EVALUATION OF A BEHAVIOR MODEL OF OCCUPANTS IN HOME BASED ON JAPANESE NATIONAL TIME USE SURVEY

Yohei Yamaguchi, Yoshiyuki Shimoda

Graduate School of Engineering, Osaka University
2-1, Yamadaoka, Suita, Osaka, 565-0871, Japan, yohei@see.eng.osaka-u.ac.jp
CREST, Japan Science and Technology Agency

ABSTRACT

This paper evaluates a stochastic model of time use of occupants in home. In this model, behaviors are divided into routine and non-routine behaviors. Routine behaviors are those undertaken routinely every day like sleeping and working. First, the duration of routine behaviors is determined and then placed on timeline of day by using probability distribution of time allocation. Then, gaps between routine behaviors are filled by non-routine behaviors. This paper evaluates the proposed model by comparing simulation result for working males and housewives with original time use data that are used to generate input data of the model.

INTRODUCTION

Existing modeling approaches

When energy demand of residential buildings is simulated, occupants' behavior is assumed by a static pattern that represents an average occupant's presence and action. This approach is easy to set up and useful to estimate total energy demand or its typical time-series pattern. However, the assumption is not a useful input to replicate energy demand at a high-temporal resolution with a realistic variability. In this context, stochastic modeling of occupants' presence and action in home has attracted attention.

There have been a number of behavior models that stochastically generate presence and action of occupants in home. There have been at least three established approaches. In the first approach, the presence and action is modelled as a Markov Chain (e.g. Richardson et al. 2010, Widén et al. 2010), in which transitions of presence and action occur randomly following transition probability from present state to another. In most basic models, transition probability only depends on the present state and the time of day. By giving a random number to transition probability, presence and action at each simulation step is sequentially decided.

Thus, an input data of Markov chain models is transition probability from one state to another at each time of day. Transition probability is often developed based on raw data of time use survey result Most of time use surveys collect one-day diaries, from several thousands of people, showing how respondents spent their time on the day. The collected data is called time use data (TUD). If raw data of TUD is available, information on the number of respondents who performed an activity i and who changed their behavior from behavior i to j (N_{ij}) can be calculated.

The second approach is proposed by Tanimoto et al (2008). The most important feature of the approach is that the model does not use raw data of TUD. Instead, the approach only uses the following statistic information developed based on TUD:

- Mean and standard deviation of duration of activities in a day
- Percentage of respondents who adopt each of behavior (PB) at a specific time of a day.

First, the duration of all considered behaviors is determined while assuming a logarithmic Gauss distribution defined by the mean and standard deviation. The list of duration for each activity is referred to discrete behaviors. Discrete behaviors are placed on the timeline of a day by considering PB in the following manner. First, the durations are adjusted to 24 hours. Then, a time step is randomly selected. One behavior is selected by giving a random number to PB at the time step. After the first behavior is placed, the behavior beginning from the time at which the first behavior ends is selected by using PB again. This process is repeated until all of the discrete behaviors are placed. For following days, the selection of the first time step is replaced with the time step at which the last discrete behavior in the previous 24 hours ended.

The third approach is proposed by Wilke et al. (2013). In the approach, first, according to the probability distribution to select behaviors at the examined time, one behavior is randomly selected. Then, the duration of the selected behavior is determined based on the probability distribution function of the duration of the selected behavior. These two processes are repeated sequentially.

Features of existing modeling approaches

In sum, Markov chain models well replicate behavior transitions. However, in Markov chain models, duration of behaviors is not necessarily well replicated, as duration is determined as a result of behavior transition. Contrary, the Tanimoto's approach well replicates duration of behaviors. However, behavior transition is not necessarily well replicated because the number of transitions depends on the number of discrete behaviors. This might be a vital weakness to apply the behavior simulation result to model energy demand of residential buildings.

Development context

As a background of the model development, this section summarizes the context of the model development in our research team.

In our research team, energy demand is estimated for a community or an urban area for which hundreds/thousands of households are simulated. In the modeling of energy demand, a variety of simulation conditions are given on direct determinants of energy demand, such as family composition, time use characteristics, appliance ownership and specification, and geometry and layout of house, in order to reflect the variety in energy demand among households (Yamaguchi et al. 2014).

This simulation context requires our occupant behavior model to have the following two characteristics. First, the model must generate a sufficient variety in presence and action among simulated households so that a variety of energy demand can be generated among simulated houses. Second, the model must generate interactions among household members as a household is the unit of simulation. For example, household members share time for meals and spaces in their house like bath room. Our occupant behavior model was designed to satisfy these two requirements for the community/urban scale energy demand modelling.

Purpose and contents of this paper

The purpose of this paper is to propose an occupant behavior model for the community/urban scale energy demand modelling and to evaluate the proposed behavior model by comparing its simulation results with the original TUD that is used to develop input data of the proposed model.

In the remaining parts of the paper, we first introduce the proposed model and input data. The next section explains the methodology of the evaluation. Then, the result of the evaluation is presented. Finally, we discuss the strength and weakness of the proposed model.

PROPOSED BEHAVIOR MODEL

As mentioned in the introduction, there are two requirements for the proposed occupant behavior model. First, the model must generate a sufficient variety in presence and action among simulated households. Second, the model must generate interactions among household members. To satisfy these requirements, the proposed model have a feature in both of the input data preparation process and the simulation algorithm.

In order to satisfy the first requirement, we prepared several categories with different time use characteristics. In the preparation of the input data for the proposed model, we classified TUD into several categories by age, occupation, gender and some other time use characteristics. By developing input data for each category, time use reflecting characteristics of each group can be generated. This point is explained in the section for data preparation.

For the second requirement, interaction among household members is considers in the simulation algorithm for having meals and bathing. Having meals and bathing have a different nature with other behaviors, since these behaviors are undertaken routinely every day in many households. Sleeping and outing for work or school are also have such characteristics. Thus, we refer these behaviors to routine behaviors and distinguish them with the other behaviors, non-routine behaviors. The simulation algorithm to determine the occurrence of routine and non-routine behaviors is different as explained below. This is the most important feature of the proposed model in the simulation algorithm.

Simulation algorithm

Figure 1 shows the modelling procedure to generate occupants' time use in home in the proposed model. The routine behaviors, which are sleeping, outing for work or school, eating, and bathing, are placed on a day prior to the rest of behaviors, non-routine behaviors. The order of placement is sleeping, working/study, breakfast, lunch, dinner and finally bathing.

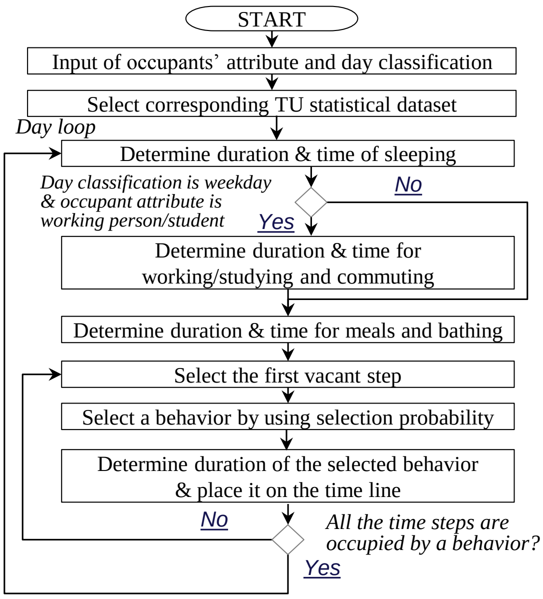

Figure 1 Procedure of the simulation model

The method to place the routine behaviors is similar to Tanimoto's model. First, the duration of the routine behaviors is determined based on a statistical information on the duration. Then, these durations of the routine behaviors are placed on timeline of day by using probability distribution of time allocation.

For sleeping, probability distribution is given for ending time, while it is given for beginning time for the other routine behaviors. This is because we observed that standard deviation of awaking time is smaller than sleep beginning time if we look at individual person.

After placing all the durations of the routine behaviors, the gaps between the routine behaviors are filled by non-routine behaviors. The method to place the routine behaviors is similar to Wilke's model. For this process, two kinds of data are used. The first data is selection probability distribution of non-routine behaviors ($n = 1$ to N), $SP_{t,n}$ at time t. At the first vacant time step after a routine behavior, a random number is given to the selection probability distribution to select a behavior after the routine behavior. Then, by using the second data, statistical data on the duration of non-routine behaviors, the duration of the selected non-routine behavior is determined. This process is repeated until all the gaps are fulfilled.

It should be noted that when the decided duration is larger than the vacant steps, the duration is shortened to fill the vacant steps.

Input data

Due to the procedure, the model uses the input data listed in Table 1. The first two data are for the routine behaviors, while the other two are for non-routine behaviors. It should be noted that any mathematic probability distribution function were not assumed for all of the input data. All the input data were simply developed based on the sample time use data.

Table 1 Input data of the proposed model

Input data	Symbol
Probability distribution of duration of a routine behavior *n*	PBD_n
Probability distribution of beginning or ending time for a routine behavior *n*	PBT_n
Percentage of people who adopt a behavior *n* at time *t*	$PB_{t,n}$
Cumulative probability at a duration *du* of non-routine behavior *n* within a time region *tr*	$Dr_{tr,n,du}$

The cumulative probability of duration, $Dr_{tr,n,du}$, is prepared for time regions, *tr*, distinguishing 24-hours of timeline. Weekdays are divided into two time regions, for people going work or school, the time region from awaking time and outing time for work or school and the time region from arriving time from work or school to sleeping time. For holidays and the other attributes, housewives and elderly male/female, the timeline is divided into four time regions by using sleeping and three meals, breakfast, lunch, and dinner. The reason why we distinguished time regions in the timeline is that the selection probability considerably differs among the time regions.

To develop the selection probability for non-routine behaviors, $SP_{t,n}$, the percentage of people who adopt the behavior *n*, $PB_{t,n}$ at time step *t* is used. Then, this probability is adjusted by the cumulative probability of duration, $Dr_{tr,n,du}$, as shown in Equation (1).

$$APB_{t,n} = \frac{\sum_{du}^{VS} PB_{t+du,n} \cdot Dr_{tr,n,du}}{VS} \tag{1}$$

The reason why $APB_{t,n}$ is adjusted is that the appropriateness of behaviors differs accordingly to the number of available vacant steps (VS) created by routine behaviors and the non-routine behavior lastly placed. Each behavior has a certain characteristic in duration defined by $Dr_{tr,n,du}$. If the number of vacant step is few, probability of selection of a behavior with relatively long duration should be small. To take into account this point, we defined $APB_{t,n}$ as in Equation (1). By using $APB_{t,n}$, $SP_{t,n}$ is defined as in Equation (2).

$$SP_{t,n} = APB_{t,n} \Big/ \sum_{i}^{N} APB_{t,i} \tag{2}$$

Due to this adjustment, $SP_{t,n}$ for a behavior with relatively long duration becomes small when the number of vacant step is few. Contrary, $SP_{t,n}$ is almost equal to $PB_{t,n}$ when there are vacant steps large enough to the duration of the behavior.

Data preparation

As shown in Table 1, the model has four input data. All of the data are developed based on the result of the national time-use survey conducted by the Statistics Bureau Japan (2014) in 2006. This survey collected time use diary from approximately 80 thousands households living in the whole country. All persons aged 10 and over in the sample households are asked to respond to the survey.

The survey used two format as time use diary, namely Diary A and B. Diary A has 20 classifications of time use. The number of sample is very large. For example, the number of sample collected from working male and female on weekdays is 27.7 and 16.0 thousands. Problem of Diary A in application to the occupant behavior model is that the classification of time use is too rough, though all the routine behaviors are covered in Diary A. Contrary, Survey B classifies time use into 85 kinds that is sufficiently detail to model behaviors with energy consumption in home. For example, household maintenance activities are divided into 14 kinds. However, a drawback of Diary B is that the number of sample is relatively small. For example, it is 1.5 thousands for working male and 1.3 thousands for working female on weekdays.

Based on the feature of Diary A and B, Diary A is used to develop the input data for the routine behaviors, while Diary B is used to develop those for non-routine behaviors.

As mentioned in the introduction, the proposed model is applied to generate data of presence and action for hundreds/thousands of people. Thus, the model must generate a variety of time use for simulated households so that a variety of energy demand can be generated among simulated households. In order to satisfy this requirement, the authors classify time use diary into 25 categories as listed in Table 2. Except working male, female and housewife, the classification was made based on gender, age, and occupation. For working male and female, a cluster analysis was conducted on the duration, beginning and ending times of the routine behaviors, which made three and five groups for working male and female respectively reflecting difference in working hours. For housewives, the need of parental care and the age of youngest children is considered. This is because time use for parental care vary significantly according to these factors.

Table 2
Classification of occupant attributes

Classification	Description
Working male	Male with a job aged from 20 to 65 with three sub-categories by working time: 1) full time worker with morning and afternoon working time, 2) long working time worker with longer working time than 1), 3) worker with afternoon and night working time.
Working female	Female with a job aged from 20 to 65 with five sub-categories: same categories as working male 1) to 3), 4) morning and 5) afternoon part time worker
Housewife	Female without a job aged from 20 to 65 with five sub-categories by age and parental care: 1) those without children younger than 45, 2) 45 and older, 3) those with children whose youngest is preschool, 4) primary to high school student, and 5) aged 18 and older
Student male and female	Primary, junior high, high school, university and collage male and female student
Elderly male and female	Male or female older than 65 with two categories for those who live alone and with family

Interaction among household members

As mentioned above, for having meals and bathing, interaction among household members is considered.

For having meals, Baptista et al. (2014) proposed a method to generate sharing of time for meals among household members, in which a male, female and their children are assumed and the female is assumed to decide the time for having meals for household members. Following the method, after durations of having meals are placed for all household members, the following adjustment is conducted. After placing time for having meals for each household member, a household member is selected who decide the time for each of meal from the person who present in home. When the decided person has the meal, the other household members in home begin to have meal together. The duration is given by those that have already been determined.

After deciding the time for meals, bathing time is determined for each household member respectively by following the method explained above. The bathing time can overlap for some of household members. In the case, the bating time of the member whose beginning time is later is shifted. For shifting the bating time, the earliest vacant time steps is searched that is larger enough for the pre-determined duration. Thus, by this process, the availability of bathing is considered as an interaction among household members.

EVALUATION METHODOLOGY

This paper evaluates the proposed model by comparing the statistical information on time use between the simulation result and the original TUD that is used to develop the input data for the model. As mentioned earlier, the Markov Chain model well replicates behavior transitions, while the Tanimoto's roulette selection model does the duration of each behavior. Based on this understanding, we evaluated the simulation result by using the following indicators:

- Time at which the routine behaviors start and end,
- Duration time of routine and non-routine behaviors,
- Percentage of simulated days on which each behavior is undertaken at each time of day,
- Number of behavior transitions per day, and
- Selection probability of behaviors from other behaviors

To calculate these indicators, we performed 2000 simulation runs for weekdays. In this paper, we only show the comparison result conducted for working male, a full time worker with morning and afternoon working time (category No. 1 in Table 2), and housewife with children whose youngest is a primary to high school student (category No. 4 in Table 2). The number of the samples in the Diary B is 1315 and 81. The number of the sample of the housewife category is very small.

RESULTS

Time at which the routine behaviors start and end

Figure 2 shows the cumulative probability of the sleeping beginning time and the ending time of the other routine behaviors of the working male. The bold lines show those of simulation results and thin dash lines show those of original TUD. The figure

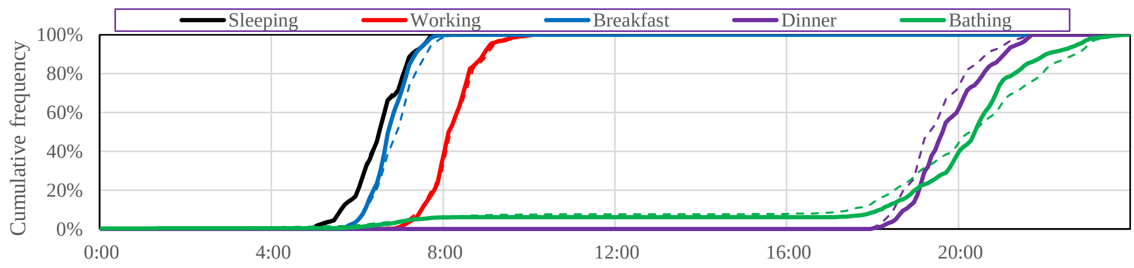

Figure 2. Cumulative frequency of sleeping ending time and beginning time of the other routine behaviors

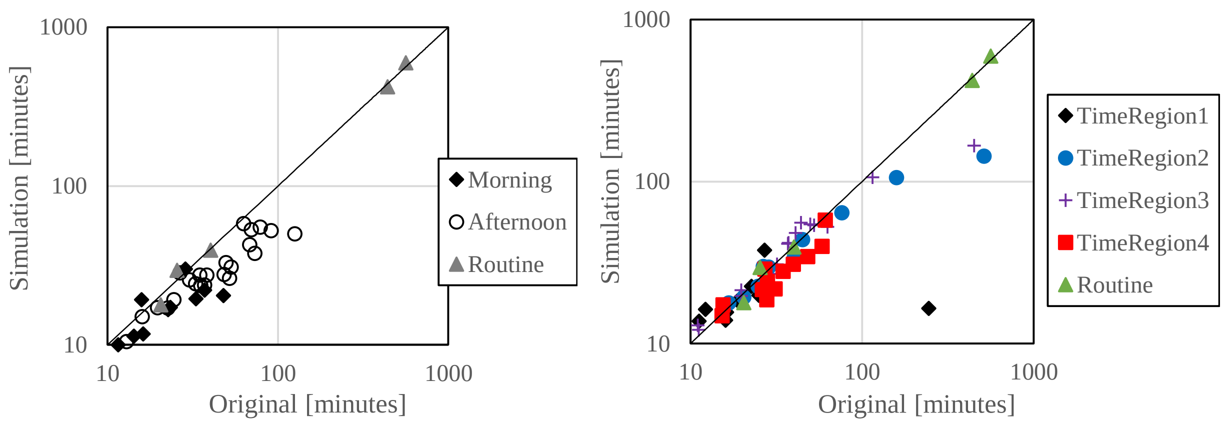

Figure 3. *Mean duration of routine and non-routine behaviors*

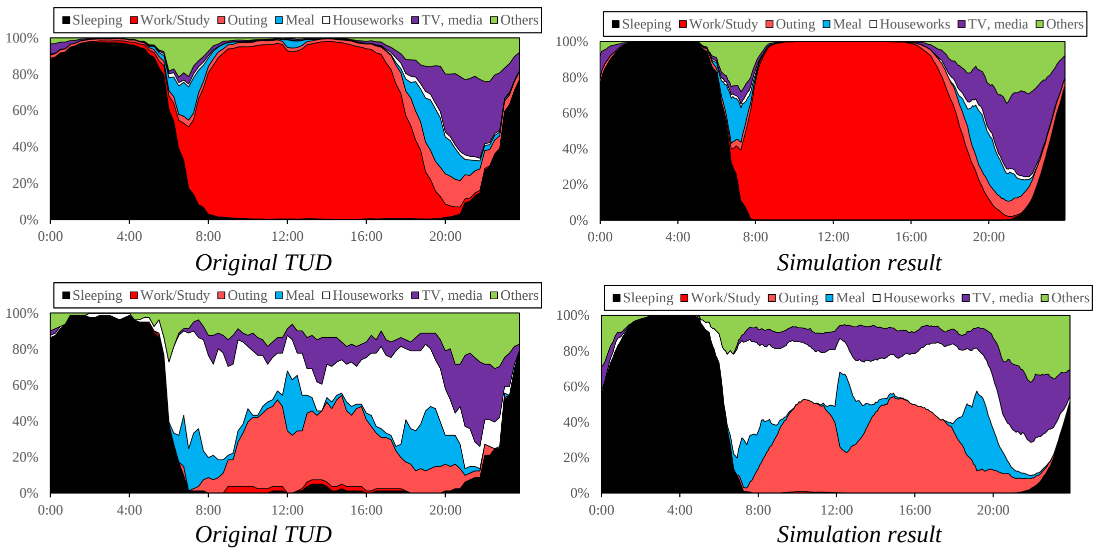

Figure 4. *Percentage of days on which each behavior is conducted at times of day*

shows that the simulation result agreed well with the original TUD. It should be noted that the beginning time of dinner and bathing have small discrepancy. This is due to the order of placement of the routine behaviors. When there are a routine behavior, such as working, placed earlier, the routine behaviors that are placed lately, like dinner and bathing, cannot be placed on the time steps. This is why simulation result of dinner beginning time is slightly delayed than the original.

Duration time of routine and non-routine behaviors

Figure 3 shows the mean durations of routine and non-routine behaviors observed for working male and housewife. Both of the axes are logarithmic. As shown in the result for the routine behaviors, the duration of the routine behaviors were well agreed between the simulation result and the original TUD. For non-routine behaviors, the duration summarized for each time region that was explained in the previous chapter, is distinguished by different marker. Although the behaviors are not specified in the figure, the result shows that the simulation result was underestimated for most of non-routine behaviors compared to the original TUD. This can be attributed to the same reason as the discrepancy observed in the beginning time of dinner and bathing. As designed in the algorithm, the routine behaviors are first placed and vacant steps between routine behaviors are filled by non-routine behaviors. When duration of selected behavior is larger than vacant steps, the duration is shortened to fill the vacant steps. Thus, cumulative probability distribution that is used for determining duration of non-routine behaviors is not fully applied when the size of vacant steps is small, which resulted in a shorter mean duration of non-routine behaviors shown in Figure 3.

Percentage of simulated days on which each behavior is undertaken at each time of day

Figure 4 shows the probability distribution showing percentage of simulation days on which each behavior is undertaken on each time of day calculated for the working male and housewife. The result agreed well with the original TUD.

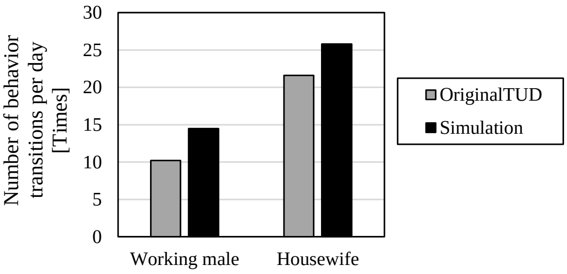

Figure 5. *Number of behavior transition per day*

Number of behavior transitions per day

Figure 5 shows the mean of the number of behavior transitions per day. As shown in the figure, the number of behavior transitions was overestimated. This overestimation corresponds to the underestimation of the durations for non-routine behaviors.

Selection probability of behaviors from other behaviors

Figure 6 shows the distribution of probability showing the share of behaviors that were selected when the simulated person finished the behaviors

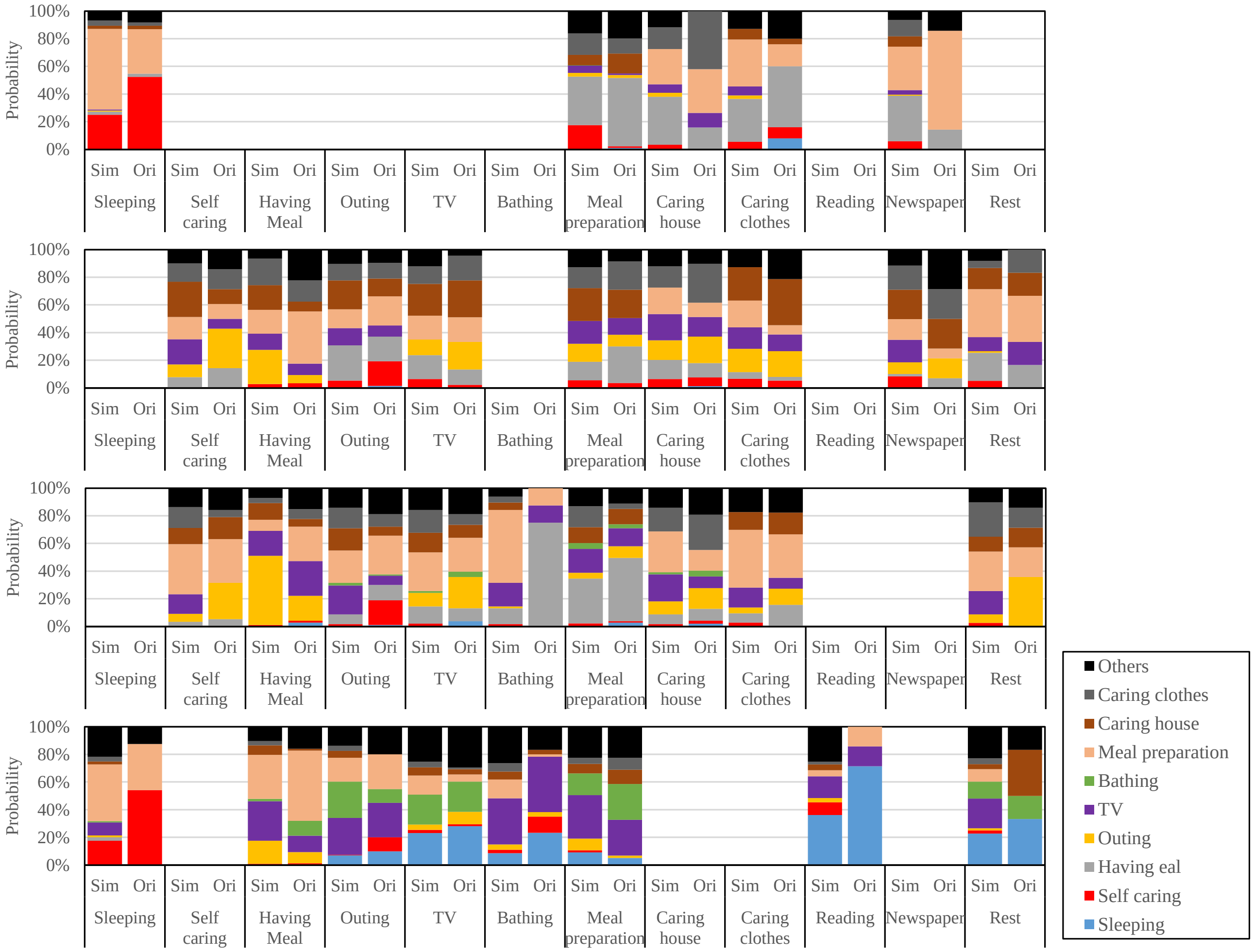

Figure 6. *Distribution of selection probability from the behaviors listed on the axes (corresponding to the time regions 1 to 4 from the top to the bottom)*

listed on the horizontal axes, which was calculated for the housewife for each of the time region 1 to 4 (from the top to the bottom). This indicator is well replicated in Markov chain models. The result only shows the behaviors from which transition probability showed a statistically significant difference from those of the original TUD.

We observed two important differences between the simulation result and the original TUD. First, the distributions for the routine behaviors, sleeping, having meals and bathing, differ significantly between the simulation result and the original TUD. In the simulation model, the behavior undertaken after these routine behaviors were selected by using $SP_{t,n}$ that is the adjusted version of $PB_{t,n}$, percentage of respondents who adopt the behavior at time t. However, the discrepancy in the selection probability implies that there is a certain sequence of behaviors after the routine behaviors that differs from those randomly selected based on $SP_{t,n}$. For example, in the result for sleeping in the time region 1, the share of the behavior with orange color, which is self-caring, is large in the original TUD. It is natural to conduct changing clothes or washing face/teeth described by self-caring soon after awaking. Contrary, the yellow part, watching TV, is the largest in the simulation result because $PB_{t,n}$ of watching TV is the largest in the time region .

The second difference cannot be observed in the figure. In the simulation result, more kinds of behaviors were selected than the original TUD. It should be noted that the sample size of the Diary B for the housewife category is only 84. Thus, there are many combinations of behaviors between which any transition was not observed in the original TUD. On the other hand, the simulation model generated a wider variety of combinations because non-routine behaviors were selected based on $PB_{t,n}$ that has a corresponding variety in the original TUD.

DISCUSSION

As mentioned in the introduction, the proposed behavior model was designed to be applied to community/urban scale modeling of the residential energy demand. For this simulation context, we classified time use diaries into several categories to generate a variety of TUD for hundreds/thousands of households. In addition to the feature in data preparation, the proposed model distinguish routine and non-routine behaviors in the simulation algorithm. In the algorithm, the routine behaviors are placed on the timeline. Then, the non-routine behaviors are placed so that the non-routine behaviors fill the gaps between the routine behaviors. It is beneficial to distinguish routine and non-routine behaviors for considering interaction among household members, even though we did not show any simulation result showing it. Times at which meal and bathing begin can be adjusted according to behaviors of other household members. In addition to this, several categories was developed for working male and female according to the characteristics in routine behaviors.

However, the simulation results shown in the previous chapter showed that there are some drawbacks in the proposed model. First, the number of behavior transitions is overestimated while the duration of non-routine behaviors per day is underestimated as shown in Figures 3 and 5. As mentioned earlier, this can be attributed to the filling process of vacant steps by non-routine behaviors. In most filling process, the last vacant steps is smaller than duration randomly selected based on its cumulative probability distribution. In such case, the duration is shortened to fit the vacant steps. This is why the difference in the number of behavior transitions per day shown in Figure 5 is close to the number of the routine behaviors.

One possible solution to this problem is that the filling of small vacant steps by a newly selected non-routine behavior is replaced by extending the last non-routine behavior to fill the vacant steps.

The second drawback of the proposed model is that the selection probabilities from one behavior to another shown in Figure 6 differ significantly from those of the original TUD. As mentioned earlier, at least, behavior sequence when routine behaviors ended must be considered. It is possible by using transition probability from the routine behaviors to non-routine behaviors.

On the contrary, the proposed model has an advantage in the selection of non-routine behaviors. The model does not use transition probability from behaviors to another behaviors. When the number of sample size of TUD is small, there are many combinations of behaviors between which any transition was observed in original TUD. There are some time use categories with small number of TUD sample since we developed 25 categories as listed in Table 2. Housewife category is one of the examples. The proposed model enables to generate behavior transitions even for occupants in a category with a small number of TUD sample. This is an important advantage to generate a variety of time use data for community/urban scale modelling of energy use of residential buildings.

CONCLUSION

This paper proposes a new approach for modelling occupants' time use which can be converted into presence and action in home. In the modelling approach, routine and non-routine behaviors are distinguished. Routine behaviors are first placed on timeline and then gaps between routine behaviors are filled by non-routine behaviors. For the filling gap process, behaviors that fill gaps are selected while considering duration of behaviors in addition to percentage of behaviors observed in sample time use data at examined time of day. This paper evaluated how this simulation procedure influences simulation results. The evaluation showed that there are two drawbacks. First, the duration of non-routine behaviors is underestimated while the number of behavior transitions per day is overestimated. Second, selection probability from one behavior to another differ from those of the original time use data. Contrary, the procedure has some functional advantages. First, it enables to generate time use data even if the sample size of time use data used to develop input dataset is small. Second, interaction among household members can be considered for routine behaviors, such as having meals together and taking bath sequentially.

As our future works, we will compare the proposed approach with the existing modeling approaches like Markov chain.

REFERENCES

Baptista, M.L., Fang, A., Prendinger, H., Prada, R., Yamaguchi, Y. 2014. Accurate Household Occupant Behavior Modeling Based on Data Mining Techniques. Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, pp. 1164-1170.

Richardson, I., Thomson, M., Infield, D., Clifford, C. 2010. Domestic electricity use: A high-resolution energy demand model. Energy and Buildings, vol. 42, pp. 1878-1887

Shimoda Y., 2007. Evaluation of city-scale impact of residential energy conservation measures using the detailed end-use simulation model, Energy 32, pp.1617-1633

Statistics Bureau, Ministry of Internal Affairs and Communications. Survey on Time Use and Leisure Activities in 2006. http://www.stat.go.jp/english/data/shakai/index.htm, accessed on August 8, 2014.

Tanimoto J., Hagishima. A., Sagara. H. 2008. Validation of methodology for utility demand prediction considering actual variations in inhabitant behaviour schedules, Journal of Building Performance Simulation, vol.1(2008), pp.31-42

Yamaguchi, Y., Shimoda Y. Validation of an Energy Demand Model of Residential Buildings. Proceedings of the ASim2014, Nagoya, Japan, pp. 625-632, 2014.11.

Widén, J., Wäckelgård, E. 2010. A high-resolution stochastic model of domestic activity patterns and electricity demand. Applied Energy, vol. 87, pp. 1880-1892.

Wilke, U., Haldi, F., Scartezzini, JL; Robinson, D. 2013. A bottom-up stochastic model to predict building occupants' time-dependent activities, Building and Environment vol. 60, pp. 254-264