Abstract

Recently, assist technologies for work staff using ICT devices is expected to be introduced into various real fields. Data science technology that collects a large amount of data from sensors and recognizes the situation in the real world plays an important role in such assist systems. Smart Voice Messaging (SVM) system has been developed as an assist system for caregiving service. SVM system comprises smartphones with application software (SVM terminals), server PC, and Bluetooth beacons located in the field. SVM system assists collaborative work such as nursing care with hands free.

As a previous research, Spatio-Temporal Situation Recognition (STSR) was proposed in a nursing home using SVM system. STSR is a technology to recognize the work situation using time series data of staff positions. In this research, behavior logs of multiple caregivers were used, and STSR for a group was proposed as a new technology. During the experiments, caregivers operate SVM terminals (smartphones), and position and movement information of caregivers are measured by Bluetooth beacons put in the tenant's room, staff station, and stairs. Regarding recognition of work situation, features are extracted from the obtained behavior logs. Then, two-phase clustering is applied to the data in order to obtain time series of class numbers, and assign care contents, such as meal care and excretion care, to each class. Then, we obtain transition of care contents. In phase 1, clustering is performed on Place Vectors obtained by extracting features from the behavior logs. Each Place Vector is a vector representing a multiple caregiver positions and movements. In phase 2, clustering is performed on the adjacency matrix generated from the result of phase 1 clustering. Regarding the assignment of care contents to each class, first the class number of phase 1 that constitutes each class number of phase 2 is identified from the adjacency matrix. Next, the movement status of the area is confirmed from the center vector of the phase 1 class. Finally, care contents are assigned by combining the information on the time at which each class of phase 2 appears and the center vector.

In the previous research, STSR technology is not sufficiently verified because the amount of experiment data was small. In addition, although the estimation of care contents was able to distinguished in working time zones, the contents itself had to be estimated only from the main work place. Therefore, in the proposal research, we first aim to verify the effect of STSR technology. We also discover situations where the effect of STSR technology will be small, and try to improve the method so that it can be effective even in such situations. In addition, regarding the experiment at the nursing home, we verify whether more accurate estimation of care contents is possible by using information other than work place.

As means for achieving the above objectives, regarding verification and improvement of the effect of STSR technology, we use a discrete event simulator that generates behavior logs of

work staff whose behavior rules are prearranged. The simulator is developed using Python's discrete-event simulation library simpy (simulation python). Regarding the verification of STSR technology, we generate the data simulating a simple model having 5 areas and 3 staffs, and an actual site that models picking work in Amazon's huge warehouse. The simulation data is generated so that the behavior rules of the staff changes according to time. We conclude STSR is effective if it is possible to detect changes in the behavior rules by applying STSR to simulation data. There are several cases in which change in the behavior rules could not be detected by the above simple model. Regarding improvement of STSR technology, it is necessary to add information to the Place Vectors or change the method of extracting features in order to handle such cases. Regarding verifying whether more accurate estimation of care contents, we use the caregiver's tweet data obtained by SVM system. We first specify keywords that indicate each care content. Labels are attached to the tweet data if the data contains the keywords. We obtain transitions of care contents by labeling, compare this results with the results of phase 2 clustering, and assign each care to the phase 2 clustering class.

We have the following results as the verification of STSR technology, in the simple model, change in the behavior rules can be detected when each behavior rule are set so that the movement range is different or the movement direction is different. The reason why it can be detected is that there is differences in the Place Vectors and the adjacency matrices. Therefore, change in the behavior rules could not be detected when each behavior rule is set so that the Place Vectors or the adjacency matrices are equal. In the case of actual site that models picking work in Amazon's huge warehouse, the change can be detected when the behavior rules gives different ranges of movement under the simple model. As a result of improvement of STSR technology, the change can be detected in the case where the Place Vectors and the adjacency matrices are equal in the simple model as described above. As an improvement, we propose to add an element representing staff movement information as a new element to the Place Vector in order to make a differences in the vectors. We can solve the problem of indistinguishable the Place Vectors by generating and clustering vectors that combines staff's departure and destination information without generating the Place Vectors from the behavior logs. As a result of verifying whether more accurate estimation of care contents, depending on the contents of tweet, there were cases in which we cannot obtain the information leading to estimation of care contents, even by comparing the results of labeling tweet with the result of clustering. However, we confirm care contents can be estimated more accurately because the amount of information increases compared to the method using only the center vector and the time at which each class appears.

On the clustering phase, we study the setting of number of cluster, effect of phase 2 clustering, and evaluation of clustering result. We also study noise resistance and dimensional reduction using the principal component analysis (PCA) for the proposal method. Regarding the

setting of number of clusters was estimated from care contents in the previous research. We try to use a method that automatically determine the number of clusters using the height of dendrogram. Regarding the effect of phase 2 clustering, we observe phase 2 clustering is effective when it is not known whether change in behavior rules can be classified correctly by looking at the dendrogram of phase 1 clustering. Regarding the evaluation of the clustering result, we compare the noise added clustering result with the normal result using DBIndex for evaluation. Regarding the noise immunity of the proposed method, noise that staff sometimes behaves without following the behavior rules is added to the proposed method and the clustering result is evaluated using DBIndex. Regarding the dimensional reduction for the proposed method, dimension of the vector combining information of a departure and a destination is reduced using PCA. For PCA, Python's machine learning library scikit-learn is used.

As a future task, this time, since we could not conduct experiments only by simulation, it is necessary to conduct experiments at nursing homes, etc., and to demonstrate the improvement of STSR technology and the estimation of work contents combining tweet data.