

# 潜在クラスモデルに基づく ビジネスチャットアプリ上の従業員コミュニケーション分析

齊藤 芙佑<sup>†</sup> 山極 綾子<sup>‡</sup> 楊 添翔<sup>‡</sup> 後藤 正幸<sup>‡</sup>

<sup>†</sup> 早稲田大学大学院創造理工学研究科 〒169-8555 新宿区大久保 3-4-1

E-mail: <sup>†</sup> fuyu-saitou@fuji.waseda.jp, <sup>‡</sup> saxophone\_0105@ruri.waseda.jp,

<sup>‡</sup> you\_tensyou@akane.waseda.jp, <sup>‡</sup> masagoto@waseda.jp

**あらまし** 近年, 多くの企業では, 簡潔な文章で密なコミュニケーションが可能なビジネスチャットアプリが浸透しつつあり, これにより蓄積されているデータを戦略的な人事施策での意思決定・実行に活用する動きがある. チャットの返信時間には, ユーザ間の関係性がコミュニケーション特性として表れやすい. よって, ビジネス展開に大きな影響を与える職場の関係性分析を目的とした, 返信時間によるコミュニケーション特性分析は有用でと考えられる. 本研究では, 発言者, 受信者, 返信時間の関係性を定量的に表現する潜在クラスモデルを提案し, 返信時間の観点から二者間のコミュニケーション特性を分析する方法論を示す. 観測変数の背後に確率的に生起する複数の潜在変数を導入することで, ユーザの多様性を考慮しつつ高次元データを解釈することができる. 最後に, 提案モデルを実データに適用し, その有用性を検証する.

**キーワード** Human Resource, 潜在クラスモデル, 確率モデル, 関係性分析, 行動評価分析

## 1. 研究背景と目的

近年, 多くの企業では, 業務上必要な社員同士のコミュニケーションや強調作業をサポートするツールとして, 端的な文章で密なコミュニケーションが可能なビジネスチャットアプリが浸透しつつある. これらのアプリに蓄積されているコミュニケーションデータも貴重な経営資源と考えることができるため, 戦略的な人事施策での意思決定・実行に活用する動きがある.

チャットでは, 電子メールなどの従来のビジネスコミュニケーションツールによるやり取りとは異なり, 受信から返信までの時間 (以下, 返信時間) がユーザごとに大きく異なっているといった点で, コミュニケーション特性が表れやすい. さらに, 「相手に信頼してもらいたいほど迅速な返信をする」, 「心地よい会話をするために発言の間隔を相手と合わせる」など, 発言者が受信者と築きたい関係性によっても返信時間は大きく変化すると考えられる. 特に, ビジネスチャットアプリでのビジネス関係についてのやり取りでは, 特定の人との情報共有を重要と考え積極的に返信するなど, 職場で築かれている関係性が大きく表れていると考えられる. 一方, 同じ職場でチームを構成する従業員間の関係性は, この職場チームの生産性やビジネス展開に大きな影響を与えると考える. 従って, 従業員間のコミュニケーション特性について, 返信時間の観点からビジネスチャットアプリ上のデータの分析を行うことで, 職場のチーム作りや組織マネジメントに有用な知見が得られると考えられる. また, 企業における人事領域での活動においては, 個人に焦点を当てた

具体的な数値に基づくミクロな分析が必要である. つまり, 関係を築き得る最小単位である二者間の関係性について詳細に把握することが必要とされているが, これは大まかな把握より難易度が高く, 作業負担も大きいと想定されるため, その効率的な実現が望まれている.

ここで, 発言者の対応は受信者が誰かということや話題に影響を受けやすく, また, 発言者の発言時の状態にも影響を受けると推測できる. つまり, あるユーザの返信時間の傾向は相手や時間帯などによって変化すると考えられ, 様々な特性を持つコミュニケーション特性が混在していると言える. このような統計的な性質の異なるグループの集合体で与えられているデータに対する有効な分析手法として, 潜在クラスモデル [1] が知られている. このモデルは, 観測データ間の背後に潜在的な離散変数の存在を仮定し, データ間の関係性を潜在クラスと紐付けて表現することで, 異質なデータを潜在クラスという軸で解釈することを可能にする. また, ユーザが複数の異なる潜在クラスに所属する度合いを所属確率により表すことで, 各ユーザの多面性を表現することも可能であり, 対象データの分析に適していると考えられる.

以上より, 本研究では, 発言者, 受信者, 返信時間の関係性を定量的に表現する潜在クラスモデルを提案し, 返信時間の観点から二者間のコミュニケーション特性を分析する方法を示す. さらに, 提案モデルを実企業のチャット履歴データに適用することで, 従業員間のコミュニケーション特性の傾向と各組の関係性を

分析し、その有用性を検証する。加えて、これらの分析結果をもとに、人事領域における具体的な施策について考察し、提案手法の有用性を示す。

## 2. 準備

### 2.1. ビジネスチャットアプリとピープルアナリティクス

ビジネスチャットアプリは、ビジネス上での利用が想定されたコミュニケーションツールである。その特徴として、複数人での端的な文章を用いた即時性のある会話や、様々なデバイスの利用により場所と時間を限定しない会話が可能であることなどが挙げられる。そのため、従来用いられてきた電子メールなどのコミュニケーションツールに比べ、双方向での効率的な情報共有が可能となる。これにより、コミュニケーションの大幅な質向上や多様な働き方の実現による業務効率化、ひいては生産性向上が期待できる新たなコミュニケーションツールとして注目され、近年多くの企業で導入されている。

従来の人事領域では、人事部門社員の主観的な判断に頼った意思決定が多く行われており、これに起因する問題も非常に多かった。一方、近年社員の労働やコミュニケーションに関するデータを収集することが技術的に可能となっており、蓄積されたデータに基づくピープルアナリティクスが注目されている。これは、

「人」に関連する多種多様な情報や数字を収集・統計解析し、理解する手法である。労働市場や経営環境の大きな変化に対応した、人事における意思決定や業務効率化、組織課題の解決などへ活用されている。さらに、単一の人事データだけではなく、社内のさまざまなデータも連携させ、より複合的なデータ活用を推進することが求められている。これら、データ分析に基づき得られる客観的な判断材料による採用の効率化、適正な人材配置、離職率の低減などにおける戦略的な人事施策の実行に繋がることが期待されている。

### 2.2. 問題設定

上述のような、近年のビジネスチャットアプリの浸透とピープルアナリティクスの注目に伴い、チャットデータを用いた企業向けの新しい分析サービスが提供されるようになってきている。しかし、その指標として用いられているのは、会話数や会話の相手の人数などの単純な統計量であることが多い。そのため、様々な特長量を複合的に取り入れるなど、チャットの特徴を活かした分析モデルが望まれている。

チャットの特徴として、場所と時間を限定せず、複数人との端的な文章で、即時性のある会話が可能という「手軽さ」がある。この点を考慮すると、チャットでは、相手への返信に対し「優先度が高く、迅速に返信するか」、あるいは「優先度が低く、後回しにするか」

は返信時間に大きな差異として表れると考えられる。すなわち、返信時間には相手との関係性が特に大きく表れやすいと言える。そこで、本研究ではこの点に注目し、返信時間という切り口から、チャットの特徴を生かした分析を行う。

### 2.3. 潜在クラスモデル

複数の異なる特性を持つデータを分析するモデルとして、潜在クラスモデルが知られている。観測変数の背後に観測できない潜在的な変数の存在を仮定することで、異質なデータが混在する場合においても観測変数間をモデル化することが可能となる。また、本研究の分析対象データのような、異質のグループが混在するデータに対して、潜在クラスという軸での解釈を可能にする。潜在クラスモデルは、購買履歴分析や文書検索問題、画像認識などの情報科学分野にとどまらず、店舗の来店人数予測やファッションコーディネート推薦、計量社会学など、幅広い分野で応用されている。

最もシンプルで代表的な潜在クラスモデルの一つである Aspect Model [1]は、2つの観測変数の間に潜在クラスを仮定し、データが潜在クラスから確率的に生起することを許容する確率モデルである。変数の共起関係を潜在クラスによる条件付確率分布で表現することで、各事象における異なる特徴の混在という多様性の表現が可能となる。ここで、 $I, J$ 個からなる観測変数集合をそれぞれ  $\mathbf{x} = \{x_i; 1 \leq i \leq I\}$ ,  $\mathbf{y} = \{y_j; 1 \leq j \leq J\}$  2つの観測変数集合の背後に仮定する  $K$  個の潜在クラス集合を  $\mathbf{z} = \{z_k; 1 \leq k \leq K\}$  と定義する。このとき、AM の確率モデルは以下の式(1)で表される。

$$p(x_i, y_j) = \sum_{k=1}^K p(z_k) p(x_i | z_k) p(y_j | z_k) \quad (1)$$

## 3. 提案手法

### 3.1. 提案モデルへの着想

上述のように、チャットの「手軽さ」に着目すると、会話は迅速に展開可能な環境下で行われており、相手への返信に対し「優先度が高く、迅速であるか」、あるいは「優先度が低く、緩慢であるか」は返信時間に大きな差異として表れると考えられる。よって、二者間の関係性が表れやすいビジネスチャットアプリ上の返信時間を用いて、その関係性の解釈を行うことが可能であると言える。例えば、迅速に返信を行うユーザほど、迅速な会話の展開を求めていると解釈できる。さらに、双方の返信時間が均等なユーザの組ほど、お互いの関係構築への意識が一致している関係であると解釈できる。そこで本研究では、会話の返信時間を分析し、二者間の関係性について評価することを考える。

なお、ユーザ A からユーザ B に発言があったもとで、その直後の B から A への発言を「B の A への返信」と見なし、返信時間は「A から B への発言」から「B の A への返信」までの所要秒数とする。発言者数・受信者数は共に多く、返信時間の取り得る範囲は広い。ため、分析対象データの取り得る組の数は膨大なものとなり、これらの関係性や多面性、傾向を直接解釈することは非常に困難である。そこで、それらの共起関係を潜在クラスで紐付けることにより表現し、相互に影響しあう発言者と受信者の関係性をグルーピングして解釈するモデルを提案する。これにより、チャットを利用する二者間の関係性について個人の多面性を考慮しつつ定量的に評価することが可能となり、職場のチーム作りや組織マネジメントに役立てることができると期待される。

### 3.2. 提案モデル

本研究では、発言者、受信者、返信時間の関係性を定量的に表現する潜在クラスモデルを提案する。潜在クラスにより、発言者、受信者、返信時間の共起関係を表現することで、提案モデルは「任意の発言者が任意の受信者に対しどのような返信を行うか」に関する分析を可能にする。

いま、発言者集合を  $\mathcal{T} = \{t_l; 1 \leq l \leq Q_{\text{talker}}\}$ 、受信者集合を  $\mathcal{R} = \{r_m; 1 \leq m \leq Q_{\text{receiver}}\}$ 、返信時間を  $s$  と定義し、発言者と受信者、返信時間の間に仮定する  $K$  個の潜在クラス集合を  $\mathcal{Z} = \{z_k; 1 \leq k \leq K\}$  と定義する。このとき、提案する確率モデルを、以下の式(2)で定義する。

$$p(t_l, r_m, s) = \sum_{k=1}^K p(z_k) p(t_l|z_k) p(r_m|z_k) p(s|z_k) \quad (2)$$

このモデルの 1 回の生起事象は、発言者  $t_l$  から受信者  $r_m$  に返信時間  $s$  の返信がされることである。

ここで、 $p(z_k)$  は離散集合  $\mathcal{Z}$  の要素を確率変数とする多項分布 ( $K$  項分布)、 $p(t_l|z_k)$  は  $\mathcal{T}$  の要素を確率変数とする多項分布 ( $Q_{\text{talker}}$  項分布)、 $p(r_m|z_k)$  は  $\mathcal{R}$  の要素を確率変数とする多項分布 ( $Q_{\text{receiver}}$  項分布) と仮定し、返信時間  $s$  の対数値  $\log s$  は正規分布に従うと仮定する。ここで、返信時間  $s$  は以下に表される式(3)に従うと仮定する。なお、 $\mu_k, \sigma_k^2$  はそれぞれ  $k$  番目の潜在クラスにおける平均と分散を示す。

$$p(s|z_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left\{-\frac{(\log s - \mu_k)^2}{2\sigma_k^2}\right\} \quad (3)$$

### 3.3. 学習アルゴリズム

$N$  件の発言者・受信者・返信時間データのうち、 $n$  番目の発言者を  $a_n (\in \mathcal{T})$ 、その時の受信者を  $b_n (\in \mathcal{R})$  とする。また、返信時間を  $c_n$  とすると、 $n$  番目の発言者・受信者・

返信時間の完全データはこれらの共起  $(a_n, b_n, c_n)$  で表現できる。このとき、全  $N$  件の発言者・受信者・返信時間データに対する対数尤度  $LL$  は以下の式(4)で表される。

$$LL = \sum_{n=1}^N \log \sum_{k=1}^K p(a_n, b_n, c_n) \quad (4)$$

各パラメータは、対数尤度関数  $LL$  を最大化するように EM アルゴリズム[2]を用いて、以下の更新式により推定する。

E-Step)

$$\frac{p(z_k|a_n, b_n, c_n)}{\sum_{k=1}^K p(z_k|a_n, b_n, c_n)} = \frac{p(z_k)p(a_n|z_k)p(b_n|z_k)p(c_n|z_k)}{\sum_{k=1}^K p(z_k)p(a_n|z_k)p(b_n|z_k)p(c_n|z_k)} \quad (5)$$

M-Step)

$$p(z_k) = \frac{\sum_{n=1}^N p(z_k|a_n, b_n, c_n)}{N} \quad (6)$$

$$p(t_l = a_n|z_k) = \frac{\sum_{n=1}^N \delta(a_n, t_l) p(z_k|a_n, b_n, c_n)}{N p(z_k)} \quad (7)$$

$$p(r_m = b_n|z_k) = \frac{\sum_{n=1}^N \delta(b_n, r_m) p(z_k|a_n, b_n, c_n)}{N p(z_k)} \quad (8)$$

$$\mu_k = \frac{\sum_{n=1}^N p(z_k|a_n, b_n, c_n) \log c_n}{\sum_{n=1}^N p(z_k|a_n, b_n, c_n)} \quad (9)$$

$$\sigma_k^2 = \frac{\sum_{n=1}^N p(z_k|a_n, b_n, c_n) (\log c_n - \mu_k)^2}{\sum_{n=1}^N p(z_k|a_n, b_n, c_n)} \quad (10)$$

ここで、 $\delta(x, y)$  は、 $x = y$  である場合に 1、そうでない場合は 0 の値を示すインジケータ関数である。

## 4. 実データ分析

### 4.1. 分析条件

本章では、実データに提案モデルを適用することにより、モデルの有用性を検証する。分析対象データは、Laboratik 社[3]より提供いただいた、某企業の職場における Slack でのチャットコミュニケーションデータセットである。なお、チャットコミュニケーションデータセットの発言日時を用いた返信時間の計算については、就業時間など返信を妨げる要因の影響により、得られる返信時間データの信頼度が下がる恐れがある。そこで、分析結果に悪影響を与えうる要因の影響を除くため、計算は 1 日単位に区切りを行い、発言者データ・受信者データ・返信時間データからなるデータセットを作成した。

分析対象期間は 2018 年 11 月 2 日から 2019 年 9 月 24 日、データ数は  $N = 98,225$ 、発言者数  $Q_{\text{talker}} = 54$ 、受信者数  $Q_{\text{receiver}} = 54$  である。なお、 $l = m$  のとき  $t_l$  と  $r_m$  は

同一人物を示す．ここで，潜在クラス数 $K$ については，解釈性の高いクラス数を探索的に求め $K = 10$ とした．

#### 4.2. パラメータの解釈

各潜在クラスのパラメータの推定結果として，対数変換した返信時間の平均 $\mu_k$ ，標準偏差 $\sigma_k^2$ ，解釈の際にわかりやすさのために秒換算した $\exp \mu_k$ と確率 $p(z_k)$ を表 1 に示す．

表 1 パラメータ

$z_k$	$\mu_k$	$\sigma_k$	$\exp(\mu_k)$	$P_{z_k}$
$z_1$	1.81	0.83	6.11	$2.36 \times 10^{-4}$
$z_2$	0.95	0.29	2.60	$3.32 \times 10^{-2}$
$z_3$	2.89	2.36	18.05	$3.49 \times 10^{-2}$
$z_4$	3.70	2.64	40.57	$4.22 \times 10^{-2}$
$z_5$	0.50	0.10	1.66	$5.70 \times 10^{-2}$
$z_6$	0.99	0.35	2.69	$9.75 \times 10^{-2}$
$z_7$	0.74	0.19	2.09	$1.11 \times 10^{-1}$
$z_8$	0.90	0.30	2.45	$1.22 \times 10^{-1}$
$z_9$	5.12	1.58	167.76	$1.94 \times 10^{-1}$
$z_{10}$	1.93	1.25	6.90	$3.08 \times 10^{-1}$

また，平均 $\mu_k$ を横軸，標準偏差 $\sigma_k^2$ を縦軸に示したグラフを図 1 に示す．なお，円の大きさは推定確率 $p(z_k)$ の値に比例している． $\exp \mu_k$ が小さいほど迅速な返信を行う潜在クラスと解釈でき， $p(z_k)$ が大きいほどその特徴を持つ組が多く存在すると言える．また，円同士の距離はそれぞれの類似度と見なすことができ，企業全体のコミュニケーション特性を大まかに把握することも可能である．

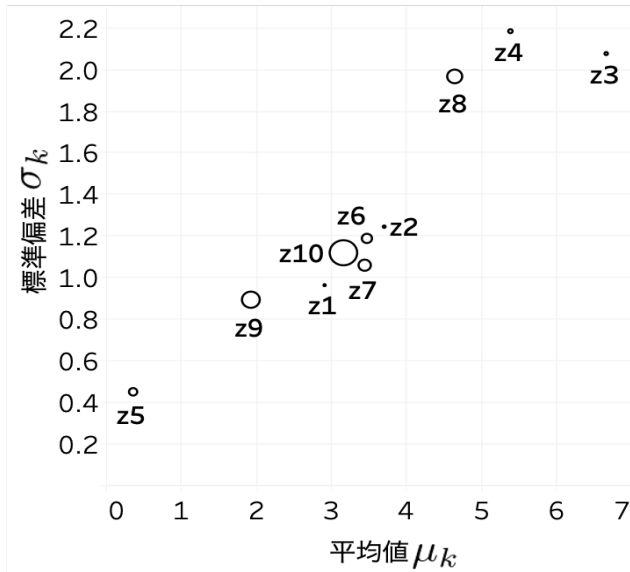


図 1 各潜在クラスに対する推定結果

#### 4.3. 二者間の関係性解釈

二者間の関係性解釈を行う方針として，まず発言者

$t_l$ から受信者 $r_m$ への返信の一方方向について特徴を抽出する．次に，発言者と受信者に当たるユーザを入れ替え，逆方向の返信の特徴を抽出する．これら双方向の特徴を比較することにより，二者間の関係性を解釈する．

$t_l$ が $r_m$ に対して行う返信の特徴が潜在クラス $z_k$ である確率は， $t_l$ が $z_k$ の特徴を持ち，かつ $r_m$ も $z_k$ の特徴を持つ同時生起確率で求められる．よって，一方方向での返信の特徴の解釈を，式(11)で表される $t_l$ と $r_m$ の $z_k$ に対する同時所属確率を計算することによって行う．

$$p(z_k|t_l, r_m) = \frac{p(t_l, r_m|z_k)p(z_k)}{p(t_l, r_m)} \quad (2)$$

発言者としてユーザ 1，受信者として 9 の同時所属確率分布を図 2 に，3 月と 11 月におけるユーザ 1 からユーザ 9 への実際の返信時間の 1 秒毎のヒストグラムを図 3 に示す．図 2 より，ある特定の発言者から受信者への返信が，複数の潜在クラスに確率的に属していることがわかる．すなわち，提案モデルでは発言者との受信者が同一であっても，図 3 に現れているユーザ 1 からユーザ 9 への実際の返信時間が 3 月と 11 月で異なるような，発言時の状況によって異なる特性を持つという多面性を表現できていると考えられる．

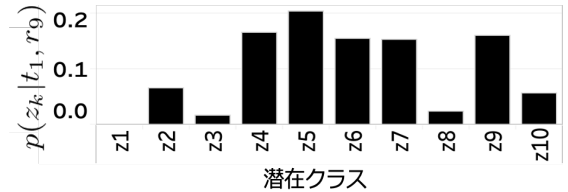


図 2 同時所属確率分布

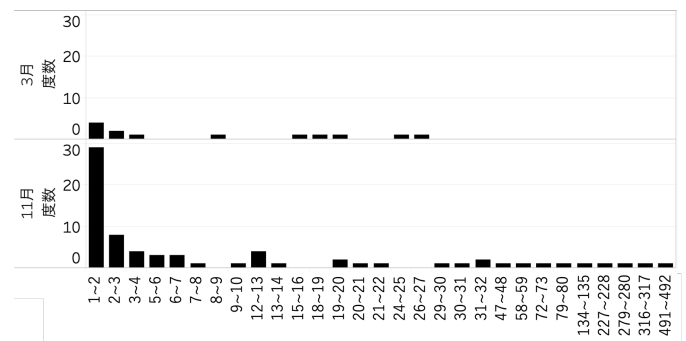


図 3 実際の返信時間の 1 秒毎のヒストグラム

本分析では，同時所属確率が最大となる潜在クラスを求め，コミュニケーション特性を評価する．ユーザ 1 と 3 からなる組とユーザ 4 と 13 からなる組の同時所属確率の結果を表 2 に表す．例えば， $p(z_k|t_1, r_3)$ が最大となる潜在クラスは $z_9$ であることから， $t_1$ は $r_3$ に対してかなり緩慢に返信を行っていると言える．このような返信の特徴解釈を $t_1$ と $r_3$ を入れ替え双方向で行うこ

とにより、以下に示す組の関係性の把握が可能となる。

表 2 同時所属確率

$p(z_k t_l, r_m)$	$z_1$	$z_2$	$z_3$	$z_4$	$z_5$
$p(z_k t_1, r_3)$	$7.8110^{-4}$	$2.5510^{-4}$	$8.7110^{-3}$	$7.0510^{-2}$	$4.1810^{-3}$
$p(z_k t_3, r_1)$	$1.3710^{-4}$	$1.1410^{-2}$	$1.3310^{-1}$	$9.6410^{-5}$	$1.5510^{-2}$
$p(z_k t_4, r_{13})$	$8.2510^{-4}$	$9.36^{-2}$	$7.8410^{-3}$	$4.1410^{-2}$	$5.2810^{-2}$
$p(z_k t_{13}, r_4)$	$4.7010^{-5}$	$2.4910^{-4}$	$2.1010^{-5}$	$1.2010^{-3}$	$2.5510^{-4}$
$p(z_k t_l, r_m)$	$z_6$	$z_7$	$z_8$	$z_9$	$z_{10}$
$p(z_k t_1, r_3)$	$2.1710^{-2}$	$2.9910^{-1}$	$2.3410^{-2}$	<b><math>5.3210^{-1}</math></b>	$3.9110^{-2}$
$p(z_k t_3, r_1)$	$2.1810^{-2}$	$3.2310^{-2}$	$4.0610^{-2}$	<b><math>4.2910^{-1}</math></b>	$3.1610^{-1}$
$p(z_k t_4, r_{13})$	$3.8010^{-2}$	<b><math>6.5710^{-1}</math></b>	$7.8710^{-3}$	$1.0010^{-1}$	$6.1010^{-4}$
$p(z_k t_{13}, r_4)$	$3.9010^{-2}$	$4.0510^{-3}$	$3.2210^{-3}$	$8.6010^{-2}$	<b><math>8.6610^{-1}</math></b>

#### (1) 双方向の関係性の特徴に差がないユーザの組

表 2 より  $p(z_k|t_1, r_3)$  が最大となる潜在クラスは  $z_9$  であり、また  $p(z_k|t_3, r_1)$  が最大となる潜在クラスは  $z_9$  である。このとき、ユーザ 1 とユーザ 3 の双方向でのコミュニケーション特性は高い確率で共通する潜在クラスに所属しており差異がなく、互いが均衡な返信を行う関係を築く組であると解釈できる。

#### (2) 双方向の関係性の特徴に差があるユーザの組

一方、表 2 より  $p(z_k|t_4, r_{13})$  が最大となる潜在クラスは  $z_7$  である。しかし、 $p(z_k|t_{13}, r_4)$  が最大となる潜在クラスは  $z_{10}$  であり異なっている。ここで、 $\exp \mu_k$  は  $z_{10}$  より  $z_7$  の方が小さい値をとるので、ユーザ 4 の方がユーザ 13 より迅速な返信を行うとわかる。すなわち、ユーザ 4 とユーザ 13 は返信の特徴に差があり、会話の展開への貢献に対する意識に差がある組であると解釈できる。

## 5. 考察

### 5.1. モデルの有用性

提案モデルを用いた実データ分析により、企業全体における返信の傾向の把握と、各特徴を持つ返信を行う組の割合の評価、および社員同士の関係性の抽出を行うことが可能となった。従来の関係性評価は評価者の主観など定性的な情報に基づいていたが、本研究では、モデルの解釈の指標として数値データである返信時間を用いることで、コミュニケーション特性を定量的に把握することが可能となっている。

また、潜在クラスモデルは、確率分布を仮定しデータが複数のクラスに所属することを許容するため、発言の対象や時間によって変化するユーザの多面性を表現することが可能である。また、推定したパラメータを用いて、ユーザ 1 人 1 人の分類だけでなく、同時所属確率の算出による各組の関係性の解釈も可能であることが示された。このように、潜在クラスモデルの導入により推定されたパラメータは、非常に応用性が高く、利用価値が高いと言える。

企業における人事領域での活動においては、個人に

焦点を当てた具体的な数値に基づくミクロな分析だけでなく、その解釈の容易さから、それらを大枠で捉えたマクロな視点での分析の両方が重宝されている。提案手法は、二者間のみに着目したコミュニケーション特性の抽出と、潜在クラス間の類似度を用いて企業全体の特徴の大まかな把握が可能となっている。さらに、潜在クラスへの所属確率による返信時間を考慮した発言者と受信者のコミュニケーション特性のクラスタリングが可能である。

### 5.2. 実務への応用

ビジネスチャット以外のデータへの適用例として、プライベートで用いる SNS ツールのチャットデータが挙げられる。個人間の関係性把握が可能となり、アプリ内での「会話が円滑な傾向の相手」の優先的な表示機能による利便性向上や、「関係性分析診断」などの新しいエンターテインメントコンテンツサービスなどにつながれると考えられる。現代は、スマートフォンやインターネットなどの爆発的な普及に伴い、チャット機能を併せ持つ様々な種類の SNS が飛躍的な進歩を遂げており、チャットデータセットの収集が比較的容易でその結果を活用する市場は大きい。このように、提案モデルは応用範囲が広く、マーケットに対するインパクトが大きい施策につながるモデルであると考えられる。

## 6. まとめと今後の課題

本研究では、チャットの返信時間にはユーザごとのコミュニケーション特性が表れやすく、またユーザ間の関係性から与えられる影響が大きいという点に着目し、発言者、受信者、返信時間の関係性を潜在クラスにより定量的に表現し、返信時間の観点からコミュニケーション特性を分析する手法を提案した。さらに、提案モデルを実データに適用し、実際に返信時間に着目した二者間のコミュニケーション特性の分析を行った。その結果として、企業に存在するコミュニケーション特性の潜在クラスを用いた傾向分析、並びに各潜在クラスへの同時所属確率を用いた各組で起こりやすい返信の特性分析が可能であることを示した。また、これらの得られた結果をもとに、人事領域などにおける具体的な施策につながる可能性を検討し、モデルの有用性を示した。

今後の課題としては、言語データなど複合的な他のデータセットとの統合によるモデルの改善や、潜在クラス数の決定方法検討、実際には会話が行われていない二者間についての関係性予測などが挙げられる。

### 謝辞

本研究を行うにあたり、貴重なデータをご提供くだ

さった Laboratik 株式会社様に深く感謝の意を表します.

#### 参 考 文 献

- [1] Hofmann, T. (1999), “Probabilistic Latent Semantic Analysis”, *Proc. of UAI '99*, pp.289–296
- [2] Bishop, C. (2010), “Pattern Recognition and Machine Learning”, *Information Science and Statics*, 2nd printing, Springer
- [3] Laboratik Inc., Laboratik. Retrieved August 8 2020, from <https://laboratik.com/>