

# UC Irvine

## Working Paper Series

### Title

The Location Selection Problem for the Household Activity Pattern Problem

### Permalink

<https://escholarship.org/uc/item/5865d8xf>

### Authors

Kang, Jee Eun  
Recker, Will W.

### Publication Date

2012

**The Location Selection Problem for the Household Activity Pattern Problem**

UCI-ITS-WP-12-1

**Jee Eun Kang  
Will W. Recker**

**Department of Civil Engineering and  
Institute of Transportation Studies  
University of California, Irvine; Irvine, CA 92697-3600, U.S.A.  
[jekang@uci.edu](mailto:jekang@uci.edu), [wwrecker@uci.edu](mailto:wwrecker@uci.edu)**

**2012**

**Institute of Transportation Studies  
University of California, Irvine  
Irvine, CA 92697-3600, U.S.A.  
<http://www.its.uci.edu>**

# **The Location Selection Problem for the Household Activity Pattern Problem**

Jee Eun Kang and Will Recker

Department of Civil and Environmental Engineering

Institute of Transportation Studies

University of California, Irvine

Irvine, CA 92697 USA

## **Abstract**

In this paper, an integrated destination choice model based on routing and scheduling considerations of daily activities is proposed. Extending the Household Activity Pattern Problem (HAPP), the Location Selection Problem (LSP-HAPP) demonstrates how location choice is made as a simultaneous decision from interactions both with activities having predetermined locations and those with many candidate locations. A dynamic programming algorithm, developed for 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 keep arrival times as functions for mathematical programming formulations of activity-based travel models that often have time variables in the objective.

## **Keywords**

Destination Choice, Location Selection, Household Activity Pattern Problem, column generation, exact dynamic programming for PDPTW

## **1. Introduction**

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 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).<sup>1</sup> 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. The scheduling and routing model we propose is based on the Household Activity Pattern Problem (HAPP) (Recker, 1995). 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). Distinct from the majority of activity-based travel demand modeling that has been 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).

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

---

<sup>1</sup> Choice of such service-provider modes as public transit that have specific routes and schedules are not included in the proposed model, since the complications introduced by their discrete temporal availability and multiple routes greatly complicate the formulation.

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) 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 paper, 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. Location Selection Problem for the Household Activity Pattern Problem**

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 for notation used):

$X_{uw}^v, u, w \in \mathbf{N}, v \in \mathbf{V}, u \neq w$	binary decision variable equal to unity if vehicle $v$ travels from activity $u$ to activity $w$ , and zero otherwise.
$T_u, u \in \mathbf{P}$	the time at which participation in activity $u$ begins.
$T_0^v, T_{2n+1}^v, v \in \mathbf{V}$	the times at which vehicle $v$ first departs from home and last returns to home, respectively.
$Y_u, u \in \mathbf{P}$	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  $\mathbf{M}_P = \{1, 2, \dots, i, \dots, n_P\}$  of  $n_P$  out-of-home activities with pre-selected (one-to-one) locations  $\mathbf{P}_P^+ = \{1, 2, \dots, i, \dots, n_P\}$  and a set  $\mathbf{A} = \{\mathbf{A}_1, \dots, \mathbf{A}_a, \dots, \mathbf{A}_m\}$  of out-of-home activities of specific types (e.g., grocery shopping)  $\mathbf{A}_a$ , each of which with  $n_{A_a}$  possible corresponding locations  $\mathbf{P}_{A_a}^+ = \{1, 2, \dots, i, \dots, n_{A_a}\}$ , using mode of travel  $v$ , can be represented by the following formulation.<sup>2</sup>

$$\text{Minimize } Z = \text{Household Disutility} \quad (1)$$

subject to:

$$\sum_{v \in \mathbf{V}} \sum_{w \in \mathbf{N}} X_{uw}^v = 1, \quad u \in \mathbf{P}_P^+ \quad (2)$$

$$\sum_{v \in \mathbf{V}} \sum_{u \in \mathbf{P}_{A_a}^+} \sum_{w \in \mathbf{N}} X_{uw}^v = 1, \quad \mathbf{A}_a \in \mathbf{A} \quad (3)$$

$$\sum_{w \in \mathbf{N}} X_{uw}^v - \sum_{w \in \mathbf{N}} X_{wu}^v = 0 \quad u \in \mathbf{P}, v \in \mathbf{V} \quad (4)$$

$$\sum_{w \in \mathbf{P}^+} X_{0w}^v \leq 1, \quad v \in \mathbf{V} \quad (5)$$

$$\sum_{u \in \mathbf{P}^-} X_{u, 2n+1}^v - \sum_{w \in \mathbf{P}^+} X_{0w}^v = 0, \quad v \in \mathbf{V} \quad (6)$$

$$\sum_{w \in \mathbf{N}} X_{wu}^v - \sum_{w \in \mathbf{N}} X_{w, n+u}^v = 0 \quad u \in \mathbf{P}^+, v \in \mathbf{V} \quad (7)$$

$$T_u + s_u + t_{u, n+u} \leq T_{n+u} \quad u \in \mathbf{P}_P^+ \quad (8-1)$$

---

<sup>2</sup> 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.

$$X_{uw}^v = 1 \Rightarrow T_u + s_u + t_{u,n+u} \leq T_{n+u} \quad u \in \mathbf{P}_A^+, w \in \mathbf{N}, v \in \mathbf{V} \quad (8-2)$$

$$\sum_{v \in \mathbf{V}} \sum_{w \in \mathbf{N}} X_{uw}^v = 0 \Rightarrow T_u = T_{u+n} = 0 \quad u \in \mathbf{P}_A^+ \quad (8-3) \quad X_{uw}^v = 1 \Rightarrow T_u + s_u + t_{uw} \leq T_w, \quad u, w \in \mathbf{P}, v \in \mathbf{V} \quad (9)$$

$$X_{0w}^v = 1 \Rightarrow T_0^v + t_{0w} \leq T_w, \quad w \in \mathbf{P}^+, v \in \mathbf{V} \quad (10)$$

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

$$a_u \leq T_u \leq b_u, \quad u \in \mathbf{P}^+ \quad (12-1)$$

$$X_{uw}^v = 1 \Rightarrow a_u \leq T_u \leq b_u, \quad u \in \mathbf{P}_A^+, u \in \mathbf{N}, v \in \mathbf{V} \quad (12-2)$$

$$X_{0w}^v = 1 \Rightarrow a_0 \leq T_0^v \leq b_0, \quad w \in \mathbf{P}^+, v \in \mathbf{V} \quad (13-1)$$

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

$$X_{u,2n+1}^v = 1 \Rightarrow a_{2n+1} \leq T_{2n+1}^v \leq b_{2n+1}, \quad u \in \mathbf{P}^-, v \in \mathbf{V} \quad (14-1)$$

$$\sum_{u \in \mathbf{P}^-} X_{u,2n+1}^v \Rightarrow T_{2n+1}^v = 0, \quad v \in \mathbf{V} \quad (14-2) \quad X_{uw}^v = 1 \Rightarrow Y_u + d_w = Y_w, \quad u \in \mathbf{P}, w \in \mathbf{P}^+, v \in \mathbf{V} \quad (15)$$

$$X_{uw}^v = 1 \Rightarrow Y_w - d_w = Y_u, \quad u \in \mathbf{P}, w \in \mathbf{P}^-, v \in \mathbf{V} \quad (16)$$

$$X_{0w}^v = 1 \Rightarrow Y_0 + d_w = Y_w, \quad w \in \mathbf{P}^+, v \in \mathbf{V} \quad (17)$$

$$Y_0 = 0, \quad 0 \leq Y_u \leq D, \quad u \in \mathbf{P}^+ \quad (18)$$

$$\sum_{v \in \mathbf{V}} \sum_{u \in \mathbf{N}} \sum_{w \in \mathbf{N}} c_{uw}^v X_{uw}^v \leq B_c \quad (19)$$

$$\sum_{u \in \mathbf{N}} \sum_{w \in \mathbf{N}} t_{uw}^v X_{uw}^v \leq B_t^v \quad (20)$$

$$X_{uw}^v = \begin{cases} 0 \\ 1 \end{cases}, \quad u, w \in \mathbf{N}, v \in \mathbf{V} \quad (21)$$

$$T_u \geq 0, \quad u \in \mathbf{P} \quad (22)$$

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

The constraints that specify that each activity location needs to be visited (performed) are split into two sets of constraints. Equations (2) impose the condition that there is one and only one path leading from each activity with pre-selected location. Equations (3) impose the condition that there is one and only one path leading from one and only one type  $A_a$  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 (4) ensure that there is a connected path among the activities (and their return trips to home) and that no activity is revisited. Equations (5) allow for the possibility that some of the vehicles in the household's stable of vehicles may not be used. Equations (6) enforce a restriction similar to that in Equations (2), but with reference to the paths leading from the origin and to the final termination (i.e., the depot). Equations (7) stipulate that the return-home trip be on the same path as its associated out-of-home activity. The original equation (8),



$T_u + s_u + t_{u,n+u} \leq T_{n+u}$ ,  $u \in P^+$ , is a restriction that the activity start times for elements of  $P^+$  precede those of corresponding elements in  $P^-$  (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 (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 (8-3). Equation (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 (10) and (11) state that restrictions similar to those imposed by Equation (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 (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,  $a_u \leq T_u \leq b_u$ ,  $u \in P$ , to be satisfied only when the node is visited for the selective locations. Equations (13), and (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 (15) through (18) impose conditions on the maximum number of sojourns allowed in any single tour. Equations (19) and (20) enforce budget constraints. Equations (21) (22) (23) add non-negativity and integer constraints.

### 3. 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. In this paper, 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 (1)-(7) form a multi-commodity problem, we can rewrite in arc-path formulation as the following:

$$\text{minimize } \sum_{r \in \Psi} c_r Y_r \quad (a)$$

$$\sum_{r \in \Psi} a_{ir} Y_r = 1, \quad i \in \mathbf{P}^+ \quad (b1)$$

$$\sum_{i \in \mathbf{P}_{Aa}^+} \sum_{r \in \Psi} a_{ir} Y_r = 1 \quad A_a \in \mathbf{A} \quad (b2)$$

$$\sum_{r \in \Psi} Y_r \leq |\mathbf{V}| \quad (c)$$

where

$\Psi$ : 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 (8) - (22). Equations (b1) and (b2) are substituted for the original constraint of PDPTW arc-path formulation for the Location Selection Problem:

$$\sum_{r \in \Psi} a_{ir} Y_r = 1, \quad i \in \mathbf{P}^+ \quad (b)$$

Equations (b1) constrain that all activities with pre-selected location need to be visited once. And Equations (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 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  $S \subseteq \mathbf{P}$ , and  $i \in S$ .  $S$  is a non-ordered set of cardinality  $k$ , where  $k$  is the iteration number.
- $(S_\alpha, i)$ : a given route  $\alpha$  to state  $(S, i)$
- $t(S_\alpha, i)$ : the arrival time at node  $i$ , following route  $\alpha$
- $c(S_\alpha, i)$ : the current cost at node  $i$ , following route  $\alpha$
- $d(S_\alpha, i)$ : the cumulative number of sojourns in a tour at node  $i$ , following route  $\alpha$

### Initialization ( $k = 1$ )

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

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

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

$$\begin{aligned} T(S_\alpha, j) &= \{ \max(a_j, a_0 + t_{0,j}) \leq t(S_\alpha, j) \leq b_j, \\ &\quad t(S_\alpha, 0) + t_{0,j} \leq t(\text{SU}\{j\}_\alpha, i) \} \\ c(S_\alpha, j) &= c_{0,j} \\ d(S_\alpha, j) &= d_i \end{aligned}$$

### Recursion ( $2 \leq k$ )

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

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

Then the states are tested for elimination criteria, and if the state  $(\text{SU}\{j\}, j)$  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:

$$\begin{aligned} T(\text{SU}\{j\}_\alpha, j) &= T(S_\alpha, i) \cup \{ \max(a_j, T_i + s_i + t_{i,j}) \leq t(\text{SU}\{j\}_\alpha, i) \leq b_j, \\ &\quad t(S_\alpha, i) + s_i + t_{i,j} \leq t(\text{SU}\{j\}_\alpha, i) \} \\ c(\text{SU}\{j\}_\alpha, j) &= c(S_\alpha, i) + c_{i,j} \\ d(\text{SU}\{j\}_\alpha, j) &= d(S_\alpha, i) + d_j \end{aligned}$$

Stop when there is no label generated at this iteration.

### Selection of Arrival Times

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

$$\text{Minimize } f(T_0, T_1, \dots, T_{2n}, T_{2n+1}) \text{ such that } T(S_\alpha, 2n + 1)$$

Algorithm 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 (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.  $\sum_{v \in V} \sum_{i \in N} \sum_{j \in N} c_{ij} X_{ij}$  or  $\sum_{v \in V} \sum_{i \in N} \sum_{j \in N} t_{ij} X_{ij}$ ), and the other as a function of arrival times (e.g.,  $\sum_{v \in V} (T_{2n+1}^v - T_0^v)$ ); e.g.,

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

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  $SU\{0\}$ , following path  $\alpha$

and during *Recursion* ( $2 \leq k$ ), a label is created with possible time windows of arrival time determined as  $\max(a_j, T_i + s_i + t_{i,j}) \leq t(SU\{j\}_\alpha, i) \leq b_j$ . Conditions respecting the path sequence are as  $t(S_\alpha, i) + s_i + t_{i,j} \leq t(SU\{j\}_\alpha, i)$ . The feasibility of arrival time windows needs to be delivered as well as previous time windows of arrival times. Then,  $c(S_\alpha, i)$ , the objective measure affected by path sequence at node  $i$  following route  $\alpha$ , is updated in the same manner as in the a original algorithm, i.e.,  $c(SU\{j\}_\alpha, i) = c(S_\alpha, i) + c_{i,j}$ . 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  $\{(SU\{j\}, j), j \in \mathbf{P}\}$  are tested relative to whether to be stored or eliminated. Some elimination criteria are based solely on the feasibility of  $(SU\{j\}, j)$ , 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  $\mathbf{A}_a$ , then any candidate location of activity  $\mathbf{A}_a$  must not have been previously visited. This elimination is tested for all selective activity types,  $\mathbf{A}_a \in \mathbf{A}$ :

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

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

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

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

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

#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:

$$\text{if } j \in \mathbf{P}^+, \text{ then } d(S_a, i) + d_j \leq D$$

#6: time constraints must be respected:

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

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

#8: for  $i \in \mathbf{P}^-, j \in \mathbf{P}^-$ , one of paths,  $i - n \rightarrow j - n \rightarrow i \rightarrow j$  or  $j - n \rightarrow i - n \rightarrow i \rightarrow j$ , 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  $i \in \mathbf{P}^+$ ,  $j \in \mathbf{P}^-$ , path  $j - n \rightarrow i \rightarrow j \rightarrow n + i$  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  $i \in \mathbf{P}^-$ ,  $j \in \mathbf{P}^+$ , path  $i - n \rightarrow i \rightarrow j \rightarrow n + j$  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:

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

#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:

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

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  $l \in \mathbf{P}_p^+$ ) and one of the selective locations (any  $l \in \overline{\mathbf{P}_{A_a}^+}, A_a \in \mathbf{A}$ ) have been visited previously, and the arrival node  $j$  is home (if  $l \in \mathbf{S}$  and  $j \in \mathbf{P}^-$ ), 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*.

$$l \in \mathbf{S} \text{ for all } l \in \mathbf{P}_p^+, \text{ and } l \in \mathbf{S} \text{ for any } l \in \overline{\mathbf{P}_{A_a}^+} \text{ for all } A_a \in \mathbf{A}, \text{ and } j \in \mathbf{P}^-$$

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.

## 4. Examples

### 4.1 Case 1 : 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.,  $\mathbf{M}_p = \{1\}$ ,  $\mathbf{P}_p^+ = \{1\}$ ;  $n_p = 1$ , 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.,  $\mathbf{A} = \{\mathbf{A}_1\}$ ,  $m = 1$ , and that there are two potential locations for this activity  $\mathbf{P}_{A_1} = \{2, 3\}$ ;  $n_{A_1} = 2$ ; 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<sup>3</sup>. In this example:

$$\mathbf{M} = \mathbf{M}_p \cup \mathbf{A} = \{1, 2, 3\}; n = n_p + n_A = 3$$

$$\mathbf{P}_p^+ = \{1\}$$

$$\mathbf{P}_A^+ = \{2, 3\}$$

$$\mathbf{P}_p^- = \{4\}$$

$$\mathbf{P}_A^- = \{5, 6\}$$

$$\mathbf{P}^+ = \mathbf{P}_p^+ \cup \mathbf{P}_A^+ = \{1, 2, 3\}$$

$$\mathbf{P}^- = \mathbf{P}_p^- \cup \mathbf{P}_A^- = \{4, 5, 6\}$$

$$\mathbf{P}_p = \mathbf{P}_p^+ \cup \mathbf{P}_p^- = \{1, 4\}$$

$$\mathbf{P}_A = \mathbf{P}_A^+ \cup \mathbf{P}_A^- = \{2, 3, 5, 6\}$$

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

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

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

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

---

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



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).

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

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,  $c(SU\{j\}_\alpha, j) = c(S_\alpha, i) + F \cdot t_{i,j}$ , where  $(S_\alpha, i)$  is the state from previous iteration.

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

Travel Time Matrix  $t_{uw}$

$v \backslash u$	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  $\mathbf{P}$ —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.

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

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

$$\begin{aligned} T(S_\alpha, j) &= \{ \max(a_j, a_0 + t_{0,j}) \leq t(S_\alpha, j) \leq b_j, \\ &\quad t(S_\alpha, 0) + t_{0,j} \leq t(SU\{j\}_\alpha, j) \} \\ c(S_\alpha, j) &= c_{0,j} \\ d(S_\alpha, j) &= d_i \end{aligned}$$

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

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

$$\{(SU\{j\}, j), j \in \mathbf{P}\} \text{ where } (S, i) \text{ is the state from previous iteration}$$

Then the states are tested for elimination criteria, and if the state  $(SU\{j\}, j)$  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:

$$\begin{aligned} T(SU\{j\}_\alpha, j) &= T(S_\alpha, j) \cup \{ \max(a_j, T_i + s_i + t_{i,j}) \leq t(SU\{j\}_\alpha, j) \leq b_j, \\ &\quad t(S_\alpha, i) + s_i + t_{i,j} \leq t(SU\{j\}_\alpha, j) \} \\ c(SU\{j\}_\alpha, j) &= c(S_\alpha, i) + c_{i,j} \\ d(SU\{j\}_\alpha, j) &= d(S_\alpha, i) + d_j \end{aligned}$$

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,  $\alpha$ , from *Final Iteration* ( $k = 2(n_p + m) + 1$ ), solve the following optimization problem, and update the final cost.

$$\text{Minimize } f(T_0, T_1, \dots, T_{2n}, T_{2n+1}) \text{ such that } T(S_\alpha, 2n + 1)$$

Algorithm 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:

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

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

**Table 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, $T(S_{\alpha}, j)$ <sup>4</sup>	Previous Path index
$k = 1$	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
$k = 2$	8	{1 4}	4	2.75	$17.22 \leq T_4 \leq 21$ $T_1 + 9 + 0.22 \leq T_4$	1
$k = 3$	14	{1 3 4}	3	4.31	$17.47 \leq T_3 \leq 2$ $T_4 + 0.25 \leq T_3$	8
$k = 4$	34	{1 3 4 6}	6	5.88	$18.72 \leq T_6 \leq 21$ $T_3 + 1 + 0.25 \leq T_6$	14
$k = 5$	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,  $c(S_{\alpha}, j) + V \cdot (T_7 - T_0)$ . For example, for label of index 46, which traveled as: 7 (Label index 46)  $\leftarrow$  6 (Label index 34)  $\leftarrow$  3 (Label index 14)  $\leftarrow$  4 (Label index 8)  $\leftarrow$  1 (Label index 1)  $\leftarrow$  0, the following problem is solved to determine arrival times.

<sup>4</sup> 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.

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

subject to:

$$6 \leq T_0 \leq 22$$

$$8 \leq T_1 \leq 9$$

$$T_0 + 0.22 \leq T_1$$

$$17.22 \leq T_4 \leq 21$$

$$T_1 + 9 + 0.22 \leq T_4$$

$$17.47 \leq T_3 \leq 2$$

$$T_4 + 0.25 \leq T_3$$

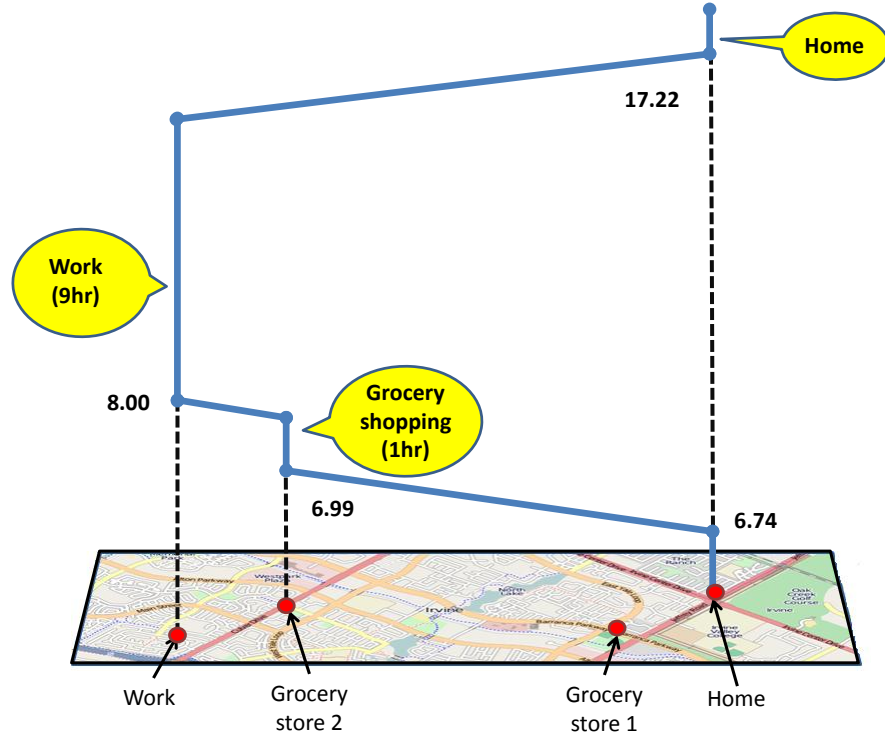
$$18.72 \leq T_6 \leq 21$$

$$T_3 + 1 + 0.25 \leq T_6$$

$$18.72 \leq T_7 \leq 22$$

$$T_6 \leq T_7$$

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 ( $T_0 = 6.74$ )  $\rightarrow$  grocery store 2 ( $T_3 = 6.99$ )  $\rightarrow$  work ( $T_1 = 8.00$ )  $\rightarrow$  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 1.



**Figure 1. Optimal Activity Pattern of Grocery Shopping Location Selection Involving a Single Vehicle**

#### 4.2 Case2: 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:

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

The household needs to complete two activities with pre-selected locations,  $N_p = \{1, 2\}$ ;  $n_p = 2$ , 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.,  $A = \{A_1\}$ ,  $m = 1$ , and that there are two potential locations for this activity  $P_{A_1} = \{3, 4\}$ ;  $n_{A_1} = 2$ ; 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 1 hour. In this example:

$$\mathbf{M} = \mathbf{M}_p \cup \mathbf{M}_A = \{1, 2, 3, 4\}; n = n_p + n_A = 4$$

$$\mathbf{S} = \mathbf{S}_p \cup \mathbf{S}_A = \{s_1, s_2, s_3, s_4\} = \{9, 0.1, 1, 1\}$$

$$\mathbf{P}_p^+ = \{1, 2\}$$

$$\mathbf{P}_A^+ = \{3, 4\}$$

$$\mathbf{P}_p^- = \{5, 6\}$$

$$\mathbf{P}_A^- = \{7, 8\}$$

$$\mathbf{P}^+ = \mathbf{P}_p^+ \cup \mathbf{P}_A^+ = \{1, 2, 3, 4\}$$

$$\mathbf{P}^- = \mathbf{P}_p^- \cup \mathbf{P}_A^- = \{5, 6, 7, 8\}$$

$$\mathbf{P}_p = \mathbf{P}_p^+ \cup \mathbf{P}_p^- = \{1, 2, 5, 6\}$$

$$\mathbf{P}_A = \mathbf{P}_A^+ \cup \mathbf{P}_A^- = \{3, 4, 7, 8\}$$

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

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

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

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

The travel time matrix is given as:

$v \backslash u$	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 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 3) form the basis of the master problem.

**Table 3. Admissible Paths of Grocery Shopping Location Selection for the master problem**

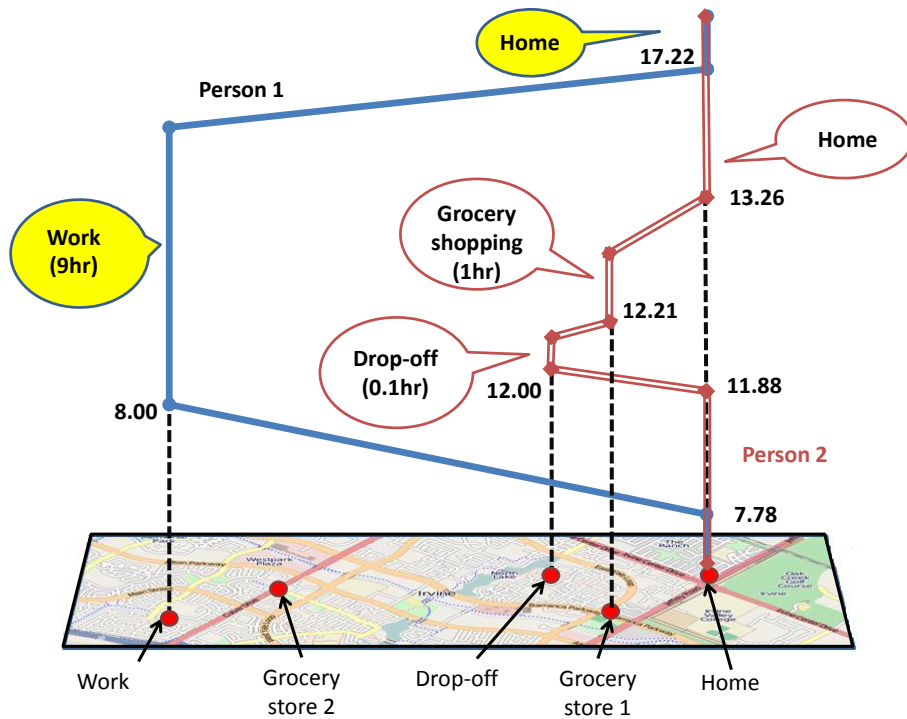
Path No., $r$	Visited nodes, S	Path Sequence and Arrival Times	Travel Disutility, $c_r$
0	{3 7 9}	home ( $T_0 = 6$ ) $\rightarrow$ grocery store 1 ( $T_3 = 6.05$ ) $\rightarrow$ home ( $T_7 = T_9 = 7.1$ )	17.13
1	{1 5 9}	home ( $T_0 = 7.78$ ) $\rightarrow$ work ( $T_1 = 8$ ) $\rightarrow$ home ( $T_5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

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

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

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

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 ( $T_0 = 7.78$ )  $\rightarrow$  work ( $T_1 = 8$ )  $\rightarrow$  home ( $T_5 = T_9 = 17.22$ ), and person 2 travel as path 7: home ( $T_0 = 11.88$ )  $\rightarrow$  drop off ( $T_2 = 12$ )  $\rightarrow$  grocery store 1 ( $T_3 = 12.21$ )  $\rightarrow$  home ( $T_6 = T_7 = T_9 = 13.26$ ). These results are depicted in Figure 2.



**Figure 2. Optimal Activity Pattern of Grocery Shopping Location Selection for a household with two vehicles**



### 4.3 Case3: 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:

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

where  $\Omega_1^V$  is the subset of activities that **cannot** be performed by vehicle/person  $v$ . Assume, for example, that person 1 is the person who needs to perform both the work as well as the grocery activities.<sup>5</sup>

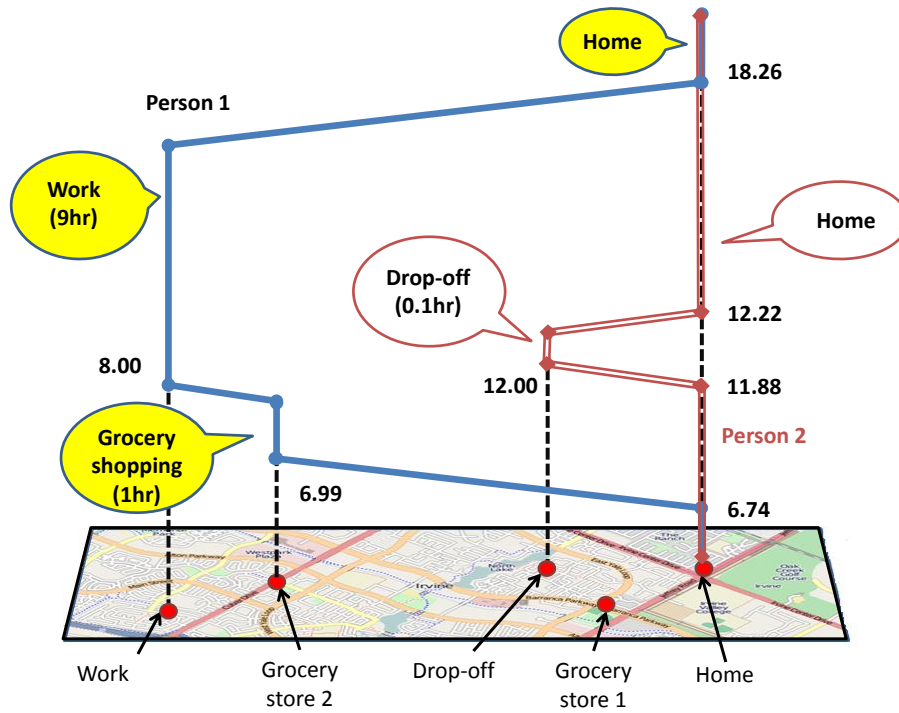
$$\Omega_0^V = \{ \}$$

$$\Omega_1^V = \{1, 3, 4\}$$

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 ( $T_0 = 6.74$ )  $\rightarrow$  grocery store 2 ( $T_4 = 6.99$ )  $\rightarrow$  work ( $T_1 = 8$ )  $\rightarrow$  home ( $T_5 = T_8 = T_9 = 17.22$ ), and person/vehicle 2 travels path 2, home ( $T_0 = 11.88$ )  $\rightarrow$  drop off ( $T_2 = 12$ )  $\rightarrow$  home ( $T_6 = T_9 = 12.22$ ), with the total cost of \$166.8 (Figure 3).

---

<sup>5</sup> Note that the notation starts from index 0.



**Figure 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 1, as shown in Algorithm 3.

### Initialization ( $k = 1$ )

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

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

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

$$\begin{aligned} T(S_\alpha, j) &= \{ \max(a_j, a_0 + t_{0,j}) \leq t(S_\alpha, j) \leq b_j, \\ &\quad t(S_\alpha, 0) + t_{0,j} \leq t(\text{SU}\{j\}_\alpha, j) \} \\ c(S_\alpha, j) &= c_{0,j} \\ d(S_\alpha, j) &= d_i \end{aligned}$$

### Recursion ( $2 \leq k$ )

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

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

Then the states are tested for elimination criteria, and if the state  $(\text{SU}\{j\}, j)$  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:

$$\begin{aligned} T(\text{SU}\{j\}_\alpha, j) &= T(S_\alpha, j) \cup \{ \max(a_j, T_i + s_i + t_{i,j}) \leq t(\text{SU}\{j\}_\alpha, j) \leq b_j, \\ &\quad t(S_\alpha, i) + s_i + t_{i,j} \leq t(\text{SU}\{j\}_\alpha, j) \} \\ c(\text{SU}\{j\}_\alpha, j) &= c(S_\alpha, i) + c_{i,j} \\ d(\text{SU}\{j\}_\alpha, j) &= d(S_\alpha, i) + d_j \end{aligned}$$

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

$$\text{if } j \in \bigcap_{v \neq r, r \in V} \Omega_r^V \text{ for any } j \in S \text{ then, } l \in \Omega_v^V, \text{ for all } l \in S, \text{ for } v \in V$$

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.

$$\begin{aligned} \text{if } j \in \bigcap_{v \neq r, r \in V} \Omega_r^V \text{ for any } j \in S \text{ then, } l \in S \text{ for all } l \in \bigcap_{v \neq r, r \in V} \Omega_r^V, l \neq j, l \in \\ P_P^+, \text{ or one of } l \in P_{A_a}^+ \text{ for } l \in \bigcap_{v \neq r, r \in V} \Omega_r^V, l \neq j, l \in P_P^+ \text{ for all } v \in V \end{aligned}$$

### Selection of Arrival Times

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

$$\text{Minimize } f(T_0, T_1, \dots, T_{2n}, T_{2n+1}) \text{ such that } T(S_\alpha, 2n + 1)$$

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

$$\text{if } j \in \bigcap_{v \neq r, r \in V} \Omega_r^V \text{ then, } l \in \Omega_v^V, \text{ for all } l \in S, \text{ for } v \in V$$

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 paper. 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 integrated in HAPP.

## 5. Case study with Orange County Travel Survey Data

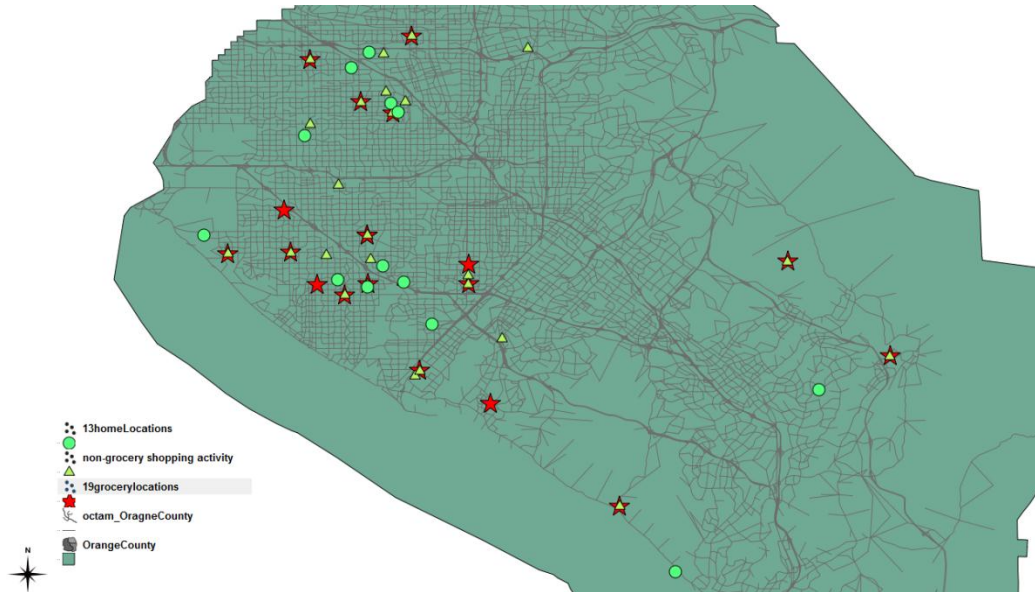
LSP-HAPP is applied to 13 households of single vehicle and single member households residing in Orange County, California, that have conducted one incidental shopping activity (includes shopping activities for grocery, medicine or house ware, but excludes such major shopping activities as furniture or 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 from Kang and Recker (2012), 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 4.



**Figure 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 from Kang and Recker (2012) 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 paper.

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.

## 6. Application of LSP-HAPP to Travel Pattern Generation in Activity-based Regional Forecasting Models

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 ( $A_1$ ) and grocery shopping ( $A_2$ )—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.,

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

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

$$c(\text{SU}\{j\}_\alpha, i) = \begin{cases} t_{i,j} & i = 0 \\ c(S_\alpha, i) + t_{i,j} & i \neq 0 \end{cases}$$

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

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  $s_{A_1} = 9$ ,  $s_{A_2} = 1$ , and time availability windows for each activity type are:

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

with corresponding return-home windows:

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

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

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

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

In this example:

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

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

$$P_p^+ = \{\}, P_p^- = \{\}$$

$$P_{A_1}^+ = \{1, 2\}, P_{A_1}^- = \{5, 6\}$$

$$P_{A_1} = P_{A_1}^+ \cup P_{A_1}^- = \{1, 2, 5, 6\}$$

$$P_{A_2}^+ = \{3, 4\}, P_{A_2}^- = \{7, 8\}$$

$$P_{A_2} = P_{A_2}^+ \cup P_{A_2}^- = \{3, 4, 7, 8\}$$

$$P_A^+ = \{1, 2, 3, 4\}$$

$$P_A^- = \{5, 6, 7, 8\}$$

$$P_A = P_p^+ \cup P_A^+ = \{1, 2, 3, 4\}$$

---

<sup>6</sup> 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)

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

$$P_P = P_P^+ \cup P_P^- = \{\}$$

$$P_A = P_A^+ \cup P_A^- = \{1, 2, 3, 5, 6, 7, 8\}$$

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

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

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  $t_{uv}$

$v \backslash u$	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  $\rightarrow$  grocery shopping at location 3 (6.05)  $\rightarrow$  work at location 1 (8)  $\rightarrow$  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 5).



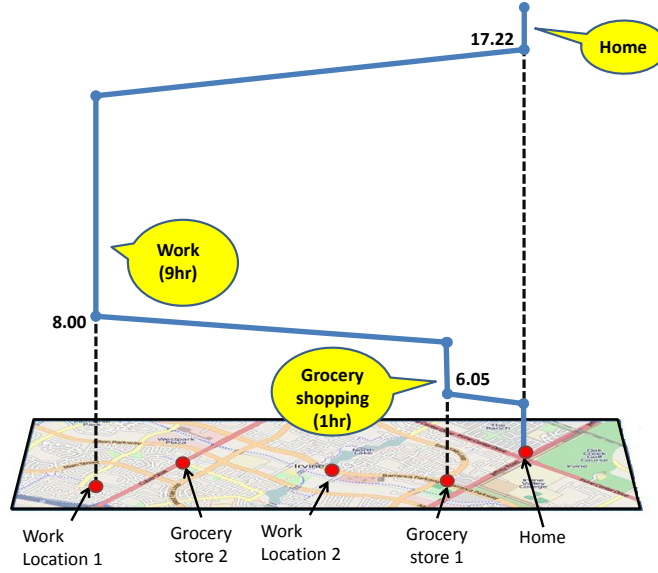


Figure 5. Synthetic Travel Pattern Generation Result

## 7. Comment on the General Column Generation Procedure for LSP-HAPP

Not only is finding the admissible path set,  $\Psi$ , 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:

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

subject to: (4) - (22)

where,

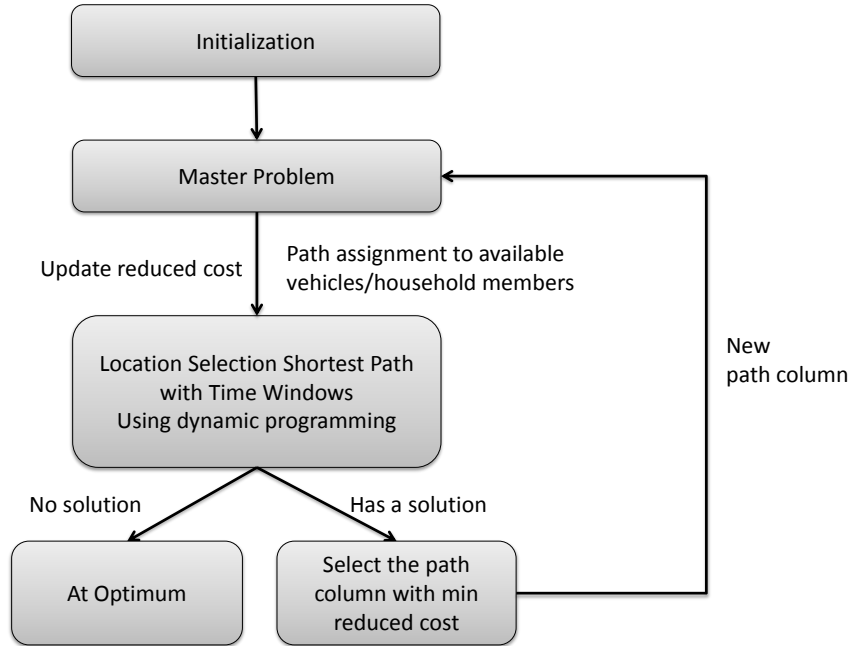
$\bar{c}_{ij}$ : the marginal cost of trip from node  $i$  to node  $j$

$\pi_i$ : the dual variables associated with (b-1),  $i \in \mathbf{P}^+$

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

Then, we can associate dual variables,  $\sigma_i$ , with each pre-selected activity node,  $\sigma_i = \pi_i, i \in \mathbf{P}^+$ . Similarly, dual variable of candidate locations of activity type  $\mathbf{A}_a$ , can be associated as,  $\sigma_i = \pi_{A_a}, a \in \mathbf{P}_{A_a}^+$ . 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  $\sigma_i = 0, i \in \mathbf{P}^-$ . 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 in this paper 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 6).



**Figure 6. Iterative Procedure of Column Generation of LSP-HAPP**

## 8. Conclusions and Discussion

In this paper, 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.

## **Acknowledgement**

This research was supported, in part, by grants from the University of California Transportation Center and the ITS Multi-Campus Research Program and Initiative on Sustainable Transportation. Their support is gratefully acknowledged.

## **References**

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, pp. 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: Policy and Practice*, 35, pp. 1-28

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

Chow, J. Y. J. and Liu, H. (2012), Generalized Profitable Tour Problems for an Online Activity Routing System, *Transportation Research Record: Journal of the Transportation Research Board*, under review

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: Methodological*, in press, doi: 10.1016/j.trb.2011.11.005

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, pp. 89-101

Desrosiers, D., 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, pp. 301-325

Desrosiers, D., Soumis, F., Desrochers, M. (1984), Routing with time windows by column generation, *Networks*, 14, pp. 545-565

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

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

Fotheringham, A. S. (1988), Consumer Store Choice and Choice Set Definition, *Marketing Science*, 7-3, pp. 299-310

- Gan, L. P. and Recker, W. W. (2008), A mathematical programming formulation of the Household Activity Rescheduling Problem. *Transportation Research, Part B: Methodological*, 42, pp.571-606
- Gan, L. P. and W. Recker (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, pp. 1-11
- Ghiani, G., Improta, G., 2000. An efficient transformation of the generalized vehicle routing problem. *European Journal of Operational Research* 122(1), 11-17.
- Kang, J. E. and Recker, W. W. (2012), Strategic Hydrogen Fuel Cell Charging Station Location Analysis with Scheduling and Routing Considerations of Individual Vehicles, ITS Working Paper Series, UCI-ITS-WP-12-2, University of California, Irvine.
- Kitamura, R. (1984), Incorporating Trip Chaining into Analysis of Destination Choice, *Transportation Research, Part B: Methodological*, 18-1, pp.67-81
- 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, pp. 291-308
- 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., pp. 37-54.
- 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, pp. 47-54
- Recker, W.W. (1995), The Household Activity Pattern Problem: General formulation and solution, *Transportation Research*, 29B, 1, pp. 61-77

- Recker, W.W. (2001), A bridge between travel demand modeling and activity-based travel analysis, *Transportation Research , Part B: Methodological*, 35, 481-506
- Recker, W.W., C. Chen and M.G. McNally (2001), Measuring the impact of efficient household travel decisions on potential travel time savings and accessibility gains. *Transportation Research, Part A: Policy and Practice*, 35, 339-369
- 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, pp. 483-501
- Recker, W. W. and Kostyniuk, L. (1978), Factors Influencing Destination Choice for the Urban Grocery Shopping Trips, *Transportation*, 7, pp. 19-33
- Recker, W. W. and Parimi, A. (1999), Development of a microscopic activity-based framework for analyzing the potential impact of TCMs on vehicle emissions, *Transportation Research, Part D: Transport and Environment*, 4, pp. 357-378
- Roy, J. R. and Thill, J. (2004), Spatial Interaction Modelling, *Regional Science*, 83, pp. 339-361
- Solomon, M. and Desrosiers, J. (1988), Time window constrained routing and scheduling problems. *Transportation Science*, 22, pp. 1-13
- Thill, J. (1992), Choice Set Formation for Destination Choice Modelling, *Progress in Human Geography*, 16, pp. 631-382

## Appendix A: Notation

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

$\mathbf{A} = \{\mathbf{A}_1, \dots, \mathbf{A}_a, \dots, \mathbf{A}_m\}$ : 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  $A_a$

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

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

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

$\mathbf{P}_{A_1}^+ = \{n_P + 1, n_P + 2, \dots, n_P + n_{A_1}\}$ : the set designating “potential” locations at which activity  $\mathbf{A}_1$  may be performed—one, and only one, may be selected.

$\mathbf{P}_{A_a}^+ = \{n_P + \sum_{a=1}^{a-1} n_{A_a} + 1, n_P + \sum_{a=1}^{a-1} n_{A_a} + 2, \dots, n_P + \sum_{a=1}^a n_{A_a}\}$ ,  $\mathbf{A}_a \in \mathbf{A}$ ,  $\mathbf{A}_a \neq \mathbf{A}_1$  : the set designating “potential” locations at which activity  $\mathbf{A}_a$  may be performed—one, and only one, may be selected for each activity  $\mathbf{A}_a \in \mathbf{A}$ ,  $\mathbf{A}_a \neq \mathbf{A}_1$ .

$\mathbf{P}_A^+ = \cup_{A_a \in \mathbf{A}} \mathbf{P}_{A_a}^+ = \{n_P + 1, n_P + 2, \dots, n_P + n_{A_1}, n_P + n_{A_1} + 1, \dots, n_P + \sum_{a=1}^m n_{A_a} = n\}$  : the set designating “potential” locations at which all activities with unspecified locations,  $\mathbf{A}_a \in \mathbf{A}$ , may be performed.

$\mathbf{P}^+ = \mathbf{P}_P^+ \cup \mathbf{P}_A^+ = \mathbf{P}_P^+ \cup \mathbf{P}_{A_1}^+ \cup \dots \cup \mathbf{P}_{A_m}^+ = \{1, 2, \dots, i, \dots, n\}$ : the set designating locations at which the combined sets of activities with predetermined and multiple candidate locations may be performed.

$\mathbf{P}_P^- = \{n + 1, n + 2, \dots, n + n_P\}$ : 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  $\mathbf{P}_P^-$  is “home”.)

$\mathbf{P}_{A_1}^- = \{n + n_P + 1, n + n_P + 2, \dots, n + n_P + n_{A_1}\}$ : the set designating the ultimate destinations of the “return to home” trips for the  $\mathbf{A}_1$  activity—each element is paired to the location selected for activity  $\mathbf{A}_1$ . (It is noted that the physical location of each element of  $\mathbf{P}_{A_1}^-$  is “home”.)



- $P_{A_a}^- = \{n + n_P + \sum_{a=1}^{a-1} n_{A_a} + 1, n + n_P + \sum_{a=1}^{a-1} n_{A_a} + 2, \dots, n + n_P + \sum_{a=1}^a n_{A_a}\}$   $A_a \in A$ ,  $A_a \neq A_1$ : set designating ultimate destination of the "return to home" trip for each activity  $A_a$ —each element is paired to the location selected for each activity  $A_a \in A$ ,  $A_a \neq A_1$ . (Note that the physical location of each element of  $P_{A_a}^-$  is "home".)
- $P_{\bar{A}} = \cup_{A_a \in A} P_{A_a}^- = \{n + 1, n + 2, \dots, n + i, \dots, n + n_P + \sum_{a=1}^m n_{A_a} = n_P + n_A = 2n\}$ : set designating all possible ultimate destinations of the "return to home" trips for "potential" locations at which all activities with unspecified locations,  $A_a \in A$ . (It is noted that the physical location of each element of  $P_{\bar{A}}$  is "home".)
- $P^- = P_P^- \cup P_{\bar{A}} = P_P^- \cup P_{A_1}^- \cup \dots \cup P_{A_m}^- = \{n + 1, n + 2, \dots, 2(n_P + n_A) = 2n\}$ : 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  $P^-$  is "home".)
- $P = P^+ \cup P^-$ : set of nodes comprising both predetermined locations and candidate locations of activities, and their corresponding "return home nodes".
- $N = \{0, P, 2n + 1\}$ : set of all nodes, including those associated with the initial and final return to home.
- $V = \{1, 2, \dots, v, \dots, |V|\}$ : 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  $w$ .
- $c_{uw}^v$ : travel cost from location of activity  $u$  to the location of activity  $w$  by vehicle  $v$ .
- $B_C$ : household travel cost budget.
- $B_t^v$ : travel time budget for the household member using vehicle  $v$ .

Appendix B

**Table A-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, $T(S_{\alpha}, j)$ <sup>7</sup>	Previous Path index
<b>k = 1</b> (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
<b>k = 2</b> (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	$17.22 \leq T_4 \leq 21$ $T_1 + 9 + 0.22 \leq T_4$	1
	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
<b>k = 3</b> (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

<sup>7</sup> 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.

	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	$17.22 \leq T_5 \leq 22$ $T_1 + 9 + 0.22 \leq T_5$	4
	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
<b>k = 4</b> <b>(12 labels)</b>	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
<b>k = 5</b> <b>(12 labels)</b>	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	$17.22 \leq T_7 \leq 22$	26

---

				$T_4 \leq T_7$	
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