

## Instagram を利用した地理情報源ツイートの発見手法 A Method for Finding Tweets as Geographic Information Sources Utilizing Instagram

川人 晋伍<sup>†</sup> 寺垣 智也<sup>‡</sup> 梅野 睦月<sup>††</sup> 木村 文則<sup>‡</sup>  
Shingo Kawahito Tomoya Teragaki Mutsuki Umeno Fuminori Kimura

### 1. はじめに

Twitter などのソーシャルメディアから地理情報と関連した情報を抽出し分析をする研究の多くは、投稿時の位置情報などを付与した「ジオタグ」を用いている。しかし、ジオタグが付与されたツイートの割合は高くなく、ジオタグ付きのツイートの収集は容易ではない。本研究では、Instagram を利用することにより、ジオタグの付与されていないツイートから、地理情報源ツイートを発見する手法を提案する。Instagram は写真を投稿するため、実際に被写体が存在する場所で撮影している可能性が高い。それと連動しているツイートも、その場所と関わる可能性が高い。本論文ではそのようなツイートを収集し、地理情報源として活用可能であるかどうかを検証した。

### 2. 関連研究

谷ら[1]はジオタグ付きツイートを集め、投稿時間を考慮しユーザが移動した移動経路を抽出した。谷らが指摘している問題点として、ジオタグ付きツイートでないと移動経路を再現することができないとある。本手法では、Instagram との連携があればジオタグが付与されていないツイートでも地理情報源として扱うことができるという点で異なる。

Lee らは[2]はジオタグを付与すべきかどうかを判定するための言語モデルの学習を行う際、学習データとして利用できるツイートを選別するために、Foursquare を外部情報として活用しフィルタリングを行った。本手法とは、ジオタグのないツイートにジオタグを付与する試みを行った点で共通しているが、我々は Instagram を利用している点で異なる。Foursquare はユーザ数が減少傾向にあることを Lee らは危惧していた。しかし、Instagram は現在もユーザ数が増加傾向にあり、本手法ではその問題点を解消できる。

### 3. 提案手法

本手法では、収集したツイートのうち、Instagram と連携しているツイートを、地理情報源ツイートであると判定する。Instagram と連携しているツイートとは、ツイートの本文中にそのユーザが投稿した Instagram へのリンクを有しているものことである。Instagram は投稿の際、必ず写真を付与するため、ユーザ本人がその場に行き投稿している可能性が高い。よって、Instagram の投稿は地理情報源を含んでいる可能性が高いと言える。それと連携しているツイートもまた地理情報源を含んでいる可能性が高いと言える。

また、地理情報源ツイートの収集精度を高めるため、2 つの方法を行った。

1 つ目は、Instagram の位置情報を付与する機能（以後、「位置情報タグ」）を利用したものである。位置情報タグは、ユーザの現在地から Instagram が近辺の地名や建物の候補をいくつか表示し、ユーザはその候補の中から該当するものを選ぶ機能である。位置情報タグが付与されていると

いうことは、ジオタグが付与されているとみなすことができる。

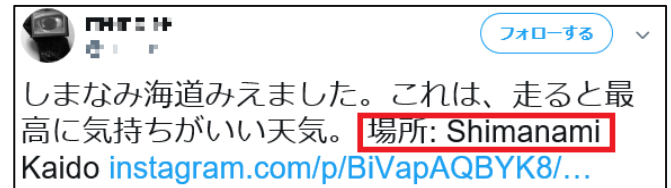


図 1 Twitter での位置情報タグの例

Twitter 上では図 1 のように「場所:○○」と表示され、Instagram 上ではユーザ名の下にその場所が表示される。

2 つ目は、ツイートの本文に出現する単語を利用する方法である。実際にその土地を行った人のツイートでは「来ました」や「一人旅してきた」など本文に過去形が見られる。一方、その土地に行っていない人のツイートでは「行きたいな」や「羨ましいな」など本文にユーザの願望が見られる。これらのように、本文を見ることで地理情報源ツイートとして活用可能かどうかを判断することができる。

### 4. 評価実験

#### 4.1 実験

提案手法の有用性を検証するため、地理的キーワードを含むツイートを収集し、提案手法で地理的情報源であると判定した際の正解率を調査した。

ブラウザ版 Twitter の検索機能を用いて「instagram.com/」および「地理的キーワード」を含むツイートを収集した。本論文で用いた地理的キーワードは、「しまなみ海道」、「太宰府天満宮」、「白川郷」、「千光寺公園」、「朱華園（広島県尾道市で有名な尾道ラーメンの店）」の 5 つである。Instagram の URL とキーワードを組み合わせ検索をすることで Instagram と連携したキーワードの場所に関するツイート、つまりキーワードの地理情報源を含む可能性のあるツイートを収集することができる。収集期間は、2018 年 4 月から 6 月の約 2 カ月間で、それぞれのキーワードに関して 100 ツイート集めた。

収集したツイートを「場所の特定可能」、「条件付きで特定可能」、「場所の特定不可能」の 3 つに分類した。「場所の特定可能」ツイートが図 2、その引用先 Instagram が図 3、「条件付きで特定可能」ツイートが図 4、その引用先 Instagram が図 5 である。

<sup>†</sup> 尾道市立大学大学院経済情報学科,

<sup>‡</sup> 尾道市立大学経済情報学部

<sup>††</sup> 株式会社エースシステムズ



図2 「場所の特定可能」ツイート



図3 「場所の特定可能」ツイート引用先 Instagram

Instagram の写真がキーワードの場所であり、文章の内容からも投稿者自身がその土地を訪れていると分かるものを「場所の特定可能」ツイートとした。

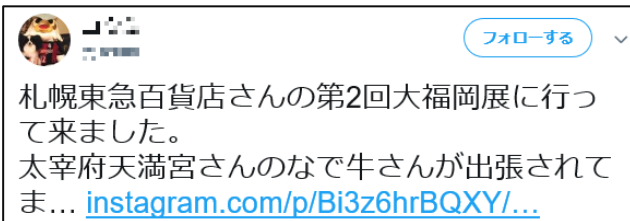


図4 「条件付きで特定可能」ツイート



図5 「条件付きで特定可能」ツイート引用先 Instagram

地理的キーワードと関連しない土地を訪れていた場合や、画像では判断できないが文章によって実際に訪れたと判断できた場合は「条件付きで特定可能」とする。

Instagram の写真では場所が分からないものであり、なおかつ文章の内容からも場所が特定できないものを「場所の特定不可能」ツイートとした。

#### 4.2 評価

それぞれのキーワードの「場所の特定可能」、「条件付きで特定可能」、「場所の特定不可能」ツイートの割合を表1に示す。

表1 場所の特定が可能なツイート割合一覧

	場所の特定可能	条件付きで特定可能	場所の特定不可能
しまなみ海道	87%	5%	8%
白川郷	91%	6%	3%
太宰府天満宮	80%	17%	3%
千光寺公園	95%	5%	0%
朱華園	92%	8%	0%
平均	89%	8.2%	2.8%

場所の特定が可能であれば、地理情報源として活用可能である。よって、「しまなみ海道」の場合「場所の特定可能」と「条件付きで特定可能」を足した93%のツイートが地理情報源ツイートであった。

同様にして、地理情報源ツイートの割合は「白川郷」と「太宰府天満宮」では97%であり、「千光寺公園」と「朱華園」では100%であった。ただし、太宰府天満宮に関しては「条件付きで特定可能」ツイートの割合が高い。これは、収集期間に「太宰府天満宮」にてイベントが行われており、Instagram と連携した、イベントに関するツイートが増加したためである。この場合は、イベントが主体であり、太宰府天満宮そのものが目的となっていないため、「条件付きで特定可能」ツイートに分類した。

#### 5. おわりに

本論文では、Instagram と連携しているツイートが地理情報源ツイートとして活用可能であるかを検証した。地理情報源ツイートの収集精度を高めるために、位置情報タグを用いる方法とツイートの本文に出現する単語を利用する2つの方法も提案した。

評価実験の結果から、Instagram と連携しているツイートの89%が地理情報源ツイートとして活用可能であることが分かった。さらに「条件付きで特定可能」ツイートも含めると97%が地理情報源ツイートとみなせることも分かった。

本手法を用いることで、ジオタグに代わる新たな地理情報源ツイートを収集することができ、ジオタグ関連の研究におけるデータ不足を補う一つの手段になると言える。

#### 参考文献

- [1] 谷直樹, 風間一洋, 榊剛史, 吉田光男, ジオタグ付きツイートを用いた交通路の抽出, DEIM Forum 2015, 2015.
- [2] Kisung Lee, Raghu K. Ganti, Mudhakar Srivatsa, Ling Liu, When twitter meets foursquare: tweet location prediction using foursquare, Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, pp. 198-207, 2014.