

ゲリラ豪雨時の社会現象に関する一考察 -twitterを用いた検討-

藤生 慎¹・高田 和幸²・富田 智之²・伊東 大輔²

¹正会員 金沢大学 理工学域環境デザイン学類 (〒921-1192 石川県金沢市角間町)

E-mail: fujju@se.kanazawa-u.ac.jp

²正会員 東京電機大学 理工学部建築・都市環境学系 (〒350-0394 埼玉県比企郡鳩山町石坂)

E-mail: takada@g.dendai.ac.jp

³非会員 株式会社NTTデータ (〒135-6033 東京都江東区豊洲3-3-3 豊洲センタービル)

E-mail: tomitatm@nttdata.co.jp

⁴非会員 株式会社NTTデータ (〒135-6033 東京都江東区豊洲3-3-3 豊洲センタービル)

E-mail: itoudi@nttdata.co.jp

災害時にtwitterデータを有効活用することを目指して、ゲリラ豪雨時のtweetデータを用いて単語の出現頻度分析、共起語分析、共起語ネットワーク分析を行い、災害時におけるtwitterへの情報発信特性の分析を行った。その結果、品詞別出現頻度の分析では状況・感情・地名・避難などの情報の発信が確認された。また、「豪雨」との共起語分析では、状況を表す単語と防災の観点からのtweetが行われていることが確認された。共起語ネットワーク分析では、ゲリラ豪雨時の被害事実、予警報情報、時間・場所の情報を取り出すことが可能であることが確認された。

Key Words : twitter, gurilla rail fall, word network analysis, consciousness analysis

1. はじめに

2013年に気象庁より発表された「記録的短時間大雨情報」は全国で76回(2012年は68回)に及び、東京では4回(2012年は0回)発表されている。2013年7月23日に目黒区付近・世田谷区付近で16時30分までの1時間に約100ミリの降水があったとして、記録的短時間大雨情報を発表した。東京都で記録的短時間大雨情報が発表されるのは3年ぶりであった。このゲリラ豪雨により、道路冠水、住宅の床上・床下浸水、鉄道の運行停止、小規模河川の氾濫などの被害が発生し、市民生活に影響が生じた。図1に2013年7月23日の世田谷区付近で生じたゲリラ豪雨時のレーダーエコー強度図を示す。12時には雨雲は確認することができないが、14時頃から雨雲が確認できるようになり、ゲリラ豪雨が観測された16時頃には、レーダー上でも、その事実を確認することが可能である。しかし、17時になると世田谷区付近にはレーダーエコーは確認することはできない。このようにゲリラ豪雨は、短時間に降雨が発生するため、屋内にいた市民はゲリラ豪雨に気がつかないことも少なくない。このため、降雨の状況を認識できる市民と被害に遭う可能性のある市民

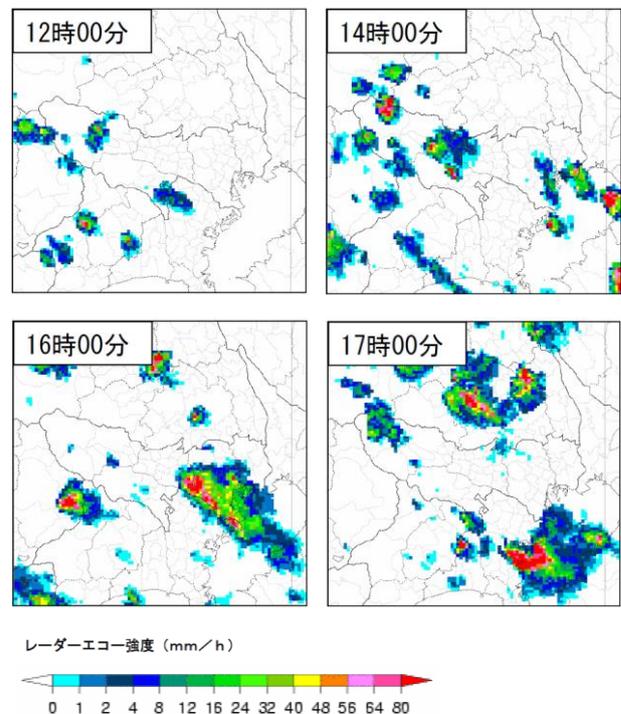


図1 ゲリラ豪雨時のレーダーエコー強度図¹⁾

は、外出をしていた者に限定されるため、事後のアンケート調査を通じて豪雨時の状況や行動について把握することは難しい。そこで本研究では、防災分野でのビッグデータの活用を目指し、まずリアルタイムな情報発信性能を有するtwitterの発信データを用いた豪雨時の発信情報の特性分析を行うことを目的とした。

2. 既往研究と本研究の位置づけの整理

Twitterデータを活用した研究は東日本大震災以降活発に行われるようになった。特に災害時の分析にtwitterを利用した研究はまだ蓄積がほとんどなく、今後の研究成果が期待される場所である。Twitterの研究は、デマの検出、予測、利用傾向に分類することができる。本研究に関連の深い分野はtwitterの利用傾向に関する既往研究である。石川ら²⁾は、twitterアカウント運用者の視点に立って分析を行っており、東日本大震災もしくは紀伊半島水害時にソーシャルメディア上で情報伝播の要として活躍した4つの組織・個人のアカウント運用者に対してヒアリング調査を行い、ソーシャルメディア運用の特徴や課題を整理している。さらに災害時の行政機関アカウントの運用指針の提案、そして課題に対して調査から明らかになった知見を踏まえた解決策の提案を行っている。さらに石川ら³⁾は、国内外の災害事例におけるソーシャルメディアの利用に関する分析を行い、得られた教訓と課題、問題点を整理している。また、2011年山陰地方豪雪災害時におけるTwitterで投稿された4,580件の災害情報の意味内容を投稿毎に判別し、詳細な意味分類による定量分析を行い、ソーシャルメディアの情報収集ツール・コミュニティツールとしての災害時利用の可能性を検討している。谷口³⁾は、災害時のtwitterの有用性について、利用アカウントの観点から分析を行い、公式アカウントは信頼性の高い情報発信を行わなくてはならない一方、非公式アカウントは相対的に責任が軽い大量の情報を発信しやすいと述べている。風間ら⁵⁾は、東日本大震災発生前後の約3億ツイートから震災に関係すると思われる単語を抽出し、その出現頻度が上位の単語から「地震」と「原発」の2つの単語を選択し、その単語と関連語及びその単語間の関係の時系列変化を分析している。宮部ら⁶⁾は、2011年3月11日に発生した東日本大震災を対象としてTwitterの利用状況の分析を行い、(1)被害の大きかった地域では直接的なメッセージのやり取りが行われる傾向がある。一方、被害の小さかった地域では、ツイートされた情報が拡散される傾向が高い。(2)特に被害の大きい地域で発信された情報については、他地域へと移動する傾向が見られたことを明らかにしている。以上のように、災害時のtwitterの利用に関する研究が行われつつ

あるがまだ研究の蓄積段階であった。さらに今後、twitterデータを災害時に有効活用するためには様々な災害を対象としてtweet内容を時系列的に分析しておくことが必要であることが確認された。

3. データ

本研究では、twitterデータベースから筆者らが設定したキーワードを用いて豪雨に関するtweetを抽出し、分析データを構築した。データの収集期間は、東京区部の注意報の発表から解除までの時間(15時30分～19時00分)とした。Tweet数は152,497(tweet)であった。図2にデータ収集期間中のtweet数を示す。赤線は10分ごとtweet数を示し、青線は、累積tweet数を示す。また、図3に分析対象期間の降雨量を示す。降雨量とtweet数の関係を考察するとtweetは東京都区部に注意報が発令された15時30分から増加している。その後、15時50分から降雨量のピークである16時30分²⁾まではほぼ一定のtweetであった。ピーク時間を過ぎるとtweet数は減少し、レーダーエコー強度図上でエコーがなくなる17時には注意報発表前と同じtweet数であった。

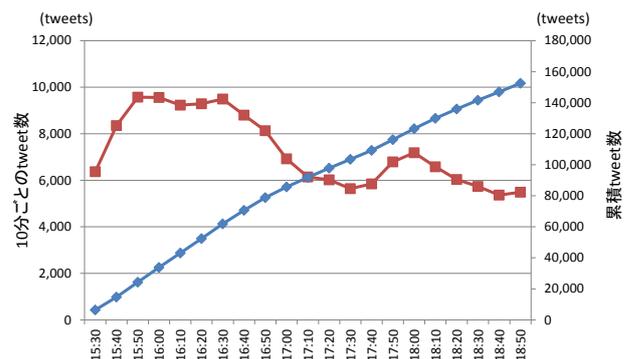


図2 tweet数の変化

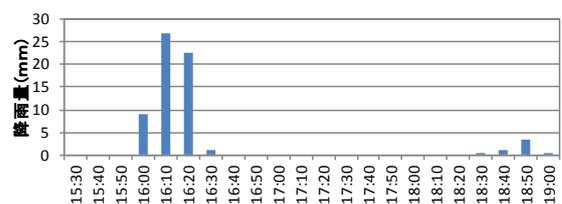


図3 降雨量の変化(世田谷区)¹⁾

4. tweet傾向の分析

(1) 単語の出現頻度分析

表1に品詞別(形容詞・副詞・動詞・名詞・固有名詞)の出現順リストの一部抜粋を示す。形容詞では、「やばい」、「すごい」、「激しい」、「ひどい」などのゲリラ豪雨の状況を表す単語の出現が確認された。また、「こわい」、「怖い」などゲリラ豪雨に見舞われた

際の感情を表す単語の出現も確認された。副詞では、「突然」や「ゴロゴロ」など、ゲリラ豪雨特有の状況を表す単語の出現が確認された。動詞では、「降る」、「濡れる」、「止む」などの降雨の状況を表す単語と「帰る」、「行く」、「出る」など豪雨時の行動を表す単語の出現も確認された。固有名詞では、東京でゲリラ豪雨が起きている様子がtweetされていることがわかる。さらに、具体的な地名として「世田谷」、「目黒」などの単語の出現も確認された。名詞では、「雨」、「豪雨」、「ゲリラ」、「大雨」、「雷雨」といったゲリラ豪雨が起きている状況を表す単語の出現が圧倒的である。以上、品詞別出現頻度を集計した結果、ゲリラ豪雨時の状況、感情、行動、場所の情報を含んだtweetが行われていることが確認された。

図2 tweet数の変化

形容詞		副詞		動詞		固有名詞		名詞	
単語	出現数	単語	出現数	単語	出現数	単語	出現数	単語	出現数
やばい	1308	ちょっと	270	する	3290	東京	1032	雨	6000
すごい	1007	これから	241	てる	2858	関東	183	豪雨	4703
ない	875	まだ	231	降る	1582	新宿	127	雷	4339
強い	384	めっちゃ	193	なる	1344	川崎	125	ゲリラ	2373
激しい	317	もう	188	いる	924	渋谷	120	雷雨	1780
怖い	290	ゴロゴロ	156	くる	871	品川	108	大雨	1335
しい	252	どう	150	いる	793	世田谷	104	雷	1020
っぽい	202	突然	147	れる	739	目黒	102	傘	969
強い	181	余計	122	ある	568			ゴリラ	665
ひどい	173	少し	121	濡れる	560			区	659
こわい	136			帰る	500			笑	603
楽しい	136			行く	476			ヤバ	510
早い	135			出る	467			中	481
凄まじい	132			思う	433			警報	443
暑い	114			止む	393			前	423

(2) 共起語の分析

ゲリラ豪雨を分析対象としているため「豪雨」をキーワードとして共起語分析を行った。表2に「豪雨」との共起語分析の結果を示す。なお表2には、共起数の高い単語を示した。表中に青色のハッチで示す「濡れる」、「雨漏り」、「滝」など状況を表す共起と表中に赤色のハッチで示す「情報」、「警報」、「警戒」、「避難」など防災の観点からの共起が確認された。状況を表す単語を含むtweetは個人アカウントから豪雨時の状況をtweetしたと考えられる。一方、防災のキーワードを含むtweetは気象庁や地方自治体などの公式アカウントからtweetされたものが多く(図4)、twitterユーザーは、そのtweetをリツイートするなどして拡散している(図5)。なお、図5はtwitter画面のキャプチャーの場合、個人アカウントが表示されることから本文の引用のみとした。

表2 共起語

帰宅	バイト	乗る	警戒	中止	嫌い
待機	安全	待つ	避難	皆	濡れる
最悪	びしょ濡れ	まだ	物	仕事	降る
滝	帰れる	学校	雲	今日	鳴る
目黒	地域	車	電車	川	
予定	雷雨	止む	窓	警報	
道路	雨漏り	ゴロゴロ	歩く	雨雲	
情報	上がる	移動	周辺	傘	

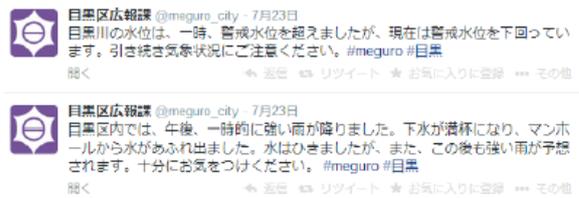


図4 twitterによる防災情報の発信(目黒区) 7)

RT @nhk_news: 都心で激しい雨 浸水や落雷に注意
http://t.co/lkkyS9eSH2 #nhk_news

RT @tenkijp: 都心に大雨警報、明日の夜も雨脚強く
http://t.co/pUxovJIJqG 関東は平野部でも雨雲や雷雲が発達中。
東京 23 区にも大雨警報と洪水警報が発表された。低い土地の浸水に注意。明日は関東は断続的に雨が降る見込み。雨の降り方は?

図5 twitterのリツイートによる防災情報の発信

(3) 共起語ネットワーク分析

災害時に被災者から発信される言語データを用いて言語解析を行った事例として石川ら³⁾があり、2011年山陰地方豪雪災害時におけるTwitterに投稿された4,580件の災害情報の意味内容を投稿毎に判別し、詳細意味分類による定量分析を行っている。また、藤生ら⁸⁾は被災者へのヒアリングデータを用いて共起語ネットワーク分析を行い、避難時の切迫感・緊迫感・危機意識を明らかにしている。一方、Williamら⁹⁾は、言語解析は、投稿文章を定量的・客観的に分析した結果として、文章の意味を精度高く分析するには自動解析では限界があることを指摘している。本研究では、既往研究での指摘を踏まえて言語解析のうち最もシンプルな手法である共起語ネットワーク解析を用いてゲリラ豪雨時のtwitterでの情報発信特性の分析を行うこととした。

図6にゲリラ豪雨時のtwitterデータを用いた共起語ネットワーク分析の結果を示す。ゲリラ豪雨時のtweetデータの共起ネットワーク分析を行うことにより生じている現象を把握することが可能であることが確認された。Tweetデータの共起語ネットワーク分析を行うことで把握することが可能な事実を図中に①から⑦で示した。①から⑦はそれぞれ被害事実、場所、予警報情報、現象、時間のいずれかを表している。①「被害事実」では、目黒川の氾濫危険性、世田谷区での豪雨、記録的短時間大雨情報の発令がわかる。②「被害事実」では、駒澤大学キャンパスが豪雨により浸水した事実がわかる。③「予警報情報」では、東京23区に大雨洪水警報は発令された事実がわかる。④「現象」では、雷を伴った豪雨であった事実がわかる。⑤「時間」については、「時」や「分」と共起ネットワークが形成されている。このことから他の被害事実や予警報情報とのネットワーク関係を見ることで、当該事象が発生した時刻を知ることができ

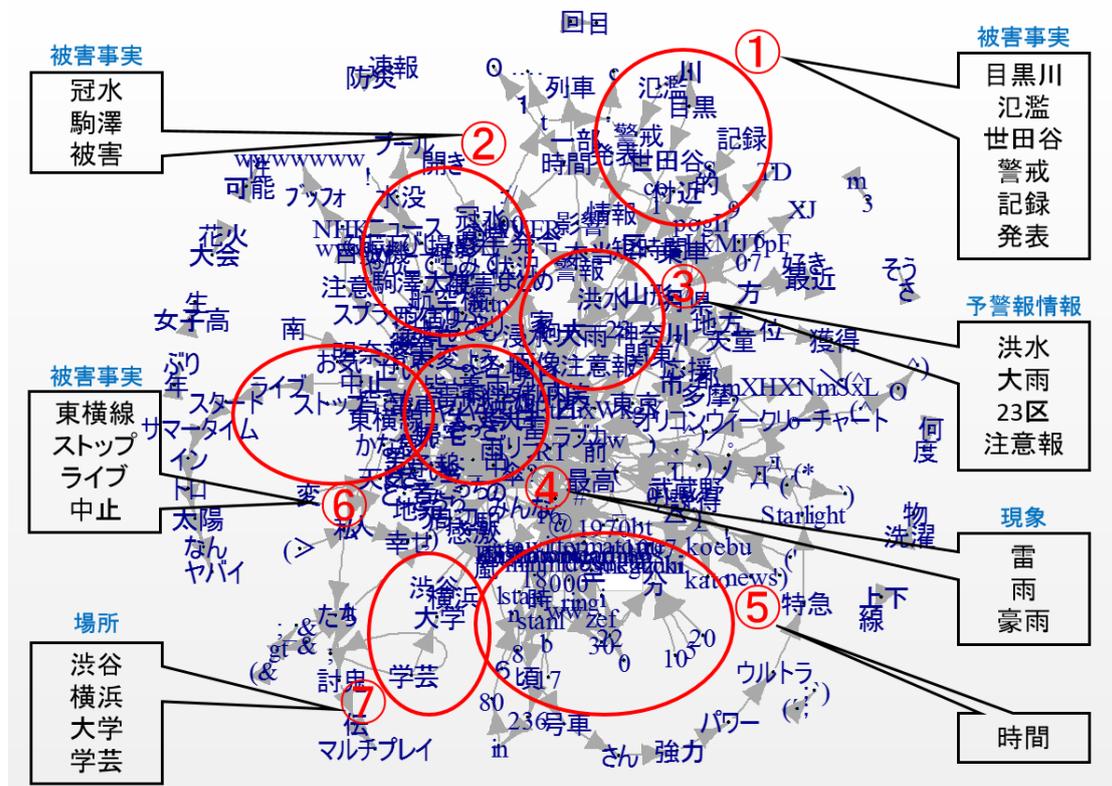


図6 共起語ネットワーク分析の結果

る。⑥「被害事実」では、東急東横線が運転見合わせになったことや、ライブイベントが中止となったことがわかる。⑦「場所」では、図1に示したゲリラ豪雨時のレーダーエコー強度図と上述した時刻の tweet の事実が一致する。このことからゲリラ豪雨に見舞われている場所がわかる。この事実は、渋谷では15時54分に「ゲリラ豪雨&雷すごいな@渋谷」と tweet があり、横浜では16時20分に「横浜市内、ゲリラ雷雨なう。」という tweet があった。

以上、tweetデータを共起語ネットワーク分析したところ、被害事実、予警報情報、時間・場所の情報を取り出すことが可能であることが確認された。

(4) 時間帯別の共起語ネットワーク分析

ゲリラ豪雨の発生中や、その前後のtweetの特性を分析するためには、前節の分析に加えてtweetデータを時間帯別に分けて共起語ネットワーク分析を行う必要がある。そこで、本節では、大雨・洪水警報/注意報が発令されていた時間帯を降雨量の最大値が観測された時間を中心に3つの時間帯に分類し、共起語ネットワーク分析を行った。分類した時間帯は、①：15:30～15:39、②：16:10～16:19、③：18:50～18:59とした。なお、分類した時間帯のうち、最大の降雨量が観測されたのは、16:10～16:19である。

①の時間帯では、地名や警報についてゲリラ豪雨を心配するtweetが多く発信されている。②の時間帯では、浸水・警報などの実際に起こっている事実以外に「怖い」

や「すごい」などの感情を表す表現が極端に増加している。このことから、ゲリラ豪雨に見舞われている最中には、感情のtweetが多く発信されることがわかった。③の時間帯は、東京での豪雨の事実、電車の遅延、注意報解除に関するtweetが増加していた。以上の結果、現象が起きている時間帯を分割したところ、時間帯別にtweetの傾向は大きく異なることが明らかとなった。

5. まとめと今後の課題

本研究では、災害時にtwitterデータを有効活用することを目指して、まずtwitterで発信されている内容の把握とその可能性を探ることを目的とした。そこで、本研究ではゲリラ豪雨時のtweetデータを用いて単語の出現頻度分析、共起語分析、共起語ネットワーク分析を行い災害時におけるtwitterへの情報発信特性の分析を行った。その結果、品詞別出現頻度の分析では状況・感情・地名・避難などの情報の発信が確認された。また、「豪雨」との共起語分析では、状況を表す単語と防災の観点からのtweetが行われていることが確認された。共起語ネットワーク分析では、ゲリラ豪雨時の被害事実、予警報情報、時間・場所の情報を取り出すことが可能であることが確認された。本研究では、分析の対象時間を注意報・警報の発令時間として取り扱うとともに、発生時間帯を3時間帯に分割して分析を行った。より詳細にtweetの特性把握を行うためには、時間を1分、5分、10分と区切り、かつ品詞ごとに共起語ネットワーク解析を

行う必要がある。さらに、対象とする災害も地震・雨・雪と拡張し、tweetによる情報発信特性の分析が必要である。これらの点は今後の課題としたい。

参考文献

- 1) 気象庁, 平成 25 年 7 月 23 日の大雨に関する東京都気象速報
(<http://www.jma-net.go.jp/>)
- 2) 石川哲也, 近藤伸也, 川崎昭如, 大原美保, 目黒公郎: 災害時における Twitter 利用の特徴と課題の整理: Twitter アカウント運用者の視点に立って, 生産研究 64(4), 545-552, 2012.
- 3) 石川哲也, 川崎昭如, 目黒公郎: 山陰地方豪雪災害時の Twitter ユーザによる情報発信行動に関する分析と考察, 地域安全学会論文集, No. 17, 2012. 07.
- 4) 谷口 慎一郎: 災害時における Twitter の有用性について: 2011 年 9 月の台風 12 号による豪雨災害を例に, 災害情報, 日本災害情報学会誌 (10), 56-67, 2012-03
- 5) 風間一洋, 鳥海不二夫, 榊剛史, 篠田孝祐, 栗原聡, 野田五十樹: 東日本大震災時の Twitter データを用いた単語間の関係の時系列変化の分析. 第 26 回人工知能学会全国大会予稿集, 2012.
- 6) 宮部真衣, 荒牧英治, 三浦麻子. 東日本大震災における Twitter の利用傾向の分析, 情報処理学会研究報告. EIP, [電子化知的財産・社会基盤] 2011. 17 (2011): 1-7.
- 7) 目黒区広報課 (@meguro_city) twitter ページ, (https://twitter.com/meguro_city)
- 8) 藤生慎, 石川哲也, 高田和幸: 言語解析を用いた東北地方太平洋沖地震による津波からの避難実態分析-宮城県気仙沼市の被災者を対象として-, 土木学会論文集 A1 (構造・地震工学), Vol. 69 (2013) No. 4 p. I_798-I_806.
- 9) Corvey, W. Vieweg, R.T. and Palmer, M.: Twitter in mass emergency: what NLP techniques can contribute, WSA '10 Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media, pp. 23-24, 2010.