

デマの訂正が社会的混乱に与える影響の分析

Analysis of the Impact of Correcting Misinformation on Social Disruption

飯塚隆介^{1*} 鳥海不二夫¹ 西口真央¹
高野雅典² 吉田光男³

¹ 東京大学

¹ The University of Tokyo

² 株式会社サイバーエージェント

² CyberAgent, Inc

³ 豊橋技術科学大学

³ Toyohashi University of Technology

Abstract: The spread of misinformation can cause social confusion. The authenticity of information on a social networking service (SNS) is unknown, and false information can be easily spread. Consequently, many studies have been conducted on methods to control the spread of misinformation on SNS. However, few studies have examined the impact of the spread of misinformation and its corrections on society. This study models the impact of the reduction of misinformation and the diffusion of corrective information on social disruption, and it identifies the features of this impact. In this study, we analyzed misinformation regarding the shortage of toilet paper during the 2020 COVID-19 epidemic, its corrections, and the excessive purchasing caused by this information. First, we analyze the amount of misinformation and corrective information spread on SNS, and we create a regression model to estimate the real-world impact of misinformation and its correction. This model is used to analyze the change in real-world impact corresponding to the change in the diffusion of misinformation and corrective information. Our analysis shows that the corrective information was spread to a much greater extent than the misinformation. In addition, our model reveals that the corrective information was what caused the excessive purchasing behavior. As a result of our further analysis, we found that the amount of diffusion of corrective information required to minimize the impact on the real world depends on the amount of the diffusion of misinformation.

1 はじめに

現在, Twitter, Facebook などの SNS は, 情報収集やコミュニケーションなど様々な用途で多くの人々に用いられている。SNS での情報収集は, 手軽に多くの情報を収集できるという点で便利だが, 情報の正確性に問題がある。また, デマや誤った情報が拡散することで, 時には社会的な混乱を招き, 重大な事態を引き起こす可能性もある。

2020 年に発生した COVID-19 のパンデミックの期間中も様々な情報が SNS 上に拡散されたが, その中には多数のデマが混じっていた。実際に出回った例をあげる

と, コロナウイルスは熱に弱く, 26 27 度のお湯をたくさん飲むと感染予防になるという内容のデマや, WHO が感染者の隔離は不要であると意見を変えたという内容のデマなどがある。また, 2020 年 2 月の後半から 3 月の前半にかけて, コロナウイルスの流行に伴い, トイレtpペーパーが不足するだろうという内容の情報, またそれを否定する内容の情報が拡散した。結果としてトイレtpペーパーが手に入りにくい状況が起きた。

SNS 上での誤情報の拡散については, 多数の研究がなされている。誤情報の拡散の程度については, 2016 年のアメリカの大統領選挙では, アメリカの成人は, 選挙前の数ヶ月間に 1 つ, または複数のフェイクニュースをみた可能性があることなどが知られている [1]。デマの拡散を減少させる方法については, シミュレーション

*連絡先: 東京大学
東京都文京区本郷 7-3-1
iizuka@torilab.net

を用いた研究がなされており、多くのフォロワーを持つユーザーに訂正情報を拡散してもらう戦略が効果的であることなどが示されている [2]. 誤った情報を訂正する情報の効果については、どのような分野の誤情報が訂正しやすいか、どのような人に訂正情報が効果的か、どのような内容の訂正情報が効果的かなどが調べられている [3][4]. その一方で、対象となる人の性質によっては、訂正する情報は逆効果になることが知られている [5]. これらの研究は、SNS 上でのデマとその訂正情報の拡散を対象としている。しかし、真に解決すべきは、それによる社会の混乱を防ぐことである。これらの研究では、誤情報や訂正情報の社会的混乱に対する影響については考慮していない。

そこで、本研究では、SNS 上でのデマやそれに関わる情報の拡散と、それによって生じる実社会における混乱との関係を明らかにすることを目的とする。本研究では、トイレットペーパー不足に関わる情報と、それによって生じたトイレットペーパーの過剰な購買を対象として分析を行った。分析の手順としては、まず、Twitter のデータを元にトイレットペーパー不足に関わるデマや訂正がどれだけ拡散したかを分析する。次にスーパーの売上データを用いて、情報の拡散と購買量の関係を回帰分析によりモデル化した。そして、購買量を小さくするためには、訂正情報がどれだけ拡散されるべきかについて、分析を行った。

本研究の貢献は以下の通りである。

- 誤情報とその訂正情報が実際に社会に与えた影響を推定し、必ずしも訂正情報が社会的混乱を低減するわけではないことを示した。
- デマとその訂正が実社会における過度な購買行動に与えた影響を推定するモデルを構築した。
- 過度な購買行動の最小化に必要な情報拡散行動の指針を示した。
- デマの拡散度合いによって、どれだけ訂正情報が拡散されるべきかを示した。

本論文は以下のように構成される。まず 2 章で関連研究について述べる。3 章では、使用データの詳細について述べる。4 章では、デマが売上高に与えた影響の分析について述べる。5 章では、訂正情報の拡散縮小策とその影響の分析について述べ、6 章では、デマ拡散率に基づく最適訂正拡散率の推定について述べる。最後に 7 章で結論を述べる。

2 関連研究

2.1 誤情報の拡散

ここでは、デマの拡散についての研究を紹介する。Allcott.H ら [1] は、2016 年のアメリカ大統領選挙時のフェイクニュースについての分析を行った。結果として、米国の成人は、選挙前の数ヶ月間に 1 つまたは複数のニュース記事を見た可能性があることが示唆された。

Chengcheng Shao ら [6] も同じ事象を対象に分析を行った。結果として、ソーシャルボットがデマの拡散に大きな役割を果たしたことを示し、デマを積極的に拡散するアカウントはボットである確率が高いことを示した。

Karishma Sharma ら [7] は COVID-19 の誤情報の拡散についての分析を行い、SNS 上で COVID-19 に関わる誤情報が国境を超えて拡散されたことなどを示した。

また、Soroush Vosoughi ら [8] は、Twitter 上での正しい情報、一部に誤りを含む情報、誤情報の拡散について分析をした。結果として、誤情報は他の情報に比べて、より速く広く拡散することがわかった。また、政治的な誤情報は他のカテゴリーの誤情報に比べて、より広く拡散されることがわかった。

上記のように、デマの拡散は様々な分野で生じていて、その背景には bot の存在がある場合もあることが示されている。また、誤った情報が他と比べて拡散されやすいことも示されている。

2.2 誤情報の拡散防止

ここでは、デマの拡散防止についての研究を紹介する。Okada.Y ら [2] は、SIR モデルをベースとした、誤情報の拡散モデルを作成し、誤情報の拡散を止める戦略を考案した。結果として、多くのフォロワーを持つユーザーに訂正情報を拡散してもらう戦略が効果的であることが分かった。

J. Kim ら [9] は、SNS におけるフェイクニュースや誤報の拡散を検知・防止するために、群衆を活用したオンラインアルゴリズムを提案した。そして、Twitter と Webio のデータを用いて、提案手法が有効である可能性を示した。

また、誤情報の拡散を防ぐためには、誤情報を検出することが求められる。デマの検出のためには、様々な特徴が用いられる。Kai shu ら [10] によると、特徴は、ニュースの内容から得られる特徴と、社会的文脈から得られる特徴がある。ニュースの内容を元にした特徴量は、言語的な特徴と視覚的な特徴に分けられる。言語的な特徴は単語や文章などで、Aditi Gupta[11] らはこのような特徴を用いて、超党的なニュースやデマを区別することを試みた。このような特徴を用いた研

究としては、他にも Sadia Afroz ら [12] による研究などがある。

視覚的な特徴とは、写真やビデオなどである。Martin Potthast ら [13] はこのような特徴を用いて、ハリケーン・サンディの偽画像と実際の画像の分類を行い、97%の精度で分類ができることを示した。このような特徴を用いた研究としては、他にも Zhiwei Jin らによる研究 [14] などがある。

社会的文脈を基にした特徴量は、ユーザーの特徴（年齢やツイート数など）[15]、投稿の特徴（デマに対する反応や意見など）[16]、ネットワークの特徴（フォローフォロワー関係など）[17] で、それぞれ多くの研究がなされている。

また、上記のように、デマの拡散には bot が関与することがあるため、bot の検出についての研究も多く行われている。bot の検出については、教師あり学習を用いた手法で良い結果が出るということが知られている [18][19]。また、教師なし学習を用いた手法としては、Nikan Chavoshi らによる研究がある [20]。この研究では、動的時間伸縮法を用いて、アカウント間の相関を基に bot 検出を行い、教師あり学習に勝る結果を残している。

上記のように拡散の防止を試みる研究は多く行われており、訂正情報の拡散によって拡散を止めるアプローチや、誤情報の検知といったアプローチがとられている。しかし訂正情報が本当にデマの拡散を止めることになるかは考慮されていない。

2.3 誤情報の訂正情報の効果

ここでは、デマの訂正情報の効果についての研究を紹介する。B. Nyhan ら [5] は、17才以下の子供を持つ親に対して、訂正情報がワクチン接種を後押しするかについての調査を行った。結果として、調査で用いたどんな訂正情報も、ワクチン接種を後押しすることはなかった。また、ワクチン接種に最も消極的だった被験者の子供は、よりワクチン接種に対して消極的になることが示された。

Carey.J.M. ら [21] は、ブラジルでのジカ熱・黄熱についてのデマに対する訂正情報が、誤った理解を解くのに効果的かどうかについて分析を行った。ジカ熱に対しての訂正情報は、ジカ熱への誤認識を解くことはできなかったが、よりよく知られている黄熱については、効果的であることが分かった。

このように、訂正情報の効果については、様々な研究が行われているが、対象となるデマの種類などの様々な要因によって効果的かどうかが変わってくる。そのため、以下のようなメタ分析が行われている。M.-p. S. Chan ら [3] は、訂正情報が、誤った情報を信じている

人の誤解を解くのに効果的となる条件について分析した。結果として、デマに対して反論する可能性が高い人々たちにとっては効果があり、誤情報を信じる理由を作ってしまった人には効果が薄れることなどが分かった。

N. Walter ら [4] は、訂正情報が誤った認識を修正する効果と、その効果に影響する要素についてのメタ分析を行った。訂正情報の効果は中程度で、健康についての誤情報よりも政治やマーケティングについての情報の方が訂正の方が困難であることが分かった。また、気候変動の否定などの現実的な誤情報の方が、飛行機の墜落などででっち上げの誤情報より訂正が困難であることが分かった。

誤情報の訂正情報が、誤認識を解くことが出来るかについては多く研究がされている。そして、対象となる誤情報の種類や、誤情報を信じている人の性質など様々な要因によって、効果が異なることが知られている。一方で、私たちは誤認識を解く効果ではなく、社会的混乱に与える影響を調べる。

3 データ概要

3.1 Twitter データ

本研究で用いた Twitter のデータは大きく分けて2つあり、ツイートデータとフォロー関係のデータである。ツイートデータは、トイレットペーパーの不足に関わるツイートを対象とした。データは関連するキーワードを元に検索、取得された。キーワードは、「トイレットペーパー」、「ティッシュ」、「生理用品」、「国民生活安定緊急措置法」の4つである。そのデータをツイートを内容に応じて、3つのカテゴリーに分類した。分類は、内容を目で見ても行った。3つのカテゴリーは、以下の通りである。

- デマやトイレットペーパーの買い占めを煽る内容（デマツイート）
- デマやトイレットペーパーの不足を否定する内容（訂正ツイート）
- トイレットペーパーが実際に売り切れていることを示す内容（売切ツイート）

データの収集期間は2月21日～3月10日である。収集期間に取得された総ツイート数は4,476,754で、その中に占める総RT数は2,945,782である。内容毎のデータの概要を表1に示す。デマツイートのみ、RT数上位6000位までを対象とした理由としては、多く拡散されたデマツイートは少ない状況で、より多くのデマツイートを集めたかったためである。

デマ情報及び訂正情報を見た可能性がある人数を推定するために、関連ツイートをリツイートしたアカウント

表 1: データ概要

内容	条件	ツイート数 (重複なし)	RTした人数
訂正ツイート	RT 数上位 1400 位以内	229	332,881
売切ツイート	RT 数上位 1400 位以内	42	18,264
誤情報ツイート	RT 数上位 6000 位以上	8	712

のフォロワーを収集した。フォロワーは、TwitterAPIを用いて、取得時点でのフォローフォロワーの取得をした。プライベートアカウントなどの情報は取れていないことに注意が必要である。その結果、97,430,525アカウントを収集した。

3.2 トイレットペーパーの売上高指数

トイレットペーパーがどれだけ売れたかを示す指標として、トイレットペーパーの売上高指数を用いた。このデータは、株式会社ナウキャストに提供してもらったデータである。トイレットペーパーの売上高指数とは、前年と比較した、トイレットペーパーの売上高の変化の程度を表す指標で、式 1 で表される。S が売上高指数、 $Sales_t$ がある日 t のトイレットペーパーの売上高、 $dt=364$ 日である。この指標は、全国のスーパー 1200 店舗の POS データから作成される。

$$S = (Sales_t / Sales_{t-dt}) - 1 \quad (1)$$

データ分析期間の売り上げ指数の変化を図 1 に示す。この期間を通じて、例年よりもトイレットペーパーの売上が大きく、2月28日にピークがあることが見て取れる。

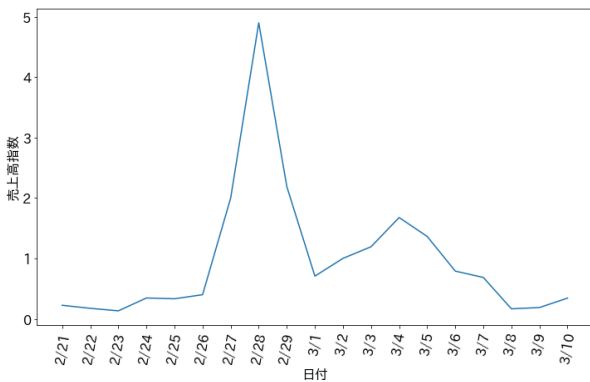


図 1: 売上高指数の推移

4 デマが売り上げ高に与えた影響の分析

4.1 閲覧可能者数

トイレットペーパー不足についてのツイートがどれだけ拡散したかを知るために、以下の 7 タイプの閲覧可能者数を調べる。

- 訂正ツイートのみ見た閲覧可能者数 (x_1)
- 誤情報ツイートのみ見た閲覧可能者数 (x_2)
- 売切ツイートのみ見た閲覧可能者数 (x_3)
- 訂正ツイートと誤情報ツイート見た閲覧可能者数 (x_4)
- 訂正ツイートと売切ツイート見た閲覧可能者数 (x_5)
- 誤情報ツイートと売切ツイート見た閲覧可能者数 (x_6)
- 全て見たユーザー数 (x_7)

閲覧可能者数は日単位で取得する。閲覧可能者数の定義は、1日の間にツイートもしくはリツイートしたユーザーとそのフォロワー数の合計である。

図 2 と表 2 は、上記の 7 タイプのユーザー数を調べた結果である。訂正ツイートのみを見た閲覧可能者数が一番多く、次いで訂正ツイートと売切ツイートを見た閲覧可能者数、売り切れツイートのみ見た閲覧可能者数が多い。デマのみ人数は訂正のみ見た閲覧可能者数と比べて非常に少なく、訂正のみ見た閲覧可能者数の 0.28%ほどである。

4.2 売上高指数の重回帰分析

4.2.1 方法概要

まずは、日別売上高指数推移を、重回帰分析によってモデル化する。説明変数には、上述した各タイプの閲覧可能者数を入力とした主成分を利用する。主成分分析をする理由としては、多重共線性を避けるため

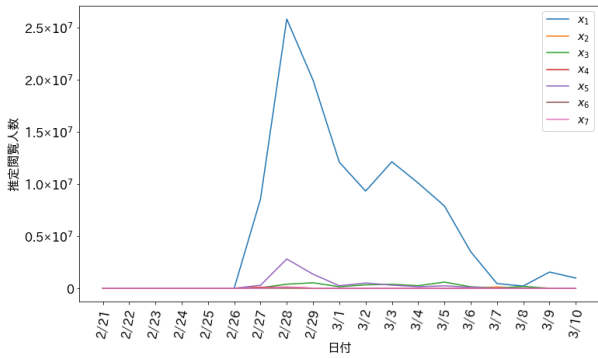


図 2: 各タイプの閲覧可能者数推移

表 2: 各タイプの述べ閲覧可能者数

	推定閲覧数
訂正ツイートのみ見た	112,440,832
誤情報ツイートのみ見た	311,345
売切ツイートのみ見た	2,974,369
訂正ツイートと誤情報ツイート見た	251,100
訂正ツイートと売切ツイート見た	5,967,030
誤情報ツイートと売切ツイート見た	4,157
全て見た	124,953

ある。表 3 は、内容毎の見た人数間の相関係数である。また、寄与率の低い第 5 主成分以降を用いず、第 4 主成分までを用いる。

表 3: 内容毎の見た人数間の相関

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.0	0.091	0.76	0.58	0.87	0.58	0.68
x_2	0.091	1.0	-0.15	0.68	0.18	0.45	0.29
x_3	0.76	-0.15	1.0	0.16	0.57	0.23	0.31
x_4	0.58	0.68	0.16	1.0	0.74	0.86	0.84
x_5	0.87	0.18	0.57	0.74	1.0	0.83	0.91
x_6	0.58	0.45	0.23	0.86	0.83	1.0	0.97
x_7	0.68	0.29	0.31	0.84	0.91	0.97	1.0

4.2.2 主成分分析

表 4 はそれぞれの主成分の固有ベクトルを示した表である。第 1 主成分の固有ベクトルは、 x_1 成分が大きく、第 2 主成分の固有ベクトルは、 x_5 成分が大きく、また、第 3 主成分の固有ベクトルは x_3 成分が大きく、第 4 主成分の固有ベクトルは x_2 成分が大きく。

表 5 はそれぞれの主成分の寄与率を示した表である。寄与率は、式 2 で表され、 C が寄与率で、 p が主成分

数、 l_i が第 i 主成分の固有値である。

$$C = \frac{l_i}{\sum_{i=1}^p l_i} \quad (2)$$

寄与率は、全情報量のうち、該当する主成分が占める情報量の割合を示す。第 1 主成分の寄与率が 9.97×10^{-1} と非常に大きく、第 1 主成分で元々のデータの大部分を説明できることがわかる。また、第 1 主成分の固有ベクトルは、 x_1 成分が大きいため、 x_1 が元々のデータの大部分を説明できることがわかる。

4.2.3 重回帰分析

表 6 は、重回帰分析の結果で、係数と p 値を示している。 a_1, a_2, a_3, a_4 は回帰係数を示し、 b は定数項を示す。第 3 主成分のみ p 値が 0.05 以上である。主成分分析をした後に、重回帰分析をしているため、係数の説明は表 7 で行い、 t 値の説明は、表 8 で行う。

表 7 は、回帰係数と固有ベクトルの内積を示した表である。回帰係数と固有ベクトルの内積は、各変数の重要度を示し、今回の場合は各人数一人当たりの売上高指数に与える影響の大きさを示す。一人当たりの影響度はデマのみ見た人数 x_2 が一番大きいことがわかる。すなわち、デマのみを見た人が増加すればトイレトペーパーの購入が加速し、品不足が生じる可能性が高いことが示された。また、次に影響力があるのは x_4 すなわち、デマと訂正情報の双方を見たアカウントであることが分かった。これは、これらのアカウントはあくまでも双方の情報を見た可能性があるアカウントであるため、実際にはデマのみ見たアカウントが多数含まれていた可能性がある。

図 3 は、回帰した売上高指数と実際の売上高指数の推移を示した図である。小さなずれはあるが、精度よく回帰できており、決定係数は 0.939、 F 値は 53.49 となっている。

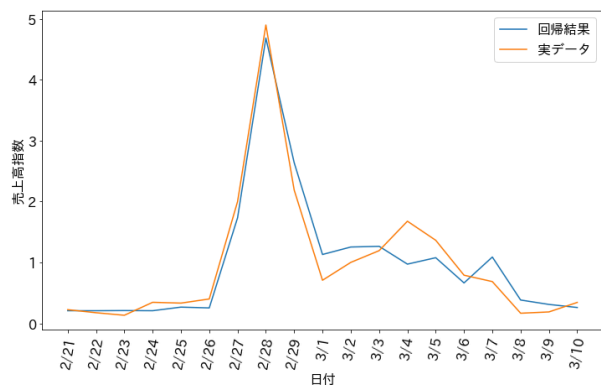


図 3: 回帰した売上高指数

表 4: 主成分の固有ベクトル

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
第 1 主成分	0.99	0.00	0.002	0.00	0.008	0.00	0.00
第 2 主成分	-0.08	0.02	-0.13	0.04	0.99	0.00	0.04
第 3 主成分	-0.03	-0.08	0.98	-0.08	0.13	0.00	-0.02
第 4 主成分	0.00	0.89	0.12	0.44	-0.03	0.00	0.05
第 5 主成分	0.00	0.41	-0.04	-0.76	0.04	-0.01	-0.51
第 6 主成分	0.00	-0.18	0.00	0.48	0.02	-0.03	-0.86
第 7 主成分	0.00	0.00	0.00	0.00	0.00	0.99	-0.03

表 5: 各主成分の寄与率

	寄与率
第 1 主成分	9.97×10^{-1}
第 2 主成分	1.93×10^{-3}
第 3 主成分	2.55×10^{-4}
第 4 主成分	2.11×10^{-5}
第 5 主成分	1.87×10^{-6}
第 6 主成分	3.70×10^{-7}
第 7 主成分	8.44×10^{-11}

表 8 は、表 2 で示した推定閲覧数と一人当たりの影響度をかけ合わせた結果である。 s_{x_i} は x_1 から x_7 それぞれの、推定閲覧数と一人当たりの影響度をかけ合わせた値を示す。この結果は、どのタイプのアカウントがどれだけ売上に影響したのかを示し、訂正ツイートのみを見たタイプのアカウントが最も売上に影響を与えていると推定されたことが分かる。またこの結果は、表 5 と表 6 とともに矛盾していない。

5 情報の拡散縮小策とその影響

5.1 訂正ユーザーの制御手法

デマとその訂正は、必ずしも訂正が浸透すれば社会的混乱を抑えられるわけではない。本章では、第 4 章で求めた、トイレットペーパーの売上を推定するモデルを用いて、社会的混乱を低減するための方針について検討を行う。

まずは、実データをもとに訂正ツイートを低減させるシミュレーションを行う。まず、実データで訂正ツイートを RT をしていたユーザーの何割かを RT をしないこととする。またその影響で、実データで訂正ツイートを RT していたユーザーのうち、RT した訂正ツイートを見る機会が消失したユーザーも RT をしないこととした。

5.2 訂正ユーザー縮小実験

5.2.1 実験条件

ここでは、実際のデータを用いて、訂正ツイートを RT したユーザーの数を減少させた場合の売上高指数の変化を調べる。この実験では、訂正ツイートを見たユーザーのうち、何割が訂正ツイートをリツイートするか（以下、訂正 RT 率）がパラメータとして存在する。本実験では、訂正 RT 率は 0.79%、0.63%、0.47%、0.32%、0.16%、0.0% の 6 通りに設定した。訂正 RT 率 0.79%、0.63%、0.47%、0.32%、0.16%、0.0% の場合の RT するユーザー数はそれぞれ、実データで訂正 RT をしたユーザー数の 100%、80%、60%、40%、20%、0% のユーザー数である。

5.2.2 売上予測

図 4 は、訂正ユーザー縮小実験の結果である。訂正 RT 率が小さいほど、売上高指数が小さく推定されており、訂正ツイートを RT する人が誰もいない場合に最も売上高指数が小さく推定されたことが分かる。これは、訂正ツイートをみた人数が多いほど、売上高指数が大きく推定されるためである。

しかし、現実的に考えると、訂正ツイートが少なくなった場合、誤情報ツイートの拡散が増大すると考えられる。以降では、これを考慮した実験を行う。

5.3 デマ不接触ユーザーの情報拡散防止

訂正ユーザー縮小実験では、ランダムに訂正ツイートをリツイートするユーザーを決めていた。しかしながら、ランダムに訂正ツイートをリツイートするユーザーが増減させるというのは施策として現実的ではない。そこで、リツイートの指針として、「訂正ツイートをリツイートする前に、デマツイートを見ていたユーザーのみが訂正ツイートをリツイートする」という指針に基づいて、売上高指数がどのようになるかを推測する。

表 6: 回帰係数と p 値と t 値

	a_1	a_2	a_3	a_4	b
係数	1.35×10^{-7}	10.13×10^{-7}	-2.494×10^{-7}	69.53×10^{-7}	0.9919
p 値	0.000	0.001	0.695	0.006	0.000
t 値	13.55	4.470	-0.400	3.211	13.47

表 7: 一人当たりの売上高指数に与える影響

	重要度
x_1	5.35×10^{-8}
x_2	624.00×10^{-8}
x_3	44.80×10^{-8}
x_4	309.00×10^{-8}
x_5	79.20×10^{-8}
x_6	2.88×10^{-8}
x_7	43.10×10^{-8}

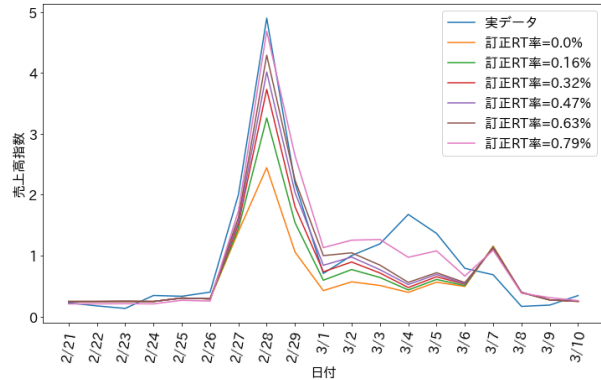


図 4: 訂正ユーザ縮小実験結果

表 8: タイプごとの影響度

	影響度
訂正ツイートのみ見た	6.0236
誤情報ツイートのみ見た	1.9412
売切ツイートのみ見た	1.3339
訂正ツイートと誤情報ツイート見た	0.7763
訂正ツイートと売切ツイート見た	4.727
誤情報ツイートと売切ツイート見た	0.0001
全て見た	0.0538

訂正 RT 率での売上高指数を比較している。提案指針では、訂正 RT 率 $\geq 0.16\%$ の場合と比べて、売上が小さく抑えられることが分かる。これは訂正ツイートをリツイートするユーザーが少ないことが理由であり、前項と矛盾しない。また元々の売上高指数 (18.85) と比較して、提案手法は 40.4% ほど売上高指数を小さく抑えられる事が分かる。

5.3.1 実験条件

この実験では、誤情報ツイートのみを見たユーザーの何割が誤情報ツイートを RT するか (デマ RT 率) というパラメータが存在する。実データで誤情報ツイートのみを見た人のうち、誤情報ツイートを RT したユーザーの割合が 0.186% であったため、本実験でのデマ RT 率は 0.186% とした。訂正 RT 率は 0.79%, 0.63%, 0.47%, 0.32%, 0.16%, 0.0% の 6 通りに設定した。訂正ツイートをリツイートするユーザーは人で元々訂正ツイートをリツイートしていたユーザーの 0.029% である。6 通りの訂正 RT 率と、提案指針の計 7 通りの条件に対して、10 回ずつ実験を行った。

表 9: デマ不接触ユーザの情報拡散防止の結果

訂正 RT 率	売上高指数の総和
0.029% (提案手法)	11.23
0.0%	9.192
0.16%	12.50
0.32%	13.34
0.47%	14.11
0.63%	14.69
0.79%	15.93

5.3.2 売上予測

表 9 が、デマ不接触ユーザの情報拡散防止の実験結果である。売上高指数の合計とは、対象期間中の売上高指数の総和を示す。デマ RT 率を 0.186% の時の、各

6 デマ拡散率に基づく最適訂正拡散率の推定

6.1 デマ訂正の拡散がデマ拡散量と売りに上げに与える影響

図5は、デマRT率が0%の場合の、訂正RT率と、デマのみ見た人数、売上高指数の関係を示した図である。右側の縦軸は、対象期間中の売上高指数の総和を示す。訂正RT率が上昇すると、デマのみ見た人数は減少するものの、売上高指数は増加することが分かる。このことから、デマの拡散度合いによっては、訂正情報を多く拡散することが、社会的混乱を抑制するのに逆効果であることが示された。そのため、以下ではデマの拡散度合いに対する、適切な訂正情報の拡散度合いを調べる。

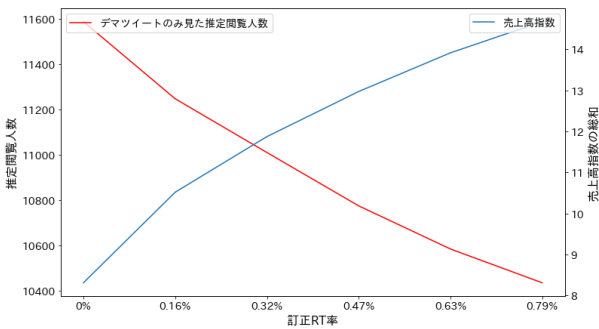


図5: デマ訂正の影響

6.2 実験条件

デマやその訂正ツイートは、その信ぴょう性の有無に基づいて拡散率が異なると考えられる。あまりにも荒唐無稽であったり重要性が認められないデマであればほとんど拡散されることはなく、また、信ぴょう性が極めて高い場合や重要な情報であるように見えるデマは拡散されやすいと考えられる。また、訂正情報にも同じことが言える。デマの拡散力によって適切な訂正情報の拡散率は変化すると考えると、デマの拡散率から訂正情報をどのように拡散すればよいかは変化するだろう。そこで、本章ではデマの拡散率と最適な訂正情報の拡散率の関係を明らかにする。

本実験では、訂正RT率は0.79%、0.63%、0.47%、0.32%、0.16%、0.0%の6通りに設定した。また、デマRT率は0%、0.186%、1.0%、2.0%、3.0%、4.0%、5.0%の7通りに設定した。これらの組み合わせで、計42通りの条件に対して実験を行った。各実験条件対

して、10回ずつ実験を行い、推定された、対象期間中の売上高指数の総和の平均を結果とした。

6.3 デマ訂正が売りに上げに与える影響の分析

図6が、デマ訂正が売りに上げに与える影響の分析の結果である。縦軸は、対象期間中の売上高指数の総和を示す。

デマRT率が小さい場合（デマRT率=0.0%、0.186%）は、訂正RT率が増加すると予想される売上が増加する。また、デマRT率が大きい場合（デマRT率 \geq 3.0%）は、訂正RT率が増加すると売上が減少することが明らかとなった。

デマの拡散が小さい場合、訂正ツイートの拡散がデマの拡散をはるかに上回り、結果として売上への影響が大きくなる。一方で、デマの拡散が大きい場合は、訂正ツイートによって誤情報ツイートの拡散を防止され売上が減少する。

現実の事象としては、デマ率が低い(0.186%)ため、訂正ツイートの拡散は逆効果であった。しかしながら、より信憑性の高くより多くの人が拡散するようなデマ情報については、訂正ツイートを拡散することが有効である。

具体的には、本研究で用いた事例においては、デマを見たユーザの3%以上がRTしていた場合は、訂正ツイートの拡散が効果的だったといえる。

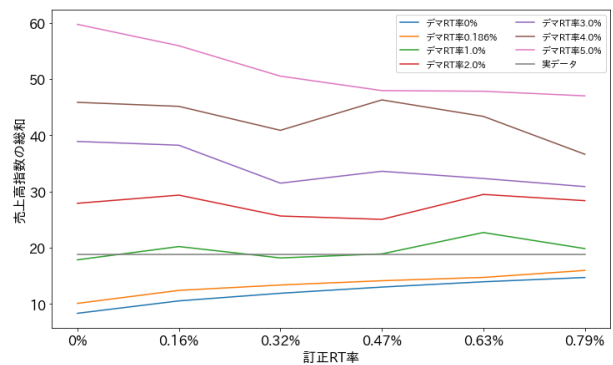


図6: デマ拡散率に基づく最適訂正拡散率の推定結果

以上より、図5で示した通り、訂正RTが増加すればデマの拡散は抑えることが可能となるが、図6で示された通り、訂正が必ずしも社会的混乱、すなわち売りに上げを減少させることに直結していないことが明らかとなった。すなわち、デマのリツイート率、すなわち信憑性に比例して、適切な訂正ツイート量は異なるといえる。

通常、デマに対しては、当該デマをできるだけ訂正することが重要であると認識されている。しかしなが

ら、過度な訂正情報はかえって社会的混乱を引き起こす要因となることが示された。デマ情報が拡散された際には、当該デマおよび訂正情報がそれぞれどの程度社会に影響を与えるのかを見積もり、どのように対応すべきかを慎重に検討する必要がある。

7 結論

本研究では、2020年のCOVID-19パンデミック下で生じた日本のトイレットペーパー不足デマについて、ツイートと売り上げのデータを用いて、デマが社会に与えた影響について分析を行った。

ツイートを内容で分類しデマや訂正情報に接触したユーザの数を算出した。結果として、訂正ツイートのみを見たユーザーが圧倒的に多く、デマのみを見たユーザーの約357倍だった。

各ツイートの閲覧可能者数を説明変数として重回帰分析を行った。その結果、決定係数は0.939、F値は53.49で、売り上げを精度良く予測できることが示された。また、回帰式の係数などから、売上高指数に最も大きい影響を与えたのは、訂正のみ閲覧可能者数が最も大きいことが分かった。

得られたモデルを用いて、訂正RTをするユーザー数が変化した場合の売上高指数を推定した。その結果、訂正ツイートをRTする人数が減少すると売上高指数が小さくなることを示し、訂正ツイートがトイレットペーパーの過剰購入の要因であることを示した。そこで、「デマを見ていないユーザーは訂正ツイートをRTしない」という施策を行った場合売り上げがどのように変化するかを確認した。その結果、施策を行わない場合と比較して売上高指数が40.4%低減できることを示した。このことから本施策が有効である事が示唆された。

最後に、提案モデルを用いて、デマの信ぴょう性、すなわちデマ拡散率（デマRT率）が変化した場合の訂正情報の最適な拡散率（訂正RT率）の推定を行った。その結果、デマの信ぴょう性が低い、すなわちデマの拡散率が小さい場合は訂正ツイートによって過剰購入が促進されることが示された。一方、デマの信ぴょう性が高い、すなわちデマ率が高い場合には、訂正ツイートによって過剰購入が低減させることができることが示された。

従来研究ではデマの訂正がデマの拡散率やデマの接触率にどのような効果をもたらすかを調べるものが多いのに対して、本研究ではデマの訂正が社会的混乱の低減に与える影響を調べた点にある。

今後は、デマの拡散率に基づく適切な訂正戦略を明らかにしていくことが必要である。また、デマが社会にもたらす混乱は過剰購入だけではない。多様なデマ

について同様に社会的影響に与える要因を分析し、訂正情報の最適化に向けた一般的な傾向を明らかにすることも必要である。さらに、デマはソーシャルメディアのみではなく、マスメディアの影響も強いと言われている (Benkler, Yochai, et al. "Mail-In Voter Fraud: Anatomy of a Disinformation Campaign." Available at SSRN (2020).). 今後、ソーシャルメディア以外の影響も考慮に入れたデマ対策手法の確立も今後の課題の一つである。

謝辞

トイレットペーパーの売上高指数データを提供していただいた株式会社ナウキャスト様に感謝申し上げます。

参考文献

- [1] Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of economic perspectives*, Vol. 31, No. 2, pp. 211–36, 2017.
- [2] Yoshiyuki Okada, Keisuke Ikeda, Kosuke Shinoda, Fujio Toriumi, Takeshi Sakaki, Kazuhiro Kazama, Masayuki Numao, Itsuki Noda, and Satoshi Kurihara. Sir-extended information diffusion model of false rumor and its prevention strategy for twitter. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 18, No. 4, pp. 598–607, 2014.
- [3] Man-pui Sally Chan, Christopher R Jones, Kathleen Hall Jamieson, and Dolores Albarraacín. Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological science*, Vol. 28, No. 11, pp. 1531–1546, 2017.
- [4] Nathan Walter and Sheila T Murphy. How to unring the bell: A meta-analytic approach to correction of misinformation. *Communication Monographs*, Vol. 85, No. 3, pp. 423–441, 2018.
- [5] Brendan Nyhan, Jason Reifler, Sean Richey, and Gary L Freed. Effective messages in vaccine promotion: a randomized trial. *Pediatrics*, Vol. 133, No. 4, pp. e835–e842, 2014.
- [6] Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Alessandro Flammini, and Filippo Menczer. The spread of fake news by social bots.

arXiv preprint arXiv:1707.07592, Vol. 96, p. 104, 2017.

- [7] Karishma Sharma, Sungyong Seo, Chuizheng Meng, Sirisha Rambhatla, Aastha Dua, and Yan Liu. Coronavirus on social media: Analyzing misinformation in twitter conversations. *arXiv preprint arXiv:2003.12309*, 2020.
- [8] Soroush Vosoughi, Deb Roy, and Sinan Aral. The spread of true and false news online. *Science*, Vol. 359, No. 6380, pp. 1146–1151, 2018.
- [9] Jooyeon Kim, Behzad Tabibian, Alice Oh, Bernhard Schölkopf, and Manuel Gomez-Rodriguez. Leveraging the crowd to detect and reduce the spread of fake news and misinformation. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, pp. 324–332, 2018.
- [10] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, Vol. 19, No. 1, pp. 22–36, 2017.
- [11] Aditi Gupta, Hemank Lamba, Ponnurangam Kumaraguru, and Anupam Joshi. Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. In *Proceedings of the 22nd international conference on World Wide Web*, pp. 729–736, 2013.
- [12] Sadia Afroz, Michael Brennan, and Rachel Greenstadt. Detecting hoaxes, frauds, and deception in writing style online. In *2012 IEEE Symposium on Security and Privacy*, pp. 461–475. IEEE, 2012.
- [13] Martin Potthast, Johannes Kiesel, Kevin Reinartz, Janek Bevendorff, and Benno Stein. A stylometric inquiry into hyperpartisan and fake news. *arXiv preprint arXiv:1702.05638*, 2017.
- [14] Zhiwei Jin, Juan Cao, Yongdong Zhang, Jianshe Zhou, and Qi Tian. Novel visual and statistical image features for microblogs news verification. *IEEE transactions on multimedia*, Vol. 19, No. 3, pp. 598–608, 2016.
- [15] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. Information credibility on twitter. In *Proceedings of the 20th international conference on World wide web*, pp. 675–684, 2011.
- [16] Zhiwei Jin, Juan Cao, Yongdong Zhang, and Jiebo Luo. News verification by exploiting conflicting social viewpoints in microblogs. In *Thirtieth AAAI conference on artificial intelligence*, 2016.
- [17] Sejeong Kwon, Meeyoung Cha, Kyomin Jung, Wei Chen, and Yajun Wang. Prominent features of rumor propagation in online social media. In *2013 IEEE 13th International Conference on Data Mining*, pp. 1103–1108. IEEE, 2013.
- [18] Clayton Allen Davis, Onur Varol, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. Botornot: A system to evaluate social bots. In *Proceedings of the 25th international conference companion on world wide web*, pp. 273–274, 2016.
- [19] Onur Varol, Emilio Ferrara, Clayton A Davis, Filippo Menczer, and Alessandro Flammini. Online human-bot interactions: Detection, estimation, and characterization. *arXiv preprint arXiv:1703.03107*, 2017.
- [20] Nikan Chavoshi, Hossein Hamooni, and Abdullah Mueen. Debot: Twitter bot detection via warped correlation. In *Icdm*, pp. 817–822, 2016.
- [21] John M Carey, Victoria Chi, DJ Flynn, Brendan Nyhan, and Thomas Zeitzoff. The effects of corrective information about disease epidemics and outbreaks: Evidence from zika and yellow fever in brazil. *Science advances*, Vol. 6, No. 5, p. eaaw7449, 2020.