

災害時のツイート分析と関連情報の同時可視化システム

山田実俊¹ 宇津圭祐¹ 内田理¹

概要: 大規模災害が発生した際に被害を最小限に食い止めるためには、迅速かつ確かな情報収集・伝達が重要であるが、そのような観点から災害時のソーシャルメディア、特に Twitter の利活用に注目が集まっている。本研究では、令和元年東日本台風や令和2年7月豪雨時に投稿されたツイートについて、共起ネットワークを用いてツイートの内容を俯瞰的に可視化した。さらにツイートに添付された画像や降水量などのツイートの内容と関連するデータを抽出し、その情報を共起ネットワークと同時に可視化するシステムを作成した。

キーワード: 災害情報, Twitter, 可視化システム, 共起ネットワーク

Visualization System of Tweet Analysis and Related Information during Disasters

SANETOSHI YAMADA¹ KEISUKE UTSU¹
OSAMU UCHIDA¹

Abstract: To minimize damage during disasters, the rapid collection and delivery of accurate information are essential. From this perspective, the use of social media, especially Twitter, in the case of a disaster has been attracted worldwide attention. In this study, we consider visualizing the contents of the tweets posted at the 2019 Typhoon Hagibis and the Kyushu flood disaster in July 2020. We also develop a system to visualize the tweets' contents and the precipitation on the day or attached images simultaneously.

Keywords: Disaster information, Twitter, Visualization, Co-occurrence network

1. はじめに

近年、日本では多数の自然災害が発生している。平成29年7月の九州北部豪雨や平成30年7月豪雨、平成30年台風第21号、令和元年東日本台風（台風第19号）、そして令和2年7月豪雨などでは、甚大な被害も発生した。このような大規模災害が発生した際に被害を最小限に食い止めるためには、迅速かつ確かな情報収集・伝達が重要であるが、そのような観点から災害時の SNS、特に Twitter の利活用に注目が集まっている[1]。2011年3月の東日本大震災の際には、多くの被災者が津波や避難所、公共交通機関の運行状況に関する情報の収集に Twitter を活用した[2]。

一方、大規模災害時には、ツイート数が爆発的に増加することが知られており、例えば、東日本大震災発生当日には、約3,300万件ものツイートが投稿されたことが判明している[3]。また、大阪府北部地震の際には、地震発生直後の6月18日午前8時から10分間の間に「地震」という単語を含むツイートが27万件以上投稿されていた[4]。そのため、災害時の Twitter の有効活用に向けて、ツイートの特徴を分析することは重要である[5]。我々は大阪府北部地震、平成30年7月豪雨、平成30年台風第21号などについて、ツイート数の変動、および利用されているハッシュタグや絵文字などを分析してきた[4][6][7]。

本研究では、ツイート分析と関連がある情報として、添付画像（動画）や降水量との関係に着目して分析を行う。特にツイート数が急上昇するタイミングとの関連が見つかれば、その時間帯での防災・避難行動の誘発や支援につながると考えられる。

そこで本研究では、令和元年東日本台風や令和2年7月豪雨時に投稿されたツイートについて、共起ネットワークを用いてツイートの内容を俯瞰的に可視化した。そして、ツイート内容と関連する添付画像や降水量の情報を同時に可視化するシステムを作成した。

2. 分析対象のツイート

本研究では令和元年東日本台風（台風第19号）と令和2年7月豪雨に関するツイートを Twitter search API によって収集した。令和元年東日本台風では、2019年10月12日～13日に投稿された「千曲川」を含むオリジナルツイート（リツイートではないツイート）を対象とした。ツイート数は64,813件、そのうち添付画像付きツイートは4,755件であった。さらにニュース記事と思われるツイート（URL付きのツイート）を取り除いた4,209件を対象データとした。令和2年7月豪雨では、2020年7月4日～9日に投稿された「大雨」という単語を含むオリジナルツイート476,084件

¹ 東海大学
Tokai University.

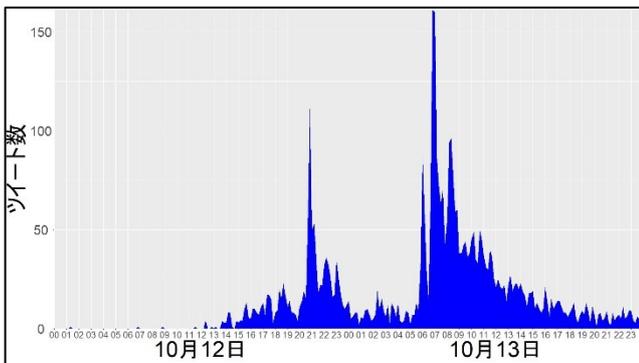


図 1: ツイート数の推移 (令和元年東日本台風)

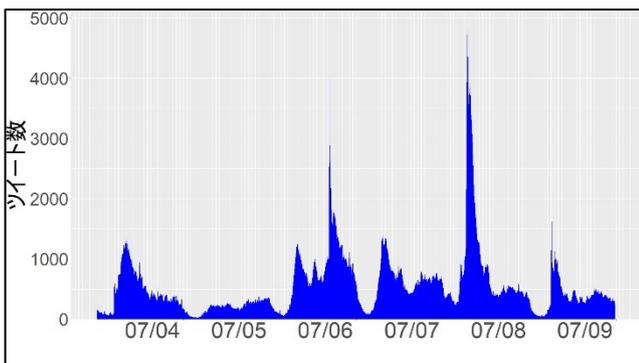


図 2: ツイート数の推移 (令和 2 年 7 月豪雨)

を対象データとした。

それぞれのツイート数を 10 分ずつ集計した結果を図 1, 図 2 に示す。令和元年東日本台風 (図 1) では, 10 月 12 日 20 時 50 分頃の千曲川の氾濫発生と, 13 日 6 時 50 分頃に朝のテレビニュースで被害の状況, 特に新幹線車両基地が水没した映像の報道が注目を浴びていた。令和 2 年 7 月豪雨 (図 2) では, 7 月 4 日 4 時頃の熊本県・鹿児島県での大雨特別警報, 6 日 16 時頃の長崎県・佐賀県・福岡県での大雨特別警報, 8 日 6 時頃の岐阜県・長野県での大雨特別警報の発表が注目を浴びており, 大きな報道に対応して投稿が増加することがわかる。

3. 分析結果

3.1 共起ネットワークによるツイート内容の可視化

ツイートの内容を把握する方法として共起ネットワークを用いた。本研究では連続する単語 A と単語 B に対して共起 ($A \Rightarrow B$) を抽出し, 出現頻度と単語 A, 単語 B それぞれに対する共起率を求めた (例: 千曲川 \Rightarrow 氾濫)。

$$\text{共起率}(A \Rightarrow B|A) = \frac{\text{共起}(A \Rightarrow B)\text{の出現頻度}}{\text{単語}A\text{の出現頻度}}$$

$$\text{共起率}(A \Rightarrow B|B) = \frac{\text{共起}(A \Rightarrow B)\text{の出現頻度}}{\text{単語}B\text{の出現頻度}}$$

共起の出現頻度や共起率に条件を付け, 共起ネットワークに表示する共起を抽出した。また, 本研究ではオリジナルツイートのみを対象としているため, ツイートに対するリツイート数・お気に入り (いいね) 数を注目度として共起の頻度を考慮した。

2019 年 10 月 12 日 13 時台のツイートデータで作成した共起ネットワークを図 3 に示す。円の大きさは単語の出現頻度, 矢印の太さは共起の出現頻度, 矢印の濃さは共起率 ($A \Rightarrow B|A$) の高さを表している。図 3 より, 「千曲川 \Rightarrow 増水」や「想像 \Rightarrow 超える」など, 川の状況についての投稿が多数あったことがわかる。千曲川の氾濫は 20 時 50 分頃とされているが, 13 時台の時点でも危険な状態であったことも推測される。

共起ネットワークでは 1 つの単語に複数の繋がりがあるとき, 前後の単語の関係が把握しづらい。特に本研究では, 「千曲川」はどのツイートにも出現するため, 「千曲川」と共起する前後の単語の繋がりの把握が困難である。そこで, 各ツイートの単語の出現頻度を用いてクラスター分析を行い, 単語のグルーピングを行った。単語は得られた共起に出現した単語を対象とした。図 3 は円の色でクラスターを表しており, 「上陸」, 「前」, 「水位」, 「上がる」が同じクラスターに属したため, 「上陸前に千曲川の水位が上がる」という繋がった文章として読み取ることができる。

3.2 添付画像の抽出と出力システム

実際にどのような画像が投稿されていたか, R の jpeg パッケージを用いて, ツイートデータから添付画像の抽出を行った。このとき, ツイート内容と添付画像との関係を見つけるために, 各クラスターの単語が使われているツイートの添付画像を抽出した。また, ツイート内容と添付画像との関係を把握しやすくするために, 共起ネットワークと添付画像を同時に出力するシステムを作成した。

2019 年 10 月 13 日 7 時台の共起ネットワークと添付画像を同時に出力するシステムの表示画面を図 4 に示す。画面の左側はその時間帯の共起ネットワークを表示し, 右側は選択したクラスターの単語が含まれるツイートの添付画像を最大 20 枚表示している。画像の上部にはツイートの投稿日時と注目度も表示している。図 4 について, 共起ネットワークでは「北陸新幹線車両センター」や「E7 系」など新幹線水没の報道について多く投稿されていることがわかる。さらに「E7 系」が含まれるツイートの添付画像には新幹線水没の様子を写した画像が多く投稿されていたこともわかる。

この可視化システムの機能として, 対象日時, 共起率の下限, 共起頻度の順位の下限, 最大クラスター数, 画像を

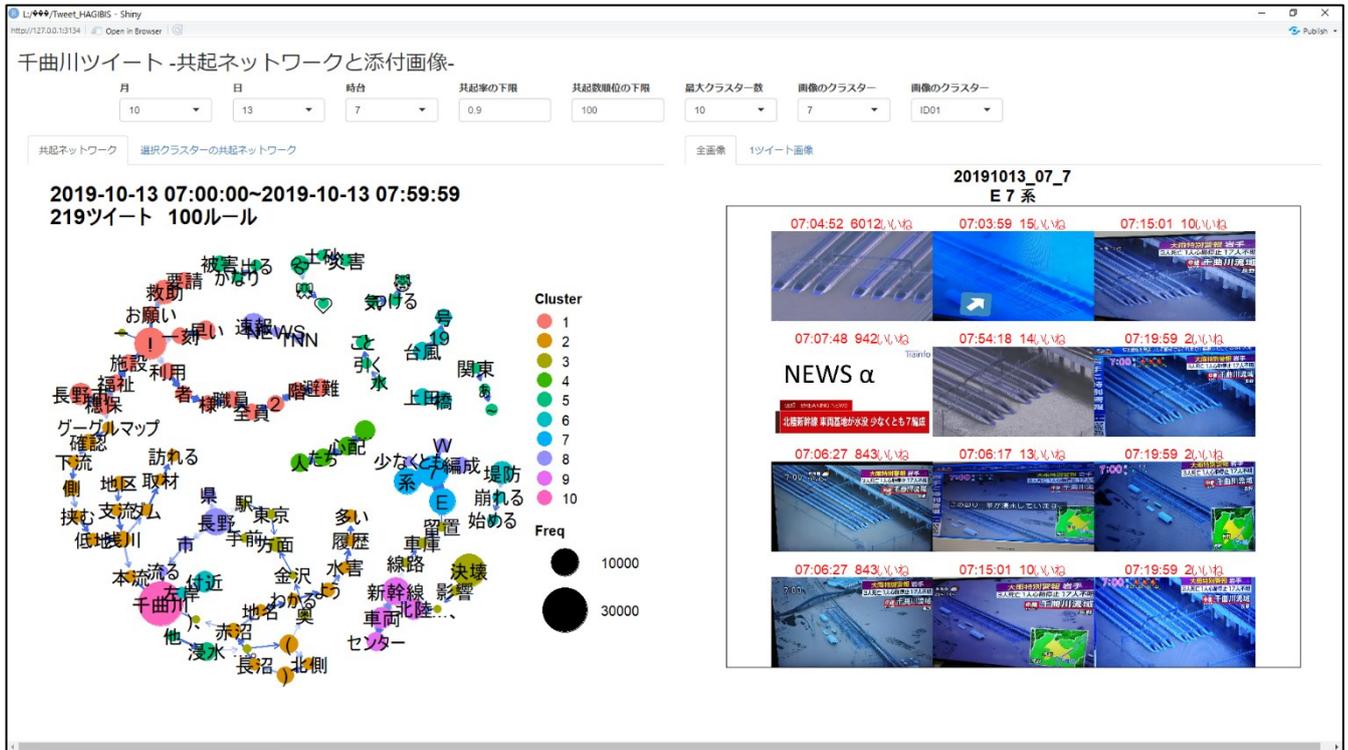


図 4: 共起ネットワークと添付画像を同時に出力するシステムの表示例 (2019 年 10 月 13 日 7 時台)

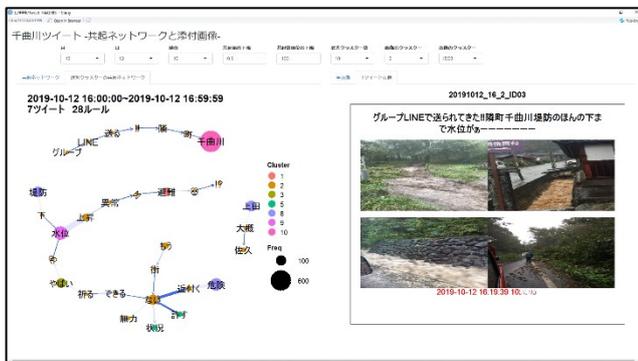


図 5: 氾濫前のツイートと添付画像 (10 月 12 日 16 時台)

本研究でわかった重要な知見としては、ツイート数が急上昇する前の時間帯に重要な情報が隠れている可能性があるということである。図 5 では川の氾濫に関する報道よりも前に、川の水位が上昇している写真、川の流れが速くなっている写真などで、氾濫の危険性を示している投稿を見つけることができた。また、図 6 では岐阜県や長野県が大雨特別警報の発表される前からかなりの大雨になっていたことも予想できる。降水量のマップだけでは九州地方だけに注目してしまうが、共起ネットワークと合わせて見ることで、大雨特別警報がまだ発表されていない地域の中でも危険な地域を発見することができた。降水量とツイート数の推移を見ることで、実際に大雨は特別警報よりも前から発生していることが確認でき、その時間帯の情報を素早く広めることができれば、市民の防災・避難行動を誘発・支

援できると考えられる。

図 6 の可視化システムでは、ある日付の共起ネットワークに出てきた都道府県について降水量とツイート数の推移を確認する分析と、一方である都道府県の降水量とツイート数の推移から注目したい日付について共起ネットワークを確認するといった探索的な分析が可能となっており、このシステムの有効性を示すことができると考えられる。この可視化システムを活用してもらえるように分析内容や可視化方法などを検討していく。

5. まとめ

令和元年東日本台風と令和 2 年 7 月豪雨時のツイートについて分析・可視化を行った。共起ネットワークを作成し、ツイート内容を俯瞰的に可視化した。また、添付画像や降水量データと組み合わせて表示する可視化システムを作成した。対象の日時に合わせたインタラクティブな可視化システムを作成した。

我々は災害時のツイート分類やマッピング[9]、効率的な情報収集[10]、Twitter を利用した災害情報共有システムの構築[11][12][13][14]などの研究に取り組んでおり、本研究で得られた知見をそれらの研究にも反映させていきたい。なお、今回構築したシステムは、収集済みのツイートデータに対する分析結果を可視化するものであるが、今後災害発生時にリアルタイムに分析結果を表示するシステムの構築を目指していきたい。

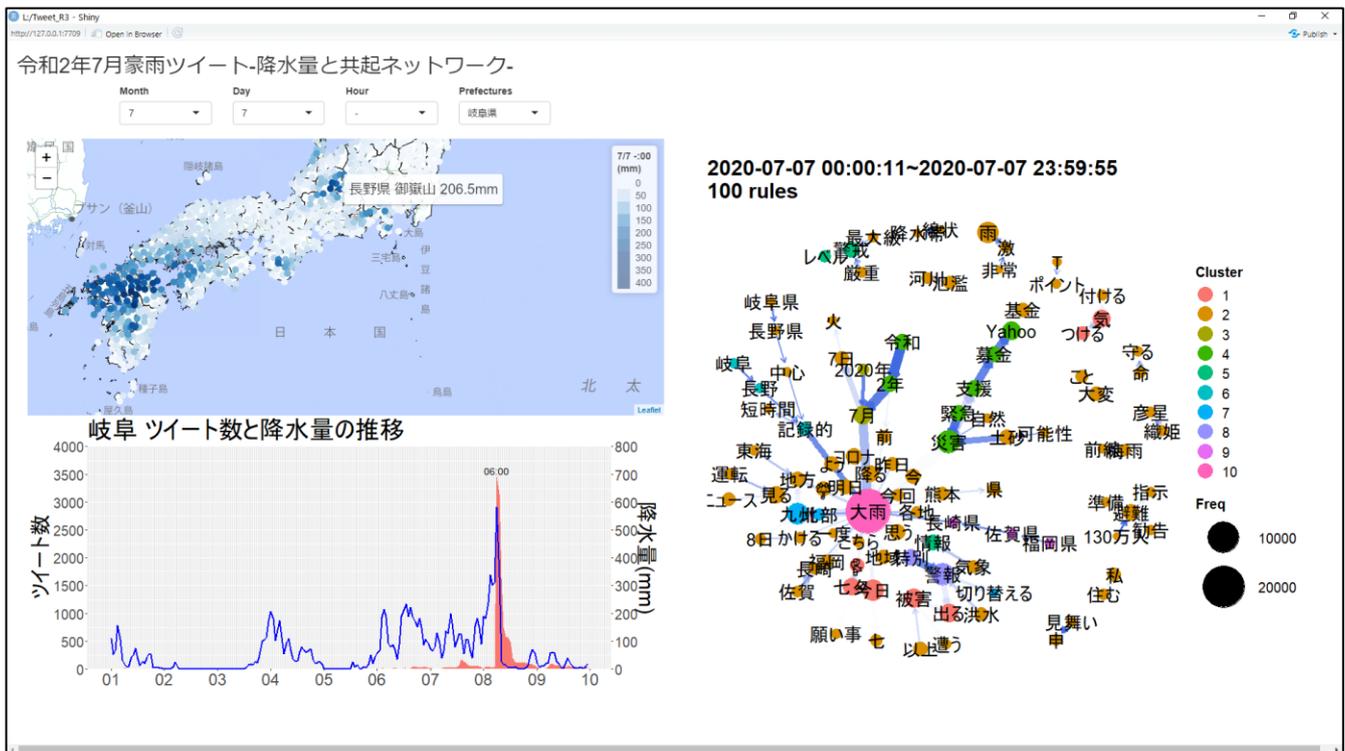


図6: 降水量と共起ネットワークの同時可視化システムの表示例 (2020年7月7日)

謝辞 本研究は、科研費基盤研究(C) 18K11553, 公益財団法人電気通信普及財団研究調査助成, 並びに東海大学総合研究機構の助成を受けて実施した。

文献

- [1] 内田理, 宇津圭祐, “災害時のソーシャルメディア活用,” 電子情報通信学会 基礎・境界ソサイエティ Fundamentals Review, vol.13, no.4, pp.301-311, 2020.
- [2] F. Toriumi, T. Sasaki, K. Shinoda, K. Kazama, S. Kurihara, and I. Noda, “Information Sharing on Twitter during the 2011 Catastrophic Earthquake.” Proc. 22nd International Conference on World Wide Web Companion, pp.1025-1028, May 2013.
- [3] NEC ビッグロブ株式会社, “東日本大震災におけるツイッターの利用状況,” ビッグロブ株式会社, <http://www.biglobe.co.jp/pressroom/release/2011/04/27-1>, 参照 June.24,2019.
- [4] S. Yamada, K. Utsu, and O. Uchida, “An Analysis of Tweets During the 2018 Osaka North Earthquake in Japan - A Brief Report,” Proc. The 5th International Conference on Information and Communication Technologies for Disaster Management, Dec. 2018.
- [5] 水沼友宏, 池内淳, 山本修平, 山口裕太郎, 佐藤哲司, 島田諭, “Twitter におけるバーストの生起要因と類型化に関する分析,” 情報社会学会誌, vol.7, no.2, pp.41-50, 2013.
- [6] S. Yamada, K. Utsu, and O. Uchida, “An Analysis of Tweets Posted During 2018 Western Japan Heavy Rain Disaster,” Proc. 2019 IEEE International Conference on Big Data and Smart Computing, Feb. 2019.
- [7] S. Yamada, K. Utsu, and O. Uchida, “Analysis of Japanese Tweets about 2018 Typhoon Jebi by Creating Co-occurrence Networks,” The 16th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2019), May 2019.
- [8] 気象庁, 過去の気象データ・ダウンロード <https://www.data.jma.go.jp/gmd/risk/obsdl/index.php>
- [9] O. Uchida, T. Rokuse, M. Tomita, Y. Kajita, Y. Yamamoto, F. Toriumi, B. Semaan, S. Robertson, and M. Miller, “Classification and Mapping of Disaster Relevant Tweets for Providing Useful Information for Victims During Disasters,” IIEEJ Transactions on Image Electronics and Visual Computing, vol.3, no.2, pp.224-232, 2015.
- [10] K. Utsu, A. Manaka, K. Nakafuri, and O. Uchida, “Web Application Prototype for Collecting Disaster-related Information Focusing on Tweets Immediately after Retweeting News Posts,” Proc. 4th International Conference on Information and Communication Technologies for Disaster Management, Dec. 2017.
- [11] O. Uchida, M. Kosugi, G. Endo, T. Funayama, K. Utsu, S. Tajima, M. Tomita, Y. Kajita, and Y. Yamamoto, “A Real-Time Information Sharing System to Support Self-, Mutual-, and Public-Help in the Aftermath of a Disaster Utilizing Twitter,” IEICE Transactions on Fundamentals, vol.E99-A, no.8, pp.1551-1554, 2016.
- [12] M. Kosugi, K. Utsu, S. Tajima, M. Tomita, Y. Kajita, Y. Yamamoto, and O. Uchida, “Improvement of Twitter-based Disaster-related Information Sharing System,” Proc. 4th International Conference on Information and Communication Technologies for Disaster Management, Dec. 2017.
- [13] 内田 理, 小杉 将史, 山口 良二, 長 幸平, “災害情報共有システム DITS・DIMS の平時共用システムへの拡張,” 電子情報通信学会技術研究報告, vol.119, no.226, Oct. 2019.
- [14] 小杉 将史, 内田 理, “みずれば: 河川情報共有システム,” 電子情報通信学会技術研究報告, vol.119, no.400, Jan. 2020.