

# 円滑な人脈形成のための SNS 投稿内容に基づく仲介者探索手法

小松 恭子<sup>†</sup> 中澤 昌美<sup>†</sup> 池田 和史<sup>†</sup> 服部 元<sup>†</sup> 滝嶋 康弘<sup>†</sup>

KDDI 研究所 〒356-8502 埼玉県ふじみ野市大原 2-1-15 <sup>†</sup>

## 1 はじめに

本稿では、面識のない 2 者を円滑に引き合わせるための仲介者を探索する手法を提案する。従来研究では、コミュニケーションを活性化させるために、同空間に居合わせた 2 者に共通する話題を提供する手法 [1] やソーシャルグラフにより共通の知人がいる 2 者を対象に、潜在的な共通話題を推定する手法 [2] が提案されている。しかし、面識のない 2 者の場合、必ずしも直接的な経路が存在するとは限らない。また、実世界における 2 者間の繋がりが親密でない場合、仲介は非現実的となる。本稿では、図 1 のように立場や環境が異なる面識のない 2 者を円滑に繋げる経路探索手法を提案する。提案手法では、ソーシャルグラフだけでなく、投稿内容から実世界での繋がりを抽出すると共に、投稿行為から 2 者の親密度合いを推定することが特徴である。提案手法における有効性を検証するため社内の SNS データを使用して評価実験を行った。

## 2 提案手法

提案手法の概要を図 2 に示す。本稿では、SNS 上の友人関係、投稿内容および投稿に対するコメントの応答を解析し、登録者との関連の強さを表す関連度  $c$  を付与した関係データ  $D$  を生成する。これを基に 2 者の仲介経路をダイクストラのアルゴリズムを用いて最短経路を提示する。

### 2.1 友人関係解析手法

従来手法  $M_\alpha$  として、ソーシャルグラフ上の繋がりがあれば関係データ  $D_\alpha$  に列挙する。表 1 に User1 の関係データ  $D$  の例を示す。ユーザ 1 がユーザ 2、3、5、N と友人関係にある場合、表 1 の関係データ  $D_\alpha$  のように列挙される。関連度  $c_\alpha$  は一定の値とする。

### 2.2 投稿内容解析手法

提案手法  $M_\beta$  として、投稿内容から投稿者と関連したキーワードを抽出し関係データ  $D_\beta$  を生成する。

SNS 上の投稿から TFIDF を用いて重要語句を抽出する。社名、部署名、個人名、技術名など実務に即した投稿者と関係のある語句を登録したユーザ関連情報辞書に基づき、抽出した重要語句から投稿者と

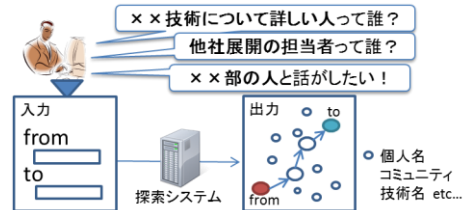


図 1 利用シーンの一例



図 2 提案手法の概要

関係のある語句を分別して選択する。選択されたキーワードは表 2 の関係データ  $D_\beta$  のように列挙する。関連度  $c_\beta$  には TFIDF 値を用いる。

### 2.3 コメント応答解析手法

提案手法  $M_\gamma$  では、コメントのやり取りがあるユーザを列挙し、関連度  $c_\gamma$  は一定の値とする。コメント頻度  $f$  (回) とその応答時間  $t$  (時間) の平均値を算出し、記録する。表 1 の関係データ  $D_\gamma$  にユーザ 1 についてユーザ 3 とのコメント回数が 20 回、その平均応答時間が 0.5 時間であった場合の例を示す。

### 2.4 関連度算出方法

探索に使用する関係データ  $D$  は  $D_\alpha$ 、 $D_\beta$ 、 $D_\gamma$  により生成される。関係データ  $D$  の関連度  $c$  は  $c_\alpha$ 、 $c_\beta$ 、 $c_\gamma$  およびコメント頻度  $f$ 、応答時間  $t$  の 5 つを用いて (1) 式により算出する。

$$c = (\alpha c_\alpha + \beta c_\beta + \gamma c_\gamma) * \begin{cases} \gamma_1 \frac{1}{f} + \gamma_2 \frac{t}{t_{Max}} & (c_\gamma = 0 \text{ のとき}) \\ \gamma_3 & (c_\gamma = 1 \text{ のとき}) \end{cases} \quad (1)$$

(1) 式において、 $f$  は 1 以上の値を取るため  $1/f$  は  $(0,1]$  となる。 $t_{Max}$  は全コメント平均応答時間の 90% を含む値とし、より応答時間の長いものは全て  $t = t_{Max}$  として扱い、 $t/t_{Max} = [0,1]$  となるよう正規化する。また、変数  $\alpha$ 、 $\beta$ 、 $\gamma$  は各手法に重み付けを行う際に用いる。例えば実関係に基づく繋がりを重視した経路提示を期待する場合、 $\beta > \alpha$ 、 $\gamma$  とする。変数  $\gamma_1$ 、 $\gamma_2$  は仲介時に求める項目により重み付けを行う際に用いる。例えばスピードを重視した経路提

Mediator search algorithm based on posted SNS messages for building a personal network

<sup>†</sup>Kyoko Komatsu, Masami Nakazawa, Kazushi Ikeda, Gen Hattori, Yasuhiro Takishima, KDDI R&D Laboratories, Inc.

表1 User1の関連データ例

	具体例
$D_\alpha$	$U_2:1.0, U_3:1.0, U_5:1.0, \dots, U_N:1.0$
$D_\beta$	A部:0.6, $U_K:0.5$ , 対話システム:0.7, S社:0.3, B部:0.5
$D_\gamma$	$U_3:1.0$ 20 0.5, $U_5:1.0$ 2 10, ..., $U_K:1.0$ 10 2
$D$	$U_1(U_2:1.0, U_3:1.0, U_5:2.0, U_K:1.5, \dots, U_N:1.0,$ A部:0.6, 対話システム:0.7, S社:0.3, B部:0.5)

表2 投稿内容解析フロー及び具体例

フロー	具体例
User1の投稿内容	A部K氏の紹介で、対話システム「商品P」に関するプレスリリースに関心を持ったS社のB部より問い合わせを受けたため、個別のデモ・説明を行った。
重要語抽出	A部、K氏、対話システム、商品P、プレスリリース、S社、B部、デモ
辞書	社名、部署名、人名、技術名
キーワード選択	A部、K氏、対話システム、S社、B部

示を期待する場合、 $\gamma_2 > \gamma_1$ と設定する。

### 3 性能評価実験

#### 3.1 実験方法

企業の社内 SNS データを対象に解析を行った。登録者数は 261 人、収集期間は 2011 年 3 月から 2013 年 4 月を利用した。各手法を均等に扱い  $\alpha = \beta = \gamma = 1$ ,  $\gamma_1 = \gamma_2 = 1$ ,  $\gamma_3 = 3$  として関連度を計算した。ソーシャルグラフのみを用いる従来手法を  $M_\alpha$ 、 $M_\alpha$  に加えて投稿内容を用いる提案手法を  $M_{\alpha+\beta}$ 、コメントの応答解析を用いる提案手法を  $M_{\alpha+\gamma}$ 、全てを合わせた提案手法を  $M_{\alpha+\beta+\gamma}$  とする。

まず、登録者 261 人における全経路を探索する定量評価を行う。従来手法では発見できなかった実務上の経路を、提案手法で発見可能か検証する。

次に、探索された経路が普段のコミュニケーションを考慮した円滑なものであるかについて、被験者による主観評価実験を行う。被験者 (from) から被験者と直接的な友人関係にない人物 (to) について、各手法を用いた経路を提示する。仲介者が増すほど、提示された経路を被験者が正確に評価することが難しくなるため、本稿では 2 ホップとなる経路を対象とする。対象 SNS 登録者 3 名に対し、条件を満たす合計 45 経路について、評価実験を行う。被験者は、提示された仲介者に関する 2 つの質問、Q1: 対応が迅速か、Q2: 実際に良く関わる人か、に対し 3 段階評価 (思わない -1・普通 0・思う 1) を行う。また、Q3: 仲介者の適切さを 4 段階でポイント付けする、ことにより各手法の有効性を比較する。

#### 3.2 実験結果

##### 3.2.1 定量評価

経路探索を行い、各手法におけるホップ数別経路数と平均ホップ数を表 3 に示す。平均ホップ数に着目すると  $M_\alpha$  の 2.60 に対し、 $M_{\alpha+\beta}$  では 2.64、 $M_{\alpha+\gamma}$  では 3.24、 $M_{\alpha+\beta+\gamma}$  では 3.17 と増加した。これは、 $M_\alpha$  で密な関係ではない仲介者を経由していた経路が、提案手法において実際の関係に基づいて仲介者を多く通した経路に変更されたためと考えられる。

表3 手法ごとの経路数と平均ホップ数

(従来手法:  $M_\alpha$  提案手法:  $M_{\alpha+\beta}$ ,  $M_{\alpha+\gamma}$ ,  $M_{\alpha+\beta+\gamma}$ )

ホップ数	$M_\alpha$	$M_{\alpha+\beta}$	$M_{\alpha+\gamma}$	$M_{\alpha+\beta+\gamma}$
1	2190	<b>2383</b>	1351	1574
2	26850	26500	12633	13668
3	33374	31911	27157	27250
:	:	:	:	:
8	—	—	—	7
経路無し	1036	<b>518</b>	1036	<b>518</b>
平均ホップ数	<b>2.60</b>	2.64	3.24	3.17

表4 手法ごとの主観評価実験結果平均値

	$M_\alpha$	$M_{\alpha+\beta}$	$M_{\alpha+\gamma}$	$M_{\alpha+\beta+\gamma}$
Q1: 応答の迅速さ [-1,1]	0.48	0.46	<b>0.66</b>	0.53
Q2: 仲介者との親密さ [-1,1]	0.40	<b>0.42</b>	0.40	<u>0.24</u>
Q3: 経路の適切さ [0,4]	2.53	2.48	<b>2.62</b>	2.57

また  $M_\alpha$  では経路が発見できなかった 1036 パターンに対し、 $M_{\alpha+\beta}$  ではその半分の 518 パターンにおいて、実務に基づく繋がりが考慮されたことにより経路を発見することができた。

##### 3.2.2 主観評価

結果を表 4 に示す。Q1 については  $M_\alpha$  0.48 に対し、 $M_{\alpha+\gamma}$  0.66 が最も高い評価を得た。また、Q2 については  $M_\alpha$  0.40 に対し、 $M_{\alpha+\beta}$  0.42 と良い評価を得たが、 $M_{\alpha+\beta+\gamma}$  は 0.24 と減少している。これより、 $M_\gamma$  は、 $M_\alpha$  では考慮されていない対応スピードを考慮した仲介者の提示が出来ていると考えられる。同時に  $M_\beta$  は、 $M_\alpha$  では得られない実関係を補足し得ると考えられる。しかしながら、それらを組み合わせた  $M_{\alpha+\beta+\gamma}$  において高評価を得られなかったのは、本実験において各手法を均等に扱い  $\alpha = \beta = \gamma = \gamma_1 = \gamma_2 = 1$  としたためであり、実際の仲介提示を行う際に重要とする項目、すなわち対応スピードや実業務での関わりの強さなどを配慮した重み付けが必要であると考えられる。また Q3 により本実験では提示経路が仲介に最も適している手法は  $M_{\alpha+\gamma}$  となった。

#### 4 まとめ

本稿では、立場や環境が異なる面識のない 2 者を円滑に繋げる経路探索手法を提案した。性能評価実験により、従来手法では発見できなかった実世界の繋がりに即した経路を発見できた。主観評価実験により、提示された仲介経路が従来手法よりも適切であることを確認した。今後の課題として各手法の重みを最適に設定することが必要と考えられる。

#### 参考文献

- [1] 藤本義治, 他, “MAKOTO: ソーシャルグラフを用いたコミュニケーション支援システムの提案”, 情報シンポジウム, 3, pp.703-706, 2011.
- [2] 與語一史, 他, “人同士の潜在的な共有関係を推定するためのソーシャルグラフ表現方法”, 信学技報, NS, 110, 372, pp. 81-84, 2011.