第Ⅳ部門　　災害時のソーシャルメディアと帰宅行動の関連性分析

神戸大学工学部　学生員　○榎本　甫
神戸大学大学院工学研究科助教　正会員　桑野　将司
神戸大学大学院工学研究科教授　正会員　小池　淳司

## 1　はじめに

近年，ソーシャルネットワーキングサービス（以降，SNS）の普及により，各個人は不特定多数の人との情報のやり取りが可能となった．東日本大震災時，各個人が発信源となるようなSNSを利用した情報（以降，ソーシャルメディア情報）が多く行き交った[1)]．ソーシャルメディア情報は，誰でも容易に情報を発信出来ることから，従来のメディアに比べ情報伝播力が強いという有用性を持っていることから，震災時の新たな情報交換ツールとして注目を集めた．

しかし一方で，情報発信源が不特定多数であり，誰でも手軽に情報を発信することが可能であることは，不確実性を孕んだ情報の拡散を誘発する要因にもなり得る．すなわち，誤った情報が拡散されやすく，情報の信頼性が必ずしも担保されるとは限らない．このようなソーシャルメディア情報が，災害時，人々の行動に及ぼす影響は明らかではない．ソーシャルメディア情報による人々への影響を定量的に把握することによって，今後，同様の災害が起こった場合，ソーシャルメディア情報を適切に扱うことで，人々の混乱を避けるような誘導策の検討が可能となることが期待される．

そこで，本研究では，東日本大震災時のソーシャルメディアにおけるテキスト情報を対象として，テキストマイニングを行い，テキスト情報の内容の時間的変化を把握する．さらに，東日本大震災時のソーシャルメディアのテキスト情報と都心部での人々の帰宅行動との関係を明らかにする．

## 2　使用データの概要

### 2.1　ソーシャルメディア情報に関するデータ

本研究では，東日本大震災時に発信された Twitter のテキストデータを用いる．詳細は以下のとおりである．

- 期間：2011 年 3 月 11 日 14 時から 2011 年 3 月 12 日 23 時
- 対象：「駒沢通り」「目黒通り」「六本木通り」をキーワードとする Tweet を 1 回以上発信したユーザーの期間中の全 Tweet

使用する Twitter テキストデータについて，1 時間当たりの Tweet 数の推移を図 1 に示す．地震発生の前後 5 時間を比較すると Tweet 数は約 6.7 倍に増加していた．このことから，震災時にソーシャルメディアが活発に利用されたことがわかる．

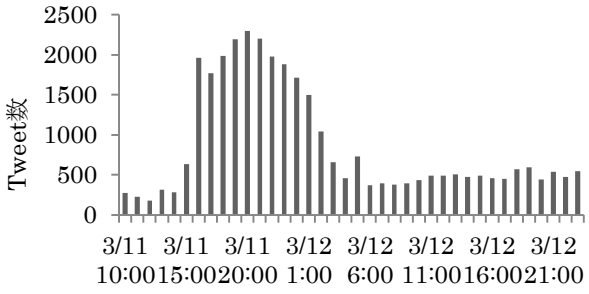


図 1　1 時間当たりの Tweet 数の推移

### 2.2　災害時の帰宅行動に関するデータ

本研究では，帰宅行動に関するデータとして東日本大震災時の人口流動メッシュデータを用いる．具体的には，2011 年 3 月 11 日，12 日の東横目黒沿線における流動人口，自宅該当者数，勤務地該当者数のメッシュデータを用いる．帰宅行動を捉えるために，震災発生の 1 週間前の 2011 年 3 月 4 日，5 日に同沿線を 1 回以上利用した人を分析対象とした．

## 3　テキストマイニングによる Tweet 内容分析

テキストマイニングとは，テキスト情報を定量的に分析する手法である．本研究では，形態素分析，共起分析を行い，1 時間ごとの単語間の関係をネットワーク図（以降，単語間ネットワーク）として可視化したうえで，その時間的推移を把握する．

3.1 節で述べた Twitter のテキストデータを対象として形態素分析を行い，動詞，助動詞，名詞，形容詞を合わせて 2392 語抽出した．抽出された単語を用いて共起分析を行い，1 時間ごとの単語間ネットワークを作成した．しかしながら，単語数が膨大なため，この状態で単語間ネットワークの時間的推移を捉えることは困難である．そこで，ネットワーク類似性を用いて，単語間ネットワークの概観を試みた．ネットワーク類似性とは，2 つのネットワークの類似度を，相関係数を用いて評価する指標である．3 月 11 日の 15 時から 24 時までの 1 時間ごとの単語間ネットワーク間の相関係数を表 1 に示す．

Hajime ENOMOTO, Masashi KUWANO, Atsushi KOIKE,　0914208t@stu.kobe-u.ac.jp

表 1　相関行列

| | 15 時 | 16 時 | 17 時 | 18 時 | 19 時 | 20 時 | 21 時 | 22 時 | 23 時 |
|---|---|---|---|---|---|---|---|---|---|
| 15 時 | - | 0.596 | 0.416 | 0.304 | 0.217 | 0.213 | 0.219 | 0.179 | 0.272 |
| 16 時 | | - | 0.617 | 0.562 | 0.479 | 0.440 | 0.429 | 0.375 | 0.464 |
| 17 時 | | | - | 0.725 | 0.653 | 0.574 | 0.572 | 0.466 | 0.538 |
| 18 時 | | | | - | 0.760 | 0.689 | 0.657 | 0.589 | 0.609 |
| 19 時 | | | | | - | 0.733 | 0.658 | 0.596 | 0.641 |
| 20 時 | | | | | | - | 0.794 | 0.724 | 0.718 |
| 21 時 | | | | | | | - | 0.717 | 0.699 |
| 22 時 | | | | | | | | - | 0.709 |
| 23 時 | | | | | | | | | - |

表 1 の 15～16 時台に着目すると，その直後の相関係数は高いものの，時間の経過とともに相関係数が小さくなっていることがわかる．一方，17 時以降に着目すると，直後は非常に相関が高く，また，その後も相関が高い状態が続いている．このことから，地震発生直後は話題が大きく変化したものの，地震発生から 3 時間程度経過するとそれ以降は話題が安定していたことが明らかとなった．

## 4　テキスト情報と帰宅行動の関連性分析

本研究では，テキスト情報と帰宅行動の関連性を分析するために正準相関分析を用いた．正準相関分析とは，2 つの変数群の間の相関関係を分析する手法である [2)]．各変数群を線形結合して得られた潜在変数（以降，正準変量）間の相関が最大になるように，係数を推定する．変数群の設定に際して，以下の 3 つの場合を想定した．

- 同時点でのテキスト情報と帰宅行動に関係がある場合(Case1)
- テキスト情報とその 1 時間後の帰宅行動に関係がある場合(Case2)
- 帰宅行動とその 1 時間後のテキスト情報に関係がある場合(Case3)

各 Case について，AIC を用いたモデル適合度の比較を行ったところ Case2 のモデル適合度が最も高かった．そこで，Case2 についての正準相関分析の結果のみを表 2 に示す．表 2 より，第 2 正準相関係数までの累積寄与率が 90%を超えたため，第 2 正準相関係数まででこの相関関係は十分説明できるとし，以降，第 1 正準変量と第 2 正準変量についての考察を示す．

表 3 の第 1 正準変量に着目すると，「運転-再開」「電車-動く」といった交通に関して楽観的な情報が少ないとき，あるいは，「電車-止まる」「通り-渋滞」といった交通に関して悲観的な情報が多いとき，自宅に帰る人の割合は低く，勤務地に滞在する人の割合は高いことが分かった．また，第 2 正準変量に着目すると，

表 2　正準相関係数

| | 正準相関係数 | 寄与率 | 累積寄与率 | $\chi^2$値 |
|---|---|---|---|---|
| $\rho_1$ | 0.576 | 46.54(%) | 46.54(%) | 73582 ** |
| $\rho_2$ | 0.560 | 44.02(%) | 90.56(%) | 38654 ** |
| $\rho_3$ | 0.259 | 9.44(%) | 100.00(%) | 6027 ** |

（**：1%有意）

表 3　各係数の推定値

| 変数群 X | $u_1$ | $u_2$ |
|---|---|---|
| $X_1$: 動く-ない | 1.869 | -0.692 |
| $X_2$: 歩く-帰る | -0.502 | 3.468 |
| $X_3$: 運転-再開 | -0.599 | -0.208 |
| $X_4$: 帰れる-ない | -2.241 | -3.084 |
| $X_5$: 通り-渋滞 | 0.683 | 1.652 |
| $X_6$: 電車-動く | -1.476 | -2.048 |
| $X_7$: 地震-速報 | 0.295 | -0.222 |
| $X_8$: 通り-歩く | 0.340 | 0.788 |
| $X_9$: 人-多い | 0.752 | -1.101 |
| $X_{10}$: 電車-止まる | 1.591 | 1.322 |
| 変数群 Y | $v_1$ | $v_2$ |
| $Y_1$:流動人口 | -0.050 | -0.833 |
| $Y_2$:自宅該当者数 | -1.371 | 0.127 |
| $Y_3$:勤務地該当者数 | 1.033 | 0.906 |

「電車-動く」「運転-再開」といった交通に関して楽観的な情報が少ないとき，移動する人の割合は低く，勤務地に滞在する人の割合は高いことが分かった．

## 5　結論および今後の課題

ネットワークの類似性を用いた分析によって，地震発生直後の話題は大きく変化し，その後しばらくすると話題は安定していたことが明らかとなった．また，正準相関分析の結果より，災害時のソーシャルメディアと帰宅行動の間の相関関係を確認した．さらに，モデル適合度の比較から，ソーシャルメディアのテキスト情報と帰宅行動には，時間的な関連性も存在する可能性も示唆された．しかし，正準相関分析はあくまでも相関関係を分析する範疇に留まっており，因果関係を明らかにするものではない．したがって，今後，両者の間の因果関係を分析する手法の開発する必要がある．また，ソーシャルメディア情報の信頼性と妥当性を検証し，ソーシャルメディア情報でのリアルタイムな交通状況を把握することで，災害時，ソーシャルメディア情報を活用した交通混乱の分散，解消を期待できる．

## 6　参考文献

1) 執行文子(2011)，東日本大震災・ネットユーザーはソーシャルメディアをどのように利用したのか，NHK 出版「放送研究と調査」，p2-13
2) 齋藤堯幸(2002)ランク回帰と冗長性分析-理論と応用-，森北出版，p12-17