

ウェブメディアがSNSに与える影響の調査

—2019年台風19号のツイートデータを対象に—

森野 穰[†] 安尾 萌[†] 松下 光範[†] 藤代 裕之^{††}

[†] 関西大学総合情報学部総合情報学科 〒569-1095 大阪府高槻市霊仙寺町 2-1-1

^{††} 法政大学社会学部メディア社会学科 〒194-0298 東京都町田市相原町 4342

E-mail: [†]{k790414,k290993,t080164}@kansai-u.ac.jp, ^{††}fujishiro@hosei.ac.jp

あらまし 本研究は、災害時の情報収集手段として SNS を活用可能にすることを旨とし、災害発生時に SNS へ投稿された情報を分析することで、情報爆発の原因を明らかにすることを目的とする。本稿では Web メディアが情報爆発に与える影響に着目し、2019 年における台風 19 号上陸時のツイートを対象とした調査を行った。分析の観点として、災害に関するツイートの総数と、Web メディアが発信したニュース記事へ遷移する URL を含むツイートの単位時間あたりの投稿数を比較し、ツイートが急増する時間帯における出現単語とその頻度を分析した。分析の結果、情報爆発が起きた時間帯におけるツイート数と、ニュース記事への URL を含むツイート数の推移には相関関係が見られた。さらに、ニュース記事への URL を含むツイートの情報源を調査した結果、特定の Web メディアや特定のニュース記事への URL が多く投稿される傾向が見られた。

キーワード ツイッター, 災害情報, 影響調査

1 はじめに

昨今、インターネットやスマートフォンの普及と共に発展を遂げたオンライン・ソーシャル・ネットワーキング・サービス(以下、SNS と記す)では個人による情報発信が盛んに行われている。発信された情報から実世界で起きた出来事を抽出・利用する研究が進んでいる。Kwak の SNS の情報伝播とその拡散に関する定量的な調査 [1] をはじめとし、McCormick は SNS で得られる情報を用いて人口統計情報の推定 [2] を行うなど、SNS には実世界における出来事の確認・調査ができる可能性がある。また、東日本大震災を機に、SNS を用いた災害被害や救援に関する情報発信が増え、個人が発信した情報を活用する動きが注目されている。SNS 情報の有効な活用の可能性として、被災者が SNS に投稿した情報に基づく人命救助や、有識者が SNS に投稿した災害解説やメディア報道を含む投稿のアーカイブを構築し、減災・防災に繋げる動きが挙げられる。一方で、SNS を大規模災害時の情報収集手段として活用する際の課題は、情報爆発とデマへの対応が必要であると指摘されている [6]。情報爆発により情報量が増加し、収集すべき情報の相対的な割合が低下することで情報収集がより困難となる [7]。

一般に、SNS における情報爆発の原因は、誰もが即時的な情報発信および拡散が可能であるという特性にある。特に、災害時においてはメディアの持つ強い発信力が情報爆発を助長させる懸念がある。こうした懸念を検証するため、本研究では Web メディアが情報爆発に与える影響に着目し、情報爆発の原因を明らかにすることを目的とする。その端緒として、本稿では代表的なマイクロブログ型 SNS である Twitter¹ を取り上げ分析

対象とし、災害時のツイート情報における Web メディアが発信したニュース記事へ遷移する URL を含むツイートの定量的調査を行った。

2 関連研究

2.1 災害時のソーシャルメディア分析

榎本らは災害時の Twitter 情報を用いてテキスト情報が人々の帰宅行動にどのように影響しているのか、定量的に調査を行った [9]。榎本らは、「既存研究は現状把握に留まっており次なる災害が訪れた時の対策までは言及されていない」と述べ、膨大なテキストデータの中から減災に有効な可能性がある情報の抽出を試みた。まず、2 つの変数群の相関性を把握するための手法である正準相関分析を用いて単語間のネットワークを時系列で分割した。ただし、正準相関分析を用いた双方向性における因果関係を特定することの困難性は、あらかじめ 3 つのケースを想定しそれらを比較することで解決している。次にネットワークの類似性と中心性を用いて単語間ネットワークの時点間推移を分析した。その結果、「Twitter のテキスト情報と関係性がある場合」を想定したケースにおいて、地震発生後 2~3 時間経過すると SNS では交通に関する話題が主となったことが明らかになり、SNS 情報が帰宅行動と時間的な関係性にあることが示唆された。具体的には、交通機関や交通状況について悲観的な情報が多い時や楽観的な情報が少ない時に帰宅行動をするユーザは少なく、何も行動せず勤務地に滞在するユーザの割合が高い。これにより、情報の信頼性が何らかの要因で担保されずとも、災害時においては SNS で発信された情報は人々の行動に際する意思決定に影響し、SNS の有効活用が可能であれば、ある程度混乱を避けられる可能性がある」と述べた。ただし、

1: <https://twitter.com>

今後の課題として、長期間取得したデータを対象とした分析手法が必要になると述べている。

2.2 SNS を用いた実世界調査

SNS をソーシャルセンサとして用いて実世界を観測する研究が数多く行われている。國友らは Twitter のリアルタイム性に着目し、土砂災害の前兆となる現象や発生を訴える情報を抽出することで実世界における前兆現象等の把握を試みた [10]。まず、ツイートの市町村推定 [11] を行い、豪雨に関するツイートを状況カテゴリによる分類を行った。次に、そのツイート群の中からキーワードによる抽出をしたツイート数と、降雨量と土砂災害の発生タイミングを比較することで、ツイート情報が前兆現象の把握に有効だと示した。しかし、ツイートの推移だけでは人口規模の小さい地域ではツイート数が少ないため、地域ごとのツイート価値を評価する必要があると述べた。

Sakaki らは Twitter 上での発言を用いて実世界における台風や地震の位置を推定し [3]、Aramaki らはインフルエンザに関する発言を集めた上で、現実世界におけるインフルエンザの患者を推定した [4]。その結果、ツイート数と実際のインフルエンザ患者に大きな差異が生じた。そのため、荒巻はそれらのデータを用いて SNS と実世界の時間的ギャップの調査を行った [5]。その結果、インフルエンザのピークが過ぎると発言量が低下するのは、話題としての価値低下が考えられると述べ、流行のピーク前またはピーク中にメディアに取り上げられることが、SNS 上で急激に発言数が増える現象（バースト）が起こる原因とした。荒巻らは今後の SNS 情報解析において、SNS で情報抽出する際には人間の欲求や不正確さを解析し、心理学的知見と合わせて考察する必要があると結論付けた。

2.3 本研究の位置付け

先行研究は、災害時の SNS において発信・拡散された投稿を収集して有効利用の可能性を調査する動きを示した。しかし、SNS において SNS が持つ性質である情報発信の即時性や拡散力の強さを主な原因とする情報爆発が発生し、SNS 情報の収集がより困難になる。本研究では、情報爆発の原因の一端を明らかにする目的で、Web メディアが SNS に与える影響を定量的に調査する。具体的には、Twitter で発信された情報とネットニュース数の比較を行う。その上で出現するツイート数と単語数の推移を観測し、ツイートに含まれる URL を調査する。

3 ツイートの収集

3.1 2019 年台風 19 号

気象庁の発表によると、台風 19 号は 2019 年 10 月 6 日南鳥島の近海で発生し、12 日 19 時前に伊豆半島に上陸した。関東地方を通過後、13 日 12 時に日本の東で温帯低気圧となった²。特に関東地方、甲信地方、東北地方で大雨となり、計画降雨量を大きく上回る記録的豪雨をもたらした。政府は 11 月 1 日、

2 : <https://www.data.jma.go.jp/obd/stats/data/bosai/report/2019/20191012/20191012.html> (2019/12/13 確認)

表 1 千曲川氾濫箇所

日時	氾濫発生箇所
12 日 20:35	上田市国分地先付近
20:50	長野市篠ノ井横田地先付近
22:10	千曲市雨宮地先付近
22:40	長野市篠ノ井小森先付近
23:18	長野市松代楽地先付近
13 日 01:15	長野市穂保地先付近
02:03	須坂市北相之島地先付近および 小布施町飯田付近
02:58	小布施町山王島地先付近
03:25	中野市立ヶ花および 栗林付近

この災害に対して激甚災害および適用措置を指定する政令を交付・施行した³。国土交通省が発表した、令和元年台風第 19 号等による被害状況等について (第 50 報)⁴によると、台風 19 号における堤防決壊は 71 河川におよび、のべ 140 箇所を超えた。本研究では、河川に関するツイートを分析対象とし、「入間川」「相模川」「千曲川」に関するツイートの収集を行った。

3.1.1 入間川の状況

入間川は埼玉県を流れる荒川水系の一つである。荒川上流河川事務所と熊谷地方気象台は、12 日 21 時 30 分入間川流域で河川氾濫が発生したと発表した⁵。

3.1.2 相模川の状況

相模川は山梨県および神奈川県を流れる相模川水系の本流である。神奈川県は 12 日 17 時から城山ダムで緊急放流を行う旨を発表し、それに伴って水害が発生する恐れがあるとして住民に避難を呼びかけた。水位が予定より低い状態だったため放流時間を 22 時に延期したが、その後の降雨状況が予測を上回ったため 21 時 30 分に緊急放流を開始した⁶。

3.1.3 千曲川の状況

千曲川は新潟県および長野県を流れる信濃川水系の本流である。新潟県域では信濃川と呼ばれ、長野県域では千曲川と呼称される。千曲川河川事務所の発表⁷によると 12 日～13 日にかけて 9 度氾濫が発生した (表 1)。

3.2 ツイートの収集方法

2019 年 10 月 10 日～10 月 16 日までの期間に投稿されたツイートの収集を行った。その際、TwitterAPI を用いて各川ごとにキーワードによる収集を行った。また、リツイート (他者のツイートを二次的に再発信する性質を持つ) と呼ばれる情報発信を排除するため、検索ワードに「-RT」を付与し収集し

3 : <http://www.bousai.go.jp/taisaku/gekijinbukko/status.html> (2019/12/14 確認)

4 : <https://www.mlit.go.jp/> (2019/12/13 確認)

5 : <http://www.news24.jp/articles/2019/10/12/07519921.html> (2019/12/14 確認)

6 : <https://www.tokyo-np.co.jp/article/kanagawa/list/201910/CK2019102002000123.html> (2019/12/14 確認)

7 : <http://www.hrr.mlit.go.jp/chikuma/bousai/shussui/index.html> (2019/12/14 確認)

表 2 期間中ツイート数

日時	入間川	相模川	千曲川
10 日	22	81	40
11 日	55	198	71
12 日	10416	31741	17277
13 日	2569	6177	47598
14 日	350	753	7027
15 日	163	451	3985
16 日	71	185	1996
合計	13676	39586	77994

た。合計取得件数は入間川 13676 件、相模川 39586 件、千曲川 77994 件であった。日付別取得件数を表 2 に示す。収集したツイートは JSON 形式で保存し、ツイート日時、ツイート本文と本文に含まれている URL 情報を抽出した。全ての川でツイート数の多かった 10 月 12 日～13 日のデータを分析対象とした。

4 取得データ

4.1 前処理

3 章で収集したツイートデータから出現する単語や URL の抽出を行うため、収集テキストの前処理を行う。日本語文は単語の区切りが明確でないため、テキストの前処理として単語の分割処理を行う必要がある。形態素解析器 MeCab (ver.0.966.2)⁸ を用いて形態素解析を行い収集したツイート本文を単語ごとに分割した。次に、ツイート本文中の URL を除く全角英数字は全て半角英数字に変換した。また、SNS では絵文字が頻繁に使われ、データ分析に影響を与えるため、絵文字の除去を行う。絵文字の除去は Unicode の文字コードを範囲指定することで取り除いた。これらの処理を行う際、Python ライブラリの neologdn (ver.0.4)⁹ を用いてテキスト処理を行った。指定範囲の文字コードは以下の通りである。

- (1) U00002190-U000021FF
- (2) U00002300-U000023FF
- (3) U000025A0-U000025FF
- (4) U00002600-U000026FF
- (5) U00002700-U000027BF
- (6) U00002B00-U00002BFF
- (7) U0001F100-U0001F1FF
- (8) U0001F300-U0001F5FF
- (9) U0001F600-U0001F64F
- (10) U0001F650-U0001F67F
- (11) U0001F680-U0001F6FF
- (12) U0001F900-U0001F9FF
- (13) U0001FA70-U0001FAFF
- (14) U0000203C
- (15) U00002049

8 : <https://taku910.github.io/mecab/> (2019/12/14 確認)

9 : <https://pypi.org/project/neologdn/> (2019/12/14 確認)

表 3 ツイート出現単語上位 20

入間川出現単語	出現数	相模川出現単語	出現数	千曲川出現単語	出現数
氾濫	1515	ダム	1923	決壊	1810
発生	388	放流	1922	長野	1500
情報	326	城山	1404	氾濫	1074
流域	322	緊急	1243	堤防	1022
埼玉	308	避難	488	被害	582
東京	287	氾濫	472	台風	499
速報	191	午後	450	新幹線	490
号	186	水位	422	救助	433
荒川	168	流域	420	無事	423
防災	159	上昇	283	ニュース	414
アプリ	142	台風	268	浸水	399
大丈夫	73	警戒	251	大丈夫	396
台風	69	大丈夫	247	大変	391
都幾川	63	神奈川	233	住宅	333
多摩川	57	心配	230	流域	307
危険	52	上流	197	心配	273
河川	50	沿い	181	濁流	270
避難	49	下流	156	北陸	269
警戒	44	相模原	156	車両	261
洪水	41	付近	139	自衛隊	259
上流	40	開始	130	糖保	255

前処理を行う前のツイート例を図 1¹⁰¹¹ に、処理を行った後の例を図 2 に示す。

4.2 名詞抽出

収集したツイートデータのうち、全ての川でツイート数の多かった 10 月 12 日～13 日のデータに対し、Python データ分析ライブラリの Pandas (ver.0.23.4) を用いて分析を行った。TwitterAPI を用いて取得したツイート日時はグリニッジ標準時 (GMT) で格納されていたため、日本標準時 (JST) に変換した。ツイートデータを 30 分単位で分割し、出現単語と出現数を調査した。それぞれの河川で最もツイート数の多かった時間帯は、入間川 : 10 月 12 日 19 時 30 分～20 時で 1908 件、相模川 : 10 月 12 日 21 時～21 時 30 分で 3139 件、千曲川 : 10 月 13 日 07 時～07 時 30 分で 4739 件となった。それぞれの河川で最もツイート数の多かった時間帯における出現単語の出現数上位 20 語を表 3 に示す。その際、SlothLib [12] に含まれるストップワードと、各河川名を除外した。

4.3 URL 抽出

ツイートに投稿された画像・動画の件数やネットニュースに遷移するリンクが含まれたツイート件数の調査を行うため、ツイートデータに含まれる URL の抽出を行った。ツイートに URL が含まれる条件は以下の通りである。

- (1) 画像または動画をつけてツイートした
- (2) 他のツイートに対してメンションを行った
- (3) 外部の Web サイトのリンクを貼った

投稿された URL は自動的に短縮化が行われるため、元の URL への展開が必要になる。本稿では本文中に含まれていた

10 : <https://twitter.com/momo04080220/status/1183034487447289858> (2019/12/19 確認)

11 : <https://twitter.com/torajiroo0508/status/1183034349752471552> (2019/12/19 確認)

Sat Oct 12 14:59:36 +0000 2019 @●●● @●●● @●●● ゆあさん ありがとうございます🙏😊 今日全面stop お店も夕方以降ほとんど お休み 今23区内警告増える😓 千曲川氾濫も心配です😓 🐱 は音恐がるから窓際で じっとしてます😓

Sat Oct 12 14:59:03 +0000 2019 @●●● 千曲川の近くだったのか!😓 長野が結構被害出てて、心配はしてたんよね😓

図 1 テキスト処理前

Sat Oct 12 14:59:36 +0000 2019 @●●● @●●● @●●● ゆあさんありがとうございます 今日全面stopお店も夕方以降ほとんどお休み今23区内警告増える 千曲川氾濫も心配です は音恐がるから窓際でじっとしてます

Sat Oct 12 14:59:03 +0000 2019 @●●●千曲川の近くだったのか! 長野が結構被害出てて、心配はしてたんよね

図 2 テキスト処理後

短縮 URL に対し正規表現を用いてツイート本文から URL を抽出し、Python ライブラリの urllib3 (ver.1.24.1)¹²を用いて展開を行った。最も URL 件数を抽出できたのは 13 日の千曲川に関するツイートに含まれる URL で全 12275 件であった。

短縮 URL を展開した元の URL の中から、ツイートに URL が含まれる条件 1, 条件 2 に該当する URL は全て Twitter へ遷移し、5475 件確認できた。写真の投稿を示す URL ([https://twitter.com/~photo/1](https://twitter.com/~/photo/1)) は 3812 件、動画投稿を示す URL (<https://twitter.com/~video/1>) は 453 件確認できた。これらの写真・動画の投稿は、渡辺らが写真付きツイートを抽出した際に利用した URL [8] を参考に抽出した。Web メディアの記事に遷移するツイートは、ツイートに URL が含まれる条件 3 に該当する。ニュース記事に遷移する URL には「news」または「article」が含まれている場合が多く、1332 件確認できた。遷移先の Web サイトの内訳は、NHK に遷移する URL (<https://www3.nhk.or.jp/news/~/>) は 276 件、産経新聞に遷移する URL (<https://www.sankei.com/premium/news/~/>) は 213 件、朝日新聞デジタルに遷移する URL (<https://www.asahi.com/article/~/>) は 33 件、Yahoo! JAPAN ニュースのアーカイブに遷移する URL (https://news.yahoo.co.jp/list?t=2019_typhoon_19) は 182 件となった。また、12 日~13 日の期間の三河川における「news」または「article」が含まれている投稿の合計は 2960 件確認できた。

ニュース記事の内訳に関しては、NHK に遷移する URL (<https://www3.nhk.or.jp/news/~/>) は 276 件、産経新聞に遷移する URL (<https://www.sankei.com/premium/news/~/>) は 213 件、朝日新聞デジタルに遷移する URL (<https://www.asahi.com/article/~/>) は 33 件、Yahoo!JAPAN ニュースのアーカイブに遷移する URL (https://news.yahoo.co.jp/list?t=2019_typhoon_19) は 182 件となった。時間帯別で見た場合の一例として、期間中最も多かった 13 日 07 時の千曲川に関するツイートを例に挙げる。これらの中から URL を含むツイートの合計は 953 件であった。URL が含まれる条件 1, 条件 2 に該当する URL は 480 件確認できた。Twitter に遷移する URL のうち、写真投稿を示す URL は 357 件、動

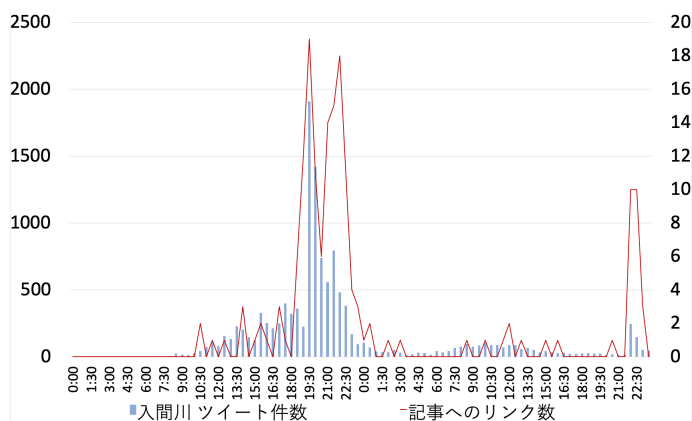


図 3 単位時間毎の入間川ツイート数と記事へのリンク数

画投稿を示す URL は 43 件確認できた。ニュースサイトや記事に遷移する URL は合計 146 件であった。そのうち NHK に遷移する URL は 29 件、産経新聞に遷移する URL は 1 件、朝日新聞デジタルに遷移する URL は 2 件、Yahoo!JAPAN ニュースに遷移する URL は 88 件となった。13 日全体のニュースの内訳と 07 時台の URL を比較すると、時間帯によって投稿される Web メディアのプラットフォームに差が出る結果となった。

5 データの参照

5.1 パーストの時間帯

入間川、相模川、千曲川、全ての河川において急激にツイート数が増えた時間帯があった。各河川における単位時間あたりのツイート数と直前 30 分のツイート数を比較したところ、入間川では 12 日 19 時 30 分から 20 時の間で 844.2%、相模川では 12 日 14 時 30 分から 15 時の間で 258.9%、同日 21 時から 21 時 30 分の間で 323.9%、千曲川では 12 日 20 時 30 分から 21 時の間で 1062.7%、13 日 06 時から 06 時 30 分の間で 368.6%の増加が各々観測された。入間川のツイート数とニュース記事へ遷移するリンク数の遷移を図 3 に、相模川のツイート数とニュース記事へ遷移するリンク数の遷移を図 4、入間川のツイート数とニュース記事へ遷移するリンク数の遷移を図 5 に各々示す。

12: <https://github.com/python/cpython/blob/3.8/Lib/urllib/request.py> (2019/12/14 確認)

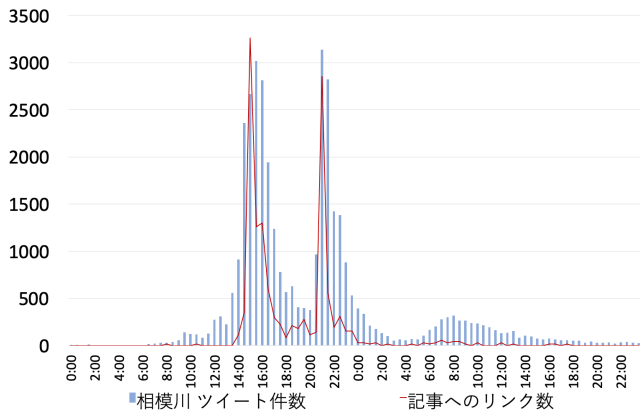


図4 単位時間毎の相模川ツイート数と記事へのリンク数

5.2 パースト時間帯の出現単語

各河川のツイート数パースト時における出現単語の出現数上位26位を調査した。パーストが発生した時間を基軸として前後30分の出現単語数を降順でソートを行った。その際「入間川」「相模川」「千曲川」および「川」は除外している。入間川12日19時30分前後を図6、相模川12日14時30分前後を図7、相模川12日20時30分前後を図8、千曲川12日20時30分前後を図9、千曲川13日06時前後を図10に示す。

5.3 出現単語上位20語におけるユニーク単語

各河川における出現単語の差とニュース記事のURLを含むツイートを比較するため、最もツイート数の多い時間帯におけ

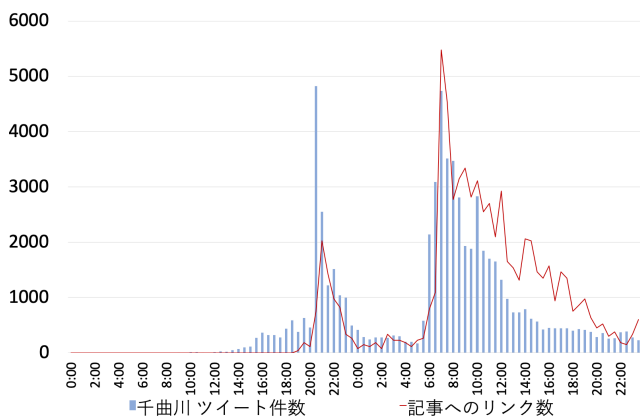


図5 単位時間毎の千曲川ツイート数と記事へのリンク数

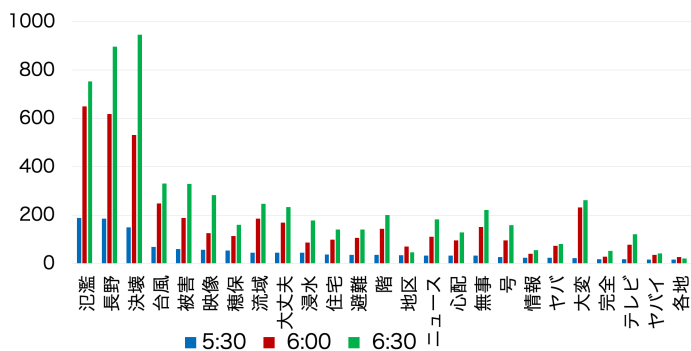


図6 入間川12日19時30分前後の頻出単語

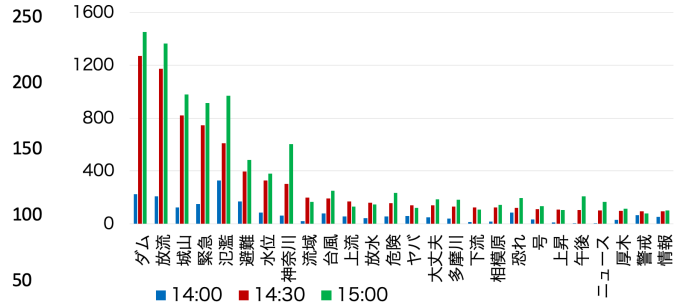


図7 相模川12日14時30分前後の頻出単語

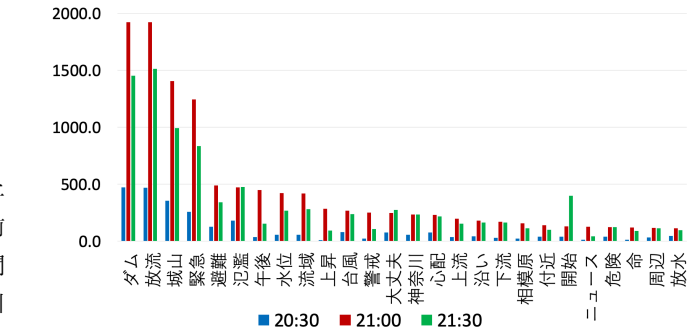


図8 相模川12日21時00分前後の頻出単語

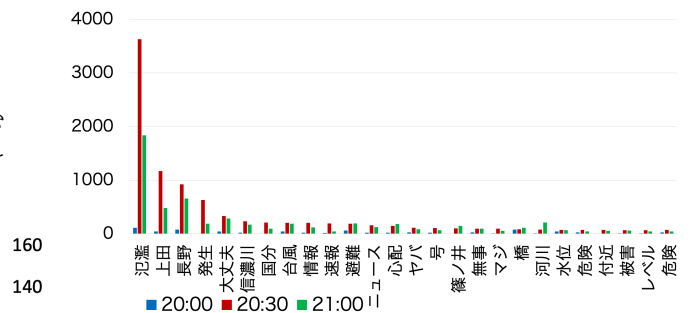


図9 千曲川12日20時30分前後の頻出単語

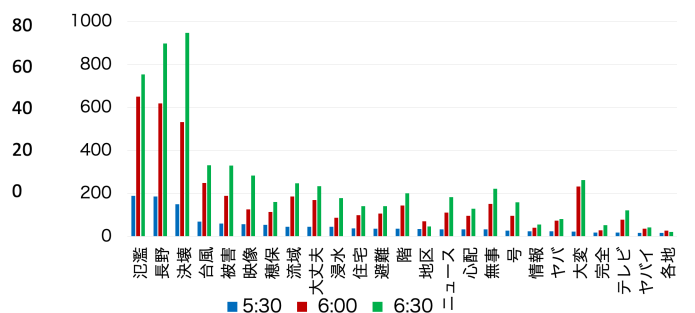


図10 千曲川13日06時00分前後の頻出単語

る各河川ごとの出現数上位20単語を参照した。各河川における最もツイート数の多い時間帯の出現単語上位20語から地名を除いた3河川に共通しない単語（以下、ユニーク単語と記す）を表4に示す。

6 考 察

6.1 パーストタイミングにおけるツイートとURL

5章で提示した各河川の単位時間毎のツイート数とニュース

表 4 出現単語上位 20 語におけるユニーク単語

出現単語	入間川	相模川	千曲川
アプリ	142	0	2
河川	50	101	114
洪水	41	53	31
ダム	33	1923	40
放流	24	1922	7
緊急	8	1243	22
午後	3	450	4
水位	37	422	59
下流	12	169	58
沿い	14	181	56
堤防	2	25	1022
新幹線	0	2	490
救助	0	0	433
無事	17	92	423
ニュース	24	128	414
浸水	10	89	399
大変	2	34	391
住宅	0	3	423
濁流	0	1	270
北陸	0	0	269
車両	0	0	261
自衛隊	0	0	259

記事数の図 3~5 から、ニュース記事へ遷移する投稿の増減に伴いツイート数も増減する様子が確認できた。入間川のバーストした時間帯のツイート数と、ニュース記事の数を比較すると、ツイートの増加とほぼ同じ動きでネットニュース数も増加している。また、バースト後はツイート数と Web ニュース数が同様に減少していることが明らかになった。入間川のバーストに関しては双方の急激な増加が見られたため、ニュース記事の増加が Twitter に影響を及ぼした可能性を示唆する。一方で、相模川における 1 度目のバーストと千曲川におけるバーストの発生では、ニュース記事へ遷移する URL を含む投稿の増加よりも早い時間にツイートの数が増加発生した。そのため、1 度目の相模川のバーストと千曲川のバーストは、ニュース記事以外の影響によって情報爆発が発生した可能性が高い。これらの可能性を調査するため、各河川におけるツイート数とニュース記事へ遷移する投稿数の相関係数を算出した。

各河川におけるツイート数と、ニュース記事へ遷移する投稿数の相関係数は入間川で 0.77、相模川で 0.780、千曲川で 0.79 となり、強い正の相関が確認された。ただし、「news」や「article」を含む URL のうち、既に記事が消されている場合はエラーとして処理されていたため、それらを手動で確認したところ NHK の記事に遷移する場所が確認できた。これは他社の記事より比較的早く NHK の記事が消される場合が多いことに起因すると考えられる。災害直後における NHK の記事に遷移する URL は、本稿で抽出できた数を大きく超えると考えられる。そのため、今後より正確に Web メディアが掲載したニュース記事の影響を計測するためには、災害発生直後の記事が削除される前に URL の検証をする必要がある。同様に、ニュース

記事に遷移する URL が含まれるツイートの拡散率を計測する必要がある。この問題は、TwitterAPI を用いて収集したデータからリツイート回数についてのデータを参照し、Web メディアの記事に遷移するツイートのリツイート件数と今回抽出したニュース記事を含むツイート件数を加算したものを分析することで解決できると考える。

6.2 バーストタイミングにおける出現単語

5 章で提示した、バースト時における頻出単語について述べる。入間川のバースト時のツイートでは、比較する際に基軸となった時間帯に最も単語出現数が多いことが確認できた。基軸となった時間帯の 30 分前に出現した単語は、基軸となった時間帯では減少傾向にあった。全ての単語において基軸となった時間帯を 30 分経過すると出現数は減少した。相模川におけるバーストは 12 日に 2 度起きた。相模川の 1 度目のバーストにおいて、基軸となった時間帯から 30 分経過した後も出現単語はほぼ全ての単語で増加している。2 度目のバーストが起きた際は、基軸となった時間帯の 30 分後には単語出現数は減少していたが、「開始」においては大幅に増加している。これは城山ダムの緊急放流開始のツイートが増えたためだと考えられる。千曲川は 12 日に 1 度、13 日早朝に 1 度、計 2 度バーストした。12 日 20 時 30 分のバーストは基軸となった時間帯の 30 分から経過すると単語出現頻度が減少した一方で、13 日 06 時のバーストでは、基軸となった時間帯から 30 分経過すると急激な単語出現数の増加が見られた。これは先に述べたバースト時間帯のツイート数およびニュース記事数が増加していることも影響していると考えられる。出現数が最も多い「氾濫」は 300%を超える増加を見せたのに対し、ツイート数およびニュースの記事の増加は、150%ほどに留まっていることから、ニュース記事以外の要因による影響も考えられる。本稿で行った調査では、ユーザの目撃情報よりメディアの報道が先かは明確ではないため、今後は実世界で起きたことがメディアを通さず拡散された可能性による影響の調査や異なるメディアの影響を複合的に調査する方法を検討する必要がある。

6.3 Web メディアの影響と放送メディアの影響

5 章で提示した、最もツイート数の多い時間帯における出現数上位 20 単語で共通しなかった単語についても考察を行う。入間川で多数出現した「アプリ」を含むツイートを確認したところ、Yahoo!防災速報アプリケーション¹³を利用して投稿されたツイートには「#防災速報アプリ」といった内容が本文に含まれており、これらは 139 件を占めた。相模川で多数出現した「ダム」「放流」「緊急」は城山ダムの緊急放流を行ったことが影響していると考えられる。千曲川でのみ出現した「救助」は実際に救援を求める声ではなく、自衛隊の救助に関する投稿が見られた。これらの投稿は、長野県が 12 日 21 時 34 分に自衛隊に災害派遣要請をし自衛隊の救助が始まったこと¹⁴に起因す

13 : <https://emg.yahoo.co.jp/> (2019/12/19 確認)

14 : <https://www.mod.go.jp/j/press/news/2019/10/13a.pdf> (2019/12/19 確認)

る可能性があると考えられる。また、「新幹線」は、千曲川の氾濫に伴い北陸新幹線車両センターが冠水した事象に基づくものであった。「新幹線」に関する投稿を、手動で参照した結果、図 11¹⁵のような投稿が多く見られた。添付された画像は「NHK ニュースおはよう日本」¹⁶で放送された映像を撮影したものだった。新幹線に関する投稿のうち、URL リンクを含む投稿は 104 件あり、そのうち前述した番組を撮影した写真または動画を添付したものは 92 件確認できた。ユーザが自ら現地に赴いて撮影した投稿は、1 種存在し (図 12)¹⁷、この投稿を引用したツイートが 2 件確認できた。その他に、図 13 のように氾濫が起こったと推測できる地域を GoogleMap¹⁸や国土地理院の Web サイト¹⁹をスクリーンショットした画像の投稿が確認できた。これらの「新幹線」を含む投稿が増えた背景には、Web メディアの影響ではなく、放送メディアの影響が考えられ、今後これらの放送メディアに関連する投稿を分類する必要がある。一方、「濁流」は朝日新聞デジタルのニュース記事²⁰のタイトルに含まれており、URL を含む投稿 240 件中 125 件を占めていた。そのうち Yahoo!ニュースの記事に遷移する投稿は 87 件であった。濁流に関する投稿が増えた要因として Web メディアが影響を与えた可能性が考えられる。ニュース記事への URL を含むツイートの情報源を参照した結果、特定のニュース記事への URL が多く投稿される傾向が見られた。この際、引用ツイートとして投稿される場合と Web サイトから直接リンクを投稿する機能を利用して投稿される場合の 2 つのケースが確認できた。

7 今後の展望

本研究では、Web メディアの記事が SNS に与える影響を定量的に調査した。しかし、SNS の影響を受けた情報を今後どのように分類し排除していくかまでは考慮できなかった。メディアの影響は一部で顕著に見られたが、メディアの影響を反映するツイートの条件を定義し、分類を行った上で定量的に調査する必要がある。また、ツイートのバースト時に出現した単語と出現数を挙げたが、これらは全てネットニュースの投稿を含んだものであり、より厳密な調査を行うには、それらを除いたツイート群で分析を行わなければならない。同様に、ネットニュースの拡散度を加味する必要もある。今後はそれらを考慮した上で、再度分析を行う必要がある。本研究において、ネットニュースの定量的調査を行ったが、相模川に関する記事にも千曲川に関する内容が記載されている可能性があるため、記事の内容を含めて定性的に調査する必要があると考える。また、放送メディアのニュース放送が SNS に対して与えた影響につ

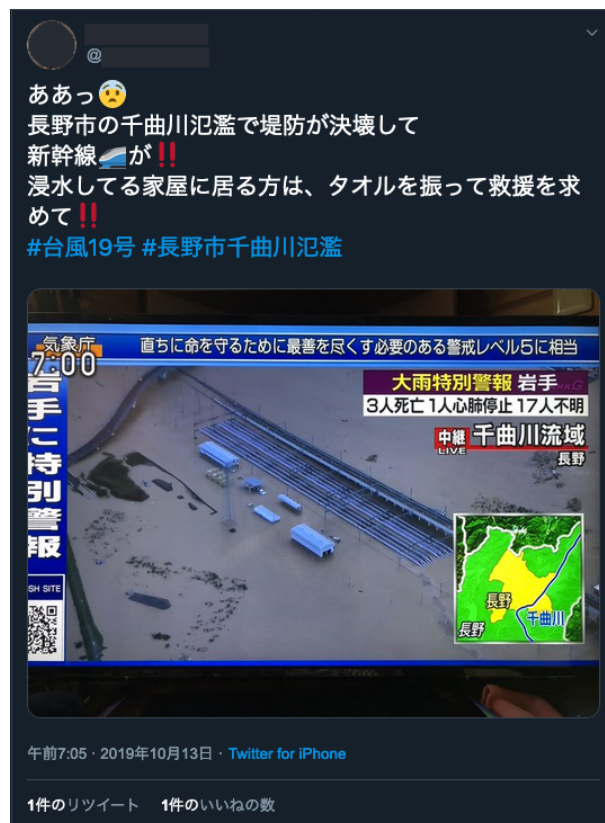


図 11 Twitter 投稿より、新幹線に関する画像付きツイート例



図 12 Twitter 投稿より、現地撮影されたツイート

15 : <https://twitter.com/hyuhypowpow/status/1183141776774709248> (2019/12/19 確認)

16 : <https://www4.nhk.or.jp/ohayou/> (2019/12/20 確認)

17 : <https://twitter.com/kasekky/status/1183137906963017728> (2019/12/19 確認)

18 : <https://www.google.co.jp/maps/> (2019/12/19 確認)

19 : <https://www.gsi.go.jp/> (2019/12/19 確認)

20 : <https://www.asahi.com/articles/ASMBF2404MBFUTIL007.html> (2019/12/20 確認)

いての調査は行っていない。テレビにおけるニュース放送の経時的データを用いて、どの程度 SNS に対して影響し、Web メディアの記事の与えた影響と比較することで、新たな課題の発見につながると考える。今後、異なるメディアの影響を複合的に調査する方法を検討する必要がある。

ツイートに URL が含まれる投稿の中から Twitter ユーザが



図 13 Twitter 投稿より, Map 画像付きツイート例

自ら投稿した画像が添付されたツイートについて確認したところ、緊急警報情報のスクリーンショットやテレビ報道をカメラで撮ったものの投稿が多く見られた。実際の被害状況を伝える投稿や川の水位が観察できるような投稿画像の存在も確認はできたが、具体的な割合や数値についての手動での調査は困難である。Web メディアの影響調査を行った結果、今後放送メディアの影響を調査する必要があることが明らかとなった。放送メディアの影響を定量的に抽出・分類する手法を検討する必要があると考える。その上で、放送メディアのニュースを撮影したものと、被害状況などのユーザが置かれている状況とを分類することができれば、SNS 情報の有効性を担保するものとなると考える。

8 おわりに

本稿では 2019 年 10 月 10 日～10 月 16 日の間に投稿された「入間川」「相模川」「千曲川」に関するツイートの収集を行い、ツイート情報とニュース記事に遷移する投稿を定量的に調査した。その結果、両者の間に正の相関が確認され、情報爆発時の Web メディアの影響が示唆された。さらに、記事に遷移する URL を含むツイートの情報源には特定の記事への URL が多く投稿される傾向が見られた。また、ユーザの投稿した写真を手動で確認したところ、放送メディアの映像を撮影したものが多く確認できた。今後、放送メディアのニュースを撮影したツイート数を調査し、ツイート数とニュース記事に遷移する投稿数との比較を行う。

文 献

[1] Kwak, H., Lee, C., Park, H., Moon, S. :What is Twitter, a social network or a news media?, Proceedings of the 19th

international conference on World wide web, pp.591-600 (2010).
 [2] McCormick, T, H., Lee, H., Cesare, N., Shojaie, A., Spiro, E, S., :Using Twitter for demographic and social science research: Tools for data collection and processing, Sociological methods & research, Vol. 46, No. 3, pp. 390-421 (2017).
 [3] Sakaki, T., Okazaki, M., Matsuo, Y. :Earthquake shakes Twitter users: real-time event detection by social sensors, Proceedings of the 19th international conference on World wide web, pp. 851-860(2010).
 [4] Aramaki, E., Maskawa, S., Morita, M., :Twitter catches the flu: detecting influenza epidemics using Twitter, Proceedings of the conference on empirical methods in natural language processing, pp1568-1576(2011).
 [5] 荒牧 英治, 若宮 翔子 :ツイート数と現実の統計量との差異に関する検討, 統計数理, Vol.64, No.2, pp.233-246(2016).
 [6] 藤代裕之, 松下光範, 小笠原盛浩: 大規模災害時におけるソーシャルメディアの活用—情報トリアージの適用可能性, 社会情報学, Vol. 6, No. 2, pp. 49-63 (2018).
 [7] 喜連川優: 情報爆発のこれまでとこれから, 電子情報通信学会誌, Vol. 94, No. 8, pp. 662-666 (2011).
 [8] 渡辺 隼矢, 桐村 喬: 写真付き SNS 投稿データからみた観光地への関心の時系列変化, 2018 年度日本地理学会発表要旨集, pp. 198 (2018).
 [9] 榎本 甫, 桑野 将司, 小池 淳司: 災害時のソーシャルメディアと帰宅行動の関連分析, 土木学会論文集 D3 (土木計画学), Vol. 70, No. 1, pp. 102-112 (2014).
 [10] 國友優, 神山嬢子: Twitter 情報を活用した土砂災害の前兆・発生状況把握の可能性, 土木技術資料, Vol. 57, No. 9, pp. 18-21 (2015).
 [11] 武田邦敬, 古川忠延, 瀧口茂隆, 山影讓, 荒木健, 伊藤顕子, 宮川健, 蒲原潤一: Twitter データを活用した土砂災害の発生推定 ~ H24 年九州北部豪雨における土砂災害の調査事例報告 ~, 平成 26 年度砂防学会研究発表会概要集 B, pp. 172-173 (2014).
 [12] 大島裕明, 中村聡史, 田中克己: SlothLib: Web サーチ研究のためのプログラミングライブラリ, 日本データベース学会 Letters, Vol. 6, No. 1, pp. 113-116 (2007)