

令和元年度 修士論文

リツイート経路分析による
リツイート数増加要因の特定

千葉大学大学院融合理工学府
地球環境科学専攻都市環境システムコース
18WM1318 中島圭佑

指導教員：塩田茂雄

令和2年3月 提出

Abstract

In recent years, with the emergence of Social Networking Service (SNS), an enormous amount of diverse and enormous information has spread rapidly through SNS, and this is an era that has a great impact on our lives. SNS is used not only as a communication tool between individuals but also in various situations such as information sharing at the time of disaster, marketing, stock price prediction, and election activities. On the other hand, it has been pointed out that negative aspects such as the spread of hoaxes and the breeding ground for individuals and groups have been pointed out. Therefore, it is required to clarify the characteristics of the information diffusion process in SNS. In this study, we tried to clarify the cause of the increase in the number of retweets by collecting and analyzing tweets by keyword search using TwitterAPI as a representative SNS. Focusing on the network between users, we created a network consisting of followers and followers, and analyzed the way information is diffused by retweets. As a result, we found that among retweet users, there were many users who did not exist in the Followee / Follower network. Based on the content of the tweet, the attributes of the user, and the diffusion speed of the information, considerations are given, and the information diffusion that occurs outside the Followee / Follower network such as TwitterTrends and external websites has a particularly small number of followers Explain that the tweets of authors are essentially important for widespread spread, and that the influence of "influencers" on the spread of retweets is limited.

概要

近年、ソーシャルネットワーキングサービス（Social Networking Service: SNS）の出現により、多様かつ膨大な情報が SNS を介して急速に拡散し、我々の生活に大きな影響を及ぼす時代となっている。SNS は個人間のコミュニケーションツールに留まらず、災害時の情報共有、マーケティング、株価の予測、選挙活動など様々な場面で活用されている。その一方で、デマが拡散される、個人や集団に対する炎上の温床となるなどの負の側面も指摘されていることから、SNS における情報拡散過程の特徴を明らかにすることが求められている。本研究では、代表的な SNS として Twitter を対象に、TwitterAPI を用いてキーワード検索によるツイートの収集・分析を行い、リツイート数が増加する要因を明らかにすることを試みた。ユーザー間のネットワークに着目してフォロワーとフォロワーから成るネットワークを作成し、リツイートによって情報が拡散していく経路を分析した。その結果、リツイートを行ったユーザーの中でフォロワー/フォロワーネットワークに存在しないユーザーが多数存在することを発見した。その上でツイートの内容やユーザーの属性、情報の拡散速度などの観点から考察を行い、Twitter トレンドや外部ウェブサイトなどのフォロワー/フォロワーネットワーク「外」で生じる情報の拡散が、特にフォロワー数が少ない投稿者のツイートが広く拡散される上で本質的に重要であること、リツイートの拡散において「インフルエンサー」の影響は限定的であることを説明する。

目次

第 1 章	序論	1
1.1	研究背景	1
1.2	研究目的	1
1.3	本論文の構成	1
第 2 章	関連研究	2
2.1	Twitter について	2
2.1.1	Twitter トレンド	3
2.1.2	ウェブサイト	4
2.2	TwitterAPI, ツイートデータについて	5
2.3	関連研究	5
第 3 章	リツイート経路の分析方法	9
3.1	ツイートデータの仕様	9
3.2	フォロワー/フォロワーネットワーク	9
3.2.1	情報元が不明の場合	12
第 4 章	データ分析	14
4.1	取得したデータの概要	14
4.2	リツイート経路: トップツイート全体の傾向	16
4.3	リツイート経路: 認証済みアカウント	19
4.3.1	認証済みの個人 (著名人)	19
4.3.2	認証済みの組織 (企業・ニュースサイト)	22
4.4	リツイート経路: 非認証アカウント	29
4.4.1	一般人	29
4.4.2	非認証の企業・ニュースサイト	37
4.5	非公開アカウント, 追跡不能なリツイート数と情報拡散速度	45
4.5.1	非公開アカウントの影響	47
4.5.2	追跡不能なリツイート数と情報拡散速度	47
第 5 章	結論	49
	参考文献	50
	謝辞	52

第 1 章

序論

1.1 研究背景

近年、ソーシャルネットワーキングサービス (Social Networking Service: SNS) の出現により、多様かつ膨大な情報が SNS を介して急速に拡散し、我々の生活に大きな影響を及ぼす時代となっている。SNS は個人間のコミュニケーションツールに留まらず、災害時の情報共有、マーケティング、株価の予測、選挙活動など様々な場面で活用されている。その一方で、デマ (フェイクニュース) が拡散される、個人や集団に対する炎上の温床となるなどの負の側面も指摘されている。このように、SNS は我々の生活に対して大きな影響力を持っており、その情報拡散過程の特徴を明らかにすることが求められている。情報拡散過程の研究対象は大きく 3 つに分類でき、「どのように拡散が進むのか」、「誰が拡散において重要な役割を果たすのか」、「どのような内容が拡散されやすいのか」という観点で研究が進められている。

1.2 研究目的

本研究では、代表的な SNS として Twitter を対象に、TwitterAPI を用いてキーワード検索によるツイートの収集・分析を行い、リツイート数が増加する要因を明らかにすることを試みた。Twitter には、自身のタイムラインに流れてきたツイートを自身のフォロワーに拡散する「リツイート」と呼ばれる機能がある。各ユーザーがツイートをリツイートしていく際、各ツイートが獲得するリツイート数には大きな差が見られる。ユーザー間のネットワークに着目してツイートデータを分析し、ツイート投稿者や「インフルエンサー」、またツイートの内容などの各種要因が情報拡散 (リツイート数の増加) にどのような影響を与えるのかについて考察することを目的とする。

1.3 本論文の構成

本論文は以下のような構成になっている。第 2 章では、本研究が対象とする SNS である Twitter の概要とデータ収集に用いた TwitterAPI について説明し、また SNS を対象にした関連研究について述べる。第 3 章では、収集したツイートデータを用いてリツイートの経路を分析する方法について述べる。第 4 章では、リツイート経路の分析結果とその考察を述べ、第 5 章では結論を述べる。

第 2 章

関連研究

第 2 章では、本研究で対象とした SNS である Twitter と Twitter に関連する Web サービス、データ収集のために用いた TwitterAPI、ツイートデータの各種パラメータについての説明を行う。また、SNS を対象とした関連研究について紹介する。

2.1 Twitter について

Web 上で社会的ネットワークの構築を可能にするサービスのことを SNS (Social Networking Service, ソーシャル・ネットワーキング・サービス) といい、Twitter はその SNS の 1 つである。Twitter の月間アクティブユーザー数は日本で 4500 万人超、全世界で 3 億 3000 万人を超えており、巨大な SNS であるといえる。Twitter には半角 280 字以内のテキストメッセージ・画像・動画・URL を投稿することができ、これを「ツイート」と呼ぶ。Twitter 上で他のユーザーを「フォロー」することで、そのユーザーのツイートを自身のタイムライン上で閲覧することが可能になる。ここで、ユーザー A から見たとき、A のフォロイー (followee) とは A がフォローしているユーザーを指し、A のフォロワー (follower) は A をフォローしている (A がフォローされている) ユーザーを指す。ユーザーは、自身や他者のツイートを自身のフォロワーに拡散することができ、これをリツイート (Retweet, RT) と呼ぶ。また、自身や他者のツイートに共感や賛同を示す「いいね (Like)」を送ることができる。

Twitter には短文を気軽に投稿でき、情報拡散のリアルタイム性・広範囲性に長けるという特徴がある。かつては学校・職場・趣味仲間などのコミュニティ間で利用されるコミュニケーションツールとしての面が強かったが、2011 年 3 月に発生した東日本大震災を契機として、大規模災害時の被害状況や避難情報などの情報伝達の場としての側面を持ち始めた [1]。

Twitter はその特徴から、企業のマーケティングや米大統領選挙 (2016) のようなプロモーション活動などにも活用事例がある。一方、「炎上」が発生する、不安を煽るような間違ったデマ情報 (フェイクニュース) が拡散してしまうなどといった危険な側面も持ち合わせている。たとえば、先述の米大統領選挙 (2016) では「ローマ法王がトランプ氏の支持を表明」や「クリントン氏を捜査中の FBI 捜査官が無理心中」等のフェイクニュースが大規模に拡散し、選挙戦終盤には主要メディアによる記事を上回るリアクションやコメントを集めている [1]。また、国内においても、2016 年 4 月に発生した熊本地震の際に、「動物園からライオンが逃走している」という虚偽の文章と写真が Twitter 上に投稿され、動物園や警察に問合せが殺到した。このように Twitter にはメリットとともに多くの課題があり、利用者のリテラシー向上や Twitter 社・政府の適切な対応が求められている。

2.1.1 Twitterトレンド

Twitter内で事件・事象が話題になり、急激に出現回数が増えると、そのキーワードが「Twitterトレンド」に表示される。トレンドはリアルタイムで更新され、アクセスしたタイミングによって、表示されているトレンドワードは異なる。Twitterのモバイルアプリでは、トレンドは[話題を検索]タブの[トレンド]セクションに表示される(図2.1)。パソコンでTwitterにログインした場合は、ホームタイムライン、通知、検索結果、プロフィールページなどの多くの場所にトレンドが表示される(図2.2)。トレンドをクリックまたはタップすると、そのトレンドのTwitter検索結果が表示され、そのキーワードやハッシュタグを含むすべてのツイートを見ることができる(図2.3)。



図 2.1: Twitterトレンド (モバイル)



図 2.2: Twitterトレンド (PC)



図 2.3: Twitterトレンドタップ後

TwitterトレンドはTwitter社のアルゴリズムによって決定されているが、詳細は明らかになっていない。トレンドの決定方法について、Twitter公式のヘルプセンターにはこのように記載されている。

トレンドはアルゴリズムによって決定され、初期設定では、フォローしているアカウント、興味関心、位置情報をもとにカスタマイズされています。ここ数日や今日1日で話題になったトピックではなく、今まさに注目されているトピックが選り出されるため、Twitterで盛り上がっている最新的话题をリアルタイムで見つけることができます [2]。

また、トレンドのルールについて、Twitter公式のヘルプセンターにはこのように記載されている。

Twitterでは、トレンドにより、健全な議論が促進されることを目指しています。そのため、特定のコンテンツがトレンドに表示されないようにする場合があります。表示されないのは次のようなコンテンツです。

- ・ 冒涔的な内容、成人向けや性的な内容をほのめかすコンテンツ

・人種，民族，出身地，性的指向，性別，性同一性，信仰している宗教，年齢，障碍，疾患を理由に嫌がらせを扇動するコンテンツ

・Twitter ルールに違反しているコンテンツ

また，違反となる可能性を判断する際に，コンテンツの話題性，つまり公共の利益が考慮される場合があります．その結果，コンテンツがそのままトレンドに表示されることがあります [2]．

Twitter ユーザーは Twitter トレンドを見ることで注目されているトピックを知ることができ，トレンドワードをクリック（タップ），リツイートすることでフォロワーではないユーザーのツイートを拡散することができる．Twitter トレンドは世間の事件・事象を如実に反映するため，Twitter トレンド入りしたキーワードを分析し，ランキング形式で発表する「Twitter トレンド大賞」が例年開催されている [3]．

2.1.2 ウェブサイト

Twitter では自身や他者のツイートをコピーし，ウェブサイトや別のツイートにリンクを貼り付けることができる（図 2.4）．このようにしてニュースサイトやまとめサイト，ブログにツイートを引用することができる．たとえば，台風や地震などの災害が発生すると，テレビ局や新聞局が「現地の人の声」として一般人のツイートを自社のニュースサイトに（許可を取った上で）引用することがある．また，Twitter 上のツイートを集めて公開できる Together（トゥギャッター）[4] というウェブサービスが存在する（図 2.5）．イベントなどの実況をまとめる，論争が起きた場合に各論者の立場をわかりやすくするなどの使い方がなされている．このようなウェブサイトを経由することで，自身のフォロワーでないユーザーのツイートに関与することができる．



図 2.4: ニュースサイトにおけるツイートの引用



図 2.5: Together

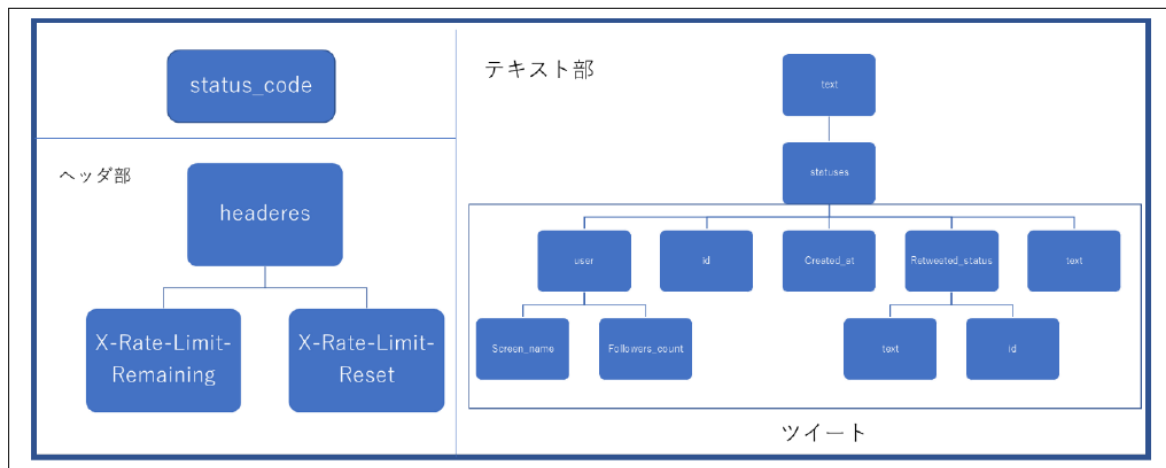


図 2.6: ツイートデータのフォーマットイメージ

2.2 TwitterAPI, ツイートデータについて

本研究で Twitter 上のデータを取得する際に用いた TwitterAPI と、取得したツイートデータについて簡単に解説する。Twitter 社は、Twitter にある情報をできるだけ広く共有するために、TwitterAPI (Application Programming Interface) を企業、開発者、利用者に提供している。TwitterAPI を利用することで、Twitter データにプログラムレベルでアクセスすることができる [5]。本研究では、TwitterAPI を利用してツイートデータを収集した。収集期間は 2019 年 10 月から 2020 年 1 月までとし、ツイートデータは適宜本論文に追加した。キーワード検索では、指定したキーワードを本文に含むツイートが収集できる。

得られたツイートデータのフォーマットイメージを図 2.6 に示す。また、取得したツイートデータが示すものを簡単に解説する。

`status_code` は API 呼び出しの正常/異常判定を行うものであり正常終了でコード 200 を返す。`X-Rate-Limit-Remaining` はアクセス可能回数を示す。カウントが 0 になると API に使用制限がかかり、`X-Rate-Limit-Reset` (アクセスリセット時間) が 0 になるまで使用できない。`text` はツイートの本文、各ツイートに固有の `id` はツイートの ID を表す。`screen_name` はユーザー表示名、`followers_count` はユーザーのフォロワー数、`created_at` はツイートが投稿された日時を示す。取得したツイートがリツイートであった場合、オリジナルのツイート情報を `retweeted_status` で得ることが出来る。

2.3 関連研究

前章で述べた通り、SNS がマーケティングや政治など社会に与える影響は拡大の一途を辿っている。したがって SNS を対象とした研究は社会的に重要とされており、様々な角度から研究が進められている。SNS を対象とした既存研究は大きく以下の 3 つに分類される。

1 つ目は「どのように拡散が進むのか」である。ネットワーク上のあるユーザーから発信された情報が他のユーザーを介して拡散された後に、やがて情報拡散は減衰し収束するが、この一連のダイナミクスを解明するものである。Kwak ら [6] は Twitter のビッグデータを収集・分析することにより、Twitter のネットワークについてフォロワーの分布がべき乗則に従わないこと、有効直径 (ネットワーク内の最も離れた 2 つのノード間

の最短距離)が短いこと, 相互性(ノード間の関係が相互フォローになっているかどうか)が低いことを発見した。これらはすべて, 人間のソーシャルネットワークの既知の特性から逸脱している。また, Twitterトレンドのトップとなったトピックの大部分(85%以上)が大見出しのニュースまたは長期的なニュースであることを示しており, リツイートの詳細を見ると, リツイートされたツイートは, 元のツイートのフォロワー数に関係なく, 平均1,000人のユーザーに到達することが明らかになっている。

[7]ではTwitter上における隔たり(何ホップで対象の人物に辿り着けるか)が平均して3.43人であると報告されている。また, Facebookは2011年にミラノ大学と共同調査を行い, 世界中のFacebookユーザーのうち任意の2人を隔てる人の数は平均4.74人であると発表している[8]。これらはMilgramが検証した[9]「六次の隔たり」より小さい数であり, SNS世界の「狭さ」を示唆しているといえる。

Twitterトレンドのアルゴリズム(どの要因がトレンドを形成するのか)の詳細は2.1.1で述べたように明らかになっていないが, Asurら[10]はTwitterトレンドに関する研究によって, トレンドの形成, 持続性, および減衰について理論的基礎を提示している。トレンドのトピックの成長を説明する確率モデルが対数正規分布を示すこと, ユーザーのフォロワー数やツイート率などの要因はトレンドに大きく寄与しないこと, ほとんどのトピックは長期間トレンドにならないことが判明した。Twitterトレンドを形成する大きな要因は他のユーザーによるリツイートであるが, ユーザーの属性よりも話題になっているコンテンツへの関連が強い。トレンドになったコンテンツの大部分は従来のメディアソース(テレビやネットニュース)から提供され, その後, Twitterで繰り返しリツイートされることでトレンドが生成されることがわかった。

Kleinebergら[11]はTwitterやFacebookなどのOSN(Online social networks)のトポロジーの進化を再現する定量的な2層モデルを提示している。これにより, 実際の社会構造, ウイルス拡散メカニズム, およびマスメディア間のカップリングがOSNの進化を支配することが示された。モデルの検証により, バイラル効果(自分が得た情報を他者に伝えたい効果)はマスメディアの影響の4~5倍強いこと, ウイルスの伝染性は社会的つながりの強さに反比例することがわかった。この結果は, 社会で情報が伝染するプロセスにおける「構造的多様性」の役割に関して示唆を含んでいるとKleinebergは述べている。

2つ目は「誰が拡散において重要な役割を果たすのか」である。企業のマーケティング, 災害時の情報共有など, 広範囲に情報を拡散させたい場合に重要となるユーザー(「インフルエンサー」)を特定することで効率的に拡散を行うことが出来ると考えられているため, その特定に関する研究が進められている。

[12]では, 情報の拡散プロセスに関してユーザーのごく一部が大きな影響を持ち, 他のほとんどのユーザーは情報の伝播中にミドルウェアとしての役割を果たす(情報拡散の速度や拡散する情報量の増減に大きな影響を与えない)ことが示されている。ユーザーエンジン(ユーザーが情報拡散プロセスの中でどれくらいユーザーに影響しているか), ユーザーの熱意, およびユーザーの継続時間の3つの測定値を提案し, 情報拡散との関係を述べている。また, 情報拡散ネットワークの構造と動的性を研究するためにコミュニティ分析を行った結果, ユーザーはクリエイター, プロモーター, サポーター, コンシューマーの4つの異なるカテゴリーに分類され, 情報拡散への関与方法と他者への影響がそれぞれ異なることを示した。大量のリツイートを獲得できるユーザーは, ユーザーエンジンの価値が高いと結論付けている。

Bakshyら[13]の研究では, Twitterフォロワーグラフで発生した7,400万件の拡散イベントを追跡することにより, 160万人のTwitterユーザーの属性と相対的な影響を調査している。影響力がありフォロワーが多いユーザーのツイート, またAmazon Mechanical Turkの労働者によってより興味深いと評価された/より肯定的な感情を引き出したツイートが広がる可能性が高いことがわかった。しかし, どの特定のユーザーまたはツイートが大きな流れを生成するかについての予測は信頼性が低かったことから, 口コミの拡散は, 多数の潜在的なインフルエンサーをターゲットにして, 平均的な効果を捉えることによってのみ活用できると結論付けている。またこの研究では, 潜在的なインフルエンサーを考慮するマーケティング戦略を検討している。その結

果、(トップクラスのインフルエンサーは高額な依頼料が必要になるが、それほど費用対効果が高くないことに対して)「普通のインフルエンサー」、つまり平均的または平均以下の影響力を発揮する個人を使用することで、最も費用対効果の高いマーケティングを行うことができるとしている。

観光地のマーケティングを行う組織(Destination marketing organizations, DMO)においては、Twitterを利用して観光地の魅力や季節ごとのイベントを宣伝することが重要な課題とされている。そのため、DMOのネットワークを分析することによって、影響力が最も大きいインフルエンサーのカテゴリーを特定する研究が行われている[14]。ネットワーク分析用のソフトウェアであるNodeXLを使用し、他のユーザーにリプライ/リツイートした回数やDMOに関する投稿で言及された回数などの尺度で、DMOのネットワーク内のアカウントにおける相対的な影響力をランク付けしている。その結果、ホテルやレストランなどのステークホルダーは、ネットワーク内での重要度が低く、DMOによる宣伝の機会を十全に利用していないことが示されている。その上で、影響力が高いと判断されたユーザー(インフルエンサー)とパートナーシップを模索すること、他のステークホルダーとコミュニケーションを取ることで、Twitter上における存在感を高めることができるとしている。Twitter上では企業アカウント同士の会話・コラボレーションが頻繁に行われており(「レッドブル」と「モンスター」、「シャープ」と「タニタ」など)、そのようなアクティビティはユーザーの注目を引くことを考えると、この研究における指摘は重要といえる。

3つ目は「どのような内容が拡散されやすいのか」に関する研究である。Vosoughiら[15]の研究では、2006年から2017年までのツイートを収集・分析することによって、真実を伝えるニュースよりもフェイクニュースのほうが速く、早く、深く、広範囲に広がっていくこと、そしてボット(自動でツイートを行うプログラム)やネットワーク構造がフェイクニュースの拡散に与える影響はほぼないことを明らかにした。この研究では、真実は1000人以上に到達することさえめったにない一方で、フェイクニュースのうち拡散度合いで上位1%にあたるものは、ほぼ常に1000~10万人もの人びとに到達していたこと、政治に関するフェイクニュースは速やかにリツイートされ、結果的に他のカテゴリーのフェイクニュースが1万人に到達するよりも3倍近い速度で2万人以上ものユーザーに到達したことなどが示された。フェイクニュースを広めやすいユーザーの特徴として、フォロワー/フォロワー数が少ないこと、ツイートの頻度が低いこと、認証ユーザーの割合は少ないこと、アカウントの保有期間が短い(Twitter歴が短い)ことが挙げられている。

三浦らの研究[16]では2011年3月11日に発生した東日本大震災の際に発生したツイートデータを用いて、特に感情を表す「感情語」に着目してその出現傾向と災害の種類の間を関連性を検討している。ネガティブ感情語あるいは活性度の高い感情語が多く含まれるツイートであるほど多数回リツイートされ、中でも極端に高い伝播性を示したツイートについては不安あるいは活性感感情語がより多く含まれていたことが示された。[15]からも同様の結果(ネガティブ感情語・活性感感情語が含まれているツイートは拡散されやすい)が示されている。また、Schäferら[17]は「2014年の日本の衆議院議員総選挙で、政治的な意見やキーワードがどのように共有・拡散がされたか」についてビッグデータ分析を行い、その結果から「あらかじめ設定された特定の単語を含む投稿を自動的にリツイートするボットによる大量の投稿があり、それらが言論の多様性を弱めている」と指摘した。ボットによって特定の情報・フレーズの量が増大され極端な方向性が生みだされることで、民主主義が健全に機能しない可能性をSchäferらは指摘している。[15]では「ボットはフェイクニュースの拡散にほぼ影響を与えない」とされているが、ボットによって生み出された極端な方向性がTwitterユーザーに影響し正当な議論を妨げる可能性が考えられる。

ツイートの内容を分析して現実の事象を分析・予測する試みとして、佐々木ら[18]はツイート内容から地震を検出する研究を行っている。地震が発生すると、人々は地震に関連してTwitterに多くのツイートを投稿するが、このツイート群を観察することで地震の発生を迅速に検出できるとしている。ツイートを監視し、リアルタイムの相互作用を調査することで地震を検出するアルゴリズムを提案している。ツイート内のキーワー

ド、単語数、およびそれらのコンテキストなどの機能に基づいて、ツイートを分類し、その上で地震について予測するための確率的時空間モデルを作成する。日本では全国各地で多数の地震が発生し、また Twitter ユーザーが多いため、ツイートを監視することで高い確率で地震を検出できるとしている（気象庁（JMA）の震度スケール 3 以上の地震の 96 %を検出している）。このシステムでは地震を検出すると登録ユーザーに電子メールで通知を行うが、その通知は JMA によってなされるアナウンスよりもはるかに速く配信される。

Petrovic ら [19] は、ツイートがリツイートされるかどうかを人間がある程度予測できることを実験によって示し、また Passive-Aggressive アルゴリズム（Passive：訓練事例を正しく分類できた場合は重みを更新しない、Aggressive：訓練事例を正しく分類できなかった場合は重みを更新するというアルゴリズム）に基づいた機械学習アプローチを使用したモデルを利用して同様の予測が行えることを示している。この研究では、ツイートの作成時間に応じて個別のモデルを構築する時間依存モデルを提案しており、通常の Passive-Aggressive アルゴリズムよりもパフォーマンスが大幅に向上するとしている。

また、Sloan らの研究 [20] では、イギリスの Twitter ユーザーのプロフィールから人口統計データ（年齢、職業、社会階級）を自動識別するための 2 つのツールを指定、設計、および評価している。専門家による検証によって自動照合プロセスの妥当性と信頼性を評価しており、2011 年国勢調査をベースラインとした比較によってイギリスの Twitter ユーザーの将来のクラス分布を予測している。2011 年の国勢調査の時点で、Twitter ユーザーの年齢分布は英国の人口分布よりもはるかに若く、16~22 歳が全ユーザーの 67.5 %を占めることがわかった。また一方で、英国では 40 歳以上の Twitter ユーザーが 50 万人以上存在する可能性があるとしている。

Bollen ら [21] は、ツイートはその内容や使用目的に関係なく、著者の気分に関する情報を伝えるという観点から感情分析を行っている。日々のタイムラインについて、6 つの気分（緊張、鬱、怒り、活力、疲労、混乱）を抽出し、株式市場・原油価格指数・米国大統領選挙・感謝祭などの主要なイベントによって記録された変動と比較した結果、社会的、政治的、文化的、経済的分野の出来事が、公共の気分のさまざまな次元に、重要で、即時かつ非常に特異的な効果をもたらすことを発見した。ユーザーが生成したツイートを分析することで、感情の集合について傾向をモデル化および予測できるが、ユーザーが置かれている社会的、経済的、文化的な観点を考慮した上で結果を議論する必要があるとしている。

第 3 章

リツイート経路の分析方法

第 3 章では、第 2 章で述べた方法で収集したツイートデータを利用してリツイートの経路を分析する方法について説明する。

3.1 ツイートデータの仕様

2.2 節で述べたように、TwitterAPI で取得したツイートのメタデータには `retweeted_status` というパラメータが設定されており、このパラメータを参照することでそのツイートが情報発信元のツイートであるかリツイートされたものであるかを調べることができる。情報発信元のツイート（オリジナルツイート）の `retweeted_status` には空白が表示され、リツイートの `retweeted_status` には情報発信元の `tweet_ID` が表 3.1 のように表示される。

表 3.1: ツイートデータ（ユーザー名, `tweet_ID`, `retweeted_status`）

ユーザー名	<code>tweet_ID</code>	<code>retweeted_status</code>
A	AAAAAAAAA	
B	BBBBBBBBB	AAAAAAAAA
C	CCCCCCCCC	AAAAAAAAA

`retweeted_status` からは「ユーザー B がユーザー A のツイートをリツイートした」ということがわかるが、例えば「ユーザー A のオリジナルツイートをユーザー B がリツイートし、それをまたユーザー C がリツイートした」場合、`retweeted_status` からそのことを読み取ることはできない（「ユーザー A のオリジナルツイートをユーザー B がリツイートした」と「ユーザー A のオリジナルツイートをユーザー C がリツイートした」ことが別々に記録される）[22]。したがって、単にツイートデータを収集するだけでは、リツイートの経路を分析することはできない。

3.2 フォロイー/フォロワーネットワーク

前節で述べたように、ツイートデータから直接リツイート経路を取得することはできない。そのことを踏まえ、本研究では、ユーザー間のフォロイー/フォロワー関係に着目してリツイート経路の分析を行った。Twitter でリツイートを受け取る時、ユーザーは自身のフォロイーからリツイートを受け取ることになる。そこで、

ツイートをリツイートしたユーザーとそのフォロイー/フォロワーで構成されるネットワーク（本研究ではこのネットワークをフォロイー/フォロワーネットワークと呼称する）を作成し、リツイート経路の分析を行った。ここで、リツイートを行ったユーザーは、自身のフォロイーのうち、自身より前の時刻にリツイートを行ったユーザーからリツイートを受け取っているとした。各ユーザーのフォロイーリストはツイートデータと同様に TwitterAPI を利用して取得した。リツイートを行ったユーザーのリストと各ユーザーのフォロイーリストを比較し、一致するユーザーが存在した場合はそのユーザーからツイート（リツイート）を受け取ったとして、ユーザーの ID を情報元としてフォロイー/フォロワーネットワークに追加する (図 3.1)。

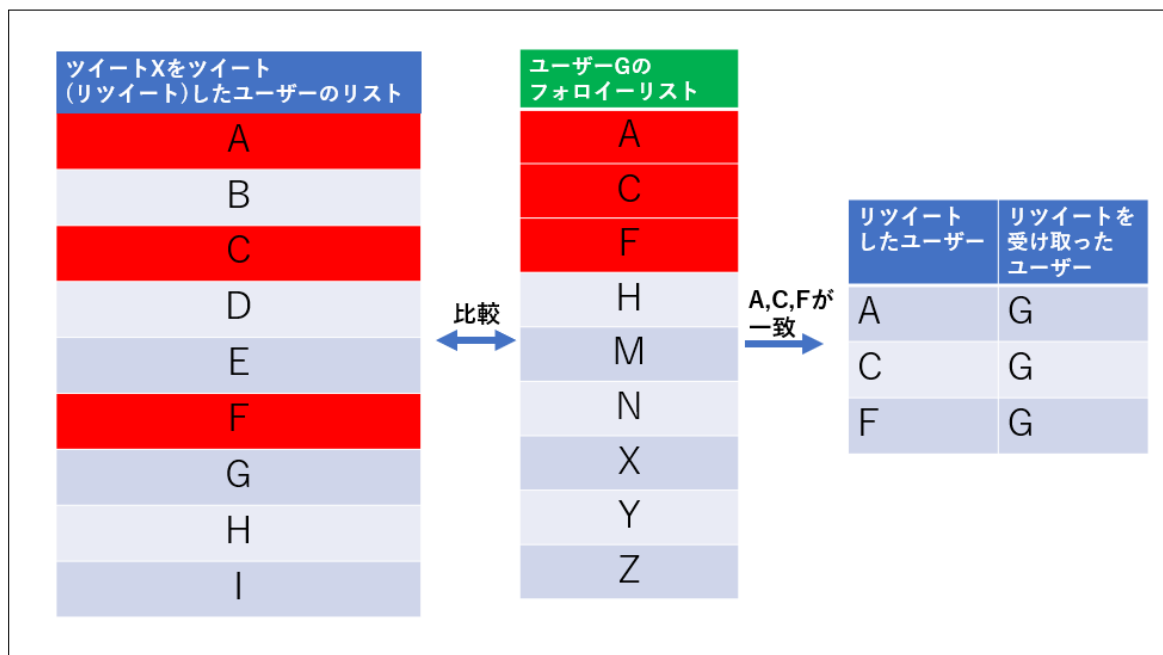


図 3.1: フォロイーフォロワーネットワーク (一致)

両リストを比較しても一人も一致しなかった場合、「情報元が不明である」としてフォロイー/フォロワーネットワークに追加する (図 3.2)。

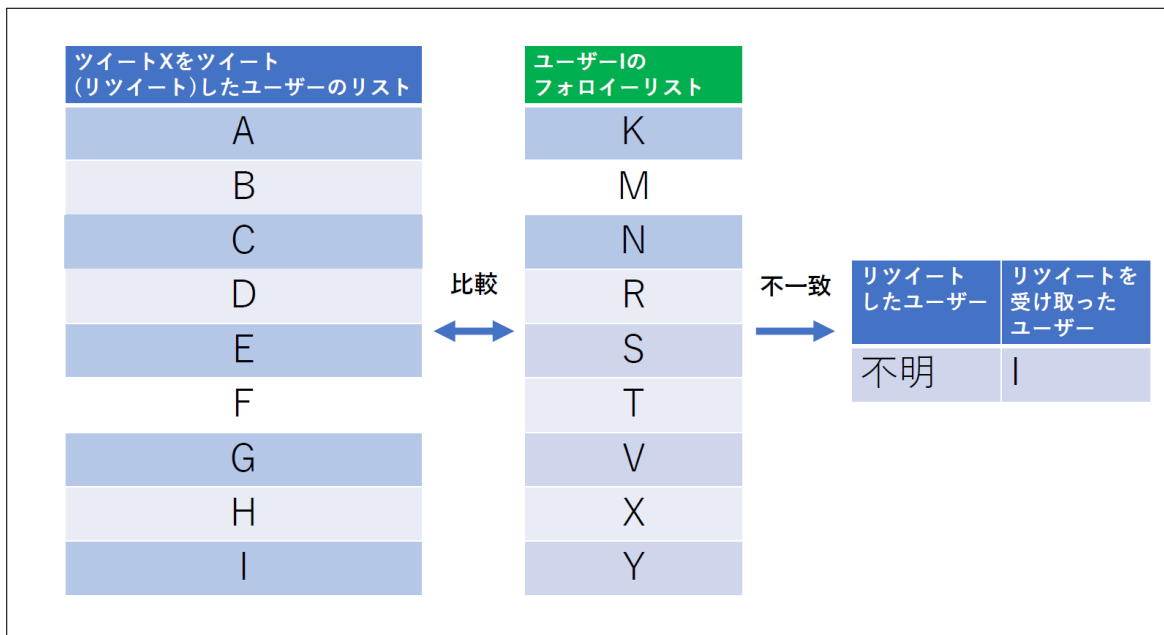


図 3.2: フォロワーフォロワーネットワーク (不一致)

(図 3.3) のようにフォロワー/フォロワーネットワークが構築された場合、情報元を辿っていくとオリジナルツイートに辿り着くリツイートを「経路を追跡可能なリツイート」、情報元を辿ってもオリジナルツイートに辿り着かないリツイートを「経路が追跡不能なリツイート」とする。

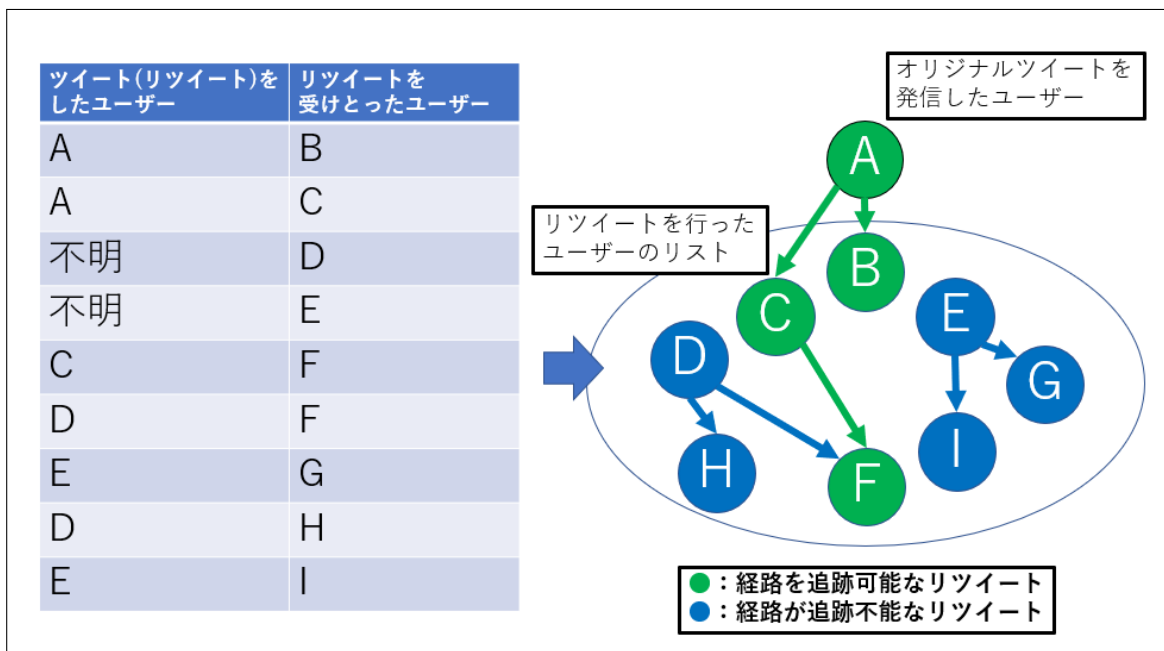


図 3.3: フォロワーフォロワーネットワーク完成図

3.2.1 情報元が不明の場合

リツイートを行ったユーザーのリストとそのユーザーのフォロワーリストの中に一致する ID がない場合、そのユーザーの情報元は 3 種類考えられる。

1. 非公開アカウントを介した拡散
2. 「いいね」のシェアによる拡散
3. フォロワー/フォロワーネットワーク外の経路での拡散 である。

非公開アカウントが公開アカウントにリツイートを行った場合、1. のケースになる。TwitterAPI の仕様上、非公開アカウントのツイート/リツイートを取得することはできない。そのため、非公開アカウントが公開アカウントにリツイートを行った場合、直接の情報元は「不明」となる。

Twitter では、フォロワーが「いいね」したツイートが自身のタイムラインに表示されることがランダムで起こる。図 3.4 では、ユーザー A のタイムラインに表示されている 3 件のツイートのうち 1 件目が A のフォロワーのツイートであり、2 件目と 3 件目は A のフォロワーが「いいね (Like)」したツイートである (A のフォロワーのツイートではない)。そのツイートをリツイートした場合 2. のケースとなり、直接の情報元は「不明」となる。Twitter の公式アプリである TweetDeck[24] では、Activity 欄を参照することでフォロワーの「いいね」したツイートを知ることができる (図 3.5)。

3. のフォロワー/フォロワーネットワーク外でツイートが拡散するケースとしては、2.1.1 節で述べた Twitter トレンド、2.1.2 節のニュースサイトやまとめサイト、そしてキーワード検索やエゴサーチなどを介してリツイートを行う場合が考えられる。



図 3.4: 「いいね」のシェアによる拡散



図 3.5: 「TweetDeck」の Activity 欄

第4章

データ分析

第4章では第3章で述べた方法を用いて行ったデータ分析の結果について述べる。本研究では、Twitter上で話題となった(Twitterトレンド入りした)事象に対してキーワード検索でツイートデータを取得し、キーワードごとのツイートデータの中でリツイート数が最も多かったもの(トップツイート)についてリツイート経路の分析を行った。#をキーワードの前につけると「ハッシュタグ」としてキーワードとは別に扱われることから(キーワードだけで取得できるツイートデータとハッシュタグだけで取得できるツイートデータは違う)、いくつかの話題についてはハッシュタグもキーワード検索に含めた。

4.1 取得したデータの概要

各キーワードごとのデータセットについて、データの取得期間、総ツイート数、リツイート数とフォロワー数の相関係数を表4.1に示す。今回取得した全47件のデータセットの半数以上でリツイート数とフォロワー数の相関係数は0.2以下であり、相関は小さいといえる。つまり、フォロワー数が多いユーザーであるからといって多くのリツイート数を獲得できるわけではないといえる。

表 4.1: 取得したツイートデータ

キーワード	データ取得期間	総ツイート数	リツイート数とフォロワー数の相関係数
アーモンドアイ	2019/10/18~2019/10/27	101666	0.149226
千葉駅	2019/10/19~2019/10/28	94087	0.0907
天皇賞	2019/10/20~2019/10/29	143730	0.083426
#地味ハロウィン	2019/10/20~2019/10/29	326006	0.13033
地味ハロウィン	2019/10/20~2019/10/29	68122	0.150244
札幌 dis OR #札幌 dis	2019/10/20~2019/10/30	100849	0.173684
Airpods Pro	2019/10/22~2019/10/31	281489	0.014485
首里城	2019/10/24~2019/11/01	1031269	0.180116
決闘罪	2019/10/25~2019/11/03	64629	0.00499
PayPay ドーム OR ペイペイドーム	2019/10/28~2019/11/03	224781	0.272776
南アフリカ AND ラグビー	2019/10/31~2019/11/09	110278	0.126131
流行語 OR #流行語大賞	2019/11/01~2019/11/10	138288	0.149933

表 4.1: 取得したツイートデータ

キーワード	データ取得期間	総ツイート数	リツイート数と フォロワー数の 相関係数
なでしこ寿司 OR #なでしこ寿司	2019/11/01~2019/11/10	177822	0.042047
香港デモ OR #香港デモ	2019/11/02~2019/11/11	108978	0.08053
京アニ放火 OR #京アニ放火	2019/11/02~2019/11/11	27119	0.420685
即位パレード OR #即位パレード	2019/11/02~2019/11/11	114105	0.140986
ターミネーター OR #ターミネーター無双	2019/11/02~2019/11/11	263942	0.00182
二宮 OR ニノ AND 結婚 OR#二宮結婚 OR#ニノ結婚	2019/11/07~2019/11/15	551435	0.151203
飯塚幸三 OR #飯塚幸三	2019/11/07~2019/11/16	122627	0.263154
松本零士 OR #松本零士	2019/11/10~2019/11/19	82121	0.369247
#バルス安倍やめろ	2019/11/13~2019/11/22	96534	0.379413
沢尻エリカ OR #沢尻エリカ	2019/11/14~2019/11/22	1300097	0.117322
GSOMIA OR #GSOMIA	2019/11/18~2019/11/27	563742	0.104239
いい風呂の日 OR #いい風呂の日	2019/11/19~2019/11/28	169556	0.081664
ローマ法王 OR ローマ教皇	2019/11/20~2019/11/28	441582	0.043196
モバイル Suica	2019/11/20~2019/11/29	9182	0.270532
次元の声	2019/11/21~2019/12/01	40916	0.011225
中曽根	2019/11/25~2019/12/04	307280	0.083296
アンタッチャブル	2019/11/26~2019/12/05	254382	0.289218
#アナ雪 2 と未知の旅へ	2019/11/26~2019/12/05	18531	0.374108
アナ雪 OR ステマ	2019/11/27~2019/12/06	575152	0.080526
#プレイステーションの日	2019/11/30~2019/12/07	187531	0.005111
サイバーマンデー OR #サイバーマンデー	2019/11/30~2019/12/09	234920	0.141319
ホームアローン OR #ホームアローン	2019/12/01~2019/12/10	71498	0.022411
忘年会スルー OR #忘年会スルー	2019/12/03~2019/12/12	51685	0.1122
ヘイトスピーチ OR #ヘイトスピーチ	2019/12/05~2019/12/14	126515	0.121004
Mac Pro OR #Mac Pro	2019/12/05~2019/12/14	121082	0.035157
今年の漢字 OR #今年の漢字	2019/12/06~2019/12/15	191198	0.190991
梅宮	2019/12/06~2019/12/15	294219	0.188244
ドイツ鉄道 OR #ドイツ鉄道	2019/12/06~2019/12/17	98886	0.055888
グレタさん OR #グレタさん	2019/12/14~2019/12/22	211818	0.040922
ミルクボーイ OR #ミルクボーイ	2019/12/18~2019/01/26	587227	0.092549
首吊り自殺 OR #首吊り自殺	2020/01/01~2020/01/09	53513	0.123858
イラン AND 撃墜	2020/01/09~2020/01/17	127608	0.171297
ゴロゴ OR #ゴロゴ	2020/01/09~2020/01/17	22162	0.497307
成人式 OR #成人式	2020/01/10~2020/01/17	2060618	0.165341
武漢 OR #武漢	2020/01/16~2020/01/24	597832	0.16409

4.2 リツイート経路：トップツイート全体の傾向

各キーワードごとのトップツイートについてリツイート経路を分析した。各 Twitter トレンドワードのトップツイートは平均 1,000 人のユーザーに到達すること [6], Twitter ユーザーにはリツイート数の多いツイートを選択的にリツイートする傾向があること [23] から, リツイート数が 1000 件に達するまでの過程が重要であると考えられる。そのため, 本研究では最初の 1000 件のリツイートに焦点を当てた。本研究ではリツイート経路分析に使用した各ツイートについて, リツイート 1000 件までのフォロワーリストを取得し, フォロワー/フォロワーネットワークを作成してリツイート経路の分析を行った。各ユーザーのフォロワーリストは TwitterAPI を利用して取得した。TwitterAPI の仕様上, 1 回のアクセスにつきユーザー 15 人のフォロワーリストを取得できるが, 1 アクセスごとに約 15 分の待ち時間が発生する。

トップツイートのリツイート経路の特徴について, 最初に全体としての傾向を示す (表 4.2)。各キーワードごとのトップツイートについて, 獲得したリツイート数, 投稿者のフォロワー数, 認証の有無 (公式アカウントかどうか), 1 次拡散者 (ツイート投稿者からツイートを受け取り, リツイートを行ったユーザー) の人数, 経路が追跡不能なユーザー数について示す。リツイート数とフォロワー数の相関は小さく (相関係数は -0.087701), フォロワー数が多いほど 1 次拡散者は多くなり (相関係数 0.528474), 追跡不能なユーザーは少なくなる (相関係数 -0.464847) 傾向にある。また, 直接の情報元が不明なリツイートユーザーは平均して約 3 割 (314.91 人) であり, 平均して約 6 割 (589.23 人) のリツイートユーザーが追跡不能であることから, 情報が拡散していく過程においてこの「フォロワー/フォロワーネットワークに含まれない追跡不能な経路」は無視できないといえる。

表 4.2: トップツイートの概観

キーワード	RT 数	フォロワー数	認証	1 次拡散者	追跡不能者
札幌 disOR #札幌 dis	9711	142767	有	253	401
PayPay ドーム OR ペイペイドーム	17170	947536	有	462	453
南アフリカ AND ラグビー	12801	273832	有	535	433
流行語 OR #流行語大賞	7920	126639	有	714	117
香港デモ OR #香港デモ	13196	274623	有	836	72
松本零士 OR #松本零士	18483	591615	有	509	242
GSOMIA OR #GSOMIA	13090	300459	有	761	78
中曽根	11683	2746349	有	611	200
アンタッチャブル	16212	252270	有	519	195
#プレイステーションの日	18658	1159735	有	894	88
サイバーマンデー OR #サイバーマンデー	41597	684677	有	987	12
忘年会スルー OR #忘年会スルー	8998	321695	有	70	912
梅宮	47938	981601	有	645	276
ミルクボーイ OR #ミルクボーイ	43541	52923	有	89	895
イラン AND 撃墜	9473	386779	有	402	339

表 4.2: トップツイートの概観

キーワード	RT 数	フォロワー数	認証	1 次拡散者	追跡不能者
ゴロゴ OR #ゴロゴ	13478	1012768	有	551	243
アーモンドアイ	1987	7008	無	401	301
千葉駅	11240	80	無	1	999
天皇賞	17760	1490	無	28	373
#地味ハロウィン	36516	81926	無	145	764
地味ハロウィン	5179	1150	無	10	975
Airpods Pro	60340	349	無	2	995
首里城	44922	1071	無	6	991
決闘罪	45865	261	無	3	974
なでしこ寿司 OR #なでしこ寿司	27117	32143	無	64	888
京アニ放火 OR #京アニ放火	2908	28549	無	201	616
即位パレード OR#即位パレード	11279	4726	無	55	846
ターミネーター OR #ターミネーター無双	50323	2964	無	16	946
二宮 OR ニノ AND 結婚 OR#二宮結婚 OR#ニノ結婚	38124	689	無	8	991
飯塚幸三 OR #飯塚幸三	16094	72545	無	161	811
#バルス安倍やめろ	3087	23136	無	531	183
沢尻エリカ OR #沢尻エリカ	65246	1337	無	10	908
いい風呂の日 OR #いい風呂の日	22143	13837	無	148	760
ローマ法王 OR ローマ教皇	47533	16871	無	17	772
モバイル Suica	1427	3376	無	74	480
次元の声	21646	2238	無	9	977
#アナ雪 2 と未知の旅へ	10229	113890	無	365	539
アナ雪 OR ステマ	21786	1415	無	34	875
ホームアローン OR #ホームアローン	15507	6594	無	10	941
ヘイトスピーチ OR #ヘイトスピーチ	11065	3868	無	21	588
Mac Pro OR #Mac Pro	9145	5330	無	10	933
今年の漢字 OR #今年の漢字	11314	77178	無	988	9
ドイツ鉄道 OR #ドイツ鉄道	27114	5353	無	9	989
グレタさん OR #グレタさん	31094	5899	無	63	503
首吊り自殺 OR #首吊り自殺	6398	62131	無	189	730
成人式 OR#成人式	42257	702	無	12	961
武漢 OR #武漢	12562	44706	無	428	120

また、トップツイートの投稿者が認証を受けているかどうかで分類して概観した。Twitter にはアカウントを認証する機能があり [25]、著名人のアカウントなどの世間の関心を集めるアカウントが本物であると示すことができる。認証済みアカウントは著名な個人や組織（企業やニュースサイト）のものである可能性が高い（図 4.1）といえる。どのようなアカウントが認証されるのかについて、Twitter のヘルプセンターにはこのよう

に記載されている。



図 4.1: 認証済みアカウントの例：ドナルドトランプ

関心を集めるアカウントであると判断された場合に、アカウントは認証を受けます。通常、音楽、演劇、ファッション、政府、政治、宗教、ジャーナリズム、メディア、スポーツ、ビジネス、その他の注目分野の利用者によって管理されているアカウントが対象となります。

認証済みバッジは、*Twitter* による支持を黙示的に表明するものではありません [25]。

今回取得したデータセット全 47 件のトップツイートのうち、16 件が認証済みアカウントから発信されたものであった。この認証済みアカウントからのトップツイート 16 件については、1 次拡散者の平均が約 552 人、追跡不能なりツイートユーザー数の平均が約 314 人であった。また、フォロワー数と 1 次拡散者数の相関係数は 0.270363、フォロワー数と追跡不能ユーザー数の相関係数は -0.24622 であった。一方、非認証アカウントによるトップツイートは 31 件であり、1 次拡散者の平均が約 130 人、追跡不能なりツイートユーザー数の平均が約 733 人であった。また、フォロワー数と 1 次拡散者数の相関係数は 0.607699、フォロワー数と追跡不能ユーザー数の相関係数は -0.40924 であった。このことから、アカウントが認証済みかどうかによって、リツイート経路の内容に差が見られること、非認証アカウントのほうがフォロワー数がリツイート経路の内容にもたらす影響が大きいことがわかる。

4.3 リツイート経路：認証済みアカウント

本節では認証済みアカウントに焦点を当て、リツイート経路を分析する。

4.3.1 認証済みの個人 (著名人)

ここでは政治家のように知名度があり、情報の拡散力が高いと考えられるユーザーが発信したツイートについて、リツイート数・投稿者のフォロワー数・1次拡散者の人数・(情報元を辿ってもオリジナルツイート者に辿り着かない) 追跡不能なリツイート数を表 4.3 に示す。

表 4.3: 著名人のリツイート経路

キーワード	RT 数	フォロワー数	1次拡散者	追跡不能者
流行語 OR #流行語大賞	7920	126639	714	117
GSOMIA OR #GSOMIA	13090	300459	761	78
香港デモ OR #香港デモ	13196	274623	836	72

今回取得したデータのうち、ツイート投稿者が「認証済みの個人」に該当したのはいずれも政治家 (国会議員・政治運動家) であった。最初のリツイート 1000 件のうち、半数以上が1次の拡散 (ツイート投稿者のフォロワーが拡散したもの) によるものであり、著名人のツイートがリツイートされていく過程では1次拡散者の割合が大きいといえる。

各ツイートについて、リツイート数の上位 10 人中に追跡不能なユーザーが何人いるか、またリツイートが 1000 件に達するまでの時間を表 4.4 に示した。

表 4.4: 著名人のリツイート経路：リツイート数上位 10 人の追跡可否，1000 リツイート到達時間

キーワード	RT 数上位 10 人中の 追跡不能なユーザー数	1000RT 到達時間 (分)
流行語 OR #流行語大賞	0	63
GSOMIA OR #GSOMIA	0	8
香港デモ OR #香港デモ	0	24

「流行語大賞」「香港デモ」「GSOMIA」のツイートについて、リツイート数の時間推移を図 4.2 と図 4.3、図 4.4 に示す。この3つの事例では、リツイート数は時間の経過に伴って直線的に増加していき、また、途中で追跡可能なリツイートと追跡不能なリツイートの総数が入れ替わることはなかった。また、追跡不能なリツイートはほぼ増加しなかった。リツイート数が 1000 件に達するまでの時間は「流行語大賞」が 63 分、「GSOMIA」が 8 分、「香港デモ」が 24 分であった。

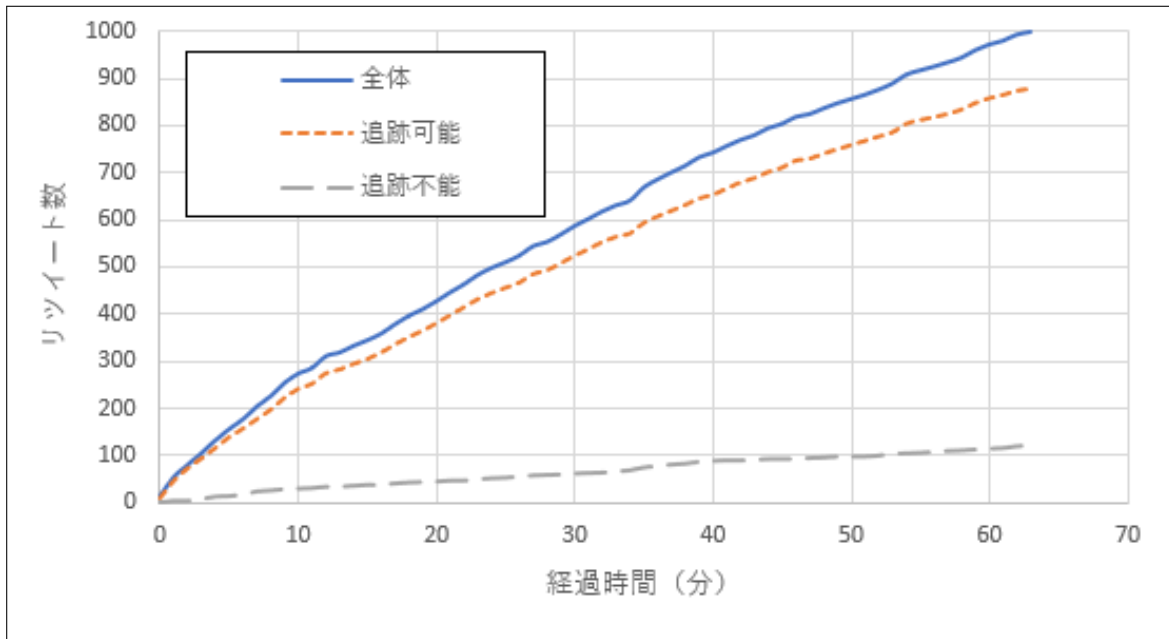


図 4.2: リツイート数時間推移：流行語大賞

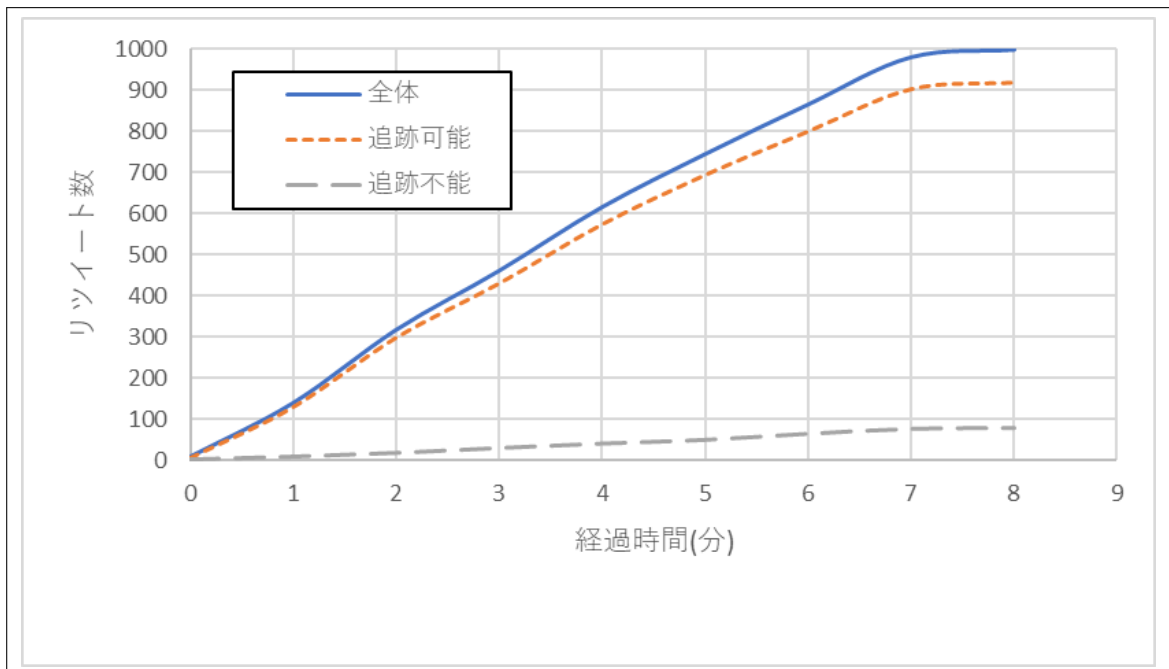


図 4.3: リツイート数時間推移：GSOMIA

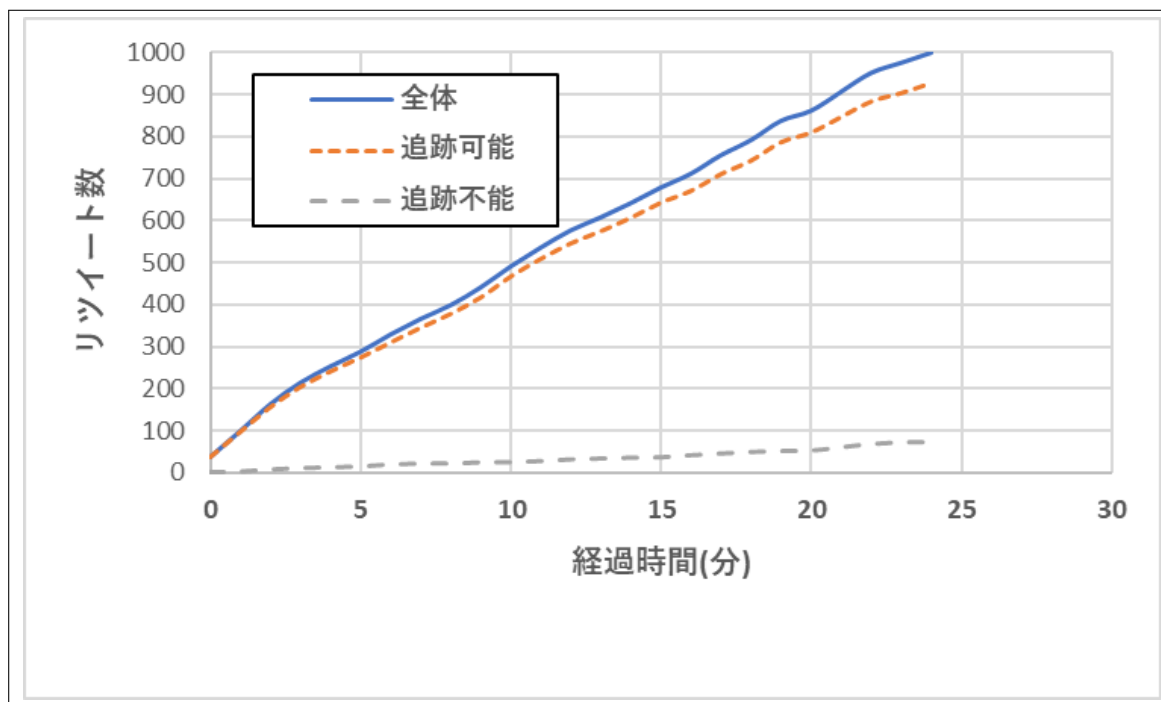


図 4.4: リツイート数時間推移：香港デモ

また、「香港デモ」のツイートにおいて、獲得したリツイート数が上位 10 位であったユーザーの「獲得したリツイート数」「(オリジナルツイートが投稿されてからリツイートするまでの) 経過時間」「リツイート経路の追跡可否」を表 4.5 に示す。

表 4.5: リツイート数上位 10 人：香港デモ

獲得 RT 数	経過時間	経路の追跡可否
837	0:00:00	可
83	0:13:06	可
55	0:02:54	可
50	0:01:39	可
49	0:01:47	可
40	0:02:44	可
40	0:10:30	可
39	0:00:50	可
33	0:07:38	可
29	0:02:37	可

表 4.5 に示されている「リツイート数」は重複を含むことに留意する必要がある。たとえば、ユーザー A がリツイートのリスト 1000 件に含まれる複数人のユーザーをフォローしていた場合、第 3 章で説明したフォロワー/フォロワーネットワーク作成方法ではその複数人からリツイートを受け取ったと表示される。表 4.5 のケースの場合、最初のリツイートユーザー 1000 人のほとんどはツイートの投稿者経由でリツイートを行った 1 次拡散者であると考えられる。また、獲得リツイート件数上位 10 人が全員経路を追跡可能であったことか

ら、初期の情報拡散がフォロワー/フォロワーネットワーク内でなされたといえる。

4.3.2 認証済みの組織 (企業・ニュースサイト)

ここでは認証を受けている企業やテレビ局・新聞局の公式アカウント、ニュースサイトなどから発信されたツイートについて、リツイート数・フォロワー数・1次拡散者の人数・追跡不能な(情報元を辿ってもオリジナルツイート者に辿り着かない)リツイート数を表4.6に示す。

表 4.6: 企業やニュースサイトのリツイート経路

キーワード	RT 数	フォロワー数	1次拡散者	追跡不能者
忘年会スルー OR #忘年会スルー	8998	321695	70	912
イラン AND 撃墜	9473	386779	402	339
札幌 disOR #札幌 dis	9711	142767	253	401
中曽根	11683	2746349	611	200
南アフリカ AND ラグビー	12801	273832	535	433
ゴロゴ OR #ゴロゴ	13478	1012768	551	243
アンタッチャブル	16212	252270	519	195
PayPay ドーム OR ペイペイドーム	17170	947536	462	453
松本零士 OR #松本零士	18483	591615	509	242
#プレイステーションの日	18658	1159735	894	88
サイバーマンデー OR #サイバーマンデー	41597	684677	987	12
ミルクボーイ OR #ミルクボーイ	43541	52923	89	895
梅宮	47938	981601	645	276

ツイート投稿者が「認証済みの組織」に該当したデータのうち、企業のマーケティングとみなせるものは「忘年会スルー」「PayPay ドーム」「#プレイステーションの日」「ミルクボーイ」のトップツイートであり、その他のトップツイートはニュースサイトによるものであった。最終的な獲得リツイート数が同じくらいであっても(「忘年会スルー」と「イラン AND 撃墜」, 「サイバーマンデー」と「ミルクボーイ」など), 1次拡散者の人数や追跡不能なリツイート数には大きな差が見られ、「どのような経路でツイートが拡散したのか」には違いがあることがわかる。

各ツイートについて、リツイート数の上位 10 人中に追跡不能なユーザーが何人いるか、またリツイートが 1000 件に達するまでの時間を表 4.7 に示した。

「忘年会スルー」「イラン AND 撃墜」「サイバーマンデー」「ミルクボーイ」の対象ツイートについて、リツイート数の時間推移を図 4.5 と図 4.6, 図 4.7, 図 4.8 に示す。

「忘年会スルー」は NHK のニュースで取り上げられたワードであるが、最もリツイートを獲得したのは LINE 株式会社のツイートであった。こちらはプレゼントのキャンペーンなどがなかったため、拡散に時間がかかり、追跡不能なリツイートが 9 割を占めていると考えられる(情報の拡散途中で追跡不能なリツイートが増えており、Twitter トレンドなどの影響が考えられる)。「イラン AND 撃墜」のトップツイートは産経ニュー

表 4.7: 企業やニュースサイトのリツイート経路：RT 数上位 10 人の追跡可否，1000RT 到達時間

キーワード	RT 数上位 10 人中の 追跡不能なユーザー数	1000RT 到達時間 (分)
忘年会スルー OR #忘年会スルー	8	560
イラン AND 撃墜	0	13
札幌 disOR #札幌 dis	1	66
中曽根	1	2
南アフリカ AND ラグビー	4	0.63
ゴロゴ OR #ゴロゴ	2	22
アンタッチャブル	0	17
PayPay ドーム OR ペイペイドーム	8	7
松本零士 OR #松本零士	2	12
#プレイステーションの日	0	47
サイバーマンデー OR #サイバーマンデー	0	5
ミルクボーイ OR #ミルクボーイ	9	6
梅宮	2	100

スが報じた「イランが旅客機の撃墜を認めた」という速報であった。「忘年会スルー」と「イラン AND 撃墜」のトップツイートはリツイート数がともに 9000 件ほどであり、ツイート投稿者のフォロワー数も 30 万人のオーダーであるが、最初のリツイート 1000 件に着目した場合の 1 次拡散者の人数、追跡不能なリツイート数、リツイート 1000 件までの到達時間は大きく異なっている。このことから、「ツイートの内容は情報拡散の内容に影響を与える」ということができる。

「サイバーマンデー」と「ミルクボーイ」のトップツイートはそれぞれ「ユナイテッド・シネマ」と「ケロログ」の公式アカウントから発信されたものであり、どちらも企業のマーケティングとしてのツイートである。最終的に獲得したリツイート数はどちらも 4 万件ほどであるが、こちらも最初のリツイート 1000 件に着目するとその情報拡散過程には違いが見られる。リツイート 1000 件までの情報拡散速度に大きな違いは見られないが、リツイートの経路を見ると逆の動きを取っている（「サイバーマンデー」では 1000 件のリツイートの 9 割以上が追跡可能なリツイートであり、「ミルクボーイ」では 9 割近くが追跡不能である）。この現象にはツイートの内容と外的要因が大きく関係している。「サイバーマンデー」では「公式アカウントをフォローした上でリツイートすると抽選でプレゼントが当たる」キャンペーンを行っており（図 4.9）、その結果追跡可能なリツイートでほとんどが占められていると考えられる。一方、ケロログ社が発信したツイートは 2019 年 12 月 22 日に開催、放送された M-1 グランプリにおけるお笑いコンビ「ミルクボーイ」のネタに関連したものであり、ミルクボーイが優勝したことに対するお祝いコメントであった（図 4.10）。そのため、番組を見ていたユーザーがミルクボーイに言及する→「ミルクボーイ」がトレンド入りする→関連ツイートとしてケロログ社のツイートが紹介される→ユーザーがリツイートする、という経路でリツイートが行われたと考えられる。このミルクボーイ（ケロログ社）の事例は、Twitterトレンドがリツイート数の増加に役立った典型的な例といえる。また、ミルクボーイの事例では最もリツイートを獲得したユーザーがケロログ社ではなかった。追跡不能なリツイートが多数を占める情報拡散ではこのようなケースが見られるが、最も情報拡散に寄与しているのは不明な情報元（この場合は Twitterトレンドと考えられる）であるため、最もリツイートを獲得したツイート発信者以外のユーザーが「インフルエンサー」であるとはいえない。

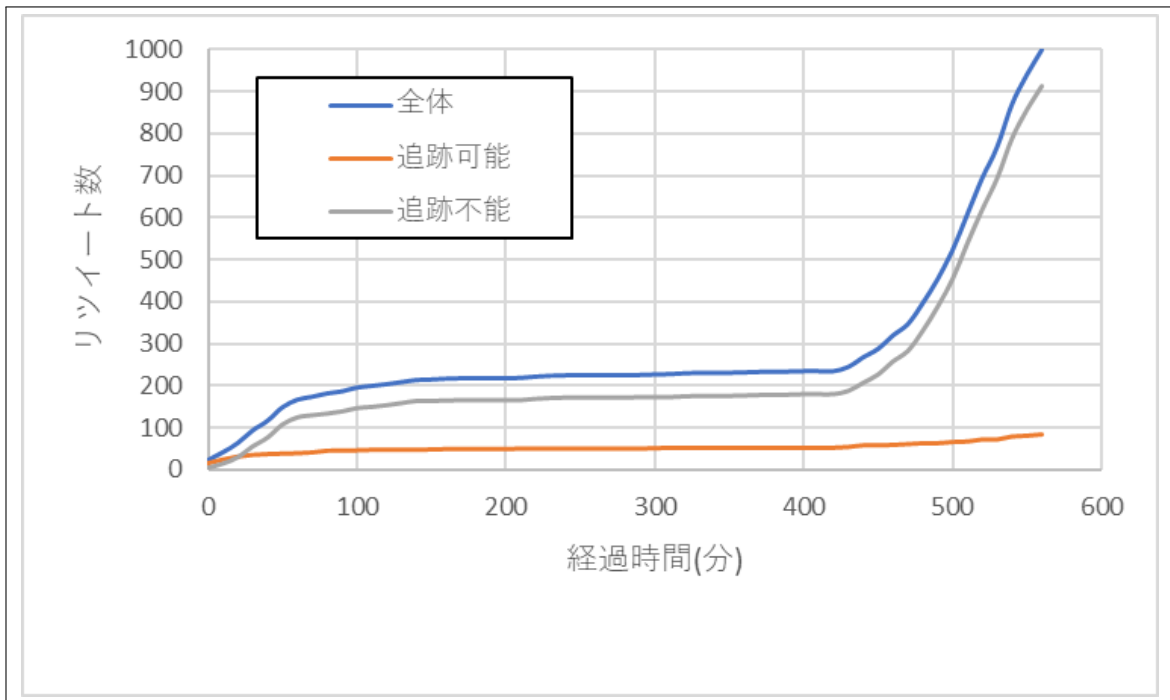


図 4.5: リツイート数時間推移：忘年会スルー

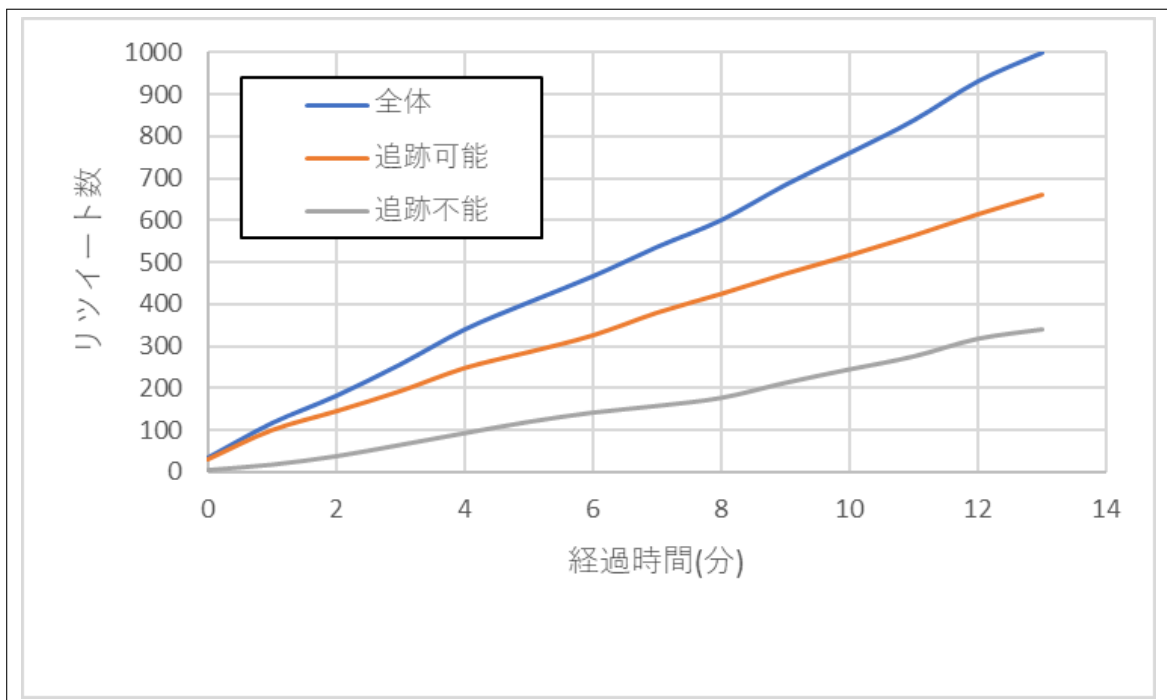


図 4.6: リツイート数時間推移：イラン AND 撃墜

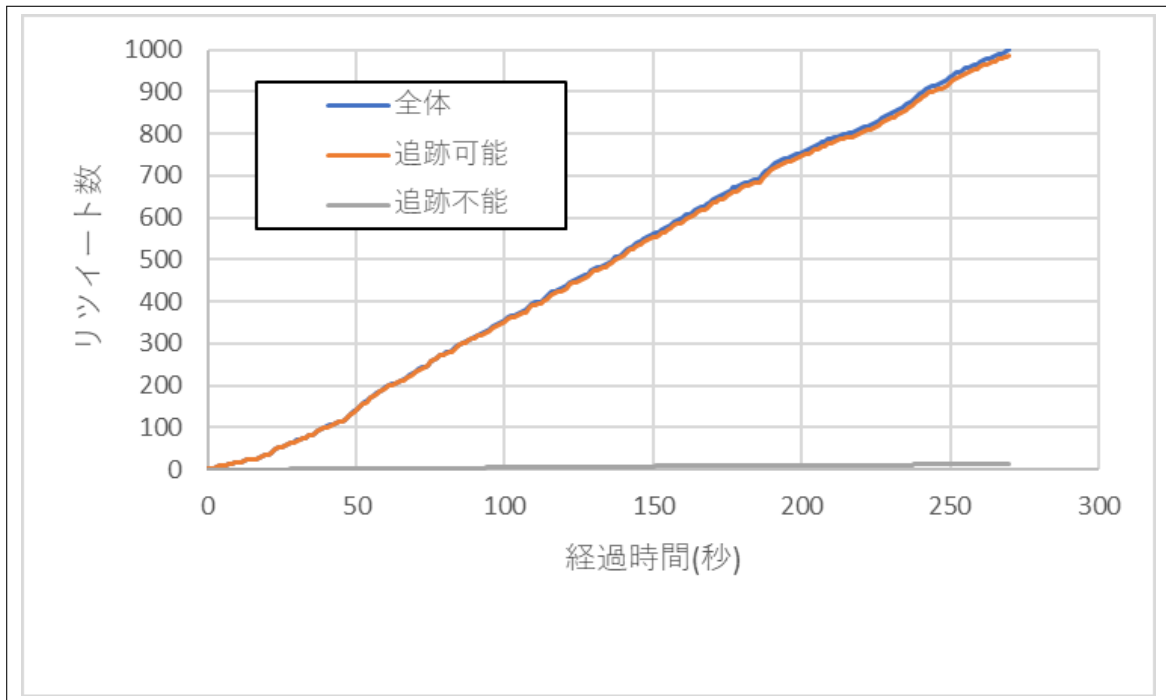


図 4.7: リツイート数時間推移：サイバーマンデー

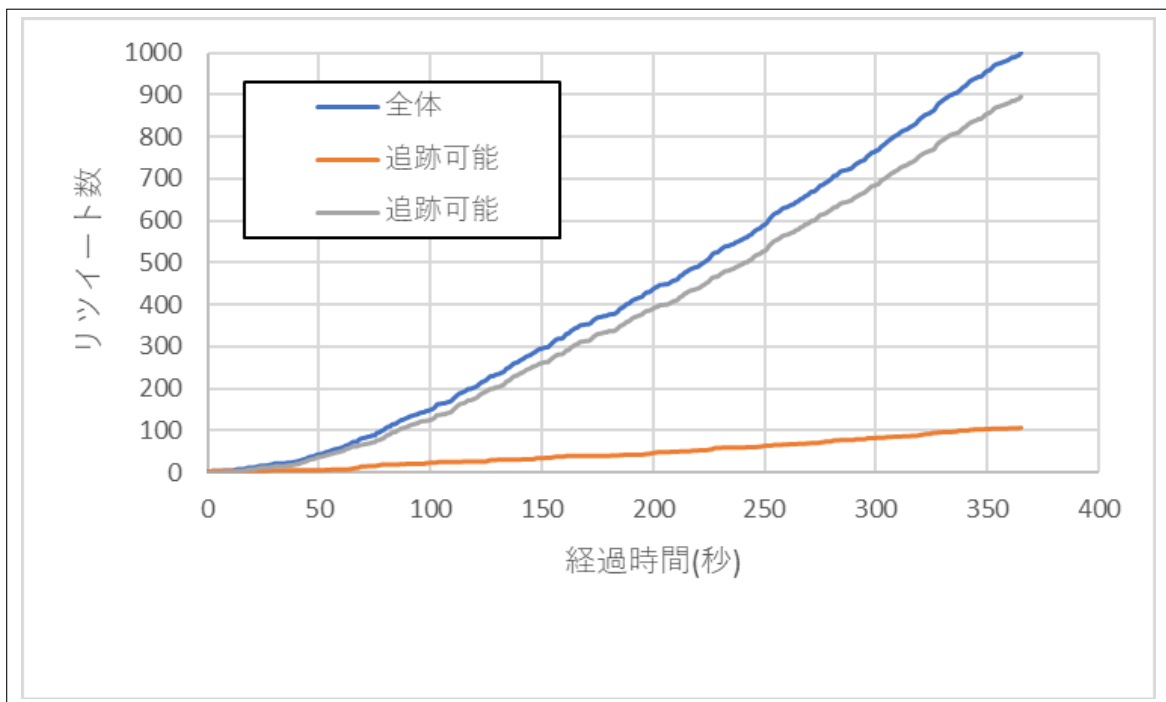


図 4.8: リツイート数時間推移：ミルクボーイ

また、「忘年会スルー」「イラン AND 撃墜」のツイートにおいて、獲得したリツイート数が上位 10 位であったユーザーの「獲得したリツイート数」「(オリジナルツイートが投稿されてからリツイートするまでの) 経過



図 4.9: 「サイバーマンデー」のトップツイート

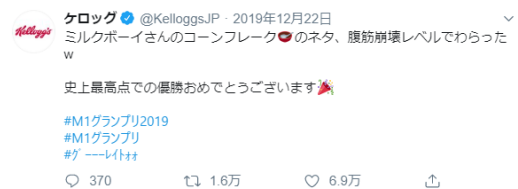


図 4.10: 「ミルクボーイ」のトップツイート

時間」「リツイート経路の追跡可否」を表 4.8 と表 4.9 に示す。

表 4.8: リツイート数上位 10 人：忘年会スルー

獲得リツイート数	経過時間	経路の追跡可否
70	0:00:00	可
9	8:15:31	否
7	7:56:11	否
6	8:54:10	否
4	8:58:16	可
3	7:59:07	否
2	9:03:38	否
2	0:45:31	否
2	8:29:48	否
2	0:43:34	否

表 4.9: リツイート数上位 10 人：イラン AND 撃墜

獲得リツイート数	経過時間	経路の追跡可否
402	0:00:00	可
44	0:06:08	可
38	0:10:48	可
31	0:05:46	可
30	0:03:02	可
12	0:07:59	可
11	0:01:49	可
10	0:06:07	可
10	0:02:55	可
10	0:05:37	可

「忘年会スルー」「イラン AND 撃墜」の 2 つの事例では、最もリツイートを獲得したユーザーがツイートの投稿者であった。「香港デモ」のケースと同じように、認証済みのアカウントから発信されるツイートについては「最多リツイート獲得者はツイートの投稿者である」「その獲得リツイート数はその他のユーザーと大きく差がある」という特徴が見られる。「忘年会スルー」の事例ではトップツイートのリツイート獲得数が 70 件であるが、それ以降のリツイートはトップツイート投稿から時間を空けてなされており、多くが追跡不能であった。「イラン AND 撃墜」のケースでは獲得リツイート数上位 10 人がいずれも追跡可能な（フォロワー/フォロワーネットワーク内の）経路でリツイートしており、トップツイートが 402 件リツイートを獲得している。この 2 つの事例においては、中間の経路で情報が広く拡散する様子は見られず、インフルエンサーの影響

は見られなかった。

キーワード「アンタッチャブル」と「梅宮」に関しては、最もリツイートを獲得したツイートだけでなく2番目のツイートに関するリツイート経路を取得した。表 4.10 と表 4.11, 図 4.11, 図 4.12, 図 4.13, 図 4.14 にそれぞれ結果を示す。

表 4.10: 獲得リツイート数 1 番目と 2 番目の比較

キーワード	リツイート数	フォロワー数	1次拡散者	追跡不能者
アンタッチャブル (リツイート数 1 番目)	16212	252270	519	195
アンタッチャブル (リツイート数 2 番目)	13549	5829	22	971
梅宮 (リツイート数 1 番目)	47938	981601	645	276
梅宮 (リツイート数 2 番目)	17760	1490	1	999

表 4.11: 獲得リツイート数 1 番目と 2 番目の比較 (時間)

キーワード	RT 数上位 10 人中の 追跡不能なユーザー数	1000 リツイート到達時間 (分)
アンタッチャブル (リツイート数 1 番目)	0	17
アンタッチャブル (リツイート数 2 番目)	7	440
梅宮 (リツイート数 1 番目)	2	1.67
梅宮 (リツイート数 2 番目)	10	27

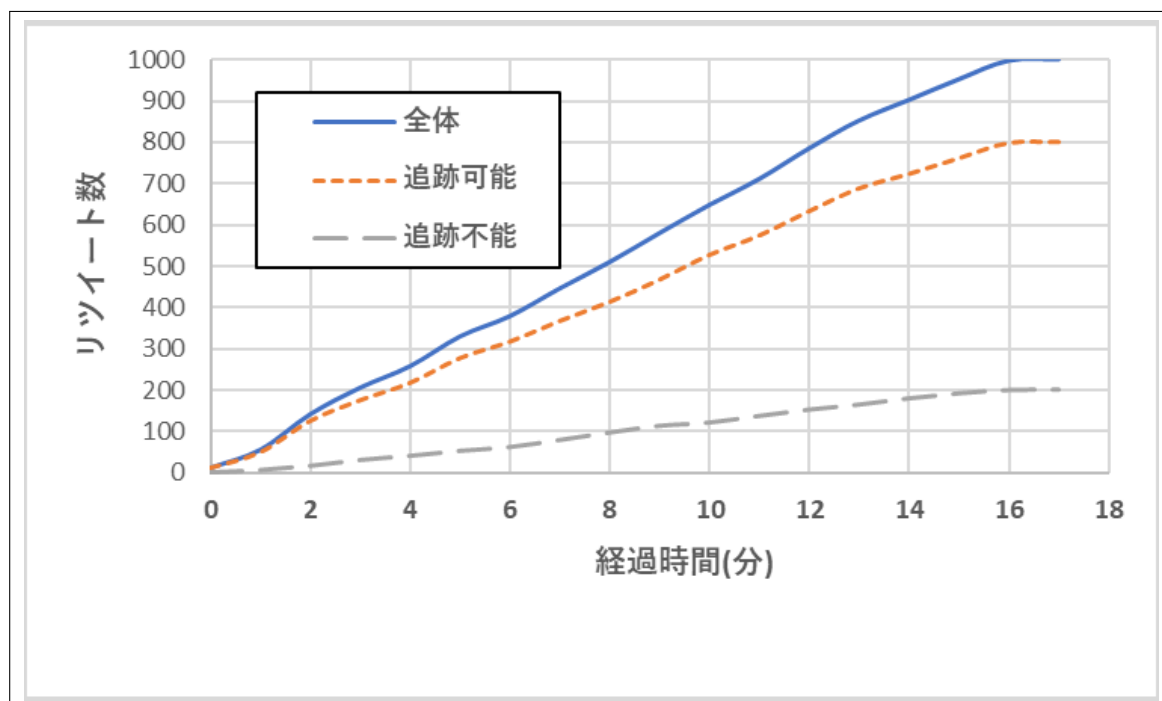


図 4.11: リツイート数時間推移 : アンタッチャブル (1 番目)

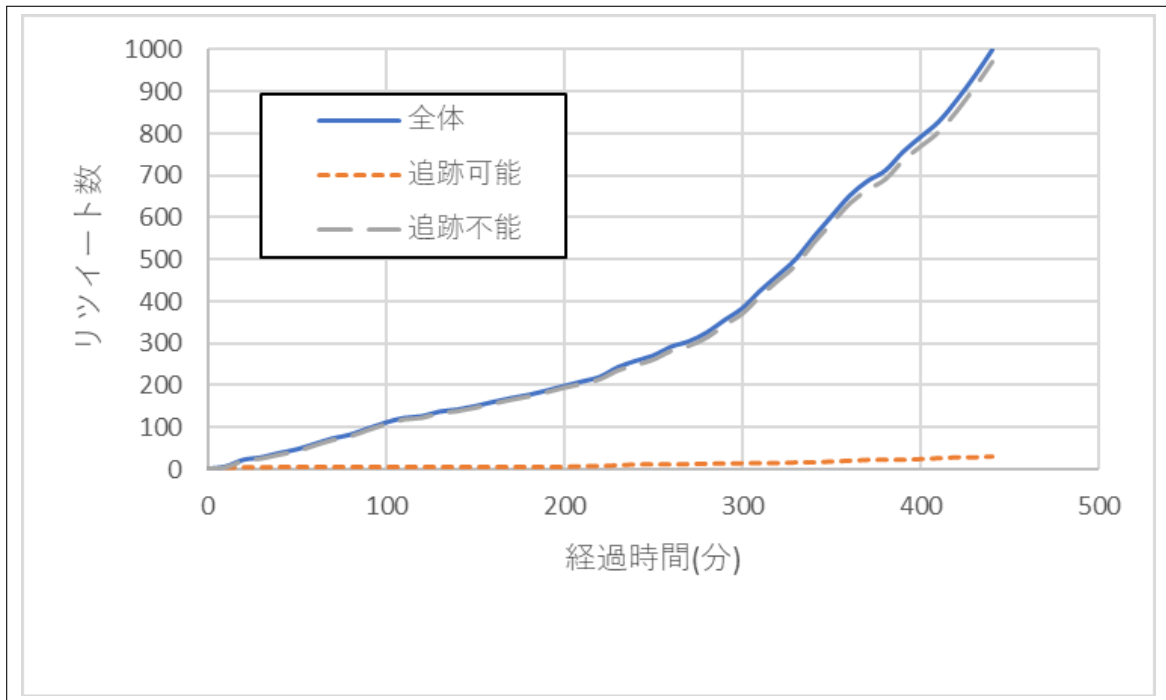


図 4.12: ツイート数時間推移：アンタッチャブル (2 番目)

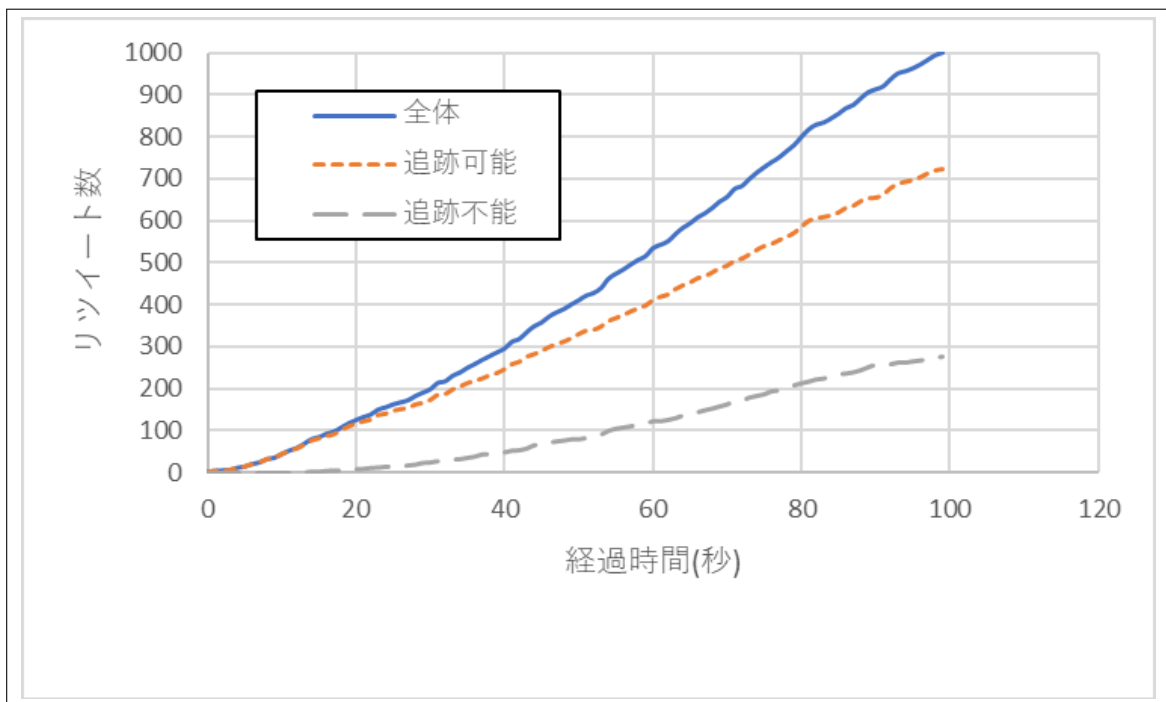


図 4.13: ツイート数時間推移：梅宮 (1 番目)

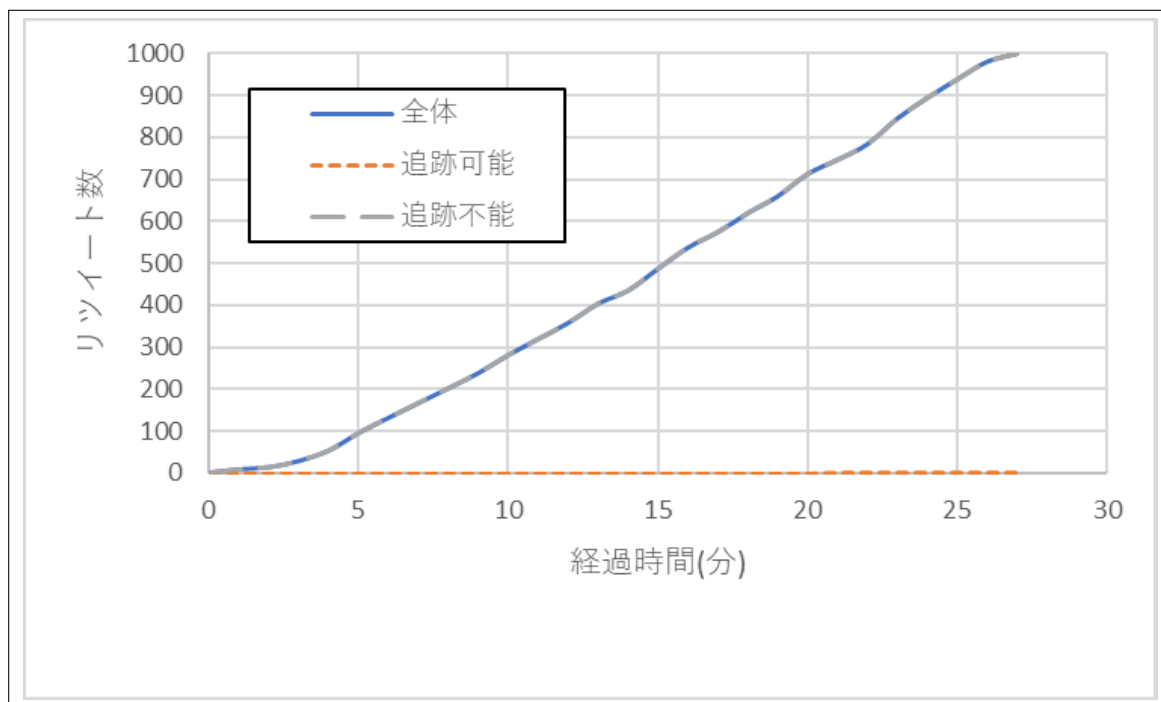


図 4.14: リツイート数時間推移：梅宮 (2 番目)

獲得したリツイート数が 2 番目に大きかったツイートについては、投稿者がいずれも非認証のユーザー (一般人) であった。追跡不能なリツイートが大半を占めており、またトップツイートと比較してリツイート 1000 件に達するまでの時間が長くなっている。「アンタッチャブル」の獲得リツイート数 2 番目では、先述の「忘年会スルー」のケースのように、情報の拡散速度が途中で速くなっている。このケースでは、テレビ番組でお笑いコンビ「アンタッチャブル」が 10 年ぶりに復活したことを受け、視聴者がツイートして Twitter トレンド入りしたと推測される。「ミルクボーイ」の事例と同じように、テレビ番組と Twitter トレンドの影響によってリツイート数が増加したといえる。

4.4 リツイート経路：非認証アカウント

Twitter 社の認証を受けていない一般のユーザーが発信したツイートについて、同様にリツイート経路を分析する。

4.4.1 一般人

Twitter 社の認証を受けていない個人、つまり一般人が発信したツイートについて、対象ツイートのリツイート数、対象ツイート者のフォロワー数、1 次拡散者 (オリジナルツイート者から直接情報を受け取り、リツイートを受け取ったユーザー) の人数、追跡不能 (情報元を辿ってもオリジナルツイート者に辿り着かないユーザー) の人数を表 4.12 に示す。

認証済みのユーザーが発信したツイートと比較すると、1 次拡散者が少ないこと、追跡不能なリツイートが多いことが傾向として挙げられる。

表 4.12: 一般人のリツイート経路

キーワード	リツイート数	フォロワー数	1次拡散者	追跡不能
アーモンドアイ	1987	7008	401	301
京アニ放火 OR #京アニ放火	2908	28549	201	616
#バルス安倍やめろ	3087	23136	531	183
地味ハロウィン	5179	1150	10	975
首吊り自殺 OR #首吊り自殺	6398	62131	189	730
Mac Pro OR #Mac Pro	9145	5330	10	933
ヘイトスピーチ OR #ヘイトスピーチ	11065	3868	21	588
千葉駅	11240	80	1	999
即位パレード OR #即位パレード	11279	4726	55	846
ホームアローン OR #ホームアローン	15507	6594	10	941
天皇賞	17760	1490	28	373
次元の声	21646	2238	9	977
アナ雪 OR ステマ	21786	1415	34	875
いい風呂の日 OR #いい風呂の日	22143	13837	148	760
ドイツ鉄道 OR #ドイツ鉄道	27114	5353	9	989
なでしこ寿司 OR #なでしこ寿司	27117	32143	64	888
グレタさん OR #グレタさん	31094	5899	63	503
二宮 OR ニノ AND 結婚 OR #二宮結婚 OR #ニノ結婚	38124	689	8	991
成人式 OR #成人式	42257	702	12	961
首里城	44922	1071	6	991
決闘罪	45865	261	3	974
ローマ法王 OR ローマ教皇	47533	16871	17	772
ターミネーター OR #ターミネーター無双	50323	2964	16	946
Airpods Pro	60340	349	2	995
沢尻エリカ OR #沢尻エリカ	65246	1337	10	908

各ツイートについて、リツイート数上位 10 人の中に追跡不能なユーザーが何人いるか、またリツイートが 1000 件に達するまでの時間を表 4.13 に示した。認証済みアカウントから発信されたツイートと比較すると、リツイートが 1000 件に達するまでの時間が長い (情報の拡散速度が遅い) 傾向にある。また、追跡不能な経路でリツイートを受け取ったユーザーが獲得リツイート数で上位 10 人に入っているケースが多い。

「千葉駅」「天皇賞」「いい風呂の日」「首里城」「沢尻エリカ」の事例について、リツイートの時間推移を図 4.15 と図 4.16, 図 4.17, 図 4.18, 図 4.19 に示す。

「千葉駅」の事例では、ツイート投稿者のフォロワーが 80 人であるが 11240 件のリツイートを獲得しており、また情報の拡散速度も (その他の一般人のツイートと比較すると) 速いといえる。このツイートは 2019 年 10 月 25 日に千葉県に降った大雨によって京成千葉駅が冠水しているという内容であり、ニュースとしての価値が高かった。当該ツイートをニュースサイトが引用している (図 4.20) [26]。千葉駅が冠水していることに多くのユーザーが言及しトレンドが形成され、その中で当該ツイートのリツイート数が伸びたといえる。「首

表 4.13: 一般人のリツイート経路：RT 数上位 10 人の追跡可否，1000RT 到達時間

キーワード	RT 数上位 10 人中の 追跡不能なユーザー数	1000RT 到達時間 (分)
アーモンドアイ	1	23
京アニ放火 OR #京アニ放火	3	400
#バルス安倍やめろ	2	240
地味ハロウィン	9	420
首吊り自殺 OR #首吊り自殺	4	210
Mac Pro OR #Mac Pro	7	280
ヘイトスピーチ OR #ヘイトスピーチ	3	840
千葉駅	10	60
即位パレード OR #即位パレード	7	27
ホームアローン OR #ホームアローン	9	800
天皇賞	3	80
次元の声	6	45
アナ雪 OR ステマ	8	30
いい風呂の日 OR #いい風呂の日	5	530
ドイツ鉄道 OR #ドイツ鉄道	10	280
なでしこ寿司 OR #なでしこ寿司	7	380
グレタさん OR #グレタさん	0	120
二宮 OR ニノ AND 結婚 OR #二宮結婚 OR #ニノ結婚	9	60
成人式 OR #成人式	8	400
首里城	10	380
決闘罪	9	220
ローマ法王 OR ローマ教皇	2	240
ターミネーター OR #ターミネーター無双	8	190
Airpods Pro	10	240
沢尻エリカ OR #沢尻エリカ	10	960

里城」のツイートも首里城の火災に際してなされたものであるが、こちらはそれに関連して以前に損壊した熊本城の現状をネタにしたものであり、ニュースとしては機能していない。同じようなネタツイートの拡散事例として「天皇賞」の事例があるが、これは追跡可能なリツイートが多数を占め、直線的なリツイート数の増加が見られる。これは一般人のツイートとしてはあまり見られない例である。このツイートは天皇賞をネタにしたジョークであり、投稿者のフォロワー数も飛びぬけて多いわけではないことから、情報拡散におけるツイート内容(面白さ)の影響が大きいと考えられる。「沢尻エリカ」のツイートも、女優の沢尻エリカ氏が麻薬取締法違反容疑で逮捕され、大河ドラマの降板が決定したことに対するジョーク(図 4.21)であり、即時性が低い(内容の面白さで拡散している)と考えられる。「いい風呂の日」の事例では、温泉の記事を書くライターがおすすめの温泉を紹介している。これは温泉のマーケティングと同時にライター自身のマーケティングとして機能している。オリジナルツイートの投稿から 400 分ほどで追跡不能なリツイート数が急激に伸びているが、これは Twitter トレンドやニュースサイトなどの外部の要因のほかに時刻の影響が大きいと考えられる。2019 年

11月26日の0時19分にツイートが投稿されており、ユーザーが反応を始めたのは約400分後(7時前)である。その際追跡不能なリツイートが増加しており、Twitterトレンドの影響が考えられる。

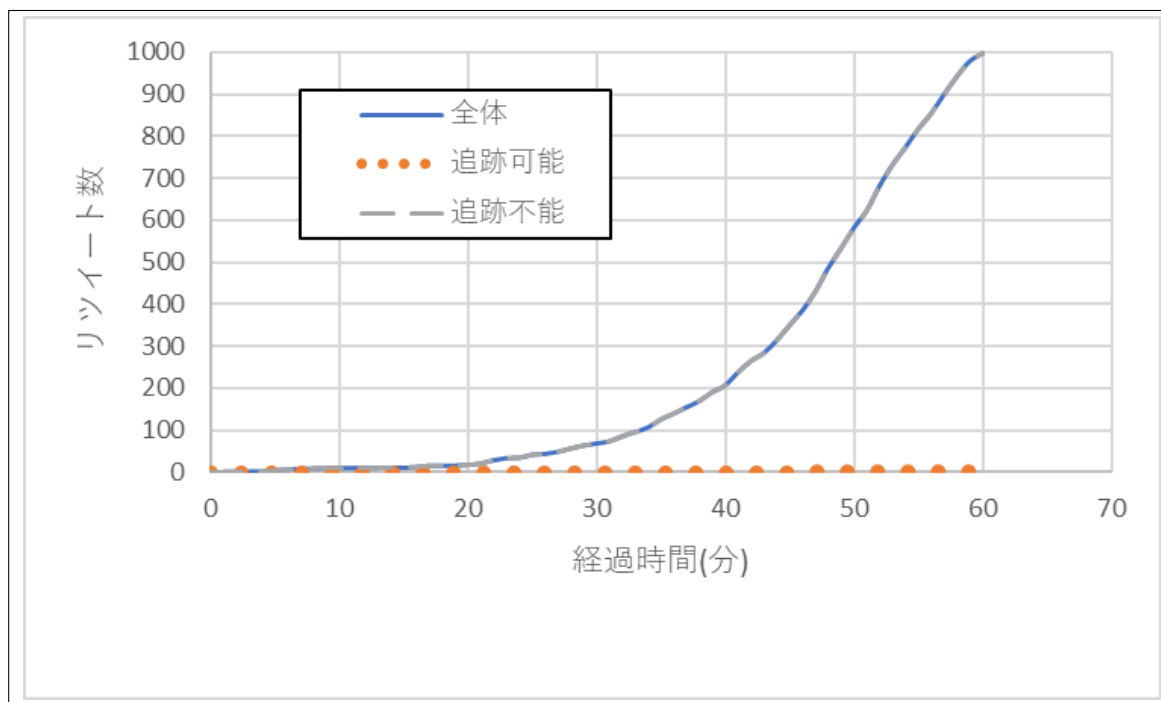


図 4.15: リツイート数時間推移：千葉駅

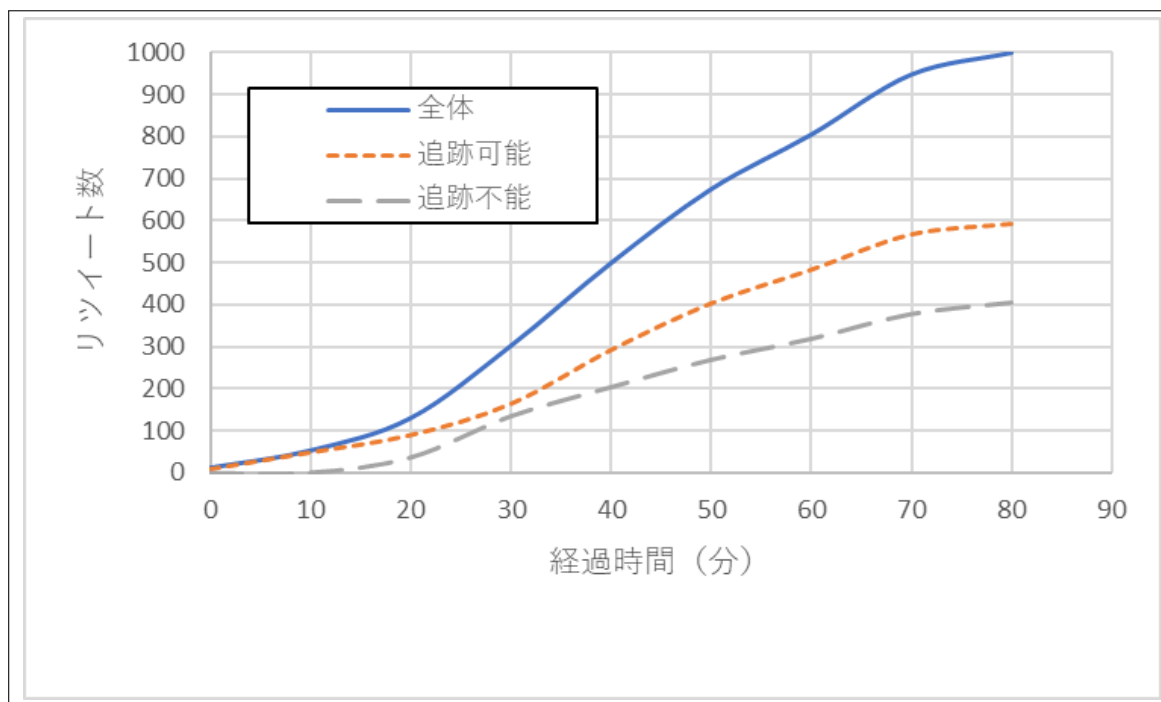


図 4.16: リツイート数時間推移：天皇賞

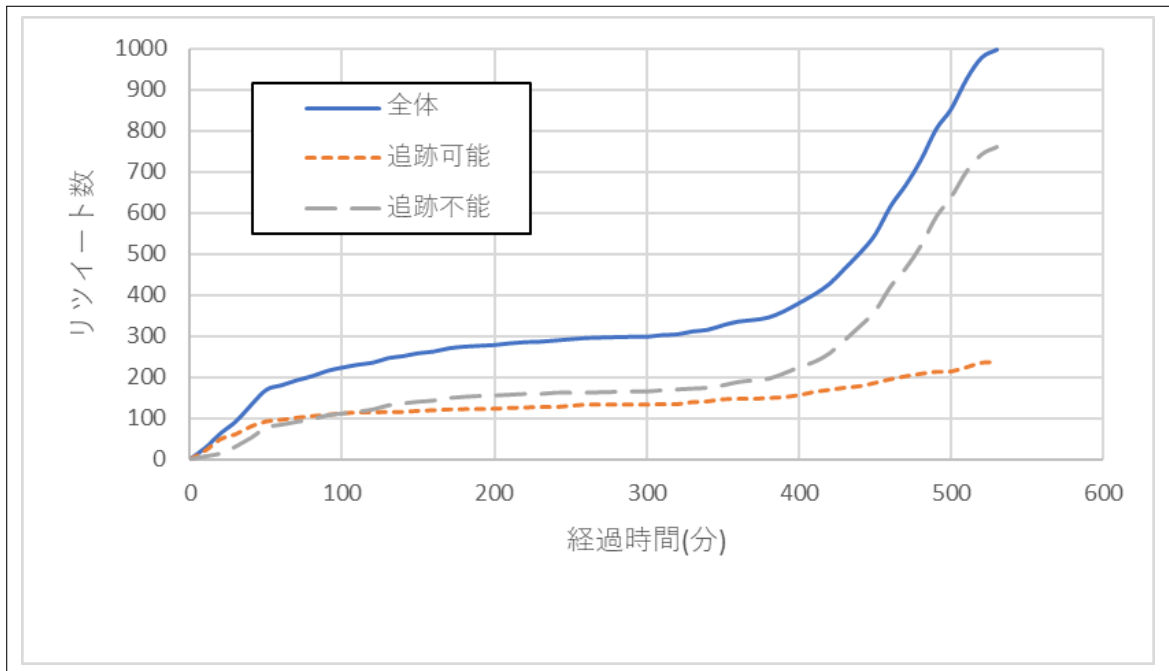


図 4.17: リツイート数時間推移：いい風呂の日

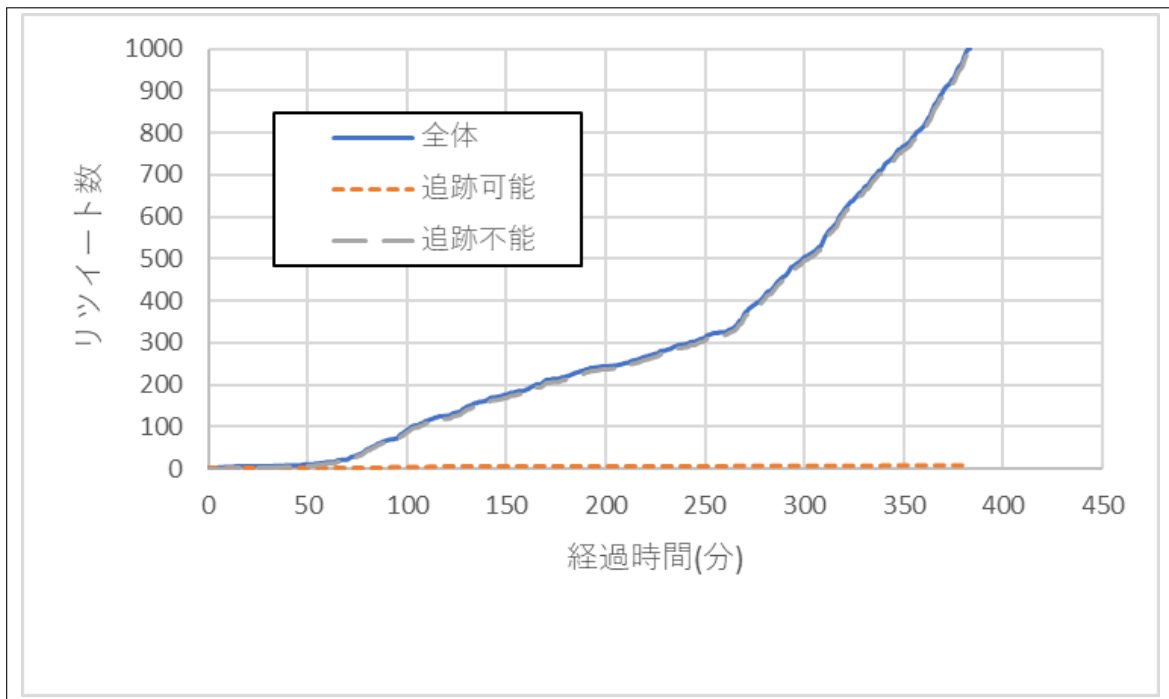


図 4.18: リツイート数時間推移：首里城

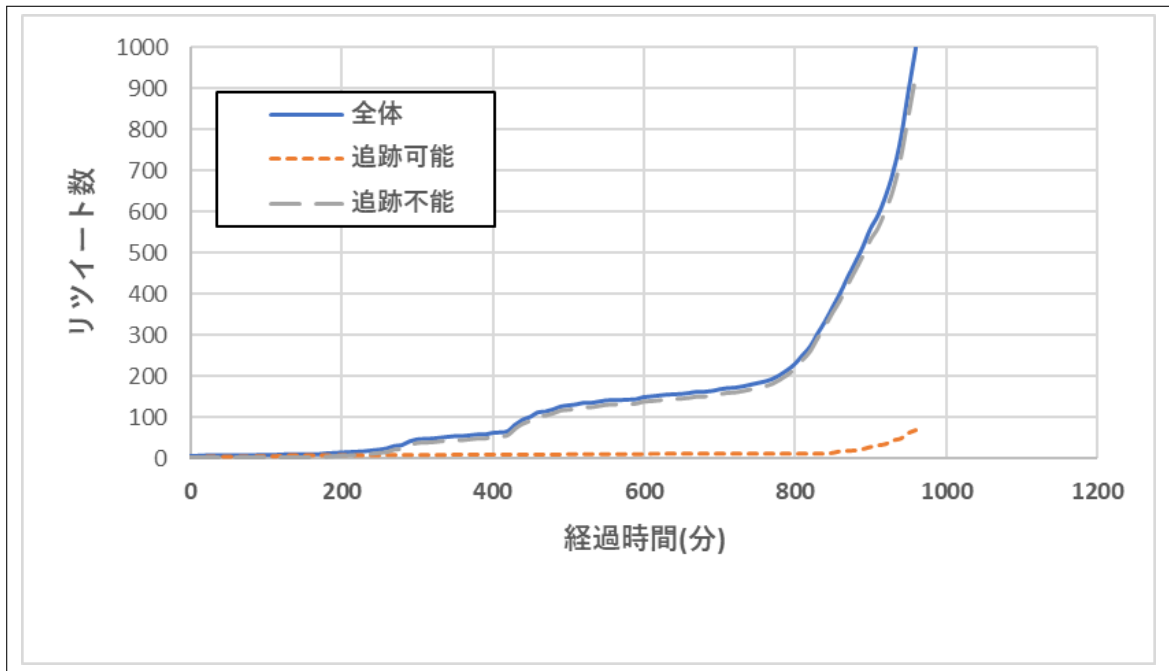


図 4.19: リツイート数時間推移：沢尻エリカ

千葉駅前などで冠水「足が浸かるくらい…」 大雨で報告相次ぐ

2019/10/25 18:24



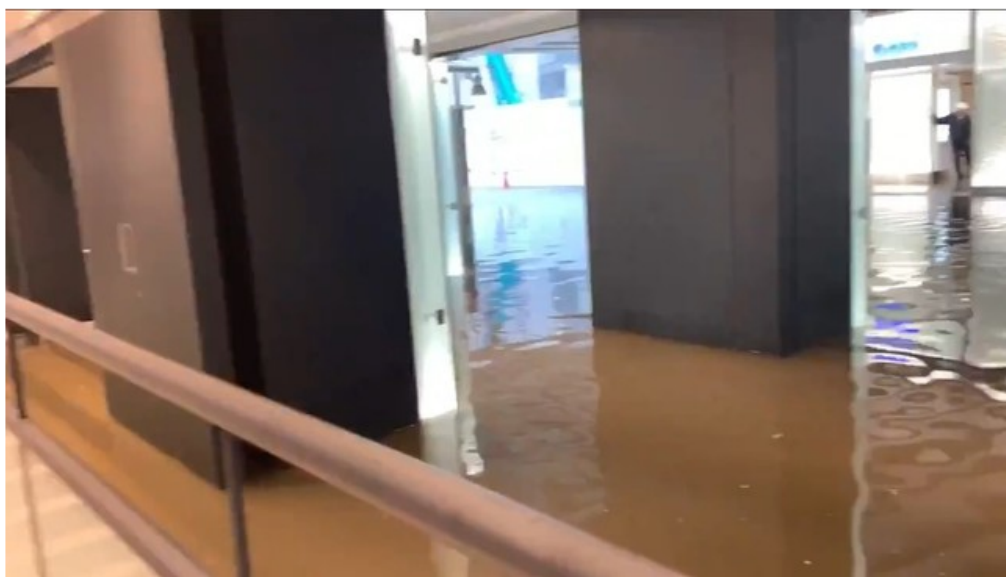
コメントを読む



PR! [エーザイが開発したサプリ『美チョコラ』無料モニター募集!](#)

気象庁は2019年10月25日、千葉県での記録的短時間大雨情報を発表した。千葉市や八街市では13時30分までに約100ミリの雨が降った。

千葉県内の駅周辺では、冠水の報告がネット上で相次いだ。J-CASTニュースでは、冠水の現場に居合わせた人らに話を聞いた。



動画の一場面 (KOUSUKE (@ok_ok_0904) さん提供)

図 4.20: ニュースサイトによる引用：千葉駅

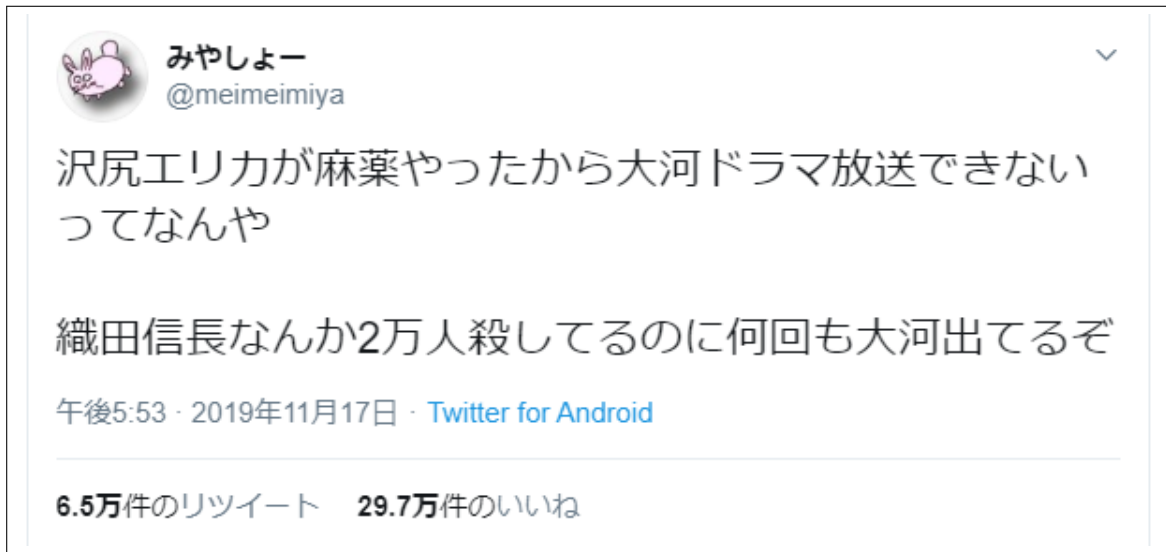


図 4.21: 「沢尻エリカ」のトップツイート

また、「天皇賞」「いい風呂の日」のツイートにおいて、獲得したリツイート数が上位 10 位であったユーザーの「獲得したリツイート数」「(オリジナルツイートが投稿されてからリツイートするまでの)経過時間」「リツイート経路の追跡可否」を表 4.14 と表 4.15 に示す。

表 4.14: リツイート数上位 10 人 : 天皇賞

獲得リツイート数	経過時間	経路の追跡可否
79	0:26:09	否
63	0:26:04	否
54	0:15:25	否
49	0:29:15	可
40	0:02:18	可
33	0:37:52	可
33	0:02:08	可
28	0:00:00	可
27	0:40:22	可
26	0:39:36	可

表 4.15: リツイート数上位 10 人 : いい風呂の日

獲得リツイート数	経過時間	経路の追跡可否
148	0:00:00	可
18	0:09:45	可
16	7:40:34	否
9	6:49:04	否
9	1:32:36	可
9	0:06:39	可
7	0:20:52	可
7	7:01:14	否
6	8:05:28	否
6	0:22:55	否

「天皇賞」の事例では、最もリツイートを獲得したユーザーはツイートの投稿者ではなかった（8 番目）。獲得リツイート数が 1-3 番目のユーザーが追跡不能な経路でリツイートを受け取っていることから、（全体としてはフォロワー/フォロワーネットワーク内のリツイートが多数を占めるものの）フォロワー/フォロワーネットワーク内でリツイートが増幅しているとはいえ、インフルエンサーの影響は小さい。認証済みのアカウントから発信されるツイートのケースとは違い、「最多リツイート獲得者がツイートの投稿者でない」「10 人の獲得リツイート数はその他のユーザーと大きく差がない」という特徴が見られる。「いい風呂の日」の事例ではトップツイートのリツイート獲得数が 148 件であるが、それ以降のリツイートについては獲得リツイート数が 6~18 件にとどまっており、また実際に（追跡不能な）リツイート数が増加したのは先述の通りツイート投

稿から7時間後である。この2つの事例においても、途中で情報が広く拡散する様子は見られなかった。

一部のデータ（「アーモンドアイ」「#バルス安倍やめろ」）について、1次拡散者と追跡可能なリツイートが多いツイートについては、最終的なリツイート数が数千件ほどに留まっているという傾向が見えた。そこで、キーワード「千葉駅」で取得したデータについて、リツイート数が100件から84件のツイートについてリツイート経路を分析した結果を表4.16に示す。

表 4.16: 千葉駅：リツイート数 100 件～84 件

リツイート数	フォロワー数	取得できたリツイート数	1次拡散者	追跡不能者
100	97520	74	41	21
98	26	67	0	67
97	159549	82	38	39
96	1728	77	20	50
95	82	83	0	83
93	20013	79	60	13
91	97520	62	36	19
89	159548	78	41	23
86	10051	66	38	15
84	196	78	1	76

追跡可能なリツイートが全リツイート数の半数を割っている、または存在しないリツイート経路が複数あり、「リツイート数が少ないツイートであっても、追跡可能なリツイートが多数を占める」とは言いえないことがわかった。

4.4.2 非認証の企業・ニュースサイト

Twitter社の認証を受けていない企業やニュースサイト、まとめサイトが発信したツイートのリツイート数、対象ツイート者のフォロワー数、1次拡散者(オリジナルツイート者から直接情報を受け取り、リツイートを受け取ったユーザー)の人数、追跡不能(情報元を辿ってもオリジナルツイート者に辿り着かない)ユーザーの人数を表4.17に示す。

表 4.17: 非認証の企業・ニュースサイトのリツイート経路

キーワード	リツイート数	フォロワー数	1次拡散者	追跡不能者
モバイル Suica	1427	3376	74	480
#アナ雪2と未知の旅へ	10229	113890	365	539
今年の漢字 OR #今年の漢字	11314	77178	988	9
武漢 OR #武漢	12562	44706	428	120
飯塚幸三 OR #飯塚幸三	16094	72545	161	811
#地味ハロウィン	36516	81926	145	764

また、各ツイートについて、リツイート数上位10人中に追跡不能なユーザーが何人いるか、またリツイー

トが 1000 件に達するまでの時間を表 4.18 に示した。

表 4.18: 非認証の企業・ニュースサイトの RT 経路 (RT 上位 10 人, 1000RT 到達時間)

キーワード	RT 上位 10 人中の 追跡不可能者	1000RT 到達時間 (分)
モバイル Suica	2	180
#アナ雪 2 と未知の旅へ	7	106
今年の漢字 OR#今年の漢字	0	12
武漢 OR#武漢	0	51
飯塚幸三 OR#飯塚幸三	8	31
#地味ハロウィン	8	32

「モバイル Suica」の事例では、リツイート数が 1427 件にとどまっていた。リツイート 1000 件のうち半数以上が追跡可能であること、1 次拡散者が 74 人であることから、フォロワー/フォロワーネットワーク内における 2 次, 3 次の拡散が見られる。「#アナ雪 2 と未知の旅へ」の事例は映画「アナと雪の女王 2」の感想についての漫画を漫画家が投稿したものであるが、ウォルト・ディズニー・ジャパンによるステルスマーケティング (ステマ) が問題となった。これ以外にも 6 人の漫画家が 12 月 3 日に感想漫画を投稿しており、漫画の内容などからステマが疑われた。12 月 5 日, 11 日にウォルト・ディズニー・ジャパンが謝罪文を発表し (図 4.22, 図 4.23), 報酬が支払われたうえでのマーケティング施策であったことが明らかになっている [27]。投稿者のフォロワー数が 10 万人を超えているが、追跡不能なリツイートが半数以上である (フォロワー/フォロワーネットワーク内の拡散が大多数でない) ことから、「複数人が漫画をほぼ同時に投稿→Twitter トrend入りする→ユーザーが漫画を読む」という流れでリツイート数が増加したと考えられる。「今年の漢字」の事例では、先述の「サイバermanデー」におけるユナイテッド・シネマ社の公式アカウントの投稿のように「アカウントをフォローした上でリツイートすると抽選でプレゼントが当たる」キャンペーンが行われており (図 4.24), 最初のリツイート 1000 件がほぼ 1 次拡散で占められているのはそれが理由であると考えられる。「飯塚幸三」の事例では、2019 年 4 月に東京・池袋で交通事故を起こし 12 人を死傷させた飯塚幸三容疑者がメディアの取材に応じたことを取り上げたまとめサイトのアカウント (図 4.25) がトップツイートを獲得した。最初のリツイート 1000 件のうち 8 割以上のリツイートが追跡不能であることから、Twitter トrendやニュースサイト経由での情報拡散が考えられる。

2019.12.05

「『アナと雪の女王2』感想漫画企画」に関するお詫び

この度は、「『アナと雪の女王2』感想漫画企画」につきまして、ご参加いただきましたクリエイターのみなさま、ファンのみなさまに多大なご心配、ご迷惑をお掛けし、深くお詫び申し上げます。

本企画は、クリエイター7名のみなさまに映画『アナと雪の女王2』をご覧いただき、ご感想を自由に表現いただいた漫画をTwitterに投稿いただく企画として実施したものです。

本企画に伴う投稿は、「PR」であることを明記していただくことを予定しておりましたが、関係者間でのコミュニケーションに行き届かない部分があり、当初の投稿において明記が抜け落ちる結果となってしまいました。

今後このような事がないよう、関係者一同、深く反省するとともに、コミュニケーション体制を見直し、再発防止策を講じてまいります。

ウォルト・ディズニー・ジャパン株式会社

図 4.22: アナ雪 2 謝罪文 (12月5日)

2019.12.11

「『アナと雪の女王2』感想漫画企画」 にご参加いただいたクリエイターのみな さま、そしてファンみなさまへ

私たちは、「『アナと雪の女王2』感想漫画企画」に関し、ご参加いただいたクリエイターのみなさまにご迷惑をお掛けしている事実を大変厳粛に受け止めています。また、この事により、ファンみなさまを失望させてしまったことを真摯に受け止めています。

ディズニーでは、マーケティング活動における社内指針を設けています。本件を含む類似の案件は、当該指針に関する周知および遵守の不徹底が招いた結果であり、ご参加いただいたクリエイターのみなさまに責任はございません。

改めまして、本件につきまして、深くお詫び申し上げます。今後は、このような事がないよう、社内指針の周知徹底を図り、再発防止に努めてまいります。これまでみなさまにいただいたディズニーに対する思いを心に留め、社員一同努力し続けてまいります。

ウォルト・ディズニー・ジャパン株式会社

図 4.23: アナ雪2 謝罪文 (12月11日)



🍇元祖ぶどうの宝石箱🍇伊藤ぶどうファーム三... · 2019年12月12日

#今年の漢字

は

【ぶどうの宝石箱】

で決まりっ♪笑

という事で
第14回Xmasプレゼント企画を開催中!!

☆応募方法
フォロー &
このツイートをRT

1 3日の金曜日 2 4時まで

抽選で3名様に

【ぶどうの宝石箱 Xmasパッケージ】

1 1 1 1 1円相当を

無料で
クリスマスプレゼント致します🍇🎁



884

1.1万

3,533



図 4.24: 「今年の漢字」のトップツイート



図 4.25: 「飯塚幸三」のトップツイート

「#アナ雪 2 と未知の旅へ」「飯塚幸三」の事例について、リツイート数の時間推移を図 4.26 と図 4.27 に示す。

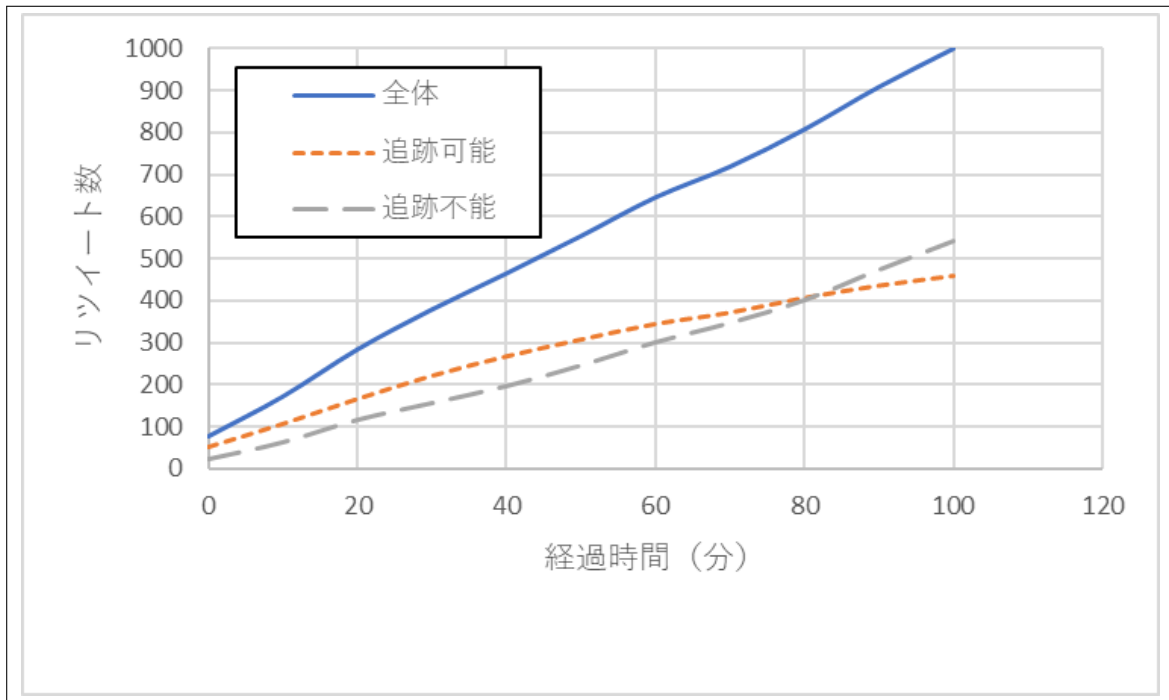


図 4.26: リツイート数時間推移：#アナ雪2と未知の旅へ

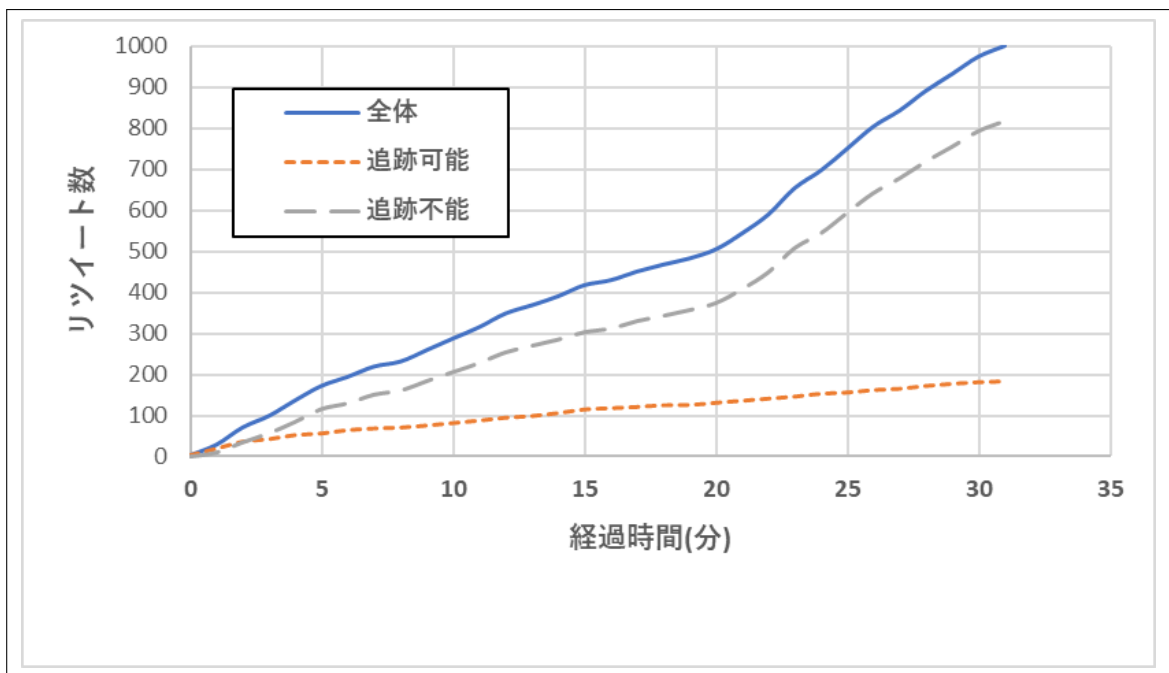


図 4.27: リツイート数時間推移：飯塚幸三

「#アナ雪2と未知の旅へ」の事例では、ツイート投稿から約80分で追跡不能なリツイートが追跡可能なリツイートを上回っている。これは先述のようにステルスマーケティングが行われ、トレンドが形成されたことによるものと考えられる。「飯塚幸三」の事例では、投稿から約20分で追跡不能なリツイートの増加速度が

上がっている。図 4.25 のように、テレビの報道によって Twitter ユーザーが大量に投稿を行い、その光景を取り上げたニュースサイト・まとめサイトのアカウントがトレンドに乗ってリツイートを集めるという流れが見られる。

「#アナ雪 2 と未知の旅へ」「飯塚幸三」のツイートにおいて、獲得したリツイート数が上位 10 位であったユーザーの「獲得したリツイート数」「(オリジナルツイートが投稿されてからリツイートするまでの) 経過時間」「リツイート経路の追跡可否」を表 4.19 と表 4.20 に示す。

表 4.19: リツイート数上位 10 人 : #アナ雪 2 と未知の旅へ

表 4.20: リツイート数上位 10 人 : 飯塚幸三

獲得リツイート数	経過時間	経路の追跡可否
365	0:00:00	可
39	0:13:35	可
14	0:20:22	否
9	0:17:31	否
8	0:05:58	否
7	0:21:07	否
7	0:15:50	可
7	0:33:03	否
6	1:03:18	否
6	0:20:59	否

獲得リツイート数	経過時間	経路の追跡可否
482	0:00:36	否
161	0:00:00	可
13	0:22:44	否
11	0:03:29	可
4	0:10:16	否
3	0:03:00	否
3	0:25:42	否
3	0:22:29	否
2	0:10:49	否
2	0:09:18	否

「#アナ雪 2 と未知の旅へ」のケースでは、投稿されたツイートが最もリツイートを獲得しているが、約 15 ~ 20 分経過すると追跡不能な経路からのリツイートが見られる。「飯塚幸三」のケースでは、対象のツイートのリツイート獲得数は 2 番目であり、1 番目は「平成を忘れない bot」(図 4.28) による拡散であった。「平成を忘れない bot」はツイート投稿者から直接ツイートを受け取っておらず、またリツイートまでの経過時間も 36 秒と短いため、Twitter トレンドを見て当該ツイートに行きついた、もしくはキーワード検索を行ってリツイートを行ったと考えられる。「平成を忘れない bot」は約 47 万人のフォロワーを有しており、実際に 1000 件のリツイート中で半数近くのリツイートに関与しているという意味ではインフルエンサーであるといえるが、あらかじめフォロワー/フォロワーネットワークに組み込まれているユーザーではない(「平成を忘れない bot」の bot プログラムに取り上げられるためにはその前に話題になっていなければならない) ことに留意する必要がある。



図 4.28: 平成を忘れない bot

4.5 非公開アカウント、追跡不能なリツイート数と情報拡散速度

本節では、非公開アカウントがリツイートに与える影響、追跡不能なリツイート数が情報拡散速度に与える影響について考察する。表 4.21 に各キーワードごとのトップツイートについての非公開アカウントの割合、追跡不能リツイート数、1000 リツイート到達時間を示す。

表 4.21: 非公開アカウントの RT 割合, 追跡不能 RT 数, 1000RT 到達時間

キーワード	非公開アカウントによる RT 数の割合	追跡不能な RT 数	1000RT 到達時間 (分)
今年の漢字 OR #今年の漢字	0.038625	9	12
サイバーマンデー OR #サイバーマンデー	0.041085	12	5
香港デモ OR #香港デモ	0.088891	72	24
GSOMIA OR #GSOMIA	0.097097	78	8
#プレイステーションの日	0.481616	88	47

表 4.21: 非公開アカウントの RT 割合, 追跡不能 RT 数, 1000RT 到達時間

キーワード	非公開アカウントによる RT 数の割合	追跡不能な RT 数	1000RT 到達時間 (分)
流行語 OR #流行語大賞	0.105177	117	63
武漢 OR #武漢	0.321048	120	51
#バルス安倍やめろ	0.302235	183	240
アンタッチャブル	0.284974	195	17
中曽根	0.171531	200	2
松本零士 OR #松本零士	0.170589	242	12
ゴロゴ OR #ゴロゴ	0.377875	243	22
梅宮	0.227502	276	1.67
アーモンドアイ	0.152994	301	23
イラン AND 撃墜	0.189803	339	13
天皇賞	0.195552	373	80
札幌 disOR #札幌 dis	0.140459	401	66
南アフリカ AND ラグビー	0.197250	433	0.63
PayPay ドーム OR ペイペイドーム	0.273034	453	7
モバイル Suica	0.015417	480	180
グレタさん OR#グレタさん	0.223162	503	120
#アナ雪 2 と未知の旅へ	0.293382	539	106
ヘイトスピーチ OR #ヘイトスピーチ	0.339423	588	840
京アニ放火 OR #京アニ放火	0.191541	616	400
首吊り自殺 OR #首吊り自殺	0.205533	730	210
いい風呂の日 OR #いい風呂の日	0.302235	760	530
#地味ハロウィン	0.391198	764	32
ローマ法王 OR ローマ教皇	0.313509	772	240
飯塚幸三 OR#飯塚幸三	0.246800	811	31
即位パレード OR#即位パレード	0.195496	846	27
アナ雪 OR ステマ	0.211488	875	30
なでしこ寿司 OR#なでしこ寿司	0.241693	888	380
ミルクボーイ OR#ミルクボーイ	0.251602	895	6
沢尻エリカ OR#沢尻エリカ	0.273411	908	960
忘年会スルー OR#忘年会スルー	0.359524	912	560
Mac ProOR#Mac Pro	0.269546	933	280
ホームアローン OR #ホームアローン	0.205004	941	800
ターミネーター OR #ターミネーター無双	0.211037	946	190
成人式 OR #成人式	0.305346	961	400
決闘罪	0.240183	974	220
地味ハロウィン	0.375169	975	420
次元の声	0.204287	977	45

表 4.21: 非公開アカウントの RT 割合, 追跡不能 RT 数, 1000RT 到達時間

キーワード	非公開アカウントによる RT 数の割合	追跡不能な RT 数	1000RT 到達時間 (分)
ドイツ鉄道 OR #ドイツ鉄道	0.193701	989	280
首里城	0.136048	991	380
二宮 OR ニノ AND 結婚 OR#二宮結婚 OR#ニノ結婚	0.401899	991	60
Airpods Pro	0.175820	995	240
千葉駅	0.198043	999	60

4.5.1 非公開アカウントの影響

非公開アカウントの割合は $1 - (\text{TwitterAPI で取得されたリツイート数} / \text{実際のリツイート数})$ で算出した。非公開アカウントがツイートをリツイートしたというデータは TwitterAPI では取得できない。例えば、20000 件のリツイートを獲得したツイート A について TwitterAPI でリツイートユーザーのデータを取得した際に 15000 件しか取得できなかったとする。この場合、ツイート A のリツイートに含まれる非公開アカウントのリツイートの割合は $1 - (15000 / 20000) = 0.25$ と推測できる。今回取得したデータセットについて、非公開アカウントによるリツイートの割合の平均は 0.230401 であった (標準偏差は 0.097950, 中央値は 0.211488)。つまり、(単純に考えて) 1000 件のリツイートに対して非公開アカウントによるリツイートが 230 件ほど含まれているといえる。「情報元が不明なリツイート」は平均で約 315 件存在し、追跡不能なリツイート数の平均は約 590 件存在すること、また非公開アカウントは Twitter ユーザーの約 3 割存在し、フォロワー数は少ない [28] という特徴を考慮すると、非公開アカウントが情報拡散に与える影響は大きくないと推測できる。非公開アカウントが最初の 1000 件のリツイートの中にほぼ同じ割合で含まれるとすると、非公開アカウントが情報拡散に与える影響はどのツイートのリツイート経路でもほぼ同じになり、追跡不能なリツイートが多数を占める経路で情報拡散の内容に差をもたらすのは別の要因であるといえる。

4.5.2 追跡不能なリツイート数と情報拡散速度

トップツイートに対する最初の 1000 件のリツイートについて、「追跡不能なリツイート数」と「1000 リツイートに到達するまでの時間」との相関係数は 0.472938 であり、高い相関を示した。この結果から、追跡不能なリツイートが多い (非公開アカウントや Twitter トレンド, 外部サイトによる情報拡散が起こっている) ツイートは情報が拡散するのに時間がかかる, といえる。また、フォロワー数と 1000 リツイート到達時間の相関係数は -0.306455 である。ツイートの投稿者が認証済みの場合、リツイートの多くが追跡可能であり情報拡散が速やかに行われる傾向にある。これは、アカウントから発信される情報、そしてアカウント自体にコンテンツとしての力がありユーザーの信頼があるからだと推測できる。ニュースサイトなどから発信された情報がそのフォロワーによって (精査されずに) 速やかに拡散される傾向が見られる。認証のないアカウントからの情報に対しては、他のソース (マスメディアなど) である程度裏付けられたものが Twitter トレンドの流れに乗って拡散されると考えられる。今回取得したデータについては、ステルスマーケティングを除くと誤情報、フェイクニュースは存在しなかった。認証済みアカウントのツイートであっても非認証アカウントのツ

イトであっても、フォロワー/フォロワーネットワーク内で中間のユーザーによる広い情報拡散の顕著な傾向は見られなかった。ツイート投稿者の属性、もしくは Twitter トレンドや外部サイトなどのフォロワー/フォロワーネットワーク「外」の要因がリツイート数増加に重要な役割を果たしていると考えられる。

第 5 章

結論

本章では本論文の総括を行い、また本研究によって発見された新たな課題や方針を示す。

本研究では TwitterAPI を用いて、Twitter 上で話題になったトレンドワードに着目してツイートデータを大量に取得し、リツイートによって情報が拡散していく経路を分析した。その際、ユーザー間のフォロワー/フォロワー関係に着目したフォロワー/フォロワーネットワークを作成し、リツイート数の増加要因について考察を行った。その結果、以下の特徴が明らかになった。

1. 各キーワードについて、最もリツイートを獲得したツイート（トップツイート）に対して行われた最初の 1000 件のリツイートにおいて、直接の情報元が不明なリツイートが平均して 300 件ほど存在する。また、リツイート経路が追跡不能な（情報元を辿ってもオリジナルツイートに辿り着かない）リツイートは平均して 600 件ほど存在する。
2. トップツイートの投稿者のフォロワー数が少ないほど、リツイート受信経路が追跡不能なリツイートの割合は大きくなる。
3. トップツイートの投稿者のフォロワー数が少ないほど、情報の拡散速度が遅くなる（リツイートが 1000 件に達するまでの時間が長くなる）。
4. トップツイートの投稿者が認証済みか否かによって、情報拡散の過程に違いが見られる。

トップツイートのリツイート経路には追跡不能なリツイートが多数存在し、Twitter トレンドや外部ウェブサイトなどのフォロワー/フォロワーネットワーク「外」で生じる情報の拡散が、特にフォロワー数が少ない投稿者のツイートが広く拡散される上で本質的に重要であることを示した。また、リツイートが拡散する過程においては最初の投稿者の影響が最も大きく、情報拡散の途中で追跡可能なリツイートが急激に増加する傾向は見られないことから、リツイートの拡散において「インフルエンサー」の影響は限定的であるといえる。しかし、本研究で用いた方法では、情報元が不明であるということが判明したとしても、(1) 非公開アカウントを介した拡散 (2) 「いいね」のシェアによる拡散 (3) フォロワー/フォロワーネットワーク外の経路での拡散のどのケースに当たるのかを特定することはできない。今後の課題としては、不明な情報元が (1) (2) (3) のどれに当たるのか、また (3) のケースの中で、その情報元が Twitter トレンド・ニュースサイト・まとめサイト等のどの要因に該当するのかを特定することが検討される。

参考文献

- [1] 総務省 令和元年度版情報通信白書 第4節 デジタル経済の中でのコミュニケーションとメディア
<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r01/pdf/n1400000.pdf> 2020/01/23 アクセス
- [2] Twitter のトレンドについてのよくある質問 <https://help.twitter.com/ja/using-twitter/twitter-trending-faqs>
2020/01/13 アクセス
- [3] Twitter トренд大賞 <https://trendaward.jp/> 2020/01/13 アクセス
- [4] Togetter <https://togetter.com/> 2020/01/25 アクセス
- [5] Twitter の API について <https://help.twitter.com/ja/rules-and-policies/twitter-api> 2020/01/25 アク
セス
- [6] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon, "What is Twitter, a Social Network or a
News Media?," *WWW 2010*, pp.591-600 (2010)
- [7] Reza Bakhshandeh, Mehdi Samadi, Zohreh Azimifar, and Jonathan Schaeffer, "Degrees of Separation in
Social Networks," *The Fourth International Symposium on Combinatorial Search*, pp.18-23 (2011)
- [8] Anatomy of Facebook, [https://www.facebook.com/notes/facebook-data-team/anatomy-of-
facebook/10150388519243859](https://www.facebook.com/notes/facebook-data-team/anatomy-of-facebook/10150388519243859) 2020/1/23 アクセス
- [9] Jon Kleinberg, "The Small-World Phenomenon: An Algorithmic Perspective," *Cornell Computer Science
Technical Report*, 99-1776 (1999)
- [10] Sitaram Asur, Bernardo A. Huberman, Gabor Szabo, and Chunyan Wang, "Trends in Social Me-
dia: Persistence and Decay," *The Fifth International AAAI Conference on Weblogs and Social Media*, pp.434-
437 (2014)
- [11] Kaj-Kolja Kleineberg and Marian Boguna, "Evolution of the digital society reveals balance between viral
and mass media influence," *PHYSICAL REVIEW X*, 4, 031046, (2014)
- [12] Xin Zhang, Ding-Ding Han, Ruiqi Yang, and Ziqiao Zhang, "Users' participation and so-
cial influence during information spreading on Twitter," *PLoS ONE* 12(9): e0183290,
<https://doi.org/10.1371/journal.pone.0183290> (2017)
- [13] Eytan Bakshy, Jake M. Hofman, Winter A. Mason, and Duncan J. Watts, "Everyone's an Influencer: Quan-
tifying Influence on Twitter," *WSDM'11: Fourth ACM International Conference on Web Search and Data
Mining*, pp.65-74, (2011)
- [14] Jane F. Bokunewicz and Jason Shulman, "Influencer identification in Twitter networks of destination mar-
keting organizations," *Journal of Hospitality and Tourism Technology*, 8(2), pp.205-219, (2017)
- [15] Soroush Vosoughi, Deb Roy, and Sinan Aral, "The spread of true and false news online," *Science*, Vol. 359,
pp.1146-1151, (2018)
- [16] 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: 東日本大震災に関する事例," *人工知能学会論文誌*, vol.31, no.1, pp.NFC-A-1-9, (2016)

- [17] Fabian Schäfer, Stefan Evert, and Philipp Heinrich, "Japan's 2014 General Election: Political Bots, Right-Wing Internet Activism, and Prime Minister Shinzō Abe's Hidden Nationalist Agenda," *Big Data*, Vol. 5, pp.294-309 (2017)
- [18] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo, "Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors," *Proc. 18th Int. World Wide Web Conference, 2010*, pp.851-860, (2010)
- [19] Sasa Petrovic, Miles Osborne, and Victor Lavrenko, "RT to Win! Predicting Message Propagation in Twitter," *Proceedings of the Fifth International AAI Conference on Weblogs and Social Media*, pp.586-589 (2011)
- [20] Luke Sloan, Jeffrey Morgan, Pete Burnap, and Matthew Williams, "Who Tweets? Deriving the Demographic Characteristics of Age, Occupation and Social Class from Twitter User Meta-Data," *PLoS ONE 10(3): e0115545*. doi:10.1371/journal.pone.0115545 (2015)
- [21] Johan Bollen, Huina Mao, and Alberto Pepe, "Determining the public mood state by analysis of microblogging posts," *Proc. of the Alife XII Conference, Odense, Denmark, 2010*, pp.667-668, (2010)
- [22] 鳥海不二夫, Twitter 上のビッグデータ収集と分析, 組織科学, vol.48, No.4, pp.47-59 (2015)
- [23] 塩田茂雄, 南川雅人, 中島圭佑, "キーワード検索で収集される Twitter データの特徴と Twitter 上の情報拡散過程," 電子情報通信学会情報ネットワーク研究会, *IN2018-64*, pp. 31-36, (2018).
- [24] TweetDeck <https://tweetdeck.twitter.com/> 2020/01/13 アクセス
- [25] 認証済みアカウントについてのよくある質問 <https://help.twitter.com/ja/managing-your-account/twitter-verified-accounts> 2020/01/25 アクセス
- [26] 千葉駅前などで冠水「足が浸かるくらい…」大雨で報告相次ぐ : J-CAST ニュース <https://www.j-cast.com/2019/10/25371034.html> 2020/01/29 アクセス
- [27] ITmedia NEWS ディズニー、アナ雪 2 のステマ騒動で再びコメント「クリエイターに責任はない」 <https://www.itmedia.co.jp/news/articles/1912/11/news103.html> 2020/01/29 アクセス
- [28] 日本人ツイッターユーザー調査 2016 年版 <http://facebook.boo.jp/twitter-user-survey-2016> 2020/01/29 アクセス

謝辞

本研究を進めるに当たり，指導教官の塩田茂雄教授からは多大な助言を賜りました。また，ゼミを通じて数々の助言を下された塩田研究室の皆様にも大変お世話になりました。厚く御礼を申し上げ，感謝する次第です。本当にありがとうございました。

研究成果

1. 中島圭佑, 南川雅人, 塩田茂雄, ”SNS における投稿件数推移分析のための情報拡散モデル,” 電子情報通信学会 コミュニケーションクオリティ研究会, CQ2017-84, 2017 年 11 月 17 日 (高松) .
2. 南川雅人, 中島圭佑, 塩田茂雄, ”強相関近似による複雑ネットワーク上の情報拡散過程の解析,” 2017 年度待ち行列シンポジウム「確率モデルとその応用」, pp. 194-195, 2018 年 1 月 19 日 (大阪) .
3. 南川雅人, 中島圭佑, 塩田茂雄, ”非反応状態を考慮した情報拡散モデルによる SNS 投稿数推移過程の解析,” 電子情報通信学会 第 5 回コミュニケーションクオリティ (CQ) 基礎講座ワークショップ, 2018 年 1 月 20 日 (東京) .
4. 塩田茂雄, 南川雅人, 中島圭佑, ”SNS 投稿件数推移分析のための情報拡散モデルと強相関近似解析,” 第二回計算社会科学ワークショップ, 2018 年 3 月 (東京) .
5. 塩田茂雄, 南川雅人, 中島圭佑, ”SNS 上の情報拡散過程と投稿数の時間推移のモデル化,” 電子情報通信学会 ソサイエティ大会, B-7-38, 2018 年 9 月 (金沢) .
6. 南川雅人, 中島圭佑, 塩田茂雄, ”SNS 上の情報拡散過程と投稿数の時間推移のモデル化,” 第一回 QoS に関する (名大, NII, 早稲田大, 電通大, 芝浦工大, 千葉大合同) 学生技術交流会, 2018 年 10 月 (東京) .
7. 塩田茂雄, 南川雅人, 中島圭佑, ”キーワード検索で収集される Twitter データの特徴と Twitter 上の情報拡散過程,” 電子情報通信学会 情報ネットワーク研究会, IN2018-64, pp. 31-36, 2018 年 12 月 (広島) .
8. 南川雅人, 中島圭佑, 塩田茂雄, ”Twitter データの特徴分析と人間の行動モデル,” 電子情報通信学会 ネットワークシステム研究会, 2019 年 3 月 (沖縄) .
9. 南川雅人, 中島圭佑, 塩田茂雄, ”リツイート数分布のべき則性の出現メカニズム,” 電子情報通信学会 総合大会, 2019 年 3 月 (東京) .
10. 塩田茂雄, 中島圭佑 ” Twitter データに見られる特徴と人間のリツイート行動,” 日本人工知能学会全国大会, 2019 年 6 月 5 日 (新潟) .
11. Shigeo Shioda and Keisuke Nakajima, ”Information spread across social network services with users’ information indifference behavior,” 11th Computer Science and Electronic Engineering Conference, 2019.
12. 中島圭佑, 塩田茂雄, ”ネットワーク炎上の事例分析：最近の Twitter での事例を元に,” 電子情報通信学会 ソサイエティ大会, B-11-8, 2019 年 9 月 11 日 (大阪) .
13. 中島圭佑, 塩田茂雄, ”ネットワーク炎上の事例分析：最近の Twitter での事例を元に,” (名大, 早大, 電通大, 広島市立大, NII, 千葉大合同) 第 2 回 学生技術交流会, 2019 年 10 月 15 日 (NII, 東京) .