Choice set generation of pedestrian route choice using data distribution of walking behavior in urban space

Sachiyo Fukuyama
National Institute for Land and Infrastructure Management

International BinN research seminar #4
June 13, 2015

1 Problem of route choice set generation

- Alternatives exist infinitely
- Observation of actual choice sets is difficult
- Individual's choice set is small and different individually

↓

The problem is
how to generate the choice set
for unbiased parameter estimation and prediction

1 Problem of route choice set generation

Characteristics of pedestrian route choice and needs for choice set

・Sensitivity to their surroundings

→Important to include **alternatives with various attributes**

・Physical and mental limitation of reach

→Important to **limit the range of the choice set** for efficiency of estimation

2 Data oriented choice set generation

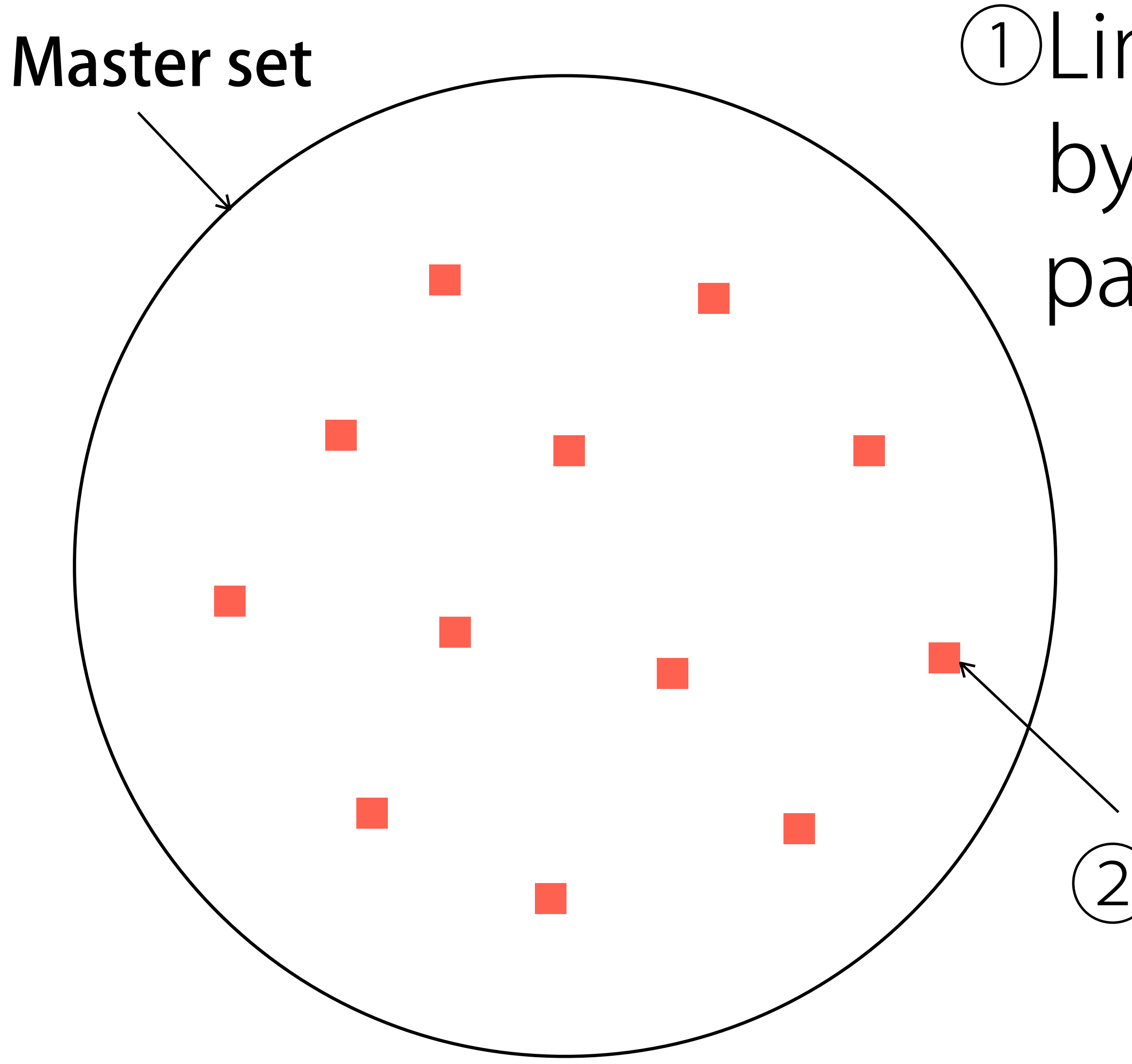
To generate the choice set
which contains **alternatives with various attribute**,

and which **reflect the pedestrian behavior**
under the difficulty of modeling from actual choice set,

→ Use the behavioral patterns obtained from
individual trace data

→ generalize the model by extracting common rules
from all of the related data

2 Data oriented choice set generation



①Limit the range of master choice set
by the threshold of the behavioral
patterns obtained from data

②Sampling of alternatives based
on the data distribution
: Importance sampling for
efficiency of model estimation

3 Route choice set generation methods

Sampling methods based on path lengths

Frejinger et al. (2009) A stochastic path generation approach

Sample each link from the origin to destination based on the link probability

Link sampling probability: $$q(\ell|\mathscr{E}_v, b_1, b_2) = \frac{\omega(\ell|b_1, b_2)}{\sum_{m\in\mathscr{E}_v} \omega(m|b_1, b_2)}$$

Weight of each link:
based on its distance
to the shortest path

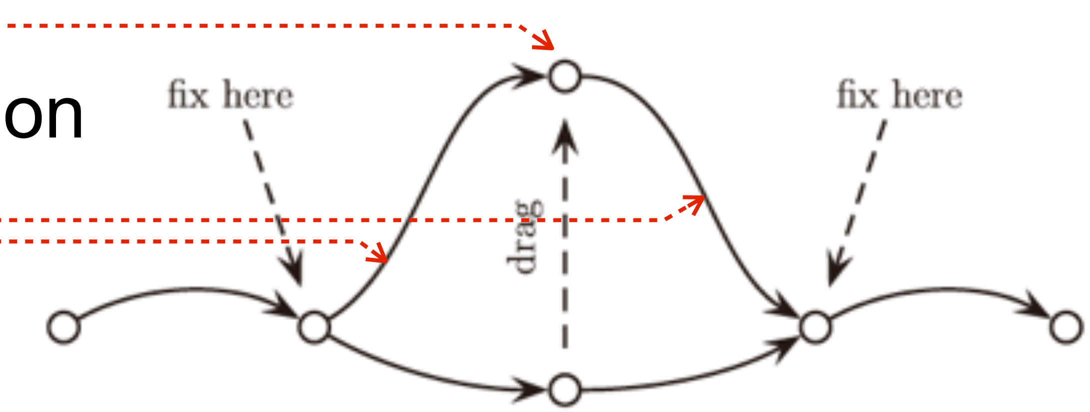
$$\omega(\ell|b_1, b_2) = 1 - (1 - x_\ell^{b_1})^{b_2}, \quad x_\ell = \frac{SP(v, s_d)}{C(\ell) + SP(w, s_d)}$$

Frötteröd and Bierlaire (2013) Metropolis-Hastings sampling of paths

Sample paths according to a given distribution from a general network

Draw an insertion node
according to pre-specified distribution

Compute shortest paths
between the insertion node
and fixed nodes



Route choice set generation methods

Choice set generation method based on attributes

Labeling approach, Ben-Akiva et al. (1984)

1. Label each route according to a criterion for which the path is optimum
2. Select the route whose labeled attribute is maximum/minimum

Guo and Loo (2013) **A modified labeling approach**

2 modification;
(1) A label is based on a linear generalized cost from all route attributes
(2) Combine different labels in the choice set generation (to reflect trade-offs)

4 Method in this research

Branch and bound

Prato and Bekhor (2006)

Path generation by constructing a connection tree
between the origin and destination of a trip
by processing sequences of links according to a branching rule

+

Sampling

Sampling of links **based on the distribution of branching rules**

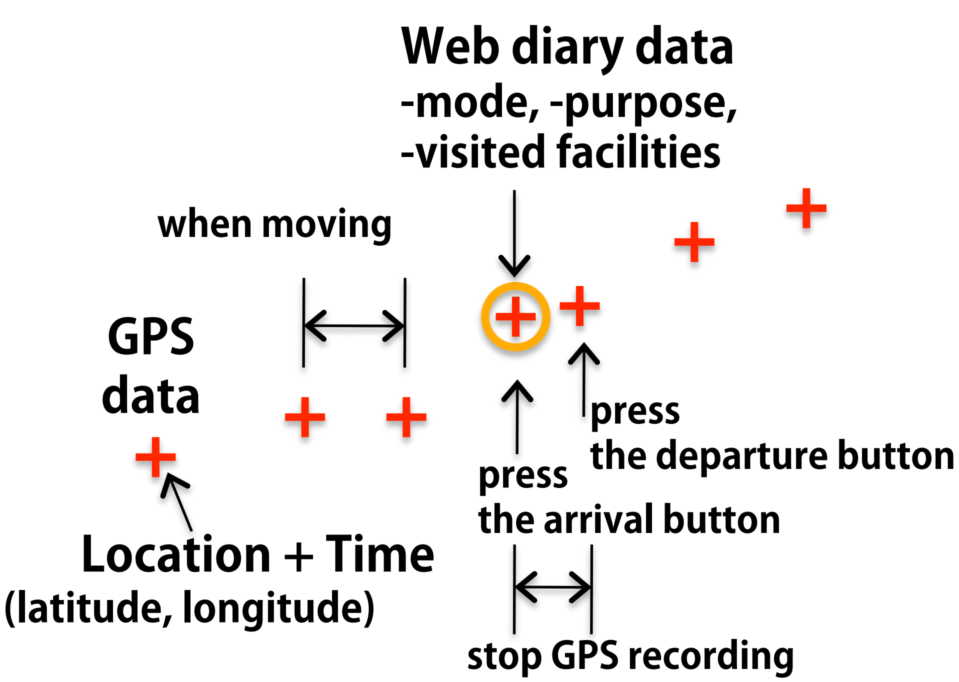
Obtaining the distribution from **observation data**
: Probe person data

5 Data

Probe person data

Probe Person survey is the method for
tracking individual travel behavior in urban space
by using **an automatic position and time recording system**
based on **GPS** and **internet communications**

→ Provides us **Individuals' precise position** and **time data**
→ Provides us long term data observed for same respondents

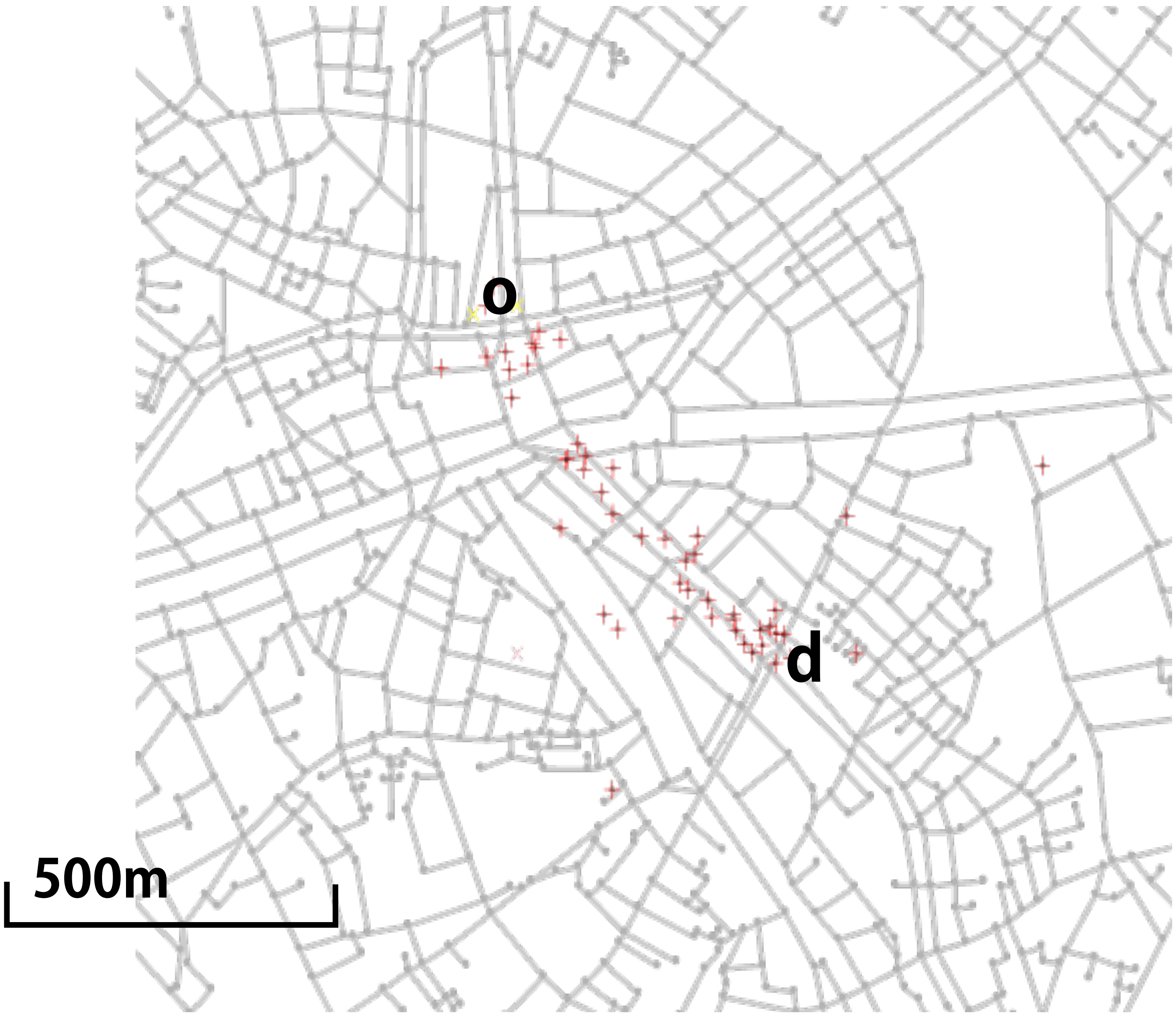


Surveillance period	2 months from 2005/11/19
Area	Shibuya, Tokyo
The number of respondents	39
The number of trips	422

5 Data

Probe person data

Example of GPS trace of a trip of an individual in Shibuya area



GPS dot data

↓

Map matching

↓

Link trace data

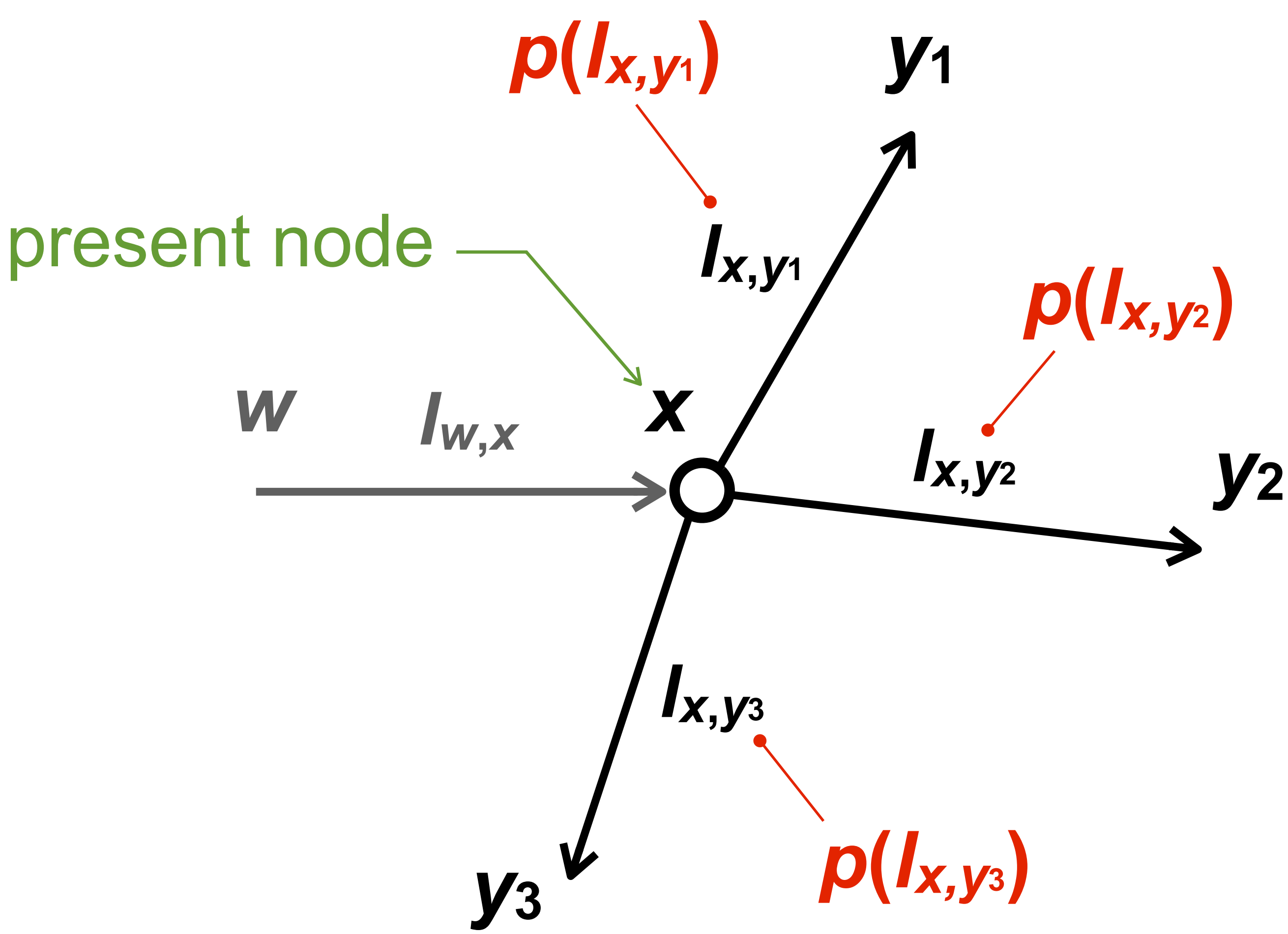
↓

Extracting branching rules

6 Link sampling method

Sampling of links based on the distribution of branching rules

$$p(l_{x,y_i}) = f_1(l_{x,y_i})\, f_2(l_{x,y_i})\, f_3(l_{x,y_i}) \,/\, Z$$



Branching rules;

① destination directivity

② keep-direction rate

③ detour rate

l_{x,y_i} : links to be chosen

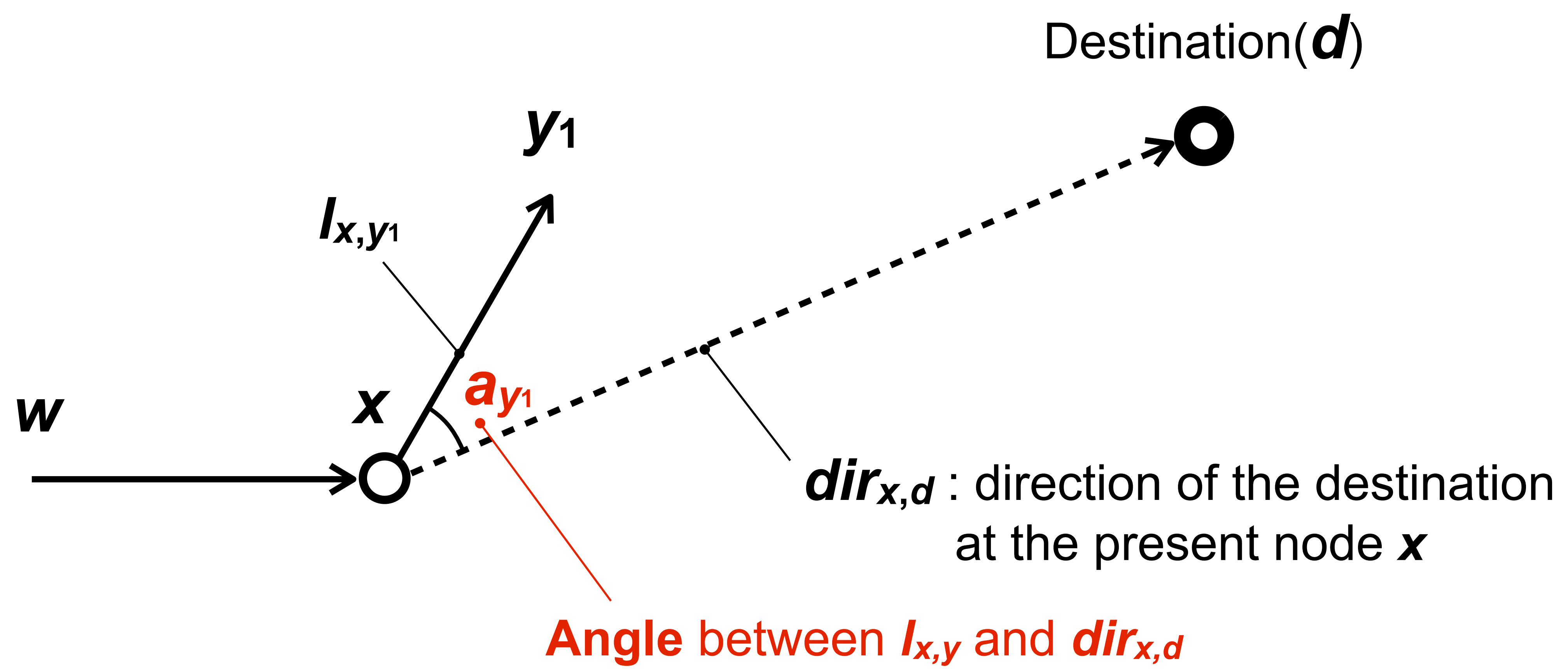
$p(l_{x,y_i})$: probability distribution of l_{x,y_i}

6 Link sampling method

① destination directivity

Tendency to head for the direction of the destination ***d***

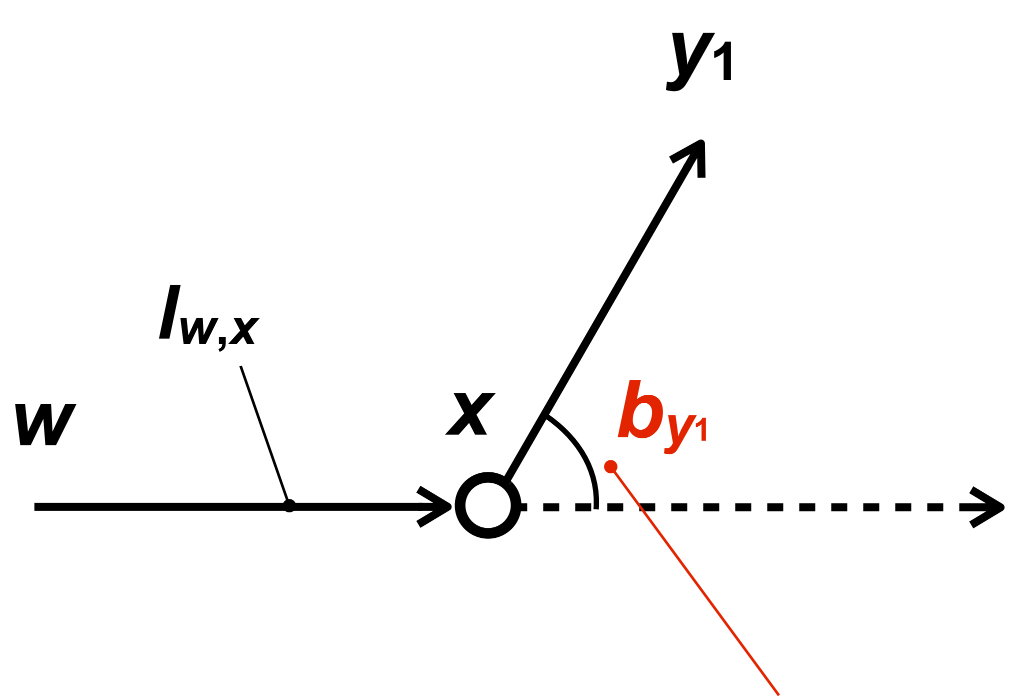
Takegami and Tsukaguchi (2006)



② keep-direction rate

Tendency to keep the same direction as the processed link $l_{w,x}$

Takegami and Tsukaguchi (2006)



Angle between the direction of $l_{w,x}$ and $l_{x,y}$

6 Link sampling method

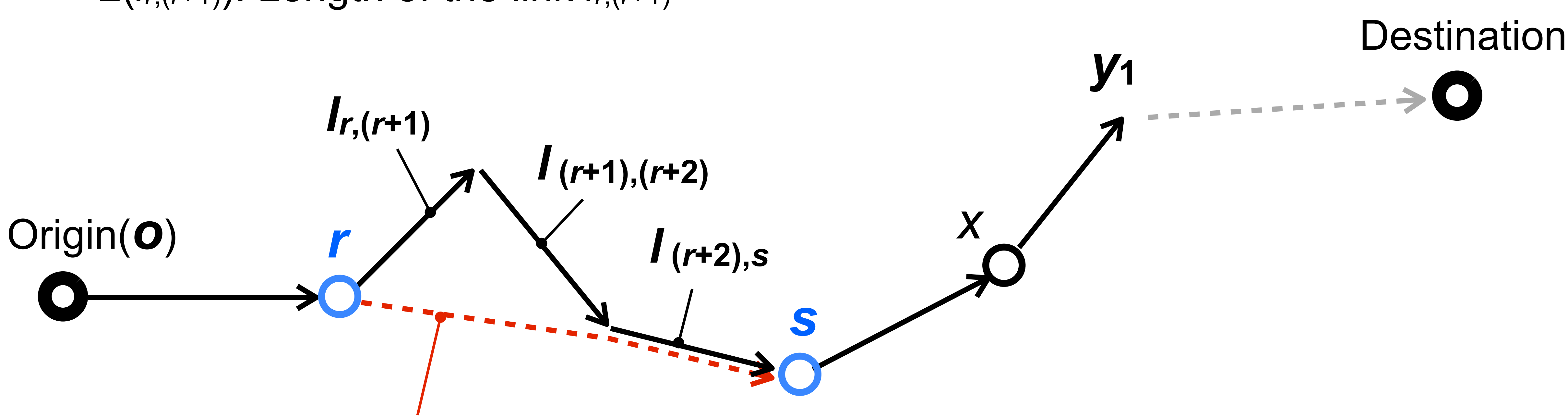
③ detour rate

The rate of **the length of each path subsegment on $P_{o,y}$** (path segment between o and y) **relative to the shortest path** between its start and end nodes

$$c_y = \max\left(\frac{\sum_{l_{i,j}\in P_{r,s}} L(l_{i,j})}{SP(r,s)}\right) \qquad \forall P_{r,s} \in P_{o,y}$$

$P_{r,s}$: Path subsegment between nodes *r* and *s* on the path segment $P_{o,y}$

$L(l_{r,(r+1)})$: Length of the link $l_{r,(r+1)}$



***SP*(*r*,*s*)**: Shortest Path between node *r* and node *s*

Slice Sampling

One of the methods of MCMC (Markov chain Monte Carlo)

Sampling from $$f(\mathbf{x}) = b \prod_{k=1}^{m} p_k(\mathbf{x}), \quad \mathbf{x} \in X$$

↓

Input auxiliary variable *y* and $$f(\mathbf{x}, \mathbf{y}) \propto \prod_{k=1}^{m} I_{\{0 \le y_k \le p_k(\mathbf{x})\}}$$

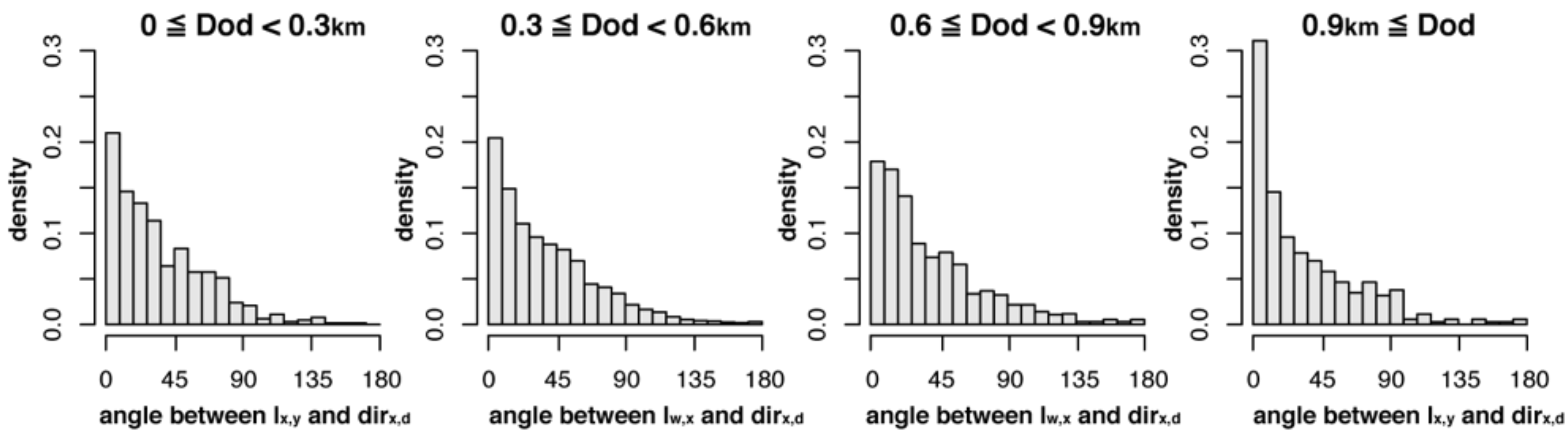
Repeat sampling from *f*(*y*|*x*) and *f*(*x*|*y*) : Gibbs sampling

Algorithm

0. $\mathbf{x}_1 \in X, t = 1$
1. Generate $\mathbf{y}$ from $f(\mathbf{y}|\mathbf{x}_t)$;
   generate $U_1, \ldots, U_m \sim U(0,1)$ and $y_k = U_k p_k(\mathbf{x}_t), \quad k = 1, \ldots, k$
2. Generate $\mathbf{x}_{t+1}$ from $f(\mathbf{x}|\mathbf{y})$;
   random choose $\mathbf{x}_{t+1}$ from $\{\mathbf{x} : p_k(\mathbf{x}) \ge y_k, k = 1, \ldots, m\} \cap \mathbf{X}$
3. If $t = T$, stop, else $t = t + 1$ and back to Step1

8 Data distribution

① destination directivity



Assume **the exponential distribution** and estimate the parameters

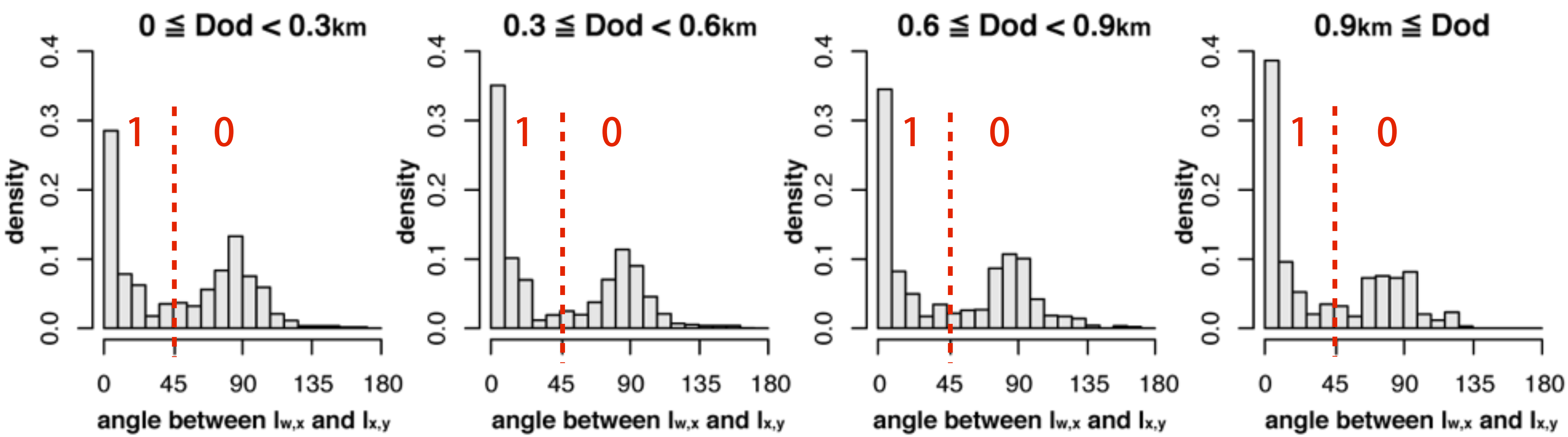
Assume **the rate parameter is proportional to *D_{od}*** (the distance between the origin and the destination)

$$f_1(a_{y_i}, \lambda_1) = \lambda_1 e^{-\lambda_1 a_{y_i}/180}$$

$$\lambda_1 = 4.6898962 + 0.2327133 * D_{od}$$

8 Data distribution

② keep-direction rate



Assume **the Bernoulli distribution**;
keep the direction (***α*** **< 45°)** or **change the direction (45°≦** ***α*** **)**

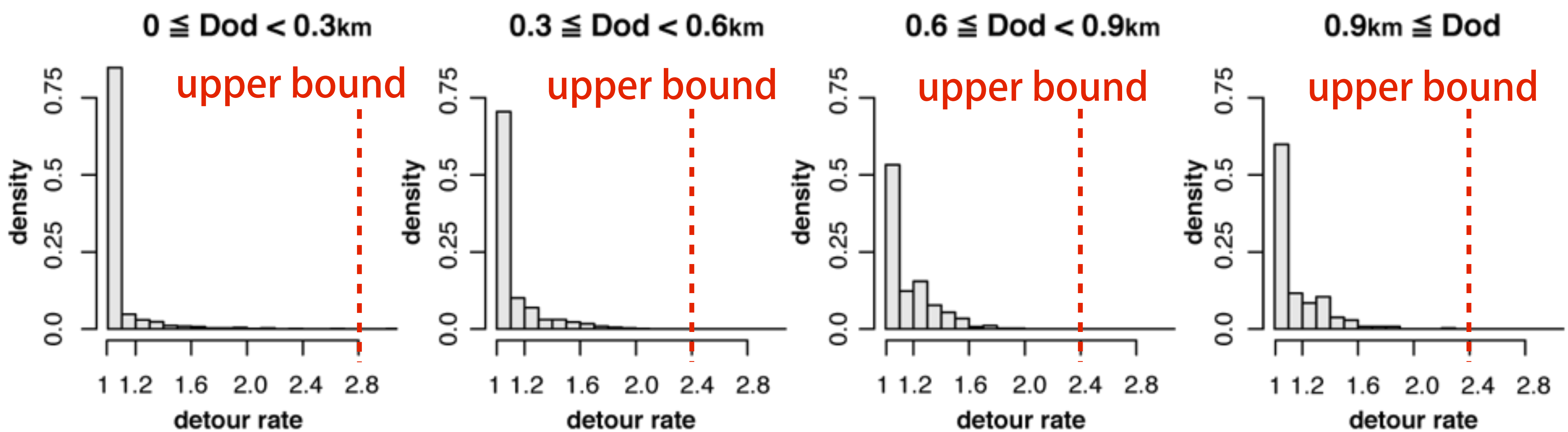
Assume **the success parameter is function of** $\boldsymbol{D_{od}}$ (the distance between the origin and the destination)

$$f_2(x, q) = q^x (1-q)^{1-x}, \quad x \in \{0, 1\}$$

$$q = 1/(1 + \exp(0.01058 + 0.34464 * D_{od}))$$

8 Data distribution

③ detour rate



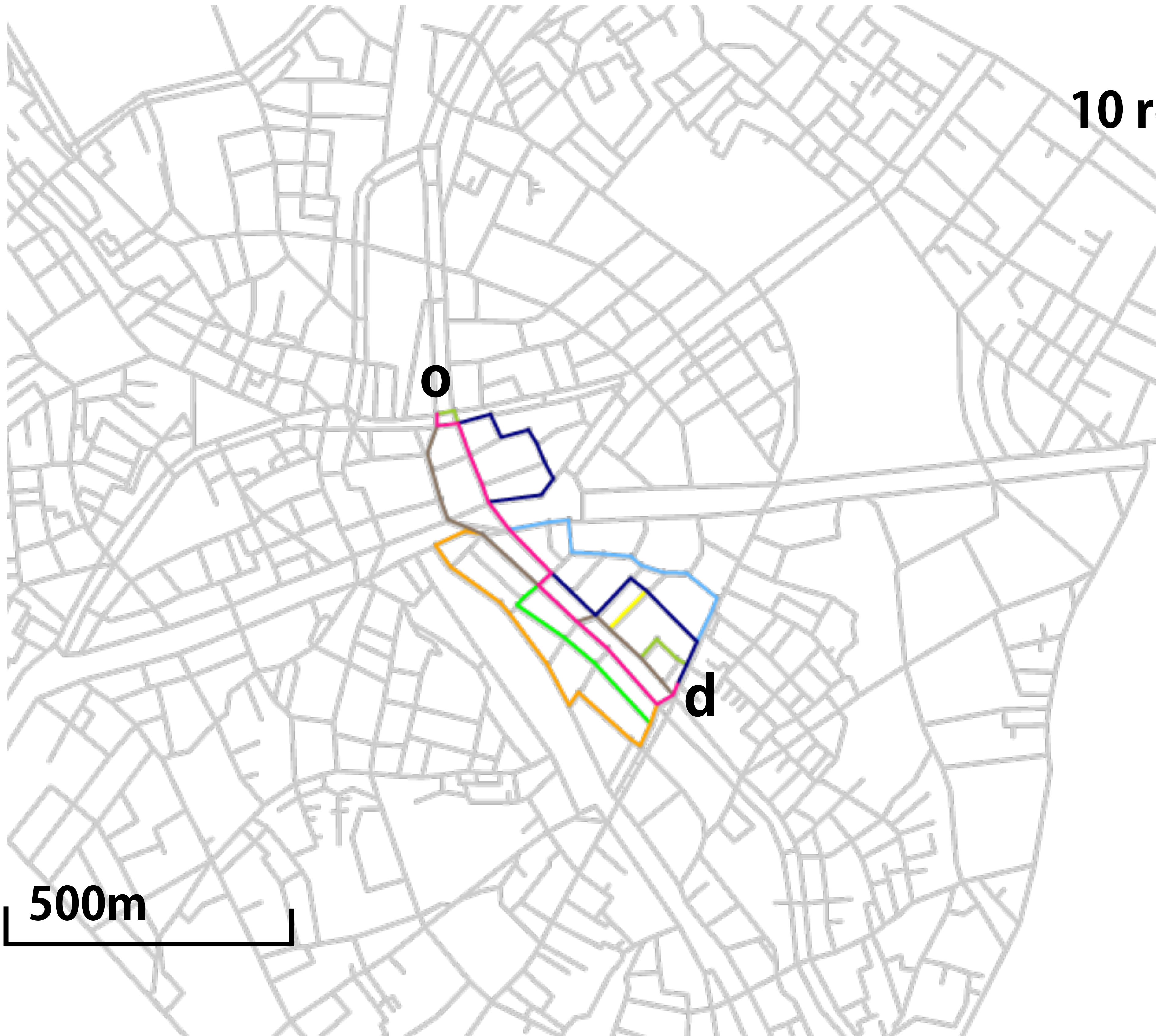
Assume **the exponential distribution** and estimate the parameters

Assume **the rate parameter is proportional to *D_{od}*** (the distance between the origin and the destination)

$$f_3(c_{y_i}, \lambda_3) = \lambda_3 e^{-\lambda_3 (c_{y_i} - 1)}$$

$$\lambda_3 = 10.958346 - 4.548481 * D_{od}$$

8 Generated choice set



10 route samples

9 Conclusion

The data oriented method for pedestrian route choice set generation is proposed;
method of link sampling according to the data distribution of branching rules.

This approach is able to be applied for individual choice set generation which reflects the characteristics of individual behavior, when we have enough amount of individual trace data.

Thank you for your attention