

災害情報のリツイート傾向に基づいたアクティブユーザのグループ化と評価

佐藤洸太郎¹ 阿部倫之¹

概要：近年、ソーシャルメディアから発信される災害情報の活用が進んでいる。ツイートを利用した災害情報の共有システムとしては、NICTのDISAANAが運用されており、自治体の防災担当者等が被災者のツイートを迅速に確認できる仕組みを提供している。ここで、被災者に近いフォロワーが少数の場合には初期段階においてツイートの拡散力が弱く、twitter内において災害情報が共有されるまでには時間がかかる。また、支援者からのツイートはスクープ性が低いため、一般に拡散されにくい。そこで、過去の震災等において災害情報の拡散に貢献したアクティブユーザを把握できれば、リプライ等によって災害情報の迅速な拡散を促すことができる。本稿では、リツイート傾向の類似性に基づいてアクティブユーザをグループ化し、グループごとに被災ツイートの緊急性等を判定する手法を提案する。また、2018年度に投稿された災害関連ツイートをを用いて提案手法の評価を実施したので報告する。

キーワード：Twitter, リツイート, クラスタリング, 災害情報, 拡散, アクティブユーザ, SVM

Clustering and its Evaluation of Active Users based on the Retweet pattern of the disaster information on Twitter

SATO KOTARO^{†1} NORIYUKI ABE^{†1}

1. はじめに

近年、ソーシャルメディアから発信される災害情報の活用が進んでいる。特にTwitterは、災害等の緊急事態における情報発信手段としてよく利用されており、熊本地震のような大災害だけではなく、集中豪雨等の警報レベルにおいても多数投稿されている。

東日本大震災の発生以降、ツイートを利用して災害状況や復興状況を把握する試みがされてきた。例えば、ジオタグ付きツイートの位置情報を利用して発信地点の投稿密度から被害状況を把握する研究[1]や、ツイート内容と帰宅困難者の行動（歩く、帰れる）との関連性を分析する試みがある[2]。また、熊本地震で投稿されたツイートの単語感情極性を利用して九州地方の被災者の心理状態を把握する研究などもある[3][4]。この熊本地震では、内閣官房のツイッター分析班が、NICTの対災害SNS情報分析システムDISAANA[5][6]を活用して情報収集にあたっている。DISAANAのエリア指定モードを利用すると、指定エリア内で投稿されたツイートが「洪水が発生している」や「浸水が酷い」などの災害カテゴリごとにリアルタイムにグループ化され、オリジナルのツイートとその投稿地点が地図上に表示される。これにより、災害情報の発信元を迅速に確認できる仕組みを提供している。

ここで、大規模災害時のように、被災報告や支援要請等のツイートが大量に投稿されると、拡散力が弱い（リツイ

ート頻度が低い）ツイートは注目されなくなる。スクープ性が高く、情緒訴求力の強いツイートが集中的にリツイートされる傾向にあるため、例えば、被災者のフォロワーが少数で関心が低い場合にはツイートの拡散力が弱く、被災情報が注目されるまでには時間がかかる。

一方、内閣官房IT総合戦略室防災班が発行している「災害対応におけるSNS活用ガイドブック」によれば、地方公共団体に対して、災害時における行政側からの情報発信および拡散手段としてSNSを十分活用することを推奨している[7]。しかし、自治体が緊急性の高い災害情報をツイートしたとき、それが迅速に拡散されるか否かについては、災害関連の情報に関心の高い熱心なフォロワーの行動力に依存している。ここで、過去の震災等において災害情報の拡散に貢献したアクティブユーザを上手くグループ化して評価できれば、リプライ等によって災害情報の迅速な拡散を促すことが期待できる。情報拡散については、デマや風評等の拡散防止や伝搬経路の把握に関する研究が多く[8][9]、リツイートユーザのグループ化や「善意の拡散者」としての利用等については検討例が少ない。

一定のリツイート頻度を持つ拡散者を「アクティブユーザ」としたとき、災害情報の緊急性や被災内容等の違いによってリツイート傾向に偏りがあれば、類似の傾向を示すアクティブユーザをグループ化できる。

本研究では、リツイート傾向の類似性を利用してアクティブユーザをグループ化し、グループの特徴や持続性、および緊急性の高い災害ツイートに対する反応状況等について分析している。本稿では、リツイート傾向の類似性に基

¹ 金沢工業大学
Kanazawa Institute of Technology

づいてアクティブユーザをグループ化する手法を提案し、2018 年度に投稿された災害ツイートをを用いて評価を実施したので報告する。

2. システム概要

提案手法を利用したシステムの概要を図 1 に示す。警報や避難などの災害関連語を災害マーカとして定義し、これを含む災害ツイートをリアルタイムに収集する。リツイートには、リツイートユーザ（拡散者）の情報とリツイート元ユーザ（投稿者）の情報がセットされているため、災害リツイートから、< 投稿者, 拡散者のリツイート頻度 > の関係を拡散者ごとに抽出する。

一定のリツイート頻度を持つ拡散者を「アクティブユーザ」として選択し、そのリツイート傾向を抽出する。これは、投稿者からのリツイート頻度を素性とする特徴ベクトルで表現する。この特徴ベクトルの類似性によりアクティブユーザをクラスタリングする。

グループは、緊急性の高いツイートのリツイート頻度等で評価し、拡散グループの評価履歴として蓄積する。災害カテゴリごとにグループの構成が変化し、さらに時間の経過に伴ってグループの求心力が低下するため、災害発生時においてグループの持続性を評価する必要がある。

災害ツイートが投稿されたとき、拡散グループの評価履歴に基づいて、グループと拡散の範囲を決定し、リプライ等により災害ツイートの拡散を支援する。本稿では、クラスタリングとアクティブユーザの評価結果について報告する。

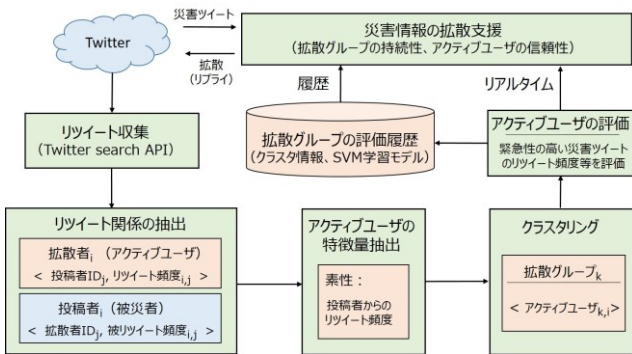


図 1 システム概要

3. リツイート関係と特徴量の抽出

提案手法を用いたリツイート関係の抽出方法を図 2 に示す。リツイートが発生したとき、拡散者のリツイート回数を投稿者の被リツイートログに記録する。また、拡散者のリツイートログにも投稿者からのリツイート回数を記録する。リツイート回数を累積する場合、過去の寄与度を割引く方法もあるが、今回は対象期間が短いため単純累積とす

る。拡散者のリツイート傾向は、ログに記録されているリツイート頻度で表現する。また、拡散者の特徴量は、投稿者 (S1, S2, S3) からのリツイート頻度を素性とする特徴ベクトルで表現する。図 2 において、拡散者の特徴ベクトルは、 $R1 = (2.0, 3.0, 0.0)$, $R2 = (0.0, 5.0, 0.0)$, $R3 = (0.0, 1.0, 2.0)$ となる。

投稿者側の被リツイート傾向については、ログに記録されている被リツイート頻度で表現できる。また、投稿者の特徴量は、拡散者 (R1, R2, R3) からの被リツイート頻度を素性とする特徴ベクトルで表現する。図 2 において、投稿者の特徴ベクトルは、 $S1 = (2.0, 0.0, 0.0)$, $S2 = (3.0, 5.0, 1.0)$, $S3 = (0.0, 2.0, 0.0)$ となる。今回は、拡散者のグループ化を目的としているため、投稿者側の特徴量はスコア (被リツイート頻度) の算出のみに使用する。

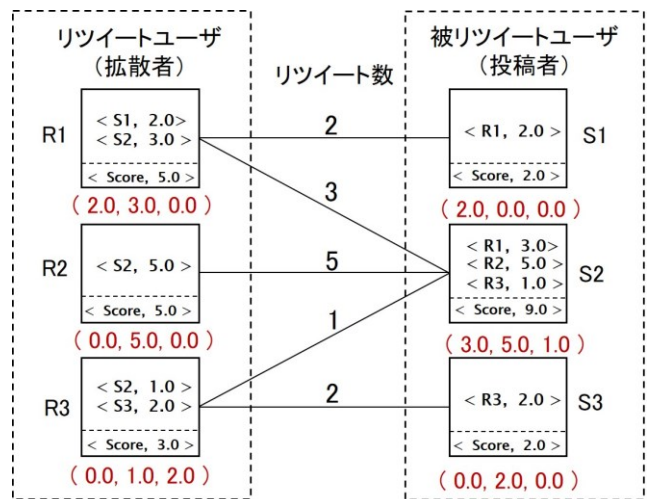


図 2 投稿者と拡散者のリツイート関係

4. 実験と評価

4.1 災害ツイートの収集

Twitter Search API を用いて、日本国内で投稿された災害関連ツイートをリアルタイムに収集している。システムは Twitter4J の Java クラスライブラリを用いて実装している。表 1 に示すような災害マーカを定義しており、これを検索クエリに指定して稼働させている。この災害マーカを含むツイートの投稿頻度は、リツイートを含めて台風 12 号で約 20TPS (tweets per second)、北海道地震で約 30TPS である。

表 1 災害マーカ

カテゴリ	災害マーカ
災害事象	災害, 洪水, 氾濫, 地震, 台風, 決壊
感情	恐怖, 怖い, 救助, 安否
インフラ	停電, 通行止, 断水, 復旧
緊急性	避難, 警報

今回の実験では、2018年に発生した4つの自然災害を取り上げる。災害の名称を以下に示す。

- ・平成30年台風第12号（以下、台風12号）
- ・大阪北部地震（以下、大阪地震）
- ・平成30年台風第21号（以下、台風21号）
- ・北海道胆振東部地震（以下、北海道地震）

ここで災害ツイートは、災害発生の前後を含む概ね24時間以内に投稿されたツイートを実験対象にする。上記の災害ツイートを確認すると、緊急性の高いツイートは、災害発生の初期段階から急速に増加しており、概ね24時間以内に収束している。各災害発生期間におけるツイートの収集状況を表2に示す。

表2 実験で使った災害ツイート

災害名	収集日	ツイート取得総数	ツイート取得数	リツイート取得数
台風12号	2018/7/29 ^{注1}	2,241,559	1,202,714	1,038,845
大阪地震	2018/6/18	2,055,709	658,500	1,397,209
台風21号	2018/9/4	2,295,746	1,242,450	1,053,296
北海道地震	2018/9/6	2,677,310	1,242,450	1,053,296

注1: 2018/7/28 16:00~24:00 までのツイートをを含む。

4.2 リツイート関係の抽出

4.2.1 ユーザ情報の抽出

災害ツイートからユーザ情報を抽出する。表3に各災害発生時におけるツイートの投稿者数（被リツイートユーザ数）と、その拡散者数（リツイートユーザ数）を示す。リツイートのみのユーザは、投稿者数に含めていない。

拡散者の内、1日あたり24回以上リツイートしているユーザを「アクティブユーザ」と呼ぶ。また、アクティブユーザから24回以上リツイートされている投稿者を「ソース（情報源）」と呼ぶ。この抽出結果を表4に示す。

表3 災害ツイートの拡散者数と投稿者数

災害名	拡散者数	投稿者数
台風12号	497,114	79,752
大阪地震	619,426	63,529
台風21号	526,833	73,808
北海道地震	690,585	80,650

表4 アクティブな拡散者数と投稿者数

災害名	アクティブユーザ数	ソース数
台風12号	2,547	509
大阪地震	2,271	530
台風21号	1,771	358
北海道地震	5,975	1107

4.2.2 リツイート関係の抽出

図2に示すようなリツイート関係を抽出する。この処理は、リツイートからユーザ情報を抽出する際に同時に実施

できる。すべての投稿者と拡散者についてリツイート関係を抽出し、その後、アクティブユーザとソースのスコア（図2:リツイート頻度）に基づいて特徴ベクトルの素性を決定する。ここで、「台風12号」のアクティブユーザのランキング結果とリツイート数およびリツイートユーザ数の関係を図3に示す。また同様に「台風21号」について図4に示す。こごグラフより、大部分のリツイートは上位ユーザに偏っていることが分かる。

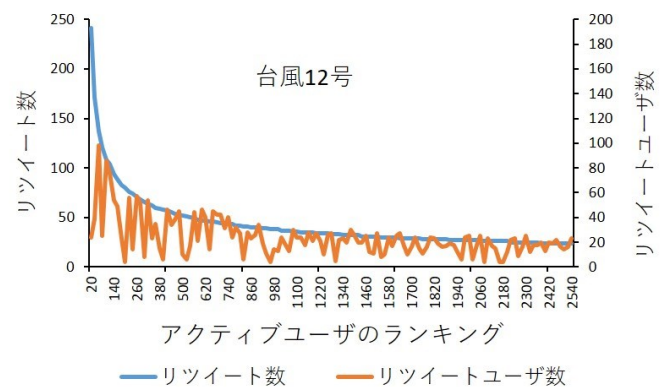


図3 台風12号のアクティブユーザのランキング

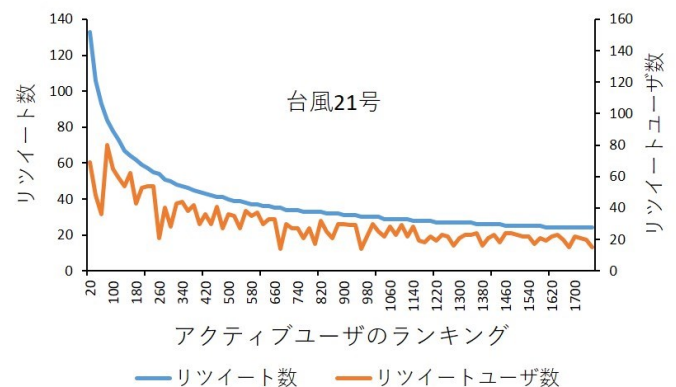


図4 台風21号のアクティブユーザのランキング

4.2.3 アクティブユーザの特徴量の抽出

アクティブユーザのリツイート傾向は、ソースとそのリツイート頻度で近似できる。そこで、ソースからのリツイート頻度を素性にして、アクティブユーザの特徴ベクトルを生成する。ベクトルのサイズはソース数であり、例えば「台風12号」の場合は509次元である（表4）。

スコア上位の投稿者は、概ねソースとして選択されているが、スコアが下位になるにつれて選択率は低下する。例えば「台風12号」のソースの内、最下位（509番目）のスコアランキングは3,449位である。一方、このスコアランキングで7位の投稿者がソースとして選択されなかった。この投稿者は「アニメの公式アカウント」であり、アクティブユーザからのリツイート条件を満足しないためにソー

スから除外されている。

4.3 アクティブユーザのグループ化と評価

4.3.1 クラスタリング

特徴ベクトルのコサイン類似度を用いてアクティブユーザをグループ化する。クラスタリング手法としては、既存の非階層型クラスタリング手法である k 平均法(k-means clustering) を使用する。今回の実験では、グループの特徴を目視で観察するため、1 グループあたり 200 ユーザ程度になるようにクラスタ数を設定した。そこで実験では、クラスタ数を 20 (k=20) に設定してアクティブユーザをグループ化した。その結果を表 5 に示す。

表 5 クラスタリング結果

クラスタ番号	台風 12 号	大阪地震	台風 21 号	北海道地震
0	41	157	42	517
1	92	152	95	193
2	124	76	61	221
3	129	119	136	287
4	94	67	124	312
5	181	93	112	300
6	137	150	51	620
7	269	43	67	145
8	198	113	68	246
9	90	102	104	333
10	37	82	70	377
11	42	339	22	170
12	185	202	206	714
13	369	92	26	138
14	60	149	131	434
15	96	65	63	245
16	46	74	107	150
17	44	61	70	225
18	121	51	50	121
19	91	79	74	199
未分類	101	5	92	28

4.3.2 グループの評価

アクティブユーザのグループを、災害情報の拡散という視点で評価する。具体的には、「緊急性」に関連したツイートに対してのリツイート頻度（反応状況）で評価する。この場合、リツイートごとに緊急性の有無を判定するため、リツイート本文から特徴量を抽出する必要がある。今回は、リツイートの特徴量を、頻出語の有無を素性とする特徴ベクトルで表現する。

この特徴量を用いて学習モデル（分類器）を作成し、2 クラス分類を実施する。この手法として、Support Vector Machine(SVM)を使用する。SVM は 2 クラス分類器として提案されたものであるが、多クラス分類問題にも適用できることからよく使用されている[11]。本研究では、緊急性の有無を判定するための 2 クラス分類器として使用する。具体的には、「台風 12 号」の災害ツイートから「緊急性あり」と「緊急性なし」の災害ツイートを目視で抽出し、学習データを作成する。この学習データから SVM 学習モデルを生成して、他の災害ツイートや災害リツイートの緊急

性を評価する。

(1) リツイートの特徴量抽出

災害ツイートとリツイートは形態素解析器 (MeCab) で品詞分析し、4 品詞 (名詞、動詞、形容詞、副詞) の出現頻度を抽出してデータベースに保存している。特徴ベクトルの素性は、「台風 12 号」の災害ツイートにおいて、よく使用されていた頻出語の上位 1,000 語を適用する。

(2) SVM 学習モデルの生成

「台風 12 号」から「緊急性あり」の災害ツイートを 500 件、「緊急性なし」の災害ツイートを 500 件、目視で抽出して SVM 学習モデルを生成した。この学習モデルの生成と評価には LIBSVM[12]を使用している。

(3) グループの評価

SVM 学習モデルを用いて災害リツイートを評価した結果、緊急性が高いと判断された災害リツイートの件数を表 6 に示す。この表は、災害ごとに各グループのリツイート件数を集計したものである。また、このリツイートを行ったアクティブユーザ数を表 7 に示す。ここで、表 5 と表 7 比較すると、グループメンバの多数が緊急性の高い災害ツイートをリツイートしていることが分かる。しかし、メンバの半数程度しかリツイートしていないグループもあるため、このようなグループは除外対象になりうる。

表 6 緊急性の高い災害リツイートの件数

クラスタ番号	台風 12 号	大阪地震	台風 21 号	北海道地震
0	95	862	655	3088
1	132	372	1230	1570
2	384	333	452	2267
3	2017	633	1207	1042
4	892	749	698	2600
5	935	604	949	2733
6	2504	914	345	3099
7	6894	311	329	1380
8	2971	443	603	2071
9	980	332	954	1626
10	147	347	1318	3910
11	1602	3480	2	1244
12	2858	1587	2302	6714
13	6811	609	51	1217
14	1013	841	238	3896
15	149	538	544	1676
16	565	433	2251	503
17	301	776	139	1943
18	1544	399	405	1427
19	1399	487	1156	1562

拡散グループを選択する場合、表 6 と表 7 の数値が指標となるが、グループの興味領域を意識する場合には、トレンドワードなどを抽出して判定する必要がある。例えば、各グループのツイートを目視で確認すると、特徴的なグループとして、「台風 12 号」のグループ 5、「大阪地震」のグループ 3、「北海道地震」のグループ 6 がある。これらのグループは政治的な話題が多く、批判的な内容のツイートが多い。また、「台風 12 号」のグループ 1 は、アニメ系の公

式アカウントからのリツイートが多かった。しかし、先に述べたアクティブユーザの条件（1日あたり24リツイートツイート以上）をクリアしているユーザの多くは、災害ツイートもリツイートしていた。

表 7 緊急性の高い災害リツイートのユーザ数

クラス 番号	台風 12号	大阪 地震	台風 21号	北海道 地震
0	24	157	42	514
1	58	130	94	192
2	111	74	61	219
3	128	117	136	275
4	90	67	118	308
5	174	93	111	300
6	137	148	44	611
7	268	43	64	144
8	180	113	64	244
9	88	100	104	323
10	21	80	68	377
11	40	339	2	169
12	185	202	202	711
13	366	92	17	138
14	59	149	105	425
15	45	64	59	243
16	42	73	107	141
17	43	61	57	225
18	117	50	49	121
19	89	79	74	199

4.4 グループの持続性の評価

災害ツイートを活発に拡散していたグループが、その後の災害においても活動が継続しているのであれば、持続性の高いアクティブユーザのグループとして評価できる。そこで本稿では、「台風 12 号」の各グループに注目し、アクティブユーザの継続状況について評価する。

4.4.1 アクティブユーザの継続性

「台風 21 号」の災害期間と「台風 12 号」のアクティブユーザの活動状況を図 5 に示す。グラフの横軸は、「台風 12 号」において、緊急性の高い災害リツイート数（表 6）のグループランキングを示している。例えば、ランキング 1 位はグループ 7 であり、2 位はグループ 13 である。縦軸は当該グループのアクティブユーザ数とその維持率を示している。ランキング上位のグループは、アクティブユーザが多いため、「台風 21 号」においても多数のアクティブユーザが継続して活動している。維持率は 30%程度であり、一定の継続性は確認できたといえる。維持率が 30%以上の下位グループも存在しており、リツイート頻度が低くても継続的な活動が期待できることを示唆している。

次に、「北海道地震」の災害期間と「台風 12 号」のアクティブユーザの活動状況を図 6 に示す。これは災害の種類が異なる場合の事例である。アクティブユーザの維持率は 40%以上であり、「台風 21 号」よりも高い。災害までの間隔は、両者とも 1 か月であるため、期間が経過したことによる影響は同じである。地震は突然発生するため、情緒的訴求力が強く、アクティブユーザの維持率が高い。またランキング下位のグループの維持率が高いことも特徴である。

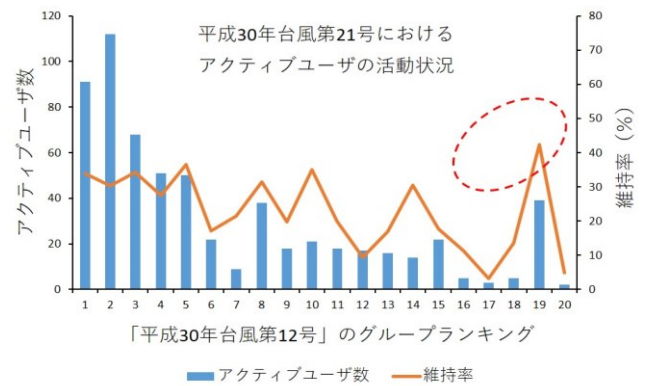


図 5 「台風 21 号」でのアクティブユーザの維持率

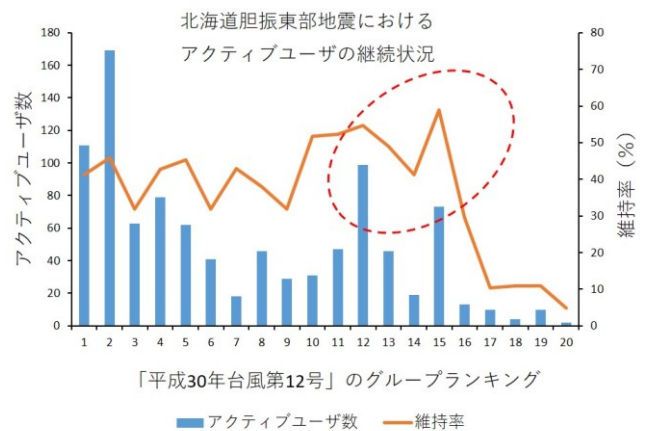


図 6 「北海道地震」でのアクティブユーザの維持率

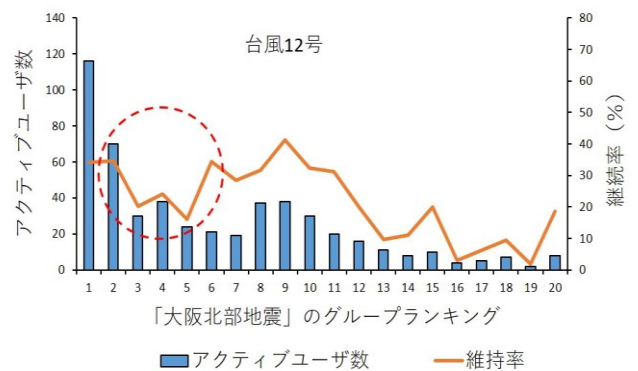


図 7 「大阪地震」のアクティブユーザの維持率

最後に「大阪地震」の災害期間と「台風 12 号」のアクティブユーザの活動状況を図 7 に示す。「大阪地震」の方が先に発生しているため、「大阪地震」のアクティブユーザの維持率を示している。この事例は、地震から台風への移行であるため、訴求力が弱く、上位ユーザの維持が低い点に特徴がある。今後、グループのツイートを目視で精査し、興味領域の詳細を把握して評価する必要がある。

4.4.2 アクティブユーザのグループ遷移

各グループの共通ユーザ数を確認することにより、他の災害におけるアクティブユーザの分散状況を把握する。「台風 12 号」と他の災害における共通ユーザ数の一部を表 8、表 9、表 10 に示す。この表には、リツイート数の上位 5 グループの共通ユーザ数を掲載している。実験ではすべてのグループについて共通ユーザ数を確認しており、リツイート数の上位グループ同士は共通ユーザが多い。特に、10 人以上の共通ユーザ数を持つグループは、一方がランキング 5 位以内に入っている。共通ユーザが 20 人以上のグループは 13 組存在しており、この内、ランキング 5 位以内のグループ同士は 8 組である。

共通ユーザが多くてランキングの低いグループは、災害にあまり関係のない特殊な話題に興味を持つグループであることが多い。例えば、共通ユーザが 30 人以上の 4 組において、唯一ランキング 5 位以内に入っていないグループのツイートを確認したところ、災害における行政の対応を批判するような政治的なツイートが多くを占めていた。これらのツイートは、SVM 学習モデルの評価により、緊急性の低いツイートと判断されているため、リツイートランキングで下位になっている。

表 8 アクティブユーザのグループ遷移 (台風 21 号)

		台風 12 号上位グループ				
		7	13	8	6	12
台風 21 号 上位 グループ	12	7	13	34	4	12
	16	23	4	0	0	2
	10	2	12	1	2	1
	1	5	14	0	3	9
	3	9	26	2	4	2

表 9 アクティブユーザのグループ遷移 (北海道地震)

		台風 12 号上位グループ				
		7	13	8	12	6
北海道 地震上位 グループ	12	22	16	8	30	15
	10	12	13	2	5	2
	14	12	48	2	1	3
	6	1	7	0	1	2
	0	10	7	4	1	3

表 10 アクティブユーザのグループ遷移 (大阪地震)

		台風 12 号上位グループ				
		7	13	8	6	12
大阪地震 上位 グループ	11	18	31	3	16	21
	12	11	16	6	4	2
	6	4	4	2	4	2
	0	2	2	1	0	1
	14	7	6	1	4	1

災害リツイートが多いグループのアクティブユーザは、他の災害においても同様のグループに所属する傾向があることを確認した。災害マーカーは、災害以外でも使用されるワードであるため、適切にフィルタリングしないとアニメや政治等の話題でグループのランキングが上位になる。今回は SVM 学習モデルを用いて「緊急性の有無」を判定しているため、特殊な興味領域を持つグループを除外できて

いる。リツイートのトレンドワードを判定して除外する方法もあるので、これらの併用も含めて検討する予定である。

5. おわりに

本稿では、リツイート傾向の類似性を利用してアクティブユーザをグループ化し、グループの特徴や持続性、および緊急性の高い災害ツイートに対する反応状況等を評価した。その結果、アクティブユーザについては、一定の継続性があることを確認した。また、リツイート傾向が類似しているアクティブユーザは、別の災害においても類似のグループを形成していることを確認した。

ある災害で活発に活動したアクティブユーザは、別の災害においても、災害の種類や日時に関係なく、一定の反応を示している。また、災害情報のリツイート頻度が高いグループのメンバは、その後の災害でも同様のグループに所属する傾向にあることを示しており、一部のグループについては持続性を確認できた。特殊な興味領域を持つグループも存在しているため、これらのフィルタリング手法も含めて検討していく予定である。

謝辞 本研究の一部は文部科学省科研費 (19K12239) による研究支援を受けて実施している。

参考文献

- [1] 岸浩稔, 中西航, 位置情報付き tweet による被害状況の逐次把握可能性の検討, 生産研究, 2013, Vol. 65, No. 4, pp. 179-182.
- [2] 榎本甫, 桑野将司, 小池淳司, 災害時のソーシャルメディアと帰宅行動の関連性分析, 土木学会論文集 D3 (土木計画学), 2014, Vol. 70, No. 1, pp. 102-112.
- [3] 横田尚己, 山田圭二郎, 熊本地震のつぶやきに見る感情極性値の時空間解析, 都市計画論文集, 2017, Vol. 52, No. 3, pp. 1081-1087.
- [4] 横田尚己, 山田圭二郎, 阿部倫之, つぶやきと感情極性の関係に着目した災害時の被害情報発見手法の検討, 第 56 回土木計画学研究発表会講演集, 2017, Vol. 56, P8-1-P8-6.
- [5] “DISAANA-対災害 SNS 情報分析システム【リアルタイム版】”, <http://disaana.jp/rtime/search4pc.jsp>, (参照 2019-08-20).
- [6] 大竹清敬, AI を用いた SNS 上の災害関連情報を要約・分析するシステムの利活用, ITU ジャーナル, 2017, Vol 47, No. 7, pp. 3-7.
- [7] “災害対応における SNS 活用ガイドブック 平成 29 年 3 月 内閣官房 情報通信技術 (IT) 総合戦略室”, https://www.kantei.go.jp/jp/singi/it2/senmon_bunka/pdf/h2903guid ebook.pdf, (参照 2019-08-20).
- [8] 宮部真衣, 梅島彩奈, 灘本明代, 荒牧英治, マイクロブログにおける流言の特徴分析, 情報処理学会論文誌, 2013, Vol. 54, No. 1, pp. 223-236.
- [9] 村上明子, 鈴木秀幸, Twitter での Retweet 情報を利用した情報拡散予測, 人工知能学会第 26 回全国大会, 2012, 4C1-R-6-9.
- [10] 渡辺研斗, 鍋島啓太, 水野淳太, 岡崎直観, 乾健太郎, Twitter における誤情報の拡散収束過程の可視化, 情報処理学会第 75 回全国大会, 2013, pp. 1-657 - 658.
- [11] 竹内一郎, 鳥山昌幸, サポートベクトルマシン, 講談社, 2016.
- [12] “LIBSVM -- A Library for Support Vector Machines”, <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>, (参照 2019-08-20)