14th Behavior Modeling in Transportation Networks, UT
Sat. 26 September. Early bird session.

A context-dependent scheduling model considering measurement errors in pedestrian network

Yuki Oyama* and Eiji Hato
* Ph.D. Candidate / The University of Tokyo
oyama@bin.t.u-tokyo.ac.jp

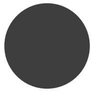
東京大学
THE UNIVERSITY OF TOKYO

Behavior in Networks Studies Unit

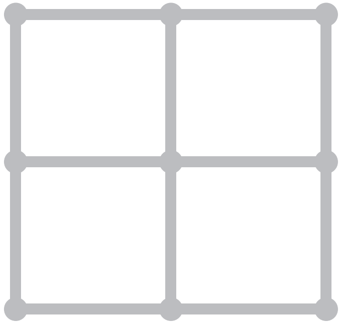
Table of Contents

Time use pattern in pedestrian network

Development around station



Urban renovations



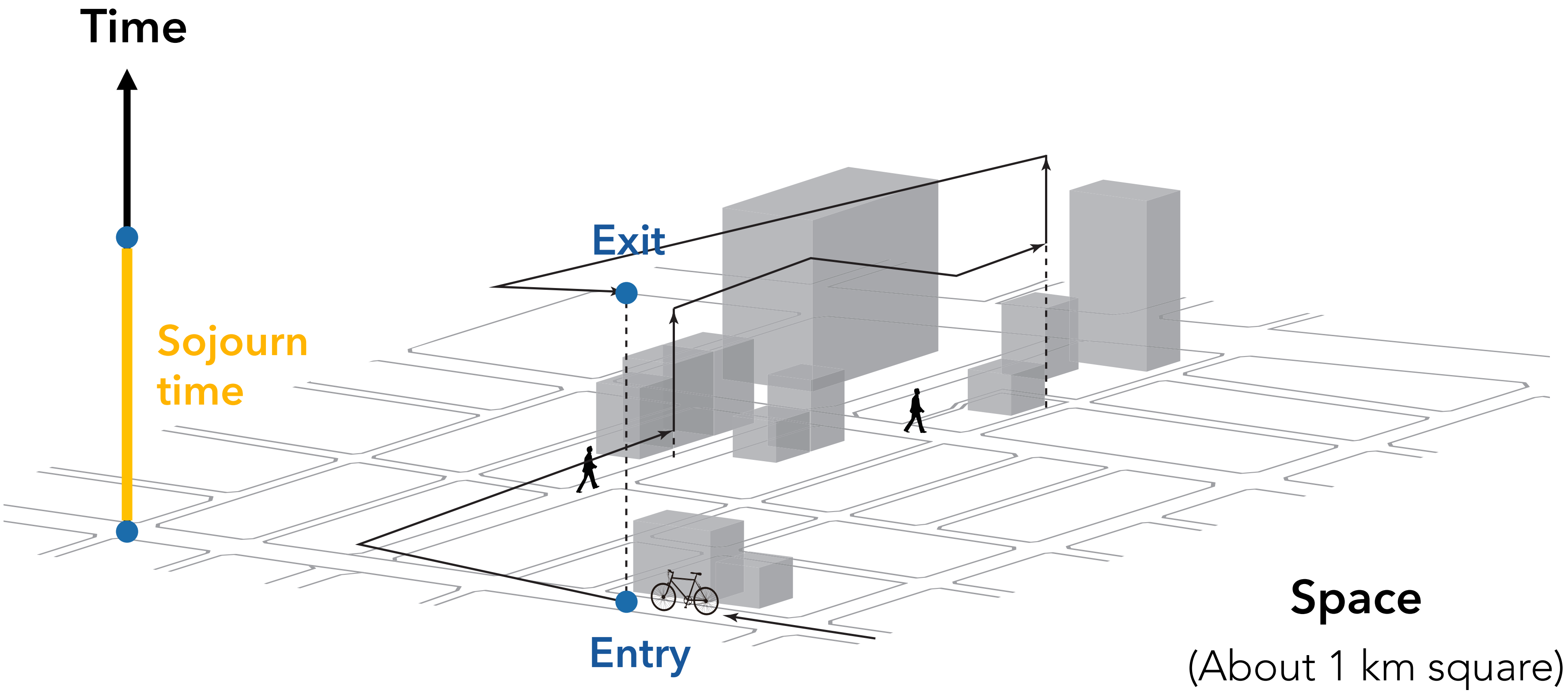
Large scale buildings
No. of visitors

Small scale projects
Frequency / Duration /…

How people spent their time in these districts?

Target | City Center Sojourn

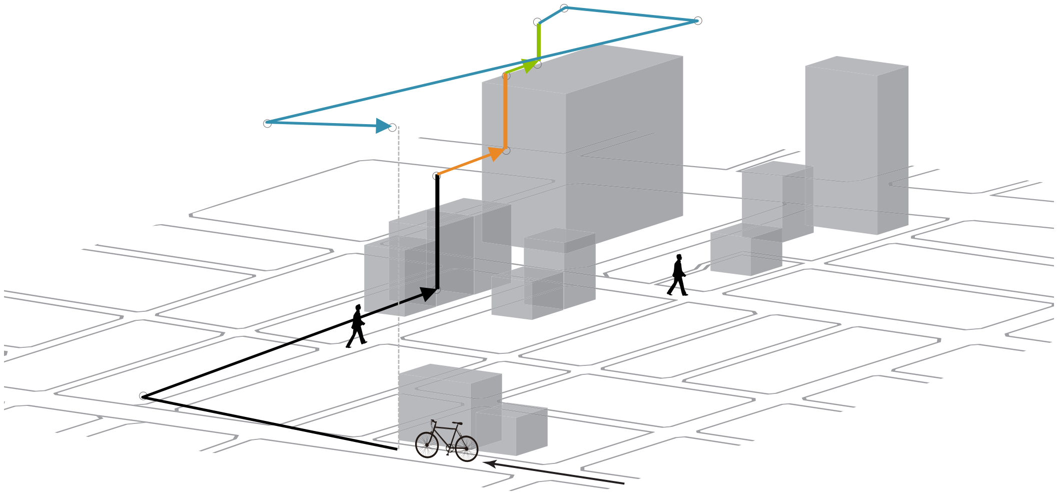
- **City Center Sojourn** refers to pedestrian scheduling behavior in city centers, which includes a sequence of moving (travel) and staying (activity) decisions.



Target | Pedestrian scheduling

Activities can be generated (walking) context-dependently

- Spatial attributes (stumbling on an attractive shop,…)
- Activity history (finding next shop for goods she wants, …)
- Social interaction (a friend says he wants to drop in a café,…) …



Pattern is not alternative but result of dynamic scheduling process

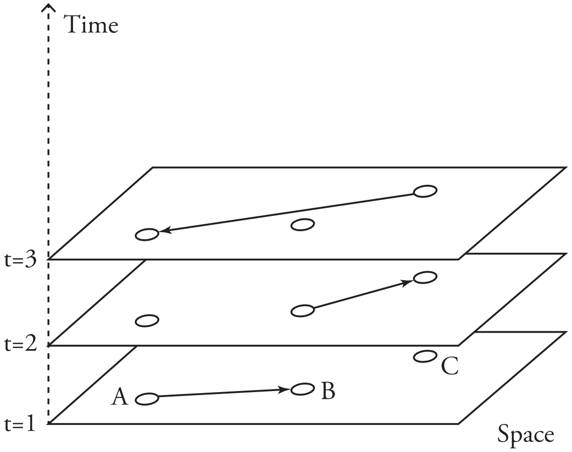
e.g.; Habib (2011)

Activities (staying) do not always decided to conduct before travels (moving)

Review | Scheduling models

1. Markov chain

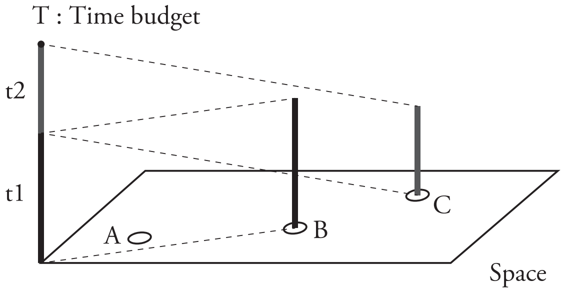
Lerman (1979),
Borgers and Timmermans (1986)



$$p_t(i,j)$$

2. Time allocation

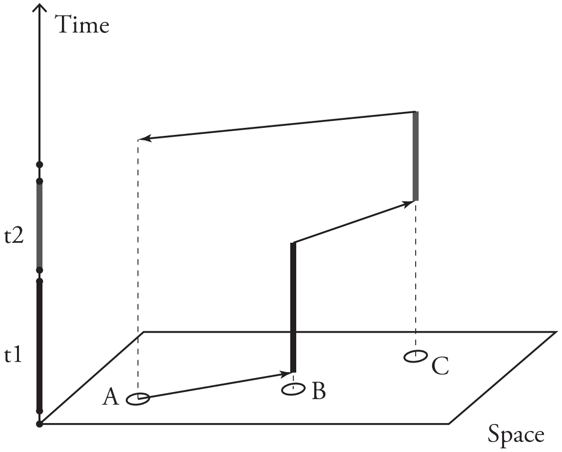
Bhat et al. (2005),
Fukuyama and Hato (2013)



$$\sum_{k=1}^{K} t_k = T$$

3. Utility maximization

Bowman and Ben-Akiva (2001),
Recker (1995)



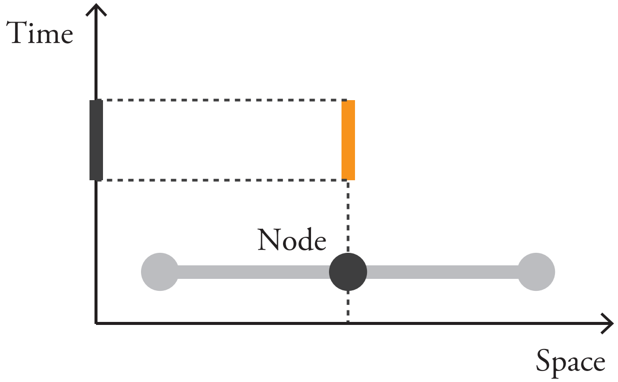
$$\max U$$

○ ordered	× random ordered	○ ordered
× separated	△ semi-separated	○ linked
○ context-dependent (but only at the time)	× independent (pre-trip)	× independent (pre-trip)

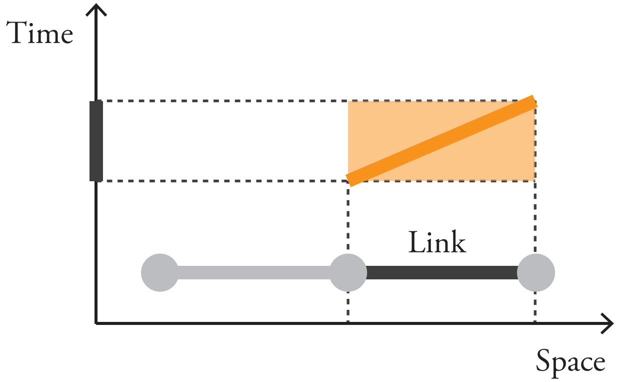
Model | Dynamic scheduling model

Activity = Time allocation behavior to a certain 'space'

- **Staying** : duration time choice in a certain **'node'** $n \in S$



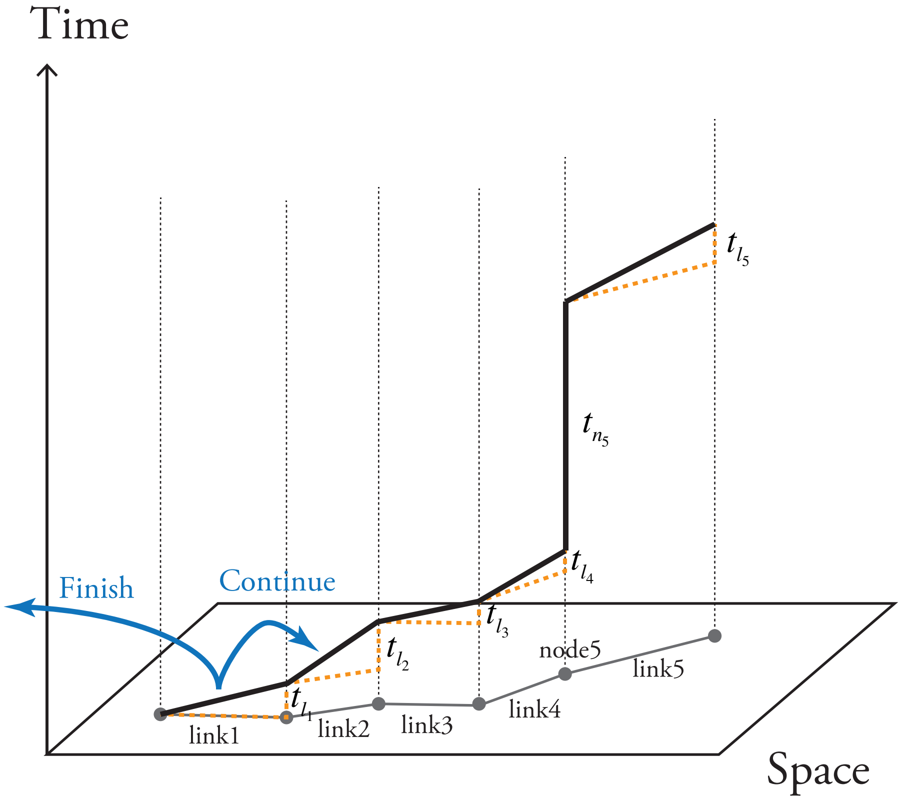
- **Moving** : duration time choice in a certain **'link'** $l \in S$



Model | Dynamic scheduling model

Activity = Time allocation behavior to a certain 'space'

- **Moving** : duration time choice in a certain **'link'** $l \in S$
- **Staying** : duration time choice in a certain **'node'** $n \in S$



Dynamic scheduling model in space

1. Activity generation model

***Continue* or *Finish* activities ?**

*'Continue' means moving next space
*'Finish' means moving out of district

What this decision is based on?

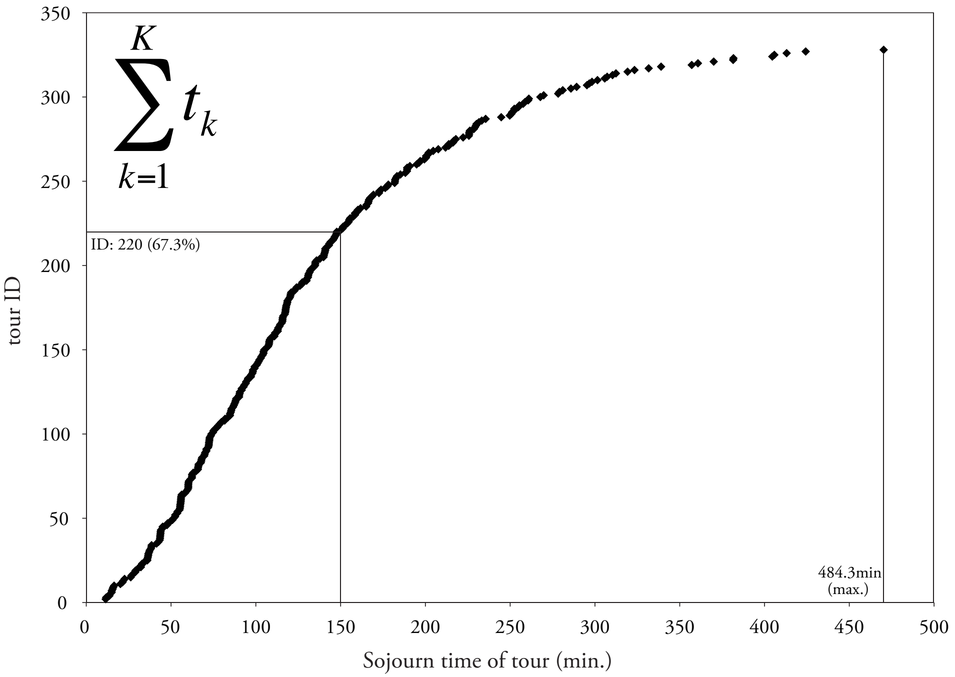
2. Time allocation model

Duration time **choice in the space**

$$\max u_k(t_k)$$

Model | Activity generation model

Is it enough with only time constraints ?



- **Non-mandatory tour** (shopping, eating, recreational, other activities are included).
- Sojourn time (cumulative duration) is **continuously distributed**.

- We **cannot explain** the sojourn time differences among tours **by only time constraint**.

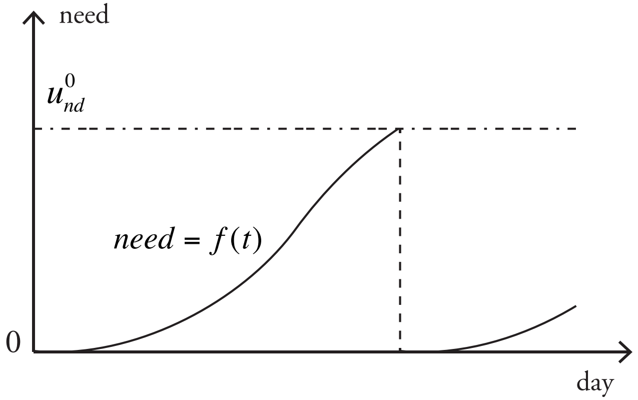
We have to consider:

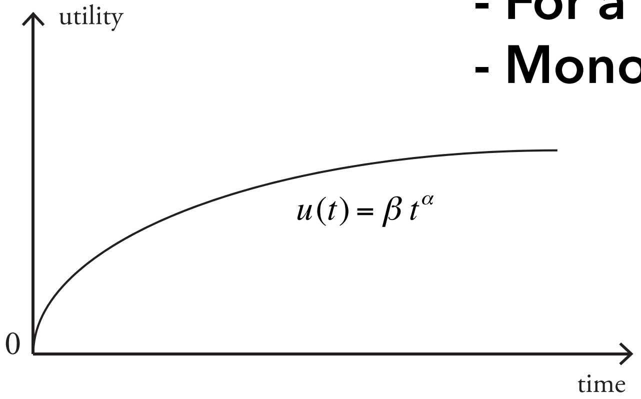
Psychological (personal) concept as resource

Model | Activity generation model

Psychological mechanisms in behavior modeling

- **Need** (Maslow, 1943; Arentze and Timmermans, 2004; Nijland et al., 2013)
- **Satisfaction** (Pattabhiraman et al., 2013)
- **Satiation** (MacAlister, 1982; Bhat et al., 2005)



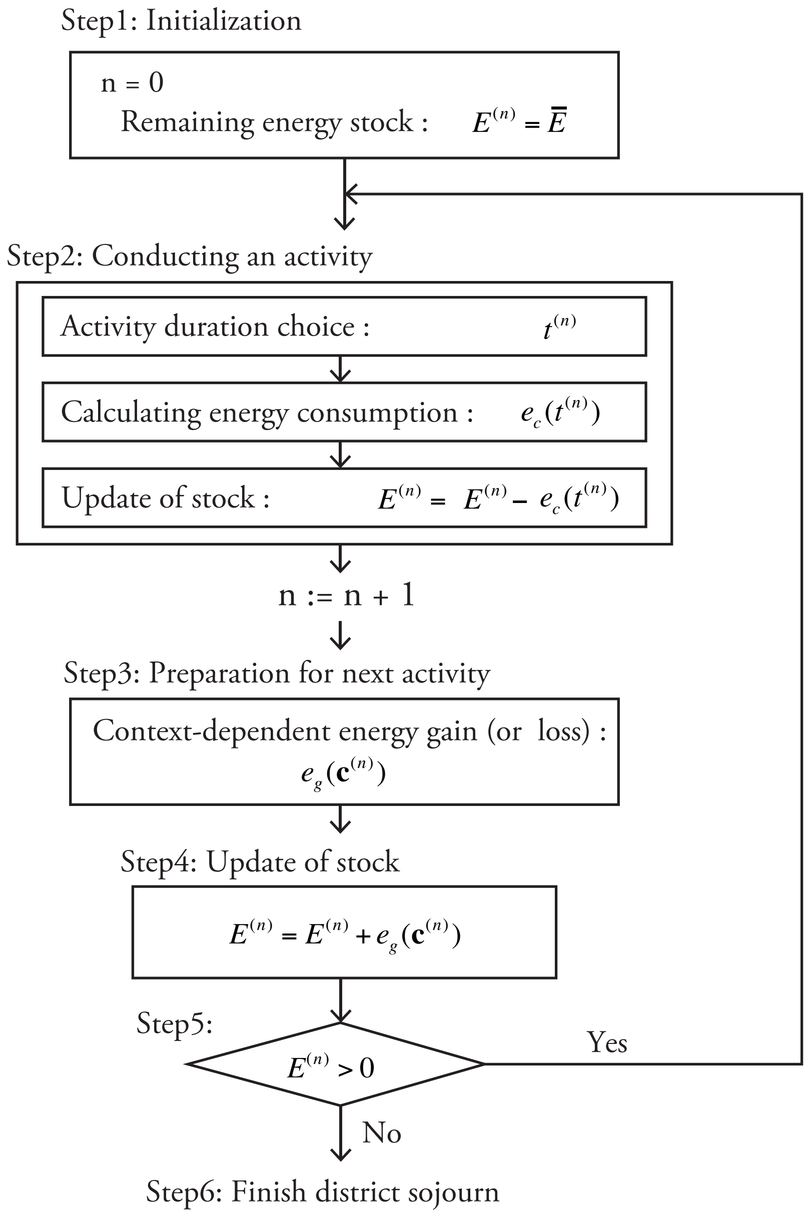


- **For a particular activity**
- **Monotonicity**

‘Energy’ : personal resource for engaging in activities.

- All of activities in **a tour have** ‘energy’ **in common**.
- Energy decreases by engaging activities, and can **increase based on context**.

Model | Activity generation model



Step1: Initialization

n = 0
Remaining energy stock : $E^{(n)} = \bar{E}$

Step2: Conducting an activity

Activity duration choice : $t^{(n)}$

Calculating energy consumption : $e_c(t^{(n)})$

Update of stock : $E^{(n)} = E^{(n)} - e_c(t^{(n)})$

n := n + 1

Step3: Preparation for next activity

Context-dependent energy gain (or loss) : $e_g(\mathbf{c}^{(n)})$

Step4: Update of stock

$E^{(n)} = E^{(n)} + e_g(\mathbf{c}^{(n)})$

Step5:

$E^{(n)} > 0$

Yes

No

Step6: Finish district sojourn

Remaining energy :

$$E^{(n+1)} = \bar{E}_{i,d} \underline{- E_c^{(n)}} \underline{+ E_g^{(n+1)}} \qquad (1)$$

Energy consumption :

$$E_c^{(n)} = \sum_{k=1}^{n} e_c(t^{(k)}, x^{(k)}) \qquad (2)$$

Energy gain (or loss) :

$$E_g^{(n)} = \sum_{k=1}^{n} \sum_{i} \gamma_{ki} I_{ki} + \sum_{j} \delta_{nj} s_{nj} \qquad (3)$$

I_{ni} : Activity history variables

s_{nj} : Spatial attributes variables

Model | Pedestrian scheduling model

Activity generation model

$$E^{(n)} = \overline{E}_{i,d} - E_c^{(n-1)} + E_g^{(n)} + \varepsilon \qquad (1)'$$

ε : random error term (i.i.d. gumbel distribution)

If energy is greater than zero, the sojourn will be continued, otherwise finished.

$$\Pr(continue) = \Pr(E^{(n)} > 0)$$

Time allocation model (Habib, 2011)

*k=1: next activity, k=2: composite activities

$$\max U(t_k) = \sum_{k=1}^{2} \frac{1}{\alpha_k} \exp\left(\psi_k z_k + \varepsilon'_k\right)\left(t_k^{\alpha_k} - 1\right) \qquad (4)$$

$$\text{s.t.,} \quad t_1 + t_2 = T \qquad (5)$$

α_k : satiation parameter (< 1) z_k : vector of variables ψ_k : vector of weights

ε'_k : random error term (i.i.d. gumbel distribution)

Model | Pedestrian scheduling model

Joint probability : Habib(2011)*

$$\Pr(continue \cap Time = t_k)$$

$$= \left(\frac{1-\alpha_1}{t_k} + \frac{1-\alpha_2}{T-t_k} \right) \cdot \frac{1}{\sigma} \exp\left(\frac{-(V'_2 - V'_1)}{\sigma} \right) \cdot \left[1 + \exp\left(\frac{-(V'_2 - V'_1)}{\sigma} \right) \right]^{-2} \times \Phi\left(\frac{J_d(\varepsilon) - \rho J_c(\varepsilon'_k)}{\sqrt{1-\rho^2}} \right) \tag{6}$$

where,

$$V'_k = \psi_k z_k + (\alpha_k - 1)\ln\left(t_k\right)$$

$J(\varepsilon)$: the inverse of CDF of standard normal distribution (Lee, 1983)

MLE (Maximum Likelihood Estimates)

$$L = \prod_{i=1}^{I} \left[\prod_{k=1}^{n} \left(\Pr_i(continue \cap Time = t_k) \right)^{\delta_{ic}} \right] \tag{7}$$

---

*Khandker M. Nurul Habib (2011). A random utility maximization (RUM) based dynamic activity scheduling model: Application in weekend activity scheduling, Transportation, Vol.38, pp.123-151.

Part 2 : Data processing

*

1. **Introduction** :
   - City Center Sojourn of pedestrians
2. **Modeling** :
   - Pedestrian dynamic scheduling model
   - Context-dependent energy
3. **Measurement Model** :
   - Probe Person data with GPS technologies
   - Detection of pedestrian activity paths

Part 2

4. **Case study** :
   - Model Estimation and Results
5. **Conclusions**

Data | Probe Person survey

Methods :

GPS (automatic) $\hat{m} = (\hat{x}, \hat{t})$

- Latitude / Longitude (a coordinate)
- Timestamp (at the interval of 5~30 s)

+ Web diary $a = (x, t^{-}, t^{+})$

- Trip purpose
- Transportation mode

+ personal information

Personal day-to-day data

Measurements : $\hat{m}_{1:J_i^d} = (\hat{m}_1, \ldots, \hat{m}_{j_i^d}, \ldots, \hat{m}_{J_i^d})$

Reported activity episodes : $a_{1:R_i^d} = (a_1, \ldots, a_{r_i^d}, \ldots, a_{R_i^d})$

where, i : an individual, d : a day

Data | Probe Person survey

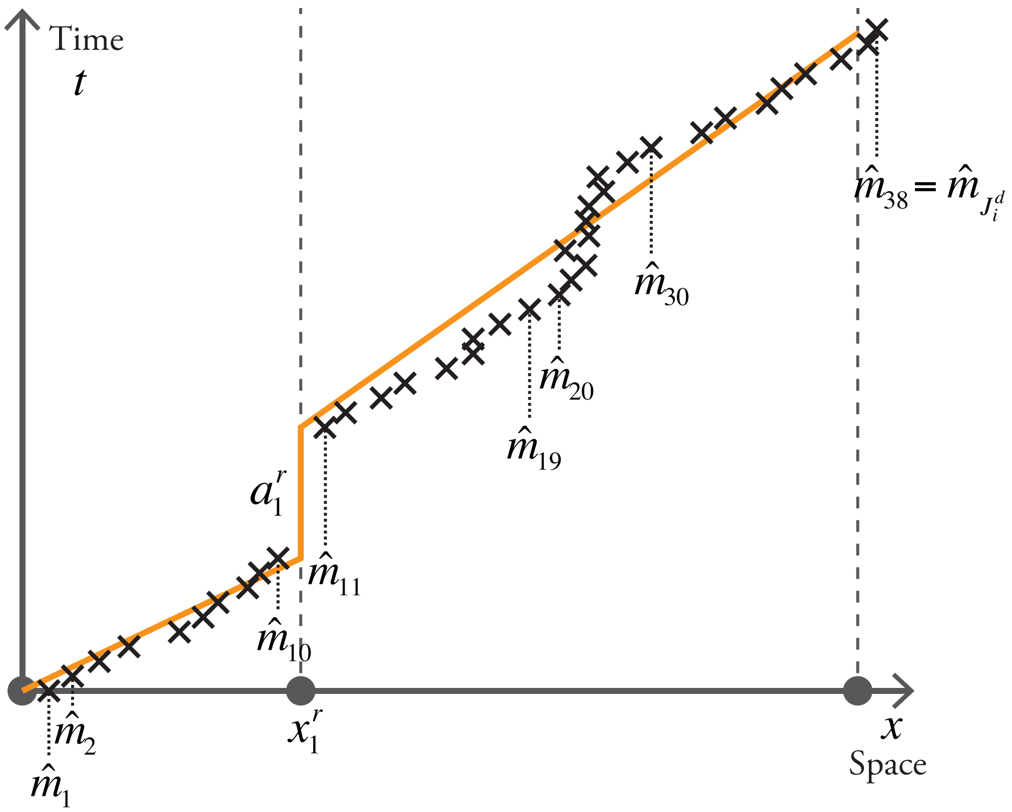
Reported path (—) :

$$\hat{m} = (\hat{x}, \hat{t})$$

$$a_r = (x, t^-, t^+)$$

- There can be dropped (non-reported) staying activity.
- Measurements have not connected with 'space' yet.

(and it has measurement error)

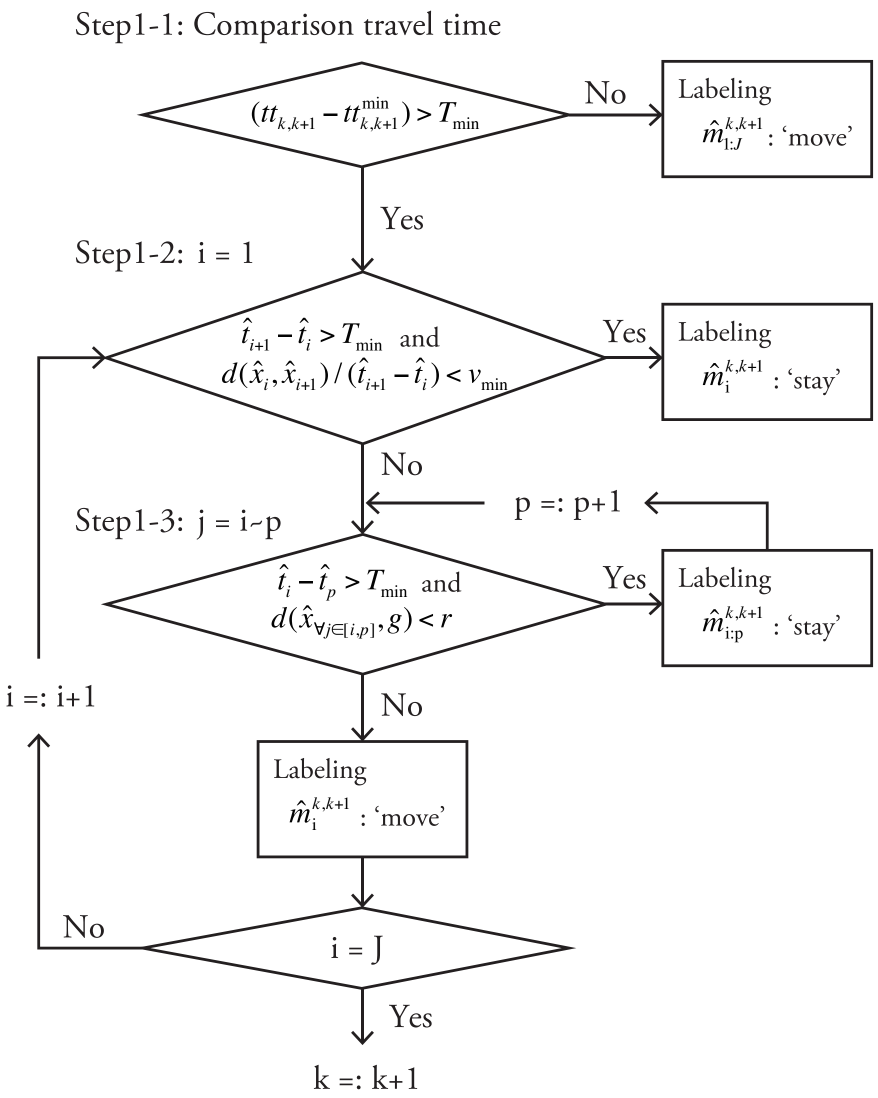


We need to

1. **Label measurements ('moving' or 'staying')**
2. **Connect measurements with 'space' (node / link)**

Data | Detection of activity paths

Step1: Classification of moving or staying



Step1-1: Comparison travel time

$(tt_{k,k+1} - tt^{\min}_{k,k+1}) > T_{\min}$

No

Labeling $\hat{m}^{k,k+1}_{1:J}$: 'move'

Yes

Step1-2: i = 1

$\hat{t}_{i+1} - \hat{t}_i > T_{\min}$ and $d(\hat{x}_i, \hat{x}_{i+1}) / (\hat{t}_{i+1} - \hat{t}_i) < v_{\min}$

Yes

Labeling $\hat{m}^{k,k+1}_{i}$: 'stay'

No

p =: p+1

Step1-3: j = i~p

$\hat{t}_i - \hat{t}_p > T_{\min}$ and $d(\hat{x}_{\forall j \in [i,p]}, g) < r$

Yes

Labeling $\hat{m}^{k,k+1}_{i:p}$: 'stay'

No

Labeling $\hat{m}^{k,k+1}_{i}$: 'move'

i = J

No

i =: i+1

Yes

k =: k+1

$tt_{k,k+1} = t^{-}_{k+1} - t^{+}_{k}$: Reported travel time

$$tt^{\min}_{k,k+1} = d(x_k, x_{k+1}) / v_w \qquad (8)$$

: Shortest path travel time

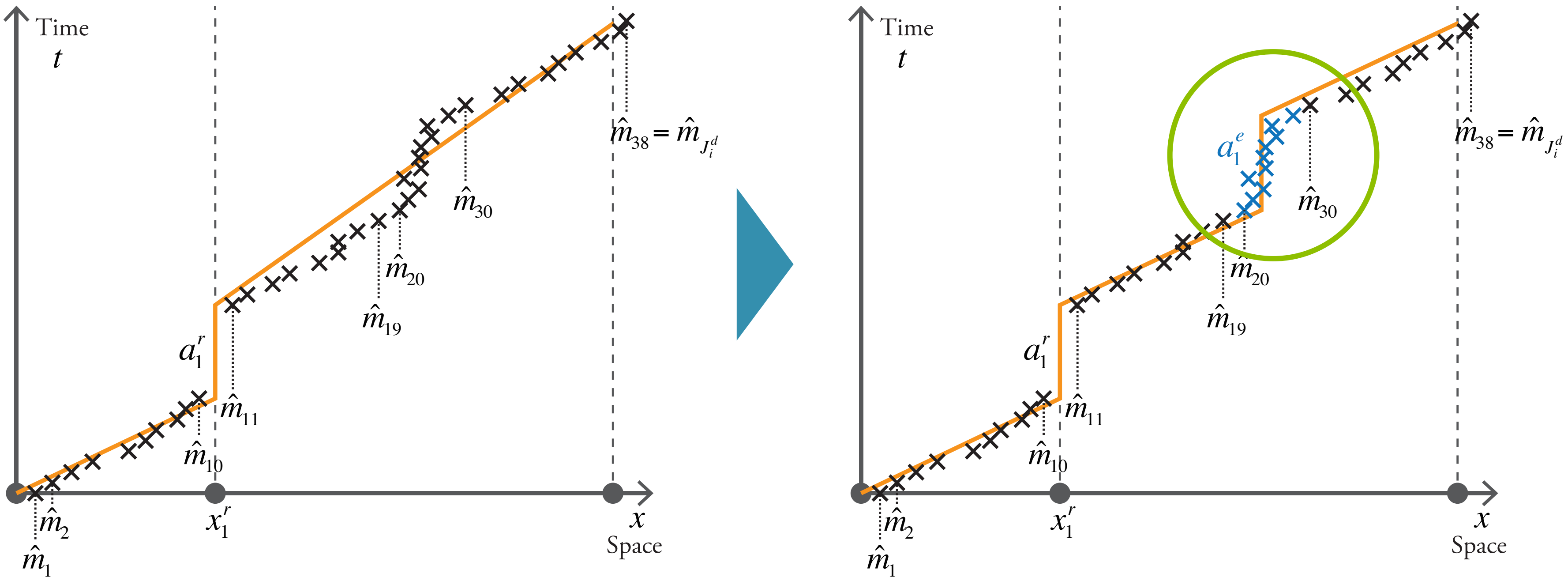
Centroid of $\hat{m}^{k,k+1}_{i:p}$

$$g_{i:p} = \frac{1}{p-i+1}\left(\sum_{j=i}^{p} \hat{x}_{jlat}, \sum_{j=i}^{p} \hat{x}_{jlon}\right) \qquad (9)$$

*$T_{\min} = 180s,\ v_w = 1.4m/s,\ r = 50m$

Data | Detection of activity paths

Step1: Classification of moving or staying



× $\hat{m} = (\hat{x}, \hat{t}, '\mathit{move}')$

× $\hat{m} = (\hat{x}, \hat{t}, '\mathit{stay}')$

Next: **Connect activities with 'space'** (move - link / stay - node)

Data | Detection of activity paths

Step2: Estimation of activity space for 'stay' data

Step2-1: Candidate set generation

Universal set: $\mathcal{U}_N = \{n : n \in S\}$

Space frequency score from day-to-day data:

$$f_{ni} = \sum_d \sum_k \delta_{k,n}^{i,d}, \quad f_n = \sum_i \sum_d \sum_k \delta_{k,n}^{i,d} \tag{10}$$

$\delta_{k,n}^{i,d}$: 1 if individual *i* stay *n* for activity *k* on day *d*, otherwise 0.

Importance Sampling using MCMC method

Adoption rate of *i* :

$$r_i = P_i / P_j = \exp(V_{in}) / \exp(V_{jn}), \quad V_{in} = \sum_j \beta_{nj} X_{nj} + w_1 f_{ni} + w_2 f_n \tag{11}$$

Finally we get a subset: $C_{iN} \subset \mathcal{U}_N$

Data | Detection of activity paths

Step2: Estimation of activity space for 'stay' data

Step2-2: Probability calculation

e.g.; Danalet et al. (2014)

Prior probability:
$$P_i(n) = \exp(V_{in}) / \sum_{m \in C_{iN}} \exp(V_{im}) \tag{12}$$

Measurement probability:

$$P(\hat{m}_{p:q} \mid n) = P(\hat{x}_{p:q} \mid x_n) = \Pi_{j=p}^{q} P(\hat{x}_j \mid x_n)$$
$$P(\hat{x}_j \mid x_n) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(- \frac{(\hat{x}_j - x_n)^2}{2\sigma^2} \right) \tag{13}$$

* We assume that measurement error is only localization σ

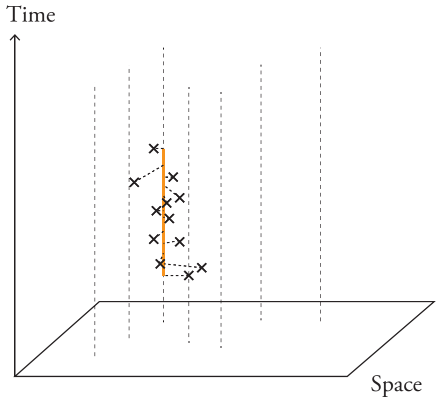
Probability of space *n* for 'stay' measurement set $\hat{m}_{p:q}$:

$$P(n \mid \hat{m}_{p:q}) = a \cdot P(\hat{m}_{p:q} \mid n) \cdot P_i(n) \tag{14}$$

Data | Detection of activity paths

Detected activity path:

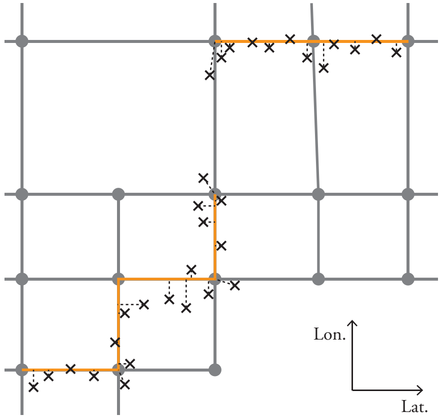
Step2: Estimation of 'stay' space



$$\hat{m} = (\hat{x}, \hat{t}, '\mathit{stay}')$$

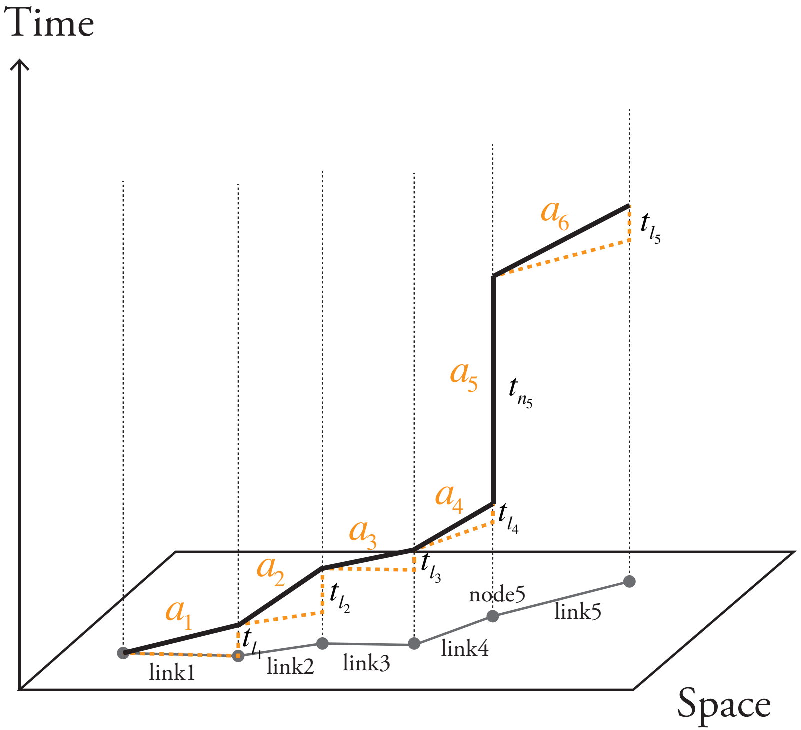
$$a = (n, t^-, t^+)$$

Step3: Estimation of 'move' space
(Map-matching)



$$\hat{m} = (\hat{x}, \hat{t}, '\mathit{move}')$$

$$a = (l, t^-, t^+)$$



Activity sequence with space

$$a_{1:M_i^d} = (a_1, \ldots, a_{m_i^d}, \ldots, a_{M_i^d})$$

Part 3 : Case study in Matsuyama city

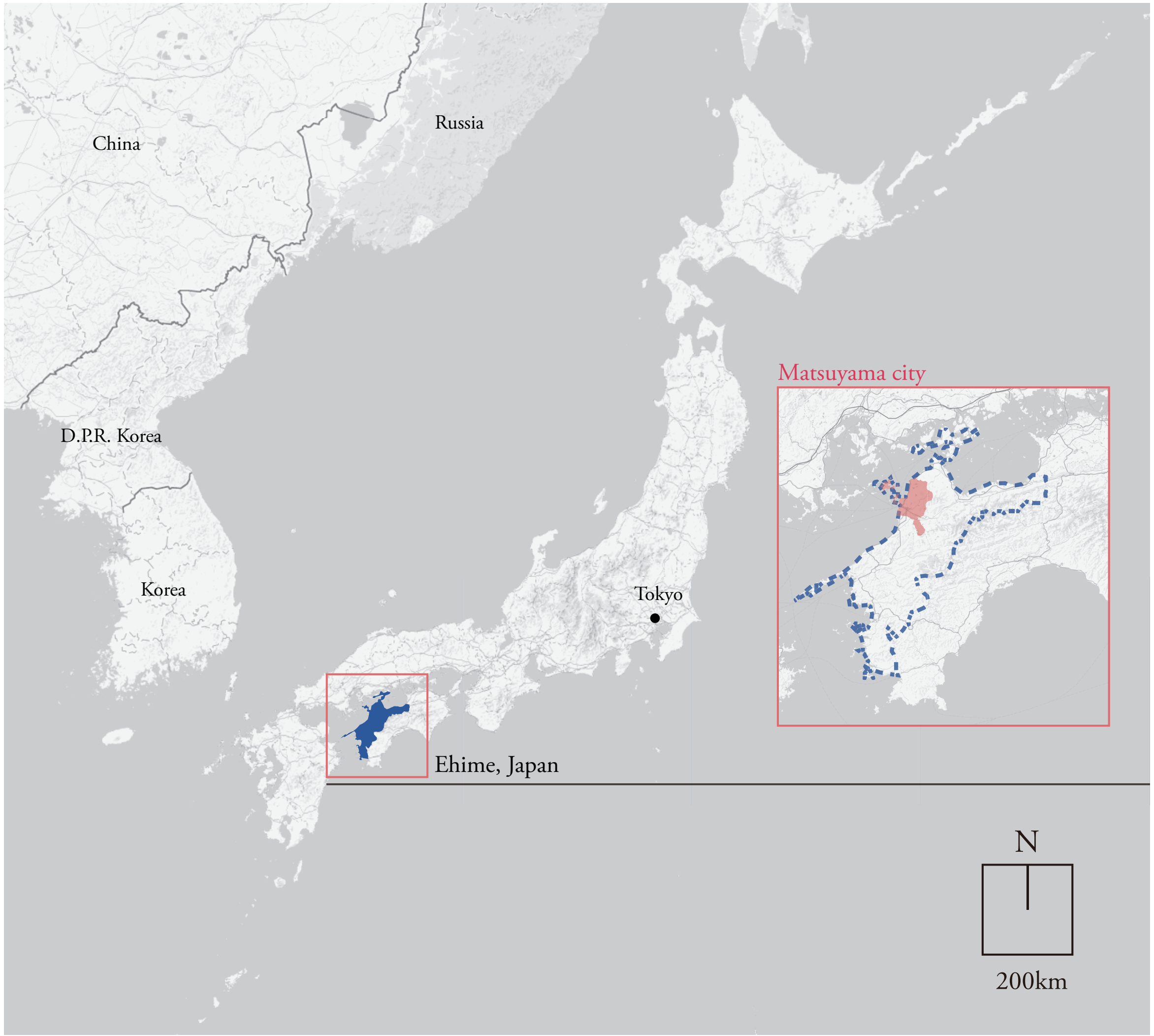
*

1. **Introduction** :
   - City Center Sojourn of pedestrians
2. **Behavior Model** :
   - Pedestrian dynamic scheduling model
   - Context-dependent energy
3. **Measurement Model** :
   - Probe Person data with GPS technologies
   - Detection of pedestrian activity paths
4. **Case study** :
   - Model Estimation and Results
5. **Conclusions**

Part 3

Case study | City center of Matsuyama

Matsuyama city :



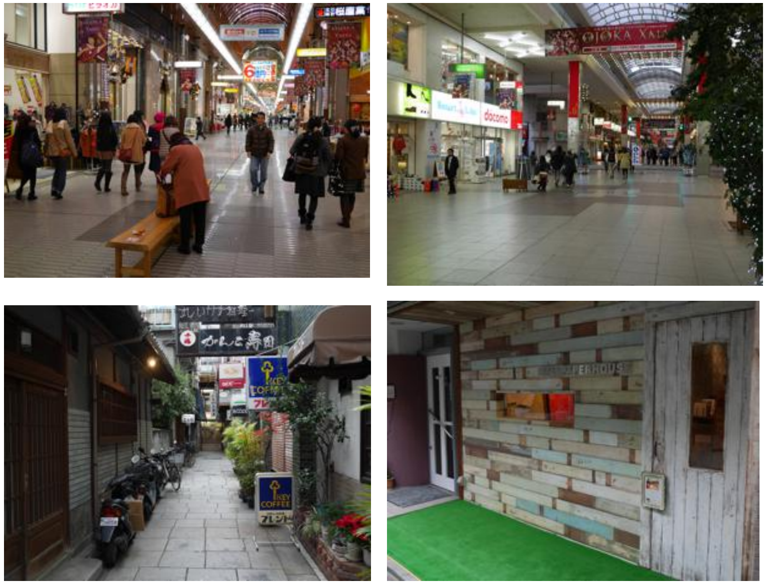
Data

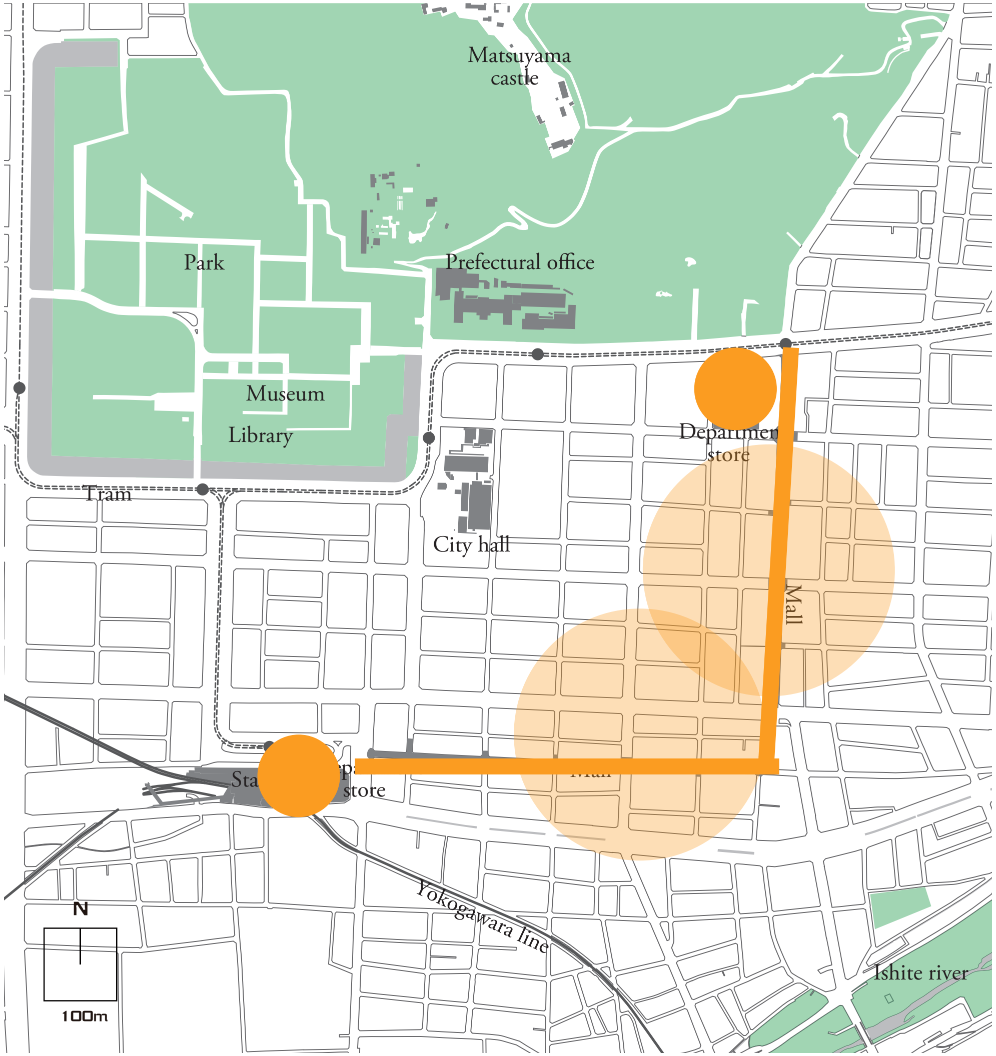
- Ehime prefecture, Shikoku region
- Population: 516,637 (December 1, 2010)
- Area: 428.86 sq. km
- Density: 1,204.68/sq. km

Case study | City center of Matsuyama

City center of Matsuyama :

- 2 department stores / 2 malls
- Various retails and restaurants are located around the streets.





About 1.5 km square

Case study | PP survey in Matsuyama

Data collection :

Survey	Period	Weeks	No. of monitors	Data (trip)
CityCenterPP2007	2007/02/19~2007/03/22	4	84	7,810
PP survey 2007A	2007/10/29~2008/01/21	12	508	17,697
PP survey 2007B	2007/10/29~2008/01/21	12	205	14,706
Bike sharing PP	2009/02/21~2009/03/07	2	15	668
Elderly PP 2010	2010/11/18~2011/01/31	12	30	1.380
Total		42	842	42,261

-> 1582 sojourn tours (non-mandatory) were observed

Estimation | Variables of Energy

Remaining energy $E^{(n+1)} = \bar{E}_{i,d} - E_c^{(n)} + E_g^{(n+1)}$

Initial stock of energy

$$\bar{E}_{i,d} = \sum_k \alpha_k x_k$$

- Female dummy (sex)
- Car inflow dummy (mode)
- Dist. between Entry point and Main (Location of entry point)

Energy consumption $E_c^{(n)} = \sum_{k=1}^{n} e_c(t^{(k)}, x^{(k)})$

By staying : $e_c^s = (\beta_{time}^s + \sum_k \beta_k^s x_k^s) \cdot t$

- Shopping purpose dummy
- How many times

By moving: $e_c^m = (\beta_{time}^m + \beta_{speed}^m s + \sum_k \beta_k^m x_k^m) \cdot t$

- Sidewalk width
- Shooing street dummy

Energy gain (or loss)

$$E_g^{(n)} = \sum_i \gamma_{ki} I_{ki} + \sum_j \delta_{nj} s_{nj}$$

- Cumulative number of activities
- Previous trip purpose

- Dist. from EP or Main facilities
- Shopping street dummy

Estimation | Results

	Variable	Parameter	t-value
ϱ	Correlation	0.048	0.90
α_c	Satiation Parameter of composite	0.594	12.98**
α_m	Satiation Parameter of moving	-3.196	-93.23**
α_s	Satiation Parameter of staying	-0.176	-8.92**
	Discrete choice		
α_1	Constant	5.354	26.03**
α_2	Female dummy	-0.170	-1.52
α_4	Log(EP-Main dist.(km) + 1)	0.164	11.07**
α_5	Car inflow dummy	0.632	5.06**
β_1	Basic parameter of time (min.)	-0.004	-6.75**
β_2	A Number of trips	0.002	6.74**
β_3	Shopping dummy	-0.003	-2.60**
β_4	Basic parameter of time (min.)	-0.264	-7.15**
β_5	Walking speed (m/s)	-0.063	-8.49**
β_6	Sidewalk width (m)	0.070	6.42**
β_7	Shopping street dummy	-0.240	-3.72**
γ_1	Cumulative stay activities	-0.986	-17.65**
γ_1	Cumulative move activities	0.638	10.78**
γ_2	Cumulative shopping stays	0.246	3.67**
γ_4	Previous activity : eating	0.364	1.82
γ_5	Previous activity : Main	-1.957	-14.24**
γ_4	Dist. from EP	-0.040	-2.60**
γ_5	Dist. from Main	-0.230	-14.52**

	Variable	Parameter	t-value
	Continuous choice (move)		
β_{mc}	Constant	-1.054	-5.17**
β_{time1}	Elapsed time (min./10)	0.006	2.55**
β_{time2}	Cumulative stay activities	-0.004	-0.16
β_{time3}	Cumulative move activities	-0.006	-1.13
β_{time4}	Link length (m)	0.024	21.56**
β_{time5}	No. of lanes	0.104	2.98**
β_{time6}	Sidewalk width (m)	-0.067	-2.49**
β_{time7}	Shopping street	0.665	2.99**
β_{time8}	Street trees	-0.057	-0.56
	Continuous choice (stay)		
β_{ms}	Constant	2.368	10.89**
β_{time9}	Elapsed time (min./10)	-0.004	-0.86
β_{time10}	Cumulative stays	-0.175	-4.87**
β_{time11}	Cumulative moves	-0.042	-4.45**
β_{time12}	Public facilities	0.113	0.99
β_{time13}	Department store	-0.564	-4.78**
β_{time14}	Shopping street	-0.606	-3.86**
β_{time15}	No. of retails	-0.029	-2.48**
	Observations		7247
	Initial Likelihood		-24949.15
	Final Likelihood		-18855.90
	Rho square (adj.)		0.243

Conclusions

- To capture **context-dependent activity generation** and **scheduling process** in pedestrian behavior,
- we incorporated "**energy**" into the scheduling model and described **a sequential time-allocation behavior to spaces**.
- And using PP data, **we detected activity paths with space**.

- As a result, it was clarified that the energy consumption and gain process are dependent on some **behavioral and spatial context variables**.

Thank you for your attentions!!
Questions?

Contact: oyama@bin.t.u-tokyo.ac.jp
or oymyk.dom@gmail.com