The aim of the multicriteria analysis method, area of use, data transformation, substance of data grouping methods

Aim: evaluation of objects and alternatives/versions (situation analysis),
facilitating decision making (decision-preparation, support),
evaluation of the impact of measures and decisions (impact assessment (IA); before-after studies)

1. Data request: measurement or statistical data

objects (e.g.: transportation companies, defined territorial units, one period)
various attributes
(must not be narrative features, only numerical indicators; e.g.: statistical, business data)

attributes and objects depend on tasks; independent attributes are more appropriate

2. Data transformation:
use of indicators transferred from their scale/magnitude and system unit

3. Aggregation:
all objects are characterized as a single attribute
e.g.: aggregate attribute calculated as (weighted) linear combination of $X_1 \dots X_p$ variables

4. Classification:
creating object groups (clusters) considering *1* or *p* attributes
(classification in 1 or p dimensional space)

1. Data request:

Objects in the rows

Analysed attributes in the columns (variables)

Object \ Attribute	X_1	…	X_j	…	X_p
O_1		…	X_{1j}	…	X_{1p}
⋮	⋮		⋮		⋮
O_i		…	X_{ij}	…	X_{ip}
⋮	⋮		⋮		⋮
O_n		…	X_{nj}	…	X_{np}
Average			$\overline{X}_j$		
Deviation			$S(X)_j$		
Minimum			X_{minj}		
Maximum			X_{maxj}		

Determination of the independence of attributes or the strengthness of relationships:
R correlation matrix

$$\mathbf{R} = (r_{k\ell})_{pp}$$

The matrix is quadratic:

$$r_{kk} = 1 \text{ (autocorrelation)}$$
$$r_{k\ell} = r_{\ell k} \text{ (symmetry)}$$

2. Data transformation:

z transformation (usually depends on temporality)
v transformation (usually depends on spatiality)

$$z_{ij} = \frac{x_{ij} - \bar{x}_i}{s(x_j)} \qquad \bar{z}_j = 0 \qquad s(z_j) = 1$$

both positive and negavtive values may occur

$$v_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \qquad 0 \le v_{ij} \le 1 \qquad \min(v_j) = 0 \qquad \max(v_j) = 1$$

Each transformation is resulted in loss of original information content but huge gain is the comparability and the aggregability.

Matrix elements after the transformation: $X_{t,ij}$

3. Aggregation:

$$_t x_i = \sum_j g_j * {}_t x_{ij} \qquad \sum_j g_j = 1 \qquad 0 \le g_j \le 1$$

During selection of weight coefficients several aspects are to be taken into account (e.g.: concept of transport policy)

4. Classification:

essence: common attributes can be assigned to the objects which belong to the same class (e.g.: quality parameters)

take *1* attribute into account – *1* dimensional space

(simplification due to the aggregation)

take *2* attributes into account – *2* dimensional space

…

take *n* attributes into account – *n* dimensional space

(„calculating distances” – sets of points – size of sets of points)

I. Calculate distance between objects

sort O_i objects and ${}_tx_i$ values in descending order

re-indexing the objects in the new order, then: for every O_i and O_{i+1} objects, ${}_tx_i \geq {}_tx_{i+1}$

$d_{i,i+1} = {}_tx_i - {}_tx_{i+1} \geq 0$ a „distance vector”

the advantage of creating distances is the easier classification *(„where the distance is high a new cluster is necessary”)*

$$d_{k,l} = \sum_{j} ({}_t x_{kj} - {}_t x_{lj})^2$$

Distance of **k** and **l** object without aggregation,

Quadratic Euclidean distance, can be represented in matrix

II. Creating object groups

objects which are „close” to each other are assigned into same group

homogeneity: the attributes of objects within a group are similar
heterogeneity: the attributes of objects among groups are different

grouping efficacy:
- calculate the distance between each object and cluster centre
- calculate the distance between cluster centres

Example: analysis, comparison and classification of vehicle types
attributes: fuel consumption, aesthetic features, purchase price, etc.

Monitoring temporal changes of attributes: **time cluster**

Time slices/intervals are the objects; statistics and transportation related attributes are assigned to the time slices/intervals (e.g.: years).

An example:

X_1 : population	[million inhabitants]
X_2 : national income	[billion HUF]
X_3 : freight transport volume	[billion t]
X_4 : freight transport performance	[billion tkm / year]
X_5 : number of transported persons	[billion persons]
X_6 : passenger transport performance	[billion pkm / year]
X_7 : passenger rolling stock	[million vehicle]

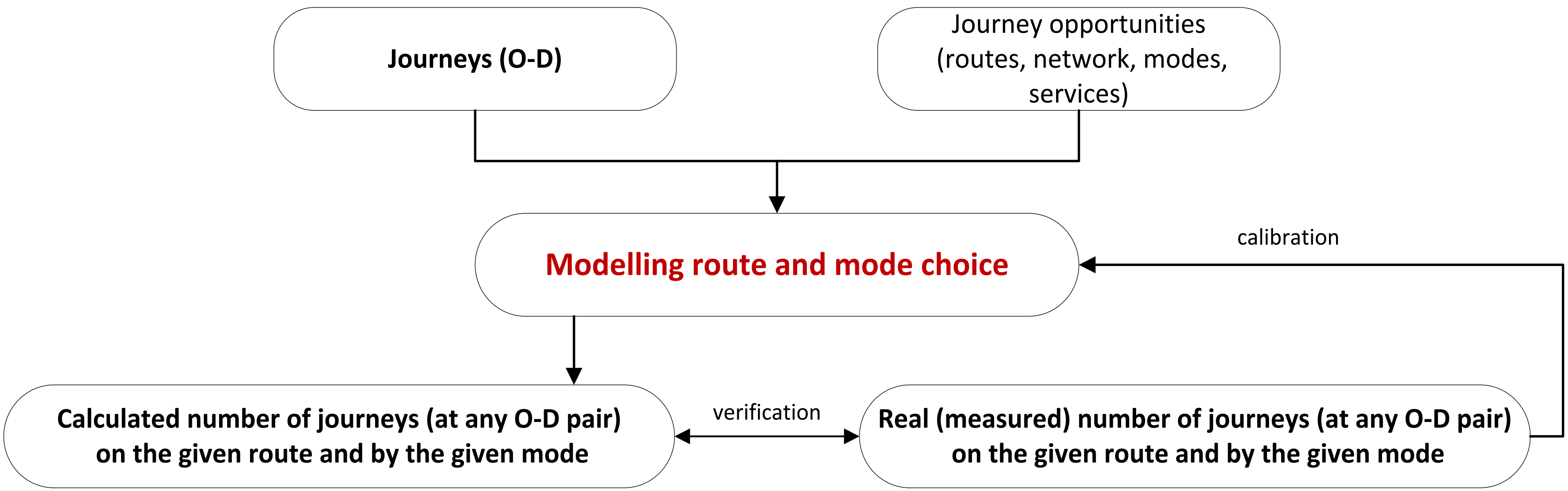
Route and mode choice modeling

Aim: determine choice probability of route alternatives (travel/journey chains).
The impact of measures can be estimated/forecasted.

Journeys in a given (O-D) relation are analysed.

I. using logit model

II. using resistances



Route and mode choice model

I. Description the transport mode and route selection using **logit model**.

The distribution among the travel chains is determined by the **quality of alternatives.**

A_1-A_n: alternatives

E_1-E_m: appraisable attributes - (e.g.: travel time, cost, distance, energy consumption)

X_{ij} : *j.* appraisable attribute at *i.* alternative

Attributes / Alternatives	E_1	E_2	E_3	E_4	…	m
A_1 (on foot)	X_{11}	X_{12}	X_{13}	X_{14}	…	⋮
A_2 (bicycle)	X_{21}	X_{22}	X_{23}	X_{24}	…	⋮
A_3 (passenger car)	X_{31}	X_{32}	X_{33}	X_{34}	…	⋮
A_4 (public transport)	X_{41}	X_{42}	X_{43}	X_{44}	…	⋮
⋮	⋮	⋮	⋮	⋮	.	⋮
n	…	…	…	…	…	X_{nm}

Data matrix of Logit model

Example: E1 = travel time, then

X_{11} on foot

X_{21} travel time with bicycle

X_{31} passenger car

X_{41} public transport

N_1-N_n: utility functions assigned to the opportunities

$$N_i = \sum_{j=1}^{m} g_{ij} * X_{ij}$$

g_{ij} : weight factors

The weight factors g_{ij} are to be determined from any traffic survey records using least square method, namely

error sum = ∑(observed value – calculated value)^2 → Min.

$$P_i = \frac{e^{N_i}}{\sum_i e^{N_i}}$$

selection probability of *i*. alternative

Nested logit model

The decision-making process is hierarchical and in every step there is two options to choose: **series of binary decisions**

$$P_1 = \frac{e^{N_1}}{e^{N_1} + e^{N_2}} = \frac{1}{1 + e^{(N_2 - N_1)}} = \frac{1}{1 + e^{g_{i1}(K_2 - K_1) + g_{i2}(T_2 - T_1)}}$$

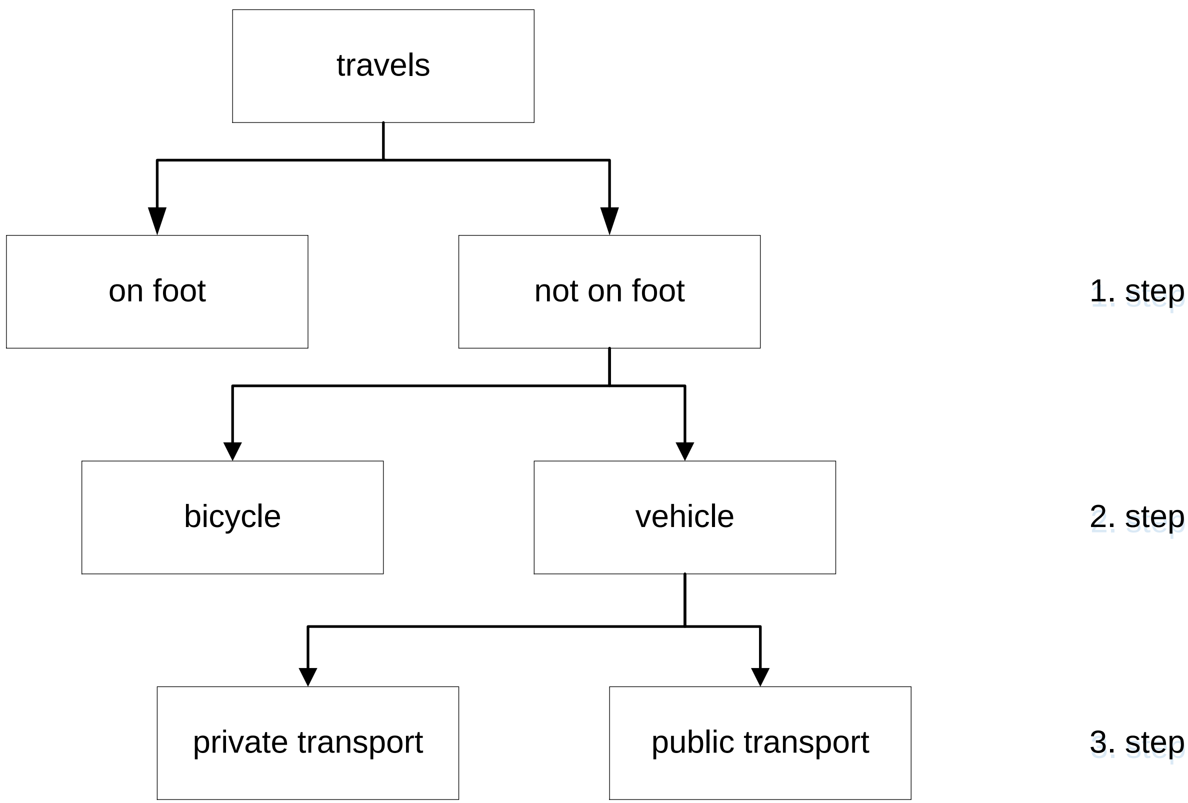
$0 < P_n < 1$ és $P_1 + P_2 = 1$

Difference of utility functions ⇒ Difference of expenses

$$N_i = g_{i1} * K_i + g_{i2} * T_i$$

The most commonly used attributes are the cost (K_i) and the time (T_i)

g_{i1}, g_{i2} estimated weight factors



Structure of the Nested logit model

The model is applied step by step

$F_m = F * P_m$

F_m = number of travels at *m* route/mode

F = number of travels

P_m = selection probability of *m* route/mode

At each step $P_1 + P_2 = 1$ (binary).

In case of *k* step the steps derived from each other:

$F^k = F^{k-1} - F_m^{k-1}$

k = number of steps (2,3)

F^k = remained number of travels in the *k*. step *(the realised travels are subtracted)*

F^{k-1} = number of travels in the *(k-1).* step

F_m^{k-1} = number of travels in the *(k-1). step* at *m* route/mode (1,2)

II. Description the transport mode and route selection using **resistances**.

Resistance function: mapping the expenditures and effects of the travel (e.g.: travel time, walking distance, cost, environmental impacts) to a common unit (e.g.: travel time, cost)

It may contain the parameters of the passenger, a specific travel or the transportation system.

- *At private transportation e.g.: it contains the parking cost*

$$P_i = \frac{w_i^{-a}}{\sum_i w_i^{-a}}$$

P_i : selection probability of i. alternative

w_i : resistance of i alternative

a : model parameter (e.g. a=1)

$$w_i = \sum_{j=1}^{n} g_{i,j} * w_{i,j}$$

$w_{i,j}$: resistance of *j.* travel phase at *i.* alternative

n : number of the phases of *i.* alternative

g_{ij} : weight factor for *j.* travel phase at *i.* alternative

The value of the resistance can be influenced by for example the terrain or the surface of the road (e.g.: bicycle, pedestrian).

The weight factor may be different through vehicle types, passenger groups and single travellers as the model is improved.