# Modeling Visit Probabilities within Space-Time Prisms of Daily Activity-Travel Patterns

Jing Lyu*[1], Feixiong Liao[2], Soora Rasouli[3]

[1]PhD candidate, Urban Planning and Transportation Group, Eindhoven University of Technology, the Netherlands
[2]Assistant professor, Urban Planning and Transportation Group, Eindhoven University of Technology, the Netherlands
[3]Professor, Urban Planning and Transportation Group, Eindhoven University of Technology, the Netherlands

## SHORT SUMMARY

Space-time prism (STP) delimits the space-time opportunities reachable by a moving object and is widely applied to measure the ability of individuals to travel and participate in activities. The majority STPs are binary measures in that all locations are considered equally accessible if within the prisms. A few probabilistic STP models discussed heterogeneous interiors, but they focus on the trip level and have not addressed daily activity programs with flexible activity sequences. This study proposes a model framework to construct and estimate the state-dependent probabilistic STP of daily activity-travel patterns based on the multi-state supernetwork representation. Utilizing GPS trajectories, the estimation and simulation results of visit probabilities in the STPs demonstrate the validity of the model framework.

**Keywords:** activity-travel patterns; multi-state supernetwork; space-time prism; visit probability

## 1. INTRODUCTION

Individuals' travel and activity participation are subject to space-time constraints. As a central time geographic concept, space-time prism (STP) delimits the space-time opportunities that can be reached by a moving object (Miller, 2017), and provides a measurement of potential mobility. The classic STP is determined by the known anchor points, time budget, the maximum attainable travel speed, and the time available for a flexible activity that can be conducted at one of multiple locations (Hägcrstrand, 1970). By constructing the STP over a transportation network, the network-time prism delimits the accessible locations with respect to the spatial network (Miller, 1991). The boundary of an STP has been widely used as a space-time accessibility measurement, indicated by a spatial location set at instant times within time intervals.

The majority STPs are deterministic binary measures such that all locations are considered equally accessible if within the prisms, otherwise not accessible. In reality, a prism does not have homogeneous interiors. To capture the variations within a prism, Winter and co-workers (Winter, 2009; Winter and Yin, 2010a, 2010b) introduced probabilistic time geography and modeled probability distributions of an individual's potential visit locations at a time moment from a stochastic perspective. Song and Miller (2013) formulated discrete and continuous stochastic models for the movement of an individual within the potential path area (PPA). In Song et al. (2016, 2017) , the visit probability within a directed STP is modeled using the approach of continuous-time semi-Markov process, describing the likelihood of visiting different locations within the prism.

In previous studies, STP is predominantly modeled at the trip level with single activities rather than for daily activity programs (APs). To capture the dependencies in an activity chain and improve realism, Chen and Kwan (2012) identified the location choice set based on STPs in the presence of all possible activity-travel chains. Kang and Chen (2016) constructed the feasible space-time region for a daily AP by intersecting a set of feasible space-time regions for single activities. Liao (2021) applied bidirectional searches of full activity-travel patterns (ATPs) in multi-state supernetworks (*SNK*) (Liao, Arentze, and Timmermans, 2010, 2013) for delineating the exact STP for an AP. However, the probabilistic characteristics of accessibility within a prism have not been investigated for an AP with multiple activities and flexible activity sequences.

The aim of this study is, therefore, to propose a model framework to construct and estimate the probabilistic STP of an AP. The model framework includes three steps. First, based on the multi-state supernetwork, we construct the activity-based STP for a daily AP with flexible activity sequences (Liao, 2021). Second, we model the visit probability within the activity-based STP using semi-Markov techniques (Howard, 1971). Third, we estimate a visit probability model using extracted GPS trajectories and simulate a typical AP to demonstrate the validity of the model framework. The visit probabilities of an AP provide quantitative descriptions of the activity-based STP interiors and evaluations of the accessibility for an individual participating in multiple activities with flexible sequences.

## 2. METHODOLOGY

### *Multi-state supernetwork (SNK) representation and STP*

Multi-state supernetworks are capable of representing ATPs of conducting an individual's AP. A daily AP's implementation is a path choice through networks of different states, including activity states specifying which activities have been conducted, and vehicle states specifying where the private vehicles are (in use or parked somewhere).

Denote a multi-state supernetwork as $SNK(N, E)$. The set of nodes $N$ indicates locations in *SNK*, including road intersections, activity locations, and parking locations. A set of links $E$ includes travel links of road segments, transaction links for conducting activities at activity locations, transition links for parking and picking up PVs, and boarding and alighting.

Given origin $\mathrm{H}_0$ and destination $\mathrm{H}_1$ as two anchors and the corresponding time budget $[t_{\mathrm{H}_0}, t_{\mathrm{H}_1}]$, the STP is constructed by deriving the potential path area (PPA) of an AP. The temporal feasibility of STP in *SNK* is formulated as follows (Liao, 2021):

$$\min\{g_{\mathrm{H}_0}(n|_s) + g_{\mathrm{H}_1}(n|_s)\} \le t_{\mathrm{H}_1} - t_{\mathrm{H}_0} \quad (1)$$

where $n|_s$ denote node $n$ at activity-vehicle state $s$, $g_{\mathrm{H}_0}(n|_s)$ and $g_{\mathrm{H}_1}(n|_s)$ are the actual activity-travel times from $\mathrm{H}_0$ to $n|_s$ and $n|_s$ to $\mathrm{H}_1$ respectively.

### *Visit probability within the STP*

We define a status as a movement starting from node $i|_s$ to node $j|_{s'}$ along link $l_{ij|ss'}$ in *SNK* that has not arrived at $j|_{s'}$ yet, $l_{ij|ss'} = (i|_s, j|_{s'})$ for $\forall s, s'$. The status space includes the movements on all possible links within the STP. We formulate the holding time density functions of *SNK* links and the visit probability of each status at a moment in time.

(1) Holding time density functions

A holding time density function describes the probability that a transition from $i|_s$ to $j|_{s'}$, corresponding to the movement on link $l_{ij|ss'}$, will take extra time $\tau$ over the minimum time on $l_{ij|ss'}$, denoted as $f_{ij|ss'}(\tau)$ when $\tau \ge 0$, otherwise $f_{ij|ss'}(\tau) = 0$ . The extra time $\tau$ on $l_{ij|ss'}$ can be calculated using

$\tau = \tau' - t_{ij|ss'}$, where $\tau'$ is the total time one traverse $l_{ij|ss'}$ and $t_{ij|ss'}$ is the minimum time expense from $i|_s$ to $j|_{s'}$. For travel link $l_{ij|ss'}$ that $i \neq j$ when $s = s'$, $t_{ij|ss'}$ is the minimum travel time. If conducting activity $\alpha \in A$ at location $i$ $(i = j)$ at state $s$ results in a new state $s'$, that is $s \neq s'$, $l_{ij|ss'}$ is an activity link and $t_{ij|ss'}$ corresponding to the minimum activity duration $d_\alpha$.

Considering parameters of $f_{ij|ss'}(\tau)$ are heterogeneous for travel and transaction links, we use the latent class models to capture the latent heterogeneity of holding times. Suppose there exist $K$ different homogeneous latent classes in the heterogeneous population of extra travel times and activity durations. Let $P_{qk}$ denote the class membership probability that an individual $q$ belongs to latent class $k$:

$$P_{qk} = \frac{\exp(\boldsymbol{\beta}_k \boldsymbol{x}_q)}{\sum_{k=1}^{K} \exp(\boldsymbol{\beta}_k \boldsymbol{x}_q)}, \qquad k = 1, \dots, K, \boldsymbol{\beta}_K = 0 \tag{2}$$

where $\boldsymbol{x}_q$ is the vector of sociodemographic variables of individual $q$, $\boldsymbol{\beta}_k$ is the parameter vector of $\boldsymbol{x}_q$ of latent class $k$. The latent class holding time density function for extra time is formulated as

$$f^q(\tau) = \sum_{k=1}^{K} P_{qk} \cdot f^q(\tau|k) \tag{3}$$

where $f^q(\tau|k)$ denotes the probability of individual $q$ that belongs to class $k$ spending extra time $\tau$ on a specific link in *SNK*, and $f^q(\tau)$ is the probability of unconditional holding time density.

(2) Visit probability formulation

For each link $l_{ij|ss'}$ in *SNK*, we can calculate a feasible time range as $\left(t^-_{i|_s}, t^+_{j|_{s'}}\right)$, where $t^-_{i|_s}$ is the earliest arrival time and $t^+_{j|_{s'}}$ is the latest departure time at $i|_s$ and $j|_{s'}$ in *SNK*, respectively. The visit probability in *SNK* as an extension over Song et al. (2016) is defined as follows.

Denote $P_{\mathrm{H}_0 \to i|_s}(t)$ as the probability $l_{ij|ss'}$ can be reached from $i|_s$ at $t$ given origin $\mathrm{H}_0$, formulated as

$$P_{\mathrm{H}_0 \to i|_s}(t) = \begin{cases} 0 & t \in \left[t_{\mathrm{H}_0}, t^-_{i|_s}\right) \\ \displaystyle\int_{t^-_{i|_s} - t_{\mathrm{H}_0}}^{t - t_{\mathrm{H}_0}} f_{ij|ss'}(\tau) d\tau & t \in \left[t^-_{i|_s}, t^+_{j|_{s'}} - t_{ij|ss'}\right) \\ \displaystyle\int_{t^-_{i|_s} - t_{\mathrm{H}_0}}^{t^+_{j|_{s'}} - t_{ij|ss'} - t_{\mathrm{H}_0}} f_{ij|ss'}(\tau) d\tau & t \in \left[t^+_{j|_{s'}} - t_{ij|ss'}, t_{\mathrm{H}_1}\right] \end{cases} \tag{4}$$

Denote $P_{j|_{s'} \to \mathrm{H}_1}(t)$ as the probability of reaching $\mathrm{H}_1$ from $j|_{s'}$ based on available departure times, formulated as

$$P_{j|_{s'} \to \mathrm{H}_1}(t) = \begin{cases} 0 & t \in \left[t^+_{j|_{s'}}, t_{\mathrm{H}_1}\right] \\ \displaystyle\int_{t_{\mathrm{H}_1} - t^+_{j|_{s'}}}^{t_{\mathrm{H}_1} - t} f_{ij|ss'}(\tau) d\tau & t \in \left[t^-_{i|_s} + t_{ij|ss'}, t^+_{j|_{s'}}\right) \\ \displaystyle\int_{t_{\mathrm{H}_1} - t^+_{j|_{s'}}}^{t_{\mathrm{H}_1} - t^-_{i|_s} - t_{ij|ss'}} f_{ij|ss'}(\tau) d\tau & t \in \left[t_{\mathrm{H}_0}, t^-_{i|_s} + t_{ij|ss'}\right) \end{cases} \tag{5}$$

The probability for visiting $l_{ij|ss'}$ at $t \in [t_{\mathrm{H}_0}, t_{\mathrm{H}_1}]$, denoted as $P(l_{ij|ss'}, t)$, is formulated as the joint probability of $l_{ij|ss'}$ being reached from $i|_s$ and arriving at $\mathrm{H}_1$ within $(t_{\mathrm{H}_1} - t)$. We normalize probabilities among all accessible links at time $t$ within the STP in *SNK*, for which all probabilities are added up to 1.

$$P\left(l_{ij|ss'},t\right)=\frac{P_{\mathrm{H}_0\to i|_s}(t)\times P_{j|_{s'}\to \mathrm{H}_1}(t)}{\sum_{l\in E}P(l,t)},t_{i|_s}^{-}+t_{ij|ss'}\le t_{j|_{s'}}^{+} \tag{6}$$

Eq. (6) can be used to derive the visit probability of each link and the discrete distribution of the status space at a time moment in *SNK*.

### *Model estimation*

We extract a transportation network and collect individuals' ATPs from GPS trajectories to estimate the holding time density functions for visit probability in the following steps.

**Step 1**: Calculate the extra available times for road segments and different types of activities. For travel links (road segments), the extra time of a sampled ATP is calculated as:

$$\tau=rec(t_{ATP})-\min(t_{ATP}) \tag{7}$$

where $rec(t_{ATP})$ is the recorded ATP time from GPS data, $\min(t_{ATP})$ is the shortest ATP time with the same activity durations calculated using a path searching algorithm. Suppose $f_{ij|ss'}(\tau)$ of travel link $l_{ij|ss'}$ follows an exponential distribution, i.e.,

$$f_{ij|ss'}(\tau)=\begin{cases}\lambda e^{-\lambda\tau}, & \tau\ge 0\\ 0, & \tau<0\end{cases} \tag{8}$$

$f^q(\tau|k)$ in Eq. (3) is substituted by Eq. (8) as $f_{ij|ss'}^{q}(\tau|\lambda_k)$, with parameter $\lambda_k$ $(k=1,2,\dots,K)$ to be estimated.

For transaction links of different types of activities, the extra activity duration of a sampled ATP is calculated as:

$$\tau_\alpha=rec(\tau_\alpha)-\min(\tau_\alpha) \tag{9}$$

where $rec(\tau_\alpha)$ is the recorded ATP's activity duration for conducting activity $\alpha$, $\min(\tau_\alpha)$ is the minimum duration of $\alpha$ among all the extracted ATPs. We suppose $f_{ij|ss'}(\tau)$ (denoted as $f_\alpha(\tau)$) of conducting $\alpha$ follows the lognormal distribution, i.e.,

$$f_\alpha(\tau)=\frac{1}{\tau\sigma\sqrt{2\pi}}\exp\left(-\frac{(\ln(\tau)-\mu)^2}{2\sigma^2}\right),\qquad \tau>0 \tag{10}$$

$f^q(\tau|k)$ in Eq. (3) is substituted by Eq. (10) as $f_\alpha^q\left(\tau|\mu_k,\sigma_k^2\right)$, with parameter $\mu_k$, $\sigma_k^2(k=1,2,\dots,K)$ to be estimated.

**Step 2**: Estimate the parameters using maximum likelihood estimation (MLE). Given $N$ the total number of extracted individuals' ATP, the likelihood for all individuals is:

$$L=\prod_{q=1}^{N}\left[\sum_{k=1}^{K}P_{qk}\cdot f^q(\tau|k)\right]=\prod_{q=1}^{N}\left[\sum_{k=1}^{K}\frac{\exp\left(\boldsymbol{\beta}_k\boldsymbol{x}_q\right)}{\sum_{k=1}^{K}\exp\left(\boldsymbol{\beta}_k\boldsymbol{x}_q\right)}\cdot f^q(\tau|\boldsymbol{\theta}_k)\right] \tag{11}$$

where $\boldsymbol{\theta}_k$ are the parameters in Eq. (8) or Eq. (10) for travel and transaction links. The log-likelihood for the sampled ATPs is:

$$\ln L = \sum_{q=1}^{N} \left( \ln \sum_{k=1}^{K} P_{qk} \cdot f^q(\tau|\boldsymbol{\theta}_k) \right) \tag{12}$$

We apply the gradient-descent method as a solution to solve the MLE for latent class models. The estimated parameters are used in simulating the visit probabilities in Eqs. (4-6).

## 3. RESULTS

The suggested model framework is implemented for an AP in a transportation network. STPs and visit probabilities are simulated for each activity state in *SNK* every 5 min in peak hours and 30 min in non-peak hours for travel links and activity locations.

### *Estimation results*

In the numerical experiment, 2714 individuals' ATPs with one fixed workplace and one flexible activity (non-daily shopping) are extracted from GPS data. We select an AP in the Eindhoven area (the Netherlands) to simulate the STP construction and the visit probabilities within the STP (**Figure** 1). The settings are as follows:

(1) The time budget is 742 min, 10% extra over the recorded ATP time. Given the recorded ATP departure time $t_{\mathrm{H}_0} = 6:20$, the time window is [6:20, 18:42]. 816 flexible activity locations are selected as alternatives for non-daily shopping. The recorded minimum duration of working and flexible activity are 558 min and 36 min respectively.
(2) The road network includes 47,901 nodes and 10,0581 directed road segments. The car speeds for non-peak hours [7:00, 9:00] and peak hours [16:30, 19:00] on 3 types of roads are <100, 80, 50> and <70, 50, 30> km/h.

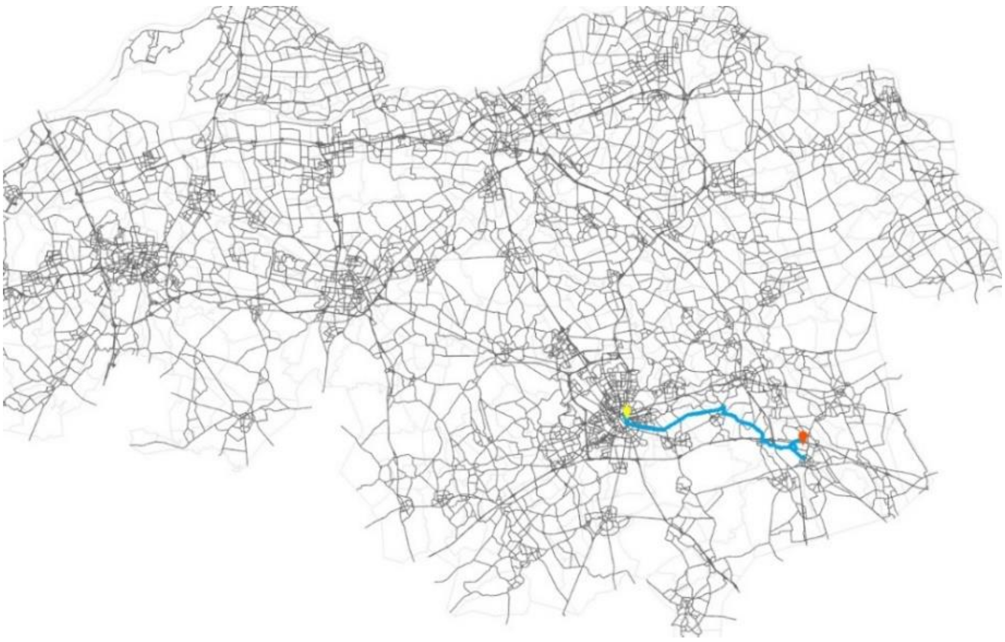


**Figure 1.** Selected ATP (anchors: red: home; yellow: workplace; blue line: GPS trajectory).

(3) The parameters of $f^q(\tau)$ for the traveling and non-daily shopping are estimated. Individual $q$'s gender and age are selected as the sociodemographic variables as $x_{q1}$ and $x_{q2}$. Let $x_{q1} = 0$ if individual $q$ is male, $x_{q1} = 1$ when $q$ is female; $x_{q2} = 0$ if individual $q$'s age is between 20 to 59, $x_{q2} = 1$ when age $\geq 60$. The estimation outcomes of the latent class models are shown in Tables 1-6 (See Tables in the Appendix).
(4) For travel links, 4 activity states are simulated based on whether work and non-daily shopping are conducted in different sequences. For transaction links, 2 states are simulated, i.e, the flexible activity is conducted before and after work.

### *Illustration of visit probabilities*

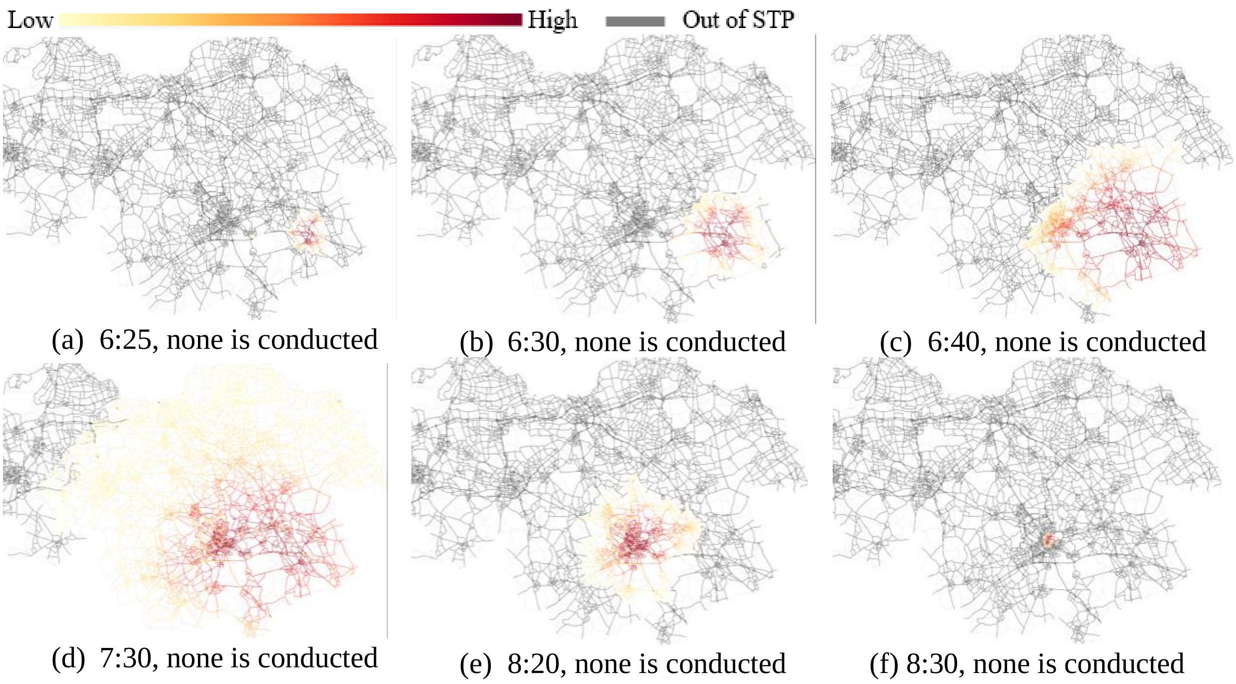
#### *Travel links*

The visit probabilities of travel links within the STPs reflect the heterogeneous interior for traveling with purpose, which is non-uniform and changing over activity states and time moments.

At activity state 0, the individual departs at 6:20 traveling to conduct the fixed and flexible activities. The number of travel links within STPs increases given available time during [6:20, 7:30], and decreases during [7:40, 8:30] since 558 min work has to be done (**Figure** 2**)**. At activity state 2 where the 36-min non-daily shopping is conducted within [6:20, 7:00], the individual has more time for visiting more travel links at the same time point during [8:00, 8:55] at state 2 compared to the situation at state 0 (see in **Figure** 3 (a) and (b) at 8:20).

There are no travel links within STP between (8:55, 16:00) since the individual is working. After 16:00 at activity state 1, the individual searches locations of non-daily shopping after work is conducted (**Figure** 4**)**. At activity state 3 that all activities are done, around 30 to 40 min are available for the individual traveling to home during [16:40, 18:40], and thus more links can be accessed compared to the situation at state 1 (see **Figure** 5 (c) and (d) at 18:00).

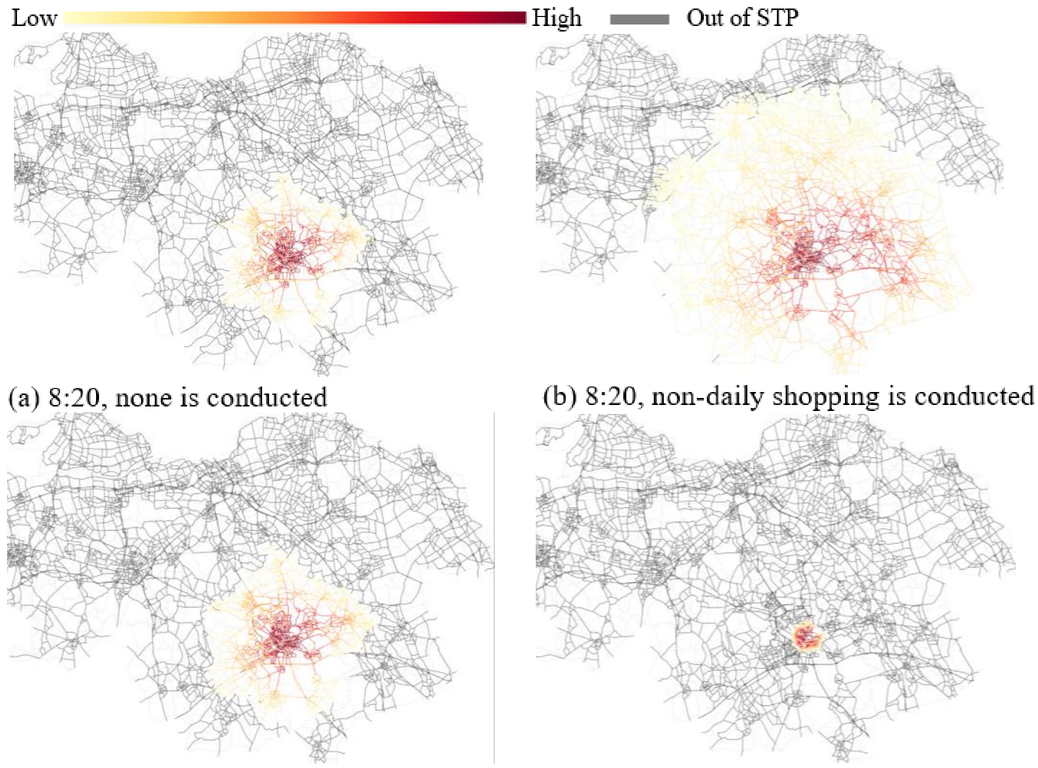
As shown in the results, the distribution of locations, activity sequences, available time, and the shortest path direction, together result in the changes of "higher visit probability travel links" (darker red).




(a) 6:25, none is conducted (b) 6:30, none is conducted (c) 6:40, none is conducted

(d) 7:30, none is conducted (e) 8:20, none is conducted (f) 8:30, none is conducted

**Figure 2.** STPs interior of travel links at activity state 0.



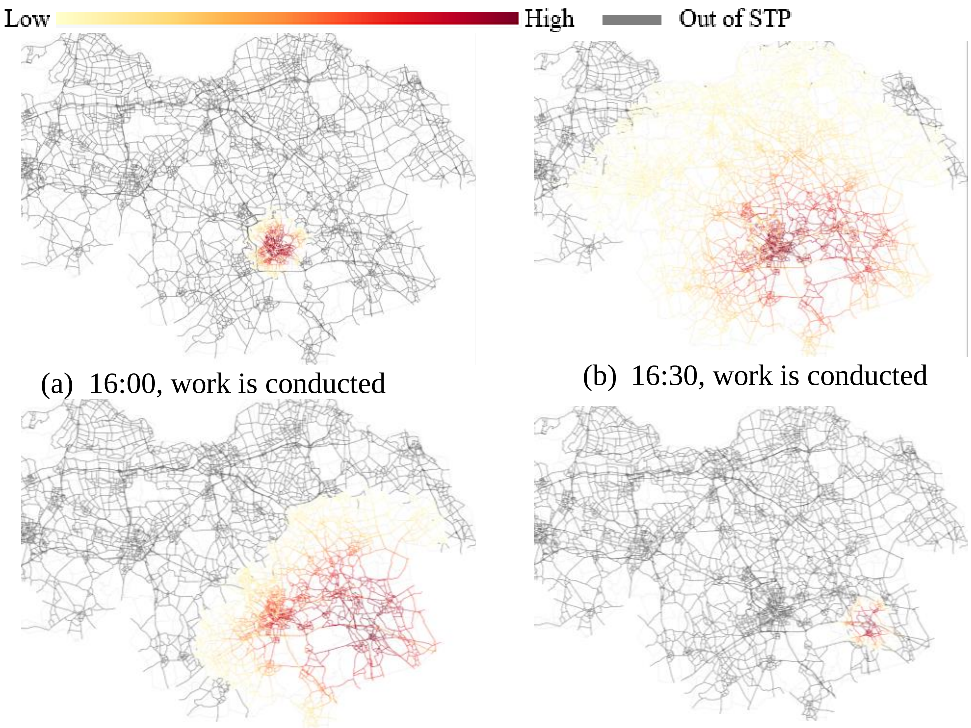

(a) 8:20, none is conducted

(b) 8:20, non-daily shopping is conducted

(c) 8:40, non-daily shopping is conducted

(d) 8:55, non-daily shopping is conducted

**Figure 3.** STPs interior of travel links at activity state 2.



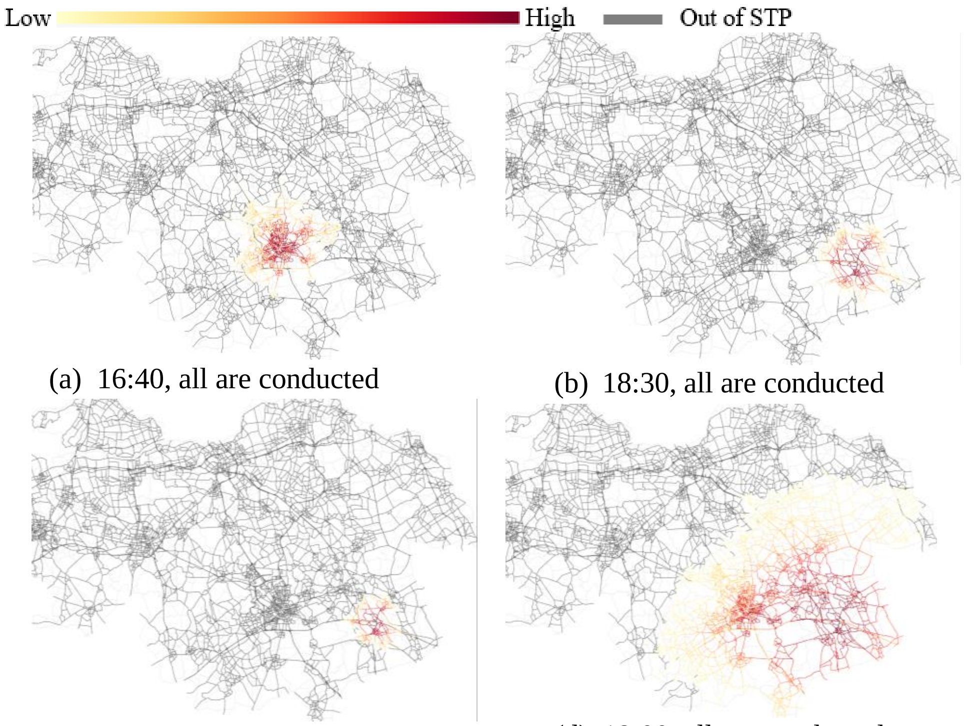

(a) 16:00, work is conducted

(b) 16:30, work is conducted

(c) 17:40, work is conducted

(d) 18:00, work is conducted

**Figure 4.** STPs interior of travel links at activity state 1.




(a) 16:40, all are conducted

(b) 18:30, all are conducted

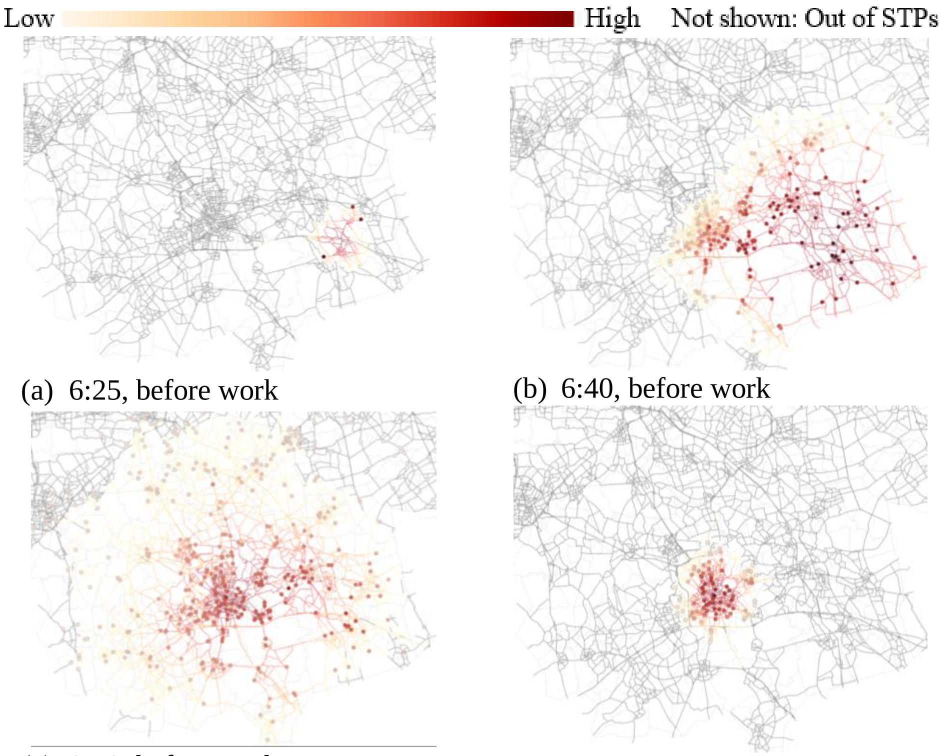
(c) 18:00, only work is conducted

(d) 18:00, all are conducted

**Figure 5.** STPs interior of travel links at activity state 3.

***Flexible activity locations***

The level of accessibilities of the flexible activity locations within the STPs are reflected by the visit probabilities, which changes over activity sequence and time moments.



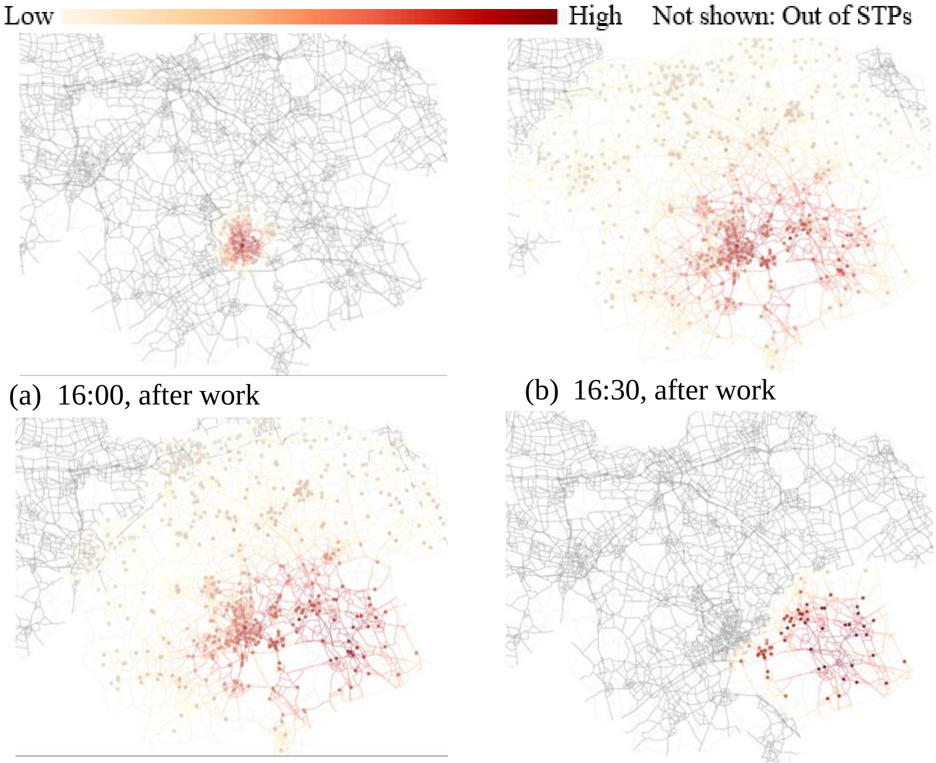

(a) 6:25, before work

(b) 6:40, before work

(c) 8:10, before work

(d) 8:50, before work

**Figure 6.** STPs interior of flexible activity locations before work.




(a) 16:00, after work

(b) 16:30, after work

(c) 17:30, after work

(d) 18:20, after work

**Figure 7.** STPs interior of flexible activity locations after work.

The number of flexible activity locations within the STP increases during [6:20, 7:40] and the locations near home have relatively higher visit probabilities. After 8:00, the closer to the workplace, the higher visit probabilities of the locations are compared to the ones near home due to the available time (**Figure** 6). After work is conducted (**Figure** 7), the accessible locations first expand from the center of the workplace with higher probabilities and finally shrink to home at 18:20, given $t_{\mathrm{H}_1}$ =18:42 and the 36-min flexible activity duration.

## 4. CONCLUSIONS

The state-dependent STPs of an AP delimit the potential mobility of an individual to access locations with limited time, but the deterministic characteristic of STP lacks the evaluation of how likely a location within the STP can be visited. This study proposes a model framework to construct and estimate the probabilistic STP of an AP based on the multi-state supernetwork representation. By assuming the distributions of holding time density functions of travel and transaction links in *SNK*, latent class models are further applied to capture and estimate the individuals' latent heterogeneity on extra time for traveling and conducting activities. The results illustrate that the visit probabilities of travel links and activity locations over different activity states and time points can describe the interior of the STPs. In the next step, we will develop proper accessibility measurements based on the quantifications of visit probabilities.

## REFERENCES

Chen, X., Kwan, M.-P. 2012. Choice set formation with multiple flexible activities under space-time constraints. *International Journal of Geographical Information Science*, Vol. 26, No. 5, pp. 941-961.

Hägcrstrand, T. 1970. *What about people in regional science?* Paper presented at the Papers of the Regional Science Association.
Howard, R. A. 1971. *Dynamic probabilistic systems , Volume II: Semi-Markov and Decision Processes* (Vol. 2). New York: Wiley.
Kang, J. E., Chen, A. 2016. Constructing the feasible space-time region of the Household Activity Pattern Problem. *Transportmetrica A: Transport Science,* Vol. 12, No. 7, pp. 591-611.
Liao, F. 2021. Exact space-time prism of an activity program: bidirectional searches in multi-state supernetwork. *International Journal of Geographical Information Science,* Vol. 35, No. 10, pp. 1975-2001.
Liao, F., Arentze, T., Timmermans, H. 2010. Supernetwork approach for multimodal and multiactivity travel planning. *Transportation Research Record,* Vol. 2175, No. 1, pp. 38-46.
Liao, F., Arentze, T., Timmermans, H. 2013. Incorporating space-time constraints and activity-travel time profiles in a multi-state supernetwork approach to individual activity-travel scheduling. *Transportation Research Part B: Methodological,* Vol. 55, pp. 41-58.
Miller, H. J. 1991. Modelling accessibility using space-time prism concepts within geographical information systems. *International journal of geographical information systems,* Vol. 5, No. 3, pp. 287-301.
Miller, H. J. 2017. Time Geography and Space-Time Prism. In *International Encyclopedia of Geography* (pp. 1-19).
Song, Y., Miller, H. J. 2013. Simulating visit probability distributions within planar space-time prisms. *International Journal of Geographical Information Science,* Vol. 28, No. 1, pp. 104-125.
Song, Y., Miller, H. J., Stempihar, J., Zhou, X. 2017. Green accessibility: Estimating the environmental costs of network-time prisms for sustainable transportation planning. *Journal of Transport Geography,* Vol. 64, pp. 109-119.
Song, Y., Miller, H. J., Zhou, X., Proffitt, D. 2016. Modeling Visit Probabilities within Network-Time Prisms Using Markov Techniques. *Geographical Analysis,* Vol. 48, No. 1, pp. 18-42.
Winter, S. 2009. Towards a probabilistic time geography. Paper presented at the Proceedings of the 17th ACM SIGSPATIAL International Conference on advances in geographic information systems.
Winter, S., Yin, Z.-C. 2010a. Directed movements in probabilistic time geography. *International Journal of Geographical Information Science,* Vol. 24, No. 9, pp. 1349-1365.
Winter, S., Yin, Z.-C. 2010b. The elements of probabilistic time geography. *GeoInformatica,* Vol. 15, No. 3, pp. 417-434.

## APPENDIX

Tables 1-3 report the estimated parameters of $f^q(\tau)$ and the percentages that the sample belongs to each class with the highest membership probability for non-daily shopping, working, and traveling, respectively. To reflect the fit of the latent class models, Tables 4-6 shows the test results of Bayesian Information Criteria (BIC) and likelihood ratio (LR) given various numbers of classes, based on the assumptions of distributions and the hypotheses of the parameters.

The estimated parameters of latent class models:

a. Flexible activity: non-daily shopping

**Table 1**. Parameters and highest $P_{qk}$ percentages of selected class number

| Class | $\boldsymbol{\beta}_k$ | | $\boldsymbol{\mu}_k, \boldsymbol{\sigma}_k$ | | Sample $N_k$ / Highest $P_{qk}$ % | |
|---|---|---|---|---|---|---|
| 1 | const | 0 | $\mu_1$ | 2.7109 | 365 | 100% |
| | gender | 0 | $\sigma_1$ | 1.9683 | | |
| | age | 0 | | | | |

b. Fixed activity: working

**Table 2**. Parameters and highest $P_{qk}$ percentages of selected class number

| Class | $\boldsymbol{\beta}_k$ | | $\boldsymbol{\mu}_k, \boldsymbol{\sigma}_k$ | | Sample $N_k$ / Highest $P_{qk}$ % | |
|---|---|---|---|---|---|---|
| 1 | const | 0.536 | $\mu_1$ | 6.2672 | 1234 | 45.47% |
| | gender | -0.6813 | $\sigma_1$ | 0.0941 | | |
| | age | -0.4178 | | | | |
| 2 | const | 0 | $\mu_2$ | 5.2213 | 1480 | 54.53% |
| | gender | 0 | $\sigma_2$ | 0.9410 | | |
| | age | 0 | | | | |

c. Travel links

**Table 3**. Parameters and highest $P_{qk}$ percentages of selected class number

| Class | $\boldsymbol{\beta}_k$ | | $\boldsymbol{\lambda}_k$ | | Sample $N_k$ / Highest $P_{qk}$ % | |
|---|---|---|---|---|---|---|
| 1 | const | 1.7586 | $\lambda_1$ | 0.0271 | 380 | 24.58% |
| | gender | -2.9839 | | | | |
| | age | -2.7812 | | | | |
| 2 | const | -0.8908 | $\lambda_2$ | 0.0346 | 842 | 54.46% |
| | gender | 2.963 | | | | |
| | age | 0.2311 | | | | |
| 3 | const | 0 | $\lambda_3$ | 0.0310 | 423 | 20.96% |
| | gender | 0 | | | | |
| | age | 0 | | | | |

The results of the fit of the latent class models:

a. Flexible activity: non-daily shopping

**Table 4**. BIC and LR with latent classes number

| Number of classes | Number of Parameters | Log-likelihood $\ln(L)$ | BIC | LR (LR-val / p-val (df)) |
|---|---|---|---|---|
| 1 | 5 | -1754.5719 | 3520.9438 | 0.5649 / 0.7539 (2) |
| 2 | 10 | -1697.8948 | 3454.7885 | 113.9192 / 8.6267e-20 (10) |
| 3 | 15 | -1681.2356 | 3450.9697 | 147.2378 / 8.5223e-24 (15) |
| 4 | 20 | -1675.6434 | 3469.2848 | 158.6689 / 1.3564e-23 (20) |
| 5 | 25 | -1665.4855 | 3478.4684 | 173.3506/ 3.6977e-24 (25) |
| 10 | 50 | -1650.1866 | 3595.3681 | 202.0330/ 3.6134e-20 (50) |

$BIC = P\ln(N) - 2\ln(L)$ ; $P$: number of $\boldsymbol{\beta}_k$ + number of $\mu_k, \sigma_k^2$

LR test: The null hypothesis: $\mu_k = \frac{1}{N}\sum_{q=1}^{N} \ln \tau_q$, $\sigma_k^2 = \frac{1}{N}\sum_{q=1}^{N}\left(\ln \tau_q - \frac{1}{N}\sum_{q=1}^{N} \ln \tau_q\right)^2$, $\boldsymbol{\beta}_k = 0$

b. Fixed activity: working

**Table 5**. BIC and LR with latent classes number

| Number of classes | Number of Parameters | Log-likelihood $\ln(L)$ | BIC | LR (LR-val / p-val (df)) |
|---|---|---|---|---|
| 1 | 5 | -18973.5117 | 37962.8319 | 4046.5781/ 0 (2) |
| 2 | 10 | -17586.0957 | 35251.2531 | 6797.3008 / 0 (10) |
| 3 | 15 | -17469.0371 | 35056.6669 | 6919.1914 / 0 (15) |
| 4 | 20 | -17285.6230 | 34729.3696 | 7024.5898 / 0 (20) |
| 5 | 25 | -17259.5254 | 34716.7052 | 7470.7734 / 0 (25) |
| 6 | 30 | -17284.9824 | 34807.1502 | 7423.2188 / 0 (30) |
| 8 | 40 | -17287.2480 | 34890.7433 | 7415.6523 / 0 (40) |
| 10 | 50 | -17285.5410 | 34966.3909 | 7414.4844 / 0 (50) |

c. Travel links

**Table 6**. BIC and LR with latent classes number

| Number of classes | Number of Parameters | Log-likelihood $\ln(L)$ | BIC | LR (LR-val / p-val (df)) |
|---|---|---|---|---|
| 1 | 3 | -6882.1519 | 13771.6471 | 0. / 1.0 (1) |
| 2 | 6 | -6877.2329 | 13813.2132 | 9.7520 / 0.2828 (6) |
| 3 | 9 | -6875.3398 | 13838.8008 | 13.7549 / 0.3166 (9) |
| 4 | 12 | -6875.2188 | 13867.9323 | 13.6055 / 0.6281 (12) |
| 5 | 15 | -6875.0942 | 13897.0570 | 14.3369 / 0.8130 (15) |
| 8 | 24 | -6874.9126 | 13984.8148 | 14.4785 / 0.9966 (24) |
| 10 | 30 | -6874.9165 | 14043.5701 | 14.6758 / 0.9999 (30) |

$P$: number of $\boldsymbol{\beta}_k$ + number of $\lambda_k$

LR test: The null hypothesis: $\lambda_k = \frac{N}{\sum_{q=1}^{N} \tau_q}, \boldsymbol{\beta}_k = 0$