

SNSを利用した災害情報分析

Analysis of disaster information using the SNS

川合康央[†] 池辺正典[†]
Yasuo Kawai[†] Masanori Ikebe[†]

[†] 文教大学 情報学部

[†] Faculty of Information and Communications, Bunkyo University.

要旨

近年、自然災害に対する防災意識の高まりとともに、情報技術の活用による情報共有の重要性が増してきている。本稿では、災害時に一定程度有用であった SNS を対象として、通常時の災害に関するキーワードを分析し、これを災害時のものと比較することによって、今後の防災計画に資する基礎的なデータの取得を目的とする。

1. 研究の目的と背景

本研究は、SNS (social networking service) を継時的にサンプリングしたデータから災害情報に関連する語をテキストマイニングによって抽出することによって、通常時と災害時における市民の災害意識を分析するものである。本稿で対象とする災害は、自然災害のうち「地震」とした。

先行研究において、相田ら[1]は東日本大震災において Twitter 上の救助要請情報をテキストフィルタリングで抽出し、類似する情報をまとめて表示する Web サイトを開発し運用を行った。また、原ら[2]は同震災時における首都圏の帰宅困難者についての分析を行っている。このように災害時において SNS は情報共有ツールとして一定程度に有効活用されている。しかし、通常時を含んだ災害意識に関する継時的なデータ分析が行われている事例はまだ少ない。

2. 研究の方法

本稿では SNS として Twitter を対象とし、この SNS 上での日本語データをサンプリングする。データ収集期間は 2016 年 10 月 1 日～10 月 31 日である。Twitter API を利用して、「地震」というキーワードを含むテキストを 3 時間ごとに収集することとした。データは、各時間帯に 100 件まで収集することとし、ここから無関係なテキストを除外したものを使用する。取得されたテキストデータは 19898 件であった。

3. 災害情報の分析

収集されたテキストデータは、MeCab を用いて形態素解析を行った。結果、テキストデータは 931660 語の品詞に分解され、そこからユニークな名詞 33881 語を抽出した。これらのうち、名詞の頻出単語上位 200 語を以下に示す (表 1)。

表 1 頻出単語上位 200 語 (名詞)

地震, 県, 熊本, 日, 震度, 鳥取, 時, 速報, 年, 地中, 発生, こと, 人, カウント, 震源, 分, 月, 情報, 日本, 雲, 累積, 被害, 原 発, グラフ, 最大, 中部, 緊急, 中, 災害, 地, 頃, 津波, 噴火, 被災, 支援, 復興, 市, さん, 前, 半年, よう, 観測, 今日, 北海道, 後, 避難, 沖, 性, 心配, ニュース, 者, 強震, 茨城, これ, 方, イタリア, 回, 時間, 的, 円, 阿蘇, 予想, モニタ, 今, 一, 阿蘇山, 事, 型, 南海, 何, 可能, 揺れ, 台風, 影響, 南, 以上, 千葉, 町, 付近, 大丈夫, 事故, 愛知, 福島, それ, ため, 大震災, 注意, 家, 岐阜, トラフ, 私, 長野, 岩手, 気象庁, 禁, 対策, 規模, そう, 国, 東京, 予算, 温泉, 東日本, 梨, 気, お願い, 地方, 城, 画像, 新潟, 警戒, 宮城, マグニチュード, 万, 断層, 三河, 数, 巨大, 弱, 増加, 韓国, みたい, 稲穂, 千, もの, 微小, 活動, 北, 寄付, 波形, 地域, 度, 放送, 拡散, 大分, 報道, 笑, 蓮, 報, 舩, 現在, 所, 億, 予知, 火山, くん, 予測, 反応, 防災, 建物, 対応, 安全, 未満, とき, 必要, 状況, 中越, 水, 政府, 昨日, 支庁, 鹿児島, 変化, 提供, 九州, 推定, 部, みんな, 皆様, 大阪, 前兆, 最近, 募 金, 生活, 不安, 公開, 静岡, 訓練, 金, データ, 観光, 福岡, 化, 機, 自分, 神奈川, 会, 全国, 発表, 発信, マップ, 今年, 以内, 関東, 神社, 写真, ベッド, 危険, 話, ここ

取得されたテキストデータは時間タグを保持している。キーワードを含むテキストデータがどれくらい
の時間で 100 件に達したかというデータをもとに、各日時にどの程度の発言があったかを推測する。
3 時間ごとの発言頻度をもとに、1 日当たりの発言数を推測したものを次に示す (図 1)。

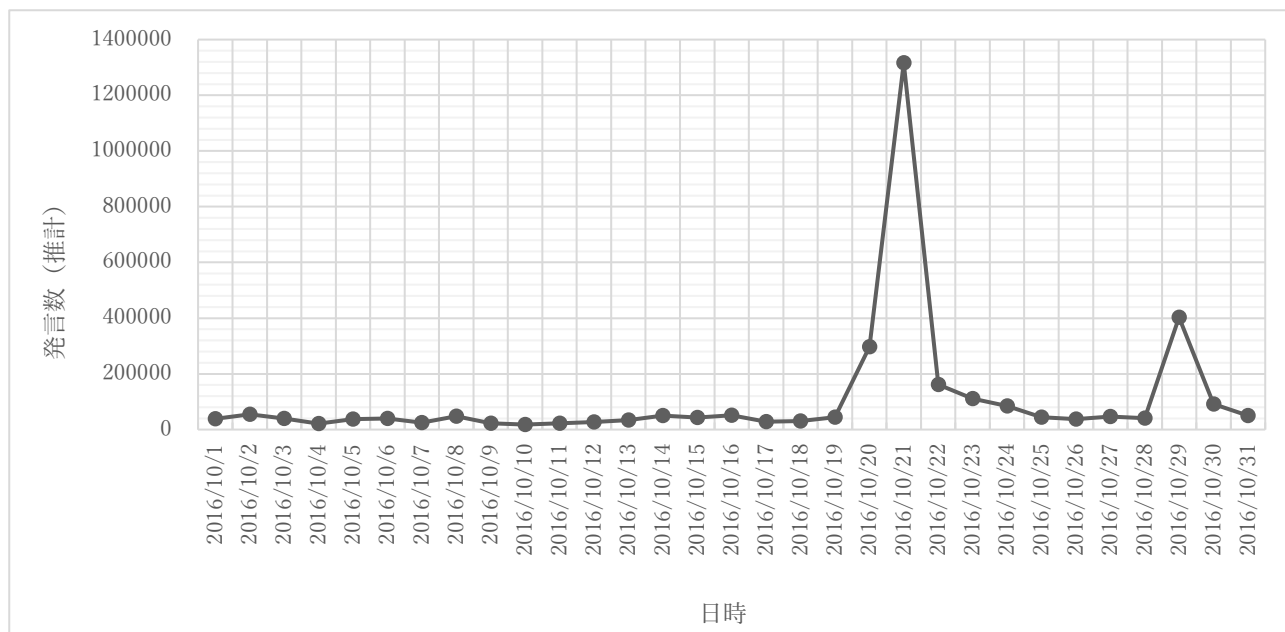


図1 「地震」に関する1日当たりの推計発言数 (2016年10月1日～10月31日)

取得したデータを推測される発言数の傾向から大きく分けて2つの期間に区分した。まず10月1日～19日を通常時のデータとし、その後発生した地震とそれ以降の期間である10月20日～31日を災害時のデータとした。「地震」というキーワードを含む発言数は通常時に平均35550件/日、災害時に平均223649件/日であったと推測される。

通常時の発言傾向としては、「熊本」や「復興」など、過去の地震に関するものが多い傾向が見られた。一方、10月20日11時50分頃、千葉県北東部においてM5.3(震度4)の地震が発生し、さらにその翌21日14時07分頃、鳥取県中部においてM6.6(震度6弱)の地震が発生した。20日の地震は関東圏の人口密集地で発生したものであり、地震そのものではなく「緊急」「速報」などについての発言が多く表れた。地震発生後の12時のデータでは、推測値で89100件/時もの発言数があった。同時間帯における通常時の平均推測値は1378件/時である。しかし、その後の15時の段階では1830件/時(通常時1153件/時)、18時には1772件/時(同1697件/時)と発言は収まっていった。

次に、21日に発生した鳥取県中部地震では、「鳥取」「震度」「震源」など地震の揺れについての発言が多くみられる。地震発生直後の15時のデータでは推測値178200件/時であり、その後も18時で162000件/時、21時で50914件/時(通常時1684件/時)となっており、1日当たりの推測発言数も1316400件/日であった。また、翌22日は161160件/日、23日は111178件/日と発言件数は減っていくものの、実際に大きな揺れを伴い、幸いにして人的被害はなかったものの住宅などに大きな被害をもたらした地震に関しては、発言の収束に時間を要している。

4. まとめ

通常時において災害を意識する発言は多くはないが、いったん大きな災害が発生すると発言数は急速に増加する。しかし、通常時の発言は少ないながらも過去の地震に関する発言が多くみられ、これらの情報に日常的に接することは防災を考えるうえで重要であると考えられる。一方で、緊急地震速報などが発信された際に実際に被害が少ない場合が繰り返されることで、地震に対する慣れと油断が蓄積することも懸念される。

参考文献

- [1] 相田慎, 新堂安孝, 内山将夫, 「東日本大震災関連の救助要請情報抽出サイト」による救助活動支援”, 自然言語処理 = Journal of natural language processing, vol.20, No. 3, 2013, pp.405-422
- [2] 原 祐輔, 「Twitter を用いた東日本大震災時の首都圏の帰宅意思決定分析」, 自然言語処理 = Journal of natural language processing, vol.20, No. 3, 2013, pp.315-334.