

# 音声感情データ解析によるコールセンターの応対評価

上野 拓武<sup>†</sup> 有次 正義<sup>††</sup>

<sup>†</sup> 熊本大学大学院自然科学教育部 〒860-8555 熊本県熊本市中央区黒髪2丁目39番1号

<sup>††</sup> 熊本大学大学院先端科学研究部 〒860-8555 熊本県熊本市中央区黒髪2丁目39番1号

E-mail: <sup>†</sup>takumu@st.cs.kumamoto-u.ac.jp, <sup>††</sup>taritsugi@cs.kumamoto-u.ac.jp

あらまし コールセンターでは顧客からのサポート対応や通信販売の申し込み電話等, 日々多くの業務を行っている。そのため, 多くのオペレータが顧客との電話を行っており, 1 センターあたりの会話数は膨大なものとなる。そこで必要となるのが, オペレータの顧客との会話に対して適切な評価を行うことであるが, 会話数が膨大であるため人力で行うことが困難であるという問題がある。そこで, 本論文では, 音声感情データの解析によって, オペレータと顧客の感情状態の遷移を抽出し, 優れた会話群をもつクラスタを発見することで, 応対評価の効率化のための課題を明確にする。

キーワード 時系列データ, 感情解析, コールセンター

## 1 はじめに

企業と顧客の接点の一つとして, 電話での応対を行うコールセンターがある。コールセンターでは, 日々多くの業務を行っており, その業務はサポート対応や通信販売の申し込み, 新規顧客の獲得や, マーケティング調査など多岐にわたる。また, 応対の良し悪しが顧客の好感度を決める大きな要因となる場合がある。商品を注文する際や, 問い合わせの際に担当するオペレータの応対が良ければ, 顧客の好感度が上がり, 再購入や知人への紹介につながるが, 応対が悪ければ, 好感度が低下し, 解約や悪い口コミの伝播にも繋がりがかねない。このように, 顧客と直接応対を行うコールセンターは, 企業イメージを作る重要な要素である。そのため, 顧客満足度向上を目的として応対品質の向上や, 受電数に合わせたオペレータ数の確保が必要となる [1]。

その中でも, 応対品質確保のためには新オペレータの教育及び既存オペレータの適切なマネジメントが不可欠である。

しかし, オペレータのマネジメントを行う際に各オペレータの応対を評価する必要があるが, 会話数が膨大で人力で評価することはコストが高い。会話の中には通信販売の申込みなど, 購入/非購入の応対結果が明確なものもあるが, すべての業務が明確な応対結果を持つわけではない。例えば, 既存顧客からの問い合わせのみの窓口などが例として上げられる。

また, 通信販売において結果として非購入になった際でも, オペレータの応対が良く顧客が満足していた場合には, 再度購入を検討し, その後購入に至る場合があるためそのような会話は適切に評価されなければならない。応対結果以外の会話の評価手法として, 顧客に会話毎にアンケートを取ることが考えられる。しかし, 会話毎にアンケートを取るとはコストがかかり, 顧客の協力を得られなかったり, アンケートを早く終わらせるために適当に評価することがあるなど, 回答がどの程度信頼できるか不明である。

そこで, 本論文では音声感情値を用いて, 感情状態の遷移に着目しオペレータの応対評価が行えないかと考えた。具体的に

は, オペレータの感情遷移と顧客の感情遷移を解析し, 似た遷移を持つオペレータのクラスタを抽出し, 各クラスタの特徴を見る。それによって, 良い応対の会話があつまったクラスタがあれば, そのクラスタに属する, 応対評価が未知のデータに対しても応対評価を行うことができるのではないかと考えた。感情遷移の抽出には隠れマルコフモデル (Hidden Markov Model, HMM [5]) を用い, 多様な感情パラメータ値の変化を隠れ状態の遷移として捉える。隠れ状態遷移パターンのクラスタリングは, 隠れ状態遷移を遷移確率行列に変換し, k-means 法を用いて行った。

本論文は以下のように構成する。第2章では, コールセンター並びに本論文で使用している技術の関連研究について述べる。第3章では, 本論文の問題定義を行う。第4章では, 提案手法について述べる。第5章では実験, 第6章では本論文のまとめについて述べる。

## 2 関連研究

近年, コールセンター業務に対して様々な観点から分析が行われている。例えば, 応対内容のテキストを用いたものでは, Mishne ら [2] などが挙げられる。Mishne らは, 会話内容を音声認識を用いて文字におこし, テキストマイニングによって会話の重要度を特定し, 問い合わせに対する解決策を検索するシステムを提案した。

また, オペレータの離職抑制を目的とした事例として, オペレータの属性や勤怠, パフォーマンス等を用いたものがある [3]。これは, コールセンターで働くオペレータの勤怠状況やオペレータの属性を用いて, 在職/退職データを教師データとして離職しそうな人を発見するモデルを構築するものである。そして, 毎月の全国の拠点の本番データに対してモデルを適用した結果から, 離職しそうな人に適切な対処をすることで離職数を抑制する。しかし, この事例ではモデルの構築のため様々なデータを収集する必要があり, データ準備のコストが高い。そこで, 本論文では, 会話の音声から自動的に生成された音声感情データの

みを使用することで、データ準備のコストを下げている。

感情解析の分野では、Ching ら [4] は、Yelp のレストランレビューに対して、アスペクトベースの感情解析と、線形回帰を用いて顧客満足度を予測し、顧客満足度の維持および改善を行った。

本論文で使用している隠れマルコフモデル (Hidden Markov Model, HMM [5]) は、観測されない隠れた状態を持つマルコフ過程を用いたものである。  $k$  個の潜在状態を持つ隠れマルコフモデル  $\theta = (A, B, \Pi)$  は次の 3 項目で定義される。

- $A = \{a_{ij}\}_{i,j=1}^k$  : 状態遷移確率分布
- $B = \{b_{i(x)}\}_{i=1}^k$  : 出力確率分布
- $\Pi = \{\pi_i\}_{i=1}^k$  : 初期状態確率分布

隠れマルコフモデルは音声合成や設備状態管理、バイオインフォマティクス、音楽予測などに使用されている。応用例として、音声合成では酒向ら [6] などがある。これは、スペクトルと基本周波数パターンを HMM によりモデル化し、自然な歌声を合成するものである。設備状態管理では、小林ら [7] は、複合的隠れマルコフ劣化モデルで道路舗装の劣化に関して表現した。バイオインフォマティクスの分野では、Manogaran ら [8] は、HMM を用いてゲノム全体の DNA コピー数の変化を検知するアルゴリズムを提案した。音楽予測では、Li ら [9] は隠れマルコフモデルを用い、音楽シーケンス予測を行った。

また、Shi ら [10] は、個人の行動パターン等が異なるときにも適用できる、個人ごとにパーソナライズされた HMM モデルの構築を行った。

本論文では、音声感情値を対象とし、その状態遷移を HMM で捉え、解析を行う。

### 3 問題定義

この章では、本論文に使用するデータセットの説明と問題の定義を行う。

表 1 に主な記号と定義を示す。今回用いるデータセットである、 $n$  会話のオリジナルデータ集合を  $\mathcal{D}_o = \{\mathcal{D}_{o1}, \dots, \mathcal{D}_{on}\}$  とする。このオリジナルデータに対して、後述する前処理した後のデータを  $\mathcal{D} = \{\mathcal{D}_1, \dots, \mathcal{D}_n\}$  とする。このとき、 $i$  番目の会話  $\mathcal{D}_i$  には、 $d$  次元のオペレータの会話データ  $C_{iOP}^{(d)}$  と顧客の会話データ  $C_{iCU}^{(d)}$  が含まれている。次元  $d$  はデータにおけるパラメータの数を表し、 $i$  番目のオペレータ・顧客の会話には  $l_{iOP}, l_{iCU}$  個のセグメントが含まれている。よって、各会話は  $l$  個のセグメントにおける観測値を持つ  $d$  次元の時系列データとして表される。この観測値は  $d$  個の感情パラメータの値を持つ。感情パラメータの一部とその説明を表 2 に示す。また、各会話はオペレータ名 (OPname) と会話開始日時 (Date) をメタデータとして持つ。

#### 定義 1(クラスタリング結果)

会話をクラスタリングした結果を  $\mathcal{F} = \{f_1, \dots, f_n\} (0 \leq f_i < c_k)$  と表し、 $f_i$  は  $i$  番目の会話がクラスタ  $f_i$  に属することを表す。例えば、1 番目の会話がクラスタ 2 に、2 番目の会話がクラ

表 1: 主な記号と定義

記号	定義
$\mathcal{D}_o$	$n$ 会話のオリジナルデータ集合 $\mathcal{D}_o = \{\mathcal{D}_{o1}, \dots, \mathcal{D}_{on}\}$
$\mathcal{D}$	$n$ 会話の前処理後のデータ集合 $\mathcal{D} = \{\mathcal{D}_1, \dots, \mathcal{D}_n\}$
$\mathcal{D}_i$	$i$ 番目の会話データ $\mathcal{D}_i = (C_{iOP}^{(d)}, C_{iCU}^{(d)})$
$C_{iOP/CU}^{(d)}$	$i$ 番目の $d$ 次元のオペレータ/顧客の会話データ $C_{iOP/CU}^{(d)} = \{s_1^{(d)}, \dots, s_{l_{iOP/CU}}^{(d)}\}$
$d$	会話データのパラメータ数 (データ次元数)
$l_{iOP}, l_{iCU}$	$i$ 番目のオペレータ/顧客の会話のセグメント数
$s_k^{(d)}$	会話内における $k$ 番目のセグメントの $d$ 次元の観測値
$\mathcal{M}$	会話のメタデータ集合 $\mathcal{M} = \{\mathcal{M}_1, \dots, \mathcal{M}_n\}$
$\mathcal{M}_i$	$i$ 番目の会話のメタデータ $\mathcal{M}_i = (\text{OPname}, \text{Date})$
$\mathcal{Z}$	HMM 状態系列集合 $\mathcal{Z} = \{(Z_{1OP}, Z_{1CU}), \dots, (Z_{nOP}, Z_{nCU})\}$
$Z_{iOP/CU}$	$i$ 番目の会話のオペレータ/顧客の HMM 状態系列 $Z_{iOP/CU} = \{z_1, \dots, z_{l_{iOP/CU}}\} (0 \leq z_k < h_k)$
$h_k$	HMM 隠れ状態数
$\mathcal{Z}_{tm}$	HMM 状態遷移確率行列集合 $\mathcal{Z}_{tm} = \{(Z_{tm1OP}, Z_{tm1CU}), \dots, (Z_{tmnOP}, Z_{tmnCU})\}$
$Z_{tm iOP/CU}$	$i$ 番目の会話のオペレータ/顧客の 大きさ $h_k \times h_k$ の HMM 状態遷移確率行列
$\mathcal{F}$	クラスタリング結果 $\mathcal{F} = \{f_1, \dots, f_n\} (0 \leq f_i < c_k)$
$c_k$	クラスタ数
$\mathcal{E}$	オペレータ・顧客間の類似度 $\mathcal{E} = \{e_1, \dots, e_n\}$

表 2: 感情パラメータの一部とその説明

パラメータ	説明
Energy	エネルギーの大きさを示す 悲しみや疲労感、快適さ、精神的なさま、などを示す
Stress	緊張を示す
Embarrassment	不快感を示す
Hesitation	快適さを示す
Uncertainty	自信の度合いを示す
Excitement	興奮や高揚を示す
Concentration	集中の度合いを示す

スタ 3 に、3 番目の会話がクラスタ 0 に分類されたとき、クラスタリング結果は、 $\mathcal{F} = \{2, 3, 0\}$  となる。

また、各会話におけるオペレータと顧客の HMM 遷移確率行列の類似度を  $\mathcal{E} = \{e_1, \dots, e_n\}$  とする。ここで、 $e_i$  は  $i$  番目の会話におけるオペレータと顧客の HMM 遷移行列の類似度を平均二乗平方根誤差 (RMSE, Root Mean Square Error) で算出したものである。

本論文の目的は、 $n$  個の会話データ集合  $\mathcal{D} = \{\mathcal{D}_1, \dots, \mathcal{D}_n\}$  が与えられたとき、会話のクラスタリング結果  $\mathcal{F} = \{f_1, \dots, f_n\} (0 \leq f_i < c_k)$  とオペレータ顧客間の遷移確率行列の類似度  $\mathcal{E} = \{e_1, \dots, e_n\}$  を求めることで、各会話の分類とクラスタの特徴を発見することである。これを、以下のように細分化して問題に取り組む。

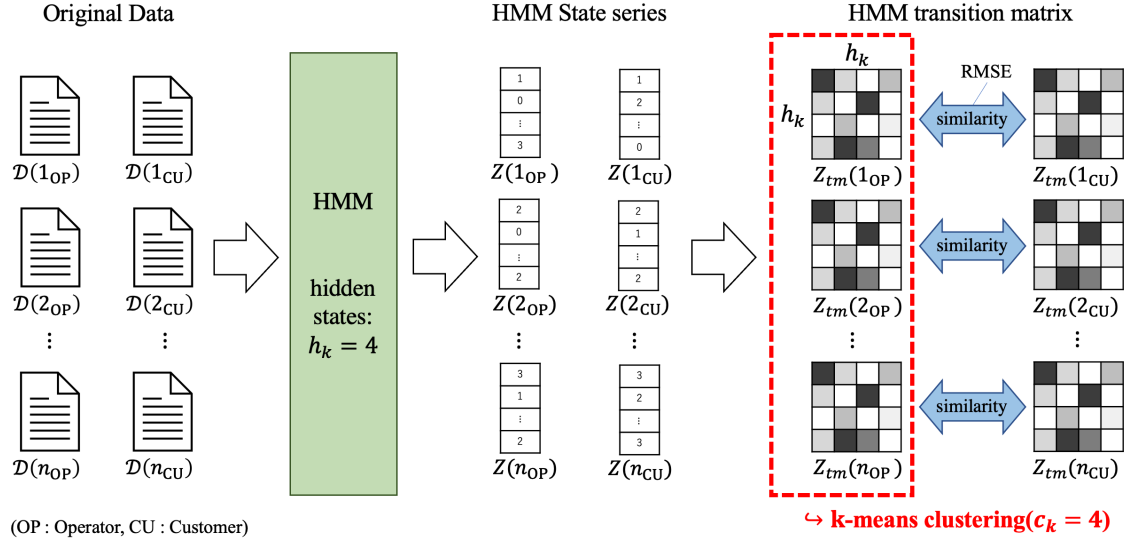


図 1: 提案手法

#### 問題 1(前処理)

オリジナルデータ集合  $\mathcal{D}_o = \{\mathcal{D}_{o1}, \dots, \mathcal{D}_{on}\}$  を  $\mathcal{D} = \{\mathcal{D}_1, \dots, \mathcal{D}_n\}$  へと変換する. 会話データに含まれる  $d$  個のパラメータはそれぞれ取りうる範囲が異なるため, 取りうる値の範囲を揃える正規化処理を行う.

#### 問題 2(データの要約 1)

会話  $C_{iOP/CU}^{(d)}$  が与えられたとき,  $h_k$  個の状態を持つ HMM を用いて HMM 状態列  $Z_{iOP/CU}$  ( $0 \leq z_k < h_k$ ) を求める. これにより, 1 次元の時系列データとなり, 各感情パラメータの値を要約することができる.

#### 問題 3(データの要約 2)

HMM 状態列  $Z_{iOP/CU}$  が与えられたとき, 大きさ  $k \times k$  の状態遷移確率  $Z_{tm_{iOP/CU}}$  を求める.  $Z_{tm_{iOP/CU}}[a, b]$  ( $0 \leq a, b < h_k$ ) は状態  $a$  から状態  $b$  への遷移確率を表す. すべての会話の状態遷移確率行列集合を  $z_{tm}$  とする. これにより, 長さの異なる会話を同じサイズの配列としてクラスタリングを行う.

#### 問題 4(会話のクラスタリング)

HMM 状態遷移確率行列集合  $z_{tm}$  が与えられたとき,  $\mathcal{Z}$  に含まれるオペレータの会話データを対象にクラスタリングを行い, クラスタリング結果  $\mathcal{F} = \{f_1, \dots, f_n\}$  ( $0 \leq f_i < c_k$ ) を得る.

#### 問題 5(オペレータ/顧客の類似度算出)

HMM 状態遷移確率行列集合  $z_{tm}$  が与えられたとき, 各会話におけるオペレータと顧客の HMM 状態列の類似度を RMSE で算出した,  $\mathcal{E} = \{e_1, \dots, e_n\}$  を得る.

## 4 提案手法

本論文では, オペレータ/顧客の音声感情データが与えられたとき, オペレータのデータで会話をクラスタリングし, 応対評価を行う手法を提案する. 第 3 章で設定した問題に沿って, 提案

手法を述べる. 図 1 に提案手法の流れを示す.

#### 4.1 手順 1(前処理)

オリジナルデータ集合  $\mathcal{D}_o = \{\mathcal{D}_{o1}, \dots, \mathcal{D}_{on}\}$  において, 各パラメータの定義域が異なることを解消するため, 各列に正規化を行う. まず, すべての会話データを縦に結合した場合の各列について最大値, 最小値を求め,  $Y = (X - x_{min}) / (x_{max} - x_{min})$  で正規化をする. ( $X$ :元データ,  $Y$ :変換後のデータ) これにより, 各列は最大値 1, 最小値 0 のデータとなる.

#### 4.2 手順 2(データの要約 1)

正規化したデータを隠れマルコフモデルで要約する. 具体的には, オペレータ/顧客の会話データをモデルに入力し, 各セグメントの観測値が  $h_k$  個の状態のうち, どの隠れ状態に属するかを示す状態系列  $Z_{iOP/CU} = \{z_1, \dots, z_{l_{iOP/CU}}\}$  ( $0 \leq z_k < h_k$ ) を得る.

#### 4.3 手順 3(データの要約 2)

HMM 状態系列集合  $\mathcal{Z} = \{(Z_{1OP}, Z_{1CU}), \dots, (Z_{nOP}, Z_{nCU})\}$  が与えられたとき, この状態系列は各会話によって長さが異なるため, さらに要約を行う. 具体的には, 状態系列における状態  $a$  から状態  $b$  への遷移確率を求め ( $0 \leq a, b < h_k$ ), 状態遷移確率行列集合  $\mathcal{Z}_{tm} = \{(Z_{tm1OP}, Z_{tm1CU}), \dots, (Z_{tmnOP}, Z_{tmnCU})\}$  を作成する. 一つの状態遷移確率行列  $Z_{tm_{iOP}}$  は  $h_k \times h_k$  の大きさを持ち,  $a$  行は状態  $a$  における各状態への遷移確率を表す.

たとえば, 状態系列  $Z_{iOP} = \{0, 0, 1, 2, 2, 3, 0\}$  が与えられた場合, 状態 0 からの状態の変化をみると,  $(0 \rightarrow 0)$  が 1 回,  $(0 \rightarrow 1)$  が 1 回である. 同様に  $(1 \rightarrow 2)$ :1 回,  $(2 \rightarrow 2)$ :2 回,  $(2 \rightarrow 3)$ :1 回,  $(3 \rightarrow 0)$  が 1 回であるから, 生成する状態遷移確率行列  $Z_{tm_{iOP}}$  は以下のようになる.

$$Z_{tmOP} = \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

#### 4.4 手順 4(会話のクラスタリング)

手順 3 で生成した状態遷移確率行列のうち、オペレータの状態遷移確率行列を用いて会話のクラスタリングを行う。まず、状態遷移確率行列を長さ  $h_k \times h_k$  の 1 次元のベクトルに平らに加工し、k-means 法によりクラスタリングを行う。このとき状態数を  $c_k$  に設定し、各会話におけるクラスタリング結果  $\mathcal{F} = \{f_1, \dots, f_n\} (0 \leq f_i < c_k)$  を得る。これにより、オペレータが似た状態遷移を取る会話群を得ることができる。

#### 4.5 手順 5(オペレータ/顧客の類似度算出)

手順 3 で生成した状態遷移確率行列において、オペレータと顧客の状態遷移確率行列の類似度を求める。具体的には、 $n$  個の各会話においてオペレータと顧客の状態遷移確率行列の各値を用いて、平均二乗平方根誤差 (RMSE)  $\mathcal{E} = \{e_1, \dots, e_n\}$  を計算する。この値が大きければ類似度が低く、異なる状態遷移を行っていることを示し、小さければ類似度が高く、類似の状態遷移パターンであることを示す。クラスタ毎のこの類似度の分布より、クラスタ分類が効果的に行えたかどうかを評価する。

## 5 実験

本章では、提案手法を用いて実データに対して実験を行った結果について述べる。

### 5.1 データセット

データセットは、コールセンターにおけるオペレータと顧客の 1,577 会話を含む音声感情値データである。データに含まれるオペレータ人数は計 12 人である。会話内容は、通信販売における定期購入の既存顧客からの入電対応業務であり、各会話には応対結果として、[(解約阻止) 成功, 問い合わせ, (解約阻止) 失敗] の 3 種類のラベルが付与されている。

### 5.2 実験条件

本実験では、隠れマルコフモデルの状態数  $h_k$  を 4, k-means のクラスタ数  $c_k$  を 4 とした。

### 5.3 結果

#### (1) クラスタ毎の RMSE 値分布

データをクラスタリングした結果を図 2, 図 3 に示す。図 2 は各クラスタにおける会話数を表し、横軸はクラスタ番号、縦軸はそのクラスタに含まれる会話数である。図 3 は各クラスタにおける RMSE のヒストグラムを表す。横軸は RMSE 値であり、縦軸はその階級に含まれる会話数である。ヒストグラムの階級幅は 0.05 である。

#### (2) クラスタ毎のラベル分布

分類された会話のクラスタ毎のラベルの分布を図 4, 表 3 に示す。縦軸は分類されたクラスタを示し、横軸はクラスタに属

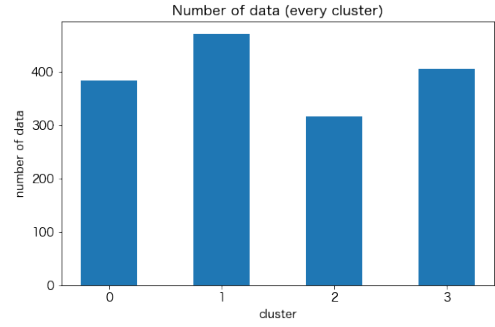


図 2: 各クラスタの会話数

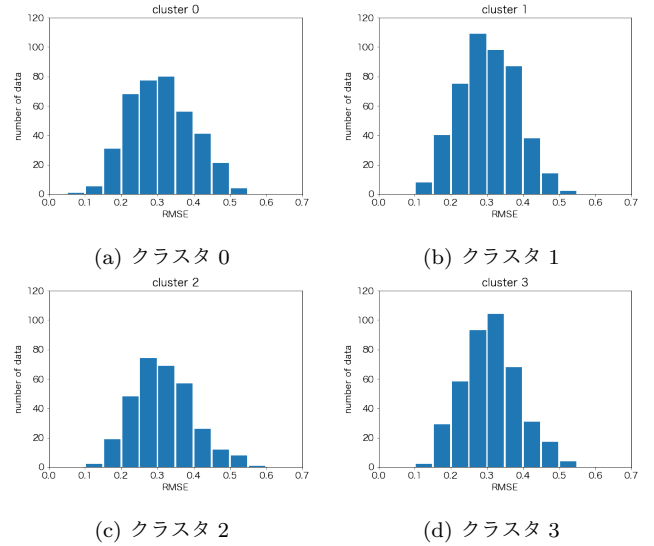


図 3: 各クラスタの RMSE ヒストグラム

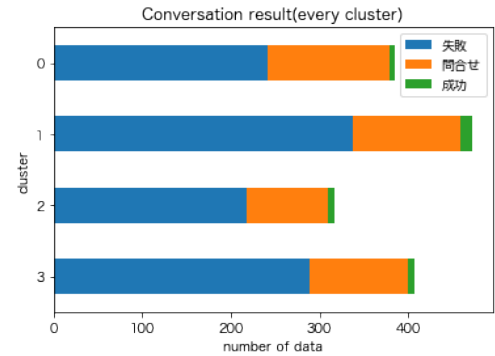


図 4: クラスタ毎のラベル分布

表 3: クラスタ毎のラベル分布

クラスタ	失敗	問合せ	成功
0	241	137	6
1	337	122	12
2	218	91	7
3	289	111	6

する会話数である。

(3) オペレータ毎のクラスタ分布/クラスタ毎のラベル分布提案手法を用いてデータをクラスタリングした結果のオペレータ毎のクラスタ分布と、クラスタ毎のラベル分布を図 5 に

示す。

図5において、1人のオペレータは2本の棒グラフで表されており、グラフの横軸は会話数である。

上段は、当該オペレータの会話が各クラスに割り当てられた数を表している。下段は、当該オペレータの各クラスに割り当てられた会話のうち、応対結果ラベル毎の分布を示す。

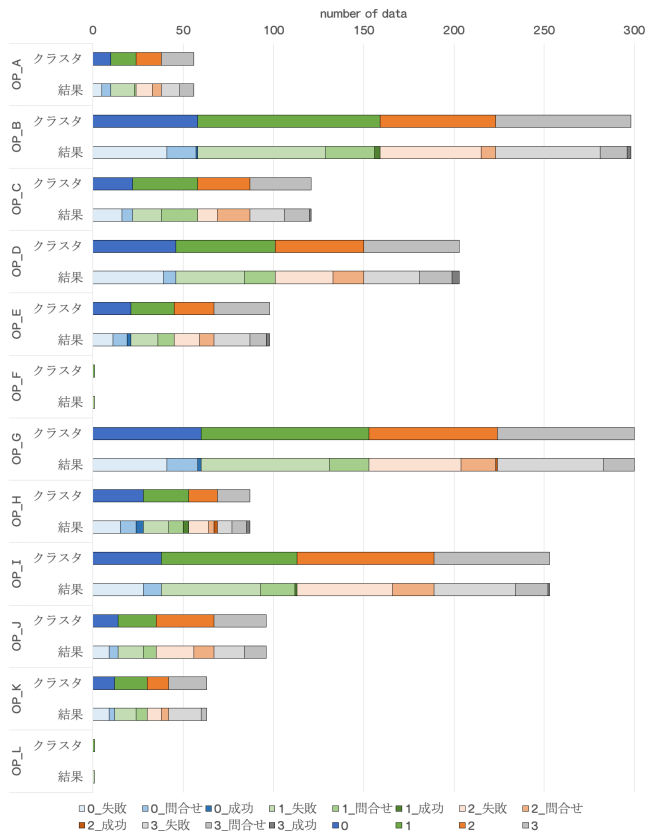


図5: オペレータ毎のクラス分布/クラス毎のラベル分布

## 5.4 考察

### (1) クラス毎の RMSE 値分布

クラスターリングにより、オペレータの状態遷移が似たものを持つ会話の集合は得ることができたが、各クラスにおける有意な RMSE 分布の差を見ることはできなかった。これは、オペレータと顧客の感情遷移の差の分布がクラス間で違いがあまりないことを示す。

原因としては次の3点が考えられる。(a) パラメータ数  $d$  が適切でなく、HMM の状態を上手く捉えられていない。(b) HMM 状態数  $h_k$  が適切でなく、感情状態の遷移が上手く捉えられていない。(c) クラス数  $c_k$  が適切でなく、会話のクラスが上手く分類できていない。

(b), (c) に関しては、 $h_k, c_k$  を 2,4,8,16 と変化させて実験をおこなったが、ヒストグラムで大きな差を得ることができなかった。(a) に関しては、値の変動が他のパラメータと比較して明確に少ないパラメータは既に削除して行ったが、感情把握に適切なパラメータ選定をさらに行う必要があると考えられる。

### (2) クラス毎のラベル分布

クラス毎のラベル分布より、どのクラスにおいてもラベル毎の割合の有意な差を見ることはできなかった。このことから、すべてのオペレータの会話を分類した結果では、解約阻止成功となった会話の感情遷移パターンが特に類似しているというわけではないことが分かる。

(3) オペレータ毎のクラス分布/クラス毎のラベル分布  
オペレータ毎のクラス分布より、各オペレータでのクラス分布差は見られなかったが、クラス毎のラベル分布より、前述のすべてのオペレータの会話での成功ラベル分布では差は見られなかったが、オペレータ毎の成功のラベル分布では特定のクラスに偏っているオペレータがいることが分かった。例えば、OP\_D の場合では、解約阻止成功となった会話はすべてクラス3に属しており、OP\_E の場合では、クラス0と3に属している。これは、オペレータ毎に解約阻止成功の際の感情遷移が類似しており、そのため成功となった会話が同じクラスに属したのではないかと考えられる。

## 6 終わりに

本論文では、コールセンターの音声感情データを解析することにより、会話の応対評価を行うという問題に取り組んだ。

提案手法では、最初に隠れマルコフモデルを用いて感情パラメータの値の遷移を隠れ状態の遷移として取得。次に、得られた隠れ状態の遷移から状態遷移確率行列を作成。最後に、作成した状態遷移確率行列を用いてオペレータの各会話の感情遷移の類似性を k-means クラスターリングを用いてクラス分類を行うことで、優れた会話集合を持つクラスを発見する手法を提案した。

実験では、提案手法を用いて実際のコールセンターの音声感情データに対して解析をおこない、3つの観点からクラスターリングの評価を行った。RMSE 分布では、クラス間の有意な差を得ることができず、また、オペレータ毎のクラス分布でも個人差は見られなかったが、オペレータのクラス毎のラベル分布では、オペレータ毎に成功会話のクラス分布に違いがあり、提案手法により分類が行えていることが分かった。

今後は、解析を行うデータ数の追加や使用パラメータの選定をはじめ、応対結果以外の管理者の主観評価などでのクラスターリング評価を行うことや、オペレータの顧客に対する共感の姿勢を隠れ状態の遷移で発見できないか検討を行い、より有用な分類となるように提案手法の改善・評価実験を行う予定である。

## 謝辞

データのご提供、並びに本研究に関して貴重なアドバイスを頂いた CENTRIC 株式会社様に感謝の意を表す。

## 文献

- [1] 伊藤稔, "コールセンターにおけるインバウンド予測", Unisys 技報, Vol.87, No.5, pp. 19-30, 2005.
- [2] Gilad Mishne, David Carmel, Ron Hoory, Alexey Roytman, and Aya Soffer, "Automatic Analysis of Call-center Conver-

sations,” In CIKM, pp. 453–459, October 2005.

- [3] ITmedia, ” コールセンターの退職予備軍を AI で予測し、半年で離職者を半分にできた理由”, ITmedia エンタープライズ, 2018 年 9 月 5 日公開, 最終閲覧日 2020 年 12 月 23 日, <https://www.itmedia.co.jp/enterprise/articles/1809/05/news001.html>
- [4] Michelle Renee D. Ching and Remedios de Dios Bulos, ”Improving Restaurants’ Business Performance Using Yelp Data Sets through Sentiment Analysis” In ICEEG, pp. 62–67, 2019.
- [5] J. G. Wilpon, L. R. Rabiner, C. H. Lee, and E. R. Goldman, ”Automatic recognition of keywords in unconstrained speech using hidden Markov models.”, IEEE Transactions on Acoustics, Speech, and Signal Processing, 38(11):1870–1878, 1990.
- [6] 酒向慎司, 宮島千代美, 徳田恵一, 北村正, ”隠れマルコフモデルに基づいた歌声合成システム”, 情報処理学会論文誌, pp.719–727, 2004
- [7] 小林潔司, 貝戸清之, 大井明, Nguyen Dinh THAO, 北浦直樹, ”データ欠損を考慮した複合的隠れマルコフ舗装劣化モデルの推計”, 土木学会論文集 E1 (舗装工学) ,71 巻, 2 号, pp. 63–80, 2015
- [8] Gunasekaran Manogaran, V. Vijayakumar, R. Varatharajan, Priyan Malarvizhi Kumar, Revathi Sundarasekar, Ching-Hsien Hsu, ”Machine Learning Based Big Data Processing Framework for Cancer Diagnosis Using Hidden Markov Model and GM Clustering”, Wireless Personal Communications volume 102, pp.2099-2116, 2018
- [9] Tao Li, Minsoo Choi, Kaiming Fu, Lei Lin, ”Music Sequence Prediction with Mixture Hidden Markov Models”, IEEE International Conference on Big Data, pp.6128-6132, 2019
- [10] Hongzhi Shi, Chao Zhang, Quanming Yao, Yong Li , Funing Sun, Depeng Jin, ”State-Sharing Sparse Hidden Markov Models for Personalized Sequences”, the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, July , pp. 1549-1559, 2019