

# ソーシャルメディアにおける議論の抽出手法の提案

星川 祐人<sup>†,a</sup> 若林 啓<sup>‡,b</sup>

<sup>†</sup> 筑波大学大学院図書館情報メディア研究科 <sup>‡</sup> 筑波大学図書館情報メディア系

a) *s1521643@u.tsukuba.ac.jp* b) *kwakaba@slis.tsukuba.ac.jp*

**概要** 人間同士の会話は、目的が明確な議論と、目的のない雑談に大別することができる。本研究では、ソーシャルメディア上で行なわれている会話から、議論を抽出する手法を提案する。議論には、会話の目的や話者の持つ知識が表れるため、ユーザ推薦やユーザの人間関係の把握に有用であると考えられる。会話の目的が決まっている程度を表す指標を議論スコアとする。会話に含まれる助詞・助動詞から話者の態度を表す発話役割をベクトルで表現し、この発話役割ベクトルと会話のツイート数を用いて、会話ごとの議論スコアを推定する。実験より、提案手法による推定結果と人手による判定との相関を評価し、提案手法の有効性を確認した。

**キーワード** 会話, 議論, ソーシャルメディア

## 1 はじめに

一人で行うのが困難だったり、複数の人間の意見を取り入れたものを完成させるとき、チームやグループを組んで行うことがある。このような行為はグループワークと呼ばれ、グループ内の人間関係が作業効率やパフォーマンスに影響を与えることがわかっている。このため、仲の悪い二者や極端に上下関係のある二者を同じチームに入れないようにするなど、プロジェクトマネージャはプロジェクトを成功させるため、チーム内の人間関係の把握や人間関係を考慮したチームづくりが求められている。しかし、人間の集合から人間関係を把握できる上限は7人前後といわれており、それ以上の集合の場合は何らかの支援が必要となる。プロジェクトを管理するプロジェクトマネージャの立場から見ると、議論をしている関係はグループワークや共同作業、協調学習を行う上で重要な関係であり、この関係の把握こそがプロジェクトを成功に導くために把握したい関係といえる [4]。このため、本研究は議論をしている関係の抽出を目指す。

また、近年ソーシャルメディアを利用するユーザが急増しており、即時性に優れている点や、時間や場所を問わず気軽にインタラクションを持つことができることから、多くの会話が存在する。Twitterのユーザは、他者のツイートに対してリプライという機能を用いてツイートすることにより、他のユーザとインタラクションを持つことができる。人間関係は会話内容に表れ [7]、会話から人間関係を把握するときは約4割が会話内容といわれている [3]。初対面同士の会話には「初めまして」のような言葉を発しやすかったり、発話に態度が表出することから、会話内容から関係性を推定することができる。また、議論には会話の目的や話者の持つ知識が表れるため、ソーシャルメディアにおける会話か

ら議論を抽出することは、ユーザ推薦やユーザの人間関係の把握に有用である。ソーシャルメディアの分析にグラフを用いることがあり、グラフにはユーザをノード、会話の関係をエッジとしたネットワークが存在する。しかしソーシャルメディアでは、ユーザは自由に他のユーザと関わりを持つため、エッジが膨大となり、グラフは複雑となる。また、エッジに対し、会話の回数などの定量的な要素では特徴付けることは難しいため、会話内容に着目し、どのような会話か推定することで、この問題を解決する。そこで本研究ではソーシャルメディアにおいて、会話を連想型会話の「雑談」と目的指向型会話の「議論」の2種類に分類し、「議論」の会話を抽出する手法を提案する。会話の目的が決まっている程度を「議論スコア」と呼ぶこととする。Twitterにおける会話を抽出し、人手による議論らしさの評価と議論スコアに相関があれば、議論スコアで議論の程度を推定でき、議論を抽出することができる。

本稿の構成は以下のとおりである。2章で関連研究として、会話の分類に関する研究と人間関係の把握に関する研究を紹介し、3章で西原らが提案する会話内容を用いた二者の親密度推定手法である仲の良さスコアについて詳しく説明する。4章で西原らの手法をもとにした本研究による議論の抽出手法を論じる。5章で実験から考察をし、6章でまとめとする。

## 2 関連研究

### 2.1 会話の分類に関する研究

小磯ら [5] は、我々が普段どのような会話を行っているのかを明らかにした。その結果、会話の形式を「雑談」、「用談・相談」、「会議・会合」、「授業・レッスン・講演」の4つに分類している。「雑談」は会話の目的や話題などがあらかじめ定められていない会話を、「用談・相談」は会話の目的はある程度決まっているが、時間や場所な

どは定められていない会話を、「会議・会合」は時間や場所などが定められている会話を、「授業・レッスン・講演」は先生や講演者など会話の流れを導く人物がいる場での会話を指す。ソーシャルメディアにおいてはユーザが分散しているため、時間や場所が定められている会話である「会議・会合」が存在しないといえる。また、文字数の制限があり、マイクロブログには長文が適さないことや、Twitter の会話では交互にリプライを送り合うため、「授業・レッスン・講演」が存在しないといえる。

また、仮屋園ら [6] は会話の種類を目標指向型会話と連想型会話の 2 種類としている。他者に対する要求や説得を目標とする会話、他者から知識を得るための会話、課題解決のための会話などが目標指向型会話にあたる。直前の発話に対する思いつきなどで継続され、続けること自体を目的とする会話が連想型会話である。本研究では「用談・相談」を「議論」と呼び、会話の目的がある程度決まっていることを「議論」の定義とする。すなわち議論スコアは会話の目的が決まっている程度を示す数値とする。

## 2.2 人間関係の把握に関する研究

西原ら [1] は、会話内容から推定される仲の良さスコアという親密度の指標を提案し、これを求めることがプロジェクトマネージャにとっての人間関係の把握につながると考えた。この指標は高い値ほど親密であり、評価実験からメール、掲示板、チャット、音声会話に対して有効であることを確認している。しかし、Twitter において仲の良さスコアが高い二者は、リプライを数多く行い、発話量が多いため、それに伴い様々な態度を示しているような二者も含まれる。このため、仲の良さスコアでは多様な人間関係の把握は難しく、会話の区別がつかないと考えられ、仲の良さスコアが高いほどグループワークでの関係に適しているとはいえない。すなわち親密度という指標では Twitter における人間関係の把握は難しいと考えられる。そこで本研究は、会話の種類に着目し、議論を行う関係こそがプロジェクトに必要な関係であるとみなし、その関係の抽出を目指す。

## 3 仲の良さスコア

西原ら [1] は発話文の内容から、二者の仲の良さを推定する手法を提案している。初対面のときは話者は控えめだが、親密になると話者の態度の種類が増えると考えられる。そこで、「仲が良い」を聞き手に対して話者が様々な態度をとれる状態とし、「仲の良さスコア」は親密な程度を示す数値とする。また、一方が発話し続けるよりは交互に発話する方が親密であるため、話者の態度（以下、発話役割とする）の種類数が多く、発話の回数が同じ程度であるほど親密とし、仲の良さスコアを推定

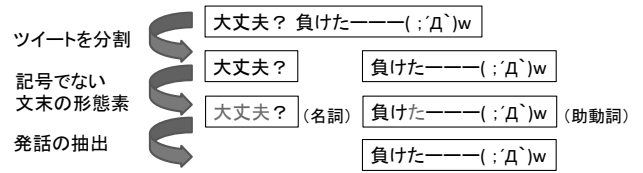


図 1 発話文抽出の流れ

する。提案手法は二者に限ったものであり、三者以上の会話においては推定が難しいことから、本研究においても二者に限った仲の良さスコアの推定を行う。

二者の発話テキストが入力されると、文末に助詞・助動詞が含まれる発話文を抽出する。発話役割と助詞・助動詞の意味の対応表を用いて、発話に含まれる助詞・助動詞の意味から発話文の発話役割ベクトルを作成し、仲の良さスコアを推定する。以下、各処理について説明する。

### 3.1 発話文の抽出

Twitter のおける 1 つのツイートには複数の発話を含んでいることがある。発話はそれぞれに別の意味を持っていると考えられるため、これを明示的に分割することが望ましい。本研究では、特定の記号が出現している箇所を発話の区切りとみなして、ツイートを発話単位に分割する。ここでは、この特定の記号の集合をセパレータと呼ぶ。セパレータは以下の記号とスペース、(笑) のような 1 文字を括弧で括ったもの、顔文字で構成される。

w w … 。 . ! ? ʘ \* ← ʘ

リプライマーク、URL を除去した上で、セパレータによって分割されたものを発話とみなし、形態素解析を行う。顔文字は三好ら [9] が提案する定義を用いた。また、顔文字等の除去がうまくいかず、記号が残る場合も考慮し、形態素解析の結果、末尾の形態素が記号であればその直前の形態素を末尾の形態素として扱う。この操作は末尾の形態素が記号でなくなるまで繰り返す。末尾の形態素が助詞・助動詞であればその発話を抽出し、後述の発話役割ベクトル生成に用いる。

図 1 に発話文抽出の流れを示す。対象のツイートを「大丈夫？負けた——( ; ʘ `)w」とすると、発話に分割した結果、「大丈夫？」と「負けた——( ; ʘ `)w」の 2 つとなる。発話の記号以外の文末の形態素はそれぞれ名詞の「大丈夫」と助動詞の「た」のため、助動詞である「負けた——( ; ʘ `)w」を発話役割ベクトル生成に用いる発話として抽出する。

### 3.2 発話役割ベクトルの生成

話者の態度は文末に用いられる助詞・助動詞の組み合わせで表されると考える。西原らが用意した発話役割と

助詞・助動詞の意味の対応表を用いることで、発話から話者の態度を推定する。対応表とは大辞林 [10] に掲載される副助詞・終助詞・係助詞・助動詞の意味を西原らが独自にまとめたものである。詳細は [1] を参照されたい。

表1の対応表を用いて、発話役割ベクトルを求める。発話役割と意味は多対多の関係になるため、発話文  $d$  の発話役割をベクトル  $V_d = (x_1, x_2, \dots, x_R)$  と表す。  $x_1, x_2, \dots, x_R$  は表1の発話役割を表し、  $R = 23$  である。ベクトル  $V_d$  の値は、発話文に含まれる助詞・助動詞の意味の数と、意味に対応する発話役割の重みから与えられる。発話文  $d$  に含まれる助詞・助動詞の集合を  $P_d$ 、  $P_d$  内の助詞・助動詞  $p_i$  の意味を  $m_{p_i}^j$  とし、式 (1) によって発話役割ベクトル  $V_d$  の要素の値  $x_{l,d}$  を求める。

$$x_{l,d} = \left( \sum_{p_i \in P_d} h(x_l, m_{p_i}^j) \right) \times w(x_l) \quad (1)$$

式 (1) において、  $h(x, m)$  は表1の中で発話役割  $x$  と意味  $m$  の対応があり、意味  $m$  が文末に存在する助詞・助動詞ならば1、それ以外なら0.5を返す関数とする。発話役割は、主に文末に含まれる助詞・助動詞によって決定されるが、文中に含まれる助詞・助動詞も発話役割を表すとす。  $w(x)$  は発話役割  $x$  の重みを返す関数である。西原らの論文では、  $w$  を1.0として発話役割を同定し、正解と比較した適合率を再び  $w$  としているが、予備実験の結果が論文からは明らかでない。このため、本研究では  $w$  を一律で平均の0.35として与える。

### 3.3 仲の良さスコアの定義

話者の態度の種類数と発話文の数から二者の仲の良さスコアを推定する。はじめに、話者の態度の種類数として、発話役割の種類数を求める。話者  $s$  のすべての発話文における発話役割の種類数  $k(s)$  は、式 (2) によって求められる。

$$k(s) = \sum_{l=1}^R \begin{cases} 1 & \text{if } \sum_{d \in D_s} x_{l,d} \geq T \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

式2は、話者  $s$  の発話文の集合  $D_s$  の中の発話役割ベクトルの各要素  $x_{l,d}$  を合計し、その値が閾値  $T$  を超える要素数を求めている。 [1] 同様、本研究においても  $T = 1.0$  とした。  $k(s)$  と入力発話テキストに含まれる話者の発話文の数  $l(s)$  から、話者  $s_a, s_b$  の仲の良さスコア  $F(s_a, s_b)$  を式 (3) によって求める。

$$F(s_a, s_b) = \frac{k(s_a) \times k(s_b)}{l(s_a) \times l(s_b)} \quad (3)$$

$F$  の値は大きいほど親密とされる。

### 3.4 仲の良さスコアの問題点

仲の良さスコアを推定し、その値が近い2つの会話を表2と表3に挙げる。表2では挨拶から始まり、1日の予定などを報告する雑談のような会話であり、対して表3は過去の競走馬について印象を話し合う議論のような会話であるといえる。これらの会話は仲の良さスコアの推定の結果、近い値を出力しているが、内容を見ると意味の異なる会話であることがわかる。このように、一言に親密といっても親密さには雑談をする親密さと、議論のように意見を交わす親密さのようにいくつか種類がある。すなわち仲の良さスコアでは、多様な会話の違いを表すことができないといえる。仲の良さスコアでは、ユーザ間の親密さの度合いであるため、会話の切れ目を考慮せず、会話をすべて一緒にしてユーザごとの発話群にまとめているが、議論スコアの場合、会話の違いを明らかにするため、会話ごとに議論スコアを推定する必要がある。また、仲の良さスコアでは「発話数が同程度ほど親密」という仮定のため、(3)式では発話数で除算しているが、ツイート数が多すぎるほど値が低くなってしまふことがある。しかし、Twitterではツイートごとに区切られており、交互にリプライが連鎖しなければ会話が成立しないため、「発話数が同程度」を考慮する必要がないと考えられる。このため、議論スコアではツイート数に影響されない工夫が必要となる。

## 4 提案手法

本章では、西原らの手法をもとに、議論を抽出するための手法を説明する。仲の良さスコアが二者間の会話すべてを用いて推定していたのに対し、議論スコアは会話ごとに求める指標である。発話文の抽出では、会話ごとに発話群をまとめること以外は仲の良さスコアと同様であるため、3.1節以降の説明をする。

### 4.1 議論に対応する発話役割

次に、3.2節にあたる発話役割ベクトルの生成であるが、仲の良さスコアに用いた発話役割には議論に必要なものも含まれていると考えられる。このため、議論に必要な発話役割を絞り込んだ上で議論スコアの推定を行う。小島ら [4] は、実際に行われた議論に対して発言意図を推定し、議論の状態を把握することで議論の支援を行った。その際推定される発言意図とは、議論に存在する発言意図であるといえる。そこで、本研究では西原らが用いた23種類の発話役割の中で、議論に関係する発話役割のみを議論スコアの推定に用いることとする。表4に [4] で付与されていた発言意図とその意味、それぞれに対応する発話役割を示す。表4より、議論スコアに用いる発話役割を9種類に限定する。発話文  $d$  の発話役割をベクトル  $V_d = (x_1, x_2, \dots, x_R)$  で表すため、本研究

表 1 発話役割と助詞助動詞の意味の対応表

	発話役割	助詞助動詞の意味
$x_1$	心情	感動・意志・自発・可能・不可能・希望・詠嘆・強調
$x_2$	事実	打消・確認・質問・添加・並立・選択・不確定・列挙・反語・程度・限定・軽視
$x_3$	補足	他との区別・添加・並立・選択・不確定・列挙・強調
$x_4$	価値判断	打消当然・不確かな気持ち・理由
$x_5$	知識獲得	確認・疑問・質問・希望・願望・不確かな気持ち・問掛
$x_6$	知識提供	強調
$x_7$	教える	類推・推量・打消推量・推定・婉曲・打消・強調
$x_8$	教えられる	確認・疑問・質問・願望・依頼・不確かな気持ち・問掛
$x_9$	依頼	疑問・質問・希望・願望・依頼・不確かな気持ち
$x_{10}$	確認	不確かな気持ち・確認
$x_{11}$	行動要求	意志・希望・願望・依頼・呼びかけ・禁止・同意要求・勧誘・不確かな気持ち・命令・許可・原因
$x_{12}$	発言要求	婉曲・確認・疑問・質問・意志・念押・呼びかけ・同意要求・勧誘・命令・問掛・許可
$x_{13}$	提案	他との区別・反語・意志・希望・強調
$x_{14}$	礼	尊敬・自発
$x_{15}$	謝罪	尊敬・意志・自発
$x_{16}$	賛成	詠嘆・適当・当然・全面否定・全面肯定
$x_{17}$	反対	打消・打消意志・打消強調・非難・全面否定・全面肯定・反対
$x_{18}$	行動要求受入	打消・打消当然・確認・質問・意志・自発・可能・不可能・打消意志・念押・願望・依頼・過去
$x_{19}$	発言要求受入	比例・様態・伝聞・例・例示・類推・推量・打消推量・推定・婉曲・打消・打消当然・確認・質問
$x_{20}$	納得	婉曲・確認・断定・不確かな断定・過去
$x_{21}$	理由陳述	理由・価値判断
$x_{22}$	保持	不確かな気持ち・確認・列挙
$x_{23}$	話題転換	依頼・意志・質問

では  $R = 9$  となる。ベクトルの成分の強さは 3.2 節と同様に求める。

## 4.2 議論スコアの定義

議論に対応する発話役割を用いて議論スコアを求める。また 3.4 節より、「発話数が同程度ほど親密」という西原らの仮定を考慮せず、ツイート数に影響を受けないようにするため、ツイート数で除算する。本研究では仲の良さスコアのように発話役割ベクトルを種類数に変換して乗算する手法と、ベクトルの成分をそのまま利用し、総和を乗算する手法の 2 種類のスコアを提案する。議論スコアの値は大きいほど議論らしく、目的が決まっているとされる。

### 4.2.1 発話役割の種類数を乗算する手法

本手法は仲の良さスコア同様に、閾値を超えるベクトルの要素の数を発話役割の種類数として議論スコアを推定する。3.3 節と同じく話者  $s$  のすべての発話文における発話役割の種類数  $k(s)$  は、式 (2) によって求められる。

$k(s)$  と会話のツイート数  $t$  から、会話  $c$  の議論スコア  $Score1(s_a, s_b)$  を式 (4) によって求める。

$$Score1(s_a, s_b) = \frac{k(s_a) \times k(s_b)}{t} \quad (4)$$

### 4.2.2 ベクトル成分の総和を乗算する手法

仲の良さスコアでは、話者がいろんな態度を示すほど親密であるという仮定があったため、閾値で発話役割の種類数を求めたが、議論スコアにおいては発話役割の成分の強さが重要とも考えられるため、本節では発話役割ベクトルの成分をそのまま利用する手法を提案する。発話役割ベクトル  $V_d$  の各要素  $x_{l,d}$  を全て足しあわせたものを  $x_{sum}(s)$  とすると、式 (5) で計算することができる。

$$x_{sum}(s) = \sum_{l=1}^R \sum_{d \in D_s} x_{l,d} \quad (5)$$

$x_{sum}(s)$  と、会話のツイート数  $t$  から会話  $c$  の議論スコア  $Score2(s_a, s_b)$  を式 (6) によって求める。

表2 雑談のような会話

皆さん♪おはようございます 今日でうちのピッチ U`ェ`U が1歳になりました♪ やったあ そして息子の高校入学式です♪今日は忙しい一日になりそうです
@A ピッチさん☆ おはよー♪息子くんご入学式おめでとうございます☆そしてピッチちゃんお誕生日おめでとう (*≧▽≦*) 今日は素敵な1日になりそうですね♪♪今週もよろしくね (*^▽^*)
@B いーちゃん♪おはよう♡ うん♡ありがとう♪♪今日はダブルで良い記念日になりそう♪入学式午後からなので行ってくるネ♪今日も楽しい一日をすごしてネ(^_^)ノ

表3 議論のような会話

思い出の皐月賞と言えば、さっきも話に出たミホノブルボンかな。スプリングSを圧勝しても距離不安説があったが、圧勝。それでもまだ距離不安の声があったが、ダービーも圧勝。坂路で鍛えられた馬体は筋肉の塊。圧倒的なスピードは弾丸のようだった。
@C 強かったですよね。坂路の申し子、戸山イズムの最高傑作と言われたのをよく覚えています。ライバルにライスシャワーの存在も大きかったですね。前年のクラシックはトウカイテイオー、翌年はBNW 3強、更にその翌年はブライアンとあの当時は各世代に今より華があったように思います。
@D サンデーサイレンスが出て来るまではクラシックレースでスローペースってまず、なかったですから、いいレースが多くて印象深いのかもかもしれませんね。外国人騎手もいなかったですから、騎手と馬のドラマというのもありましたからねえ。

表4 議論に対応する発話役割

発言意図	意味	対応する発話役割
提起	問題・話題の提案等、新しい話題の提示	提案・話題転換
確認	yes・no で答えられる質問	確認
説明	現在の話題に関する付加情報や詳細情報の提示	知識提供
同意	賛成の立場表明	賛成
非同意	反対の立場表明	反対
質問	他者に対する質問	発言要求
補足	直前の発言に対する付加情報	補足
回答	質問に対する答え	教える・知識提供
その他	上記のいずれにも分類不可能な発言	該当なし

を本研究で用いるデータセットとする。ここでは、二者による往復するリプライを行う行為を Twitter における会話とする。起点となるツイートからリプライの連鎖が終了するまでが1つの会話である。会話の最後のツイートの投稿時刻が2013年4月中に存在していれば、会話のどのツイートの投稿時刻が期間外であっても、そのツイートを含めた会話をコーパスとして用いる。各会話に対し、人手によりどれだけ目的が決まっているかを1~5の5段階で評価し、正解データとして扱う。なお、アノテータは第一著者1名とする。

### 5.1 比較手法

提案手法の有効性を検証するため、ツイート数、ツイートあたりの平均文字数、仲の良さスコアを比較手法とする。ツイート数が多い場合、またツイートの文字数が多い場合に議論だと感じることも考えられるため、これらのスコアと比較する。仲の良さスコアは3.4節で指摘した、仲の良さスコアでは議論か雑談かの違いを表すことができないという問題を改善できているかの確認のため比較手法とした。

### 5.2 実験結果

各手法の人手による評価との相関係数を表5に示す。実験結果から提案手法のいずれもどの比較手法より高い値を示していることにより、提案手法の有効性を確認できる。Score1が最も高く、その値は0.573であった。

### 5.3 考察

提案手法が比較手法より高い相関係数を示した理由は2つ考えられる。1つは仲の良さスコアよりも相関係数

$$Score2(s_a, s_b) = \frac{x_{sum}(s_a) \times x_{sum}(s_b)}{t} \quad (6)$$

## 5 実験

本章では、前章にて提案した手法の有効性を検証するための実験を行い、その結果から考察を行う。

Twitter における 2013 年 4 月 1 日~2013 年 4 月 30 日の期間に存在する会話からサンプリングした 100 会話

表5 実験結果

Score1	Score2	ツイート数	平均文字数	仲の良さスコア
0.573	0.555	0.185	0.498	0.447

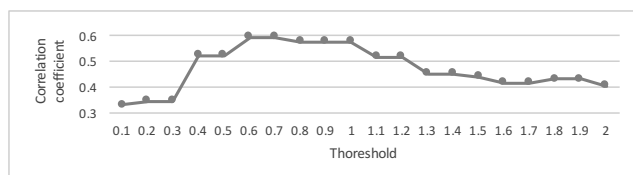


図2 閾値と相関係数の変化

が高いことから、議論に合わせた発話役割を選択していることが要因である。また、ツイート数で除算することで、ツイート数の長さに影響されず、各ツイートが議論らしいかを計算することができている。Score1において、実験では閾値  $T = 1.0$  としたが、これは西原らが発話役割の種類数を計算する上で定めた閾値であるため、閾値を変化させたときの相関係数の変化を図2に示す。図2より、 $T = 0.6$  及び  $T = 0.7$  のときに相関係数は0.593となり、実験で用いた閾値よりも低く設定した方が相関係数は高いことを確認した。これは、仲の良さスコアがユーザごとに会話を全てまとめて発話群にしていたのに対し、議論スコアでは会話ごとの指標であるため、発話量が仲の良さスコアを求めるときよりも少ないことが多く、発話役割ベクトルの成分もそれに伴い、平均的に低下したためと考えられる。しかし、閾値を低くしすぎると、助詞・助動詞が存在するだけで種類数とみなされてしまい、人が感じる議論らしさとは離れてしまうことがわかる。逆に高くしすぎると種類数とみなされる発話役割も限られてしまい、相関係数は低下してしまうことが確認できた。このため、閾値は0.6または0.7として、ノイズをある程度除去することで、最も相関のあるスコアとなる。

また、議論の研究には議論をモデル化し、議論を円滑に進める支援をする研究や、議論をするエージェントを開発する研究、議論から参加者の領域知識を明らかにする研究などが存在するが、いずれの研究においてもまず対象となる議論が必要となる。これまでの研究では、実際に課題を与えて被験者にグループワークを課しており、コストがかかることが問題であった。本研究により意図的に会話から議論を抽出できるため、容易に議論を行うコーパスを獲得するという応用も期待できる。

## 6 まとめ

本研究ではTwitterで行われた会話から議論を抽出する手法を提案した。本研究はソーシャルメディアにおけるユーザの推薦や、エッジに対する特徴付け、人間関係の把握に対して有用だと考えられる。このため、Twitterにおける会話を雑談と議論の2種類に分類し、議論の抽出を目指した。議論の程度を表す指標を議論スコアとし、発話に含まれる助詞・助動詞から発話役割ベクトルを生成し、これとツイート数を用いて推定する。仲の良さスコアと同様に、発話役割の種類数を用いる手法と、発話役割ベクトルの成分の総和を用いる手法の2種類を提案した。実験において、いずれの手法もツイート数、平均文字数、仲の良さスコアの比較手法より、相関係数が高いことから、提案手法の有効性を確認した。特に、発話役割ベクトルの種類数を用いた場合が最も高く、閾値を仲の良さスコアの時より少し低くした場合に、さらに相関係数が上昇することも確認した。

## 謝辞

本研究の一部は、JSPS 科研費（課題番号 16H02904）の助成によって行われた。

## 参考文献

- [1] 西原陽子, 砂山渡, 谷内田正彦: 発話テキストからの人間の仲の良さや上下関係の推定. 電子情報通信学会論文誌, Vol. J91-D, No. 1, pp. 78-88, 2008.
- [2] 綱倉久永, 岡田正大, 内田恭彦, 日本企業のトップマネジメント・チーム-デモグラフィー・コミュニケーション・意思決定: 上智大学経済学会ディスカッション・ペーパー・シリーズ, No. 3, 2007.
- [3] 木村昌紀, 大坊郁夫, 余語真夫: 社会的スキルとしての対人コミュニケーション認知メカニズムの解明. 日本心理学会第47回大会, pp. 122-123, 2006.
- [4] 小島圭一, 岡本敏雄: CSCW の対話における発言意図の推定に関する研究 (3). 電子情報通信学会技術研究報告, Vol. AI95-53, pp. 87-94, 1996.
- [5] 小磯花絵, 土屋智行, 渡部涼子, 横森大輔, 相澤正夫, 伝康晴: 均衡会話コーパス設計のための一日の会話行動に関する基礎調査. 国立国語研究論集, Vol. 10, pp. 85-106, 2016.
- [6] 仮屋園昭彦, 伊藤克浩, 廣瀬等: 会話の種類が会話の記憶に及ぼす影響. 心理学研究, Vol. 64, No. 5, pp. 389-395, 1993.
- [7] K.Matsumoto, J.Minato, F.Ren, S.Kuroiwa: Estimating human emotions using wording and sentence patterns, IEEE Conference on Information Acquisition, Vol. 2005, No. 1, pp. 421-426, 2005.
- [8] J.Y.Bak, C.Y.Lin, A.Oh: Self-disclosure topic model for classifying and analyzing Twitter conversations. ACL 2014, 2014.
- [9] 三好辰明, 太田学: ツイートに出現する顔文字等の文字と記号に着目した感情分類. DEIM Forum 2013, 2013.
- [10] 大辞林. <http://dictionary.infoseek.ne.jp/>