

E-066

コールセンターのコールメモと通話を対象とした差分マイニング Using Differential Text Mining to Analyze Speech Dialogues and Call-memos in a Call Center

田村 晃裕† 石川 開† 安藤 真一†
Akihiro Tamura Kai Ishikawa Shinichi Ando

1. はじめに

コールセンターは、企業にとって顧客と直接コンタクトする窓口であるため、近年、様々な企業が、顧客の声を獲得する絶好の場としてコールセンターを重要視している。実際に、多くの企業が、コールセンターのデータを分析する事で、製品やサービスの問題、潜在リスク、顧客ニーズや評価などを抽出し、コールセンターの応対や、製品、サービスの改善を試みている。

通常、コールセンターでは、応対の内容を、定型データ、エージェントにより作成されるコールメモ、通話の録音などの形で記録している。日々蓄積されるこれらのデータは、膨大な量に上るため、人手で分析するのは困難な状況にある。そこで、多くの企業において、膨大なデータから有用な情報を抽出する手段であるテキストマイニングの導入が急速に浸透しつつある[1]。

これまで、コールセンターにおけるテキストマイニングの分析対象は、定型化され処理しやすいという理由もあり、コールメモが中心であった。しかし、コールメモは、エージェントが通話応対をまとめたものであるため、応対内容の要点など、限られた情報しか記述されない。そのため、顧客の生の声や、エージェントの応対上の問題といった、より詳細な内容を分析するには、コールメモの内容だけでは不十分であった。そこで、最近では、顧客とエージェント間の通話を直接分析対象とする方式の研究が盛んになってきている[2][3][4]。しかし、通話自体は冗長性が高く、また、多種多様な情報が混在しているため、通話を単純にマイニングする方法では、欲しい情報が埋もれてしまい獲得できない可能性がある。

そこで、我々は、コールメモから得られる情報は、従来のコールメモを対象にしたマイニングで獲得し、顧客の生の声等のコールメモからは得られない通話特有の情報を、通話からマイニングして獲得する、というアプローチが有効であると考えた。具体的には、通話からその通話に対応するコールメモの内容を除いた差分を生成し、その差分を分析対象としてマイニングを実施する。これにより、通話特有の情報に注目した分析を行う事ができる。本稿では、この二つのテキスト間の差分を対象にマイニングを行う差分マイニングの方法を提案する。そして、提案手法を用いる事で、通話を単純にマイニングする従来の手法では抽出できない、通話特有の知見を獲得できることを示す。

2. テキストマイニング

テキストマイニングとは、大量のテキストデータから新規性のある有用な情報を抽出するための技術の総称であるが、本章では、本稿内の説明、及び、実験で使用す

る具体的なテキストマイニング技術とその分析手順を説明する。この分析は、[5]における“Rule Analysis”である。以降、「特徴語分析」と呼ぶ。

分析対象は、テキスト情報（例：コールメモの「受付内容」、「コメント」）と定型情報（例：コールメモの「受付日時」、「顧客名」）を持つデータを1事例とする事例集合である。分析者は、この特徴語分析により、ある定型情報によって規定されるカテゴリ（例えば、「顧客A」）に特徴的な知見を、あるテキストから発見する事が可能である。

特徴語分析を用いて知見を獲得するまでの手順は以下の通りである。

1. 分析者が分析したいカテゴリとテキスト情報を指定。
2. 分析対象に指定されたテキスト群を構成要素（単語や係り受け等）に分割。本稿では、単語とする。
3. 分割した各構成要素と分析カテゴリの相関度を計算。分析カテゴリに頻出するものほど、また、分析カテゴリ以外のカテゴリに出現しないものほどスコアが高くなる尺度を用いる。実験では ESC[6]を使用した。
4. 3 で計算したスコアの高い構成要素からランキングした表を分析者に提示。したがって、順位が上位の構成要素ほど、分析カテゴリに特徴的なものとなる。
5. 分析者は、順位が上位の構成要素から順に、その構成要素を含む原文を実際に参照しながら、新たな知見の発見を試みる。

上記の 2、3、4 が機械処理部分で、1 と 5 が分析者自身が行う事である。

3. 提案手法

本章では、提案手法である差分マイニング手法について説明する。以下では、分析データを n 個の事例とし、各事例は、通話テキスト（通話をテキスト化したもの）とその通話に対応するコールメモの 2 つのテキストを持つものとする。 i 番目の事例の通話テキストを D_i 、コールメモのテキストを M_i とする。そして、通話テキストの集合を $D=\{D_1, \dots, D_n\}$ 、コールメモの集合を $M=\{M_1, \dots, M_n\}$ とし、通話特有の情報を得るための差分マイニング手法について説明する。提案手法は、事例ごとにテキスト間の差分 $D_i - M_i$ を生成し、その差分集合 $D-M=\{D_1-M_1, \dots, D_n-M_n\}$ を分析対象のテキストとしてテキストマイニングを行う。

以降では、まず、提案手法により通話 D 特有の知見が得られるメカニズムを説明した後、具体的な差分の生成法について述べる。

3.1 差分マイニングのメカニズム

2 節で述べた通り、分析者は、特徴語分析により得られたランキング表の上位の語から順に、それに関する知見を探索していく。この過程において、分析時間は有限で

† 日本電気株式会社 共通基盤ソフトウェア研究所

表1 「至急」に関する表現の不一致の例

通話	コールメモ
[大至急]やってください	[至急]対応依頼
[急ぎ]の件	[至急]連絡
[早く]してほしい	[早め]の対応
[緊急の]対応お願いしたい	[クイック]対応
出来るだけ[早く]	[急ぎ]の対応希望

あるため、順位が上位の語に関する知見は発見できる可能性が高いが、下位になるほど、その語に関する知見は発見しづらくなる。したがって、以下では、差分 **D-M** を対象にマイニングする事で、通話特有の情報を表す構成要素が、特徴語のランキング表において順位が上昇するメカニズムを説明する。

・共通語の消失による差分語のランキング上昇

D_i と M_i の双方に出現する語は、差分 $D_i - M_i$ には出現しない。このような共通語は、**D** 単独を対象としたマイニングで順位が上位であっても、差分 **D-M** のランキング表からは消失する。これにより、共通語ではない **D** 特有の語の順位が繰り上がり、上位に出現する。

・カテゴリに相関する差分語のランキング上昇

ある語が D_i に出現しかつ M_i に出現しない事例、すなわち、差分 $D_i - M_i$ への出現が、分析対象のカテゴリと相関する場合、特徴度は高くなる。このため、分析対象のカテゴリに特徴的な **D** 特有の語の順位が、分析対象のカテゴリと相関の無い差分語と比較して相対的にランクが上昇する。

これら2つのメカニズムにより、従来の **D** 単独を分析対象としたマイニングでは得られなかった、**D** 特有の知見を、**D-M** の差分マイニングで発見することが可能となる。

3.2 差分生成法

本節では、通話テキスト D_i とコールメモ M_i から、それらの差分 $D_i - M_i$ を生成する具体的な手法について説明する。

最も単純に二つのテキスト間の差分を生成する手法として、語の照合に基づく生成法が考えられる。具体的には、まず、テキスト D_i と M_i をそれぞれ形態素解析し、 D_i 中、 M_i 中に含まれる単語を求める。そして、 D_i 中の単語から、 M_i 中の単語と表層が完全一致するものを除いた単語のリストを差分 $D_i - M_i$ とする手法である。以降、本手法を Matching Word Subtraction(MWS) と呼ぶ。

しかし、差分を生成する対象の二つのテキスト(コールメモと通話)間には、多くの同義表現が存在する。実際、顧客とエージェントで使用する表現が異なっている場合が多い。また、同一エージェントにおいても、通話内では、顧客の理解や状況に応じた表現を用いる傾向がある一方、コールメモ内では、関係者向けの簡潔な表現を用いる傾向がある。例として、「至急」に関する、通話とコールメモ内での表現の不一致を表1に示す。MWSでは、このような表現のずれが存在すると、本来同一視されるべき表現が別々の表現として扱われるため、本来ならば差分ではないものが差分として生成されてしまう。その結果、正しい差分が生成できず、その差分を用いる差分マイニングの分析結果に悪影響を与えてしまう。

したがって、MWSで、この同義語問題に対処するためには、二つのテキスト間で使用される表現に対して、同

義表現と正規表現の対を網羅した同義語辞書を用意し、同義表現の正規化を行い、差分を生成する必要がある。以降、同義語辞書を用いて同義表現の正規化を行った上で、MWSにより差分を生成する手法を MWS_{syn} と呼ぶ。

前述のように、同義表現のずれの問題への対策の一つとして、同義語辞書の利用が考えられる。しかし、実際、同義表現を網羅するためのコストは膨大である。そこで、我々は、共起関係に基づいた差分生成方法を提案する。次節で、提案手法である差分生成法について説明する。

3.2.1 語の共起関係に基づく差分生成法(CWS)

本節では、語の共起関係に基づき差分を生成する手法 Co-occurring Word Subtraction(CWS)を提案する。CWSは、差分を生成するテキストの組(D_i, M_i)の集合($i=1, \dots, n$)において、一定以上共起する関係にある単語を共通部分と判定する。これにより、同義語辞書を用意する事なく同義語問題を解決する。 D_i 中の単語 v が、差分 $D_i - M_i$ の構成要素かどうかを判定するための手順は以下の通りである。(ステップ1)

差分を生成するテキストの組(D_i, M_i)の集合 ($i=1, \dots, n$) に対して、単語 v, w の出現に関する相互情報量 $I(v; w)$ を次のように計算する。

$$I(v; w) = \sum_{x \in \{v \in D_i, v \notin D_i\}} \sum_{y \in \{w \in M_i, w \notin M_i\}} p(x, y) \log \frac{P(x, y)}{P(x)P(y)}$$

ここで、 x は、単語 v が D_i 中に出現する、または、出現しないという状態を取り、 y は、単語 w が M_i 中に出現する、あるいは、出現しないという状態を取る。また、 $P(x)$ 、 $P(y)$ 、 $P(x, y)$ は、各状態の出現確率であり、全 n 事例における各状態の出現頻度から最尤推定によって求める。

(ステップ2)

D_i 中の単語 v の差分スコアを次のように計算する。

$$D_Score(v, M_i) = \exp \left[-\beta \sum_{w \in M_i^N} I(v; w) \right] \dots (1)$$

ここで、 β は正の値を持つパラメータを表す。また、 M_i^N は、 M_i 中の単語の内、相互情報量 $I(v; w)$ の値の大きい上位 N 個の単語集合である。 M_i 中の全単語との相互情報量の和を基に差分スコアを計算すると、差分スコアは M_i 中の単語数に比例して小さくなりやすい。それゆえ、実際は差分であるものが、 M_i のテキスト長が長いために差分とならない現象が生じる。そこで、 M_i^N を導入し、テキスト長の違いによる影響を受けないようにする。

この差分スコアが高いほど、単語 v の情報が M_i 中に含まれていない可能性が高い事を意味し、低いほど M_i 中に含まれている可能性が高い事を意味する。したがって、 D_i 中の単語 v の差分スコアが、閾値 θ 以上である場合、この単語 v を差分 $D_i - M_i$ の構成要素とする。

この θ の値が大きいほど、 D_i と M_i における同義表現を差分から削除できる可能性が高まるが、同時に、同義表現以外の、 D_i と M_i の共通トピックの関連語なども差分から削除される可能性が高まる。このため、 θ は、分析対象のデータや分析目的に応じて適宜調整する必要がある。

3.2.2 MWS+CWS

CWSで差分を生成する際、分析データであるテキストの組(D_i, M_i)における出現回数が少ない単語の場合、相互情報量が適切に推定されない場合がある。そのため、 D_i

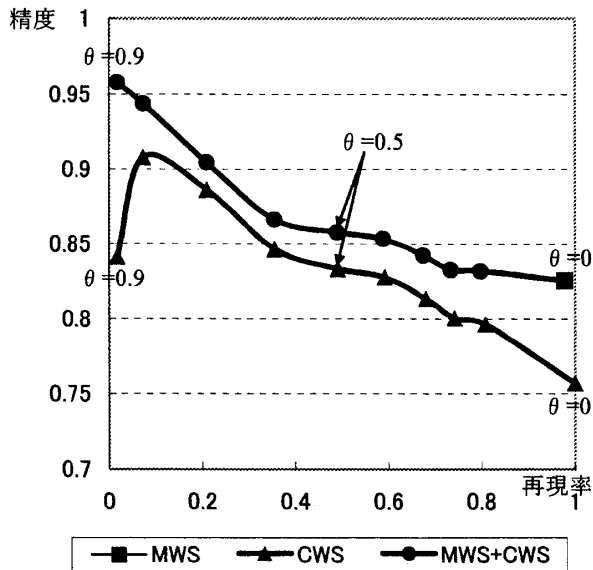


図1 差分の品質評価

と M_i の双方に、表層が完全一致している単語が出現していても、それが差分と判定される場合がある。したがって、表層が完全一致の単語については、共通要素と判断し差分には含めないことで、生成する差分の精度を向上させる。この手法を MWS+CWS と呼ぶ。これは、MWS と CWS の差分の和集合を差分集合とする事に他ならない。

4. 評価実験と考察

本章では、MWS、CWS、MWS+CWS で生成される差分の品質や各差分マイニング手法の有効性を検証する。以降の実験では、単語として自立語のみを使用する。

実験データとして、国内コールセンターにおけるコールメモ 4,596 件と、各コールメモに対応する顧客とエージェント間の通話音声を書き起こした通話データを用いた。コールメモは、「お客様名」、「エージェント名」、「受付日時」、「装置名」等の定型情報と「症状」、「その他コメント」といったエージェントが自由に記述することができるテキスト部分から構成されている。

4.1 生成差分の品質評価

本節では、MWS、CWS、MWS+CWS で生成される差分の品質を比較、評価する。評価セットは、実験データより無作為に抽出した 40 事例から人手で作成した。具体的には、通話データ D_i とコールメモ M_i を比較しながら、 D_i の中で、 M_i に書かれていない部分を特定した。そして、形態素解析を行い、その部分に含まれる自立語リストを生成し、それを、差分 $D_i - M_i$ の正解データとした。

CWS および MWS+CWS のパラメータは、経験的に $\beta=100, N=3$ を使用した。また、 θ は、0 から 0.9 までの範囲で、0.1 刻みで値を変えて差分を作成した。

各差分生成法により、差分と判定された単語に対して、精度と再現率(40 事例の平均値)をプロットした結果を図 1 に示す。縦軸は精度を示し、横軸は再現率を表す。

MWS は、パラメータが存在しないため、評価値は 1 点となる。一方、CWS と MWS+CWS では、各 θ の値ごとに差分の評価を行い、R-P 曲線を描いている。右端の点から順に、 $\theta=0, 0.1, \dots, 0.9$ の時の評価値を示す。CWS にお

表2 Dc の特徴語のランキング

順位	特徴語	ESC
1	注文する	0.844
2	キャンセル	0.505
3	どうやって	0.384
4	次第	0.379
5	寿命	0.368
:	:	:
1174	ロック	0.050
:	:	:

表3 Dc-M の特徴語のランキング

順位	特徴語	ESC
1	注文する	0.582
2	どうやって	0.384
3	次第	0.379
4	箱	0.323
5	メニュー	0.309
:	:	:
14	ロック	0.193
:	:	:
149	キャンセル	0.111
149	寿命	0.111
:	:	:

いて、 $\theta=0$ 時は、全要素を差分と判定した場合の評価値である。また、MWS+CWS において、 $\theta=0$ 時は、MWS の評価値と一致する(図では黒丸が黒四角に隠れている)。

図 1 より、CWS および MWS+CWS は、差分スコアの閾値 θ を上げることで、差分の再現率は低くなるが、精度の高い差分を生成できることが分かる。また、MWS+CWS は、CWS と比較して、精度、再現率共に高い、品質の良い差分を生成できる事が分かる。

4.2 差分マイニング MWS の有効性検証

テキストの差分を対象としてマイニングを行う差分マイニングというアイデアの有効性を、MWS によるマイニング結果と従来のマイニング結果とを比較し検証する。

実験データの内、ある機器に対する事例(事例数: 1,466)を用いて、2 節で記述した手順で、「ある機器の応対の中で、故障修理ではなく問い合わせの案件で顧客が発言する特徴的な事」を調べた。つまり、定型情報として記入されている「問い合わせ」という値を利用し、分析対象のカテゴリを問い合わせの事例とし、分析対象のテキストを通話の顧客側発話 **Dc** として、特徴語分析を行った。その結果得られた、問い合わせの案件に特徴的な **Dc** 中の自立語のランキング表を表 2 に示す。この結果は、従来のマイニング手法による結果に相当する。また、**Dc** ではなく、事例ごとに MWS により生成した顧客側発話とコールメモの差分 **Dc-M** を分析対象のテキストとして、同様に分析した結果得られたランキング表を表 3 に示す。

表 3 より、**Dc-M** において、上位(14 位)に「ロック」がランクインしている。これらの事例のコールメモを見ると、「ロック」という単語は使われず、代わりに「ある部品の交換方法の問い合わせ」といった汎化された表現が使われていた。このため、「ロック」が差分 **Dc-M** として現れていた。これらの事例の通話データを分析した

結果、顧客がある部品を交換する際、ロックの仕方でトラブルを生じていることが分かった。このことから、エージェントが操作方法を説明することで解決できている事例として、「ある部品のロックの仕方に関するトラブル」の存在が発見されたといえる。この知見を基に、例えば、ロックの仕方を詳細にマニュアルに載せる、といった対応策をとることで、CSの向上やコールセンターのコスト削減を実現できる可能性がある。一方、表2では、「ロック」は1174位であり、前述の知見が発見できない可能性が高い。

以上の事から、差分マイニングによって、従来のDcのみを対象としたマイニングでは得られなかった、コールセンターにとって重要な知見を獲得できる事が分かり、差分マイニングの有効性を確認できる。

4.3 同義表現の差分マイニングに対する影響

3.2節で、同義表現の存在が差分マイニングに悪影響を及ぼすことを説明したが、本節では、その影響と、共起関係を用いる事で、その問題が改善できることを示す。

4.2節では、通話特有の情報に着目しマイニングしたが、一方、コールメモには、エージェントの追加コメント等、通話にない情報が含まれている。そこで、本節では、コールメモから通話の内容を除いた差分を作成し、コールメモ特有の情報に注目してマイニングを行う。つまり、事例ごとに、コールメモ M_i から通話データ D_i を引いた差分 $M_i - D_i$ を生成し、その差分集合 $M-D$ を対象にテキストマイニングを行う。

具体的には、実験データの内、あるエージェントに対する事例(事例数:441)を用いて、「あるエージェントに対して、通話には含まれないがコールメモには記述された情報について、ある応対月の特徴」を調べる。つまり、定型情報である受付日を利用し、ある応対月を分析対象のカテゴリとし、下記の差分生成法により生成した差分を分析対象のテキストとして、特徴語分析を行った。本実験で用いる差分生成法は、MWS、 MWS_{syn} (表1のような「至急」に関する18種類の同義表現をまとめた同義語辞書を使用)、CWS、MWS+CWSである。CWSおよびMWS+CWSにおけるパラメータは、経験的に $\beta=100, N=3, \theta=0.45$ を使用する。

特徴語分析の結果、 MWS_{syn} を用いた差分マイニングでは、「至急」が21位にランクインした。これらの事例を分析すると、「通話の中で顧客が至急の対応を要望していない場合でも、エージェントは、トラブルの影響度や切迫度が高いと判断した場合には、コールメモ中に至急の対応が必要という記載を行う」という知見を得られた。この知見は、従来のコールメモからのマイニングでは、顧客が至急の対応を要望している事例の中に埋没してしまい得られないため、差分マイニングを用いる事で初めて得られる知見である。

一方、MWSを用いた差分マイニングでは、「至急」は1056位となり、特徴的なものにはならなかった。これを分析すると、コールメモと通話データの間、表1に示したような「至急」に関する表現の不一致が存在するため、本来差分ではない事例が、差分の中に多数含まれていることが分かった。以上より、MWSを用いた差分マイニングを有効に機能させるには、同義表現を正規化し、見かけ上の差分を取り除く必要がある事が実験で示された。

また、CWSを用いた差分マイニングでは、「至急」は186位、MWS+CWSを用いると50位となった。両手法とも同義語辞書を用いた正規化は行っていないが、MWSの1056位と比較すると、順位が改善されている事が分かる。実際に差分を見ると、MWSでは差分と判定されていた見かけ上の差分(例: D_i 中の「大至急」と M_i 中の「至急」)が、CWSでは除かれていた。また、MWS+CWSではさらに、CWSで差分と判定されていた表層が同一の単語(例: D_i 中と M_i 中に存在する「早急に」)が、同一視されていた。この事より、共起関係を用いる事によって、同義表現がある程度同一視でき、差分マイニングの効果を改善できることが確認できた。

ここで、CWSやMWS+CWSは、「至急」の順位が MWS_{syn} には及んでいないため、 MWS_{syn} より劣っているように見える。しかし、これらの手法は、「至急」以外の同義表現も考慮しているため、「至急」が本来の妥当な順位になっているとも考えられ、 MWS_{syn} に劣っているとは一概にはいえない。したがって、「至急」以外の語についても分析する必要があり、今後行う方針である。

5. おわりに

本稿では、テキスト間の差分に着目してマイニングを行う、差分マイニングの方法を提案した。ベースラインとして、語の照合に基づいた差分生成法MWSを導入し、国内コンタクトセンターの実データを対象に評価実験を行うことにより、差分マイニングの有効性を示した。さらに、差分マイニングにおいて同義表現が見かけ上の差分を生じ、適切なマイニング結果が得られないという問題に対して、その解決法として、分析対象における共起関係に基づいた差分生成法CWSおよびMWS+CWSの提案を行い、その有効性を示した。また、各差分生成法により生成される差分の品質評価から、MWS+CWSは、CWSと比較して、より品質の良い差分を生成することができることが分かった。今後は、本提案手法を、通話音声に対する音声認識結果に適用し、音声認識誤りを含む通話音声データに対しても適切なマイニング結果が得られる手法に発展させていく予定である。

参考文献

- [1] Morinaga, S., Arimura, H., Ikeda, T., Sakao, Y., and Akamine, S., "Key semantics extraction by dependency tree mining", In *Proceedings of KDD-05*, 666-671, 2005.
- [2] Mishne, G., Carmel, D. and Hoory, R., "Automatic analysis of call-center conversations", In *Proceedings of CIKM-05*, 453-459, 2005.
- [3] Takeuchi, H., Subramaniam, L. V., Nasukawa, T. and Roy, S., "Automatic Identification of Important Segments and Expressions for Mining of Business-Oriented Conversations at Contact Centers", In *Proceedings of EMNLP-07*, 458-467, 2007.
- [4] 那須川 哲哉, 宅間 大介, 竹内 広宣, 荻野 紫穂. "コールセンターにおける会話マイニング". 言語処理学会 第13回年次大会, 2007.
- [5] Li, H., and Yamanishi, K., "Mining from open answers in questionnaire data", In *Proceedings of KDD-01*, 443-449, 2001.
- [6] Yamanishi, K., "A decision-theoretic extension of stochastic complexity and its applications to learning", *IEEE Transactions on Information Theory*, 44(4):1424-1439, 1992.