Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia
Social and Behavioral Sciences

Procedia - Social and Behavioral Sciences 80 (2013) 157 – 185

International Symposium on Transportation & Traffic Theory 20

On activity-based network design problems

Jee Eun Kang[a], Joseph Y.J. Chow[b*], Will W. Recker[a]

[a]*Institute of Transportation Studies, University of California, Irvine, 4000 AIR Building, Irvine, CA 92697, USA*

[b]*Department of Civil Engineering, Ryerson University, 350 Victoria Street, Toronto, ON, M5B 2K3, Canada*

Abstract

This paper 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 compared to a conventional model, 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.

Selection and peer-review under responsibility of Delft University of Technology

Keywords: activity based model; network design; location routing problem; HAPP; pickup and delivery problem; bi-level problem

1. Background

Network design problems (NDPs) are a class of optimization models related to strategic or tactical planning of resources to manage a network *[1]*. Even for purposes of improving road networks for commuters *[2]* and despite the complexity of traveler choices *[3]*, 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

* Corresponding author. Tel.: +1-416-979-5000 ext. 7618 ; fax: +1-416-979-5122 .
E-mail address: joseph.chow@ryerson.ca

Selection and peer-review under responsibility of Delft University of Technology
doi:10.1016/j.sbspro.2013.05.011

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 *[4]*, Heydecker's *[5]* NDP with dynamic user equilibrium (DUE), the linear DUE-NDP from Ukkusuri and Waller *[6]*, dynamic toll pricing problem with route and departure time choice *[7]*, and the reliability maximizing toll pricing problem with dynamic route and departure time choice *[8]*. 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 *[9]* 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 *[10]* formulated such a model with heterogeneous traveler values of time. Poon *et al. [11]* presented a dynamic equilibrium model for schedule based transit assignment. Hamdouch and Lawphongpanich *[12]* 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 *[13]*. 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 (*[3]*, *[14]*). 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. *[15]*, *[16]*). 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 *[17]*, with subsequent studies on dynamic rescheduling/rerouting of those itineraries *[18]* and calibration of the activity route assignment models (*[19]*, *[20]*). The second is the aggregate time-dependent activity-based traffic assignment model of Lam and Yin *[21]*. 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 *[22]*. 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. [23]* and a more recent contribution by Nagy and Salhi *[24]*. Problem types developed over the years that may be applicable to activity-based network design in transportation planning include: stochastic LRP (*[25]*), where there is more than one planning horizon with time-dependent customer locations and demand; LRP with a mixed fleet *[26]* for multimodal network consideration; location-routing-inventory *[27]* for modeling activity types as inventory-based needs that are fulfilled periodically; and LRP with nonlinear costs *[28]* 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 by Kang and Recker *[29]*. 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.

In Section 2, several examples and insightful paradoxes are used to illustrate why an activity-based approach is necessary at the tactical and planning level NDP. Section 3 introduces the formulation as a bi-level structure with shortest path allocation and disaggregated subproblems per household. While the inspiration of the formulation is from Perl and Daskin's LRP, key differences are also noted. An alternative model with activity/destination choice is also provided. Since the problem is nonconvex and NP-hard, Section 4 presents a heuristic solution method and suggestions for meta-heuristics, using a simple test network to demonstrate the method and the sensitivity of underlying assumptions. Section 5 presents a larger-scale case study of the Orange County, California region as a test network to demonstrate the model's practical application to systematic improvement.

2. Motivating examples

The argument that we provide here, much like Perl and Daskin *[22]* 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 *[17]* 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 [20], 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.

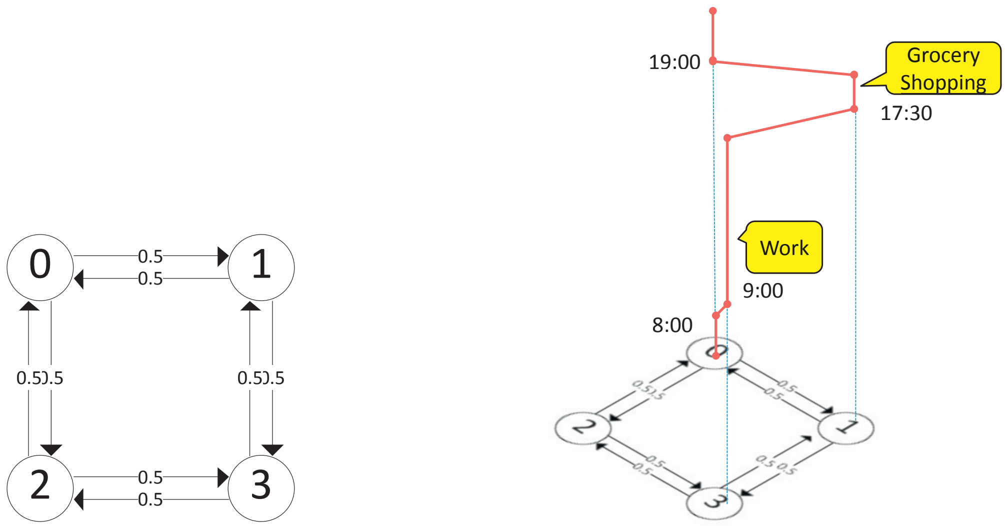
2.1. Departure time choice and itinerary re-sequencing

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 [17] 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.

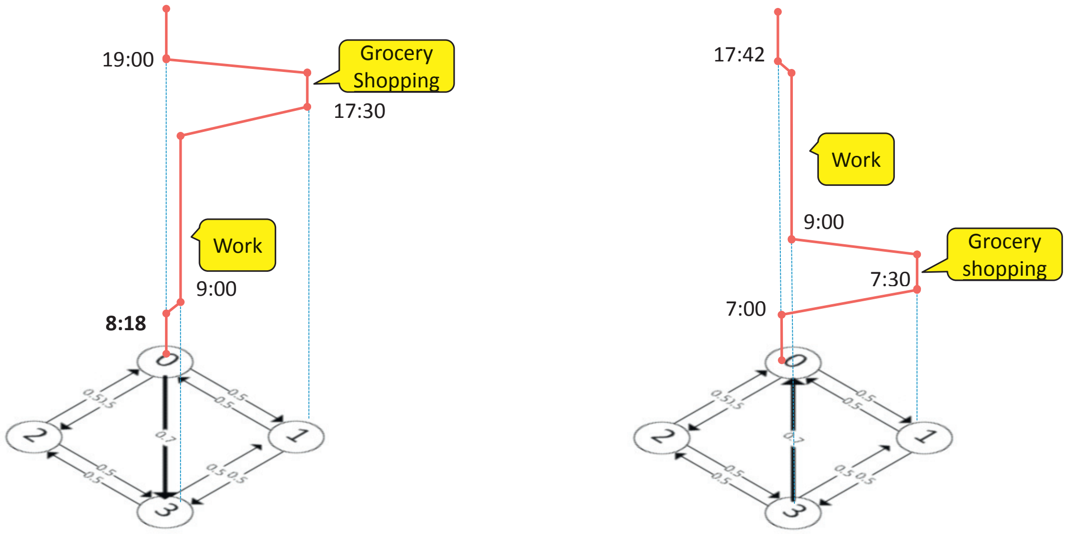
Table 1. Case 1 household characteristics

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

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



(a) Current Network (b) Optimal Household Activity Pattern



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

Fig. 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 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 1-(d), the optimal itinerary results in a re-sequence of activities as well as an adjustment in departure times.

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 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 *[17]* and calibrated empirically by Chow and Recker *[20]* 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 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 2. Case 2 household characteristics

Household	**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

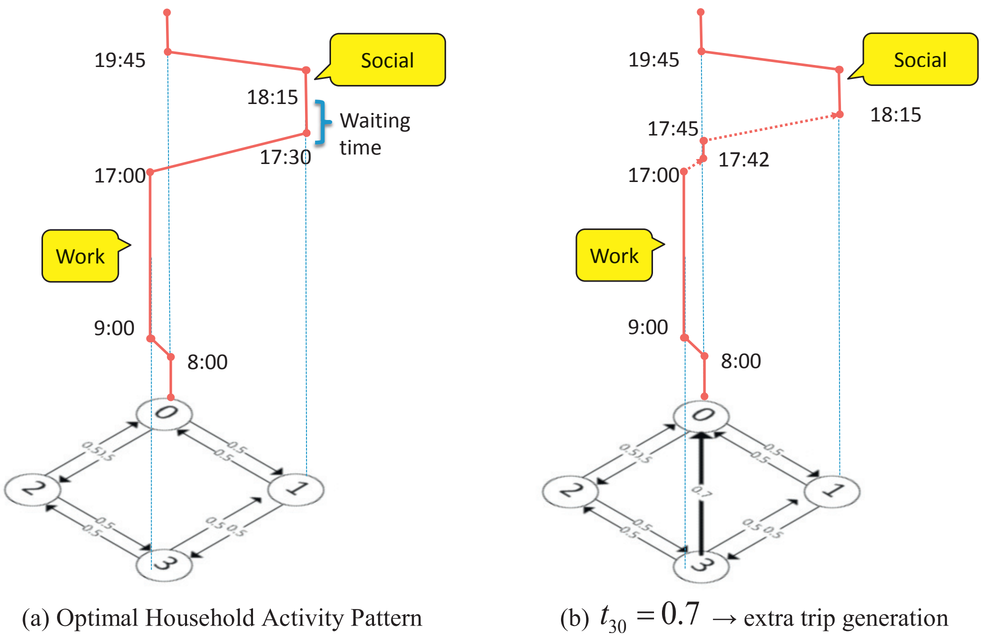


Fig. 2. The optimal household activity patterns for Case 2

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 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.

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 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 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.

3. Proposed NDP-HAPP model

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:

N set of all nodes in the analysis
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:
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.
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 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. *[30]*, *[31]*). 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 *[17]*, we define the following sets and parameters that are specific for each household, $h \in \mathbf{H}$:

$\boldsymbol{\beta}_h = \{\beta_h^a, \beta_h^b, ...\}$ 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".)

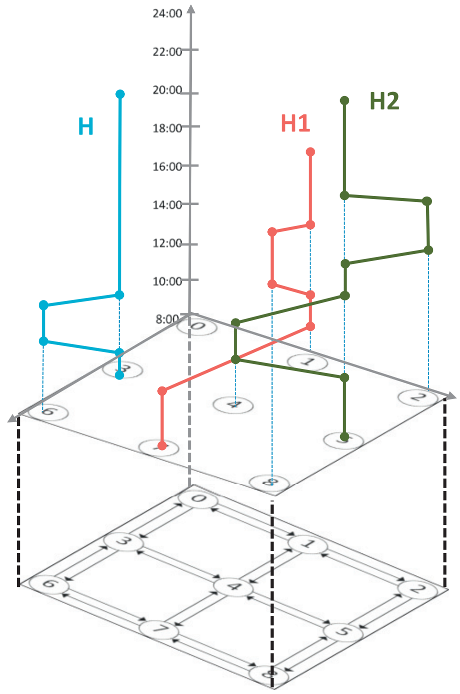


Fig. 3. Bi-level interactions between the infrastructure and activity networks

$\left[a_u^h, b_u^h\right]$	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 $\left(\delta_{uw,ij}\right)^*$ is the solution of a shortest path problem for each activity link (i, j), i.e.

Shortest Path Allocation Problem

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

subject to

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

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

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.

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

$c_{uw}^{v,h} = c_{uw}^{v,h}(\mathbf{z})$ 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.

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

$f_{ij} = f_{ij}(\mathbf{X})$ 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.

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

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 (1) – (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 *[2]*. 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 (7).

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

where

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

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 *[1]*. 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)

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

subject to:

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

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

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

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

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

where

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

Equations (9) – (10) require each path $(u,w) \in \mathbf{K}$ to satisfy the given OD demand. Equations (11) simply show the conservation of flows for intermediate nodes. Equation (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 paper, 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 (15) instead of having to solve Equations (1) – (3) separately.

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

The decomposed lower-level HAPP (dHAPP) problem is shown in Equations (16) – (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 *[17]* if travel time/cost factors are not functions of the allocated shortest path. More complex variations presented in *[17]* 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

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

Subject to
(A1) – (A26)

where

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

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

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

As discussed in other HAPP model studies, the objective shown in Equation (16) is just one example of 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 (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.

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 *[32]*) and generalized vehicle routing problem *[33]* 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 (*[29]*, *[34]*), and a "profitable tour" variation of this approach was developed for activity-based traveler information systems *[35]* and for testing algorithms in scenario analysis *[36]*.

In GHAPP, the constraints in Equation (A1) are modified to Equation (A1-1). 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.

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

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.

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

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

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

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

3.4. Decomposition solution algorithm

There are many different types of solution algorithms developed for LRPs *[24]*, 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., *[37] – [39]*), the suggested decomposition method is highly adaptable to different types of activity-based models.

The decomposed problems remain computationally challenging, particularly the NP-hard (*[17]*) HAPP. Because these problems are widely studied, there are various methods available. Geoffrion and Graves *[40]* are referred for network design problems, and Cordeau and Laporte *[41]* 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 *[22]* in the context of Location Routing Problems and the Iterative Optimization Assignment (IOA) algorithm in Yang and Bell *[2]* in the context of bi-level Network Design Problems. Perl and Daskin *[22]* 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 *[42]* 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 in this paper, 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 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 *[21]* 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 (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 (4) – (6) are maintained throughout this iterative process. The same algorithm can be applied to NDP-GHAPP.

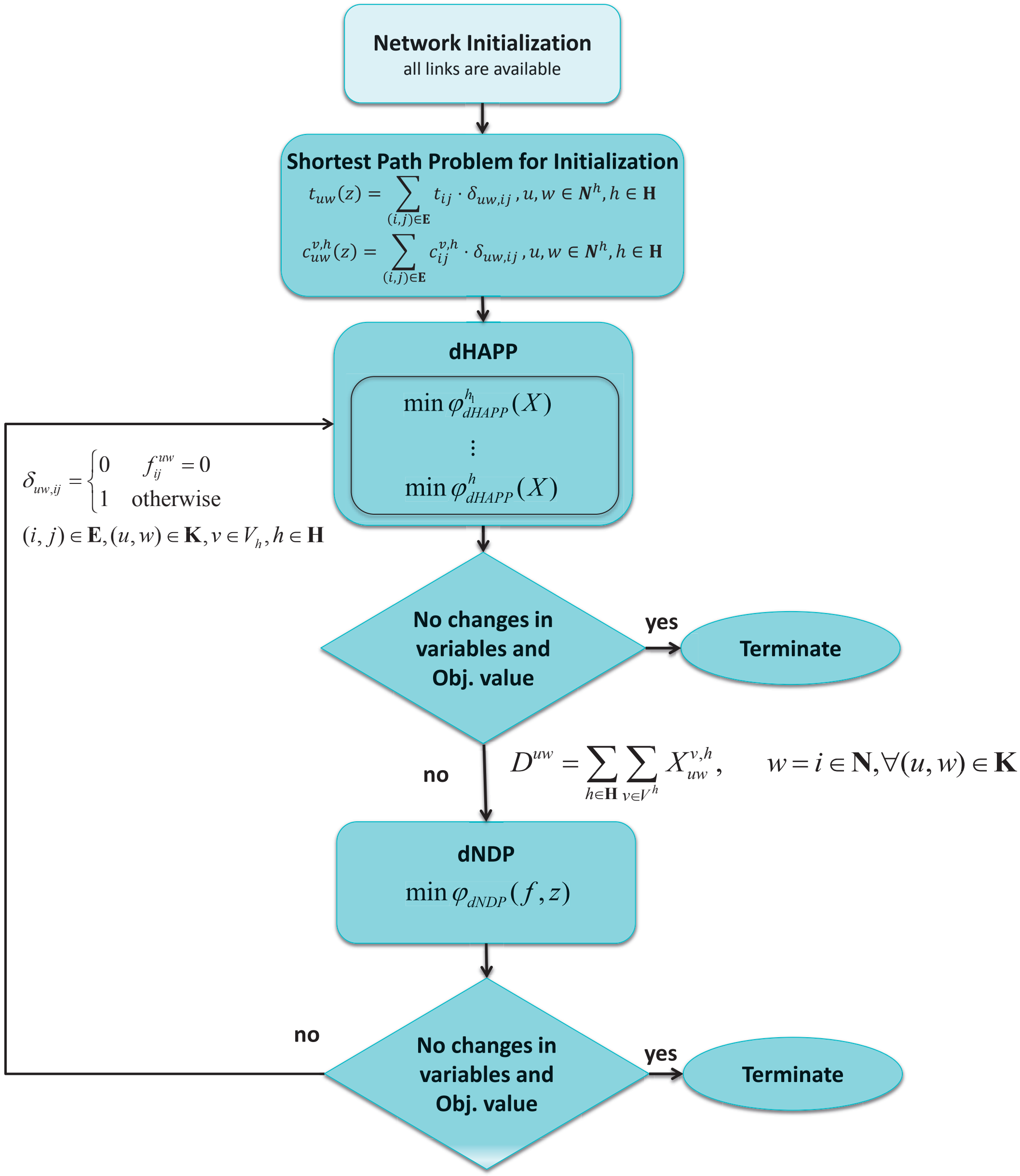


Fig. 4. A decomposition solution method for NDP-HAPP

4. Numerical examples

4.1. Simple example with known optimal solution

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 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

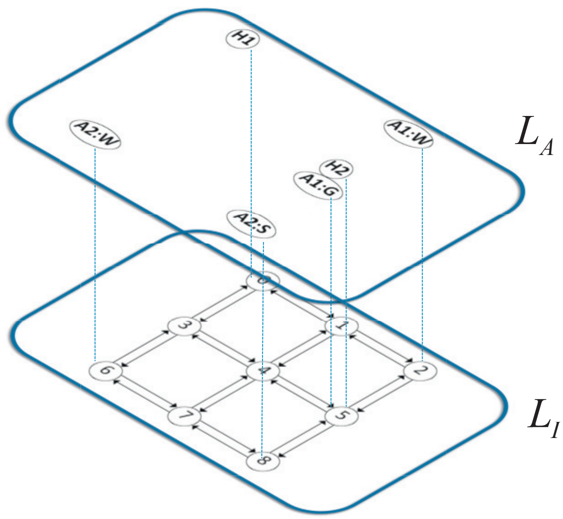
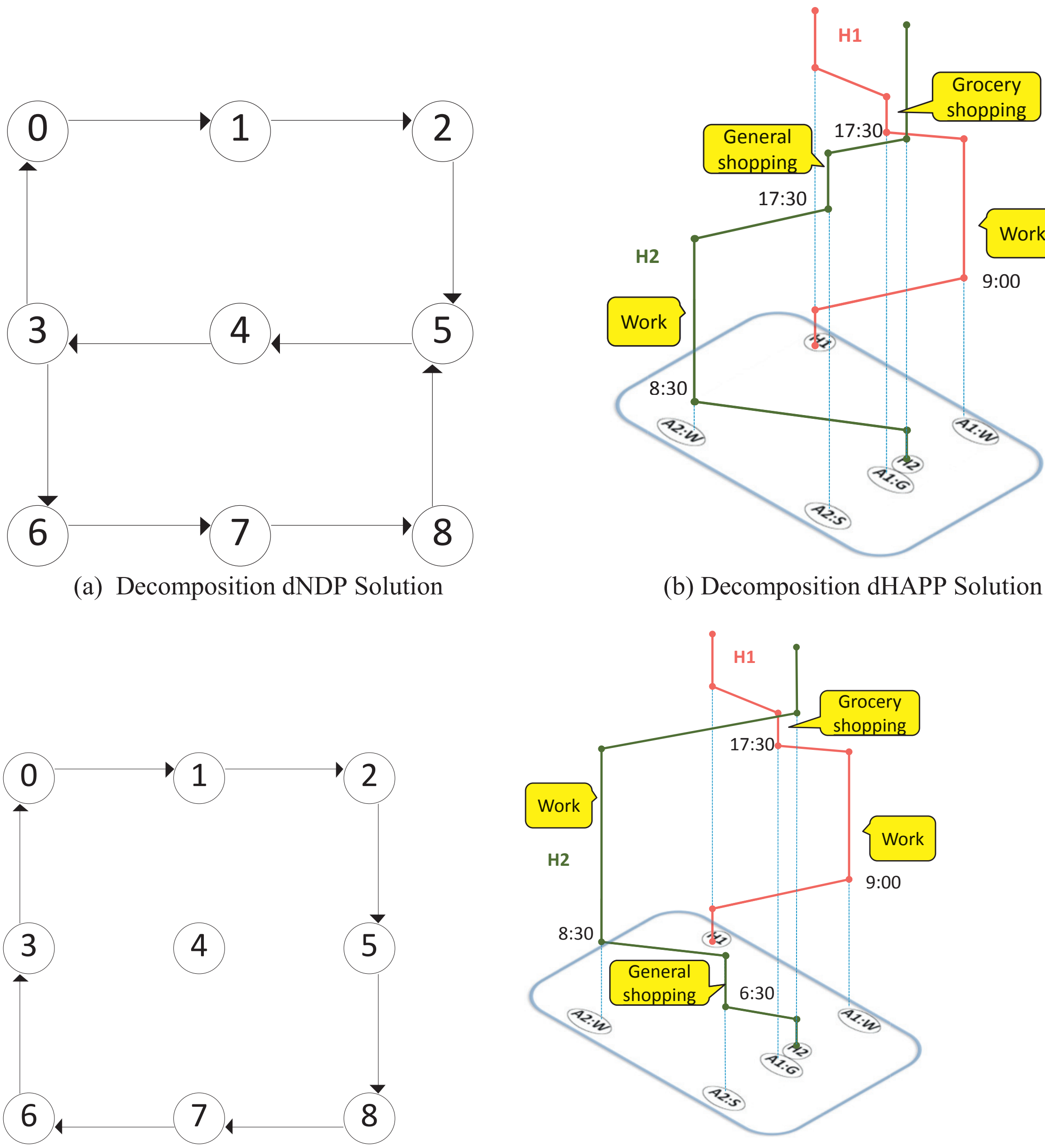


Fig. 5. Supernetwork depiction.

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 6 shows the solution from the proposed method (6-(a), 6-(b)) and the actual optimal solution (6-(c), 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 4.



(a) Decomposition dNDP Solution (b) Decomposition dHAPP Solution

(c) Optimal dNDP Solution (d) Optimal dHAPP Solution

Fig. 6. NDP-HAPP decomposition solution comparison to enumerated exact solution

Table 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.

4.2. Simple example: generalized HAPP

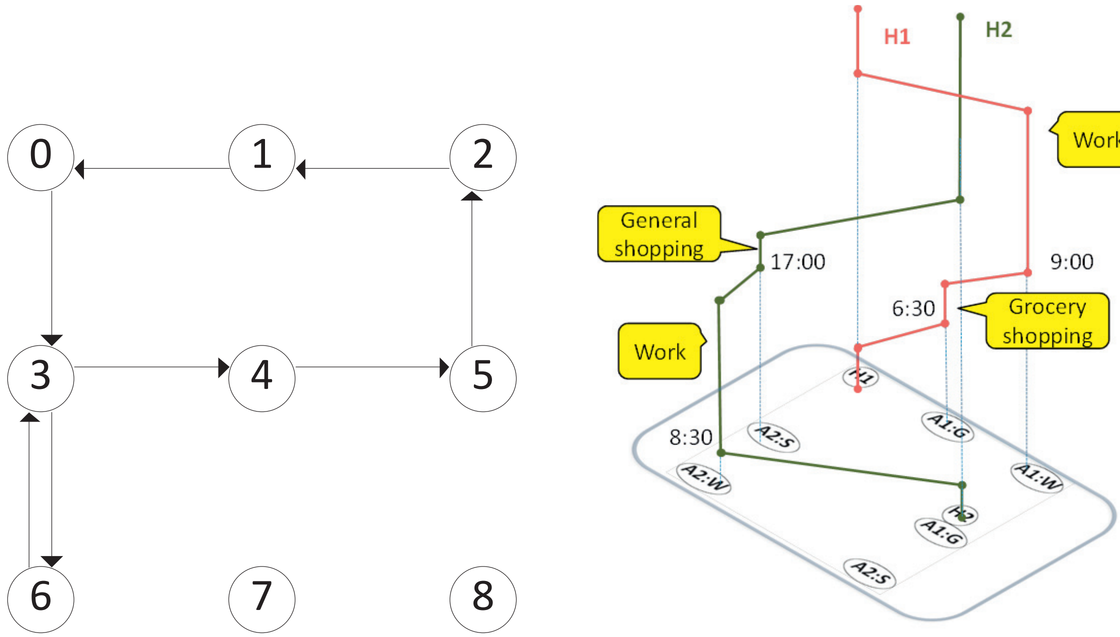
Using the generalized model allows us to include behavioral changes in destination choice as well as routing/scheduling of activities with respect to network design decisions. Following the example in Section 4.1, assume that there are two grocery shopping locations, node 1 and node 5, $P^{+}_{A_a=GroceryShopping}=\{1,5\}$, 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. 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. Detailed computational illustration of the NDP-SHAPP 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 objective values: HH1: 2 HH2: 3	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) Update each dHAPP objective values: HH1: 4 HH2: 4	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 objective values: HH1: 3 HH2: 4	NA[3]
Final Objective	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

Fig. 7. NDP-SHAPP example enumerated optimal solution

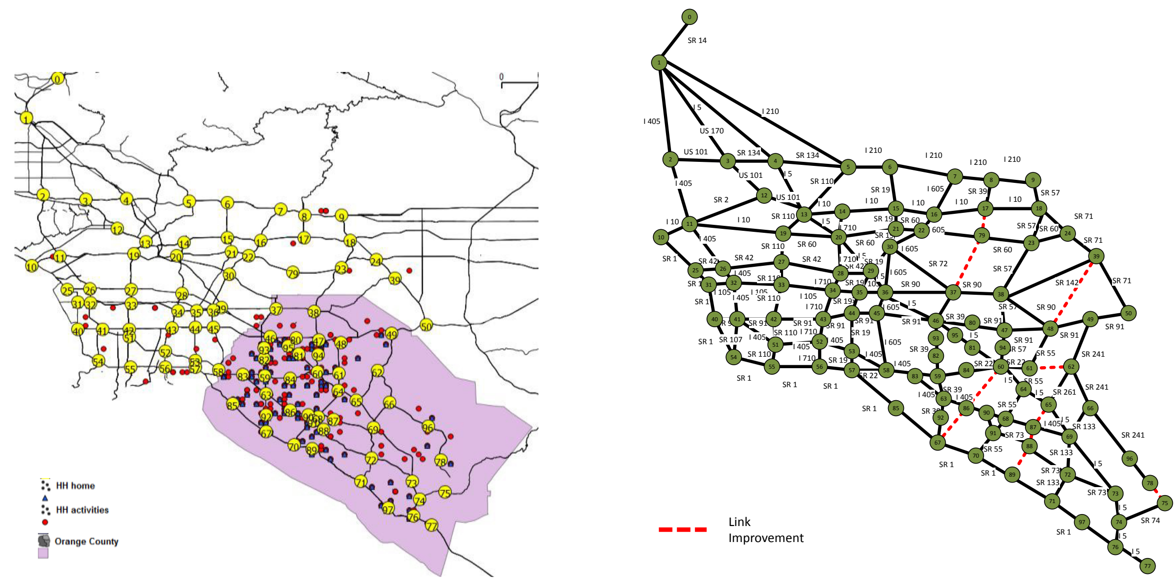
4.3. Large network example: NDP-HAPP

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 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 8-(b).

Specifications of each candidate link are in Appendix 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 *[43]* 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 $\beta_h^E, \beta_h^T, \beta_h^D$:

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



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

Fig. 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 *[20]*. 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 *[20]*. 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 from Kang and Recker *[34]*, which adopted the method from Recker and Parimi *[44]* with slight modifications.

In Table 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 9.

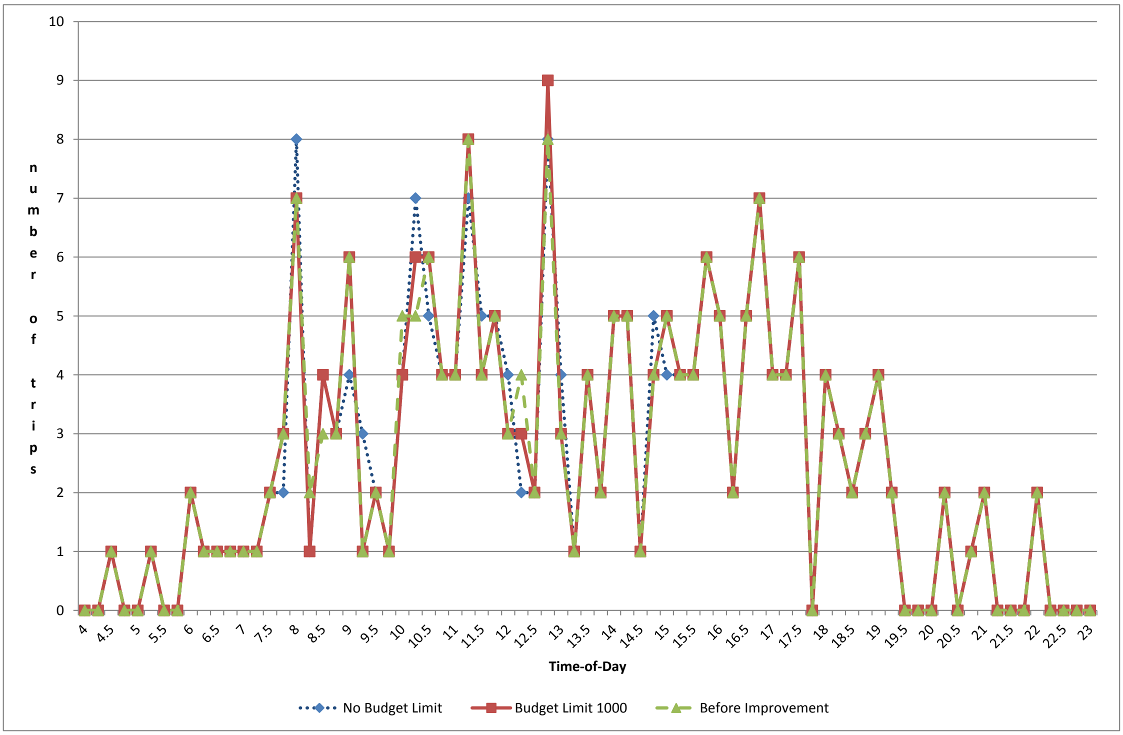


Fig. 9. Comparison of activity arrival time histograms

As shown in Figure 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.

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.

Table 6. Large-scale NDP-HAPP results

Budget	NDP-HAPP							Conventional NDP	
	# iterations	Link Construction Decision	dNDP objective	dHAPP objective	# total trips (# intrazonal)	# HHs affected	Computation time (seconds)[1]	Link Construction Decision	NDP objective
Before Improvement	NA	NA	27.02	616.49	199 (76)	NA	NA	NA	27.02
1000	2	8988, 7875, 7578	25.99	609.58	199 (76)	5/60	306	8988, 7875, 7578	25.99
2000	2	8988, 7875, 7578, 7937, 8660, 6786, 8887	25.30	606.51	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	604.49	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, 8765, 8788	24.79	604.12	199 (76)	17/60	196	8988, 7875, 7578, 7937, 8660, 6786, 8887, 6086, 8667, 8889, 6162, 6589, 8765, 8788	24.79
5000	1	8988, 7875, 7578, 7937, 8660, 6786, 8887, 6086, 8667, 8889, 6162, 6589, 8765, 8788, 6261	24.79	604.11	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	604.11	199 (76)	17/60	215	All	24.79

[1] 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.

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 *[21]* 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.

Acknowledgments

This research was supported, in part, by grants from the University of California Transportation Center and the Institute of Transportation Studies Multi-Campus Research Program and Initiative on Sustainable Transportation. Dr. Joseph Chow was partially supported by the Canada Research Chairs program. Their support is gratefully acknowledged.

References

[1] Magnanti, T.L., & Wong, R.T. (1984). Network design and transportation planning: models and algorithms. Transportation Science, 18(1), 1-55.
[2] Yang, H. & Bell, M. G. H. (1998). Models and algorithms for road network design: a review and some new development. Transport Review, 18(3), 257-278.
[3] Recker, W.W. (2001). A bridge between travel demand modeling and activity-based travel analysis. Transportation Research Part B, 35(5), 481-506.
[4] Waller, S.T., & Ziliaskopoulos, A.K. (2001). Stochastic dynamic network design problem. Transportation Research Record, 1771, 106-113.
[5] 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.
[6] Ukkusuri, S.V., & 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.
[7] Joksimovic, D., Bliemer, M.C.J., Bovy, P.H.L., & Verwater-Lukszo, Z. (2005). Dynamic road pricing for optimizing network performance with heterogeneous users. Proc. IEEE Networking, Sensing and Control, 407-412.
[8] Li, H., Bliemer, M.C.J., & Bovy, P.H.L. (2007). Optimal toll design from reliability perspective. Proc., 6th Triennnial Conference on Transportation Analysis, Phuket, Thailand.
[9] Crainic, T.G. (2000). Service network design in freight transportation. European Journal of Operational Research, 122(2), 272-288.
[10] Tong, C.O., & Wong, S.C. (1998). A stochastic transit assignment model using a dynamic schedule-based network. Transportation Research Part B, 33(2), 107-121.
[11] Poon, M.H., Wong, S.C., & Tong, C.O. (2004). A dynamic schedule-based model for congested transit networks. Transportation Research Part B, 38(4), 343-368.
[12] Hamdouch, Y., & Lawphongpanich, S. (2008). Schedule-based transit assignment model with travel strategies and capacity constraints. Transportation Research Part B, 42(7-8), 663-684.
[13] Hamdouch, Y., & Lawphongpanich, S. (2010). Congestion pricing for schedule-based transit networks. Transportation Science, 44(3), 350-366.
[14] Pinjari, A.R., & 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.
[15] Lin, D.Y., Eluru, N., Waller, S.T., & Bhat, C.R. (2008). Integration of activity-based modeling and dynamic traffic assignment. Transportation Research Record, 2076, 52-61.
[16] Konduri, K.C. (2012). Integrated model of the urban continuum with dynamic time-dependent activity-travel microsimulation: framework, prototype, and implementation. Unpublished Ph.D. dissertation, Arizona State University.
[17] Recker, W.W. (1995). The household activity pattern problem: general formulation and solution. Transportation Research, 29B(1), 61-77.
[18] Gan, L.P., & Recker, W.W. (2008). A mathematical programming formulation of the household activity rescheduling problem. Transportation Research Part B, 42(6), 571-606.
[19] Recker, W.W., Duan, J., & Wang, H. (2008). Development of an estimation procedure for an activity-based travel demand model. Computer-Aided Civil and Infrastructure Engineering, 23(7), 83-501.
[20] Chow, J.Y.J., & Recker, W.W. (2012). Inverse optimization with endogenous arrival time constraints to calibrate the household activity pattern problem. Transportation Research Part B, 46(3), 463-479.
[21] Lam, W.H.K., & Yin, Y. (2001). An activity-based time-dependent traffic assignment model. Transportation Research Part B 35(6), 549-574.
[22] Perl, J., & Daskin, M.S. (1985). A warehouse location-routing problem. Transportation Research, 19B(5), 381-396.
[23] Min, H., Jayaraman, V. & Srivastava, R. (1998). Combined location-routing problems: a synthesis and future research directions. European Journal of Operations Research, 108(1), 1-15.
[24] Nagy, G., & Salhi, S. (2007). Location-routing: issues, models and methods. European Journal of Operations Research, 177(25), 649-672.
[25] Laporte, G., & Dejax, P. J. (1989). Dynamic location-routing problems. Journal of Operational Research Society, 40(5), 471-482.
[26] Wu, T.H., Low, C., & Bai, J.W. (2002). Heuristic Solutions to multi-depot location-routing problems. Computers and Operations Research, 29(10), 1393-1415.
[27] Liu, S.C., & 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.
[28] Melachovsky, J., Prins, C. & Wolfler Calvo, R. (2005). A metaheuristic to solve a location-routing problem with non-linear costs. Journal of Heuristics, 11, 375-391.
[29] Kang, J.E., & Recker, W.W. (2012a). 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.
[30] TRANSIMS (accessed in 2012), TRansportation ANalysis SIMulation System, http://code.google.com/p/transims/
[31] Arentze, T., & Timmermans, H. (2004). Multistate supernetwork approach to modeling multi-activity, multimodal trip chains. Geographical information science, 18(7), 631-651.

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

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

[34] Kang, J.E., & Recker, W.W. (2012b). The location selection problem for the household activity pattern problem. ITS Working Paper Series, UCI-ITS-WP-12-1, University of California, Irvine.

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

[36] Chow, J.Y.J. (2013). Activity-based travel scenario analysis with routing problem reoptimization. Computer-Aided Civil and Infrastructure Engineering, in press.

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

[38] Bhat, C.R., Guo, J.Y., & Srinivasan, S. (2004). Comprehensive econometric microsimulator for daily activity-travel. Transportation Research Record, 1894, 57-66.

[39] Balmer, M., Axhausen, K.W., & Nagel, K. (2006). Agent-based demand-modeling framework for large-scale microsimulations. Transportation Research Record, 1985, 125-134.

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

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

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

[43] CalTrans (2001), 2000-2001 California Statewide Household Travel Survey.

[44] Recker, W.W., & 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, 4(6), 357-378.

Appendix A: dHAPP Constraints (same as original HAPP constraints in Recker, 1995)

$$\sum_{v \in V_h} \sum_{w \in \mathbf{Q}_h} X_{uw}^{v,h} = 1, \quad u \in P_h^+, h \in \mathbf{H} \tag{A1}$$

$$\sum_{w \in \mathbf{Q}_h} X_{uw}^{v,h} - \sum_{w \in \mathbf{Q}_h} X_{wu}^{v,h} = 0, \quad u \in P_h, v \in V_h, h \in \mathbf{H} \tag{A2}$$

$$\sum_{w \in P_h^+} X_{0w}^{v,h} = 1, \quad v \in V_h, h \in \mathbf{H} \tag{A3}$$

$$\sum_{u \in P_h^-} X_{u,2n+1}^{v,h} - \sum_{u \in P_h^+} X_{0,u}^{v,h} = 0, \quad v \in V_h, h \in \mathbf{H} \tag{A4}$$

$$\sum_{w \in \mathbf{Q}_h} X_{wu}^{v,h} - \sum_{w \in \mathbf{Q}_h} X_{w,n+u}^{v,h} = 0, \quad u \in P_h^+, v \in V_h, h \in \mathbf{H} \tag{A5}$$

$$T_u^h + s_u^h + t_{uw}^h = T_{n+u}^h, \quad u, w \in P_h^+, h \in \mathbf{H} \tag{A6}$$

$$X_{uw}^{v,h} = 1 \Rightarrow T_u^h + s_u^h + t_{uw}^h = T_w^h, \quad u, w \in P_h, v \in V_h, h \in \mathbf{H} \tag{A7}$$

$$X_{0w}^{v,h} = 1 \Rightarrow T_0^{v,h} + s_u^h + t_{0w}^h = T_w^h, \quad w \in P_h^+, v \in V_h, h \in \mathbf{H} \tag{A8}$$

$$X_{u,2n+1}^{v,h} = 1 \Rightarrow T_u^h + s_u^h + t_{u,2n+1}^h = T_{2n+1}^h, \quad u \in P_h^-, v \in V_h, h \in \mathbf{H} \tag{A9}$$

$$a_u^h \le T_u^h \le b_u^h, \quad u \in P_h, h \in \mathbf{H} \tag{A10}$$

$$a_0^h \le T_0^{v,h} \le b_0^h, \quad v \in V_h, h \in \mathbf{H} \tag{A11}$$

$$a_{2n+1}^h \le T_{2n+1}^{v,h} \le b_{2n+1}^h, \quad v \in V_h, h \in \mathbf{H} \tag{A12}$$

$$X_{uw}^{v,h} = 1 \Rightarrow Y_u^h + d_w = Y_w^h, \quad u \in P_h, w \in P_h^+, v \in V_h, h \in \mathbf{H} \tag{A13}$$

$$X_{uw}^{v,h} = 1 \Rightarrow Y_w^h = 0, \quad u \in P_h, \mathrm{w} \in P_h^-, v \in V_h, h \in \mathbf{H} \tag{A14}$$

$$X_{0w}^{v,h} = 1 \Rightarrow Y_0^h + d_w = Y_w^h, \quad w \in P_h^+, v \in V_h, h \in \mathbf{H} \tag{A15}$$

$$Y_0^h = 0, \quad 0 \le Y_0^h \le D, \quad u \in P_h^+, h \in \mathbf{H} \tag{A16}$$

$$\sum_{v \in V_h} \sum_{u \in \mathbf{Q}_h} \sum_{w \in \mathbf{Q}_h} c_{uw}^{v,h} \cdot X_{uw}^{v,h} \le B_C^{v,h}, \quad h \in \mathbf{H} \tag{A17}$$

$$\sum_{v \in V_h} \sum_{u \in \mathbf{Q}_h} \sum_{w \in \mathbf{Q}_h} t_{uw}^h \cdot X_{uw}^{v,h} \le B_T^{v,h}, \quad v \in V_h, h \in \mathbf{H} \tag{A18}$$

$$X_{uv}^{v,h} \in (0,1), \quad u, w \in \mathbf{Q}_h, v \in V_h, h \in H \tag{A19}$$

Appendix 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