

Twitterデータを用いた水害時の 住民の反応と行動の時系列分析

柳森 和真¹・井料 隆雅²

¹学生会員 神戸大学 大学院工学研究科市民工学専攻 (〒657-0058 神戸市灘区六甲台町1-1)
E-mail:159t144t@stu.kobe-u.ac.jp

²正会員 神戸大学 大学院工学研究科市民工学専攻 (〒657-0058 神戸市灘区六甲台町1-1)
E-mail:iryoy@kobe-u.ac.jp

水害が多発する日本では、それらによる経済損失や「逃げ遅れ」による人的被害も甚大である。水害時の被害を軽減するための方策として、緊急時に発令される避難勧告等の避難情報が挙げられる。近年では、避難情報の発令手段として、携帯電話等のエリアメールや、SNSなど情報技術を用いたものが活用されている。これらの新しい情報伝達手段の効果評価や改善のためには住民の反応を知ることが重要である。本研究では、平成26年8月に発生した対象水害前後に収集したTwitterデータと避難情報発令状況を用い、Twitterユーザが避難情報にどう反応したかを分析した。結果として避難情報発令時のTwitterの発言数は避難情報の発令対象者数の時系列変化に類似し、Tweetの内容に関してはTwitter特有のバイアスが観測されただけでなく、避難行動に関する過去の知見¹⁾と整合性の取れるものとなった。

Key Words : *Twitter, typhoon, evacuation orders, evacuation advisories, evacuation behavior*

1. はじめに

日本は毎年平均して約3つの台風が上陸²⁾する世界的に水害が多い国である。水害による被害も甚大で、2014年8月に広島市近郊で発生した大雨による土砂災害など、多大な被害をもたらすだけでなく人々の生活基盤に打撃を与えるものも少なくない。

水害による人的被害を軽減するためには住民に避難勧告等の災害情報を迅速かつ確実に伝達することがコスト面では安価かつ有力である。情報通信技術の発展は迅速な情報伝達を可能にしてきた。近年運用が始まった携帯電話による災害情報同報システム(エリアメール³⁾、緊急速報メール⁴⁾等)はこの好例であるといえよう。

災害情報伝達システムが人命被害を軽減させるためには、「災害情報に対する住民の反応(フィードバック)」を知ることが欠かせない。具体的には情報が被災地の住民にいつどのように認知され、それがどのような行動を招いたかなどである。これらを知る代表的手段として、住民に対するアンケート調査が行われてきた。しかし、記憶違いによる精度の低下や高額な調査費用などのため、必ずしも万能ではない。他の災害と同様に、水害においても「逃げ遅れ」による人的被害が多く報告されている⁵⁾。その原因と対策を分析するためにも、アン

ケートでは得られないような十分な時間精度を持つデータを安価に入手することは有益である。

十分な時間精度を持ち、安価に入手可能なデータとしてTwitterデータを用いる研究が近年盛んである。Twitterデータは突発的な事象に対する人々の反応を観測する手段としての有用である。これは既往研究からも既に示されている。例としては、地震発生がTwitterデータで検出できるとした研究⁶⁾や、地震発生時のTwitterテキストデータから帰宅困難者の帰宅行動分析を行った研究⁷⁾、地震や台風災害時に特定のTweet発言について時系列分析を行うことで、災害時の被災者の興味関心を推測を分析し更に、ジオデータ付きのTweetが多く吐かれた集中地から地震の震源や台風のコースと比較した研究⁸⁾等が挙げられる。

本研究では平成26年8月の台風11号を対象に、Twitterのデータを用いて、避難情報に対する住民反応の分析を試みる。「避難」というキーワードを含むTweetデータを部分的に取得し、各Tweetに含まれる災害情報に対する住民の反応の種類に応じ、機械的および人間の手によってそれを分類した。どのような反応が得られているか、Tweetに出現する地名の分布、発言の時系列の分析を行い、Twitterデータにより災害情報に対する住民反応がどれだけ取得できるかを検討する。

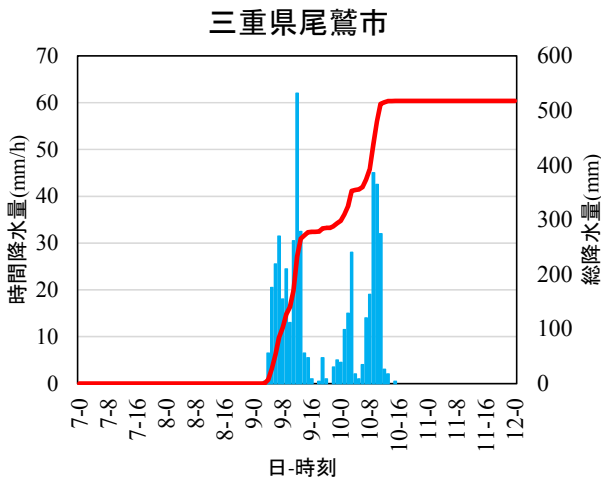


図-1 三重県尾鷲市の時間降水量及び総降水量⁹⁾

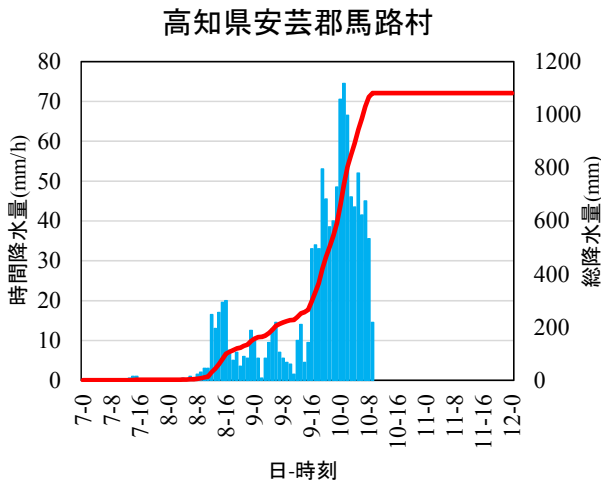


図-2 高知県安芸郡馬路村の時間降水量及び総降水量⁹⁾

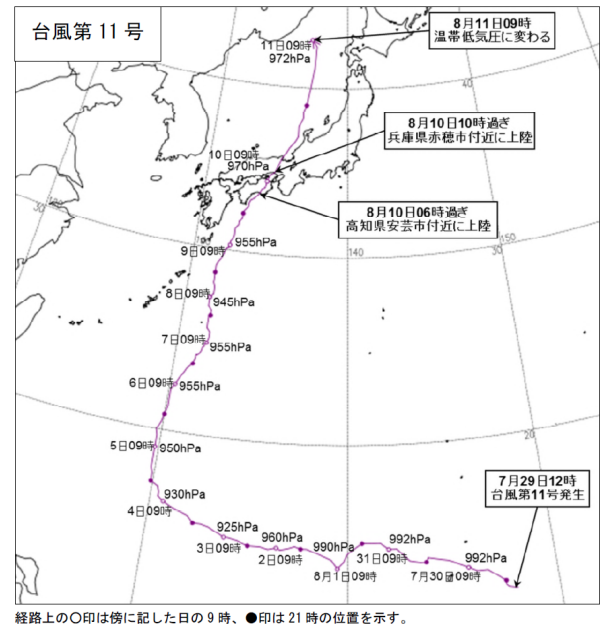


図-3 台風11号の進路及び中心気圧の概要図⁹⁾

2. 分析対象概要

(1) 対象期間

平成26年8月9日から11日を対象期間とする。当該期間中には平成26年台風11号が日本列島に上陸し、三重県尾鷲市で85.0mm⁹⁾、高知県安芸郡馬路村で80.5mm⁹⁾の時間降水量を記録し、西日本を中心に各地で猛烈な大雨や強風による被害が発生した。図-1、図-2に上記市町村の雨量状況を、図-3には台風11号の進路及び中心気圧を示す。図-1から図-3より今回対象とした台風は四国の南岸から日本列島を縦断し、9日の朝から10日の朝にかけて降水量が急激に増加傾向にあることが分かる。

(2) 避難勧告、避難指示の発令状況

後述するTwitterデータと比較分析を行うために、台風11号上陸時に発令された避難勧告、避難指示などの避難情報の発令の時系列的な把握、及び発令対象者数から発令の規模について把握する。避難情報に関するデータとしては、内閣府が公表したもの¹⁰⁾を用いる。図-4、図-5には対象期間中に発令された避難勧告、避難指示の発令対象者数及び、累積人数の時間変化を示す。下図より避難勧告に関しては、9日9時～21時頃及び10日6時～15時頃に、避難指示に関しては9日15時～9日21時頃に発令が集中し、それによる避難情報の発令対象者数が増加していることが分かる。また対象期間中の8月9日17時20分に三重県四日市市全域約31万人対象、同日18時30分には三重県鈴鹿市全域約20万人対象の避難指示が発令された。四日市市においては、8月9日17時台には39.5mm/hの時間降水量を観測していることから、急激な雨量を事前に予測して、迅速に避難情報を発令していると考えられる。

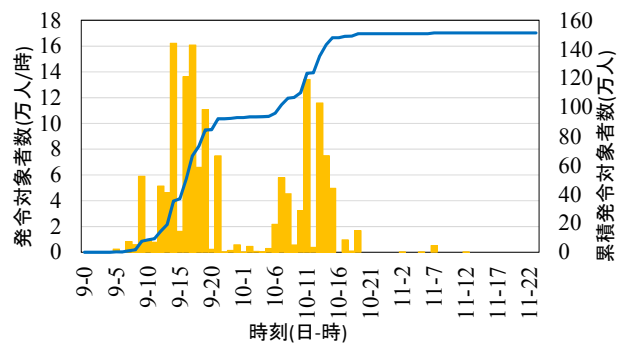


図-4 避難勧告発令対象者及びその累積人数の時間変化

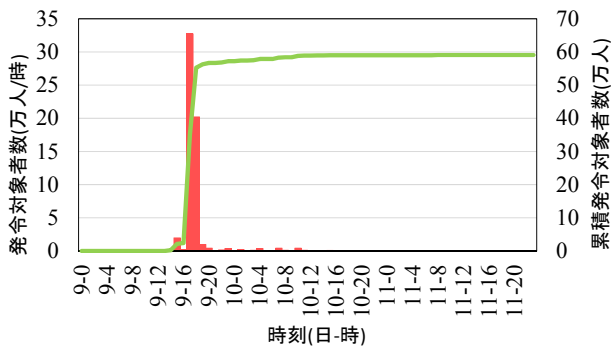


図5 避難指示発令対象者及びその累積人数の時間変化

(3) 対象とするTwitterデータ

9日18時14分から11日23時59分に「避難」というフレーズを含むTweetのうち、実際に収集できた65,168Tweet対象とする。これらのTweetは30秒に1回の間隔で、取得時点から遡って最大20 tweetを取得した。今回は簡便な手法で取得したので、その技術的制約の関係で30秒間に20 tweetを超えた場合は取得できていない。よって、時間帯によっては全ての対象Tweetを取得出来ていない。

3. 分析方法

(1) Tweet数と避難情報発令履歴との関係分析

1時間ごとの Tweet 数と避難指示・避難勧告が発令されている地域の対象人口(本研究では避難情報発令対象者数と定義する)を時系列で比較し傾向を分析した。時間あたり Tweet 数 v_t (tweet/hour)は、隣接する Tweet の間隔から推定した。具体的には以下に示す式(1)で計算した。

$$v_t = \left(\frac{\Delta t}{n} \right)^{-1} * 60^2 \quad (\text{tweet/hour}) \quad (1)$$

ただし、 n は取得された Tweet の間隔数、 Δt はその Tweet 間隔の平均(sec)である。

(2) 抽出した一部の Tweet の内容の分析

取得された Tweet に含まれる避難情報に対する住民の反応を知るために、取得された Tweet を予め定めた基準に従って分類した。分類は手動で行った。具体的な手順を以下に示す。

1. Tweet のサンプリング：手作業の量を削減するため、各 30 秒で取得された最大 20 tweet の中で最初に取得されたもののみをサンプリングした。これにより 10,533 tweet が抽出された。
2. 住民の反応と無関係な Tweet の分類：他者への返信、リツイート、リンクを含む Tweet を機械的に、水害と関連ない Tweet や同一ユーザーにより反復される

同一内容の Tweet を手作業で選んだ。以上で選ばれた Tweet は水害による住民の反応と無関係とし、それにラベル Z をつけた。

3. 住民の反応と関係がありそうな Tweet の分類：表-1 に示す分類基準に従って手作業で分類をし、それぞれ、A~Eの主ラベルと、1~5の副ラベルをつけた。1個のTweetには複数のラベルがつけられることもある。表-1には主ラベルA~E、副ラベル1~5の分類のラベル付けの基準を示し、図-6には主ラベル分類の手順をフローチャートで示した。

表-1 Tweet分類のラベル付けの基準とその結果

ラベル		主分類の基準	副分類の基準
主	副		
A	-	避難情報を得ている	-
B	1	避難情報の入手手段がわかる	エリアメール等
	2		防災無線・行政広報
	3		インターネット
	4		TV・ラジオ
	5		その他
C	-	避難情報発令地区を把握している	-
D	1	避難情報への反応に関する情報を含む	避難している(た)
	2		避難を検討中
	3		避難に疑問や不安を感じている
	4		避難していない
E	-	避難情報への感想を述べている	-
Z	-	水害の避難と無関係	-

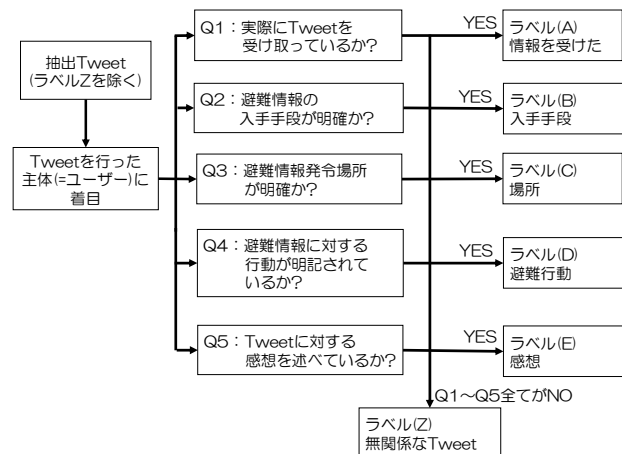


図-6 主ラベル分類の方法

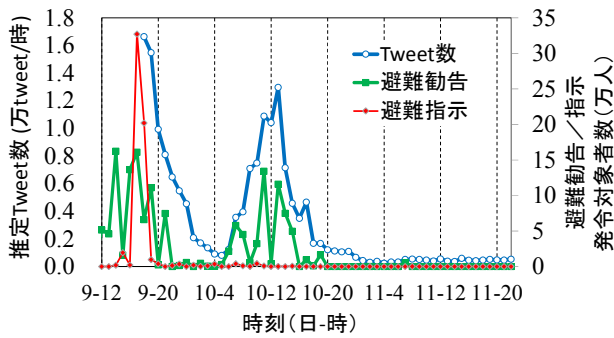


図-7 避難情報発令対象者数と推定Tweet数の時間変化

4. 分析結果

(1) Tweet数と避難情報発令履歴との関係分析

図-7には時間ごとの避難指示及び避難勧告の発令対象者数と推定Tweet数の時系列比較を示す。図-7より、避難情報が多く発令されている9日午後と10日12時前後に推定Tweet数の増加が見られる。特に後者の時間帯ではその増加が明確であり、推定Tweet数の増加は避難情報対象者数の増加とほぼ同一の変化をしているようにみられる。しかし推定Tweet数のピーク時刻は、必ずしも避難情報発令対象者数のそれと一致しておらず、減少に関しては、およそ1時間から3時間程度、変化に遅れが見られることから、Twitter利用者は避難情報発令後もそれらに対して反応をしていると考えられる。また、対象期間中に発生したTweet数は約17万tweetと推定できる。本研究ではその約17万tweetを推定Tweetと定義する。

(2) Tweet内容のラベル分類による集計結果

3章(2)で述べた分類基準にのっとり分類を行った、対象Tweetの結果を表-2に示す。表の値は述べTweet数であることに注意すると、最も多く貼られたラベルはZ（無関係なTweet）であり全体の86%を占める。それ以外のTweet（1,487 tweet）のうち、避難情報の入手手段を示すラベルB、発令場所に関するラベルC、避難行動・避難情報への反応に関するラベルDに分類された避難情報に対する住民の反応として重要な情報を含むTweetはそれぞれ300 tweet前後あった。Tweetの取得期間における「避難」というキーワードを含むTweet数の総数は約17万tweetと推定されるので、今回は「避難」を含むTweetの約6%（10,533 tweet）にラベル付けしたと推定できる。「避難」を含むTweetの全てを分析すれば、拡大係数をKとすると、式(2)より、これらについてそれぞれ5,000程度のTweetが得られたと期待できる。

$$K = \frac{\text{期間中の総Tweet数}}{\text{ラベル分類を行ったTweet数}} \quad (2)$$

表-2 主ラベル、副ラベルに分類されたのべTweet数

ラベル		主分類の基準	副分類の基準	ラベルされたTweetの数
主	副			
A	-	避難情報を得ている	-	1402
B	1	避難情報の入手手段がわかる	エリアメール等	219
	2		防災無線・行政広報	32
	3		インターネット	18
	4		TV・ラジオ	13
	5		その他	4
C	-	避難情報発令地区を把握している	-	384
D	1	避難情報への反応に関する情報を含む	避難している(た)	33
	2		避難を検討中	32
	3		避難に疑問や不安を感じている	88
	4		避難していない	135
E	-	避難情報への感想を述べている	-	406
Z	-	水害の避難と無関係	-	9046

各ラベル別の割合に関しては、ラベルBの避難情報の入手手段においては、副ラベル1の「エリアメール等」に約76.6%に当たるTweetが分類され、ラベルBの大半のTweetはエリアメールで避難情報を入手していることが分かる。次いで副ラベル2の「防災無線・行政広報」が約11.1%となった。ラベルDについては、副ラベル4の「避難行動をしていない」が約46.9%あり、集計されたTweetの大半は避難情報が発令されている状況下でも関わらず避難行動を行っていないことが分かった。避難行動に積極的な副ラベル1, 2に関してはラベルDの全数に対してそれぞれ約11.4%, 約11.1%となり、Tweetの内容を踏まえると全体的に避難行動には積極的であることが観測される。

(3) 各ラベルの時系列分析結果

(2)の集計方法は、実際に観測されたTweetのみを数え上げた。そこで避難情報に対する住民の反応として重要な情報を含むTweetを含む、避難行動及び、避難情報の伝達、入手に関するラベルB, C, Dの時系列変化を分析する。本研究では2時間毎の各ラベルのTweet数を推定Tweet数との比を用いた(3)式を用いて推定した。

$$x_i = \frac{l_i * e_i}{m_i} \quad (3)$$

ここで、 x_i は推定されるラベルの2時間ごとのTweet数(tweet/2hour)、 l_i は当該ラベルの2時間ごとの実測値(tweet/2hour)、 e_i は式(1)で求めた推定Tweet数の2時間毎のTweet数、 m_i は2時間ごとの実測Tweet数とする。図-8にこれらの結果を示す。図-8より、各主ラベルのピークの時間帯はほぼ一致する。しかし、ピーク時刻においては、避難情報の発令地域を発言しているラベルCが最も多く、次いで避難情報を入手したことを発言しているラベルBが多い。避難情報に対する反応を示すラベルDにおいては、他の二つに比べて激しい増減をしていない。

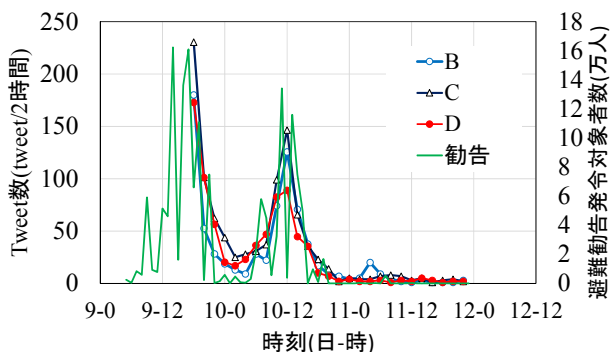


図-8 主ラベル(B-D)の時間変化

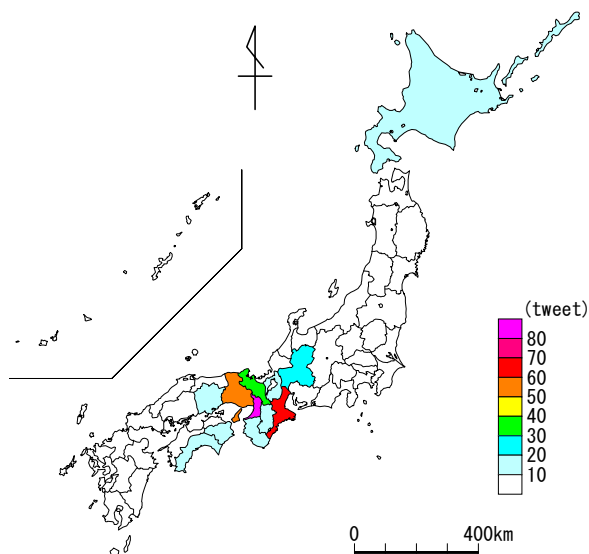


図-9 ラベルCの都道府県別のラベル分布

(4) ラベルCの都道府県別分布と実際の避難情報の比較

避難情報発令場所に関するラベルCに関して、呟き主体がどこで発令された避難情報を把握しているのか、都道府県別に集計した。図-9はそれらの分布を示したもの、図-10は避難情報発令者とラベルCのそれぞれの全数に対する割合を示した。避難情報発令対象者が最多であった三重県や大阪府ではラベルCの割合の全数の約4割程度

のTweet数を占めていたことから、Tweetは避難情報発令対象者が多い地域か、人口密集地で増加しやすいことが分かる。しかしながら上記の2府県や他都道府県を見ると、Tweetによる都道府県別の分布と避難情報発令対象者の分布は必ずしも完全には一致していない。このことから、Tweet利用者による地域間のバイアスの存在が考えられる。

5. 結論とおよび今後の課題

本研究では、台風による水害時において、Twitterデータを用いて避難情報に対する住民反応を観測するために、データの分類方法や抽出方法に特定の基準を設け、分析を行った。データを分類し「避難情報を入力したか」「避難情報の入手手段」「発令場所の把握」「避難行動」「避難情報に対する感想」の5つのラベルに関して避難情報に対する住民反応の特徴を得た。その結果、Twitterにより水害時の避難情報に対する住民の反応が相当数観測できることが分かった。これらを時系列ごとに集計し、行政から発令された避難情報発令の履歴や対象人数と比較すると、Twitterデータと避難情報発令の対象人数の分布は同様の傾向が見受けられた。これらの観測結果において、特徴的な知見が見受けられたラベルB, C, D) についての考察を述べる。

ラベルBについては、エリアメールからの入手の割合が最も多かった。この結果はTwitterというSNSの性質上、スマートフォンなどの携帯電話から投稿される場合が多く、エリアメールを携帯電話で受け取った直後にそのままTwitterに書き込みを投稿するというTwitter利用者の個人特性のバイアスを反映していると考えられる。

ラベルCに関しては、大阪府と三重県のように必ずしもラベル内でのTweet数の割合と避難情報発令対象者数内での割合の順位が一致しているとは限らない場合があった。また、ラベルCは、あくまで呟きの内容のみの分

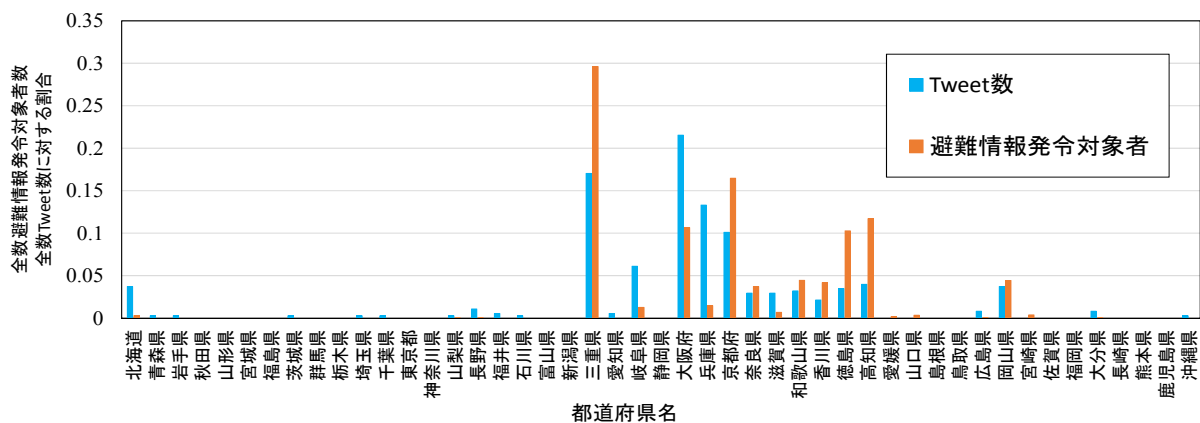


図-10 都道府県別の全数に対する割合

析であるために、地名が不明なTwitterデータの位置情報を分析できていない。また都道府県別のTwitter利用者数¹¹⁾においては大阪府が東京都に次いで第2位、兵庫県が第8位、京都府が第10位であることから、Twitter利用者の属性や地域間の分布によるバイアスがあったと考えられる。しかしTwitter利用者数では第25位である三重県がラベルCでは相当数分類されたことから、局所的に大規模な避難情報が発令されるとTwitterデータに反応が見られることが分かる。

避難行動に消極的である、という既存の知見（逃げ遅れの問題）と整合的な結果が得られた。ラベルDに関しては、副ラベルの割合から、「避難行動していない」が半数近く観測された。また、副ラベル3の「避難に疑問や不安を感じている」においては、避難所の立地に関する疑問や、夜間に避難情報が発令されたために避難行動を行うことの不安などを発言するTwitter利用者が多く観測された。、実際に避難を実行しているわけではないが、「避難行動」に対しては興味、関心を抱く住民が多かったと考えられる。

本研究ではTweetの内容分析は30秒に1度取得される20tweetの中で最初に取得されたTweetのみを対象としたが、より精度の高い住民反応を得るためには、更に正確な分析方法で期間中における対象Twitterデータの拡大を行う必要がある。そのためTweet取得による性能的な制約がないように取得可能なTweet数に余裕のあるインフラを整備する必要がある。また手作業で行ったTwitterデータの分類も自動化を施して人的な作業量を削減した上で更なる分析を実施するべきであろう。Twitter利用者の分布などによるバイアスについても定量的な検討を行った上で、住民反応の推定値等を算出することも必要であると考えられる。

本研究では、住民への避難情報伝播過程や住民が避難行動を実施するまでの思考過程等に関する知見を得ることで、行政などによる災害避難情報発信の課題分析を目的とした。分析の結果、比較的安価なデータであるTwitterデータから既存の知見と整合的な結果を得た。更に時系列的、地理的に分析を行うことで、アンケート調査では顕在化しないような避難情報に対する住民反応の傾向を見出せる可能性があることが分かった。これらは今後、災害時の人々の行動分析を行う上で更なる分析手法の発展を図る意義となりうるだろう。

最終的な目的達成のためには、今回の分析では考慮しなかった様々な要素を盛り込む必要がある。発令当時に住民が見舞われた状況として当時の天候データや、ライフラインの状況、交通状況を把握し、行政が集計した実際の避難者数などの他データを考慮して分析を行うことも有望だろう。各種要素を踏まえて避難情報が発令された時と発令されない時との比較を行い住民の避難行動に

対する意思決定の過程などを明らかにし、最終的に行政等による災害避難情報発信の際の課題分析を行うことが本研究において最も重要な点となるだろう。

参考文献

- 1) 田崎篤郎：自然災害の行動科学，福村出版株式会社，pp. 79-83, 1988.
- 2) 国土交通省気象庁：知識・解説，台風の発生，接近，上陸経路，<http://www.jma.go.jp/jma/kishou/known/typhoon/1-4.html> (2014/12/26閲覧).
- 3) NTTドコモ：緊急速報「エリアメール」，<https://www.nttdocomo.co.jp/service/safety/areamail/> (2015/1/7閲覧).
- 4) KDDI：災害時・緊急時対策「緊急速報メール」，<http://www.au.kddi.com/mobile/anti-disaster/kinkyu-sokuho/> (2015/1/7閲覧).
- 5) 天野和信，辻本晋吾，井料隆雅：既往研究及び事例調査に基づいた災害時避難行動の特徴分析，土木計画学研究発表会・講演集，Vol. 48, 2013.
- 6) Avvenuti, M., Cresci, S., La Polla, M. N., Marchetti, A., and Tesconi, M.: Earthquake emergency management by social sensing, *Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on IEEE*, 2014.
- 7) 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性分析，土木学会論文集D3（土木計画学），Vol. 70, No. 1, pp. 102-112, 2014.
- 8) Sakaki, T., Okazaki, M., and Matsuo, Y.: Earthquake shakes Twitter users: real-time event detection by social sensors, *Proceedings of the 19th international conference on World Wide Web, ACM*, 2010.
- 9) 気象庁：台風第12号，第11号と前線による大雨と暴風 平成26(2014)年7月30日～8月11日(速報)，http://www.data.jma.go.jp/obd/stats/data/bosai/report/2014/20140730/jyun_sokujij20140730-0811.pdf (2015/1/7閲覧).
- 10) 内閣府：台風第12号及び第11号による被害状況等について(第22報)，<http://www.bousai.go.jp/updates/h26typhoon12/> (2015/1/7閲覧).
- 11) 都道府県別統計とランキングで見る県民性[とどラン]：Twitterユーザー数，<http://todo-ran.com/t/kiji/13528> (2015/1/30閲覧)