

## 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～

株式会社富士通研究所 ○武田邦敬 古川忠延 瀧口茂隆<sup>※1</sup> 山影譲  
日本工営株式会社 荒木健 伊藤顕子 宮川健  
国土交通省 国土技術政策総合研究所 蒲原潤一

### 1 はじめに

土砂災害が発生した可能性がある箇所に関する情報を明示的に提示することで、発生箇所および周辺地域の避難情報として活用することが検討されている<sup>1)</sup>。著者らは、災害を伝えるソーシャルメディア情報が災害情報として利活用できるとの仮説を立て、Web 上のマイクロブログサービスの一つである Twitter\*のデータに対して統計的な手法を適用し、浸水・冠水事象を対象として災害の発生を推定する研究を行った<sup>2)</sup>。この研究において、Twitter データを用いて災害の発生時刻と場所を推定できる可能性を示した。しかし、災害の発生位置の推定は都道府県単位であり、災害の発生情報を避難等に活用するためには、より詳細に災害の発生位置を特定することが課題となる。また、この手法を土砂災害へ適用することが可能であるか、検証が必要である。

本研究では、土砂災害を対象に、Twitter データを活用して、市町村単位で災害の発生を推定するための手法を提案し、評価を行った。以下では、2012 年 7 月に発生した九州北部豪雨災害（以下、H24 年九州北部豪雨災害とする）における熊本県内で発生した土砂災害に対して、提案手法を適用した結果を示す。

### 2 Twitter データの予備調査

Twitter 社が提供する検索用 API を用いることで、特定のキーワードを含むツイート<sup>†</sup>をほぼリアルタイムに収集することが可能である。著者らは、災害に関連する Twitter データの収集を行っている。予備調査のために、土砂災害に関連のあるキーワード「土砂崩れ」「土砂くずれ」「山崩れ」「山くずれ」「土石流」「地滑り」「地すべり」「崖崩れ」「崖くずれ」「がけ崩れ」を含む Twitter データの内、H24 年九州北部豪雨災害の発災日前後のツイート数の推移を調査した。

調査の結果、発言者の居住地が熊本県であると推定したデータにおいて、7 月 12 日の朝 7 時前後で、ツイート数が急増していることがわかった。居住地の推定は、発言者のプロフィール情報を基に推定する方法をとった(3.2 節参照)。しかし、熊本県内での主な土砂災害は、同日の朝 5 時から 6 時頃に集中して発生しており<sup>§</sup>、これと比較してツイート数の増加がやや遅れる結果となった。このため、土砂災害の発生を早期に推定するためには、プロフィールを用いて居住地が推定できなかったツイートに対しても居住地情報を付与し、ベースとなるツイート数を増やすことが課題となる。

また、7 月 12 日の熊本県内の土砂災害関連キーワードを含むツイートの内、GPS 情報が付与されていたのは、2 件のみであった。このため、災害が発生した市町村を推定するためには、GPS 以外の情報を用いて発言者の位置を推定することが課題となる。

以下では、これらの課題を解決するための手法を提案し、H24 年九州北部豪雨災害事例に対して評価した結果を示す。

### 3 Twitter データを活用した災害事象の推定手法

#### 3.1 アプローチおよび全体の流れ

本研究は、Twitter データを用いて災害の発生日時と市町村単位で場所を推定することを目的とする。この目的の達成のため、特定の災害事象のキーワードを含む Twitter の発言数の急激な増加を検出して都道府県単位で災害の発生を推定した後、Twitter データを活用して市町村を推定するアプローチを用いた。

土砂災害関連キーワードを含む Twitter データを取得した後のプロセスを以下に示す。

- (1) Twitter データから伝聞等のノイズを排除する
- (2) 発言者の居住地を推定し、Twitter データを都道府県別に仕分けする
- (3) 発言数の急増を捉えて都道府県単位で災害の発生を推定する
- (4) 災害推定時刻以降の Twitter データを用いて、市町村を推定する

上記プロセスの(1)、(3)については、文献 2)に示す方法を用いた。一方、(2)、(4)については、プロフィールを用いる方法<sup>‡</sup>だけでは十分な情報が得られなかったため、本研究では新たな手法を適用した。

#### 3.2 発言者の場所の推定手法

本研究では、災害の発生場所を推定するために、ツイート発言者の位置を都道府県および市町村単位で推定するための方法として、下記 4 つの手法を適用した。

**GPS** ツイートに含まれる GPS 情報を住所情報へ変換し、発言者の位置を求める。

**プロフィール** 発言者のプロフィールの住所情報から、発言者の居住地を推定する。

**ツイート本文** ツイートに含まれるランドマーク情報から、発言者の位置を推定する。

**過去発言** 発言者の過去の発言に含まれるランドマーク情報を集計し、統計的な手法を用いて発言者の居住地を推定する<sup>‡</sup>。

土砂災害関連キーワードを含む、2012 年 7 月 4 日から 2012 年 7 月 31 日までの Twitter データ 14,997

※1 現：国土交通省

\* <https://twitter.com/>

† Twitter 上に投稿された発言をツイートと呼ぶ。

‡ ソーシャルメディアの発言内容から、発言者の居住地を推定できる可能性があるという仮定<sup>§</sup>に基づく。

件に対して上記手法を適用した。この調査において、都道府県および市町村を推定できたツイート数を表 1 に示す。表 1 より、複数の推定手法を組み合わせることで、居住地が推定できるツイート数が増加することがわかった。ただし、発言者の真の居住地は不明であるため、推定結果の精度を測ることができない。そのため、本研究では、本手法によって出力された災害発生市町村と実際に災害が発生した市町村を比較することによって、評価を行った。

表 1 居住地を推定できたツイート数および割合

| 推定手法                      | 都道府県推定         | 市町村推定         |
|---------------------------|----------------|---------------|
| GPS                       | 47 (0.3%)      | 47 (0.3%)     |
| GPS, プロフィール               | 9,182 (61.2%)  | 3,158 (21.1%) |
| GPS, プロフィール, ツイート本文, 過去発言 | 12,845 (85.7%) | 7,228 (48.2%) |

※( )内は、全件 14,997 件に対する割合を示す

#### 4 評価

前節で提案した手法を、H24 年九州北部豪雨災害における熊本県内で発生した土砂災害に適用し、災害の発生時刻と場所の評価を行った。評価実験では、ツイートを 1 時間毎に集計した結果を用い、1 時間間隔で災害の発生時刻と場所を推定し、出力した。2012 年 7 月 12 日前後の Twitter データに対して、本手法を適用した結果を示す。

##### 4.1 都道府県単位での災害発生の推定

評価実験の結果、熊本県において、2012 年 7 月 12 日の朝 6 時の時点で、土砂災害の発生を推定することができた。同日のツイート件数の推移を図 1 に示す。

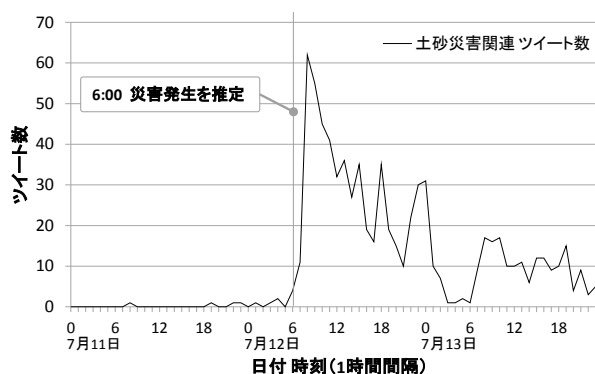


図 1 H24 年九州北部豪雨災害前後のツイート数の推移

都道府県単位で災害を推定した 2012 年 7 月 12 日の朝 6 時集計以降のツイート 545 件に対して、本手法の(4)を適用した結果を次節に示す。

##### 4.2 市町村単位での災害発生の推定

同日熊本県内では、11 市町村で実際に土砂災害が発生しており、これを正解として評価した<sup>§</sup>。評価実験の結果、正解事例 11 件の内、8 件を出力することができた(再現率 0.73)。一方、実験で出力した 19 市町村の内、正解は 8 件であった(精度 0.42)。表 2 に、本手法により出力した市町村、市町村毎のツイ

ト件数、初めに出力した時刻、実際の災害有無および災害発生時刻を示す。表 2 より、ツイート数が多い市町村では初出時刻が比較的早くなる傾向にあることがわかった。一方、必ずしもツイート数が少ない市町村が不正解となるとはいえないため、出力候補から正解と不正解を分離する方法を検討する必要がある。

表 2 出力した市町村および発災有無(ツイート数順)

| 市町村  | ツイートからの推定結果 |       | 実際の災害発生時刻 <sup>§</sup> |
|------|-------------|-------|------------------------|
|      | ツイート数       | 初出時刻  |                        |
| 熊本市  | 183         | 6:00  | 10:00 頃発生              |
| 阿蘇市  | 119         | 6:00  | 5:00 頃発生               |
| 菊池市  | 27          | 7:00  | 5:30 頃発生               |
| 山鹿市  | 11          | 7:00  | -                      |
| 玉名市  | 9           | 7:00  | -                      |
| 人吉市  | 6           | 11:00 | -                      |
| 南阿蘇村 | 6           | 8:00  | 5:00 頃発生               |
| 合志市  | 6           | 9:00  | -                      |
| 八代市  | 4           | 14:00 | 9:30 頃発生               |
| 水俣市  | 3           | 12:00 | -                      |
| 西原村  | 2           | 8:00  | -                      |
| 球磨村  | 2           | 20:00 | 9:00 頃発生               |
| 荒尾市  | 1           | 6:00  | -                      |
| 山都町  | 1           | 21:00 | -                      |
| 宇城市  | 1           | 22:00 | -                      |
| 上天草市 | 1           | 9:00  | -                      |
| 天草市  | 1           | 17:00 | 9:00 頃発生               |
| 南小国町 | 1           | 10:00 | -                      |
| 益城町  | 1           | 17:00 | 5:00 頃発生               |

※災害発生時刻“-”は発災記録がないことを示す

#### 5 まとめ

本研究では、土砂災害を例として、Twitter データを用いて災害事象の発生を市町村単位で推定する手法を提案した。本手法により、H24 年九州北部豪雨災害事例での評価において、市町村単位で災害発生地域の一部を推定することができた。ただし、推定市町村の精度や網羅性において課題がある。市町村単位での推定精度・網羅性を向上すること、および、より詳細に災害発生地域を推定することが今後の課題である。

##### 参考文献

- 1) 国土交通省：「土砂災害への警戒の呼びかけに関する検討会」報告書，2013
- 2) 武田邦敬，瀧口茂隆，高橋哲郎，山影譲，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
- 3) 高橋哲郎：実世界のセンサーとしての 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
- 4) 安田宜仁，平尾努，鈴木潤，磯崎秀樹：ブログ作者の居住地の推定，言語処理学会年次大会発表論文集，vol.12，pp.512-515，2006

<sup>§</sup> 国土交通省から提供していただいた土砂災害に関するデータを用いて求めた。