

国土技術政策総合研究所資料

TECHNICAL NOTE of
National Institute for Land and Infrastructure Management

No.1013

February 2018

共同研究報告書

ソーシャルメディア分析によるリアルタイム
災害発生情報検知手法の確立に関する研究

Research on the establishment of a method for detecting disaster occurrence
information in real-time by social media analysis

国土交通省 国土技術政策総合研究所

株式会社富士通研究所

National Institute for Land and Infrastructure Management
Ministry of Land, Infrastructure, Transport and Tourism, Japan
Fujitsu Laboratories Ltd.

共同研究報告書
ソーシャルメディア分析によるリアルタイム
災害発生情報検知手法の確立に関する研究

國友 優^{*1} 神山嬢子^{*2} 野呂智之^{*3}
山影 譲^{*4} 武田邦敬^{*5}

概要

毎年のように土砂災害による犠牲者が出ている我が国において、人的被害の軽減を図るためには、警戒・避難体制の強化が不可欠である。そのため、避難勧告等の判断の目安の一つとされている土砂災害の前兆現象や発生情報を早期に把握し、警戒・避難に役立てることが重要である。

本共同研究は、近年普及がめざましいソーシャルメディア情報から災害が発生する現地でしか知り得ない土砂災害の前兆現象や発生状況を迅速に把握し、これらの情報を警戒・避難システムに活用する技術の確立に向けた研究を行った。本資料は共同研究の内容を取りまとめたものである。

キーワード：土砂災害、警戒避難、SNS、Twitter

- *1 国土交通省 国土技術政策総合研究所 土砂災害研究部 土砂災害研究室 室長
(現：水管理・国土保全局 砂防部 保全課 土砂災害対策室 室長)
- *2 国土交通省 国土技術政策総合研究所 土砂災害研究部 土砂災害研究室 主任研究官
- *3 国土交通省 国土技術政策総合研究所 土砂災害研究部 土砂災害研究室 室長
- *4 株式会社富士通研究所 知識情報処理研究所 主管研究員
- *5 株式会社富士通研究所 知識情報処理研究所
(現：富士通株式会社 AIサービス事業本部 AIインテグレーション事業部)

Joint research report
**Research on the establishment of a method for detecting
disaster occurrence information in real-time by social media analysis**

Masaru KUNITOMO*¹ Joko KAMIYAMA*² Tomoyuki NORO*³
Yuzuru YAMAKAGE*⁴ Kunihiro TAKEDA*⁵

Synopsis

Victims of sediment-related disasters occur every year in Japan, therefore warning/evacuation systems need to be improved in order to reduce the impact on humans. For that reason, it is important to grasp the precursory phenomena and occurrence information of sediment-related disasters as one of the criteria for making evacuation recommendations, etc. at an early stage, and to use it for warning/evacuation.

This joint research is exploring technologies for quickly grasping the precursory phenomena and occurrence situation of sediment-related disasters that can be known only in the local area where disasters occur from social media, which has spread remarkably in recent years, and to utilize this information for warning / evacuation systems. This paper summarizes the contents of this cooperative research.

Key Words : Sediment Disaster, early warning, SNS, Twitter

- *1 Head, Sabo Risk-Management Division, Sabo Department, NILIM, MLIT
Current Affiliation: Land Conservation Division, Sabo Department, Water and Disaster Management Bureau, MLIT
- *2 Senior Researcher, Sabo Risk-Management Division, Sabo Department, NILIM, MLIT
- *3 Head, Sabo Risk-Management Division, Sabo Department, NILIM, MLIT
- *4 Research manager, Knowledge Information Processing Laboratories, Fujitsu Laboratories Ltd.
- *5 Knowledge Information Processing Laboratories, Fujitsu Laboratories Ltd.
Current Affiliation: AI Services Business Unit, Fujitsu Ltd.

共同研究報告書

ソーシャルメディア分析によるリアルタイム災害発生情報検知手法の確立に関する研究

目 次

1	はじめに.....	1
1.1	共同研究の概要.....	1
1.2	研究の背景.....	1
1.3	研究の目的.....	6
2	Twitter 情報分析による土砂災害の前兆現象等情報の収集・把握手法の検証.....	7
2.1	Twitter 情報の収集・把握手法.....	7
2.1.1	Twitter データ.....	7
2.1.2	関連研究.....	8
2.1.3	本研究のアプローチ.....	9
2.2	災害事例におけるつぶやきの内容と豪雨・災害事象との関係.....	10
2.2.1	平成 24 年 7 月九州北部豪雨.....	10
2.2.2	平成 26 年 8 月広島市土砂災害.....	12
2.3	災害発生場の状況把握の可能性検討.....	13
2.3.1	分析手法.....	13
2.3.2	対象とする災害および Twitter 情報の概要.....	14
2.3.3	発言内容に基づくカテゴリー分けとキーワード設定.....	14
2.3.4	設定したキーワードによる状況把握.....	16
2.3.5	まとめ.....	17
2.4	警戒期から発災前における状況把握の可能性.....	18
2.4.1	警戒期における Twitter データの定量的調査.....	18
2.4.2	雨発言を利用した潜在的ユーザー分布の可視化.....	21
2.5	土砂災害の前兆現象等情報の収集・把握手法の可能性と課題.....	33
2.5.1	Twitter データに対する統計的アプローチにおける課題.....	33
3	社会実装に向けた適用条件等の整理.....	36
3.1	SNS 情報の警戒避難システムへの組み込みに対する社会的ニーズ.....	36
3.2	Twitter を活用した災害情報収集システムの運用モデル.....	36
3.3	DIGSUSS ユーザーインターフェースの検討.....	38
3.3.1	インターフェース仕様.....	38
3.3.2	収集キーワード.....	40
3.4	試作版システムの評価.....	41
3.4.1	有効な点.....	41
3.4.2	改善点.....	42
3.4.3	DIGSUSS 利用における留意点.....	43
3.4.4	評価結果の考察.....	43

4	まとめ.....	44
---	----------	----

謝辞

参考文献

巻末資料

1. DIGSUSS 仕様書 (案)

1 はじめに

1.1 共同研究の概要

土砂災害に対し適切な避難行動をとることにおいて、「山鳴り」や「流木の流出」等の前兆現象の発生の把握が重要である。災害が発生する現地でしか知り得ない情報を、広い国土の中で網羅的かつ迅速に把握する手段の一つとして、住民等が発するソーシャルメディア上の投稿情報を活用することが考えられた。そこで、この「つぶやき情報」を用いて、土砂災害の「予兆」や「発生」を検知する手法の実用化に向けた研究を行うこととした。研究項目と概要を以下に示す。

1. ソーシャルメディア分析による土砂災害の発生検知手法の検証

ソーシャルメディア分析による土砂災害の予兆・発生検知の適用可能性を検討するため、テキストマイニング技術を用いて過去の災害時におけるソーシャルメディア上の投稿情報を利用し、危険が迫る地域を絞り込んだ上で、予兆や災害発生検知を行うプログラムを作成し、災害検知の実施、検知結果の評価等を行う。

2. ソーシャルメディア分析による土砂災害の発生検知手法の適用可否条件等の整理

ソーシャルメディア分析による土砂災害の予兆・発生検知の実用化に向けて、検知手法の適用可否条件の整理、情報提供時の表現方法等の検討を行う。

本研究は、国土技術政策総合研究所と（株）富士通研究所が、平成 26 年 7 月～平成 28 年 3 月に共同で実施したものである。

1.2 研究の背景

毎年のように土砂災害によって甚大な被害が発生している。土砂災害は毎年 1,000 件程度の報告があるが、これらによる死者・行方不明者の数は、大地震・大津波を除くと自然災害による犠牲者の 41%を占めると言われている¹⁾。このような状況において、人的被害の軽減を図るためには、警戒・避難対策の強化が不可欠である。しかしながら、市町村による避難勧告等の発令や住民自らの避難を判断するための状況把握は難しく、避難が遅れるといった課題が見られる。

土砂災害が急迫した際に、人々にその危険性を伝えるために提供される情報として、土砂災害警戒情報がある。平成 27 年 1 月に施行された土砂災害防止法の一部改正により、都道府県知事は、避難勧告等の判断に資するため、土砂災害警戒情報を関係市町村や一般に提供することが義務づけられた。また「土砂災害警戒避難ガイドライン」において、市町村は、土砂災害警戒情報が発表された場合、直ちに避難勧告等を発令することが基本とされており²⁾、「避難勧告等の判断・伝達マニュアル作成ガイドライン」においても、内閣府は自治体に対し、土砂災害警戒情報の発表を避難勧告の判断基準とするよう求めている³⁾。土砂災害警戒情報は、過去の土砂災害実績と降雨量の推移から、対象としている地域全体の土砂災害の危険度を評価し発表されるものであり、個別斜面の地形・地質及び植生等の影響までは考慮されていない。したがって、より具体的にどの溪流、どの斜面の危険性が高まっているかを知るためには、それぞれの箇所における変状を知る必要がある。

そのため、土砂災害警戒情報が発表されたのちに、個別の箇所における更なる切迫性の高まりを知るには、「地鳴り」「土臭い」など土砂災害の前兆現象が重要となる。

住民が土砂災害の前兆現象を見つけ、家族や近隣住民とともに避難し、人的被害を回避した事例は度々報告されている（図 1-1）。表 1-1 に示すような「地鳴り」「土臭い」などの土砂災害の前兆現象は、内閣府の「避難勧告等の判断・伝達マニュアル作成ガイドライン」において、自治体が発令する避難指示や、住民自らが直ちに身を守る行動をとるための重要な判断指標であるとされている。

■前兆現象が避難勧告発令や避難につながった事例

平成24年九州北部豪雨災害
（福岡県うきは市）

平成24年7月13日
23:59 土砂災害警戒情報発表
深夜 自主避難

- ・ **地域住民からの情報**
- ・ **浸水情報**
- ・ **人家裏の崖から小石が落ちてくる**

6:45 避難勧告発令
9:00頃 土石流発生



土石流による土砂の氾濫と家屋等の被害

※国土交通省砂防部資料より

図 1-1 事前に避難した事例（福岡県うきは市）

表 1-1 土砂災害の前兆現象

	土石流	がけ崩れ	地すべり
			
視覚	<ul style="list-style-type: none"> ・川の水が濁る ・降雨継続中に川の水位が下がる ・落石 ・濁水に流木が混じる ・溪流内の火花 	<ul style="list-style-type: none"> ・がけに割れ目・緩み ・小石が落ちる ・表面流が生じる ・斜面上での湧水 ・湧水の濁り ・樹木の傾斜 	<ul style="list-style-type: none"> ・地面ひび割れ、陥没・隆起 ・沢、井戸の水の濁り ・斜面上の湧水 ・池沼水位の急減 ・樹木の傾斜 ・家屋、擁壁の亀裂や傾斜
聴覚	<ul style="list-style-type: none"> ・地鳴り、山鳴り、 ・転石同士の衝突音 	<ul style="list-style-type: none"> ・樹木の揺れる音 ・樹木の根が切れる音 ・地鳴り 	<ul style="list-style-type: none"> ・樹木の根が切れる音
嗅覚	<ul style="list-style-type: none"> ・腐った土の匂い 	-	-

土砂災害の前兆現象をとらえ、それを警戒・避難に活かそうとする取り組みは、比較的早くから行われてきた。1999年度より導入された土砂災害110番制度は、まさに住民によって感知された前兆現象を他の地域の避難に活かそうという取り組みに当たる。しかしながら、この前兆現象を住民が感知したとしても、行政にまで通報されるケースは多くはなく、土砂災害110番制度は必ずしも効果を発揮しているとは言えない。理由は多々あるであろうが、地域毎に見ればかなりまれな出来事である土砂災害の前兆現象を、いざ感知した時に土砂災害110番に連絡しなければならないという考えが浮かばない、また仮に連絡しようとしても、直ちに連絡先を思い出すことができない、すなわち土砂災害110番は普段使いしない連絡先であるということが、主な理由ではないかと考える。そのため、前兆現象等を把握しても情報の伝達範囲が家族や近隣住民に限られ、情報を迅速に収集し、他の地域に伝達することに課題があったと考える。

また、加藤ら(2008)は、前兆現象を防災システムに取り込もうと、消防へ寄せられる通報を用いて土砂災害の危険度レベルを評価するシステムの研究を行っている。その結果、その有効性は認めつつも、寄せられた通報をリアルタイムに分析するシステムの構築に課題が残るとしている⁴⁾。この手法もこれら課題が解決されない限り、前兆現象の防災への活用の決め手とはならない。

このように、土砂災害に対する警戒・避難体制の強化には、土砂災害の前兆現象の発生をとらえ、避難に活かすことは有効であるが、いかに迅速にかつ網羅的に収集できるかが、大きな課題である。

一方で、近年のインターネット接続が可能な携帯端末や Social networking service (以下「SNS」という。)の普及により、文字や写真による情報発信や共有、収集が容易になった。主なSNSの日本におけるユーザー数の推計値を見ると、Twitter、Facebookともに平成24年に1,400万人程度まで増加している⁵⁾(図1-2)。Twitter社が公表した平成27年12月の日本国内のTwitter月間アクティブユーザー数は、3,500万人⁶⁾と言われており、日本におけるユーザー数の増加が著しい。

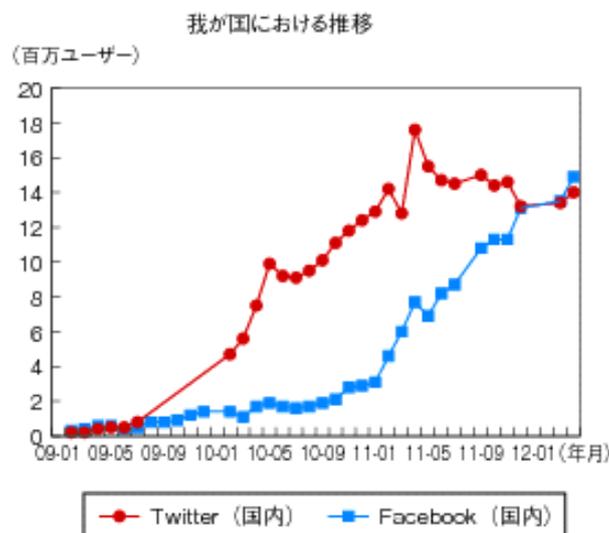


図1-2 ソーシャルメディア利用者数の推移 (Facebook, Twitter の例)⁵⁾

SNSの中でも、Twitterは、文字数が140字に制限される代わりに通信のリアルタイム性が高く、ほかのユーザーとの情報の交換や情報の転送が容易であり、一般的なSNSよりもユーザー同士が気軽・容易につながることによって情報が広範囲に急速に拡散される⁷⁾。

Twitterは、東日本大震災の際、携帯電話の通話ができない状態が続いたなかで、安否情報等、緊急性の高い情報を交換することができたほぼ唯一の手段であったと言われており、防災ツールとして大きな注目を集めた。Twitterのリスクコミュニケーションツールとしての利用には、賛否両論あるようである。しかしながら谷口(2012)は、デマの拡散といった負の効果を考慮したとしても、全体像が把握しづらい災害初期の段階において、Twitterは情報共有ツールとして有用であると指摘している⁸⁾。

また、Twitterに関しては、リスクコミュニケーションツールとしての使用法以外にも様々な研究が進められており、震源地の推定やインフルエンザの流行予測のように、そのリアルタイム性を生かした自然現象の動きそのものをとらえる研究や様々な社会現象を把握する研究もなされている。榊(2012)は、Twitterユーザーが感知したことを文字情報等としてソーシャルメディア上に投稿(出力)する行為は、物理センサ(地震計や雨量計など)の構造と良く似ており、Twitterユーザー自体をソーシャルセンサと位置づけ、これを活用することによって、これまで物理センサでは観測できなかった多様な現象を観測できるようになると指摘し、今後何らかの工夫を施すことによって、現時点では物理センサには及ばない信頼性・安定性をも将来には獲得できる可能性がある⁹⁾。また阿部(2014)は、Twitterのようなソーシャルメディア情報と、携帯電話の位置情報、カーナビ、車両センサーデータ等のいわゆるビッグデータを重ねるあわせることで、災害の全貌が見えてくるとも指摘している¹⁰⁾。

防災分野に係るソーシャルセンサと物理センサを試行的に分類した例を図1-3に示す。

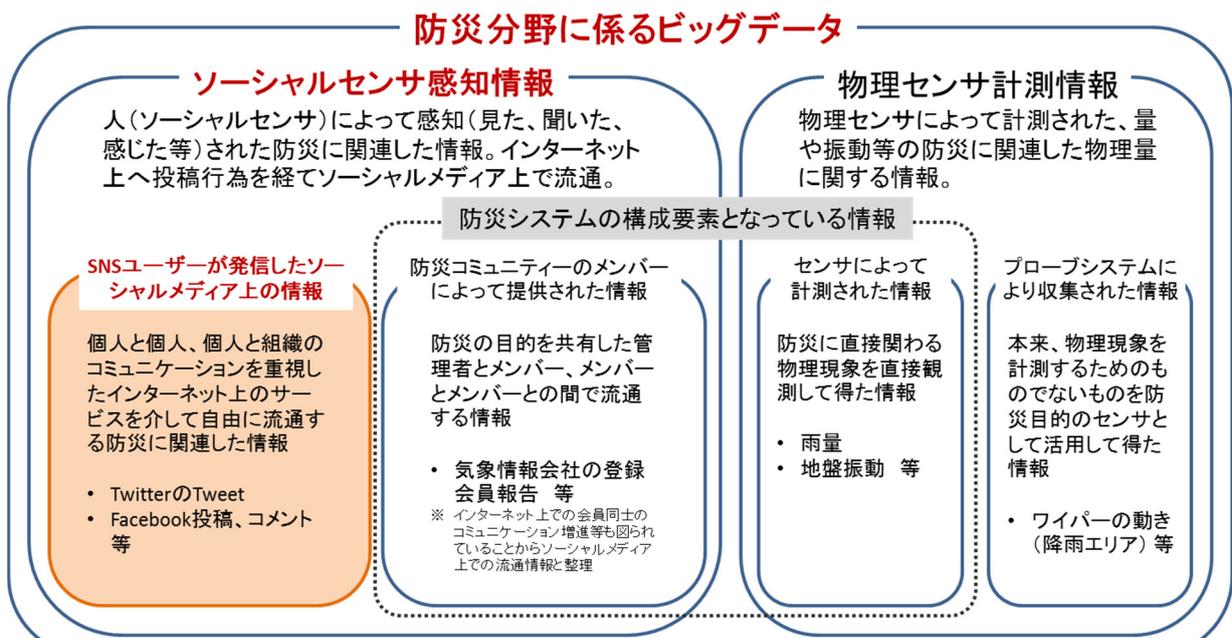


図1-3 防災分野に係るソーシャルセンサと物理センサの分類(國友, 2015¹¹⁾を一部改変)

このような Twitter の防災利用に関する研究動向をまとめると、Twitter は発災直後の他の通信手段が機能していない段階でも情報共有ツールとして機能し、信頼性・安定性は物理センサには及ばないものの、物理センサには感知することが難しい自然現象もとらえることが可能なものもある。また、物理センサ情報と相互補完することにより、今まで分からなかった事柄を可視化するのに有用なツールであることが分かる。このことは、「地鳴り」「土臭い」等、これまで物理センサではとらえることが難しかった事象が、すでに Twitter ユーザーの手によって直接ソーシャルメディア上に意図せず出力されている可能性があることを示唆している。

これは、長年の課題であった前兆現象の迅速かつ網羅的な把握は、日々つぶやかれる膨大な量の Twitter 情報をうまく活用することによって可能となり、防災情報として活用できるようリアルタイムで可視化することが現実のものとなる可能性があることを意味している。

また、IT 総合戦略本部により開催された「防災・減災における SNS 等の民間情報の活用に関する検討会」は、Twitter 等の SNS は広域においても小地域においても情報収集やコミュニケーションが可能であり、適時性が高い情報や、履歴情報も収集可能であること、普段使いのツールであり国の行政機関や地方公共団体に設備投資のために著しい予算負担がないことをメリットとしてあげ、今後信頼性を補完するための取組を交えながら活用することによって、行政機関が防災上の課題を克服するツールとなりうることを指摘している¹²⁾。

1.3 研究の目的

本研究では、ソーシャルメディア上の投稿情報として Twitter 社の twitter 情報等を利用し、土砂災害発生情報を検知する際の適用可能性を検討するため、ソーシャルメディア分析技術の精度検証と実用化に向けた課題整理を行う。

2 Twitter 情報分析による土砂災害の前兆現象等情報の収集・把握手法の検証

2.1 Twitter 情報の収集・把握手法

2.1.1 Twitter データ

(1) 基本事項

Twitter は、米国の Twitter 社が提供する、Web 上のマイクロブログサービスのひとつである。Twitter の利用者は、ツイートと称する 140 文字の短文を Web 上に投稿することができ、投稿されたツイートは、投稿者が制限を掛けない限りにおいては、一般に公開される。全世界で、3 億 2000 万人の月間アクティブユーザー¹を抱える巨大な Web サービスであり、国内でも広く利用されている。また、Twitter 上の情報は、Web ブラウザやスマートフォンアプリを利用して閲覧することができる他、Twitter 社が提供する Web 上の Application Programming Interface (以下、「API」という)を用いて取得することも可能である。したがって、ソーシャルメディア情報として、2 次的な活用を考えることができるものである。一方、匿名で利用できるサービスであることから、必ずしも正しい情報が流通しているとは限らないことに注意する必要がある。

Twitter の情報としての特徴は、以下の 2 点である。

● 投稿される情報にリアルタイム性がある

ツイートは 140 文字以内に制限されており、利用者は、長文を投稿することができない。この制約から、利用者は通常のプロログなどと比較して、自分自身の意見や観察した情報を即座に投稿する傾向がある。

● 情報の拡散が早い

リツイートと呼ばれる、他人の発言を複製して簡単に再投稿する機能を有している。利用者は、リツイート機能を用いて、他の利用者が投稿したツイートに対して、関心を持ったり、何かしらの意見を持ったりした場合に、当該ツイートを簡単に投稿することができる。このため、Twitter は、他のソーシャルメディアと比較して、インパクトのある情報が拡散しやすい傾向にある。一方、ツイートデータに対して統計的な処理を行う場合には、リツイートによって同じ情報が複製されて水増しされることを考慮する必要がある。

(2) データの収集方法

Twitter 上に投稿されたツイートを、データとして収集するための手段としては、大きく分けて、(a)無償公開されている Web 上の Twitter API を利用する方法と、(b)Twitter 社のデータを販売する業者を通じて有償で購入する方法がある。

(a)のメリットは、無償でデータを入手できることと、Twitter API へのアクセス処理を工夫することで、リアルタイムに近い形でデータを収集できることである。一方、デメリットとしては、Twitter 上に存在するツイートデータを、もれなく入手できるわけではないこと

¹ Twitter 社の日本国内 Web ページより引用。 <https://about.twitter.com/ja/company>

である。Twitter API は、用途に応じて複数用意されており²、代表的なものとして、検索キーワードを指定して当該キーワードを含むツイートを収集できる「Search API」がある。この API は、指定したキーワードを含むツイートを、できる限り提供するものである。無償であることから、本 API を活用したスマートフォンアプリや、Web サービスが数多く存在している。しかしながら、Twitter 社が定める取得上の制限³があり、必ずしもすべてのデータを取得できるわけではない。また、キーワードを指定する方式であるため、取得したいツイートに含まれると想定されるキーワードを、あらかじめ調査して確定させておく必要がある⁴。

一方、(b)による方法としては、大きく分けて、i)日付やキーワードを指定して条件に該当するデータを一括で購入する方法と、ii)随時リアルタイムに全量データを入手できる契約を結ぶ方法がある。i)は、日付やキーワードを指定することで、当該条件を満足する全ツイートを入手できることが特徴だが、リアルタイム性は期待できない。このため、ケーススタディのように、過去の何らかの事例に対してデータを俯瞰的に調査する用途に適している。本研究においても、一部の調査を行うために、過去データの購入を実施した。一方、ii)の方法は、非常に高コストであり、主に大規模な Web サービスを展開する企業が話題分析やマーケティングに利用されていることが考えられる。このため、研究用途としては利用されていないものと想定される。

2.1.2 関連研究

Twitter に代表されるマイクロブログに対するデータマイニング技術に関しては、奥村 (2012)¹³⁾にて概説されている。この中で、ソーシャルメディアの情報から特定の話題を検出する技術として、Kleinberg (2002) によるバースト検知¹⁴⁾が紹介されている。Kleinberg は、Web 上で通常ランダムに出現する特定のキーワードが、ある時に急増する現象を「バースト」と呼んだ。このバーストを捉えることにより、特定のキーワードに関連する話題やイベントの発生を捉えることができるという仮説がある。

高橋 (2011) は、Twitter 上に投稿されるツイートは、実世界で起きている事象を捉えるためのセンサ (ソーシャルセンサ) であると仮定し、Twitter データを用いて花粉症の広がりを捉える手法を提案している¹⁵⁾。

これらの研究成果を踏まえ、著者らは、防災分野への適用を検討するため、Twitter を用いて災害の発生をいち早く捉えるための技術を開発した¹⁶⁾¹⁷⁾。この研究により、浸水災害や土砂災害といった事象においては、Twitter からリアルタイムに発災情報を捉えることができる可能性があることがわかった。

² Twitter 社の開発者向け Web ページを参照。https://dev.twitter.com/rest/public

³ Twitter の Search API に対して通常のユーザー認証方式にてアクセスした場合、15分に180回までアクセス可能であり、1回のアクセスで標準15件、最大100件まで取得可能となっている。これらの仕様については、今後予告なく変更される可能性がある。

⁴ 富士通研究所においては、研究用途での利用を前提として、2012年7月から災害や防災に関連するキーワードを含むツイートを収集し、データ調査や研究開発に活用してきた。

2.1.3 本研究のアプローチ

本研究では、土砂災害が発生する前の警戒に資する情報を Twitter データから抽出することを目的として、データの調査と手法の検討を行った。本研究のアプローチを以下に示す。

1. 特定の災害事例に対する詳細なデータ調査（初期調査）による災害発生状況把握の可能性検討

特定の災害事例を対象として、気象条件や災害の発生事実と Twitter データを突き合わせることで、災害の発生前後でどのようなツイートが投稿されているか調査し、情報の特性と利活用の可能性を考察した。（2.2, 2.3 節）

2. 俯瞰的なデータ調査による警戒期から発災前における状況把握の可能性検討

土砂災害の発災前後の Twitter データの全体像を掴み、Twitter データから抽出できる可能性がある情報を整理するため、統計解析手法を用いた俯瞰的なデータ調査を実施した。（2.4, 2.5 節）

3. 前兆現象等情報の収集・把握のための統計手法の検討

以上の調査結果を踏まえ、土砂災害が発生する前の警戒的な情報を Twitter データから自動的に取得するための手法と課題を考察した。（2.6 節）

2.2 災害事例におけるつぶやきの内容と豪雨・災害事象との関係

2.2.1 平成 24 年 7 月九州北部豪雨

平成 24 年 7 月九州北部豪雨は、阿蘇地域（阿蘇市・南阿蘇村・高森町）において、土砂災害が 85 件発生し、死者・行方不明者 25 名をもたらした甚大な被害が発生した災害である¹⁸⁾。

本研究では、武田らが構築した分析プロセス¹⁷⁾を用い、さらに詳細に災害時の分析を進めるため、まず熊本県内全域に大雨注意報が発令された直前の 2012 年 7 月 11 日 16 時 0 分から翌 12 日 12 時まで日本全国で投稿された Twitter データ約 2,850 万件を取得した。これらを分析した結果、熊本県内で投稿されたと推定されるツイートは約 18 万件あり、その内市町村推定までできたものは約 2.4 万件であった。

県内でツイートされたと推定されたもののうち、「大雨」「豪雨」もしくは「土砂降り」を含むツイート数は約 4,100 件、「土砂崩れ」もしくは「土石流」を含むツイート数は約 600 件あった。図-2.1 は、降雨強度 (mm/h) の空間分布の変化に伴う土砂災害発生箇所（熊本県から国土交通省に報告のあったもので、人家等に影響のなかったものは除く）の変化と各時間帯においてつぶやかれた主なツイートを示したものである。

当該災害の際、阿蘇市・南阿蘇村に土砂災害警戒情報が発表されたのは 7 月 12 日 2 時 40 分であった。この時の Twitter データを見てみると、その直前の 2 時 32 分に阿蘇市に隣接する菊池市において既に小規模な土砂崩れ（添付されていた写真により確認）が発生していたことを示すツイートがあることが分かる（図 2-1 (a)）。また、3 時 59 分には玉名市・熊本市（田原坂）で土砂崩れが発生したことをうかがわせるツイートがある（図 2-1 (b)）。阿蘇地域において、土石流等が集中的に発生したのは 4 時台後半から 5 時台にかけてであった（図 2-1 (c) 及び (d)）。このとき、仮に Twitter データがリアルタイム分析できていれば、近隣地域での前兆現象を、発災の約 30 分から 2 時間前に把握することができ、より早い時間帯に避難指示等を発令することができた可能性がある。

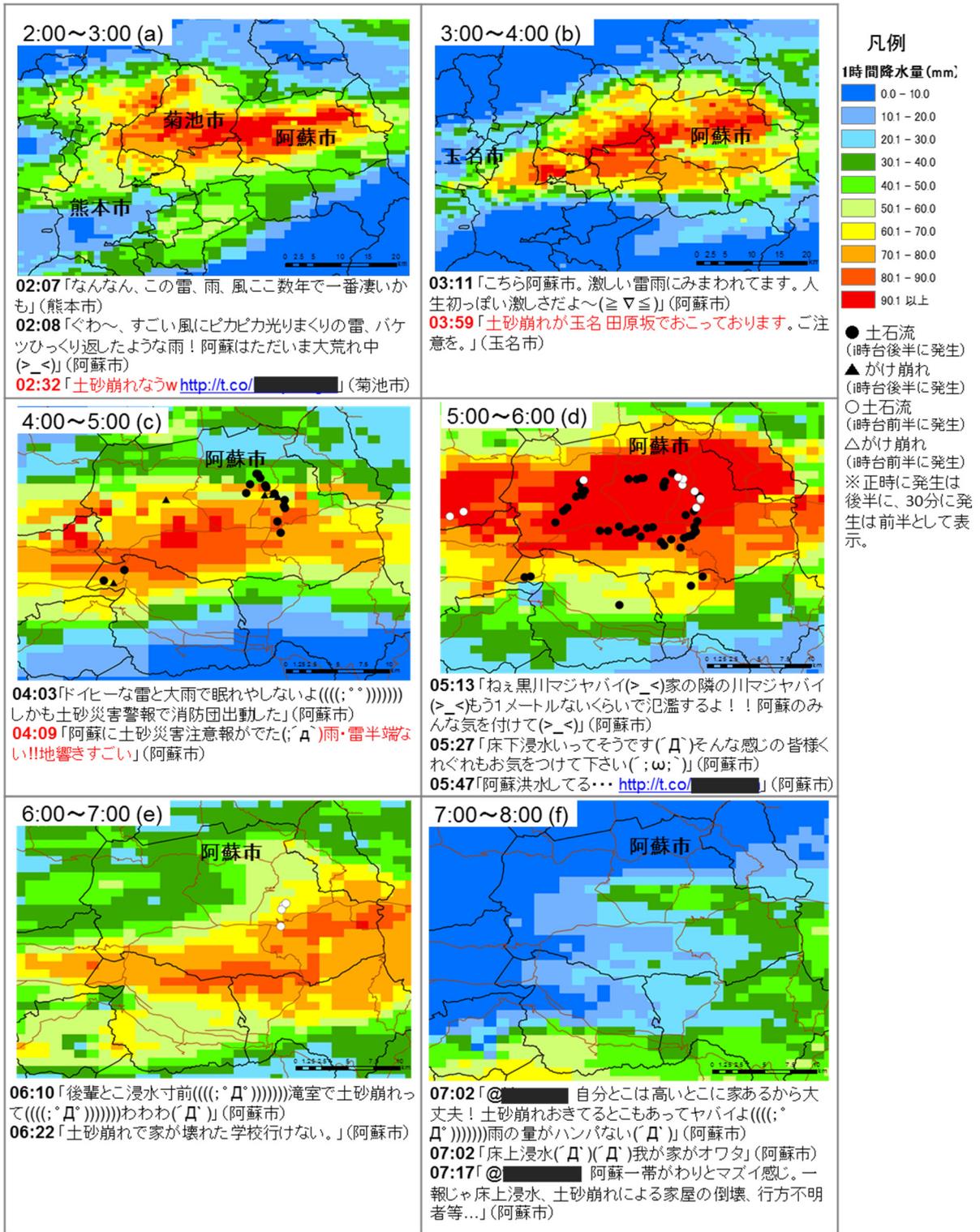


図 2-1 降雨強度 (mm/h) の空間分布の変化に伴う土砂災害発生箇所の変化と各時間帯においてつぶやかれた主なツイート

ただし、酒井ら (2013) は、土砂災害警戒情報が発表された段階で、既に猛烈な豪雨が降っており、仮に早期に避難指示が発令されていたとしても、現実的には避難行動をとること

は困難な状況にあったと考察している¹⁹⁾。このようなケースの場合、前兆現象に関する情報を得られたとしても、住民に対し実際どのような行動を促すかは課題として残る。

土砂災害が集中的に発生し始める直前の4時9分には、土砂災害の前兆現象ともとれる「地響き」を含むツイートがある。このツイートは共起語として「雷」を含んでおり、雷による震動を感じ「地響き」として投稿している可能性がある。しかしながら、酒井ら(2013)の調査によると、4時半頃に「地鳴り」を聞いたとの証言を複数得ており¹⁹⁾、直ちに雷によるものであると決めつけるのは早計のようにも思われる。他の土砂災害の事例を見ても、「地響き」等は比較的良好に投稿されており、「地響き」等の前兆現象に関連するキーワードとしての取り扱いは今後の課題となる。

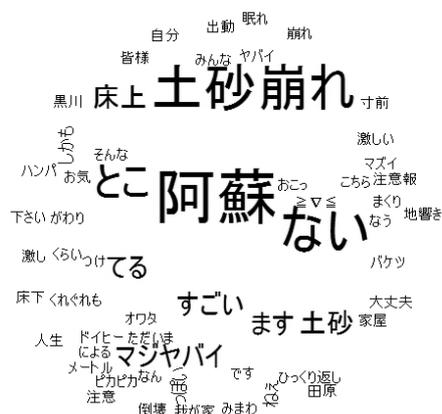


図 2-2 平成 24 年 7 月九州北部豪雨災害（阿蘇地域土砂災害）の際の
 主なツイートをワードクラウド化したもの
 ※形態素解析を行い頻度が高い語を大きなフォントで表現

2.2.2 平成 26 年 8 月広島市土砂災害

平成 26 年 8 月 19 日から 20 日未明にかけて広島県広島市で発生した豪雨により、広島市の安佐北区、安佐南区を中心に土砂災害が 167 件発生し²⁰⁾、死者 74 名をもたらす甚大な被害が発生した²¹⁾。

8 月 19 日～20 日の気象状況、および 20 日未明の行政対応状況と、筆者らがツイート本文やプロフィール等を見て広島県に関するツイートと判断したものを収集し、時系列に整理した(図 2-3)。災害発生当日、土砂災害警戒情報が発表された 1 時 15 分より以前から、雷を伴う豪雨による「地鳴り」や「冠水」といったツイートが見られている。それ以降、「地響き」や「川を石がごろごろ転がるときみたいな音」などの、土砂災害の前兆現象とも考えられる内容のツイートが投稿されている。そして、災害が集中して発生しはじめたと考えられる 3 時頃を過ぎると、「土砂崩れ」や「土臭い」というツイートが次々に投稿され、未明の外が薄暗い状況の中でも人家周辺に流れ込んだ土砂やその臭いで異常を感じられていたことが推測された。

図 2-3 に記載したツイートのうち、簡易的な方法により投稿場所を推定できるツイートは僅かである。そのため、災害が発生しているおそれがある地域における有用な情報を収集するためには、Twitter ユーザーのプロフィールや過去の発言も利用する方法²²⁾や、方言など言葉の地域性を考慮した推定を行うことが考えられる。

気象状況 ^{1) 2)}	行政対応等周辺状況 ²⁾	ツイート状況
<p>01:15 土砂災害警戒情報 03:00 安佐北区:累積雨量が警戒基準雨量超 03:49 記録の短時間大雨情報 04:00 1時間雨量101mm 05:00 雨がほぼやむ</p>	<p>01:32 市)防災情報メールで土砂災害の注意喚起 01:35 市)災害警戒本部を設置</p> <p>02:50 市)防災行政無線で大雨の注意喚起</p> <p>03:00 安佐北区などで累積雨量が「警戒基準雨量」を超える</p> <p>03:21 安佐南区山本の住民から人命救助を求める最初の119番通報 03:30 市)災害対策本部を設置</p> <p>04:15 市)安佐北区の一部に避難勧告 04:30 市)安佐南区の一部に避難勧告 安佐北区可部地区で3歳男児を救助活動中の消防士が土石流に巻き込まれる</p> <p>05:25 安佐北区の避難勧告が拡大 06:30 知事が陸上自衛隊に災害派遣要請 07:00 安佐北区上原で19日午前9時からの累積雨量が287mmに 08:00 安佐南区八木地区で避難所開設</p>	<p>01:08 「ガチで音やばそうw地鳴りすごい」 01:19 「これだけ広範囲に冠水したのは始めて。(後略)」</p> <p>02:05 「地鳴りすごい！ゆれる！」 02:26 「こっちは雨は降ってないよー。山2つ向こうくらいでドドドドッという雷の地響きみたいな怖い音は聞こえる(T_T)」 02:50 「川を石がごろごろ転がるときみたいな音がずっとしてる」 02:31 「家の目の前の川が初めて見るペースで増水中。」 02:51 「雨の音+雷の音+地響き+救急車の音とかやめて」 03:01 「は、わや。家揺れたと思ったら、家の前土砂崩れ。ほんまにわや。」 03:09 「結構やばい、横の山が落ちてきとる」 03:20 「八木の方土砂崩れおきとるしやばいな」</p> <p>03:32 「やば土砂崩れおきたヤバいこわい」 03:33 「絶対土砂崩れだよ119つながらん」 03:35 「自宅手前で土石流を見かけるという悪夢。」 03:36 「土砂崩れで家の周りの家がない、、、」 03:47 「風が土臭い」 03:51 「やばい震える怖い死ぬまじで土砂崩れが家まで入ってきてとる」</p> <p>以降、土砂崩れや土の臭い等のツイート多数</p>

【引用文献】 1) 広島地方気象台 気象速報(平成26年8月20日14時現在) 2) 毎日新聞 2014年9月1日より

図 2-3 平成 26 年広島災害時の対応等時系列とツイート状況
(筆者らが広島に関するツイートと判断したものを掲載)

2.3 災害発生場の状況把握の可能性検討

前節の事例から、災害時に土砂災害の前兆や発災に関する情報をツイートから得られる可能性があることが分かった。

そこで、九州北部豪雨災害(2012年7月)及び広島豪雨災害(2014年8月)の際のツイートを分析し、Twitter 情報を活用した土砂災害発生場の状況把握の可能性について検討を行った。

2.3.1 分析手法

Web 上に流れる膨大な Twitter 情報からリツイートを除き、ツイートの内容がどこの市町村での出来事を指しているのか絞り込みを行った上で(以下「市町村推定」という。)、土砂災害に関連するものを抽出し、その発言内容から状況を把握するため、土砂災害に関連する発言内容をカテゴリー分けし、さらに、各カテゴリーを代表するキーワードの設定を試みた。

設定したキーワードが含まれるツイートを抽出し、各カテゴリーのツイート数の変化に基づく状況の推察及び、状況を直接的に示す単語による、さらに的確な現象の把握の可能性を考察した。なお、市町村推定は武田ら(2014)の手法¹⁷⁾を一部簡略化して行った。推定の流れは次の通りである。

- ① キーワードを設定する。
- ② 公開されている GPS 情報やユーザープロフィール、ツイート本文中の地名・ランドマー

ク等からユーザーの居住地を推定し、Twitter データを都道府県別に仕分ける。

- ③ 同種の「つぶやき」の急増をとらえて都道府県単位での異常事態の発生を推定する。
- ④ 異常事態の発生を推定結果を踏まえ、対象とする都道府県内で投稿されたと推定される Twitter データに絞り込んだ上で②の分析を再度行い、市町村を推定する。

2.3.2 対象とする災害および Twitter 情報の概要

対象災害は、土砂災害発生場の状況を表す諸現象が顕著であり、土砂災害の前兆現象についても住民により把握されていたことが分かっている（例えば、酒井ら、2013）¹⁹⁾、2012 年 7 月九州北部豪雨（以下「九州北部豪雨」という。）及び 2014 年 8 月豪雨（以下「広島豪雨」という。）によって、熊本県阿蘇市・南阿蘇村、広島県広島市で発生した土砂災害とした。分析対象期間は土砂災害発生前後の時間帯とし、九州北部豪雨では 2012 年 7 月 11 日 16 時～12 日 12 時、広島豪雨では 2014 年 8 月 20 日 1 時～5 時 30 分とした。双方とも、深夜に降雨が強まって明け方にかけて土砂災害が発生し、多数の人的被害等をもたらした。分析対象ツイートは、分析対象期間に投稿されたツイートのうち、発言者の位置推定ができたもののうち、九州北部豪雨では阿蘇市および南阿蘇村（以下「阿蘇」という。）と推定された 2207 件を、広島豪雨では広島市（以下「広島」という。）と推定された 5813 件とした。

2.3.3 発言内容に基づくカテゴリ分けとキーワード設定

本検討に先立って、「避難勧告等の判断・伝達マニュアル作成ガイドライン」（内閣府、2014）³⁾等を参考に土砂災害の現象を示す専門的な用語（表面流発生、小石がばらばら落下等）をキーワードとして使用することが可能かどうかを確認したが、ガイドラインに掲載されているような表現のままでツイート中に出現する事例は極めて少なかったことから、実際のツイートの記載内容を読み解いて、改めてキーワードとして設定する方針とした。

設定の流れを以降に示す。

- ① ツイートの内容に基づき、土砂災害の前兆現象や災害発生に関連するツイートを抽出
- ② 抽出したツイートを内容からカテゴリに分類
- ③ 各カテゴリを構成する語を抽出
- ④ 土砂災害に関連する語をキーワードとして設定

ツイートの内容に基づき、広島で 1,617 件、阿蘇で 968 件を土砂災害に関連するツイートとして抽出した。さらに、抽出したツイートを表 2-1 の「状況カテゴリ」に分類した。

表 2-1 状況カテゴリー

項目	内容
気象関係（雨、雷）	雨、雷等天気の状態
警報、避難、消防活動関係	警報発表や避難勧告等の発令、消防活動等の状況
河川関係（増水、流量）	流量の増加等、河川に関する状況
水害関係（氾濫、浸水）	氾濫、浸水等の水害に関する状況
土砂災害関係	土石流、がけ崩れなど土砂災害に関する状況
休校、休業関係	休校、休業などの対応に関する状況
交通関係（通行止め）	通行止め、交通機関の遅延など交通に関する状況
不安感等	災害発生の危険等に対する懸念や感想
その他	上記に分類されないが災害に関する状況

テキスト型データを統計的に分析するソフトウェアである KH Coder を利用し、状況カテゴリーごとに、ツイートに登場する語を品詞分解した上で自動抽出を行った。抽出した語の中から、各状況カテゴリーを表し得る語（例：土砂災害では、「土砂崩れ」「災害」など、不安では、「やばい」「怖い」など）をキーワードとして設定した。土砂災害カテゴリーの例を表 2-2 に示す。ここで、品詞の区分は KH Coder の分類ルールに基づくものであり、「名詞」は漢字を含む 2 文字以上の語を、名詞 C は漢字 1 文字の語を示す。

表 2-2 設定したキーワード例（土砂災害カテゴリー）

九州北部豪雨

広島豪雨

抽出語	品詞	出現回数
土砂崩れ	名詞	34
生き埋め	名詞	20
土砂	名詞	8
災害	名詞	5
土石流	名詞	3
崩れ	名詞	3
崩れる	動詞	3
埋まる	動詞	3
がけ崩れ	名詞	1
なぎ倒す	動詞	1
安否	名詞	1
溢れる	動詞	1
壊れる	動詞	1
地響き	名詞	1
倒木	名詞	1
被害	名詞	1

抽出語	品詞	出現回数
土砂崩れ	名詞	107
土砂	名詞	29
災害	名詞	26
崩れる	動詞	11
裏山	名詞	9
生き埋め	名詞	4
流れ込む	動詞	4
崖	名詞C	3
山	名詞C	3
がけ崩れ	名詞	2
流れる	動詞	2
土砂降り	名詞	1
土石流	名詞	1
被害	名詞	1

表 2-1 の各カテゴリーにツイート内容を読み解いて分類したツイート数（図 2-4「内容分類による抽出」）及び、各状況カテゴリーのキーワードを含むツイート数（図 2-4「キーワードによる抽出」）の時系列変化を比較すると、両者の傾向は概ね一致し、設定したキーワードで各状況カテゴリーのツイート数の変化を捕捉することが確認できた（図 2-4）。

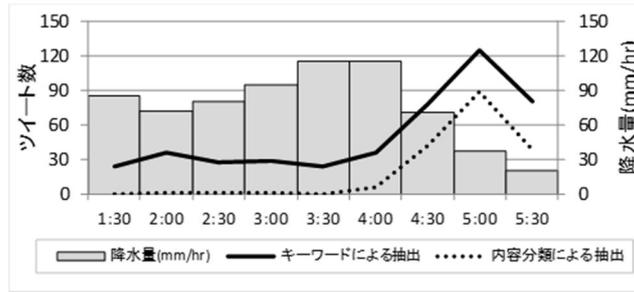


図 2-4 土砂災害に関連するツイート数の時系列変化（広島）

2.3.4 設定したキーワードによる状況把握

阿蘇市における災害の状況と、設定したキーワードで抽出したカテゴリーごとのツイート数の変化及び状況を直接的に示す内容のツイート例を図 2-5 に示す。

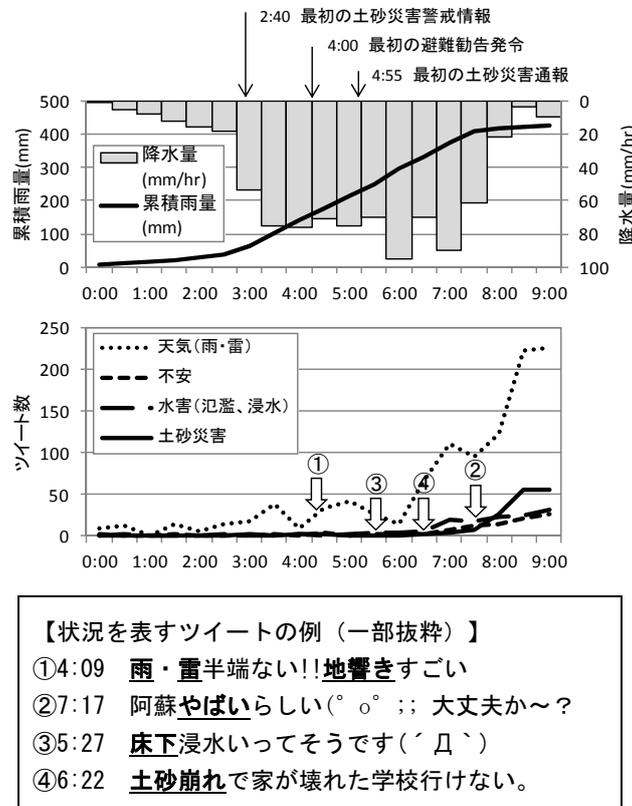


図 2-5 ツイートによる災害情報の把握（九州北部豪雨）

天気カテゴリーのツイート数は継続して増加しており、激しい降雨が継続している状況が推察できる。また、7 時頃から水害や土砂災害カテゴリーのツイート数が急増しており、災害が発生又は危険が認識されている状況が推察できる。一方、激しい降雨や実際に土砂崩れが起きた状況等を具体的に示す内容のツイートも抽出された。これらの発言からはより具体的な現場の状況を把握できる。広島の結果を同様に図 2-6 に示す。阿蘇と同様、ツイート数の変化及びツイートの内容から、激しい降雨が続いている様子や災害が発生している状況が推察できる。

また、広島の場合は、阿蘇と異なり、土砂災害の発生より早い時間帯でツイート数の急増（例えば気象関連の2時30分頃や土砂災害関連の2時頃）が確認できる。要因として、人口規模が大きい都市部であり、分析対象となるツイート数が多いことが挙げられる。

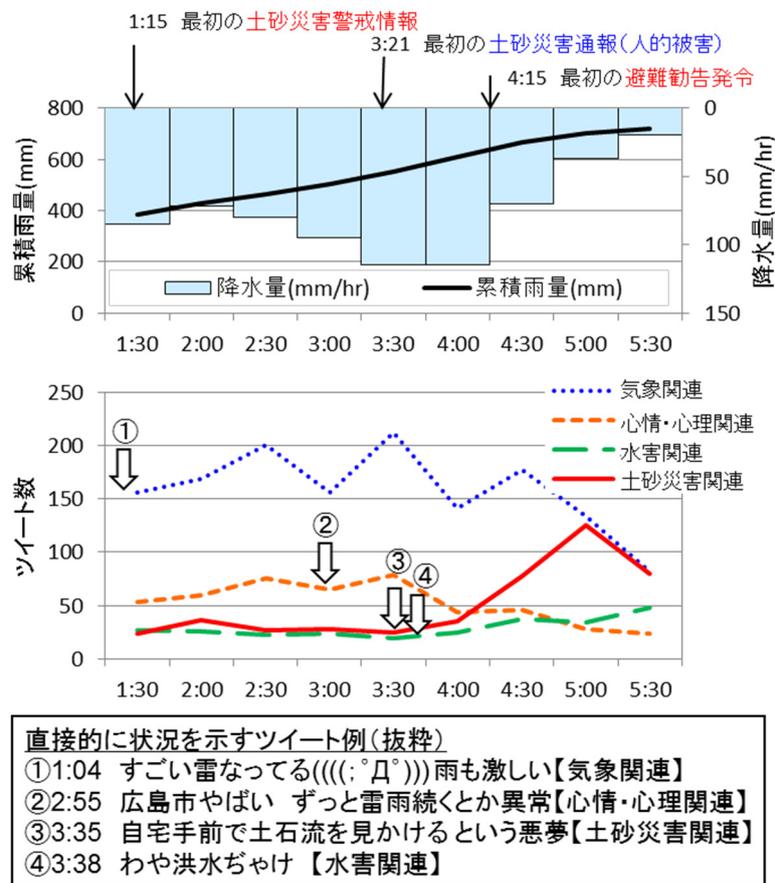


図 2-6 ツイートによる災害情報の把握（広島豪雨）

2.3.5 まとめ

本検討の結果、土砂災害に関連するツイート情報のカテゴリ分けおよびキーワード設定を行うことで、設定したキーワードから抽出したツイートの件数の変化や発言内容をもとに危険性の検知や災害状況の把握が可能であることが確認できた。

検討の対象外とされた位置推定ができていないツイートには「地鳴り」等の重要な語が含まれていることから、位置推定精度の向上による分析対象となる Twitter 情報の量及び正確性の確保、及び、設定したキーワードの組合せ（共起語等）により、災害に関連する発言内容をよりの確に把握する技術や、その他強雨域の分布等、物理センサでの計測情報と重ね合わせて表示させることにより、防災情報としての有効性の向上等の検討が課題である。

2.4 警戒期から発災前における状況把握の可能性

前節までの調査結果から、Twitter 上では、発災に関する目撃情報や、警戒期において特徴的な情報が投稿されている場合があることがわかった。こうした情報を防災現場で有効に活用するためには、災害に関係するツイートを収集して可視化する一連の処理が、自動化されることが望ましい。一方、Twitter 上に投稿されるデータは多種多様であるため、防災に関連するツイートに絞り込む必要がある。また、災害に関連するデータに限ったとしても、災害規模が大きな場合においてはデータ量が増加する場合もあると想定される。このため、統計解析手法等を用いて、システムが自動的に情報を選別し、状況を可視化するような仕組みが必要だと考えられる。

そこで、本研究では、統計的手法を用いて、Twitter データから災害が発生する前の状況を把握できるかどうか調査した。具体的な調査項目は以下のとおりである。

- 状況が把握できる可能性を定量的に評価するため、特定の災害事例を対象として、Twitter 全量データに対して統計的な手法を適用し、警戒期から発災前に出現するデータの特徴を調査した。
- ソーシャルセンサとしての特性を整理するために、長期間のデータを用いて、雨に対する Twitter を介した人の反応を調査した。
- 以上の調査結果を踏まえ、統計的な手法の適用を行う場合の課題を整理した。

2.4.1 警戒期における Twitter データの定量的調査

(1) 調査の目的・アプローチ

本調査では、土砂災害が発生する前の警戒的な情報を Twitter データから抽出することを目的として、土砂災害の発災前後の Twitter データの全体像を掴み、Twitter データから抽出できる可能性がある情報を整理した。この調査のため、過去の土砂災害事例を対象として、Twitter の全量データを用いた俯瞰的な分析を実施し、統計的な観点から警戒期の情報を整理するアプローチをとった。警戒期に利用される気象データやカメラ画像等の多種多様な情報に加え、大量に存在する Twitter データから前兆現象等の情報を収集し、迅速に防災対応に役立つには、将来的に処理を自動化することが望ましい。そのため、大量に存在する Twitter データからシステムが自動的に防災上有用な情報を抽出することを念頭においたものである。俯瞰的な分析としては、ツイート本文に対して話題分析を行って警戒期の話題の傾向を押さえた後、特定話題の傾向を分析するために共起語の調査を実施した。

(2) データ準備

本調査では、以下に示す2つの災害事例を対象として、Twitter データの分析を行った。

事例 (a) 2012 年 7 月に発生した九州北部豪雨災害における熊本県で発生した土砂災害、2012 年 7 月 11 日 16 時から 12 日 12 時までの Twitter データを調査

事例 (b) 2014 年 8 月に発生した広島市における土砂災害、2014 年 8 月 20 日 1 時から 5 時 30 分までの Twitter データを調査

Twitter データは、Twitter データを販売する業者から購入したもので、当該期間に日本でツイートされている全データを対象としたものである。本データに対してリツイートを除き、GPS、プロフィールおよびツイート本文の情報を元に都道府県を推定し、データの母集団とした。

(3) 話題分析

土砂災害の発災前後における、Twitter 上の話題を俯瞰し、出現する話題の経時的変化を調査した。具体的には、都道府県および時間帯毎にツイート本文中に出現する単語のランキングを作成し、話題分析を行った。ランキングの作成にあたっては、特定の場所と時間帯において特徴的に出現する単語を捉えるため、TF-IDF によるスコアを用いた。

TF-IDF は、単純に単語の出現回数の大小でスコアリングするのではなく、文書全体で出現しやすい単語のスコアを小さくし、特定の文書に集中して出現する単語のスコアを大きくする計算手法である。本調査では、全国のツイートを母集団とし、都道府県単位に仕分けた上で 1 時間毎にデータを集約し、TF-IDF の計算における文書の単位とした。この方法により、全データと相対的に比較した上で、特定の都道府県および時間帯に集中して出現した単語の傾向を可視化することができる。このため、災害が発生した都道府県において、災害が発生する前に限って特徴的に出現するキーワードを調査することができ、後の調査の土台となる情報を得られると考えた。

事例 (a), (b) の調査結果を図 2-7 へ示す。調査の結果、災害発生前には、災害に関連するキーワードとして雷、停電、豪雨等のキーワードが頻出する傾向があることがわかった。

事例(a)H24熊本県土砂災害								事例(b)H26広島市土砂災害																	
2012-07-11				2012-07-12				2014-08-20				2014-08-20													
23:00-24:00		00:00-01:00		01:00-02:00		02:00-03:00		03:00-04:00		04:00-05:00		05:00-06:00		06:00-07:00		01:00-02:00		02:00-03:00		03:00-04:00		04:00-05:00		05:00-05:30	
1	ハニラジ	雷	雷	ジャミルさま	雷	生活用品	添加剂	雷	雷	雷	雷	雷	雷	雷	雷	雷	1	何処	停電	午	ばなそにつく	せんばあ			
2	絵描き枠	一瞬停電	味増汁	光る	マモ	れいくん	バナンパリン	雷	雷	雷	雷	雷	雷	雷	雷	雷	2	譲る	雷	停電	避難	避難勧告			
3	雷	ポスト数	雷	果報	カミナリ	雑貨	インテリア	雷	雷	雷	雷	雷	雷	雷	雷	雷	3	チケット	せんばあ	雷	避難勧告	ラグビー			
4	おやすみなさい	雷	雷	果報	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	4	停電	シロシロ	ファンネルミサイル	停電	景ちゃん			
5	熊クマイケメン	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	5	雷	雷	中華	占領	発令			
6	ブサイカーズ	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	6	ピカピカ	ソロソロ	落雷	のびる	わしわし			
7	恭一郎	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	7	警報	ピカピカ	雷	氾濫	開設			
8	寫字時の筆跡	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	8	シャロ	光る	せんばあ	消防車	雷雨			
9	マムージャ	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	9	ノヤ	豪雨	あとアイチ	せんばあ	氾濫			
10	家族軍兵力報告	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	10	光る	まき	けんじる	京葉線	新			
11	トロール	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	11	ずばり	一印象	バグ	避難場所	マツピオ			
12	人生放棄	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	12	入浴時間	落雷	冠水	四時	避難場所			
13	特性	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	13	願う	自発	サイレン	水位	根谷川			
14	しるぽッチャマ	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	14	やや	殺生丸様	鳴る	安佐南区	バラダイステレビHD			
15	ロボ製作所	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	15	ゴロゴロ	離子	避難	ハトカー	可部小学校			
16	通信欄	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	16	豪雨	雷鳴	きいちごババ	vahrehvahさん	テレビ王国			
17	Androidアプリ	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	17	一印象	大雨	竜巻注意報	幼児語	大林小学校			
18	死者	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	18	ビートルズ	ゴロゴロ	竜巻	びおか	微笑			
19	マモル	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	19	扱い	信号	カミナリ	一次体制	可部学区			
20	光る	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	20	雨雲	カミナリ	ピカピカ	根谷川	大林学区			

図 2-7 TF-IDF スコアのランキング (話題分析)

事例 (a) 熊本県と推定されたツイートで集中して出現した単語

事例 (b) 広島県と推定されたツイートで集中して出現した単語

事例 (a) を対象に、武田ら (2013), 武田ら (2014) で Twitter データの検索に用いたキーワード群と、土砂災害の前兆現象である土の臭いや地鳴りに関連するキーワード¹⁶⁾¹⁷⁾、および、前節での調査結果を元に選定したキーワードを用いて、それぞれのキーワードを含むツ

ツイート数の推移を調査した（図 2-8）。この調査結果から、量的な面では「怖い、やばい、すごい」「音」「停電」「注意報・警報」「強い雨の観察」に関する発言が発災前に増加することがわかった。また、ツイート数の増加率に関する調査結果から、土砂災害に関する前兆的ワード(小石, 土臭い, 地鳴り, 地響き)が増加する時間帯があることがわかった。

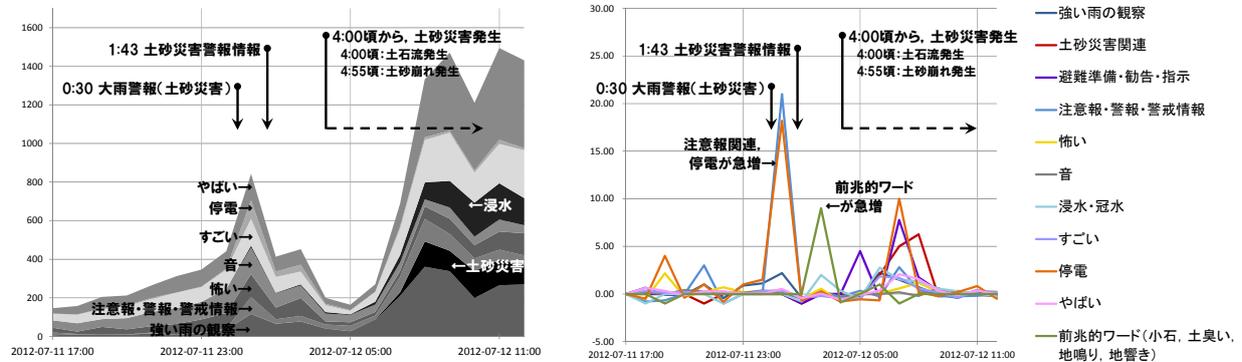


図 2-8 事例(a) 熊本県内のツイート数の推移
 (左)ツイート数の推移, (右)増加率の推移 ※1 時間単位で集計

(4) 雨発言に対する共起語の調査

前節の調査結果の中で、「怖い、やばい、すごい」「強い雨の観察」「音」に関する発言が警戒期に増加することに着目し、ツイートに含まれる共起語の推移を調査した。事例(a)に対する、検索キーワード別の共起語ランキングの推移を、図 2-9 へ示す。調査の結果、発災の直前から、災害に関する単語やネガティブな単語が、共起語の上位にランクインすることがわかった。また、全般的に「雨」や「雷」が共起語として頻出する傾向にあり、前節の調査と合わせて Twitter 上では「強い雨の観察」に関する情報を抽出できる可能性があると考えられる。降雨の情報については、雨量レーダの情報を用いると正確な数値情報を入手できるが、Twitter 上の「強い雨の観察」は、住民による観察において普段とは違う、もしくは近年の経験から見て相対的に強弱を判断した上での発言と想定できる。より定性的な観点では、夜中に目を覚まして雨の恐怖をツイートするような状況であれば、十分警戒すべき状況であると考えることができる。

事例(a) H24熊本県土砂災害											事例(b) H26広島市土砂災害										
2012-07-11											2014-08-20										
2012-07-12											2014-08-20										
		22:00-23:00		23:00-24:00		00:00-01:00		01:00-02:00		02:00-03:00		03:00-04:00		04:00-05:00		05:00-06:00		06:00-07:00		07:00-08:00	
雨	1	降る	降る	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷
	2	くる	くる	降る	すごい	すごい	すごい	降る	猛烈	寒い	すごい	寒い	すごい	寒い	すごい	寒い	すごい	寒い	すごい	寒い	すごい
	3	明日	明日	すごい	明日	すごい	明日	すごい	みる	すごい	みる	すごい	みる	すごい	みる	すごい	みる	すごい	みる	すごい	みる
	4	今日	すごい	くる	風	怖い	怖い	怖い	猛烈	解折	降る	寒い	寒い								
	5	ない	雷	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る
怖い	1	私	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷
	2	笑	女子	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷
	3	すぎる	ない	寝る	光る	雨	雨	雨	眠れる	話	今	やばい	やばい								
	4	一番	笑	私	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る
	5	そう	思う	すぎる	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る
すごい	1	思う	雨	雷	雨	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷
	2	いい	ない	雨	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷
	3	ない	思う	雨	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷
	4	やる	言う	今	今	今	今	今	今	今	今	今	今	今	今	今	今	今	今	今	今
	5	気	可愛い	いい	思う	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る
やばい	1	笑	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷
	2	すぎる	行く	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷
	3	そう	ない	すぎる	笑	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る
	4	くる	雨	笑	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る
	5	いい	いう	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ	死ぬ
音	1	やる	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷
	2	感じる	聞こえる	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷	雷
	3	時	よう	怖い	聞こえる	雨	雨	雨	雨	雨	雨	雨	雨	雨	雨	雨	雨	雨	雨	雨	雨
	4	立てる	くる	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい	うるさい
	5	思う	すごい	いい	考える	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る	寝る

図 2-9 キーワード別共起語ランキング

(5) まとめ

本調査では、過去の災害事例を中心に俯瞰的なデータ分析を実施し、Twitter 上で警戒期から発災前に特徴的に出現する情報を整理した。調査結果から、Twitter 上では警戒期に特徴的な情報が出現する可能性があることがわかった。

2.4.2 雨発言を利用した潜在的ユーザー分布の可視化

(1) 調査の目的・アプローチ

Twitter データをソーシャルセンサと仮定して防災用途に利活用することを考えた場合、適用範囲を見極めるため、センサとしての網羅性や感度を把握しておく必要がある。具体的には、気象や災害等の事象の発生に対して、Twitter を介して人がどのような反応を示すのか、定量的な評価を行う必要がある。一方、前節までの調査で、Twitter 上では、雨に関するツイートが比較的多く存在することがわかった。このため、全国の雨に関するツイート数と雨量との関係を統計的に整理することで、全国の潜在的な Twitter ユーザーの分布や感度を可視化できるのではないかと、仮説を持った。この仮説を検証するため、著者が収集した 2012 年 7 月から 2015 年 6 月までのツイートを対象として、市町村毎の雨量とツイート数の関係を、統計解析手法を用いて考察した。

(2) データ準備

前処理として、Twitter データと雨量データをそれぞれ以下のように加工・集約した上で、場所（市町村単位）と日付をキーとして、データ結合を行った。総データ件数は、1,791,878 件となった。

1. Twitter データの準備

「雨」を検索キーワードとして Search API を用いて収集した、2012 年 7 月 4 日から 2015 年 6 月 30 日のデータを母集団とした。母集団に対して、(1)GPS、プロフィールおよ

びツイート本文を用いて、ツイートの場所を市町村単位で推定し、(2)リツイート、伝聞やニュース等の目撃情報以外のツイートをテキスト解析処理・機械学習処理を用いて排除した。このデータ群に対して、市町村別に1日単位で集計し、集計テーブルを作成した。

2. 雨量データの準備

気象庁が提供する解析雨量データを雨量データとして用いた。前処理として、緯度経度情報から住所情報を推定する逆ジオコーディング処理を用いて、全てのメッシュを市町村へ対応付けた⁵。その上で、市町村毎に、領域内に含まれる雨域（雨量が0でないメッシュ）の平均雨量を当該市町村における平均雨量（時間雨量）として算出し、更に、平均雨量の24時間分の合計値を、当該市町村の1日雨量として算出した（図2-10）。集約することで、雨の降り方についての詳細な情報は消失するが、Twitterデータとの突き合わせを考慮し、ある程度まとまった単位での集計とした。

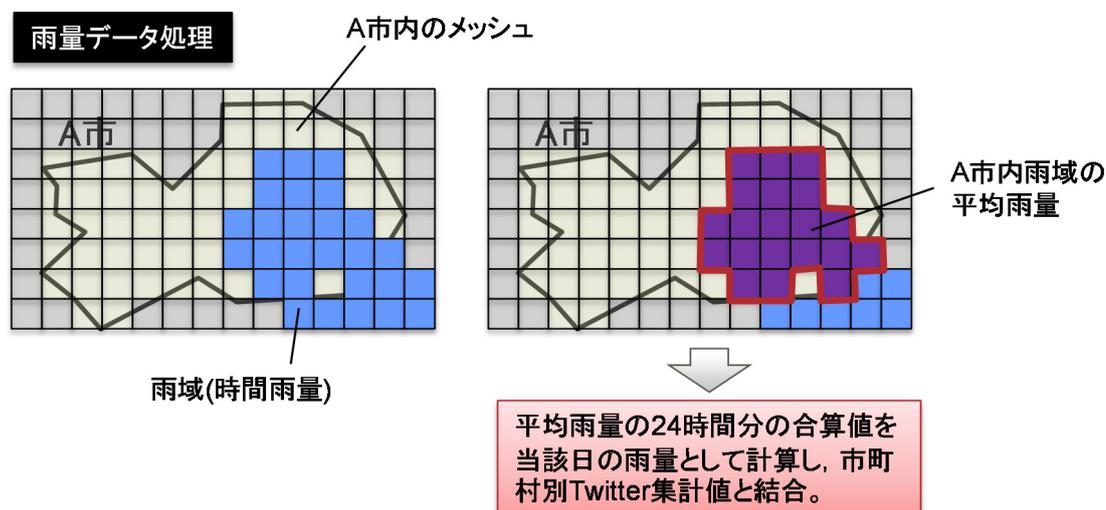


図 2-10 雨量データの準備方法

⁵ この際、海上のメッシュは排除する処理を行った。住所起点の情報からの距離を用いた処理を行ったため、ある程度のノイズを含む。

(3) データ予備調査

雨に関するツイート数を変動させる要因として、①雨量の大きさ、②地域毎の人口、③雨量・人口以外の地域差の3つが関係するのではないかとの仮説を持った。この仮説を確かめるために、データを可視化し関係性を確認した。データの可視化においては、以下2つの観点から調査した。

1. 「①雨量の大きさ」の影響を確かめるために、特定の地域に絞って②および③の要因を排除して、雨量とツイート数の関係を可視化し傾向を調査した。ここで、雨量の大きさとツイート数に明確な関係がみられる場合、ソーシャルセンサとしての Twitter データが、雨量の大きさに対して反応が変わるものと想定される。
2. 「②地域毎の人口」による影響を確かめるために、市町村毎の人口差を是正したデータを用いて、雨量とツイート数の関係を可視化し傾向を調査した。ここで、1. で考察した結果と同様の傾向が見られた場合、地域毎の反応の違いが「②地域毎の人口」、すなわちソーシャルセンサの数のみの影響を受けると推測される。一方、1. で考察した結果と異なる傾向が見られた場合、「③雨量・人口以外の地域差」が存在するものと想定される。

「①雨量の大きさ」を確かめるために、鹿児島市のデータを対象として調査した。鹿児島市の全データの散布図を図 2-11 に示す。図より、雨量が大きくなると、雨に関するツイート数が増加する傾向にあると想定される。

次に、「②地域毎の人口」の影響を分析するための調査を実施した。用意した全データ（全国市町村のデータ）に対する散布図を図 2-12 に示す。図（左）は、1日雨量とツイート数の関係を示したものである。一方、図（右）は、市町村毎の人口差を是正するために、ツイート数を市町村の人口で除して調整したものである（以下、発言率という）。本調査では、各集計値の1日単位で集計しているため、動的な昼夜の人口変動を考慮する必要はないが、人

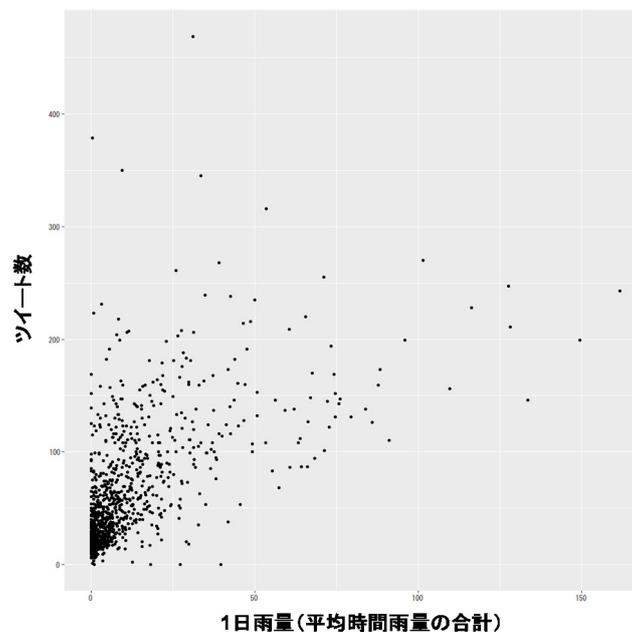


図 2-11 鹿児島市における集計データの散布図

口のベースラインの算出方法を定めておく必要がある。そこで、総務省より公開されている統計情報を用いて、昼間人口と夜間人口の平均値を算出し、当該市町村の人口のベースラインとした⁶。

この図 2-12（右）をみると、雨量の大きさが大きくなったとしても、必ずしも発現率が増加するとは言えないことがわかった。また、雨量が小さい範囲においては、発現率のばらつきが大きい傾向にあった（図中の赤丸で囲った部分）。これらの傾向は、図 2-11 から考察された傾向とは異なるものである。したがって、雨に対するツイート数は、市町村人口や雨量以外の何らかの要因、すなわち「③雨量・人口以外の地域差」の影響を受けている可能性がある。これらの差異は、Twitter 上のデータを用いて統計処理を行う上で考慮すべき観点であり、定量的な評価が必要と考えられる。

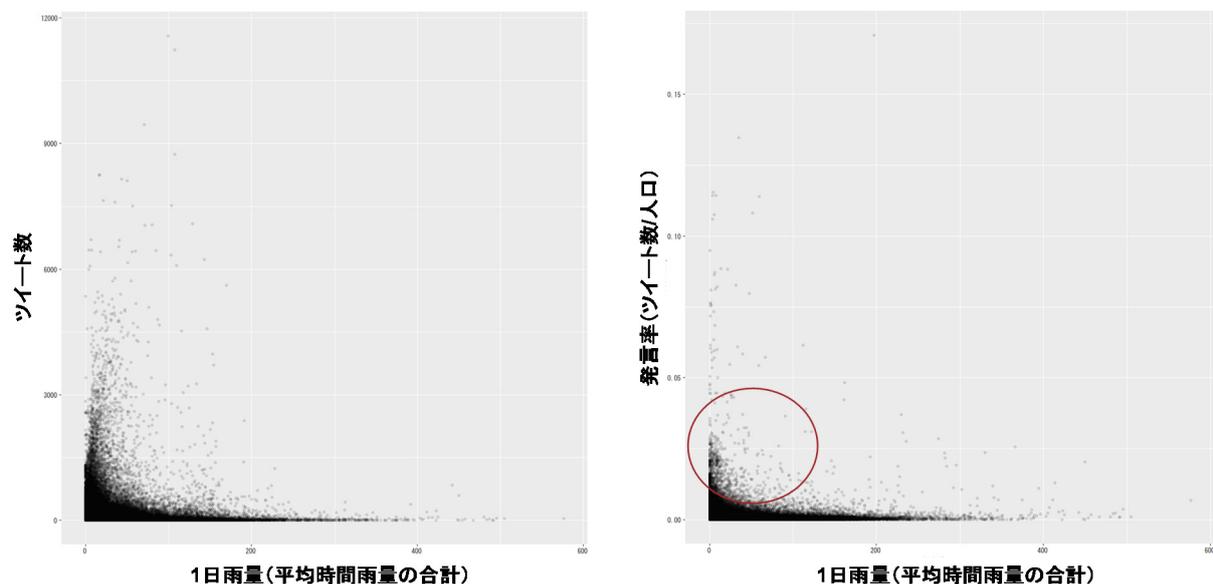


図 2-12 全データの散布図

(左) 1日雨量とツイート数の関係を示した図

(右) 1日雨量と人口調整後のツイート数（発言率）の関係を示した図

以上を踏まえ、定量的な評価を行うためには、市町村別の人口格差を排除した上で、統計モデルによる市町村の比較を行う必要がある。上の散布図では、ツイート数の人口調整のために、除算処理を実施した。しかしながら、被除数が小さく、かつ、人口の格差が大きい場合には、極端に人口が小さい地域において、調整後の値が過剰に評価されるという問題が発生する。この問題を、Small Number Problem といい、疫学等の分野において統計解析処理を

⁶ これは、人の自宅外での活動時間帯のピークを朝 8 時～夜 8 時頃と想定したものであるが、あくまで仮説に過ぎない。全国市町村の人口変動を正確に把握することは現時点で困難であるため、本調査ではこの想定にしたがった。

行う際の留意点とされている⁷。この解消のための一つの方法として、階層的なベイズモデルを用いて、地域別の真の相対度数（人口調整後のツイート数）を潜在的なパラメータとして推定する手法が提案されている⁸。

本調査においても、階層的なベイズモデルを用いた推定処理を1日毎に実施し、全国市町村の人口調整後のツイート数を推定した上で、雨量データと突き合わせた。用意した全てのTwitterデータを対象に、2つの異なる方法で人口調整を行った場合の比較を図2-13に示す。図より、本調査で用いた集計データにおいても、単純な除算を用いて人口調整を行った場合、人口が小さい市町村では、集計値が1-2件となるような場合に、過剰に評価されていることがわかった。一方、階層的なベイズモデルを用いて推定した場合には、左記の問題は解消されているため、以後の解析では本データ処理を施したデータを用いた。

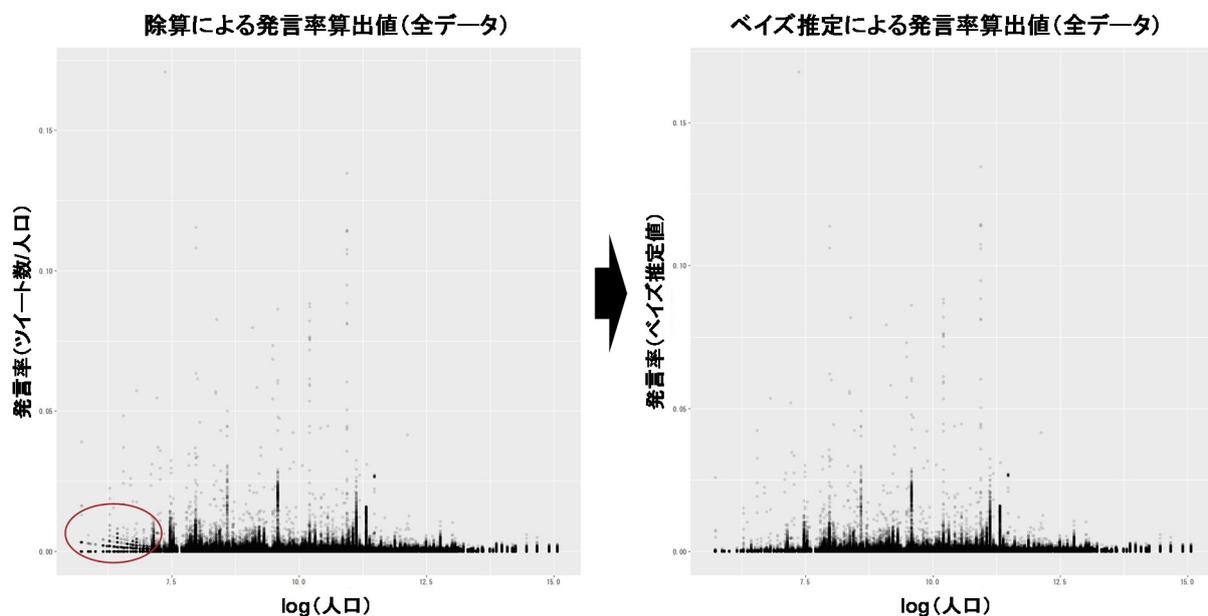


図 2-13 全ツイート数集計値における人口と人口調整後ツイート数（発言率）の関係

（左）除算によって人口調整した場合 ※赤丸箇所 Small Number Problem が発生

（右）階層的なベイズモデルを用いた推定によって人口調整した場合

(4) 統計モデルを用いた解析

市町村別の雨に対する Twitter 上の反応をモデル化するにあたり、ツイート数を目的変数、雨量を説明変数とする回帰モデルの適用を検討した。統計モデルとして、以下2つの方法を比較した。

⁷ 丹後ら (2007)²³、瀬谷ら (2014)²⁴を参照。

⁸ 丹後ら (2011)²⁵を参照。本調査においても、参考文献に記載されたモデル式を用いて推定した。モデルパラメータの推定には、Gelman et al.²⁶を利用した。

モデル① 全国と市町村の2階層を仮定した階層的な線形回帰モデル（一般化階層線形モデル）を適用し、それぞれの市町村における全国共通パラメータからの差分を地域差とみなす方法。

モデル② 市町村それぞれに対して独立かつ非線形な回帰モデルを仮定し、独立に得られたパラメータを比較して地域差を推定する方法。

モデル①は、全国的に共通性を仮定してパラメータの交換可能性を考慮したモデルを想定することで、地域差そのものをパラメータとして推定することを狙ったアプローチである。このアプローチでは、全国の平均的なモデルをベースとして、地域差が傾きや切片のバラツキとして表現することができる。このため、市町村別の観測データ数が極端に偏る場合などにおいては、市町村別に独立的なモデルを仮定するよりも推定値が安定することが期待できると考えた。図 2-14 に、モデル①を用いて解析を行った例を示す。

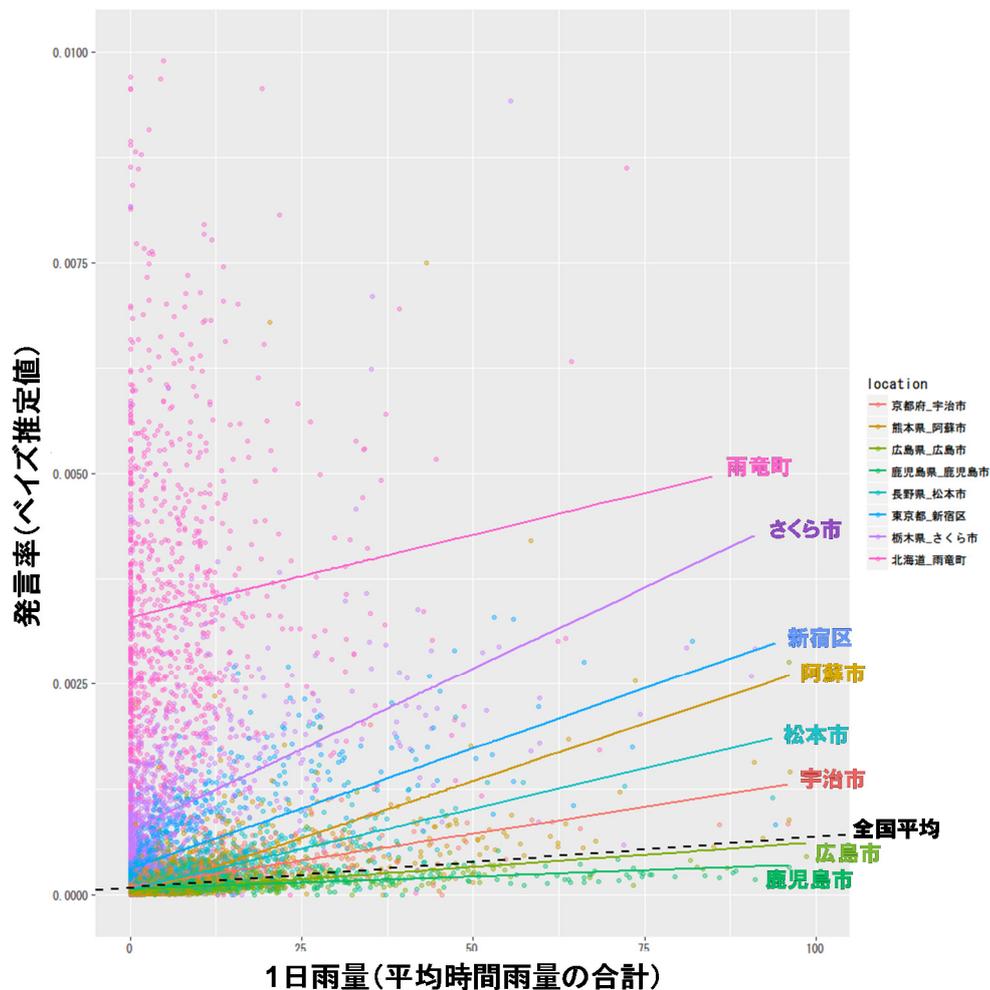


図 2-14 モデル①による解析例

(実線は市町村別、破線は全国レベルでの回帰モデル予測値を示す)

上の例では、線形モデルの傾き、切片ともにバラツキを持つモデルをあてはめた結果を示している。例えば、鹿児島市においては、全国平均と比較して傾きが小さく、Twitter を介した雨に対する反応が相対的に小さいものと考えられる⁹。しかしながら、他の事例を含めて予備調査を行った結果、地域別に雨の降り方にもバラツキや偏りがあるため、すべての市町村で線形モデルを仮定するのは難しいことがわかった。

次に、モデル②の非線形な回帰モデルを用いた解析を検討した結果を示す。非線形な回帰モデルとしては、べき乗関数等の特定の関数を固定して考えるパラメトリックなモデルと、それらを仮定しないノンパラメトリックな手法がある。データ可視化等の予備調査の結果、市町村別に必ずしも同じような非線形な関係を持つとは限らないことが観察された。そこで、ノンパラメトリックな非線形モデルとして、市町村それぞれについて、独立な一般化加法モデル¹⁰ を用いた。特定の市町村に対して、一般化加法モデルをあてはめた結果を図 2-15 に示す。

⁹ 3章に述べる現場実証実験におけるヒアリングにおいて、鹿児島県下では少々の雨では驚かないとのコメントを得ており、統計解析の結果とも一致する傾向となった。

¹⁰ 一般化加法モデルは、それぞれの説明変数に対して非線形な基底関数を仮定して適用することで、非線形な統計モデルを構築する手法である。これにより、線形でなく平滑化されたモデルを構築することができる。平滑化には、平滑化スプライン関数やテンソル積が用いられる²⁷⁾。

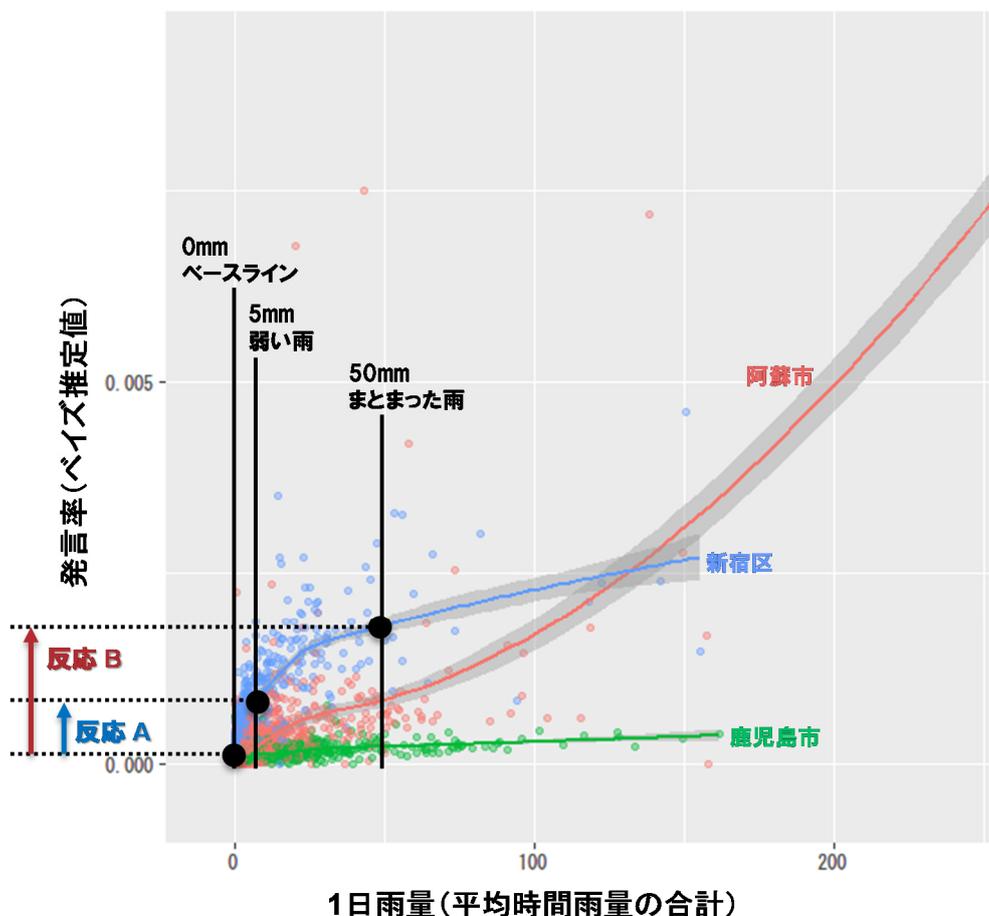


図 2-15 モデル②による解析例

(曲線は一般化加法モデルによる予測値，灰色の領域は信頼区間を示す)

推定結果より，1日雨量とツイート数の関係として非線形な関係があり，かつ，市町村別に異なる関係性を持つと考察した。一方，市町村別に独立なノンパラメトリックなモデルを仮定すると，単純な方法では地域差を定量的に比較することができない。そこで，各市町村の推定モデルにおける，1日雨量0mm，5mm，50mmにおける人口調整後ツイート数のモデル予測値を使用して，1日雨量5mmと0mmの予測値の差分を反応A（弱い雨に対する反応），1日雨量50mmと0mmの予測値の差分を反応B（まとまった雨に対する反応）として算出する方法をとった。この方法を用いて，全国市町村の雨に対する反応の違いを数値化し，地図上に可視化したものを図 2-16、図 2-17 に示す。図では，色が濃い程反応が大きいことを示している（等量区分にて10段階に色分けした）。

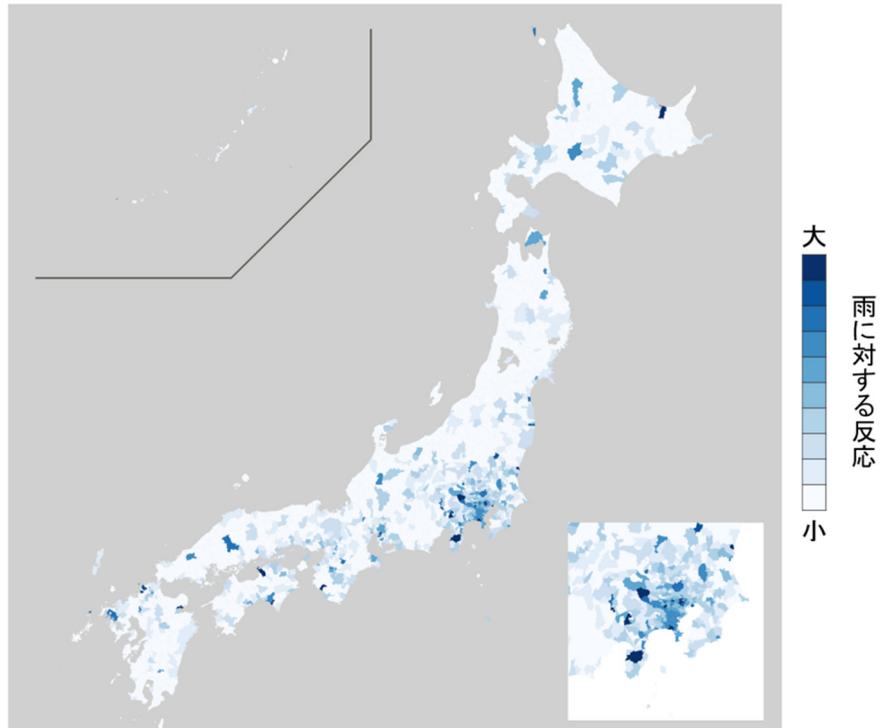


図 2-16 反応 A (弱い雨) に対する反応の分布

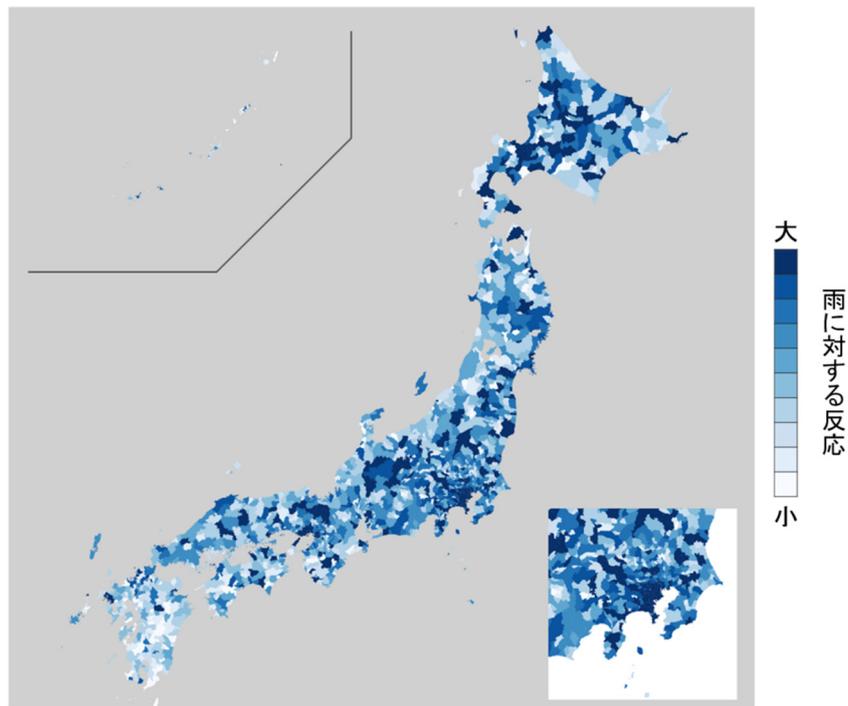


図 2-17 反応 B (まとまった雨) に対する反応の分布

図より、雨に対する反応は地域によって差異があることがわかった。九州南部や四国の一部地域では相対的に反応が小さいく、日常的に雨がよく降る地域では、雨に対して大きな関心を寄せない可能性がある。また、反応の大小はあるものの、全国的にみると、多くの地域で何らかの Twitter 上の反応があることが分かった。このため、統計的な前処理を行うこと

で、人口が小さい地域においても、武田ら（2013），武田ら（2014）で用いた統計解析手法¹⁶⁾¹⁷⁾を適用して、Twitter 上の反応を定量的な情報として抽出できる可能性がある。

なお、ここで示した地域差には、地域別の人の反応の差に加えて、テキストデータを用いて場所を推定する上での曖昧性を含んでいると想定される。本調査では、プロフィールやツイート本文に含まれる住所やランドマークの情報を用いて場所を推定した。この推定処理においては、単語と場所をマッチングさせる辞書の影響を受けることに加え、自然言語処理の特性から一部確率的な処理を含んでおり、出力結果に偏りが発生する可能性がある。したがって、本調査で得られた地域差に相当する解析結果は、純粹に人の反応の差のみを示すものとは言えないが、一連の処理プロセスの結果として捉えた場合に、システムの出力結果の偏りを示すものであると考察した。このため、統計的な処理を用いて、全国のツイート数を評価する場合においては、地域の人口差に加えて、本調査で得られた潜在的な反応の差（例えば、降雨に対する地域別の人の反応の差）を考慮する必要があると想定される。

(5) 本調査で得られた解析結果の活用例

過去の災害事例として、2014年8月20日に発生した広島市の土砂災害事例を取り上げる。本事例では、2.4.1節に示すように、土砂災害が発生する数時間前から、Twitter 上で強雨に関して言及するツイートが投稿される傾向にあった。そこで、「雨」を含むツイート群に対して、更に「強い」「怖い」「やばい」等の単語を共起語として含むツイートに絞り込み、2.4.1節で示したノイズ排除処理および場所推定処理を行ったところ、広島県内で01:00から02:00の間に65件のツイートが存在することがわかった。

こうした傾向を、統計的な処理を用いて自動的に抽出できるとするならば、災害発生前の現場で起きている状況を、システムが自動的に検出してアラートを上げることができる可能性がある。ただし、実際の運用を考えた場合には、どの場所で日常と異なる状況が発生しているか不明であることから、全国のデータを横並びで把握する必要があると想定される。そこで、当該時間帯の全国集計値を比較したグラフを、図 2-18 へ示す。



図 2-18 強雨観察ツイート数 (2014年8月20日 01:00-02:00)

上図より、確かに全国的に見ても広島県のツイート数が突出していることがわかる。一方、全国のツイート数を、相対的な観点で比較する場合には、人口による調整を実施する必要がある。そこで、本節で議論した方法を用いて、当該時間帯の全国ツイート数の人口格差を階層的なバイズモデルを用いて是正した結果を図 2-19 (上) に、独立な一般化加法モデルを用いて地域の反応差を考慮して調整した結果を図 2-19 (下) に示す。

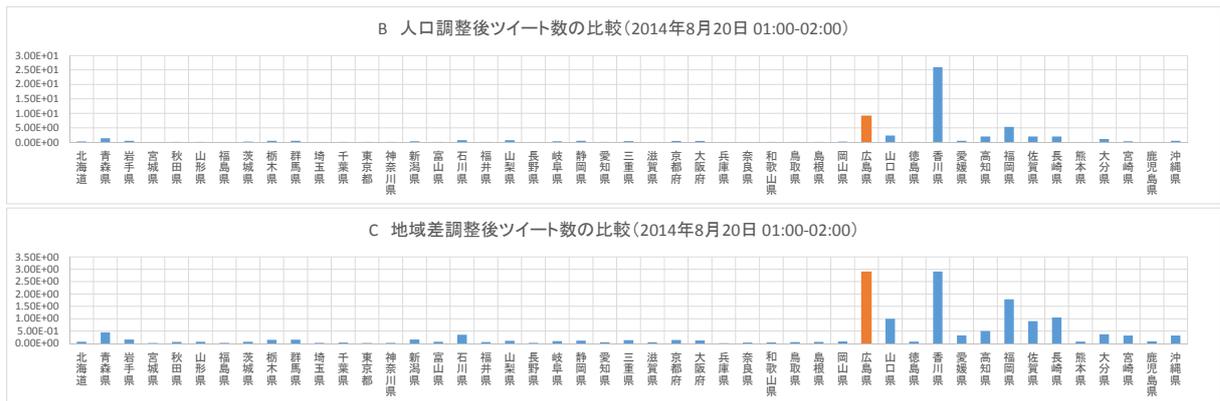


図 2-19 調整後の強雨観察ツイート数 (2014年8月20日 01:00-02:00)

調整後の強雨観察ツイート数を見ると、単純に人口調整のみを適用した場合には、広島県よりも香川県が、相対的にツイート数が多くなる結果となった。この結果に対して、地域差を考慮した調整を行うと、広島県、香川県がそれぞれ相対的に大きく評価され、統計処理を行った場合に、どちらの場所においても異常検知等の手法¹¹を用いて情報を抽出できるものと想定される。なお、香川県におけるツイートは、当時、まんのう町やさぬき市、東かがわ市で1mm全後の雨量が観測されており、これに対する反応と考えられる。しかし、強雨が観測された広島県と比べ、雨量が小さかったにも関わらず同程度のツイート数となっている。この要因としては、場所の推定処理の誤りや偏りによって、広島県のツイートの一部が香川県と推定されたことや、伝聞や意見等の直接的な観察ではないノイズとなるツイートが存在していたことの影響が大きかった可能性が考えられる。

(6) まとめ

本調査では、統計モデルを用いて、雨という自然現象に対する Twitter を介した人の反応をモデル化し、ソーシャルセンサとしての Twitter の特性を調査した。また、全国市町村の潜在的な Twitter ユーザーの分布を可視化することを試みた。

この調査の結果、Twitter を介した人の反応は、人口以外の隠れた要因によって差異が生じる可能性があることがわかった。また、調査の過程において、統計的な観点から人口調整を行う場合には、Small Number Problem への考慮が必要ながわかった。

これらの観点は、Twitter データをソーシャルセンサと見なして統計解析処理を行う場合において、考慮すべき事項であると考えた。具体的には、地域別に Twitter 上の話題(災害、気象等)の発生を自動的に抽出することを考える場合、人口格差と隠れた地域差を考慮することで、全国的に偏ることなく情報を抽出できると想定される。この仮説は、特定の事例において効果があることがわかった。また、Twitter 上では全国的に自然現象等に関するツイ

¹¹ データ群において、通常のデータとは発生量や振る舞いが異なるデータを検出することを異常検知という。統計的・確率的な手法が数多く提案されている^{28), 29), 30)}。

ートの発生が期待できることがわかった。

以上の考察は、約3年間のデータを使った実験より導いたものである。しかしながら、気象現象の周期性や土砂災害の発生頻度を考慮すると、短い期間のデータに基づく調査結果である。降雨に対するTwitterユーザーの反応は、雨期（6月～10月）とそれ以外の時期における違いや、強雨の程度による差異なども考えられる。また、隠れた地域差には、地域やTwitterユーザーの防災意識などが影響することが考えられる。そのため、ソーシャルセンサの防災分野の活用を検討するにあたっては、社会実装を行いながらデータを収集し、今後も継続的な分析調査を行っていく必要があると考える。

2.5 土砂災害の前兆現象等情報の収集・把握手法の可能性と課題

2.5.1 Twitter データに対する統計的アプローチにおける課題

(1) アプローチ

本共同研究では、土砂災害警戒期の Twitter データを用いて、統計的な観点で情報を抽出するための基礎的な調査を実施し、以下の考察を得た。

- ① 土砂災害の発生前において、Twitter に投稿されるツイート文に、特徴的な表現（キーワード）が出現する可能性がある。【2.4.1 節より】
- ② Twitter データをソーシャルセンサと仮定した場合、センサの感度は人口以外の隠れた要因により変動する推定されるが、その要因を統計的にモデル化することで、センサ感度の全国的なバラツキを是正できる可能性がある。【2.4.2 節より】

上記①から、Twitter データを活用することにより、土砂災害発生前の特徴的な情報を抽出できる可能性があることが分かった。このため、Twitter 上の関連ツイートをソーシャルセンサと見立て、武田ら（2013）、武田ら（2014）に示すような異常検知の手法¹⁶⁾¹⁷⁾を適用することにより、土砂災害が発生する前の警戒的な情報を自動的に抽出することが可能であると考える。この考えに基づいた、Twitter データ活用のアプローチを以下に示す。

- 上記①で明らかになった特徴的なキーワードを元に Twitter データの収集を常時行う。
- 収集したデータに対して、武田ら（2013）に示すノイズの除去処理¹⁶⁾および武田ら（2013）、武田ら（2014）に示す場所の推定処理¹⁶⁾¹⁷⁾を経て地域別に集計する。その後、上記②のセンサ感度のバラツキを是正する処理を実施し、異常検知処理における特徴量とする。
- 上記特徴量を入力として統計的な異常検知手法を適用し、平常時と異なる状態を自動的に検知し、システムの利用者等にアラートを送信する。異常検知手法としては、一例として、平常時のデータを用いてモデル化し、時間・地域（空間）的に突出したデータを捉える手法の適用が想定される¹⁶⁾。
- アラートが発生した場合、システムの利用者がデータの内容を把握できるようにするため、異常検知に寄与したツイートや関連する投稿写真、キーワード別の集計値の推移等をアラート情報と紐づけて蓄積する。

上記アプローチにより、土砂災害に関連のある Twitter データを常時収集しながら、統計的手法を用いて自動的に監視するシステムを構築することができると想定する。

(2) 課題

前項で示したアプローチを防災業務に適用するには、以下の課題を解決する必要がある。

1. 異常検知手法の継続的な検討

2.4.1 節の調査で明らかになった特徴的なキーワードの中には、雨などの自然現象に対する観察を意味するものが多く見受けられた。このように、人が自然現象を観察して目撃した際のつぶやきを利用する異常検知では、過去に経験したことがないような豪雨や災害に遭遇した場合の言葉の表現、防災意識等の変化による災害に関するツイート数や言葉の変化などが影響する可能性が考えられる。このため、異常検知手法のアルゴリズムを検討する上でその妥当性を担保するためには、より長期間のデータを用いて継続的な調査を行う必要があると考える。また、本研究で取り扱った土砂災害の発生間隔は、これまでソーシャルメディアが利用されてきた期間と比べて長いため、アルゴリズムの精度検証は継続的に実施されるべきである。

2. 場所推定精度・細かさの改善

Twitter 上に投稿されるツイートの多くは位置情報 (GPS) が付与されていない場合が多く、ユーザーのプロフィールやツイート本文から投稿者の所在地を推定する必要があるため、リアルタイムで把握できる情報の細かさは現時点では市町村レベルの情報となっている¹⁷⁾。一方、ツイートから得られた警戒的な情報を防災情報として活用するためには、より詳細な場所の情報があるのが望ましい。このため、各ツイートもしくはツイート群から得られた情報の発生源をより細かく推定するための技術開発が必要である。この課題の解決に向けては、仮説として、Twitter データ以外の気象情報、土壌雨量指数など動的に変化するメッシュ情報を活用し、場所を絞り込むなどのアプローチが考えられる。

3. 適用条件の整理

武田ら (2015) の調査により、人口の少ない地域であっても、ある程度 Twitter の利用者が存在することがわかっている¹⁷⁾。しかしながら、人口が著しく少ない地域や、起きている人が相対的に少なくなる深夜時間帯などにおいては、災害の警戒期に外界の状況を観察する人の数が減少するため、ソーシャルセンサの網羅性や感度に影響すると想定される。

また、災害の外力や災害そのものの規模によって、ソーシャルセンサの感度も異なると考えられる。このため、多種多様な災害事例を通して、本アプローチを検証し、適用条件を整理することが課題である。適用条件の整理においては、異常検知手法の汎用的な性能を検証するために、多くの事例を用いて統計的な整理をするべきである。

上に示した課題事項は、主として Twitter データを軸とした災害情報抽出手法やその適用範囲に関する課題事項である。一方、他の防災情報との関連を考えた場合、以下のような課題がある。

4. 防災情報としての表現方法の検討

防災分野における新たな情報ソースとして Twitter データを位置付けた場合、他の防災情報とどのように関連付けて活用するか、統合的な情報管理の方法を検討する必要がある。この検討のためには、適用条件の整理結果を起点として、Twitter から得られた情報を効果的かつ効率的に防災関係者が状況を把握できるようにするような新たな表現方法を検討することが必要である。具体的には、Twitter から得られた情報が他の防災情報と並列で提示された場合に、防災関係者にその位置付けをわかりやすく、かつ、誤解のないように提示することを目的とするものである。特に、Twitter の情報は曖昧性を含むため、統計的手法を用いて情報の確からしさを数値的に表現したり、ツイート本文に含まれる一次情報をコンパクトに提示したりするなどの工夫が必要と想定する。

5. 統合的な防災データの解析手法の開発

長期的な観点として、Twitter データ、気象情報、土壌雨量指数等のデータを災害に係る統合的なデータ群に位置付け、新たなデータ解析手法を駆使し、一連の情報から土砂災害が発生する前の警戒的な情報をより精度よく抽出するための検討が考えられる。一方、情報の性質、観測時間の間隔のズレ、場所の細かさが異なる多種多様なデータを活用することになるため、本共同研究や関連研究で示した手法とは別に、新たな手法を開発する必要があると想定する。

3 社会実装に向けた適用条件等の整理

3.1 SNS 情報の警戒避難システムへの組み込みに対する社会的ニーズ

内閣府は、広島市の災害を含む近年の局所的豪雨による土砂災害の発生状況を踏まえ、「総合的な土砂災害対策検討ワーキンググループ（以下「WG」と言う。）」を設置し、政府一体となって土砂災害対策を推進するための検討を行い、WGとしての提言を取りまとめている。

この提言は、以下の5つの柱で構成されている。

- ①土砂災害の特徴と地域の災害リスクの把握・共有
- ②住民等への防災情報の伝達
- ③住民等による適時適切な避難行動
- ④まちづくりのあり方と国土保全対策の推進
- ⑤災害発生直後からの迅速な応急活動

この柱の一つとして「住民等への防災情報の伝達」が掲げられているが、ここでは市町村自身が「住民等からの情報を集約したり、SNS等の情報を活用して危険が高まっている箇所を推定したりする仕組みについても強化」すること、また土砂災害の前兆現象が発生している場合や、隣接自治体で発生した場合であって、市町村が避難勧告等の判断に迷う場合には、国や都道府県は積極的に市町村に対して助言を行うことの必要性について言及されている³¹⁾。

また、消防庁は、広島市の災害を受けて「突発的局地的豪雨による土砂災害時における防災情報の伝達のあり方」について検討を行い、その結果を報告書に取りまとめている。当該報告書においては、「土砂災害の前兆現象等をもとに適時的確に避難勧告等を発令するためには、住民から市町村へ前兆現象や被害情報等の通報を受けることにより、より早期かつ確実に把握することも重要である」とした上で、「現状、電話などが主な手段であるが、災害前兆現象等の通報が集中した場合、避難の判断等の対応に支障をきたす可能性」があるとし、「SNSへの投稿内容をリアルタイムに分析し、土砂災害の前兆現象等や災害の危険性が高まっている地域を把握する技術・・・が実用化された際には、市町村において導入や活用の検討を行う」ことが重要であると指摘している³²⁾。

SNS情報を活用し土砂災害の危険性の高まりについて把握する技術については、2.3節において、広島市のように人口規模が大きな地域においては、前兆現象等の把握に相当程度有効であること、また、避難指示の判断に役立つ可能性のある住民の心情・心理等についてもかなり捉えることができ、警戒・避難システムへの応用に期待がもてることを示した。

このように、SNS情報の警戒避難システムへの組み込みに対する社会のニーズの高まりが見られることを受けて、Twitter情報を活用した災害情報収集システム（Disaster Information Gathering System Using Social Sensors）（以下「DIGSUSS」という。）の試作版を開発し、ユーザーインターフェース等のあり方について検討を行った。

3.2 Twitter を活用した災害情報収集システムの運用モデル

DIGSUSS が直接・間接を問わず、避難勧告等の意思決定に活用されることを念頭に置き、運用モデルを考えるに先立ち、意思決定に必要な情報種の分類を試みた。

避難勧告等の意思決定に必要な情報種を分類した事例を見てみると、医療分野では血圧や脈拍などを「客観的な情報」、患者自身の言語的表現を「主観的な情報」と分類されて

いることが分かった³³⁾。また、医療分野においては、「主観的な情報」は、医師が意思決定を行う際に不可欠な非常に重要な情報として取り扱われている³⁴⁾。

災害に関する情報として「どこで」「どのような状況になっているか」をツイートから把握する場合、得られる情報の内容に応じて、情報の客観性や主観性が異なることが考えられる。そこで、以下のようにツイートから得られる情報の分類を試みた（表 3-1）。

①客観的な情報

ツイート場所と状況が具体的に分かる情報として、位置情報と写真の両方が添付されたツイートを「客観的な情報」に分類した。

また、具体的なツイート場所や状況は分からなくても、それらが推定できるものを「準客観的な情報」に分類した。これは、場所推定されたツイートや状況を描写したものが該当する。なお、場所推定ができない場合でも、利用者が読めば関係する地域の者の投稿であることが大凡推測できる場合は、場所が推定できるものとして、「準客観的な情報」とした。

②主観的な情報

一方、避難勧告等の判断に活用できる可能性がある「住民の心情・心理等」を表すツイートは、場の状況を表す主観的な情報であると位置づけた。このうち、位置情報付きのものを「主観的な情報」、場所推定したもの（場所推定はできない場合でも、利用者が読めば関係する地域の者の投稿であることが大凡推測できるものを含む）を「準主観的な情報」と分類した。（表 3-1）。

表 3-1 Twitter 情報の分類

ツイート	画像	状況描写	心情・心理描写
	客観的な情報		主観的な情報
位置情報添付	客観的な情報	準客観的な情報	主観的な情報
場所推定	準客観的な情報	準客観的な情報	準主観的な情報

次に、当該システムの運用主体は、避難勧告等を発令する立場にある市町村と、市町村に助言する立場にある国・都道府県の二つが考えられる。なお、運用法が確立されるまでは当面行政内部での利用とし、一般住民の利用は検討の対象としない。

また、利用者については、それぞれの運用主体における意思決定者層（以下「リーダー」という。）と、意思決定者を補佐する者の層（以下「フォロワー」という。）の二階層が考えられる。具体的には、市町村の場合、避難勧告等の発令の意思決定を行う市町村長もしくはこれに準ずる副市町村長や部課長クラスがリーダーに当たり、これを補佐するため意思決定に必要となる様々な情報を収集し意思決定者に提供する役割を担う担当者クラスがフォロワーに当たる。これら運用モデルの型式は図 3-1 のようになる。

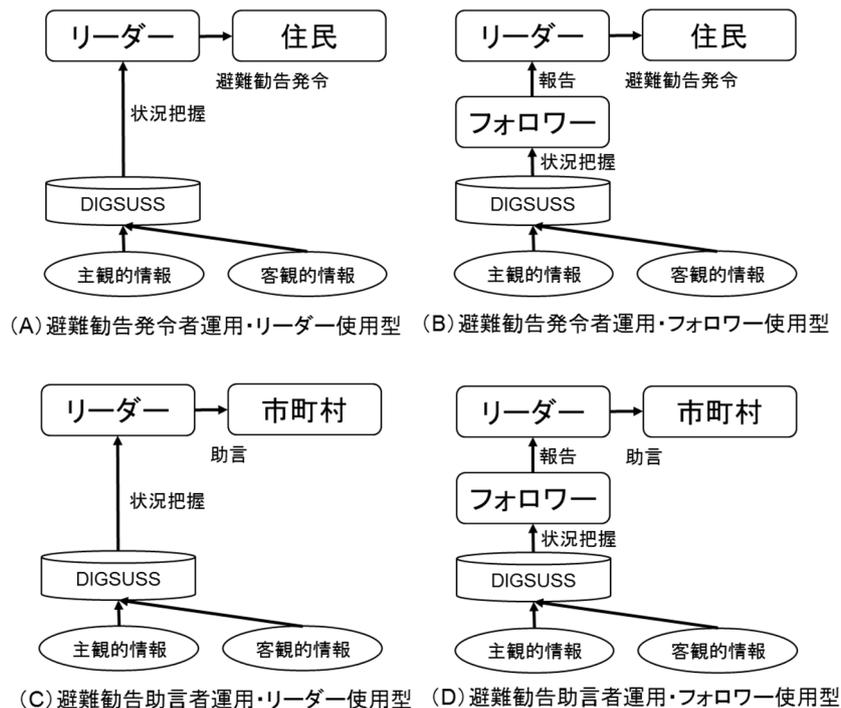


図 3-1 運用モデル形式の分類

3.3 DIGSUSS ユーザーインターフェースの検討

前章で検討した 4 つの運用モデルを想定し、客観的情報・準客観的情報と、主観的情報・準主観的情報を視覚的に確認しやすいと考えられる画面構成を、まずは試験的に 1 つ作成し、モデルユーザーに試用してもらいユーザービリティについて聞き取り調査を行い、改善点を確認する手法を採用した。

モデルユーザーは、市町村の防災担当者は人員が限られている¹⁾こと、当該システムが集中豪雨時の使用を想定していることから、市町村単位で試用調査を行う場合、調査期間が短いと対象事象に巡り会わない可能性があるため、今回は九州地方整備局、および鹿児島県に試用調査（避難勧告助言者運用型の試用）に協力頂くこととした。また、利用者については、最も頻繁に利用することが想定されるフォロワーを試用の対象とした。

3.3.1 インターフェース仕様

試用調査に用いた試作版システムの仕様を以下に記述する。システムの開発にあたり、神山ら（2014）により、雨量や災害に関連する情報を重ね合わせ一元的に表示することが災害時の状況把握に有効とされている³⁵⁾ことを参考に、地図上にレーダ雨量とツイート情報を重ね合わせて表示することとした。また、リアルタイムにツイートをタイムラインで表示する画面と地図表示画面を並べた構成とした（図 3-2）。

(1) 地図表示画面

客観的情報・準客観的情報を視覚的に確認しやすいように、位置情報付きのツイート、都道府県・市町村単位で場所推定したツイートは、画面の地図上の位置にマーカで表示した。また、降雨の状況が視覚的に分かるように、地図上にレーダ雨量を重ねて表示した。主なマ

ーカや表示情報について以下に示す。

- ・GPS マーカ

ツイートに位置情報が付与されている場合に、地図上の該当位置に GPS マーカ（バルーン型のアイコン）を表示する。投稿時刻からの経過時間に応じてマーカの表示色を変化させ、12時間前までの投稿を表示することとした。また、地図上の GPS マーカをクリックすると、地図表示画面上でツイート本文が吹き出しで表示され、地図上からツイートを直接確認することができる機能を設けた。

- ・場所推定マーカ

本文やプロフィールからツイートの投稿場所が推定された場合に、地図上の該当場所に場所推定マーカ（円形のアイコン）を表示する。場所推定は都道府県・市町村レベルで実施する。過去1時間の投稿数に応じてマーカの表示色を変化させ、投稿数が増加した際に利用者に気付きを与えるためのブザー音が鳴る機能を設けた。また、地図上の場所推定マーカをクリックすると、地図表示画面上でツイート本文が吹き出しで表示され、地図上からツイートを直接確認することができる機能を設けた。

- ・レーダ雨量

ツイート場所を示すマーカ（GPS マーカ・場所推定マーカ）とともに、レーダ雨量（国土交通省 X バンドレーダ・C バンドレーダを選択）を重ねて表示した。

(2) タイムライン表示機能

特に客観性が高い画像が添付されているツイートを視覚的に認識しやすいように、タイムライン表示画面では、ツイート本文に並べて画像アイコンを表示した。また、タイムライン表示画面で確認した位置情報付きのツイート、場所推定したツイートの場所を地図上で確認できるように、GPS アイコン、場所推定アイコンを表示した。

主観的情報については、現在の検索エンジンでは場所推定が難しいものの、利用者が読めば関係する地域の者の投稿であることがおおよそ推定できるツイートも多数存在することから、場所推定は出来ていないが、キーワードのみで抽出されたツイートも表示することとした。一方で、関心がある地域に絞ったツイートを利用者が閲覧できるように、ツイートを表示する地域を「全国（場所推定できなかったツイートを含む）」、「選択した都道府県（位置情報付き、場所推定できたツイートのみ）」、「選択した市町村（位置情報付き、場所推定できたツイートのみ）」を切り替え表示できる機能を設けた。以下に、主なアイコンの機能を示す。

- ・画像アイコン

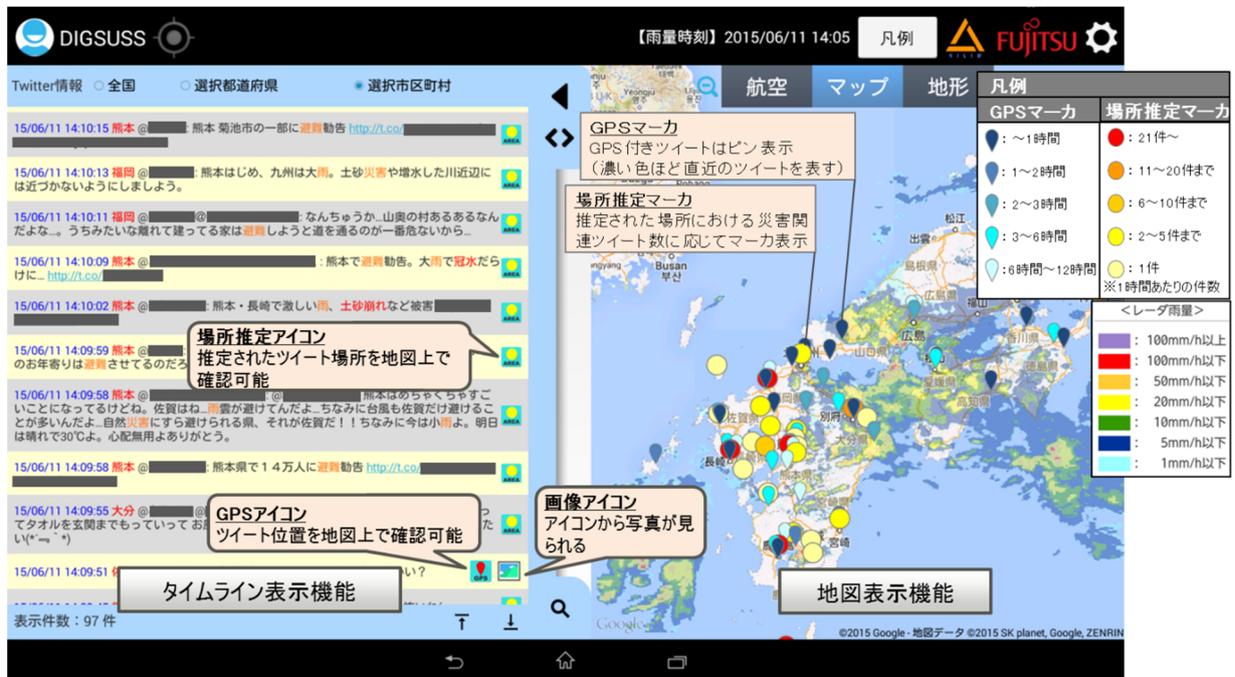
ツイートに画像が添付されている場合に表示した。アイコンをクリックすると、添付されている画像が表示される。

- ・GPS アイコン

ツイートに位置情報が付与されている場合に表示した。アイコンをクリックすると、地図上の該当位置に地図中心が移動し、GPS マーカが地図上でフォーカスされる機能を設けた。

- ・場所推定アイコン

本文やプロフィールからツイートの投稿場所が推定された場合に表示した。場所推定アイコンをクリックすると、地図上の該当位置（都道府県・市町村）に地図中心が移動し、場所推定マークが地図上でフォーカスされる機能を設けた。



※タイムライン表示画面上の「GPSアイコン」、「画像アイコン」は、機能説明のために画面を加工し追加している。

図 3-2 DIGSUSS の画面、および主な機能概要

3.3.2 収集キーワード

試作版システムでは、災害に関連するキーワードを Twitter API に対して検索することで、ツイートの収集を行っている（図 3-3）。システム上の制約によりツイートを収集可能なキーワード数が 10 個に限られるため、本試用調査では、災害発生前の状況や前兆現象を捉えること、および災害事象や災害発生後の状況をいち早く捉えることに着目し、表 3-2 に示すキーワードを設定した。収集したツイートからリツイートなどの伝聞等を排除し、ツイートの場所推定を行った。

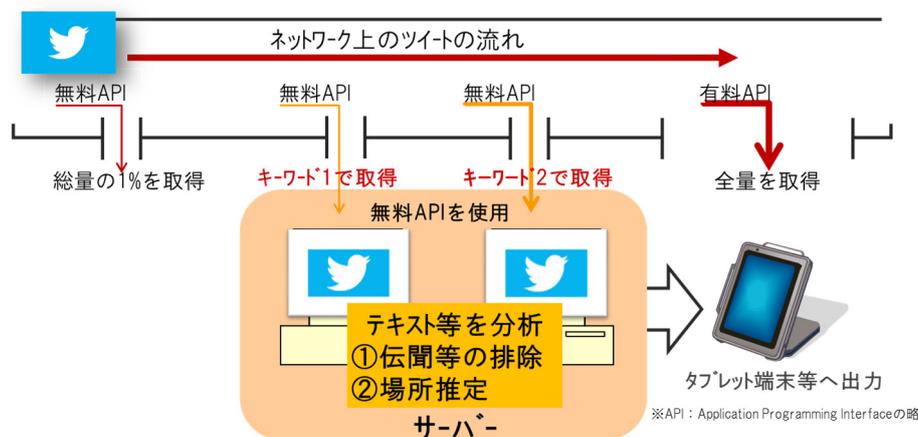


図 3-3 ツイート収集方法（イメージ）

表 3-2 収集キーワード一覧

捉えたい事象	収集キーワード(10個)	
災害発生前の状況 や前兆現象	・雷 ・雨	・地響き ・濁流
災害事象や災害発 生後の状況	・洪水 ・浸水 ・冠水	・土砂崩れ ・災害 ・避難

3.4 試作版システムの評価

試用調査は、九州地方整備局、および鹿児島県の防災担当者（フォロワー）を対象に、梅雨から台風の時期を含む平成 27 年 6 月下旬～10 月に実施した。利用者に試作版システムを閲覧可能なタブレット端末を配布し、主に防災体制に入った際にシステムを閲覧して頂き、Twitter 情報から災害関連情報を収集することの活用可能性や、試作版システムのユーザーインターフェース等について聞き取り調査を行った。

以下、聞き取り調査により得られた主な意見を記述する。

3.4.1 有効な点

(1) レーダ雨量の重ね合わせ表示

試作版システムの地図表示画面上で、レーダ雨量とツイート情報と重ね合わせて表示したことについて、「雨の多さとともにツイート情報が多くなり、状況が時々刻々とわかる」との意見があった。レーダ雨量という客観的情報とツイートによる住民が知覚した情報とを組み合わせることで、時系列的な状況変化の把握が容易になった事が考えられる。

(2) 写真付きのツイート

「写真付きのツイートは情報として確度が高く有効」との意見があった。ツイート本文だけでは必ずしも状況把握が明快では無いため、状況が客観的に示される情報として写真が添付されていることで、情報の確度や信頼度が高いと判断され、防災対応において活用されやすい事が考えられる。

(3) 迅速な情報収集・共有

期間中に発生した土砂災害において、現場からの情報伝達がなされる前に試作版システム上で関連する写真付きのツイート（準客観的な情報）が確認され、いち早く状況把握ができた事例が見られ、情報収集・共有の迅速性について効果が確認された。

その他、リアルタイムに情報を閲覧できる点や、地図上で投稿場所を閲覧できる点、ツイートのタイムラインと地図の画面を並べて表示したことが分かりやすいといった、試作版システムの有効な点に関する意見があった。

3.4.2 改善点

(1) 情報の絞り込み・要約する機能

試作版システムでは限定されたキーワードについて Twitter 情報を収集しているものの、収集されるツイート数が多いため、災害状況把握につながる重要な情報の見逃しにつながる可能性が懸念された。

また、各種情報の監視や様々な対応が求められる防災対応時に、タイムラインに表示されるツイートを読み続けることは難しいといった意見があった。

そのため、災害に関係がないツイートを可能な限り排除し、災害に関係のある人が直接感知した情報に絞り込む必要があると考える。そこで、武田ら (2013) によりが伝聞の排除方法が提案されている¹⁶⁾ように、リツイートと呼ばれる他人のツイートを再びツイートしたもののや、伝聞や報道・ニュースなどに関するツイートをなるべく排除することで、情報を絞り込むことが考えられる。なお、本試用調査に用いた試作版システムではリツイートの排除を行ったが、ツイートから得られるどのような情報が防災対応上、役に立つかを把握するために、伝聞や報道・ニュースを排除せずに広く情報を表示するかたちとした。

また、情報収集時の利用者側での柔軟な運用を支援する観点では、初めは広く情報を収集するためにツイート数が多いと思われるキーワードを設定し、時間の経過や状況に応じて利用者がキーワードを狭めるといった情報の検索・表示が可能な機能を設けることが考えられる。

また、防災担当者に気づきを与える機能として、状況を要約して表示する機能が必要であると考えられる。そこで、災害の前兆現象や災害発生後の発生状況に関するツイートを統計処理し、災害発生状況を自動的に推定して表示する機能を設けることが考えられる。

ツイートから災害発生状況を推定する手法については、武田ら (2013) により浸水・冠水を例として、災害に関するツイート数の急増を捉えて災害事象の発生を検知する手法の提案がなされており、ノイズ情報の除去技術によって伝聞と人が直接知覚した情報を分離することで、推定精度が高まる可能性があることが示唆されている¹⁶⁾。このような手法をシステムに導入することも方法として考えられる。

一方、災害の前兆現象については、前兆現象を示すツイートが統計量として意味のある変化の傾向を示すかどうか、引き続き検討すべき課題である。

その他の主な意見を表 3-3 にまとめる。

表 3-3 改善点

<ul style="list-style-type: none">・写真のサムネイル（多数の画像を一覧表示するために縮小した画像）を表示する機能（ツイートに添付された写真が有用な情報かどうかを確認するため）・任意でツイートのチェック、ソート、記録、保存ができる機能（更新される情報量が多く、既読や振り返って確認したいツイートを簡単に確認するため。また、得られた情報を整理、報告する際に必要な機能）・ツイートから、その Twitter ユーザーのサイトが閲覧できる機能（有効なツイートがあった場合、そのユーザーの過去の発言や今後の発言も閲覧できると有効な情報が得られる可能性がある）

3.4.3 DIGSUSS 利用における留意点

DIGSUSS から得られる情報の利用における留意点を以下に示す。

収集した Twitter 情報には、リツイートと呼ばれる他人のツイートを再びツイートする方法や、伝聞や報道・ニュースなどのノイズが含まれる。リツイートのテキスト解析処理や、伝聞や報道等を機械学習させることなどにより、ノイズと思われるツイートを機械的に除去することがある程度可能であるが、完全に除去することは難しい。また、災害が発生していない日常においては、災害に関連するものとして設定したキーワードであっても、災害とは関係のない発言に使われることもあるため、防災対応上は不必要な情報が含まれることがあることに留意する必要がある。

また、Twitter は文字数が 140 文字に制限されている代わりに、通信のリアルタイム性が高く、ユーザー同士が気楽・容易につながることができるといった特徴があることから、そのような情報の特性上、収集したツイートから得られる情報の事実関係について、その都度、確認を行うといった使い方にはそぐわないものとする。

3.4.4 評価結果の考察

試用調査の結果、フォロワーは客観的情報の収集を行う傾向にあることが分かった。これは、フォロワーは、多くの災害対応業務を行いながら DIGSUSS の使用に割くことができる時間が限られているためと考えられる。また、一般的にリーダーに対して報告を行う際、客観的情報に比べて主観的情報を伝えることが難しいことも影響しているのではないかと考える。

このため、フォロワーの利用を想定した場合、客観的情報の抽出精度をより一層向上するほか、客観的情報の視覚的な確認のしやすさ、また、把握した情報を整理・報告しやすくするための機能を追加することが望ましいと考える。

今回は、リーダーでの試用調査を行うことが出来なかったため、リーダーに対して簡単な聞き取り調査を行ったところ、主観的情報は非常時の意志決定において有用であるとの主旨の回答を得た。このような結果から、DIGSUSS の開発を進めるに当たっては、リーダー、フォロワーの双方に対して汎用性の高いものを目指すのではなく、リーダーには主観的・客観的情報の両方を俯瞰的に確認できる機能を有したもの、フォロワーには客観的情報を迅速に確認しリーダーに報告することが可能な機能を有したものといった、それぞれの利用者層に特化したものとする方が有効である可能性が高いことが分かった。

リーダーのニーズとしてある主観的情報のシステム上での表現方法については、今後の課題である。

4 まとめ

本研究では、土砂災害による被害軽減に向けて、ソーシャルメディア情報の警戒・避難システムにおける活用の可能性、社会実装のためのシステム開発とその効果について検討を行った。その結果、以下の事が分かった。

- ある程度人口規模が大きい地域では、ツイートによる前兆現象等の把握は、有効である。また、近隣地域の前兆現象（崩壊）が集中的な土砂災害の前に把握できる可能性がある。
- ツイートの投稿数の変化、ツイート本文から、激しく降雨が続く様子や災害の発生状況と、それらに対する住民の不安感を推察でき、災害に対する切迫した状況を捉えられる。
- 住民の通報や消防団等の現地を確認した者からの報告と比べて、個々の Twitter 情報の信頼性は劣るが、迅速性に優れた Twitter 情報は、現場からの第一報を受ける前に豪雨時の地域状況を把握することが出来る。
- Twitter 上では、全国的に自然現象等に関するツイートが発生することが期待でき、また、土砂災害に対する警戒期には、「豪雨」や「停電」、「怖い、やばい」などの特徴的な情報が現れる可能性がある。
- Twitter を介した人の反応は、人口格差や潜在的な地域差（例えば、降雨に対する地域別の人の反応の差）によって差異が生じる可能性があることが分かった。また、これらを考慮した統計的な処理を行うことで、Twitter 上の反応を定量的な情報として抽出し、地域別に災害等の発生を自動的に抽出できる可能性があることが分かった。
- 災害情報収集システム（DIGSUSS）では、降雨状況と Twitter 情報を重ねて地図上に表示することで、現地状況の理解が促され、また Twitter 情報の信頼性の確認に有効である。
- DIGSUSS の利用者に応じて、利用者の役割やニーズに特化した機能や表示方法をとる方が有効である可能性が高い。

このように、ソーシャルメディア情報を警戒・避難システムに活用できる可能性があることが分かった一方で、土砂災害の危険箇所が多く分布する山間地域では、住民の高齢化が進んでおり、本研究で災害関連情報の収集に用いる SNS を使用する年代の人口が少ないことが想定される。平成 27 年情報通信白書によると、Twitter の年代別の利用率は、60 代以上が 15.5%と言われており、20 代以下（52.8%）との差は 40 ポイント近くにも及び³⁶⁾、現時点では、若年層に比べ高齢層の利用率は非常に低い状況である。ただし、一段若い世代である 50 代の利用率を見てみると 24.3%と 10 ポイント近く上昇する。この点をとらえ、今後世代間の IT リテラシー格差は小さくなるであろうとする報告³⁷⁾もあり、スマートフォンの高齢層への浸透が進みつつある状況を勘案すると、今後、高齢層への Twitter ユーザー数は増加する可能性がある。しかし、情報発信の観点からみると、その主役はあくまでも 20 代以下の世代である³⁸⁾。豪雨に関連するツイートを発信しているユーザーの年齢層に関する調査・分析は行っていないが、見たこと、感じたことを反射的にツイートしているのは、若い世代のユーザーのように感じる。発信された情報を利用して頂く観点からは、現時点では単に高齢率が高い地域というよりは、滞留・移動人口も含めて若い世代の人口が少ない地域の情報の取得は難しいものとする。

本手法による情報収集は、ユーザー数の多寡に大きく影響を受けるばかりでなく、大雨の

発生頻度が高い九州南部地域などは小さい降雨に対するツイート数が相対的に少ないなど、地域特性によっても発信される情報の内容や多寡が変化する可能性があることが分かってきた。今後、Twitter データを定量的に評価し、統計解析処理によって地域別に災害発生を自動的に抽出するためには、このようなユーザー数や地域による潜在的な反応の差を考慮し、災害に関する一つのツイートの重みを評価する必要がある。また、災害に関する発言内容や使用される単語に地域による差異が見られる場合には、地域に応じた収集キーワードの設定等、地域特性を踏まえた災害情報収集システムのカスタマイズが必要になってくるものと考えられる。

そのためには、今後、継続的に Twitter データを収集し、より長期間のデータ、多種多様な災害事例、異なる地域における統計的な分析を通じて、災害をもたらす外力や災害の規模、地域による潜在的な反応の差を明らかにし、本研究のアプローチを検証する必要がある。また、これと並行して複数の地域で本システムを試用しながら、災害関連情報の収集に有効なキーワードや収集した情報の表現方法などを改善していくことが必要と考える。

また、これまでに述べたような課題はあるものの、近年、防災においてソーシャルセンサを活用する取り組みが進められている。高度情報通信ネットワーク社会推進戦略本部防災・減災分科会では、SNS 等民間情報の防災活用について検討が行われ、これを受けて SNS を利用した情報収集を行うための「検索用語集」が提案されている³⁷⁾。また、国立研究開発法人情報通信研究機構では、災害時に SNS 上の膨大な災害関連情報に対して、平易な質問を入力することで情報を容易に取得するシステムが開発されており³⁸⁾、平成 28 年熊本地震の際に政府の被災者生活支援チームに設置された Twitter 分析班において、被災状況の把握に活用された³⁹⁾。

国土交通省においても、情報が不足しがちな災害対応の初動時に災害の兆候や発生状況に関する推定情報を収集し、TEC-FORCE 派遣等の自治体支援などの判断に活用するための取り組みが始められている⁴⁰⁾。

今後は、ソーシャルメディア情報を実際の防災対応において活用しつつ、効果的に活用するための技術的な検討が進められることで、警戒・避難システムの高度化の一助となることが期待される。

謝辞

本研究の実施にあたり、実証実験にご協力いただいた九州地方整備局、鹿児島県はじめ関係の方々に、深く謝意を表す。

参考文献

- 1) 土砂災害への警戒の呼びかけに関する検討会：「土砂災害への警戒の呼びかけに関する検討会」報告書－土砂災害に対する実効性の高い呼びかけに向けて－，2012
- 2) 国土交通省砂防部：土砂災害警戒避難ガイドライン，2015，
<http://www.mlit.go.jp/common/001087388.pdf> (accessed March 19 2017)
- 3) 内閣府（防災担当）：避難勧告等の判断・伝達マニュアル作成ガイドライン，2014，
http://www.bousai.go.jp/oukyu/hinankankoku/guideline/pdf/140922_honbun.pdf
(accessed March 19 2017)
- 4) 加藤誠章，菊井稔宏，宮瀬将之，酒谷幸彦，西井洋史：前兆現象による土砂災害の発生危険度の評価手法について，砂防学会誌（新砂防），Vol.60，No.6，p.11-19，2008
- 5) 平成24年度版 情報通信白書：ソーシャルメディア利用者数の推移（Facebook，Twitterの例），
<http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h24/html/nc123220.html>
(accessed March 19 2017)
- 6) The Huffington Post：Twitterが国内ユーザー数を初公表「増加率は世界一」
http://www.huffingtonpost.jp/2016/02/18/twitter-japan_n_9260630.html (accessed March 19 2017)
- 7) 風間一洋：Twitterにおける情報伝播，人工知能学会誌27巻1号，p.35-42，2012
- 8) 谷口慎一郎：災害時におけるTwitterの有効性について-2011年9月の台風12号による豪雨災害を例に，災害情報No.10，p.56-67，2012
- 9) 榊剛史：ソーシャルセンサとしてのTwitter-ソーシャルセンサは物理センサを凌駕するか？-，人工知能学会誌27巻1号，p.67-74，2012
- 10) 阿部博史：震災ビッグデータ，NHK出版，2014
- 11) 國友優：土木用語解説「ソーシャルセンサ」，土木技術資料，Vol.57，No.9，pp.54，2015
- 12) 防災・減災におけるSNS等の民間情報の活用に関する検討会：報告書，2014，
http://www.kantei.go.jp/jp/singi/it2/senmon_bunka/bousai/dai6/houkokusyo.pdf
(accessed March 19 2017)
- 13) 奥村学：マイクロブログマイニングの現在，第3回集合知シンポジウム，2012
- 14) J. Kleinberg: Bursty and hierarchical structure in streams, In Proc. of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1-25, 2002
- 15) 高橋哲朗：実世界のセンサとしてのTwitterの可能性，電子情報通信学会技術研究報告，NLC，110(400)，pp.43-48，2011

- 16) 武田邦敬, 瀧口茂隆, 高橋哲郎, 山影讓, 渡部勇: 豪雨時の Twitter データを活用した災害事象の検知, 平成 25 年度砂防学会研究発表会概要集 B, pp. 218-219, 2013
- 17) 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
- 18) 阿蘇地域土砂災害対策検討委員会 報告書, 2013,
http://www.pref.kumamoto.jp/common/UploadFileOutput.ashx?c_id=3&id=6973&sub_id=1&flid=2&dan_id=1 (accessed 20 June 2016)
- 19) 酒井敦章, 道畑亮一, 菊井稔宏: 平成 24 年 7 月九州北部豪雨による土砂災害発生時の住民の行動実態, 砂防学会誌 (新砂防), vol. 66, No. 2, p. 57-63, 2013
- 20) 内閣府: 平成 26 年 8 月 19 日からの大雨による被害状況等について (11 月 6 日 12:00 現在), 2015, <http://www.bousai.go.jp/updates/h260819ooame/pdf/h260819ooame35.pdf> (accessed 20 June 2016)
- 21) 8.20 豪雨災害における避難対策等検証部会: 平成 26 年 8 月 20 日の豪雨災害避難対策等に係る検証結果, 2015,
<http://www.city.hiroshima.lg.jp/www/contents/1476873330360/files/01honpen.pdf>, (accessed 20 June 2016)
- 22) 武田邦敬, 山影讓, 國友優, 神山嬢子, 松下一樹: 土砂災害発生前の Twitter データの俯瞰的分析, 平成 27 年度砂防学会研究発表会概要集 A, pp. 248-249, 2015
- 23) 丹後俊郎, 横山徹爾, 高橋邦彦: “空間疫学への招待”, 朝倉書店, 2007
- 24) 瀬谷創, 堤盛人: “空間統計学”, 朝倉書店, 2014
- 25) 丹後俊郎, Taeko Becque: “ベイジアン統計解析の実際”, 朝倉書店, 2011
- 26) A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, D. B. Rubin: “Bayesian Data Analysis Third Edition”, CRC Press, 2014
- 27) S. N. Wood: “Generalized Additive Models”, CRC Press, 2006
- 28) 井出剛, 杉山将: “異常検知と変化検知”, 講談社, 2015
- 29) C. C. Aggarwal: “Outlier Analysis”, Springer, 2013
- 30) J. Han, M. Kamber, J. Pei: “Data Mining: Concepts and Techniques, Third Edition”, Morgan Kaufmann, 2011
- 31) 内閣府「総合的な土砂災害対策検討ワーキンググループ」: 総合的な土砂災害対策の推進について(報告), 2015, <http://www.bousai.go.jp/fusuigai/dosyaworking/index.html> (accessed 20 June 2016)
- 32) 消防庁「突発的局地的豪雨による土砂災害時における防災情報の伝達のあり方に関する検討会」: 突発的局地的豪雨による土砂災害時における防災情報の伝達のあり方に関する

- る検討会報告書, 2015,
http://www.fdma.go.jp/neuter/about/shingi_kento/h26/saigai_dentatsu/index.html
(accessed 20 June 2016)
- 33) 茂野香おる：基礎看護技術〈1〉—基礎看護学〈2〉(系統看護学講座 専門分野), 医学書院
- 34) 日本医師会：患者に関する情報(前編), DOCTAR-ASE,
http://www.med.or.jp/doctor-ase/vol14/14page_ID03main4.html (accessed 20 June 2016)
- 35) 神山嬢子, 森田直志, 水野正樹, 蒲原潤一：スマートフォンによる効果的な土砂災害関連情報の収集・提供手法の開発, 土木技術資料, Vol. 56 No. 10, pp. 12-15, 2016
- 36) 総務省：情報通信白書, 2015
<http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h27/html/nc242220.html>
(accessed March 19 2017)
- 37) 内閣官房情報通信技術(IT)総合戦略室：SNSを活用した情報収集 検索用語集—防災・減災—, 2015,
http://www.kantei.go.jp/jp/singi/it2/senmon_bunka/pdf/kensakuyougosyu.pdf
(accessed 20 June 2016)
- 38) 国立研究開発法人情報通信研究機構：対災害 SNS 情報分析システム「DISAANA (ディサーナ) を Web 上に試験公開, 2014, <http://www.nict.go.jp/press/2014/11/05-1.html>
(accessed 20 June 2016)
- 39) 総務省：大規模災害時の非常用通信手段の在り方に関する研究会 第5回, 2015,
http://www.soumu.go.jp/main_content/000422662.pdf (accessed 20 June 2016)
- 40) 国土交通省：国土交通省重点政策 2015 「2. 国民の安全・安心の確保」参考資料, 2015
<http://www.mlit.go.jp/common/001101578.pdf> (accessed 20 June 2016)

卷末資料

1. DIGSUSS 仕様書 (案)

DIGSUSS 仕様書 (案)

1 タブレットアプリ機能仕様

(基本仕様) ※3.3 節において構築した DIGSUSS 仕様

1.1 地図表示機能

1.1.1 地図表示

- 画面を2ペインで表示し、右側の領域に地図を表示する
- 地図を拡大・縮小して表示できる
- 地図・地形図・航空写真を切り替えて表示する
- 凡例ボタンをタップしたときに地図領域に凡例(レーダ雨量)を表示する

1.1.2 GPS マーカ

- 緯度経度情報が付与されているツイートに関して地図の該当位置にマーカ表示する
- GPS マーカはツイートされてからの経過時間に応じてマーカ色を変更して表示する(5段階)
- GPS マーカの形状はバルーン型とする
- GPS マーカをクリックしたときにツイート内容を表示する

1.1.3 場所推定マーカ

- ツイート本文等から場所推定された該当場所を地図にマーカ表示する(5段階)
- 場所推定マーカは各市町の代表地点(市町村庁舎の位置)に表示する
- 場所推定マーカは過去1時間の同場所におけるツイート件数に応じてマーカ色を変更して表示する(5段階)
- 場所推定マーカの形状は丸形とする
- 場所推定マーカをクリックしたときに、同場所に推定されたツイート内容を表示する

1.1.4 レーダ雨量

- 国土交通省 C バンドレーダ雨量/X バンドレーダ雨量を地図に重ねて表示する
- 設定により C バンドレーダ雨量/X バンドレーダ雨量の表示切り替えが可能
- 地図の拡大縮小に合わせてレーダ雨量画像も拡大・縮小表示する

1.2 タイムライン表示機能

1.2.1 タイムライン表示

- 画面を2ペインで表示し、左側の領域にツイートをタイムライン表示する
- ツイートの日付・名前・本文・推定場所を表示する
- ツイートの収集キーワード(検索キーワード)を強調色で表示する
- 収集したツイートを上から下へ最新データを順に挿入して表示する
- 新規ツイートを取得した場合は、上からデータを追加する
- スクロールによりタイムラインエリアの下端に達すると過去データを取得し、タイムラインの下に追加して時間順に表示する

1.2.2 画像アイコン

- 画像アイコンをクリックしたときに画像を表示する

1.2.3 GPS アイコン

- 該当する GPS マーカをクリックしたときに MAP 中央に移動する

1.2.4 場所推定アイコン

- 該当する推定マーカをクリックしたときに MAP 中央に移動する

1.2.5 表示フィルタ

- 選択市町村・選択都道府県の選択で、設定画面で選択した県・市町村のツイートデータのみフィルタリングして表示する

1.2.6 検索機能

- 検索期間を指定して検索できる
- 検索条件として、GPS・推定・なしの複数選択が可能
- 検索条件として、画像の「有り」・「無し」・「両方」を選択が可能
- 入力できるフィールドを備えること。(起動時は空白)
- ユーザー名・本文を対象として検索が可能
- 検索結果はタイムライン時と同様に表示が可能

1.3 設定機能

1.3.1 監視エリア設定

- 都道府県別の市町村選択可能
※2015 年度実証実験時は「九州」地方のみ設定

1.3.2 雨量エリア設定

- 国土交通省 C バンドレーダ雨量/X バンドレーダ雨量を設定可能

2 収集サーバ機能仕様

2.1 ツイート収集機能

2.1.1 ツイート収集

- 設定された検索キーワードを用いて、TwitterAPI からツイートを収集する
- 検索キーワードは 10 個のキーワードを設定可能
- 検索キーワードにアンド条件(最大3ワード)を設定可能
- 除外文字を設定可能

2.1.2 場所推定

- ツイート本文中の地名からツイート内容/発言場所の推定を行う
- 都道府県・市町村のレベルで場所推定を行う

2.2 レーダ雨量収集解析機能

- 国土交通省 C バンドレーダ雨量/X バンドレーダ雨量データを受信する
- レーダ雨量データを解析しタブレットアプリで表示できる形式に変換する

2.3 アプリ連携機能

- 収集したツイート情報・レーダ雨量データをタブレットアプリに配信する

3 追加検討仕様

- 写真のサムネイル表示機能
- 期間を指定してツイートデータをダウンロードする機能
- ツイートから Twitter ユーザーのサイトが閲覧できる機能
- ツイート数の変化から災害発生状況を推定して表示する機能(発災推定)

国土技術政策総合研究所資料

TECHNICAL NOTE of N I L I M

No. 1013 February 2018

編集・発行 ©国土技術政策総合研究所

本資料の転載・複写の問い合わせは

〒305-0804 茨城県つくば市旭1番地

企画部研究評価・推進課 TEL 029-864-2675