

D-018

## マイクロブログにおける発言者と投稿文の信頼度推定手法

岩井 一晃<sup>†</sup> 鈴木 優<sup>‡</sup> 石川 佳治<sup>‡††</sup><sup>†</sup> 名古屋大学大学院情報科学研究科<sup>‡</sup> 名古屋大学 情報基盤センター<sup>††</sup> 国立情報学研究所

## 1 はじめに

現在, Twitter をはじめとしたマイクロブログの発展, 普及に伴い情報の発信を気軽に行うことができるようになった. マイクロブログの発展, 普及によって, 従来のマイクロブログが普及していなかった場合に比べ, 様々な情報が早く広がるようになった. ところが利用者に伝わる情報が信頼できるかどうかは利用者の判断に任されており, 正しくない情報も数多く広まっている. 正しくない情報には利用者が一見して明らかに誤まっている情報もあれば, 一見内容だけを見れば真実であるように思われる情報も存在し, 内容分析による判別は難しいと考えられる. 本研究では, 情報が信頼できるか否かの判断は利用者に依存していると考え, 利用者の選択に着目して投稿文の信頼度判別を行う. ここで利用者の選択とは, 利用者が投稿文の引用を行った場合, その投稿文を選択したと考える. 多くの信頼できる投稿文を選択する利用者は信頼度が高く, 信頼できない投稿文を引用する利用者は信頼度が低いと仮定できると考えた. 信頼度が高い利用者が選択した投稿文は正しい可能性が高い. これらの仮定を用いて利用者と投稿文の信頼度を算出し, 利用者に提示することによって, 利用者が判断するための支援を行う.

情報の信頼度に関する研究が多く行われている. Castillo ら [1] は Twitter で得られる投稿文内に URL が含まれているか否かをはじめ, 投稿文の拡散や, 利用者の属性などの情報を利用し投稿文の信頼度を決定している. 一方, マイクロブログ以外の情報に関する信頼度を求める手法として鈴木ら [2] の手法がある. この手法は Wikipedia の編集履歴から信頼情報を算出する. 信頼情報の算出は, 利用者が行った過去の編集に着目, 同一の利用者が行う編集の信頼度はどの記事も一定であるとして行われる. 本手法はこの手法と同様に, 同一の利用者が選択する信頼度は一定であるとして, 利用者の行った過去の選択に着目し, 投稿文の信頼度を算出する. また Page[3] によって Web ページの評価手法として提案された PageRank という手法がある. この手法では良質なページから張られているリンクは同様に良質であるという仮定に基づき, ウェブ

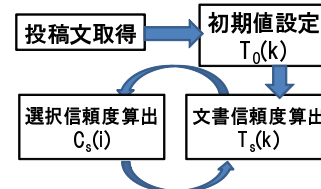


図1: 本研究のシステムの流れ

ページの質を評価する. 本手法も同様に, 高い信頼度を持つ利用者が選択する投稿文は同様に高い信頼度を持ち, 高い信頼度を持つ投稿文を選択している利用者は高い信頼度を持っているという前提に基づき投稿文の信頼度を算出する.

## 2 提案手法

本研究では引用が行われた投稿文の信頼度を算出する. 本研究のシステム構成を図1に示す.

本研究では選択された投稿文の信頼度と利用者の選択に関する信頼度を利用する. 投稿文の信頼度を文書信頼度とし,  $T_s$  と表記する. 利用者の選択に関する信頼度を選択信頼度とし,  $C_s$  と表記する. 信頼度はそれぞれ  $-1$  から  $1$  の値を取り,  $1$  に近いほど信頼できるとする. また  $s$  ( $s = 0, 1, \dots, S$ ) は値の算出まで計算が繰り返し行われた回数である. これらをくり返し使い, 文書信頼度を算出する. それぞれの信頼度の算出は以下の式に従う.

$$L_c(U_i) = \{D | D \text{ は } U_i \text{ が選択した投稿文}\} \quad (1)$$

$$L_t(D_k) = \{U | U \text{ は } D_k \text{ を選択した利用者}\} \quad (2)$$

$$C_{s+1}(U_i) = \frac{1}{|L_c(U_i)|} \sum_{D \in L_c(U_i)} T_s(D_k) \quad (3)$$

$$T_s(D_k) = \frac{1}{|L_t(D_k)|} \sum_{U \in L_t(D_k)} C_s(U_i) \quad (s \geq 1) \quad (4)$$

式 (1), 式 (2) はそれぞれ利用者が引用した投稿文の集合, 投稿文を引用した利用者の集合である. これらを後述の式 (3), 式 (4) で利用することで文書信頼度, 選択信頼度を算出する.  $U_i$ ,  $D_k$  はそれぞれ利用者, 引用された投稿文の実体であり,  $k$  ( $k = 0, 1, \dots, K$ ),  $i$  ( $i = 0, 1, \dots, I$ ) は, それぞれ利用者, 選択した投稿文に付与される ID である.  $L_c(U_i)$  は利用者  $U_i$  が引用した投稿文の集合,  $L_t(D_k)$  は投稿文  $D_k$  を引用した利用者の集合である. 式 (3) を用いて, 選択信

A Method for Evaluating Credibility of Users and Messages on Microblogs

Kazuaki Iwai<sup>†</sup>, Yu Suzuki<sup>††</sup>, Yoshiharu Ishikawa<sup>†† ‡</sup>

<sup>†</sup> Graduate School of Information Science, Nagoya University

<sup>††</sup> Information Technology Center, Nagoya University

<sup>‡</sup> National Institute of Informatics

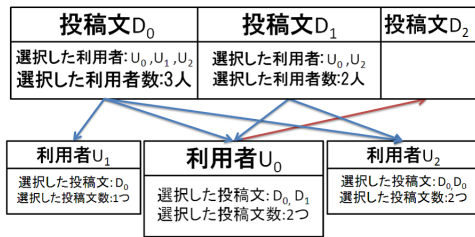


図2: 文書信頼度と選択信頼度の伝播図

信頼度を算出する． $C_{s+1}(U_i)$  は選択信頼度， $T_s(D_k)$  は文書信頼度を示す． $s$  は値の算出まで計算が繰り返した回数である．また  $|L_c(U_i)|$  は式1で求められる利用者が引用した投稿文の集合が有する要素数である．投稿文数を用いて正規化を行うことによって，選択した投稿文の数による差異を無くす．式(4)を用いて，文書信頼度を算出する． $T_s(D_k)$  は文書信頼度， $C_s(U_i)$  は選択信頼度を示す． $D_k, U_i$  はそれぞれ引用された投稿文，利用者の実体である．また  $|L_t(D_k)|$  は式(2)で求められる投稿文を引用した利用者の集合が有する要素数である．式(3)と同様に， $T_s(D_k)$  を算出する．投稿文を選択した利用者の数を用いて正規化を行うことによって選択した人数による差異を無くす．また  $T_s(D_k)$  の初期値， $T_0(D_k)$  を定める必要がある．初期値はシステム利用者によって決定される．システム利用開始時に利用される投稿文集合の中で影響力が高い複数の投稿者が投稿し，引用された投稿文それぞれ一つずつ提示する．システム利用者が提示された投稿文に対し，文書信頼度の初期値を設定することによって決定される．初期値は-1から1で入力する．ここで影響力が高い利用者とは，投稿文が多くの利用者に引用される利用者とする．

図2は利用者 $U_0$ に注目すると，利用者 $U_0$ が投稿文 $D_0, D_1$ を引用している図である．利用者 $U_0$ が新たに投稿文 $D_2$ を引用した場合を考える．この場合，投稿文 $D_2$ の文書信頼度に応じて利用者 $U_0$ の選択信頼度が再計算される．同時に，利用者 $U_0$ が既に引用している投稿文 $D_0, D_1$ の文書信頼度が再計算される．これは利用者 $U_0$ の選択信頼度が変化するためである．さらに，利用者 $U_0$ が引用している投稿文 $D_0, D_1, D_2$ を，利用者 $U_0$ が投稿文 $D_2$ 引用する前に引用していた利用者の選択信頼度も変化する．これは，新たに投稿文を選択した利用者の選択信頼度が変化することによって，文書信頼度が変化が起これば，この値の変化を受け選択信頼度の再計算を行うからである．このように利用者の新たな選択を受け，関わりのある文書信頼度と選択信頼度が収束するまで再計算され，値が更新され続ける．二種類の信頼度をそれぞれ更新し続けることにより，高精度な文書信頼度を利用者に表示することができる．

### 3 実験

Twitterより取得した引用を行った投稿文をデータセットとして作成，実験を行う予定である．本研究では文書信頼度は-1から1の値をとり，数値が1に近いほど信頼度が高いとしている．信頼度が正の投稿文を信頼できる，信頼度が負の投稿文を疑わしいとして評価する．実験を行ったデータセットに対し，信頼できるか疑わしいかを人手で決定し，正解データセットを作成する．正解データセットと本手法が算出した結果が一致した場合，適合とする．そうでない場合は不適合とする．これらより精度，再現率，F値を算出し，提案手法を評価する．また，どのような投稿文が不適合になり易いかについての考察を引用した利用者，投稿者の他に引用された投稿文から考察し，本手法の改善を図る．

### 4 まとめ

本稿では，マイクロブログの投稿文を選択した利用者に着目し，投稿文の信頼度と利用者の選択に関する信頼度を繰り返し計算することで文書信頼度を算出する手法を提案した．今後はシステムの実験評価を行い，結果の考察，手法の改善を行う予定である．考えられる問題として，投稿文の一回しかしていない利用者には信頼度が伝播せず計算できない可能性が高い．また信頼度の伝播について問題でシステム利用者に入力してもらった初期値の数も適切な個数が不明である．全てに初期値を入力するのは全てを評価できるが，システムとして意味を為さないため，適切な数を評価する必要がある．信頼度の収束に関しても，適切な収束誤差を求める必要がある可能性が高い．

### 謝辞

本研究の一部は，文部科学省科学研究費補助金(22300034, 23700113)による．ここに記して謝意を表す．

### 参考文献

- [1] Carlos Castillo, Marcelo Mendoza, and Barbadra Pobelete "Information Credibility on Twitter" WWW2011, pp.675-684, Hyderabad, India., (2011).
- [2] 鈴木 優, 吉川 正俊: 「Wikipediaにおけるキーパーソン抽出による信頼度算出精度および速度の改善」, 情報処理学会論文誌: データベース, Vol.3, No. 3 (TOD47), pp. 20 - 32, (2010)
- [3] Larry Page "PageRank: Bringing Order to the Web", Stanford Digital Library Project, talk. August 18, 1997 (archived 2002)