Materials Research Proceedings **27** (2023) 103-110 https://doi.org/10.21741/9781644902455-13

Modeling and utilizing habits using process mining for building spatial design systems

Naohiro Haraguchi[1, a *] and Ami Ogawa[2, b]

[1]School of Science for Open and Environmental Systems, Keio University, 3-14-1 Hiyoshi, Kohoku, Yokohama 223-8522, Japan

[2]Department of System Design Engineering, Keio University, 3-14-1 Hiyoshi, Kohoku, Yokohama 223-8522, Japan

[a]naohiro1711@keio.jp, [b]ogawa@sd.keio.ac.jp

Keywords: Smart Home, Process Mining, Habit Model, Spatial Design

Abstract. Residents need to change their habitual behaviors following living space changes, such as moving or remodeling, and that may occur mental stress. This stress is a major problem, especially for the elderly, who are less able to cope with changes in their environment. To reduce this stress, a system that reflects the living information of the original houses in new houses, where habits can be retained in the new environments is needed. Many studies have been conducted to quantify life information as a habitual model using data mining and pattern recognition methods. “Process Mining” is a theory developed to visualize and improve processes in the business field and applied to lifestyle information, and it is possible to create a habit model. In recent years, several studies on habit models using process mining have been reported. However, there are no studies in which these process mining-based habit models have been adopted to design architectural spaces such as living spaces. Therefore, the purpose of this study is to investigate the relationship between habit and architectural space by utilizing a process mining-based habit model. Specifically, we propose the automatic extraction and visualization of habit behaviors through process mining and the use of habit models. The data acquisition experiment was conducted in an experimental smart home. This smart home is a mobile trailer house built by a multi-company project and is equipped with many sensors that can automatically acquire many daily living data. Subjects were recruited randomly and lived alone in this smart home for one week. An input matrix was created from the acquired data set and process mining was adapted to create habit models. In this study, two habit models were created: (1) a habit model based on behavioral information and (2) a habit model based on location information. Each input matrix consisted of (1) 16 types of behavior record data manually entered by the subject and (2) ground reaction force data in the house divided into 7 areas. We investigated the relationship between habitual behaviors and spatial conditions by integrating these two models.

Introduction

People experience a great deal of stress when their living environment changes, such as when they move or renovate their homes [1]. The elderly often renovate their homes as their bodies age. While physical aspects such as functionality and accessibility are often considered during such home renovations, it is also important to reflect on personal experiences such as personal habits [2]. In other words, reflecting on personal life information from the previous home during the design phase of a new home and tailoring the home design to the residents will reduce this stress and allow for a smooth home move. Barry developed a set of guidelines for designing spaces that reflect his habits for the renovation of his father's house [3]. While these guidelines reflect personal experience and help smooth home changes, they are problematic in that they do not quantitatively automate the acquisition of habitual behaviors or the stages of interior space design. On the other

Materials Research Proceedings **27** (2023) 103-110 https://doi.org/10.21741/9781644902455-13

hand, although many studies have been conducted to automatically create optimal space design [4], there are still no studies that automatically form indoor spaces using habit information as input. Since actual individual habits are complex, it is considered difficult to incorporate habit information at the spatial design stage. Therefore, in order to spatial design with habit as an input, it is first necessary to quantitatively clarify the relationship between architectural space and habit. In recent years, several studies have reported on the application of "process mining" to lifestyle information to create a habit model. Process mining is a theory for visualizing and improving processes in the business field. Dimaggio et al. created habit models by applying process mining to log data obtained from PIR (Passive Infra-Red) sensors [5]. Ma'arif created habit models by applying labeled events [6]. These studies only used data from a single type of sensor acquired in the smart home, were aimed at creating habit models, did not discuss the use of habit models, and did not utilize them in spatial design.

Therefore, the purpose of this study is to investigate the relationship between habits and architectural space by utilizing habit models created using process mining. This study uses experimental data from a smart home located in Japan, where habit models were created by applying process mining to data acquired from multiple types of sensors. This study is positioned as the first step in the construction of a space design system using the habit model as input.

Materials and methods

・*Experimental setup and subjects*

The data acquisition experiment was conducted under the experimental environment of a smart home that simulates a living space, called the "Mirai no ie". This smart home was built for the verification of cutting-edge systems as a part of a collaborative project. As shown in Figure 1, this smart home is a mobile trailer house, and the experimental environment was designed for a single person living alone. This smart home is equipped with many sensors, including IoT home appliances, and can automatically acquire daily life data such as the usage status of home appliances and physiological data of the subjects. In addition, subjects can control door locks and lightings from their smartphones.

Subjects were required to live alone in this smart home for one week. Subjects did not have to stay in the house all day but could go to school or work as usual and lead a natural life. Subjects were also given a tablet and tasked to press "start" when they performed an activity in the home and "end" when they completed the activity. The experiments were approved by the ethical committee of Keio University Faculty of Science and Technology (approval number: 2020-73) and the Kanagawa Institute of Industrial Science and Technology (ethical committee on Research Development and Demonstration of Robots for Human Subjects).

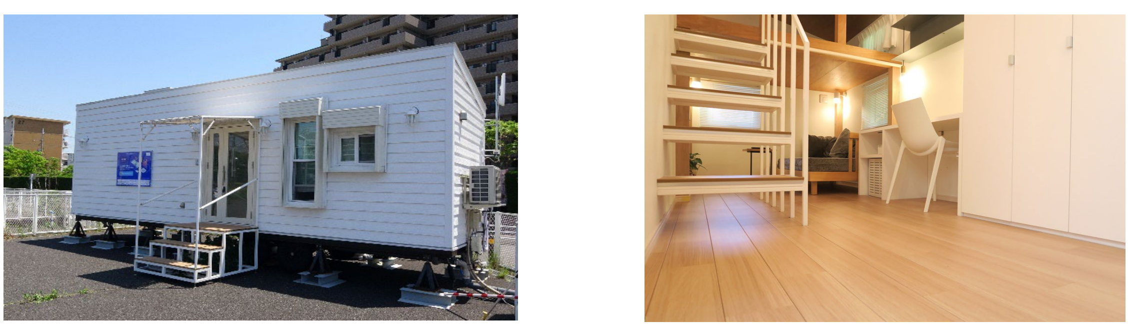


Figure 1. Photos of the exterior and interior of the smart home named "Mirai no ie"

Materials Research Proceedings **27** (2023) 103-110 https://doi.org/10.21741/9781644902455-13

・***Data analysis: Process mining***

In this study, an input matrix was created using two types of data from the acquired data, activity and location information, and a habit model was created by applying process mining.

- *Extraction of Event Logs*

Activity information was data manually entered by the subjects. Subjects were given a tablet and tasked to press "start" when they performed an activity in the house and "end" when they finished the activity. The tablet displayed "start" and "end" buttons for the 16 activities shown in Table 1. The location information was obtained from the ground reaction force sensors (Locationfloor, Toppan Inc., Tokyo, Japan) installed all over the floor in the smart home. These sensors are integrated with flooring and can detect the resident's positions. In this study, the entire house was divided into seven zoning areas as shown in Figure 2, and the data was stored. The ground reaction force data continued to be stored automatically if the subject was in the house.

Table 1. 16 different activities displayed on the tablet

TV: watching TV	**Cook**: cooking a meal
Outing: getting out of the house	**Medicine**: taking the medicine
Smartphone: using a smart phone	**Laundry**: doing laundry
Toilet: using the toilet	**Sleep**: going to bed
Work: doing work	**Preparation**: preparation for outing
Hygiene: brushing teeth, washing hands, etc.	**Exercise**: exercise one's body
Bath: taking the bath	**Eat**: eating food
Cleaning: cleaning home	**Dishwashing**: washing the dishes

Figure 2. The locations of the ground reaction force sensors and defined seven zoning areas

- *Event Log Filtering*

Because the amount of data in the location information recorded automatically is huge compared to the manually input activity information, filtering was performed on the location information log data to align the granularity of the event logs. Specifically, time frame filtering was created based on the time from the start to the end of the activity information, and location information data was reduced.

Materials Research Proceedings **27** (2023) 103-110 https://doi.org/10.21741/9781644902455-13

- *Creation of Input Matrix*

To perform process mining, three columns are required: Subject ID, Timestamp, and Event log. In this study, both the activity data and location data were included in the Event log columns as shown in Table 2. This means that the event log includes the start and end of each activity, as well as location information.

Table 2. Input Matrix Example

Subject ID	Timestamp	Event log
1	2020-12-08 09:31:44	TV_start
1	2020-12-08 09:31:54	Sofa
1	2020-12-08 09:51:31	TV_end
⋮	⋮	⋮

- *Creation of habit model*

Process mining was applied to the input matrix to create a habit model. This study used Heuristic Miner, one of the process mining algorithms. Heuristic Miner can create process models by considering the frequency and is suitable for irregular data such as human habit models [7].
Heuristic Miner calculates the degree of dependence, meaning the connection between each event [8]. The degree of dependence is calculated between 0.00~1.00, and the closer to 1.00, the stronger connections. Only those with this degree of dependence exceeding the threshold value were formed as habit models.
Among the several formats for displaying process models, we used the Heuristic Net to visualize the created habit models. Since Heuristic Nets are models of connected only event logs, they are a suitable representation method for visualizing complex human habits. The activity and location information in the event logs are represented by transitions, respectively, and the processes that exceed the threshold values are connected in the model. This model extracts only the processes that exceed the threshold value, and it is possible to create habit models with different granularity by changing the threshold value.
The open source pm4py (Fraunhofer Institute for Applied Information Technology FIT, Sankt Augustin, DE) was used to execute the process mining [9]. We also used Anaconda3 (Anaconda Inc., Austin, TX, USA) to perform the analysis.

・***Evaluation***

After Process mining, to investigate the relationship between activities and location, from the created habit model, we extracted only those parts of the model where there was a connection between activity and location information for each subject. This is necessary because the initial habit model also has activity-to-activity processes, so we extracted only the processes that location-activity in the habit model.

Results and discussions

The subjects were 4 students (M:2, F:2) aged 22.5±0.5, and the data acquisition experiment ran from November 17, 2020, to December 15, 2020. They lived in this smart home alone for a week and collected data from check-in to check-out. Process mining was applied to the input matrix to extract processes for each subject’s data.

Materials Research Proceedings **27** (2023) 103-110 https://doi.org/10.21741/9781644902455-13

・***Habit models***

A habit model was created for each subject. As an example of a habit model, Figures 3 display subject A's habit models at two threshold values; (1) 0.80 and (2) 0.50. In Figures 3, we can see the connections between each activity and location and read the visualized personal habits. Comparing (1) and (2), at higher threshold, the number of event logs extracted and the number of paths connecting them are reduced, making the habit model simpler. In other words, the application of the Heuristic Miner algorithm requires the determination of the threshold value according to the purpose of the study since granularity of the habit model changes significantly depending on the threshold value. This method has the potential to approach a variety of problems such as daily monitoring of the elderly and personal habit change by adjusting the threshold value of the habit model.

In this case, a threshold value of 0.50 is appropriate because the purpose of this study is not to look at the connection between specific activities and location data, but rather to conduct a broad survey. In addition, duration of this experiment was only one week, and the number of data was small. Usually, the degree of dependence on Heuristic Miner increases as the number of data increases, so a high threshold value should not be set in this case; 0.5 is appropriate.

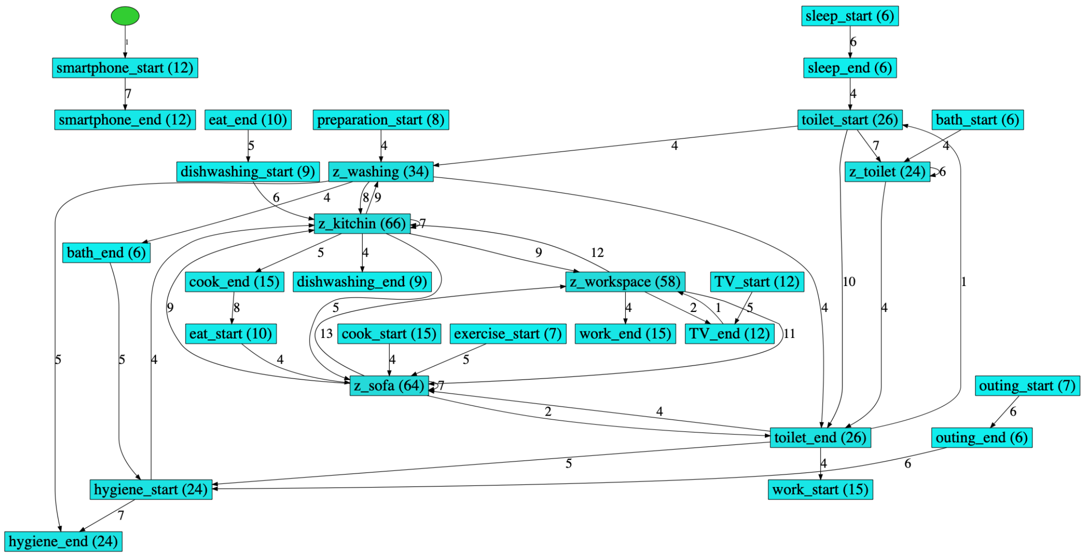
・***Relationship between habit and architectural space***

Table 3 shows the names of the activity associated with location information and their degree of dependence that extracted from the habit model when the threshold value is 0.50. These activities indicate that they took place at the corresponding locations. Activities with high degrees of dependence are more strongly tied to the location.

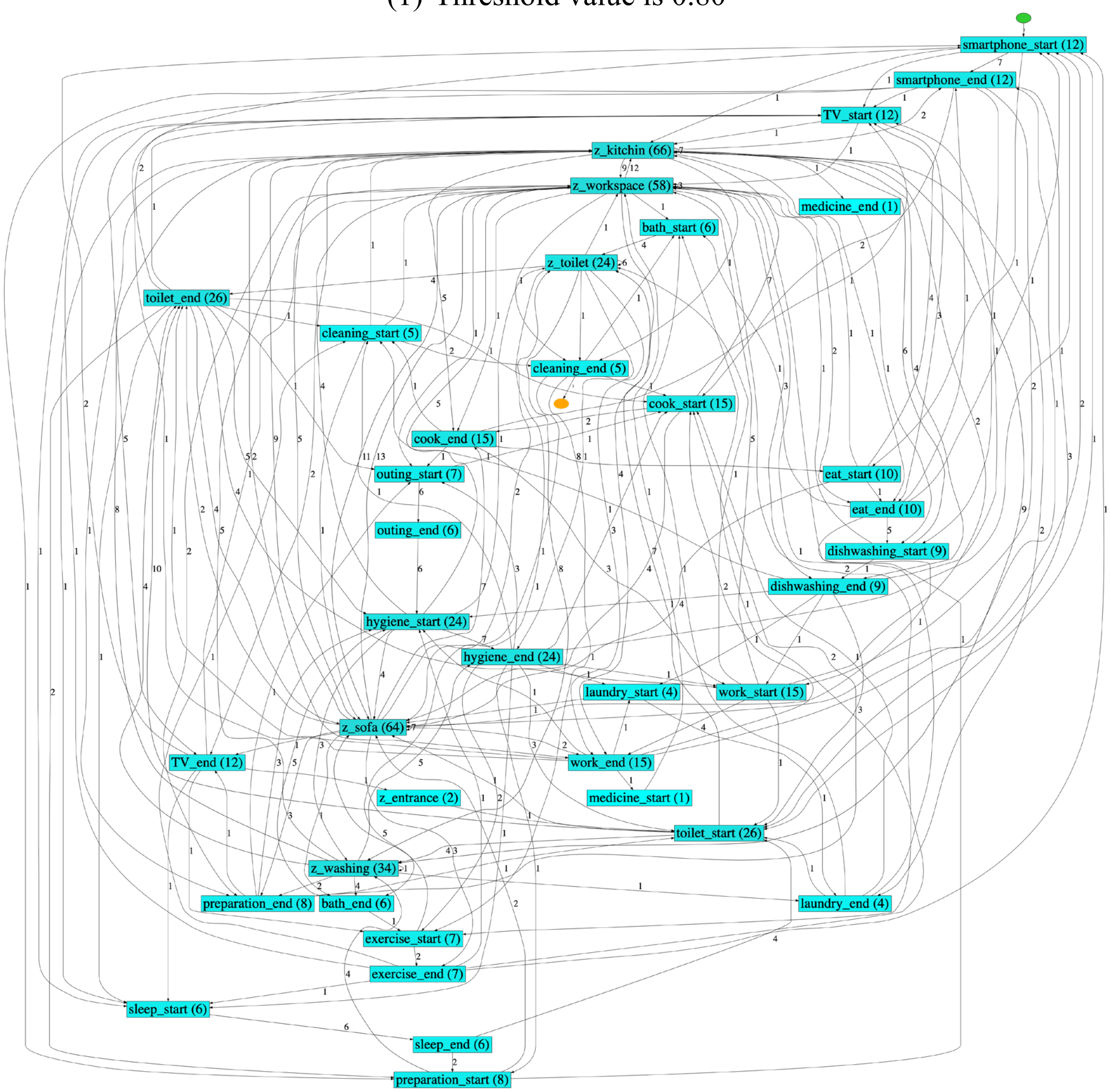
Focusing on the seven areas we can see that multiple activities were taking place in each area. For example, focusing on subject A, the type of activity performed in the workplace area and in the washing area is different. This means that this subject has a habit of performing certain activities in certain architectural spaces. This suggests the need to design spaces that are tailored to habits. At the same time, the values of the dependence differ for each activity, allowing us to see each of the activities with the strongest ties to each area.

A comparison of the results of the 4 subjects shows that some results were common, and some were not in the relationship between habits and architectural space. For example, focusing on the workplace area, we see that all subjects are performing "work" activity, but only subjects B and C were performing "smartphone" activity at workspace. Furthermore, a common activity to all subjects is "hygiene" activity in the washing area and "cook" and "eat" activities in the kitchen area. On the other hand, there was no common activity to all in the toilet, sofa, entrance and closet areas, but individual characteristics were evident. The fact that individual differences appeared in the relationship between activity and location information while living in the same house strongly suggests the need to create architectural spaces that are tailored to the habits of each resident.

Materials Research Proceedings **27** (2023) 103-110 https://doi.org/10.21741/9781644902455-13



(1) Threshold value is 0.80



(2) Threshold value is 0.50

Figure 3. An example of a subject A's habit model of Heuristic Net in Pm4py

Table 3. Activity information and the degrees of dependence strongly related to the 7 areas (Threshold value is 0.50)

Area	Subject A		Subject B		Subject C		Subject D	
workplace	**work**	**0.80**	**work**	**0.90**	preparation	0.80		
	dishwashing	0.75	hygiene	0.75	hygiene	0.75		
	eat	0.66	eat	0.66	sleep	0.66	**work**	**0.50**
	preparation	0.66	cook	0.50	toilet	0.66	hygiene	0.50
	cook	0.50	sleep	0.50	smartphone	0.66	bath	0.50
	cleaning	0.50	toilet	0.50	outing	0.66	outing	0.50
			outing	0.50	**work**	**0.66**	smartphone	0.50
			preparation	0.50	dishwashing	0.66		
					bath	0.50		
					laundry	0.50		
					eat	0.50		
					cook	0.50		
washing	**hygiene**	0.83			**hygiene**	0.83	bath	0.75
	bath	0.80			preparation	0.75		
	toilet	0.80						
	preparation	0.66	sleep	0.50	bath	0.50	TV	0.66
	laundry	0.50	**hygiene**	0.50	work	0.50	cook	0.66
							hygiene	0.50
							preparation	0.50
toilet	toilet	0.80	bath	0.75				
	bath	0.50	cook	0.50	toilet	0.66	bath	0.50
	hygiene	0.50			smartphone	0.50	smartphone	0.50
	cleaning	0.50						
	work	0.50						
sofa	preparation	0.75			smartphone	0.75		
	exercise	0.75						
	cook	0.66	dishwashing	0.66	eat	0.66	smartphone	0.66
	work	0.60	eat	0.50	exercise	0.66	cook	0.66
	bath	0.50	bath	0.50	hygiene	0.66	eat	0.66
	TV	0.50			TV	0.66	work	0.50
	smartphone	0.50			sleep	0.50	TV	0.50
					bath	0.50		
					preparation	0.50		
kitchen	**cook**	0.83	**eat**	0.83	**eat**	0.88	**eat**	0.75
	dishwashing	0.80	**cook**	0.83	toilet	0.80	**cook**	0.75
					bath	075	hygiene	0.75
					cook	0.75		
	smartphone	0.66	laundry	0.66	smartphone	0.50	TV	0.66
	work	0.66	hygiene	0.66	laundry	0.50	work	0.50
	laundry	0.60	bath	0.50	hygiene	0.50	preparation	0.50
	eat	0.50	toilet	0.50	outing	0.50	outing	0.50
	preparation	0.50						
	cleaning	0.50						
	medicine	0.50						
entrance								
			outing	0.50	outing	0.50	outing	0.50
			eat	0.50				
			work	0.50				
closet								
			work	0.50			outing	0.50

*Cells of high degrees of dependence are colored in blue.
**Bold shows common activities through all subjects.

Conclusion and Future plan

By applying process mining to the data acquired in the smart home, a model of habits was created, and investigated the relationship between habits and architectural space. In particular, we applied process mining by incorporating both activity and location information in the same input data. From the created habit model, we visualized where and which activities are performed in the house, and the results showed that the relationship between habits and architectural space differs from person to person, suggesting the need to create architectural space tailored to each person's habits. In the future, we will construct a space design system using the created habit model as input.

Acknowledgment
The authors thank the Mirai no ie project for providing the experimental environment and dataset and subjects who participated positively in the experiment

References

[1] A. Ballantyne, "Architecture, life, and habit," *The Journal of Aesthetics and Art Criticism*, vol. 69, no. 1, pp. 43-49, 2011. https://doi.org/10.1111/j.1540-6245.2010.01445.x

[2] B. Tanner, C. Tilse, and D. De Jonge, "Restoring and sustaining home: The impact of home modifications on the meaning of home for older people," *Journal of Housing for the Elderly*, vol. 22, no. 3, pp. 195-215, 2008. https://doi.org/10.1080/02763890802232048

[3] J. E. Barry, "Everyday habits and routines: Design strategies to individualize home modifications for older people". *PhD diss., Washington State University*, 2008.

[4] D. H. Dorrah, and M. Marzouk, "Integrated multi-objective optimization and agent-based building occupancy modeling for space layout planning," *Journal of Building Engineering*, vol. 34, 2021, 101902. https://doi.org/10.1016/j.jobe.2020.101902

[5] M. Dimaggio, F. Leotta, M. Mecella, and D. Sora, "Process-based habit mining: Experiments and techniques," *2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld),* IEEE, pp. 145-152, 2016. https://doi.org/10.1109/UIC-ATC-ScalCom-CBDCom-IoP-SmartWorld.2016.0043

[6] M. R. Ma'arif, "Revealing daily human activity pattern using process mining approach," *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, IEEE, pp. 1-5, 2016. https://doi.org/10.11591/eecsi.v4.1066

[7] T. Iijima, K. Tabata, and S. Saito, "A survey on business process mining (05: Case study), *Journal of Information Systems Society of Japan*, vol. 13, no. 2, pp. 78-86, 2018. (in Japanese)

[8] A. Weijters, W. van der Aalst, and A. Alves de Medeiros, "Process mining with the heuristics miner-algorithm," Technische Universiteit Eindhoven, Tech. Rep. WP 166, 2006.

[9] A. Berti, S. J. van Zelst, W. M. P. van der Aalst, "Process mining for python (PM4Py): bridging the gap between process-and data science," P*roceedings of the ICPM Demo Track 2019, Co-located with 1st International Conference on Process Mining (ICPM 2019)*, pp. 13-16, 2019.