University of California Transportation Center
UCTC Dissertation UCTC-DISS-2013-05

Integration of Locational Decisions with the Household Activity Pattern Problem and Its Applications in Transportation Sustainability

Jee Eun Kang
University of California, Irvine
2013

UNIVERSITY OF CALIFORNIA,

IRVINE

Integration of Locational Decisions with the Household Activity Pattern Problem and Its Applications in Transportation Sustainability

DISSERTATION

submitted in partial satisfaction of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in Civil Engineering

by

Jee Eun Kang

Dissertation Committee:

Professor Will W. Recker, Chair

Professor Michael G. McNally

Professor G. Scott Samuelsen

2013

DEDICATION

To my family

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

ACKNOWLEDGEMENTS

This has been a long, sometimes painful but fun and rewarding experience for me, and I would like to thank a number of people who made this journey possible.

This dissertation would have not been possible without support and advising from Professor Will Recker. I could not have asked for anything more than the academic environment he has provided. I am particularly thankful that he was so open to support my research pursuits and that he allowed me to complete tasks on my own schedule. This nurturing approach allowed me to mature at my own pace and by my own motivation. I believe that motivation driven from within oneself is strong and powerful, and it has made me find my passion for research which I consider the biggest achievement during my time as a Ph.D student. I truly feel that everything I have achieved in my academic career to this point and even beyond, I owe to Professor Recker. I sincerely wish him the best for the last phase of his long and revered career.

I would like to express my gratitude for the committee members, Professor Michael McNally and Professor Scott Samuelsen for their comments and support which provided me with critical ideas and thoughts for the dissertation. I would also like to thank Professor R. Jayakrishnan and Professor Amelia Regan for their encouragement especially the last year which has been an emotional roller coaster for me.

I also feel that I was very fortunate to have been involved in research collaboration with the Advanced Power and Energy Program (APEP) at University of California, Irvine. Through the collaboration, we were able to merge our expertise in transportation systems and energy systems. This not only resulted in numerous journal articles but has also produced impactful outcomes for the real world implementation of sustainable transportation, which rarely comes for academic researchers. I am particularly grateful for interactions with Dr. Tim Brown, Dr. Shane Stephens-Romero, Li Zhang, and Professor Scott Samuelsen. Their expertise in energy systems provided me with noble ideas for transportation sustainability work in this dissertation.

I also benefited significantly from the active graduate student community at the Institute of Transportation Studies at the University of California, Irvine. Through our times spent together working on homework, exercising at the ARC, taking camping trips, and enjoying dinner parties, not only did I make great memories that I will have for a long time, but I was able to build many lifelong friends and research partners. I will truly miss our times together in Irvine, but we have many years of TRB trips in the future and I will always look forward to UCI reunion night.

I would like to thank my family, my parents who have supported me with unconditional love and care for the last 30 years, and my sister Hee Eun, for the sisterly bond and relationship we have. I would also like to show my gratitude and love for my husband, Brandon Choi, for bringing the better self out of me personally and academically, and for going through my successes and failures with me for the last seven years. I hope I am being the rock for you as you always have been for me. Lastly, I am always proud and grateful that I have such a tight extended family – I have not seen any other family as close as ours and your prayers and blessings are the reason how I was able to get to this point in my life.

This research was supported, in part, by grants from the University of California Transportation Center, the ITS Multi-Campus Research Program and Initiative on Sustainable Transportation, and the Graduate Dean's Dissertation Fellowship from the University of California, Irvine. I am very grateful to the faculty and staff at the University of California

Transportation Center and the Institute of Transportation Studies for all their financial and academic support.

CURRICULUM VITAE

EDUCATION

2013	Ph.D. in Civil Engineering
	University of California, Irvine, CA
2008	M.S. in Civil Engineering
	University of California, Irvine, CA
2006	B.S. in Urban Engineering and Planning (Transportation)
	Yonsei University, Seoul, Korea

PEER-REVIEWED PUBLICATIONS

- **JE Kang**, WW Recker (2013), Measuring the inconvenience of operating an alternative fuel vehicle, *Transportation Research Part D: Transport and Environment*, *Under Review*
- **JE Kang**, TM Brown, WW Recker, and GS Samuelsen (2013), Refueling hydrogen fuel cell vehicles with 68 proposed refueling stations in California: measuring deviations from daily travel patterns, *International Journal of Hydrogen Energy*, *Under Review*
- **JE Kang**, WW Recker (2013), Strategic hydrogen refueling station locations with scheduling and routing considerations of individual vehicles, *Transportation Science*, *Revision Under Review*
- **JE Kang**, WW Recker (2013), The Location Selection Problem for the Household Activity Pattern Problem, *Transportation Research Part B: Methodological*, 55, 75-97

- **JE Kang**, JYJ Chow, WW Recker (2013), On Activity-based Network Design Problems, *Transportation Research Part B: Methodological*, *forthcoming*
- SD Stephens-Romero, TM Brown, M Carreras-Sospedra, **JE Kang**, J Brouwer, D Dabdub, WW Recker, and GS Samuelsen (2011), Projecting full built-out environmental impacts and roll-out strategies associated with viable hydrogen fueling infrastructure strategies, *International Journal of Hydrogen Energy*, 36-22, 14309-14323
- SD Stephens-Romero, TM Brown, **JE Kang**, WW Recker, and GS Samuelsen (2010), Systematic planning to optimize investments in hydrogen infrastructure deployment, *International Journal of Hydrogen Energy*, 35-10, 4652-4667
- **JE Kang** and WW Recker (2009), An activity-based assessment of the potential impacts of plug-in hybrid electric vehicles on energy and emissions using 1-day travel data, *Transportation Research Part D: Transport and Environment*, 14-8, 541-556

CONFERENCE PRESENTATIONS

- "Measuring the inconvenience of Refueling/Recharging for Alternative Fuel Vehicles", UCTC Student Conference, UCLA, March 2, 2013.
- "Integration of locational decisions with the household activity pattern problem and its applications in transportation sustainability", Doctoral Student Research in Transportation Modeling, 2013 Transportation Research Board 92nd Annual Meeting, Washington DC, Jan 13-17, 2013.
- "Hydrogen Charging Station Location Analysis with Scheduling and Routing Considerations", INFORMS 2012, Phoenix, AZ, April 24-25, 2012.

- “The location selection problem for the household activity pattern problem”, UCTC Student Conference, UC Davis, April 20, 2012.
- “Truck traffic as a function of traffic flow”, UCTC Student Conference, UC Berkeley, February 24-15, 2011.

TEACHING EXPERIENCE

University of California, Irvine – Irvine, CA

- CEE 111 Engineering Decision Analysis (Winter Quarter, 2007, 2008, 2009, 2010)
- CEE 11 Probability and Statistics (Spring Quarter, 2007)

HONORS AND FELLOWSHIPS

- Eno Fellow (2013), Eno Center for Transportation
- Graduate Student Research and Travel Grant (2013), The Henry Samueli School of Engineering, University of California, Irvine
- Graduate Scholarship (2012), Women’s Transportation Seminar Orange County Chapter
- Graduate Dean’s Dissertation Fellowship (2012), University of California, Irvine
- Graduate Student Community Award (2011), University of California, Irvine
- Dissertation Fellowship (2010), University of California Transportation Center

TECHNICAL SKILLS

- Programming language: Java
- Mathematical, statistical tools: CPLEX, MatLAB, R, AMPL, SPSS

- Transportation tool: GIS, TransCAD

SERVICES

- Journal referee: Networks and Spatial Economics [1], Transportation Research: Part A [1]

ABSTRACT OF THE DISSERTATION

Integration of Locational Decisions with the Household Activity Pattern Problem

And Its Applications in Transportation Sustainability

By

Jee Eun Kang

Doctor of Philosophy in Civil Engineering

University of California, Irvine, 2013

Professor Will W. Recker, Chair

This dissertation focuses on the integration of the Household Activity Pattern Problem (HAPP) with various locational decisions considering both supply and demand sides. We present several methods to merge these two distinct areas—transportation infrastructure and travel demand procedures—into an integrated framework that has been previously exogenously linked by feedback or equilibrium processes. From the demand side, travel demand for non-primary activities is derived from the destination choices that a traveler makes that minimizes travel disutility within the context of considerations of daily scheduling and routing. From the supply side, the network decisions are determined as an integral function of travel demand rather than a given fixed OD matrix.

First, the Location Selection Problem for the Household Activity Pattern Problem (LSP-HAPP) is developed. LSP-HAPP extends the HAPP by adding the capability to make destination choices simultaneously with other travel decisions of household activity allocation, activity sequence, and departure time. Instead of giving a set of pre-fixed activity locations to visit, LSP-

HAPP chooses the location for certain activity types given a set of candidate locations. A dynamic programming algorithm is adopted and further developed for LSP-HAPP in order to deal with the choices among a sizable number of candidate locations within the HAPP modeling structure. Potential applications of synthetic pattern generation based on LSP-HAPP formulation are also presented.

Second, the Location – Household Activity Pattern Problem (Location-HAPP), a facility location problem with full-day scheduling and routing considerations is developed. This is in the category of Location-Routing Problems (LRPs), where the decisions of facility location models are influenced by possible vehicle routings. Location-HAPP takes the set covering model as a location strategy, and HAPP as the scheduling and routing tool. The proposed formulation isolates each vehicle's routing problem from those of other vehicles and from the master set covering problem. A modified column generation that uses a search method to find a column with a negative reduced price is proposed.

Third, the Network Design Problem is integrated with the Household Activity Pattern Problem (NDP-HAPP) as a bilevel optimization problem. The bilevel structure includes an upper level network design while the lower level includes a set of disaggregate household itinerary optimization problems, posed as HAPP or LSP-HAPP. The output of upper level NDP (level-of-service of the transportation network) becomes input data for the lower level HAPP that generates travel demand which becomes the input for the NDP. This is advantageous over the conventional NDP that outputs the best set of links to invest in, given an assumed OD matrix. Because the proposed NDP-HAPP can output the same best set of links, a new OD matrix and a detailed temporal distribution of activity participation and travel are created. A decomposed heuristic solution algorithm that represents each decision makers' rationale shows optimality gaps of as much as 5% compared to exact solutions when tested with small examples.

Utilizing the aforementioned models, two transportation sustainability studies are then conducted for the adoption of Alternative Fuel Vehicles (AFVs). The challenges in adopting AFVs are directly related to the transportation infrastructure problems since the initial AFV refueling locations will need to provide comparable convenient travel experience for the early adopters when compared to the already matured gasoline fuel based transportation infrastructure. This work demonstrates the significance of the integration between travel demand model and infrastructure problems, but also draws insightful policy measurements regarding AFV adoption.

The first application study attempts to measure the household inconvenience level of operating AFVs. Two different scenarios are examined from two behavioral assumptions – keeping currently reported pattern and minimizing the inconvenience cost through HAPPR or HAPPC. From these patterns, the personal or household inconvenience level is derived as compared to the original pattern, providing quantified data on how the public sector would compensate for the increases in travel disutility to ultimately encourage the attractiveness of AFVs.

From the supply side of the AFV infrastructure, Location-HAPP is applied to the incubation of the minimum refueling infrastructure required to support early adoption of Hydrogen Fuel Cell Vehicles (HFCVs). One of the early adoption communities targeted by auto manufacturers is chosen as the study area, and then three different values of accessibility are tested and measured in terms of tolerances to added travel time. Under optimal conditions, refueling trips are found to be toured with other activities. More importantly, there is evidence that excluding such vehicle-infrastructure interactions as well as routing and scheduling interactions can result in over-estimation of minimum facility requirement.

CHAPTER 1 INTRODUCTION

Activity-based travel demand models were developed to incorporate more fundamental travel decision making processes for transportation planning. Compared to the conventional trip-based models, activity-based models offer greater detail regarding activities, trips, and linkages between the two as well as a better resolution of time and space. In addition, activity-based models are more policy-sensitive due to their derivation from fundamental travel decision making processes and are better able to capture the detailed interactions that occur between elements of the transportation system. Activity-based travel demand models have matured over last several decades, and still continue to mature. In fact, some models have been developed to the stage that they are implemented for regional travel forecasting.

In this dissertation, we focus on the household activity pattern problem (HAPP) proposed by Recker (1995), in particular. HAPP formulates the time-space geography (Hagerstrand, 1970) under which an individual's activity participation and travel takes place, where time and space are treated as a single continuous path through temporal and spatial dimensions rather than as discrete trips. HAPP is an interpretation of personal- or household-level daily activity scheduling based on an extension of the pickup and delivery problem with time windows (PDPTW). In this dissertation, HAPP is further developed to include locational decisions, for both the demand and supply sides. On the demand side, HAPP is extended to include destination choice, making this model a tool by which to link travel demand models with the transportation network. On the supply side, HAPP is integrated with two well-known network problems – the set covering problem and the network design problem. Previously, for many network problems, origin-destination (OD) decision, referred to here simply as demand, is given statically, exogenous from the model although individual's travel demand, and therefore, OD are also influenced by the network's level-of-service. By incorporating the travel demand procedure within the network problem, interactions between the network and travel demand is better captured.

These extended models are then applied to tackle transportation sustainability issues. The transportation sector is responsible for about one third of greenhouse gas emissions and the majority of foreign oil consumption. Adoption of Alternative Fuel Vehicles (AFVs) is expected to solve both issues; however, there are large hurdles to overcome in order for AFVs to become the norm. Due to inconvenience factors related to limited infrastructure supply for refueling/recharging and the requirement for drivers' to invest in new vehicle technology, AVF adoption success requires heavy analysis into how these factors should be addressed. Accordingly, the models developed in this dissertation are targeted tools for evaluating these issues. For example, drivers' inconvenience can be better analyzed using an activity-based travel demand model, and limited infrastructure supply can be viewed as a network problem. Since HAPP combines individual driver's travel decisions with the spatial and temporal physical constraints, it is possible to predict feasible travel patterns which are not previously observed, such as travel patterns of drivers using hypothetical AFVs.

1.1 THE HOUSEHOLD ACTIVITY PATTERN PROBLEM

Distinct from the majority of activity-based travel demand models are based on either econometric or simulation approaches, HAPP is a network-based mathematical programming approach that can offer explanations to a variety of transportation behaviors not directly amenable to either econometric or simulation approaches (Gan and Recker, 2012, Chow and Recker, 2011; Recker et al., 2008; Gan and Recker, 2008; Recker, 2001; Recker, et al, 2001; Recker and Parimi, 1999; Recker, 1995).

The structure of the formulation is as follows.

$$\min \mathbf{Z}^h = \text{Travel Disutility of Household } h = f(\mathbf{X}^h, \mathbf{T}^h)$$

s.t.

$$\mathbf{A}^h \begin{bmatrix} \mathbf{X}^h \\ \mathbf{T}^h \\ \mathbf{Y}^h \end{bmatrix} \leq = \mathbf{b}^h \ ;$$

$$\mathbf{X}^h = \left[X_{u,w}^{v,h} \ , u, w \in \mathbf{N}_h, v \in \mathbf{V}_h \right], \ \mathbf{T}^h = \left[T_u^h \ , u \in \mathbf{P}_h \right], \mathbf{Y}^h = \left[Y_u^h \ , u \in \mathbf{P}_h \right]$$

where $\mathbf{Z}^h$ is the travel disutility associated with the travel pattern adopted by household h, $\mathbf{N}_h$ is the set of all nodes associated with household h, $X_{u,w}^{v,h}$ is a binary decision variable equal to unity if vehicle v of household h travels from activity u to activity w, and zero otherwise, T_u^h is the time at which participation in activity u of household h begins, Y_u^h is the total accumulation of either sojourns[1] or time spent away from home on any tour, of household h on a particular tour immediately following completion of activity u, $\mathbf{V}_h$ is the set of vehicles available to the household, and $\mathbf{A}^h$ is a matrix of spatial, temporal constants as well as the tour length limit. For more details regarding HAPP, readers are referred to Recker (1995) or Appendix 5-A.

[1] We have used the total accumulation of sojourns, and a maximum capacity of 4 (D = 4).

Throughout the dissertation, minor notations may differ across chapters. Depending on the type of integration or expansion from the base HAPP formulation, different notations are required. The chapter specific notations are included in the appendices. Data also differ across chapters. Although all data are drawn from Caltrans Travel Survey (Caltrans, 2001), selection of households and regional scope (Orange County and Los Angeles County) vary. Each chapter includes descriptions of its data set.

1.2 ADOPTION OF ALTERNATIVE FUEL VEHICLES

Since the adoption in 1990 of the California Air Resources Board's legislation of Low Emission Vehicle (LEV) and Zero Emission Vehicle (ZEV) mandates, there have been positive expectations of Alternative Fuel Vehicle (AFV) adaptation. Recently, concerns about rising gasoline cost, technical feasibility of the 'green' AFVs, the success of the Hybrid Electric Vehicles (HEVs) in the automobile market, and the government's effort in reduction of Greenhouse Gas (GHG) Emissions such as California's SB 32, achieving sustainable transportation system has never seemed more promising. Many recent assessments on energy and emission of AFVs have suggested positive outcomes. Two major advantages of the AFVs are their significant reductions in energy use and harmful emissions, considered to be two of the more significant automobile externalities (Parry et al., 2007). Obviously, the degree of these positive outcomes is dependent on the extent of AFV adoption.

However, even with these positive expectations of further adoption of AFVs, the survival of AFVs in the automobile market is not guaranteed. It has been postulated that there is a "sustainable" AFV market penetration threshold below which AFVs will not survive in the market. This is the so-called "chicken-and-egg" problem that explains the vehicle demand (purchase)–fuel infrastructure interaction. Such interactions are tested in Stephan and Sullivan (2004a, 2004b) and Schwoon (2007), and the condition of fuel infrastructure provision that would determine the survival of AFVs explored. The concept of "infrastructure" can be expanded more broadly to include such "social" factors as: word of mouth, social exposure, marketing, scale and scope economics, learning from experience, R&D, and innovation spillover (Struben and Sterman, 2008). In addition to the demand-infrastructure relationship, there are various factors that would determine the survival and the extent of AFV adoption. There are various stakeholders such as government, auto manufacturers, collaborators, competition, and activist groups, and impediments including regulatory barriers, resources, infrastructure and vehicle characteristics, as stated by Byrne and Polonsky (2001). With respect to all of the factors mentioned above, how effectively and economically we can overcome this "chicken-and-egg" status and get to the threshold point and maximize social benefits is the key.

In this dissertation, we take a different approach to incorporating driver's behavior by proposing an activity-based travel demand model to produce feasible travel patterns within the physical constraints of AFVs operations. On the demand side, inconvenience bounds of operating AFVs are assessed and on the supply side, vehicle-infrastructure interactions such as refueling location decisions are evaluated.

1.3 RESEARCH SUMMARY

The objective of this research is to expand the HAPP to include various locational decisions. As stated earlier, HAPP is a mathematical programming model, and it provides a convenient integration since many locational problems are formulated in mathematical programming. Through these integrations, it is possible to capture the interaction between individual driver's travel decisions and transportation infrastructure. A secondary objective is to apply the developed models to study transportation sustainability, specifically the adoption of AFVs.

In Chapter 2, the Location Selection Problem for the Household Activity Pattern Problem (LSP-HAPP) is developed. Destination choice is incorporated in the model where the destination choice decision is made from the interactions of other types of travel decisions made from HAPP. With this addition, five types of travel decisions are made simultaneously: activity allocation between household members, sequence of activities, departure times, some level of mode choice, and destination choice. A dynamic programming algorithm, developed for the Pick-up and Delivery Problem with Time Window Constraints (PDPTW) is adapted to handle a potentially sizable number of candidate locations. It is shown to be efficient for HAPP and LSP-HAPP applications. The algorithm is extended to retain arrival times as functions for mathematical programming formulations of activity-based travel models that often have time variables in the objective.

In Chapter 3, household-level inconvenience cost of operating an AFV is measured. Assuming that people do not change their participation in activities as done with their current conventional ICEVs, two scenarios are tested. For the upper bound of minimizing the travel inconvenience, LSP-HAPP is used for refueling application to observe the inconvenience of limited refueling infrastructure for HFCVs, and HAPP is extended to include battery engine inventory for BEVs. Regarding the lower bound, analysis of refueling trip insertion and analysis of delay – which keep the currently reported travel patterns, are used for HFCVs and BEVs respectively.

In Chapters 4 and 5, we focus on the supply side integration. In Chapter 4, Location-HAPP is developed, a Location Routing Problem which uses a set covering model as a facility location strategy, and HAPP as individual's travel demand procedure – vehicle routing strategy. It is shown that ignoring the interaction between service facilities and driver's travel demand procedure, leads to the minimum facility requirement given a guaranteed level of accessibility. This problem is specifically developed for a hydrogen refueling station siting application although such application can be applied to other types of services. Finally, the Location-HAPP model is applied to an early-adoption hydrogen community.

Chapter 5, we propose a general framework of demand-network integration, Network Design Problem integrated with the HAPP (NDP-HAPP). This is assumed by that OD is *not a priori* but a subject of responses of individual travelers to network level-of-service. A decomposed solution method is developed and found to have reasonable error rates (0 – 5%) and computation times. Suggested formulation is then applied to Southern California freeway network, and it captures temporal shifts of travel demand.

Chapter 6 summarizes all the chapters and presents the contributions and future work.

1.4 RESEARCH CONTRIBUTIONS

The primary contributions of this research include:

- The destination choice is integrated with HAPP and five types of travel decisions are made simultaneously: allocation of activities between household members, sequence of activities, departure time, some level of mode choice, and destination choice.

- The exact solution method is extended from the PDPTW to handle the computational burden of HAPP and LSP-HAPP.
- Inconvenience of operating AFVs is quantified as monetary cost, using full-day activity and travel data. A simulation of two behavioral assumptions create lower and upper bounds of inconvenience.
- Using the inconvenience measurement, a few policy guidelines are suggested.
- Location-HAPP is developed as a Location Routing Problem which takes a set covering model as a facility location strategy and HAPP as individual's travel demand procedure
- A search method that does not require the full matrix information is developed and tested for Location-HAPP.
- Location-HAPP shows that ignoring the interaction between service facilities and driver's travel demand procedure, leads to the minimum facility requirement given a guaranteed level of accessibility.
- Location-HAPP is applied to one of early-adoption hydrogen communities.
- NPD-HAPP proposes a general framework of demand-network integration with NDP-HAPP problem where OD is *not a priori*
- A decomposed solution method for NDP-HAPP is developed and found to have reasonable error rates (0 – 5%) and computation times.

CHAPTER 2 THE LOCATION SELECTION PROBLEM FOR THE HOUSEHOLD ACTIVITY PATTERN PROBLEM

Individual- or household-level destination choice is not an output of optimizing a single objective but rather is a complex decision-making process involving a multitude of issues related to such aspects as type of activity, personal preference, accessibility, time-of-day, trip chaining, mode choice and etc. For this reason, destination choice modeling has been studied within the context of associations with those influencing factors. Although there are other approaches to model destination choice (Gärling and Axhausen, 2003; Louviere and Timmermans, 1990), most of the work in this area has modeled destination choice using discrete choice analysis based on random utility theory.

Many trip-based single destination choice studies have focused on the influences of type of activity. A few of the papers in this category are Bhat et al. (1998) – work and shopping, Fotheringham (1988), Recker and Kostyniuk (1978) – grocery shopping, and Pozsgay and Bhat (2001) – recreational trip destination. In more fundamental approaches relative to how travel decisions are made, discrete choice models of destination choice have been integrated into tour-based approaches, involving such considerations as proximity to other activity locations, travel time and duration, etc. Such considerations are particularly important in analyzing destination choice associated with non-primary activities that people tend to include in tours with other activities. Kitamura (1984) included a zone attraction component within trip chaining behavior that included considerations of locations of home and other activities within trip chains, but his approach was limited in that trip chaining sequence, time-of-day, and selection of activities in a tour are static. Bowman and Ben-Akiva (2000) proposed integrated activity-based demand modeling including destination choice as well as types of pattern, travel mode, time-of-day, etc.

Here, we propose an integrated approach similar to Bowman and Ben-Akiva (2000), based on a scheduling and routing framework, HAPP, for daily activities that includes a capability of modeling the selection of activity locations, time-of-day, pattern types, and choice of personal travel modes (e.g., automobile, bicycle, walk). In the formulation, destination choices for certain activities (i.e., those without fixed locations) are viewed not as a primary choice that travelers make, but rather as an auxiliary choice made within their daily schedule and routing.

There are a number of potential practical advantages that the properties mathematical programming models, compared to discrete choice analysis, offer in application to activity-based travel demand. Principal among these is that such temporal constraints as the open hours of a particular shopping destination, or such spatial-temporal constraints as the space-time prism associated with an activity at particular location is insufficient to permit performance of a subsequent activity, that may be placed on travel/activity decisions can be incorporated explicitly, rather than be implied in the predefined specification of the set of discrete alternatives. For example, in the nested logit model example from Bowman and Ben-Akiva (2000), each decision nest needs pre-defined alternative choice sets, leading to 54 possible outcomes (discrete alternatives). Although infeasible decisions need to be addressed via constraints (which implicitly may nonetheless be enumerated as part of the solution algorithm), it is not required to pre-define all sets of actions—such as types of activity patterns, time-of-day, destination choice, composition of activities in each tour, and etc.—that are possible. Another (obvious) advantage of mathematical programming models is their ability to handle decisions involving both continuous (time) as well as discrete (location) variables. Additionally, because discrete choice model estimation allows for only a relatively small number of alternatives, with the alternative destination set universal for all individuals (although specific individuals typically may not include all alternatives in their respective choice sets), specification must be defined either to meet pre-specified requirements, or be randomly sampled. This aspect makes discrete choice analysis in application to destination choice particularly limiting in its ability to represent individual choices. For more discussion and literature review on choice-set generation sub-problem of destination choice modeling based on discrete choice analysis, refer to Thill (1992).

Of course, there are also significant disadvantages associated with the current state of mathematical programming approaches to activity-based travel/activity modeling, many of which are enumerated by Recker (2001) who showed that conventional discrete transportation choice models (e.g., destination, route, mode) can be represented as a special case of the HAPP family of mathematical programming models. In essence, both approaches are based on utility maximization principles applied at the individual (or disaggregate level), the principal differences being that the discrete choice case involves an unconstrained optimization of discrete choices based on specification of utility in terms of continuous and/or discrete variables with a specified error structure, while the mathematical programming case involves a constrained optimization of both continuous and discrete variables based on specification of utility in terms of continuous and/or discrete variables with no assumed error structure. The specification of the error structure in discrete choice models is conducive to estimation by standard maximum likelihood techniques, while the lack of such has presented a challenge to moving mathematical programming approaches toward being descriptive (and, ultimately, predictive) from being merely proscriptive; recent advances based on inverse optimization techniques (Chow and

Recker, 2011) and genetic algorithms (Recker et al., 2008) have made progress toward estimation. And, as a constrained generalization of the discrete choice case, the mathematical programming modeling approach actually generally greatly increases the dimension of the choice set alternatives over that of discrete modeling approaches, but shifts the burden of the increased dimensionality to the solution algorithm rather than to the specification of the model choice alternatives; this can present a serious obstacle since mathematical programming models such as HAPP are known to be np-hard. Despite these disadvantages, the advantages that mathematical programming models offer in guaranteeing the internal consistency of the linkages dictated by time-space constraint considerations are deemed an avenue of research of potential benefit in modeling complex travel choices.

In this chapter, we extend the basic HAPP formulation to the case involving a choice of selecting a location from many candidate locations for performance of a desired activity. As described above, a structural advantage that HAPP provides is a flexible form for incorporating new behavioral aspects while maintaining the consistency of inviolable rules governing construction of activity patterns that are ensured by the mathematical formulation of the basic HAPP model—extensions can be easily built from the basic formulation. Although the basic formulation for the Location Selection Problem (LSP) is easily obtained from the HAPP formulation by expanding the constraints that specify that only one location of each activity type is to be visited, the size and the complexity of the problem become an issue due to the various possible locations within the range of one's spatial and temporal accessibility—computational limitations have been an obstacle that makes it difficult for even the basic HAPP model to reflect realistic travel behaviors in the model. Fortunately, the PDPTW on which the model is based has been studied extensively, and numerous algorithms to handle large-scale problems have been offered. Here, we adopt methodology incorporating dynamic programming algorithms with path eliminations developed by Desrosiers et al. (1986) and Dumas et al. (1991), with suitable modifications to meet the requirements of the Location Selection Problem. The Location Selection Problem for the Household Activity Pattern Problem presented here can handle a larger number of alternative locations, without the additional step of generation of specific alternative destination sets.

2.1 LSP-HAPP FORMULATION

In the most general formulation of the Location Selection Problem for the Household Activity Pattern Problem (LSP-HAPP), we presume that there are activities with specified locations, as well as activities with no specific location—there exist a number of candidate sites for each such activity type (total of m activities), that are scheduled to be completed by the household. Specifically, we assume that among the activities scheduled for completion by the household are those for which the locations are predetermined (e.g., work, school) and some for which the location can be selected from a number of candidate locations (e.g., grocery shopping). In the HAPP analogy to the PDPTW, activities are viewed as being "picked up" by a particular household member (who, in this basic case, is uniquely associated with a particular vehicle) at the location where performed and, once completed (requiring a service time s_i) are "logged in" or "delivered" on the return trip home. Multiple "pickups" are synonymous with multiple sojourns on any given tour. The scheduling and routing protocol relative to some household objective produces the "time-space diagram" commonly referred to in travel/activity analysis.

Decision variables, directly analogous to those of the PDPTW, are defined as (see Appendix 2-A for notation used):

$X_{uw}^{v},\ u,w \in \boldsymbol{N}, v \in \boldsymbol{V}, u \neq w$ binary decision variable equal to unity if vehicle υ travels from activity u to activity w, and zero otherwise.

$T_u,\ u \in \boldsymbol{P}$ the time at which participation in activity u begins.

$T_0^v, T_{2n+1}^v,\ v \in \boldsymbol{V}$ the times at which vehicle υ first departs from home and last returns to home, respectively.

$Y_u,\ u \in \boldsymbol{P}$ the total accumulation of either sojourns or time (depending on the selection of D and d_u) on a particular tour immediately following completion of activity u.

With these definitions, the LSP-HAPP (the Location Selection Problem for the Household Activity Pattern Problem) for a household's completion of a set $\boldsymbol{M_P} = \{1,2,\dots,i,\dots,n_P\}$ of n_P out-of-home activities with pre-selected (one-to-one) locations $\boldsymbol{P_P^+} = \{1,2,\dots,i,\dots,n_P\}$ and a set $\boldsymbol{A} = \{\boldsymbol{A}_1,\dots,\boldsymbol{A}_a,\dots,\boldsymbol{A}_m\}$ of out-of-home activities of specific types (e.g., grocery shopping) $\boldsymbol{A_a}$, each of which with $n_{\boldsymbol{A_a}}$ possible corresponding locations

$\boldsymbol{P}^+_{\boldsymbol{A_a}} = \{1,2,\dots i,\dots,n_{\boldsymbol{A_a}}\}$, using mode of travel υ, can be represented by the following formulation.[2]

$$Minimize\ Z = \ Household\ Disutility \tag{2-1}$$

subject to:

$$\sum_{\upsilon\in V}\sum_{w\in N} X^{\upsilon}_{uw} = 1, \quad u \in \boldsymbol{P}^+_{\boldsymbol{P}} \tag{2-2}$$

$$\sum_{\upsilon\in \boldsymbol{V}}\sum_{u\in \boldsymbol{P}^+_{A_a}}\sum_{w\in \mathbf{N}} X^{\upsilon}_{uw} = 1, \quad \boldsymbol{A_a} \in \boldsymbol{A} \tag{2-3}$$

$$\sum_{w\in \mathbf{N}} X^{\upsilon}_{uw} - \sum_{w\in \mathbf{N}} X^{\upsilon}_{wu} = 0 \quad u \in \boldsymbol{P}, \upsilon \in \mathbf{V} \tag{2-4}$$

$$\sum_{w\in \mathbf{P}^+} X^{\upsilon}_{0w} \le 1 \ , \quad \upsilon \in \mathbf{V} \tag{2-5}$$

$$\sum_{u\in \mathbf{P}^-} X^{\upsilon}_{u,2n+1} - \sum_{w\in \mathbf{P}^+} X^{\upsilon}_{0w} = 0 \ , \quad \upsilon \in \mathbf{V} \tag{2-6}$$

$$\sum_{w\in \mathbf{N}} X^{\upsilon}_{wu} - \sum_{w\in \mathbf{N}} X^{\upsilon}_{w,n+u} = 0 \quad u \in \boldsymbol{P}^+, \upsilon \in \mathbf{V} \tag{2-7}$$

$$T_u + s_u + t_{u,n+u} \le T_{n+u} \quad u \in \boldsymbol{P}^+_{\boldsymbol{P}} \tag{2-8-1}$$

$$X^{\upsilon}_{uw} = 1 \Rightarrow T_u + s_u + t_{u,n+u} \le T_{n+u} \quad u \in \boldsymbol{P}^+_{\boldsymbol{A}}, w \in \mathbf{N}, \upsilon \in \mathbf{V} \tag{2-8-2}$$

$$\sum_{\upsilon\in \boldsymbol{V}}\sum_{w\in \mathbf{N}} X^{\upsilon}_{uw} = 0 \Rightarrow T_u = T_{u+n} = 0 \quad u \in \boldsymbol{P}^+_{\boldsymbol{A}} \tag{2-8-3}$$

[2] LSP-HAPP is different from selective pickup and delivery problem in that there is no utility associated with visiting a location, and that only one of the same types of location can (and must) be visited.

$$X_{uw}^{v} = 1 \Rightarrow T_u + s_u + t_{uw} \leq T_w \;\;,\;\; u, w \in \mathbf{P},\, v \in \mathbf{V} \tag{2-9}$$

$$X_{0w}^{v} = 1 \Rightarrow T_0^{v} + t_{0w} \leq T_w \;\;,\;\; w \in \boldsymbol{P}^{+},\, v \in \mathbf{V} \tag{2-10}$$

$$X_{u,2n+1}^{v} = 1 \Rightarrow T_u + s_u + t_{u,2n+1} \leq T_{2n+1}^{v} \;\;,\;\; u \in \boldsymbol{P}^{-},\, v \in \mathbf{V} \tag{2-11}$$

$$a_u \leq T_u \leq b_u \;\;,\;\; u \in \boldsymbol{P}_{\boldsymbol{P}}^{+} \tag{2-12-1}$$

$$X_{uw}^{v} = 1 \Rightarrow a_u \leq T_u \leq b_u \;\;,\;\; u \in \boldsymbol{P}_{\boldsymbol{A}}^{+}, u \in \mathbf{N},\, v \in \mathbf{V} \tag{2-12-2}$$

$$X_{0w}^{v} = 1 \Rightarrow a_0 \leq T_0^{v} \leq b_0 \;\;,\;\; w \in \boldsymbol{P}^{+},\, v \in \mathbf{V} \tag{2-13-1}$$

$$\sum_{w \in \mathbf{P}^{+}} X_{0w}^{v} = 0 \Rightarrow T_0^{v} = 0 \;\;,\;\; v \in \mathbf{V} \tag{2-13-2}$$

$$X_{\mathrm{u},2n+1}^{v} = 1 \Rightarrow a_{2n+1} \leq T_{2n+1}^{v} \leq b_{2n+1} \,,\;\; u \in \boldsymbol{P}^{-}, v \in \mathbf{V} \tag{2-14-1}$$

$$\sum_{u \in \mathbf{P}^{-}} X_{\mathrm{u},2n+1}^{v} \Rightarrow T_{2n+1}^{v} = 0 \;\;,\;\; v \in \mathbf{V} \tag{2-14-2}$$

$$X_{uw}^{v} = 1 \Rightarrow Y_u + d_w = Y_w, u \in \boldsymbol{P},\;\; w \in \boldsymbol{P}^{+}\; v \in \mathbf{V} \tag{2-15}$$

$$X_{uw}^{v} = 1 \Rightarrow Y_w - d_w = Y_w,\;\; u \in \boldsymbol{P},\;\; w \in \boldsymbol{P}^{-}\; v \in \mathbf{V} \tag{2-16}$$

$$X_{0w}^{v} = 1 \Rightarrow Y_0 + d_w = Y_w,\;\; w \in \boldsymbol{P}^{+},\, v \in \mathbf{V} \tag{2-17}$$

$$Y_0 = 0 \;\;,\;\; 0 \leq Y_u \leq D,\;\; u \in \boldsymbol{P}^{+} \tag{2-18}$$

$$\sum_{v \in \boldsymbol{V}} \sum_{u \in \mathbf{N}} \sum_{w \in \mathbf{N}} c_{uw}^{v} X_{uw}^{v} \leq B_c \tag{2-19}$$

$$\sum_{u \in \mathbf{N}} \sum_{w \in \mathbf{N}} t_{uw}^{v} X_{uw}^{v} \leq B_t^{v} \tag{2-20}$$

$$X_{uw}^{v} = \begin{cases} 0 \\ 1 \end{cases},\;\; u, w \in \mathbf{N}, v \in \mathbf{V} \tag{2-21}$$

$T_u \geq 0, u \in \boldsymbol{P}$ (2-22)

$T_0^v, T_{2n+1}^v \geq 0, v \in \mathbf{V}$ (2-23)

The constraints that specify that each activity location needs to be visited (performed) are split into two sets of constraints. Equations (2-2) impose the condition that there is one and only one path leading from each activity with pre-selected location. Equations (2-3) impose the condition that there is one and only one path leading from one and only one type $\mathrm{A_a}$ out-of-home activity location. This can be viewed as a Generalized Vehicle Routing Problem suggested by (Ghiani and Improta, 2000). The rest of the formulation follows the classical PDPTW, and the base case HAPP, except for a few conditional constraints to relax constraints on unselected candidate nodes. Equations (2-4) ensure that that there is a connected path among the activities (and their return trips to home) and that no activity is revisited. Equations (2-5) allow for the possibility that some of the vehicles in the household's stable of vehicles may not be used. Equations (2-6) enforce a restriction similar to that in Equations (2-2), but with reference to the paths leading from the origin and to the final termination (i.e., the depot). Equations (2-7) stipulate that the return-home trip be on the same path as it's associated out-of-home activity. The original equation (2-8), $\mathrm{T_u} + \mathrm{s_u} + \mathrm{t_{u,n+u}} \leq \mathrm{T_{n+u}}, \ \ \mathrm{u} \in \boldsymbol{P}^+$, is a restriction that the activity start times for elements of $\boldsymbol{P}^+$ precede those of corresponding elements in $\boldsymbol{P}^-$ (the end point, home location, of the connected graph defining the path from the location of performance of an activity to the ultimate trip to the home location). However, for LSP-HAPP, this constraint needs to be satisfied only if the solution includes visiting that specific node among many candidates as in (2-8-2). Similarly, when the objective function involves time variables, the time variables for the unvisited activity nodes need to be constrained in order not to affect the objective function as in (2-8-3). Equation (2-9) is the restriction that the commencement time of the activity associated with any trip end w, i.e., T_w, requiring travel from another trip end u can occur no sooner than the termination time of the corresponding activity at u plus the travel time from the site of activity u to the site of activity w. Equations (2-10) and (2-11) state that restrictions similar to those imposed by Equation (2-9) hold for travel from the origin node, 0, to any activity, as well as for travel from any activity to its "return home" activity. Equations (2-12) state that each activity and the selected node needs to start within its given time windows. This equation is modified from the original constraint, $a_u \leq T_u \leq b_u$, $u \in \mathbf{P}$, to be satisfied only when the node is visited for the selective locations. Equations (2-13), and (2-14) add restrictions regarding the time windows available for activity completion. For the case in which the vehicle does not operate for the given day, its time windows need to be set to zero, so as not to affect the objective function. Equations (2-15) through (2-18) impose conditions on the maximum number of sojourns allowed in any single tour. Equations (2-19) and (2-20) enforce budget constraints. Equations (2-21) (2-22) (2-23) add non-negativity and integer constraints.

2.2 SOLUTION METHODOLOGY

As noted, HAPP is an NP-hard problem; for a total number of all activities— with pre-selected locations plus the number of candidate locations for activities with alternative candidate locations—of n, the number of flow decision variables is $(2n+2)^2$. As such, its application faces significant challenges imposed by computational limitations. All HAPP cases examined previously in the literature have had only a few activities. Application of LSP-HAPP to cases involving multiple vehicles with numbers of activities having a large number of candidate locations within one's spatial and temporal accessibility seriously stretches this computational limitation.

Numerous algorithms have been developed to solve large-size PDPTW (see, e.g., Cordeau and Laporte, 2003), and problems with locations up to 2,500+ have been successfully solved. Here, we follow the solution method proposed by Dumas et al. (1991), which was used to solve large scale PDPTW, and modify it to meet the specifications of LSP-HAPP problem. In their approach, an exact dynamic forward programming routine in a sub-problem is used to generate possible and feasible paths, and then combinations of these paths are decided in the master problem to assign each path to each vehicle.

It has been shown that the arc-path notation's sub-problem to generate admissible paths in the multi-commodity problem is the shortest path problem (Ford and Fulkerson, 1958). Since LSP-HAPP equations (2-1)-(2-7) form a multi-commodity problem, we can rewrite in arc-path formulation as the following:

$$minimize \sum_{r\in\Psi} c_r Y_r \qquad (2-a)$$

$$\sum_{r\in\Psi} a_{ir} Y_r = 1, \qquad i \in \boldsymbol{P_P^+} \qquad (2-\mathrm{b1})$$

$$\sum_{i\in\boldsymbol{P_{A_a}^+}} \sum_{r\in\Psi} a_{ir} Y_r = 1 \quad \boldsymbol{A}_a \in \boldsymbol{A} \quad (2-\mathrm{b2})$$

$$\sum_{r\in\Psi} Y_r \leq |\mathrm{V}| \qquad (2-\mathrm{c})$$

where

Ψ: the set of admissible paths

$$Y_r = \begin{cases} 1 & \text{if path } r \text{ is used} \\ 0 & \text{otherwise} \end{cases}, \quad r \in \Psi$$

$$a_{ir} = \begin{cases} 1 & \text{if path } r \text{ includes activity node } i \\ 0 & \text{otherwise} \end{cases}, \quad i = 1, \dots, n, \quad r \in \Psi$$

c_r: the cost of route r, $r \in \Psi$

Here, r is an admissible path for a given vehicle/household member, v, that satisfies all of the properties of the problem as specified in the remaining Equations (2-8) - (2-22). Equations (2-b1) and (2-b2) are substituted for the original constraint of PDPTW arc-path formulation for the Location Selection Problem:

$$\sum_{r \in \Psi} a_{ir} Y_r = 1, \qquad i \in \boldsymbol{P}^{+} \qquad (2-\mathrm{b})$$

Equations (2-b1) constrain that all activities with pre-selected location need to be visited once. And Equations (2-b2) constrain that one and only one of the candidate locations for each activity type with multiple candidate locations needs to be visited once and only once.

Variable a_{ir} shows whether each activity node i is on path r. Then the column vector $[a_{ir}, a_{ir}, \dots, a_{ir}]^T$ shows all the activity nodes that the path r covers. Therefore, by finding an admissible path r, we are performing the column generation, which is widely used for large-scale combinatorial optimization problems. For arc-path formulation of PDPTW, the sub-problem (the dual problem) to find admissible path r is the shortest path problem with time windows. For LSP-HAPP, the sub-problem becomes LSP-adaptation of the shortest path problem with time windows

This sub-problem of finding r of LSP-adaptation from the shortest path problem with time windows can be solved by the following dynamic programming algorithm (Algorithm 2-1, shown below), which is adapted from Dumas et al. (1991) and Desrosiers et al. (1986), and follows notations used in Desrosiers et al. (1986), i.e.,

state (S, i): a feasible route to node i, the terminal node, that visits all the nodes in S $\subseteq$ **P**, and $i \in$ S. S is a non-ordered set of cardinality k, where k is the iteration number.

$(\mathrm{S}_\alpha, \mathrm{i})$: a given route α to state (S, i)

$t(\mathrm{S}_\alpha, \mathrm{i})$: the arrival time at node i, following route α

$c(\mathrm{S}_\alpha, \mathrm{i})$: the current cost at node i, following route α

$d(\mathrm{S}_\alpha, \mathrm{i})$: the cumulative number of sojourns in a tour at node i, following route α

Initialization ($k = 1$)

A set of states of routes visiting one activity node from home location are generated.

$$\{(\{j\}, j),\ j \in \boldsymbol{P}^{+}\}$$

Corresponding arrival time, cumulative cost, and cumulative number of sojourns in a tour are updated as:

$$T(\mathrm{S}_{\boldsymbol{\alpha}}, j) = \{\ \max\left(a_j, a_0 + t_{0,j}\right) \le t(\mathrm{S}_{\boldsymbol{\alpha}}, j) \le b_j,$$
$$t(\mathrm{S}_{\boldsymbol{\alpha}}, 0) + t_{0,j} \le\ t(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j})\ \}$$
$$c(\mathrm{S}_{\boldsymbol{\alpha}}, j) = c_{0,j}$$
$$d(\mathrm{S}_{\boldsymbol{\alpha}}, j) = d_i$$

Recursion ($2 \le k$)

New states are constructed by adding one node, j, to the total visited at the preceding iteration:

$\{(S\cup\{j\}, j),\ j \in \boldsymbol{P} \cup \{2n+1\}\}$ where (S, i) is the state from previous iteration

Then the states are tested for elimination criteria, and if the state $(S\cup\{j\}, j)$ is not eliminated, its label set will be created. Its corresponding arrival time, cumulative cost, and cumulative number of sojourns in a tour are updated as:

$$T(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, j) = T(\mathrm{S}_{\boldsymbol{\alpha}}, j) \cup \{\ \max\left(a_j, T_i + s_i + t_{i,j}\right) \le t(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j}) \le b_j,$$
$$t(\mathrm{S}_{\boldsymbol{\alpha}}, \mathrm{i}) + s_i + t_{i,j} \le\ t(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j})\}$$
$$c(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j}) = c(\mathrm{S}_{\boldsymbol{\alpha}}, i) + c_{i,j}$$
$$d(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j}) = d(\mathrm{S}_{\boldsymbol{\alpha}}, i) + d_j$$

Stop when there is no label generated at this iteration.

Selection of Arrival Times

For all completed paths, $\boldsymbol{\alpha}$, solve the following optimization problem, and update the final cost.

Minimize $f(T_0, T_1, \dots, T_{2n}, T_{2n+1})$ such that $T(\mathrm{S}_{\boldsymbol{\alpha}}, 2n+1)$

Algorithm 2-1 LSP-HAPP path generation algorithm for objective function involving time variables

Here we have extended the algorithm so that only one of the candidate locations is visited for activity types without pre-selected locations as constrained in LSP-HAPP Equation (2-3), and introduce new elimination criteria to support such patterns—a method that works well for large-scale problems. Although similar to the shortest path problem addressed by the algorithm presented by Dumas et al. (1991) and Desrosiers et al. (1986), the problem considered by LSP-HAPP (as well as by other HAPP-based formulations) differs in an important aspect that requires attention before the algorithm can be applied. It is often the case that the actual time selected for performance of an activity (within an acceptable time window) influences the net utility (utility of the activity less the travel disutility) one experiences. In the algorithm proposed by Desrosiers et al. (1986) and Dumas et al. (1991), the earliest possible arrival time is selected for T_u. To the contrary, arriving at an activity at its earliest possible arrival time may result in out-of-home wait time delays (waiting for the next activity window to become available) in completing other scheduled activities that may lead to reduced utility. This aspect is more critical for LSP-HAPP than for PDPTW since activity start (return home) time windows are not homogeneous compared to pick up (delivery) time windows of PDPTW. Indeed, such factors as time being outside of home, or delay time in starting an activity have been found to play a role in personal activity patterns (Chow and Recker; 2012, Recker et al; 2008).

To address these issues, first the objective function is separated into two parts—one as a function of flow decisions (e.g. $\sum_{v\in \boldsymbol{V}}\sum_{i\in \boldsymbol{N}}\sum_{j\in \boldsymbol{N}} c_{ij}X_{ij}$ or $\sum_{v\in \boldsymbol{V}}\sum_{i\in \boldsymbol{N}}\sum_{j\in \boldsymbol{N}} t_{ij}X_{ij}$), and the other as a function of arrival times (e.g., $\sum_{v\in \boldsymbol{V}}(T_{2n+1}^v - T_0^v)$); e.g.,

$$Minimize\ Z = \sum_{v\in \boldsymbol{V}}\sum_{i\in \boldsymbol{N}}\sum_{j\in \boldsymbol{N}} c_{ij}X_{ij} + \sum_{v\in \boldsymbol{V}} f(T_0^v, T_1, \dots, T_{2n}, T_{2n+1}^v)$$

The first part of the objective which is affected by path sequence is updated according to the original algorithm. The other part, which is dependent on activity start (arrival) times cannot be updated at each iteration because the optimal arrival time may not be determined during the process of creating paths, and also because variables may not have been defined yet; this part is left to be assessed by a final procedure. Instead, we define a new set to represent arrival times as a function,

$T(S_{\alpha}, j)$: set of arrival times windows of all activities in $\mathrm{S} \cup \{0\}$, following path $\boldsymbol{\alpha}$

and during *Recursion* $(2 \leq k)$, a label is created with possible time windows of arrival time determined as $\max(a_j, T_i + s_i + t_{i,j}) \leq t(\mathrm{S} \cup \{j\}_{\alpha}, \mathrm{j}) \leq b_j$. Conditions respecting the path sequence are as $t(\mathrm{S}_{\alpha}, \mathrm{i}) + s_i + t_{i,j} \leq t(\mathrm{S} \cup \{j\}_{\alpha}, \mathrm{j})\}$. The feasibility of arrival time windows needs to be delivered as well as previous time windows of arrival times. Then, $c(\mathrm{S}_{\alpha}, \mathrm{i})$, the objective measure affected by path sequence at node i following route $\boldsymbol{\alpha}$, is updated in the same manner as in the a original algorithm, i.e., $c(\mathrm{S} \cup \{j\}_{\alpha}, \mathrm{j}) = c(\mathrm{S}_{\alpha}, i) + c_{i,j}$. For the elimination criteria involving possible time window violations, T_i is assumed to be the earliest possible time.

Once all feasible paths are created, arrival times are decided by minimizing the objective function while respecting time windows created along the paths generated. This is a problem of finding the optimal value with arrival time decisions given a path sequence, in the form of a linear programming problem of n variables, and solved very easily. By leaving time variables as a function, the algorithm loses some of the simplicity since several linear programs need to be solved during the final step of the procedure, but it allows specification of objectives in terms of time variables. (If objectives are in terms of load variables, the same approach can be used). Then, the path with the smallest objective function is selected, and the same limit of problem size is guaranteed as those of the original algorithms.

State elimination criteria are employed to efficiently reduce the size of path combinations needed to explore. At the beginning of each recursion iteration k, all combinations of $\{(S\cup\{j\}, j), j \in \boldsymbol{P}\}$ are tested relative to whether to be stored or eliminated. Some elimination criteria are based solely on the feasibility of $(S\cup\{j\}, j)$, and some elimination criteria also consider the terminal node i of the previous path (S, i) from previous iteration $k-1$:

Elimination criteria

#1: node j must not have been previously visited:

$$j \in \bar{S}$$

#2: if node j is one of the candidate locations for activity type $\boldsymbol{A_a}$, then any candidate location of activity $\boldsymbol{A_a}$ must not have been previously visited. This elimination is tested for all selective activity types, $\boldsymbol{A}_a \in \boldsymbol{A}$:

For all $\boldsymbol{A}_a \in \boldsymbol{A}$, if $j \in \boldsymbol{P}^+_{\boldsymbol{A_a}}$, then $l \in \bar{S}$, for all $l \in \boldsymbol{P}^+_{\boldsymbol{A_a}}$ and $l \neq j$

#3: if node j is one of the return home locations for activity type $\boldsymbol{A_a}$, then any return home location for activity type $\boldsymbol{A_a}$, must not have been previously visited. This elimination is tested for all selective activity types, $\boldsymbol{A}_a \in \boldsymbol{A}$:

For all $\boldsymbol{A}_a \in \boldsymbol{A}$, if $j \in \boldsymbol{P}^-_{\boldsymbol{A}_a}$, then $l \in \bar{S}$, for all $l \in \boldsymbol{P}^-_{\boldsymbol{A}_a}$ and $l \neq j$

#4: if node j is a return home node, then the activity node, $j - n$ must have been previously visited (precedence constraint):

if $j \in \boldsymbol{P}^-$, then $j - n \in S$

#5: if node j is an activity node, total number of sojourns (cumulative time away from home) must not exceed the maximum number of sojourns (time away from home) allowed in a tour:

if $j \in \boldsymbol{P}^+$, then $d(\mathrm{S}_{\boldsymbol{\alpha}}, i) + d_j \leq D$

#6: time constraints must be respected:

$T_i + s_i + t_{i,j} \leq b_j$

#7: for $i \in \boldsymbol{P}^+$, $j \in \boldsymbol{P}^+$, one of paths, $i \to j \to n + \ i \to n + j$ or $i \to j \to n + j \to n + i$, must be feasible with time $T_i = a_i$, which is the earliest time at which node i can be visited.

#8: for $i \in \boldsymbol{P}^-$, $j \in \boldsymbol{P}^-$, one of paths, $i - n \to j - n \to \ i \to j$ or $j - n \to i - n \to \ i \to j$, must be feasible with time $T_i = a_i$, and $T_j = a_j$, which is the earliest time at which node i, j can be visited.

#9: for $i \in \boldsymbol{P}^+$, $j \in \boldsymbol{P}^-$, path $j - n \to i \to j \to n + i$ must be feasible with time $T_{j-n} = a_{j-n}$, which is the earliest time at which node $j - n$ can be visited.

#10: for $i \in \boldsymbol{P}^-$, $j \in \boldsymbol{P}^+$, path $i - n \to i \to j \to n + j$ must be feasible with time $T_{i-n} = a_{i-n}$, which is the earliest time at which node $i - n$ can be visited

#11: if node i is the final home node, then cannot expand a path from this path:

$i \neq 2n + 1$

#12: if node j is the final home node, then the final visited node i must be one of the return home nodes:

if $j = 2n + 1$, then $i \in \boldsymbol{P}^{-}$

#13: if node j is the final home node, then for all the activity location nodes that are visited, l, all of the corresponding return home nodes must have been visited:

if $j = 2n + 1$, then $n + l \in \boldsymbol{S}$ for all $l \in \boldsymbol{P}^{+}$ and $l \in \boldsymbol{S}$

Criteria #2 and #3 are introduced to meet the specifications of LSP- HAPP. The rest of the label generating criteria are from Dumas et al. (1991) and Derosiers et al. (1986). Criteria #7 - #10 tighten criteria #6 with possible time window violations to reduce the number of label generations. The efficiency of dynamic programming is dependent on how efficient these elimination criteria are.

Additionally, since the physical location of all return nodes is home for the LSP-HAPP application, it is not meaningful to identify the order of visiting those nodes during Recursion. This drastically reduces the number of labels to be created.

14: if all pre-selected locations (all $l \in \boldsymbol{P}_{\boldsymbol{P}}^{+}$) and one of the selective locations (any $l \in \overline{\boldsymbol{P}_{\boldsymbol{A}_a}^{+}}, \boldsymbol{A}_{\boldsymbol{a}} \in \boldsymbol{A}$) have been visited previously, and the arrival node j is home (if $l \in S$ and $j \in \boldsymbol{P}^{-}$),

create the new label and terminate *Recursion* from this label, add the rest of return home trips of all the visited nodes if missing, and pass the label to *Final Iteration*.

$l \in S$ for all $l \in \boldsymbol{P}_{\boldsymbol{P}}^{+}$, and $l \in S$ for any $l \in \boldsymbol{P}_{\boldsymbol{A_a}}^{+}$ for all $\boldsymbol{A}_a \in \boldsymbol{A}$, and $j \in \boldsymbol{P}^{-}$

Patterns generated by the algorithm are now introduced to the master problem, (a) – (c), and solved. It is noted that the information on path cost, arrival time, and load are not carried onto the master problem. Those data need to be stored separately.

2.3 EXAMPLES

2.3.1 Case1: Grocery Shopping Location Selection Involving a Single Vehicle

As an example of the application of this basic LSP-HAPP formulation, we consider the case of a household with one vehicle that is available for travel to any activity beginning at 6:00 and ending at 20:00, but must return to home from any activity no later than 21:00. The household has one work activity with a fixed location, i.e., $\mathbf{M}_P = \{1\}$, $\boldsymbol{P}_{\boldsymbol{P}}^{+} = \{1\}$; $n_P = 1$, with duration of $s_1 = 9$ hours and start time availability windows between 8:00 and 9:00 and no additional constraint on returning home from the work activity. Assume further that the household also has a grocery shopping trip to be scheduled; i.e., $\boldsymbol{A} = \{\boldsymbol{A_1}\}$, $m = 1$, and that there are two potential locations for this activity $\boldsymbol{P}_{\boldsymbol{A_1}} = \{2, 3\}$; $n_{A_1} = 2$; the operation hours for both stores is assumed to be from 6:00 to 22:00 and the duration of the shopping activity at either location is 1 hour[3]. In this example:

$$\boldsymbol{M} = \boldsymbol{M_p} \cup \boldsymbol{A} = \{1, 2, 3\};\ n = n_P + n_A = 3$$

$$\boldsymbol{P_P^+} = \{1\}$$

$$\boldsymbol{P_A^+} = \{2, 3\}$$

$$\boldsymbol{P_P^-} = \{4\}$$

$$\boldsymbol{P_A^-} = \{5, 6\}$$

$$\boldsymbol{P^+} = \boldsymbol{P_P^+} \cup \boldsymbol{P_A^+} = \{1, 2, 3\}$$

$$\boldsymbol{P^-} = \boldsymbol{P_P^-} \cup \boldsymbol{P_A^-} = \{4, 5, 6\}$$

$$\boldsymbol{P_P} = \boldsymbol{P_P^+} \cup \boldsymbol{P_P^-} = \{1, 4\}$$

$$\boldsymbol{P_A} = \boldsymbol{P_A^+} \cup \boldsymbol{P_A^-} = \{2, 3, 5, 6\}$$

with time availability windows, and corresponding return-home windows:

[3] Although assumed identical in this particular example, durations and/or time widows at the various locations need not be.

$$[a_i, b_i] = \begin{bmatrix} a_1, b_1 \\ a_2, b_2 \\ a_3, b_3 \end{bmatrix} = \begin{bmatrix} 8, 9 \\ 6, 21 \\ 6, 21 \end{bmatrix} \quad , [a_{n+i}, b_{n+i}] = \begin{bmatrix} a_4, b_4 \\ a_5, b_5 \\ a_6, b_6 \end{bmatrix} = \begin{bmatrix} 6,\ 21 \\ 6,\ 22 \\ 6,\ 22 \end{bmatrix} \quad ,$$

$$[a_0, b_0] = [6, \quad 20]$$

$$[a_{2n+1}, b_{2n+1}] = [a_{13}, b_{13}] = [6, \quad 21] \quad .$$

In this example, the household's objective function is assumed to be that of minimizing the total monetary cost—that is, total travel time multiplied by fuel cost (first term)—plus the value of the extent of the travel day (second term).

$$F \cdot \sum_{u \in \boldsymbol{N}} \sum_{w \in \boldsymbol{N}} t_{uw}^{v} X_{uw}^{v} + V \cdot [T_{2n+1} - T_0]$$

where V, F respectively are the monetary value of the temporal extent of the travel day, fuel cost per hour (derived from assumed average speed and miles per gallon). For purposes of illustration, in our example, we arbitrarily set $V = \$15/hr$, $F = \$6.25/hr$.

During recursion iterations, cost is simply updated as, $c(\mathrm{S} \cup \{j\}_{\boldsymbol{\alpha}}, \mathrm{j}) = c(\mathrm{S}_{\boldsymbol{\alpha}}, \mathrm{i}) + F \cdot t_{i,j}$, where $(\mathrm{S}_{\boldsymbol{\alpha}}, \mathrm{i})$ is the state from previous iteration.

We additionally assume the following travel time matrix associated with the three locations:

Travel Time Matrix t_{uw}

u \ v	0	1	2	3
0	0	0.22	0.05	0.25
1	0.22	0	0.22	0.01
2	0.05	0.22	0	0.2
3	0.25	0.01	0.2	0

For this case involving a single vehicle, some simplifications of the general solution procedure outlined in the previous section can be made—it is not necessary to assign admissible paths to each vehicle since there is only one vehicle. Rather, efficiently finding the best admissible path that tours all of the nodes that need to be traversed in one path is the key. In this case, the algorithm suggested for the sub-problem of the shortest path problem with time windows can be used; however, a few adjustments can be made in order to render the solution method more efficient. These adjustments exclude paths that do not visit all activity nodes that are required to be completed since there is only one path for single vehicle households. First, the recursion step occurs for iterations $2 \leq k \leq 2(n_p + m)$, and the new node to be added is only from $\boldsymbol{P}$—thereby excluding labels to add the final depot nodes during this step—adding the final node at the final iteration, $k = 2(n_p + m) + 1$. These changes ensure that all required nodes are visited in this tour before the final return home. The algorithm for this case is as follows:

Initialization $(k = 1)$

A set of states of routes visiting one activity node from home location are generated.

$$\{(\{j\}, j),\ \ j \in \boldsymbol{P}^{+}\}$$

Corresponding arrival time, cumulative cost, and cumulative number of sojourns in a tour are updated as:

$$T(\mathrm{S}_{\boldsymbol{\alpha}}, j) = \{\ \max\left(a_j, a_0 + t_{0,j}\right) \le t(\mathrm{S}_{\boldsymbol{\alpha}}, j) \le b_j,$$
$$t(\mathrm{S}_{\boldsymbol{\alpha}}, 0) + t_{0,j} \le\ t(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j})\ \}$$
$$c(\mathrm{S}_{\boldsymbol{\alpha}}, j) = c_{0,j}$$
$$d(\mathrm{S}_{\boldsymbol{\alpha}}, j) = d_i$$

Recursion $(2 \le k \le 2(n_p + m))$

New states are constructed by adding one node, j, to the total visited at the preceding iteration:

$\{(S\cup\{j\}, j),\ \ j \in \boldsymbol{P}\}$ where (S, i) is the state from previous iteration

Then the states are tested for elimination criteria, and if the state $(S\cup\{j\}, j)$ is not eliminated, its label set will be created. Its corresponding arrival time, cumulative cost, and cumulative number of sojourns in a tour are updated as:

$$T(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, j) = T(\mathrm{S}_{\boldsymbol{\alpha}}, j) \cup \{\ \max\left(a_j, T_i + s_i + t_{i,j}\right) \le t(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j}) \le b_j,$$
$$t(\mathrm{S}_{\boldsymbol{\alpha}}, \mathrm{i}) + s_i + t_{i,j} \le\ t(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j})\}$$
$$c(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j}) = c(\mathrm{S}_{\boldsymbol{\alpha}}, i) + c_{i,j}$$
$$d(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j}) = d(\mathrm{S}_{\boldsymbol{\alpha}}, i) + d_j$$

Final Iteration $(k = 2(n_p + m) + 1)$

There is only one state to be generated. All activity nodes and corresponding return home nodes have been visited, and the terminal node is the final depot node:

$$\{(\{1,2, \dots, 2n, 2n + 1\}, 2n + 1)\}$$

Corresponding arrival time, cumulative cost, and cumulative number of sojourns in a tour are updated as previous.

Selection of Arrival Times

For all completed paths, $\boldsymbol{\alpha}$, from *Final Iteration* $(k = 2(n_p + m) + 1)$, solve the following optimization problem, and update the final cost.

Minimize $f(T_0, T_1, \dots, T_{2n}, T_{2n+1})$ such that $T(\mathrm{S}_{\boldsymbol{\alpha}}, 2n + 1)$

Algorithm 2-2 Single vehicle LSP-HAPP path generation algorithm for objective function involving time variables

For this, criteria #11- #13 are not necessarily useful since the final return home node is not added to the labels until all of the other nodes are added. In order to increase the efficiency of the algorithm, the following criteria can be included in the elimination test:

#15: given the arrival time at node j, T_j, it must be possible to visit each subsequent unvisited preselected node $l \in \overline{S \cup \{j\}}$ while respecting the time constraint:

$$T_j + s_j + t_{j,l} \leq b_l, \text{ for all } l \in \boldsymbol{P_P}, \text{ and } l \in \overline{S \cup \{j\}}$$

The full results of label generation for this example of LSP-HAPP is presented in Table A-1 in Appendix 2-B. A summary for label of index 46 is presented in Table 2-1.

Table 2-1 Label generation procedure of grocery shopping location selection: Single vehicle

Iteration	Index	Visited nodes, S	Terminal node, j	Current cost, $c(S_\alpha, j)$	Time window constraints, $T(S_\alpha, j)$[4]	Previous Path index
$\boldsymbol{k = 1}$	1	{1}	1	1.38	$6 \leq T_0 \leq 22$ $8 \leq T_1 \leq 9$ $T_0 + 0.22 \leq T_1$	0
$\boldsymbol{k = 2}$	8	{1 4}	4	2.75	$17.22 \leq T_4 \leq 21$ $T_1 + 9 + 0.22 \leq T_4$	1

[4] This column only shows arrival time windows that are newly added during the iteration. Constraints from previous paths carry on, but due to space limit, they are not shown in this table. The full set of constraints can be constructed by tracking down previous indexes.

$\boldsymbol{k = 3}$	14	{1 3 4}	3	4.31	$17.47 \leq T_3 \leq 2$ $T_4 + 0.25 \leq T_3$	8
$\boldsymbol{k = 4}$	34	{1 3 4 6}	6	5.88	$18.72 \leq T_6 \leq 21$ $T_3 + 1 + 0.25 \leq T_6$	14
$\boldsymbol{k = 5}$	46	{1 3 4 6 7}	7	5.88	$18.72 \leq T_7 \leq 22$ $T_6 \leq T_7$	34

For all 12 completed labels, time variables are determined according to delivering the optimal value of the objective function, $c(\mathrm{S}_{\alpha}, \mathrm{j}) + V \cdot (T_7 - T_0)$. For example, for label of index 46, which traveled as: 7 (Label index 46) ← 6 (Label index 34) ←3 (Label index 14) ← 4 (Label index 8) ← 1 (Label index 1) ← 0, the following problem is solved to determine arrival times.

$minimize\ Z = V \cdot (T_7 - T_0)$

subject to:

$6 \leq T_0 \leq 22$

$8 \leq T_1 \leq 9$

$T_0 + 0.22 \leq T_1$

$17.22 \leq T_4 \leq 21$

$T_1 + 9 + 0.22 \leq T_4$

$17.47 \leq T_3 \leq 2$

$T_4 + 0.25 \leq T_3$

$18.72 \leq T_6 \leq 21$

$T_3 + 1 + 0.25 \leq T_6$

$18.72 \leq T_7 \leq 22$

$T_6 \leq T_7$

Once the time variables for all 12 final labels are chosen to achieve the optimum, the cost is updated to represent the full objective function value. Then, the label with the lowest value is the optimal solution. In the current example, it is label 35. The optimal path is: home ($T_0 = 6.74$) → grocery store 2 ($T_3 = 6.99$) → work ($T_1 = 8.00$) → home ($T_4 = T_6 = T_7 = 17.22$), with total cost of \$160.2. The activity and routing of the optimal path is visualized in Figure 2-1.

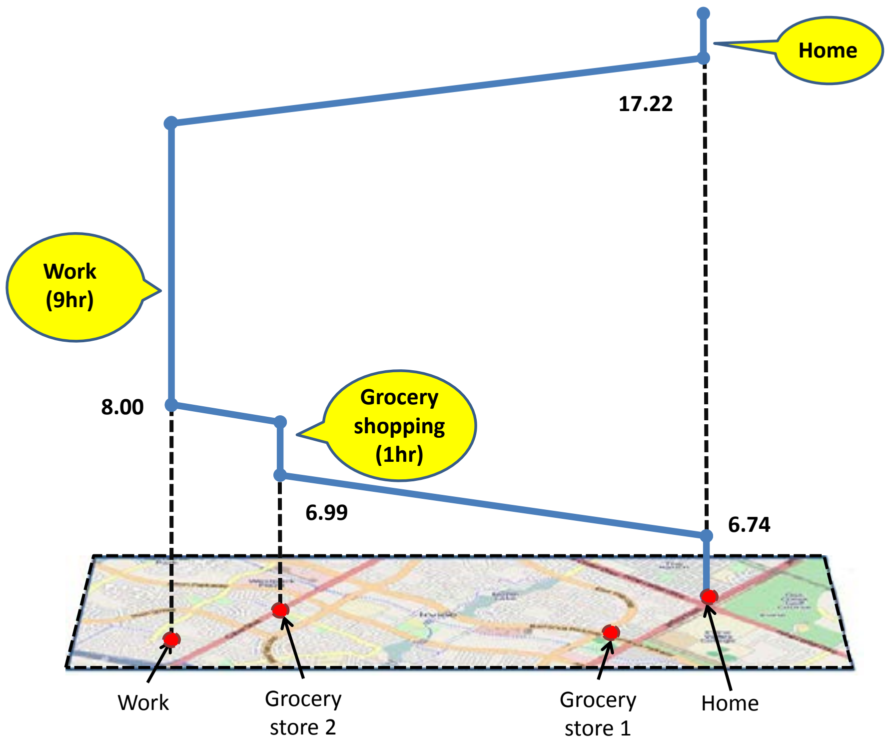


Figure 2-1 Optimal activity pattern of grocery shopping location selecting involving a single vehicle

2.3.1 Case 2: Grocery Shopping Location Selection for a Household with Two Vehicles

Similar to the previous example of grocery shopping location selection, assume a household with two vehicles and two household members, each with its vehicle exclusively available. The travel disutility is simply expanded to multiple vehicles as:

$$F \cdot \sum_{v \in \boldsymbol{V}} \sum_{u \in \boldsymbol{N}} \sum_{w \in \boldsymbol{N}} t_{uw}^{v} X_{uw}^{v} + V \cdot \sum_{v \in \boldsymbol{V}} [T_{2n+1}^{v} - T_{0}^{v}]$$

The household needs to complete two activities with pre-selected locations, $\mathbf{N}_P = \{1,2\}$; $n_P = 2$, which are work (node 1), with duration of $s_1 = 9.0$, and a drop-off activity (node 2), with duration of $s_2 = 0.1$. As in the previous example, the household also has a grocery shopping trip to be scheduled; i.e., $\boldsymbol{A} = \{\boldsymbol{A_1}\}$, $m = 1$, and that there are two potential locations for this activity $\boldsymbol{P_{A_1}} = \{3, 4\}$; $n_{A_1} = 2$; the operation hours for both stores is assumed to be from 6:00 to 22:00 and the duration of the shopping activity at either location is 1hour. In this example:

$\boldsymbol{M} = \boldsymbol{M_p} \cup \boldsymbol{M_A} = \{1, 2, 3, 4\}$; $n = n_P + n_A = 4$

$\boldsymbol{S} = \boldsymbol{S_p} \cup \boldsymbol{S_A} = \{s_1, s_2, s_3, s_4\} = \{9, 0.1, 1, 1\}$

$\boldsymbol{P_P^+} = \{1, 2\}$

$\boldsymbol{P_A^+} = \{3, 4\}$

$\boldsymbol{P_P^-} = \{5, 6\}$

$\boldsymbol{P_A^+} = \{7, 8\}$

$\boldsymbol{P^+} = \boldsymbol{P_P^+} \cup \boldsymbol{P_A^+} = \{1, 2, 3, 4\}$

$\boldsymbol{P^-} = \boldsymbol{P_P^-} \cup \boldsymbol{P_A^-} = \{5, 6, 7, 8\}$

$\boldsymbol{P_P} = \boldsymbol{P_P^+} \cup \boldsymbol{P_P^-} = \{1, 2, 5, 6\}$

$\boldsymbol{P_A} = \boldsymbol{P_A^+} \cup \boldsymbol{P_A^-} = \{3, 4, 7, 8\}$

with time availability windows, and corresponding return-home windows:

$$[a_i, b_i] = \begin{bmatrix} a_1, b_1 \\ a_2, b_2 \\ a_3, b_3 \\ a_4, b_4 \end{bmatrix} = \begin{bmatrix} 8, 9 \\ 12, 12.5 \\ 6, 21 \\ 6, 21 \end{bmatrix} \quad , [a_{n+i}, b_{n+i}] = \begin{bmatrix} a_5, b_5 \\ a_6, b_6 \\ a_7, b_7 \\ a_8, b_8 \end{bmatrix} = \begin{bmatrix} 6, \; 21 \\ 6, \; 21 \\ 6, \; 22 \\ 6, \; 22 \end{bmatrix} \quad ,$$

$$[a_0, b_0] = [6, \quad 20]$$

$$[a_{2n+1}, b_{2n+1}] = \left[a_{17}, b_{17}\right] = [6, \quad 21] \quad .$$

The travel time matrix is given as:

u \ v	0	1	2	3	4
0	0	0.22	0.12	0.05	0.25
1	0.22	0	0.13	0.22	0.01
2	0.12	0.13	0	0.11	0.1
3	0.05	0.22	0.11	0	0.2
4	0.25	0.01	0.1	0.2	0

The dynamic programming procedure with respect to time variables, Algorithm 2-1, generated 4 $(k = 1)$, 12 $(k = 2)$, 20 $(k = 3)$, 16 $(k = 4)$, 16 $(k = 5)$, label sets of feasible paths, and there are total of 20 completed paths (terminal node at final home depot). Each of these completed paths is a candidate route column. However, if there exists a label with same visited set that dominates in travel disutility (objective function), loads and arrival times, that label can be dropped. Of these, 14 paths (paths numbered 5, 6, 8, 9, 10, 12, 13, 15 – 19) are not used for the master problem of finding the optimal combination because there exists a different path(s) that traverses the same set of nodes (albeit with a different order) and end at the same node with either lower or same travel disutility, and with either earlier or same arrival time at the final node. The remaining paths (shown in Table 2-3) form the basis of the master problem.

Table 2-2 Admissible paths of grocery shopping location selection for the master problem

Path No., r	Visited nodes, S	Path Sequence and Arrival Times	Travel Disutility, c_r
0	{3 7 9}	home ($T_0 = 6$) → grocery store 1 ($T_3 = 6.05$) → home ($T_7 = T_9 = 7.1$)	17.13
1	{1 5 9}	home ($T_0 = 7.78$) → work ($T_1 = 8$) → home ($T_5 = T_9 = 17.22$)	144.35
2	{2 6 9}	home ($T_0 = 11.88$) → drop off ($T_2 = 12$) → home ($T_6 = T_9 = 12.22$)	6.60
3	{4 8 9}	home ($T_0 = 6$) → grocery store 2 ($T_4 = 6.25$) → home ($T_8 = T_9 = 7.5$)	25.63
4	{2 3 6 7 9}	home ($T_0 = 11.88$) → drop off ($T_2 = 12$) → grocery store 1 ($T_3 = 12.21$) → home ($T_6 = T_7 = T_9 = 13.26$)	22.45
7	{2 4 6 8 9}	home ($T_0 = 10.65$) → grocery store 2 ($T_4 = 10.9$) → drop off ($T_2 = 12$) → home ($T_6 = T_7 = T_9 = 12.22$)	26.49
11	{1 4 5 8 9}	home ($T_0 = 6.74$) → grocery store 2 ($T_4 = 6.99$) → work ($T_1 = 8$) → home ($T_5 = T_8 = T_9 = 17.22$)	160.20
14	{1 3 5 7 9}	home ($T_0 = 6.73$) → grocery store 1 ($T_3 = 6.78$) → work ($T_1 = 8$) → home ($T_5 = T_7 = T_9 = 17.22$)	160.41

Then, the matrix presentation of master problem (a) - (c) is

$$r\ [c_0 \quad c_1 \quad c_2 \quad c_3 \quad c_4 \quad c_7 \quad c_{11} \quad c_{14}]$$

$$\begin{array}{l} a_{1r} \\ a_{2r} \\ a_{3r} \\ a_{4r} \end{array} \begin{bmatrix} & 1 & & & & & 1 & 1 \\ & & 1 & & 1 & 1 & & 0 \\ 1 & & & & 1 & & & 1 \\ & & & 1 & & 1 & 1 & \end{bmatrix} \begin{bmatrix} Y_0 \\ Y_2 \\ \vdots \\ Y_{14} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \left.\begin{matrix} \\ \\ \end{matrix}\right\} + = 1 \end{bmatrix}$$

The master program, which is an integer programming problem, concludes ($Y_1 = Y_4 = 1$) that paths 1 and 4 bring the minimum cost of \$166.8 for this household. The grocery store 1 located at node 3 is selected over the grocery store at node 4. By tracking the previous indices, we find that person 1 travels path 1: home ($T_0 = 7.78$) → work ($T_1 = 8$) → home ($T_5 = T_9 = 17.22$), and person 2 travel as path 7: home ($T_0 = 11.88$) → drop off ($T_2 = 12$) → grocery store 1 ($T_3 = 12.21$) → home ($T_6 = T_7 = T_9 = 13.26$). These results are depicted in Figure 2-2.

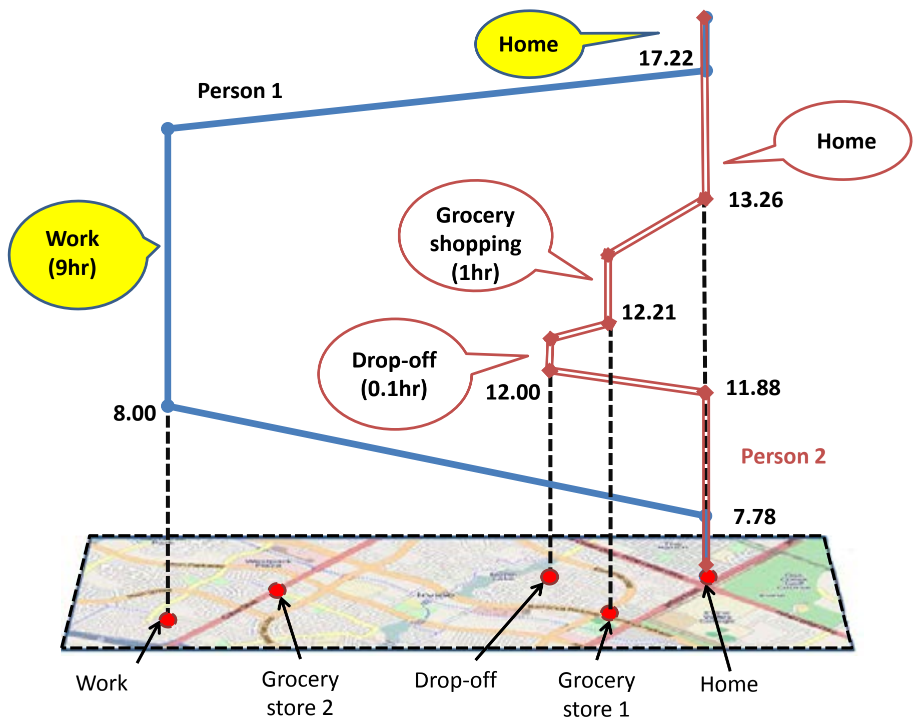


Figure 2-2 Optimal activity pattern of grocery shopping location selecting for a household with two vehicles

2.3.2 Case 3: Grocery Shopping Location Selection for a Household with Two Vehicles with Restricted Activity Participation

The above example places no restrictions on which members of the household perform the scheduled activities. For more realistic assignment of household activities, we can add restrictions:

$$\sum_{w \in \boldsymbol{\Omega}_v} \sum_{u \in \boldsymbol{P}} X_{u,w}^{v} = 0, v \in \boldsymbol{V}$$

where $\boldsymbol{\Omega}_1^{\boldsymbol{V}}$ is the subset of activities that **cannot** be performed by vehicle/person υ. Assume, for example, that person 1 is the person who needs to perform both the work as well as the grocery activities.[5]

$\boldsymbol{\Omega}_0^{\boldsymbol{V}} = \{ \quad \}$

$\boldsymbol{\Omega}_1^{\boldsymbol{V}} = \{1, 3, 4\}$

Here, we can eliminate terminated paths which include only one of work and grocery shopping activities. In the example, $Y_0 = Y_2 = 0$ and these paths do not enter the master problem as a candidate path column, or are constrained to be zero. The optimal assignment combination is decided among paths r =2, 3, 4, 7, 11, 14, and found to be $Y_2 = Y_{11} = 1$: person/vehicle 1 travels path 11, home ($T_0 = 6.74$) → grocery store 2 ($T_4 = 6.99$) → work ($T_1 = 8$) → home ($T_5 = T_8 = T_9 = 17.22$), and person/vehicle 2 travels path 2, home ($T_0 =$

[5] Note that the notation starts from index 0.

11.88) → drop off ($T_2 = 12$) → home ($T_6 = T_9 = 12.22$), with the total cost of \$166.8 (Figure 2-3).

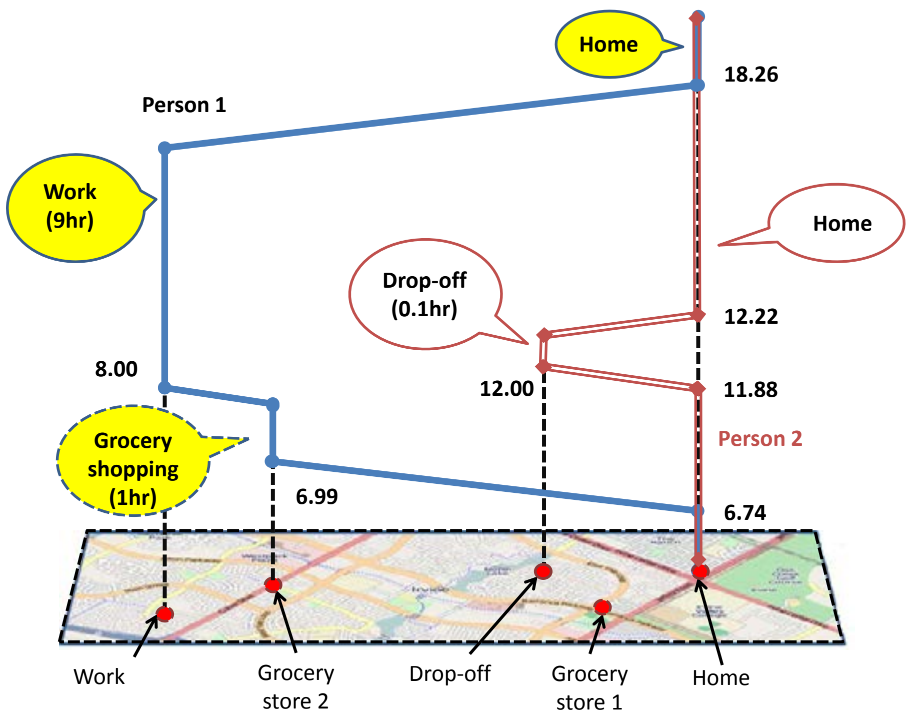


Figure 2-3 Optimal activity pattern of grocery shopping location selection for a household with two vehicles with activity assignment restrictions

The process of path removal that violates personal restrictions can be imbedded at the end of recursion from Algorithm 2-1, as shown in Algorithm 2-3.

Initialization ($k = 1$)

A set of states of routes visiting one activity node from home location are generated.

$$\{(\{j\}, j),\ j \in \boldsymbol{P}^{+}\}$$

Corresponding arrival time, cumulative cost, and cumulative number of sojourns in a tour are updated as:

$$T(\mathrm{S}_{\boldsymbol{\alpha}}, j) = \{\ \max\left(a_j, a_0 + t_{0,j}\right) \le t(\mathrm{S}_{\boldsymbol{\alpha}}, j) \le b_j,$$
$$t(\mathrm{S}_{\boldsymbol{\alpha}}, 0) + t_{0,j} \le\ t(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j})\ \}$$
$$c(\mathrm{S}_{\boldsymbol{\alpha}}, j) = c_{0,j}$$
$$d(\mathrm{S}_{\boldsymbol{\alpha}}, j) = d_i$$

Recursion ($2 \le k$)

New states are constructed by adding one node, j, to the total visited at the preceding iteration:

$\{(S\cup\{j\}, j),\ j \in \boldsymbol{P} \cup \{2n+1\}\}$ where (S, i) is the state from previous iteration

Then the states are tested for elimination criteria, and if the state $(S\cup\{j\}, j)$ is not eliminated, its label set will be created. Its corresponding arrival time, cumulative cost, and cumulative number of sojourns in a tour are updated as:

$$T(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, j) = T(\mathrm{S}_{\boldsymbol{\alpha}}, j) \cup \{\ \max\left(a_j, T_i + s_i + t_{i,j}\right) \le t(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j}) \le b_j,$$
$$t(\mathrm{S}_{\boldsymbol{\alpha}}, \mathrm{i}) + s_i + t_{i,j} \le\ t(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j})\}$$
$$c(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j}) = c(\mathrm{S}_{\boldsymbol{\alpha}}, i) + c_{i,j}$$
$$d(\mathrm{S}\cup\{j\}_{\boldsymbol{\alpha}}, \mathrm{j}) = d(\mathrm{S}_{\boldsymbol{\alpha}}, i) + d_j$$

Stop when there is no label generated at this iteration.

Removal of Paths based on Restrictive Activity Participation

For all generated paths, if any activity node $j \in \boldsymbol{P}^{+}$ in its visited node set, $j \in \mathrm{S}$, is an activity that can only be performed by one specific vehicle/household member v, $v \in \boldsymbol{V}$, then any of the other visited nodes cannot be the activity that is restricted for v:

if $j \in \bigcap_{v \neq r, r \in \boldsymbol{V},} \boldsymbol{\Omega}_{\boldsymbol{r}}^{\boldsymbol{V}}$ for any $j \in \mathrm{S}$ then, $l \in \boldsymbol{\Omega}_{\boldsymbol{v}}^{\boldsymbol{V}}$, for all $l \in \mathbf{S}$, for $v \in \boldsymbol{V}$

And all activities (all pre-selected activities and one of the selective locations) that need to be performed by v, needs to be in the visited set.

if $j \in \bigcap_{v \neq r, r \in \boldsymbol{V},} \boldsymbol{\Omega}_{\boldsymbol{r}}^{\boldsymbol{V}}$ for any $j \in \mathrm{S}$ then, $l \in \mathrm{S}$ for all $l \in \bigcap_{v \neq r, r \in \boldsymbol{V},} \boldsymbol{\Omega}_{\boldsymbol{r}}^{\boldsymbol{V}}, l \neq j, l \in P_P^{+}$, or one of $l \in P_{A_a}^{+}$ for $l \in \bigcap_{v \neq r, r \in \boldsymbol{V},} \boldsymbol{\Omega}_{\boldsymbol{r}}^{\boldsymbol{V}}, l \neq j, l \in P_P^{+}$ for all $v \in \boldsymbol{V}$

Selection of Arrival Times

For all completed paths, $\boldsymbol{\alpha}$, solve the following optimization problem, and update the final cost.

Minimize $f(T_0, T_1, \ldots, T_{2n}, T_{2n+1})$ such that $T(\mathrm{S}_{\boldsymbol{\alpha}}, 2n+1)$

Algorithm 2-3 LSP-HAPP path generation algorithm with restrictive activity participation for objective function involving time variables

The first part of the condition can be imposed as an additional elimination rule, #16, during the recursion process to increase the efficiency, but the second condition needs to be performed for completed paths.

#16: For all generated paths, if any activity node $j \in \boldsymbol{P}^{+}$ in its visited node set, $j \in$ S, for is an activity that can only be performed by one specific vehicle/household member v, $v \in \boldsymbol{V}$, then any of the other visited nodes cannot be the activity that is restricted for v:

$$\text{if } j \in \bigcap_{v \neq r, r \in \boldsymbol{V},} \boldsymbol{\Omega}_{r}^{\boldsymbol{V}} \text{ then, } l \in \boldsymbol{\Omega}_{v}^{\boldsymbol{V}}, \text{ for all } l \in \mathbf{S}, \text{ for } v \in \boldsymbol{V}$$

For HAPP Case 4 and HAPP Case 5, the same changes as in Equations (2)-(3) can be made; however, the solution process overcoming the computational difficulties is not developed in this dissertation. Because these cases require generation of person-based and vehicle-based patterns and matching of these two, it is highly related to mode choice problem which has not yet been fully integrated in HAPP.

2.4 CASE STUDY

LSP-HAPP is applied to 13 households of single vehicle and single member households residing in Orange County, California, that have conducted one incidental shopping activity (includes shopping activities for grocery, medicine or house ware, but excludes such major shopping activities as furniture or automobile shopping) during the survey day. The data are drawn from the California Travel Survey (2001). For this example, individual household's travel disutility is specified by the linear combination of the total extent of the day, the travel times, and the delay of return home caused by trip chaining for each of out-of-home activities by the individual weights of such measurements, $\beta^E, \beta^T, \beta^D$:

$$\min Z = \beta^E \sum_{v \in V} (T_{2n+1}^v - T_0^v) + \beta^D \sum_{w \in P^+} (T_{w+n} - T_w) + \beta^T \sum_{v \in V} \sum_{u \in N} \sum_{w \in N} t_{uw}$$

The weights of these households are empirically estimated using the inverse optimization calibration process in Chow and Recker (2012). Time windows of activities are separately generated using the methodology described in Chapter 4, which adopted the method from Recker and Parimi (1999) with relaxation of return home activity's time windows.

Candidate shopping locations are derived from the reported shopping locations in the study area, which numbered a total of 19. For practical implementation of the model, there would need to be a zoning procedure for aggregating candidate locations within a geographical area, but with the limited number of survey data used in this example, exact locations are spatially sparse enough to be individually located for the purpose of testing LSP-HAPP. These locations along with household home locations and their other activity locations are shown in Figure 2-4.

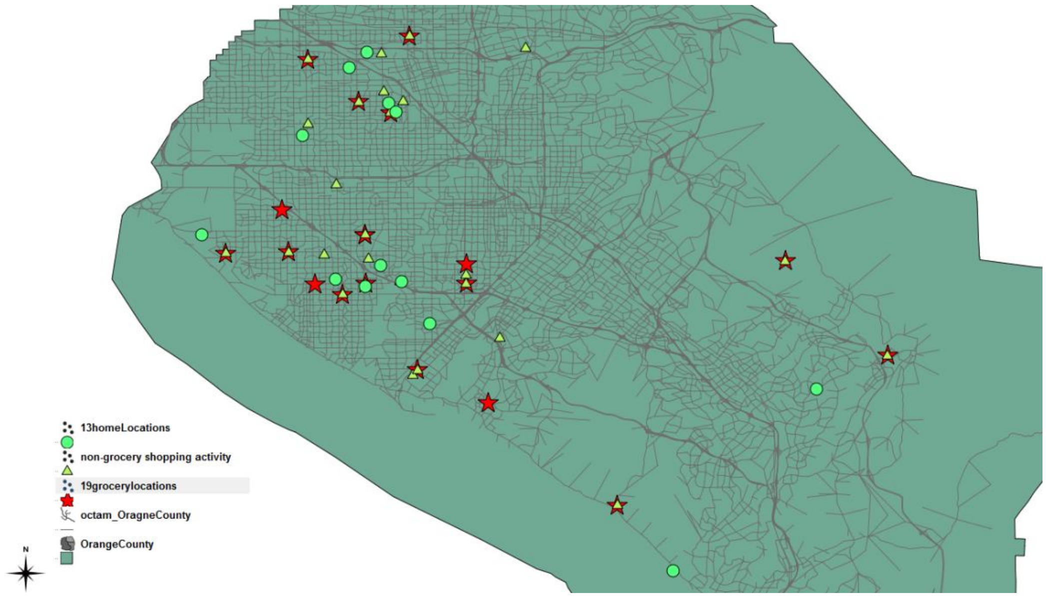


Figure 2-4 Case study area

Of the test sample of 13 households, application of the LSP-HAPP model resulted in the destination choice of 8 households being the exact same location as the reported shopping location. For the remaining five households, the distance/travel time differences between the outcome of the model and the reported locations are 2.4 miles (0.15 hours), 1.5 miles (0.12 hours), 2.5 miles (0.13 hours), 4.2 miles (0.22 hours), and 1.65 miles (0.09 hours). The average absolute difference between the model output and real data of start times of these shopping activities is 1.67 hours, with a maximum deviation of 4.16 hours, and a minimum of 0. It is noted that the activity start times determined by the model are highly dependent on how accurately the estimates of time windows are generated. In this application, the method we have adopted as explained in Chapter 4, which is based on Recker and Parimi (1999) provides fairly accurate arrival time selection but in a number of cases leads to infeasible cases for the reported pattern due to discrepancies in reported travel times and the actual shortest-path based travel time matrix, especially when it includes a tour that traverses many activities. While refining and improving this time window generation is an important issue for the practicality of the HAPP models in general, it is not the scope of this dissertation.

The performance of the suggested algorithm is also found to be competitive. Solving LSP-HAPP directly by calling the CPLEX library took on average of 2,910 seconds, maximum case at 12,730 seconds, and minimum case at 180 seconds. Alternatively, Algorithm 1 took on average of 614 seconds (maximum at 3625 seconds, minimum at 25 seconds) which includes the generation of 577 (maximum of 2778, minimum of 28) labels, and average 148 runs (maximum of 718 runs, minimum of 2 runs) of "easy linear programming" of selecting the activity start (arrival) times via CPLEX library.

2.5 APPLICATION IN TRAVEL PATTERN GENERATION

For activity-based transportation planning, synthetic pattern generation and assignment of those patterns over space are fundamental steps for travel forecasting. HAPP has been shown to be a useful tool for synthesizing daily activity patterns on a household basis. With the capability of choosing locations, LSP-HAPP can work as a pattern synthesizer as well as a tool for linking spatial information with such patterns, given activities and their durations for a household. In these two aspects, such application is similar to the approach proposed by McNally (1997), although the specifications of models are different. McNally (1997) selected a representative pattern that includes a set of activities and durations, given household characteristics, and matched the pattern with spatial information, whereas the LSP-HAPP model creates a pattern simultaneously linking to spatial information, given a modeler's desired goal and a set of activities to be performed along with their durations, possibly generated from household characteristics.

As an illustration, assume that the modeler's goal is to select activity locations and generate travel patterns for a one-vehicle household that, either from direct survey data or from regional models, is assigned two activities—work ($\boldsymbol{A_1}$) and grocery shopping ($\boldsymbol{A_2}$)—and a travel of $\bar{t}$ minutes for the day. Then the objective function within the planning model context is to minimize the error between desired and generated travel times, i.e.,

$$minimize \left| \bar{t} - \sum_{v \in \boldsymbol{V}} \sum_{u \in \boldsymbol{N}} \sum_{w \in \boldsymbol{N}} t_{uw}^{v} X_{uw}^{v} \right|$$

During recursion ($1 \leq k \leq 2(n_p + m)$), we can store cost as the cumulative travel times updated as:

$$c(\mathrm{S} \cup \{j\}_{\alpha}, \mathrm{j}) = \begin{cases} t_{i,j} & i = 0 \\ c(\mathrm{S}_{\alpha}, \mathrm{i}) + t_{i,j} & i \neq 0 \end{cases}$$

and in the final iteration ($k = 2(n_p + m) + 1$), we can select the optimal path as path $\boldsymbol{\alpha}$ with the smallest difference between the desired and observed total travel time, $|t - c(\mathrm{S} \cup \{2n + 1\}_{\alpha}, 2n + 1)|$[6].

[6] Because the objective function is not related to arrival time variables, Algorithm 2 without the final step of selecting arrival times, is used to solve this problem. Arrival times are selected as possible earliest time during the initialization and recursion as in Desrosiers et al. (1986)

Because the goal is matching locations with the modeling objective while generating the travel patterns, there is no activity location that has been pre-defined; i.e., $N_P = \{\}$; $n_P = 0$. Suppose activity durations of work and grocery shopping are $s_{\boldsymbol{A_1}} = 9$, $s_{\boldsymbol{A_2}} = 1$, and time availability windows for each activity type are:

$$\begin{bmatrix} a_{\boldsymbol{A_1}}, b_{\boldsymbol{A_1}} \\ a_{\boldsymbol{A_2}}, b_{\boldsymbol{A_2}} \end{bmatrix} = \begin{bmatrix} 8, 9 \\ 6, 22 \end{bmatrix}$$

with corresponding return-home windows:

$$\begin{bmatrix} a_{\boldsymbol{A_1}+n}, b_{\boldsymbol{A_1}+n} \\ a_{\boldsymbol{A_2}+n}, b_{\boldsymbol{A_2}+n} \end{bmatrix} = \begin{bmatrix} 6, 21 \\ 6, 22 \end{bmatrix}$$

and with initial departure and end-of-travel day windows:

$$\begin{aligned} [a_0, b_0] &= [6, \quad 20] \\ [a_{2n+1}, b_{2n+1}] &= [6, \quad 21] \quad . \end{aligned}$$

Assume also that there are two central business district locations for work ($\boldsymbol{A_1}$), and also two possible locations for grocery shopping ($\boldsymbol{A_2}$) in the area. $\boldsymbol{P_{A_1}} = \{1, 2\}$; $n_{\boldsymbol{A_1}} = 2$, $\boldsymbol{P_{A_2}} = \{3, 4\}$; $n_{\boldsymbol{A_2}} = 2$, $n_{\boldsymbol{A}} = 4$ and $n = n_{\boldsymbol{P}} + n_{\boldsymbol{A}} = 4$.

In this example:

$\boldsymbol{N} = \boldsymbol{N_p} \cup \boldsymbol{N_A} = \boldsymbol{N_p} \cup \boldsymbol{N_{A_1}} \cup \boldsymbol{N_{A_2}} = \{1, 2, 3, 4\}$; $n = n_P + n_A = 4$

$\boldsymbol{S} = \boldsymbol{S_p} \cup \boldsymbol{S_A} = \{s_1, s_2, s_3, s_4\} = \{9, 9, 1, 1\}$

$\boldsymbol{P_P^+} = \{\}$, $\boldsymbol{P_P^-} = \{\}$

$\boldsymbol{P}_{\boldsymbol{A}_1}^{+} = \{1, 2\}$, $\boldsymbol{P}_{\boldsymbol{A}_1}^{-} = \{5, 6\}$

$\boldsymbol{P}_{\boldsymbol{A}_1} = \boldsymbol{P}_{\boldsymbol{A}_1}^{+} \cup \boldsymbol{P}_{\boldsymbol{A}_1}^{-} = \{1, 2, 5, 6\}$

$\boldsymbol{P}_{\boldsymbol{A}_2}^{+} = \{3, 4\}$, $\boldsymbol{P}_{\boldsymbol{A}_2}^{-} = \{7, 8\}$

$\boldsymbol{P}_{\boldsymbol{A}_2} = \boldsymbol{P}_{\boldsymbol{A}_2}^{+} \cup \boldsymbol{P}_{\boldsymbol{A}_2}^{-} = \{3, 4, 7, 8\}$

$\boldsymbol{P}_{\boldsymbol{A}}^{+} = \{1, 2, 3, 4\}$

$\boldsymbol{P}_{\boldsymbol{A}}^{-} = \{5, 6, 7, 8\}$

$\boldsymbol{P}^{+} = \boldsymbol{P}_{\boldsymbol{P}}^{+} \cup \boldsymbol{P}_{\boldsymbol{A}}^{+} = \{1, 2, 3, 4\}$

$\boldsymbol{P}^{-} = \boldsymbol{P}_{\boldsymbol{P}}^{-} \cup \boldsymbol{P}_{\boldsymbol{A}}^{-} = \{5, 6, 7, 8\}$

$\boldsymbol{P}_{\boldsymbol{P}} = \boldsymbol{P}_{\boldsymbol{P}}^{+} \cup \boldsymbol{P}_{\boldsymbol{P}}^{-} = \{\}$

$\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{P}_{\boldsymbol{A}}^{+} \cup \boldsymbol{P}_{\boldsymbol{A}}^{-} = \{1, 2, 3, 5, 6\ 7, 8\}$

with time availability windows, and corresponding return-home windows:

$$\begin{bmatrix} a_1, b_1 \\ a_2, b_2 \\ a_3, b_3 \\ a_4, b_4 \end{bmatrix} = \begin{bmatrix} 8, 9 \\ 8, 9 \\ 6, 22 \\ 6, 22 \end{bmatrix}, \qquad \begin{bmatrix} a_5, b_5 \\ a_6, b_6 \\ a_7, b_7 \\ a_8, b_8 \end{bmatrix} = \begin{bmatrix} 6, 21 \\ 6, 21 \\ 6, 22 \\ 6, 22 \end{bmatrix}$$

We additionally assume that the total travel time desired to be matched is $t = 0.5$, and the following travel time matrix associated with the four locations is as:

Travel Time Matrix t_{uw}

v / u	0	1	2	3	4
0	0	0.22	0.17	0.05	0.17
1	0.22	0	0.18	0.22	0.13
2	0.17	0.18	0	0.12	0.17
3	0.05	0.22	0.12	0	0.10
4	0.17	0.13	0.17	0.10	0

For this scenario, the algorithm generated as the optimal solution path home → grocery shopping at location 3 (6.05) → work at location 1 (8) → home (17.22) as depicted in the Figure 4, and the total travel time for this pattern is 0.49 hours, yielding an error between desired and generated travel times of 0.01 (Figure 2-5).

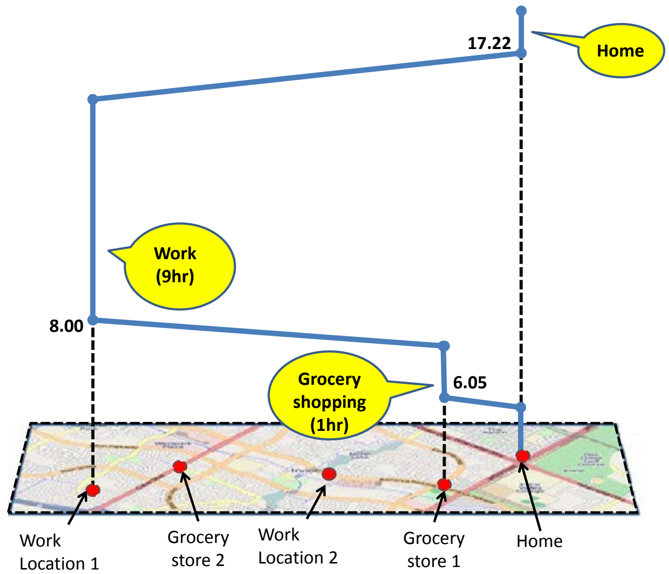


Figure 2-5 Synthetic travel pattern generation results

2.6 COMMENTS ON THE GENERAL COLUMN GENERATION PROCEDURE

Not only is finding the admissible path set, $\boldsymbol{\Psi}$, a combinatorial problem, but finding path combinations for each vehicle/household member is also an exponential combinatorial problem. Compared to the general pick-up and delivery problem with time windows, the total number of household members and the total number of vehicles are rather limited for the case of HAPP. Yet, it is still helpful to examine how the iterative procedure of column generation can be applied to LSP-HAPP. There exist other algorithms and methodologies, but the structural property that each routing path forms a column, has resulted in column generation as a technique widely used in vehicle routing problems (Desrosiers et al., 1984) as well as PDPTW.

In the previous example, all possible paths are introduced to the master problem; however if there are a large number of paths created, computational issues can become critical even for the master problem. Dumas et al. (1991) developed and tested iterative column generation procedures for multiple vehicle PDPTW. The same master and the sub-problem framework can be applied to LSP-HAPP with small adjustments.

The sub-problem finds one path column with the most negative reduced cost to add to the master problem, and then the master problem is solved to find the best combinations of paths. The sub-problem that finds this one column path with the smallest marginal cost can be written as:

$$minimize \sum_{i \in \mathbf{N}} \sum_{j \in \mathbf{N}} \bar{c}_{ij} X_{ij} = \sum_{i \in N} \sum_{j \in N} (c_{i,j} - \sigma_i) \cdot X_{i,j} \qquad (2-d)$$

subject to: (2-4) - (2-22)

where,

$\bar{c}_{ij}$: the marginal cost of trip from node i to node j

π_i: the dual variables associated with (b-1), $i \in \boldsymbol{P_P^+}$

π_{A_a}: the dual variables associated with (b-2), $a \in \boldsymbol{A}$

Then, we can associate dual variables, σ_i, with each pre-selected activity node, $\sigma_i = \pi_i, i \in \boldsymbol{P^+}$. Similarly, dual variable of candidate locations of activity type $\boldsymbol{A_a}$, can be associated as, $\sigma_{\boldsymbol{i}} = \pi_{\boldsymbol{A_a}}, a \in \boldsymbol{P_{A_a}^+}$. Lastly, dual variable values associated with departure home node, final home node, and return home nodes are all zeroes. $\sigma_1 = 0, \sigma_{2n+1} = 0$, and $\sigma_i = 0, i \in \boldsymbol{P^-}$. To find dual values from the master problem, the master problem is relaxed to be non-integer. Set partitioning problems, which the arc-path formulation of maximum multi-commodity problem forms, often achieve optimum at binary values even when relaxed.

For PDPTW (and therefore also for HAPP), there exists an efficient dynamic programming procedure that generates shortest paths with time windows, which means that this sub-problem does not have to be solved as a network formulation of a linear programming problem. Also, for LSP-HAPP, the dynamic programming algorithms developed can be the solution method for the sub-problem. At each iteration, the path cost of new reduced cost is simply updated to all paths generated from the dynamic programming algorithm. Then, the master program is rerun with a new path column with the most negative reduced cost until there is no path that can deliver better objective function value. The iterative procedure is shown as the following diagram (Figure 2-6).

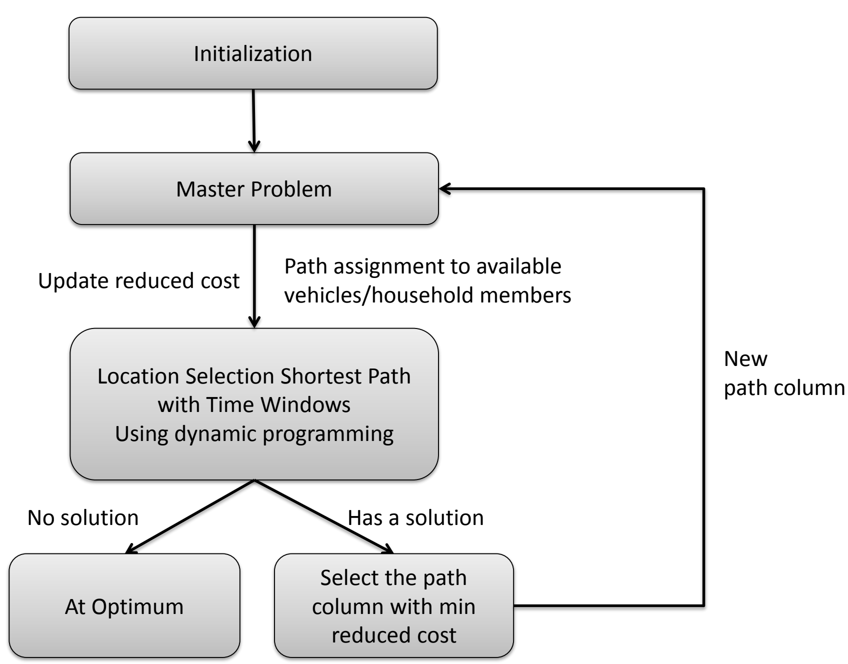


Figure 2-6 Iterative procedure of column generation of LSP-HAPP

2.7 CONCLUSION AND DISCUSSION

In this chapter, the Location Selection Problem extending the Household Activity Pattern Problem (LSP-HAPP) is presented. This is accomplished by relaxing the constraints that specify the condition that all nodes need to be visited. In the LSP-HAPP formulation, only one of possible locations for each activity with no pre-selected location is traversed. This formulation demonstrates how location choice for certain activities is made within the tours and scheduling of pre-selected activities and other activities with many candidate locations.

A dynamic programming algorithm, developed for PDPTW, is adapted for LSP-HAPP in order to deal with choice from among a sizable number of candidate locations within the HAPP modeling structure. The algorithm generates labels of all possible paths and selects the best path in the final step. The efficiency of the algorithm is determined by path elimination criteria that rule out illogical paths, and is shown to be efficient both in the literature on PDPs as well as in this application. Additionally, by the properties of label generation that updates time and sojourn variables and the objective function values, we are able to accommodate some level of nonlinearities in time, sojourn and cost. Lastly, an improvement is made to the algorithm in that arrival times are kept as functions, not parameters. This is because HAPP cases often have travel disutility measures involving time variables but the previous algorithms assume that travel disutility (costs) and arrival times are independent. From the case study, we can conclude that the formulation provides reasonable results in location selection as well as activity start times, and the solution method is superior in terms of computation time.

In developing the model, it is assumed that destination choice associated with non-primary activities is an auxiliary choice made within the scheduling of other, primary activities, and other activities that can be completed by visiting one of many candidate sites. It is arguable that LSP-HAPP ignores socio-economic influences, personal preference or habitual travel behaviors, but if such are measurable and quantifiable in the objective function, they can be easily reflected in the model. Estimation results from choice models (Bhat et al., 1998; Fotheringham, 1988; Pozsgay and Bhat, 2001; Recker and Kostyniuk, 1978) might be helpful in determining those influences. Once candidate factors are selected and measured, we can estimate the HAPP (Chow and Recker, 2011; Recker et al., 2008), determine their effects, and use them for LSP-HAPP models. However, in order to fit real data for destination choices within the structure of LSP-HAPP, new estimation schemes need to be developed and evaluated.

Finally, an application of LSP-HAPP that generates synthetic patterns and links with spatial information in a single model for activity-based forecasting models is presented. In transportation forecasting, microscopic travel patterns need to be aggregated and at an aggregated level, destination choice can be viewed as a category in spatial interaction models (Roy and Thill, 2004). For this example to be integrated into regional transportation forecasting models, further investigation on how to aggregate it to meet certain data, such as traffic counts or OD tables, is needed.

CHAPTER 3 MEASURING THE INCONVENIENCE OF OPERATING AN ALTERNATIVE FUEL VEHICLE

The objective of this research is to measure household-level inconvenience of operating an Alternative Fuel Vehicle (AFV). It is expected that households with one or more AFVs would experience some level of inconvenience, or increased travel disutility, caused by limited refueling opportunities— e.g., in the case of Hydrogen Fuel Cell Vehicles (HFCVs)—as well as by shorter ranges and longer charging times—e.g., in the case of Battery Electric Vehicles (BEVs). The results are derived from simulations replicating/changing actual vehicle usage patterns reported in the California Statewide Household Travel Survey. A key assumption is that people do not change their participation in activities as accomplished with their current conventional Internal Combustion Engine Vehicles (ICEVs), but may change travel decisions of how to perform those activities. More specifically, different sets of scenarios involving completion of respondents' stated activities based on behavioral assumptions associated with refueling and recharging are tested. Results indicate that with limited fuel infrastructure, operating a HFCV costs about \$19 - \$39 worth of inconvenience for the day refueling is needed. For travelers with more than 60 miles, operating a BEV additionally imposes average of \$47 - \$50 with level 1 charging infrastructure, and \$6 - \$10 with level 2 charging infrastructure.

3.1 BACKGROUND AND MOTIVATION

Since the adoption in 1990 of the California Air Resources Board's legislation of Low Emission Vehicle (LEV) and Zero Emission Vehicle (ZEV) mandates, there have been positive expectations of Alternative Fuel Vehicle (AFV) adoption. Recently, owing to concerns about rising gasoline cost, increased technical feasibility of the 'green' AFVs, the success of the Hybrid Electric Vehicles (HEVs) in the automobile market, and the government's efforts in reduction of Greenhouse Gas (GHG) emissions (such as California's SB 32), achieving sustainable transportation system has never seemed more promising. Many recent assessments of energy use and emissions of AFVs have suggested positive outcomes. Major advantages of AFVs are their significant reductions in energy use and harmful emissions, considered to be two of the more significant automobile externalities (Parry et al., 2007). Obviously, the degree of these positive outcomes is dependent on the extent of AFV adoption.

However, even with these positive expectations of further adoption of AFVs, the survival of AFVs in the automobile market is not guaranteed. It has been postulated that there is a "sustainable" AFV market penetration threshold below which AFVs will not survive in the market. This is the so-called "chicken-and-egg" problem that explains the vehicle demand (purchase)–fuel infrastructure interaction. Such interactions are tested in Stephan and Sullivan (2004a, 2004b) and Schwoon (2007), and the condition of fuel infrastructure provision that would determine the survival of AFVs explored. The concept of "infrastructure" can be expanded more broadly to include such "social" factors as: word of mouth, social exposure, marketing, scale and scope economics, learning from experience, R&D, and innovation spillover (Struben and Sterman, 2008). In addition to the demand-infrastructure relationship, there are various factors that would determine the survival and the extent of AFV adoption. There are various stakeholders such as government, auto manufacturers, collaborators, competition, and activist groups, as well as impediments including regulatory barriers, resources, infrastructure and vehicle characteristics, as stated by Byrne and Polonsky (2001). With respect to all of the factors mentioned above, how effectively and economically we can overcome this “chicken-and-egg” status and get to the threshold point and maximize social benefits is the key.

From the consumer (driver) demand side, one of the more widely studied topics in the area of AFV demand is in the formulation of vehicle purchase choice models. Many studies have used stated preference survey data to identify factors influencing purchase of different vehicle fuel types via discrete choice models (Calfee, 1985; Bunch et al., 1993; Brownstone et al., 1996; Ewing and Sarigollu, 2000; Dagsvil et al., 2002). Factors found to influence the decision of AFV purchase include: vehicle price, operating cost, range, fuel availability, environmental benefit, and etc. There are well-known data unreliability issues associated with stated preference surveys, but with recent Hybrid Electric Vehicle (HEV) phenomenon, it can be expected that revealed preference data sets will become available for such models in the near future.

In this chapter, we take a completely different approach to analyzing potential barriers of AFV adoption and operation. Instead of identifying the factors that influence vehicle purchase decisions, we examine the additional inconvenience caused by operating AFVs as opposed to ICEVs. This inconvenience (or negligibility or even non-existence of such inconvenience) may not be directly used for predicting future vehicle adoption rates since the vehicle purchase decision is influenced by perception of possible inconvenience rather than by real inconvenience. However, it is possible to evaluate future policy directions or the extent of subsidies for AFV operations by understanding the inconvenience of AFV operations. In this paper, we examine the full-day travel patterns associated with current ICE vehicle use and simulate the same activity participation using an AFV. We test two behavioral scenarios in this simulation: 1) keeping the currently reported travel pattern intact (but subject to identified limitations of the particular AFV being simulated), and 2) possibly rearranging the reported tours (again subject to the AFV limitations) in order to minimize the total travel disutility. Using these two behavioral assumptions as the lower and the upper bounds, it is possible to get some quantified understanding of the magnitudes of inconvenience of owning and operating AFVs.

The importance of using a full-day activity and travel data for AFV studies has been stressed in previous works, mainly regarding the uses of BEVs or PHEVs (Jing Dong, 2012; Axsen et al., 2011; Zhang et al. 2011; Kang and Recker, 2009; Gondor et al., 2007)—the principal reason being that they give temporal estimates of power limitations and electricity demand increases. Such an activity-based approach provides greater detail, including distribution along the time dimension, and therefore more accurate spatial and temporal energy profiles are derived. In this paper, we expand the use of full-day activity and travel data to also include connectivity of activities and trips; using this information, we demonstrate that it is possible to analyze potential changes in travel decisions imposed by different characteristics of vehicles.

3.2 METHODOLOGY

Our analysis is based on a very simple assumption: the revealed activity/travel pattern (i.e., that reported in the travel survey) executed with conventional ICEVs represents the "optimal" pattern for the household, subject only to constraints imposed by activity location, travel time, and activity availability, and is in no way constrained by such vehicle characteristics as either range or availability of refueling/recharging opportunities. In the simulations, we replace one of the ICEVs owned by the household with an AFV, with the stipulation that individuals do not change their participation in their reported daily activities with an ICEV—in cases involving BEVs, certain range limitations and recharging rates are examined; in cases involving HFCVs, we insert a refueling trip, subject to certain densities of refueling stations[7].

Two sets of scenarios of AFV drivers' behaviors are analyzed. The first scenario considered is based on the behavioral assumption that each individual's (vehicle's) daily activities and travel scheduling/sequencing maintains the reported sequence. In this scenario, for a given fixed sequence of activities and trips reported from the California Statewide Household Travel Survey, we perform the following simulations: 1) for HFCVs, one refueling trip is inserted, and 2) for BEVs, charging between trips (both while performing an activity and also waiting for the vehicle to have sufficient charge to extend the range to make it to the next destination) is performed. Then, for each simulation, the best travel pattern is selected for that vehicle. For HFCVs, the goal is to simply assess the inconvenience of one added refueling trip with limited refueling infrastructure since the range (up to 400 miles), and refueling times (10 minutes) are similar to those of Internal Combustion Engine Vehicles (ICEVs). However, for Battery Electric Vehicles (BEVs), the battery status is monitored, and multiple charging events are allowed in order to maintain the driving range consistent with the daily driving distances. Results of this scenario provide an upper bound of inconvenience of operating an AFV, representing the worst case.

Under a second behavioral assumption, each driver's travel decisions are presumed to be made according to the behavior that would be optimal relative to minimizing their travel inconvenience, subject to completing their reported activities. The assumption is that drivers adapt to AFVs' physical characteristics and refueling/recharging, and change their travel

[7] We understand that a more justified sample here would be to restrict the analysis to the substitution of a hydrogen refueling activity only for those households that exhibited an ICEV refueling in their pattern. Our reasoning for using the entire sample is twofold: one practical—such restriction greatly reduces the sample size on which to perform the analysis; the other, more contextual—because of the ubiquitous nature of gasoline stations, refueling an ICEV likely causes only nominal (or no) adjustment to planned activities.

decisions to lower the inconvenience level since the refueling/recharging of AFVs generally increase the travel inconvenience. We assume that each individual tries to minimize the total inconvenience; this second scenario provides a lower bound of inconvenience—representing the best case.

These optimal behavior results are obtained running the Household Activity Pattern Problem (HAPP) which was first proposed in Recker (1995). Among a number of applications, HAPP has been applied to assess the impacts of "environmentally-optimal" behavior of carpooling and trip chaining (Parimi and Recker, 1999). The HAPP problem is a set of full-day activity-based travel demand models with a formulation that captures the spatial-temporal constraints first proposed by Hagerstrand (1970). Since HAPP formalizes this fundamental travel decision-making scheme under spatial and temporal physical constraints, it is possible to synthesize "feasible" travel patterns not previously observed, or even possible, under current conditions that exist. This is a particular advantage for this study, or for any study focusing on a new type of vehicle with different characteristics (e.g., BEVs, HFCVs) for which neither travel nor vehicle usage data are not generally available. Although it is presumed that some changes in travel behavior would occur, there are no data supporting this assumption. We note, however, that there has been some effort to examine refueling behaviors associated with limited refueling opportunities (Kelly and Kuly, 2012; Kitamura and Sperling, 1987). The basic structure of the HAPP model is given as:

$$
\begin{aligned}
&\min \mathbf{Z}^h = \text{Travel Disutility of Household } h = f(\mathbf{X}^h, \mathbf{T}^h) \\
&\text{s.t.} \\
&\mathbf{A}^h \begin{bmatrix} \mathbf{X}^h \\ \mathbf{T}^h \\ \mathbf{Y}^h \end{bmatrix} \leq= \mathbf{b}^h \ ; \\
&\mathbf{X}^h = \left[X_{u,w}^{v,h} \ , u, w \in \mathbf{N}_h, v \in \mathbf{V}_h \right], \ \mathbf{T}^h = \left[T_u^h \ , u \in \mathbf{P}_h^A \cup \mathbf{P}_h^R \right], \mathbf{Y}^h = \left[Y_u^h \ , u \in \mathbf{P}_h^A \cup \mathbf{P}_h^R \right]
\end{aligned}
\tag{3-1}
$$

where, using the same notation as (Recker, 1995): $\mathbf{Z}^h$ is the travel disutility associated with the travel/activity pattern adopted by household h; $\mathbf{N}_h$ is the set of all abstract nodes associated with household h (including those associated with refueling); $X_{u,w}^{v,h}$ is a binary decision variable equal to unity if vehicle v of household h travels from activity u to activity w, and zero otherwise; T_u^h is the time at which participation in activity u of household h begins; Y_u^h is the total accumulation of either sojourns[8] or time spent away from home on any tour, of household h on a particular tour immediately following completion of activity u; $\mathbf{V}_h$ is the set of vehicles available to the household; and $\mathbf{A}^h$ is a matrix of spatial, temporal constants as well as the tour length limit. For the details of Equation (1), readers are referred to Recker (1995). We modify the original HAPP for refueling (HAPPR) and for charging (HAPPC) for HFCVs and BEVs, respectively. (Details of the modifications can be found in the Appendices).

For this study, we ignore direct costs and define the travel disutility as a linear combination of three measurements: 1) the total extent of the travel day, 2) the delay of return to home by trip chaining multiple out-of-home activities, and 3) the travel times, using respective weights $\beta_h^E, \beta_h^T, \beta_h^D$, as estimated in Chow and Recker (2012) on the same sample used in this study:

[8] We have used the total accumulation of sojourns, and a maximum capacity of 4 (D = 4).

$$Z_h = \beta_h^E \cdot \sum_{v \in \mathbf{V}^h} (T_{2n_h+1}^{v,h} - T_0^{v,h}) + \beta_h^T \cdot \sum_{w \in \mathbf{P}_h^+} (T_{w+n_h}^h - T_w^h) + \beta_h^D \cdot \sum_{w \in \mathbf{N}_h} \sum_{u \in \mathbf{N}_h} \sum_{v \in \mathbf{V}^h} t_{u,w}^h \cdot X_{u,w}^{v,h}$$

Although these weights are estimated individually in Chow and Recker (2012) and used to account for the heterogeneity of individual travelers, average weights from that study are used in this paper to quantify the travel disutility in an objective manner: $\bar{\beta}_h^E = 0.84, \bar{\beta}_h^D = 0.74, \bar{\beta}_h^T = 3.45$. Also, in order to represent this travel disutility as monetary cost, we use a "value of time" of \$30 per hour for time spent on traveling (Brownstone and Small, 2005). Keeping the same ratio, this translates into the monetary rates of $\beta_h^E = \$7.3/hr$, $\beta_h^D = \$6.4/hr$, and $\beta_h^T = \$30/hr$. We define the "inconvenience" as the increased monetary value of travel disutility from the reported travel patterns caused by operating an AFV.

Travel pattern data are drawn from the California Statewide Travel Survey for households that reside in Orange County and Los Angeles County. Of five "early adoption" clusters identified (CaFCP, 2012), three are in these southern California counties. A total of 392 travel patterns are selected as the sample that show exclusive use of each vehicle for each member. On average, these travel patterns had a total of 3.75 (minimum of 2 and maximum of 12) trips and traveled for 1.07 hours (minimum of 0 and maximum of 6.31 hours) and 29.69 miles (minimum of 0.03 and maximum of 278.38 miles). Travel disutility in monetary terms (based on the value of time assumed above) for the existing patterns was \$163.84 (minimum of \$3.33 and maximum of \$514.00).[9]

[9] We, of course, recognize that these values are fictitious, based on the premise that all measures of time expended are costs, and not benefits. Certainly, the 'extent of the travel day' includes time spent on activities, which would be beneficial. Our argument here is that, since we assume that the time spent on activities is constant among all scenarios, the differences in time costs between activity patterns for all scenarios represents only differences in travel inconvenience, i.e., the time benefits of activity participation cancel in comparisons.

3.3 HFCVs: INSERTION ANALYSIS HAPPR

For HFCVs, we assume that (at least initially) refueling infrastructure is minimally provided. In fact, providing the initial refueling opportunities has been the biggest challenge for the adoption of HFCVs. The cost of building one hydrogen refueling station is estimated to be between $1and $5 million (CaFCP, 2009) in the early-adoption stage, and to support these vehicles' mobility, multiple stations need to be in service. Many studies have focused on the strategic provision of minimum refueling infrastructure (Kang and Recker, 2013b; Stephens-Romero et al., 2010, 2011; Kuby et al., 2009; Nocholas and Ogden, 2006; Melaina, 2003). In practice, the California Fuel Cell Partnership (2012) adopted the strategy from Stephens-Romero, et al. (2010, 2011), and proposed 68 hydrogen stations—36 in southern California (Figure 3-1) — for the "pre-commercial" period. In this paper, we focus on southern California and assume that the 36 of these 68 stations that are located in southern California are the only refueling opportunities of the pre-commercial period.

In this analysis we insert one refueling trip per vehicle, under the assumption that the daily travel diaries reported in the survey represent a "typical" travel/activity day for the respondent. This is because the range of HFCVs is between 190 miles – 430 miles[10], and daily travel distance is found to be less than 60 miles (Kang and Recker, 2009). Therefore, analysis based on fuel inventory is not appropriate for one-day data. The focus here is on the determination of inconvenience of limited refueling opportunities within empirical daily routines. However, we note that fuel inventory can be included, as shown in Kang and Recker (2013a), and also in Appendix 3-C.

[10] Chevy Equinox FC (190 miles), Honda FCX Clarity (240 miles), Toyota FCHV-adv (Fuel Cell Hybrid, 431 miles)

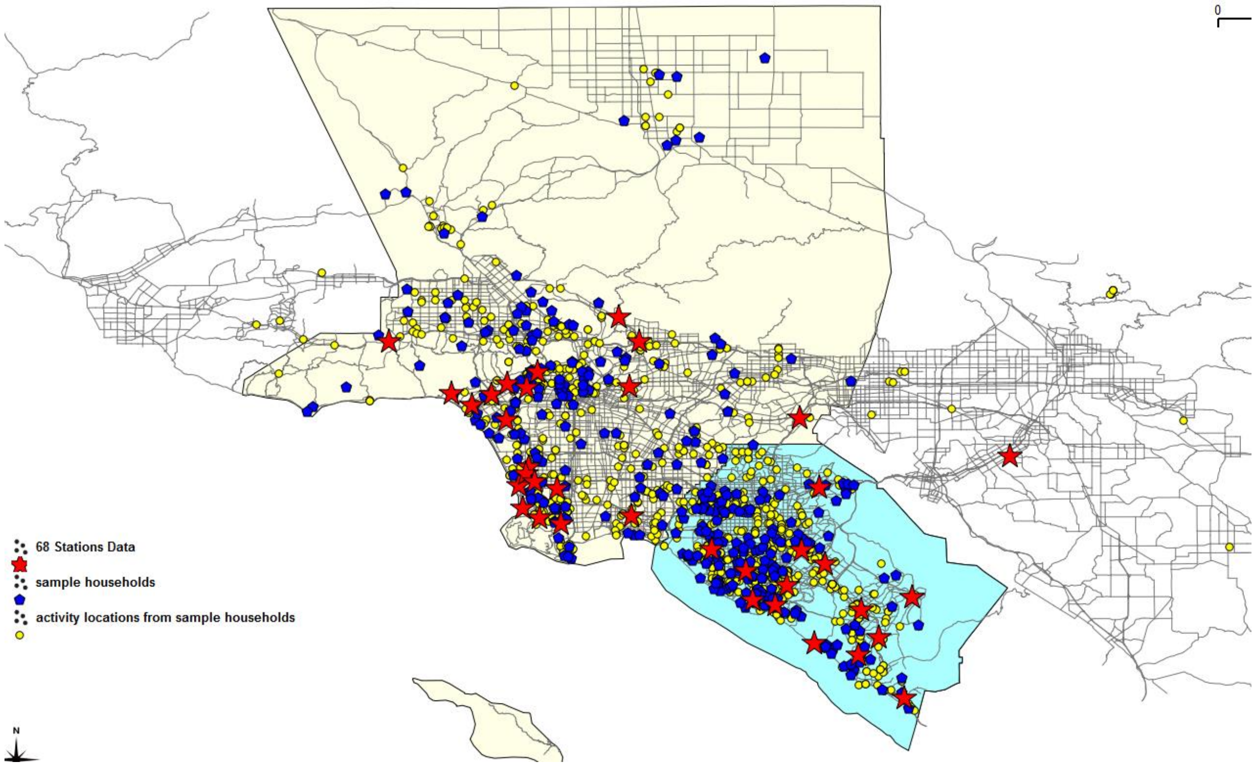


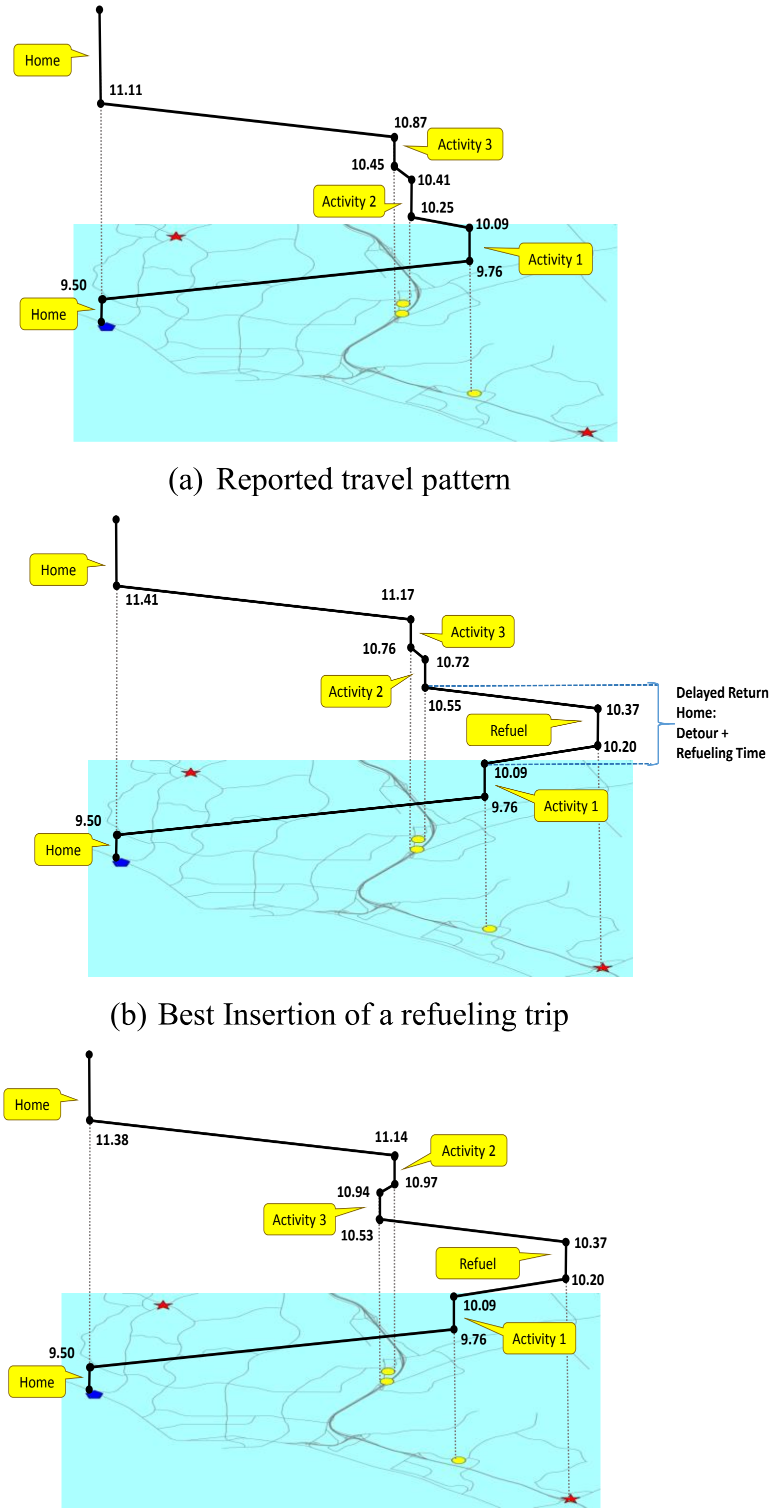
Figure 3-1 Sample data and proposed 36 hydrogen refueling stations in Southern California

The first analysis, based on inserting a refueling trip while keeping the reported travel pattern is as follows: for every possible insertion between trips, k, insert a trip to each refueling stations, r, and add the following trip: 1) select r with the least detour cost; 2) then add the remaining trips to complete the travel patterns; and 3) for the k_r full-day travel patterns, for which a refueling trip is inserted, select the pattern among k_r with least delayed completion of travel day. This process is equivalent to selecting a station r and an insertion k with the smallest detour time of $t^h_{k,r} + t^h_{r,k+1} - t^h_{k,k+1}$. This process is presented in Kang et al. (2013) and also in Appendix 3-A. Here, we are only minimizing the time of detouring between activities but this search process can be categorized as a very simple case of dynamic programming for the Pickup and Delivery Problem with Time Windows (PDPTW), or Vehicle Routing Problem (VRP), used in Desrosiers et al. (1986), Kang and Recker (2013a), where the path is predetermined. The search process of checking all of the possible routes and selecting the best is the core of the process.

The second analysis is allowing the households to revise their travel decisions involving travel sequence, tour construction and departure time if it results in lowering their inconvenience. This is accomplished using HAPP-Refueling (HAPPR) which extends the Location Selection Problem for the Household Activity Pattern Problem (Kang and Recker, 2013a). A set of potential refueling activities (each with specific location common to all households), is given to the model, $\mathbf{P}_h^{R+}$, only one of which is to be completed. The model decides which refueling location to visit while simultaneously making decisions of travel sequences and departure times to perform a set of compulsory activities with predetermined locations, $\mathbf{P}_h^{A+}$. This model has been used for an individual refueling application in Kang and Recker (2013b). The model minimizes the total travel disutility as defined earlier and the full model is in Appendix 3-B. Time windows are generated in the same manner as Kang and Recker (2013b) which extends the method from Recker and Parimi (1999).

An illustrative example of these procedures is shown in Figure 3-2. Household #2150230 drawn from the sample has three out-of-home activities, trip-chained in a single tour. The travel/activity pattern has the monetary travel disutility value of \$50.79 and travel of a total of 0.69 hours (22.89 miles). The best refueling insertion for this household is from activity 1 to activity 2, and the detour time is $t_{1,r=23}^{h=2150230} + t_{r=23,2}^{h=2150230} - t_{1,2}^{h=2150230} = 0.13$, where $t_{u,w}^{h}$ refers to travel time from activity location u to activity location w of household h, resulting in a 0.3-hour delay in return-to-home (including the refueling time of 0.17 hours). Its monetary travel disutility is

$71.86, with travel of 0.83 hours (29.06 miles). When travel decision revisions are allowed, this household "optimally" performs activity 3 before participating in activity 2, resulting in travel time savings from utilizing the freeway connection between the activity 1 location and activity 3 activity location. By changing the sequence of activities, the travel time saving is 0.03 hours. Its total monetary travel disutility of this revised pattern is $ 63.59, with travel time of 0.80 hours and distance of 26.6 miles. Adding flexibility in travel decisions shows how this driver can potentially adjust to reduce the inconvenience of $21.07 to $12.80; we note that a portion of this reduction (i.e., the travel time savings) could also be achieved with the ICEV under similar sequencing, but under our primary assumption, this would presumably also result in sub-optimal activity participation—whether or not the individual would be willing to accept a lower value of activity participation in order to accommodate use of an AFV remain an open question, not addressed here. The lower and upper bound inconvenience caused by operating a HFCV is thus calculated as $12.08 and $21.07, respectively, for this driver.



(a) Reported travel pattern

(b) Best Insertion of a refueling trip

(c) Rescheduled with a refueling trip using HAPPR

Figure 3-2 Travel patterns with a refueling trip

HAPP is a variation of the Vehicle Routing Problem (VRP) which is NP-hard and its computation is not always easily implemented. For HAPPR, more activity locations are put into the model, consequently placing additional computational burden. For solution, we follow the label generation method from Kang and Recker (2013a) which extends the work from Desrosiers, et al (1986) for HAPP with destination choice. The method generates all feasible sequences of activities (called "label") with elimination rules to remove infeasible labels during the process. Due to the objective function of HAPP, that includes temporal variables, $\mathbf{T}^h$, as well as flow variables, $\mathbf{X}^h$, arrival and departure times are not decided during the process of label generation as in Desrosiers, et al. (1986). Instead, earliest and latest time windows are imposed during the process of label generation and each completed path label is put to a simple linear program (LP) to minimize cost associated with temporal variables, $\mathbf{T}^h$, as in Kang and Recker (2013a).

There are 36 planned hydrogen refueling stations within the scope of this study, and even with the method from Kang and Recker (2013a), the computation is a challenge. To ease the computational burden, a set of non-inferior stations is selected for each travel pattern to be added to HAPPR. For each station, if there exists any other station with smaller travel times to/from all locations including out-of-home compulsory activities and home, that station is removed from the non-inferior set. The average size of non-inferior set is 3.08, with a minimum of 1 and maximum of 16. These numbers also represent the number of stations that are assumed to be considered by the driver for refueling throughout the day since all activity locations and potential paths throughout the day are considered when creating the non-inferior set.

A total of 7 of 392 HAPPR problems reached "out-of-memory" error while generating labels. These HAPPR problems have an average of 6.1 out-of-home compulsory activities, while those HAPPR problems that did not reach “out-of-memory” error have an average of 2.4 activities. A total of 83 of 392 HAPPR problems were found to be infeasible with respect to the estimated time windows imposed—these are the additional, unobserved, constraints that are not part of the insertion analysis. For these travel patterns, patterns from the insertion analysis are used for both upper and lower bounds[11]. For 302 travel patterns, an average of 155 (minimum of 3 and maximum of 3225) complete labels are put as LP with computation time of 153 seconds (minimum of 2 seconds and maximum of 2907). Most of the computation time is used to solve simple LPs, taking about 1 second per LP.

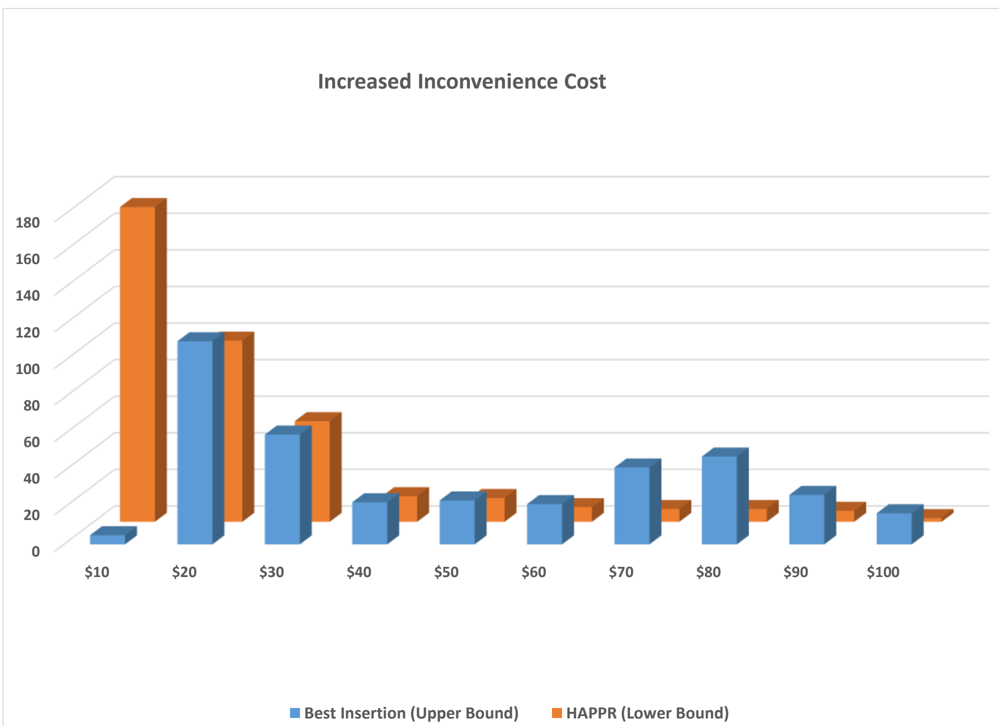
For the sample, on average, when a refueling trip in inserted to minimize the deviation time while keeping the sequence of reported travel patterns, the travel disutility cost is \$219.25 (minimum of \$17.99 and maximum of \$600.13); this represents an additionally-imposed inconvenience of \$39.04 (minimum of \$8.85 and maximum of \$208.48). This is an average of 0.43% increase in the inconvenience, and increases of 0.15 hours in travel time (and, an increase of 8.12 miles in travel distance).

When travel decisions are revised to minimize the total travel disutility cost through HAPPR, a total of 293 travel patterns out of 316 were rescheduled. On average, the travel disutility cost is \$188.69 (minimum of \$12.89 and maximum of \$600.13). Additionally imposed inconvenience is \$19.38 (minimum of -\$75.66 and maximum of \$208.48). The reason for negative inconvenience cost is that the rescheduling allowed a lower travel disutility cost than that of the originally-reported travel pattern. This is caused by using the average weights in the

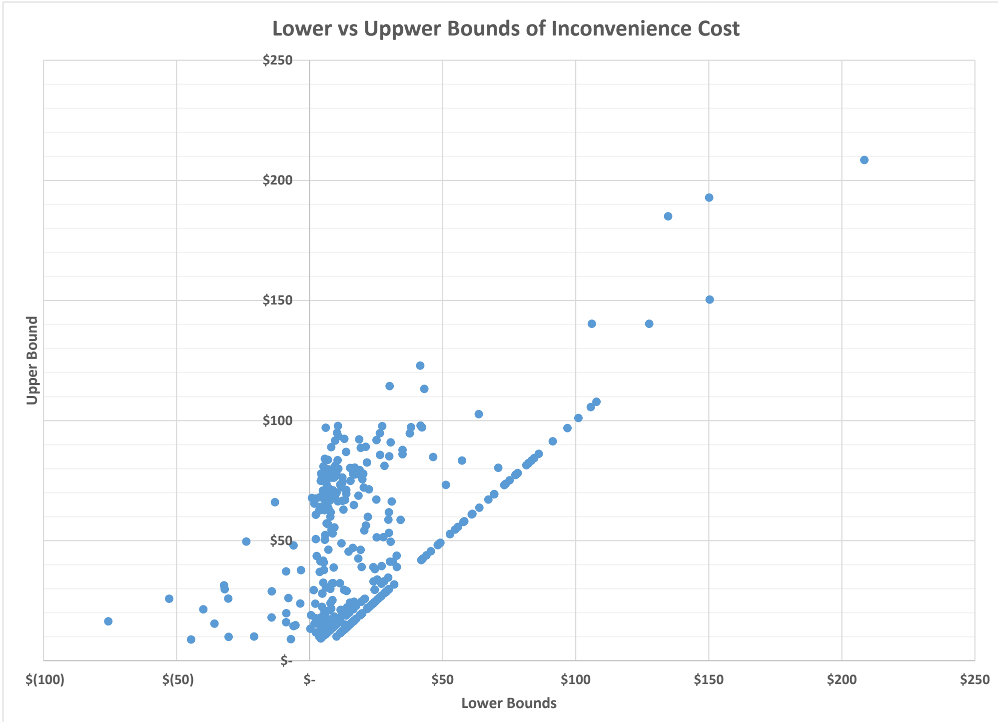
[11] It is noted that insertion analysis finds the temporal shift of reported patterns, but does not consider whether new patterns respect time windows. And therefore no insertion pattern is infeasible. While it does not account for temporal constraints, it provides estimates of potential inconvenience via delayed return home time, instead of reaching infeasibility.

disutility functions of the individual drivers—we assume that drivers have their own individually-perceived costs which they try to minimize and which resulted in the reported pattern. These "more" optimal results are caused by using average values; however they give insights of "average" or "norm" costs, showing an average of 0.11% increases in the travel disutility, and an average increase in travel times of 0.22 hours (and, an increase of 8.70 miles).

The histogram of lower and upper bounds of inconvenience cost is in Figure 3-(a). In Figure 3-(b), the inconvenience costs are plotted between lower and upper bounds. The straight line represents 90 households with same lower and upper bounds.



(a) Histogram of Inconveniences



(b) Lower vs Upper Bounds of Inconvenience

Figure 3-3 Lower and upper bounds of inconvenience cost of operating a HFCV

3.4 BEVs: DELAY ANALYSIS AND HAPPC

The challenges of BEV adoption are rather different from the infrastructure provision problem of HFCVs. Although there has been some effort looking into the optimal provision of public charging stations as in the efforts associated with hydrogen refueling station provision, Level 1 and level 2 charging can be done at any electricity outlet. Here, the challenges are due to the vehicle's limited range and relatively long charging times compared to ICEV or HFCV refueling, which can impose significant inconvenience for drivers if the daily driving range (or even any particular tour) is greater than the range of the battery. While most of daily driving is less than the fully charged BEV range of 60 – 250 miles, for a subset of days/vehicles for which travel is more than the given range, battery and charging status needs to be monitored.

To assess delays in the travel pattern and the minimum inconvenience caused by the delay of individuals' charging we assume that charging can occur during all activities regardless of activity types, locations, or durations; i.e., we impose charging behavior of "charge everywhere whenever," even for travelers driving less than the full range. We note, however, that charging behaviors are expected to be different for drivers, and charging behavior determines the charging profiles and therefore increased electricity demand (Zhang et al., 2011; Kang and Recker, 2009; Gondor et al, 2007).

For the delay analysis in which travelers replicate their respective reported travel patterns (but now with a BEV), we define E_k as the electric battery inventory upon the arrival at the location of activity *k,* for every activity location. Then the battery status is updated as $E_{k+1} = \max\{E_k + r \cdot (b_k - a_k), R\} - e_{k,k+1}$ where a_k , b_k are the arrival/departure time at k^{th} activity, R is the battery capacity, r is the charging rate, s_k is the reported duration of activity *k*, and $e_{k,k+1}$ is the battery consumption for traveling from *k* to *k+1* (here, this is equivalent to travel distance, $d_{k,k+1}$). The maximum possible charging amount responds to $(b_k - a_k)$, and the charged amount at the time of departure is the maximum of $E_k + r \cdot (b_k - a_k)$ and the capacity.

Delay occurs when the battery status at the end of an activity is less than that needed to complete the length of a trip that follows. If this is the case, the driver waits for the vehicle to be charged to make that specific trip as:

$$b_k = \begin{cases} a_k + \dfrac{e_{k,k+1} - E_k}{r} & ,E_k + r \cdot s_k \le e_{k,k+1} \\ a_k + s_k & E_k + r \cdot s_k \ge e_{k,k+1} \end{cases}$$

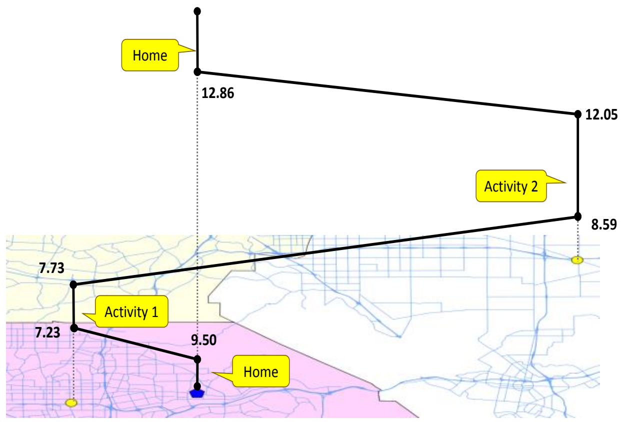
The detailed algorithm for this computation is presented in Appendix 3-D. Compared to the HFCV insertion algorithm (Kang et al., 2013), the algorithm includes the additional feature of battery status minus a refueling trip insertion. For the analysis, we assume a vehicle capacity of 20 kWh, 60 miles[12]. As stated earlier, two charging options are tested since they require relatively little economic infrastructure investment compared to level 3 public fast charging, and are considered to be relatively easily accessible. Level 1 charging assumes 15-hour charging time for the full battery, equivalent charging rate of 4 miles per hour. For Level 2 charging, the corresponding time to full charge is 5.6 hours, which translates to a charging rate of 10.7 miles per hour (Schroeder and Traber, 2012). (In the analysis, battery capacity is kept in the unit of miles.) For presentation purposes, we assume Level 1/ Level 2 charging only and that Level 2 charging is available everywhere.

In the second analysis, changes in travel decisions regarding travel sequences, including departure times are allowed to change from the reported patterns. HAPP-Charging (HAPPC) is extended from the original HAPP (Recker, 1995) with battery status tracking as shown in Appendix 3-E. An additional decision variable set, $\mathbf{E}^h$, is introduced to keep track of the electric battery status. Charging behavior for HAPPC is assumed to follow the protocol that the battery status can be charged up to the capacity, or by the charging time (which may or may not include intentional waiting for charging following the completion of an activity).

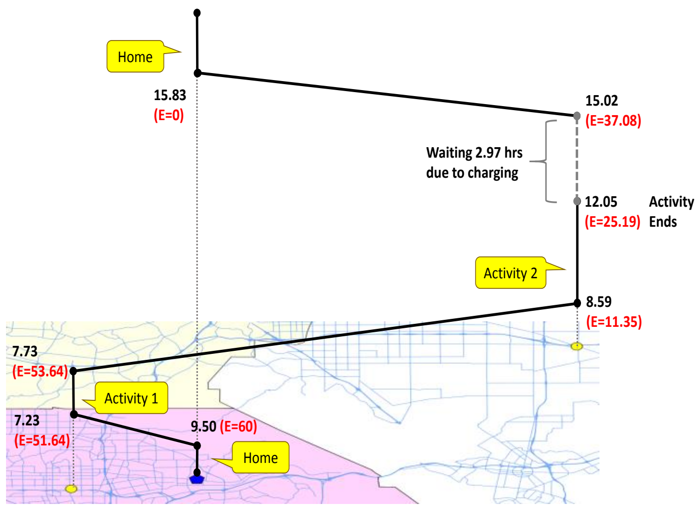
Figure 3-5 presents an illustrative example of the results of the procedures for household # 20113556 of the sample. This driver engaged in two activities, grouped together in a single tour; the current travel disutility cost of the pattern is $163.12, and total travel distance is 87.72 miles, which is greater than the vehicle range of 60 miles. If this traveler keeps his/her current

[12] This is a conservative measure of many EPA estimated driving ranges: Nissan Leaf (24Kw, 73 miles), Mitsubishi i MiEV (16kWh, 62 miles), Ford Focus (23kWh, 72 miles)

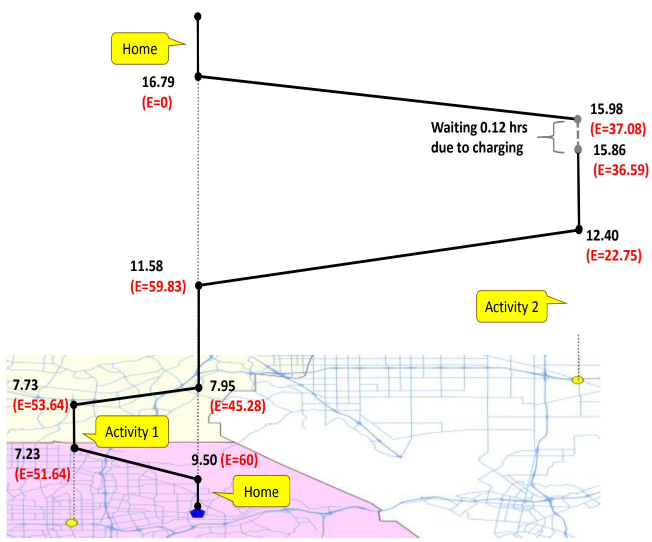
travel pattern with a BEV, return home is delayed by 2.79 hours, waiting for the vehicle to be charged to make the last trip home, assuming Level 1 charging during participating in activities and waiting time. For the BEV case, the travel disutility cost is $201.06 which is an increase of $37.94 over the reported travel pattern. Under the assumption that this driver would attempt to further lower his/her travel inconvenience cost within spatial and temporal constraints, the optimal travel pattern for this traveler is shown in Figure 3-5-(c). Instead of trip-chaining two activities, the vehicle is charged at home, lowering the cost associated with return home delay caused by trip chaining, $\beta_h^T \cdot \sum_{w \in \mathbf{P}_h^+} (T_{w+n_h}^h - T_w^h)$, which may also be a crude surrogate for waiting time or time outside home in this particular example. The extra waiting time for vehicle charging following participation in activity 2 is 0.12 hours. This pattern produces a travel disutility cost of $166.48, which is increased by only $3.36 over the original pattern completed with the ICEV.



(a) Reported travel pattern



(b) Delay analysis



(c) Rescheduled using HAPPC

Figure 3-4 Travel patterns with level 1 charging between trips

A total of 81 of the 392 travel patterns exhibited travel in excess of 60 miles, as shown in Figure 6. The average travel disutility cost for these patterns are $250.05 (minimum of $90.73 and maximum of $514.00). Although drivers with travel distances of consistently more than 60 miles may not be the most likely customers for BEVs at an early-adoption stage, they nonetheless present perhaps the most interesting case for evaluation in that in many cases it can be presumed that such drivers may have numerous days, other than the day of the survey, in which their travel does not exceed this threshold. (Travel patterns in which vehicles traveled fewer than 60 miles are not of interest in this analysis since they can operate without any inconvenience.)

Using Level 1 charging, only 38 vehicles experience waiting time caused by charging; 43 vehicles are able to keep their reported patterns just by charging during activities. By allowing schedule changes, 2 vehicles among these 43 vehicles are able to lower their inconvenience from $66 to $28 and from $38 to $3, respectively. On average, the lower and upper inconvenience costs are $47 and $50, respectively, for vehicles that traveled more than 60 miles (average total travel disutility costs of $297.33 and $300.42 for lower and upper bounds). If we account for travel patterns that traveled less than 60 miles (no delay occurred for these patterns and therefore inconvenience cost of $0), the average inconvenience cost is $10 for both the lower and upper bounds. It is noted that in 6 cases, the optimal travel patterns lowered the reported travel disutility.

Using Level 2 charging, only 13 vehicles experience waiting time increases with an average of lower and upper values of $6.7 and $9.8 for 81 patterns (average total travel disutility cost of \$256.72 and \$259.88 for lower and upper bounds) with more than 60 miles, and \$2 (both the lower and upper) for all 392 patterns.

Compared to the results of HFCVs, apparently there is little that individual drivers can do to ease the inconvenience of operating BEVs. Rather, this inconvenience is highly dictated by charging rates. By switching from Level 1 Charging to Level 2 Charging strategy, the average inconvenience cost, as well as the number of affected travel patterns, dropped significantly. If public charging options become available, it is expected to remove most of the inconvenience.

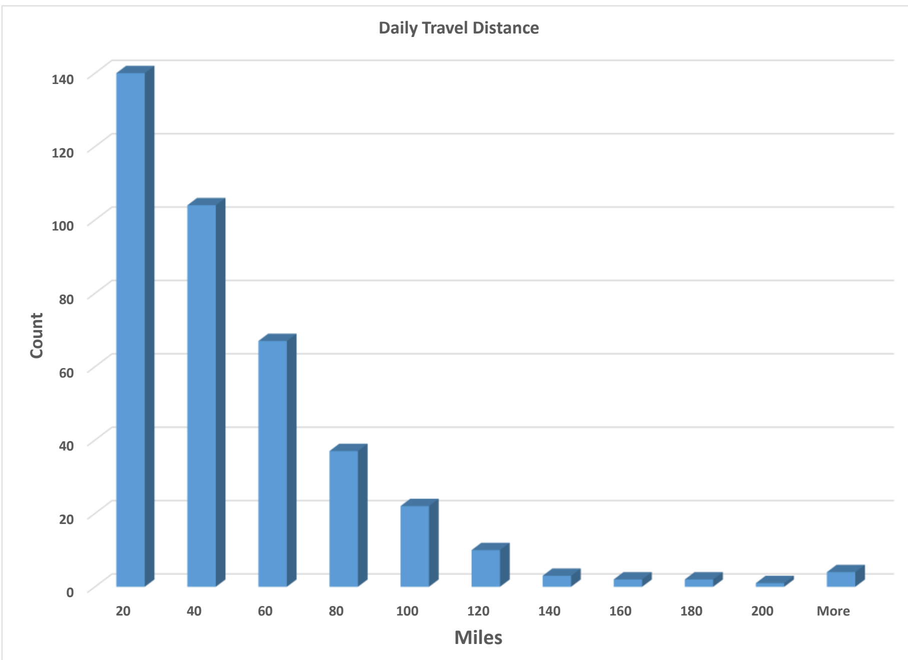
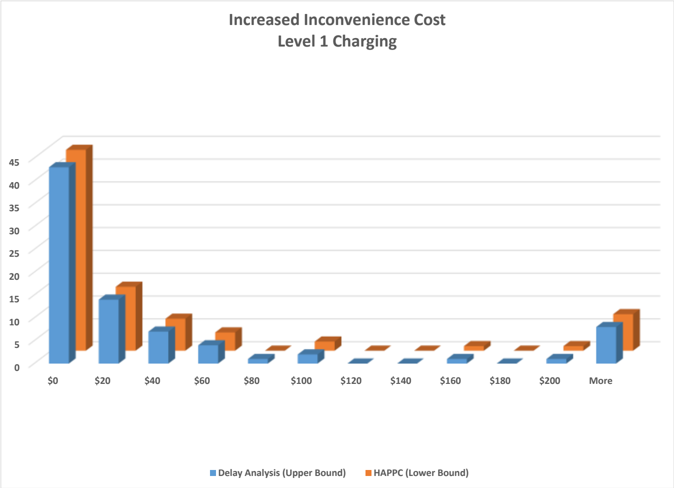
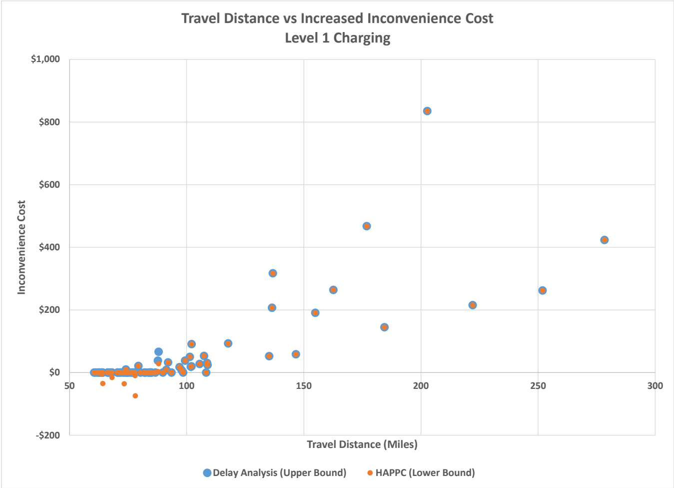


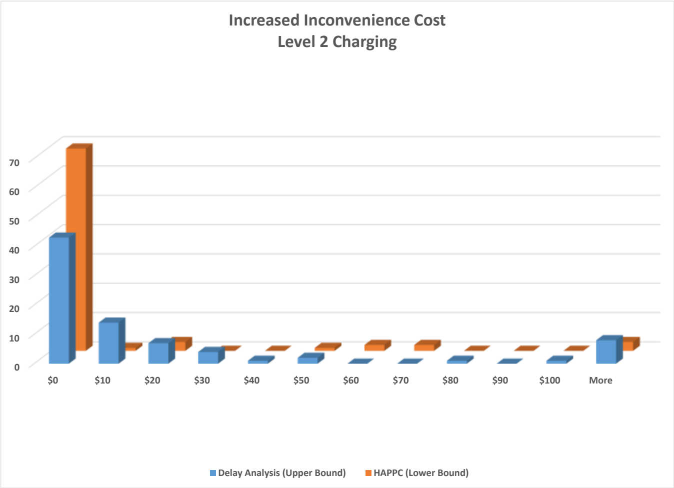
Figure 3-5 Reported daily travel distance



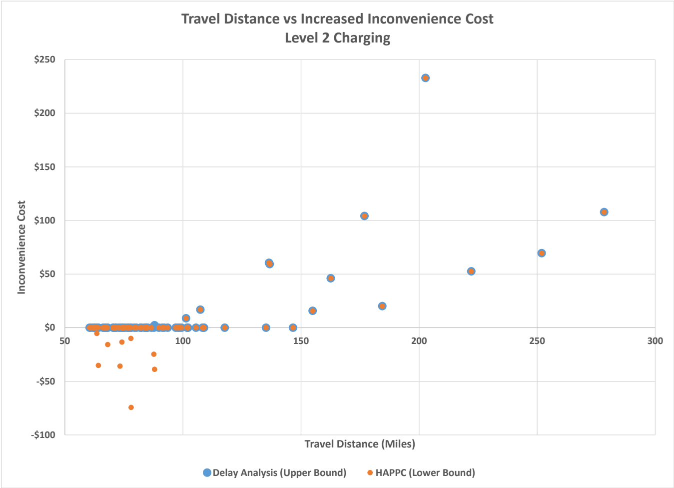
(a) Histogram of Inconveniences: Level 1 Charging



(b) Lower vs Upper Bounds of Inconvenience: Level 1 Charging



(c) Histogram of Inconveniences: Level 2 Charging



(d) Lower vs Upper Bounds of Inconvenience: Level 2 Charging

Figure 3-6 Lower and upper bounds of inconvenience cost of operating a BEV

3.5 POLICY IMPLICATIONS

Inconvenience as measured in this paper can be used to guide some useful policy implications. From the scenarios resulting the lower and upper bounds of two types of vehicles, it is found that the inconvenience cost of operating a HFCV is between \$19 and \$38 for the day of refueling. The inconvenience cost of operating a BEV is between \$47 and \$50 using level 1 charging infrastructure, or between \$6 and \$10 using level 2 charging infrastructure for drivers with more than 60 miles of travel in a day. From these results, we find that it is more critical to deploy faster charging infrastructure for the promotion of BEVs than it is for simply making Level 1 outlets more commonly available. On the other hand, for the promotion of HFCVs, it may benefit from such rather “soft” strategies as awareness campaigns or providing drivers with information on the locations of refueling stations in an effort to mobilize the behavioral changes that will deliver the lower-bound, best, usage case.

These results can also be used as guidelines for state and federal government subsidy policies since we have derived the inconvenience in units of monetary cost. Many state governments offer various subsidies on fuel cost, household charging infrastructure installment, vehicle purchase, and etc. NCSL (2013). It can be argued that the inconvenience cost needs to be additionally subsidized—at least during the early adoption stages—when the inconvenience is apparent as verified in this analysis. For example, inconvenience of limited refueling opportunities can be subsidized as \$991 per year, assuming one refueling trip per week[13] and the lower bound inconvenience cost of \$19. Although there is travel disutility cost associated with ICEVs, we assume that it is zero since a refueling trip to a gasoline station typically does not require the preplanning as for an alternative fuel with limited refueling opportunities (Kitamura and Sperling, 1987; Kelly and Kuby, 2012). For BEVs, it may be argued that inconvenience cost

[13] Daily driving range of 29.69 miles and 250 mile fuel range result in one refueling activity in 8.4 days.

of \$3,431 needs to be subsidized if a vehicle is driven more than 60 miles for 20% of the time (as seen in the data—81/392 patterns exceeded the full-charge range capability of a BEV) when only the basic level 1 charging infrastructure is provided. With widespread installment of Level 2 charging infrastructure, which is more economical than either Level 3 or fast charging (Schroeder and Traber, 2012), this subsidy can be reduced to \$453.

Using the analysis described in this paper, similar policy guidelines can also be extracted to other units such as travel times or travel distance to compensate for increases in such units.

3.6 CONCLUSION

In this chapter, we presented analyses to help quantify the monetary cost of the inconvenience of operating an HFCV and a BEV. Assuming that people would continue to participate in the same set of activities done with ICEV, additional conditions that are specific to each vehicle type are imposed. For HFCV, a refueling trip is inserted for all travel patterns to measure the inconvenience of limited refueling infrastructure. For BEVs, battery status is monitored along with consumption and charging actions. The inconvenience is defined as the difference between the monetary cost of travel disutility of the reported travel pattern and that of the travel pattern with an AFV.

Two scenarios are tested representing the upper bound, the worst case and the lower bound, the best case. Although these two extreme scenarios may not likely be the only feasible travel patterns, they provide the bounds of inconvenience of operating AFVs for which quality data are currently unavailable due to their limited market penetration. The first scenario tries to keep the currently reported travel pattern; a refueling insertion that deviates the least is generated via deviation analysis for HFCVs, and delay is assumed to occur when charge status at the time of completion of activities is not sufficient to make the next trip. The second scenario derives the "best" solution given the set of activities to perform and also the conditions of vehicle types. HAPPR and HAPPC are developed from the original HAPP to accommodate those conditions. Results suggest that for HFCVs, the inconvenience of refueling is \$19 - \$38, and for drivers with more than 60 miles would experience monetary inconvenience of average \$47-\$50 and \$6 - \$10 with level 1 and level 2 charging infrastructure respectively. These values provide insight to policy guidelines and strategies for AFV promotions.

Calculating the lower bound is possible using the property of HAPP that it includes temporal and spatial constraints a traveler faces in completing various activities throughout the day. This property allows simulation of travel associated with previously unobserved circumstances when some level of behavioral change is expected. For AFV applications in this paper, it is possible to impose additional vehicle physical characteristics along with spatial and temporal constraints. Although not addressed in this paper, it is noted that with the incorporation of activity participation redress decisions (Gan and Recker, 2012), the results would be more encompassing when the circumstances dictate changes not only in the itinerary decisions but also in the activity participation decisions.

CHAPTER 4 LOCATION-HAPP PROBLEM

A facility location strategy that considers individual vehicle's scheduling and routing is presented. By coupling a location strategy of the Set Covering Problem and a routing and scheduling strategy of the Household Activity Pattern Problem, this problem falls into the category of Location Routing Problem. This problem also introduces a tour-based approach to facility location siting, with tour-construction capability within the model. There are multiple decision makers in this problem: the public sector as the service provider, and the collection of individual households that make their own routing decisions to perform a given set of out-of-home activities together with a visit to a refueling location. A solution method that does not require the full information of the coverage matrix is developed to reduce the computational burden. When compared with the point-based Set Covering Problem, the results indicate that the minimum infrastructure requirement may be overestimated when vehicle (refueling demand) - infrastructure (refueling supply) interactions are excluded.

In the dissertation, we specifically set the service as "hydrogen refueling" service and apply the model to site initial hydrogen stations as the hydrogen station siting itself is an area of great interest in recent years.

4.1 BACKGROUND

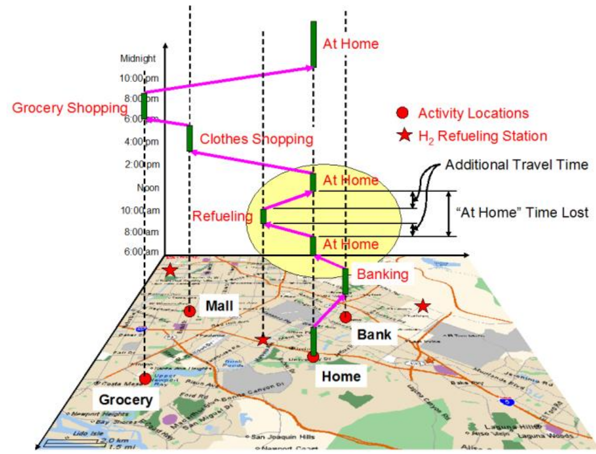
Hydrogen fuel cell vehicles (HFCVs) operate on an electric motor powered by a hydrogen fuel cell, producing zero emissions during vehicle operation. Their fuel energy efficiency is 40-60% while that of Internal Combustion Engine Vehicles (ICEVs) is around 20%.Because a variety of energy sources can be used to produce the hydrogen vary, increased adoption of HFCVs may lead to lower fossil energy dependency, and may ultimately draw more from renewable energy sources. The main advantage of HFCVs over Battery Electric Vehicles (BEVs) or Plug-in Hybrid Electric Vehicles (PHEVs) is their similarity to conventional ICEVs. Because the full driving range of HFCVs is competitive with ICEVs, as is the time required for refueling as low as 3 minutes, use of HFCVs will not require significant change to a driver's previous behavior.

Despite these advantages for success in the automobile market, HFCVs face a tremendous obstacle against widespread early adoption: an enormous investment is needed to provide the refueling infrastructure critical for HFCV operations. According to California Fuel Cell Partnership (2009), one hydrogen refueling station is built at between $ 1.5 and 5.5 million. The so-called "chicken-and-egg" problem due to insufficient consumer demand needed to support the building of hydrogen refueling stations—versus not having in place enough stations to enable the consumers to purchase the HFCVs remains largely unsolved. Therefore, It has been presumed that there likely will be a public sector role in making the initial investment likely needed to "jump start" the hydrogen infrastructure that could lead to practicality of HFCVs. Identifying the minimum requirement for initial siting of hydrogen stations, and maximizing the effect of the public investment likely has received increasing attention, albeit mainly on a general scale.

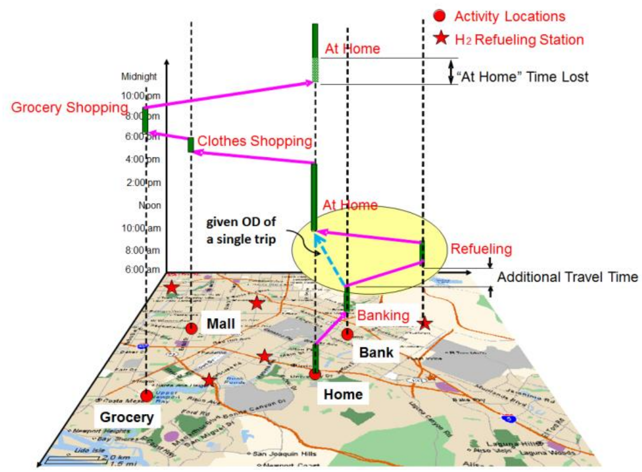
Melaina (2003) examined two stage initial conditions for the US road system: stage 1 to support early adopters' travel and stage 2 to support initial mass production; the minimum numbers of stations according to the criteria were estimated to be 4,500 for stage 1 and 17,700 stations for stage 2. A number of studies have taken facility location siting approach for optimizing the initial investment of HFCVs refueling infrastructure. Nicholas and Ogden (2006) and Nicholas et al. (2004) applied the p-median problem (where p is the number of facilities to be located) to minimize the total travel times to the nearest refueling stations for a set of trips originating from each TAZ in the California metropolitan areas of Sacramento, San Francisco, Los Angeles and San Diego. From various scenarios, the density of stations to achieve certain levels of average driving time to nearest station and the population density are shown to have an inverse relationship. For example, the Los Angeles metropolitan area requires 6.8 % of current gasoline stations to provide hydrogen refueling service to achieve an average of 3 minutes driving time while the Sacramento metropolitan area requires 15.8%. Stephens-Romero et al. (2010) and Stephens-Romero et al. (2011) applied a Set Covering Problem for a series of early adoption communities in Southern California targeted by auto-manufacturers, and found the percentages of current gas stations "refitted for hydrogen refueling" that would guarantee the tolerable travel time from all nodes in the areas. Wang and Lin (2009) took the same approach of Set Covering Problem with the set comprised of long distance (inter-city) path demands. Kuby and Lim (2005) developed the Flow-Refueling Location Problem from the Flow Capturing Location Model (Hodgson, 1990) assuming that vehicles can be refueled between the origin and the destination subject to a range limit. Based on this model, a "clustering and bridging" strategy of hydrogen infrastructure investment plan was suggested (Kuby et al., 2009). A recent development allows deviation from the shortest path (Kim and Kuby, 2012) of a given OD pair. In addition to the four different types of models of refueling station locations cited here, there are a number of other models, many of which extending the Flow Capturing Location Model. MirHassani and Ebrazi (2012) categorized these existing models into three different groups: node-based, arc-based, and path-based. Readers are referred to their study for a detailed literature review of up-to-date hydrogen refueling station siting studies.

By the scope of applications, Flow Refueling Location Problem variations deal with long distance, inter-city trips Nicholas and Ogden (2006), Nicholas et al. (2004), Stephens-Romero et al. (2010), and Stephens-Romero et al. (2011) on metropolitan/city regions. Melaina (2003) considers both metropolitan areas and inter-city approaches.

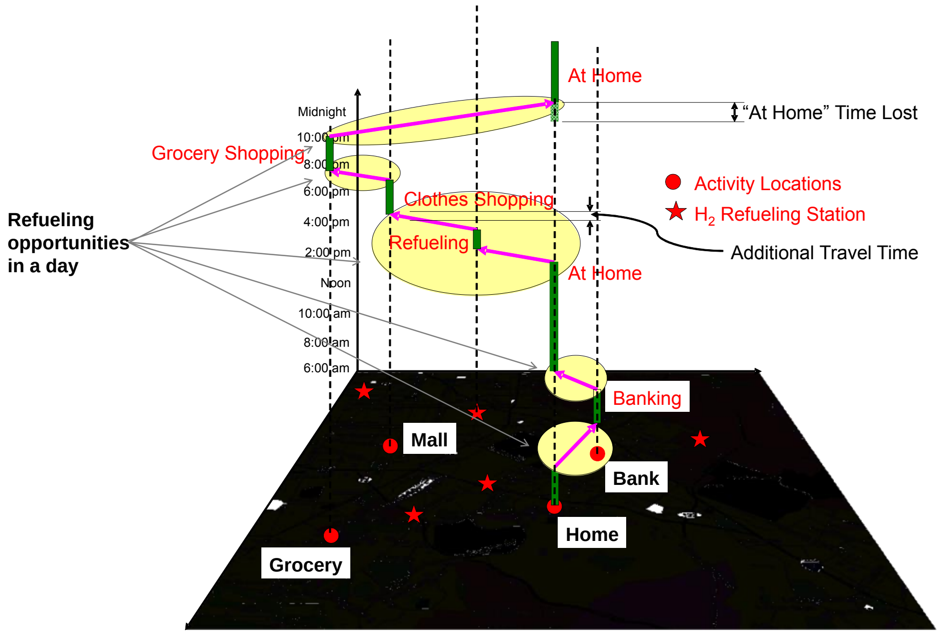
Here, we argue that accessibility to refueling is best measured within the context of the touring of trips in which drivers commonly insert such activities in planned tours with other routine activities as the need arises (typically with a frequency of once in several days, depending on driving behavior), for daily travels which are most likely to be traveling within the metropolitan area. From a travel behavior modeling perspective, we can visualize such accessibility measurements as in Figure 1, which illustrates vehicle-infrastructure interactions associated with different types of models, together with routing and scheduling interactions. Figure 4-1-(a) represents models using a single point, and Figure 4-1-(b) using a single trip (OD). Although the representation of demands in the latter case works well for long distance trips (Kuby et al., 2009; Kim and Kuby, 2012), about 70-80% of daily cumulative travel distance is less than 40 miles (Kang and Recker, 2009), which is not anywhere near the full range of HFCVs. In order to investigate daily travel comprised of a several tours of rather short distance trips, here we expand the temporal scope of the problem to a full day and address the "optimal" selection of the refueling activity (an individual decision based on the particular travel-activity pattern of that individual), together with the "optimal" siting of the refueling locations (a decision of the provider based on the collective decisions of the individuals in executing their travel-activity patterns) as in Figure 4-1-(c).



(a) Refueling trip starting from home ending at home



(b) Refueling trip deviating from a given trip



(c) Refueling trip opportunities throughout the day

Figure 4-1 Alternative HFCV usage patterns for a single day

Specifically, based on the concept that refueling trips are mostly linked to other primary activities, we propose a facility location problem with full-day household scheduling and routing considerations. This is in line with Location-Routing Problems, where decisions of facility locations are influenced by possible vehicle routings. The model we propose takes the Set Covering Problem (with the "set" comprising household travel patterns with HFCVs) as a location strategy, and the Household Activity Pattern Problem as the scheduling and routing algorithm for performance of the households' activity schemes. The proposed model solves the location problem simultaneously with multiple routing problems that include visitation to one of the available locations for refueling. Due to the computational complexity of the problem, a solution algorithm specific to the model is also developed.

4.2 LOCATION-ROUTING PROBLEM AND LOCATION-HAPP PROBLEM

The Location-Routing Problem (LRP) refers to locational problems that consider optimal locations of facilities from a vehicle routing perspective. It evolved from the management point of view that the location of distribution centers and the routing of vehicles to visit all customers from each distribution center are closely interrelated. Location Routing Problems take possible routing patterns around these depots, and the costs associated with them, into account at the time of locating distribution centers. Typically, an LRP formulation includes three parts: location, routing, and allocation. During the last few decades, the Location-Routing Problem has been studied widely, resulting in various problem formulations and numerous methodological advances (Min et al., 1998; Nagy and Salhi, 2007). The practical applications addressed by LRPs are not limited to the private sector. While most cases involve the shipping industry or decision making of private firms related to product/goods distribution or plant locations, several papers present applications, such as medicine (Or and Pierskalla, 1979; Chan et al, 2001), waste/hazardous materials (List et al., 1991), or, more generally, undesirable activities (Cappanera et al, 2004), of interest to the public sector.

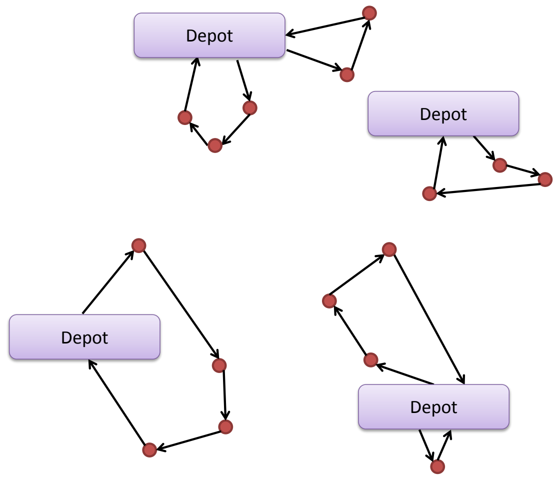
To categorize by structure, the standard LRP minimizes the overall cost, comprised of depot cost, and vehicle routing cost, which includes tour planning by which a set of vehicles (one or more) traverses customer locations from/to the depots. Many previous works, both in the public sector as well as in the private sector, have this structure. Some of the non-standard structures include waste/hazardous materials applications (List et al., 1991) that often use multiple objective functions and replace tour problems with transportation problems, the many-to-many LRP of the shipping industry that includes customers sending goods to others (Nagy and

Salhi, 1998), and vehicle routing allocation problem (Labbe et al., 2005). Although an important aspect for many public sector infrastructure problems (and even for the private sector), there has not been substantial effort with a focus on covering problems.

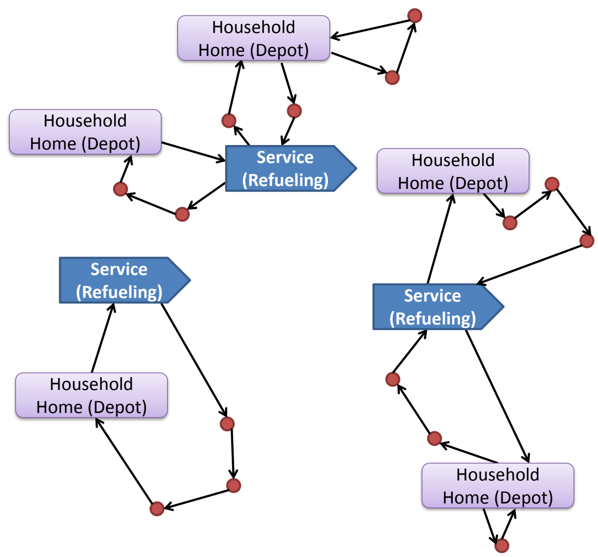
Here, we formulate a Location–Household Activity Pattern Problem (Location-HAPP) that can be identified as an LRP in a broad sense. The goal of the problem is to help identify the minimum hydrogen refueling service infrastructure—which may be either fully provided or subsidized by the public sector—that is necessary to support the initial phases of HFCV growth, i.e., before a sustainable private sector market is generated. The Location-HAPP takes the set covering problem as the basis for the location part in the LRP and couples it to an activity-based modeling framework to estimate the basic coverage of refueling stations needed to support early deployment of HFCVs. Rather than defining basic coverage by a set covering problem based on the range of HFCVs, which varies from 190 to 430 miles, the term "basic coverage" here is defined in terms of "the maximum tolerable inconvenience" level to drivers for hydrogen refueling. A requirement likely to result in much denser packing than one that supports only maximum travel to a station. In addition, we stress that coverage in this context is not merely in terms of the direct distance (cost function) between home and the service location, but rather in terms of an additional distance within the tours of existing activities for any given day, i.e., detours.

For the routing part of standard LRPs, the Traveling Salesman Problem or its variations are often used. For the hydrogen refueling stations case, however, a completely different routing structure needs to be used. Assuming that the refueling activity of drivers tends to tour with other activities when visiting the refueling service location, the routing part of this problem needs to describe how each customer visits the service location within the confines of their daily schedules and within the constraints imposed by the touring of other multiple activities. We use the Household Activity Pattern Problem (Recker, 1995) as a tool that optimizes personal and household travel behaviors within the context of a Vehicle Routing Problem (VRP). This routing structure of HAPP is the reason we emphasize the problem as an LRP rather than an extension from the point- or path-based models to a tour-based model. A tour comprised of trips is actually an output of a decision-making scheme; specifically, travel decisions determining the sequence of activities (trips), departure times, and locations for certain activity types – including refueling in the application considered in this problem. It may be argued that, in a conventional setting (e.g., ICEV gasoline refueling), a vehicle refueling activity either may or need not influence the sequence of trips and departure times; however, that it has significant impact on travel behavior when fuel availability is limited has been verified (Kitamura and Sperling, 1987).

In contrast to LRPs, the Location-HAPP formulation features multiple decision makers in the problem: a public/private agency that makes facility location decisions, and a collection of individuals that make their own independent daily travel decisions but with consideration of the refueling locations common to all. Although this may well be formulated as a bi-level problem, we instead have treated it as a single problem by conveniently parameterizing the individual travel problem utilizing the structure of the set covering problem shown in the next section. Additional notable differences are: 1) we are locating service locations that need to be visited by multiple customers, 2) the "depots" are customers' home locations in the routing problem, and 3) activity locations that are neither home (depot) nor service locations are visited as well. The inclusion of locations that are not directly connected to the facilities that are being located allows each household's location selection choice for the service type to be an output of interactions among other activities and schedules included in the analysis. These differences between the classical LRP models and what we propose in this chapter are conceptually depicted in Figure 4-2.



(a) A single-stage Location-Routing Problem



(b) A single-stage Location-HAPP problem

Figure 4-2 Conceptual diagram of classical location-routing problem vs location-HAPP problem

Problem Formulation

Consider a set of $|\mathbf{H}|$ households, $\mathbf{H} = \{h_1, h_2, ..., h, ..., |\mathbf{H}|\}$, in possession of one or more HFCVs, each of which on a particular day has an agenda, $\mathbf{A}_h = \{1, 2, ..., n_A^h\}$, comprised of n_A^h routine out-of-home activities with specific activity locations scheduled for completion, together with a need to refuel/service an HFCV at any one of a set of n_R candidate locations, $\mathbf{A}_R = \{1, 2, ..., n_R\}$—for example, in the case of hydrogen refueling stations, candidate locations $\mathbf{A}_R$ might logically be current gasoline stations. The problem is to find the optimal locations of the hydrogen refueling stations both with respect to the supply side costs of the refueling stations (via the objective function of minimizing the cost) as well as with respect to the corresponding demand side optimal activity patterns of the households (via constraints that guarantee that a certain level of accessibility is ensured for everyone in the specified area).

Define $Z_j, j \in \mathbf{A}_R$ as the binary locational decision variable of service type R, and $C_j, j \in \mathbf{A}_R$ as the corresponding stationary cost associated with operating the service location j. We assume that each household's , $h \in \mathbf{H}$, travel decisions are made so as to minimize the travel disutility subject to temporal and spatial constraints specified by LSP-HAPP developed in chapter 2. LSP-HAPP extended the original model (Recker, 1995) to include the capability of selecting one location for one or more activity type(s) from many candidate locations. In the case of hydrogen refueling, we specify that one and only one candidate locations for service type R needs to be visited. Then, each household has the following form of minimizing the total disutility:

$$\min O^h = \text{Travel Disutility of Household } h = f(\mathbf{X}^h, \mathbf{T}^h)$$

s.t.

$$\mathbf{A}^h \begin{bmatrix} \mathbf{X}^h \\ \mathbf{T}^h \\ \mathbf{Y}^h \end{bmatrix} \leq = \mathbf{b}^h \; ; \tag{4-1}$$

$$\mathbf{X}^h = \left[X_{u,w}^{v,h}, u, w \in \mathbf{N}_h, v \in \mathbf{V}_h \right], \; \mathbf{T}^h = \left[T_u^h, u \in \mathbf{P}_h^A \cup \mathbf{P}_h^R \right], \mathbf{Y}^h = \left[Y_u^h, u \in \mathbf{P}_h^A \cup \mathbf{P}_h^R \right]$$

where O^h is the travel disutility associated with the travel pattern adopted by household h, $\mathbf{N}_h$ is the set of all nodes associated with household h, (including those associated with refueling), $X_{u,w}^{v,h}$ is a binary decision variable equal to unity if vehicle v of household h travels from activity u to activity w, and zero otherwise, T_u^h is the time at which participation in activity u of household h begins, Y_u^h is the total accumulation of either sojourns[14] or time spent away from home on any tour, of household h on a particular tour immediately following completion of activity u, $\mathbf{V}_h$ is the set of vehicles available to the household (including one or more HFCVs), $\mathbf{P}_h^{A+}$ is the set of activities with predetermined locations, $\mathbf{P}_h^{R+}$ is the set of potential refueling activities (each with specific location common to all households), only one of which is to be completed, and $\mathbf{A}^h$ is a matrix of spatial, temporal constants as well as the tour length limit. (The details of Equation (4-1) are presented in the Appendix.).

Unlike in most general Vehicle Routing Problem (VRP) applications in which the objective function is a well-defined quantity (e.g., minimize total cost), individual travel behavioral problems typically specify the goal of individual choice in terms of utility maximization principles—or, more commonly, in terms of minimizing "travel disutility." The

[14] We have used the total accumulation of sojourns, and a maximum capacity of 4 (D = 4).

term “travel disutility” is a personal measurement that is assumed to be comprised of many observable (e.g., travel time, waiting time, total time spent outside home, and etc.) and unobservable (e.g., preferences, routing choice, spontaneous activity participation decisions, and etc.) factors. Even for choices involving a single trip, it is not straightforward to define what each individual is trying to minimize, and this has been the subject of numerous studies on estimation of such random utility models. For choices involving the interactions among activity/travel decisions over the course of some time period, say a day, the estimation problem is vastly more complex (see, e.g., Recker, 2001). Recker et al. (2008) and Chow and Recker (2012) focused on identifying the weights of the objective function of the HAPP mathematical program formulation through estimation processes based on genetic algorithms and inverse optimization, respectively. It was found in Chow and Recker (2012) that individual household’s travel disutility can be adequately represented as the weighted linear combination of the total extent of the day, the travel times, and the delay of return home caused by trip chaining multiple out-of-home activities. In Recker et al. (2008), the objective function is defined as the linear combination of eight different similarity measurements. One big difference of HAPP objective function compared to VRPs is that it includes temporal element (determined from the time variables) additional to cost (determined by the spatial path variables) element, which makes the computation more difficult as seen in Chapter 2.

In this study, however, in order to keep the computation simple, we assume that the individual travel disutility is measured solely by the cumulated travel time: is specified by the simple linear relationship.

$$f(\mathbf{X}^h) = \sum_{v \in \mathbf{V}_h} \sum_{u \in \mathbf{N}_h} \sum_{w \in \mathbf{N}_h} t_{u,w}^h \cdot X_{u,w}^{v,h} \qquad (4\text{-}2)$$

where $t_{u,w}^{h}$ denotes the travel time from the location of activity u to the location of activity w of household h.

The objective function, (4-1) is to minimize the sum of travel disutility of all households. However, because the objective functions and constraints are completely separable into each household problem since they do not share any of the variables, parameters or constraints across different households.

The refueling service provider's objective is to minimize infrastructure setup costs. Adding this objective and constraints, the Covering Location-HAPP formulation is as follows.

$$\min Z = \sum_{j \in \mathbf{A}_R} C_j \cdot Z_j \qquad (4\text{-}3)$$

Subject to

$$X_{u,j}^{v,h} \le Z_j, \quad u \in \mathbf{N}_h, j \in \mathbf{A}_R, v \in \mathbf{V}_h, h \in \mathbf{H} \qquad (4\text{-}4)$$

$$\sum_{v \in \mathbf{V}_h} \sum_{u \in \mathbf{N}_h} \sum_{w \in \mathbf{N}_h} t_{u,w}^{h} \cdot X_{u,w}^{v,h} \le O_{\min}^{h} + L, \quad h \in \mathbf{H} \qquad (4\text{-}5)$$

and conditions (4-A2) – (4-A20), $h \in \mathbf{H}$, contained in the Appendix C.

Where,

$O_{\min}^{h}$:The travel disutility without visiting the service location of type R. This value can be obtained either by solving (4-A1) subject to (4-A2), (4-A4) – (4-A20), independently from the proposed problem, or the current value from existing survey data can be used.

L : The maximum tolerable inconvenience, or the minimum level of service for service type R

The objective of the Location-HAPP to minimize the cost of providing the infrastructure available to everyone in the region, $\sum_{j \in \mathbf{A}_R} C_j \cdot Z_j$. Equations (4-4) constrain each visit to a service location can only be made at a location that provides refueling service. These are the linking constraints between the master problem of locating service centers and the sub-problem of household activity and routing decisions. Conditions (4-A2)-(4-A20), together with conditions (4-3) and (4-4) describe the scheduling and routing possibilities for each household's activities[15]. In the original HAPP and typical location-routing problems, the objective minimizes routing cost as well, but for Location-HAPP, optimal routing is not necessarily of the model's interest. Rather, identifying whether obtaining the feasible region as constrained by condition (4-5) is possible or not is the key. Conditions (4-5) guarantee the minimum level of service to every vehicle/person or household in the analysis area. In (4-5), the inconvenience is specified in terms of travel times, $t^h_{u,w}$, but other travel disutility functions are equally substitutable.

(4-A2)-(4-A20) constrain the daily movement of a personal vehicle to perform a given set of out-of-home activities at the given locations and a refueling stop at one of the refueling locations. The assumption is that the vehicle is replaced with an HFCV, but all individuals do not change their participation of the daily activities that they performed with an ICEV and a refueling trip per day. Because of the general unavailability of travel diaries over sufficient time periods to capture activities of the likely frequency of refueling (say, once in several days), we make certain assumptions regarding the refueling activity. In this analysis we insert one refueling trip per vehicle, under the assumption that the daily travel diaries reported in the survey represent a "typical" travel/activity day for the respondent. This, of course, in no way is meant to imply

[15] When two or more activities – both refueling and compulsory – are physically at the same location, each activity is labeled separately. This case, the optimization problem is highly likely to output the back-to-back scheduling of these two activities since the travel time between these two activities is 0 or a very small nominal number.

that such a refueling trip/activity occurs on every day, but rather that on the day that the refueling does occur the activity schedule of the individual is that reported in the survey. We note that, as an alternative synthetic activity/travel pattern generation may be utilized as in Xi et al. (2012), however this approach is not taken in this dissertation because of the unreliability of such generations in forecasting behavior (e.g., refueling at sparsely populated locations) for which there is no empirical base. The focus here is on the determination of inconvenience of limited refueling opportunities within empirical daily routines. However, we note that information on fuel inventory can be easily included in the base formulation described above without changing its basic structure simply by adding the additional set of constraints in Appendix 2.

While the formulation, (4-3) – (4-5), (4-A2) – (4-A20), represents the development from the travel behavior modeling perspective, the structure of the set covering model can be utilized for the actual computation. By introducing the binary coverage parameters, a_{hj}, the Location-HAPP can also be written as the standard covering problem formulation as follows:

$$\min Z = \sum_{j \in \mathbf{A}_R} C_j \cdot Z_j \tag{4-3}$$

s.t

$$\sum_{j \in A_R} a_{hj} \cdot Z_j \geq 1, \quad h \in H \tag{4-6}$$

Where $a_{hj} = \begin{cases} 1 & \text{if household } h \text{ is within the tolerable service level from refueling location } j \\ 0 & \text{otherwise} \end{cases}$

This alternative formulation specified by (4-3), (4-6) separates each household's travel decision from each other as well as the master facility location problem. It is noted that the first formulation, (4-3) – (4-5), (4-A2) – (4-A20), and the second formulation, (4-3), (4-6), produce the same results.

Computationally, the Location-HAPP Problem differs from the Location-Routing Problems in the following ways. First, in the routing sub-problem, each activity point (except refueling locations) is already assigned to households while in the classical LRPs' sub-problem each customer location can be visited by any vehicle from any depot. Second, the facility node is not the depot, but an intermediate location that each vehicle needs to traverse once and only once. These properties make the computational complexity of Location-HAPP much less intensive since the numbers of nodes each vehicle visits are rather limited, and the decision of choosing one of the open service locations is much easier than the decision of allocation of all nodes to depots and vehicles.

The computational burden of HAPP is usually not as great as that of general VRPs since there are empirical limits to the number of household members and the number of activities performed. However, time windows are not as constraining for some activities compared to PDPTW—especially the return home time windows—which leads to higher computational cost, as shown in Desrosiers et al. (1984) and Desrochers et al. (1991). Moreover, efficient solution methods have not yet been developed for more complicated cases that include interactions among household members as in HAPP formulations 4 and 5 (Recker, 1995), or that include location selection as in Chapter 2. Possible solution methods include using heuristics (Chow and Liu, 2012) or employing route length estimators (Beardwood et al., 1959; Bruns et al., 2000) with bounded region of one's mobility.

4.3 SOLUTION ALGORITHM

With (4-6), each vehicle's routing problem has been completely isolated from those of other vehicles, as well as from the master problem. However, it requires $|\mathbf{H}| \times n_R$ number of HAPP problems in which each visits one of specified candidate locations that are consistent with the parameter matrix, a_{hj}. This parameter matrix can be calculated from solving (4-A1)-(4-A20) with a condition that the household visits refueling location j. Finding this matrix becomes an issue in the Location-HAPP Problem—calculating $|\mathbf{H}| \times n_R$ sub problems that are NP-hard[16] makes the computational set-up for the master problem important.

The Set Covering and Set Partitioning problems are well-known binary combinatorial problems, and there exist many algorithms to handle the associated computational complexity (Caprara et al., 2000; Christofides and Korman, 1975). For the problem addressed here, the key is to have the least number of actual a_{hj} parameters to be added in the master problem of the Location – HAPP Problem. Column Generation is a well-known technique that does not require all variables and parameters to be in the master problem, which greatly reduces the number of a_{hj} parameters actually needed to solve the problem. Instead, column generation solves a sub problem of finding a variable that would reduce the objective function value of the master problem. This procedure is found to be efficient for some types of large combinatorial integer problems (Barnhart et al., 1998), including the Set Covering problem. Moreover, the fact that initialization is not difficult and that linear relaxation of the master problem is stable makes column generation suitable for application to the problem considered here.

[16] However, there are many methodologies to handle HAPP or its original form of PDPTW.

For standard column generation, the sub problem is another optimization problem that finds the most negative reduced cost. However, this sub problem requires the full information of the coverage matrix, which we wish to avoid. Instead, we deploy a new search procedure that finds an entering column with a negative reduced cost, not necessarily a minimum. This keeps the structure of column generation without requiring full knowledge of the coverage parameters.

The overall iterative procedure we propose for the Location-HAPP is shown in Figure 3. After initialization, the master problem is solved. It then passes the dual values to the sub-problem that finds the next entering station location variable, j, with a negative reduced cost by our search algorithm. If no dual variable exists with negative cost, or no stations that are not in the master problem can deliver a negative dual cost, we conclude that it is optimual. Otherwise, the entering variable and its coverage column are added to the master problem. Details of initialization and the third box are changed from the standard procedure of Column Generation. π_h is the dual value associated with the coverage constraint of household h. $\overline{C_j}$ is the marginal cost of locating a station at candidate node j.

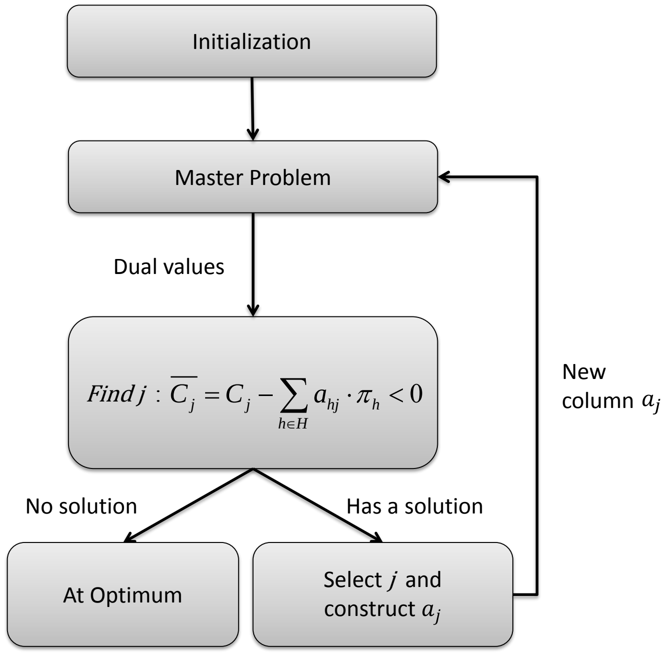


Figure 4-3 Schematic of iterative procedure

For the initial setup of the problem, it is first assumed that each household is covered by the refueling station closest to the home location (home is always one of the nodes in the household's pattern). If that station is not within the tolerance level, the next closest station can be tested. Although it is likely that the household is covered by a station closer to home location, this proximity rule may further be developed to be more sophisticated, and refined to be appropriate for routing considerations; for example, by the order of the distance to any of its activity nodes. Once it is decided that station j is able to cover household h, the remaining a_{hj} parameters for all households are assumed to be 0 for station j, regardless of whether or not station j covers household h within the tolerance level. This is purely for the computational convenience that there is no practical advantage to constructing and verifying all parameters, and that the number of HAPPs to solve should be kept to the lowest possible. This way, we reduce the initialization cost of the problem. For example, assume a case of 5 households and 3 refueling candidate locations, with full coverage matrix shown in (M1):

$$\begin{array}{c} \\ HH1 \\ HH2 \\ HH3 \\ HH4 \\ HH5 \end{array} \begin{array}{c} \begin{array}{ccc} c1 & c2 & c3 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \end{array} \tag{M1}$$

If it is the case that the closest refueling candidate station to households 1, 2, and 3 is station candidate 1, and the closest refueling candidate station to households 4 and 5 is station candidate 2—all within the tolerance level—then the initial parameter matrix is:

$$\begin{array}{c} \\ HH1 \\ HH2 \\ HH3 \\ HH4 \\ HH5 \end{array} \begin{array}{c} \begin{array}{cc} c1 & c2 \end{array} \\ \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \end{array} \tag{M2}$$

Thus, we do not check for the coverage parameters, $a_{HH4,c1}, a_{HH5,c1}, a_{HH1,c2}, a_{HH2,c2}, a_{HH3,c2}$, and eliminate the need of solving those 5 HAPPs.

We also note, however, that degeneracy occurs for cases in which all actual solutions are a subset of the initial (synthetic) solution and the real coverage of those stations is superior to others not in the initial master problem. In such cases, at any iteration stations already in the master problem cannot be the entering variable, and therefore the initial synthetic coverage columns cannot be replaced with real ones, which are superior to the synthetic ones. For example, assume that the algorithm initialized using (M2), and the original full coverage matrix would have been (M3):

$$\begin{array}{c} \\ HH1 \\ HH2 \\ HH3 \\ HH4 \\ HH5 \end{array} \begin{array}{c} \begin{array}{ccc} c1 & c2 & c3 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \end{array} \qquad \text{(M3)}$$

The optimal solution is building one station at site c1, which covers all households. Because our initial coverage matrix is (M2), c3 is the only column to be considered for bringing into the problem. However, the master problem would decide that c1, c2 are the solution, which will not allow a chance for real columns of those stations to be entered. This issue can be avoided simply by constructing the actual coverage matrix at initialization, but this would require calculating $|\mathbf{H}| \times (\text{number of initial columns})$ different HAPP models to be solved.

Following the standard Column Generation procedure, the sub-problem finds one station's coverage column with the most negative reduced cost to be added to the master problem, and then the master problem is solved to find the best combinations of stations to be built. As mentioned earlier, finding a station with the most negative reduced cost, full knowledge of the coverage matrix, $\mathbf{A}$ is needed. On the other hand, the sub-problem does not necessarily have to be an optimization problem since it is going to lower the objective function value in the master problem as long as the entering variable has a negative reduced cost. This may increase the number of iterations since we are not looking for the best variable, but it reduces the search cost of the column which presents the greatest difficulty in this particular problem. Therefore, the sub-problem becomes:

$$Find\ j \quad \text{s.t. } \overline{C_j} = C_j - \sum_{h \in \mathbf{H}} a_{hj} \cdot \pi_h < 0 \qquad \text{(4-a)}$$

Where,

$\overline{C_j}$: the marginal cost of locating a station at candidate node j

π_h : the dual variables associated with the coverage constraint of household h, $h \in \mathbf{H}$

The search method employed here to find station j with a reduced cost is as follows. First, there is no need to search for coverage parameters for constraints with zero dual values, $\pi_h = 0$; only those of positive dual values are generated by solving HAPP. Therefore, we only go over households with positive dual values, and update the cumulative term, $\sum_h a_{hj} \cdot \pi_h$ as the search progresses. And, as soon as $\sum_h a_{hj} \cdot \pi_h > C_j$ has been established, there is no need to keep searching since $a_{hj} \cdot \pi_h$ is non-negative for any household constraint, so it can be concluded that j would be included in the master problem. Once it has been decided that j is to be the entering variable, all column parameters of a_j need to be generated. The most efficient order of searching is to check the parameters from the highest π_h to the lowest, for instance, using a priority queue data structure. In the example provided in the next section, a uniform unit cost of 1 is used; so, all dual variable values are either 1 or 0. If $\sum_h a_{hj} \cdot \pi_h \le C_j$, even after all household coverage constraints have positive dual values, that variable is discarded and a new station is tested.

When choosing a station j to test, one random household with positive dual values is selected. Then, a priority queue that stores all candidate locations in increasing order of the direct distance to that household is created. (We note, however, that this procedure can be developed to be more sophisticated; for example, by the order of the distance to any of activity nodes for all households with positive dual values.)

With the following notations defined for each iteration, i, this search method can be summarized as the following:

$\mathbf{Z}^i$: the set of stations that are in the solution from the master problem of previous iteration. This is the subset of $\mathbf{A}_R$

$\mathbf{M}^i$: the set of households with coverage constraints of dual values from previous iteration that are greater than 0.

$\mathbf{I}_j^i$: the subset of $\mathbf{M}^i$ of which the coverage parameter a_{hj} has been checked for the possibility of station j being the next variable. If $I^i = \mathbf{M}^i$, all components of $\mathbf{M}^i$ have been checked.

If $\mathbf{M}^i = \emptyset$, it is at optimum.

Else,

For all $j \in \overline{\mathbf{Z}^i}$ (loop 1)

For all $h \in \mathbf{M}^i$, (loop 2)

If, $a_{hj} = 1$, update $\sum'_{h \in \mathbf{I}_j^{\,i}} a_{hj} \cdot \pi_h \leftarrow \sum'_{h \in \mathbf{I}_j^{\,i}} a_{hj} \cdot \pi_h + a_{hj} \cdot \pi_h$, $\mathbf{I}_j^{\,i} \leftarrow \mathbf{I}_j^{\,i} \cup h$

If $\sum'_{h \in \mathbf{I}_j^{\,i}} a_{hj} \cdot \pi_h > C_j$, select j as the entering variable and break both (loop 1) and (loop 2)

Else if $\mathbf{M}^i \neq \mathbf{I}_j^{\,i}$, select a different $h \in \mathbf{M}^i$, $h \notin \mathbf{I}_j^{\,i}$ and continue (loop 2)

Else, Select a different $j \in \overline{\mathbf{Z}^i}$ and initiate a new (loop 2)

If $\sum'_{h \in \mathbf{I}_j^{\,i} = \mathbf{M}^i} a_{hj} \cdot \pi_h \leq C_j$ for all $j \in \overline{\mathbf{Z}^i}$, it is at optimum.

Algorithm 4-1 The sub-problem of finding an entering variable with a negative reduced cost

4.4 CASE STUDY

Southern California is anticipated to be one of the early adoption areas of HFCVs. It is the location of major auto manufacturers' US headquarters, and three target areas (Torrance, Santa Monica, Irvine and Newport Beach) in Southern California have been identified as early adoption communities (CaFCP, 2009; CaFCP, 2010). As of 2011, sixteen refueling stations are under operation and 40 more stations are planned, and the number of HFCVs deployed in this area is expected to be in the thousands by 2013, and tens of thousands by 2016 (CaFCP, 2012).

In the case study presented here, we focus on Irvine and Newport Beach as the study area (Figure 4-4). In the example, we presume that the candidate sites for future hydrogen refueling stations are drawn only from existing gasoline stations—there currently are 34 gasoline stations in the area. Existing or planned hydrogen stations are not considered in this analysis[17]. Further, when there are two or more stations at an intersection, they are considered as one for simplicity. These existing gasoline stations are the only candidate sites.

The analysis is based on a subset of household samples from the study region drawn from the California Statewide Household Travel Survey (CalTrans, 2001). The Travel Survey contains the household daily travel activities and their full location data. Each trip contains information on departure and arrival times, trip/activity durations, and geo-coded information on longitude/latitude of the activity locations. From the full set of 500 Orange County households in the dataset, we identified a suitable subset of 134 full-day vehicle patterns (households) with the requisite complete location information and vehicle types that can be substituted by HFCVs, excluding motorcycles, bicycles, etc. Based on these data, person-based trip chains and activity ordering are converted to equivalent vehicle-based chains in order to simulate vehicle routing patterns. We assume that this sample is representative of travel behavior in the county.

For each vehicle/household, a travel time matrix is generated associated with all possible combinations of $X_{u,w}^{v,h}$ for the out-of-home activities performed. In the formulation, for every household, the first 34 nodes, 1 – 34, are specified as candidate locations for refueling—therefore universal for the entire sample—while nodes labeled 35 and greater are specified as the nodes unique to each individual household's out-of-home activities that were reported as performed. All-to-all travel time matrices are generated by calling *MapQuest API*.

Our principal behavioral assumption is that, with replacement of households' current conventional vehicles by HFCVs, each new vehicle will continue to perform activities that were

[17] Two current hydrogen refueling stations in the area are: National Fuel Cell Research Center at University of California, Irvine and Shell (Newport Beach, CA) which is annexed to a gasoline station

reported in the survey, plus an additional refueling activity at one of the candidate locations. This vehicle usage model is equivalent to Case 1 in Recker (1995). For the construction of time windows, the method in Recker and Parimi (1999) is adopted. Work, school, meals between two work activities, and pick-up/drop-off activity types are constrained to have exact times as reported. The open/close windows of other activity types are specified as the minimum/maximum of respondent's reported activity start/end time and sample mean activity start/end time for the activity. For each individual, the vehicle/household start time window is the minimum of respondent's reported travel start time for his/her initial activity and mean reported travel start time for initial activity for the sample. In general, latest return to home of an individual vehicle/household is taken as the max of respondents' reported return-to-home time for his/her final activity and mean reported travel return-to-home time for final activity for the sample. However, when the sequence of activities performed by a household comprises only activities with exact start/end times, or in cases where the travel time generated from *MapQuest* is larger than the reported travel time, the last return home is relaxed since even the basic HAPP of reported activities becomes infeasible, due to an additional refueling trip will also need to be inserted somewhere at some time.

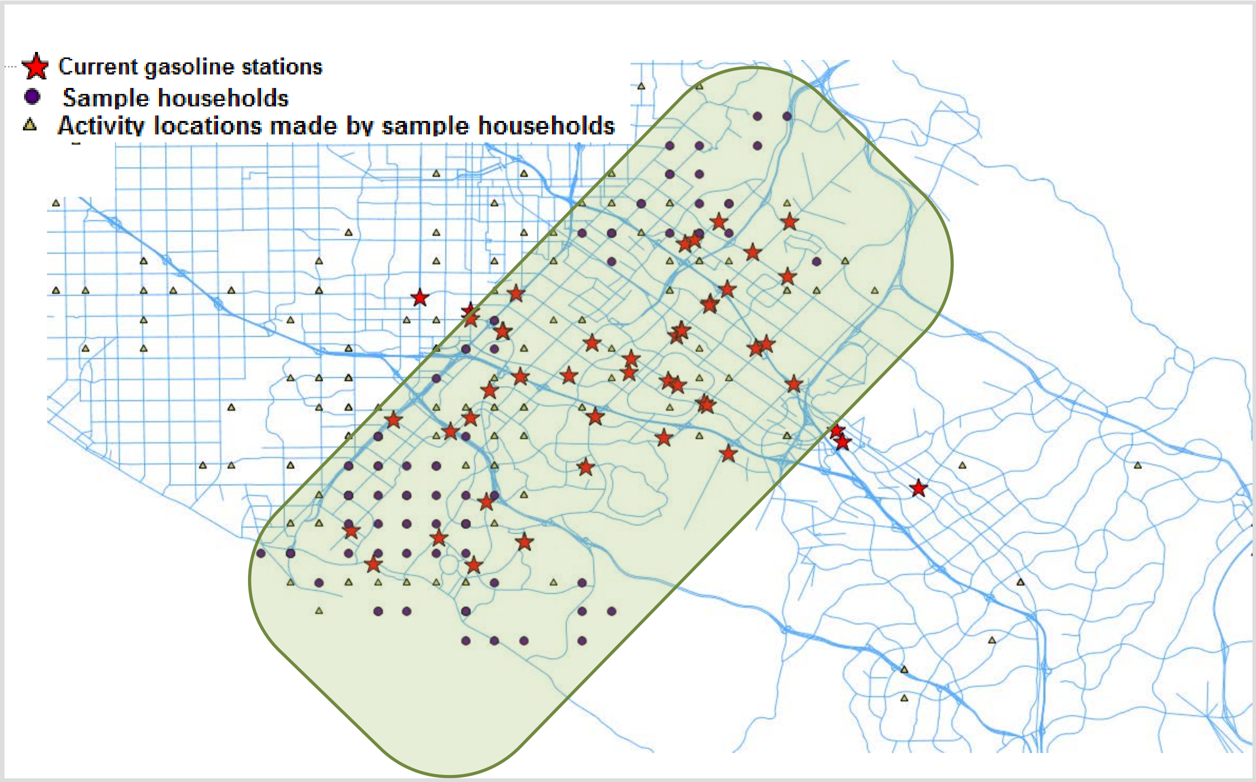


Figure 4-4 Study area: Irvine and Newport Beach in California with sample households

The objective function for each household is assumed to be minimizing the total travel time throughout the day; i.e.,

$$\min O_h = \sum_{u \in N_h} \sum_{w \in N_h} \sum_{v \in V_h} t_{u,w}^{h} \cdot X_{u,w}^{v,h} \;\; ; \;\; \forall h \in H \qquad (4\text{-}7)$$

When checking for a_{hj}, the comparison measurement of $O_{\min}^{h}$ is generated from HAPP without a visit to one of the refueling stations but with visits to all activities given by the survey responses for that household. When solving a HAPP, the exact dynamic program developed by Desrosiers et al. (1986) is used; however, direct calculation using an optimization tool, or any other suitable algorithm (e.g., branch and bound) would work as well (Cordeau and Laporte, 2003). Households in the data set not covered by any of the refueling candidate locations within the given tolerance, or households that would require violation of given constraints to visit any of the candidate locations, are omitted.

The detailed results for the base case involving 122 households with a refueling activity subject to the additional travel time tolerance of $L = 0.2hr$ are shown in Table 4-1. Optimality is reached after 7 iterations, solving 966 out of 4,148 HAPP cases, to find the a_{hj} parameters needed to perform the proposed search algorithm.

Table 4-1 Results of location-HAPP problem (L=0.2 hour)

	Initialization $i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$
Columns in MP	10	11	9	8	5	5	4	NA
Optimal Solution	10	8	7	4	4	3	3	NA
# of HAPPs solved [18]	122	2	2	2	6	4	3	93
# constraints with (-) dual	NA	9	9	7	6	3	2	2

[18] This is counting only HAPPs solved to find the next entering variable, or to find that it is at optimum. Once an entering variable is found, 122 additional HAPPs are solved to construct the parameter vector, a_j of the entering parameter column at each iteration.

Four different travel time tolerance level cases are tested and compared in Table 4-2. Cases 2, 3 and 4 are run using the proposed initialization method. For Case 1, one of the initial columns is found to be highly superior, causing degeneracy, and therefore it is necessary to construct the full initial columns. This increases the number of HAPP models that needed to be solved to 35% of total HAPPs. The best case we tested was case 4, for which only 7% of total number of possible HAPP computations were required.

The results of same levels of accessibility using the point-based Set Covering Problem are compared in Table 4-2. Here, the "Set" refers to "home" locations—that guarantees the same levels of access based on the households' home locations. The coverage matrices are significantly sparser for point-based Set Covering Problem, leading to a larger number of stations for a particular level of access. It is found that we cannot guarantee the maximum accessibility times of 0.4, 0.3 hour for all households' home locations, and it requires 5 and 2 stations to guarantee the maximum accessibility times of 0.2 and 0.15 hour. When the concept of accessibility is expanded out to tours, it is found that we can guarantee the maximum deviation times of 0.4, 0.3, 0.2, 0.15 hour with 5, 4, 3, 1 stations, respectively. From the results, it is argued that the point-based Set Covering Problem may significantly overestimate the number of stations required. Considering the cost of building and operating a station, it can be concluded that by including routing and scheduling considerations as a part of "accessibility," the overall cost of infrastructure supply can be lowered significantly as seen from the results of the Location-HAPP model.

Table 4-2 Results of Location-HAPP problem vs Point-Based Set Covering Problem

	Tolerance level	Case 1	Case 2	Case 3	Case 4

	(Guaranteed maximum deviation time of L)	($L = 0.15hr$)	($L = 0.2hr$)	($L = 0.3hr$)	($L = 0.4hr$)
	# Households in Analysis	116	122	124	124
LRP – HAPP	**Optimal Solution**	5	4	3	1
	% of coverage of the full parameter matrix[19], A	1,498 / 3,944 (37.98 %)	2,048/4,148 (49.37 %)	2,888/4,216 (68.50 %)	3,108/4,216 (74.93 %)
	# HAPPs solved[20]	1,385[21] / 3,944 (35.11 %)	966 / 4,148 (22.29 %)	1,077 / 4,148 (25.96 %)	285 / 4,148 (6.87 %)
Point-based Set Covering Problem	**Optimal Solution**	Infeasible	Infeasible	5	2
	% of coverage of the full parameter matrix[22], A	96 / 3,944 (2.43 %)	264 / 4,148 (6.36 %)	856 / 4,216 (20.30 %)	1768 / 4,216 (41.94 %)

Since we used uniform station construction cost, there are a number of different combinations of locations that satisfy the optimality conditions for this particular example. The results for one set of optimal refueling locations are shown in Figure 4-5 for four different additional travel time tolerances, together with a comparison to the solution to the point-based Set Covering Problem (for the case of 0.3 hr and 0.4 hr additional travel time) for this application. The results indicate that, based on the reported daily activity agendas (assumed repetitive) of households in the study area, Location-HAPP identified that full coverage within a

[19] This is calculated by solving all HAPPs for parameters a_{hj}

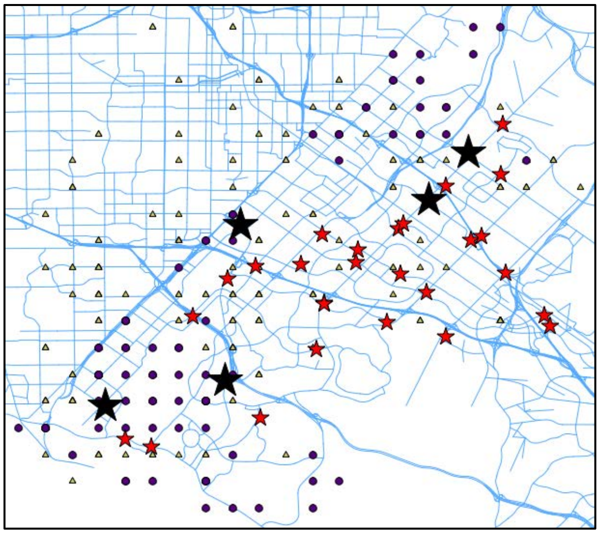
[20] This includes the construction of each $a_j = [a_{1j}, a_{2j}, ..., a_{hj}, ..., a_{|\mathbf{H}|j}]^T$ column vector at each iteration: total of $|\mathbf{H}|$ number of HAPPs. This also include thes number of HAPPs that are solved during the search method in order to decide whether column j can be the next entering column. Therefore, some of a_{hj} 's are run twice for both the search and the construction.

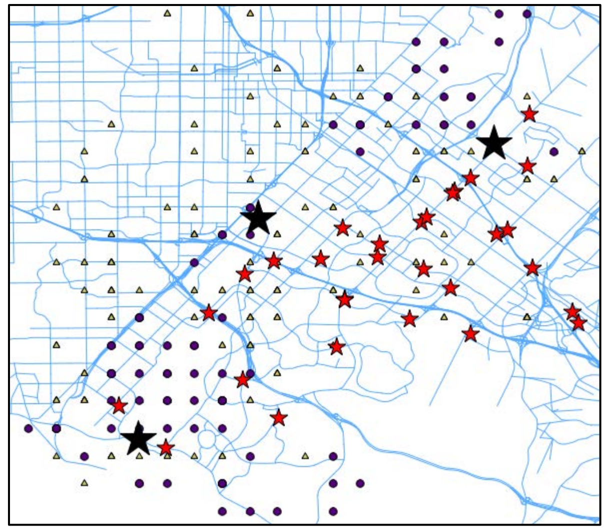
[21] Due to degeneracy, all initial column parameters a_{hj} are checked.

[22] This is based on direct distance from home. Accounting for a round trip, tolerance/2 is used measure the coverage.

0.4 hr additional travel time window could be achieved with a single refueling station—tolerances of 0.15 hr, 0.2 hr, and 0.3 hr could be achieved with 2, 3 and 5 stations, respectively. In contrast, ignoring these patterns of travel and using a traditional "point-based" Set Covering Model based only on the home locations of the residents, full coverage within a 0.4 hr travel time required 2 refueling stations and 0.3 hr required 5 refueling stations. There are no feasible solutions of the traditional point-based Set Covering Problem for tolerances of 0.15 hr and 0.2 hr. These results are consistent with previous studies comparing LRP results to results of more static approaches. In Salhi and Rand (1989), Nagy and Salhi (1996), Balakrishnan et al. (1987), the results of LRPs are compared to a location-first routing-second approach. The findings are consistent with previous works (Salhi and Rand, 1989; Nagy and Salhi, 1996) in that ignoring the routing structure for location strategy results in higher cost.

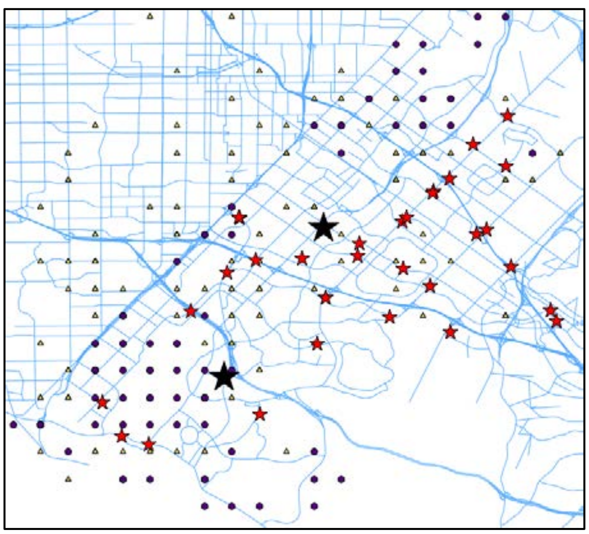
Another observation is that the Location-HAPP favors areas of high volumes of activities (such as central business district or shopping centers) whereas the point-based Set Covering Model favors the residential areas. By definition, we only include home locations to be covered in the point-based Set Covering Model whereas in the Location-HAPP, we expand the concept of "set" to the path connecting activity locations, giving the flexibility to cover any trip to any activity location. If there exist certain activity/travel patterns induced by land use, it is reasonable that the master problem in the Location –HAPP would favor a high activity volume area in a collective sense.

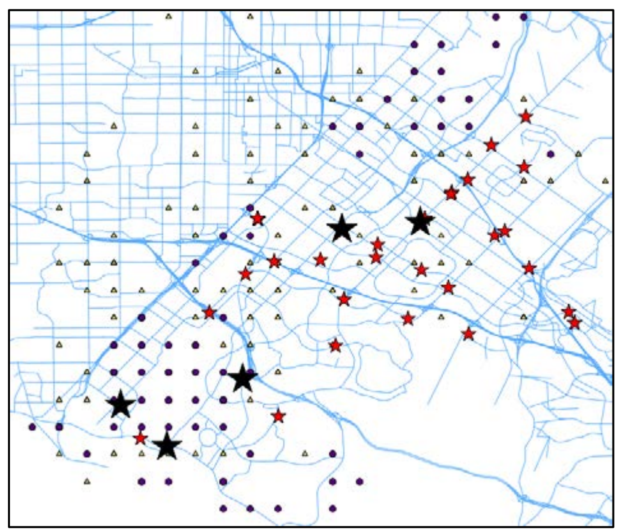




(a) Case 1 ($L = 0.15hr$)

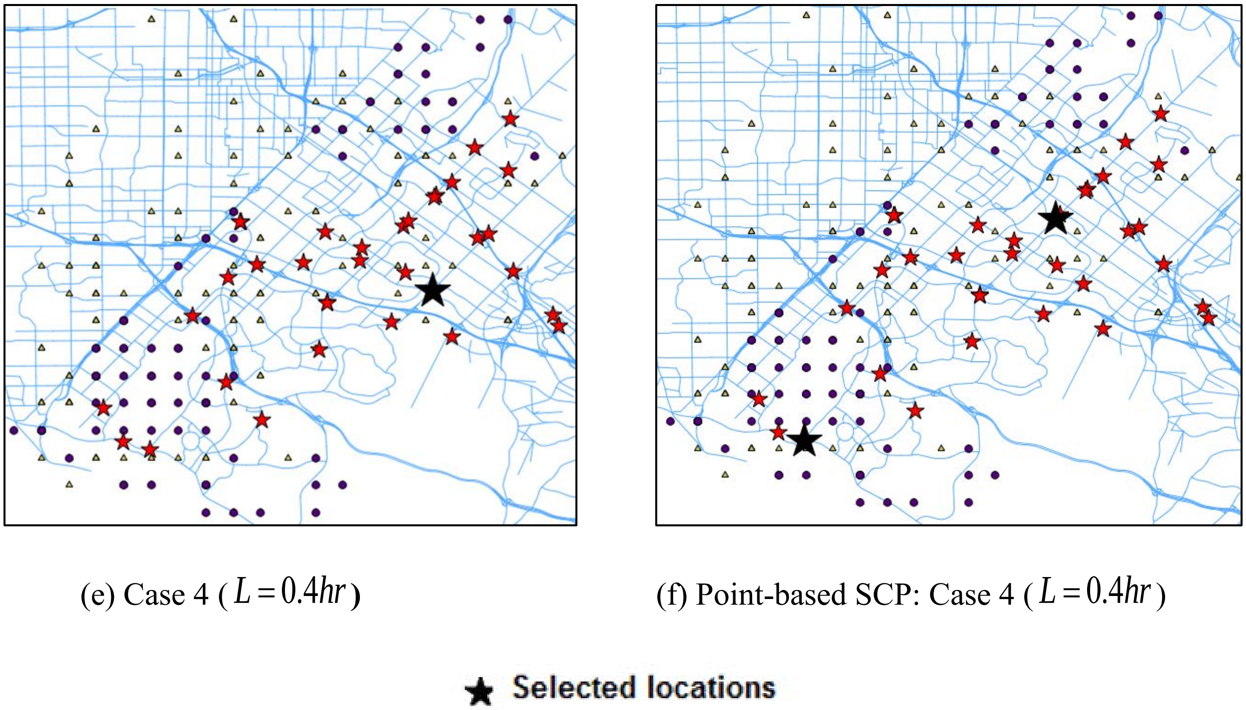
(b) Case 2 ($L = 0.2hr$)





(c) Case 3 ($L = 0.3hr$)

(d) Point-based SCP: Case 3 ($L = 0.3hr$)

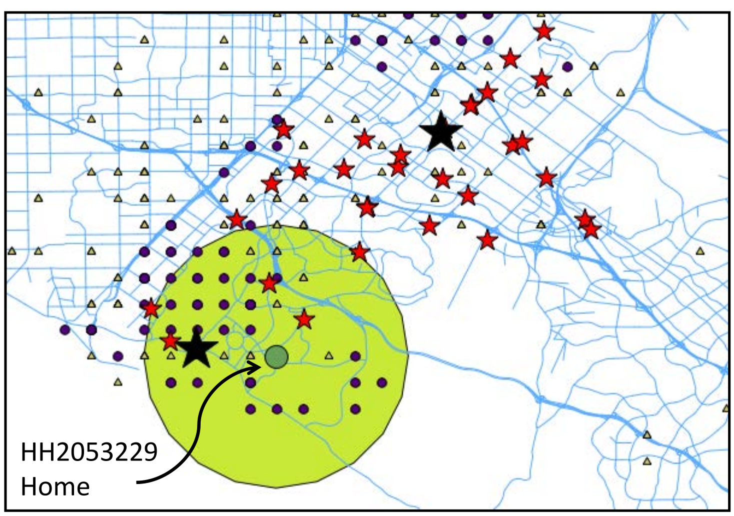


(e) Case 4 ($L = 0.4hr$) (f) Point-based SCP: Case 4 ($L = 0.4hr$)

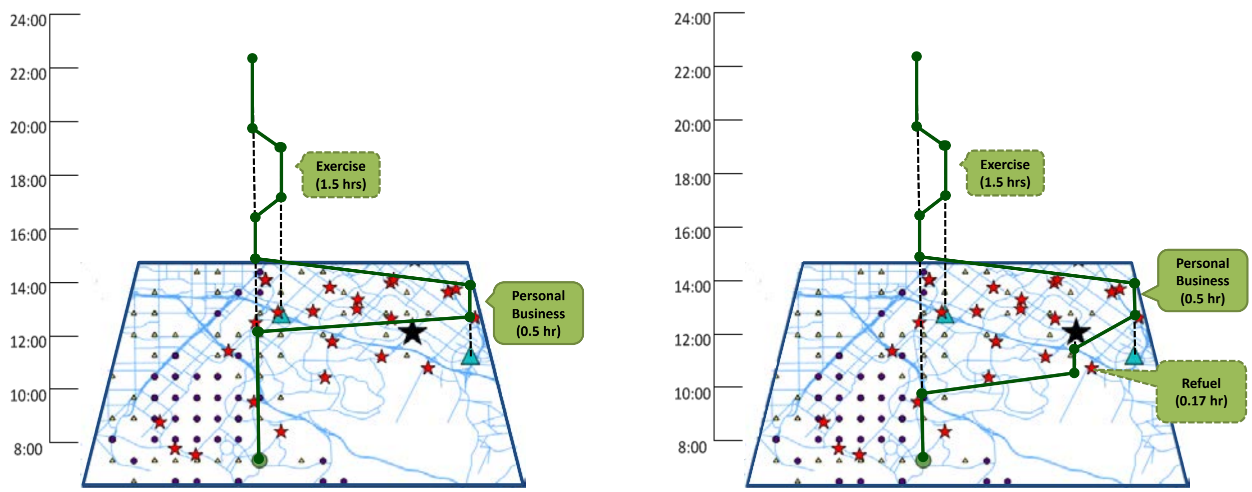
Figure 4-5 Location-HAPP and Point-Based Set Covering Problem results

We can also conceptualize the accessibility in the Location-HAPP with an example. Because the point-based Set Covering Model focuses only on the accessibility of given points (for our case, home locations), the accessibility range is limited to the home location, as seen in Figure 6a, for the case of household #253229 in the Travel Survey, under the presumption that refueling will take place at the station located within the shaded area. In the Location-HAPP model, accessibility is not limited to home but includes accessibility along the path over the space throughout the day, expanding the area of coverage, albeit subject to the temporal constraints imposed by the available time windows for completion of each household's activity agenda. This also presumes grouping/touring refueling with other compulsory activities is an acceptable alternative. Although there appears to be evidence that drivers' tend to prefer to refuel "near home" (Kitamura and Sperling, 1987), and some studies assume this preference

(Nicholas and Ogden, 2006; Nicholas et al., 2004), refueling activity between origins and destinations have been observed at the early adoption stage where there are limited refueling opportunities (Kelly and Kuby, 2012). For household #253229, the reported travel pattern is that shown in Figure 4-6b. The optimal (minimization of travel time) travel/activity pattern for this particular household obtained from HAPP is shown in Figure 4-6b. Since one of the compulsory activities is near one of the refueling candidate locations, this household completes the refueling activity on the way to "personal business" activity at the refueling station that is located close to that "personal business" activity location (Figure 4-6c), rather than at the station closest to home. Were the reported activity agenda for this household on the day of the survey repetitive (representative of this particular household's daily routine), the service provider's need to locate a facility near its home location could been substituted by a refueling location that would have been not considered as accessible by the point-based Set Covering Model constraints.



(a) Accessibility of HH 2053229 within 0.4 hr Using Direct Distance Measurement



(b) Reported Travel Pattern (c) Travel Pattern with an added Refueling Trip

Figure 4-6 Refueling pattern of Household 2053229 (Case 4 = 0.4 hr)

4.5 CONCLUSION AND FUTURE RESEARCH

In this chapter, we developed a facility location problem with full-day scheduling and routing considerations. The model is a type of Location-Routing Problem (LRP), where the decisions of facility location models are influenced by possible vehicle routings. The model we propose takes the classic coverage model as a location strategy, and the Household Activity Pattern Problem (HAPP) as the scheduling and routing tool. The Location-HAPP includes multiple decision makers: the public/private sector provides the refueling service, and each individual makes his/her own routing decisions, including a visit to one of the service locations. It is a LRP that addresses public/private sector's refueling service provision using a non-standard structure that uses an exact solution method via decomposition. Following the LRP categories from Nagy and Salhi (2007), this problem has: (a) hierarchical structure, (b) deterministic data input, (c) static period planning, (d) exact solution method, (e) objective function of cost minimization where cost is facility installation cost, (f) discrete solution space, (g) multiple depots (h) heterogeneous vehicles, and (i) route structure of HAPP. It also is an extension from node-based and path-based models to a tour-based model—to be exact, a special case of a tour-based model in a way that allows changes in the sequence within the specified time windows.

Although here we apply the Location-HAPP Problem to refueling station siting, the models can be extended to other applications that require coverage and in which customers (drivers) can travel to within a tour of other activities (e.g., access to such basic public services as post office, public health care). Alternatively, for services that are visited directly to demand points from supply points, or that are considered to be primary (e.g., emergency service, education service), the point-based Set Covering Model can be sufficiently effective.

The proposed formulation isolates each vehicle's routing problem from those of other vehicles as well as the master set covering problem. However, its coverage matrix requires the solving of $|H| \times n_R$ (number of households x number of candidate locations) HAPPs. A modified column generation that finds a column with a negative reduced price, but not necessarily the most negative, is developed. This way, only partial knowledge of the full coverage matrix is needed. A search method is developed for finding such columns in order to reduce the number of HAPPs to be solved. The performance of the methodology is described by the percentage of HAPPs that are actually solved to the total number of HAPPs in the full coverage matrix.

Although general to the location problem of any service facility that can be considered as ancillary to the spatio-temporal movement of households as they complete their daily routines, the specific application developed here relates to the incubation of the minimum refueling infrastructure that might be required to support early adoption of HFCVs. The proposed model and methodology are applied to a case study of HFCV refueling stations in Irvine/Newport Beach community—one of four early hydrogen vehicle adoption communities targeted by auto manufacturers—subject to three different values of accessibility measured in terms of tolerances to added travel time. Under optimal conditions, refueling trips are found to be mostly toured with other activities, and this traveling behavior is captured by HAPP. We have tested four different levels of service provisions, and the suggested method shows that only 6% – 35% of sub problem are need to be solved, compared to direct calculation. More importantly, from the results and the coverage matrices, there is evidence that station location approaches that do not allow such vehicle-infrastructure interactions as well as routing and scheduling interactions can result in over-estimation of the minimum number of facilities.

This model makes a simplification of imposing one refueling trip per day per a vehicle mainly due to limitations imposed by the scarcity of travel diary data covering more than a single day. On the other hand, we are trying to define the individual inconvenience level with respect to spatial, temporal constraints generated by participating in daily compulsory activities. In that context, it serves the goal to insert one additional refueling trip to be performed by every individual. More generally, however, data limitation has been one of the major issues for conducting activity-based analyses. This may be overcome using simulations to synthesize activity/travel patterns – possibly for multi-days – as they are developed more fully. For refueling studies, it will certainly help in aggregating results to a whole population, as well as multi-day analysis of the refueling needs based on fuel inventory. This way, queuing (waiting time) at the refueling service locations can be evaluated. Another shortcoming relative to travel behavior is that we have ignored the intra-household interactions of travel decisions. It is expected that if an AFV is one of the vehicle sets in a household, some of the activities may be shifted to other vehicles depending on the length, property, and exchangeability of activities; such an extension could logically be based on HAPP cases 2 – 5.

In addition to the improvements to the activity-based analyses mentioned above, this work can further be extended to more sophisticated facility location strategies. For example, the time frame can be extended to multi-period as in Laporte and Dejax (1989) which would ostensibly yield more realistic investment plans since the hydrogen infrastructure is likely to be implemented over a long period of time (CaFCP, 2009, 2010, 2012). Station's fuel capacity limitations can also be added following inventory considerations (Liu and Lee, 2003). Finally, an extension of LRP with nonlinear costs (Melachovsky et al., 2005) can be used to address nonlinear queuing functions frequently associated with location-queuing problems (Berman et al., 2007; Berman and Drezner, 2007).

CHAPTER 5 NDP-HAPP

As shown in Chapter 4, activity-based travel demand models can be integrated into the network problems in order to capture the demand changes at the time of decision making. In this Chapter, we generalize such integration into a framework in a bi-level structure. The network problem is generalized as a Network Design Problem (NDP) as well. Although roadway NDPs focus on congestion effect from the traffic assignment, in this dissertation, we focus on the design of network system with the effects of demand changes.

This chapter examines network design where OD demand is not known *a priori*, but is the subject of responses in household or user itinerary choices to infrastructure improvements. Using simple examples, we show that falsely assuming that household itineraries are not elastic can result in a lack in understanding of certain phenomena; e.g., increasing traffic even without increasing economic activity due to relaxing of space-time prism constraints, or worsening of utility despite infrastructure investments in cases where household objectives may conflict. An activity-based network design problem is proposed using the location routing problem (LRP) as inspiration. The bilevel formulation includes an upper level network design and shortest path problem while the lower level includes a set of disaggregate household itinerary optimization problems, posed as household activity pattern problem (HAPP) (or in the case with location choice, as generalized HAPP) models. As a bilevel problem with an NP-hard lower level problem, there is no algorithm for solving the model exactly. Simple numerical examples show optimality gaps of as much as 5% for a decomposition heuristic algorithm derived from the LRP. A large numerical case study based on Southern California data and setting suggest that even if infrastructure investments do not result in major changes in link investment decisions the results provide much higher resolution temporal OD information to a decision maker. Whereas a conventional model would output the best set of links to invest given an assumed OD matrix, the proposed model can output the same best set of links, the same daily OD matrix, and a detailed temporal distribution of activity participation and travel from which changes in peak period OD patterns can be observed.

5.1 BACKGROUND

Network design problems (NDPs) are a class of optimization models related to strategic or tactical planning of resources to manage a network (Magnanti and Wong, 1984). Even for purposes of improving road networks for commuters (Yang and Bell, 1998) and despite the complexity of traveler choices (Recker, 2001), NDPs generally assume either static demand at a node (elastic or not) or trip-based origin-destination demand. While this assumption is sufficient in many applications, there is increasing recognition that explicit consideration of travelers' schedules, choices, and temporal decision factors is needed. This need has grown in parallel to three related research trends in network design in the past few years: (operational) network design with dynamic assignment considerations when considering only peak period effects, (tactical) service network design with schedule-based demand under longer periods of activity, and (planning) facility location problems that explicitly consider the effects that they have on decisions related to routing and scheduling of vehicles. At the planning level, these NDPs have often been based on private firm decisions, rather than on household-based urban transportation planning considerations.

The rationale behind dynamic network design problems is rooted in bi-level NDPs that feature congestion effects. These NDPs operate primarily in civil infrastructure systems, as other types of networks do not generally share the same "selfish travelers" assumptions. In this paradigm, the performance of infrastructure improvements is assumed to depend primarily on the route choices of travelers (the commuter) during peak periods of travel, which in turn depend on the choices of other travelers. The dynamic component further allows modelers to assess intelligent transportation systems (ITS) that require more realistic modeling of traffic propagation obeying physical queuing constraints and information flow. Some examples include the stochastic dynamic NDP from Waller and Ziliaskopoulos (2001), Heydecker's (2001) NDP with dynamic user equilibrium (DUE), the linear DUE-NDP (Ukkusuri and Waller, 2008), dynamic toll pricing problem with route and departure time choice (Joksimovic *et al.*, 2005), and the reliability maximizing toll pricing problem with dynamic route and departure time choice (Li *et al.*, 2007). Although these NDPs are especially useful for ITS evaluation and operational strategies, they focus primarily on choices made over a single trip.

Tactical level NDPs tend to place more emphasis on time use and scheduling over congestion effects. Tactical service NDPs (Crainic, 2000) are a specific class used to manage fleets of vehicles with such temporal decision variables as service frequency. However, most of these NDPs focus on the schedules of the service being provided, rather than on incorporating the demand-side schedules of the travelers/users as endogenous elements of the design. Despite the incorporation of temporal effects, most service NDPs assume trip-based demand. There has been a surge of research in schedule-based transit assignment (as opposed to NDP), where travelers' departure time choices are handled explicitly. Tong and Wong (1998) formulated such a model with heterogeneous traveler values of time. Poon *et al.* (2004) presented a dynamic equilibrium model for schedule based transit assignment. Hamdouch and Lawphongpanich (2008) developed a schedule-based transit assignment model that accounts for individual vehicle capacities. They proposed one of the few schedule-based service network design problems, in the form of a transit congestion pricing problem that models passengers' departure time choices (Hamdouch and Lawphongpanich, 2010). Their model uses a time-expanded network and considers fare pricing to optimize the distribution of travelers within specific capacitated transit vehicles. The origin-destination (OD) demand remains as fixed trips, and not as linked itineraries.

Despite having the greatest need for such consideration, there are no NDP models at the planning level that consider routing and scheduling choices of travelers. It has long been acknowledged that models of traveler activities and time use are much more accurate than statistical trip-based approaches (Recker, 2001; Pinjari and Bhat, 2011). Activity consideration can bring about a tighter integration of infrastructure investment with land use planning and demand management strategies. Activity-based models can capture realistic impacts on travelers that are not limited to single trips but rather to chains of trips and activities forming detailed daily itineraries. Historically, the bulk of activity-based models have been designed as econometric models that do not account for network routing and scheduling mechanisms. The emerging trend in seeking to integrate network characteristics has been to force an interaction with a dynamic traffic assignment problem (e.g. Lin *et al.*, 2008; Konduri, 2012). However, this approach still ignores the network constraints present in scheduling and selection of activities for a household. There have been two primary exceptions to this approach. The first is the disaggregate activity route assignment model (HAPP) pioneered by Recker (1995), with subsequent studies on dynamic rescheduling/rerouting of those itineraries Gan and Recker (2008) and calibration of the activity route assignment models (Recker *et al.*, 2008; Chow and Recker, 2012). The second is the aggregate time-dependent activity-based traffic assignment model (Lam and Yin, 2001; Fu and Lam, 2013). Both modeling frameworks address the issue of activity scheduling, although Lam and Yin's model gives up disaggregate itinerary route choices and trip chains in favor of capturing congestion effects.

Although the transportation planning field has not seen any significant NDP research that models traveler routing and scheduling, the private logistics field has. One such model is the location routing problem (LRP), formulated and solved by Perl and Daskin (1985). The LRP is a set of inter-related problems that includes a facility location problem. What distinguishes LRPs from other facility location problems is that it doesn't assume that demand to a node is accessed through a single round trip. Instead, a lower level vehicle routing problem is embedded in the model to satisfy demand nodes in the most efficient manner, subject to where the facilities are located. In essence, it is an integrated NDP that accounts for responsive routing and scheduling. Numerous studies have been conducted on variants of the problem or on applications in industry. Several literature reviews have been published, including one from Min *et al.* (1998) and a more recent contribution by Nagy and Salhi (2007). Problem types developed over the years that may be applicable to activity-based network design in transportation planning include: stochastic LRP (Laporte and Dejax, 1989), where there is more than one planning horizon with time-dependent customer locations and demand; LRP with a mixed fleet Wu *et al.* (2002) for multimodal network consideration; location-routing-inventory (Liu and Lee, 2003) for modeling activity types as inventory-based needs that are fulfilled periodically; and LRP with nonlinear costs (Melachovsky *et al.*, 2005) that may provide means to incorporate congestion effects at link or activity node level. Readers are referred to Nagy and Salhi's paper for further details. One direct application of LRP with a truck fleet replaced by household travelers is shown in Chapter 4.

They use HAPP as a routing subproblem in a hydrogen vehicle refueling station LRP that allows the behavioral impacts of households' responses to located facilities to reflect siting decisions.

Given the increasing realization that transportation planning needs to reflect travelers' preferences at the level of the activity, we make a parallel observation to Perl and Daskin—that in the transportation planning field there is also a need for integrated NDPs that feature explicit consideration of travelers' tour patterns that include trip chaining, scheduling, time windows and even destination choice. At the activity-based level, we are concerned more with tactical and planning level policies, and less so with such operational technologies as ITS and information flow (hence foregoing congestion effects for now). In essence, we propose to change the conventional NDP, with a given OD matrix, to a new class of activity-based NDPs. This new problem accounts for a population of travelers with demand for activities at particular locations and at particular times, which are fulfilled via calibrated activity routing models. Like the LRP, the activity-based NDP is a set of integrated models. Unlike the conventional NDP, the OD matrix is not given *a priori*, but rather depends on the scheduling choices of households, which in turn depend on travel impedances. The solution of this set of models is a corresponding set of infrastructure link investments as well as the resulting optimal itineraries decided by the households in response to changes in link travel characteristics. The itineraries can then be aggregated to obtain the final OD matrix resulting from the NDP.

5.2 MOTIVATING EXAMPLES

The argument that we provide here, much like Perl and Daskin (1985) did for locating warehouses, is that the choice of which element of a network to improve can have a significant impact on how households set their itineraries each day. Trip-based (even dynamic ones) or fixed schedules ignore such changes as departure time, sequence of activities, or routing that each driver/household makes according to the changes made in the network. The following three cases demonstrate the influence that network designs can have on a household, which would be unaccountable under trip-based circumstances. For these examples, the utility maximization framework from Recker (1995) is assumed: households are multi-objective decision makers with their own sets of objectives with respective weights that dictate how they choose to schedule and route their activities. This has been demonstrated empirically by Chow and Recker (2012), where a population of households were fitted with heterogeneous sets of objective weights and desired arrival times to activities such that each of their observed itineraries were considered optimal to them.

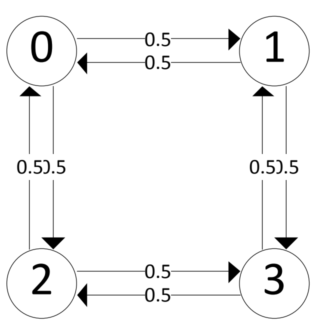
5.2.1 Departure time choice and itinerary re-sequencing

Assume a household has one member and one vehicle, and two activities to perform for the day: a work activity and a grocery shopping activity. Specifications of start (a_u, b_u) and completion (a_{n+u}, b_{n+u}) time windows and activity durations (s_u) are shown in Table 1, in units of hours. Here and throughout, the notation used in Recker (1995) is followed. Assume also that the household objective is solely to minimize the length of their itinerary, i.e., $\min Z = \sum_{v \in V} (T_{2n+1}^v - T_0^v)$, where T_u^v is the arrival time to node u via vehicle v, and node 0 is the home starting point while node $2n$+1 is the home ending point.

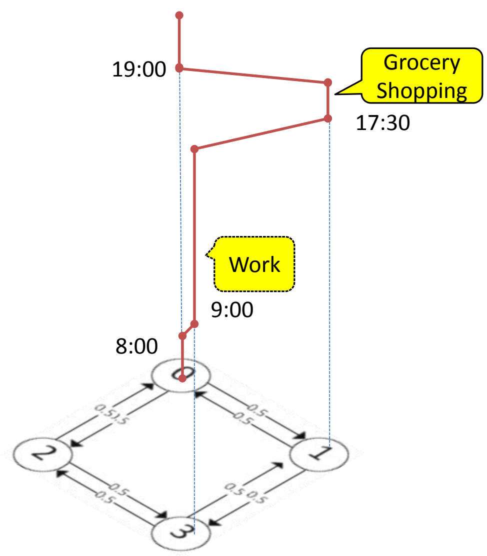
Table 5-1 Case 1 household characteristics

	Location	$\left[a_u, b_u\right]$	$\left[a_{n+u}, b_{n+u}\right]$	s_u
Home	Node 0	$\left[a_0, b_0\right]=\left[6, 21\right]$	$\left[a_{2n+1}, b_{2n+1}\right]=[10, 22]$	NA
Work activity	Node 3	$\left[9, 9\right]$	$\left[10, 22\right]$	8
Grocery Shopping activity	Node 1	$\left[5, 20\right]$	$\left[6, 22\right]$	1

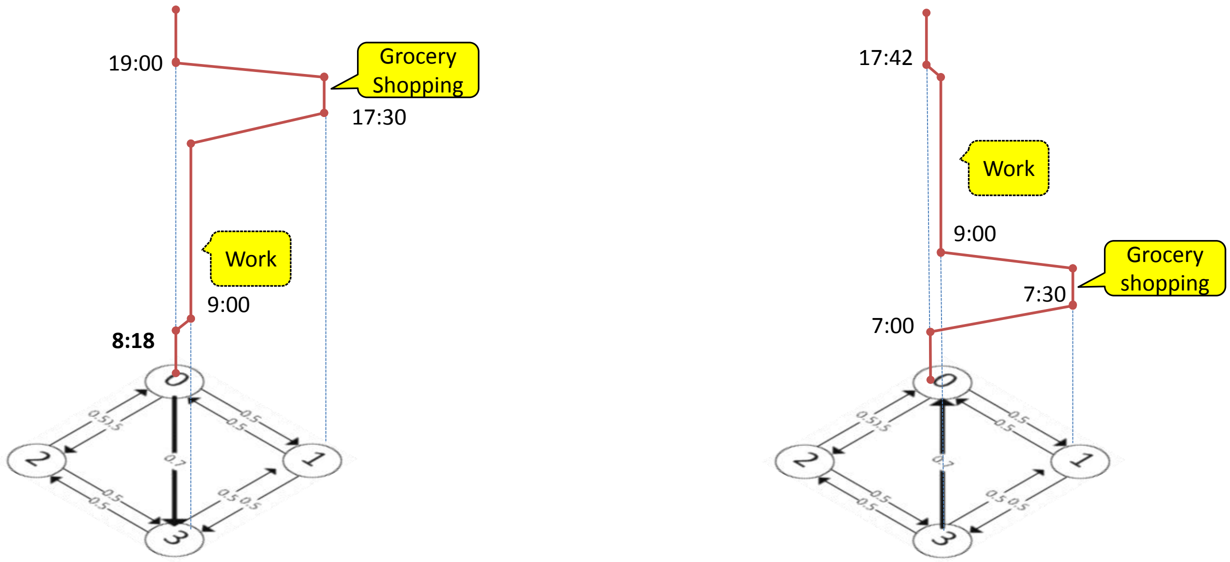
Assume a grid network with four nodes, and network connections as shown in Figure 5-1-(a). Travel time on each link, t_{ij}, is 0.5 hours. Figure 5-1-(b) shows the optimal pattern if no investment is made.



(a) Current Network



(b) Optimal Household Activity Pattern



(c) $t_{03} = 0.7 \rightarrow$ Time Adjustment   (d) $t_{30} = 0.7 \rightarrow$ Activity Sequence and Time Adjustment

Figure 5-1 The optimal household activity patterns for Case 1

Even in this simplest case, two types of schedule responses can be observed for standard link investments which would be ignored in conventional NDPs. If link {0,3} is constructed with travel time of 0.7 hours as shown in Figure 5-1-(c), the household member would now be able to delay their departure time from 8AM to 8:18AM. Alternatively, if link {3,0} is instead constructed with travel time of 0.7 hours as shown in Figure 5-1-(d), the optimal itinerary results in a re-sequence of activities as well as an adjustment in departure times.

5.2.2 Trip Chaining Trade-offs

A paradoxical consequence of considering elastic itineraries in network design is that it is possible to evaluate a *link investment that generates traffic without any increase in economic activity*. Traditionally, the argument made with elastic demand considerations is that improving infrastructure may result in additional trips made to fulfill latent demand between an OD pair. However, exceptions can also exist if travel is viewed as a way of achieving objectives while constrained within a space-time prism. By relaxing some of those constraints through network improvements, we may observe only increased trips due to untangling of less desired travel patterns within the tighter constraints. This can result in more trips made if it improves the overall objective of the household but would not contribute in any way to economic demand because the household may be reconfiguring the same itinerary without adding new destinations to visit. This occurrence can be best illustrated with a household with activities that have very strict time windows.

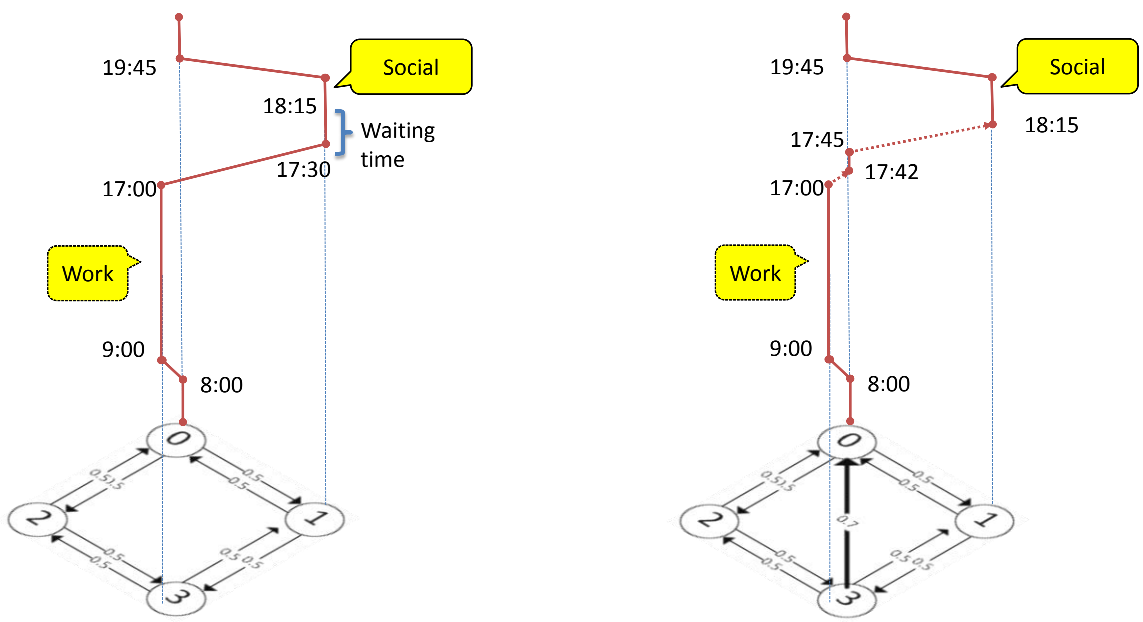
We consider the same activity agenda as in the previous section, but with both activities having strict start time windows as in Table 5-2. Both activities require the household member to be at the respective locations at a specific time, which is often quite a realistic assumption. Assume also that this particular household has two potentially conflicting objectives: to minimize the travel time with weight β_T, and to minimize delay from returning home after an activity, with weight β_C. The delay from the returning home objective represents the desire of the household to minimize the duration of any particular activity period away from home, as discussed by Recker (1995) and calibrated empirically by Chow and Recker (2012) for a set of households. The higher the weight of this objective relative to travel time, the more likely it is that a household would not want to trip chain. Then the objective function becomes:

$$\min Z = \beta_T \cdot \sum_{v \in V} \sum_{w \in \mathbf{N}} \sum_{u \in \mathbf{N}} t_{uw} \cdot X_{uw}^{v} + \beta_C \cdot \sum_{u \in \mathbf{N}} (T_{u+n} - T_u)$$

where X_{uw}^{v} is a binary variable representing a route taken between node u and node w with vehicle v, and the weights are assumed to be $\beta_T = 1$ and $\beta_C = 1$. The optimal solution on the base network is shown in Figure 5-2-(a), with an objective function travel disutility of 14.25 and a total of three trips made. Due to the time windows, the household traveler is constrained to trip chain from the work activity to the social activity.

Table 5-2 Case 2 household characteristics

	Location	$[a_u, b_u]$	$[a_{n+u}, b_{n+u}]$	s_u
Home	Node 0	$[a_0, b_0] = [6, 21]$	$[a_{2n+u}, b_{2n+u}] = [10, 22]$	NA
Work activity	Node 3	$[9, 9]$	$[10, 22]$	8
Social activity	Node 1	$[18.25, 18.25]$	$[18.5, 22]$	1



(a) Optimal Household Activity Pattern (b) $t_{30} = 0.7 \rightarrow$ extra trip generation

Figure 5-2 The optimal household activity patterns for Case 2

Now consider a link addition {3,0} with travel time of 0.7 hours. Because the household can now return home immediately after work and still make the social activity in time, they do so for an improved travel disutility of 12.9. The result is not only a change in trip ODs (due to re-sequence in a tour), but one extra trip is also created as shown in Figure 5-2-b (4 trips). Essentially a trip has been added without adding a new non-home destination to visit, but the household sees an improvement in travel disutility because of the relaxation of spatial-temporal constraints that were binding before the network improvement. A conventional trip-based approach, or even a fixed schedule approach, would miss such a response altogether.

5.2.3 Increasing travel disutility

If we consider a continuous link improvement (in which a route travel time is improved), then another counterintuitive situation can occur. Consider the household in Table 5-2 again, but in this case let's assume that the household seeks to minimize idle time. Idle time is defined as the extent of the travel day that is not used in performing activities or traveling—such tradeoffs are similar to studies comparing values of in-vehicle travel time against out-of-vehicle access or idle/wait time. The potential for conflict between the two objectives is not immediately apparent; however, in the presence of strict time windows it is possible that improving travel times can result in increasing idle time. Consider the following:

$$\min Z = (\beta_T - \beta_W) \cdot \sum_{v \in V} \sum_{w \in \mathbf{N}} \sum_{u \in \mathbf{N}} t_{uw} \cdot X_{uw}^v + \beta_W \cdot \sum_{v \in V} (T_{2n+1}^v - T_0^v)$$

where $\beta_T = 1$ and $\beta_W = 1.5$. The durations of the activities s_u are not included because they are constant and drop out. In the base case shown in Figure 5-2-(a), the disutility under this new objective is 16.625 instead of 14.25.

If a continuous improvement is made to link {3,1} such that travel time improves from 0.5 hours to 0.25 hours (e.g. repaving, lane expansion), then due to time window constraints there are no other alternative routes and the household would still have to follow the same schedule. However, this results in a direct trade-off between travel time and idle time. *If a household values idle time minimization more than travel time minimization, then such an improvement can result in a paradoxically higher disutility, even without considering congestion effects.* The travel time improvement simply results in a decrease in the travel time objective of 0.25 but a direct increase in idle time of 0.25. Since $\beta_W > \beta_T$, the disutility actually increases from 16.625 to 16.75. Effects such as this would be completely ignored if NDPs were applied without considering their effect on household scheduling. However, explicitly incorporating household scheduling mechanisms into the NDP allow paradoxes such as this to be avoided.

We have presented three scenarios that can arise from network improvements when realistically considering the effects they have on household scheduling and planning. Network changes can cause significant reshaping of temporal /spatial constraints for households that result in changes in their trip patterns. We argue that these effects should not be ignored when considering NDPs at the tactical or planning level.

5.3 NDP-HAPP

5.3.1 Definitions

The activity-based NDP using HAPP subproblems to address household schedule response to network changes is here designated as NDP-HAPP. As a kernel activity-based NDP, the NDP-HAPP is formulated using the simplest structure. More complex formulations that explore link capacities, vehicle and household member interactions, multimodal networks, or congestion effects will be explored in future research. The kernel formulation is first presented as a set of multiple subproblems, and then further modified to consider activity choice in cases with non-compulsory activities. There are two distinct types of networks in this problem: an infrastructure network where changes can actively be made, and a responsive activity network that represents the routing and scheduling decisions made at the household level. Assume an infrastructure network layer L_I, and the following parameters for the infrastructure network system:

$\mathbf{N}$ set of all nodes in the analysis

$\mathbf{E}$ set of all direct links in the analysis

F_{ij} fixed link design costs

c_{ij} operational per unit link routing costs

B total budget for the network system

t_{ij} travel time between the direct link from node i to node j

$c_{ij}^{v,h}$ personal travel cost for vehicle v of household h, between the direct link from node i to node j

Variables related to the infrastructure network system are:

f_{ij} flow on the direct link (i, j)

z_{ij} binary decision variable that indicates whether or not link (i, j) is chosen as part of the network's design

Assume also an activity layer L_A, and the following parameters for the activity network system:

$\mathbf{P}$ set of all activity nodes in the analysis. It is a subset of the node set from the infrastructure network, $\mathbf{N}$.

(u, w), $u, w \in \mathbf{P}$ route from activity point u to activity point w. Its connectivity is derived from L_I.

$\mathbf{H}$ set of households using on the activity nodes $\mathbf{P}$ in the analysis.

Although their physical locations are the same, the two sets of networks operate in a bi-level fashion. This bi-level property of NDP-HAPP, together with its unfolding in the time-space dimension, can be depicted conceptually in Figure 5-3. Such separation of networks—a supernetwork approach—has been used widely in activity-based transportation networks, mainly concerning various modal choices and their specific networks (e.g. TRANSIMS, 2012; Arentze and Timmermans, 2004). However, an optimization-based routing and scheduling procedure, to our knowledge, has never been applied to the activity layer in response to infrastructure changes.

Following the notation of Recker (1995), we define the following sets and parameters that are specific for each household, $h \in \mathbf{H}$:

$\boldsymbol{\beta}_h = \{\beta_h^a, \beta_h^b, \ldots\}$ set of relative weights for different travel disutility terms for household h

A_h set of out-of-home activities to be completed by travelers in household h

V_h set of vehicles used by travelers in household h to complete their scheduled activities.

$n_h = |A_h|$ number of activities to be performed by household h

$P_h^+ \subset \mathbf{P}$ set designating location at which each assigned activity is performed by travelers in household h; the set of activities and their physical locations are different for each household.

$P_h^- \subset \mathbf{P}$ set designating the ultimate destination of the "return to home" trip from out-of-home activities to be completed by travelers in household, h. (Note: the physical location of each element of P_h^{A-} is "home".)

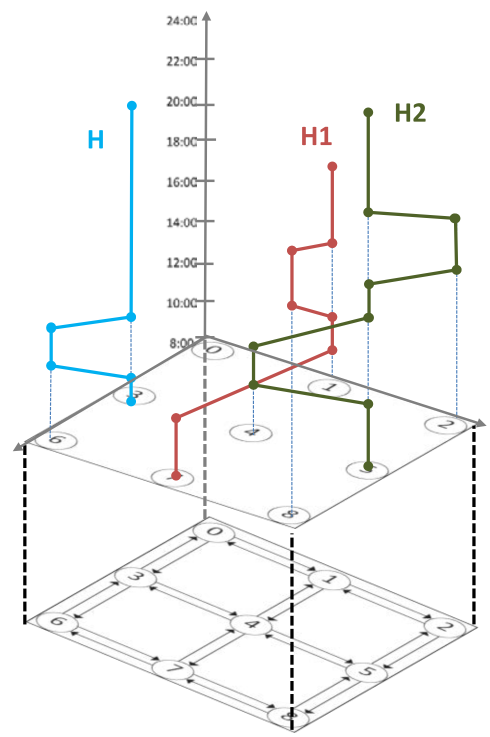


Figure 5-3 Bi-level interactions between the infrastructure and activity networks

$\left[a_u^h, b_u^h\right]$ time window of available start times for activity u for household h.

$\left[a_{n_h+u}^h, b_{n_h+u}^h\right]$ time windows for the "return home" arrival from activity u of household h. (Note: b_u^h must precede $b_{n_h+u}^h$ by an amount equal to or greater than the duration of the activity.)

$\left[a_0^h, b_0^h\right]$ departure window for the beginning of the travel day for household h.

$\left[a_{2n_h+1}^h, b_{2n_h+1}^h\right]$ arrival window by which time all members of the household h must complete their travel.

s_u^h: duration of activity u of household h.

t_{uw}^h travel time from the location of activity u to the location of activity w.

$c_{uw}^{\upsilon,h}$ travel cost for household h, from location of activity u to the location of activity w by vehicle v.

B_C^h travel cost budget for household h.

$B_T^{\upsilon,h}$ travel time budget for the household h's member using vehicle v.

$\mathbf{P}_h = \mathbf{P}_h^+ \cup \mathbf{P}_h^-$ set of nodes comprising completion of the activities of household h.

$\mathbf{Q}_h = \{0, \mathbf{P}_h, 2n_h + 1\}$ set of all nodes for household h, including those associated with the initial departure and final return to home. This is a subset of $\mathbf{P}$.

In the following, i, j are used to refer to nodes in the infrastructure layer, and link (i, j) refers to the direct link connecting those two nodes. Notation u, w are used to refer to activity nodes in the activity layer, and it is not necessarily a direct infrastructure link but rather a path between (u, w). Such path information as travel time and travel cost are passed onto the activity layer from the infrastructure layer, but the connectivity data of the path needs to be drawn from the infrastructure layer.

The household-specific decision variables are:

$X_{uv}^{v,h}, \quad u, w \in \mathbf{Q}_h, v \in V_h, h \in \mathbf{H}$	binary decision variable equal to one if vehicle v travels from activity u to activity w, and zero otherwise.
$T_u^h, \quad u \in P_h, h \in \mathbf{H}$	time at which participation in activity u of household h begins.
$T_0^{v,h}, T_{2n_h+1}^{v,h}, \quad u \in P_h, h \in \mathbf{H}$	times at which vehicle v from household h first departs from home and last returns to home, respectively
$Y_u^h, \quad u \in P_h, h \in \mathbf{H}$	total accumulation of either sojourns or time (depending on the selection of D and d_u) of

household h on a particular tour immediately following completion of activity u.

Variables connecting the infrastructure network, L_I, and the activity network system, L_A, are:

$\delta_{uw,ij} = \delta_{uw,ij}(\mathbf{z})$ binary indicator variable whether route (u, w) in the activity network uses link (i, j) in L_I. Assuming the shortest cost path is used between two activity nodes, the design variables determine the connectivity of nodes in L_I. If link (i, j) is not constructed, $z_{ij} = 0$, $\delta_{uw,ij}$ is automatically 0, and otherwise, it can be identified by solving a shortest path problem between the origin and the destination, (u, w).

$$\delta_{uw,ij} = \delta_{uw,ij}(\mathbf{z}) = \begin{cases} 0 & z_{ij} = 0 \\ (\delta_{uw,ij})^* & z_{ij} = 1 \end{cases}$$

where $(\delta_{uw,ij})^*$ is the solution of a shortest path problem for each activity link (i, j), i.e.

Shortest Path Allocation Problem

$$\min \sum_{(i,j)\in\mathbf{E}} t_{ij} \cdot \delta_{uw,ij} \tag{5-1}$$

subject to

$$\sum_{j\in\mathbf{N}} \delta_{uw,ji} - \sum_{j\in\mathbf{N}} \delta_{uw,ij} = \begin{cases} 1 & i = u \\ 0 & i \neq u, w \\ -1 & i = w \end{cases} \tag{5-2}$$

$$\delta_{uw,ji} \in (0,1) \tag{5-3}$$

The problem is defined for all households and their activity routes, $u, w \in \mathbf{Q}_h, v \in V_h, h \in \mathbf{H}$.

$t_{uw} = t_{uw}(\mathbf{z})$ travel time from the location of activity *u* to the location of activity *w*. It is a function of the decision variable vector $\mathbf{z}$, and the given network $(\mathbf{N}, \mathbf{E})$ since the connectivity decision variables z_{ij} determine the travel times.

$$t_{uw} = t_{uw}(z) = \sum_{j\in\mathbf{E}} \sum_{i\in\mathbf{E}} t_{ij} \cdot \delta_{uw,ij}(\mathbf{z}), \quad u, w \in \mathbf{Q}_h, h \in \mathbf{H} \tag{5-4}$$

$c_{uw}^{v,h} = c_{uw}^{v,h}(\mathbf{z})$ travel cost from the location of activity u to the location of activity w for vehicle v of household h. It is a function of the decision variable vector $\mathbf{z}$, and the given network $(\mathbf{N}, \mathbf{E})$ since the connectivity decision variables z_{ij} determine the travel costs.

$$c_{uw}^{v,h} = c_{uw}^{v,h}(z) = \sum_{j \in \mathbf{E}} \sum_{i \in \mathbf{E}} c_{ij}^{v,h} \cdot \delta_{uw,ij}(\mathbf{z}), \quad u, w \in \mathbf{Q}_h \tag{5-5}$$

$f_{ij} = f_{ij}(\mathbf{X})$ link flow on direct link *ij*. It is a function of the household activity decision variable vector, $\mathbf{X}$, and connects the path flow on layer L_A to the link flow on layer L_I. It is a function of the decision variable vector $\mathbf{z}$, and the given network $(\mathbf{N}, \mathbf{E})$ since the connectivity decision variables z_{ij} determine the link flows.

$$f_{ij}(X) = \sum_{h \in \mathbf{H}} \sum_{u \in \mathbf{Q}_h} \sum_{w \in \mathbf{Q}_h} \sum_{v \in V^h} \delta_{uw,ij}(\mathbf{z}) \cdot X_{uw}^{v,h}, \quad (i,j) \in \mathbf{E} \tag{5-6}$$

5.3.2 Decomposed Formulation of NDP-HAPP

Typically, the LRP formulation includes three parts: location, routing, and allocation. This property applies to NDP-HAPP as well, where the upper level "location" is the network design variables and the lower level routing part is the HAPP model. Allocation refers to assignment of the activity link impedance from the shortest path problem in the infrastructure network, shown in Equation (5-1) – (5-3). The objective function of the upper problem in the LRP is to minimize the overall cost, which is comprised of depot cost and vehicle cost.

Similarly, NDP-HAPP in the most basic form is decomposed into two models solved as a bi-level problem: NDP (upper) and HAPP (lower). There are two sets of decision makers, so the solution can be classified as a leader/follower Stackelberg equilibrium, as described in Yang and Bell (1998). Instead of a traffic equilibrium lower level problem, the NDP-HAPP has a set of household scheduling problems in the lower level for each household. Considering the network design problem as the upper level decision and the household activity/scheduling/routing decisions (HAPP) as reactions to the network design, we can express the problem most generally in Equations (5-7).

$$\begin{aligned} &\min_{z,f} G(z, f(X)) = \varphi_{dNDP}(z, f) \\ &\text{subject to} \\ &H(z, f(X)) \le 0 \end{aligned} \tag{5-7a}$$

where

$$\begin{aligned} &\min_{X,T} g(X(\delta(f,z)), T(\delta(f,z))) = \varphi_{dHAPP}(X,T) \\ &\text{subject to} \\ &h(z, X(\delta(f,z)), T(\delta(f,z))) \le 0 \end{aligned} \tag{5-7b}$$

where G is the objective function, z is the decision vector, and H is the constraint set of the upper level problem. In the lower level problem, g is the objective function, X,T are the decision vectors, and h is the constraint set.

The kernel network design problem we present is a modified version of the unconstrained multicommodity case of the formulation in Magnanti and Wong (1984). The formulation minimizes the design cost while satisfying the given flow demands at origin and destination nodes. The formulation is in terms of direct links and link flows only, whereas the integrated NDP-HAPP includes path flows which are connected by $\delta_{uw,ij}$ to direct link flows, f_{ij} ,. In order for the OD pairs to be assigned to sequences of direct links, we treat each OD pair (u,w) as a commodity as in the case of multicommodity flow problems, i.e., we define a single commodity $f_{ij}^{uw}, \forall (u,w) \in \mathbf{K}$ where $f_{ij} = \sum_{(u,w)\in \mathbf{K}} f_{ij}^{uw}$, and where $\mathbf{K}$ is the set of all OD (u,w) pairs.

We formulate this decomposed NDP (dNDP) in terms of direct link flows only, and each OD pair is represented as a commodity. The demand values are calculated as shown in Equation (14). They take household sequence decisions and aggregate them into origin-destination pairs.

Upper Level NDP (dNDP)

$$\min \varphi_{dNDP}(z, f) = \sum_{(i,j)\in \mathbf{E}} F_{ij} \cdot z_{ij} + \sum_{(i,j)\in \mathbf{E}} c_{ij} \cdot f_{ij} \tag{5-8}$$

subject to:

$$\sum_{j\in \mathbf{N}} f_{ji}^{uw} - \sum_{l\in \mathbf{N}} f_{il}^{uw} \geq D^{uw}, \quad \forall i = u \in \mathbf{N}, \forall (u,w) \in \mathbf{K} \tag{5-9}$$

$$\sum_{j\in \mathbf{N}} f_{ij}^{uw} - \sum_{l\in \mathbf{N}} f_{li}^{uw} \geq D^{uw}, \quad \forall i = u \in \mathbf{N}, \forall (u,w) \in \mathbf{K}, \tag{5-10}$$

$$\sum_{j\in \mathbf{N}} f_{ji}^{uw} - \sum_{j\in \mathbf{N}} f_{ij}^{uw} = 0, \quad \forall i \in \mathbf{N}, i \neq u, i \neq w, \forall (u,w) \in \mathbf{K} \tag{5-11}$$

$$f_{ij}^{uw} \leq D^{uw} \cdot z_{ij}, \quad \forall (i,j) \in \mathbf{E}, (u,w) \in \mathbf{K} \tag{5-12}$$

$$z_{ij} \in (0,1), \quad (i,j) \in \mathbf{E} \tag{5-13}$$

where

$$D^{uw} = \sum_{h \in \mathbf{H}} \sum_{v \in V^h} X_{uw}^{v,h}, \qquad w = i \in \mathbf{N}, \forall (u,w) \in \mathbf{K} \tag{5-14}$$

Equations (5-9) – (5-10) require each path $(u,w) \in \mathbf{K}$ to satisfy the given OD demand. Equations (5-11) simply show the conservation of flows for intermediate nodes. Equation (5-12) constrains flow variables to be on the links that are built in a manner that does not exceed the capacity. Because we do not consider cases in which the capacity of links is exceeded in this problem, only the shortest path will be loaded with flows. As such, the shortest path information is provided directly by the f_{ij}^{uw} variable. We can implicitly obtain the shortest path variables for each OD pair as shown in Equation (5-15) instead of having to solve Equations (5-1) – (5-3) separately.

$$\delta_{uw,ij} = \begin{cases} 0 & f_{ij}^{uw} = 0 \\ 1 & \text{otherwise} \end{cases}, \qquad \forall (i,j) \in \mathbf{E}, (u,w) \in \mathbf{K}, v \in V_h, h \in \mathbf{H} \tag{5-15}$$

The decomposed lower-level HAPP (dHAPP) problem is shown in Equations (5-16) – (5-19). It is composed of the set of constraints in the Appendix which would be equivalent to the original constraints from Case 1 in Recker (1995) if travel time/cost factors are not functions of the allocated shortest path. More complex variations presented in Recker (1995) can be substituted if household member interactions and carpooling effects are desired. Also, each household can be treated separately since all of the constraints and objective functions are separable by household. With constant travel times/costs, i.e., without congestion effects, each household's dHAPP is solved separately.

Lower Level HAPP (dHAPP) for Each Household

$$\min \varphi_{dHAPP}(X,T) = \sum_{h\in\mathbf{H}}\sum_{v\in V^h}\beta_h^T\cdot(T_{2n_h+1}^{v,h}-T_0^{v,h})+\sum_{h\in\mathbf{H}}\sum_{u\in\mathbf{Q}_h}\sum_{w\in\mathbf{Q}_h}\sum_{v\in V^h}\beta_h^C\cdot c_{uw}^{v,h}\cdot X_{uw}^{v,h} \tag{5-16}$$

Subject to

(5-A-1) – (5-A-26)

where

$$t_{uw}(z)=\sum_{(i,j)\in\mathbf{E}} t_{ij}\cdot\delta_{uw,ij},\quad u,w\in\mathbf{Q}_h, h\in\mathbf{H} \tag{5-17}$$

$$c_{uw}^{v,h}(z)=\sum_{(i,j)\in\mathbf{E}} c_{ij}^{v,h}\cdot\delta_{uw,ij},\quad u,w\in\mathbf{Q}_h \tag{5-18}$$

$$\delta_{uw,ij}=\begin{cases}0 & f_{ij}^{uw}=0\\ 1 & \text{otherwise}\end{cases},\qquad \forall(i,j)\in\mathbf{E},(u,w)\in\mathbf{K},v\in V_h,h\in\mathbf{H} \tag{5-19}$$

As discussed in other HAPP model studies, the objective shown in Equation (5-16) is just one example of a multi-faceted objective problem. Others can be specified and estimated using the method from Chow and Recker (2012). The process of specifying the multiple components of the objectives and calibrating their coefficients with desired arrival times can be thought of as a confirmatory modeling process that seeks to fit a hypothesis of how household travelers behave onto a data set. Fitness of an objective is determined by the significance of its estimated coefficient relative to other objectives. For example, a data set might reveal that Equation (5-16) results in a length of day coefficient (first term) equal to 0.0001 relative to a weight of 1 for the travel cost objective. In that case, it would suggest that the first objective is not very important in the travelers' scheduling choices, and removing it might result in smaller variances in the remaining objectives when re-calibrated.

NDP-HAPP as presented in Section 3.2 differs conceptually from the LRP in two primary ways. First, the LRP has a single decision maker involved in both planning and tactical strategic design, whereas the NDP-HAPP has a single decision maker involved in planning and multiple household decision makers responding to the plan at a tactical level. Second, the node demand for the upper level problem in the LRP is known *a priori*, but the cost of delivering service to the demand node is not known. Instead, it is derived from the output of the VRP. Alternatively, the NDP-HAPP does not have OD demand known *a priori*, but costs between nodes are given; the OD demand is derived from the output of the HAPP.

5.3.3 Generalized NDP-HAPP (NDP-GHAPP)

The NDP-HAPP model is extended to include the capability for households to choose locations for such non-primary activities as grocery shopping and refueling. This is done by relaxing the condition in the HAPP that requires each household to visit every location, determined exogenously; rather in NDP-HAPP each household visits one candidate location from a cluster of such activity types. This is similar to the generalized traveling salesman problem (e.g., the E-GTSP in Fischetti *et al.*, 1997) and generalized vehicle routing problem Ghiani and Improta (2000) in the logistics literature, where visits to nodes are modified to visits to single nodes from each cluster. The generalized HAPP (GHAPP) has been formulated and applied in the earlier chapters, and a "profitable tour" variation of this approach was developed for activity-based traveler information systems (Chow and Liu, 2012) and for testing algorithms in scenario analysis (Chow, 2013).

In GHAPP, the constraints in Equation (5-A1) are modified to Equation (5-A1-1) as shown in Chapter 2. Instead of requiring each node to have a flow, the generalized formulation instead requires one node from a cluster of nodes to be visited. Compulsory activity types would have only one node in the cluster, whereas such non-primary activities as grocery shopping or refueling could have multiple candidate nodes from which to choose.

$$\sum_{u \in \mathbf{P}_{A_a}^{+}} \sum_{v \in \mathbf{V}_h} \sum_{w \in \mathbf{Q}_h} X_{uw}^{v,h} = 1, \quad A_a \in \mathbf{A}, h \in \mathbf{H} \tag{5-A1-1}$$

where

$\mathbf{A} = \{A_1, A_2, ..., A_a, ..., A_m\}$ set of m different activity types with unspecified locations

$P_{A_a}^{+}$ set designating "potential" locations at which activity A_a may be performed

When integrated with NDP, GHAPP becomes infeasible if one or more candidate nodes are not connected to the network; constraints in (A7), (A11) also need to be modified to be conditional such that the temporal constraints are imposed only when there is a visit to that candidate location. This allows having one or more of unconnected candidate nodes, which have infinite travel times.

$$\sum_{v \in V_h} \sum_{w \in \mathbf{Q}_h} X_{wu}^{v,h} = 1 \Rightarrow T_u^h + s_u^h + t_{uw}^h(\mathbf{z}) = T_{n+u}^h, \quad u \in P_h^{+}, h \in \mathbf{H} \tag{5-A7-1}$$

$$\sum_{v \in V_h} \sum_{w \in \mathbf{Q}_h} X_{wu}^{v,h} = 1 \Rightarrow a_u^h \le T_u^h \le b_u^h, \quad u \in P_h, h \in \mathbf{H} \tag{5-A11-1}$$

Similarly, when the objective function involves time variables, those of the unvisited activity nodes need to be constrained. For example:

$$\sum_{v \in V_h} \sum_{w \in \mathbf{Q}_h} X_{wu}^{v,h} = 0 \Rightarrow T_u^h = 0, \quad u \in P_h, h \in \mathbf{H} \tag{5-A7-2}$$

5.3.4 Decomposition Solution Algorithm

There are many different types of solution algorithms developed for LRPs (Nagy and Salhi, 2007) and discrete or mixed NDPs (e.g. Luathep *et al.*, 2011; Wang *et al.*, 2013), and they can potentially be adopted for NDP-HAPP. However, the iterative method proposed here decomposes the problem into several blocks that actually represent each decision maker's rationale in this complex problem. Additionally, this kind of decomposition does not necessarily require the problem to be formulated in the structure of mathematical optimization as long as the drivers' response to the network design is captured and updated. This means that different types of integrated activity-based approaches can be used to model individuals' routing/scheduling behavior. Because the majority of these activity-based models are based on discrete-choice or simulation-based models (e.g., Bowman and Ben-Akiva, 2000; Bhat *et al.*, 2004; Balmer *et al.*, 2006), the suggested decomposition method is highly adaptable to different types of activity-based models.

The decomposed problems remain computationally challenging, particularly the NP-hard (Recker, 1995) HAPP. Because these problems are widely studied, there are various methods available. Geoffrion and Graves (1974) are referred for network design problems, and Cordeau and Laporte (2003) are referred for a survey of algorithms for the Pickup and Delivery Problem with Time Windows (PDPTW), on which the simplest HAPP is based. The decomposition proposed here is comparable to Perl and Daskin (1985) in the context of Location Routing Problems and the Iterative Optimization Assignment (IOA) algorithm in Yang and Bell (1998) in the context of bi-level Network Design Problems. Perl and Daskin (1985) used three decomposed models to tackle the warehouse location routing problem: the complete multi-depot vehicle-dispatch problem (MDVDP), the warehouse location-allocation problem (WLAP), and the multi-depot routing-allocation problem (MDRAP). The location-allocation and muti-depot routing allocation blocks are in parallel with dNDP and dHAPP. For NDP, the iterative optimization-equilibrium in Friesz and Harker (1985) includes similar blocks of Equilibrium Assignment Program and Design Optimization, in line with dHAPP and dNDP. Since there is no congestion in the dHAPP model, the issue of having IOA converge to a Cournot-Nash equilibrium is not relevant here. For the examples and case studies presented, the overall processes are coded in Java calling a CPLEX library for dHAPP and dNDP problems.

An iterative solution algorithm for the NDP-HAPP is depicted in Figure 5-4. First, the initial network decision solution is assumed to use all links, $z_{ij}^0 = 1, (i, j) \in \mathbf{E}$. Then, $\delta_{uw,ij}^0$ can be derived from z_{ij}^0 using the standard shortest path problem—for example, Floyd's Algorithm can be used to efficiently update the travel time matrix. Based on the updated travel times, dHAPP is solved independently for each household since no congestion effect is present. Hypothetically, if congestion is incorporated in future research (perhaps through integration with Lam and Yin's (2001) framework to provide feedback on the skim table in a fashion similar to how the Trip Distribution step can be updated with Trip Assignment results in the Four Step Model), this framework should still be feasible. After the travel decisions are made by each household, supply and demand are updated from Equations (5-15), and dNDP can then be solved as the conventional NDP. The proposed iterative process continues until there is no improvement in the objective function. The implicit shortest path allocation from the upper level problem and the path-link conversion conditions in Equations (5-4) – (5-6) are maintained throughout this iterative process. The same algorithm can be applied to NDP-GHAPP.

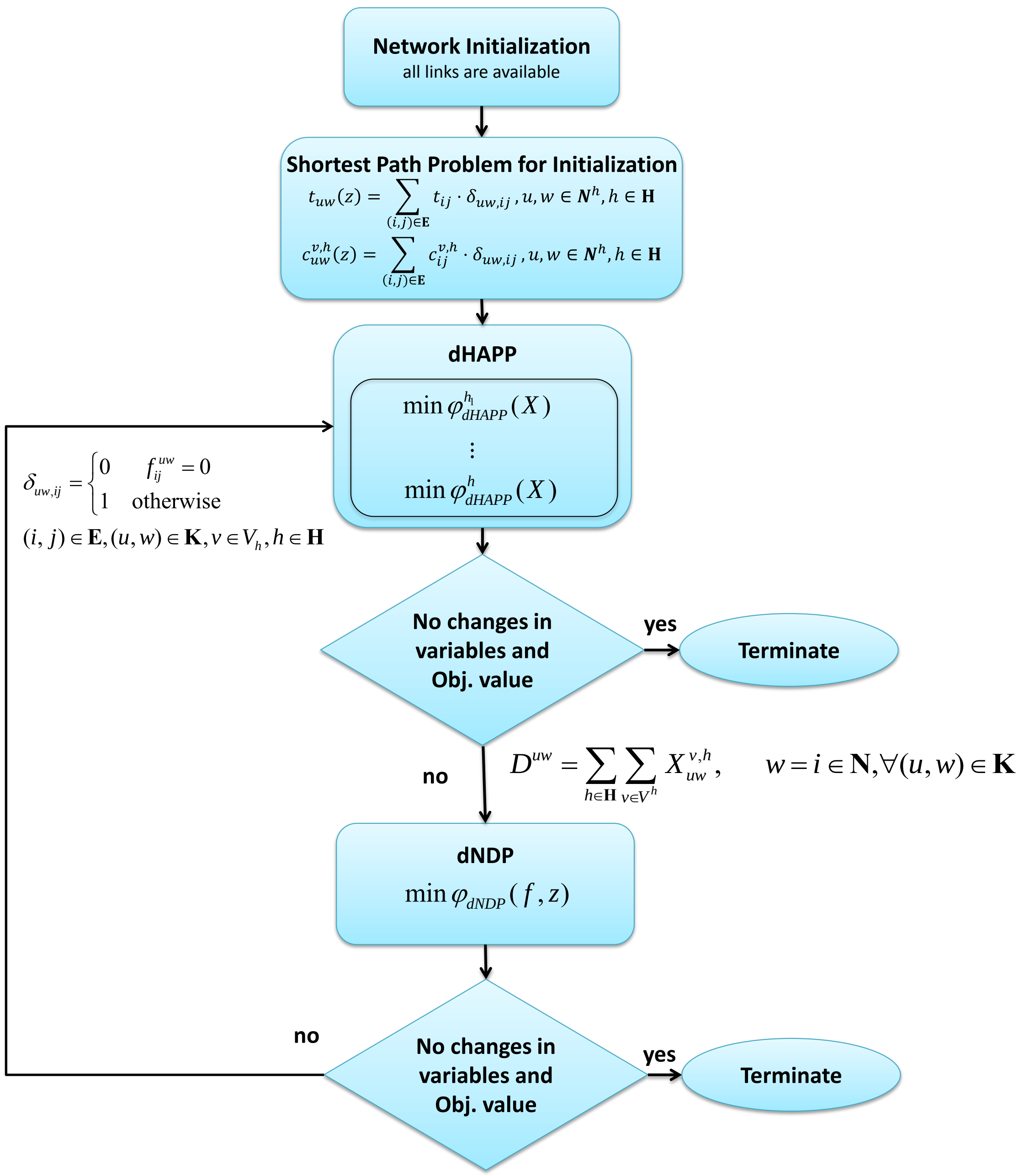


Figure 5-4 A decomposition solution method for NDP-HAPP

5.4 Numerical Examples

5.4.1 Simple example: NDP-HAPP

Assume a grid network with nine nodes, with possible link construction as shown in Figure 5. When constructed, travel time for each link is 0.5. Construction cost for each link is 3, and operating cost per link flow is 0.5, i.e., $\min \varphi_{dNDP}(z, f) = \sum_{(i,j)\in\mathbf{E}} 3 \cdot z_{ij} + \sum_{(i,j)\in\mathbf{E}} 0.5 \cdot f_{ij}$.

Assume two households, $\mathbf{H} = \{h_1, h_2\}$, with one vehicle each, $V_1 = \{1\}$, $V_2 = \{1\}$, and their activities $A_1 = \{\text{work, grocery shopping}\}$, $A_2 = \{\text{work, general shopping}\}$ to perform. These activities' locations in the infrastructure layer are shown in Figure 5, and their activity start/end time windows, activity durations are all shown in Table 3. Except for the activity start time windows of work activities, time windows are not necessarily constraining, leaving some room to explore different path sequences. Both households' objective functions are assumed to be minimizing total travel cost only (households consider the travel costs for all direct links), $\min \varphi^h_{dHAPP}(X) = \sum_{u\in\mathbf{Q}_h} \sum_{w\in\mathbf{Q}_h} \sum_{v\in V^h} c^{v,h}_{uw} \cdot X^{v,h}_{uw}$.

Table 5-3 Simple example household characteristics

	Location on L_I	$\left[a^h_u, b^h_u\right]$	$\left[a^h_{n_h+u}, b^h_{n_h+u}\right]$	s^h_u
h_1 home	Node 0	$\left[a^h_0, b^h_0\right] = [6, 21]$	$\left[a^h_{2n_h+1}, b^h_{2n_h+1}\right] = [10, 24]$	NA
h_1 work activity	Node 2	$[9, 9.5]$	$[10, 22]$	8
h_1 grocery shopping activity	Node 5	$[5, 22]$	$[10, 22]$	1
h_2 home	Node 5	$\left[a^h_0, b^h_0\right] = [6, 21]$	$\left[a^h_{2n_h+1}, b^h_{2n_h+1}\right] = [10, 22]$	NA
h_2 work activity	Node 6	$[8.5, 9]$	$[10, 22]$	8
h_2 general shopping activity	Node 8	$[5, 21]$	$[10, 22]$	1

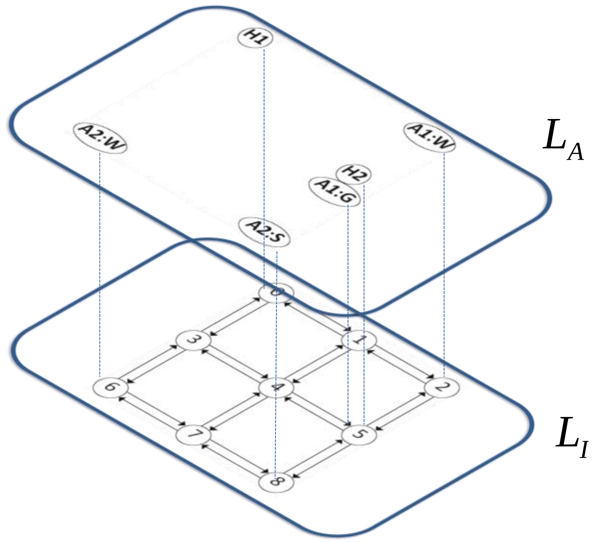


Figure 5-5 Supernetwork depiction

Because the NDP-HAPP is not a simple problem to check for optimality, all possible combinations of household decisions are enumerated and given to dNDP, and its objective value combined with the objective value of corresponding household decision combination is used to derive the true optimal solution value. Figure 5-6 shows the solution from the proposed method (5-6-(a), 5-6-(b)) and the actual optimal solution (5-6-(c), 5-6-(d)). The decomposition solution converged after one iteration and is 5% worse than the actual optimal solution, 40, for this simple example. Detailed illustration of the computational process is available in Table 5-4.

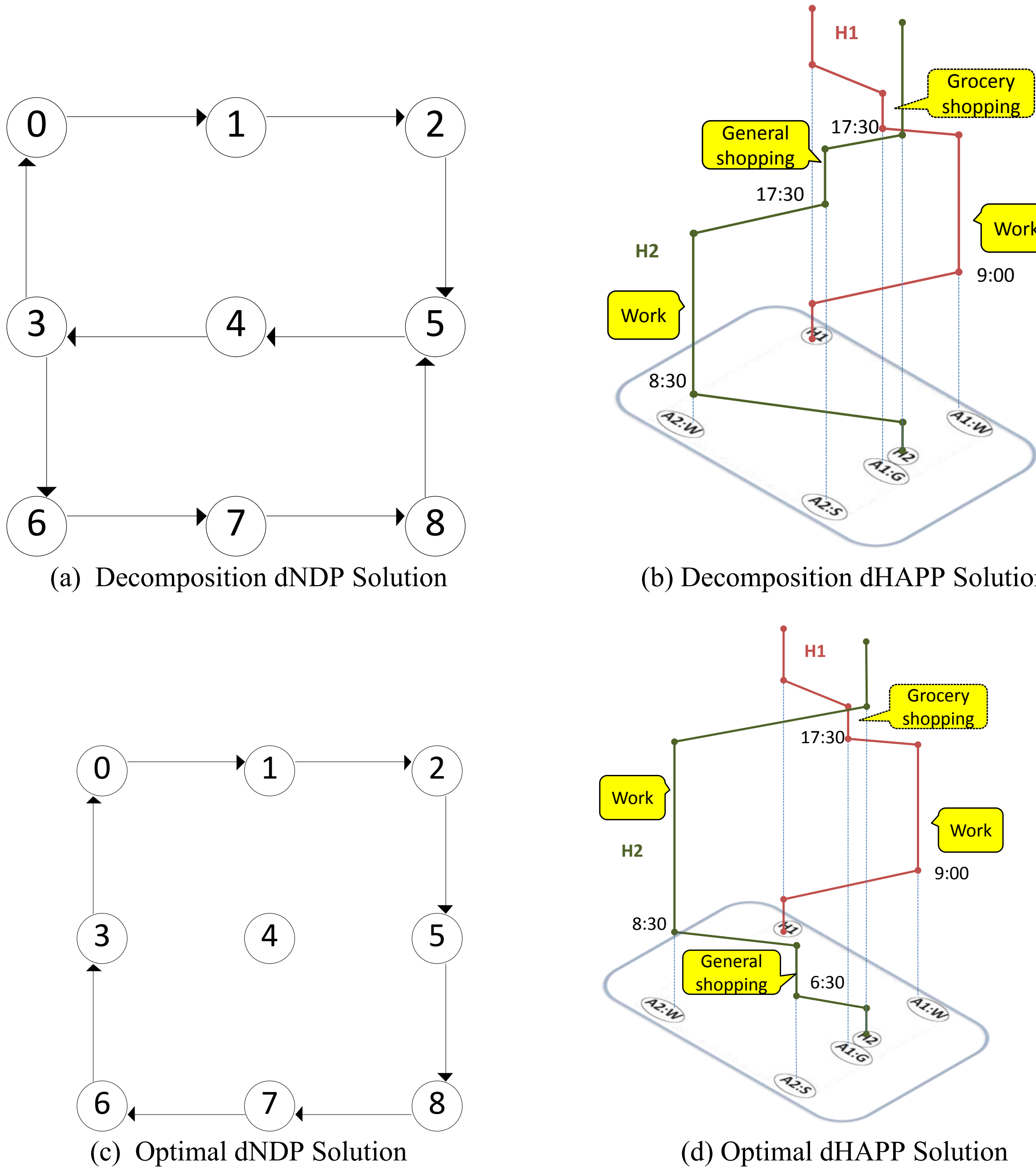


Figure 5-6 NDP-HAPP decomposition solution comparison to enumerated exact solution

Table 5-4 Detailed computational illustration of the basic NDP-HAPP example

	Iteration 1	Iteration 2
dHAPP1	Path[1]: Home (0) → work (2) → grocery shopping (5) → home (0) Objective Value: 3	Path: Home (0) → work (2) → grocery shopping (5) → home (0) Objective Value: 3
dHAPP2	Path[2]: Home (5) → work (6) → general shopping (8) → home (5) Objective Value: 3	Path: Home (5) → work (6) → general shopping (8) → home (5) Objective Value: 3
dNDP	Network Design Decisions: Z01, Z12, Z25, Z30, Z36, Z43, Z54, Z36, Z78, Z85 dNDP objective value: 36 HH1 Paths link Flows: • Home (0) → (1) → Work (2) • Work (2) → Grocery Shopping (5) • Grocery Shopping (5) → (4) → (3) → Home (0) HH2 Paths link Flows: • Home (5) → (4) → (3) → Work (6) • Work (6) → (7) → General Shopping (8) • General Shopping (8) → Home (5) Update each dHAPP objective values: HH1: 3 HH2: 3	NA[3]
Final Objective	42	42

[1]These paths are based on the assumption that all links are available.

[2]These paths are based on the assumption that all links are available.

[3]No changes in variables and objective function value. Therefore the algorithm stopped after this iteration.

5.4.2 Simple example: NDP-GHAPP

Using the generalized model allows us to include behavioral changes in destination choice as well as routing/scheduling of activities with respect to network design decisions. Following the example in Section 5.4.1, assume that there are two grocery shopping locations, node 1 and node 5, $P^{+}_{A_a=GroceryShopping}=\{1,5\}$, and two general shopping locations, node 3 and node 8, $P^{+}_{A_a=GeneralShopping}=\{3,8\}$ —each household is required to visit one, and only one, of the candidate locations to perform the shop activity.

Here, NDP-GHAPP optimality is checked in the same way as in the previous example, i.e., by comparing to the results of dNDP for all possible combinations of household decisions, including the destination choice as well as path sequence decisions and arrival time decisions to return. The solution from the iterative method reached the true optimal value after three iterations, shown in Figure 7. The intuition is that the flexibility introduced by NDP-GHAPP allows the method to search for many different options. Detailed illustration of the computational process of the proposed algorithm is shown in Table 5-5. In this simple example, changes in activity sequence, link level flow in dNDP, and dNDP network design decisions are shown, resulting in a *joint output* of infrastructure link investments, destination choices for each household, and schedule choices for each household.

Table 5-5 Detailed computational illustration of the NDP-GHAPP example

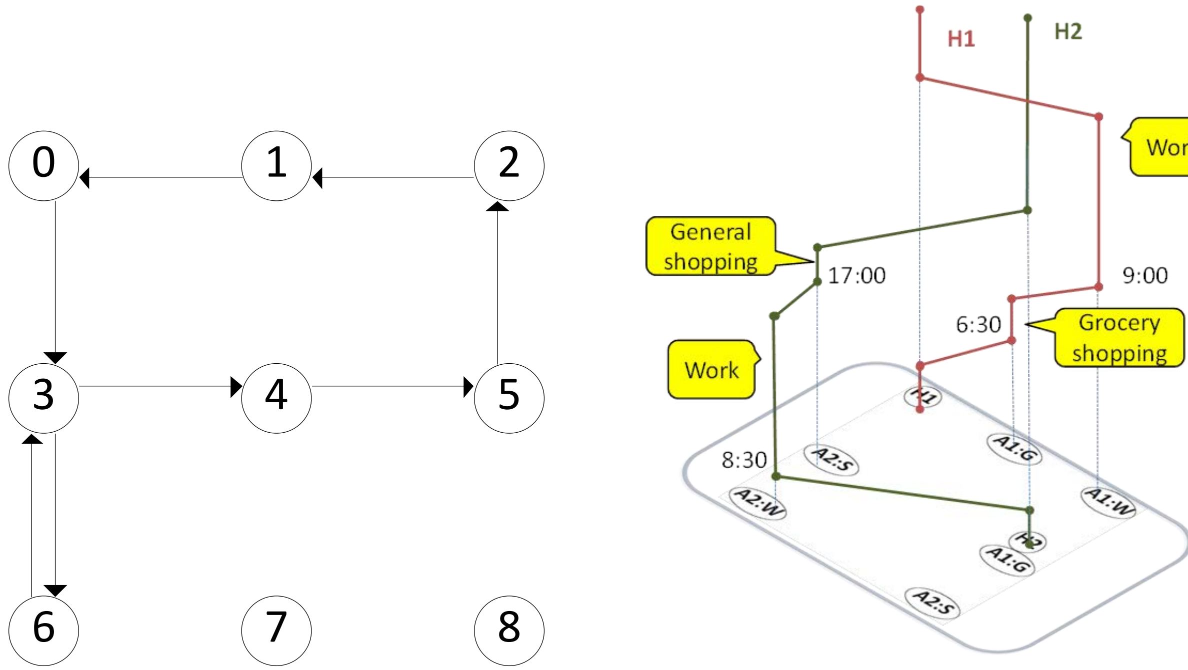
	Iteration 1	Iteration 2	Iteration 3	Iteration 4
dHAPP1	Path[1]: Home (0) → grocery shopping (1) → work (2) → home (0) Objective Value: 2	Path: Home (0) → work (2) → grocery shopping (1) → home (0) Objective Value: 2	Path: Home (0) → grocery shopping (5) → work (2) → home (0) Objective Value: 4	Path: Home (0) → grocery shopping (5) → work (2) → home (0) Objective Value: 3
dHAPP2	Path[2]: Home (5) → work (6) → general shopping (8) → home (5) Objective Value: 3	Path: Home (5) → work (6) → general shopping (8) → home (5) Objective Value: 3	Path: Home (5) → work (6) → general shopping (3) → home (5) Objective Value: 4	Path: Home (5) → work (6) → general shopping (3) → home (5) Objective Value: 4
dNDP	Network Design Decisions: Z01, Z10, Z12, Z21, Z58, Z67, Z76, Z78, Z85, Z87 dNDP objective value: 35 HH1 Paths link Flows: • Home (0) → Grocery Shopping (1) • Grocery Shopping (1) → Work (2) • Work (2) → (1) → Home (0) HH2 Paths link Flows: • Home (5) → (8) → (7) → Work (6) • Work (6) → (7) → General Shopping (8) • General Shopping (8) → Home (5) Update each dHAPP	Network Design Decisions: Z03, Z10, Z21, Z36, Z52, Z67, Z78, Z85 dNDP objective value: 32 HH1 Paths link Flows: • Home (0) → (3) → (6) → (7) → (8) → (5) → Work (2) • Work (2) →Grocery Shopping (1) • Grocery Shopping (1) → Home (0) HH2 Paths link Flows: • Home (5) → (2) → (1) → (0) → (3) → Work (6) • Work (6) → (7) → General Shopping (8) • General Shopping (8) → Home (5)	Network Design Decisions: Z03, Z10, Z21, Z34, Z36, Z45, Z52, Z63 dNDP objective value: 31 HH1 Paths link Flows: • Home (0) → (3) → (4) → Grocery Shopping (5) • Grocery Shopping (5) → Work (2) • Work (2) → (1) → Home (0) HH2 Paths link Flows: • Home (5) → (2) → (1) → (0) → (3) → Work (6) • Work (6) → General Shopping (3) • General Shopping (3) → (4) → Home (5) Update each dHAPP	NA[3]

	objective values: HH1: 2 HH2: 3	Update each dHAPP objective values: HH1: 4 HH2: 4	objective values: HH1: 3 HH2: 4	
Final Obj	40	40	38	38

[1]These paths are based on the assumption that all links are available.

[2]These paths are based on the assumption that all links are available.

[3]No changes in variables and objective function value. Therefore aborted after this iteration.



(a) Optimal Solution (b) Optimal Solution

Figure 5-7 NDP-GHAPP example enumerated optimal solution

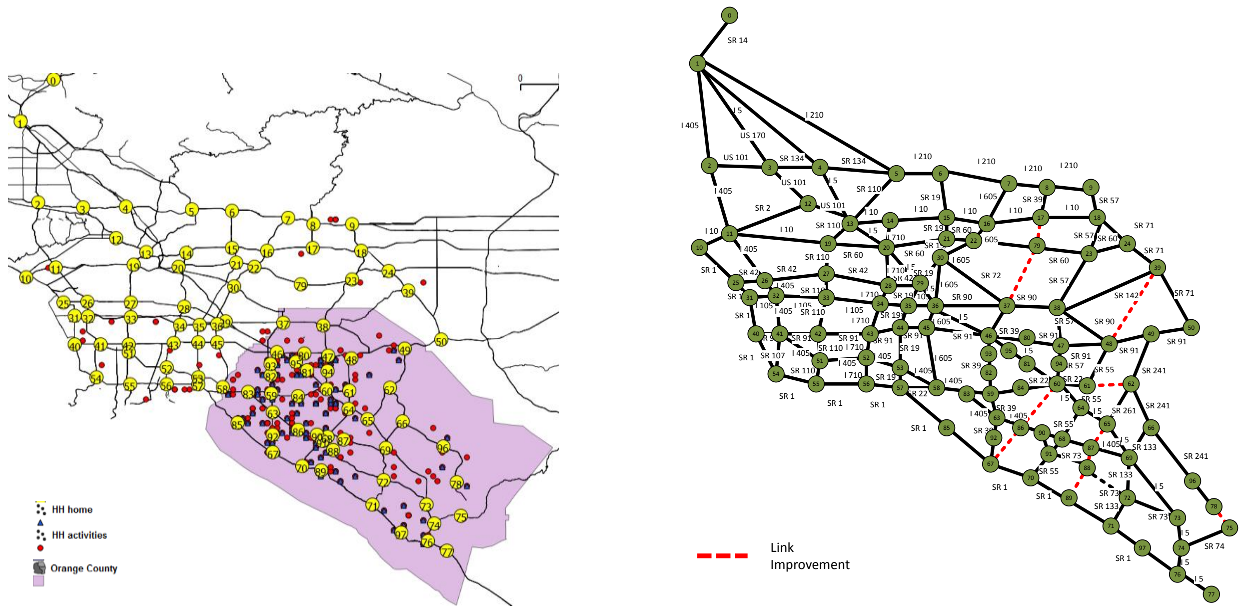
5.4.3 Large Network example: NDP-HAPP

This case study illustrates the performance of the NDP-HAPP solution algorithm for a major roadway system located in Orange County, California, a subsystem of the Los Angeles metropolitan roadway network. The base network with household locations and their activities throughout the day are shown in Figure 5-8-(a). We assume that the network design decision maker is a public agency from Orange County, and its goal is to provide the best mobility for Orange County residents, where the mobility is expressed in terms of total travel times. Hypothetically suggested candidate improvements on the network system are extensions of SR 39, SR 57, SR 55, SR 22, SR 261, and SR 241 as seen in dashed red lines in Figure 5-8-(b).

Specifications of each candidate link are in Appendix 5-B. Speeds are drawn from the average speed for all links on the same facility, and construction cost for each link is assumed to be proportional to both average speed and distance.

A sample of 60 single-member, single-vehicle households residing in Orange County drawn from the 2001 California Household Travel Survey (Caltrans, 2001) is used to reflect fairly realistic trip patterns of this class of households. The objective function for dNDP is to minimize the total travel time for the system, $\min \varphi_{dNDP}(z, f) = \sum_{(i,j)\in \mathbf{E}} t_{ij} \cdot f_{ji}$, and the objective function for each dHAPP is to minimize its own travel disutility. For this example, individual household's travel disutility is defined by the linear combination of the total extent of the day, the travel times, and the delay of return home caused by trip chaining multiple out-of-home activities using weights $\beta_h^E, \beta_h^T, \beta_h^D$:

$$\min \varphi_{dHAPP}(X,T) = \sum_{h\in\mathbf{H}} \sum_{v\in V^h} \beta_h^E \cdot (T_{2n_h+1}^{v,h} - T_0^{v,h}) + \sum_{h\in\mathbf{H}} \sum_{w\in\mathbf{P}_\mathbf{h}^+} \beta_h^D \cdot (T_{w+n_h}^h - T_w^h) + \sum_{h\in\mathbf{H}} \sum_{u\in\mathbf{Q}_h} \sum_{w\in\mathbf{Q}_h} \sum_{v\in V^h} \beta_h^T \cdot t_{uw}^{v,h} \cdot X_{uw}^{v,h}$$



(a) Sample households and their activity locations (b) Extracted network

Figure 5-8 Large-scale case study area

The weights of these 60 households are individually estimated from the inverse optimization calibration process in Chow and Recker (2012). For the households in the sample, the estimated results have average values of $\overline{\beta_h^E} = 0.84, \overline{\beta_h^D} = 0.74, \overline{\beta_h^T} = 3.45$, which means that on average these household decision makers value a minute of travel time savings about 4 times more than a minute of total extent of the day savings, and about 5 times more than a minute delay in returning home caused by trip chaining from out-of-home activities. The values were based on having the same set of arrival time penalties for all activity types, with 0.613 early penalty and 2.396 late penalty, similar to Chow and Recker (2012). The correlations from the 60 samples were close to zero for $\rho_{E,T}$ and $\rho_{D,T}$, although the correlation between extent of day savings and return home delay was $\rho_{E,D} = 0.248$. Time windows of activities are separately estimated using the methodology as in Chapters 2, 3 and 4, which adopted the method from Recker and Parimi (1999) with slight modifications.

In Table 5-6, results of NDP-HAPP are compared to conventional NDP solutions that take the O/D matrix derived from the optimal HAPP results with current network as an input. Six different budget limits are tested. The results indicate that both dNDP and dHAPP objective function values improved with increasing budget limits, together with more households benefiting from the improvements. These households experience shorter travel times, but given the coarse geographic network and the limited activity participation from the 60 single-member households sampled, these improvements are not sufficiently large for the sample households to change their activity sequences For example, 38% of the total trips are intrazonal trips; although we have assigned a nominal travel time for intrazonal trips, it is highly unlikely that one would change the sequence of trips in a way that shifts intrazonal trips to interzonal trips. Another explanation can be the activity participation. More than 70% of households in the sample performed only one out-of-home activity: these cases can never change activity sequences regardless of network level-of-service. The O/D table stays the same, and therefore the conventional NDP delivers what appears to be the same results as NDP-HAPP. However, a view of the time of day distribution of all activity participation reveals changes that can be captured as a result of the NDP-HAPP, as shown in Figure 5-9.

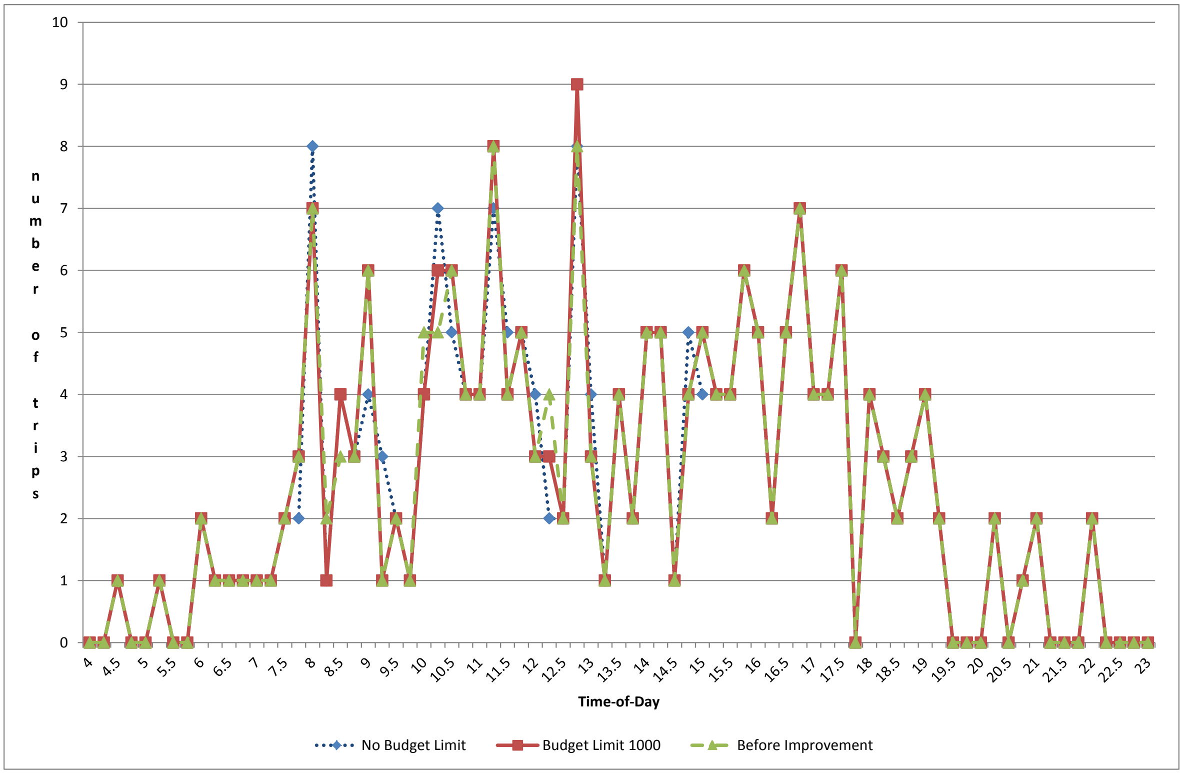


Figure 5-9 Comparison of activity arrival time histogram

As shown in Figure 5-9, the schedules of most households did not change much towards the evening, but shifts in arrival times can be seen as a consequence of changes in the network. There is a noticeable shift, particularly in the morning periods, as a result of the network improvements and the structure of the time windows defined for the households' activities. In other words, even though the total daily OD patterns did not change, the morning peak period OD patterns shifted in this simplest of examples drawn from a small sample of households. We can expect that as we evaluate larger samples of households with more members and interactions among them, and joint destination and schedule choices, the differences between NDP and NDP-HAPP would only be greater.

Table 5-6 Large-scale NDP-HAPP results

Budget	NDP-HAPP							Conventional NDP	
	# iter	Link Construction Decision	dNDP obj	dHAPP obj	# total trips (# intra zonal)	# HHs affec-ted	Time (sec)[23]	Link Construction Decision	NDP obj
Before Improvement	NA	NA	27.02	617	199 (76)	NA	NA	NA	27.02
1000	2	8988, 7875, 7578	25.99	610	199 (76)	5/60	306	8988, 7875, 7578	25.99
2000	2	8988, 7875, 7578, 7937, 8660, 6786, 8887	25.30	607	199 (76)	13/60	294	8988, 7875, 7578, 7937, 8660, 6786, 8887	25.30
3000	2	8988, 7875, 7578, 7937, 8660, 6786, 8887, 6086, 8667, 8889	24.88	605	199 (76)	14/60	326	8988, 7875, 7578, 7937, 8660, 6786, 8887, 6086, 8667, 8889	24.88
4000	1	8988, 7875, 7578, 7937, 8660, 6786, 8887, 6086, 8667, 8889, 6162, 6589,	24.79	604	199 (76)	17/60	196	8988, 7875, 7578, 7937, 8660, 6786, 8887, 6086, 8667, 8889, 6162, 6589,	24.79

[23] For calculating this computation time, we did not use any decomposition or heuristic method. All calculations are done by calling a CPLEX library directly from proposed formulations.

		8765, 8788						8765, 8788	
5000	1	8988, 7875, 7578, 7937, 8660, 6786, 8887, 6086, 8667, 8889, 6162, 6589, 8765, 8788, 6261	24.79	604	199 (76)	17/60	191	8988, 7875, 7578, 7937, 8660, 6786, 8887, 6086, 8667, 8889, 6162, 6589, 8765, 8788, 6261	24.79
No Limit	1	All	24.79	604	199 (76)	17/60	215	All	24.79

5.5 Conclusion

Given the arguments for considering activity behavior in transportation planning, it is logical to consider the applicability of activity scheduling in network design problems. Conventional NDPs studied previously focused on congestion issues, such as Braess' Paradox, and at best considered only supply side schedules. This research takes a step in advancing NDP theory where OD demand is not known *a priori*, but rather is the subject of responses in household itinerary choices that depend on the infrastructure improvements. Using simple examples, we show that falsely assuming that household itineraries are not elastic can result in a lack of understanding in certain phenomena; e.g., increasing traffic even without increasing economic activity due to relaxing of space-time prism constraints, or worsening of utility despite infrastructure investments in cases where household objectives may conflict.

An activity-based network design problem is proposed using the location routing problem as inspiration. The kernel problem is a bilevel formulation that includes an upper level network design and shortest path problem while the lower level includes a set of disaggregate household itinerary optimization problems, posed as HAPP (or in the case with location choice, as generalized HAPP) models. By using the simplest case HAPP model to represent the kernel problem, conclusions made with it can be extended to more complex variations. As a bilevel problem with an NP-hard lower level problem, there is no algorithm for solving the NDP-HAPP exactly. Nonetheless, the simple numerical examples demonstrate the sufficient accuracy of the decomposition heuristic algorithm derived from the LRP. The large numerical example based on Southern California data shows that the solution algorithm can handle medium-sized applications. The computation times were found to be reasonable considering the complexity of the problem posed. The setting also suggests that even if infrastructure investments do not result in major changes in itineraries (or any, in this particular example due to the small sample of simple 1-member households that do not have many activities in their itineraries), the results provide much higher resolution information to a decision maker. Whereas a conventional NDP would output the best set of links in which to invest given an assumed daily OD matrix, the NDP-HAPP can output the same best set of links, the same daily OD matrix, plus a detailed time-of-day temporal distribution of activity participation and travel that shows changes in OD patterns for peak periods of interest.

Beyond the most obvious extensions and future research applicable to this work (improved heuristics, adding uncertainty, dynamic policies, etc.), there are a number of important issues that need further study. Congestion effects certainly fall among the top of that list, and the interplay between congestion effects and schedule effects will be an interesting challenge to tackle. The kernel NDP-HAPP currently handles planning and tactical considerations, but expansions of the problem are needed to include such operational network design strategies as optimal toll pricing, ramp metering, or signal timing. There are actually two levels of congestion for consideration. The first is the effect on the infrastructure layer, which is what Lam and Yin (2001) or a dynamic traffic assignment integration could achieve. Congested links in the upper level problem would result in multiple paths between each pair of nodes, which means some weighting of travel times is needed to translate over a single perceived travel time matrix for the lower level household scheduling problems. We suspect that although incorporating congestion effects in HAPP should be straightforward by using feedback loops with a connected DTA model, the consideration of congestion along with demand scheduling within an NDP framework will be not be so simple. Recent advances in activity-based travel simulations can at least provide a convenient platform to generate synthetic data for a whole population, which is required for integration of network design problems with both congestion and demand.

The other congestion effect is at the activity layer, and more generally speaking refers to both negative (congestion) and positive (bandwagon) effects. For example, the time-dependent utility of some activities may depend heavily on how popular they are with multiple individuals. Another effect that can be incorporated is the link/node capacity in the upper level problem. Since only the shortest path between all nodes is being allocated to the households, adding capacity would require some weighted average path travel times similar to the link congestion effect. Also, by adding node capacity, we can impose upper limits of each facility location which ultimately leads to utility changes of activity participation at a given time segment.

We believe that there is a tremendous opportunity to apply activity-based NDPs to networks where demand scheduling is a more important consideration than congestion effects: one obvious application is multimodal transport system network design. Because there are only supply-side scheduling considerations in the state-of-the-art NDPs for multimodal systems, each system can only be evaluated on its own. However, a unified demand scheduling platform is needed so that alternative systems and their effects on households' schedules can be evaluated in a NDP framework. Due to the lack of such a unified treatment using demand scheduling, many modern multimodal transport systems (park-and-ride, carsharing, demand responsive transit, parking pricing, etc.) cannot be evaluated by public planning agencies—the historical attention on only congestion has left out the importance of demand scheduling which is more critical for developing these modern systems.

Another important consideration is the number of new types of NDPs that can benefit from having activity or itinerary response, not just from transportation planning perspective. For example, private firms can integrate supply chain demand networks with their facility networks in a similar fashion so that supply chain patterns (modeled as activity schedules) can be output along with facility investment decisions in an integrated fashion.

CHAPTER 6 CONCLUSION

This dissertation proposes various methods to integrate the travel demand procedure with transportation infrastructure from both the demand and supply sides. The particular travel demand model we use in this dissertation is the HAPP problem. Its structure of mathematical programming provides a convenient platform for integration with network problems since network problems are often formulated as mathematical programming. Then these models are applied to transportation sustainability regarding the adoption of AFVs.

6.1 SUMMARY OF CONTRIBUTIONS

The LSP-HAPP is developed in Chapter 2. This model adds destination choice to the HAPP, which originally made simultaneous travel decisions of activity allocation between household members, sequence of activities, departure time, and mode choice. Destination is not an independent decision but rather a complex decision made from interactions with other travel decisions such as accessibility, time-of-day, trip-chaining, or travel mode. An exact dynamic programming method is adopted to solve LSP-HAPP. When applied to a real data set, the results generated reasonable destination and activity start time predictions. Computational time is also found to be superior to direct calculation. An application of using LSP-HAPP as a synthetic full-day travel pattern is also presented. This model can work as a bridge to link between travel demand models with transportation network, more generally, with infrastructure as seen in the later chapters.

In Chapter 3, HAPPR and HAPPC are developed to represent the best usages of HFCVs and BEVs, respectively. For the worst usages of HFCVs and BEVs, refueling trip insertion analysis and charging delay analysis are used. This analysis derives several useful policy implications. This kind of upper and lower bound analysis can also be utilized for other circumstances with no apparent data such as infrastructure service level changes, vehicle (or personal) constraint changes, policy changes, or fuel cost changes. This is possible due to the HAPP structure that it formulates the physical spatial and temporal constraints. We can impose the circumstantial constraints in addition to the original spatial and temporal constraints.

In Chapter 4, a facility location problem with full-day scheduling and routing considerations (Location-HAPP) is formulated as a Location-Routing Problem (LRP). Column generation technique is adapted to solve this problem. The sub-problem of finding an entering column vector is found via a new search method we developed to ease the computation of calculating many np-hard HAPP problems. Although this formulation can be applied to services that require coverage, we specify its application to hydrogen refueling station siting in this dissertation. From the perspectives of refueling station siting works, it introduces a tour-based approach with tour construction capability. The results and the coverage matrices indicate that excluding vehicle-infrastructure interactions as well as routing and scheduling interactions may over-estimation of the minimum number of refueling stations to guarantee the same level of accessibility.

In Chapter 5, a generalized framework for travel demand models and network problems is proposed. OD demand is assumed not known *a priori*, but rather is the subject of responses in household travel decisions that depend on the network level of service. A solution method that decomposes the problem into each decision maker's rationale is developed and applied to small examples and a case study. Computational time is reasonable given the complexity of the problem, and the error rate is between 0 – 5% when tested with examples with known optimal solutions. Although the particular travel demand model we use is HAPP, and the particular network problem is NDP (NDP-HAPP), these sub-problems can be substituted with other travel demand and network models.

6.2 FUTURE RESEARCH

One important feature of activity-based travel demand models is that it can be used as a micro-simulation tool for the whole population in the study area. In this dissertation, we only used a sample data set from the California Travel Survey, and the applications simply utilize the reported data set. Activity-based travel demand model needs to be able to synthesize the feasible activity/travel patterns to represent the whole population. Not only would this approach output better forecasting results, but the data set size of real demand leads us to evaluate various infrastructure capacities. For example, we can test service quality (waiting time) via queuing at early stage refueling stations as mentioned in Location-HAPP. This can also be tied to the congestion effect mentioned in NDP-HAPP to address the interplay between congestion effects and schedule effects.

The HAPP needs many further methodological and practical developments to be functional as a simulation tool, such as CMEM (Bhat et al., 2004), ALBATROSS (Arentze and Timmermans, 2004), or TASHA (Miller and Roorda, 2003). As noted throughout this dissertation, HAPP formulates the physical constraints of time and space dimensions and therefore can provide feasible synthetic patterns of certain phenomena with no preceding data or observations. Also, its unique structure can be used to tackle some transportation issues that cannot be explained by other types of models. Major developments that need to be made is, first, the full integration of mode choice. Although the original HAPP includes some level of mode choice – vehicles owned by the household, along with carpooling decisions, it needs to be able to include other types of travel modes such as public transit, walking and biking. Second, activity participation decisions need to be endogenous. Some level of activity participation is addressed in Gan and Recker (2012), but activity duration decisions need to be made within the model. Third, more efficient computational methods need to be explored since computation has been one of the issues regarding the practicality of HAPP. The primary goal of activity-based travel demand models is travel demand forecasting to represent the population, and therefore their process times need to be highly efficient.

Extensions and future research on network problems in this dissertation include adding uncertainty and multi-period approach. We have presented the problems with rather simple network problems to focus on the presentation of the concept, but it is unrealistic that network decisions are made for one period with deterministic parameters. For the extension of more sophisticated network problems, we can utilize various ideas, concepts, and heuristics from the rich literature in this area.

BIBLIOGRAPHY

Arentze, T. and Timmermans, H. (2004). A learning-based transportation oriented simulation system, Transportation Research Part B, 38, 613-633

Arentze, T. and Timmermans, H. (2004). Multistate supernetwork approach to modeling multi-activity, multimodal trip chains, International Journal of Geographical Information Science, 18(7), 631-651

Axsen, J., Kurani, K. S., McCarthy, R., and Yang, C. (2011). Plug-in hybrid vehicle GHG impacts in California: Integrating consumer-informed recharge profiles with an electricity-dispatch model, Energy Policy, 39-3, 1617-1629

Balmer, M., Axhausen, K.W., and Nagel, K. (2006). Agent-based demand-modeling framework for large-scale microsimulations, Transportation Research Record: Journal of the Transportation Research Board, 1985, 125-134

Barnhart, C., Johnson, E. L., Nemhauser, G. L., Savelsbergh, M. W. P., Vance, P. H. (1998). Branch-and-price: column generation for solving huge integer programs, Operations Research, 46, 316-329

Beardwood, J., Halton, J. H. and Hammersley, J. M. (1959). The shortest path through many points, Mathematical Proceedings of the Cambridge Philosophical Society, 55, 299 – 327

Bhat, C.R., Guo, J.Y., Srinivasan, S., and Sivakumar, A. (2004). Comprehensive econometric microsimulator for daily activity-travel. Transportation Research Record: Journal of the Transportation Research Board, 1894, 57-66

Bhat, C. R., Govindarajan, A., Pulugurta, V., 1998. Disaggregate attraction-end modeling: formulation and empirical analysis. Transportation Research Record: Journal of the Transportation Research Board, 1654 (1), 60-68

Bowman, J.L. and Ben-Akiva., M.E. (2000). Activity-based disaggregate travel demand model system with activity schedules, Transportation Research Part A, 35(1), 1-28

Brownstone, D., Bunch, D. S., Golob, T. F., and Ren, W. (1996). Transaction choice models for forecasting demand for alternative fuel vehicles, Research in Transportation Economics, 4, 87-129

Bruns, A., Klose, A. and Stähly, P. (2002). Restructuring of Swiss parcel delivery services, OR Spektrum, 22, 285-302

Bunch, D. S, Bradley, M., Golob, T. F., Kitamura, R., and Occhiuzzo, G. P. (1993). Demand for clean-fuel vehicles in California: A discrete-choice stated preference pilot project, Transportation Research Part A, 27-3, 237-253

Byrne, M. R. and Polosky, M. J. (2001). Impediments to consumer adoption of sustainable transportation: alternative fuel vehicles, International Journal of Operations & Production Management, 21-12, 1521-1538

Calfee, J. E. (1985). Estimating the demand for electric automobiles using fully disaggregated probabilistic choice analysis, Transportation Research, 19B, 287-302

California Fuel Cell Partnership (CaFCP) (2009). Hydrogen fuel cell vehicle and station deployment plan: A strategy for meeting the challenge ahead, Action plan, West Sacramento, CA

California Fuel Cell Partnership (CaFCP) (2010). Hydrogen fuel cell vehicle and station deployment plan: A strategy for meeting the challenge ahead, Progress and next steps, West Sacramento, CA

California Fuel Cell Partnership (CaFCP) (2012). Hydrogen fuel cell vehicle and station deployment Plan: A California road map: The commercialization of hydrogen fuel cell vehicles, West Sacramento, CA

CalTrans (2001). 2000-2001 California Statewide Household Travel Survey

Caprara, A., Toth, P., Fischetti, M. (2000). Algorithms for the set covering problem, Annals of Operations Research, 98, 353-371

Chan, Y., Carter, W. B., Burnes, M. D. (2001). A multiple-depot, multiple-vehicle, location-routing problem with stochastically processed demands, Computers and Operations Research, 28, 803-826

Chow, J. Y. J. and Liu, H. (2012). Generalized profitable tour problems for an online activity routing system, Transportation Research Record, 2284, 1-9.

Chow, J. Y. J. and Recker, W. W., (2011). Inverse optimization with endogenous arrival time constraints to calibrate the household activity pattern problem. Transportation Research Part B, 46 (3), 463-479

Chow, J. Y. J. (2013). Activity-based travel scenario analysis with routing problem reoptimization, Computer-Aided Civil and Infrastructure Engineering, in press, doi: 10.1111/mice.12023

Christofides, N. and Korman, S. (1975). A computational survey of methods for the set covering problem, Management Science, 21, 591-599

Cordeau, J. F. and Laporte, G. (2003). The dial-a-ride problem: Variants, modeling issues and algorithms, 4OR: A Quarterly Journal of Operations Research, 1-2, 89-101

Crainic, T.G. (2000). Service network design in freight transportation. European Journal of Operational Research, 122(2), 272-288

Dagsvik, J. K., Wennemo, T., Wetterwald, D. G., and Aaberge, R. (2002). Potential demand for alternative fuel vehicles, Transportation Research Part B, 36-4, 361-384

Desrochers, M. Desrosiers, J., and Solomon, M. (1992). A new optimization algorithm for the vehicle routing problem with time windows, Operations Research, 40 (2), 342-354

Desrosiers, J., Dumas, Y., Soumis, F. (1986). A dynamic programming solution of the large-scale single-vehicle dial-a-ride problem with time windows, American Journal of Mathematical and Management Sciences, 6, 301-325

Desrosiers, J., Soumis, F., and Desrochers, M. (1984). Routing with time windows by column generation, Networks, 14 (4), 545-565

Dong, A. and Lin, Z. (2012). Within-day recharge of plug-in hybrid electric vehicles: energy impact of public charging infrastructure, Transportation Research Part D, 17-5, 405-412

Dumas, Y., Desrosiers, J., Soumis, F. (1991). The pickup and delivery problem with time windows. European Journal of Operational Research, 54 (1), 7-22

Ewing, G. and Sarigollu, E. (2000). Assessing consumer preferences for clean-fuel vehicles: A discrete choice experiment, Journal of Public Policy & Marketing, 19-1, 106-118

Fischetti, M., González, J.J.S., and Toth, P. (1997). A branch-and-cut algorithm for the symmetric generalized traveling salesman problem, Operations Research, 45(3), 378-394.

Ford, L. R. and Fulkerson, D. R. (1958). A suggested computation for maximal multi-commodity network flows. Management Science, 5 (1), 97-101

Fotheringham, A. S. (1988). Consumer store choice and choice set definition. Marketing Science, 7 (3), 299-310

Friesz, T. L. and Harker, P. T. (1985). Properties of the iterative optimization-equilibrium algorithm, Civil Engineering Systems, 2(3), 142-154

Fu, X. and Lam, W. H. K. (2013) A network equilibrium approach for modelling activity-travel pattern scheduling problems in multi-modal transit networks with uncertainty. Transportation, in press, doi: 10.1007/s11116-013-9470-9

Gan, L. P. and Recker, W. W. (2008). A mathematical programming formulation of the household activity rescheduling problem, Transportation Research Part B 42(6), 571-606

Gan, L. P. and Recker, W.W. (2012). Stochastic pre-planned household activity pattern problem with uncertain activity participation (SHAPP). Transportation Science, In press.

Gärling, T. and Axhausen, K. W. (2003), Introduction: habitual travel choice. Transportation, 30 (1), 1-11

Geoffrion, A. M. and Graves, G. W. (1974). Multicommodity distribution system design by Benders decomposition, Management Science 20(5), 822-844

Ghiani, G. and Improta, G. (2000), An efficient transformation of the generalized vehicle routing problem, European Journal of Operational Research, 122(1), 11-17

Gondor, J., Marke, T., Simpson, A., and Thornton, M. (2007). Using GPS Travel Data to Assess the Real World Driving Energy Use of Plug-in Hybrid Electric Vehicles, National Renewable Energy Laboratory Conference Paper, NREL/CP-540-40858

Hagerstrand, T. (1970). What about people in regional science?, Papers in Regional Science, 24-1, 7-24

Hamdouch, Y. and Lawphongpanich, S. (2008). Schedule-based transit assignment model with travel strategies and capacity constraints, Transportation Research Part B, 42(7-8), 663-684

Hamdouch, Y. and Lawphongpanich, S. (2010). Congestion pricing for schedule-based transit networks, Transportation Science, 44(3), 350-366

Heydecker, B. (2001). Dynamic equilibrium network design. In: Transportation and Traffic Theory in the 21st Century, ed. Taylor, M.A.P., Oxford: Pergamon Press, 349-370

Hodgson, M. J. (1990). A flow capturing location-allocation model, Geographical Analysis, 22, 270–279

Joksimovic, D., Bliemer, M. C. J., Bovy, P. H. L. and Verwater-Lukszo, Z. (2005). Dynamic road pricing for optimizing network performance with heterogeneous users. In: Proc. IEEE Networking, Sensing and Control, 407-412

Kang, J. E., Brown, T. M, Recker, W. W., and Samuelsen, G. S. (2013), "Refueling Hydrogen Fuel Cell Vehicles with 68 proposed refueling stations in California: measuring deviations from daily travel patterns", Under Review

Kang, J. E. and Recker, W. W. (2009). An activity-based assessment of the potential impacts of plug-in hybrid electric vehicles on energy and emissions using 1-day travel data, Transportation Research Part D: Transport and Environment, 14, 541–556

Kelly S., and Kuby M. (2012). Spatial refueling patterns of compressed natural gas (CNG) drivers in Southern California. In: Informatics rising: INFORMS 2012 annual meeting, Phoenix, USA, 14-17 October; 2012; 195

Kim, J-G. and Kuby, M. (2012). The deviation-flow refueling location model for optimizing a network of refueling stations, International Journal of Hydrogen Energy, 37, 5406-5420

Kitamura, R. (1984). Incorporating trip chaining into analysis of destination choice, Transportation Research Part B, 18 (1), 67-81

Kitamura, R. and Sperling, D. (1987). Refueling behavior of automobile drivers, Transportation Research Part A: Policy and Practice, 3, 235-245

Konduri, K.C. (2012). Integrated model of the urban continuum with dynamic time-dependent activity-travel microsimulation: framework, prototype, and implementation. Ph.D. dissertation, Arizona State University

Kuby, M. and Lim, S. (2005). The flow-refueling location problem for alternative-fuel vehicles, Socio-Economic Planning Sciences, 39, 125-148

Kuby, M., Lines, L., Schultz, R., Xie, Z., Kim, J.K. and Lim, S. (2009). Optimization of hydrogen stations in Florida using the flow-refueling location model, International Journal of Hydrogen Energy, 34, 6045-6064

Lam, W.H.K. and Yin, Y. (2001). An activity-based time-dependent traffic assignment model, Transportation Research Part B, 35(6), 549-574

Laporte, G., and Dejax, P. J. (1989). Dynamic location-routing problems. Journal of Operational Research Society, 40 (5), 471-482

List, G. F., Mirchandani, P. B., Turnquist, M. A. and Zografos, K. G. (1991). Modeling and analysis for hazardous materials transportation: risk analysis, routing/scheduling and facility location, Transportation Science, 25, 100-114

Li, H., Bliemer, M. C. J., and Bovy, P. H. L. (2007). Optimal toll design from reliability perspective. In: Proc. 6th Triennnial Conference on Transportation Analysis, Phuket, Thailand

Lin, D. Y., Eluru, N., Waller, S. T. and Bhat, C. R. (2008). Integration of activity-based modeling and dynamic traffic assignment, Transportation Research Record: Journal of the Transportation Research Board, 2076, 52-61

Liu, S. C. and Lee, S. B. (2003). A two-phase heuristic method for the multi-depot location routing problem taking inventory control decisions into considerations. International Journal of Advanced Manufacturing Technology, 22, 941-980

Louviere, J. J., and Timmermans, H. J. P., 1990. Using hierarchical information integration to model consumer responses to possible planning actions: recreation destination choice illustration. Environment and Planning A, 22 (3), pp. 291-308

Luathep, P., Sumalee, A., Lam, W. H. K., Li, Z. C., and Lo, H. K. (2011). Global optimization method for mixed transportation network design problem: a mixed-integer linear programming approach, Transportation Research Part B, 45(5), 808-827

Magnanti, T. L. and Wong, R. T. (1984). Network design and transportation planning: models and algorithms, Transportation Science, 18(1), 1-55

McNally, M. G. (1997). An Activity-Based Micro-Simulation Model for Travel Demand Forecasting. Activity-Based Approaches to Travel Analysis, Ettema., D. F. and Timmermans, H. J. P., 37-54.

Melachovsky, J., Prins, C. and Wolfler Calvo, R. (2005). A metaheuristic to solve a location-routing problem with non-linear costs. Journal of Heuristics, 11, 375-391

Melaina, M. (2003). Initiating hydrogen infrastructures: preliminary analysis of a sufficient number of initial hydrogen stations in the US, International Journal of Hydrogen Energy, 28-7, 743-755

Miller, E. J. and Roorda, M. J. (2001), Prototype models of household activity-travel scheduling, Transportation Research Record: Journal of the Transportation Research Board, 1831, 114-121

Min, H., Jayaraman, V. and Srivastava, R. (1998). Combined location-routing problems: A synthesis and future research directions, European Journal of Operations Research, 108-1, 1-15

MirHassani S. A. and Ebrazi, R. (2012). A flexible reformulation of the refueling station location problem, Transportation Science, doi: 10.1287/trsc.1120.0430

Nagy, G. and Salhi, S. (1996). Nested heuristic methods for the location-routeing problem, Journal of Operations Research, 39, 150-156

Nagy, G. and Salhi, S. (1998). The many-to-many location-routing problem, TOP, 6, 261-275

Nagy, G. and Salhi, S. (2007). Location-routing: issues, models and methods, European Journal of Operations Research, 177-25, 649-672

National Conference on Stage Legistlatures (NCSL), accessed in September, 2012, http://www.ncsl.org/issues-research/energyhome/state-electric-vehicle-incentives-state-chart.aspx

National Fuel Cell Research Center (NFCRC), University of California, Irvine (accessed in 2012). http://www.nfcrc.uci.edu/2/ACTIVITIES/PROJECTS/hydrogen/HydrogenFuelingStations.aspx#top

Nicholas, M. A. and Handy, S., and Sperling, D. (2004). Using geographic information systems to evaluate siting and networks of hydrogen stations, Transportation Research Record, 1880, 126-134

Nicholas, M. A. and Ogden, J. (2006). Detailed analysis of urban station siting for California hydrogen highway network, Transportation Research Record, 1883, 121-128

Or, I. and Pierskalla, W. P. (1979). A transportation location-allocation model for regional blood banking, AIIE Transactions, 11, 86-94

Perl, J. and Daskin, M. S. (1985). A warehouse location-routing problem, Transportation Research Part B, 19(5), 381-396

Parry, I. W. P., Wallas, M., and Harrington, W. (2007). Automobile externalities and policies, Journal of Economic Literature, 45-2, 373-399

Pinjari, A.R. and Bhat, C.R. (2011). Activity-based travel demand analysis. In: Handbook of Transport Economics, eds. A. de Palma, R. Lindsey, E. Quinet, R. Vickerman, Edward Elgar Publishing.

Pozsgay, M. A. and Bhat, C. R. (2001). Destination choice modeling for home-based recreational trips: analysis and implications for land use, transportation, and air quality planning. Transportation Research Record: Journal of the Transportation Research Board, 1777 (1), 47-54

Poon, M.H., Wong, S. C. and Tong, C. O. (2004). A dynamic schedule-based model for congested transit networks, Transportation Research Part B, 38(4), 343-368

Recker, W.W., Chen, C. and McNally, M. G., 2001. Measuring the impact of efficient household travel decisions on potential travel time savings and accessibility gains. Transportation Research Part A, 35 (4), 339-369

Recker, W. W. and Kostyniuk, L. (1978). Factors influencing destination choice for the urban grocery shopping trips. Transportation, 7 (1), 19-33

Recker, W.W. (1995). The household activity pattern problem: general formulation and solution, Transportation Research, 29B, 1, 61-77

Recker, W.W. (2001). A bridge between travel demand modeling and activity-based travel analysis. Transportation Research, 35B, 481-506

Recker, W. W., Duan, J. and Wang, H., (2008). Development of an estimation procedure for an activity-based travel demand model. Computer-Aided Civil and Infrastructure Engineering, 23 (7), 483-501

Recker, W. W and Parimi, A. (1999). Development of a microscopic activity-based framework for analyzing the potential impacts of transportation control measures on vehicle emissions, Transportation Research Part D: Transport and Environment, 4, 357-378

Roy, J. R. and Thill, J. (2004). Spatial interaction modelling. Regional Science, 83 (1), 339-361

Salhi. S., and Rand, G. K. (1989). The effect of ignoring routes when locating depots, European Journal of Operations Research, 39, 150-156

Schroeder, A. and Traber, T. (2012). The economics of fast charging infrastructure for electric vehicles, Energy Policy, 43, 136-144

Schwoon, M. (2007). A tool to optimize the initial distribution of hydrogen filling stations, Transportation Research Part D, 12-2, 70-82

Solomon, M. and Desrosiers, J., 1988. Time window constrained routing and scheduling problems. Transportation Science, 22 (1), 1-13

Stephan, C. and Sullivan, J. (2004a). An agent-based hydrogen vehicle/infrastructure model. IEEE, Proceedings of the 2004 Congress on Evolutionary Computation, Portland, OR

Stephan, C. and Sullivan, J. (2004b). Growth of a hydrogen transportation infrastructure. In Proceedings of the Agent 2004 Conference on Social Dynamics: Interaction, Reflexivity and Emergence, Chicago, IL, University of Chicago and Argonne National Laboratory.

Stephens-Romero, S. D., Brown, T. M., Carreras-Sospedra, M., Kang, J. E., Brouwer, J., Dabdub, D., Recker, W. W., and Samuelsen, G. S. (2011). Projecting full build-out environmental impacts and roll-out strategies associated with viable hydrogen fueling infrastructure strategies, International Journal of Hydrogen Energy, 36, 14309-14323

Stephens-Romero, S. D., Brown, T. M., Kang, J. E., Recker, W. W., and Samuelsen, G. S. (2010). systematic planning to optimize investments in hydrogen infrastructure deployment, International Journal of Hydrogen Energy, 35-10, 4652-4667

Struben, J. and Sterman, J. D. (2008). Transition challenges for alternative fuel vehicle and transportation systems, Environment and Planning. C, Government & Policy, 35-6, 1070

Thill, J., 1992. Choice set formation for destination choice modelling. Progress in Human Geography, 16 (3), 631-382

Tong, C. O. and Wong, S. C. (1998). A stochastic transit assignment model using a dynamic schedule-based network, Transportation Research Part B, 33(2), 107-121

TRANSIMS (accessed in 2012), TRansportation ANalysis SIMulation System, http://code.google.com/p/transims/

Ukkusuri, S. V. and Waller, S. T. (2008). Linear programming models for the user and system optimal dynamic network design problem: formulations, comparisons and extensions, Networks and Spatial Economics, 8(4), 383-406

Waller, S. T. and Ziliaskopoulos, A. K. (2001). Stochastic dynamic network design problem, *Transportation Research Record: Journal of the Transportation Research Board*, 1771, 106-113

Wang, S., Meng, Q., and Yang, H. (2013). Global optimization methods for the discrete network design problem, *Transportation Research Part B*, 50, 42-60

Wang, Y. W. and Lin, C. C. (2009). Locating road-vehicle refueling stations, *Transportation Research Part E: Logistics and Transportation*, 45, 821-829

Wu, T. H., Low, C. and Bai, J. W. (2002). Heuristic Solutions to multi-depot location-routing problems, *Computers and Operations Research*, 29(10), 1393-1415

Xi, X., Sioshansi, R., and Marano, V. (accessed in 2013). A simulation-optimization model for locating of a public electric vehicle charging infrastructure, *http://ise.osu.edu/isefaculty/sioshansi/papers/charge_infra.pdf*

Yang, H. and Bell, M. G. H. (1998). Models and algorithms for road network design: a review and some new development, *Transport Reviews*, 18(3), 257-278

Zhang, L., Brown, T. M., and Samuelsen, G. S. (2011). Fuel reduction and electricity consumption impact of different charging scenarios for plug-in electric vehicles, *Journal of Power Sources*, 196-15, 6559-6566

APPENDIX 2-A LSP-HAPP NOTATION

The following notations (extended from those in Recker, 1995) are used in the formulation:

$\boldsymbol{A} = \{\boldsymbol{A}_1, \dots, \boldsymbol{A}_a, \dots, \boldsymbol{A}_m\}$: the set of m different activity types with unspecified locations that the household needs to complete in a given day. The household needs to choose one, and only one, location from among many candidate locations for each activity in this set.

n_{A_a} : the number of alternative locations for conducting activity type A_a

$\boldsymbol{M_P} = \{1, 2, \dots, i, \dots, n_P\}$: the set of those out-of-home activities, each with a single "predetermined" location, to be completed by travelers in the household.

$\boldsymbol{M} = \boldsymbol{M_P} \cup \boldsymbol{A}$: the combined set of all out-of-home activities scheduled for completion by the household

$\boldsymbol{P_P^+} = \{\boldsymbol{1}, \boldsymbol{2}, \dots, \boldsymbol{n_P}\}$: the set designating the respective locations at which activities with predetermined locations are performed.

$\boldsymbol{P_{A_1}^+} = \{\boldsymbol{n_P} + \boldsymbol{1}, \boldsymbol{n_P} + \boldsymbol{2}, \dots, \boldsymbol{n_P} + \boldsymbol{n_{A_1}}\}$: the set designating "potential" locations at which activity A_1 may be performed—one, and only one, may be selected.

$\boldsymbol{P_{A_a}^+} = \left\{\boldsymbol{n_P} + \sum_{a=1}^{a-1} \boldsymbol{n_{A_a}} + \boldsymbol{1}, \boldsymbol{n_P} + \sum_{a=1}^{a-1} \boldsymbol{n_{A_a}} + \boldsymbol{2}, \dots, \boldsymbol{n_P} + \sum_{a=1}^{a} \boldsymbol{n_{A_a}}\right\}$, $\boldsymbol{A_a} \in \boldsymbol{A}$, $\boldsymbol{A_a} \neq \boldsymbol{A_1}$: the set designating "potential" locations at which activity A_a may be performed—one, and only one, may be selected for each activity $A_a \in A,\ A_a \neq A_1$.

$\boldsymbol{P_A^+ = \bigcup_{A_a \in A} P_{A_a}^+ = \{n_P + 1, n_P + 2, \ldots, n_P + n_{A_1}, n_P + n_{A_1} + 1, \ldots, n_P + \sum_{a=1}^{m} n_{A_a} = n\}}$: the set designating "potential" locations at which all activities with unspecified locations, $\boldsymbol{A}_a \in \boldsymbol{A}$, may be performed.

$\boldsymbol{P^+ = P_P^+ \cup P_A^+ = P_P^+ \cup P_{A_1}^+ \cup \ldots \cup P_{A_m}^+ = \{1, 2, \ldots, i, \ldots, n\}}$: the set designating locations at which the combined sets of activities with predetermined and multiple candidate locations may be performed.

$\boldsymbol{P_P^- = \{n + 1, n + 2, \ldots, n + n_P\}}$: the set designating the ultimate destinations of the "return to home" trips for activities with predetermined locations. (It is noted that the physical location of each element of $\boldsymbol{P_P^-}$ is "home".)

$\boldsymbol{P_{A_1}^- = \{n + n_P + 1, n + n_P + 2, \ldots, n + n_P + n_{A_1}\}}$: the set designating the ultimate destinations of the "return to home" trips for the $\boldsymbol{A_1}$ activity—each element is paired to the location selected for activity $\boldsymbol{A_1}$**.** (It is noted that the physical location of each element of $\boldsymbol{P_{A_1}^-}$ is "home".)

$\boldsymbol{P_{A_a}^- = \left\{n + n_P + \sum_{a=1}^{a-1} n_{A_a} + 1, n + n_P + \sum_{a=1}^{a-1} n_{A_a} + 2, \ldots, n + n_P + \sum_{a=1}^{a} n_{A_a}\right\}\ A_a \in A,}$ $\boldsymbol{A_a \neq A_1}$**:** set designating ultimate destination of the "return to home" trip for each activity $\boldsymbol{A_a}$—each element is paired to the location selected for each activity $\boldsymbol{A_a \in A,\ A_a \neq A_1}$**.** (Note that the physical location of each element of $\boldsymbol{P_{A_a}^-}$ is "home".)

$\boldsymbol{P_A^- = \bigcup_{A_a \in A} P_{A_a}^- = \{n + 1, n + 2, \ldots, n + i, \ldots, n + n_P + \sum_{a=1}^{m} n_{A_a} = n_P + n_A = 2n\}}$: set designating all possible ultimate destinations of the "return to home" trips for "potential" locations at which all activities with unspecified locations, $\boldsymbol{A_a} \in \boldsymbol{A}$. (It is noted that the physical location of each element of $\boldsymbol{P_A^-}$ is "home".)

$\boldsymbol{P^-} = \boldsymbol{P_P^-} \cup \boldsymbol{P_A^-} = \boldsymbol{P_P^-} \cup \boldsymbol{P_{A_1}^-} \cup \ldots \cup \boldsymbol{P_{A_m}^-} = \{\boldsymbol{n+1}, \boldsymbol{n+2}, , \ldots, \boldsymbol{2(n_P + n_A) = 2n}\}$: set designating all possible ultimate destinations of the "return to home" trips for the combined set of activities. (Note that the physical location of each element of $\boldsymbol{P^-}$ is "home".)

$\boldsymbol{P = P^+ \cup P^-}$: set of nodes comprising both predetermined locations and candidate locations of activities, and their corresponding "return home nodes".

$\boldsymbol{N} = \{\boldsymbol{0}, \boldsymbol{P}, \boldsymbol{2n+1}\}$: set of all nodes, including those associated with the initial and final return to home.

$\boldsymbol{V} = \{\boldsymbol{1}, \boldsymbol{2}, \ldots, \boldsymbol{v}, \ldots, |\boldsymbol{V}|\}$: set of vehicles used by travelers in the household to complete their scheduled activities.

$[a_i, b_i]$: time window of available start times for activity i. (Note: b_i must precede the closing of the availability of activity i by an amount equal to or greater than the duration of the activity.)

$[a_{n+i}, b_{n+i}]$: time windows for the "return home" arrival from activity *i*.

$[a_0, b_0]$: departure window for the beginning of the travel day.

$[a_{2n+i}, b_{2n+i}]$: arrival window by which time all members of the household must complete their travel.

s_i: duration of activity *i*.

t_{uw}: travel time from the location of activity u to the location of activity w.

c_{uw}^{υ}: travel cost from location of activity u to the location of activity w by vehicle υ.

B_C: household travel cost budget.

B_t^{υ}: travel time budget for the household member using vehicle υ.

APPENDIX 2-B LABEL GENERATION PROCEDURE OF GROCERY SHOPPNG LOCATION SELCTION: SINGLE VEHICLE

Iteration	Index	Visited nodes, S	Terminal node, j	Current cost, $c(S_{\alpha}, j)$	Time window constraints, $T(S_{\alpha}, j)$[24]	Previous Path index
$k = 1$ (3 labels)	1	{1}	1	1.38	$6 \leq T_0 \leq 22$ $8 \leq T_1 \leq 9$ $T_0 + 0.22 \leq T_1$	0
	2	{2}	2	0.31	$6 \leq T_0 \leq 22$ $6.05 \leq T_2 \leq 21$ $T_0 + 0.05 \leq T_2$	0
	3	{3}	3	1.56	$6 \leq T_0 \leq 22$ $6.25 \leq T_3 \leq 21$ $T_0 + 0.25 \leq T_3$	0
$k = 2$ (7 labels)	4	{1 2}	1	1.69	$8 \leq T_1 \leq 9$ $T_2 + 1 + 0.22 \leq T_1$	2
	5	{1 3}	1	1.63	$8 \leq T_1 \leq 9$ $T_3 + 1 + 0.01 \leq T_1$	3
	6	{1 2}	2	2.75	$17.22 \leq T_3 \leq 21$ $T_1 + 9 + 0.2 \leq T_3$	1
	7	{1 3}	3	1.44	$17.01 \leq T_3 \leq 21$ $T_1 + 9 + 0.01 \leq T_1$	1
	8	{1 4}	4	2.75	$17.22 \leq T_4 \leq 21$	1

[24] This column only shows arrival time windows that are newly added during the iteration. Constraints from previous paths carry on, but due to space limit, they are not shown in this table. The full set of constraints can be constructed by tracking down previous indexes.

					$T_1 + 9 + 0.22 \leq T_4$	
	9	{2 5}	5	0.63	$7.1 \leq T_5 \leq 22$ $T_1 + 1 + 0.05 \leq T_2$	2
	10	{3 6}	6	3.13	$7.5 \leq T_6 \leq 22$ $T_3 + 1 + 0.25 \leq T_6$	3
$k = 3$ **(12 labels)**	11	{1 2 5}	1	2.00	$8 \leq T_1 \leq 9$ $T_5 + 0.22 \leq T_1$	9
	12	{1 3 6}	1	4.50	$17.22 \leq T_1 \leq 9$ $T_6 + 0.22 \leq T_1$	10
	13	{1 2 4}	2	3.06	$17.27 \leq T_2 \leq 21$ $T_4 + 0.05 \leq T_2$	8
	14	{1 3 4}	3	4.31	$17.47 \leq T_3 \leq 2$ $T_4 + 0.25 \leq T_3$	8
	15	{1 3 4}	4	3.00	$17.22 \leq T_4 \leq 21$ $T_1 + 9 + 0.22 \leq T_4$	5
	16	{1 3 4}	4	3.00	$18.26 \leq T_4 \leq 21$ $T_1 + 9 + 0.25 \leq T_4$	7
	17	{1 2 4}	4	3.06	$17.22 \leq T_4 \leq 21$ $T_1 + 9 + 0.22 \leq T_4$	4
	18	{1 2 4}	4	3.06	$18.27 \leq T_4 \leq 22$ $T_1 + 9 + 0.05 \leq T_4$	6
	19	{1 2 5}	5	3.06	$17.22 \leq T_5 \leq 22$	4

				$T_1 + 9 + 0.22 \leq T_5$	
	20	{1 2 5 }	5	3.06	$18.27 \leq T_5 \leq 22$ $T_2 + 1 + 0.25 \leq T_5$	6
	21	{1 3 6}	6	3.00	$17.22 \leq T_6 \leq 22$ $T_1 + 9 + 0.22 \leq T_6$	5
	22	{1 3 6}	6	3.00	$18.26 \leq T_6 \leq 22$ $T_3 + 1 + 0.25 \leq T_6$	7
$k = 4$ (12 labels)	23	{1 2 4 5}	4	3.38	$17.22 \leq T_4 \leq 21$ $T_1 + 9 + 0.22 \leq T_4$	11
	24	{1 3 4 6}	4	3.00	$17.22 \leq T_4 \leq 21$ $T_6 \leq T_4$	21
	25	{1 3 4 6}	4	3.00	$18.26 \leq T_4 \leq 21$ $T_6 \leq T_4$	22
	26	{1 3 4 6}	4	5.88	$17.22 \leq T_4 \leq 21$ $T_1 + 9 + 0.22 \leq T_4$	12
	27	{1 2 4 5}	4	3.06	$17.22 \leq T_4 \leq 21$ $T_5 \leq T_4$	19
	28	{1 2 4 5}	4	3.06	$18.27 \leq T_4 \leq 21$ $T_5 \leq T_4$	20
	29	{1 2 4 5}	5	3.06	$18.27 \leq T_5 \leq 22$ $T_4 \leq T_5$	18
	30	{1 2 4 5}	5	3.06	$17.22 \leq T_5 \leq 22$ $T_4 \leq T_5$	17

	31	{1 2 4 5}	5	3.38	$18.32 \leq T_5 \leq 22$ $T_2 + 1 + 0.05 \leq T_5$	13
	32	{1 3 4 6}	6	3.00	$17.22 \leq T_6 \leq 22$ $T_4 \leq T_6$	15
	33	{1 3 4 6}	6	3.00	$18.26 \leq T_6 \leq 22$ $T_4 \leq T_6$	16
	34	{1 3 4 6}	6	5.88	$18.72 \leq T_6 \leq 21$ $T_3 + 1 + 0.25 \leq T_6$	14
$k = 5$ **(12 labels)**	35	{1 2 4 5 7}	7	3.00	$17.22 \leq T_7 \leq 22$ $T_4 \leq T_7$	24
	36	{1 3 4 6 7}	7	3.00	$17.22 \leq T_7 \leq 22$ $T_6 \leq T_7$	32
	37	{1 3 4 6 7}	7	3.00	$18.26 \leq T_7 \leq 22$ $T_4 \leq T_7$	25
	38	{1 2 4 5 7}	7	3.06	$17.22 \leq T_7 \leq 22$ $T_5 \leq T_7$	30
	39	{1 3 4 6 7}	7	3.00	$18.26 \leq T_7 \leq 22$ $T_6 \leq T_7$	33
	40	{1 2 4 5 7}	7	3.06	$18.27 \leq T_7 \leq 22$ $T_4 \leq T_7$	28
	41	{1 2 4 5 7}	7	3.06	$18.27 \leq T_7 \leq 22$ $T_5 \leq T_7$	29
	42	{1 2 4 5 7}	7	5.88	$17.22 \leq T_7 \leq 22$ $T_4 \leq T_7$	26
	43	{1 2 4 5 7}	7	3.38	$18.32 \leq T_7 \leq 22$ $T_5 \leq T_7$	31

44	{1 3 4 6 7}	7	3.38	$17.22 \leq T_7 \leq 22$ $T_4 \leq T_7$	23
45	{1 3 4 6 7}	7	3.06	$17.22 \leq T_7 \leq 22$ $T_4 \leq T_7$	27
46	{1 3 4 6 7}	7	5.88	$18.72 \leq T_7 \leq 22$ $T_6 \leq T_7$	34

APPENDIX 3-A HFCV: INSERTION ANALYSIS

Let n^h be the number of out-of-home and at-home activities performed by household h: each household's activity locations are $\mathbf{P^h} = \{P_0^h, P_1^h, ..., P_i^h, ..., P_{n^h}^h, P_0^h\}$ where P_0^h denotes the home location of household h, and P_i^h denotes the location for activity i of household h. It is noted that in-home activities during the day are also considered as separate activities. Corresponding arrival times are $\mathbf{a^h} = \{NA, a_1^h, ..., a_i^h, ..., a_{n^h}^h, a_0^h\}$ and departure times are $\mathbf{b^h} = \{b_0^h, b_1^h, ..., b_i^h, ..., b_{n^h}^h, NA\}$. There are a total of $n^h + 1$ trips for household h.

For $\forall k$ s.t $0 \le k \le n^h + 1$

Initially, start with a sequence of activities of home:

$$\mathbf{P_k^h} = \{P_0^h\}, \mathbf{a_k^h} = \{NA\}, \mathbf{b_k^h} = \{b_0^h\}$$

For all activities before refueling, $0 \le i \le k$

Add activities and corresponding times:

$$\mathbf{P_k^h} = \mathbf{P_k^h} \cup P_i^h,\ \mathbf{a_k^h} = \mathbf{a_k^h} \cup a_i^h,\ \mathbf{b_k^h} = \mathbf{b_k^h} \cup b_i^h$$

For $\forall r \in \mathbf{R}$

Add a refueling activity at refueling station r:

$$\mathbf{P_{k,r}^h} = \mathbf{P_k^h} \cup P^r$$

$$\mathbf{a_{k,r}^h} = \mathbf{a_k^h} \cup (a_k^{r,h} = a_k^h + t_{k,r}^h)$$

$$\mathbf{b_{k,r}^h} = \mathbf{b_k^h} \cup (b_k^{r,h} = b_k^h + t_{k,r}^h + s_{\mathbf{R}})$$

where

P^r denotes the refueling location of station r

$t^h_{k,r}$ denotes the travel time from activity location k of household h to refueling station r

$a^{r,h}_k / b^{r,h}_k$ denote arrival/departure time at refueling station r when the refueling trip is following activity k of household h

$s_{\mathbf{R}}$ denotes refueling time

Then, add a trip from the refueling station to the next activity location:

$$\mathbf{P^h_{k,r}} = \mathbf{P^h_{k,r}} \bigcup P^h_{k+1}$$

$$\mathbf{a^h_{k,r}} = \mathbf{a^h_{k,r}} \bigcup (a^h_{k+1} = a^h_k + t^h_{k,r} + t^h_{r,k+1} - t^h_{k,k+1} + s_{\mathbf{R}})$$
$$\mathbf{b^h_{k,r}} = \mathbf{b^h_{k,r}} \bigcup (b^h_{k+1} = b^h_k + t^h_{k,r} + t^h_{r,k+1} - t^h_{k,k+1} + s_{\mathbf{R}})$$

where

$t^h_{r,k+1}$ denotes the travel time from refueling station r to activity location $k+1$ of household h

$t^h_{k,k+1}$ denotes travel time from activity location k to $k+1$.

$t^h_{k,r} + t^h_{r,k+1} - t^h_{k,k+1}$ is the deviation time caused by refueling at station r between activities k and $k+1$

$s_{\mathbf{R}}$ refers to refueling duration

Then select the least deviated pattern among r different generated patterns:

$$\mathbf{P^h_k} = \mathbf{P^h_{k,r*}},\ \mathbf{a^h_k} = \mathbf{a^h_{k,r*}},\ \mathbf{b^h_k} = \mathbf{b^h_{k,r*}}$$

where $r*$ represents a refueling station with the least deviation:

$$r^* = \arg\min_r (t_{k,r}^h + t_{r,k+1}^h - t_{k,k+1}^h)$$

$d_k^h = t_{k,r^*}^h + t_{r^*,k+1}^h - t_{k,k+1}^h$ refers to the deviation time when visiting r^* between activities k and $k+1$

Then, for the remaining activities of the sequence, $k+1 < i \leq n+1$

Add activities and corresponding times:

$$\mathbf{P_k^h} = \mathbf{P_k^h} \cup P_i^h,\ \mathbf{a_k^h} = \mathbf{a_k^h} \cup (a_i^h = a_i^h + d_k^h + s_{\mathbf{R}}), \mathbf{b_k^h} = \mathbf{b_k^h} \cup (b_i^h = b_i^h + d_k^h + s_{\mathbf{R}})$$

Following activities arrival/departure times are delayed by the deviation and refueling time

Select the smallest deviant insertion and the largest deviant insertion and measure by the earliest and the latest return home times:

$\mathbf{P^h_{k\_best}}$ where $k\_best = \arg\min_k (\mathbf{a_k^h}(a_0^h))$

APPENDIX 3-B HAPPR

Define the following sets that are specific for each household, $h \in \mathbf{H}$

$\mathbf{V}_h = \{1, 2, ..., |\mathbf{V}_h|\}$: set of vehicles available to travelers in household h to complete their scheduled activities, one or more of which is a HFCV.

$\mathbf{P}_h^{R+} = \{1, 2, ..., n_R\}$: set designating candidate location at which each service type R can be performed. Index numbers and physical locations of this set are identical for all households, but the set is defined specifically for each household since index numbers of "return home" nodes from this set are different across households.

$\mathbf{P}_h^{A+} = \{n_R + 1, n_R + 2, ..., n_h = n_R + n_A^h\}$: set designating location at which each assigned activity is performed for household, h. Each activity and the physical location is different for each household.

$\mathbf{P}_h^{+} = \mathbf{P}_h^{R+} \cup \mathbf{P}_h^{A+} = \{1, 2, ..., n_R, n_R + 1, n_R + 2, ..., n_h = n_R + n_A^h\}$: set designating location at which each activity is performed for household h.

$\mathbf{P}_h^{R-} = \{n_h + 1, n_h + 2, ..., n_h + n_R\}$: set designating the ultimate destination of the "return to home" trip from candidate locations of service type R. Physical locations of this set are identical for all households, but the set is defined specifically for each household since index numbers are different across households. (It is noted that the physical location of each element of $\mathbf{P}_h^{R-}$ is "home".)

$\mathbf{P}_h^{A-} = \{n_h + n_R + 1, n_R + 2, ..., 2n_h = 2(n_R + n_A^h)\}$: set designating the ultimate destination of the "return to home" trip from out-of-home activities to be completed by travelers in household, *h*. (It is noted that the physical location of each element of $\mathbf{P}_h^{A-}$ is "home".)

$\mathbf{P}_h^{-} = \mathbf{P}_h^{R-} \cup \mathbf{P}_h^{A-} = \{n_h + 1, n_h + 2, ..., n_h + n_R, n_h + n_R + 1, n_R + 2, ..., 2n_h = 2(n_R + n_A^h)\}$: set designating the ultimate destination of the "return to home" trip for each activity for household *h*. (It is noted that the physical location of each element of $\mathbf{P}_h^{-}$ is "home".)

$\left[a_i^h, b_i^h\right]$: time window of available start times for activity *i* for household *h*. (Note: b_i^h must precede the closing of the availability of activity *i* of household *h*, by an amount equal to or greater than the duration of the activity.)

$\left[a_{n_h+i}^h, b_{n_h+i}^h\right]$: time windows for the "return home" arrival from activity *i* of household *h*.

$\left[a_0^h, b_0^h\right]$: departure window for the beginning of the travel day for household *h*.

$\left[a_{2n_h+i}^h, b_{2n_h+i}^h\right]$: arrival window by which time all members of the household *h* must complete their travel.

s_i^h: duration of activity *i* of household *h*.

$t_{u,w}^h$: travel time from the location of activity *u* to the location of activity *w*.

$c_{u,w}^{\upsilon,h}$: travel cost for household *h*, from location of activity *u* to the location of activity *w* by vehicle υ.

B_C^h: travel cost budget for household *h*.

$B_T^{\upsilon,h}$: travel time budget for the household h's member using vehicle υ.

$\mathbf{P}_h = \mathbf{P}_h^{+} \cup \mathbf{P}_h^{-}$: set of nodes comprising completion of all the activities of household h.

$\mathbf{P}_h^{A} = \mathbf{P}_h^{A+} \cup \mathbf{P}_h^{A-}$: set of nodes comprising completion of the activities household h. This does not include trips related to service type R

$\mathbf{P}_h^{R} = \mathbf{P}_h^{R+} \cup \mathbf{P}_h^{R-}$: set of nodes comprising completion of service type R of household h.

$\mathbf{P}_h = \mathbf{P}_h^{A+} \cup \mathbf{P}_h^{A-} \cup \mathbf{P}_h^{R+} \cup \mathbf{P}_h^{R-} = \mathbf{P}_h^{A} \cup \mathbf{P}_h^{R}$: set of nodes comprising completion of the household's scheduled and service type R activities.

$\mathbf{N}_h = \{0, \mathbf{P}_h, 2n_h + 1\}$: set of all nodes for household h, including those associated with the initial departure and final return to home.

$$Z_h = \beta_h^E \cdot \sum_{v \in \mathbf{V}^h} (T_{2n_h+1}^{v,h} - T_0^{v,h}) + \beta_h^T \cdot \sum_{w \in \mathbf{P}_h^+} (T_{w+n_h}^h - T_w^h) + \beta_h^D \cdot \sum_{w \in \mathbf{N}_h} \sum_{u \in \mathbf{N}_h} \sum_{v \in \mathbf{V}^h} t_{u,w}^h \cdot X_{u,w}^{v,h} \tag{3-A1}$$

Subject to

$$\sum_{v \in \mathbf{V}_h} \sum_{w \in \mathbf{N}_h} X_{u,w}^{v,h} = 1, \quad u \in \mathbf{P}_h^{A+} \tag{3-A2}$$

$$\sum_{u \in \mathbf{P}_h^{R+}} \sum_{w \in \mathbf{N}_h} X_{u,w}^{v,h} = 1, \quad v \in \mathbf{V}_h \tag{3-A3}$$

$$\sum_{w \in \mathbf{N}_h} X_{u,w}^{v,h} - \sum_{w \in \mathbf{N}_h} X_{w,u}^{v,h} = 0, \quad u \in \mathbf{P}_h, v \in \mathbf{V}_h \tag{3-A4}$$

$$\sum_{w \in \mathbf{P}_h^+} X_{0,w}^{v,h} \le 1, \quad v \in \mathbf{V}_h \tag{3-A5}$$

$$\sum_{u \in \mathbf{P}_h^-} X_{u,2n_h+1}^{v,h} - \sum_{u \in \mathbf{P}_h^+} X_{0,u}^{v,h} = 0, \quad v \in \mathbf{V}_h \tag{3-A6}$$

$$\sum_{w \in \mathbf{N}_h} X_{w,u}^{v,h} - \sum_{w \in \mathbf{N}_h} X_{w,n_h+u}^{v,h} = 0, \quad u \in \mathbf{P}_h^+, v \in \mathbf{V}_h \tag{3-A7}$$

$$T_u^h + s_u^h + t_{u,w}^h = T_{n_h+u}^h, \quad u, w \in \mathbf{P}_h^+ \tag{3-A8}$$

$$X_{u,w}^{v,h} = 1 \Rightarrow T_u^h + s_u^h + t_{u,w}^h = T_w^h, \quad u, w \in \mathbf{P}_h, v \in \mathbf{V}_h \tag{3-A9}$$

$$X_{0,w}^{v,h} = 1 \Rightarrow T_0^{v,h} + s_u^h + t_{0,w}^h = T_w^h, \quad w \in \mathbf{P}_h^+, v \in \mathbf{V}_h \tag{3-A10}$$

$$X_{u,2n_h+1}^{v,h} = 1 \Rightarrow T_u^h + s_u^h + t_{u,2n_h+1}^h = T_{2n_h+1}^h, \quad u \in \mathbf{P}_h^-, v \in \mathbf{V}_h \tag{3-A11}$$

$$a_u^h \le T_u^h \le b_u^h, \quad u \in \mathbf{P}_h \tag{3-A12}$$

$$a_0^h \le T_0^{v,h} \le b_0^h, \quad v \in \mathbf{V}_h \tag{3-A13}$$

$$a_{2n_h+1}^h \le T_{2n_h+1}^{v,h} \le b_{2n_h+1}^h, \quad v \in \mathbf{V}_h \tag{3-A14}$$

$$X_{u,w}^{v,h} = 1 \Rightarrow Y_u^h + d_w = Y_w^h, \quad u \in \mathbf{P}_h, w \in \mathbf{P}_h^+, v \in \mathbf{V}_h \tag{3-A15}$$

$$X_{u,w}^{v,h} = 1 \Rightarrow Y_w^h = 0, \quad u \in \mathbf{P}_h, \mathrm{w} \in \mathbf{P}_h^-, v \in \mathbf{V}_h \tag{3-A16}$$

$$X_{0,w}^{v,h} = 1 \Rightarrow Y_0^h + d_w = Y_w^h, \quad w \in \mathbf{P}_h^+, v \in \mathbf{V}_h \tag{3-A17}$$

$$Y_0^h = 0, \quad 0 \le Y_u^h \le D, \quad u \in \mathbf{P}_h^+ \tag{3-A18}$$

$$\sum_{v \in \mathbf{V}_h} \sum_{u \in \mathbf{N}_h} \sum_{w \in \mathbf{N}_h} c_{u,w}^{v,h} \cdot X_{u,w}^{v,h} \le B_C^{v,h} \tag{3-A19}$$

$$\sum_{v \in \mathbf{V}_h} \sum_{u \in \mathbf{N}_h} \sum_{w \in \mathbf{N}_h} t_{u,w}^{h} \cdot X_{u,w}^{v,h} \le B_T^{v,h} \tag{3-A20}$$

The household-specific decision variables in (A1)-(A20) are:

$X_{u,w}^{v,h}$ binary decision variable equal to unity if vehicle υ travels from activity u to activity w, and zero otherwise.

T_u^h the time at which participation in activity u of household h begins.

$T_0^{v,h}, T_{2n_h+1}^{v,h}$ the times at which vehicle υ from household h first departs from home and last returns to home, respectively

Y_u^h the total accumulation of either sojourns or time (depending on the selection of D and d_u) of household h on a particular tour immediately following completion of activity u.

APPENDIX 3-C HAPPR: FUEL INVENTORY EXTENSION

$$X_{u,j}^{v,h} = 1 \Rightarrow F_w^{v,h} = I^{v,h}, \quad u \in \mathbf{N}_h, j \in \mathbf{P}_h^{R+}, v \in \mathbf{V}_h \tag{3-A21}$$

$$X_{u,w}^{v,h} = 1 \Rightarrow F_w^{v,h} = F_u^{v,h} - f_{u,w}^{v,h}, \quad u, w \in \mathbf{N}_h, v \in \mathbf{V}_h \tag{3-A22}$$

$$0 \le F_u^{v,h} \le I^v, \quad u \in \mathbf{N}_h, v \in \mathbf{V}_h \tag{3-A23}$$

$$X_{u,w}^{v,h} = 1 \Rightarrow F_u^{v,h} - f_{u,w}^{v,h} \ge 0, \quad u, w \in \mathbf{N}_h, v \in \mathbf{V}_h \tag{3-A24}$$

where the variable $F_w^{v,h}$ is defined as:

$F_u^{v,h}$ the fuel inventory of vehicle v of household h at the arrival at the location for activity u.

and related parameters of

$I^{v,h}$ the full fuel inventory of vehicle v of household h.

$f_{u,w}^{v,h}$ fuel usage from the location of activity u to the location of activity w using vehicle v of household h.

In the above, (A21) represents the refueling activity. We assume that a vehicle is refueled to the maximum. (A22) represents the fuel consumption, and updates the fuel inventory value. (3-A23) constrains that vehicles can only be operated when the fuel inventory is not empty. (3-A24) constrains the fuel inventory to be non-negative while traveling since $F_u^{v,h}$ is defined only upon the arrival at activity u.

If the fuel inventory constraints are employed, we can relax the constraint of one visitation to a refueling trip. A driver can choose not to refuel or to refuel one or more times (for long distance trips) as in (A2-1) instead of one refueling trip per day as in (A2).

$$\sum_{j \in \mathbf{P}_h^{R+}} \sum_{w \in \mathbf{N}_h} X_{j,w}^{v,h} \geq 0, \quad v \in \mathbf{V}_h \tag{3-A2-1}$$

APPENDIX 3-D BEV: DELAY ANALYSIS

Let n^h be the number of out-of-home and at-home activities performed by household h: each household's activity locations are $\mathbf{P^h} = \{P_0^h, P_1^h, ..., P_i^h, ..., P_{n^h}^h, P_0^h\}$ where P_0^h denotes the home location of household h, and P_i^h denotes the location for activity i of household h. It is noted that in-home activities during the day are also considered as separate activities. Corresponding reported arrival times are $\mathbf{a^h} = \{NA, \bar{a}_1^h, ..., \bar{a}_i^h, ..., \bar{a}_{n^h}^h, \bar{a}_0^h\}$ and departure times are $\mathbf{b^h} = \{\bar{b}_0^h, \bar{b}_1^h, ..., \bar{b}_i^h, ..., \bar{b}_{n^h}^h, NA\}$. There are a total of $n^h + 1$ trips for household h. Define set of battery status $\mathbf{E}^h = \{R, E_1^h, ..., E_i^h, ..., E_{n^h}^h, E_0^h\}$. It is assumed that the vehicle starts with a fully charged battery.

Initially, start with a sequence of activities of home:

Add the home node departure at the reported departure time.

$$\mathbf{P^h} = \{P_0^h\}, \mathbf{a^h} = \{NA\}, \mathbf{b^h} = \{b_0^h = \bar{b}_0^h\}$$

$$\mathbf{E}^h = \{R\}$$

Repeat for $1 \le k \le n^h + 1$

$$\mathbf{P^h} = \mathbf{P^h} \cup P^r$$

$$\mathbf{a^h} = \mathbf{a^h} \cup (a_k^h = b_{k-1}^h + t_{k-1,k}^h)$$

$$\mathbf{b^h} = \mathbf{b^h} \cup b_k^h \text{ where } b_k = \begin{cases} a_k + \dfrac{e_{k,k+1}^h - E_k}{r} & , E_k + r \cdot s_k \le e_{k,k+1}^h \\ a_k + s_k & E_k + r \cdot s_k \ge e_{k,k+1}^h \end{cases}$$

$$\mathbf{E}^h = \mathbf{E}^h \bigcup (E_k^h = \max\{E_{k-1}^h + r \cdot (b_k^h - a_k^h), R\} - e_{k,k+1}^h)$$

where $s_k = \overline{b}_k - \overline{a}_k$ denotes the reported duration of activity k of household h

APPENDIX 3-E: HAPPC

Define the following sets that are specific for each household, $h \in \mathbf{H}$

$$Z_h = \beta_h^E \cdot \sum_{v \in \mathbf{V}^h} (T_{2n_h+1}^{v,h} - T_0^{v,h}) + \beta_h^T \cdot \sum_{w \in \mathbf{P}_h^+} (T_{w+n_h}^h - T_w^h) + \beta_h^D \cdot \sum_{w \in \mathbf{N}_h} \sum_{u \in \mathbf{N}_h} \sum_{v \in \mathbf{V}^h} t_{u,w}^h \cdot X_{u,w}^{v,h} \quad \text{(3-A1)}$$

Subject to

(3-A2), (3-A3) – (3-A20)

$$X_{0,w}^{v,h} = 1 \Rightarrow E_w^{v,h} = R^{v,h} - e_{0,w}^{v,h}, \quad \mathrm{w} \in \mathbf{N}_h, v \in \mathbf{V}_h \quad \text{(3-A25)}$$

$$X_{u,w}^{v,h} = 1 \Rightarrow E_w^{v,h} \le E_u^{v,h} + r \cdot (T_w - T_u - t_{u,w}^{v,h}) - e_{u,w}^{v,h}, \quad u,\mathrm{w} \in \mathbf{P}_h, v \in \mathbf{V}_h \quad \text{(3-A26-1)}$$

$$X_{u,w}^{v,h} = 1 \Rightarrow E_w^{v,h} \le R^{v,h} - e_{u,w}^{v,h}, \quad u,\mathrm{w} \in \mathbf{P}_h, v \in \mathbf{V}_h \quad \text{(3-A26-2)}$$

$$X_{u,w}^{v,h} = 1 \Rightarrow E_w^{v,h} \ge E_u^{v,h} - e_{u,w}^{v,h}, \quad u,\mathrm{w} \in \mathbf{P}_h, v \in \mathbf{V}_h \quad \text{(3-A27)}$$

$$0 \le E_u^h \le R^{v,h}, \quad u \in \mathbf{N}_h \quad \text{(3-A28)}$$

$$X_{u,2n_h+1}^{v,h} = 1 \Rightarrow E_u^{v,h} + r \cdot (T_{2n_h+1}^v - T_u - t_{u,w}^{v,h}) \ge e_{u,2n_h+1}^{v,h}, \quad u \in \mathbf{P}_h, v \in \mathbf{V}_h \quad \text{(3-A29)}$$

where the variable $E_u^{v,h}$ is defined as:

$E_u^{v,h}$ the electric battery inventory of vehicle v of household h right before the beginning of activity u.

and related parameters of

$I^{v,h}$ the full electric battery inventory (capacity) of vehicle v of household h.

$e_{u,w}^{v,h}$ the electric battery consuption from the location of activity u to the location of activity w using vehicle v of household h. In this study, travel distance is used for electric battery usage.

r the rate at which the electric battery is recharged

(3-A25) represents the battery consumption, and updates the battery inventory value for the first trip. A vehicle is assumed to be fully charged at the start of the travel day. (3-A26-1) constrains that if the electric battery is recharged, the amount is dependent on time and rate, if it is less than the capacity of the battery. If the battery reaches the capacity while performing an activity, charged amount is the capacity (3-A26-2). This is equivalent to taking the battery status as the maximum of $E_k + r \cdot (b_k - a_k)$ and the capacity.

(3-A26-1) (3-A26-2) along with (3-A27) states the upper and lower bounds of the battery status. If the vehicle is charged, the upper bound of charged battery is decided by time and charging rate (3-A26-1), or the capacity (3-A26-2). If the vehicle is not charged at all, the battery status is subtracting the consumption (3-A28). (3-A29) constrains the battery status of the final return home that it needs to be greater than the final battery consumption. Although the final return home has no distance value since it is one of the return home trips to final return home trip, this condition is imposed to earlier trips through (3-A26) – (3-A28).

APPENDIX 4-A NOTATIONS AND FORMULATIONS OF HAPPR [25]

$$\min O^h = \text{Travel Disutility of Household } h \quad (4\text{-A1})$$

$$\sum_{v \in \mathbf{V}_h} \sum_{w \in \mathbf{N}_h} X_{u,w}^{v,h} = 1, \quad u \in \mathbf{P}_h^{A+} \quad (4\text{-A2})$$

$$\sum_{j \in \mathbf{P}_h^{R+}} \sum_{w \in \mathbf{N}_h} X_{j,w}^{v,h} = 1, \quad v \in \mathbf{V}_h \quad (4\text{-A3})$$

$$\sum_{w \in \mathbf{N}_h} X_{u,w}^{v,h} - \sum_{w \in \mathbf{N}_h} X_{w,u}^{v,h} = 0, \quad u \in \mathbf{P}_h, v \in \mathbf{V}_h \quad (4\text{-A4})$$

$$\sum_{w \in \mathbf{P}_h^{+}} X_{0,w}^{v,h} \le 1, \quad v \in \mathbf{V}_h \quad (4\text{-A5})$$

$$\sum_{u \in \mathbf{P}_h^{-}} X_{u,2n_h+1}^{v,h} - \sum_{u \in \mathbf{P}_h^{+}} X_{0,u}^{v,h} = 0, \quad v \in \mathbf{V}_h \quad (4\text{-A6})$$

$$\sum_{w \in \mathbf{N}_h} X_{w,u}^{v,h} - \sum_{w \in \mathbf{N}_h} X_{w,n_h+u}^{v,h} = 0, \quad u \in \mathbf{P}_h^{+}, v \in \mathbf{V}_h \quad (4\text{-A7})$$

$$T_u^h + s_u^h + t_{u,w}^h = T_{n_h+u}^h, \quad u, w \in \mathbf{P}_h^{+} \quad (4\text{-A8})$$

$$X_{u,w}^{v,h} = 1 \Rightarrow T_u^h + s_u^h + t_{u,w}^h = T_w^h, \quad u, w \in \mathbf{P}_h, v \in \mathbf{V}_h \quad (4\text{-A9})$$

$$X_{0,w}^{v,h} = 1 \Rightarrow T_0^{v,h} + s_u^h + t_{0,w}^h = T_w^h, \quad w \in \mathbf{P}_h^{+}, v \in \mathbf{V}_h \quad (4\text{-A10})$$

$$X_{u,2n_h+1}^{v,h} = 1 \Rightarrow T_u^h + s_u^h + t_{u,2n_h+1}^h = T_{2n_h+1}^h, \quad u \in \mathbf{P}_h^{-}, v \in \mathbf{V}_h \quad (4\text{-A11})$$

$$a_u^h \le T_u^h \le b_u^h, \quad u \in \mathbf{P}_h \quad (4\text{-A12})$$

$$a_0^h \le T_0^{v,h} \le b_0^h, \quad v \in \mathbf{V}_h \quad (4\text{-A13})$$

$$a_{2n_h+1}^h \le T_{2n_h+1}^{v,h} \le b_{2n_h+1}^h, \quad v \in \mathbf{V}_h \quad (4\text{-A14})$$

$$X_{u,w}^{v,h} = 1 \Rightarrow Y_u^h + d_w = Y_w^h, \quad u \in \mathbf{P}_h, w \in \mathbf{P}_h^{+}, v \in \mathbf{V}_h \quad (4\text{-A15})$$

---

[25] It is noted that this is the same as HAPPR in APPENDIX 3-A, except for the objective function.

$$X_{u,w}^{v,h} = 1 \Rightarrow Y_w^h = 0, \quad u \in \mathbf{P}_h, \mathrm{w} \in \mathbf{P}_h^-, v \in \mathbf{V}_h \tag{4-A16}$$

$$X_{0,w}^{v,h} = 1 \Rightarrow Y_0^h + d_w = Y_w^h, \quad w \in \mathbf{P}_h^+, v \in \mathbf{V}_h \tag{4-A17}$$

$$Y_0^h = 0, \quad 0 \le Y_u^h \le D, \quad u \in \mathbf{P}_h^+ \tag{4-A18}$$

$$\sum_{v \in \mathbf{V}_h} \sum_{u \in \mathbf{N}_h} \sum_{w \in \mathbf{N}_h} c_{u,w}^{v,h} \cdot X_{u,w}^{v,h} \le B_C^{v,h} \tag{4-A19}$$

$$\sum_{v \in \mathbf{V}_h} \sum_{u \in \mathbf{N}_h} \sum_{w \in \mathbf{N}_h} t_{u,w}^{h} \cdot X_{u,w}^{v,h} \le B_T^{v,h} \tag{4-A20}$$

The household-specific decision variables in (4-A1)-(4-A20) are:

$X_{u,w}^{v,h}$ binary decision variable equal to unity if vehicle υ travels from activity u to activity w, and zero otherwise.

T_u^h the time at which participation in activity u of household h begins.

$T_0^{v,h}, T_{2n_h+1}^{v,h}$ the times at which vehicle υ from household h first departs from home and last returns to home, respectively

Y_u^h the total accumulation of either sojourns or time (depending on the selection of D and d_u) of household h on a particular tour immediately following completion of activity u.

APPENDIX 5-A dHAPP CONSTRAINTS

$$\sum_{v \in V_h} \sum_{w \in \mathbf{Q}_h} X_{uw}^{v,h} = 1, \quad u \in P_h^+, h \in \mathbf{H} \tag{5-A1}$$

$$\sum_{w \in \mathbf{Q}_h} X_{uw}^{v,h} - \sum_{w \in \mathbf{Q}_h} X_{wu}^{v,h} = 0, \quad u \in P_h, v \in V_h, h \in \mathbf{H} \tag{5-A2}$$

$$\sum_{w \in P_h^+} X_{0w}^{v,h} = 1, \quad v \in V_h, h \in \mathbf{H} \tag{5-A3}$$

$$\sum_{u \in P_h^-} X_{u,2n+1}^{v,h} - \sum_{u \in P_h^+} X_{0,u}^{v,h} = 0, \quad v \in V_h, h \in \mathbf{H} \tag{5-A4}$$

$$\sum_{w \in \mathbf{Q}_h} X_{wu}^{v,h} - \sum_{w \in \mathbf{Q}_h} X_{w,n+u}^{v,h} = 0, \quad u \in P_h^+, v \in V_h, h \in \mathbf{H} \tag{5-A5}$$

$$T_u^h + s_u^h + t_{uw}^h = T_{n+u}^h, \quad u, w \in P_h^+, h \in \mathbf{H} \tag{5-A6}$$

$$X_{uw}^{v,h} = 1 \Rightarrow T_u^h + s_u^h + t_{uw}^h = T_w^h, \quad u, w \in P_h, v \in V_h, h \in \mathbf{H} \tag{5-A7}$$

$$X_{0w}^{v,h} = 1 \Rightarrow T_0^{v,h} + s_u^h + t_{0w}^h = T_w^h, \quad w \in P_h^+, v \in V_h, h \in \mathbf{H} \tag{5-A8}$$

$$X_{u,2n+1}^{v,h} = 1 \Rightarrow T_u^h + s_u^h + t_{u,2n+1}^h = T_{2n+1}^h, \quad u \in P_h^-, v \in V_h, h \in \mathbf{H} \tag{5-A9}$$

$$a_u^h \le T_u^h \le b_u^h, \quad u \in P_h, h \in \mathbf{H} \tag{5-A10}$$

$$a_0^h \le T_0^{v,h} \le b_0^h, \quad v \in V_h, h \in \mathbf{H} \tag{5-A11}$$

$$a_{2n+1}^h \le T_{2n+1}^{v,h} \le b_{2n+1}^h, \quad v \in V_h, h \in \mathbf{H} \tag{5-A12}$$

$$X_{uw}^{v,h} = 1 \Rightarrow Y_u^h + d_w = Y_w^h, \quad u \in P_h, w \in P_h^+, v \in V_h, h \in \mathbf{H} \tag{5-A13}$$

$$X_{uw}^{v,h} = 1 \Rightarrow Y_w^h = 0, \quad u \in P_h, \mathrm{w} \in P_h^-, v \in V_h, h \in \mathbf{H} \tag{5-A14}$$

$$X_{0w}^{v,h} = 1 \Rightarrow Y_0^h + d_w = Y_w^h, \quad w \in P_h^+, v \in V_h, h \in \mathbf{H} \tag{5-A15}$$

$$Y_0^h = 0, \quad 0 \le Y_0^h \le D, \quad u \in P_h^+, h \in \mathbf{H} \tag{5-A16}$$

$$\sum_{v\in V_h}\sum_{u\in \mathbf{Q}_h}\sum_{w\in \mathbf{Q}_h} c_{uw}^{v,h}\cdot X_{uw}^{v,h} \le B_C^{v,h}, \quad h\in \mathbf{H} \tag{5-A17}$$

$$\sum_{v\in V_h}\sum_{u\in \mathbf{Q}_h}\sum_{w\in \mathbf{Q}_h} t_{uw}^{h}\cdot X_{uw}^{v,h} \le B_T^{v,h}, \quad v\in V_h, h\in \mathbf{H} \tag{5-A18}$$

$$X_{uv}^{v,h} \in (0,1), \quad u,w\in \mathbf{Q}_h, v\in V_h, h\in H \tag{5-A19}$$

APPENDIX 5-B CASE STUDY LINK IMPROVEMENT

ID	A node	B node	Facility	Distance (miles)	Travel Time (minutes)	Avg Speed (MPH)	Cost
3779	37	79	SR 39	6.03	13.13	27.56	166.19
7937	79	37	SR 39	6.03	13.13	27.56	166.19
7917	79	17	SR 39	5.5	11.98	27.56	151.58
1779	17	79	SR 39	5.5	11.98	27.56	151.58
6086	60	86	SR 57	6.36	6.50	58.75	373.65
8660	86	60	SR 57	6.36	6.50	58.75	373.65
8667	86	67	SR 57	4.77	4.87	58.75	280.24
6786	67	86	SR 57	4.77	4.87	58.75	280.24
4839	48	39	SR 55	12.27	15.68	46.96	576.20
3948	39	48	SR 55	12.27	15.68	46.96	576.20
6162	61	62	SR 22	5.29	6.71	47.28	250.11
6261	62	61	SR 22	5.29	6.71	47.28	250.11
6587	65	87	SR 261	4.6	4.30	64.20	295.32

8765	87	65	SR 261	4.6	4.30	64.20	295.32
8788	87	88	SR 261	2.53	2.36	64.20	162.43
8887	88	87	SR 261	2.53	2.36	64.20	162.43
8889	88	89	SR 261	4.07	3.80	64.20	261.29
8988	89	88	SR 261	4.07	3.80	64.20	261.29
7875	78	75	SR 241	5.65	5.21	65.09	367.76
7578	75	78	SR 241	5.65	5.21	65.09	367.76