Proceedings of ICServ2016, pp.281-287 (2016)

Proposing an interactive label attaching system for supervised service operation estimation

Karimu Kato[1*], Takashi Okuma[2] and Takeshi Kurata[1,2]

[1] Graduate School of Systems and Information Engineering, Tsukuba University, Japan

[2] Human Informatics Research Institute, National Institute of Advanced Industrial Science and Technology, Japan

* E-mail: s1230187@u.tsukuba.ac.jp

Abstract

In this paper, we propose an interactive label attaching system that helps users make training data for a supervised learning method, named Service-Operation Estimation (SOE), which makes a human behavior observation for service improvement effective. SOE is a technique based on supervised learning for estimating staffs' Service-Operations (SOs) using various data taken in the service field. SOE learns and estimates SOs using training data that are obtained by attaching SOs to the staff data. The attaching operations are the most expensive work in our method of which purpose is cost reduction for human behavior observation. Therefore, we have been developing the interactive SO label attaching system. This system estimates the temporal bias of the attached SO labels, and allow users to prepare the training data by providing a user interface that show the calculated bias. This system is expected to contribute to reducing costs of creating training data. Through our plenary evaluation, we figured out some issues of the current implementation of our proposing system. Some of issues are caused by low accuracy of the current implementation of SOE method.

Keywords

Service-process analysis, Service Operation, Machine Learning

1 INTRODUCTION

Deep Learning is in the public limelight so that the artificial intelligence's long period of depression comes to an end. The machine learning will be the major method for data analysis. However, whenever supervised learning is used, training data have to be prepared by human. As machine learning becomes more general, the costs cannot be ignored. Service Operation Estimation (SOE) (Tenmoku, Ueoka, Makita, Shimmura, Takehara, Tamura, Hayamizu, and Kurata. 2011) is a method for estimation of Service Operations (SOs) based on supervised learning for service process improvement. It seems difficult that labeling Service Operations as training data to time series of data because we have to check various types of sensor data, order of Operations and varying definition of each SO that the cost of creating training data for SOE is higher than that for computer vision based object recognition and voice recognition.

This paper proposes an interactive label attaching system for supervised service operation estimation. SOE estimates corresponding SO label to non-learned data using discrimination functions obtained from the training data. Therefore, if training data have too biased distribution, adequate accuracy cannot be obtained. So this study focuses on work occurrence times as the primary factor in SO process variations, defines a model for parent population generation and calculates the temporal bias of SO labels by comparing it to the distribution of SO labels that have already been attached. Proposing system provides a user interface that shows the calculated bias for effective labeling. This interface will allow for efficient attaching SO labels, and is expected to contribute to reducing costs of creating training data.

2 SERVICE OPERATION ESTIMATION'S SUMMARY AND PROBLEM

As we mentioned above, SOE is defined as a technique based on supervised learning for estimating staffs SOs. The purpose of this method is the automation of the observation and description of staffs' SOs. There has been a technique for observing staffs work called 'time study'. This technique uses observation results to improve the service processes. The technique is pairing up staff with observer, and the observer notes down staff's SO while working. This technique can obtain accurate SO descriptions. However, staffs are aware of being observed. Thus, staffs feel the psychological stress. SOE is difficult to obtain accurate SO descriptions as in time studies, but SOE is no need for observers. It promises the following benefits: staffs are relieved from the psychological stress of an observer's presence; the technique's application to work which is difficult to check by observers such as service industry work; long-term measuring of staffs; and reduced costs of SO description.

SOE's training data are obtained by combining staff data and service-operations. The staff's data are consisted of the

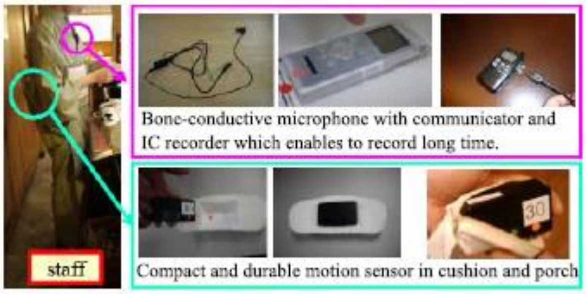

Figure 2 :A human behavior sensor, microphone and recorder. (Takehara, Kato, Tamura, Tenmoku, Kurata, and Hayamizu 2014)

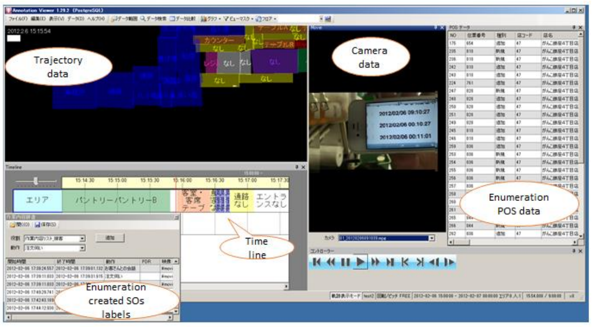

Figure 3 : The service-process analysis support system's main screen

staff's behavior measure data and the work data. This staff's data are extracted feature. SOE can use camera's data for feature extraction. However, our study considers an environment that is difficult to set the camera. Thus, our study is not used the camera's feature. Figure 2 shows the devices for use in the staff behavior measurement. There are Pedestrian Dead Reckoning (PDR), a voice recorder, and a Bone-conductive microphone for recording speech. The staff's position, orientation, action and Voice Activity Detection (VAD) are obtained through these devices. In the measurement, the services staffs wear a small PDR at the waist. We considered that they don't feel of strangeness by the small device. PDR measure a relative coordinates by barometer, magnetometer, accelerometer, gyroscope, and thermometer. The staff's orientation and position are obtained by Sensor Data Fusion (SDF) .SDF combines Active Radio Frequency Identification (active RFID)'s absolute positioning and PDR's relative coordinates. Bone-conductive microphone's voice data are recorded by the voice recorder. This device considers exclusion of customer's talk. As the work data, SOE can use the work schedule and the shift data. In this study, the work data is Point of Sale (POS) data. The ordering start time and the receiving payment start time are recorded to POS data. The staff's data are separated by optional time interval. The training data are obtained by linking it with SO labels.

SOE is based on supervised learning. Thus, this method needs to prepare SO labels. Linking the staff data with SO labels is termed 'attaching SO labels'. This has a problem in terms of personnel costs. In attaching SO labels, the analysts use the service-process analysis support system (Figure 3). This system shows the staff behavior measurement data and the work data while doing time synchronization. In attaching SO label, the analysts confirm

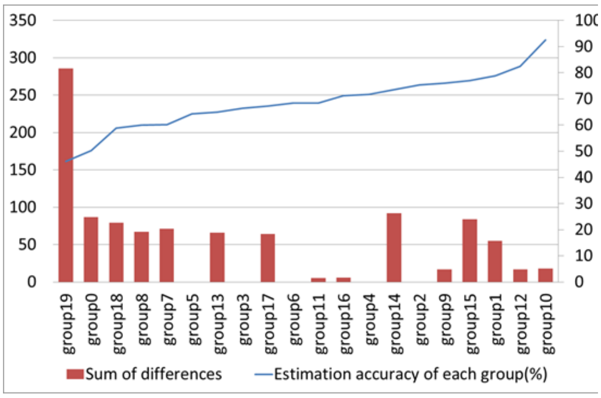

Figure 1 : Sum of differences between estimation data and training data by SO and cross-validation results

the staff data, and attach label by the information (e.g. the staff does SO_x between start time and end time.). The staff data is time-series data. Thus, the analysts need to check SOs context for attaching SO label. The fast-forward is difficult because the analysts cannot confirm the voice recorder data. Therefore, the checking staffs' SOs are high costs. Moreover, the analysts don't consider the temporal distribution bias. Thus, the improving estimation accuracy can't expect from the increasing SO labels. And, there exist SOE's unique problem. This problem is the variation of SOs process. In same SO's name, SOs process finely vary due to various reason (e.g. the location, the time, the staff, and the customer). For example, we present process of SO: "Greeting and guiding customers to the table". This process greets the customer entering from an entrance, and guides the customer to a table. However, the guiding routes vary due to each entrance location and empty seat location. If the customer is steady clientele, the process will be complex for the chat. Thus, variation of SOs processes promotes an increasingly attaching SO label's costs, and is major hurdle of creating the training data.

3 RELATION OF THE SO LABELS TEMPORAL DISTRIBUTION AND THE ESTIMATION ACCURACY

For classification, the supervised learning learns the optimal combination of the feature and threshold from the training data. If the feature is designed sufficiently, High estimation accuracy is obtained by the smallness of training data's distribution bias. The bias of training data expresses the difference between the distribution of SO labels already attached and the distribution of its parent population. If there is no bias in the distribution of SO labels, it decreases the possibility that the data subject to estimation falls outside of the training data distribution. Thus, it increases estimation accuracy of the discriminant functions learned through the training data. However, SO processes vary due to various reasons. It affects the various data, and complicates the parent set distribution. To obtain better learning results, the authors consider it important that the distribution of training data is analyzed by the place, the time, the staff, and the customer. Figure 1 shows relation of the SO label's temporal distribution bias and the estimation accuracy. For cross-validation, the training data are divided into 20 groups, and are used to learning and

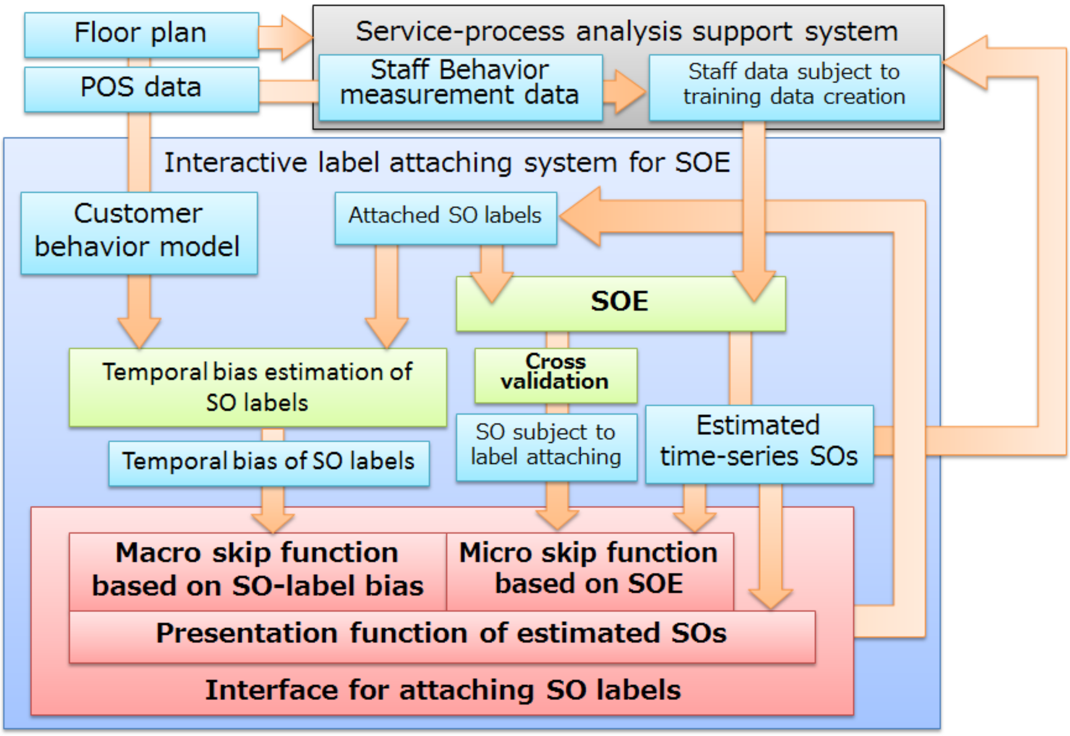

Figure 2 : Interactive label attaching system's conceptual diagram

estimation. In the result, the estimation accuracy of each group had large difference. For investigation into the cause of difference, the differences are calculated based on the number of data that are grouped by the group, SO, and the data’s hour. And, the sum of differences is totalized the differences. The differences are the number of data subject to estimation is large. If the sum of difference's value is large, the discriminant functions will not learn the difference's hour. This figure shows correlation between the sum of difference and the estimation accuracy. Thus, the smallness of training data temporal distribution bias has a relationship of increasing the estimation accuracy. In the present study, we have assumed that times of work occurrence play an important role in the variation of SO processes.

4 EFFICIENT CREATING SO LABELS

In the present study, the temporal distribution bias of SO label is calculated based on the times of work occurrence. If number of the attached SO labels is surplus in the specific hour, the attaching SO labels in the hour cannot expect increasing of the estimation. Thus, the present study develops the interactive label attaching system for SOE. This system liaises with the service-process analysis support system. By the system, the SO labels of the hour of reducing the bias are attached on a priority basis. And, the system provides efficient increasing of the estimation accuracy. This system creates the training data by the attached SO labels, estimates the staff data. When the analysts confirm the staff data, this system shows the estimated result to the analysts. Furthermore, the system combines the estimated result and the temporal distribution bias of SO labels, provides the interface for attaching SO labels.

5 PRECEDING STUDY

The method of increasing number of the labels is put to practical use widely as the system. (E.g. Microsoft Office IME in the character translation software's field, Picasa in the image management software's field, and the Auto-complete method in the search engine's field (Pereira, and Chen 2006)). These case studies have the efficient method of increasing number of the labels. This method obtains the high estimation accuracy by the creating labels of the users, feeds back to the users.

In academic field, Zhang proposed the photo management algorithm of using the facial recognition (Zhang, Chen, Li and Zhang 2003). This study developed the family album management system. This system estimates the character of the photo's faces using Bayesian Framework, and shows the character's candidate to the users. This study laid the foundation of photo management system. Goto proposed the PodCastle that is the voice information search system (Goto, and Ogata 2011). The PodCastle has the speech-to-text method by the voice recognition. The method is used to Japanese podcast in the web. The estimation's error point is attached the label by anonymous users. By this method, this system is expected to increase the estimation accuracy of voice recognition. In environment of the voice corpus insufficiency, this study challenged to practical use of the voice recognition. It was epoch-making experiment. However, in the service field, the system of efficient increasing the number of label does not exist.

6 INTERACTIVE LABEL ATTACHING SYSTEM 'S SUMMARY

The system is expected to reduce the work rate of attaching SO labels.

Figure 2 shows the conceptual diagram. The system is constructed by the three functions that are the interface for attaching SO labels, the temporal bias estimation of SO

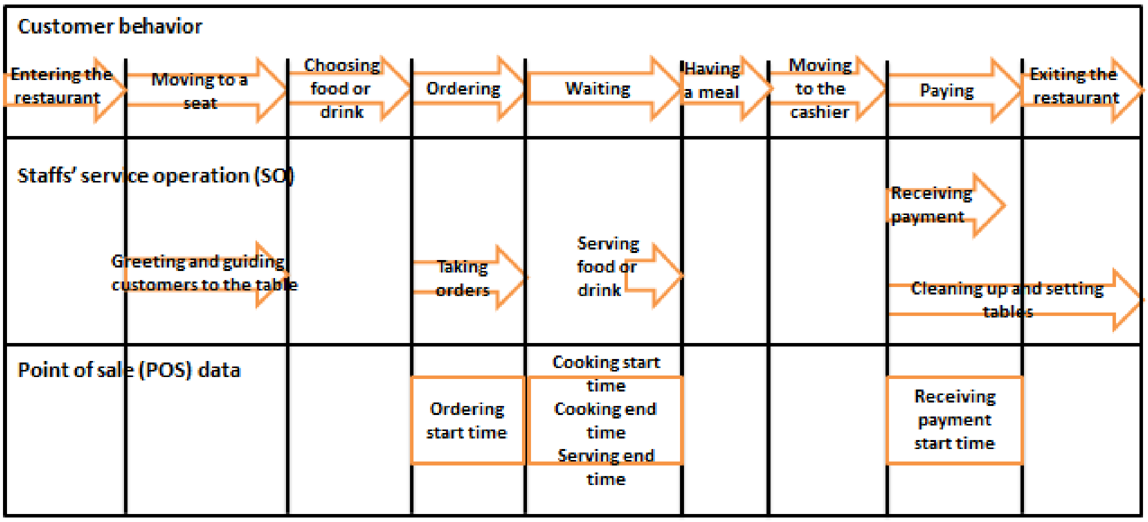

Figure 5: Customer behavior model and correlation between staff SOs and POS data

labels, and SOE. , The temporal bias estimation of SO labels calculates the temporal distribution bias of attached SO labels. This function's result is one of the indexes. SOE estimates the staff data by using attached SO labels. This data are the staff data subject to training data creation. SOE outputs the indexes that are the estimated result and the SO subject to label attaching. The interface for attaching SO labels is using the indexes of other functions. This function includes following functions. Macro skip function based on SO-label bias. Micro skip function based on SOE. Presentation function of estimated SOs.

7 TEMPORAL BIAS ESTIMATION OF SO LABELS

Calculating the SO labels temporal bias needs the parent populations of the temporal distribution. For creating the parent set distribution, we defined the customer behavior model, and created the SOs event probability distribution by it. We assumed that the staff SOs and POS data works closely with the customer's behavior in the restaurant. Figure 5 shows relation between the staff's SOs, POS data, and the customer's behavior. The several staffs work in the restaurant. Thus, it's not always true that the staffs' SOs are exercised by the staff subject to training data creation.

The SO event probability distributions are based on the POS data's ordering start time and the POS data's receiving payment start time. Figure 6 shows the daytime's model of the greeting and guiding customers to the table. This SO happens before the ordering start time. Thus, this SO's start time and end time are calculated by combining the ordering average time, the quality criticism time, and the traveling time between a seat and an entrance. The ordering average time uses average time of the attached SO labels. We conducted the Q&A to the business manager. Thus, we defined the quality criticism time as the day time (3 min) and the night time (5min). The traveling time is calculated by combining the distances and the walking speed. The distances between the seats and the entrances are calculated by POS data's table number and floor plan. The walking speed is 1.6m/s. We referenced the paper of IATSS (Tsujimura, Nagayama, Nagamachi, Takizawa, Suzuki, Morita, and Nakai 1979). The SO event probability

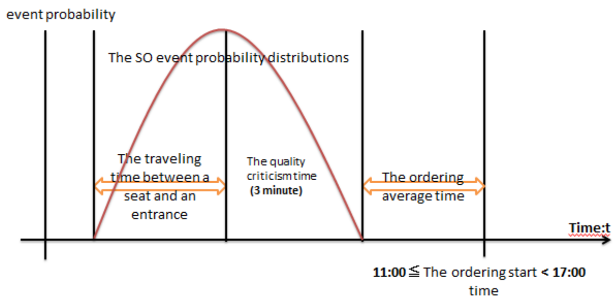

Figure 6 : The SO event probability distributions (The greeting and guiding customers to the table of the daytime)

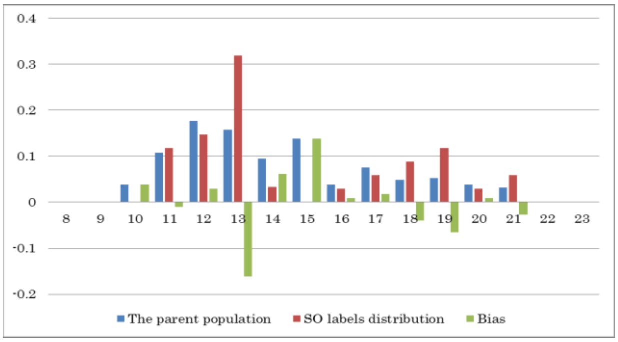

Figure 7 : The SO labels temporal bias estimation (The greeting and guiding customers to the table of the daytime)

distributions are defined by the above elements. The parent population is created by combining POS data and it. And, the temporal bias of SO labels is calculated by differential between the parent population and the temporal distribution of attached SO labels.

Figure 7 shows the calculated result of temporal bias of SO labels. This figure shows the temporal bias of "the greeting and guiding customers to the table". This figure shows a state where the attached labels are focused in thirteen. Moreover, by comparing the parent population, this figure shows that the attached labels in fifteen are the fewest.

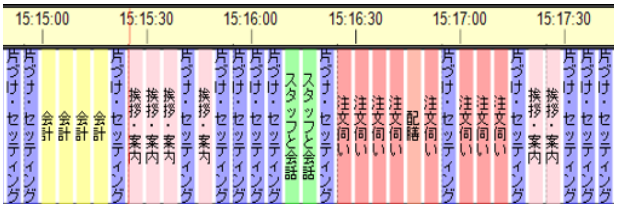

Figure 8 : The showing estimated SOs function

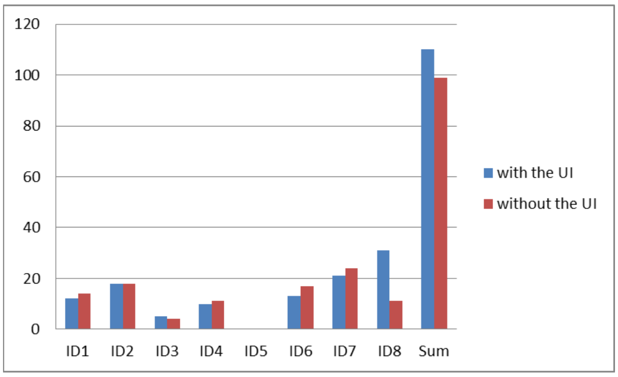

Figure 9 : The accumulated number of the created SOs labels

8 INTERFACE FOR ATTACHING SO LABELS

This interface provides the skip function that uses the three indexes. The skip functions narrow down the SO and the hour. Thus, this interface enables efficient checking of the staff's SOs. First, the users select the SO by SOE's cross-validation result. Next, using the SO's temporal bias , the users skip to the large bias's hour. It is macro skip function based on the SO-label bias. Last, the users finely skip using the SO's estimated result and the large bias's hour. It is micro skip function based on SOE. Figure 8 shows presentation function of estimated SOs. The users skip while checking it. The red vertical line is the current time that is showed in the service-process analysis support system. The skip functions are executed by the registered

Table 1 : The SOs definitions

ID	SO
1	**Dialogue with customers**
2	**Dialogue with other staff**
3	**Greeting and offering customers to the table**
4	**Moving and carrying something**
5	**Accounting**
6	**Taking orders**
7	**Serving food or drink**
8	**Cleaning up and setting tables**

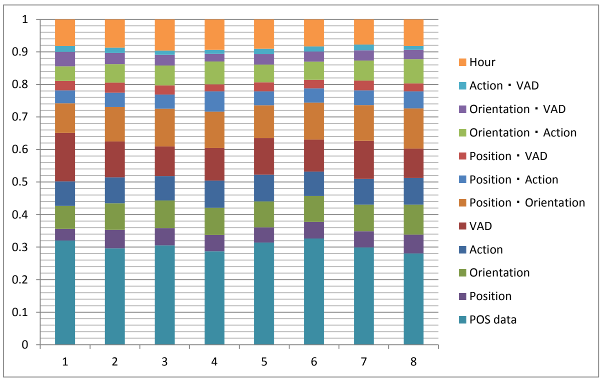

Figure 10: Contribution ratio of each category data

macro key, enable skipping to the SO subject to label attaching. In this figure, the first target SO's hour is between the current time and the 15:15:50. The second target SO's hour is between 15:17:20 and 15:17:30. This interface is expected that the users avoid checking between 15:15:50 and 15:17:20.

9 PRELIMINARY EXPERIMENT

For accurate usability evaluation of the system, the user interface should have the optimal implementation. In order to raise a usability of whole system, it is important that detail implementations are determined through usability tests with the situation which imitates the real use. In this section, we describe a preliminary experiment which has been conducted for improving implementations of our proposed user interface design. We compared attaching operation by two subjects using user interfaces with and without the simple implementation of our design in order to make it clear how our method works well. The table 2 shows the combination of the task condition. The order of the experimental task was counterbalanced. To reduce the differences of the attaching results that are caused by the subjectivity of the subjects, before they start the attaching task, they had conducted a pre-training in which they were explained the definition of each SO (Table 1) and see the attached SO labels in the training data created beforehand for 30 minutes.

In order to acquire data as much as possible under the consideration of load of the subjects, the time for the experimental task was set to 90 minutes at once. Between two tasks, subjects took a break. Target SO is set to the “Greeting and Offering Customers to the Table (ID 3).” After they finished the tasks, we conducted interviews about the usability of the user interfaces. The following is the summary of the data that is used for the preliminary experiment;

- The details of the staff's data
  - ➢ Measurement Field: a Japanese restaurant

Table 2 : the combination of experiment's condition

	A	B	Confirming hours	SO
1st	with the UI	without the UI	12:00~13:00	ID 3
2nd	without the UI	with the UI	13:00~14:00	ID 3

- Measurement Period: February 4, 2011 09:56:00 ~ 23:47:00
- Role of the measured staffs: waitresses

10 EVALUATION AND DISCUSSION

Figure 9 shows the comparison result of the number of newly attached SO label with and without the proposed user interface. No statistical significant difference was observed (Mann-Whitney U test, p=0.3248 > 0.05). Subjects reported there were too many unreliable estimated SO labels for the skep function, so that they could not use it enough. Then, we analyses the accuracy of the SOE that is implemented in the test user interface using random forest method. We use the true or false matrix shown in the Table 3 for the analysis. The numbers 1 to 8 indicated in the table as the SO labels and estimated result are corresponding to the IDs in the Table 1 that shows the definitions of each SO. The diagonal components of the matrix indicate the number of correctly estimated SOs, so that we can consider the accuracy is high when these values are large relative to the other values in the matrix.

From these results, we could not figure out if the proposed interface contributes to the increase of the number of the attached SO labels for the training data. The main reason why is the subjects could not use the skip function appropriately because the accuracy of the implemented SOE in the test interface was insufficient.

In other words, we figured out that the micro skip function could not work appropriately under the condition in which the precision of the SOE is 62.14%, which is calculated using data from this preliminary experiment. Therefore, we have to determine the required precision for the user

Table 3: The true or false matrix

		SO labels								Precision (%)
		1	2	3	4	5	6	7	8	
Estimated result	1	158	1	2	0	0	50	20	0	68.40
	2	9	592	96	13	18	77	24	54	67.04
	3	2	81	709	25	46	47	76	155	62.14
	4	0	2	23	155	0	27	14	124	44.93
	5	2	18	0	0	358	10	42	6	82.11
	6	79	52	66	9	8	866	33	85	72.29
	7	58	41	59	37	11	95	1211	214	70.16
	8	9	92	259	143	11	108	152	1806	70.00
Recall(%)		49.84	67.35	58.40	40.58	79.20	67.66	77.04	73.90	68.56

Table 4: The difference of matrix

		SO labels							
		1	2	3	4	5	6	7	8
Estimated result	1	7	3	−1	0	0	8	7	0
	2	20	84	12	7	−7	3	57	95
	3	1	−5	−28	−8	−13	21	15	116
	4	0	4	−13	20	0	−1	−2	9
	5	−2	0	0	0	15	4	7	36
	6	−7	21	57	31	7	254	103	68
	7	4	7	2	−9	1	98	39	57
	8	1	48	26	−11	18	−6	47	2

Table 5: Accuracy of each SO and the difference

SOID	without the newly added training data	with the newly added training data	difference
1	57.66	55.37	−2.30
2	67.20	61.59	−5.60
3	60.21	54.28	−5.93
4	42.64	45.22	2.58
5	80.63	76.99	−3.64
6	69.90	66.02	−3.88
7	73.44	66.31	−7.13
8	71.89	65.37	−6.53

interface in our future work. Meanwhile, we cannot guarantee the required precision during the number of the training data is insufficient, so we have to consider new interface design for the effective SO labeling for such condition.

In addition, subjects reported an issue that they lost the context of the behavior of the staff when they used the skip functions. Even the context information is important for attaching each SO label to data in particular duration, the interface skips to the start time of the estimated SO directly. We have to improve the skip function to show the context, for example, using fast-forward play of the staffs' data.

In this study, we mainly focus on the user interface design and not on the improvement the accuracy of the SOE itself. However, our interface design aimed to accelerate the efficiency of the SO labelling operation as increase the number of attached SO labels and training data with incremental learning, so that the user interfaces includes the function that feedbacks estimated results. Therefore, we can predict it easily that the accuracy of the estimation affects to the efficiency of the label attaching operation. We cannot ignore the accuracy of the SOE using the created training data with our proposed user interface. So we compared the accuracy of the SOE with and without the training data created in the preliminary experiment as additional training data to the data prepared beforehand. Basically, we only add the training data that two subjects attached same SO labels for each time slot. Table 5 shows the change of the accuracy without and with the newly added training data. In the preliminary experiment, accuracy for all SO labels expect for "Moving and carrying something (ID 4)" reduced. Table 4 shows the difference of matrix. In many cases, "Cleaning up and Setting Tables (ID 8)" is estimated as "Greeting and Offering Customers to the Table (ID 3)" and "Serving Food and Drinks (ID 7)" and "Taking Orders (ID 6)" are mixed up each other. These estimation errors cause the decrease of the accuracy.

From interviews, we found out that subjects attached the "Cleaning up and Setting Tables" label to the scene in which they can imagine staff is taking dishes away while customers are sitting at the table. Actually, it is reasonable that staffs were chatting with the customers during they taking the dishes away in such case. However, the training data prepared beforehand only include the scene in which they have no conversation. Therefore, these kind of "Cleaning up and Setting Tables" scenes are labeled to "Greeting and Offering Customers" in cross validation.

In addition, "Taking Orders" label is strongly affected by presence of the POS data. Figure 10 shows contribution ratio of each category data for identifying each label and indicates contribution ratio of the POS data is very high for identifying "Taking Orders" scene. However, in actual scene, "Taking Orders" scene without POS data occurs often. In this plenary experiment, because this kind of "Taking Order" scene is newly added to the training data, the total accuracy of the SOE was decreased.

Even these are the issues of SOE algorithms, we have figured out that the accuracy of the SOE is more strongly affect the effectiveness of the labelling operations than we assumed. In future, we have to evaluate our user interface after the SOE algorithm improved in order to show the effectiveness of our method.

11 CONCLUSIONS AND FUTURE WORKS

In this study, we have been developing an interactive label attaching system for creating training data of SOE. We designed the user interface that have skip functions based on the estimated result of the supervised SOE. Through our plenary evaluation, we figured out some issues of the current implementation of our proposing system. Some of issues are caused by low accuracy of the current implementation of SOE method. Therefore, the future works are the following; we will determine the required precision for the user interface that the user can use this skip function appropriately; We also have to improve the skip function to show the context, for example, using fast-forward play of the staffs' data. After these improvement, we will evaluate our user interface for each function and total effectiveness of our interactive label attaching system.

References

Tenmoku, R., Ueoka, R., Makita, K., Shimmura, T., Takehara, M., Tamura, S., Hayamizu, S., and Kurata,T. (2011). Service-Operation Estimation in a Japanese Restaurant Using Multi-Sensor and POS Data. Proceeding of International Conference Advances in Production management Systems, vol.3-4,1

Takehara, M., Kato, K., Tamura, S., Tenmoku, R., Kurata, T., and Hayamizu, S. (2014). Improvement of service-operation estimation using voice activity detection of employee's speech, IEICE, vol.J97-D No.10, pp.1563-1571.

Pereira, J., and Chen, L. (2006). Server-deployed cache list management for presenting an auto-complete list, US, Patent US20060242109 A1.

Zhang, L., Chen, L., Li, M., and Zhang, H. (2003). Automated annotation of human faces in family albums, Proceedings of the 11th ACM International Conference on Multimedia, pp. 716–723.

Goto, M., and Ogata, J. (2011). PodCastle: Recent advances of a spoken document retrieval service improved by anonymous user contributions, Proceedings of the 12th Annual Conference of the International Speech Communication Association, pp.3073-3076.

Tsujimura, A., Nagayama, Y., Nagamachi, M., Takizawa, K., Suzuki, H., Morita, T., and Nakai, M. (1979). Regional Characteristics and Driver Behavior –Indexing Social Speed-, Proceedings of International Association of Traffic and Safety Sciences, Vol.5 No.4, pp.231-24