

情報への危機意識向上のための流言注意喚起ボットの提案

Proposal of a Groundless-Rumor Alert Bot to Raise Awareness of the Danger of Information

西村 涼太[†] 吉野 孝[†] 平林 (宮部) 真衣[‡]
 Ryota Nishimura Takashi Yoshino Mai Miyabe Hirabayashi

1. はじめに

近年, SNS の普及により個人の情報発信が増加している [1]. ソーシャルメディアは, 投稿や共有を気軽にでき, また誰でも発信が可能であるという特性から, 正しい情報だけでなく, 流言¹を伝播してしまうという問題がある. 東日本大震災時には多数の流言が伝えられ, 中には特定の人種を差別するものや, 誤った避難情報など, 人命に関わる深刻な事例もあった [2]. また, 2019 年末に発生した新型コロナウイルスの感染が拡大する中, 多数の流言が拡散し, 物品の買い占めなど社会的に大きな影響を及ぼした. これに対して, マイクロブログサービスの一つである Twitter²では, 新型コロナウイルスに関する偽情報ツイートを削除する対策を講じた [3] が, 偽情報とする情報の線引きの困難さや, 言論の自由に関する問題が指摘されている. このように, 災害等で社会情勢が不安定な状況では, 特に多くの流言が伝えられる.

また, Twitter などの SNS で拡散した流言は, その利用者の間でのみ共有されるわけではない. 総務省のデータによれば, Twitter ユーザは 10 代および 20 代が中心 [4] であるが, 2020 年 2 月に話題となった「トイレトペーパーが品薄になる」という流言は, 世代を超えて人々の行動に影響を与えた [5]. これは, 流言を信じた人々が, 他の人に伝えたり, 実際に購買行動を起こしたりすることで, それがメディアで報道され, さらに多くの人に広まるといったことを繰り返した結果である. このように, SNS で拡散している流言は, それを直接閲覧していない人々にも広がる可能性がある.

そこで我々は, 日常的に情報への危機意識を高める仕組みが必要であると考え, 流言の注意喚起を行うシステムの提案を行う. このシステムは, 自然言語で応答可能なチャットボットが, ユーザからの問いかけに関連する流言の提示や, 深刻度の高い流言の注意喚起を行うものである. 本稿では, 事前に実施した, 情報共有と流言の注意喚起に関するアンケートの結果や, 関連研究をもとに, 流言の注意喚起を行うシステムを提案する.

2. 関連研究

流言の拡散を防止する研究として, 流言の検出を行う研究がある [7]. 鳥海らは, Twitter 上の流言注意喚起ツイートを利用し, ツイートが流言であるか判定するアルゴリズムを提案した. 東日本大震災時のツイートデータを用いて

表 1: 回答者の年代と性別の構成

年代	男性 (名)	女性 (名)	合計 (名)	割合
10 代	5	12	17	10 %
20 代	67	89	156	88 %
30 代	0	2	2	1 %
40 代	0	0	0	0 %
50 代	0	1	1	1 %
合計	72	106	178	100 %

※割合は小数点第 1 位で四捨五入

実験を行った結果, 提案した手法は, 80 %以上の精度で流言を検出可能であることがわかった. また一方で, 判定は Twitter 上のツイートにもとに行われるため, 流言判定の根本的な部分は人の判断に依存していると述べている. また, 高精度な検出が可能であっても, 最終的にその情報を信頼するかどうかは, ユーザが自ら判断を行う必要がある. したがって, 情報への危機意識を高める仕組みが必須であると我々は考える.

柿本らは, Twitter ユーザ 108 人にアンケートを行い, リツイート機能³を利用する際に, 情報の真偽を確認していると答えた人の割合が, 40 %を下回ることを示した [6]. また, Web ページ上の流言を強調表示することで, 情報を閲覧するユーザに流言に対する気づきを与える仕組みを提案している. 強調表示に利用する流言情報は, あらかじめ Twitter 上から収集したものである. 実験の結果, システムはユーザの情報拡散行動前に, 流言に関する気づきを提供するという目的を満たしていることを示した. しかし, 1 章でも述べたように, 流言は Twitter や Web ページ上の情報を閲覧する若者の間でのみ広まるのではなく, 口頭やテキストメッセージなどのコミュニケーションを介して, 幅広い年齢の人々に拡散していくものであると考えられる. したがって, Web ページの閲覧時に気づきを提供するだけでなく, 幅広い年代の人々が日常的に利用し, 情報への危機意識を高められるシステムを構築する必要がある. そこで, 本研究では, 身近な SNS 上で動作する, 流言の注意喚起を行うボットを構築する.

3. アンケート

流言の注意喚起を行う際の要件を明らかにするために, SNS の利用頻度や情報共有の実態に関するアンケートを実施した. アンケートは Web 上で回答してもらった. 表 1 にアンケート回答者の年齢と性別の構成を示す. 回答者 178 名中, 年代は 20 代が最も多く, 全体の 88 %となった. 男女比は 4:6 である.

[†] 和歌山大学 システム工学部, Faculty of Systems Engineering, Wakayama University

[‡] 東京大学大学院 医学系研究科, Graduate School of Medicine and Faculty of Medicine, The University of Tokyo

¹本研究では, 十分な根拠がなくその真偽が不明, または真偽が人々に疑われている情報を流言と定義し, 発生過程での悪意の有無は問わないものとする.

²<https://twitter.com>

³他のユーザが発信したツイートを再発信することができる機能.

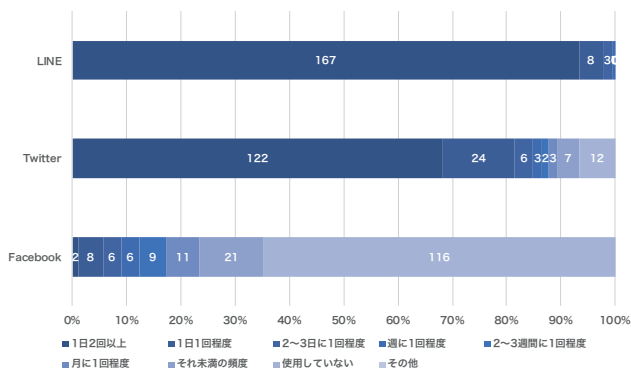


図1: SNSの利用頻度

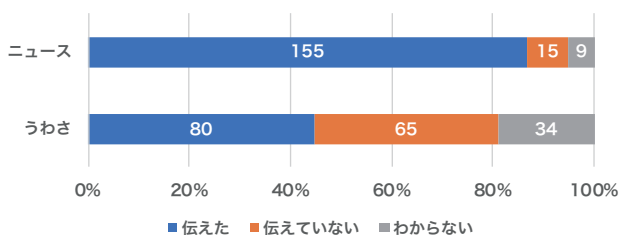


図2: 【ニュース/うわさ】を伝えたか

LINE¹, Twitter, Facebook²の利用頻度に関するアンケート結果を図1に示す。LINEおよびTwitterは、1日に2回以上と答えた人とが大部分を占めており、LINEでは93%、Twitterでは69%である。一方で、Facebookは使用していないと答えた人が65%となった。これらの傾向は、回答者の88%が20代だったことが影響していると考えられる。また、総務省のデータ[4]によれば、主要なSNSであるLINE, Twitter, Facebookのなかで、最も幅広い年齢層の人々が利用しているのはLINEであることが示されている。実施したアンケート結果と総務省のデータを考慮し、日常的に様々な年齢層の人々が利用することを想定した本システムは、LINE上で動作するチャットボットとして構築することとした。使用頻度の高いLINE上で動作するシステムにすることで、システム利用の敷居が低くなるように試みる。

次に、情報共有に関するアンケートの結果を図2に示す。このアンケートでは、テレビやWebなどで報道、掲載されたものを「ニュース」、事実かどうか明確でないものを「うわさ」としている。1ヶ月以内に、人にニュースを伝えた人と答えた人は87%、うわさを伝えた人と答えた人は45%となった。どちらも、伝えていないと答えた人を上回る結果となった。また、うわさについて、(伝えたかどうか)わからないと答えた人が19%となり、ニュースを大きく上回った。この結果から、伝える情報の真偽を意識していない人が一定数いる可能性が示唆される。

次に、ニュースおよびうわさを伝えた相手を図3に示す。ニュースでは「親」および「実社会の友人」がそれぞれ40%を占めており、うわさでは「親」が34%、「実社会の友人」

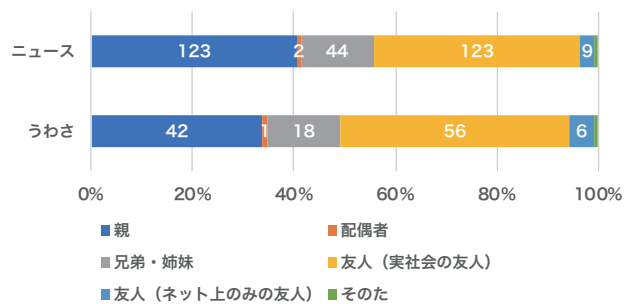


図3: 【ニュース/うわさ】を伝えた相手

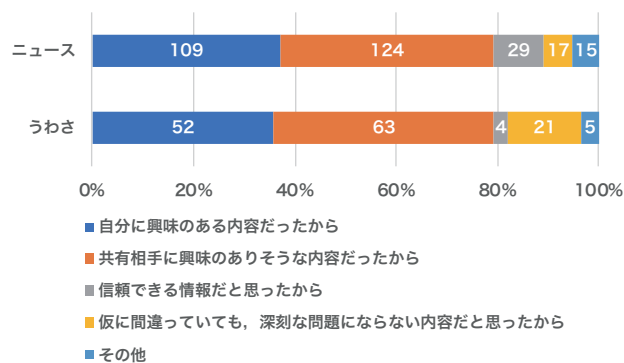


図4: 【ニュース/うわさ】を伝えた理由

が46%となった。この結果から、アンケート回答者の多くは、得た情報を身の回りの親しい間柄の人々に伝えていることが明らかになった。つまり、Twitterを利用していない人であっても、Twitter以外の伝達経路を介して、Twitterで発信された流言を受け取る可能性がある。このことから、特定のSNSで発信された流言に関しても、当該SNSの利用者だけでなく、非利用者に対して、注意喚起を行う必要があると考えられる。

また、ニュースおよびうわさを伝えた理由を図4に示す。どちらも、「自分に興味のある内容だったから」と「共有相手に興味のある内容だったから」がそれぞれ40%前後を占めている。これは、Wangらの情報共有に関する聞き取り調査の結果[8]と同様である。Wangらの調査では、この結果を受けて、相手の役に立ちたいという社会的欲求から情報を伝えるのだと結論づけている。善意の情報伝達であっても、情報の真偽に目が向けられないと、流言を拡散してしまうことがある。また、うわさについて、「仮に間違っていたとしても、深刻な問題にならない内容だと思った」が14%あり、これはニュースよりも多い。これより、誤情報である可能性に気づきながらも、自己の判断で伝えることもあると考えられる。

これらを踏まえて本論文では、日常的に情報への危機意識を高める仕組みを構築することを目的とし、流言情報の注意喚起と、手軽な流言情報の確認が行えるシステムの提案を行う。

¹ <https://line.me>

² <https://www.facebook.com>

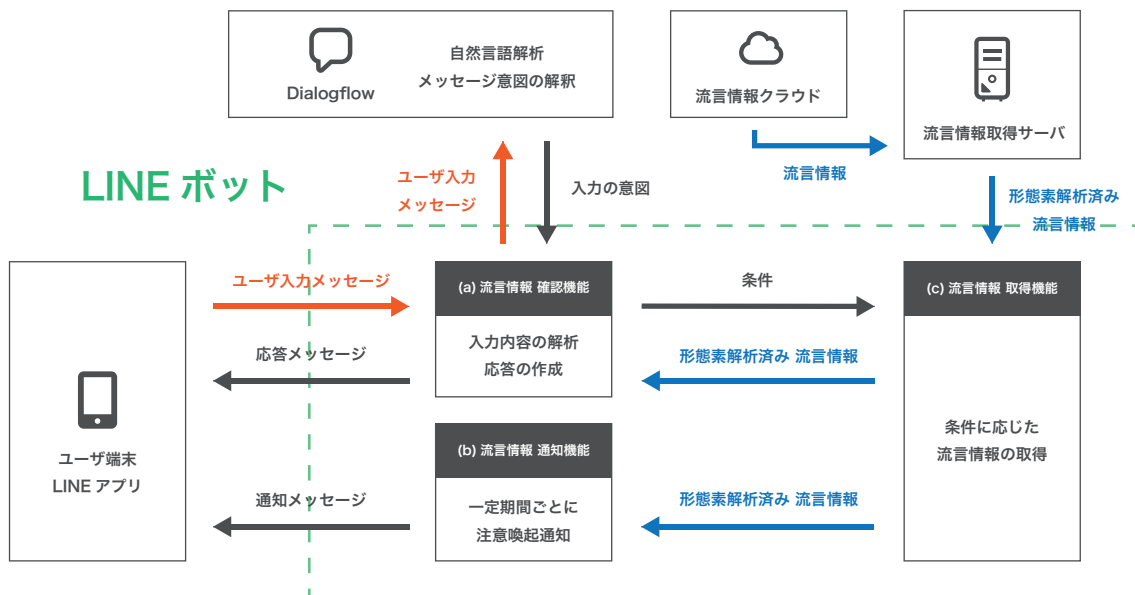


図 5: システム構成図

4. 提案システム「ちるも」

4.1 システム概要

本稿で提案するシステム「ちるも⁴⁾」は、流言の注意喚起を行うチャットボットである。3章のアンケート結果から、日常的な利用者の多い SNS である LINE 上で動作するボットとして実装することとした。ユーザは、LINE 上で本システムを「友だち追加」することで、システムの利用を始められる。LINE の動作する端末であれば、PC、スマートフォン、タブレットなどデバイスを問わず利用することが可能である。

ユーザはシステムと自然言語を用いて対話することができる。例えば、人間に対して問いかけるように、「最近出回っているデマはある？」とメッセージを送信すると、それに対してシステムが応答する。また、ユーザの問いかけがなくとも、システムからプッシュ通知²⁾でメッセージを送ることも可能である。

4.2 設計方針

2章でも述べたとおり、流言の注意喚起は幅広い属性の人々に行う必要がある。システムはより多くの人が気軽かつ日常的に利用可能であることが望ましい。システムを継続して利用してもらうためには、まず、流言の確認に関するユーザの負担を少なくすることが不可欠である。そこで、国内において多くの人々が利用する LINE 上で、自然な言葉で話しかけて流言を確認することができるチャットボットを構築することで、ユーザに流言への気づきを与え、ユーザの情報への危機意識向上を目指す。

(1) 日常的な利用を可能にする

情報の真偽確認を行うことのできる Web サイトは、

Factcheck.org³⁾や PolitiFact⁴⁾など、複数存在する。しかし、Wang らが指摘するように、ファクトチェックを行う人は少数であり [8]、これら既存の Web サイトのような真偽確認を行う仕組みは、日常的に活用されていないと考えられる。そこで、ユーザが日頃の生活のなかで自然にシステムを利用できるように、日常的な利用の多い LINE 上で動作するシステムを構築する。

(2) 流言の確認を容易にする

提案システム「ちるも」はチャットボットとして構築する。自然言語処理技術を利用し、ユーザの入力内容と関連する流言の抽出と提示を行う。これにより、ユーザは、Twitter などで見つけた情報をそのままシステムに入力するだけで、容易に流言の確認ができる。

(3) 流言への気づきを与える

従来研究で、流言を検出する手法が提案されている [7] が、検出精度が 100% にならない限り、情報を信頼するかどうかの最終的な判断は、ユーザ自ら行う必要がある。そこで本研究では、提示する情報を、流言の“可能性がある”ものとし、ユーザに「閲覧している情報が流言かもしれない」という気づきを与えるようにする。

4.3 システム構成

システム構成図を図 5 に示す。ユーザからの問いかけがあった場合、(a) 流言情報確認機能により、入力内容の解析および、応答メッセージの作成を行う。この際、ユーザが入力したメッセージを Google 社⁵⁾が提供する自然言語処理プラットフォームである、Dialogflow⁶⁾によって解析する。

³⁾ <http://factcheck.org>

⁴⁾ <https://www.politifact.com/>

⁵⁾ <http://www.google.co.jp/>

⁶⁾ <https://cloud.google.com/dialogflow>

¹⁾ Check Rumors (チェックルーモア) から「ちるも」と命名した。

²⁾ アプリケーションを起動していなくても、システムが自動的に情報を通知する仕組み。



図 6: システム動作画面例

これによって、ユーザがどのような情報を求めているかといった情報を、メッセージから「意図」として抽出する。この「意図」およびメッセージに含まれるキーワードを用いて、(c) 流言情報取得機能で、条件に一致、もしくは類似する流言情報を取得し、ユーザに応答メッセージとして返す。また、(b) 流言情報確認機能では、一定期間ごとに(c) 流言情報取得機能で取得した流言情報を、ユーザに通知し、注意喚起する。これらの機能により、ユーザに流言への気づきを与え、情報への危機意識を高められるように試みる。

それぞれの機能で利用する流言情報は、2012 年から運用している流言情報クラウド [9] と連携して取得している。流言情報取得サーバでは、取得した流言情報に形態素解析を行い、形態素解析結果を付与した流言情報を (c) 流言情報取得機能に提供している。

システムの動作画面例を図 6 に示す。(a) のように、システムはユーザのメッセージ内容に類似する流言情報を応答することができる。システムが返すメッセージには、ユーザのメッセージと類似する流言が、類似度の高い順に最大 3 件含まれる。加えて、(b) に示すように、Twitter 上のツイートの URL を送信することで、そのツイートの内容に類似する流言情報を応答することができる。この機能により、ユーザは Twitter 閲覧時も、気になるツイートを直接本システムに送信し、類似する流言を確かめることが可能である。また (c) に、流言の注意喚起通知の様子を示す。(c) では、メッセージ送信時点で確認されている、最新の流言を 5 件提示している。また (d) には、ショートカットボタンの例を示す。これは、ボタンを押すことで、それぞれ対応する機能を利用することができるものである。「流言情報クラウド」ボタンを押すと、流言情報クラウドの Web ページ¹⁾にアクセスできる。「最新の流言」ボタンを押すと、最新の流言が提示される。「ヘルプ」ボタンを押すと、どのような問

いかけに対応しているかを提示するようになっている。

5. おわりに

本研究では、情報への危機意識を高めることを目的とし、流言の注意喚起を行うシステムを提案した。本稿では、情報共有に関するアンケートを実施し、流言の注意喚起を行うチャットボットを構築した。今後は、提案システムが目的を満たすかを検証するために実験を行い、その結果について考察する。

謝辞

本研究の一部は、JSPS 科研費 19H04221 の助成による。

参考文献

- [1] 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 (2019) .
- [2] 萩上チキ：検証 東日本大震災の流言・デマ，光文社新書 (2011) .
- [3] 「Twitter、新型コロナ関連で削除対象とする偽情報ツイート範囲を拡大」ITmedia NEWS, 2020 年 03 月 19 日 (最終閲覧日：2020 年 7 月 11 日) .
- [4] 「平成 30 年度情報通信メディアの利用時間と情報行動に関する調査報告書概要」総務省 情報通信政策研究所, 2019 年 09 月 13 日 (最終閲覧日：2020 年 7 月 16 日) .
- [5] 「高齢者は朝のドラッグストアへ本当に「殺到」したか 購買データで解明」ITmedia ビジネス ONLiNE, 2020 年 05 月 07 日 (最終閲覧日：2020 年 7 月 16 日) .
- [6] 柿本大輔, 宮部真衣, 荒牧英治, 吉野孝: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) .
- [7] 鳥海不二夫, 篠田孝祐, 兼山元太:ソーシャルメディアを用いたデマ判定システムの判定精度評価; 情報処理学会, デジタルプラクティス, Vol3, No.3, pp.201–208(2012).
- [8] Luping Wang,Susan R. Fussell : More Than a Click: Exploring College Students' Decision-Making Processes in Online News Sharing.Proceedings of the ACM on Human-Computer Interaction,Vol.4, No.9, pp.1–20(2020).
- [9] 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).

¹⁾<http://mednlp.jp/miyabe/rumorCloud/rumorlist.cgi>