Authors

Wainaina Simon

Cracow University of Technology
Institute of Roads and Railways
Ul. Warszawska 24
31-155 Kraków, Poland
simon@transys.wil.pk.edu.pl
Tel. 048-12-632-5360

Richter Matthias
Kirchheim Alfred
Bauhaus-University
Faculty of Civil Engineering
Coudraystrasse 13b
D-99421 Weimar, Germany
matthias.richter@bauing.uni-weimar.de
Tel. 049-3643-58-4298

The Use of Activity Chain Models to Analyse Stochastic Travel Demand

1. Introduction

Activity chain analysis consists of spatial location and activity type participation of an individual throughout 24-hr period. These chains consist of activity patterns, which can be defined as an ordered sequence of activities participated during the day by members of a household. The three-dimensional (time, space, activity type) unlike mono-dimensional (trip) approach not only better defines human movements in time / space continuum but it also allows the use of stochastic models to effectively capture the problematic activity chains which deviate from the dominant home - work - home chain. This deviation is largely due to high degree of temporal and spatial substitutability of non-work activities. It is also caused by flexibility and hence complexity of the travel options available to non-work travellers. These options typically include the frequency, destinations and time of departure, among other factors. By considering activity engagement decisions, which are made sequentially, conditioned only on the $r-1$ previous destinations in the activity chain but not on the destinations still ahead a stochastic model can be formulated to describe travel demand. The model used here treats activity participation frequency (and hence travel frequency), destination choice, travel time to destination and multi-destination of activities within a unified random stochastic framework. Since the model is based on activities, the next chapter gives an outline of characteristics of activity types.

2. Characteristics of activity types

Depending on the location of activity participation, activities can be divided into home and out-of-home activities. While the former incurs no spatial displacement throughout its duration the latter is characterised by space / time continua, which corresponds to spatial displacement during travel. Figure 1 shows how out-of-home activities differ from each other.

3. Daily activity participation decisions

Since the conceptual basis of a stochastic model can be captured by considering a potential trip maker sequential decisions, Figure 2 assumes that he/she begins the day at home after a given dwell time. This is defined as a starting after midnight (in the morning) when the individual decides whether to pursue out-of-home activity or to remain at home. If an out-of-home activity is chosen, he/she must decide whether to pursue fixed activity or flexible activity. If either the former or the latter is chosen, the individual is met by spatial constraints which he/she must overcome (through mode choice) in order to participate in activity. Having arrived at some destination by a given mode, the decision on time to spend in the activity will depend on its temporal flexibility. After participation of first activity (either fixed or flexible),

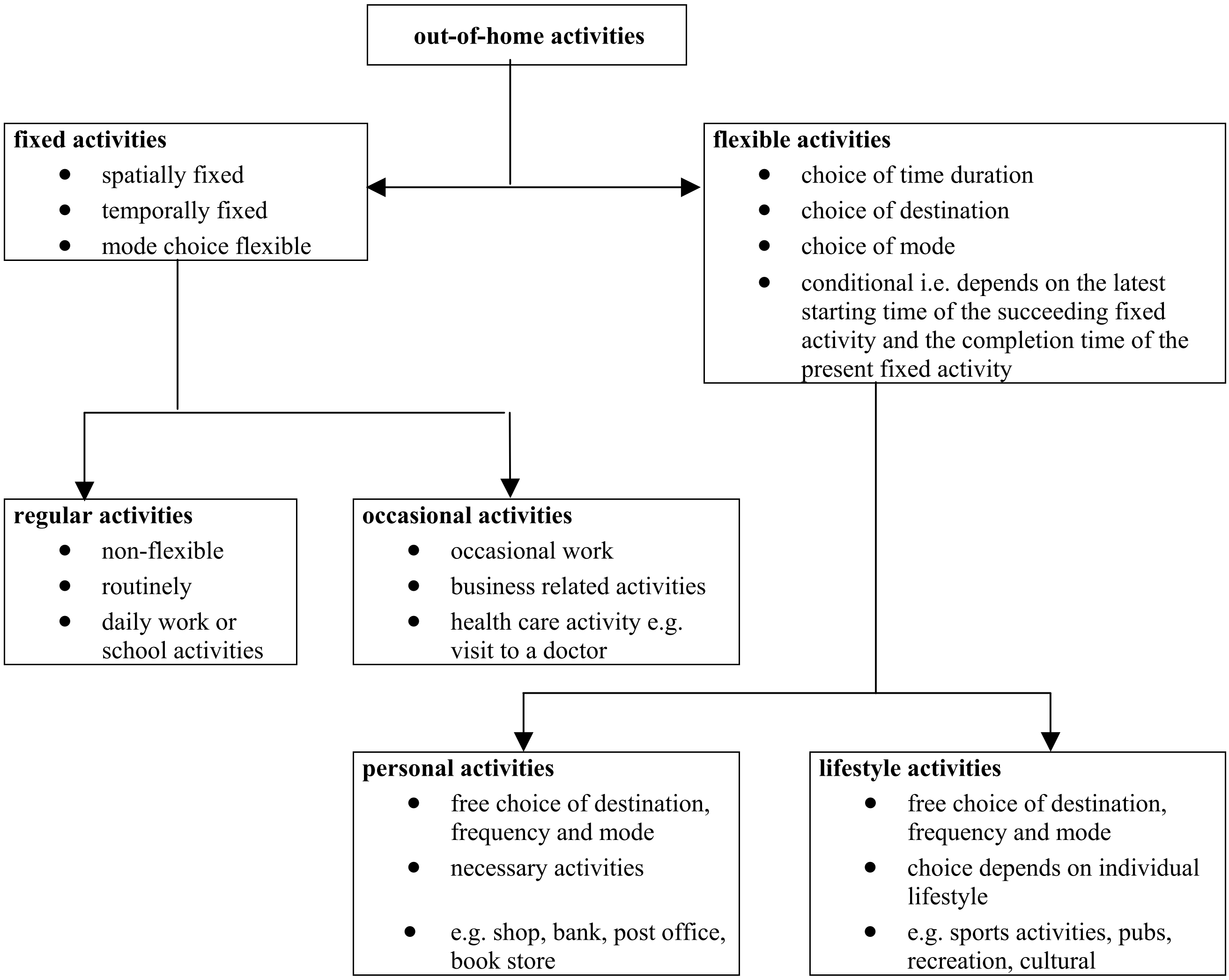


Figure 1: characteristics of activity types

an individual can decide to participate in another out-of-home activity (through mode - destination choice) or to go back home (mode choice). If an individual returns home and decides to leave again, the entire process begins once again with exception; that he/she had already completed one cycle.

4. Representation and characterisation the data of survey

Households travel survey questionnaire delivers data about travel behaviour of the members of the household. Such models can be found in [3, 4 and 6]. The KBR (comprehensive travel survey) of Katowice / Siemianowice Sląskie cities was administered through personal interview and generated among other survey data, the data of out-of-home activities participated by members of the household. The survey covered 3074 households (6399 individuals) in Katowice and 1357 households (2654 individuals) in Siemianowice Sląskie representing 2% and 4% of the population in the respective cities.
From this data, trip chains were computed and by use of VISEM their probabilities were estimated. The probabilities of the chains conducted by several homogeneous behavioural groups were used by [3] to estimate the number of trips generated in a given zone and therefore trip distribution between the zones in the surveyed region. In this paper, characteristics of travel demand i.e. the number of daily out-of-home activities, conditional transition probabilities and the distribution of departure times are examined with the help of

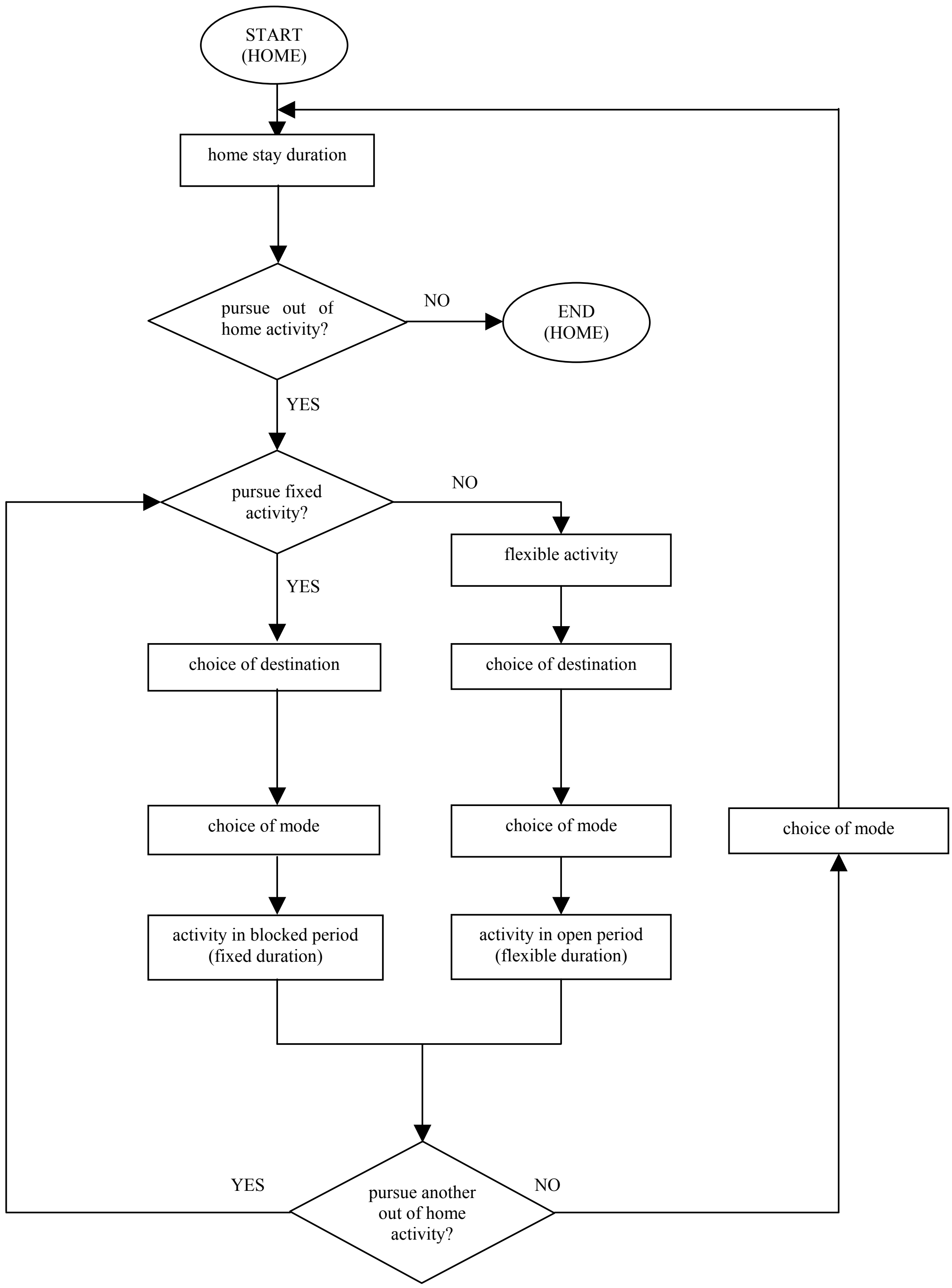


Figure 2: a framework of activity participation decisions

the original survey data. To model the daily travel demand from activity chains, the locations of activities are denoted by an universal set of possible activities which individual can participate in during the day,

$$A := \{a_1, a_2, \ldots, a_m\}.$$

The set A is termed as *activity space*. From the survey data, the number of distinguished states is equivalent to $m = 8$, that is;

a_1:	D	-	Home	a_5:	O	-	Official activities
a_2:	P	-	Work	a_6:	R	-	Private activities
a_3:	S	-	School	a_7:	Z	-	Retail shopping
a_4:	U	-	University	a_8:	H	-	Wholesale shopping

The set of activities available to an individual is considered to be finite and many topologically distinct activity patterns are likely to be similar for homogeneous behavioural groups. Every individual who tends to satisfy her/his needs and desires out of home during the day takes part in an activity chain. For a given individual let $A^{(r)}$ denote the state which is ordered a the r^{th} position of the chain of activities, $A^{(r)} \in A$, $r = 0,1,\ldots,l$. Therefore, l is the number of all activities participated during the day (including home stays, with the exception of the first home stay). The length l can also be termed as the *total- or cross-length* of the corresponding activity chain and it corresponds to the number of trips, which the individual executes during the day and hence it is an important variable to describe travel demand. Since every activity chain is assumed to start and end at home, it then means that for each individual holds $A^{(0)} = a_1$ and $A^{(l)} = a_1$.

The value of l depends essentially on the *number of out-of-home activities* participated by individual during the day, which is denoted by n. In addition, let k denote the number of *cycles*, whereby a cycle is defined as a sequence of activities beginning and ending at home (between these home stays there are assumed to be only out-of-home activities). An activity chain can be composed of one or more cycles. In the case of $k = n = 0$, it implies that individual did not leave home during the day.

For example:

Consider a typical activity chain in which an individual leaves home to go to work, comes back home and leaves once again for private activities. The activity chain of interest in this case is D-P-D-R-D. The values of l, n and k are $l = 4, n = 2$ and $k = 2$, respectively.

For an individual randomly chosen from a homogeneous behavioural group, the number of out-of-home activities, the number of cycles and the gross-length of the activity chain have to be modelled as random variables, denoted by N, K and L, respectively.

To describe transitions between several states, A is used as the state-space of a stochastic point process $A^{(r)}$ (the state which is ordered at the r^{th} position in the activity chain). Transition probabilities can be described by the conditional probabilities of equation (1),

$$\mathbf{P}\left(A^{(r+1)} = a_{i_{r+1}} | A^{(r)} = a_{i_r}, \ldots, A^{(1)} = a_{i_1}\right), \quad i_1,\ldots,i_r,i_{r+1} = 1,\ldots,m\,, \quad r = 1,\ldots,L-1\,, \quad (1)$$

and the initial probabilities

$$\mathbf{P}\left(A^{(1)} = a_{i_1}\right), \quad i_1 = 2,\ldots,m\,. \quad (2)$$

(Note that it holds $\mathbf{P}\left(A^{(0)} = a_1\right) = 1$, i.e. $\mathbf{P}\left(A^{(0)} = a_{i_0}\right) = 0$ for $i_0 = 2,\ldots,m$.)

Because of the complexity, it is hard to work with this general model, particularly with regard to the estimation of the transition probabilities by use of statistical data. Therefore assumptions for simplification have to be introduced.

Finally, the distribution of departure times can be considered. In [5] it was observed, that there are three kinds of different time distributions, firstly the distribution of the time of first departure from home since an initial time (which is here defined as midnight), denoted by $T^{(0)}$, secondly the distribution of non-home dwell time and finally the distribution of dwell time at home after returning from some trip.

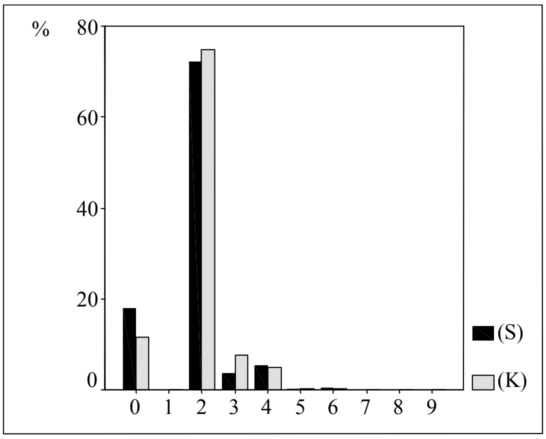
From the above considerations it is clear, that "home" should be treated as a special state. For instance the sequence D-D can not occur. In [1] and [2] another notation is used, whereby random variables R_i, $i = 1,\ldots,n$, are considered, which have the value 1 if the individual returns home after the i^{th} out-of-home activity and otherwise 0. In that case, the spatial activity chain for an individual is completely described by the values of R_i and the sequence of out-of-home activities $\widetilde{A}^{(i)} \in \{a_2, a_3, \ldots, a_m\}, i = 1, \ldots, n.$

5. Number of out-of-home activities and gross-length of activity chains

Using the notations introduced in Section 4 the following relationship between the considered random variables L, N and K is derived,

$$\mathbf{P}(L = N + K) = 1. \tag{3}$$

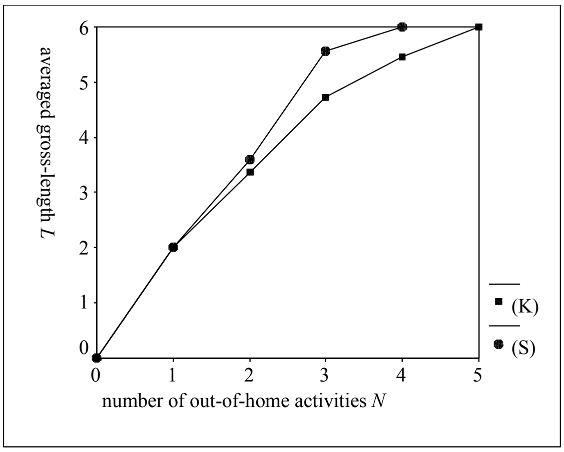
Figure 3 shows the empirical frequencies of the variable L of Siemianowice Sląskie (S) and Katowice (K) inhabitants. In both cases the dominating value of L is $L = 2$, which means that the corresponding individuals take part in only one out-of-home activity during the day.



L	(S)	(K)
0	0.1804	0.1169
1	----	----
2	0.7206	0.7491
3	0.0373	0.0774
4	0.0539	0.0493
5	0.0019	0.0036
6	0.0054	0.0034
7	0.0004	0.0002
8	----	----
9	----	0.0002

Figure 3: gross-length L of the observed activity chains

The stochastic relationship between the variables L, N and K can be examined in more details. Figure 4 shows the averaged gross-length L of the activity chain in dependence on the number N of daily out-of-home activities. A declining increase can be observed. The reason for this empirical observation is the tendency of the individuals to chain up their out-of-home activities and this behaviour is the base for the considerations of this paper. By going directly from one out-of-home activity to another without returning home, the number of trips per activity decreases. This tendency is visible particularly for the Katowice survey data, which is especially suited because of the large number of cases.



	Trips per out-of-home activity L/N	
Number of out-of-home activities N	(S)	(K)
1	2.00	2.00
2	1.80	1.68
3	1.85	1.57
4	1.50	1.36
5	----	1.20

Figure 4: averaged values of L and L/N in dependence on N

Additionally, Figure 5 shows the results for the empirical distribution of the number of out-of-home activities. Because of the low number of cases with $N \geq 5$, only the observations with $N \leq 4$ are shown. The results are split according to homogeneous behavioural groups. The groups considered here are; unemployed people (ue), primary school pupils (pr), full time secondary school students (se), full time post secondary and university students (un) and employed people (em) for both Katowice (K) and Siemianowice Sląskie (S) towns. A remarkable difference between the empirical distributions of N with regard to unemployed people (who tend to have less out-of-home activities than other behavioural groups) is observed.

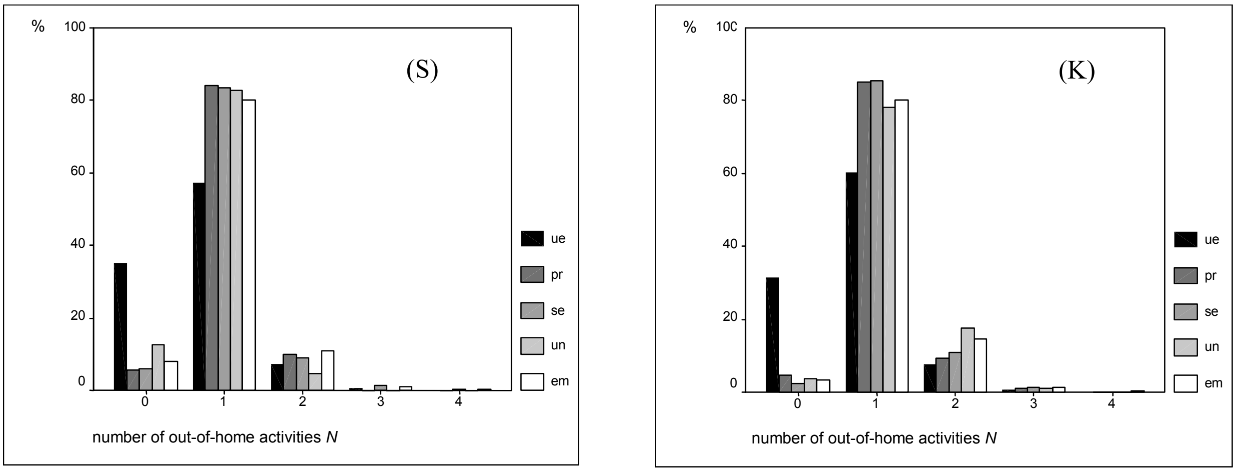


Figure 5: empirical distribution of N

The empirical results of Figure 5 suggest the suitability of Poisson-distributions to model the stochastic behaviour of N split by behavioural groups (see also [2]). In such a case, the distribution function would be described by $\mathbf{P}(N = n) = \frac{\lambda^n \mathbf{e}^{-\lambda}}{n!}, \quad n \geq 0.$

For the expectation of N, it then holds that $\mathbf{E}\, N = \lambda.$ The parameter λ should be chosen separately for every behavioural group. Furthermore it would be possible to describe the influence of other individual factors $x = (x_1, x_2, x_3, \ldots)$, like sex, age, availability of a car, for instance by use of $\lambda(x) = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots)$, where the significance of the several factors reflects the parameters $\beta_1, \beta_2, \beta_3, \ldots$ However, the assumption of Poisson distribution for N can not be kept. This phenomenon was also observed by [2]. Figure 6 shows the estimated expectations and variances of N for the considered homogeneous behavioural groups mentioned above.

	ue		pr		se		un		em	
	(S)	(K)	(S)	(K)	(S)	(K)	(S)	(K)	(S)	(K)
$\bar{x}$	0.73	0.77	1.04	1.07	1.07	1.11	0.92	1.16	1.06	1.16
s^2	0.37	0.37	0.16	0.17	0.25	0.17	0.17	0.22	0.25	0.26

Figure 6: estimated expectations and variances of N

In case of Poisson distribution, expectation and variance should be equal, but from empirical data a clear underdispersion is observed. The results of corresponding Kolmogorov-Smirnov tests for Poisson distribution confirm this. The reason for this behaviour is on the one hand due to the dominance of the case $N = 1$, and on the other hand the Poisson model presumes a time-constant rate of activities during the day while this assumption is not fulfilled in reality. However, Figure 6 shows a remarkable correspondence between the results for the estimated expectations and variances of N which were kept in Katowice and Siemianowice Śląskie surveys. This shows that the inhabitants of both towns have a similar behaviour with respect to the number of daily out-of-home activities. Finally, one can refer to [1], where models for the description of the conditional probability distribution of the cross-length L with respect to N are considered.

6. Transition probabilities

In this chapter, only individuals, who leave home at least once per day are considered. According to the probability $\mathbf{P}(N > 0)$ these are 81.96% of the queried persons of Siemianowice Śląskie (S) and 88.31% of those of Katowice (K), respectively.
In order to analyse transition probabilities, the initial probabilities $\mathbf{P}(A^{(1)} = a_i)$, $i = 2,\ldots,8$, are considered first. Figure 7 shows these probabilities for the whole population of (S) and (K) (on the left hand side) as well as split by the considered homogeneous behavioural groups (only Katowice survey data in this case). A very natural dependence on the behavioural groups is observable.

		$\mathbf{P}(A^{(1)} = a_i)$	
i	state	(S)	(K)
2	P	0.4034	0.4551
3	S	0.1870	0.1724
4	U	0.0223	0.0836
5	O	0.0232	0.0165
6	R	0.2003	0.1505
7	Z	0.1600	0.1185
8	H	0.0038	0.0033

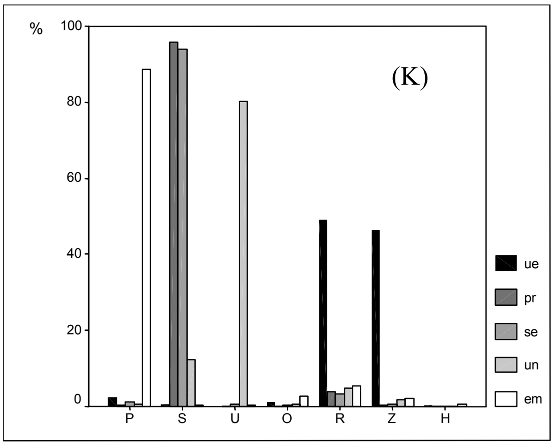


Figure 7: initial probabilities $\mathbf{P}(A^{(1)} = a_i)$

To study travel demand for the whole population, transition from $A^{(1)}$ to $A^{(2)}$ is considered. Thereby, the considerations are restricted again to the Katowice survey data. Figure 8 shows the estimated transition probabilities $\mathbf{P}(A^{(2)} = a_j | A^{(1)} = a_i)$ for $i, j = 1,\ldots,8$. As mentioned before, "home" is a special state and the case $A^{(1)} = a_1$ can not occur (marked by the symbol $*$ in the table). From the results of Figures 7 and 8 the probability, that an individual returns immediately home after first out-of-home activity $\mathbf{P}(A^{(2)} = a_1)$ can be computed by 90.24%.

$A^{(1)}$ \ $A^{(2)}$	D	P	S	U	O	R	Z	H	total
D	*	*	*	*	*	*	*	*	*
P	0.8970	0.0044	0.0012	0.0008	0.0121	0.0271	0.0553	0.0020	1
S	0.9424	0.0011	0.0011	0	0.0011	0.0405	0.0139	0	1
U	0.8791	0.0044	0	0.0110	0	0.0549	0.0484	0.0022	1
O	0.7444	0.0778	0	0	0.0444	0.0333	0.0889	0.0111	1
R	0.8694	0.0366	0.0073	0.0037	0.0024	0.0232	0.0574	0	1
Z	0.9519	0	0.0016	0	0	0.0388	0.0078	0	1
H	0.6667	0.1667	0	0	0.0556	0.0556	0	0.0556	1

Figure 8: transition probabilities $\mathbf{P}(A^{(2)} = a_j | A^{(1)} = a_i)$

In case $r \geq 2$ the following situation arises. The probability, that a transition between $A^{(r)}$ and $A^{(r+1)}$ occurs at all corresponds to the probability $\mathbf{P}(L \geq r+1)$. Under the condition, that such an transition takes place, we have to analyse the transition probabilities $\mathbf{P}\left(A^{(r+1)} = a_j \mid A^{(r)} = a_i\right)$for $i, j = 1,\ldots,8$. Figure 9 shows the corresponding estimated probabilities in case of the transition between $A^{(2)}$ and $A^{(3)}$ (i. e. $r = 2$).

$A^{(2)}$ \ $A^{(3)}$	D	P	S	U	O	R	Z	H	total
D	0	0.0441	0.0271	0.0136	0.0203	0.5661	0.3186	0.0102	1
P	0.7963	0.0185	0	0	0.0185	0.1481	0.0185	0	1
S	0.9091	0	0.0909	0	0	0	0	0	1
U	0.8000	0	0	0	0	0.1000	0.1000	0	1
O	0.6579	0.1842	0	0	0.1316	0.0263	0	0	1
R	0.9494	0	0	0	0	0.0393	0.0112	0	1
Z	0.9655	0.0043	0	0	0	0.0172	0.0129	0	1
H	0.6250	0.1250	0	0	0	0.1250	0	0.1250	1

Figure 9: transition probabilities $\mathbf{P}\left(A^{(3)} = a_j \mid A^{(2)} = a_i\right)$

A remarkable difference to the transition between $A^{(1)}$ and $A^{(2)}$ is visible, which means, that the corresponding state process is not homogeneously. Similar estimations for $r \geq 3$ are possible. Nevertheless, it should be mentioned, that the estimation of the corresponding probabilities becomes more and more sophisticated because of the decreasing number of observations, this problem arises even in case $r = 2$.

The above considerations are only of descriptive nature. Because the probabilities $\mathbf{P}\left(A^{(r)} = a_{i_r}, A^{(r-1)} = a_{i_{r-1}}, \ldots, A^{(1)} = a_{i_1} \mid L = r\right)$ are of a special interest, it has to be analysed, whether (under the condition $L = r$) the equation

$$\mathbf{P}\left(A^{(r)} = a_{i_r}, A^{(r-1)} = a_{i_{r-1}}, \ldots, A^{(1)} = a_{i_1}\right) = \mathbf{P}\left(A^{(1)} = a_{i_1}\right) \mathbf{P}\left(A^{(2)} = a_{i_2} \mid A^{(1)} = a_{i_1}\right) \ldots \mathbf{P}\left(A^{(r)} = a_{i_r} \mid A^{(r-1)} = a_{i_{r-1}}, \ldots, A^{(1)} = a_{i_1}\right) \quad (4)$$

by use of

$$\mathbf{P}\left(A^{(\tau)} = a_{i_\tau} \mid A^{(\tau-1)} = a_{i_{\tau-1}}, \ldots, A^{(1)} = a_{i_1}\right) = \mathbf{P}\left(A^{(\tau)} = a_{i_\tau} \mid A^{(\tau-1)} = a_{i_{\tau-1}}\right) \quad (5)$$

for $\tau = 2, \ldots, r$ (Markov property) can be reduced to

$$\mathbf{P}\left(A^{(r)} = a_{i_r}, A^{(r-1)} = a_{i_{r-1}}, \ldots, A^{(1)} = a_{i_1}\right) = \mathbf{P}\left(A^{(1)} = a_{i_1}\right) \mathbf{P}\left(A^{(2)} = a_{i_2} \mid A^{(1)} = a_{i_1}\right) \ldots \mathbf{P}\left(A^{(r)} = a_{i_r} \mid A^{(r-1)} = a_{i_{r-1}}\right). \quad (6)$$

In particular the influence of the special state “home” has to be investigated more detailed.

7. Distribution of departure times

As mentioned in Chapter 4, in [5] three kinds of distributions of the departure times are suggested. To be brief, here only the time $T^{(0)}$ of the first departure from home since midnight is considered. Figure 10 shows the empirical results for the Katowice (K) survey data, on the left hand side regarding to the whole population, on the right hand side according to the consideration of homogeneous behavioural groups exemplarity regarding to the group of full time post secondary and university students (un).

According to the considerations of [5] by inspection of the histograms and subsequent comparison of the observed and predicted values a shifted gamma distribution was fitted to

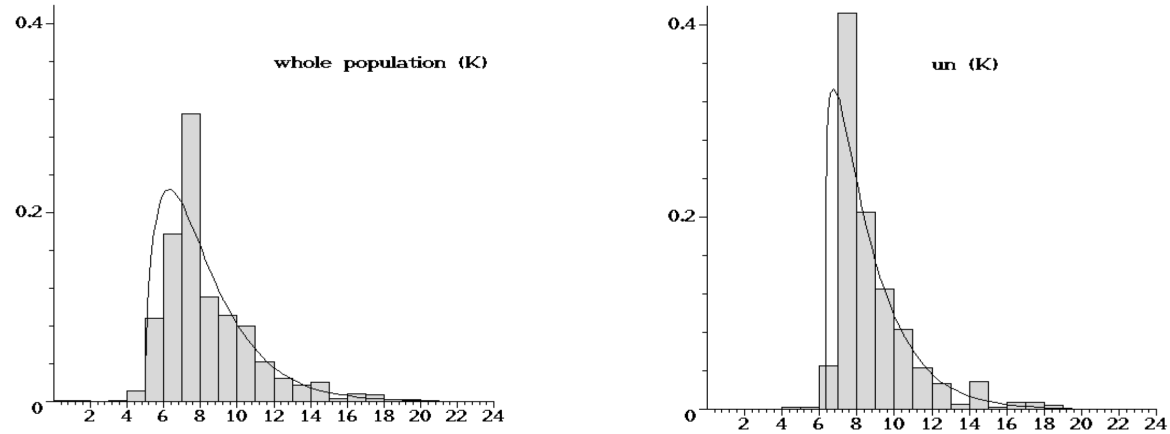


Figure 10: distribution of departure time $T^{(0)}$

the actual data. Thereby, the expectation and variance of the shifted gamma distribution and the corresponding estimated values from the actual data coincide. Therefore, the parameters (b, p) of the gamma distribution and the shift value t_s are chosen as $b = 0.0087$, $p = 1.6766$, $t_s = 300$ min (for the whole population) and $b = 0.0089$, $p = 1.2250$, $t_s = 380$ min (in case of the "un" behavioural group). Similar considerations for the other behavioural groups are possible. However, it shows, that for the time after the visible "peak" a remarkable correspondence exists, whereas the "peak" itself can not been described satisfyingly. As a consequence a χ^2-test of goodness of fit rejects the hypothesis of shifted gamma distribution. Nevertheless, for orientating considerations, this distribution seems to be quite suitable.

8. Conclusion

Various aspects in the description of travel demand, which base on activity chain models, were presented. It has been shown, that through the application of a sequential decision model the mobility of individuals can be estimated. Statistical computations regarding the travel survey of Katowice / Siemianowice Śląskie cities have been done. By using several characteristics of activity chains consequences with respect to the description of travel behaviour can be drawn, whereby the inclusion of Markov processes property can improve the model.

References

1] Hautzinger H., Aktivitätenbezogene Verkehrserzeugungsmodelle -Ein neues Konzept zur Personenverkehrs-prognose-, Zeitschrift für Verkehrswissenschaft, 53. Jg., 1982.
2] Hautzinger H., Pfeiffer, M., Gesetzmäßigkeiten des Mobilitätsverhaltens, Verkehrsmobilität in Deutschland zu Beginn der 90er Jahre, Berichte der Bundesanstalt für Straßenwesen -Mensch und Sicherheit-, Heft M57, 1996.
3] KBR KATOWICE / SIEMIANOWICE, Complex Travel Survey, 1998.
4] KONTIV '89, Activity-Based Travel Survey, organized by the Federal Ministry of Transport, Germany, 1989.
5] Lerman S. R., The Use of Disaggregate Choice Models in Semi-Markov Process Models of Trip Chaining Behavior, Transportation Science, Vol. 13, No. 4, 1979.
6] System repräsentativer Verkehrsumfragen (SrV-Plus 1991), Institut für Stadtbauwesen der TU Dresden, 1991.