

科学研究費助成事業 研究成果報告書

平成 28 年 5 月 16 日現在

機関番号：33302

研究種目：挑戦的萌芽研究

研究期間：2013～2015

課題番号：25540142

研究課題名(和文) ソーシャルメディアにおけるデマ訂正情報の拡散を支援するシステムの開発

研究課題名(英文) A Method and Its Support System for diffusing Rumors Correction Information in Social Media

研究代表者

阿部 倫之 (Abe, Noriyuki)

金沢工業大学・工学部・准教授

研究者番号：60231966

交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)：近年、ソーシャルメディアにおいてデマ情報が拡散されており、社会問題に発展している。特にtwitterでは、リツイート機能によって迅速かつ広範囲にデマ情報が拡散されが、デマを打ち消す訂正ツイートの拡散力は弱い。そこで、デマ訂正ツイートの拡散を支援するための手法とその支援システムが求められている。本研究では、動的リツイート関係に基づいた支持度と貢献度による拡散手法を提案する。実験結果より、本手法を用いることによって訂正リツイートを広範囲に拡散できる可能性を示した。

研究成果の概要(英文)：Recently, rumors information have been diffused in social media. On twitter, the rumors information is spread quickly and extensively by using the retweet function. However, there is a problem that the rumors correction information is spread slowly and narrowly by the lazy retweet others. The support system becomes important to be diffused eagerly the rumors correction retweet. This paper proposes a method of using the support score and contribution score based on the dynamic retweet relations. In the experimental result, this method showed the ability to diffuse extensively correction retweets.

研究分野：Web情報学

キーワード：リツイート フォロー 拡散 中心ユーザ ツイート Webマイニング 発信源 デマ

1. 研究開始当初の背景

近年、ソーシャルメディアから発信されるデマ情報が大きな社会問題となっている。特に匿名性の高い Twitter から発信されるデマツイートはリツイートの仕組みによって急激に拡散される一方、熱狂に水を差すデマ訂正ツイートは拡散されにくい。したがって、デマ訂正ツイートの拡散を支援する仕組みが求められていた。

先行研究の多くは、デマ訂正ツイートを発見する方法に主眼が置かれており、ツイートの拡散に貢献しているユーザの発見手法等については検討されていない。デマ訂正ツイートの拡散力はデマよりも低いため、ツイート拡散の中心的ユーザを発見する手法を明らかにする必要がある。

2. 研究の目的

本研究では、観察対象を Twitter に絞り、デマ訂正ツイートの拡散を支援するための仕組みを明らかにする。ここでデマ訂正ツイートの発見と、そのツイートからデマ部分を抽出する手法については先行研究で検討しているため、本研究ではツイートを拡散している中心的ユーザを発見するための手法を検討し、フィールドテストを実施して提案手法の評価を試みる。

3. 研究の方法

(1) リツイートの収集

テレビ番組の視聴ツイートを題材にして、中心ユーザの発見手法とその有効性を確認する。テレビ番組の視聴ツイートは Twitter streaming API を用いてリアルタイムに収集し、リツイートを分離する。ユーザの視聴傾向に与えるリツイートの影響を観察することにより、ノイズツイートとして除外すべきリツイートを把握する。リツイートは、引用元のユーザ(ソースユーザ)に対する関心を表しているため、ノイズツイートが適切に除外できれば、リツイート回数で関心の強さを測ることができる。

(2) 中心ユーザの発見手法

支持度と貢献度の測定
リツイートが発言されたとき、発信者とソースユーザ情報を抽出し、データベースに保存する。このとき、リツイートユーザにはソースユーザの支持度を記録する。またソースユーザにはリツイートユーザの貢献度を記録する。支持度と貢献度はリツイート回数であり、過去の数値を割引いた上で累積する。これにより、時間の経過に伴って変化していくリツイートユーザとソースユーザ間の関連性の強さを適切に測ることができる。

リツイートパターンの抽出

支持度を用いてリツイートパターンを抽出する。これは支持度のスコア列であり、各スコアはリツイートしたソースユーザに対応

している。

リツイートユーザのグループ化
リツイートパターンの類似しているユーザを k-平均法を用いてグループ化する。

ツイート発信源の発見
グループのメンバーがよく支持しているソースユーザを貢献度のスコアでランキングすることにより、グループ内で指示されている代表的なツイート発信源を把握する。

(3) 実験と評価

本手法の実験プログラムを Java 言語で実装し、リアルタイムに収集した視聴リツイートから中心的ユーザの発見を試みる。なお実験評価には視聴率の高い人気番組を選定する。

4. 研究成果

ソースユーザに対する支持度の類似性に基づいてリツイートユーザをグループ化し、話題の発言と拡散に貢献している中心的ユーザを把握する手法を示す。またテレビ番組の視聴ツイートを題材にして、提案手法の有効性を評価したのでその結果を示す。

最初にツイートを利用した TV 視聴傾向の評価手法と実験結果を示すことで、番組放送中のツイートやリツイートの傾向などを中心に評価対象の特徴を把握する。その後、リツイートユーザの支持度や貢献度を累積していく方法、リツイートパターンによる類似性の判定方法などを示し、最後に実験結果と今後の展開について述べる。

(1) ツイートによるテレビ視聴傾向の把握

実験システムの概要

本実験システムでは、ハッシュタグと番組セッションを設定すると、視聴ツイートをリアルタイムに取得してツイートとユーザ情報を保存する。このとき、ハッシュタグとよく共起しているワードを視聴マーカとして抽出する。視聴マーカは、視聴ツイートの絞り込みやハッシュタグを含まない視聴ツイートの取得に利用する。

ここで、テレビ番組の放送時間帯に前後 30 分程度を加えた期間を「番組セッション」と定義する。この番組セッション中のツイートの内、ハッシュタグを含むツイートを反響ツイートと呼ぶ。

実験システムの構成を図 1 に示す。Twitter Streaming API のフィルタ機能を利用して視聴ツイートをリアルタイムに取得する。トラックワードにはハッシュタグを指定しており、ツイートは形態素解析器 (MeCab) を経由してデータベース (PostgreSQL) に保存する。視聴ツイートを取得するプログラムは、Twitter4J の Java クラスライブラリを使用して実装している。この実験システムでは、取得した全ツイートオブジェクトを JSON 形式で保持しており、そこからユーザオブジェ

クトを分離して視聴ユーザを獲得している。リツイートオブジェクトの場合には、リツイートユーザとリツイート元のユーザを分離して保持することで、動的なリツイート関係を構築できるように実装している。

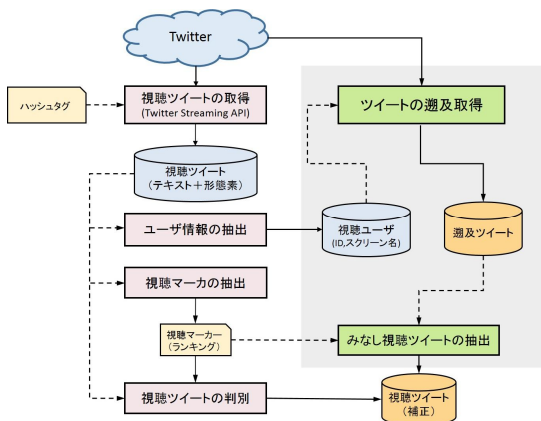


図1 システム構成

最近のキーワードを認識できるようにするために、MeCabの辞書に「はてなキーワード」約32万語を組み込んで使用している。トレンドワードが明らかな場合には強制抽出語として登録し、必要に応じて除外ワードを設定している。視聴マーカは名詞（固有名詞）を中心に抽出しており、ハッシュタグとよく共起しているワードをスコア化してランキング上位の視聴マーカを選択する。ハッシュタグとの共起スコアを自己相互情報量PMIで算出し、ランキング上位のワードを当該番組の視聴マーカとする。この視聴マーカを利用してツイートを絞り込むことができるが、ツイートの網羅性も低下するため工夫が必要である。本実験システムでは、視聴マーカのみを含むツイートを番組セッション終了後に遡及取得することで視聴ツイートの網羅性を維持している。なお、ドラマのような連続性の高い番組では、視聴マーカを次のハッシュタグとして直接利用することも考えられるが、視聴マーカを目視で精査する必要があるため導入していない。

視聴者の網羅性を維持するために、視聴ユーザの番組セッション中のツイートを過去に遡って取得する。取得したツイートの内、視聴マーカを含むツイートを視聴ツイートに追加する。この視聴ツイートにはハッシュタグを含まない視聴ツイートが含まれており、これを「みなし視聴ツイート」として視聴傾向の評価に反映させる。

(2) テレビ視聴傾向の評価事例

年末年始放送されたテレビ番組について、視聴ツイートの発言状況を表1と表2に示す。番組セッションは、放送時間の前後10分とした。表1のツイート数はリツイートを除いてカウントしたものである。また表2のツイートユーザ数はリツイートユーザを除いてカウントしている。この表より、リツイート

ユーザ数も含めてリツイートの比率は高いことが分かる。「相棒」以外の4番組は、年末年始恒例の人気番組であり、ツイートの規模も大きい。「相棒」もスペシャルではあるものの、毎週放送されている連続性の高い番組であるため、放送時間の延長分のみツイート数が増加している。

表1 視聴ツイート数

番組名	ハッシュタグ	ツイート数	リツイート数
第66回 紅白歌合戦	紅白, 紅白歌合戦	98,777	234,757
ガキの使いやあらへんで!	ガキ使, ガキの使い	214,776	84,094
相棒 元旦スペシャル	相棒	13,546	7,297
箱根駅伝 往路	駅伝, 箱根駅伝	55,770	57,943
箱根駅伝 復路	駅伝, 箱根駅伝	50,180	60,548

表2 視聴ユーザ数

番組名	ツイートユーザ数	リツイートユーザ数	リツイート元(ソース)ユーザ数
第66回 紅白歌合戦	41,391	140,484	6,615
ガキの使いやあらへんで!	48,988	58,682	5,028
相棒 元旦スペシャル	3,155	2,714	600
箱根駅伝 往路	13,186	33,703	2,564
箱根駅伝 復路	12,742	30,132	2,896

次に「第66回紅白歌合戦」について、ツイート数の毎分推移を図2に示す。グラフのツイート数にはリツイート数を含めていない。この視聴傾向から4か所の反応ポイント(AからD)が確認できる。AからCの反応ポイントでは、ツイートの急上昇に伴ってリツイートが発生し、時間の経過に伴ってリツイートの連鎖が緩やかに収束している。この状況は十分に予想できるリツイートの反応であるが、Dの反応ポイントは大きく異なっている。ツイートが収束傾向にある状況での大きなリツイート反応であるため、発言内容の精査を実施した。その結果、20:42から20:47において、「まだだ、まだ終わらんよ! #紅白歌合戦」というツイートが約5500件リツイートされていることが判明した。番組セッションでは、13,403件のリツイートを観測している。このリツイートを除外した結果を図3に示す。このグラフでは、図2の反応ポイントDが完全に除外されているのが分かる。このようなリツイートは、視聴傾向の把握という観点からはノイズに分類することも考えられるが、数分間の統一行動に参加している点を考慮すると、リツイートユーザの類似性を測るための指標としては積極的に利用することを検討したい。

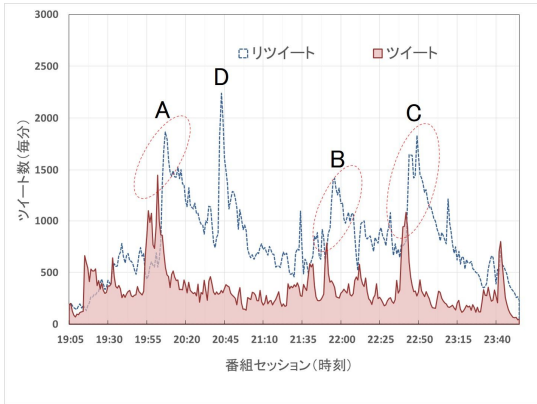


図2 紅白歌合戦の視聴傾向（除外なし）

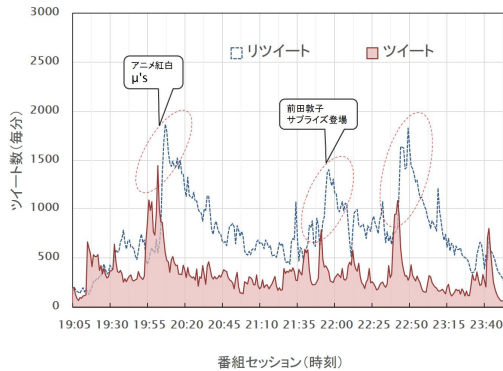


図3 紅白歌合戦の視聴傾向（除外あり）

最後に、視聴マーカを利用して小さな反応ポイントを発見する事例を示す。ハッシュタグ「紅白」や「紅白歌合戦」とよく共起していた上位のワードは「ニコファーレ、森進一、向井秀徳、津軽海峡、天童よしみ、ギンギラギン、高橋真梨子、スクリーム、五木ひろし、桃色吐息、ラガーマン」などであった。これらの視聴マーカを用いて反応ポイントを絞り込んだ結果を図4に示す。目視による視聴マーカの精査も実施しているため、最大の反応ポイントでも300ツイート程度までに絞り込んでいる。網羅性は大きく低下するが、小さな反応ポイントを把握するための手法としては有効である。リツイートを含めると視聴マーカのノイズが増加するため、リツイートを除外してワードの共起スコアを算出している。

本研究では、ソースユーザに対する貢献度の類似性に基づいてリツイートユーザをグループ化する方法を検討している。この視聴マーカを含むツイート数でリツイートユーザとソースユーザをランキングすることにより、分類対象のユーザを絞り込むことができる。なお、番組内でのリツイート数に基づいてソースユーザに対する貢献度を算出するため、ユーザのグループは番組ごとに生成する。

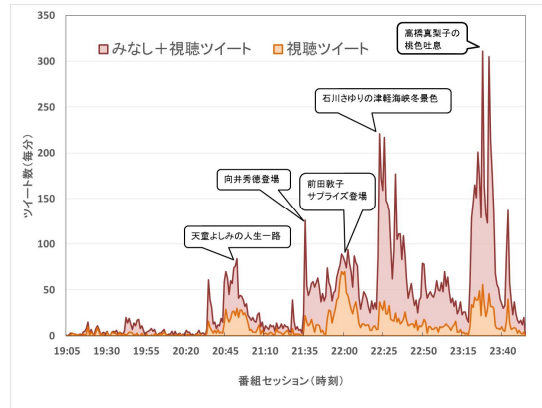


図4 紅白歌合戦の視聴傾向（補正）

(3) 中心的ユーザの発見

ソースユーザと関連性の高いリツイートユーザのグループを発見し、中心ユーザとして把握するための手法を示す。また、テレビ番組の視聴ツイートを収集して実験評価を行ったので、その結果について述べる。

支持度と貢献度

支持度と貢献度の関係を図5に示す。リツイートが発生したとき、ソースユーザにリツイートユーザの貢献度を記録する。また、リツイートユーザにはソースユーザの支持度を記録する。支持度と貢献度は累積値で測るため、前回の更新時点からのリツイート回数を加算する。なお、ツイートの賞味期限の短さに対応するため、番組セッション開始前に累積値の割引処理を実施し、番組セッション終了後にリツイート回数を加算する。これにより、リツイート無しのユーザについては、前回までの支持度と貢献度の割引処理のみが実施される。割引率を1/2に設定すると、4話連続でリツイートが無かった場合には、当該ユーザの支持度と貢献度は1/16になる。時間の経過に伴って過去の寄与度を割引く手法は、機械学習の分野でよく用いられている。

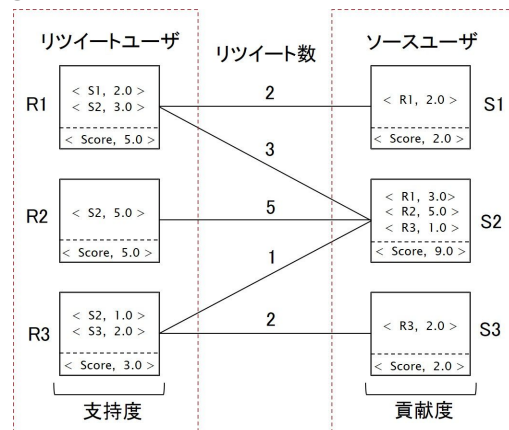


図5 支持度と貢献度の関係

ユーザの特徴抽出

ユーザの類似性を判定するために、特徴抽出を行う。本手法では、リツイートユーザの

特徴をソースユーザの支持度で表現し、ソースユーザの特徴をリツイートユーザの貢献度で表現する。図5の例では、リツイートユーザの特徴は (S1, S2, S3) , 各特徴量の数値列は、特徴ベクトル
 $R1 = (2.0, 3.0, 0.0)$,
 $R2 = (0.0, 5.0, 0.0)$,
 $R3 = (0.0, 1.0, 2.0)$
で表現する。同様にソースユーザの特徴は (R1, R2, R3) , 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。この特徴ベクトルのサイズは、特徴表現の対象ユーザ数で決まるため、例えば、紅白歌合戦のリツイートユーザの特徴ベクトルはソースユーザ数 (6,615) , ソースユーザの特徴ベクトルはリツイートユーザ数 (140,484) でサイズが確定する。なお予備実験により、特徴表現の対象範囲が広がると相関性が低くなることを確認しているため工夫が必要である。本手法では、
・支持度および貢献度のスコア、
・視聴マーカを含むリツイート数、
を利用してユーザランキングを実施することにより、ランキング上位のユーザで特徴ベクトルを生成する。

中心的ユーザの把握

特徴ベクトルを利用した分類手法としては、既存の k-平均法を用いる。類似度の計算には、コサイン類似度を用いており、特徴ベクトルは、ノルムが1になるように正規化して使用する。クラスタリングでは、主にリツイートユーザのグループ化を実施する。中心的ユーザの把握においては、各グループにおいて上位のリツイートユーザを「拡散ユーザ」とする。また拡散ユーザが共通に支持しているソースユーザの内、ランキング上位のソースユーザを「発信源」とする。この拡散ユーザと発信源を併せて当該グループの「中心的ユーザ」と推定する。したがって、中心的ユーザはグループごとに存在する。中心的ユーザの把握では、ソースユーザについてもグループ化を実施することで、「発信源」が所属するグループのメンバーも中心的ユーザに関連したソースユーザとして把握できる可能性がある。

(4)実験と評価

年末のテレビ番組「第 66 回紅白歌合戦」と「ガキの使いやあらへんで！」の視聴ツイートを利用して評価を実施した。1 回完結の番組であるため、支持度と貢献度の割引処理は発生しない。クラスタリング等の実験評価システムは Java で実装している。まず支持度と貢献度でランキングを実施し、上位 100 人のユーザを特徴表現の対象ユーザとした。これにより、特徴ベクトルサイズは 100 となる。リツイートユーザのランキング結果を表

3 と表 4 に示す。支持度はリツイートしたソースユーザ数、フォロー率はソースユーザをフレンドしている割合である。このように直接フォローしていないソースユーザをリツイートする傾向が高いため、動的なリツイート関係で発信源と関連性を把握することは重要である。次にリツイートユーザの上位 1000 人、ソースユーザの上位 100 人の特徴ベクトル生成し、k-平均法でグループ化を実施した。各グループの代表ベクトルは所属メンバーの重心で算出している。グループ数は 10 (k=10) である。

表 3 拡散ユーザ (紅白歌合戦)

ランク	略称	支持度	支持数	フォロー率
1	Ishi...	220	194	0.025
2	Take...	192	3	0.667
3	Kyu...	167	7	0.429
4	nuru...	147	12	0.75
5	roku...	138	116	0.009
6	HAN...	133	7	0.285
7	mknh...	113	3	-
8	mom...	110	4	0.75
9	wby...	102	20	0.6
10	asnh...	100	18	0.389

表 4 拡散ユーザ (ガキの使い)

ランク	略称	支持度	支持数	フォロー率
1	EKI...	307	278	0.014
2	TOK...	299	265	0.004
3	Caro...	159	120	0.058
4	Bud...	153	104	-
5	ora1...	133	95	0.0315
6	5k...	122	9	-
7	noah...	101	85	0.047
8	moto...	97	21	0.571
9	Ice...	83	18	0.388
10	div...	82	66	0.03

「紅白歌合戦」の結果を表 5 に、「ガキの使いやあらへんで！」の結果を表 6 に示す。各グループに所属している上位メンバーのランク値を主要メンバー欄に記載した。表 5 ではグループ 1 とグループ 9、表 6 ではグループ 2 とグループ 7 に上位のリツイートユーザが集中している。よくリツイートされている上位 100 人のソースユーザに対する支持度で特徴ベクトルを生成しているため、リツイート頻度が高いユーザが同じグループに集まる傾向がある。

発信源の欄には、グループ内の上位ユーザがリツイートしたソースユーザの内、貢献度ランキング 10 位以内のソースユーザを記載した。数字はソースユーザのランク値を表している。グループ内の上位ユーザから順番にピックアップしているため、ソースユーザのランク順には並んでいない。被リツイート頻度 1 位のソースユーザが多いのは明らかであるが、比較的少数グループについては、第 4 位や第 7 位のソースユーザに支持が集まっている。特に表 5 の「紅白歌合戦」では、

第4位のソースユーザに対する支持が大きい。目視での調査の結果、ORICON STYLEの公式アカウントであることが分かった。今回は視聴マーカを用いないでランキングを実施しているが、「森進一」、「ギンギラギン」など具体的な場面を連想させる視聴マーカも多いため、今後、実験評価を継続して進めていきたい。

表5 グループの主要メンバー(紅白歌合戦)

グループ	主要メンバー (ランク)	総スコア	人数	発信源 (ランク)
0	57, 70, 82, 103, 119	1743	93	1, 3, 4, 2, 5
1	2, 4, 5, 6, 7, 9, 11, 12	8961	371	1, 3, 4, 6
2	10, 43, 59, 69, 140	972	46	1, 3, 2, 4, 6
3	78, 88, 127, 289, 299	480	27	4, 1, 3,
4	16, 98, 120, 141, 147	1227	66	1, 3, 2, 5, 6
5	75, 83, 196, 421	520	31	1, 4, 7
6	128, 195, 300, 421	347	22	2, 6, 3, 1
7	211, 244, 252, 265	389	24	1, 4, 6, 3, 5
8	76, 97, 120, 138, 145	622	34	1, 4, 2, 9, 6
9	1, 3, 8, 14, 17, 18, 19	6955	262	1, 3, 4

表6 グループの主要メンバー(ガキの使い)

グループ	主要メンバー (ランク)	総スコア	人数	発信源 (ランク)
0	15, 17, 28, 39, 42	977	75	1, 4, 2, 9
1	21, 35, 56, 58, 61	2290	236	2, 9, 1, 5
2	1, 2, 3, 4, 6, 16, 24	1829	63	1, 2, 7, 3, 9
3	34, 41, 44, 55, 81	590	60	1, 2, 4, 6, 7
4	5, 7, 23, 25, 27, 32	1362	108	1, 2, 3, 4, 7
5	8, 22, 38, 51, 65, 68	903	73	1, 4, 3, 6, 2
6	98, 107, 117, 202	287	32	9, 4, 5, 7, 2
7	9, 10, 11, 12, 13, 14	1310	67	1, 2, 4, 6, 9
8	36, 70, 152, 176, 196	469	54	1, 3, 2, 7, 4
9	19, 54, 66, 69, 76	1659	182	1, 4, 3, 7, 2

(4)まとめ

本研究では、ソースユーザに対する貢献度の類似性に基づいてリツイートユーザをグループ化し、話題の発言と拡散に貢献している中心的ユーザを把握する手法について検討した。またテレビ番組の視聴ツイートを題材にして、番組放送中における視聴ツイートとリツイートの発言傾向を示し、リツイートが視聴傾向に与える影響について考察した。本手法を用いてリツイーターユーザをグループ化し、支持度の高いソースユーザを当該グループのツイート発信源として把握できる可能性を実験的に示した。

今回はランキング上位のユーザを対象に実験評価を実施したが、今後は中位から下位にランキングされているユーザを対象に実験と評価を進める予定である。中位以下のユーザはリツイート数も少ないため、視聴ツイートを利用する場合には、視聴マーカによる対象ユーザの絞り込みが必要と考えている。ドラマなどの連続性の高い番組において、形成されたグループの継続性なども含めて検討を進めたい。

話題性の高いツイートがデマ訂正ツイートの場合には、デマ訂正情報を拡散しようとしているグループとその発信源が把握できるため、当該グループの情報を継続的に発信することにより、訂正情報の拡散支援が可能である。

5. 主な発表論文等

〔学会発表〕(計 4 件)

阿部 倫之, 米谷 優也, テレビ番組の視聴ツイートを利用した中心ユーザの発見支援, 情報処理学会第136回情報システムと社会環境研究発表会, 2016年6月, 國學院大学渋谷キャンパス(東京)

平岡 照久, 中村 和裕, 荒木 伸也, 阿部 倫之, ツイートを利用したテレビ視聴傾向の評価手法, 第77回情報処理学会全国大会, 2015年3月, 京都大学吉田キャンパス(京都)

荒木 信也, 阿部 倫之, 服部 進実, 反響ツイートを利用したテレビ番組の評判システムに関する一考察, 第76回情報処理学会全国大会, 2014年3月, 東京電機大学東京千住キャンパス(東京)

荒木 信也, 阿部 倫之, 服部 進実, 反響ツイートの拡散情報を利用したテレビ視聴傾向の予測に関する一考察, 第76回情報処理学会全国大会, 2014年3月, 東京電機大学東京千住キャンパス(東京)

6. 研究組織

(1)研究代表者

阿部 倫之 (NORIYUKI, Abe)

金沢工業大学・工学部・准教授

研究者番号: 60231966