

# 災害時のソーシャルメディアにおける情報拡散後の投稿に着目した有益情報抽出システム

代表研究者

宇津 圭祐

東海大学 情報通信学部 准教授

## 1 はじめに

近年、災害時における情報共有の手段としてソーシャルメディアの利活用が注目されている。特に Twitter は即時性が高く利用者数が多いという特徴を有していることから、発災直後の情報発信・収集に活用できるメディアと言われている。また、2011 年の東日本大震災や 2012 年の Hurricane Sandy, 2013 年の Typhoon Haiyan などの際に活用されたことが報告されている[1]~[3]。例えば、東日本大震災発生当日（2011 年 3 月 11 日）、日本国内において約 3,300 万件ものツイートが投稿され、そのうち約 7 割が震災について述べたものであったと報告されている[4]。また、投稿された情報をもとに被災者が救出された事例も存在する。しかし、大規模災害発災時にはツイート数が爆発的に増加する傾向にあり、被害状況の把握や被災者支援のための情報収集は容易ではない。例えば、単純なキーワード検索では災害に関連しない情報（ノイズ）も多く出力される結果となる。このような問題の解決を図るべく、近年、災害時の Twitter 活用を指向した研究・開発が活発に行われている。例えば、水野ら[5] は、災害時における投稿分析・活用システム DISAANA を開発し、Web アプリケーションとして公開している。また、Kitajima et al. [6] は、災害関連ツイートの有益性についてニューラルネットワークを用いて判定する手法を提案した。

我々は、災害に関するニュース投稿をリツイートした直後のツイートに着目して検証を進めてきた。本研究では、まずこの方法の有効性を示すため、平成 28 年（2016 年）熊本地震を対象として検証し、災害関連情報の収集に有用であることを示す。また、本方法を用いて災害に関連する情報を収集するシステムの試作を行った。

## 2 災害関連情報の収集

### 2-1 収集方式の概要

我々は塚本ら[7]の検討を参考に、Twitter におけるニュース投稿リツイート直後の投稿に着目し、災害に関連する情報を収集する方法について検討している[8][9]。図 1 を用いて説明する。まず、災害発災後に、Twitter に災害に関するニュース投稿  $n$  が投稿される（図 1 の 1）。その後、 $n$  を見たアカウントの一部 ( $a_1, a_2, \dots, a_m$ ) が  $n$  をリツイートする（図 1 の 2）。リツイートを行ったアカウントがその直後に投稿を行う（図 1 の 3）。そのリツイート直後の投稿を収集対象とする。

ニュース投稿とそのリツイートを行ったアカウントにおける直後の投稿の例を図 2 に示す。このニュース投稿は、2015 年小笠原諸島西方沖地震発生後に、Yahoo!ニュース (@YahooNewsTopics) に投稿されたものである。サンプルの例を例 1~3 に示している。例 1 では、地震の震度を知り、その大きさについて驚きを表すような投稿となっている。例 2 では、ニュース投稿に掲載されていた地震発生時の震度分布図についての感想が述べられている。また、投稿にはリツイートに対しての感想であることを示す“>RT”や“rt”が含まれている。例 3 はニュース投稿が示す地震についての感想は述べておらず、鉄道運行状況についての内容である。この例では、ニュース投稿の文面のみでは知りえない情報が含まれている。具体的には、ニュース投稿の文面からは、新幹線の運転停止について知ることはできず、ニュース投稿と例 3 の両方を見ることによって、地震の発生とともに新幹線の運転停止が生じている事実を知ることができる。

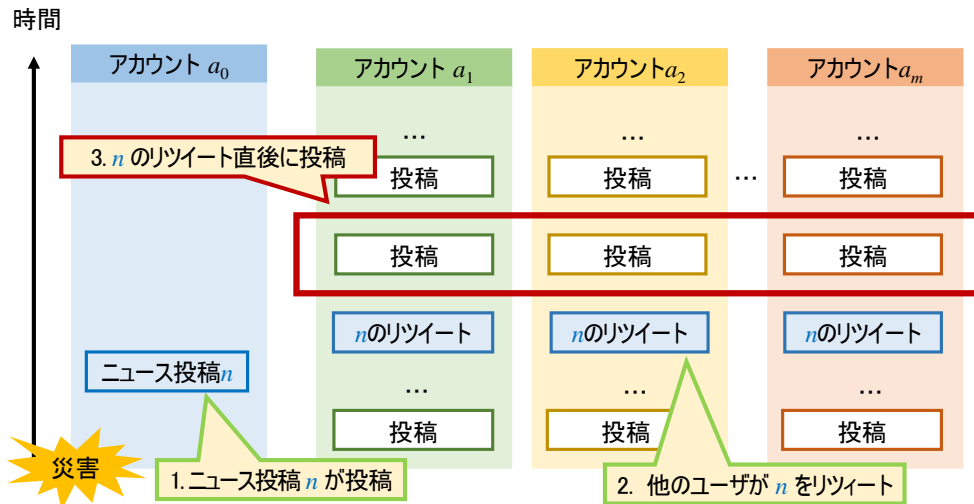


図1 対象のツイート

ニュース投稿



例1

小笠原で震度5強の地震が、東京で震度5弱ってすごくないか？

例2

なんだこの地震図((((;°;Δ;°;)))かかかか なんか恐ろしいぞ> rt

例3

新幹線止まってる(´θ`ⅢⅢ)どーするの日帰りできるの  
[twitter.com/YahooNewsTopic...](https://twitter.com/YahooNewsTopic)

図2 ニュース投稿のリツイート後のツイートの例

2-2 収集方式の有効性

平成 28 年（2016 年）熊本地震に関する 9 件のニュース投稿（表 1 の  $n_1 \sim n_9$ ）を対象に提案手法を適用し、有用性の検証を行う。リツイート直後のツイート収集にはウェブサービス[10] を利用した。表 1 のカッコ内の数字は収集されたツイートの数を示している。

表1 着目するニュース投稿

ニュース投稿 (収集されたツイートの数)	アカウント 投稿日時 投稿内容
$n_1$ (92)	Yahoo!ニュース, @YahooNewsTopics 2016年4月14日午後9:28 【地震情報】 yahoo.jp/eEWiyO
$n_2$ (64)	Yahoo!ニュース, @YahooNewsTopics 2016年4月14日午後9:38 【熊本で震度7の地震】 yahoo.jp/BQt8u8
$n_3$ (79)	Yahoo!ニュース, @YahooNewsTopics 2016年4月14日午後9:42 【熊本で震度7の地震津波なし】 14日午後9時26分ごろ、熊本県熊本地方で震度7を観測する強い地震があった。この地震による津波の心配はない。
$n_4$ (69)	Yahoo!ニュース, @YahooNewsTopics 2016年4月15日午後10:28 【熊本地震物資の支援焦らずに】地震が発生した熊本県に「支援物資を送りたい」との問い合わせが相次ぐ。担当者は「態勢が整うまで待ってほしい」と呼びかけ。 <a href="https://twitter.com/Yahoo_weather/status/721014639643926528">https://twitter.com/Yahoo_weather/status/721014639643926528</a>
$n_5$ (76)	Yahoo!天気・災害, @Yahoo_weather 2016年4月16日午前1:37 【4月16日熊本県で震度6強の地震津波に注意】 2016年4月16日1時25分ごろ、熊本で震度6強の地震がありました。津波注意報が出ています。海岸には近づかないで下さい。 yahoo.jp/KgH1ty
$n_6$ (72)	Yahoo!ニュース, @YahooNewsTopics 2016年4月16日午前4:52 【熊本16日の地震が「本震」か】 気象庁は16日午前1時25分ごろ発生したM7.3の地震が本震で、14日の最大震度7の地震は前震と考えられるとの見解。 <a href="http://yahoo.jp/zupVA7">http://yahoo.jp/zupVA7</a>
$n_7$ (98)	Yahoo!ニュース, @YahooNewsTopics 2016年4月16日午後2:24 【熊本16日の地震が「本震」】 気象庁は16日午前1時25分ごろ発生したM7.3の地震が、14日夜から熊本地方で起きている一連の地震の「本震」だと発表。
$n_8$ (69)	Yahoo!ニュース, @YahooNewsTopics 2016年4月16日午前9:54 【阿蘇山小規模な噴火が発生】 気象庁によると、阿蘇山がきょう午前8時30分に噴火。気象庁は「一連の地震とは直接関連がないとみている」。 <a href="http://yahoo.jp/zupVA7">http://yahoo.jp/zupVA7</a>
$n_9$ (66)	Yahoo!ニュース, @YahooNewsTopics 2016年4月17日午後8:38 【熊本地震空き巣など通報20件】 警察庁によると、熊本市を中心に空き巣や事務所荒らしの通報が約20件あった。益城町や南阿蘇村では同様の通報はないという。

ここで、11人の分析参加者がもとのニュース投稿と併せて見たうえで、収集された各ツイートについて、以下に示すカテゴリーA, Bに該当するかどうか判断する。

- カテゴリーA：ツイートはニュース投稿の内容に関連している。
- カテゴリーB：ツイートはニュース投稿の内容のみでは知り得ない、付加的な情報を含む。

上記のカテゴリーA, B の一方,あるいは両方に該当するツイートを「災害に関連している」とみなす。各ニュース投稿に対して収集された全ツイートのうち,分析参加者の過半数(6人以上)がカテゴリーA, B に該当すると判断したものの割合を図1に示す。ここでは,カテゴリーA, B の一方,または両方に該当するツイートの割合(赤の棒グラフ)に着目する。最高となったのは  $n_1$  の場合で 87.0%, 最低となったのは  $n_4$  の場合で 49.3%であった。すべてのニュース投稿での平均は 69.8%であった。すなわち,収集されたツイートの 7割近くが災害に関連する内容であった。また,分析参加者の過半数がカテゴリーBに該当すると判断されたツイートには,例えば以下のようなものが該当した。

- 例1 ( $n_1$ に関して収集されたツイート) :  
「愛媛県にも緊急地震速報出たけど、東予の新居浜市では揺れを感じなかった。南予では震度 3。震源地の熊本では震度 7。」
- 例2 ( $n_3$ に関して収集されたツイート) :  
「列車運行情報アプリから、地震の影響で山陽新幹線が小倉～博多間で運転見合わせ中って通知があって、また慎重な対応をするもんやなと読み流してそのまま忘れてたけど、とんでもないことが起こった。」
- 例3 ( $n_3$ に関して収集されたツイート) :  
「LINE 電話がつながりやすいというのはすごいな。」
- 例4 ( $n_4$ に関して収集されたツイート) :  
「ニュースでも、持病がある方の薬が足りないって言ってたな...???? 高血圧や糖尿病、心臓病などの毎日服用する薬。でも、こればかりは素人判断で送ったりしないで！」
- 例5 ( $n_6$ に関して収集されたツイート) :  
「熊本・西原村、大切畑ダムが決壊したらしい・・・(°\_°) #jishin」

以上の結果より,災害関連のニュース投稿をリツイートした直後のツイートに着目することにより,災害に関連するツイートを高い確率で収集できることが検証できた。さらにはニュース投稿からだけでは知り得ない情報の収集について可能性を示した。

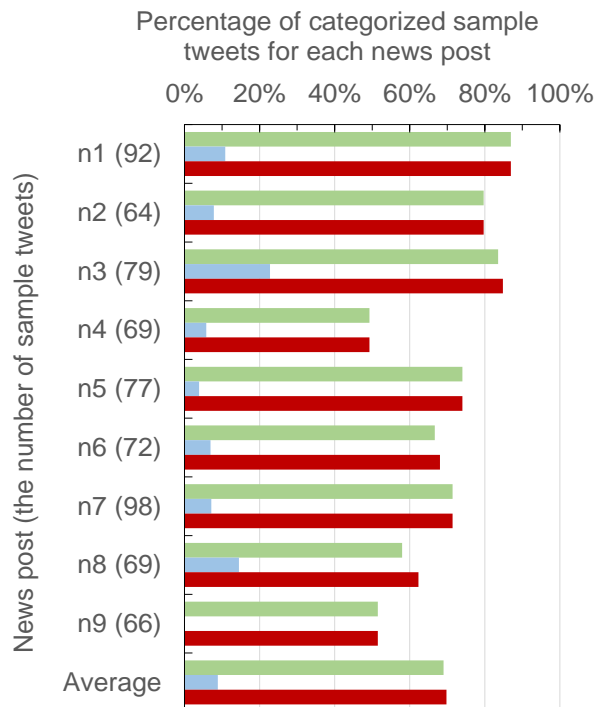


図3 収集されたツイートの分類結果

### 2-3 収集されたツイートにおける災害関連情報の判定

サンプル投稿が災害関連情報であるか否かを判別する方法を提案した。具体的には、まず、ニュース投稿のテキストに対し[9]を適用することによりキーワードを抽出する。次に、抽出されたキーワードに形態素解析を適用し、キーワードに含まれている語句（特徴語句）を抽出する（ここでは、形態素解析には[10]を利用した）。そして、この特徴語句のいずれかを含むサンプル投稿は災害関連情報であると判別する。本研究では、2016年熊本地震に関するニュース投稿9件（ $n_1 \sim n_9$ ）に関するサンプル投稿にこの手法を適用し、有用性を評価する。表2にニュース投稿と抽出された特徴語句の例を示す。ただし、以下の例外を適用する。名詞と接尾辞・接頭辞の組み合わせは単語に分割しない（表2の\*1）。数字の連続については分割しない（表2の\*2）。そして、地震のマグニチュードを示す“M”とその後の数字は分割しない（表2の\*3）。

表2 ニュース投稿（ $n_1 \sim n_9$ ）のキーワードと特徴語句

	キーワード	特徴語句（形態素解析の結果）
$n_1$	地震情報	地震, 情報
$n_2$	震度, 熊本, 地震, 震度7	震度, 熊本, 地震, 震度, 7
$n_3$	震度, 熊本県熊本地方, 震度7, 地震津波, 強い地震, 14日午後9時26分ごろ, 心配	震度, 熊本, 県, 熊本, 地方, 震度, 7, 地震, 津波, 強い, 地震, 14日*2, 午後, 9時*2, 26分*2, ごろ, 心配
$n_4$	熊本地震物資, 問い合わせ, 支援物資, 熊本県, 担当者, 態勢	熊本, 地震, 物資, 問い合わせ, 支援, 物資, 熊本, 県, 担当, 者, 態勢
$n_5$	津波注意報, 震度6強, 熊本県, 地震津波, 海岸, 2016年4月16日1時25分ごろ	津波, 注意, 報, 震度, 6強*1, 熊本, 県, 地震, 津波, 海岸, 2016年*2, 4月*2, 16日*2, 1時*2, 25分*2, ごろ
$n_6$	本震, M7.3, 地震, 前震, 気象庁, 見解, 最大震度7, 16日午前1時25分ごろ, 熊本16日, 14日	本震, M7.3*3, 地震, 前震, 気象, 庁, 見解, 最大, 震度, 7, 16日*2, 午前, 1時*2, 25分*2, ごろ, 熊本, 16日*2, 14日*2
$n_7$	本震, M7.3, 地震, 気象庁, 熊本地方, 一連, 16日午前1時25分ごろ, 熊本16日, 発表, 14日夜	本震, M7.3*3, 地震, 気象, 庁, 熊本, 地方, 一連, 16日*2, 午前, 1時*2, 25分*2, ごろ, 熊本, 16日*2, 発表, 14日*2, 夜
$n_8$	阿蘇山, 気象庁, 噴火, 一連, きょう午前8時30分, 地震, 発生, 関連	阿蘇, 山, 気象, 庁, 噴火, 一連, きょう, 午前, 8時*2, 30分*2, 地震, 発生, 関連
$n_9$	南阿蘇村, 益城町, 熊本地震空き巣, 警察庁, 事務所荒らし, 熊本市, 通報, 中心, 通報20件	南阿蘇, 村, 益城, 町, 熊本, 地震, 空き巣, 警察, 庁, 事務所, 荒らし, 熊本, 市, 通報, 中心, 通報, 20件*2

上記の判別手法を[1]のニュース投稿9件に適用した結果について述べる。人手による分類により災害関連情報であると判断されたサンプルの集合を  $X$ , 特徴語句のいずれかを含むサンプルの集合を  $Y$  としたときの  $|X|, |Y|, |X \cap Y|$  の値について表3に示す。そして、判別手法の適合率, 再現率,  $F$  値を以下のように求める。

- 適合率 =  $|X \cap Y| / |Y|$
- 再現率 =  $|X \cap Y| / |X|$
- $F$  値(F-measure) =  $(2 \times \text{適合率} \times \text{再現率}) / (\text{適合率} + \text{再現率})$

求められた判別手法の適合率, 再現率,  $F$  値を図4に示す。適合率については、最高となったのは  $n_1$  と  $n_2$  の場合で 1.000, 最低となったのは  $n_9$  の場合で 0.714 であった。すべてのニュース投稿（ $n_1 \sim n_9$ ）での平均は 0.935 であり、特徴語句のいずれかを含むサンプル投稿は、高い確率で災害関連情報であることが示された。これにより、災害関連情報を収集する上で、ニュース投稿の特徴語句を含むサンプル投稿を収集することの

有用性が示せた。一方、すべてのニュース投稿 ( $n_1 \sim n_9$ ) について、再現率の平均は 0.491, F 値の平均は 0.627 にとどまった。これは、収集されたツイートにおいて、ニュース投稿の特徴語句が含まれているかどうかで判定を行っていることが原因である。これは、複数のニュース投稿から提案の収集方式を適用することにより多くのツイートを収集することで対処できると考えられるが、再現率と F 値の向上は今後の検討課題である。

表 3 収集されたツイートの詳細

	災害関連情報であると判断されたサンプルの数 $ X $	特徴語句のいずれかを含むサンプルの数 $ Y $	XとYの両方に該当するサンプルの数 $ X \cap Y $
$n_1$	16	16	16
$n_2$	20	20	20
$n_3$	49	50	49
$n_4$	14	15	14
$n_5$	30	31	30
$n_6$	30	31	30
$n_7$	45	47	45
$n_8$	26	29	26
$n_9$	10	14	10

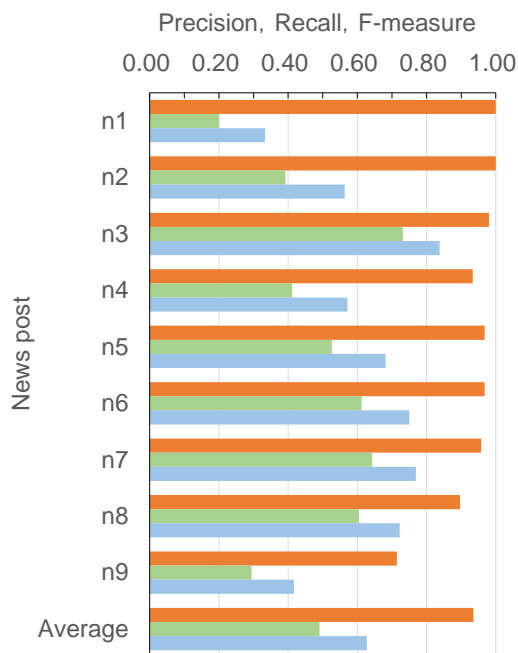


図 4 収集されたツイートの判定結果

### 3 収集システムの試作

コンテンツプロバイダや地方自治体職員あるいは一般ユーザが、Twitter を用いて情報を簡単に収集するシステムの実現を目的とし、2 章で述べた方法を用い、ニュース投稿に関連する投稿を収集するためのベースとなるシステムについて試作した。本システムは PHP5 によって作成したプログラムを Web サーバ上に設置することにより利用できるようにした。

#### 4-1 システムの構成・利用方法

本システムでは、ニュース投稿のリツイート直後の投稿を収集したい対象のニュース投稿  $n$  の URL を入力することにより、 $n$  のリツイート直後の投稿を収集し、 $n$  の内容に関連すると判別されたサンプル群  $o(n)$  を出力する。

具体的に図 5 に示す。入力されたニュース投稿  $n$  の URL をもとに、Twitter API により、 $n$  のテキストと、 $n$  をリツイートしたアカウントの一覧を取得する。これをもとに、リツイートしたアカウントのリツイート直後の投稿を取得する（図 5 の 1）。収集されたツイートの群  $s(n)$  と表記している。ただし、収集されたサンプルが、ニュース投稿のリツイートを行った日時よりも 30 分以上経過している場合、そのサンプルは除外する（すなわち、リツイート直後の投稿とみなさない）。続いて、 $n$  のテキストを、Yahoo! テキスト解析：キーフレーズ抽出[9]を用いて、 $n$  のキーフレーズの抽出を行う（図 5 の 2）。そして、 $n$  のキーフレーズを Yahoo! テキスト解析：日本語形態素解析[11]を用いて、 $n$  の特徴語句を抽出する（図 5 の 3）。最後に、 $s(n)$  の各サンプルのうち、 $n$  の特徴語句を 1 個以上含むサンプルの群  $o(n)$  を出力する。

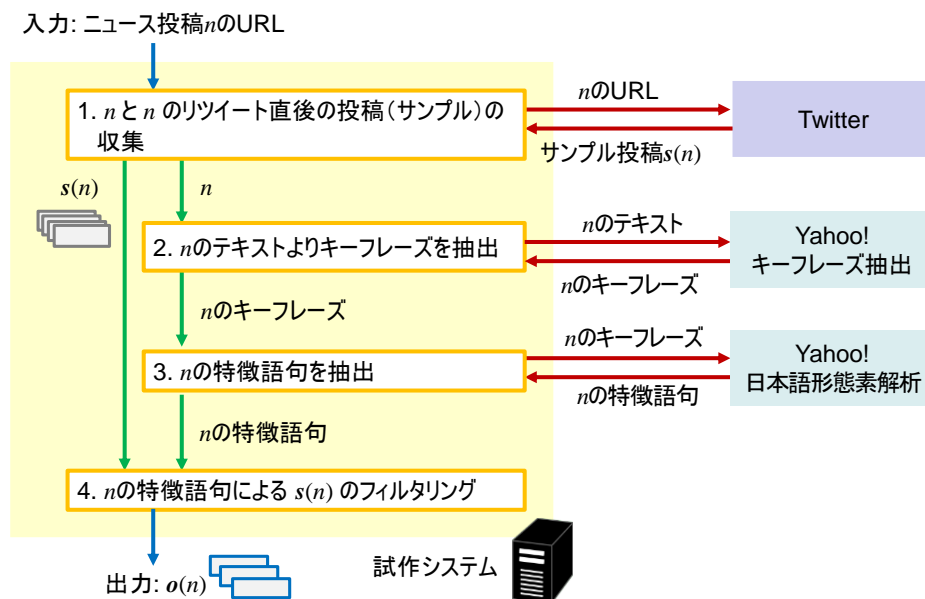


図 5 収集のためのベースシステムの概要

システムの使用方法について説明する。ユーザはウェブブラウザを用いて、本システムに接続すると、図 6 のように初期画面が表示される。この画面で入力する項目は以下(1)~(5)の通りである。すべての項目について、入力あるいは選択が必須である。

- (1)ニュース投稿  $n$  の URL :  
リツイート直後の投稿を収集したい対象のニュース投稿の URL を入力する。
- (2)ニュース投稿  $n$  のリツイート直後の投稿を検索するアカウント数:  
初期値は 100 である。Twitter API の制限があるため、最大値は 100 としている。
- (3)収集された投稿が(1)とは別の投稿に関するリツイートである場合に表示するかどうか:  
「Yes」を選択すると、リツイート直後の投稿が、 $n$  以外の別の投稿に関するリツイートで合った場合にも検索対象となる。初期値は「Yes」である。
- (4)ニュース投稿に含まれている特徴語句を含む投稿のみ絞込み表示かどうか:  
初期値は「Yes」である。「Yes」を選択すると、検索されたサンプル投稿のうち、ニュース投稿の特徴語句を 1 個以上含むもののみをフィルタリングして表示される。なお、図 6 に示す画面上では、「特徴語句」ことを理解しやすいよう、「キーワード」と表示している。
- (5) ユーザあたりの最新の投稿から遡っての取得する件数:  
 $n$  のリツイートを行ったユーザのタイムラインにおいて、 $n$  のリツイート直後の投稿を、そのユーザ

の最新投稿から何件を上限に遡って検索するかについて指定する。初期値は 100 である。Twitter API の制限があるため、最大値は 3200 としている。

**システムの概要 /Outline**  
Twitterにおけるニュース投稿のURLを入力することで、そのニュース投稿をリツイートしたアカウントにおけるリツイート直後の投稿を収集できる。  
ニュース投稿のリツイート直後の投稿が、リツイートした後10分以上経過していた場合は除外される。  
This system can collect Twitter posts after retweeting of the news post.  
Tweets submitted after more than 10 minutes of retweeting are not displayed.

**(1)ニュース投稿のURLを入力**  
Enter a URL of the news post on Twitter.

**(2)ニュース投稿のリツイート後の投稿を検索するユーザ数を設定**  
Set the number of users who posted after retweeting of the news post will be found by the system.  
初期値 /Default:100, 最小 /Min:10, 最大 /Max:100  
 ユーザ数 /The number of users : 100

**(3)収集された投稿が(1)とは別の投稿に関するリツイートである場合に表示する？**  
Display if sample posts are retweets of a post other than the news post in (1).  
初期値 /Default: Yes  
 Yes  No

**(4)ニュース投稿に含まれているキーワードを含む投稿のみ絞込む？ ※日本語用**  
Filtering posts which do NOT include any keywords of the news post. The option is for Japanese.  
初期値 /Default: Yes(Filtering available)  
 Yes (Filtering enabled)  No

**(5)ユーザあたりの最新の投稿から遡っての取得する件数**  
Set the number of posts will be found from a latest post at a user.  
初期値 /Default:100, 最小 /Min : 100, 最大 /Max : 3200  
 件数 /The number of posts : 100

図 6 試作システムの初期画面

#### 4-2 システムの動作例

システムの動作例について以下に述べる。ここでは図 7 に示す、2017 年 2 月 28 日（火）に福島県沖で発生したマグニチュード 5.6 の地震に関する、Yahoo!ニュースの投稿を例に用いた。投稿の URL は、<https://twitter.com/YahooNewsTopics/status/836485438319677441>、投稿時刻は 2017 年 2 月 28 日（火）午後 4 時 57 分、内容は「2 月 28 日 16 時 49 分ごろ宮城県などで震度 5 弱を観測する地震が発生。この地震による津波の心配はありません。→ <http://yahoo.jp/BYekzA>（外部ページへのリンク URL）」であった。

このニュース投稿の URL を本システムに入力した。また、(2)~(5)の各項目については初期値とした。収集は同日午後 6 時頃実行した。出力結果の表示ページの一部を図 8 に示す。ページ上部には、ニュース投稿について、投稿したアカウント、投稿の本文、リツイート数、投稿日時(a)が表示されている。また、絞込オプションの(3)、(4)の設定、ニュース投稿のキーフレーズ抽出結果、特徴語句の抽出結果が表示されている。ページ下部には、出力サンプルが表形式で表示されている。表示項目は、アカウント名、投稿の本文、サンプルが投稿された日時(b)、ニュース投稿をリツイートした日時(c)、(c)から(b)の経過時間（単位は秒）、そして(a)から(b)の経過時間（単位は秒である）。なお、図 8 において、プライバシー保護のため、投稿アカウントについてはモザイク処理を施している。

出力されたサンプルの件数は 27 件であった。このうち、出力されたサンプルのうち、特徴語句を 1 個以上含むものとして出力されたものは、7 件であった。これらのサンプルはクリーム色の背景で表示されている。また、別の投稿のリツイートであった（投稿の本文が RT からはじまっている）ものは、20 件であった。これらのサンプルは白の背景で表示されている。





図7 ニュース投稿の例

以下は出力表示されたツイートの例である。

- 例1：特徴語句を1個以上含むものとして出力されたもの  
 【#東日本大震災 の悪夢、再び】 16:49 分頃、福島県沖で震度 5 弱の地震が発生。この地震による津波の心配はない。 (hashtags and URL) 近い将来、震度 7 クラスの 地震の可能性は
- 例2：別の投稿のリツイートであったもの  
 RT @nhk\_seikatsu: 【福島第一・第二原発 新たな異常無し】 東京電力によりますと、双葉町と大熊町にまたがる福島第一原子力発電所と、楡葉町と富岡町にまたがる福島第二原子力発電所では、今回の地震による新たな異常は発生していないということです。 (URL).

今回の試作システムでは、サンプルの収集について、ニュース投稿が掲載されてからある程度の時間が経過した後に行うことを想定したシステムとなっている。より実用化するために、ツイートをリアルタイムで収集する必要があり、今後はこれを実現するためシステムの改良を行う。

アカウント / Account	ニュース投稿(1) / Text of news post (1)	リツイート / Retweets	投稿日時 / Posted date-time
Yahoo!ニュース	2月28日16時49分ごろ宮城県などで震度5弱を観測する地震が発生。この地震による津波の心配はありません。→ <a href="https://t.co/CJJGgEz70">https://t.co/CJJGgEz70</a> <a href="https://t.co/ZLAC8Dtebb">https://t.co/ZLAC8Dtebb</a>	393	2017-02-28 16:57:24

【オプション設定 / Options】

収集された投稿が(1)とは別の投稿に関するリツイートである場合に示す？  
Display if sample posts are retweets of a post other than the news post in (1). : Yes

ニュース投稿に含まれているキーワードを含む投稿のみ絞込む？  
Filtering posts which do NOT include any keywords of the news post. : Yes

【キーワード / Key phrases】

地震, 津波, 震度5弱, 2月28日16時49分ごろ宮城県, 発生, 心配,

【絞り込めるキーワード / Key words for filtering】

地震, 津波, 震度, 5弱, 2月, 28日, 16時, 49分, ごろ, 宮城県, 発生, 心配,

【条件に一致する投稿 / Posts match the options】

収集時刻 / Date-time: 2017-02-28 17:59:13

アカウント Account	ニュース投稿(1)のRT直後の投稿 Posts after retweeting of the news post (1)	RT直後の投稿日時 posted date-time	RT時刻 date-time of RT	RTからの経過時間 Time btw RT - the post	ニュース投稿からの経過時間 Time btw the sample - the news
	RT @e_m__e: みんな地震大丈夫? <a href="https://t.co/wSsI3z8mh">https://t.co/wSsI3z8mh</a>	2017-02-28 17:48:14	2017-02-28 17:47:56	18	3050
	みんな地震大丈夫?	2017-02-28 17:46:14	2017-02-28 17:46:01	13	2930
	RT @value_investors: 東日本大震災のときは、直前に震度5弱5強の地震が何回も起きたんですね。	2017-02-28 17:38:35	2017-02-28 17:37:56	39	2471
	RT @nhk_seikatsu: さきほどの地震で震度3の揺れが観測された仙台市青葉区の映像です。宮城県岩沼市、福島県相馬市、福島県相馬市、福島県楡葉町、福島県双葉町では震度5弱の揺れを観測しました。この地震による津波の心配はありません。 <a href="https://t.co/ovZA...">https://t.co/ovZA...</a>	2017-02-28 17:33:50	2017-02-28 17:33:44	6	2186
	RT @nhk_seikatsu: 【福島第一・第二原発 新たな異常はなし】 東京電力によりますと、双葉町と大熊町にまたがる福島第一原子力発電所と、楡葉町と富岡町にまたがる福島第二原子力発電所では、今回の地震による新たな異常は発生していないということです。 <a href="https://t...">https://t...</a>	2017-02-28 17:31:14	2017-02-28 17:30:57	17	2030
	RT @marukarichan11: 東北地震、、、大丈夫ですか(〜)地震、最近お話し、みなさん気をつけてくださいませ。	2017-02-28 17:30:54	2017-02-28 17:30:51	3	2010
	【#東日本大震災の悪夢、再び】 16:49分頃、福島県沖で震度5弱の地震が発生。この地震による津波の心配はない。#福島 #宮城 #地震 #jshin #拡散希望RT <a href="https://t.co/SKViCgZt5H">https://t.co/SKViCgZt5H</a> 近い将来、震度7クラスの地震の可能性は	2017-02-28 17:34:34	2017-02-28 17:29:51	283	2230
	RT @UN_NERV: 【地震情報 2017年2月28日】 16時49分頃、福島県沖で震源とする地震がありました。震源の深さは約50km。地震の規模はM5.6。最大震度5弱を宮城県、福島県で観測しています。この地震による津波の心配はありません。 <a href="https://t.co/b0...">https://t.co/b0...</a>	2017-02-28 17:29:35	2017-02-28 17:29:09	26	1931
	RT @nhk_seikatsu: 震度5弱では、固定していない家具が移動し、中には倒れるものもあります。また、補強していないブロック塀が倒れることもあります。壁からの落下物や、割れたガラスなどに十分注意してください。	2017-02-28 17:27:24	2017-02-28 17:27:18	6	1800
	明日また東京に出張するんだけど全然地震にあったことないから怖い。	2017-02-28 17:27:24	2017-02-28 17:26:26	58	1800
	震度2の地域揺れが感じず	2017-02-28 17:25:12	2017-02-28 17:24:06	66	1668
	RT @marukarichan11: 東北地震、、、大丈夫ですか(〜)地震、最近お話し、みなさん気をつけてくださいませ。	2017-02-28 17:23:41	2017-02-28 17:23:37	4	1577
	RT @value_investors: 東日本大震災のときは、直前に震度5弱5強の地震が何回も起きたんですね。	2017-02-28 17:21:46	2017-02-28 17:21:37	9	1462
	RT @nhk_seikatsu: 震度5弱では、固定していない家具が移動し、中には倒れるものもあります。また、補強していないブロック塀が倒れることもあります。壁からの落下物や、割れたガラスなどに十分注意してください。	2017-02-28 17:21:05	2017-02-28 17:21:02	3	1421
	RT @news24ntv: #速報 宮城南部、福島県沖で震度5弱の#地震。津波の心配なし。 <a href="https://t.co/pcW5riQzJv">https://t.co/pcW5riQzJv</a>	2017-02-28 17:18:08	2017-02-28 17:17:56	12	1244
	RT @tbs_news1: 【速報】JR東北新幹線は地震の影響で新白河と古川の間の上下線が一時運転を見合わせていたが、運転を再開した。	2017-02-28 17:18:02	2017-02-28 17:17:53	9	1238
	@riogurande051 地震大丈夫でした?	2017-02-28 17:18:44	2017-02-28 17:17:23	81	1280
	RT @mainichi: 28日午後4時49分ごろ、宮城県南部と福島県沖で震度5弱を観測する#地震がありました。この地震による	2017-02-28	2017-02-28	13	1100

図8 収集された投稿の表示画面 (一部)

#### 4 まとめと今後の検討課題

本研究では、災害に関するニュース投稿をリツイートした直後のツイートに着目して検証を進めてきた。本研究では、まずこの方法の有効性を示すため、平成28年(2016年)熊本地震を対象として検証した。その結果、災害関連情報の収集に有用であることを示した。災害に関連するツイートを高い確率で収集できることが検証できた。また、特徴語句による判定を行い、高い適合率で災害に関連するツイートを判別できることを示した。検討課題として、再現率およびF値の改善が挙げられた。災害発災時には多くの情報を収集することも必要になると考えられるため、フィルタリングにより除外されたサンプルの内容を吟味し、フィルタリング方法が適切であるかを検討する必要がある。その結果に応じて、他の判別方法あるいは絞り込みオプションについて検討する予定である。

続いて、コンテンツプロバイダや地方自治体職員あるいは一般ユーザが、Twitter を用いて情報を簡単に収集するシステムの実現を目的とし、上記方法を用い、ニュース投稿に関連する投稿を収集するためのベースとなるシステムの試作を実施した。具体的には、Twitter における災害に関するニュース投稿の URL を入力することにより、そのニュース投稿のリツイート直後の投稿を収集し、ニュース投稿に関連する投稿を出力するシステムを試作した。今後の検討課題は以下のとおりである。今回の試作システムでは、サンプルの収集について、ニュース投稿が掲載されてからある程度の時間が経過した後に行うことを想定したシステムとなっている。より実用化するために、ツイートをリアルタイムで収集する必要があり、今後はこれを実現するためシステムの改良を行う。

## 【参考文献】

- [1] P. Meier, Digital Humanitarians, “How Big Data Is Changing the Face of Humanitarian Response,” CRC Press, 2015
- [2] T. Simon, A. Goldberg, and B. Adini, “Socializing in emergencies -A review of the use of social media in emergency situations,” International J. Information Management, vol.35, no.5, pp.609–619, 2015
- [3] B. D. M. Peary, R. Shaw, and Y. Takeuchi, “Utilization of Social Media in the East Japan Earthquake and Tsunami and its Effectiveness,” J. Natural Disaster Science, vol.34, no.1, pp.3–18, 2012.
- [4] 三浦麻子, 鳥海不二夫, 小森政嗣, 松村真宏, 平石界, “ソーシャルメディアにおける災害情報の伝搬と感情：東日本大震災に際する事例,” 人工知能学会論文誌, vol.31, no.1, pp.1- 9, 2015.
- [5] 水野淳太, 後藤淳, 大竹清敬, 川田拓也, 鳥澤健太郎, クロエツェージュリアン, 田仲正弘, 橋本力, 奥村明俊, “モダリティ解析の導入による対災害 SNS 情報分析システム DISAANA の質問応答性能の改善,” 情処学論コンシューマ・デバイス&システム, vol.6, no.1, pp.106- 120, 2016.
- [6] R. Kitajima, R. Kamimura, O. Uchida, F. Toriumi, “Neural Potential Learning for Tweets Classification and Interpretation”, Proc. 7th International Conference on Soft Computing and Pattern Recognition, 2015
- [7] 塚本悠馬, 笹野遼平, 高村大也, 奥村学, “マイクロブログ上の告知投稿に対する非明示的な関連投稿の収集”, 情報処理学会研究報告, Vol.2013-NL-214 No.14, pp.1-8, 2013 年
- [8] Ayami Manaka, Shiori Kodama, Akio Ogata, Osamu Uchida, Yoshiro Yamamoto, Hiroshi Ishii, Keisuke Utsu, “Collection of Disaster-related Information by Focusing on Twitter Posts Immediately after Retweeting Announcement Posts”, Proceedings of the International Conference on 2016 IEEE Region 10 Conference, pp. 2253-2257, 2016
- [9] Yahoo! テキスト解析：キーフレーズ抽出, <http://developer.yahoo.co.jp/webapi/jlp/keyphrase/v1/extract.html>
- [10] 形態素解析ウェブアプリ Unidic-MeCab, <http://www4414uj.sakura.ne.jp/Yasanichi1/unicheck/>
- [11] Yahoo! テキスト解析：日本語形態素解析, <http://developer.yahoo.co.jp/webapi/jlp/ma/v1/parse.html>

## 〈発表資料〉

題 名	掲載誌・学会名等	発表年月
Twitter におけるニュース投稿のリツイート直後の投稿に着目した情報収集システムの試作	電子情報通信学会技術研究報告, Vol. 117, no. 67, ICTSSL2017-1, pp. 1-6,	2017 年 5 月
A Disaster-related Information Gathering Method by Focusing on Tweets Posted Immediately after Retweeting of News Posts about Kumamoto Earthquake	Proceedings of the 2017 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'17), pp. 196-202, Jul. 2017	2017 年 7 月
Web Application Prototype for Collecting Disaster-related Information Focusing on Tweets Immediately after Retweeting News Post	The 4th International Conference on Information and Communication Technologies for Disaster Management	2017 年 12 月