

# Spatio-Temporal Situation Recognition in Service Fields

- Validation by Discrete-Event Simulation -

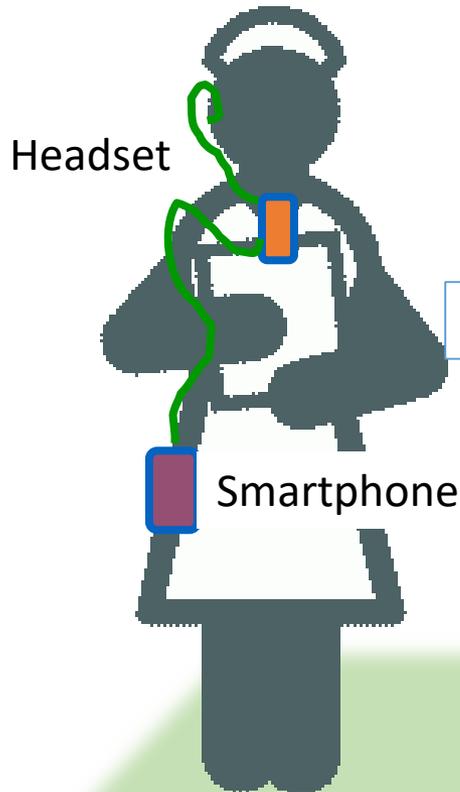
Kunihiko Hiraishi

School of Information Science,  
Japan Advanced Institute of Science and Technology  
(JAIST)

# Smart Voice Messaging System



Bluetooth Beacon



Headset

Smartphone

WiFi/4G



Voice Message

This is ○○○. I'll start drip soon.

Location Data (Bluetooth Sensor)  
Action Data (Acceleration Sensor)

Daily Reports



Speech recognition &  
Keyword extraction



時間	位置	動作	速度	加速度	その他
10:00	10.0000	0.0000	0.0000	0.0000	
10:01	10.0000	0.0000	0.0000	0.0000	
10:02	10.0000	0.0000	0.0000	0.0000	
10:03	10.0000	0.0000	0.0000	0.0000	
10:04	10.0000	0.0000	0.0000	0.0000	
10:05	10.0000	0.0000	0.0000	0.0000	
10:06	10.0000	0.0000	0.0000	0.0000	
10:07	10.0000	0.0000	0.0000	0.0000	
10:08	10.0000	0.0000	0.0000	0.0000	
10:09	10.0000	0.0000	0.0000	0.0000	
10:10	10.0000	0.0000	0.0000	0.0000	
10:11	10.0000	0.0000	0.0000	0.0000	
10:12	10.0000	0.0000	0.0000	0.0000	
10:13	10.0000	0.0000	0.0000	0.0000	
10:14	10.0000	0.0000	0.0000	0.0000	
10:15	10.0000	0.0000	0.0000	0.0000	
10:16	10.0000	0.0000	0.0000	0.0000	
10:17	10.0000	0.0000	0.0000	0.0000	
10:18	10.0000	0.0000	0.0000	0.0000	
10:19	10.0000	0.0000	0.0000	0.0000	
10:20	10.0000	0.0000	0.0000	0.0000	
10:21	10.0000	0.0000	0.0000	0.0000	
10:22	10.0000	0.0000	0.0000	0.0000	
10:23	10.0000	0.0000	0.0000	0.0000	
10:24	10.0000	0.0000	0.0000	0.0000	
10:25	10.0000	0.0000	0.0000	0.0000	
10:26	10.0000	0.0000	0.0000	0.0000	
10:27	10.0000	0.0000	0.0000	0.0000	
10:28	10.0000	0.0000	0.0000	0.0000	
10:29	10.0000	0.0000	0.0000	0.0000	
10:30	10.0000	0.0000	0.0000	0.0000	
10:31	10.0000	0.0000	0.0000	0.0000	
10:32	10.0000	0.0000	0.0000	0.0000	
10:33	10.0000	0.0000	0.0000	0.0000	
10:34	10.0000	0.0000	0.0000	0.0000	
10:35	10.0000	0.0000	0.0000	0.0000	
10:36	10.0000	0.0000	0.0000	0.0000	
10:37	10.0000	0.0000	0.0000	0.0000	
10:38	10.0000	0.0000	0.0000	0.0000	
10:39	10.0000	0.0000	0.0000	0.0000	
10:40	10.0000	0.0000	0.0000	0.0000	
10:41	10.0000	0.0000	0.0000	0.0000	
10:42	10.0000	0.0000	0.0000	0.0000	
10:43	10.0000	0.0000	0.0000	0.0000	
10:44	10.0000	0.0000	0.0000	0.0000	
10:45	10.0000	0.0000	0.0000	0.0000	
10:46	10.0000	0.0000	0.0000	0.0000	
10:47	10.0000	0.0000	0.0000	0.0000	
10:48	10.0000	0.0000	0.0000	0.0000	
10:49	10.0000	0.0000	0.0000	0.0000	
10:50	10.0000	0.0000	0.0000	0.0000	
10:51	10.0000	0.0000	0.0000	0.0000	
10:52	10.0000	0.0000	0.0000	0.0000	
10:53	10.0000	0.0000	0.0000	0.0000	
10:54	10.0000	0.0000	0.0000	0.0000	
10:55	10.0000	0.0000	0.0000	0.0000	
10:56	10.0000	0.0000	0.0000	0.0000	
10:57	10.0000	0.0000	0.0000	0.0000	
10:58	10.0000	0.0000	0.0000	0.0000	
10:59	10.0000	0.0000	0.0000	0.0000	
11:00	10.0000	0.0000	0.0000	0.0000	

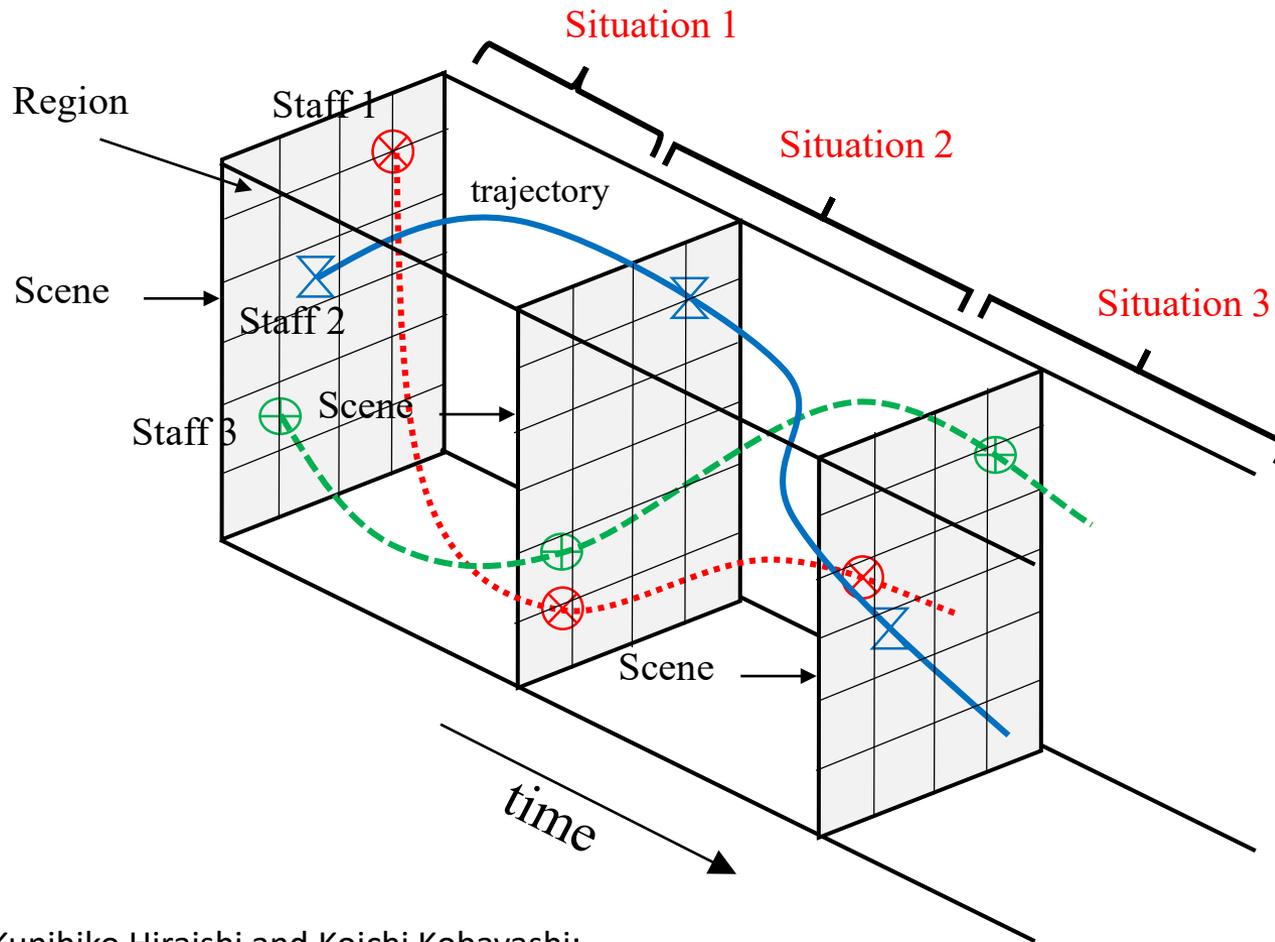
Event Logs

Server on Cloud

Distribution  
Rules

Distribution of  
Voice/Text messages  
Asynchronous Communication

# Spatio-Temporal Situation Recognition



Kaoru Sato, Kunihiro Hiraishi and Koichi Kobayashi:  
Spatio-Temporal Situation Recognition for Groups in Caregiving Services,  
IEEE CogSIMA2016, pp.78-82 (2016/3/21-25, San Diego, USA)

# Spatio-Temporal Situation Recognition

- Phase I: Recognition of Scenes

1. From the event log, we make *place vectors*. Place vectors indicates how persons are arranged in the space.

**1 2 3 4 5 6 7 8 9 10**  
[0, 0.5, 0, 3, 2.5, 0, 0, 0, 0, 0]

There are 10 regions. Three persons are in region 4, two persons are in region 5, and one person is moving from region 2 to region 5 (0.5 is added to region 2 and region 5).

2. A clustering algorithm is applied to the set of place vectors and we obtain a set of clusters  $C = \{ c_1, \dots, c_n \}$ . Each  $c_i$  is called *a scene*.

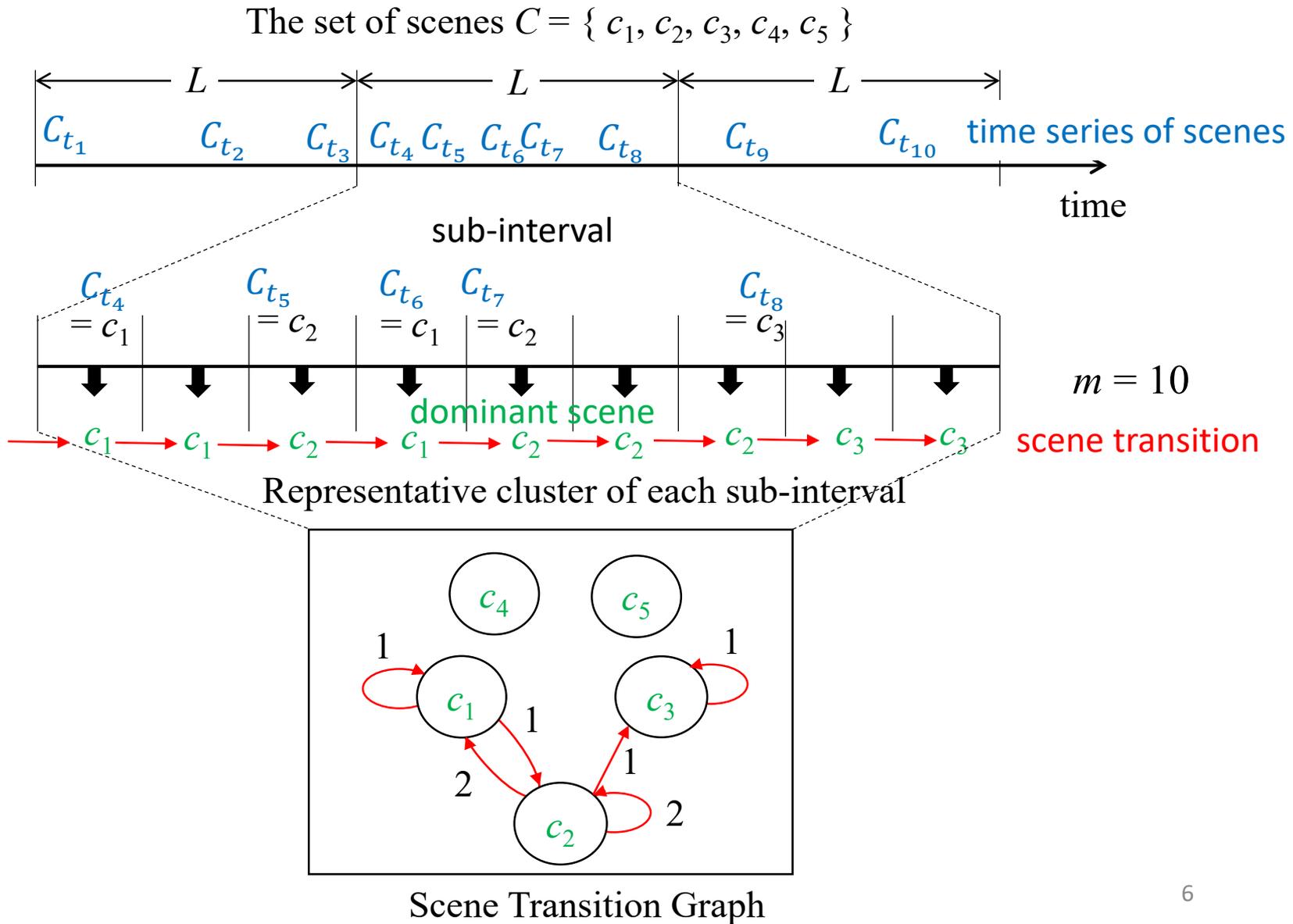
# Spatio-Temporal Situation Recognition

- Phase II: Recognition of Scene Transition

1. A time series of place vectors  $v_{t_1}, v_{t_2}, \dots$  ( $t_1 < t_2 < \dots$ ) is transformed into a time series of scenes  $C_{t_1}, C_{t_2}, \dots$ .
2. Time axis is divided into intervals with length  $L$ .
3. Each interval is further divided into  $m$  sub-intervals.
4. Decide a dominant scene for each subinterval.
5. Make *a scene transition graph* for each interval.
6. Clustering of the set of scene transition graphs. Each cluster is called *a situation*.

**Situation = How scenes change with time**

# Spatio-Temporal Situation Recognition



# Aim

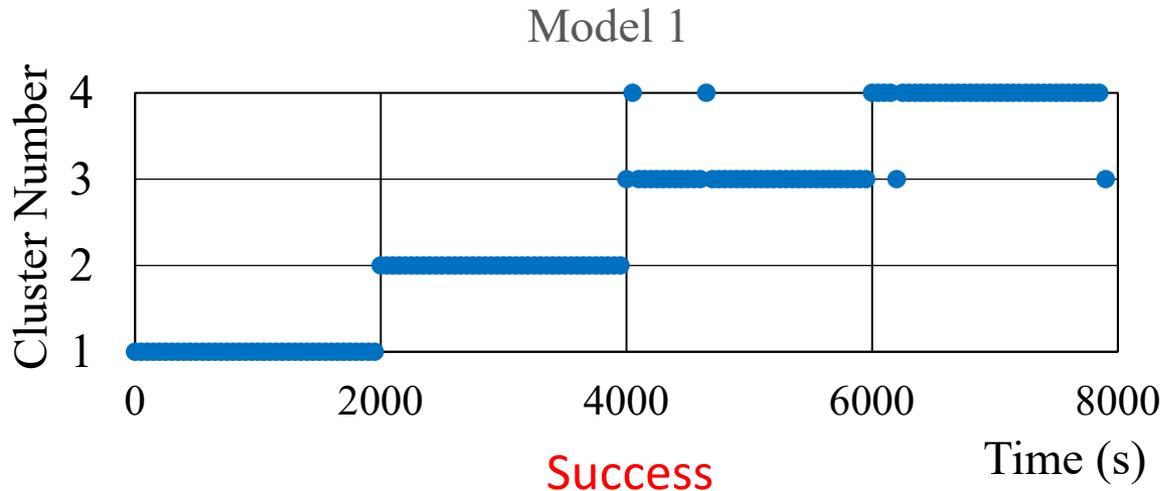
- The proposed method is applied to a small number of real data obtained in a nursing home.
- In this paper, we validate the method by artificial logs generated by discrete event simulation.
  - to find cases that the proposed method does not work well
  - to find idea toward improvement

# Simple Model 1

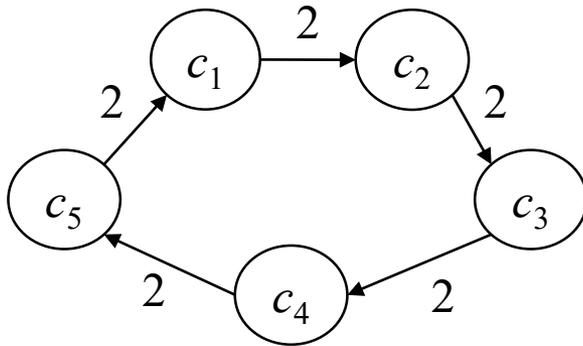
- There are five rooms 1 - 5 and three persons  $P_1, P_2, P_3$ .

Model 1	
Time(sec.)	Rules
0 ~ 1999	$R_1. P_1 \sim P_3: 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 1 \rightarrow \dots$
2000 ~ 3999	$R_2. P_1 \sim P_3: 1 \rightarrow 5 \rightarrow 4 \rightarrow 3 \rightarrow 2 \rightarrow 1 \rightarrow \dots$
4000 ~ 5999	$R_3. P_1 \sim P_3: \text{Move rooms } 1 \sim 5 \text{ randomly.}$
6000 ~ 7999	$R_4. P_1 \sim P_3: \text{Move room } 1 \sim 3 \text{ randomly.}$

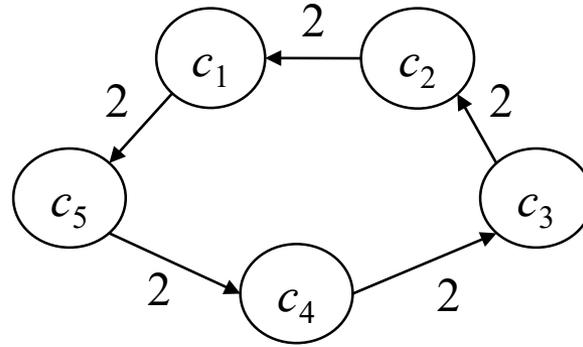
$n_1$  (the number of Phase I clusters) = 5,  $n_2$  (the number of Phase II clusters) = 4,  
 $L = 50$  (sec.),  $m = 10$ .



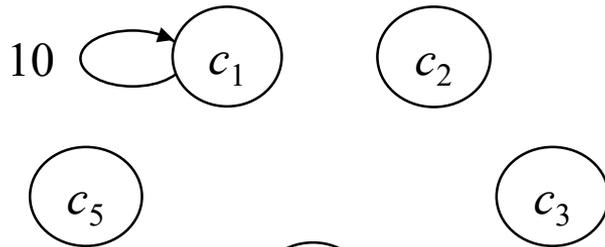
# Simple Model 1



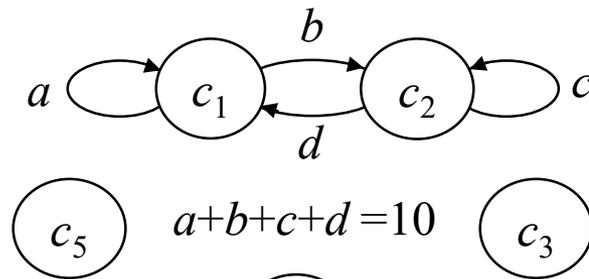
$g_1$



$g_2$



$g_3$

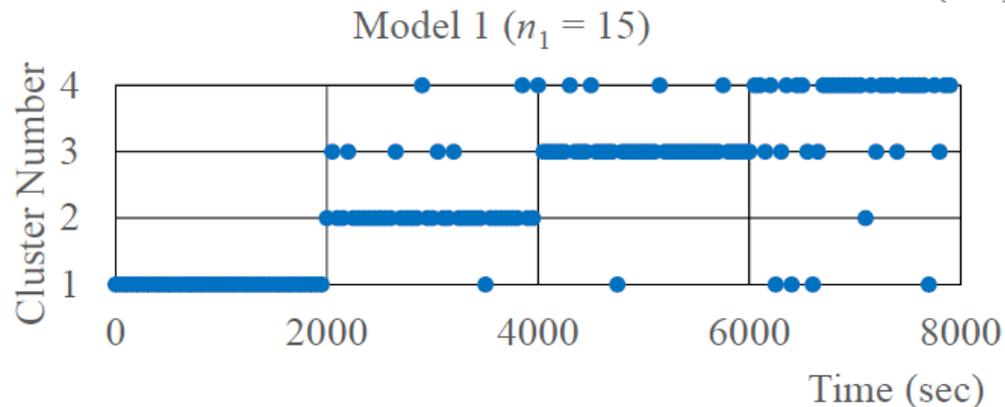
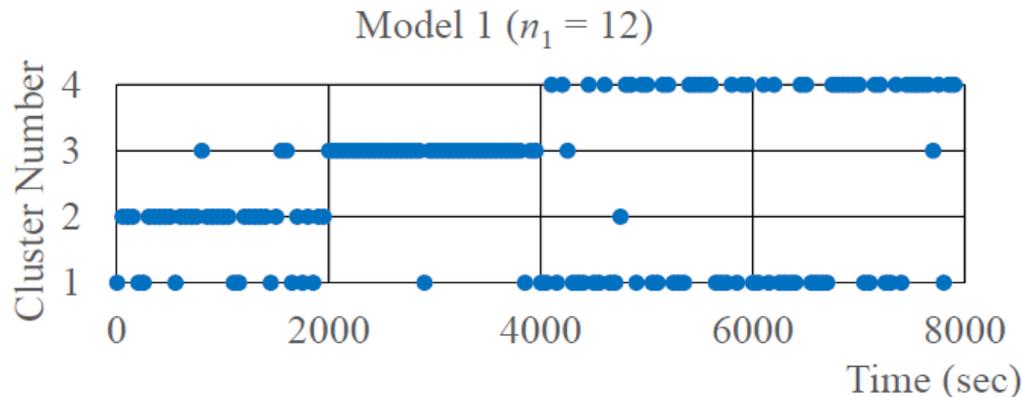


$g_4$

Scene Transition Graphs at Phase II

# Simple Model 1 with Noise

- 70% → Main rule, 30% → Other three rules  
 $n_1$  (the number of Phase I clusters)  $\geq 5$ ,  $n_2$  (the number of Phase II clusters) = 4,  
 $L = 50$  (sec.),  $m = 10$ .



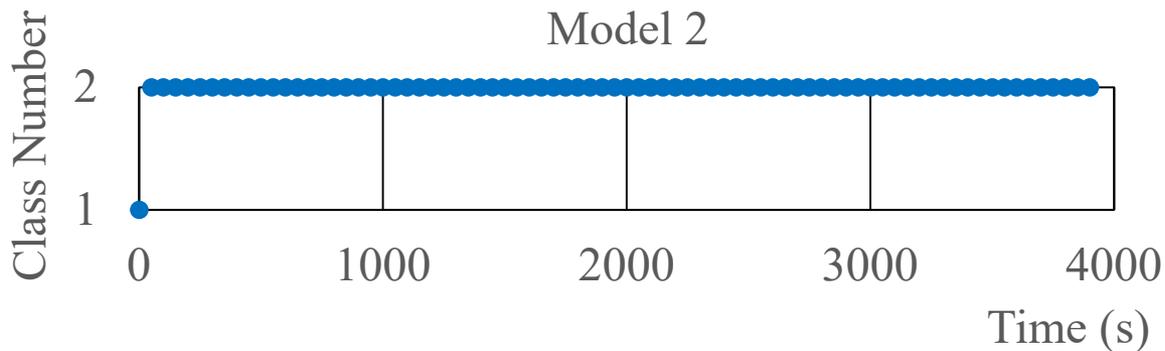
Larger  $n_1$  is needed.

# Simple Model 2

- There are five rooms 1 - 5 and three persons  $P_1, P_2, P_3$ .

Model 2	
Time(sec.)	Rules
0 ~ 1999	$P_1: 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 1 \rightarrow \dots$ $R_5. P_2: 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 1 \rightarrow 2 \rightarrow \dots$ $P_3: 3 \rightarrow 4 \rightarrow 5 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow \dots$
2000 ~ 3999	$P_1: 1 \rightarrow 3 \rightarrow 3 \rightarrow 5 \rightarrow 5 \rightarrow 2 \rightarrow \dots$ $R_6. P_2: 2 \rightarrow 2 \rightarrow 4 \rightarrow 4 \rightarrow 1 \rightarrow 1 \rightarrow \dots$ $P_3: 3 \rightarrow 4 \rightarrow 5 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow \dots$

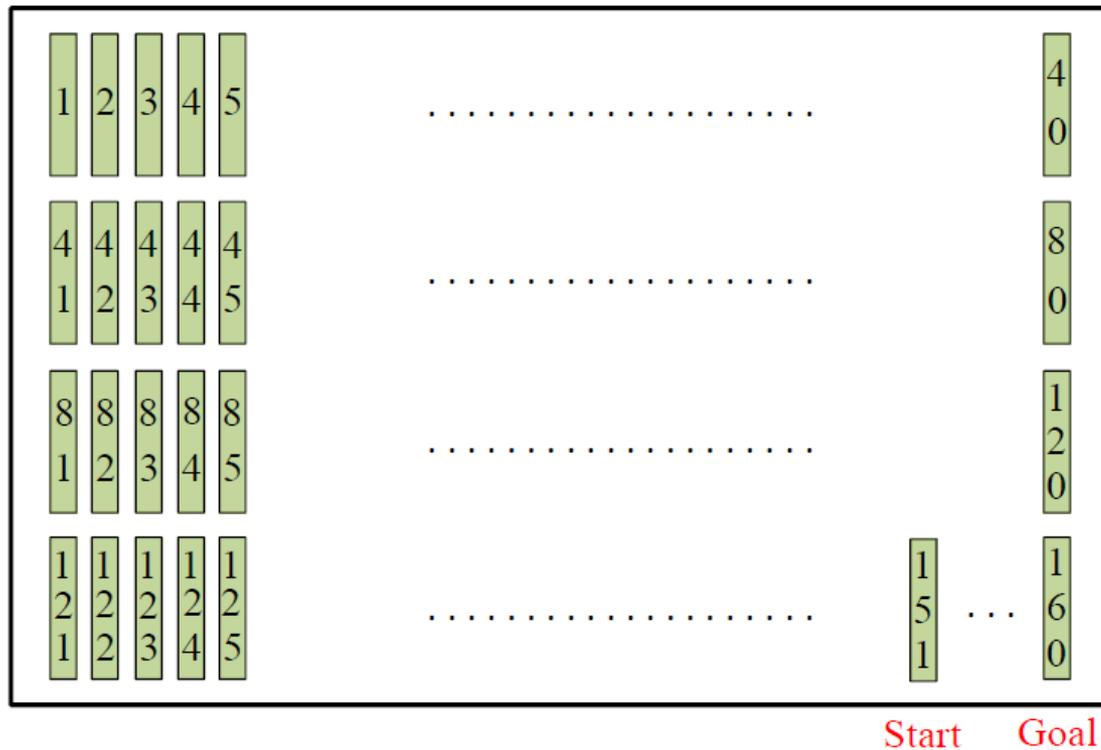
$n_1$  (the number of Phase I clusters) = 5,  $n_2$  (the number of Phase II clusters) = 2,  $L = 50$  (sec.),  $m = 10$ .



Failure. Two rules give the same time series of place vectors.

# Picking in Warehouse

- There are 30 persons working in a warehouse. At each round, each person starts from START, picks goods on shelves, and return to GOAL.



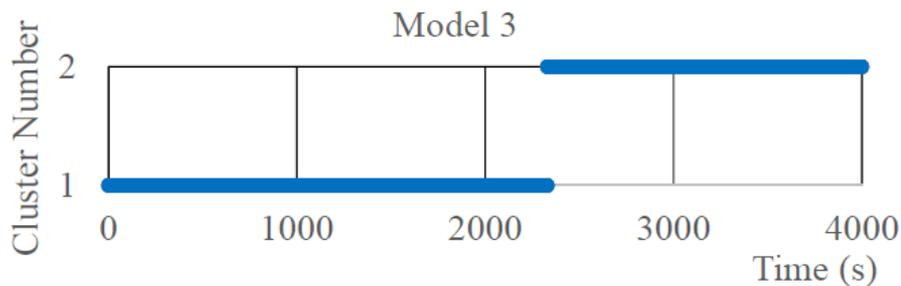
# Picking in Warehouse 1

- Rules

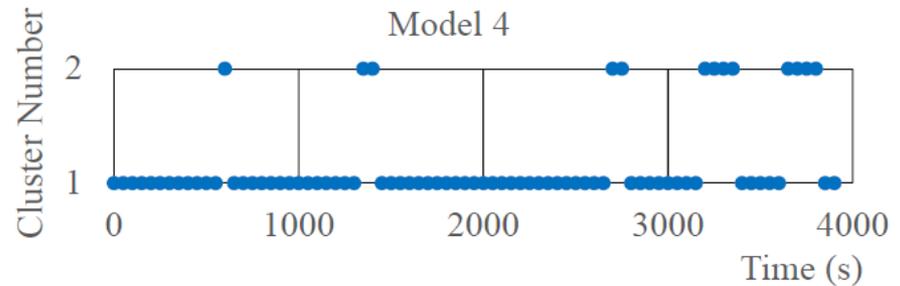
Model 3	
Time(sec.)	Rules
0 ~ 1999	$R_7$ . All: At each round, choose 5 shelves randomly from 1-80 and visit them by a random route.
2000 ~ 3999	$R_8$ . All: At each round, choose 5 shelves randomly from 81-160 and visit them by a random route.

Model 4	
Time(sec.)	Rules
0 ~ 1999	$R_9$ . All: At reach round, choose 5 shelves randomly and visit them by shortest routes.
2000 ~ 3999	$R_{10}$ . All: At reach round, choose 5 shelves randomly and visit them by random routes.

$n_1$  (the number of Phase I clusters) = 10,  $n_2$  (the number of Phase II clusters) = 2,  $L = 50$  (sec.),  $m = 10$ .



Success



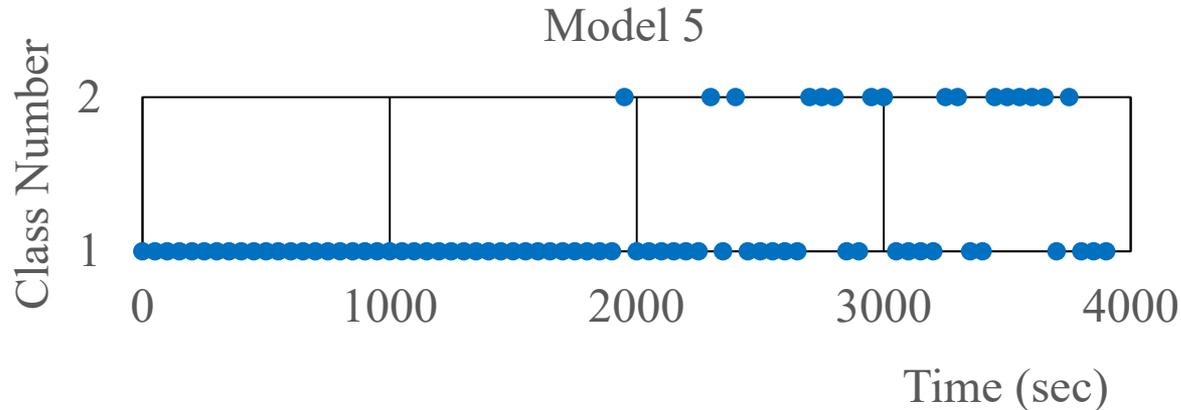
Failure

# Picking in Warehouse 2

- Rules

Model 5	
Time(sec.)	Rules
0 ~ 1999	$R_{11}$ . All: Move to 1 ~ 150 randomly.
2000 ~ 3999	$R_{12}$ . All: Move to a shelf within 3 shelves from the current location randomly.

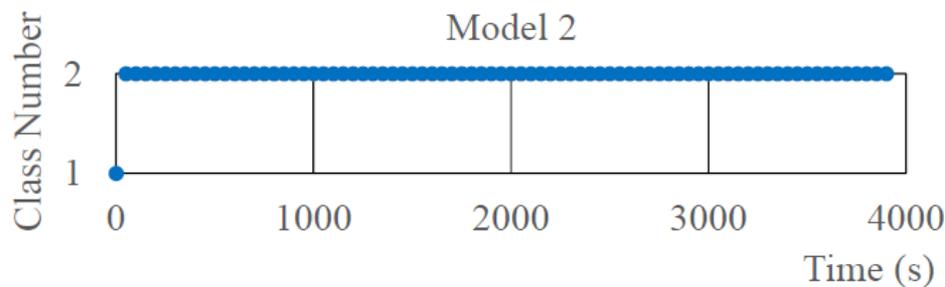
$n_1$  (the number of Phase I clusters) = 10,  $n_2$  (the number of Phase II clusters) = 2,  $L = 50$  (sec.),  $m = 10$ .



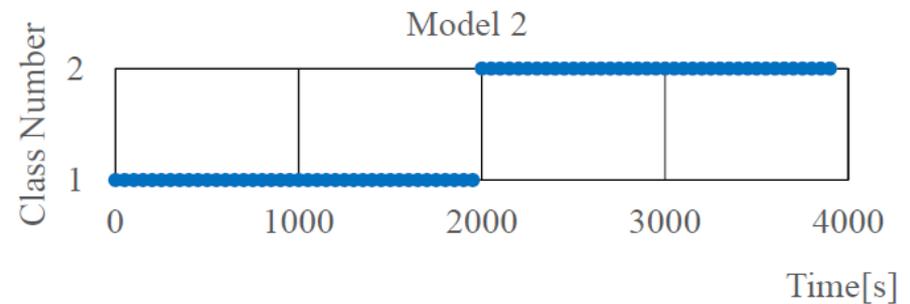
Two rules are not separated clearly, but the change point is detected.

# Improvement

- **Direction vector**  $v = [v_{ij}]$ , where  $v_{ij}$  is the number of persons that move from region  $i$  to region  $j$  if  $i \neq j$ , and the number of persons stay at region  $i$  if  $i = j$ .
- Since the dimension of the vector is  $r^2$ , where  $r$  is the number of regions, we may apply PCA(Principal Component Analysis) to the vectors.



Place vector



Direction vector

# Conclusion

- The proposed method does not work well if the action rules give similar sequences of place vectors → Direction vectors + PCA.
- Other improvements
  - Using Staff ID (= distinguishing individual persons) with data structure that is invariant for exchanging IDs.
  - Adding other statistics on movement

Thanks to Shun Hayakashi!