.

5 Conclusions

5.1 Summary

Fully-autonomous vehicles (AVs) and the recent emergence of shared-use AV mobility services (SAMSs) are likely to significantly impact passenger transportation systems and the behavior of travelers using these systems. This study aims to provide a methodology and modeling framework to support the joint re-design of multimodal transit networks and SAMS fleets to explore and plan for likely future AV-enabled mobility scenarios. Accordingly, this study introduces the joint transit network redesign and SAMS fleet size determination problem (JTNR-SFSDP), along with a solution approach demonstrated on an actual large-scale network.

The JTNR-SFSDP formulation introduced in this research is a bi-level mathematical program. The upper-level is a modified transit network frequency setting problem (TNFSP) formulation that incorporates SAMS fleet size as a decision variable and allows transit pattern frequencies to be set to near-zero, effectively allowing transit patterns to be eliminated. A nonlinear programming solver is employed to obtain solutions of the upper-level problem. The lower-level formulation is a dynamic combined mode choice—traveler assignment problem (DCMC-TAP). Because the lower-level problem formulation is analytically intractable, an agent-based simulation model with three integrated components – mode choice, transit assignment-simulation, and AV fleet simulation – is used to solve the DCMC-TAP.

In the iterative heuristic solution approach, the lower-level model returns pattern-level transit demand and time-dependent SAMS demand to the upper-level model. Given this modal demand, the upper-level model solves the modified TNFSP and outputs transit pattern headways and SAMS fleet size to the lower-level module. The lower-level module re-solves the DCMC-TAP and outputs new modal demands to the upper-level module. This iterative process repeats until the solution converges.

This research presents a case study using the transit traveler demand and the multimodal transit network in the metropolitan area of Chicago to demonstrate the modeling framework and solution procedure. The results indicate that the JTNR-SFSDP modeling framework can improve the travel experiences of current transit users, in terms of average wait times, relative to the existing transit system. Importantly, the results suggest that improved traveler experience can be obtained without increasing the transit agency's budget. This is possible through the reallocation of resources away from transit patterns currently serving few travelers toward both SAMS vehicles and transit patterns with high demand.

5.2 Contributions

5.2.1 Chapter 4: DCMC-TAP

This study makes several unique contributions to the academic literature that also have significant and immediate value to transportation practitioners. First, this study integrates a SAMS simulation model with a dynamic transit assignment-simulation model. These simulation models are high-resolution microsimulation models that capture congestion and agent interactions, such as crowding on transit vehicles and long-wait times for SAMS, if demand exceeds supply. Second, the modeling framework embeds the dynamic transit assignment-simulation and SAMS simulation models within a mode choice model. This integrated modeling framework provides a planning and forecasting tool for transportation planners. The model explicitly captures the unique attributes of the SAMS mode and determines SAMS performance as a function of the spatio-temporal demand for the mode.

5.2.2 Chapter 3: JTNR-SFSDP

This study presents conceptual, theoretical, and methodological contributions. First, there are no other studies that define, model, or solve a joint transit network redesign and mobility service fleet size determination problem (JTNR-SFSDP). The study models the problem using a bi-level mathematical programming formulation. The problem definition and modeling framework represent a timely contribution to the existing literature, especially with the emergence of AVs, rapid growth of mobility services, and their interaction with transit services.

The second and third contributions relate to the formulation of the upper- and lower-level problems in the bi-level formulation, respectively. This study modifies a transit network frequency setting problem (TNSFP) formulation for the upper-level problem via allowing transit frequencies to be set to near-zero and incorporating SAMS fleet size as a decision variable. For the lower-level problem, this study employs a dynamic combined mode choice—traveler assignment problem (DCMC-TAP) formulation. This appears to be the first study to incorporate time-dependent mode (and route) choice in the lower-level problem of a transit network design problem. Hence, the modeling framework captures the modal split response to the frequency setting of transit patterns and operation/subsidization of SAMS fleets.

The fourth contribution is the use of a detailed agent-based simulation tool with three components (from Chapter 4) to address the lower-level problem. The three component models include a multinomial logit mode choice model, a transit traveler assignment-simulation model, and a SAMS fleet assignment-simulation model. Although data-intensive, the agent-based model for the lower-level provides information about individual travelers, the transit network, and the SAMS fleet. This information is valuable for model verification, model validation, and understanding the complex interactions between design decisions and travel behavior.

Relative to the small but growing literature modeling the intersection and integration of SAMSs with public transit, the proposed modeling framework imposes fewer restrictions. For example, this study makes no a priori assumptions about efficient joint designs of transit networks and SAMSs; whereas existing studies define scenarios in which SAMSs (i) replace specific transit routes (Pinto et al. 2018; Winter et al. 2018); (ii) replace transit systems (Basu et al. 2018), (iii) are implemented instead of new transit lines (Mendes, Bennàssar, and Chow 2017), or (iv) act as a transit feeder mode (Meyer et al. 2017; Scheltes and de Almeida Correia 2017). Similarly, the modeling framework in the current study endogenously determines modal flows (demand) for SAMSs and public transit based on the performance of the two systems, rather than exogenously (e.g. (Shen, Zhang, and Zhao 2018)).

5.3 Applications and Implications

The framework represents a powerful tool to address transit planning and design problems in the era of shared-use mobility and coming era of connected, automated, and shared mobility systems. Transit agencies can benefit from models and solution methods such as those presented in this dissertation that explicitly model travelers' responses to changes in transit networks and SAMS fleet sizes.

The results section illustrates the benefits of using an agent-based modeling framework for the lower-level DCMC-TAP because the output provides highly-detailed information about individual travelers, the transit network, and the SAMS fleet. This approach provides valuable insights to researchers and transit agencies trying to understand traveler responses to system-wide changes to transit networks. Moreover, the module-based modeling framework allows modelers and transit agencies to employ different behavioral models for mode and route choice.

Several policy insights can be drawn from the application of this framework and further assessed in future applications. The integration of SAMS with transit as a public service is expected to attain social and environmental benefits, and should be prioritized over the use of private AVs. SAMS should feed demand to high-capacity transit on reliable and high level of service routes in dense areas and fill accessibility gaps in low-density areas. There is great opportunity for SAMS to enhance accessibility to/from economically disconnected areas due to longer trips that are more likely to be shared.

SAMS also offers flexibility that should be leveraged. One way to do it in the context of integrated SAMS and transit services is by using them to test and observe demand for potential new transit routes before making decisions to allocate resources to what can be a very costly fixed route service. With use of collected ridesourcing data from transportation network companies (TNC), identifying such potential new routes or services to supply and areas to target can be trivial based on observed (TNC) traveler behavior trends.

5.4 Limitations

The JTNR-SFSDP subject to user-equilibrium constraints at the mode and route choice levels presented in this study is a multi-layered complex problem. The study represents a methodological foundation on which to build future refinements and extensions to both formulation and solution procedures, to overcome existing limitations and expand the scope of applicability. The remainder of this section discusses these limitations and potential future research areas.

First, using discrete agents to address the lower-level DCMC-TAP problem precludes an analytical expression connecting the leader's decisions to the followers' response. Response functions are typically a key component in bi-level solution approaches. As a first step, future work should investigate a heuristic response function that captures the lower-level mode and route choice response to the upper-level decision problem (setting transit pattern headways and SAMS fleet size) through an elastic demand function.

Second, the modeling framework in this study does not incorporate the impacts of SAMS vehicles on roadway traffic. As the size of SAMS fleets increase, and individual SAMS fleets significantly impact congestion on roadways, modeling road networks and road network congestion will be important. Liang et al. (2018) present a model that combines trip network assignment and dynamic routing for an automated taxi fleet that explicitly considers the impact of congestion on the efficiency of the automated taxi service. However, the focus of this work is on the important elements of the JTNR-SFSDP from the perspective of a transit agency. As such, additional model complexity related to the road network (in an already complex modeling framework) would likely not provide much added value, while considerably increasing the computational burden. Furthermore, the operational impact of AVs on roadway operation in mixed traffic (human-driven and autonomous) remains a topic of ongoing research in its own right.

Third, the mode choice component of the modeling framework includes several parameter values that significantly impact the modal split of travelers, including price elasticity and wait time elasticity for SAMSs. While this study used parameter values from existing mode choice models calibrated for the study area, uncertainty remains regarding these parameters for future modes and conditions. Future research and transit agencies should consider conducting behavioral studies, possibly in a virtual reality environment, to gain greater insight into these important parameters and dimensions of technology adoption.

Similarly, although the upper-level model incorporates the important transit and SAMS cost components, several refinements would increase realism in the modeling framework. For example, the model in this study considers a very simple fixed operational cost per SAMS vehicle, versus a per-mile or per-minute based cost. Research in the area of cost modeling for SAMS fleets would improve the realism of the modeling framework.

Lastly, the JTNR-SFSDP modeling framework only allows transit routes and route patterns to be removed from the transit network, it does not allow new routes or route patterns to be added. Although this limits the true solution space for the JTNR-SFSDP, the results in this study illustrate that despite this limitation the existing modeling framework can substantially improve transit service quality under capital and operational cost constraints. Future research could allow new transit routes and route patterns to be added.

In terms of implementation of the solution approach used, there is opportunity to evaluate whether adding a pre-SAMS step approach will provide solutions with more high level of service routes, so that the addition of SAMS can be targeted to low-demand or low-density areas. This can be performed in two steps: (1) optimize upper-level objective function for a No-SAMS scenario (minimize wait time and rejection penalties assuming transit-only services) to obtain high transit level of service corridors; then (2) apply JTNR-SFSDP framework to the solution from step (1), where the transit demand is given for the high level of service corridors and SAMS will seek to fill accessibility gaps as well as feeding demand to transit corridors.

Shifting to the application of the proposed framework, an interesting adaptation to be tested would be having autonomous shuttles (higher in-vehicle capacity) instead of limiting the SAMS fleet to rides that can be shared by only 2 people, which is the case in this study. This would also require a different SAMS routing strategy for picking up and dropping off passengers. An idea for this would be to consider shortest paths between zones and assuming passengers are willing to walk to/from a centroid/station of the zone where they are located/heading to.

A limitation of the application performed in this study is that the SAMS coverage area is much smaller than the area covered by the transit network, yet the upper-level design decisions include transit patterns that belong in the entire Chicago metropolitan region. This implies that the optimization being performed in the upper-level also adjusts for inefficiencies in the existing transit services that have nothing to do with the implementation of SAMS. If the transit and SAMS coexist in the same coverage area, the solution will translate more directly to a reallocation of resources between the two types of services. This shortcoming is dealt with in this study by providing performance metrics that are specific for the area where transit and SAMS coexist (Evanston). Future work would benefit from assessing SAMS services that cover the entire transit network area to ensure that the opportunity is broadly given to reallocate resources between the two modes.

6 References

Abdelghany, Khaled F., and Hani S. Mahmassani. 2001. "Dynamic Trip Assignment-Simulation Model for Intermodal Transportation Networks." In *Transportation Research Record*. https://doi.org/10.3141/1771-07.

Baaj, M.Hadi, and Hani S. Mahmassani. 1995. "Hybrid Route Generation Heuristic Algorithm for the Design of Transit Networks." *Transportation Research Part C: Emerging Technologies* 3 (1): 31–50. https://doi.org/10.1016/0968-090X(94)00011-S.

Basu, Rounaq, Andrea Araldo, Arun Prakash Akkinepally, Bat Hen Nahmias Biran, Kalaki Basak, Ravi Seshadri, Neeraj Deshmukh, Nishant Kumar, Carlos Lima Azevedo, and Moshe Ben-Akiva. 2018. "Automated Mobility-on-Demand vs. Mass Transit: A Multi-Modal Activity-Driven Agent-Based Simulation Approach." *Transportation Research Record: Journal of the Transportation Research Board*, May, 036119811875863. https://doi.org/10.1177/0361198118758630.

Bösch, Patrick M., Felix Becker, Henrik Becker, and Kay W. Axhausen. 2018. "Cost-Based Analysis of Autonomous Mobility Services." *Transport Policy* 64 (May): 76–91. https://doi.org/10.1016/J.TRANPOL.2017.09.005.

Cea, J. de, and E. Fernandez. 1993. "Transit Assignment for Congested Public Transport Systems: An Equilibrium Model." *Transportation Science* 27 (2): 133–47. https://doi.org/10.1287/trsc.27.2.133.

Ceder, Avishai, and Yechezkel Israeli. 1998. "User and Operator Perspectives in Transit Network Design." *Transportation Research Record: Journal of the Transportation Research Board* 1623 (January): 3–7. https://doi.org/10.3141/1623-01.

Ceder, Avishai, and Nigel H M Wilson. 1986. "Bus Network Design." *Transportation Research Part B*. https://doi.org/10.1016/0191-2615(86)90047-0.

Cepeda, M., R. Cominetti, and M. Florian. 2006. "A Frequency-Based Assignment Model for Congested Transit Networks with Strict Capacity Constraints: Characterization and Computation of Equilibria." *Transportation Research Part B: Methodological*. https://doi.org/10.1016/j.trb.2005.05.006.

Cervero, Robert. 1984. "Cost and Performance Impacts of Transit Subsidy Programs." *Transportation Research Part A: General*. https://doi.org/10.1016/0191-2607(84)90015-3.

Clarens, Gérard C, and V F Hurdle. 1975. "An Operating Strategy for a Commuter Bus System ." *Transportation Science*. https://doi.org/10.1287/trsc.9.1.1.

Cominetti, Roberto, and Jose Correa. 2001. "Common-Lines and Passenger Assignment in Congested Transit Networks." *Transportation Science* 35 (3): 250–67. https://doi.org/10.1287/trsc.35.3.250.10154.

Constantin, I., and M. Florian. 1995. "Optimizing Frequencies in a Transit Network: A Nonlinear Bi-Level Programming Approach." *International Transactions in Operational Research* 2 (2): 149–64. https://doi.org/10.1111/j.1475-3995.1995.tb00011.x.

Correia, Gonçalo Homem de Almeida, Bart Wiegmans, Kara Kockelman, and Meng Qiang. 2018. "Editorial of the TR Part C Special Issue on: 'Operations with Automated Vehicles (AVs): Applications in Freight and Passenger Transport.'" *Transportation Research Part C: Emerging Technologies*. https://doi.org/10.1016/j.trc.2018.07.014.

Daganzo, Carlos F., and Yosef Sheffi. 1977. "On Stochastic Models of Traffic Assignment." *Transportation Science* 11 (3): 253–74.

Dial, R B. 1967. "Transit Pathfinder Algorithm." *Highway Research Record* 205: 67–85. http://onlinepubs.trb.org/onlinepubs/out_of_print.htm.

Fagnant, Daniel J., and Kara M. Kockelman. 2014. "The Travel and Environmental Implications of Shared Autonomous Vehicles, Using Agent-Based Model Scenarios." *Transportation Research Part C: Emerging Technologies* 40: 1–13. https://doi.org/10.1016/j.trc.2013.12.001.

Fagnant, Daniel J., Kara M. Kockelman, and Prateek Bansal. 2015. "Operations of a Shared Autonomous Vehicle Fleet for the Austin, Texas Market." *Transportation Research Record: Journal of the Transportation Research Board* 2536: 98–106.

Fan, Wei, and Randy Machemehl. 2011. "Bi-Level Optimization Model for Public Transportation Network Redesign Problem." *Transportation Research Record: Journal of the Transportation Research Board* 2263 (December): 151–62. https://doi.org/10.3141/2263-17.

Fisk, Caroline. 1980. "Some Developments in Equilibrium Traffic Assignment." *Transportation Research Part B: Methodological*. https://doi.org/10.1016/0191-2615(80)90004-1.

Furth, P G, and F B Day. 1985. "TRANSIT ROUTING AND SCHEDULING STRATEGIES FOR HEAVY-DEMAND CORRIDORS (ABRIDGMENT)." *Transportation Research Record*.

Furth, P G, and Nigel H M Wilson. 1981. "Setting Frequencies on Bus Routes:Theory and Practice." *Transportation Research Record*.

Gao, Ziyou, Huijun Sun, and Lian Long Shan. 2004. "A Continuous Equilibrium Network Design Model and Algorithm for Transit Systems." *Transportation Research Part B: Methodological* 38 (3): 235–50. https://doi.org/10.1016/S0191-2615(03)00011-0.

Gendreau, Michel, and Jean Yves Potvin. 2005. "Metaheuristics in Combinatorial Optimization." *Annals of Operations Research*. https://doi.org/10.1007/s10479-005-3971-7.

Gentile, Guido. 2018. "New Formulations of the Stochastic User Equilibrium with Logit Route Choice as an Extension of the Deterministic Model." *Transportation Science*.

https://doi.org/10.1287/trsc.2018.0839.

Greenblatt, Jeffery B., and Susan Shaheen. 2015. “Automated Vehicles, On-Demand Mobility, and Environmental Impacts.” *Current Sustainable/Renewable Energy Reports* 2 (3): 74–81. https://doi.org/10.1007/s40518-015-0038-5.

Guan, J.F., Hai Yang, and S.C. Wirasinghe. 2006. “Simultaneous Optimization of Transit Line Configuration and Passenger Line Assignment.” *Transportation Research Part B: Methodological* 40 (10): 885–902. https://doi.org/10.1016/J.TRB.2005.12.003.

Guihaire, Valérie, and Jin-Kao Hao. 2008. “Transit Network Design and Scheduling: A Global Review.” *Transportation Research Part A: Policy and Practice* 42 (10): 1251–73. https://doi.org/10.1016/J.TRA.2008.03.011.

Hall, Jonathan D, Joseph Price, and Craig Palsson. 2017. “Is Uber a Substitute or Complement to Public Transit ?” *Working Paper*, 1–19. http://individual.utoronto.ca/jhall/documents/Uber_and_Public_Transit.pdf%0Ahttps://www.economics.utoronto.ca/public/workingPapers/tecipa-585.pdf.

Hamdouch, Younes, H. W. Ho, Agachai Sumalee, and Guodong Wang. 2011. “Schedule-Based Transit Assignment Model with Vehicle Capacity and Seat Availability.” *Transportation Research Part B: Methodological* 45 (10): 1805–30. https://doi.org/10.1016/j.trb.2011.07.010.

Hamdouch, Younes, and Siriphong Lawphongpanich. 2008. “Schedule-Based Transit Assignment Model with Travel Strategies and Capacity Constraints.” *Transportation Research Part B: Methodological* 42 (7–8): 663–84. https://doi.org/10.1016/j.trb.2007.11.005.

Hickman, Mark, and Kelly Blume. 2001. “Modeling Cost and Passenger Level of Service for Integrated Transit Service.” In , 233–51. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-56423-9_14.

Hyland, Michael, and Hani S. Mahmassani. 2018. “Dynamic Autonomous Vehicle Fleet Operations: Optimization-Based Strategies to Assign AVs to Immediate Traveler Demand Requests.” *Transportation Research Part C: Emerging Technologies* 92 (July): 278–97. https://doi.org/10.1016/j.trc.2018.05.003.

Ibarra-Rojas, O.J., F. Delgado, R. Giesen, and J.C. Muñoz. 2015. “Planning, Operation, and Control of Bus Transport Systems: A Literature Review.” *Transportation Research Part B: Methodological* 77 (July): 38–75. https://doi.org/10.1016/J.TRB.2015.03.002.

Iliopoulou, Christina, Konstantinos Kepaptsoglou, and Eleni Vlahogianni. 2019. “Metaheuristics for the Transit Route Network Design Problem: A Review and Comparative Analysis.” *Public Transport*. https://doi.org/10.1007/s12469-019-00211-2.

Kamel, I, S Hasnine, A Shalaby, Habib KN, and B Abdulhai. 2019. “Evaluating Time-Based

Transit Fares: Integrated Model of Departure Time and Travel Mode Choice within a Dynamic Multimodal Assignment Framework." In *98th Annual Meeting Transportation Research Board*.

Keil, Roger, and Jean-Paul D. Addie. 2015. "'It's Not Going to Be Suburban, It's Going to Be All Urban': Assembling Post-suburbia in the Toronto and Chicago Regions." *International Journal of Urban and Regional Research* 39 (5): 892–911.

Lam, W. H.K., Z. Y. Gao, K. S. Chan, and H. Yang. 1999. "A Stochastic User Equilibrium Assignment Model for Congested Transit Networks." *Transportation Research Part B: Methodological*. https://doi.org/10.1016/S0191-2615(98)00040-X.

Liang, Xiao, Gonçalo Homem de Almeida Correia, and Bart van Arem. 2016. "Optimizing the Service Area and Trip Selection of an Electric Automated Taxi System Used for the Last Mile of Train Trips." *Transportation Research Part E: Logistics and Transportation Review* 93 (September): 115–29. https://doi.org/10.1016/j.tre.2016.05.006.

Liu, Tao, and Avishai Avi Ceder. 2017. "Integrated Public Transport Timetable Synchronization and Vehicle Scheduling with Demand Assignment: A Bi-Objective Bi-Level Model Using Deficit Function Approach." In *Transportation Research Procedia*, 23:341–61. https://doi.org/10.1016/j.trpro.2017.05.020.

Lu, Chung-Cheng, Hani S. Mahmassani, and Xuesong Zhou. 2009. "Equivalent Gap Function-Based Reformulation and Solution Algorithm for the Dynamic User Equilibrium Problem." *Transportation Research Part B: Methodological* 43 (3): 345–64. https://doi.org/10.1016/j.trb.2008.07.005.

Lu, Zhongming, Rundong Du, Ellen Dunham-Jones, Haesun Park, and John Crittenden. 2017. "Data-Enabled Public Preferences Inform Integration of Autonomous Vehicles with Transit-Oriented Development in Atlanta." *Cities*. https://doi.org/10.1016/j.cities.2017.01.004.

Malalgoda, Narendra, and Siew Hoon Lim. 2019. "Do Transportation Network Companies Reduce Public Transit Use in the U.S.?" *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2019.09.051.

Marguier, Philippe H. J., and Avishai Ceder. 1984. "Passenger Waiting Strategies for Overlapping Bus Routes." *Transportation Science* 18 (3): 207–30. https://doi.org/10.1287/trsc.18.3.207.

Martin, Elliot W., and Susan a. Shaheen. 2014. "Evaluating Public Transit Modal Shift Dynamics in Response to Bikesharing: A Tale of Two U.S. Cities." *Journal of Transport Geography*. https://doi.org/10.1016/j.jtrangeo.2014.06.026.

Mendes, Lucas Mestres, Manel Rivera Bennàssar, and Joseph Y. J. Chow. 2017. "Comparison of Light Rail Streetcar Against Shared Autonomous Vehicle Fleet for Brooklyn–Queens Connector in New York City." *Transportation Research Record: Journal of the*

Transportation Research Board 2650 (January): 142–51. https://doi.org/10.3141/2650-17.

Meyer, Jonas, Henrik Becker, Patrick M. Bösch, and Kay W. Axhausen. 2017. "Autonomous Vehicles: The next Jump in Accessibilities?" *Research in Transportation Economics* 62 (June): 80–91. https://doi.org/10.1016/J.RETREC.2017.03.005.

Minocha, Inshu, P. S. Sriraj, Paul Metaxatos, and Piyushimita [Vonu] Thakuriah. 2008. "Analysis of Transit Quality of Service and Employment Accessibility for the Greater Chicago, Illinois, Region." *Transportation Research Record: Journal of the Transportation Research Board* 2042: 20–29.

Newell, G. F. 1971. "Dispatching Policies for a Transportation Route." *Transportation Science* 5 (1): 91–105. https://doi.org/10.1287/trsc.5.1.91.

Nguyen, S., and S. Pallottino. 1988. "Equilibrium Traffic Assignment for Large Scale Transit Networks." *European Journal of Operational Research* 37 (2): 176–86. https://doi.org/10.1016/0377-2217(88)90327-X.

Oliker, Nurit, and Shlomo Bekhor. 2018. "A Frequency Based Transit Assignment Model That Considers Online Information." *Transportation Research Part C: Emerging Technologies* 88 (March): 17–30. https://doi.org/10.1016/j.trc.2018.01.004.

Owen, Andrew, and David Levinson. 2014. "Access Across America : Transit 2014."

Peeta, Srinivas, and Athanasios Ziliaskopoulos. 2001. "Foundations of Dynamic Traffic Assignment: The Past, the Present and the Future." *Networks and Spatial Economics*. https://doi.org/10.1023/A:1012827724856.

Pinto, Helen Karla Ramalho de Farias, Michael F Hyland, İ. Ömer Verbas, and Hani S Mahmassani. 2018. "Integrated Mode Choice and Dynamic Traveler Assignment-Simulation Framework to Assess the Impact of a Suburban First-Mile Shared Autonomous Vehicle Fleet Service on Transit Demand." In *Transportation Research Board Annual Meeting*. Washington, D.C.

Pucher, John. 1982. "Discrimination in Mass Transit." *Journal of the American Planning Association*. https://doi.org/10.1080/01944368208976181.

Reeven, Peran Van, and Vladimir Karamychev. 2016. "Subsidies and US Urban Transit Ridership." *Journal of Transport Economics and Policy*.

Salzborn, Franz J. M. 1972. "Optimum Bus Scheduling." *Transportation Science* 6 (2): 137–48. https://doi.org/10.1287/trsc.6.2.137.

Sbayti, Hayssam, Chung-Cheng Lu, and Hani Mahmassani. 2007. "Efficient Implementation of Method of Successive Averages in Simulation-Based Dynamic Traffic Assignment Models for Large-Scale Network Applications." *Transportation Research Record: Journal of the Transportation Research Board* 2029: 22–30. https://doi.org/10.3141/2029-03.

Schéele, Siv. 1980. “A Supply Model for Public Transit Services.” *Transportation Research Part B*. https://doi.org/10.1016/0191-2615(80)90039-9.

Scheltes, Arthur, and Gonçalo Homem de Almeida Correia. 2017. “Exploring the Use of Automated Vehicles as Last Mile Connection of Train Trips through an Agent-Based Simulation Model: An Application to Delft, Netherlands.” *International Journal of Transportation Science and Technology* 6 (1): 28–41. https://doi.org/10.1016/j.ijtst.2017.05.004.

Schmöcker, Jan Dirk, Achille Fonzone, Hiroshi Shimamoto, Fumitaka Kurauchi, and Michael G.H. Bell. 2011. “Frequency-Based Transit Assignment Considering Seat Capacities.” *Transportation Research Part B: Methodological* 45 (2): 392–408. https://doi.org/10.1016/j.trb.2010.07.002.

Shen, Yu, Hongmou Zhang, and Jinhua Zhao. 2018. “Integrating Shared Autonomous Vehicle in Public Transportation System: A Supply-Side Simulation of the First-Mile Service in Singapore.” *Transportation Research Part A: Policy and Practice* 113 (July): 125–36. https://doi.org/10.1016/J.TRA.2018.04.004.

Spiess, Heinz, and Michael Florian. 1989. “Optimal Strategies: A New Assignment Model for Transit Networks.” *Transportation Research Part B* 23 (2): 83–102. https://doi.org/10.1016/0191-2615(89)90034-9.

Stiglic, Mitja, Niels Agatz, Martin Savelsbergh, and Mirko Gradisar. 2018. “Enhancing Urban Mobility: Integrating Ride-Sharing and Public Transit.” *Computers and Operations Research*. https://doi.org/10.1016/j.cor.2017.08.016.

Sun, Huijun, Ziyou Gao, and Jianjun Wu. 2008. “A Bi-Level Programming Model and Solution Algorithm for the Location of Logistics Distribution Centers.” *Applied Mathematical Modelling* 32 (4): 610–16. https://doi.org/10.1016/J.APM.2007.02.007.

Tilahun, Nebiyou, and Moyin Li. 2015. “Walking Access to Transit Stations.” *Transportation Research Record: Journal of the Transportation Research Board* 2534: 16–23. http://dx.doi.org/10.3141/2534-03.

Transportation Research Board, and Engineering and Medicine National Academies of Sciences. 2016. “Shared Mobility and the Transformation of Public Transit.” Edited by Sharon Feigon and Colin Murphy. Washington, DC: The National Academies Press. https://doi.org/10.17226/23578.

Verbas, I., and Hani Mahmassani. 2013. “Optimal Allocation of Service Frequencies over Transit Network Routes and Time Periods.” *Transportation Research Record: Journal of the Transportation Research Board* 2334: 50–59. https://doi.org/10.3141/2334-06.

Verbas, İ. Ömer, and Hani S. Mahmassani. 2015. “Integrated Frequency Allocation and User Assignment in Multimodal Transit Networks.” *Transportation Research Record: Journal of the Transportation Research Board* 2498: 37–45. https://doi.org/10.3141/2498-05.

Verbas, İ. Ömer, Hani S. Mahmassani, and Michael F. Hyland. 2015. "Dynamic Assignment-Simulation Methodology for Multimodal Urban Transit Networks." *Transportation Research Record: Journal of the Transportation Research Board* 2498: 64–74. https://doi.org/10.3141/2498-08.

Verbas, Omer, Hani S Mahmassani, Michael F Hyland, and Hooram Halat. 2016. "Integrated Mode Choice and Dynamic Traveler Assignment in Multimodal Transit Networks: Mathematical Formulation, Solution Procedure, and Large-Scale Application." *Transportation Research Record: Journal of the Transportation Research Board* 2564: 78–88.

Vuchic, Vukan R. 1999. *Transportation for Livable Cities.*

Winter, Konstanze, Oded Cats, Gonçalo Correia, and Bart van Arem. 2018. "Performance Analysis and Fleet Requirements of Automated Demand-Responsive Transport Systems as an Urban Public Transport Service." *International Journal of Transportation Science and Technology* 7 (2): 151–67. https://doi.org/10.1016/J.IJTST.2018.04.004.

Wong, K. I., S. C. Wong, J. H. Wu, Hai Yang, and William H.K. Lam. 2004. "A Combined Distribution, Hierarchical Mode Choice, and Assignment Network Model with Multiple User and Mode Classes." In *Urban and Regional Transportation Modeling: Essays in Honor of David Boyce*. https://doi.org/10.4337/9781845420536.00008.

Wu, Jia Hao, Michael Florian, and Patrice Marcotte. 1994. "Transit Equilibrium Assignment: A Model and Solution Algorithms." *Operations Research Society of America. Transportation Science Section. Transportation Science* 28 (3): 193–203. https://doi.org/10.1287/trsc.28.3.193.

Wu, Liyu, Weihua Gu, and Wenbo Fan. 2018. "A Continuum Model for Optimal Joint Designs of Bike-Sharing and Transit Networks." *Transportation Research Record*.

Yan, X., Jonathan Levine, and X. Zhao. 2019. "Integrating Ridesourcing Services with Public Transit: An Evaluation of Traveler Responses Combining Revealed and Stated Preference Data." *Transportation Research Part C: Emerging Technologies*. https://doi.org/10.1016/j.trc.2018.07.029.

Yoo, Gyeong Sang, Dong Kyu Kim, and Kyung Soo Chon. 2010. "Frequency Design in Urban Transit Networks with Variable Demand: Model and Algorithm." *KSCE Journal of Civil Engineering* 14 (3): 403–11. https://doi.org/10.1007/s12205-010-0403-2.

Yu, Bin, Zhongzhen Yang, and Jinbao Yao. 2010. "Genetic Algorithm for Bus Frequency Optimization." *Journal of Transportation Engineering* 136 (6): 576–83. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000119.

Zhang, Kuilin, Hani S Mahmassani, and Peter Vovsha. 2011. "Integrated Nested Logit Mode Choice and Dynamic Network Micro-Assignment Model Platform to Support Congestion and Pricing Studies: The New York Metropolitan Case." In *Transportation Research Board*

90th Annual Meeting. Washington, DC: Transportation Research Board.

Zhou, Xuesong, Hani S. Mahmassani, and Kuilin Zhang. 2008. "Dynamic Micro-Assignment Modeling Approach for Integrated Multimodal Urban Corridor Management." *Transportation Research Part C: Emerging Technologies*. https://doi.org/10.1016/j.trc.2007.07.002.