

Title	作業スタッフの行動ログを用いた時空間状況認識の性能向上について
Author(s)	早貸, 舜
Citation	
Issue Date	2018-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/15440
Rights	
Description	Supervisor:平石 邦彦, 先端科学技術研究科, 修士 (情報科学)

修士論文

作業スタッフの行動ログを用いた時空間状況認識の性能向上について

1610149 早貸 舜

主指導教員 平石 邦彦
審査委員主査 平石 邦彦
審査委員 内平 直志
緒方 和博
上原 隆平

北陸先端科学技術大学院大学

先端科学技術研究科[情報科学]

平成 30 年 2 月

概要

近年、情報技術の発達により、生産、医療、メンテナンスなどの様々な分野において、ICT デバイスを用いた作業スタッフのアシスト技術の導入が期待されている。この時、データサイエンスの技術は重要であり、我々の開発した SVM(Smart Voice Messaging)システムが有効であると考えている。SVM システムはターミナルとなるアプリケーションがインストールされたスマートフォン、つぶやき内容と話者の位置情報を蓄積するサーバ、位置情報を取得するための Bluetooth ビーコンから構成される。このシステムは介護などの共同作業のアシストをハンズフリーで実現可能である。

本研究の先行研究は上記の SVM システムを用いて、介護現場にて複数の介護士の位置情報を含む行動ログから作業状況を認識する技術を開発したが、使用できる実験データの量が少ないため、認識技術の十分な検証が行われていないという問題があった。また、作業内容の推定は作業時間帯で分けはできていたが、内容自体は主たる作業場所から推定するしかなかった。そこで、本研究では、あらかじめ行動のルールを決められた作業スタッフの行動ログを生成する離散事象シミュレータを作成し、シミュレーションデータを用いることで認識技術の効果を検証する。また、認識技術の効果が小さくなるような状況を発見し、そのような状況でも効果を発揮できるよう改善する。さらに、先行研究の介護現場での実験について、介護士のつぶやきデータを従来の情報と組み合わせることでケア内容の推定精度が向上するか検証する。

シミュレーションによる検証の結果、作業によってスタッフが行動する範囲が異なる場合には、認識技術が有効であることが確認できた。また、認識技術の効果が小さくなる場合でも、行動ログから特徴抽出後に情報を加える、もしくは先行研究とは別の特徴抽出方法を用いることで改善できたことを確認した。つぶやきデータを用いた作業内容の特定については、つぶやき方にも依存はするが、行動ログのみを用いるよりも情報量が増えるため効果はあることが確認できた。

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.

目次

第1章	はじめに	1
1.1	背景.....	1
1.2	目的.....	1
1.3	構成.....	1
第2章	関連研究	3
2.1	アプリケーション.....	3
2.1.1	顧客行動解析.....	3
2.1.2	流動支援制御.....	3
2.1.3	つぶやきのテキストマイニング.....	3
2.1.4	看護業務行動の推定.....	4
2.2	データマイニング手法.....	4
2.2.1	外れ値検知.....	4
2.2.2	アノマリ検知.....	4
2.2.3	バースト解析.....	4
2.2.4	つぶやきデータの自動タグ付け.....	5

第3章 先行研究および時空間状況認識	6
3.1 先行研究	6
3.2 時空間状況認識	7
第4章 従来法の検証	10
4.1 離散事象シミュレータ	10
4.2 簡単なモデルを用いたシミュレーション	10
4.3 Amazon の巨大倉庫でのピッキングを模したシミュレーション	14
第5章 提案法による改善	22
第6章 作業場所の情報とつぶやきデータを組み合わせたケア内容推定の検証	28
第7章 考察	36
7.1 クラスタリングのクラスタ数	36
7.2 フェーズ2 クラスタリングを行う意味.....	36
7.3 提案法のノイズ耐性	41
7.4 クラスタリング結果の評価	43
7.5 提案法におけるベクトルの次元増加に対しての主成分分析による次元削減	43
第8章 むすび	47

第1章 はじめに

1.1 背景

近年、情報技術の発達により、生産、医療、メンテナンスなどの様々な分野において、ICTデバイスを用いた作業スタッフのアシスト技術の導入が期待されている。このとき、センサなどを用いて大量のデータを集め、実世界の状況を認識するデータサイエンス技術は重要である。本研究の先行研究として、本研究グループが開発した SVM (Smart Voice Messaging) システム[1][2]を用いて、介護現場における時空間状況認識を行った研究[3]がある。SVM システムはターミナルとなるアプリケーションがインストールされたスマートフォン、つぶやき内容と話者の位置情報を蓄積するサーバ、位置情報を取得するための Bluetooth ビーコンから構成される。この研究では、複数の介護士の位置情報を含む行動ログから作業状況を認識する技術を開発しており、グループに対する時空間状況認識という新しい技術を提案した。しかし、使用できる実験データの量が少ないため、認識技術の十分な検証が行われていないという問題があった。また、作業内容(介護のケア内容)の推定は作業時間帯で区分けはできていたが、内容自体は主たる作業場所から推定するしかなかった。

1.2 目的

本研究の目的は、問題点であったデータ量の少なさを克服し、時空間状況認識手法の効果を検証することである。また、当該手法では行動ログの推定精度が低くなるような状況を発見し、そのような状況に対しても、高い精度で状況推定が可能ないように当該手法を改良することである。さらに、先行研究の介護現場での実験について、介護のケア内容を作業場所以外の情報も用いてより精度高く推定できるか検証する。

上記の目的を達成する手段として、時空間状況認識手法の検証および改善については、あらかじめ行動のルールを決められた作業スタッフの行動ログを生成する離散事象シミュレータを作成し、シミュレーションデータを先行研究の当該手法に適用することで行う。また、介護のケア内容の推定については、SVM システムで取得した介護士のつぶやきデータを作業場所以外の情報として作業場所の情報と組み合わせて推定に役立つか検証する。

1.3 構成

本論文は以下のように構成される。

- 2章：関連研究の調査
- 3章：先行研究および時空間状況認識の紹介
- 4章：シミュレーションデータを用いた従来法の検証
- 5章：シミュレーションデータを用いた改善方法の提案
- 6章：作業場所の情報とつぶやきデータを組み合わせたケア内容推定の検証

- 7章：4, 5章の考察
- 8章：本論文のまとめ

第2章 関連研究

本章では、本研究に関連する研究を紹介する。

2.1 アプリケーション

2.1.1 顧客行動解析[4]

人間をステレオカメラで捉え、動線を抽出し、動線の時間変化を混合分布で表現することにより、人流データの長期間保存を可能にする方法を提案し、実際に大型複合施設において得られる1年単位の動線を比較・可視化する実験で提案法の有効性を確認した。

本研究との類似点は人の動きをセンサで捉え特徴量を抽出しクラスタリングしている点で、相違点はデータの長期保存を目的とした手法であり、作業状況を認識するものではない点である。

2.1.2 流動支援制御[5]

大規模群衆流動の制御に向けた総合支援サービスとして、見物客に自発的に混雑緩和に協力してもらうための情報配信、見物客の計数や経路選択を調査するための人流計測を用いて、数万人規模の来場者が集まる関門海峡花火大会で実験を行った。実験で取得したデータにより、群衆流動シミュレータを用いた実際の群衆流動のモデル化の正しさを評価し、より現実に近い群衆流動シミュレーションを実装可能にした。

本研究との類似点は見物客の位置情報を得ている点である。相違点は群衆が行動する際の状況把握、群衆流動のモデル化のためのデータ取得が目的であり、介護施設や倉庫内でのピッキングなどの作業とは種類が異なる点である。

2.1.3 つぶやきのテキストマイニング[6]

2011年3月の東日本大震災発生前日から同年4月4日までにTwitterで投稿されたつぶやきに対して、Information Value (IV)と呼ばれる統計量を用いて震災の発生によりバースト(2.2.3項参照)したと考えられる単語を機械的に抽出した。その後バースト状態となった単語を震災に関連するものとししないものに分類するために単純ベイズ分類器を用いる方法を提案した。また、キーワードを基に被災地の現状をリアルタイムに把握することが可能であることを示した。さらに、つぶやきの分類において、単純ベイズ分類器がSVM(Support Vector Machine)と同程度の分類能力があることを検証実験により実証した。

IV はつぶやきに出現する単語の相対的な出現頻度, その単語が東日本大震災の発生と因果関係があるかを考慮可能な統計量と考えられており, Weight Of Evidence (WOE)を Kullback and Leibler により拡張されたものである.

本研究との類似点はつぶやきデータに含まれるキーワードを用いて状況を把握する点である. 相違点はつぶやきデータが音声認識で得たものではなく Twitter 上のものである点, キーワードを機械的に抽出する手法であり, つぶやき自体のタグ付けは行わない点である.

2.1.4 看護業務行動の推定[7]

センサデータと業務知識を用いた業務推定手法を提案し, 実際に病院での実験を通じて提案手法の有効性を評価した. 提案手法ではセンサデータから看護師の基本行動(乗り物, 走行, 静止, 歩行)を教師あり学習で推定した. その後, 業務行動(食事介助や会議など)を教師なし学習で推定した. 業務行動の推定にはトピックモデルを適用することでセンサデータをトピックに分類し, 業務の順番などの業務知識を用いることにより得られたトピックに業務を割り当てた.

本研究との類似点はセンサで得た情報を基に作業状況の推移を取得している点である. 相違点は作業状況の作業内容の推定に事前知識を用いている点である,

2.2 データマイニング手法

2.2.1 外れ値検知[8]

時系列データやデータストリームの中から類似度計測およびクラスタリングやk近傍法などを用いて普段は起こり得ない値を検知する手法

本研究の時空間状況認識手法との類似点は個人の異常行動に対しての検知などに用いることができる点である. 相違点は本研究の手法は集団への異常検知に用いられるという点である.

2.2.2 アノマリ検知[9]

データ通信などにおいて, 正常な通信やデータのパターンを定義しておき, そのパターンから外れたものがあれば異常として検知する手法

本研究の手法との類似点, 相違点は上の外れ値検知と同様である.

2.2.3 バースト解析[10]

[6]で用いられていた手法. 大規模なイベント等の発生により, ある時点を境に特定のトピックについての記述が急激に増加する現象をバーストと呼ぶ. [10]においては, バースト度というパラメータを設け, 時系列に沿った各キーワードのバースト度の変化やバーストしているかどうかの判定, バースト度によるキーワードのランク付けを

する手法

本研究ではつぶやきデータのラベル付けに、事前にラベル毎にキーワードを設定し、キーワードに該当した場合ラベル付けを行う手法を用いた。そのため、想定されるキーワードが大量で事前にキーワードを設定できない場合や[6]のようにキーワードそのものを抽出したい場合などに有効であると考えられる。

2.2.4 つぶやきデータの自動タグ付け[11]

[1][2]で紹介した SVM システムを用いている。つぶやいた人の位置情報、つぶやきの内容、つぶやいた人に装着している加速度計からの情報が含まれるつぶやきデータに対して、エスティメーターとソルバを用いて自動でタグ付けする。エスティメーターは3種類あり、位置情報から部屋番号、つぶやき内容から患者の名前やケアの名前、加速度情報からケアの状況などのタグを信頼度付きで生成する。ソルバはサーバ上にあり、業務知識を用いて最適なタグのセットを検索する。

つぶやき内容からのタグ付けにだけ着目した場合は、患者の名前やケアの名前といった特定の情報に対してタグを作成するため、本研究のつぶやきデータのラベル付けの手法と類似している。相違点は本研究では位置情報は時空間状況認識の中で用いているが、加速度情報は用いていない点である。

第3章 先行研究および時空間状況認識

本章では、1章で述べた先行研究とその研究で用いられた時空間状況認識について述べる。

3.1 先行研究[3]

介護現場において、複数の介護士の移動情報を含む行動ログを取得する実験を行い、行動ログに時空間状況認識を適用することで作業状況の認識を行った。この研究では、複数の介護士の位置情報を含む行動ログから作業状況を認識する技術を開発しており、グループに対する時空間状況認識という新しい技術を提案した。実験は介護施設を図3-1のように10区画に分け、入居者の部屋やスタッフステーション、階段にBluetoothビーコンを設置し、介護士にICTデバイス(スマートフォン)を装着してもらうことにより位置と移動の情報などを取得した。

行動ログの解析フローを図3-2に示す。取得した行動ログから特徴抽出を行い、その後2回のフェーズに分けてクラスタリングをすることにより、時系列のクラス番号を取得した。さらに、各クラスに食事ケアや排泄ケアなどのケア内容を当てはめ、ケア内容の推移を取得する。フェーズ1クラスタリングでは、行動ログから特徴抽出をして得たプレイスペクトルに対してクラスタリングを行う。プレイスペクトルとは介護士の行動ログから位置と移動を特徴として抽出したもので、複数の介護士の位置と移動を表すベクトルである。ベクトルの要素数はエリア数分あり、各要素が各エリアに対応し、エリアにいる介護士の数を表す。例えば、エリア数が10、介護士が6人で、3人の介護士がエリア4、2人の介護士がエリア5、1人の介護士がエリア2から5へ移動しているとき、プレイスペクトルは以下ようになる。

Place vector : [0, 0.5, 0, 3, 2.5, 0, 0, 0, 0, 0]

フェーズ2クラスタリングでは、フェーズ1クラスタリングの結果から隣接行列を生成しクラスタリングを行う。隣接行列はインターバルとサブインターバルを設定し、サブインターバル毎に最頻出するクラス番号のエッジに対応する要素に値を加えることで生成する。例えば、インターバルが9秒、サブインターバルが3秒、フェーズ1クラスタリングの結果が表3-1のとき、隣接行列は以下の通り、有向荷重グラフは図3-3のようになる。

Adjacency Matrix : [0, 1, 0, 1]

クラスタリングは、フェーズ 1, 2 共にベクトル間の差をユークリッド距離で定義し、ワード法でクラスタリングしている。また、クラスタ数はデンドログラムを見て、介護のケア内容の種類を基にクラスタ数がいくつ必要か予想し決定している。

最後の各クラスにケア内容をあてはめる際は、まず、フェーズ 2 の各クラス番号を構成するフェーズ 1 のクラス番号を隣接行列から判断する。次にそれらのフェーズ 1 の中心ベクトルから介護士のエリアの移動状況を確認する。最後に、フェーズ 2 の各クラスが現れる時間の情報と前述の中心ベクトルの情報を組み合わせ、既知のケア内容（食事のケア、排泄処理のケア、口腔ケアなど）がどのクラスに該当するかを考えてあてはめる。

3.2 時空間状況認識

時空間状況認識とは、作業をしているスタッフの位置などの時系列データを用いて、作業状況を認識する技術である。集団に対しての時空間状況認識のイメージは図 3-4 であり、ある空間の中でスタッフが集団で働いていて、スタッフの位置が時間によって変わるときを考える。スタッフの作業状況を認識するには空間の中でスタッフがどう配置されているか、時間とともに配置がどう変わるかを考える必要がある。この技術は上記の介護現場における介護作業状況の認識[3]などで応用され、類似した手法が大型複合施設における顧客行動解析[4]やイベントにおける流動制御支援[5]などで用いられている。

時空間状況認識の手順は以下の通りである。

1. 実世界に存在する情報を、センサなどを用いて集める。
2. その情報から特徴を抽出する。
3. 抽出した特徴に対してクラスタリングを用いて解析する。
4. クラスタリングされたデータの各クラスに既知の作業状況に当てはめる。

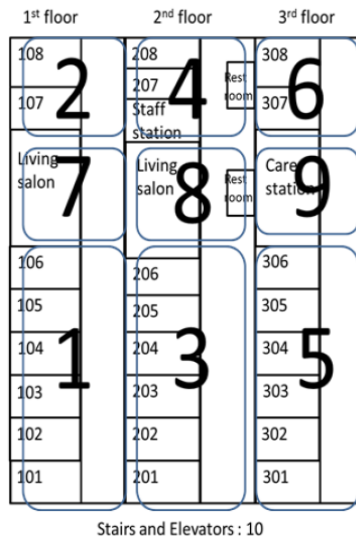


図 3-1. 介護施設の区画設定図

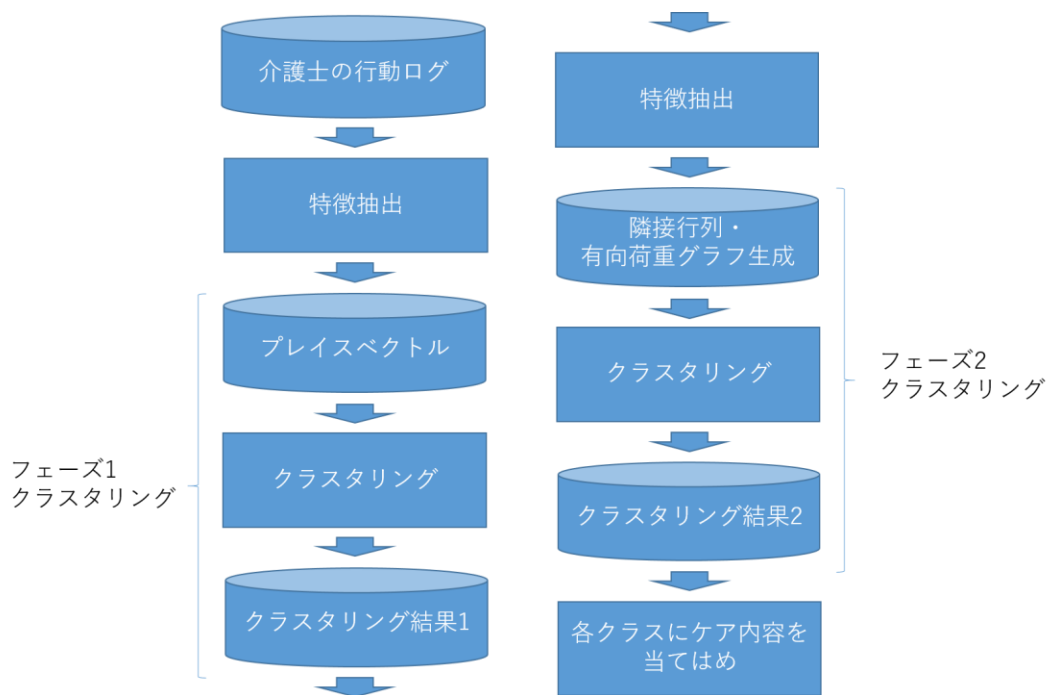


図 3-2. 先行研究の時空間状況認識の処理フロー図

表 3-1. フェーズ 1 クラスタリング結果のサンプル

Time[s]	Class Number
1	1
2	1
3	1
4	1
5	2
6	2
7	2
8	2
9	2

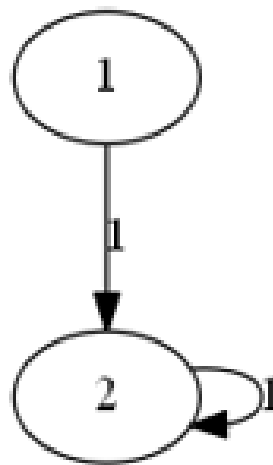


図 3-3. 有向荷重グラフのサンプル

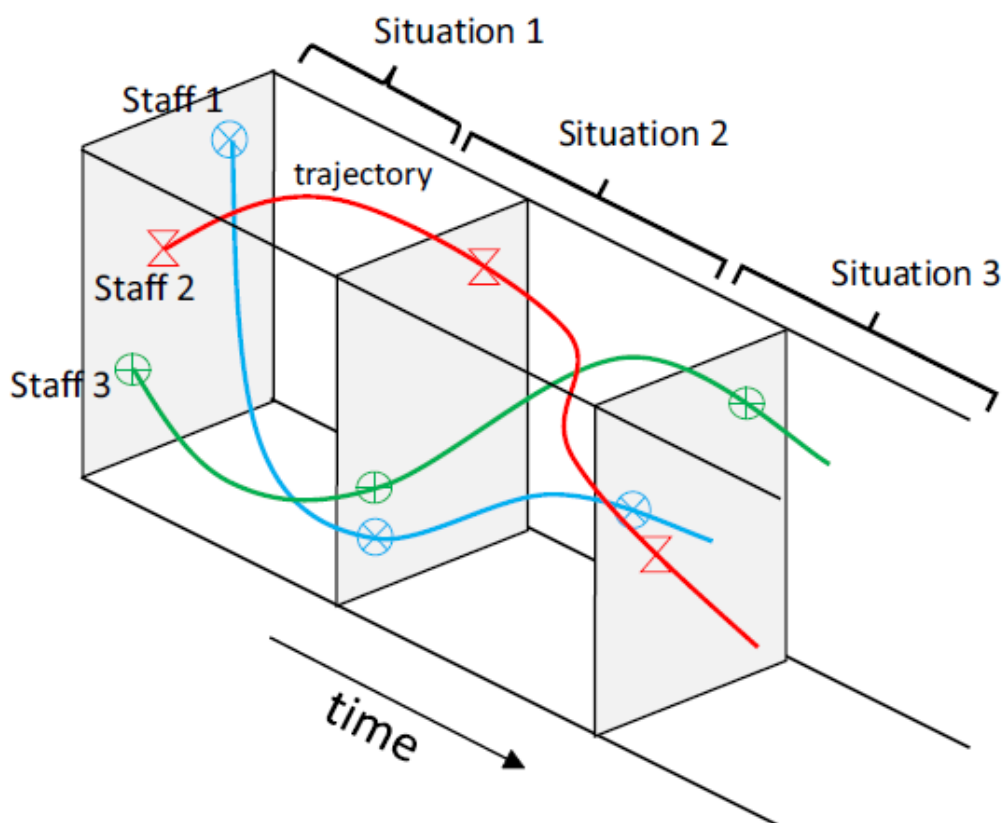


図 3-4. 集団に対しての時空間状況認識のイメージ([3]より引用)

第4章 従来法の検証

本章では、離散事象シミュレータを用いた従来法の検証について述べる。検証は、時間によってスタッフの行動ルールが変わるように離散事象シミュレータを用いてシミュレーションデータを作成し、3章で述べた時空間状況認識手法を適用することで行う。この時クラスタリングの結果からルールの変化が切り替わるタイミングを正しく検知できていれば、従来法の効果あり、検知できなければなしとする。

4.1 離散事象シミュレータ

本研究を行うにあたり、エリア数やスタッフ数などのパラメータ、スタッフの行動ルールを決めることで、プレイベクトルなどの特徴抽出後の行動ログを出力可能なシミュレータを Python で作成した。作成には Python の離散事象シミュレーションパッケージである SimPy を用いた。SimPy を用いることで容易に単位時間毎の各スタッフの移動情報を得ることができる。

4.2 簡単なモデルを用いたシミュレーション

本節では、いくつかの条件を設定し、従来法の検証を行った結果を示す。スタッフの行動ルールは時間によって変わり、クラスタリングにより、ルールが切り替わるタイミングが検知できれば、従来法の効果あり、検知できなければなしとする。

- 条件 1

図 4-1 のように 5 つのエリアがあり、スタッフが各エリア間を移動する状況を想定する。シミュレーションの各パラメータを表 4-1 にまとめ、各スタッフの行動ルールを表 4-2 にまとめる。また、このシミュレーションのフェーズ 2 クラスタリングの結果を図 4-2 に示す。図 4-2 より、表 4-2 で設定した行動ルール毎にクラス分けされていることが確認できる。

- 条件 2

エリアマップは図 4-1 と同じ、各パラメータ、行動ルールを表 4-3、4-4 にまとめ、図 4-3 にフェーズ 2 クラスタリングの結果を示す。図 4-3 より、表 4-4 で設定した行動ルール毎にクラス分けされていないことが確認できる。表 4-4 よりプレイベクトルは

$[0.5, 1, 1, 0.5, 0], [0, 0.5, 1, 1, 0.5], [0.5, 0, 0.5, 1, 1], [1, 0.5, 0, 0.5, 0.5], [1, 1, 0.5, 0, 0.5]$

の 5 パターンの繰り返しであるため、これをクラスタリングしたフェーズ 1 の結果に対して、隣接行列を生成してもすべて同じ値になってしまい、クラス分けができない。

- 条件 3

エリアマップは図 4-1 から 2 エリア減らし、3 エリアとした状況を想定する。各パラメータ、行動ルールを表 4-5、4-6 にまとめる。ここで、フェーズ 2 以降のクラスタリングを行わないのは、表 4-6 からすべての時間においてプレイスペクトルの値が等しくなることが分かるので、クラス分けできないことが自明なためである。

条件 1~3 の結果より、隣接行列およびプレイスペクトルが等しくなるような状況下においては、従来法は効果を発揮しないことが確認できた。続く 4.3 節で、隣接行列およびプレイスペクトルが異なるような実際の環境に近い状況を想定し、従来法の効果および重要性を検証する。

表 4-1. 条件 1 のパラメータ

エリア数	5 区画
スタッフ数	3 人
移動に要する時間 (一律)	5 秒
待機に要する時間	5 秒
フェーズ 1 クラスタ数	5
フェーズ 2 クラスタ数	4
隣接行列生成インターバル	50 秒
隣接行列生成サブインターバル	5 秒

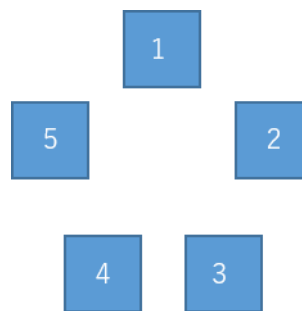


図 4-1. 条件 1, 2 のエリアマップ

表 4-2. 条件 1 のスタッフの行動ルール

スタッフ 1~3	
0~1999[s]	エリア番号が大きくなるように移動する (1→2→3→4→5→1→…)
2000~3999[s]	エリア番号が小さくなるように移動する (1→5→4→3→2→1→…)
4000~5999[s]	エリア 1~5 から行き先をランダムで選んで移動する (既にいるエリアと同じエリアが行き先となった場合は待機)
6000~7999[s]	エリア 1~3 から行き先をランダムで選んで移動する

La_Sm_Al_Li

P1_C5_P2_S5_C4_M5W5

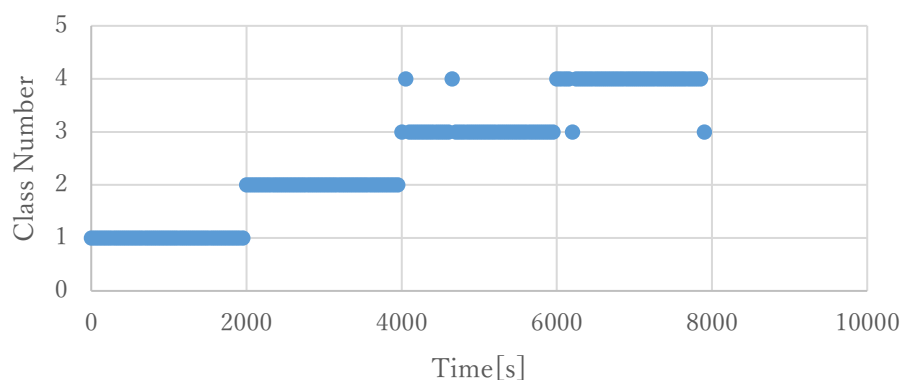


図 4-2. 条件 1 のフェーズ 2 クラスタリング結果

表 4-3. 条件 2 のパラメータ

エリア数	5 区画
スタッフ数	3 人
移動に要する時間 (一律)	5 秒
待機に要する時間	5 秒
フェーズ 1 クラスタ数	5
フェーズ 2 クラスタ数	2
隣接行列生成インターバル	50 秒
隣接行列生成サブインターバル	5 秒

表 4-4. 条件 2 のスタッフの行動ルール

	スタッフ 1	スタッフ 2	スタッフ 3
0~1999[s]	エリア番号が大きくなるように移動する (1→2→3→4→5→1→…)	エリア番号が大きくなるように移動する (2→3→4→5→1→2→…)	エリア番号が大きくなるように移動する (3→4→5→1→2→3→…)
2000~3999[s]	エリア番号が大きくなるように 2 エリア移動し, 1 度待機を繰り返す (1→3→3→5→5→2→…)	エリア番号が大きくなるように 2 エリア移動し, 1 度待機を繰り返す (2→2→4→4→1→1→…)	エリア番号が大きくなるように移動する (3→4→5→1→2→3→…)

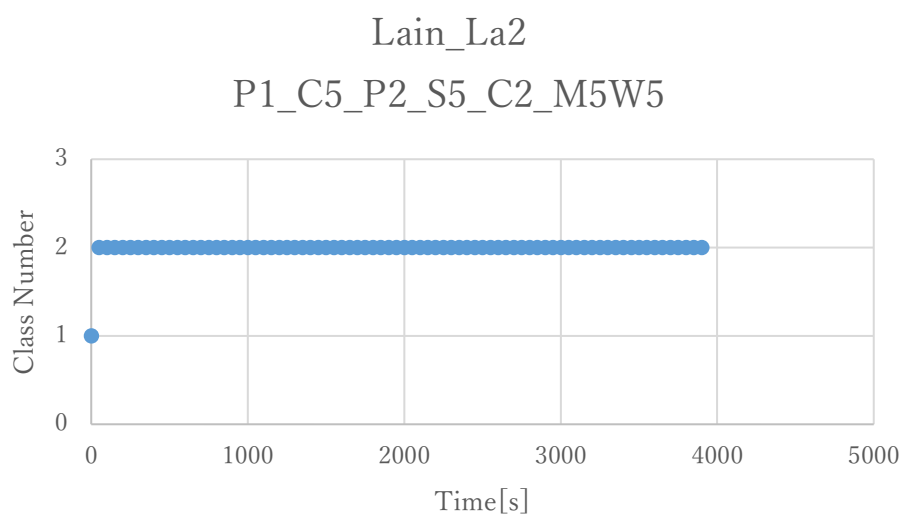


図 4-3. 条件 2 のフェーズ 2 クラスタリング結果

表 4-5. 条件 3 のパラメータ

エリア数	3 区画
スタッフ数	3 人
移動に要する時間 (一律)	5 秒

表 4-6. 条件 3 のスタッフの行動ルール

	スタッフ 1	スタッフ 2	スタッフ 3
0~1999[s]	エリア番号が大きく なるように移動する (1→2→3→1→…)	エリア番号が大きく なるように移動する (2→3→1→2→…)	エリア番号が大きく なるように移動する (3→1→2→3→…)
2000~3999[s]	エリア番号が小さく なるように移動する (1→3→2→1→…)	エリア番号が小さく なるように移動する (2→1→3→2→…)	エリア番号が小さく なるように移動する (3→2→1→3→…)

4.3 Amazon の巨大倉庫でのピッキングを模したシミュレーション

本節も前節と同様、いくつか条件を設定し、スタッフの行動ルールの切り替わるタイミングをクラスタリングにより検知できれば従来法は効果あり、なければ効果なしとする。

- 条件 4

図 4-4 のようにエリアマップを想定し、スタッフが各エリア間を移動する状況を想定する。シミュレーションの各パラメータを表 4-7 にまとめ、各スタッフの行動ルールを表 4-8 にまとめる。また、このシミュレーションのフェーズ 2 クラスタリングの結果を図 4-5 に示す。図 4-5 より、表 4-8 で設定した行動ルール毎にクラス分けされていることが確認できる。切り替わるタイミングが 2000 秒より少し後ろにずれているのは、条件 4 では、1 エリア移動するのに 5 秒かかり、前のルールの移動を終えてから、次の行動ルールで動くという設定にしたためである。

- 条件 5

エリアマップ、各パラメータは条件 4 と共通、行動ルールを表 4-9 にまとめ、図 4-6 にフェーズ 2 クラスタリングの結果を示す。図 4-6 より表 4-9 で設定した行動ルール毎にクラス分けされていないことが確認できる。これは、2000 秒以降はスタッフが 50% の確率で行動しないが、移動範囲がルール毎に変わらないため、プレイスペクトルの差が小さいことがクラス分けできない原因であると考えられる。ただし、スタッフの行動しない確率を上げてゆけば 2000 秒以降のプレイスペクトルが同じクラスに分類されるようになるため、クラス分けは可能であると推測できる。(詳細は 7.2 節参照)

- 条件 6, 7

図 4-7 のように図 4-4 にピッキングのスタートとゴールの地点を追加した状況を想定する。条件 6, 7 の各パラメータは共通で、表 4-10 にまとめる。条件 6 の行動ルールを表 4-11 にまとめ、図 4-8 にフェーズ 2 クラスタリングの結果を示す。また、条件 7 の行動ルールを表 4-12 にまとめ、図 4-9 にフェーズ 2 クラスタリングの結果を示す。図 4-8, 4-9 より表 4-11, 4-12 で設定した行動ルール毎にクラス分けされていないことが確認できる。これは柵を回る順番が異なっても、移動範囲がルール毎に変わらないため、プレイスペクトルの差が小さいことがクラス分けできない原因であると考えられる。

- 条件 8

エリアマップは図 4-7 と同じ、各パラメータ、行動ルールを表 4-13, 4-14 にまとめ、図 4-10 にフェーズ 1 クラスタリングの結果を示す。図 4-10 より表 4-14 で設定した行動ルール毎にクラス分けされていることが確認できる。切り替わるタイミングが 2000 秒より少し後ろにずれている理由は条件 4 と同じである。ここで、フェーズ 2 クラスタリングを行わないのは、後の 7.2 節で詳しく述べるが、フェーズ 1 クラスタリングのデンドログラムを見て、2 クラスに分類できると判断したからである。

条件 4~8 の結果より、プレイスペクトルおよび隣接行列が異なっても、移動範囲が同じである場合には従来法の効果が小さいことが確認できた。また、従来法の重要性については、先行研究の舞台である介護現場のように、ケア内容が時間によって変わり、移動範囲が異なる状況であれば異常検知などに貢献したが、今回のピッキング作業のような時間によって移動範囲が変わらないと想定される状況では、異常検知やピッキングチームの成績評価などに用いるのは難しいと考えられる。対応策として、プレイスペクトルに 1 回のピッキングでスタッフが移動したエリア数の合計を追加する、もしくはプレイスペクトルとは別に移動したエリア数の合計を行動ルールの切り替わりを検知するために用いれば、条件 6, 7 には有効であると考えている。

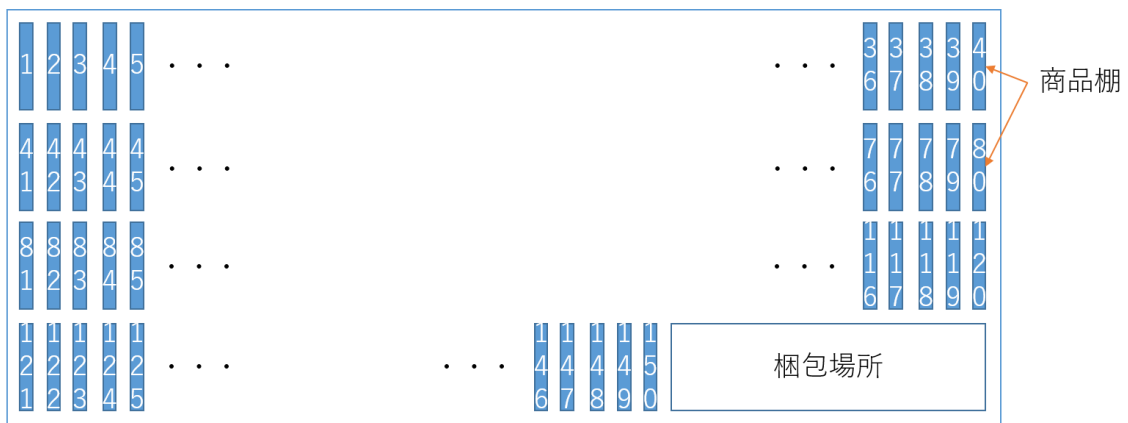


図 4-4. 条件 4, 5 のエリアマップ

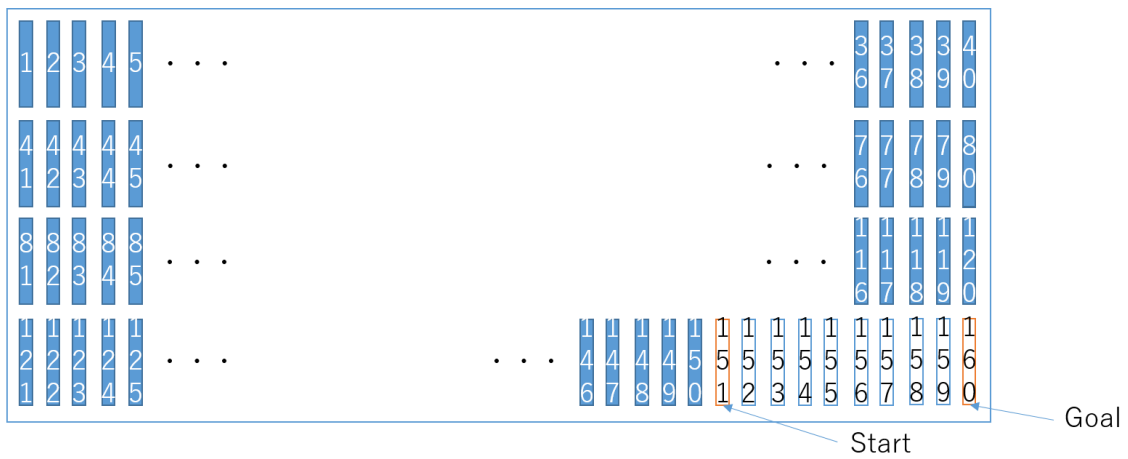


図 4-7. 条件 6, 7, 8 のエリアマップ

表 4-7. 条件 4, 5 のパラメータ

エリア数	150 区画
スタッフ数	30 人
移動に要する時間 (1 エリア毎)	5 秒
待機に要する時間	5 秒
フェーズ 1 クラスタ数	10
フェーズ 2 クラスタ数	2
隣接行列生成インターバル	50 秒
隣接行列生成サブインターバル	5 秒

表 4-8. 条件 4 のスタッフの行動ルール

スタッフ 1~30	
0~1999[s]	エリア 1~150 から行き先をランダムで選んで移動する
2000~3999[s]	1 つ前にいたエリアから 3 エリア以内のエリアをランダムで選んで移動する

All_Limit_moverange3

P1_C10_P2_S5_C2_M5W5

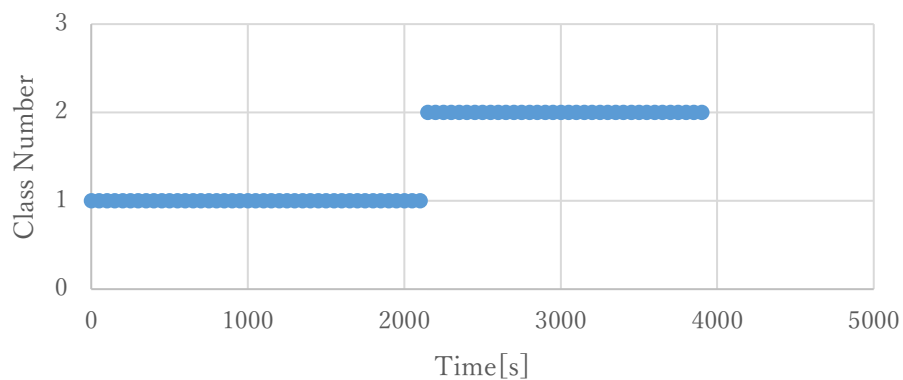


図 4-5. 条件 4 のフェーズ 2 クラスタリング結果

表 4-9. 条件 5 のスタッフの行動ルール

スタッフ 1~30	
0~1999[s]	エリア 1~150 から行き先をランダムで選んで移動する
2000~3999[s]	エリア 1~150 から行き先をランダムで選んで移動する ただし、50%の確率でスタッフは行動しない (現在いるエリアと同じエリアを行き先を選ぶ)

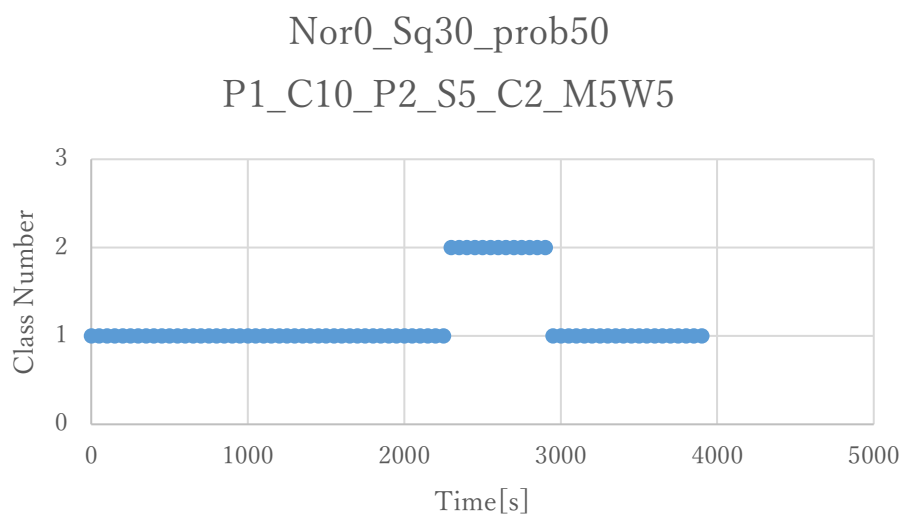


図 4-6. 条件 5 のフェーズ 2 クラスタリング結果

表 4-10. 条件 6, 7 のパラメータ

エリア数	150 区画
スタッフ数	30 人
移動に要する時間 (1 エリア毎)	5 秒
待機に要する時間	5 秒
フェーズ 1 クラスタ数	10
フェーズ 2 クラスタ数	2
隣接行列生成インターバル	50 秒
隣接行列生成サブインターバル	5 秒
一度のピッキングで回る棚の数	5 か所

表 4-11. 条件 6 のスタッフの行動ルール

スタッフ 1~30	
0~1999[s]	エリア 1~150 から一度のピッキングで回る棚の数分だけ行き先をランダムで選び、スタート地点→棚→ゴール地点が最短経路になるように移動する
2000~3999[s]	エリア 1~150 から一度のピッキングで回る棚の数分だけ行き先をランダムで選び、棚を回る順番はランダムに決め、スタート地点→棚→ゴール地点と移動する

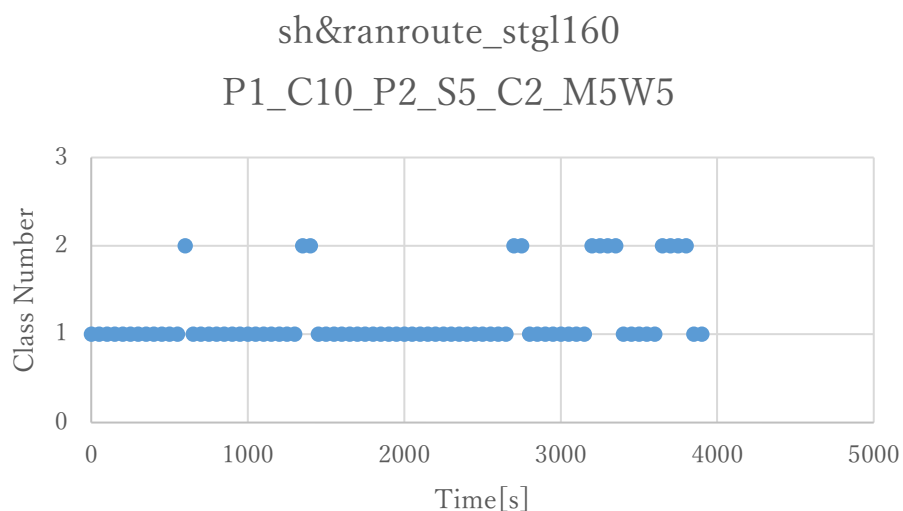


図 4-8. 条件 6 のフェーズ 2 クラスタリング結果

表 4-12. 条件 7 のスタッフの行動ルール

スタッフ 1~30	
0~1999[s]	エリア 1~150 から一度のピックアップで回る棚の数分だけ行き先をランダムで選び、棚を回る順番は棚番号昇順とし、スタート地点→棚→ゴール地点と移動する
2000~3999[s]	エリア 1~150 から一度のピックアップで回る棚の数分だけ行き先をランダムで選び、棚を回る順番は棚番号降順とし、スタート地点→棚→ゴール地点と移動する

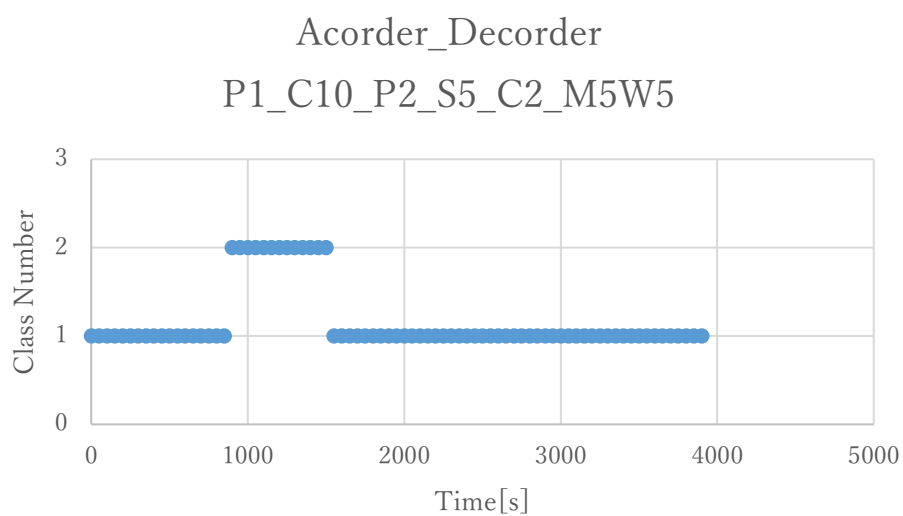


図 4-9. 条件 7 のフェーズ 2 クラスタリング結果

表 4-13. 条件 8 のパラメータ

エリア数	150 区画
スタッフ数	30 人
移動に要する時間 (1 エリア毎)	5 秒
待機に要する時間	5 秒
フェーズ 1 クラスタ数	2
一度のピッキングで回る棚の数	5 か所

表 4-14. 条件 8 のスタッフの行動ルール

スタッフ 1~30	
0~1999[s]	エリア 1~80 から一度のピッキングで回る棚の数分だけ行き先をランダムで選び、棚を回る順番はランダムに決め、スタート地点→棚→ゴール地点と移動する
2000~3999[s]	エリア 81~150 から一度のピッキングで回る棚の数分だけ行き先をランダムで選び、棚を回る順番はランダムに決め、スタート地点→棚→ゴール地点と移動する

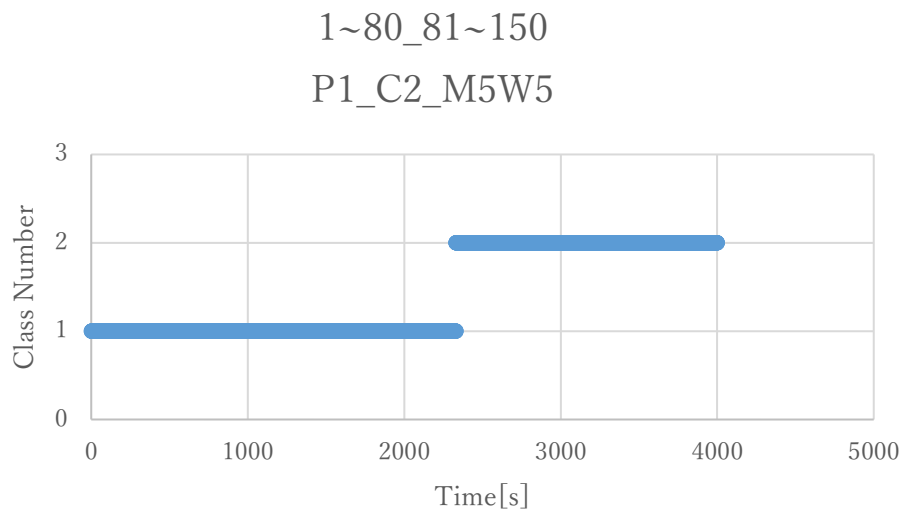


図 4-10. 条件 8 のフェーズ 1 クラスタリング結果

第5章 提案法による改善

本章では4.2節で取り上げたプレイスペクトルの値がすべての時間において等しくなり従来法の効果が小さくなる場合の改善について述べる。提案手法はプレイスペクトルの値に差を生じさせるため、プレイスペクトルに情報を追加、またはプレイスペクトルの代わりにスタッフの移動元と移動先の情報を組み合わせたベクトルを用いる。これらの手法の効果の検証は4章と同様に、時間によってスタッフの行動ルールが変わるように作成したシミュレーションデータに対し、ルールの切り替わるタイミングが検知できれば効果あり、できなければなしとする。

- 改善案1：プレイスペクトルに新しく情報を追加

プレイスペクトルにスタッフの移動方向を示す要素を追加することですべての時間において等しくなることを防ぐ。移動方向の表現の仕方としては、各スタッフにつき、エリア番号が大きくなるように移動する場合は1、変わらない（待機）場合は0.5、エリア番号が小さくなるように移動する場合は0を追加する要素に加える。例えば、エリア数が3、スタッフ数が2で、1人のスタッフがエリア1から2、もう1人のスタッフがエリア3で待機しているとき、移動情報を表す要素の値は1.5となり、プレイスペクトルは[0.5,0.5,1,1.5]となる。

この方法を4.2節の条件1~3に適用した結果を図5-1~5-3に示す。図5-1に示すように、情報を追加したことによる劣化は見受けられない。また、図5-2、5-3に示すように、この方法を用いると、それぞれ表4-4、4-6で設定した行動ルール毎にクラス分けされていることが確認できる。条件2では、プレイスペクトルに要素を追加することにより、フェーズ1クラスタリングの結果が変化したことが原因である。条件3も同様に、従来法ではすべての時間においてプレイスペクトルは[1,1,1]であったのに対し、この方法では、0~1999秒では[1,1,1,3]となり、2000~3999秒では、[1,1,1,0]となったのが原因である。なお、フェーズ2クラスタリングを行わない理由は4.3節の条件8と同様である。ただし、クラスタリングのパラメータとして、プレイスペクトルのパターンが5から増えているので、フェーズ1クラスタリングのクラスタ数を増やす必要があると判断し、条件1において、いくつかのパターンを試した結果クラスタ数を10とした。他のパターンの結果として、クラスタ数5、100の結果をそれぞれ図5-4、5-5に示す。

- 改善案2：移動元と移動先の情報を組み合わせたベクトルを使用

プレイスペクトルの代わりに移動元と移動先の情報を組み合わせたベクトルを用いる。移動元のエリアと移動先のエリアのすべての組み合わせを表現するため、

このベクトルの要素数はエリア数の二乗である。例えば、エリア数が3、スタッフ数が2で、1人のスタッフがエリア1から2、もう1人のスタッフがエリア3で待機しているとき、1→2を示す2番目の要素と3→3を示す9番目の要素に値が追加され、ベクトルは[0,1,0,0,0,0,0,0,1]となる。

この方法を条件1~3に適用した結果を図5-6~5-8に示す。図5-6より、劣化は見られず、図5-7、5-8より表4-4、4-6で設定した行動ルール毎にクラス分けされていることが確認できる。行動ルール毎にクラス分けされた原因は、改善案1と同じである。ただし、図5-6のみはフェーズ1クラスタスタリングのクラスタ数が10では不足したため、12とした。

上記により、プレイスペクトルが等しくなる場合やプレイスペクトルが異なるが隣接行列が等しくなる場合にプレイスペクトルに情報を追加、またはプレイスペクトルの代わりの特徴抽出方法を用いることで改善できることを示した。改善案1の移動方向の表現の仕方について、図4-4の巨大倉庫でのピッキングのような場合には表現の仕方を工夫する必要があると考えられるが、図4-1の介護施設のような場合は現状の方法で効果を発揮できると思われる。また、改善案2について、従来のプレイスペクトルを用いる方法に比べベクトルの要素数が増えてしまうのがネックである。これについては後の7.5節で主成分分析を用いた次元削減について考察を行う。

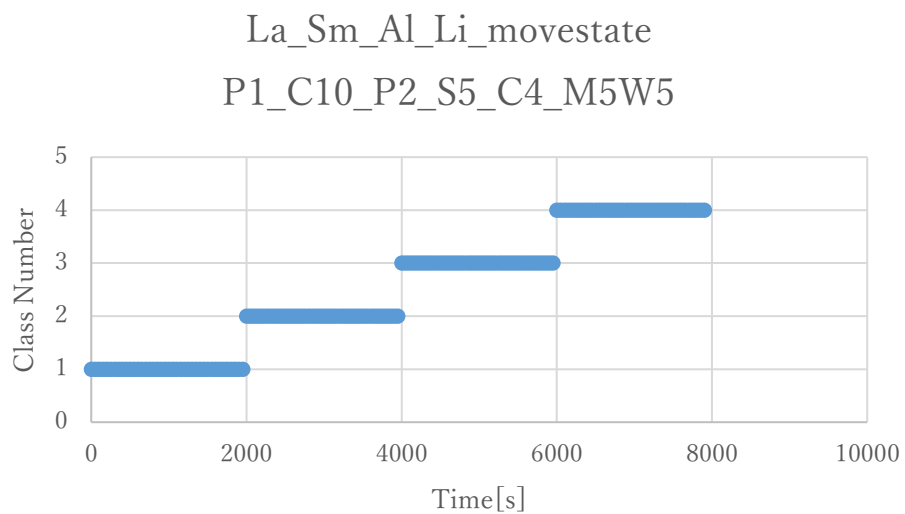


図 5-1. 改善案 1 による条件 1 のフェーズ 2 クラスタリング結果

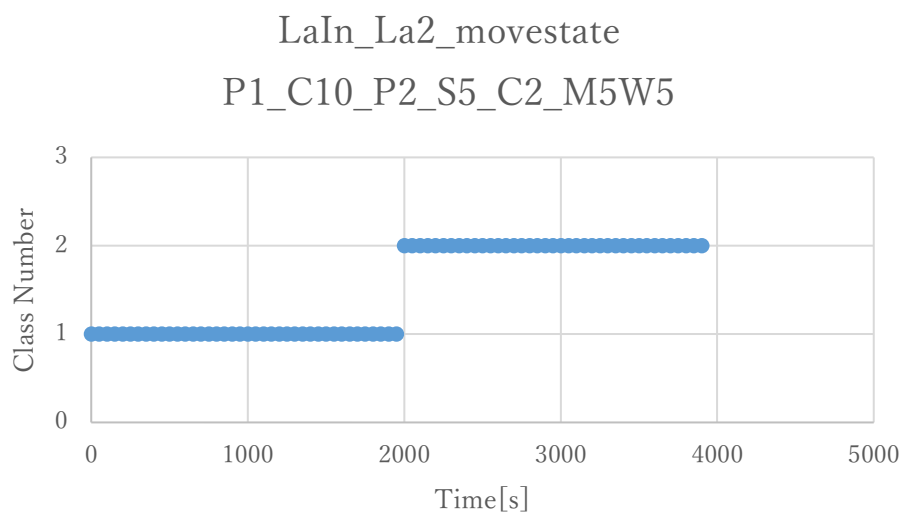


図 5-2. 改善案 1 による条件 2 のフェーズ 2 クラスタリング結果

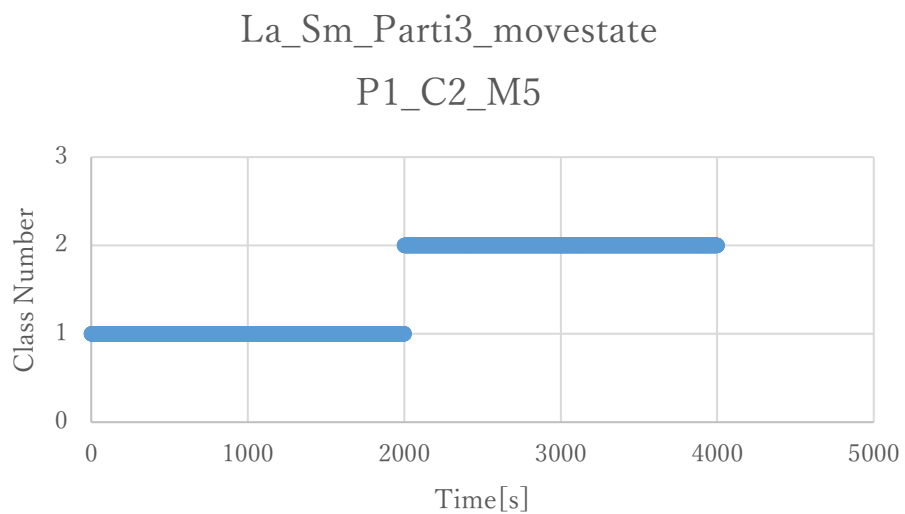


図 5-3. 改善案 1 による条件 3 のフェーズ 1 クラスタリング結果

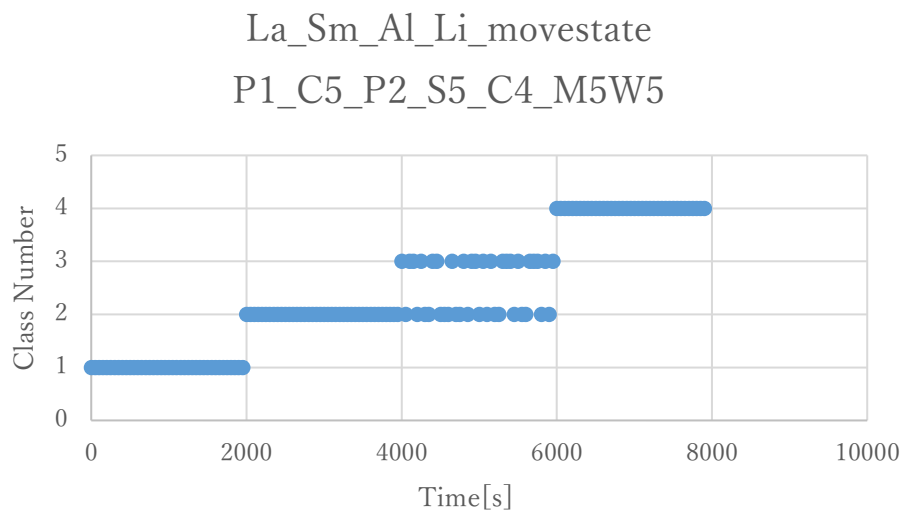


図 5-4. 改善案 1 による条件 1 のフェーズ 2 クラスタリング結果(クラス数 5)

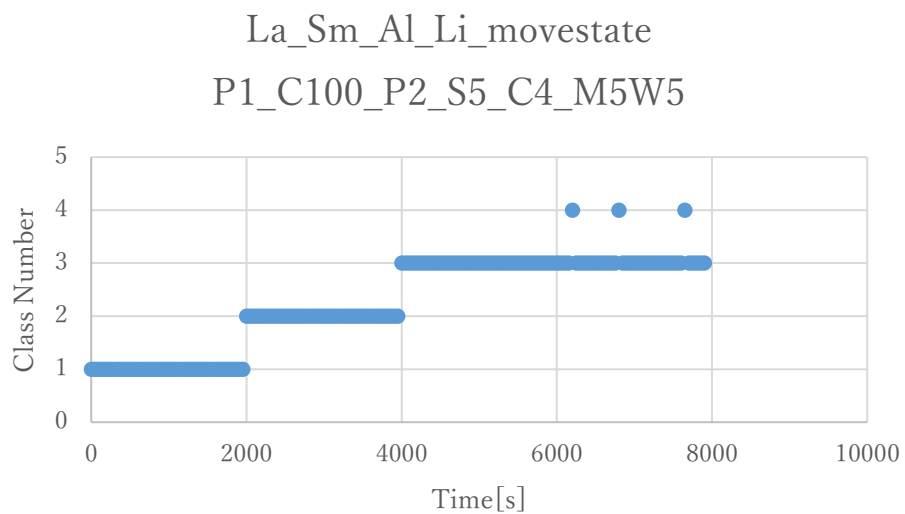


図 5-5. 改善案 1 による条件 1 のフェーズ 2 クラスタリング結果(クラス数 100)

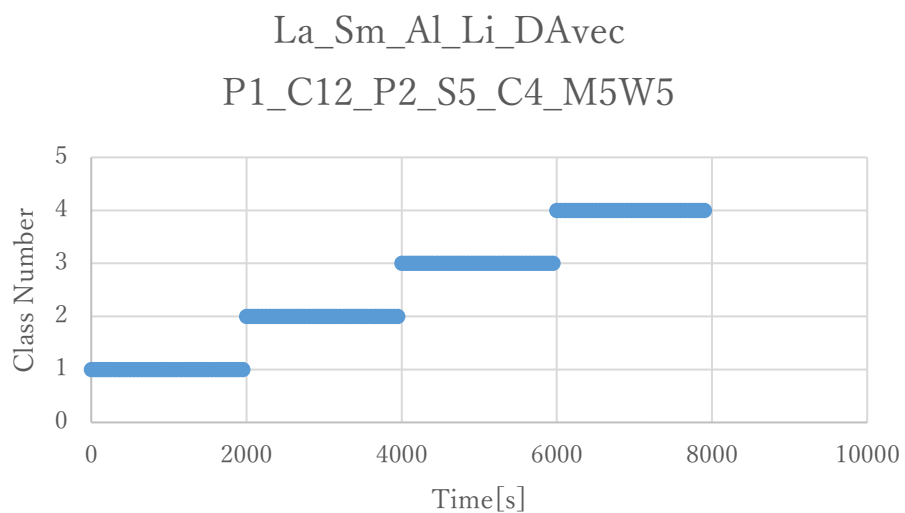


図 5-6. 改善案 2 による条件 1 のフェーズ 2 クラスタリング結果

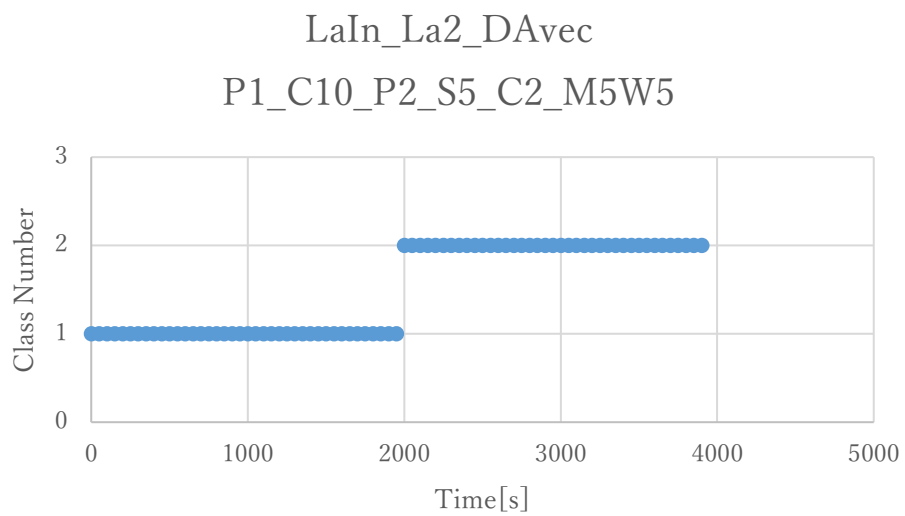


図 5-7. 改善案 2 による条件 2 のフェーズ 2 クラスタリング結果

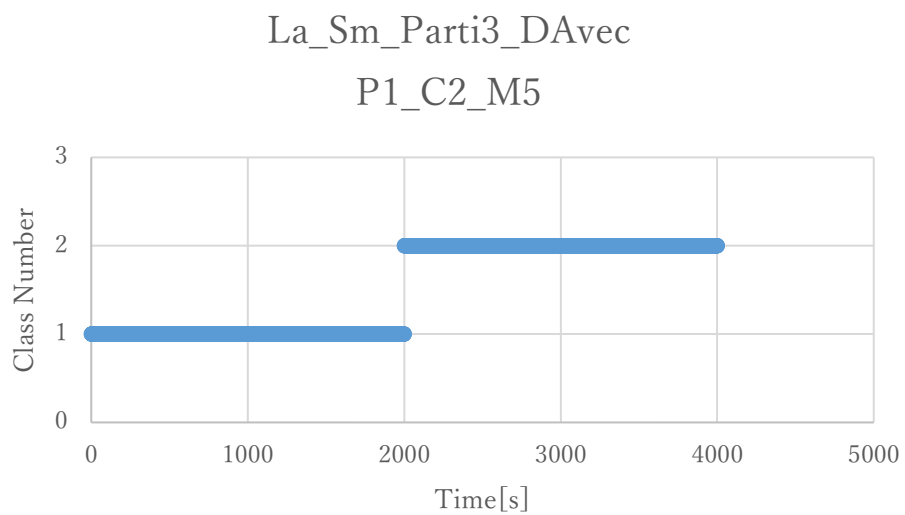


図 5-8. 改善案 2 による条件 3 のフェーズ 1 クラスタリング結果

第6章 作業場所の情報とつぶやきデータを組み合わせたケア内容推定の検証

本章では3.1節で紹介した介護施設での実験結果[3]に対し、作業の主たる位置情報に加え、実験で取得した介護士をつぶやきデータを用いてケア内容の推定を行った場合に推定精度が上昇するか検証する。検証は、まず各ケアに対応したキーワードを設定し、そのキーワードが含まれるつぶやきデータにラベル付けを行った結果とクラスタリングの結果を見比べて各クラスにケア内容を当てはめる。次に、この結果と従来の結果(作業の位置情報を基に推定したケア内容の推定結果)を組み合わせ、従来の結果では推定できなかった部分についての情報が上がっているかを確認することで行う。

介護士をつぶやきは、1.1節で述べた[1][2]のSVMシステムを用いており、図6-1のようにサーバにプールされる。あらかじめ各ケアに関するキーワードを設定しておき、つぶやきをラベル付けした結果とフェーズ2クラスタリングの結果を見比べることでクラス内容の特定を行う。設定した各ケアのキーワードを表6-1にまとめ、昼食と夕食それぞれ3日分の実験データのフェーズ2クラスタリング結果とつぶやきをラベル付けした結果を図6-2~6-7に示す。図6-2~6-7からクラス1が排泄ケア、2が食事ケアであると推定できる。また、先行研究での中心ベクトルを用いたクラス内容の推定結果を表6-2にまとめる。表6-2において、クラス5が口腔ケアであるが、図6-7を見るとクラス5は移動中を示しているのは、口腔ケアが各部屋に移動して行うケアであるからだと思われる。今回の結果からはラベル付けしたつぶやきデータを用いてクラスタリング結果のクラス3、4の内容を特定することは難しかったが、クラス1、2については表6-2と合わせて推定の裏付けに貢献できていると思われる。

つぶやきデータを各ケアにラベル付けするにあたり、介護施設というある程度使用されるワードが決まっている環境であることを考慮し、本研究ではあらかじめキーワードを設定するという手法を用いた。このような環境以外で、時間によって作業内容が変化する場合には、2.2節で紹介したバースト解析[10]などが有効ではないと思われる。



図 6-1. SVM システム概略図([2]より引用)

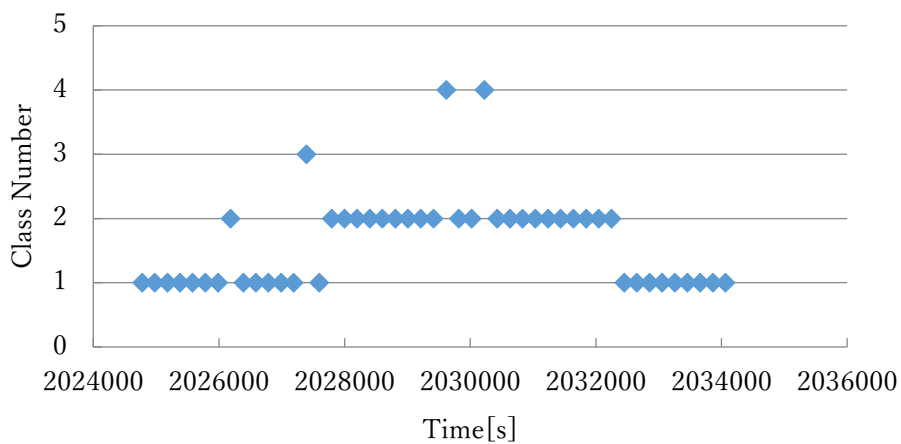
表 6-1. 各ケアのキーワード設定

ケア内容	キーワード
クラス 1:排泄ケア	排泄, 排尿, 排便, トイレ, 座り, 立ち上がり
クラス 2:食事ケア	食事, 食介, 薬, 食後, 全量, おかず, 食べこぼし, 野菜, ダイニング, お肉, 箸, 食器, ご飯, おいしい
クラス 3:口腔ケア	口腔ケア, 歯磨き
クラス 4:イブニングケア (着替え, 体拭き等)	イブニングケア, 床, 着替え, 介助, 手, 頭, お腹, 腹部, 背中, 胸, 足
クラス 5:移動中	1階, 2階, 3階, 4階, 誘導, 居室, 行きます, 戻ります, 入ります, 移送, 送り
クラス 0:該当なし	上記のどれにも該当しない場合

表 6-2. 中心ベクトルを用いたクラス内容の推定結果

クラス番号	内容
クラス 1	昼食または夕食前の排泄ケア
クラス 2	食事ケアまたは昼食か夕食後のアクティビティーのケア
クラス 3	ケアスタッフ同士の打ち合わせ
クラス 4	食事ケアの準備
クラス 5	各部屋での口腔ケア

3日目昼食 クラスタリング結果



3日目昼食 つぶやきラベル付け結果

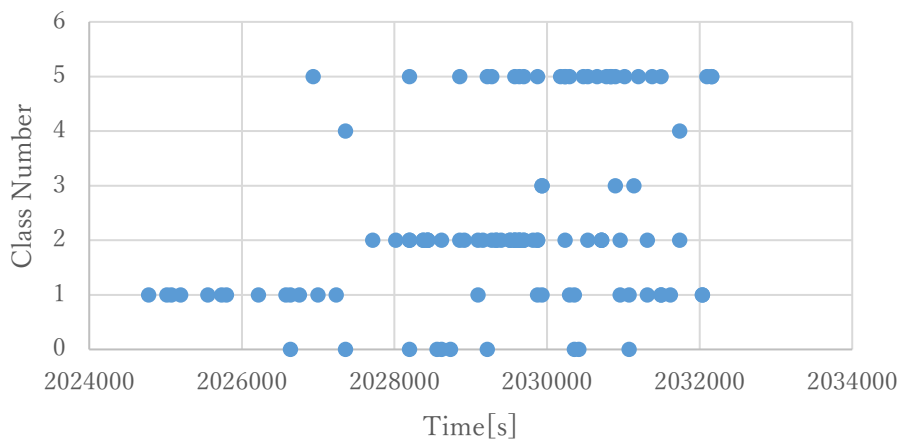
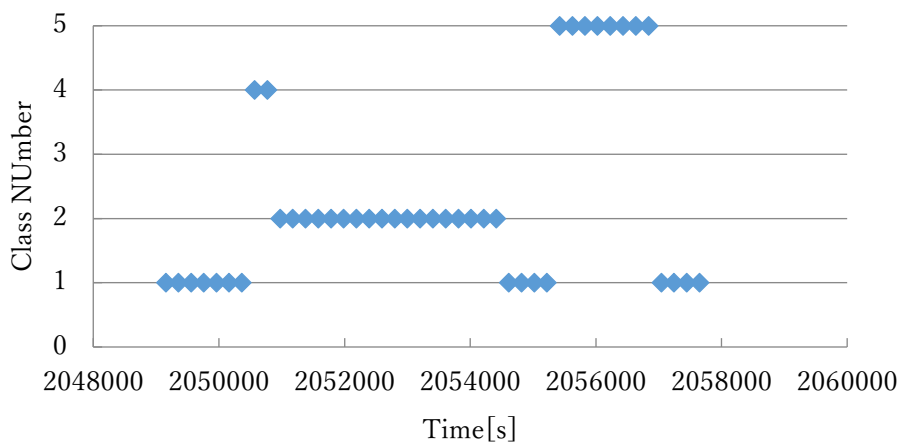


図 6-2. クラスタリング結果とつぶやきラベル付け結果(昼食 1)

3日目夕食 クラスタリング結果



3日目夕食 つぶやきラベル付け結果

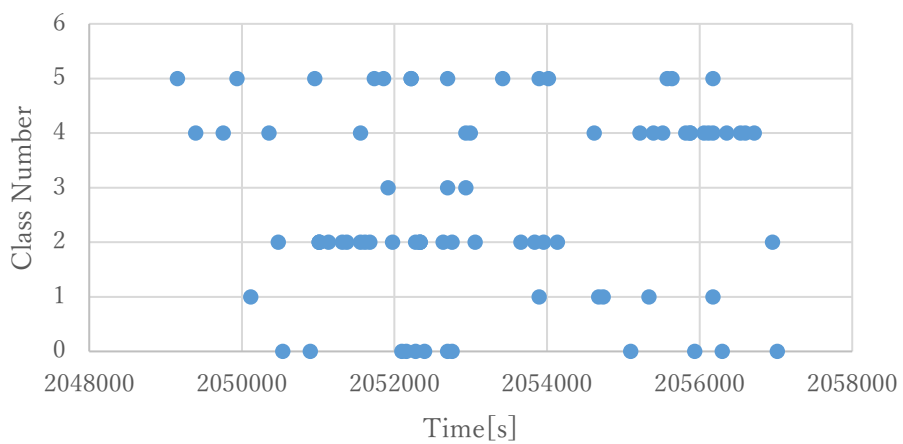
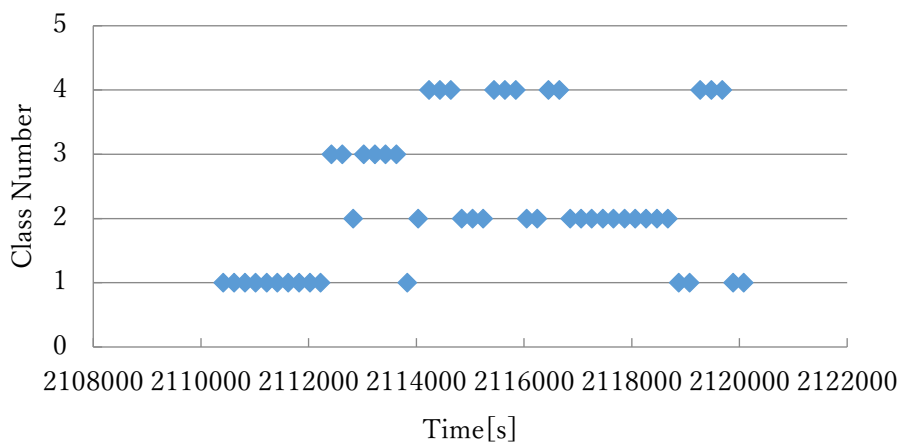


図 6-3. クラスタリング結果とつぶやきラベル付け結果(夕食1)

4日目昼食 クラスタリング結果



4日目昼食 つぶやきラベル付け結果

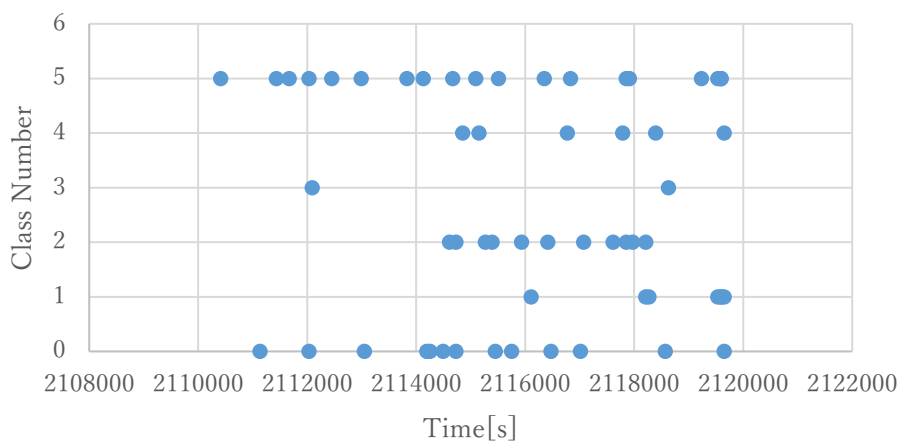
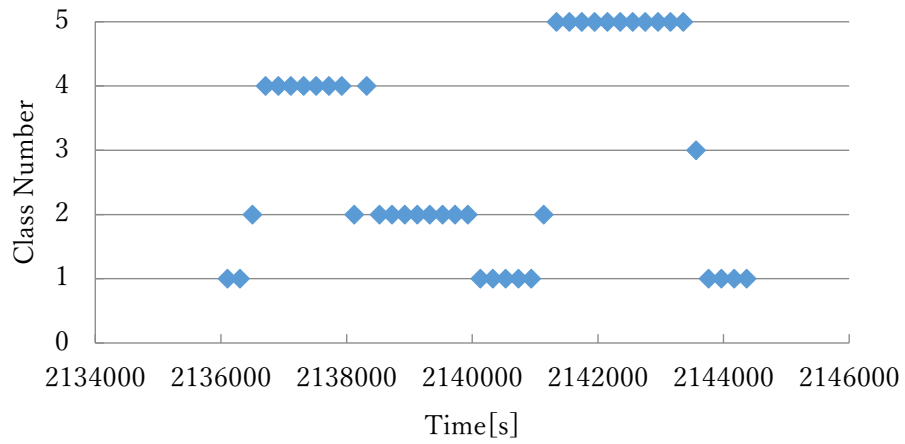


図 6-4. クラスタリング結果とつぶやきラベル付け結果(昼食 2)

4日目夕食 クラスタリング結果



4日目夕食 つぶやきラベル付け結果

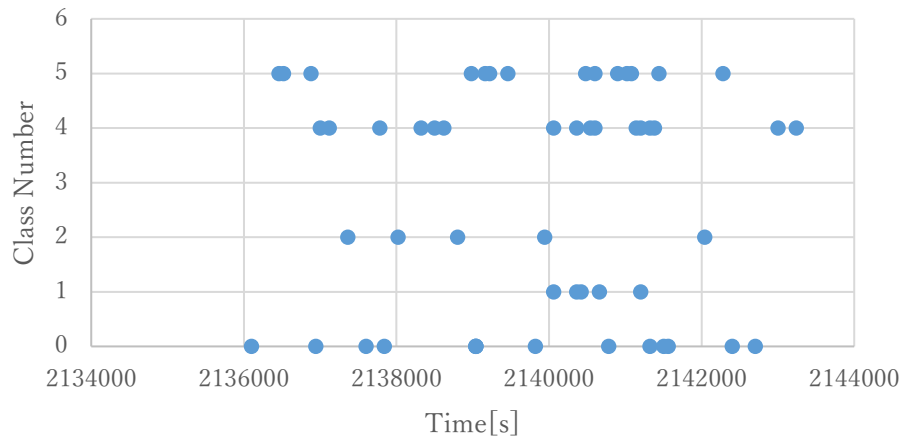
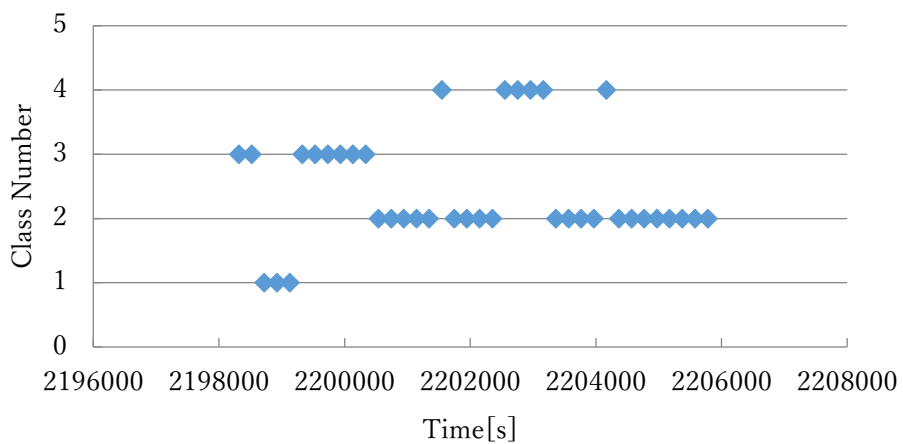


図 6-5. クラスタリング結果とつぶやきラベル付け結果(夕食2)

5日目昼食 クラスタリング結果



5日目昼食 つぶやきラベル付け結果

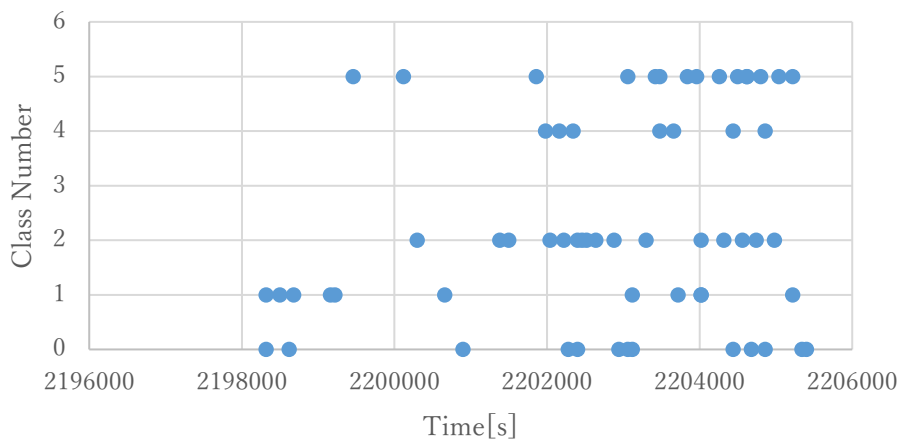
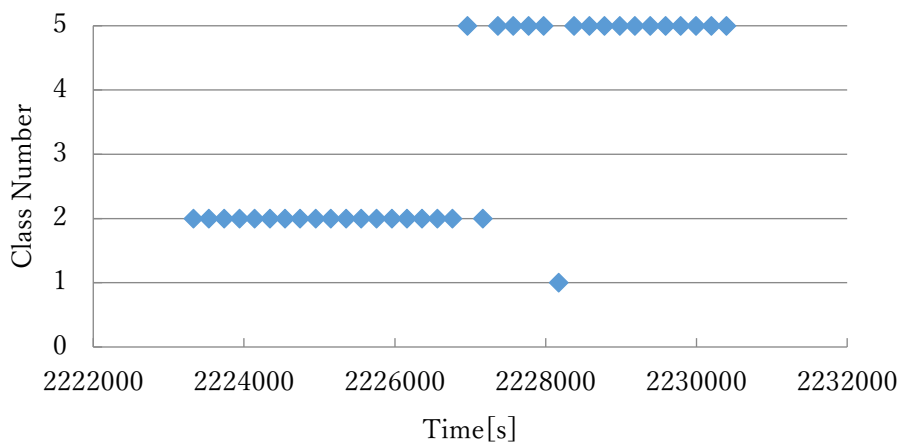


図 6-6. クラスタリング結果とつぶやきラベル付け結果(昼食 3)

5日目夕食 クラスタリング結果



5日目夕食 つぶやきラベル付け結果

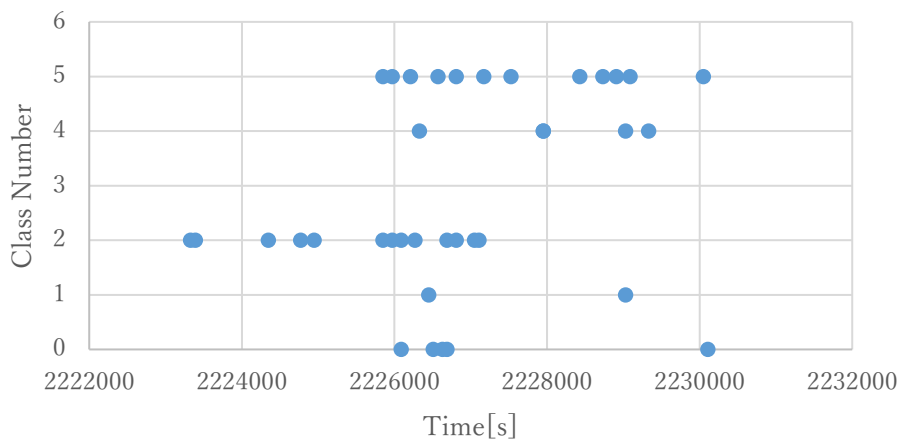


図 6-7. クラスタリング結果とつぶやきラベル付け結果(夕食3)

第7章 考察

本章では、第4, 5章の内容についての考察を述べる。

7.1 クラスタリングのクラスタ数

3.1節で述べたように、従来法ではクラスタリングにおけるクラスタ数はフェーズ1, 2共にデンドログラムを見て、介護のケア内容の種類を基にクラスタ数がいくつ必要か予想し、決定していた。また、従来法の検証において、フェーズ1クラスタリングのクラスタ数はデンドログラムを見て、フェーズ2クラスタリングでクラスタ数が不足しないように決定している。そこで、本節ではクラスタ数を自動で決定した場合のクラスタリングの効果を考察する。デンドログラムにおいてマージされている高さの最も大きい値に対して割合でクラスタリングする高さを決めることで、クラスタ数を自動で決定する。

4.2節で述べた条件1について、従来法で解析を行ったときのフェーズ1クラスタリングのデンドログラムを図7-1に示す。図7-1において、マージされている高さの最も大きい値は2704.499である。4, 5章よりフェーズ1クラスタリングのクラスタ数は不足するとフェーズ2クラスタリングの効果が小さくなることが分かっていたので、クラスタ数がある程度大きく設定するため、今回はこの値の3割と1割の値の高さでフェーズ1クラスタリングを行った。その後、フェーズ2クラスタリング（クラスタ数4）を行った結果をそれぞれ図7-2, 7-3に示す。この時、フェーズ1クラスタリングのクラスタ数はそれぞれ6, 9であった。図7-2, 7-3より、図7-2ではクラス3と4の境界が少しわかりにくいものの行動ルール毎にクラス分けされているのが確認できる。そのため、大量のデータをクラスタリングする際など、デンドログラムを見てクラスタを人手で決定する手間をかけられない場合にはこの自動決定の手法が使えるのではないかと思われる。

7.2 フェーズ2クラスタリングを行う意味

本節ではフェーズ2クラスタリングの効果について考察する。3.1節で述べた通り、フェーズ2クラスタリングはフェーズ1クラスタリングの結果から隣接行列を生成し、それをクラスタリングする。そのため、フェーズ1クラスタリングのクラス番号の変化を表していると言える。これにより、フェーズ1クラスタリングのデンドログラムを人の目で見ると、一見、行動ルール毎にクラス分けできるか不明な場合でも、フェーズ2でクラス分けすることができる効果を持つと考えられる。クラス分けできた例として、4.3節の条件4のフェーズ1クラスタリングのデンドログラムを図7-4に示し、クラスタリング結果を図7-5に示す。クラスタ数は10としていたのでクラスタリングをした高さを破線で示す。図7-4中のオレンジ色の枠で囲まれた部分がフェーズ1クラスタリングにおいて、クラス10にあたる。

図 7-5 から分かるように、クラス 10 が連続して 2200 秒以降続き、隣接行列を生成した際に中心ベクトルが他の時間帯の隣接行列と大きく異なる。そのため、クラスタリングした際に、図 4-5 から分かるように 2200 秒を境にクラスが分かれる。一方で、クラス分けできなかった例として、条件 6 のデンドログラムおよびクラスタリング結果を図 7-6, 7-7 に示す。図 7-7 より、900 秒から 1600 秒あたりまでクラス 6 が続いたため、隣接行列を生成し、フェーズ 2 で 2 クラスに分けようとした際に、図 4-9 のように 900~1600 秒の部分が別クラスとクラスタリングされる。このように、人の目でフェーズ 1 クラスタリングのデンドログラムを見て、行動ルール毎にクラス分けできるかを確認できない場合に、フェーズ 2 クラスタリングが効果を発揮する。

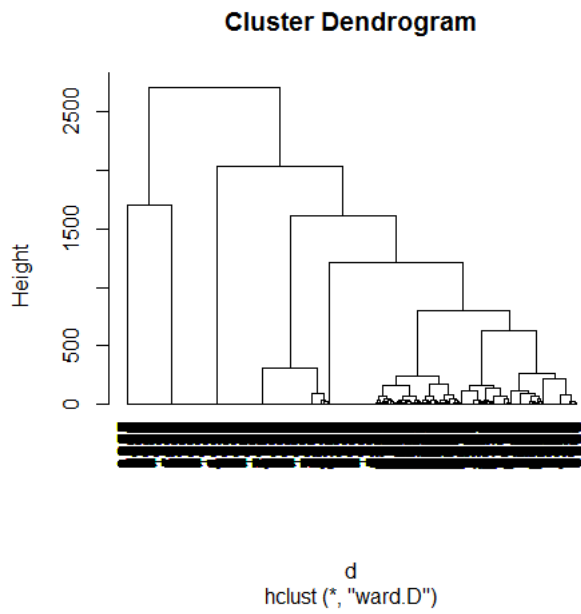


図 7-1. 条件 1 のフェーズ 1 クラスタリングのデンドログラム

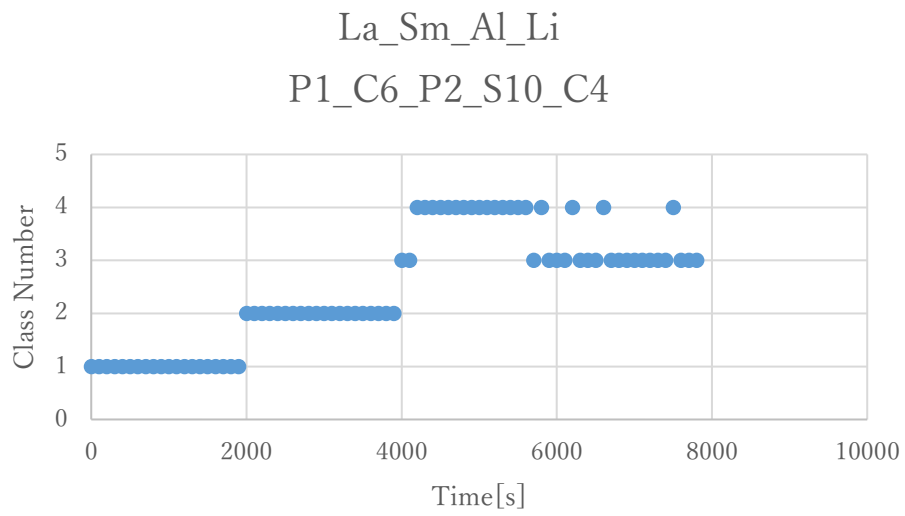


図 7-2. 条件 1 のフェーズ 2 クラスタリング結果(3 割)

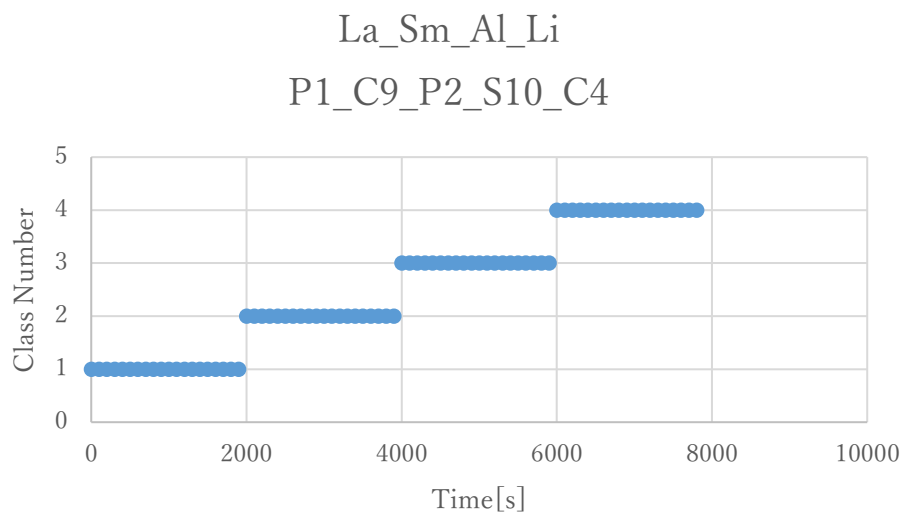


図 7-3. 条件 1 のフェーズ 2 クラスタリング結果(1 割)

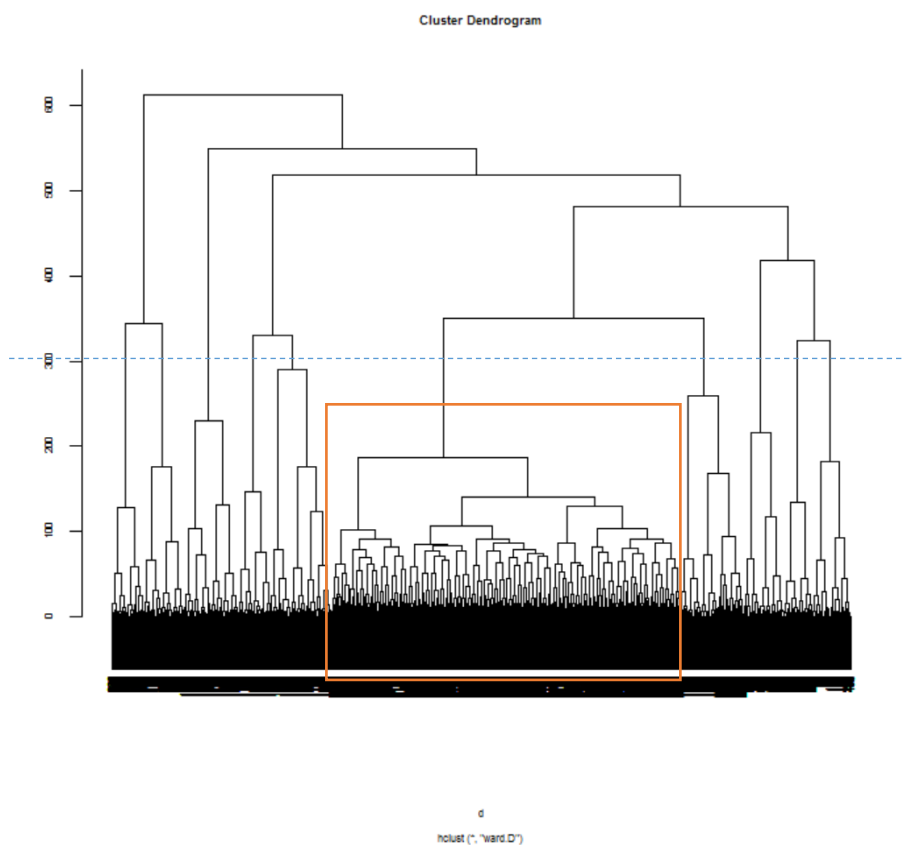


図 7-4. 条件 4 のフェーズ 1 クラスタリングのデンドログラム

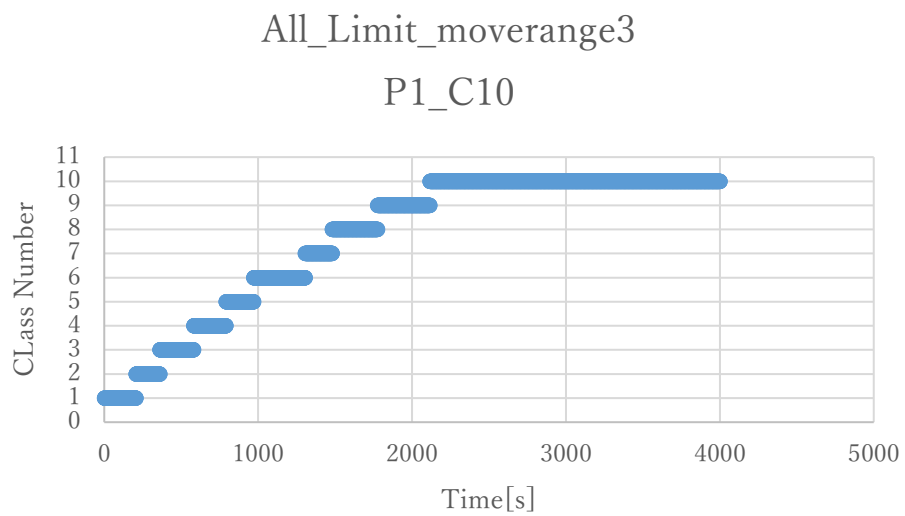


図 7-5. 条件 4 のフェーズ 1 クラスタリング結果

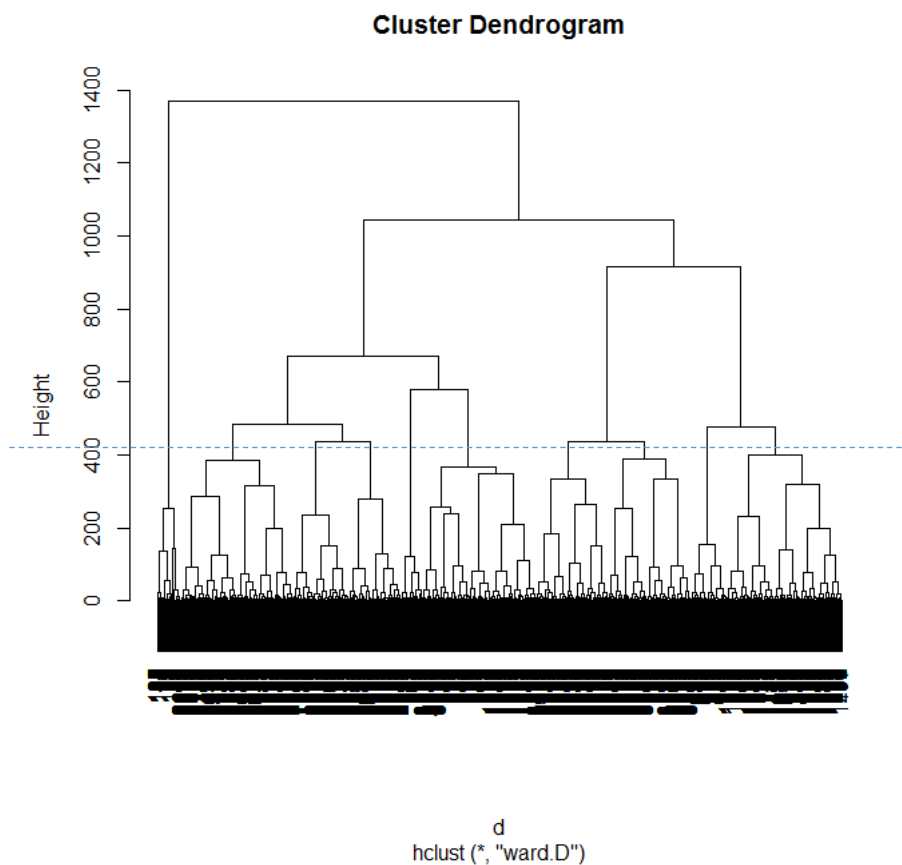


図 7-6. 条件 6 のフェーズ 1 クラスタリングのデンドログラム

Acoder_Decoder P1_C10

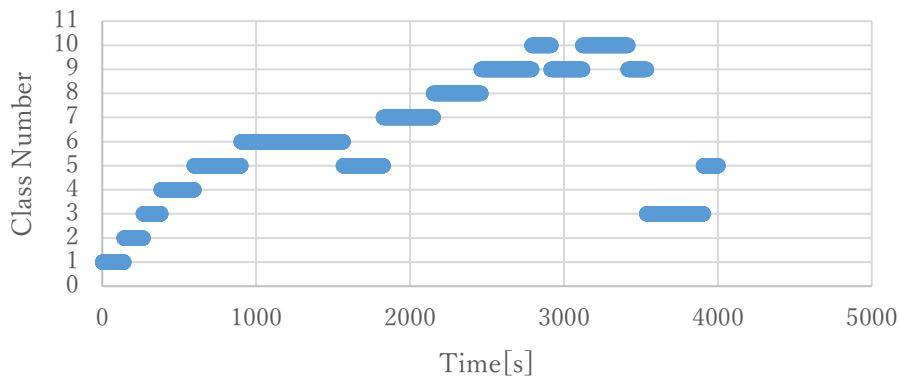


図 7-7. 条件 6 のフェーズ 1 クラスタリング結果

7.3 提案法のノイズ耐性

本節では、4.3 節で述べた提案法の改善案 2 について、スタッフが一定の確率で行動ルールに沿った行動をとらないというノイズを加えた場合について条件 1 と 3 を用いて考察する。ただし、スタッフの移動中および待機中にノイズは発生せず、次の行き先を選ぶときのみノイズ発生の判定を行う。

表 4-2 より条件 1 では行動ルールが 4 つある。ノイズは 70% の確率で規定のルールに従い、10% ずつの確率で規定以外の 3 つのルールそれぞれに従うように発生させる。このノイズを与えた状態でフェーズ 2 クラスタリングまで行った条件 1 の結果を図 7-8 に示す。図 7-8 より、行動ルール毎にクラス分けできていないため、フェーズ 1 クラスタリングのクラス数を 12 から 15 に変更し、フェーズ 2 クラスタリングを行った結果を図 7-9 に示す。図 7-9 より、クラス 3, 4 の境界が少しわかりづらいがクラス分けできていることが確認できる。これは、ノイズによるプレイスペクトルのパターン増加にクラス数を増やしたことで対応したためであると考えられる。また、条件 3 についても同様にノイズを与えた場合のフェーズ 2 クラスタリング結果を図 7-10 に示す。図 7-10 より、こちらはフェーズ 1 クラスタリングのクラス数が 10 でもクラス分けできているのが確認できる。これは、ノイズを付与しない場合のプレイスペクトルのパターンが 2 つのみで、ノイズの付与によるプレイスペクトルの増加に結果が乱されなかったためであると考えられる。

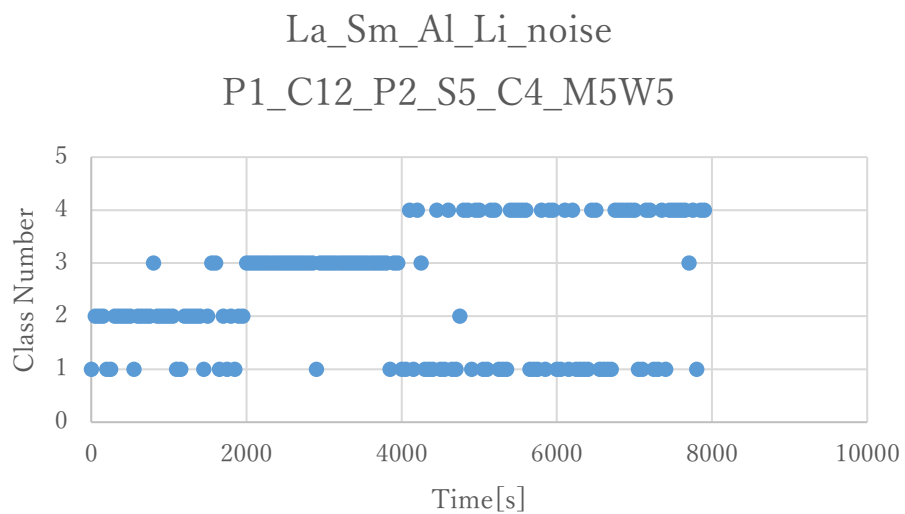


図 7-8. 提案法による条件 1 のフェーズ 2 クラスタリング結果(ノイズ有, クラスタ数 12)

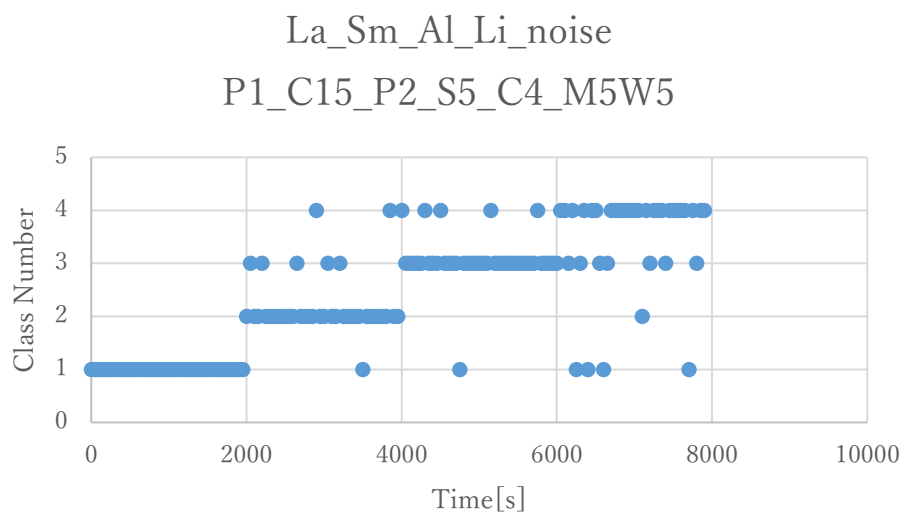


図 7-9. 提案法による条件 1 のフェーズ 2 クラスタリング結果(ノイズ有, クラスタ数 15)

La_Sm_Part3_noise
P1_C10_P2_S5_C2_M5W5

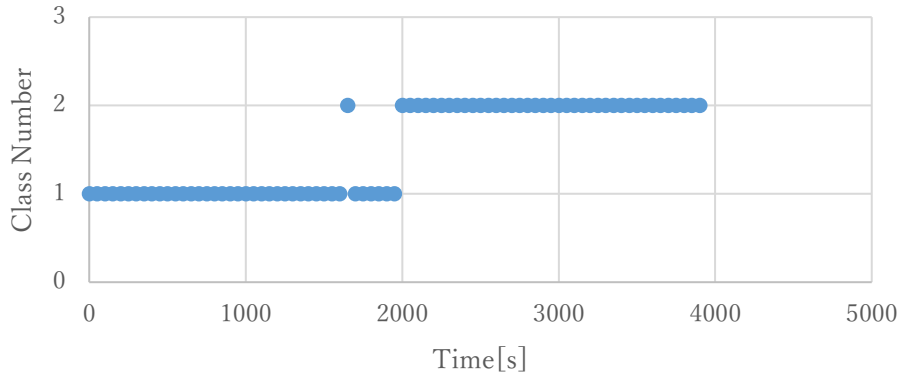


図 7-10. 提案法による条件 3 のフェーズ 2 クラスタリング結果(ノイズ有)

7.4 クラスタリング結果の評価

本節ではクラスタリングの結果の評価について考察を行う。クラスタリングの評価方法については、DBIndex[12]や DunnIndex[13]などがあり、今回は DBIndex による評価を行う。DBIndex は次の式で表され、 c_i はクラス i のセントロイド、 $d(a, b)$ は a と b のユークリッド距離を表す。

$$D(Davis)B(Bouldin)Index = \frac{1}{k} \sum_{i=1}^k R_i$$

$$R_i = \max_{1 \leq j \leq k, j \neq i} \left\{ \frac{S_i + S_j}{d_{ij}} \right\}$$

$$d_{ij} = d(c_i, c_j)$$

$$S_i = \left(\frac{1}{|C_i|} \sum_{x \in C_i} (d(x, c_i))^2 \right)^{\frac{1}{2}}$$

この式より、各クラス同士のスセントロイド間の距離が大きく、各クラス内の分散が小さいほど、DBIndex の値は小さくなる。

図 5.1 の結果では、DBIndex は 0.033、図 7-9 の結果では、1.553 となる。ノイズを入れた場合は各クラス同士のセントロイド間の距離が小さくなり、各クラス内の分散が大きくなっているため、DBIndex はクラスタリング結果の評価に用いることができると言える。

7.5 提案法におけるベクトルの次元増加に対しての主成分分析による次元削減

4 章で移動元と移動先の情報を組み合わせたベクトルを使用する方法では、ベクトルの要素数がエリア数の二乗になってしまい、次元数が増加する問題があると述べた。そこ

で、本節では、Python の機械学習ライブラリである scikit-learn に含まれる主成分分析 (PCA; Principal Component Analysis) の機能を用いた次元削減を考察する。

条件 1 において、PCA を用いてベクトルの要素を 9, 15, 20 としたときのフェーズ 2 クラスタリングの結果をそれぞれ図 7-11, 7-12, 7-13 に示す。図 7-11 から 9 まで次元を削減してしまうとクラス 3 と 4 の境界が分からなくなってしまうが、図 7-12, 7-13 のように 15 や 20 までの削減であればルールの切り替わりが目視で確認できるため、PCA を用いた次元削減は有効であると言える。

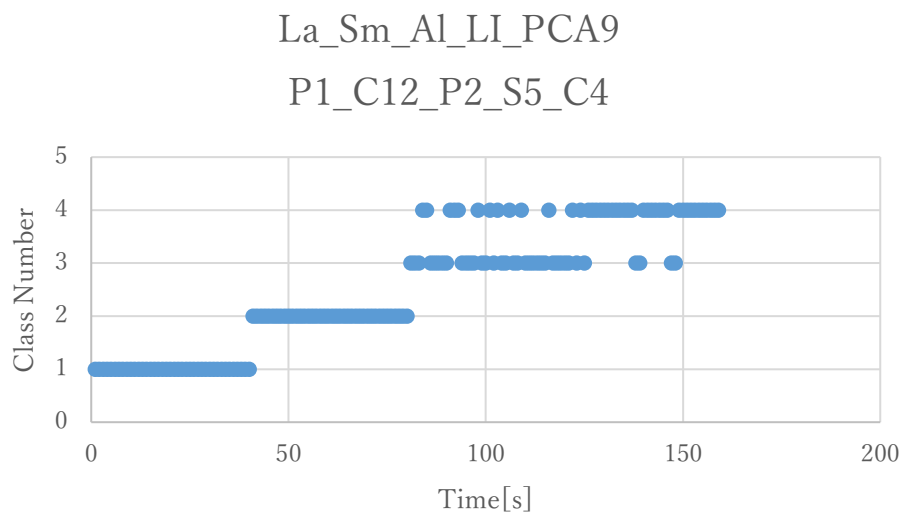


図 7-11. 提案法による条件 1 のフェーズ 2 クラスタリング結果(PCA9)

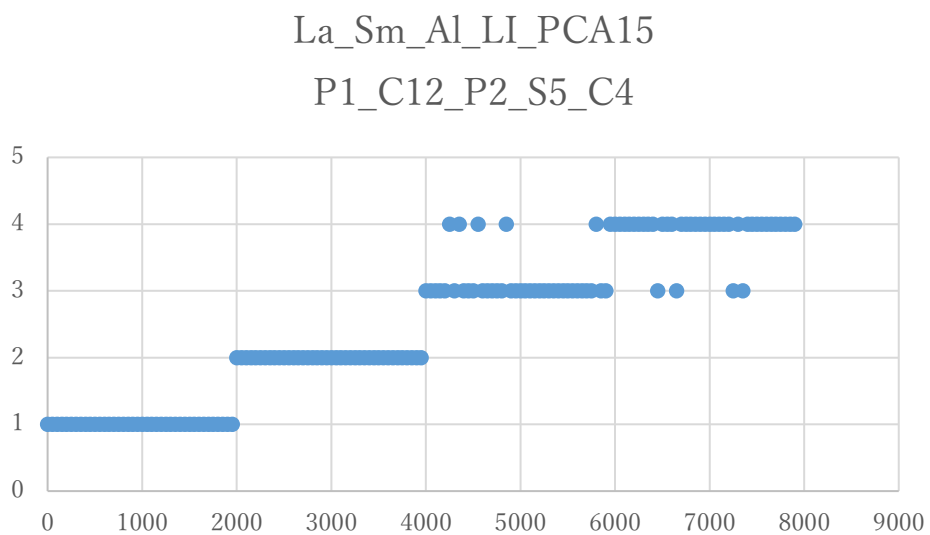


図 7-12. 提案法による条件 1 のフェーズ 2 クラスタリング結果(PCA15)

La_Sm_Al_LI_PCA20

P1_C12_P2_S5_C4

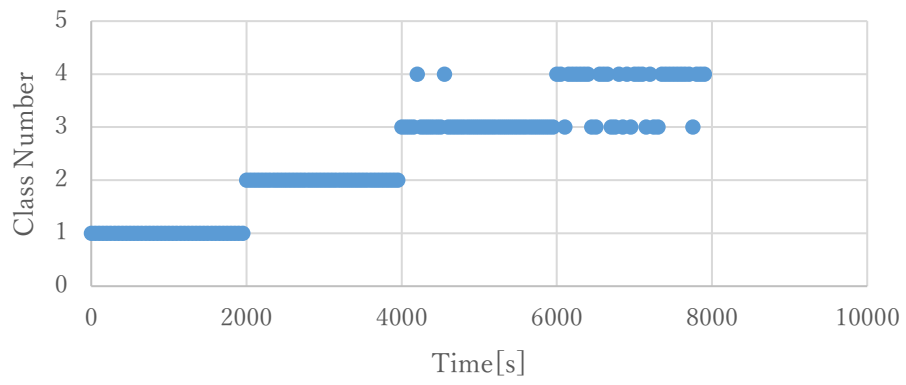


図 7-13. 提案法による条件 1 のフェーズ 2 クラスタリング結果(PCA20)

第8章 むすび

本論文では、十分な検証が行われていなかった先行研究の時空間状況認識手法に対して、シミュレーションを用いて、簡単なモデルを用いた場合とアマゾンの巨大倉庫のような実際の現場を想定した場合において時空間状況認識手法の効果の検証を行った。また、簡単なモデルを用いた検証において、時空間状況認識技術の効果が小さくなるような状況を見出し、そのような状況でも効果を発揮できるように改善した時空間状況認識手法を提案した。さらに、先行研究の介護現場での実験について、食事ケアや排泄ケアといった各ケアに関連するキーワードを設定することで介護士のつぶやきデータを分類し、分類されたつぶやきデータが作業内容の推定に役立つかどうかの検証を行った。

今後の課題として、今回はシミュレーションのみで実験を行えていないため、介護施設などで実験を行い、5章で述べた提案法および6章で述べたつぶやきデータを用いたクラス内容の推定の実証を行う必要がある。先行研究での実験ではいつ、どのようなケアをしていたかという詳細な記録はとっていない。そのため、実験を行うにあたり、介護施設であれば、ケア内容の推移の記録、その他の施設であれば作業状況の推移を記録する必要があると考えている。

謝辞

本論文の作成にあたり、主指導教員である平石邦彦教授にはいつも貴重な時間を割いていただき、丁寧な御指導とご助言を賜りました。この場を借りまして心よりの感謝を申し上げます。大変ありがとうございました。また、副テーマ指導教員としてご指導いただきました白井清昭准教授、審査委員として貴重なご意見をいただきました内平直志教授、緒方和博教授、上原隆平教授の各先生方にも改めて感謝の意を表します。

参考文献

- [1] N. Uchihira, S. Choe, K. Hiraishi, K. Torii, T.Chino, Y.,Hirabayashi, T.Sugihara (2013)“Collaboration Management by Smart Voice Messaging for Physical and Adaptive Intelligent Services”, Proceedings of PICMET 2013,Technology Management for Emerging Technologies, pp. 251-258
- [2] K. Hiraishi, K. Kobayashi, S. Choe, N. Uchihira (2014), “Behavior Modeling in Physical and Adaptive Intelligent Services”, Proc. of the 2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods inSituation Awareness and Decision Support, pp. 221-226, San Antonio
- [3] Karou Sato, Kunihiro Hiraishi, Kouichi Kobayashi (2016): Spaito – Temporal Situation Recognition for Groups in Caregiving Services, Proc. IEEE CogSIMA2016, pp.78-82
- [4] 大西正輝・依田育士 (2010) 「大型複合施設における長期間にわたる人流比較と可視化手法」, 『電子情報通信学会論文誌』, D Vol.J-93D No.4 pp.486-493
- [5] 山下倫央・荒牧英治・宮部真衣・磐田敦之・大西正輝・依田育士・野田五木樹 (2014) 「大規模群集流動の制御に向けた統合支援サービスの提案」, Joint Agent Workshops & Symposium 2014
- [6] 坂巻英一・亀井悦子 (2014) 「Twitter 上のつぶやきに関するテキストマイニングの事例研究 - 大規模災害発生時の被災地における現状把握への応用 - 」, 『日本経営工学会論文誌』, Vol.65 No.1 pp.39-50
- [7] 村上知子・鳥居健太郎・長健太・内平直志 (2014) 「センサデータと業務知識からのトピックモデルを用いた看護業務行動の推定」, 『人工知能学会論文誌』, Vol.29, No.5A pp.427-435
- [8] Manish Gupta, Jing Gao, Charu C.Aggrawal, Jiawei Han (2014) “Outlier Detection for Temporal Data: A Survey”, *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, VOL.26, NO.9 pp.2250-2267
- [9] VARUN CHANDOLA, ARINDAM BANERJEE, VIPIN KUMAR (2009) “Anomaly Detection: A Survey”, *ACM Computing Surveys*, pp.1-72
- [10] J. Kleinberg (2002) “Bursty and Hierarchical Structure in Streams”, Proc. of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining
- [11] Kentaro Torii, Naoshi Uchihira, Yuji Hirabayashi, Testuro Chino, Takanori Yamamoto, Satoko Tsuru (2016) “Improvement of Sharing of Observations and Awareness in Nursing and Caregiving by Voice Tweets”, in T.Maeno et al.(eds.), *Serviceology for Designing the Future*, Springer Japan, pp.161-175

[12] David L. DAVIES, DONALD W. BOULDIN (1979) "A Cluster Separation Measure".
IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. PAMI-1, No.2
pp.224-227.

[13] Dunn, J. C. (1973). "A Fuzzy Relative of the ISODATA Process and Its Use in
Detecting Compact Well-Separated Clusters". Journal of Cybernetics 3 (3): 32-57

研究業績

1. 早貸舜・平石邦彦・内平直志
作業スタッフの行動ログを用いた時空間状況認識の性能向上について,
MSS 研究会, 2018-3-14(発表予定)