

情報伝播ネットワークの推定と災害による変化

The Impact of Disaster toward Information Flow Network on Social Media

上子 優香*¹ 鳥海 不二夫*¹ 吉田 光男*² 大橋 弘忠*¹
Yuka Kamiko Fujio Toriumi Mitsuo Yoshida Hirotada Ohashi

*¹東京大学大学院工学系研究科 システム創成学専攻

Department of Systems Innovation, School of Engineering, The University of Tokyo

*²豊橋技術科学大学 情報・知能工学系

Department of Computer Science and Engineering, Toyohashi University of Technology

Social media have grown to be one of the most popular media under crisis. While previous studies focused on the scale of information diffusion on social media, this study focuses on the path of information dissemination and describes the structural changes occurred to both the information flow network and the reply network on Twitter by the 2016 Kumamoto earthquakes. As a result, we clarified three facts. First, users tend to characterize their friends as either who to reply or who to share information. Second, network density and assortativity index shift in the opposite direction in the short and long run. Third, on the information flow network, edges from verified users are more persistent than those from general users.

1. 序論

災害時における情報の共有手段として、ソーシャル・メディアが浸透しつつある。これまでも災害時にユーザによって多くの情報がソーシャル・メディアを通じて発信、拡散され、災害情報の共有に役立てられることが Vieweg ら [Vieweg 10], Heverin ら [Heverin 10], 篠田ら [篠田 13] などによる研究で示されている。しかしながら従来の研究では災害時の情報拡散規模に注目されることが多く、情報伝播経路に重きをおいて情報伝播ネットワークを作成し、災害に因るネットワーク構造の変化を明らかにした例は見られない。情報の伝播経路を明らかにすることは、ミクロな視点で情報拡散を捉える上で必要不可欠であり、今後ソーシャル・メディアの利活用を促進させることを目標に、ユーザの理想的な利用方法を議論する上で重要である。

本研究では平成 28 年熊本地震を災害事例として取り上げ、ツイッターデータを用いて災害前後の情報伝播ネットワークおよびリプライネットワークの変化を、1ヶ月スパンの長期的変化と1日スパンの短期的変化の2つの時間幅で捉え分析する。さらに情報伝播ネットワークとリプライネットワークを比較することによって、共有行動とリプライ行動の違いについても考察する。

2. 分析に用いるデータとネットワーク

2.1 データ概要

分析には代表的なソーシャル・メディアであるツイッターのデータを用いる。まず 541,134 ユーザに関するフレンド情報を収集し、それらを対象ユーザとする。情報伝播ネットワーク作成にはリツイートデータ、リプライネットワークにはリプライデータを使用する。リツイートデータ、リプライデータは共に Twitter API を利用して収集し、リツイートデータは 2016/3/15 - 5/13 に対象ユーザが関与し投稿から 1 週間以内

表 1: フレンド, リツイート, リプライデータの重複ユーザ数

比較データ	重複ユーザ数
フレンド & リツイート	50,157
フレンド & リプライ	167,733
フレンド & リツイート & リプライ	41,099

に 10 回以上リツイートされた投稿に関する 381,159,908 件のデータを収集した。リプライデータは同期間に対象ユーザが行った 16,218,262 件のデータを収集した。なおフレンド, リツイート, リプライデータの重複ユーザ数は表 1 の通りである。

これらのデータを基に、平成 28 年熊本地震では 2016/4/14, 4/16 に震度 7 を記録していることから、長期的なネットワークの変化分析のために

- 災害前 (2016/3/15 - 4/13) 情報伝播ネットワーク
- 災害後 (2016/4/14 - 5/13) 情報伝播ネットワーク
- 災害前 (2016/3/15 - 4/13) リプライネットワーク
- 災害後 (2016/4/14 - 5/13) リプライネットワーク

の 4 種類のネットワークを、短期的なネットワークの変化分析のために

- 4/12 - 18 における各日の情報伝播ネットワーク
- 4/12 - 18 における各日のリプライネットワーク

の 14 種類のネットワークを用意し、計 18 種類のネットワークの比較分析を行う。なお平成 28 年熊本地震の概要については後述する。

2.2 災害前・災害後情報伝播ネットワーク作成

災害前・災害後情報伝播ネットワークとは、災害前後 1ヶ月において情報の伝播経路がどのように変化したかを分析する為のネットワークを指す。両ネットワークはフレンド情報を入手している 541,134 ユーザをノードとして、図 1 のフローチャートに従い有向リンクを作成した。なおリンクが 1 つもないノードは作成しない。図 1 に記された提案手法とは、[Kamiko 16] で提案した同一投稿の共有時刻の差に注目した手法である。

連絡先: 東京大学大学院工学系研究科システム創成学専攻

〒 113-8656 東京都文京区本郷 7-3-1

Email: tori@sys.t.u-tokyo.ac.jp

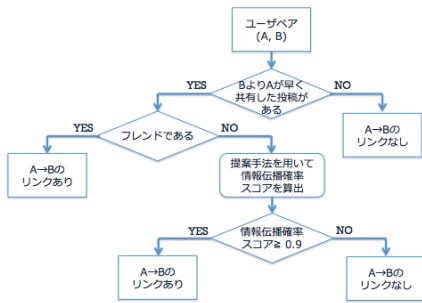


図 1: 情報伝播ネットワークのリンク推定フロー

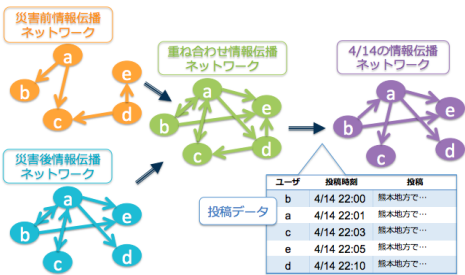


図 2: 各日の情報伝播ネットワークの作成方法

2.3 各日における情報伝播ネットワークの作成

災害直前から直後の 4/12 - 18 各日における情報伝播ネットワークを作成する。初めに前述した災害前情報伝播ネットワークと災害後情報伝播ネットワークのどちらか一方にでも存在したノード、リンクを作成し、重ね合わせ情報伝播ネットワークとする。重ね合わせる理由は、4/12 - 18 は災害前情報伝播ネットワークから災害後情報伝播ネットワークへの移行期間にあたりと考えられるからである。この重ね合わせ情報伝播ネットワークにあるリンクが、情報が伝播する可能性のあるリンクであると仮定する。そして重ね合わせ情報伝播ネットワークに存在するリンクの内、各日において情報が流れているリンクを投稿データを用いて調査し作成する。

図 2 を用いて具体的方法を説明する。右下に「熊本地方で…」という同一投稿に関する投稿データがあるが、投稿時刻順からは

$(b \rightarrow a), (b \rightarrow c), (b \rightarrow e), (b \rightarrow d), (a \rightarrow c),$

$(a \rightarrow e), (a \rightarrow d), (c \rightarrow e), (c \rightarrow d), (e \rightarrow d)$

への情報伝播が可能性として考えられる。そこでこの内、重ね合わせ情報伝播ネットワークにリンクが存在する

$(b \rightarrow a), (b \rightarrow e), (a \rightarrow c), (a \rightarrow e), (a \rightarrow d)$

の 5 つのリンクを作成する。このように対象とする日に行われた投稿データ全てに対して同処理を施すことで、各日の情報伝播ネットワークを作成する。

2.4 リプライネットワークの作成

情報伝播ネットワーク同様、対象とした 541,134 ユーザをノードとして、対象期間に 1 度でもリプライがあったノード間に有向リンクを作成する。災害前・災害後リプライネットワー

表 2: 情報伝播ネットワークとリプライネットワークの重複ノード・リンク数および重複率

NW: ネットワークの略。

括弧内は、比較ネットワーク A&B における、 $\frac{|A \cap B|}{|A|}$ を表す。

比較ネットワーク (NW)	重複ノード数	重複リンク数
災害前情報伝播 NW & 災害前リプライ NW	30,802 (75%)	6,061 (1.6%)
災害後情報伝播 NW & 災害後リプライ NW	30,130 (79%)	9,663 (2.9%)

クには 2016/3/15 - 4/13 と 2016/4/14 - 5/13 のリプライ情報を、4/12 - 18 の各日リプライネットワークには各日のリプライ情報を利用した。なお情報伝播ネットワーク同様、リンクが 1 つもないノードは作成しない。

2.5 情報伝播・リプライネットワークのノード・リンク重複率

前述の方法により作成された災害前・災害後情報伝播ネットワークと災害前・災害後リプライネットワークのノード、リンクの重複率について分析する。重複率は表 2 の通りで、災害前後の情報伝播ネットワークとリプライネットワークのノード重複率は 75 - 80% と高いが、リンク重複率はわずか 2 - 3% に過ぎないことが判明した。これは多くのユーザは情報伝播目的のユーザとリプライ目的のユーザを別に有していることを意味する。つまりソーシャル・メディアにおける情報伝播を分析する際には、全てのリンクを同質のものとして扱うことは不適であり、リンクの主たる目的が情報伝播かリプライか、を区別すべきことを示している。

3. 情報伝播・リプライネットワークの災害による変化

3.1 平成 28 年熊本地震概要

本研究で災害事例として取り上げた平成 28 年熊本地震は、2016/4/14 以後に発生した熊本県熊本、阿蘇地方及び大分県中部を震源とした一連の地震である。国から激甚災害に指定され、4/14 21:26 及び 4/16 01:25 に最大震度 7 を記録している。

3.2 出次数相関、入次数相関、出入次数相関、入出次数相関 [Newman 03]

作成したネットワークを比較分析する指標値として、ノード数、リンク数、ネットワーク密度および次数相関を用いる。出次数相関、入次数相関、出入次数相関、入出次数相関の定義は次の通りである。有向ネットワークでは各ノードに出次数および入次数が与えられる。ノード間の次数相関を考える際には、リンク元・リンク先ノードそれぞれに対して出次数・入次数の選択が可能であるため、 2^2 種類の次数相関が定義可能である。

- 出次相関：リンク元の出次数とリンク先の出次数
- 入次相関：リンク元の入次数とリンク先の入次数
- 出入次相関：リンク元の出次数とリンク先の入次数
- 入出次相関：リンク元の入次数とリンク先の出次数

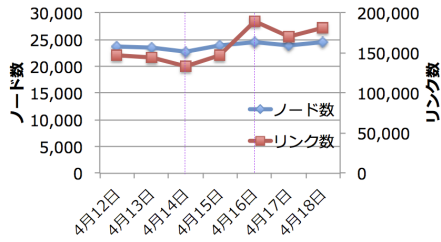


図 3: 短期的な情報伝播ネットワークのノード数, リンク数の変化

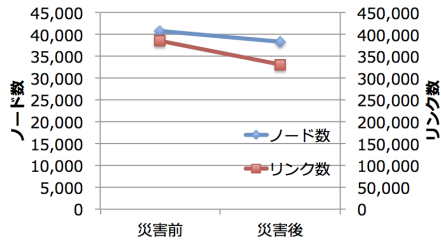


図 4: 長期的な情報伝播ネットワークのノード数, リンク数の変化

次数相関の定義は, ネットワークのリンク数を M , リンク元の次数 k_m , リンク先の次数 l_m とおくと,

$$\frac{M^{-1} \sum_m k_m l_m - \{M^{-1} \sum_m \frac{1}{2}(k_m + l_m)\}^2}{\frac{1}{2} M^{-1} \sum_m \{(k_m)^2 + (l_m)^2\} - \{M^{-1} \sum_m \frac{1}{2}(k_m + l_m)\}^2}$$

であり, 目的とする次数 (出次数, 入次数) にあわせて k_m, l_m の値を定めることで, 4 種類の次数相関が算出可能である.

3.3 ノード数, リンク数, ネットワーク密度変化

図 3, 図 4 は情報伝播ネットワークのノード数とリンク数の変化を比較したものである. 図 3 は各日のネットワーク比較であり, 図 4 は災害前 (2016/3/15 - 4/13), 災害後 (2016/4/14 - 5/13) の比較である. 図 3 より情報伝播ネットワークは短期的にはノード数はほとんど変化しないが, リンク数が増加するため, ネットワーク密度が上昇することがわかる. しかし長期的には図 4 よりリンク数が減少しネットワーク密度が低下する. 情報伝播ネットワークの密度が短期的に上昇することは, 災害時に通常時よりも情報を発信・受信する相手が増加する, ということである. 発信側から見れば情報の受け手が増える, 受信側から見れば入手する情報の発信者の多様性が高まると解釈できる. これは平均入出次数が 4/15 - 16 にかけて 12.3 から 15.5 に増加することからも確認できる.

次に図 5 と図 6 よりリプライネットワークの変化を分析する. 図 5 より短期的には前震が発生した 14 日, 本震が発生した 16 日にノード数, リンク数が増加しており, 災害時に通常時以上に人々のリプライに対する積極性が高かったことが窺える. 情報伝播ネットワークでは変化の少ない 14 日に最も多いノード数, リンク数を記録していることから, 人々がリプライ行動を情報伝播行動よりも早く起こすということが考えられる. すなわち前震が発生した時刻が 21 時 26 分と夜遅く, 緊急性の高い会話がリプライで発生直後に行われるのに対し, 情報収集行動はそれよりも遅れて 00 時 00 分を回って以降に主に行われたため, リンク数の増加が見られない可能性が高い. 長期的には図 6 よりノード数は変わらずリンク数が微増したことがわかる.

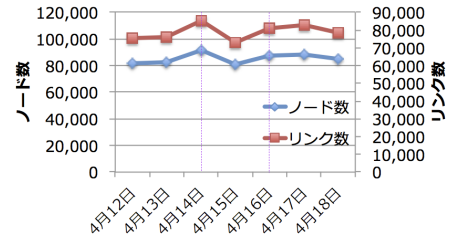


図 5: 短期的なリプライネットワークのノード数, リンク数の変化

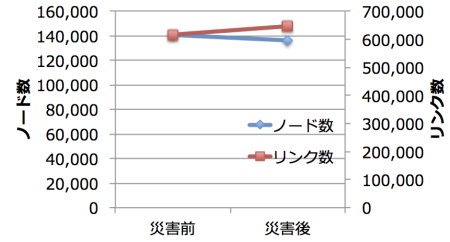


図 6: 長期的なリプライネットワークのノード数, リンク数の変化

3.4 出次数相関, 入次数相関, 出入次数相関, 入出次数相関の変化

図 7 から図 10 はネットワーク別の次数相関推移であり, 短期的な日別の次数相関の値推移と, 長期的な災害前後 1ヶ月間の次数相関の値比較を示す. 図 7 と図 8 より情報伝播ネットワークは短期的には次数相関が低下するが, 長期的には上昇することが読み取れる. 次数相関が低下するということは異質の者との交流が増えるということを示す. すなわち災害時には短期的には普段は情報伝播しないような異質なユーザーの間でも情報伝播を行うため短期的には次数相関が低くなる. しかし長期的には災害時の交流を通して, それらの新たに繋がった人との繋がりを全体的に強めるために次数相関が高くなる方向へ働いていると解釈できる. 他にもデマの発生によってユーザーの信頼性を求めるようになり, 近しい者以外を排除する方向へ進むため, 次数相関が高くなるとも解釈可能である.

リプライネットワークでは短期的には図 9 より特に入出次数相関で日ごとの変化が大きい. しかしながら 4/12 - 13 にかけての入出次数相関の上昇, 入次数相関の上昇など, 災害を要因としない変化が観測されている. 図 10 より長期的には入出次数相関と入出次数相関の値に変化が見られるが, 短期的変化の不規則性と併せて考えると災害に起因する変化とは言い難い.

3.5 情報伝播と認証済アカウント

最後に認証済アカウントからの情報伝播が一般アカウントからよりも発生しやすいかを分析する. まずリンクを一般ユーザー間のものであるか, 認証済アカウントが関わっているかによって 4 種類に分類し, 種類別にリンク継続率を算出する. 認証済アカウントとはツイッターから企業や著名人の公式アカウントであることを認証されたアカウントであり, これらを認証ユーザーと表記する. なお対象ノードが認証ユーザーであるかは, ツイッターのユーザーページにアクセスすることで判定した. 4 種類の分類及び災害前後の情報伝播ネットワークにおけるリンク数, リンク継続率は表 3 の通りである. 括弧内は各ネットワークにおける全体のリンク数を占める割合を示す.

一般 → 認証へのリンク数および認証 → 認証のリンク数は全体の 1%未満であり, 総数として非常に限定的である. これ

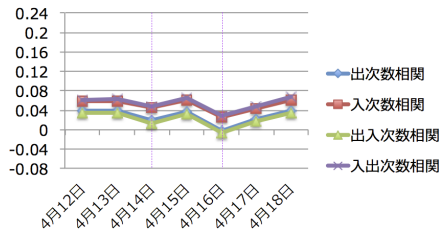


図 7: 短期的な情報伝播ネットワークの次数相関の変化

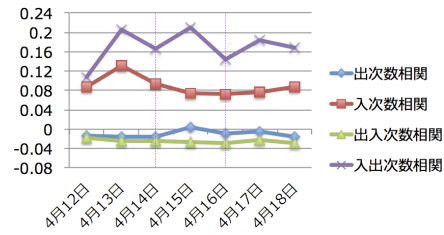


図 9: 短期的なプライベートネットワークの次数相関の変化

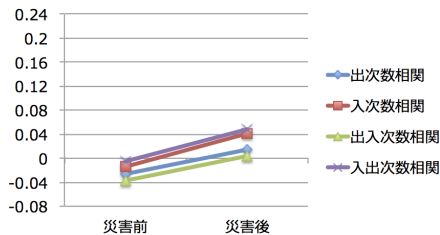


図 8: 長期的な情報伝播ネットワークの次数相関の変化

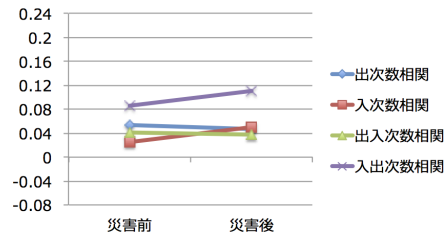


図 10: 長期的なプライベートネットワークの次数相関の変化

表 3: 災害前後の情報伝播ネットワークにおける認証ユーザのリンク数およびリンク継続率

NW: ネットワークの略

リンクの種類	災害前 情報伝播 NW	災害後 情報伝播 NW	リンク 継続率
一般 → 一般	369,839(96.1%)	313,712(94.6%)	46.8%
一般 → 認証	706(0.2%)	775(0.2%)	53.3%
認証 → 一般	13,878(3.6%)	16,460(5.0%)	57.5%
認証 → 認証	543(0.1%)	676(0.2%)	58.4%
全体	384,966(100.0%)	331,623(100.0%)	47.2%

は認証ユーザが情報発信により慎重であることが原因と考えられ、認証ユーザが特定の一般ユーザからの情報を安易に伝播させることは少ないとわかる。リンク継続率に注目すると、一般 → 一般への継続率が 46.8%であるのに対して、認証 → 一般でリンク継続率は 57.5%と一般ユーザ同士に比べて 10%以上継続率が高いことが明らかになった。つまり認証ユーザからの情報伝播リンクは一般ユーザからに比べ安定的であり、両リンクを区別することによって、より正確に情報拡散現象を捉えることができる可能性が高いことが明らかになった。

4. 結論

本研究では、平成 28 年熊本地震発生時におけるツイッターの情報伝播ネットワークおよびプライベートネットワークを作成し災害によるそれらの変化を分析した。情報伝播ネットワークの作成には著者ら [Kamiko 16] の手法を採用した。これまでの研究との差異は情報の拡散規模に注目し、投稿作成者とリツイート者を結んだネットワークを作成するのではなく、情報の伝播経路に注目し、ユーザ間で情報が伝播したかを推定し情報伝播ネットワークを作成した点にある。結果として多くのユーザにとって情報を入手する相手とリブライをする相手は別であること、災害による情報伝播ネットワークの密度と次数相関の変化が短期的視点と長期的視点で逆転すること、認証ユーザからの情報伝播が一般ユーザよりも継続して発生し易いこと、

が判明した。他の災害事例に対しても同様の手法を適用し、同様の変化が見られるかを調査することは今後の課題である。

5. 謝辞

本研究は日本学術振興会 課題設定による先導的人文・社会科学推進事業「リスク社会におけるメディアの発達と公共性の構造転換」プロジェクトの一部として行われた。

参考文献

- [Heverin 10] Heverin, T. and Zach, L.: *Microblogging for crisis communication: examination of Twitter use in response to a 2009 violent crisis in the Seattle-Tacoma, Washington, area*, ISCRAM (2010)
- [Kamiko 16] Kamiko, Y., Yoshida, M., Ohashi, H., and Toriumi, F.: Uncovering information flow among users by time-series retweet data: Who is a friend of whom on Twitter?, in *Big Data (Big Data), 2016 IEEE International Conference on*, pp. 2500–2504 IEEE (2016)
- [Newman 03] Newman, M. E.: Mixing patterns in networks, *Physical Review E*, Vol. 67, No. 2, p. 026126 (2003)
- [Vieweg 10] Vieweg, S.: Microblogged contributions to the emergency arena: Discovery, interpretation and implications, *Computer Supported Collaborative Work*, pp. 515–516 (2010)
- [篠田 13] 篠田 孝祐, 榎 剛史, 鳥海 不二夫, 風間 一洋, 栗原 聡, 野田 五十樹, 松尾 豊: 東日本大震災時における Twitter の活用状況とコミュニケーション構造の分析, *知能と情報*, Vol. 25, No. 1, pp. 598–608 (2013)