

# RumorFinder : 流言拡散防止のための 情報確認行動促進システム

RUMOR FINDER: INFORMATION-CREDIBILITY VERIFICATION-BEHAVIOR  
FACILITATION SYSTEM FOR PREVENTING FALSE RUMORS SPREADING

柿本大輔<sup>1</sup>・宮部真衣<sup>2</sup>・荒牧英治<sup>3</sup>・吉野孝<sup>4</sup>

Daisuke KAKIMOTO, Mai MIYABE, Eiji ARAMAKI and Takashi YOSHINO

<sup>1</sup>システム工学研究科, <sup>2</sup>公立諏訪東京理科大学工学部講師,

<sup>3</sup>奈良先端科学技術大学院大学研究推進機構特任准教授, <sup>4</sup>システム工学部教授

SNSの一つであるマイクロブログサービスでは、ユーザがリアルタイムに多種多様な情報を取得・発信することが可能であり、近年その普及が著しい。マイクロブログでは手軽に情報を発信できるが、その手軽さゆえに、流言が伝播されるという問題も起こっている。流言はユーザ間の適切な情報共有を阻害し、特に災害時などでは、深刻な問題が引き起こされる可能性があり、流言の拡散を防止する仕組みが必要である。本研究では、流言である可能性が含まれる情報を閲覧している際、ユーザに気づきを与え、情報の真偽確認行動を促進することで、流言の拡散防止を支援するシステムを構築した。システムの一般公開による、利用状況の分析および評価の結果、システムはユーザに情報の真偽確認を促すことができる可能性があることがわかった。

キーワード : SNS, Twitter, 流言, 情報確認行動, インタフェース

## 1. はじめに

近年、SNSおよびマイクロブログサービスが普及しており、ユーザはリアルタイムに多種多様な情報を取得・発信することが可能である。代表的なマイクロブログサービスの一つに、Twitter<sup>1)</sup>がある。Twitterでは、情報発信時の入力可能文字数が140字に制限されていることなどから、ユーザの情報発信に対するハードルが大きく下がっている<sup>2)</sup>。2011年3月に発生した東日本大震災の際には、Twitterは、速報性の高さを発揮し、重要な情報インフラとして活用されていた<sup>3)~6)</sup>。

一方、情報の取得・発信時に、誰もが情報の信頼性を正しく判断することができるとは限らず、流言が伝播されるという問題も起こった<sup>7)</sup>。東日本大震災の際に、Twitterにおいて多く拡散された流言の一つに、「放射性物質にはうがい薬が効く」という内容のものがある。この流言に対し、後に独立行政法人放射線医学総合研究所から訂正と注意喚起がなされた<sup>7)</sup>。このような流言は、専門知識を持たないユーザにとっては真偽の判断が難しく、その情報を鵜呑みにして行動に至ると、人間の身体

に有害な影響を及ぼす可能性がある。流言は、ユーザ間の適切な情報共有を阻害し、その伝播が深刻な問題を引き起こす可能性がある。そのため、流言の拡散を防止する仕組みが必要である。

これまでに我々は、流言の拡散防止を目的とした、情報確認行動促進システムの開発を行ってきた<sup>8)</sup>。本システムは、流言である可能性が含まれる情報を閲覧している際、ユーザに気づきを与えることで、情報の真偽確認行動を促進し、流言の拡散防止を支援する。WebページおよびTwitterタイムライン(フォローしているユーザの発信したツイートが一覧で表示された画面)閲覧場面对象としたシステムの評価実験の結果、本システムは、ユーザのページ閲覧を妨げることなく、流言に関する気づきを与えられる可能性があることがわかった<sup>8)</sup>。しかし、ユーザに対し、流言に関する気づきをより効果的に与えるためには、システムのインタフェースの改善が必要であることがわかった。

本稿では情報確認行動促進システムと、そのインタフェースの改善について述べる。また、システムの一般公開による、ユーザの利用状況を分析し、システムの効果について評価および考察を行う。

## 2. 関連研究

Twitterでは情報発信の手軽さゆえに、流言が多数発信・拡散されており、Twitterにおける流言に関する研究は盛んである。Twitterにおける流言拡散防止に関する研究としては、流言の検出および信頼性の自動推定に着目したものがあつた。Takahashiら<sup>9)</sup>は、Twitter上の流言の出現自体を抑えることは困難であるが、訂正ツイートが発信されることで、流言の拡散を抑制することができることを述べている。そこでTakahashiら<sup>9)</sup>は、訂正ツイートの早期発信に役立てるため、ツイートの固有表現、リツイート比率、キーワードにもとづくフィルタリングによる、Twitterにおける流言検出手法を提案した。Castilloら<sup>10)</sup>は、ソーシャルメディアの中には、ユーザが情報の信頼性判断をすることができるシグナルがあると考え、ソーシャルメディア上に存在する情報のみを用いて情報の信頼性を自動評価する手法を提案した。Yang<sup>11)</sup>らは、Sina Weibo<sup>12)</sup>を対象に、Castillo<sup>10)</sup>らと同様の手法を用いて、情報の信頼性を自動評価する手法を提案した。

また、Twitterにおけるリツイート機能(他のユーザが発信したツイートを再発信する機能)は、その手軽さからTwitter上での情報の伝播に影響を与えやすいと考えられ、特にリツイート機能に着目して、情報の信頼性推定を行った研究がある。中原ら<sup>13)</sup>は、Twitterにおけるリツイートのうち、非公式リツイート(引用リツイート)に付加されたコメント、およびリツイートされた回数から、あるツイートに対し訂正ツイートと危険度という2つの情報を取得し、ユーザに提示するシステムを提案した。

このように、流言自体の検出および信頼性の自動推定を試みる研究は多数行われており、流言として検出あるいは信頼性が低いと推定された情報は、流言の拡散防止の一助となる可能性がある。しかし、これらの研究はあくまでも流言と思われる情報を検出することに主眼を置いたものであり、検出した情報をユーザに提供する方法や、ユーザの不用意な拡散を踏みとどまらせる方法など、インタフェース面での検討は行われていない。検出した情報を人々に適切かつ効果的に提供する方法が実現されることで、流言の拡散防止につながると考えられる。そのため、情報を閲覧している段階で、情報確認行動を促すような仕組みおよびインタフェースが必要である。そこで本研究では、ユーザに対し流言に関する気づきを提供し、情報の真偽確認行動を促進することで、流言拡散防止を目指すシステムを検討する。さらに、Twitter上拡にとどまらず、より広い場面においてユーザの流言拡散を防ぐためのシステムを目指す。

また、流言に関する気づきの提供および情報確認行動の促進においては、これまでに我々が提案した流言情報蓄積システム<sup>14)</sup>により収集した流言情報を用いて、ユーザに情報提供を行う。

## 3. 情報確認行動促進システム

流言が拡散する要因の一つとして、人間による情報の真偽確認不足が考えられる。これまでに我々は、Twitterにおける情報確認行動に関する調査を行った。その結果、リツイート対象とするツイート内容の真偽を意識せずにリツイート機能を利用しているTwitter利用者が多く、流言が拡散されやすい可能性があることがわかつた<sup>7)</sup>。これまでに、Twitter上で流言の拡散防止を支援する手法およびシステムは提案されている<sup>9) 10) 11) 13)</sup>が、これらは、ユーザの能動的なシステムの利用が前提になっている。情報の真偽に関する関心の低さを鑑みると、流言拡散防止支援システムを能動的に利用してまで、ツイート内容の真偽を確認しようとする人々は少ないと考えられる。流言の拡散を防止するためには、このような人々に情報の真偽確認行動を促すことが重要であると考えられる。

そこで我々は、流言である可能性が含まれる情報を閲覧している際、ユーザに気づきを与えることで、情報の真偽確認行動を促進するシステムを開発した。

### (1) 設計方針

Twitterユーザは情報の真偽を確認することなくリツイートすることが多いことを考慮すると、システムの構築においては、ユーザによる真偽確認負担を軽減することが重要であると考えられる。また、流言はTwitterだけでなく、その他のSNSやWebページ上にも掲載される可能性があるため、Twitterに限定せず、インターネット上の幅広い情報閲覧場面において、流言拡散を防止できることが望ましい。

システムの設計方針を以下に示す。

1. 真偽確認作業の負荷を軽減する  
ユーザが目にする全ての情報に関して、ユーザ自身が能動的に真偽確認することは難しい。そこで、ユーザが閲覧中のページに含まれる情報に、流言が含まれるかどうかをシステムが判定し、流言が含まれる場合、ユーザが受動的に情報を受け取ることができる仕組みを構築する。
2. 幅広い場面での対応を可能にする  
流言はTwitterだけでなく、その他のSNSやWebページなどにも掲載される可能性がある。そこで、Webブラウザのアドオン機能として実装することにより、ブラウザ上での様々なページ閲覧時に、システムから情報が提供されるようにする。
3. 最新の流言情報に基づいて情報を提供する  
ユーザに注意を促す際には、新しく発生した流言情報などにも対応できることが望ましい。そこで、2012年から運用している流言情報クラウド<sup>14)</sup>と連携し、最新の流言情報を提供可能にする。

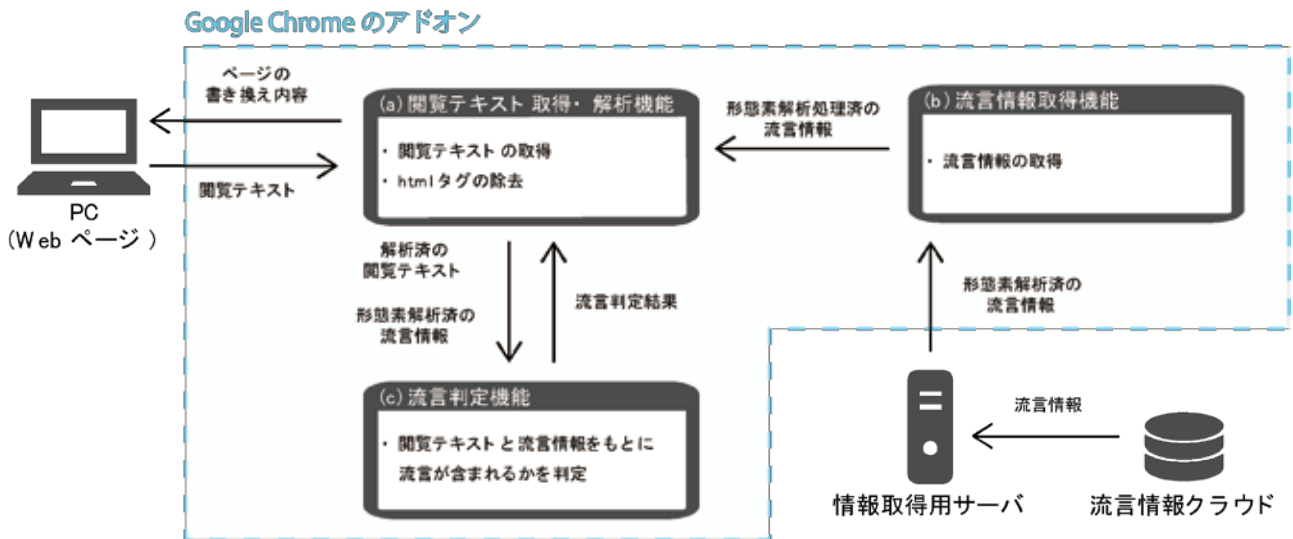


図-1 システムの構成

## (2) システムの構成

本システムは、Google<sup>15)</sup>社が提供するWebブラウザである、Google Chrome<sup>16)</sup>のアドオンとして実装した。現在、様々なブラウザが普及しているが、Google Chrome(バージョン: 69.0)の全世界シェアは、2018年11月20日時点で47.0%であり、1位となっている<sup>17)</sup>。また、アドオンなど、機能の拡張がしやすいという特徴もある。そこで、今回はGoogle Chromeのアドオンとしてシステムの開発を行うことにした。

システムの構成を図-1に示す。まず、(a)閲覧テキスト取得・解析機能により、Webブラウザで閲覧しているコンテンツのテキスト(閲覧テキスト)を取得する。同時に、(b)流言情報取得機能により、データ取得用サーバを介して、流言情報を取得する。次に、取得した閲覧テキストおよび流言情報をもとに、(c)流言判定機能により、流言判定を行う。最後に、その結果をもとに閲覧中Webページ上でのテキストの強調表示と、該当部分のマウスオーバーによる吹き出しの表示という形式でページの書き換えを行い、ユーザーに流言に関する気づきを提供する。

システムの動作画面例を図-2に示す。システムは、ユーザーが閲覧中のWebページ内に、流言である可能性のある情報が含まれる場合、該当テキストの強調表示を行う。また、吹き出しの表示についてはユーザーのページ閲覧の妨げにならないよう、強調表示箇所のマウスオーバーにより、ユーザーが任意で表示させられるようにした。吹き出しには、流言に関する情報(判定された流言内容)、訂正情報(流言情報に対する訂正ツイート)、訂正数(訂正情報の数)、Web検索結果へのリンク(判定された流言に含まれる名詞をクエリとしたWeb検索ページへのリンク)、および詳細リンク(これまでの訂正数および直近一週間の訂正数の推移を表すグラフを確認できるページへのリンク)が含まれる。

## (3) 各機能の実装

3章(2)節で述べた(a)閲覧テキスト取得・解析機能、(b)流言情報取得機能および(c)流言判定機能の実装手法を以下に示す。

- 閲覧テキスト取得・解析機能**  
アドオンを起動している場合、Webページに挿し込まれるスクリプトファイルにより、ユーザーが閲覧しているテキストを取得し、不要なhtmlタグの除去を行う。また、流言判定の結果をもとに、閲覧ページの書き換えを行う。
- 流言情報取得機能**  
バックグラウンドで常時動作しているスクリプトファイルにより、情報取得用サーバとの通信を介して、流言情報クラウドから最新の流言情報の取得を行う。また、情報取得用サーバでは定期的に流言情報クラウドから流言情報を取得し、Juman<sup>18)</sup>による形態素解析を行う。
- 流言判定機能**  
取得した閲覧テキストと形態素解析済みの流言情報をもとに、閲覧中のテキストに流言が含まれているかどうかを判定する。この際、閲覧テキストを一行ごとに分割し、流言の形態素解析結果のうち、名詞に分類された形態素が全て含まれていれば、そのテキストに流言が含まれているとみなす。

## 4. インタフェースの改善

1章で述べたように、これまでの評価実験の結果、システムは、ユーザーのページ閲覧を妨げることなく、流言に関する気づきを与えられる可能性があることがわかっている。しかし、ユーザーに対し、流言に関する気づきを



流言判定箇所の強調表示：流言が含まれる可能性のある部分を判定し、強調表示。  
 マウスオーバーによる吹き出しの表示：強調表示部分のマウスオーバーで、吹き出しを表示。  
 流言内容：判定された流言。  
 訂正情報：流言情報に対する訂正ツイート。  
 訂正数：訂正情報の数。  
 Web検索結果へのリンク：判定された流言情報における名詞をクエリとしたWeb検索ページへのリンク。  
 詳細リンク：これまでの訂正数および直近一週間の訂正数の推移を表すグラフを確認できるページへのリンク。

図-2 システムの動作画面

より効果的に与えるためには、「Twitterタイムライン閲覧場面のような様々な情報が混在する環境においても、ユーザが流言に関する情報を見落とさないようにする」「システムがユーザの情報閲覧の妨げにならないよう、ユーザ側の設定に応じて情報提供方法を変更可能にする」などの、システムのインターフェースの改善が必要であることがわかった。そこで我々は、新たに以下の機能を本システムに追加した。

#### 機能1. トースト通知機能

Twitterのタイムラインのような環境では、様々なユーザの発信した情報が混在しているため、流言に関する情報が提示されても、見落としてしまう可能性がある。そこで、このような環境でもユーザが流言に関する情報を見落とさないよう、トースト通知機能を追加した。トースト通知機能の動作例を、図-3に示す。この機能は、ページを読み込んだ際に、テキストを強調表示するだけでなく、Webブラウザの画面右上に、検出した流言情報リストを提示する機能である。

#### 機能2. カスタマイズ機能

普段Webを閲覧している際と同様に、ユーザにシステムを利用してもらうためには、強調表示箇所などに応じて、流言に関する気づきの提供方法を変更できることが望ましい。そこで、カスタマイズ機能を追加した。カスタマイズ機能の動作例を、図-4に示す。カスタマイズ機能により、Webペー

ジ閲覧時、本システムにおけるテキストの強調表示、トースト通知および吹き出し表示の各機能を動作させるかどうかを、ユーザは自由に切り替えることができる。また、カスタマイズ機能はWebブラウザの画面右上のアイコンから起動することができ、図-4上部に示すように、検出した流言情報のリストアップも行う。

## 5. 利用状況の分析と考察

### (1) 利用状況

我々は、Chromeウェブストアを介して、2018年7月4日より本システムを一般公開している(システム名は「Rumor Finder (ルーマーファインダー)」としている)。また、本システムがどのように利用されるのかを分析するために、25人の大学生および大学院生にシステムを利用してもらうよう呼びかけた。呼びかけの際、システムの概要を紹介するWebページ<sup>19)</sup>のURLを伝え、利用ログの送信を許可してもらうよう伝えた。その他の指示は与えず、自由に使ってもらったこととした。

利用ログについては、閲覧ページ、検出流言、強調表示箇所のマウスオーバー、詳細リンクへの遷移およびWeb検索リンクへの遷移について、2018年8月2日より継続して取得している。そこで本稿では、2018年8月2日～2018年11月19日の110日間の利用ログについて分析を行





図-3 トースト通知機能



図-4 カスタマイズ機能

う。なお、2018年11月19日時点でのシステムのユーザ数は59人であり、期間中に利用ログの送信を一度でも行ったユーザ数は47人であった。ただし、利用ログは端末ごとに取得しているため、同一ユーザによる複数端末におけるシステムの利用および複数ユーザによる同一端末におけるシステムの利用が考えられる。本稿では、利用ログを取得した端末1台を1ユーザとみなす。また、利用ログの送信を行った47人のユーザのうち、11人のユーザは、閲覧ページ数の合計が10件未満であった。本システムは、一度インストールされると、ユーザに対して流言情報の検出が自動通知される仕組みを取っており、システムからの通知に対し、確認行動を取るかどうかはユーザの自由である。ただし、本システムが流言の拡散防止に貢献するためには、インストール直後に起こりがちな一過性の利用ではなく、積極的な確認行動をユーザが継続できるようにすることが求められる。一過性の利用のみで、その後利用が見られないユーザについては、一般公開によりシステムをダウンロードしたユーザであるため、なぜ利用しないのか等を直接調査することが難しい。一方、

継続した利用が見られるユーザについては、その利用状況を確認することが可能である。継続利用ユーザがどのように利用しているのかを分析することにより、継続利用につながる特徴などが確認できる可能性がある。そこで、本稿では、閲覧ページ数の合計が10件未満であった11人のユーザを除外し、継続してシステムを利用していると考えられる残りの36人のユーザを対象として利用状況を分析する。

## (2) 分析結果と考察

5章(1)節で述べた36人のユーザの利用ログにおける、2018年8月2日～2018年11月19日の110日間の閲覧ページ数の合計は348,774件であった。そのうち、閲覧ページ数が多かった上位10人のユーザにおける同期間の閲覧ページ数の合計は323,662件であり、約92.80%を占めている。そのため本稿では、これらのユーザの利用状況について議論する。

閲覧ページ数が多かった上位10人のユーザのシステムの利用状況(閲覧ページ数、流言検出ページ数、検出割合および吹き出し表示回数)について、表-1に示す。なお、表-1では、システムの概要を紹介するWebページおよび流言情報クラウドのWebページにおける動作ログは除外している。表-1より、(F)流言検出割合については、ユーザ1～10において、5%未満となった。また、全てのユーザにおける平均は約2.02%であった。この結果より、ユーザがWebページ閲覧時、本システムにより流言の検出が通知される頻度は低いことがわかった。この理由としては、(1)そもそも、災害時など以外の平常時では、ユーザが流言情報を目にする機会がさほど多くないこと、(2)システムが流言として検出すべき情報を、正しく検出できていない可能性があることが考えられる。

(G)吹き出し表示割合については、ユーザ1～10において、平均約31.61%となった。ただし、吹き出しの表示を複数回行うユーザがいるものの、全く行わないユーザが1人いることもわかった。また、全てのユーザにおける平均は、約22.78%となった。(H)詳細リンク遷移割合については、ユーザ1～10において平均約5.85%、(I)Web検索リンク遷移割合については、ユーザ1～10において平均約3.51%となった。また、36人の全てのユーザにおける平均は、(H)詳細リンク遷移割合については約3.09%、(I)Web検索リンク遷移割合については、約5.78%となった。この結果より、システムが行う強調表示により、ユーザに対して吹き出し表示による情報確認行動を促進できる可能性があることがわかった。しかし、吹き出し表示を全く行わないユーザもいることから、流言に関する情報の提供方法の追加・変更が必要である可能性がある。そのため、トースト通知機能およびカスタマイズ機能のような、テキストの強調表示および吹き出し表示以外の情報の提供方法について、今後検討する必要があると考えられる。さらに、詳細リンクおよびWeb検索リン

表-1 システムの利用状況

	ユーザ									
	1	2	3	4	5	6	7	8	9	10
(A)閲覧ページ数(件)	249411	17186	11430	11324	8111	7097	6668	6175	3205	3055
(B)流言検出ページ数(件)	31	182	70	67	40	37	22	274	12	21
(C)吹き出し表示回数(回)	42	53	33	5	18	2	2	34	3	0
(D)詳細リンク遷移回数(回)	0	5	4	1	2	0	0	0	0	0
(E)Web検索リンク遷移回数(回)	0	9	3	0	1	0	0	0	0	0
(F)流言検出割合(%)	0.01	1.06	0.61	0.59	0.49	0.52	0.33	4.44	0.37	0.69
(G)吹き出し表示割合(%)	135.48	29.12	47.14	7.46	45.00	5.41	9.09	12.41	25.00	0.00
(H)詳細リンク遷移割合(%)	0.00	9.43	12.12	20.00	11.11	0.00	0.00	0.00	0.00	-
(I) Web検索リンク遷移割合(%)	0.00	16.98	9.09	0.00	5.56	0.00	0.00	0.00	0.00	-

- (A)閲覧ページ数において、同一のWebページの重複に関する考慮は行っていない。  
 (B)流言検出ページ数は、1件以上の流言情報が検出されたWebページの数である。  
 (F)流言検出割合は、(B)流言検出ページ数(A)閲覧ページ数\*100の値である。  
 (G)吹き出し表示割合は、(C)吹き出し表示回数(B)流言検出ページ数\*100の値である。  
 (H)詳細リンク遷移割合は、(D)詳細リンク遷移回数(C)吹き出し表示回数\*100の値である。  
 (I)Web検索リンク遷移割合は、(E)Web検索リンク遷移回数(C)吹き出し表示回数\*100の値である。

表-2 検出した流言情報

	流言情報	検出数(回)
1	個人情報が抜き取られた	412
2	携帯で1、1、0を押して電話を掛ける	83
3	NYTIMESは韓国を批判している	77
4	中国人が.....	70
5	できてすらいない本を手売りしてたらしい	53
6	この人は虫としゃべれる	46
7	沖縄は基地で食っている	45
8	この花火大会見たらわかれまーす!	45
9	#BTSがユニセフに原爆看板を持ち込み	36
10	学校では決して教えてくれない情報	29

クのような他ページへの遷移を行ってまで、情報の真偽確認行動を行うユーザは少ない可能性があることがわかった。そのため、吹き出しの表示による情報確認行動以外の、他ページへの遷移を行うなどの情報確認行動を促進するため、システムの提供する情報およびそのインタフェースについて今後再検討する必要があると考えられる。

### (3) 検出流言と誤検出

2018年8月2日～2018年11月19日の110日間に、システムが検出した流言情報について、検出回数が多かった上位10件の流言情報を表-2に示す。表-2より、システムは同期間中に、流言情報2「携帯で1、1、0を押して電話を掛ける」および流言情報9「#BTSがユニセフに原爆看板を持ち込み」のような、実際にインターネット上で拡散された流言情報を検出した。しかしその一方で、流言情報1「個人情報が抜き取られた」のような、議論の対象が不明瞭なものおよび流言情報6「この人は虫としゃべ

れる」のような、そもそも真偽を確認する必要のないような流言情報も検出された。5章(2)節で述べたように、詳細リンクおよびWeb検索リンクのような他ページへの遷移を行ってまで、情報の真偽確認行動を行うユーザは少ない可能性があるが、このような流言情報をシステムが検出した際、ユーザは吹き出し表示以外の確認行動を行う必要がない。そのため、流言情報の重要度や拡散時の深刻度の違いなどについて検討し、流言情報の内容に合わせて情報確認行動の促進方法を変更するなどの改善が必要である可能性がある。

また、本システムは「不具合の報告」として、本システムの問題点および要望などを自由に記述し、送信できる機能を備えている。本稿で述べた36人のユーザのうち2人のユーザから「携帯電話番号の文字が流言として検出される」などの、「あるページの閲覧時、流言情報が含まれていないにもかかわらず、システムが特定の流言情報を検出した」という旨の報告が得られた。実際にはWebページの中に含まれていなかったにもかかわらず、システムが誤検出したと報告された流言情報は、表-2における流言情報2「携帯で1、1、0を押して電話を掛ける」であった。システムの流言判別手法には、流言情報の形態素解析結果における名詞が、判別対象とするテキスト内に全て含まれるかどうか、という基準が含まれる。

「携帯で1、1、0を押して電話を掛ける」という流言情報の形態素解析の結果、名詞は“携帯”および“電話”のみとなった。そのため、ユーザが閲覧しているWebページ内に、“携帯”および“電話”の両方が存在している場合、その箇所を流言情報として判別する可能性がある。これらの結果より「流言として判別すべきでない情報を、システムが流言として判別することがある」という問題があることがわかった。また、表-2における流言

情報1「個人情報が抜き取られた」に関しては、検出数が412回となっており、他の流言情報と比較して検出回数が多い。この流言情報の形態素解析の結果、名詞は“個人”および“情報”のみとなった。流言情報2「携帯で1、1、0を押して電話を掛ける」と同様、システムが誤検出を行ったことが考えられる。5章(2)節で述べたように、システムがユーザに流言検出を通知する頻度は低い。しかし、ユーザがシステムを利用する際、流言情報とは関係のない情報が強調表示されると、ユーザのWebページ閲覧を妨げ、システムが利用されなくなる可能性がある。そのため、システムによる流言情報の判別手法について、再検討する必要がある。

## 6. おわりに

我々はこれまでに、ユーザがWebページを閲覧している際、流言が含まれる可能性のある部分を強調表示し、吹き出しにより流言に関する情報を提示する情報確認促進システムを開発してきた。情報の真偽に対して関心の低いユーザにシステムを利用してもらうためには、ユーザに対し、流言に関する気づきをより効果的に与えることが重要である。

本稿では、システムのインタフェースに関する改善を行った。Twitterのタイムラインのような、様々なユーザの発信した情報が混在している環境でも、ユーザが流言に関する情報を見落とさないよう、トースト通知機能を追加した。また、普段Webを閲覧している際と同様に、ユーザにシステムを利用してもらうために、カスタマイズ機能を追加し、流言に関する気づきの提供方法を変更可能にした。

さらに、システムの利用状況の分析および評価の結果、システムが流言の検出をユーザに通知する頻度は低いこと、流言情報と判別すべき部分を、システムが流言情報として判別できない場合があることがわかった。また、流言情報と判別すべきでない部分を、システムが流言情報として判別する場合(誤検出)があることがわかった。そのため、今後はシステムによる流言情報の判別手法を再検討する必要がある。さらに、本システムはユーザに対し、吹き出し表示による情報確認行動を促進することができる可能性があるものの、吹き出し表示を全く行わないユーザもいること、他ページへの遷移を行ってまで情報確認を行うユーザは少ないことがわかった。そのため、システムの提供する情報およびインタフェースについて、今後さらに検討する必要がある。

今後は、システムによる流言情報の判別手法の再検討およびシステムの提供する情報およびインタフェースの再検討を行う。また、ユーザの利用ログをさらに収集し、システムがユーザの流言拡散防止を支援することができるかどうかを検証する。

謝辞：本研究は、JSPS科研費 15H05317の助成による。

## 付録 RumorFinderのインストール

本システム「Rumor Finder(ルーマーファインダー)」は、Chromeウェブストア上で一般公開している。インストール用URL(本システムの機能を概説するWebページのURL)は以下である。

[http://www2.yoslab.net/~kakimoto/addon\\_intro/index.html](http://www2.yoslab.net/~kakimoto/addon_intro/index.html)  
また、インストール手順については付録図-1に示す。



付録図-1 システムのインストール手順

## 参考文献

- 1)Twitter, <<http://twitter.com>>, 2018年12月4日アクセス
- 2)垂水浩幸: 実世界インタフェースの新たな展開:4. ソーシャルメディアと実世界, 情報処理学会誌, Vol.51, No.7, pp.782-788 (2010).
- 3)三浦麻子, 鳥海不二夫, 小森政嗣, 松村真宏, 平石界: ソーシャルメディアにおける災害情報の伝播と感情:東日本大震災に際する事例, 人工知能学会論文誌, Vol.31, No.1, pp. NFC-A 1-9 (2016).
- 4)インプレスR&D インターネットメディア総合研究所: インターネット白書2011, インプレスジャパン, pp.44-47 (2011).
- 5)西谷智広: I 見聞録:Twitter 研究会, 情報処理学会誌, Vol.51, No.6, pp.719-724 (2010).
- 6)立入勝義: 検証東日本大震災そのときソーシャルメディアは何を伝えたか?, ディスカヴァー・トゥエンティワン,

- pp.20-26 (2011).
- 7) 荻上チキ: 検証東日本大震災の流言・デマ, 光文社, pp.27-28 (2011).
  - 8) 柿本大輔, 荒牧英治, 宮部真衣: 流言拡散防止のための情報確認行動促進システムの構築, ヒューマンインタフェース学会論文誌, Vol.20, No.1, pp.1-11 (2018).
  - 9) Takahashi, T. and Igata, N.: Rumor detection on twitter, Soft Computing and Intelligent Systems (SCIS) and 13th International Symposium on Advanced Intelligent Systems (ISIS), 2012 Joint 6th International Conference on, IEEE, pp.452-457 (2012).
  - 10) Castillo, C., Mendoza, M. and Poblete, B.: Information Credibility on Twitter, Proceedings of the 20th international conference on World wide web, ACM, pp. 675-684 (2011).
  - 11) Yang, F., Liu, Y., Yu, X. and Yang, M.: Automatic detection of rumor on Sina Weibo, Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics, No. 13, pp. 1-7 (2012).
  - 12) Sina Weibo, <<http://weibo.com>>, 2018年12月4日アクセス
  - 13) 中原英美, 富永一成, 牛尾剛聡: リツイート構造を用いたデマ拡散防止支援手法, DEIM Forum 2012, F2-3, pp.1-7 (2012).
  - 14) 宮部真衣, 灘本明代, 荒牧英治: 人間による訂正情報に着目した流言拡散防止サービスの構築, 情報処理学会論文誌, Vol.55, No.1, pp.563-573(2014).
  - 15) Google, <<http://www.google.co.jp/>>, 2018年12月4日アクセス
  - 16) Google Chrome, <<http://www.google.co.jp/chrome/browser/desktop/index.html>>, 2018年12月4日アクセス
  - 17) WebブラウザシェアランキングTOP10(日本国内・世界), <[https://webrage.jp/techblog/pc\\_browser\\_share/](https://webrage.jp/techblog/pc_browser_share/)>, 2018年12月4日アクセス
  - 18) Kawahara, D. and Kurohashi, S.: A Fully-Lexicalized Probabilistic Model for Japanese Syntactic and Case Structure Analysis, In Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL2006), pp.176-183(2006).
  - 19) RumorFinder, <[http://www2.yoslab.net/~kakimoto/addon\\_intro/index.html](http://www2.yoslab.net/~kakimoto/addon_intro/index.html)>, 2018年12月4日アクセス
  - 20) 「110」を入力すると携帯電話の通信速度が速くなるというデマが拡散 - ライブドアニュース, <<http://news.livedoor.com/article/detail/10512608/>>, 2018年12月11日アクセス
  - 21) BTSがユニセフに「原爆看板を持ち込み」はデマ ネットで拡散, <<https://www.buzzfeed.com/jp/kotahatachi/bts-debunk>>, 2018年 12月11日アクセス

(2018.12.14受付)