

# マイクロブログからの社会的影響力を持つ情報カスケードの早期検知

川本 貴史<sup>†</sup> 豊田 正史<sup>††</sup> 吉永 直樹<sup>††,†††</sup>

<sup>†</sup> 東京大学大学院情報理工学系研究科 〒113-0033 東京都文京区本郷 7-3-1

<sup>††</sup> 東京大学生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1

<sup>†††</sup> 国立研究開発法人 情報通信研究機構 〒184-8795 東京都小金井市貫井北町 4-2-1

E-mail: †{kawamoto,toyoda,ynaga}@tkl.iis.u-tokyo.ac.jp

あらまし マイクロブログではユーザ間での情報共有が連鎖することによる情報カスケードがしばしば観測され、その中には商品の回収につながり得る風評の拡散や災害への対処方法の共有など社会的影響力を持つものも存在する。このような情報カスケードに迅速に対応するために、本稿では Twitter における情報カスケードから社会的影響力を持つ情報カスケードを早期に検知するというタスクを新たに提案し、これを教師あり学習に基づく分類器により解く手法を提案する。我々はまず、複数人によるアノテーションを行い提案するタスクの妥当性の検証を行った。その後、本文のテキスト情報や情報カスケードのグラフ構造、情報カスケード毎のユーザ分布、情報カスケードに対してメンションを行ったユーザやそのメンションのテキスト特徴量といった多様な特徴量を設計し、カスケードが大規模に成長するか分類を行ったあとで社会的影響力の有無を分類する二段階の分類手順を踏む方法とそれらを同時に分類する方法の二種類の手法を提案し比較を行った。実験では、実際のツイートデータセットにおいて検知する対象の情報カスケードのサイズ（初期リツイート数）を変化させて社会的影響力を有し 600RT 以上に拡散する情報カスケードを検知する実験を行い、どれだけ早期に検知が可能かを明らかにし、提案した特徴量の有効性を評価した。

キーワード マイクロブログ, Twitter, ソーシャルグラフ, カスケード, 情報伝播, データマイニング

## 1. はじめに

近年、オンライン上でユーザが日常的な情報をリアルタイムに発信する Twitter などのマイクロブログが盛んに用いられており、その上で友人関係をバーチャルに表すソーシャルネットワークが大規模化している。このソーシャルネットワークでは友人間でのコミュニケーションが行われるが、単なる友人間でのやりとりにとどまらず、友人から受け取った情報をさらに他の友人へと発信することが日常的に行われる。このような情報共有が連鎖することによって引き起こされる情報拡散を情報カスケード（以下、誤解なき限りカスケード）という [3]。

近年では、商品を批判する意見が拡散し、商品の回収につながるようなケースや、事故や災害への対処方法を共有する動きなど社会的影響力を持つ情報カスケードも観測されている。特にマイクロブログにおける情報カスケードは、マイクロブログのリアルタイム性の高さによって急速に広がるため、政府、マスメディア、企業などにとって社会的影響力の高い情報カスケードを早期に発見することは風評被害に対する未然の対処、世論動向、報道、商品に対するフィードバックとして重要である。

以上のように、マイクロブログにおける情報カスケードには早期に対応、あるいは認知すべき情報が含まれるが、一方でアフェリエイトリンクへと誘導するスパムやジョーク、有名人の日常のつぶやき、広告など社会的影響力の少ない情報も多い。注目すべき情報カスケードを検知するための研究としては、将来的に拡大する可能性の高い情報カスケードを検知する手法 [7]

や、スパムツイートを検知する手法 [6] があるが、3.4 節で確認するように、前者は広く拡散する情報カスケードが必ずしも社会的影響力を持つとは限らない点で、また後者はスパム以外にも社会的影響力のない情報カスケードが存在する点で、社会的影響力を持つ情報カスケードを検知する上では不十分である。

この問題点に対して我々は、マイクロブログにおける情報カスケードの中から社会的影響力を持つものを検知するタスクを提唱した [20]。この研究では、大規模な情報カスケードを分析し、その分析をもとに社会的影響力を持つ情報カスケードの定義を行っている。またそれに基づいたアノテーションを行い、作成したデータセットを用い教師あり学習によって解く手法を提案した [20]。しかしこの研究には 1) アノテーションが著者一人によって行われておりタスク設定の妥当性、実験結果の信頼性の担保が不十分であるという問題点と、2) 600 回以上 RT が観測された情報カスケードのみを対象にしており、情報カスケードが広まるかどうか、という前問題を解いていないため早期検知にどの程度有効か分からない、という 2 つの問題点が存在する。そこで本研究では 1) アノテーションを複数人で行い、その一致率を元にタスク設定の妥当性を確認し、2) 情報カスケードが拡大化するかという問題と組み合わせることで、社会的影響力を持つ情報カスケードの早期検知という目的に即した問題設定を行い、それを具体的に解く手法を提案する。

実験では、複数人の被験者によるアノテーションにより構築されたデータセットを基に SVM を用いて分類器を学習し、分類器の示す  $F_1$  値によって提案手法の有効性の評価を行った。

## 2. 関連研究

情報カスケードの性質を明らかにすることを目的として、マイクロブログにおけるカスケードの予測に関する研究 [9] [2] [17] [7], マイクロブログにおけるカスケード分類の研究 [16] [5] [11], マイクロブログにおけるスパム検出などが行われている [6] [10]. 以下で、これらのタスクについて、我々が提案するタスクとの関連性を明らかにするとともに、情報カスケードの性質の分析に用いられている特徴量を紹介する.

### 2.1 カスケード予測に関する研究

カスケード予測に関する研究としては、カスケードの成長 [7] を予測する研究に限らず、実際につぶやくユーザを予測する研究 [9] や、ユーザの影響力を定量化する研究 [2] など幅広く存在する. また、対象とするカスケードもミームを Twitter のハッシュタグ [17] や URL [2] [9] とする研究や、あるいは Facebook の投稿拡散機能であるシェアによる情報カスケードを扱う研究 [7] など多岐に渡る.

カスケードの成長を予測する研究 [7] は、社会の中で広く拡散される投稿を当てる問題を解いており、ある意味では社会的影響力の高い投稿を当てているともいえる. しかしながら、1. 節で述べたように、広く拡散される投稿だからといって、必ずしも社会的影響力を持つ投稿とは限らない. そのため、本研究ではこれらの研究のように単に拡散するかどうかだけでは無く、その上社会的影響力を持つかどうかまで分類することを目指す. また、[6] を参考に成長の予測と組み合わせる手法も行い、精度の比較を行う.

### 2.2 ツイートやカスケードの分類に関する研究

マイクロブログにおけるカスケード分類の研究にはトピックによる意味的分類やグラフパターンによる構造的分類の他にツイートの信頼性判定を行うものなどがある. Sriram らはユーザが読むツイートを絞るための手助けとしてツイートを News, Opinions, Deals, Events, Private Messages の 5 つに分類する手法を提案している [16]. その際、分類の特徴量として Bag of words を用いている. また、Ren らは各ツイートに対し Web Forum や質問応答システムなどで一般的なラベル付けである複数の階層的な意味ラベル付けを行うことを提案している [14]. しかし、これらのタスクではツイートの持つ情報の社会的影響力については考慮されていない.

また、Castillo らはあるトピックのツイート集合において News クラス, Chat クラス, 判断できない, のどのクラスに属するかを判定した後、自動でそのトピックの信頼性を判定している [5]. その際の特徴量としてはユーザの特徴量、トピックの特徴量、リツイートの特徴量を用いている. この研究においては、ツイートの信頼性判定に重点が置かれており、前段階の News クラス分類においても、特定の出来事に関するニュースかどうかという観点で判定が行われている. そのため、本研究で考える、情報が社会的に影響を持つかどうかという判断基準とは異なる.

社会的影響力に関係する分類の研究としては、Zheng らがツイートに対して情報の公共性の有無に基づく分類を行っている

[12]. ツイートの内容の公共性に基づき、News, Commercial, Private の 3 種類の区分に分割するという問題に取り組んでいるが、彼らはツイート本文の特徴量のみを用いているが、我々が取り組む社会的影響力の影響力の検知にはテキスト特徴量だけではなく、グラフ特徴量やユーザ特徴量が有効である [20].

Topic Detection and Tracking と呼ばれる研究分野では特定のイベントを表すトピックを時系列データの中から抽出している [1]. その中で、ユーザ毎のトピックや話題の時間変化をモデル化する研究や [4], 近年では新しいトピックの出現の検知として、SNS 上でのイベント検知を行う研究 [18] などが行われている. これらの研究で対象とする話題やイベント自体の抽出は、情報カスケード自体の抽出という点で関連するが、どの研究も情報の社会的影響力は考慮していない. 重要な話題の早期検知という観点では SNS 上で長期間流行する話題の早期検出を行う研究 [19], Twitter における政治的なトピックの発生を早期発見する研究 [15] など存在するが、これらと本研究とは注目する情報カスケードの性質が異なるといえる.

### 2.3 マイクロブログにおけるスパム検出に関する研究

近年マイクロブログでは悪意のある投稿を自動で行うやスパムユーザ・投稿が増加しており、これらを自動検出することが広く求められている. Chen らはインタラクションの構造に着目し、クラスタリング係数や推移性等の指標が有効であるとしている [6]. Gao らは投稿に注目し、投稿毎にスパム判定をするシステムを提案している [10].

しかし、1 節で述べた通り、社会的影響力の有無を判定するに際しては、カスケードがスパムでないとは判定するだけでは、3.4 節で取り上げる例のように、スパムとはいえないが、社会的影響力を持たない情報カスケードが多く存在するため不十分である. また、スパム検出は基本的にユーザや投稿に対して行っておりカスケードは対象としていない.

## 3. 情報カスケードへのアノテーション

本章では、提案手法の評価に用いる社会的影響力の有無をアノテーションした情報カスケードのデータセットの構築方法とその内容の分析について述べる. アノテーション対象の情報カスケードとして、まず、Twitter API によるツイートの収集を行い、リツイート数に基づいて情報カスケードの抽出を行う. このようにして得られた情報カスケードについて、被験者が社会的影響力の有無をアノテーションすることで評価用のデータセットを得る. その後、社会的影響力を持つ情報カスケードの内容を分析する.

### 3.1 Twitter データセット

情報カスケードの分析、評価を行うためのデータセットとしては、著者らの研究室において 2011 年 3 月より継続的に収集している Twitter のデータセットを用いた. 本データセットは、150 万人程度の公開ユーザからタイムラインを継続的に収集したもので、2015 年 8 月時点で約 250 億のツイートが蓄積されている. 収集対象のユーザは、2011 年 3 月に 30 名程度の著名な日本人ユーザを選択し、それらのユーザに対してメンション (以下では Mt) やリツイート (以下では RT) を行ったユーザを

表 1 情報カスケードの統計量

|                   | 1月      | 2月      |
|-------------------|---------|---------|
| カスケード数 (50RT 以上)  | 31,479  | 16,817  |
| カスケード数 (600RT 以上) | 1,130   | 475     |
| カスケードへの参加総ユーザ数    | 407,034 | 338,640 |

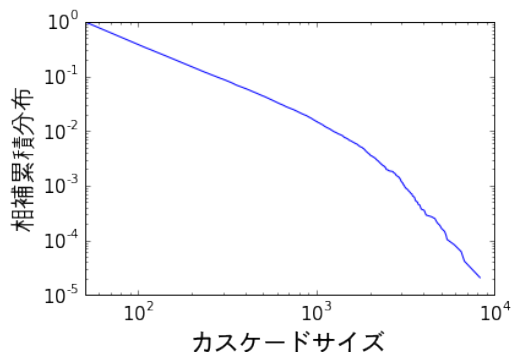


図 1 1月,2月のカスケードサイズの分布

さらに収集対象として順次拡大していったものである。この中から、2012年1月から2013年2月の間につぶやかれたツイートを用いて評価用データセットの構築を行った。

### 3.2 情報カスケードの抽出

本稿における情報カスケードは Twitter API による公式 RT によって拡散されたツイート (元ツイート) とその公式 RT の集合とする。分析対象となる情報カスケードは、次で述べるインタラクショングラフに含まれるユーザを観測対象のユーザセットと限定した上で 2013 年 1 月、2 月のツイートそれぞれから 50 回以上 RT が観測された日本語を含む元ツイートとその RT を抽出することによって作成した。また、非公式 RT が起点となって RT され拡散した情報カスケードは今回分析対象から外した。結果抽出された情報カスケードの統計量は表 1、図 1 の通りである。

また、社会的影響力の有無を判断する情報カスケードは 600 回以上 RT が観測されたものでフィルタし、さらに、情報カスケードの社会的影響力の有無を判断するには (ツイート収集時に取得していなかった) リンクされている画像等も参照する必要があったため、リンク先を復元できなかったの 29 の情報カスケードは分析対象から外した。

インタラクショングラフの構築情報カスケードを抽出する期間以前の 2012 年 1 月から 12 月のユーザ間の投稿のやり取りを元にユーザ間の関係を表す有向グラフ (インタラクショングラフ)  $G$  を作成し、情報カスケードの経路を推定する。なお、推定した経路は次節で述べる提案手法で分類の手がかりの一つとして用いる。RT と Mt はどちらもユーザ間の情報のやり取りを表しており、このようにして得られるユーザ間のつながりはカスケードの情報伝播の主要な経路となると考えられる。そこで各ユーザをノードとして (過去に) 情報が流れる方向と同方向となるよう、RT に関しては RT 元から RT したユーザへ情報が流れるため同方向のエッジを、Mt に関しては Mt を送る際は送り先のユーザの投稿を見て送ったと考えられるため Mt

表 2 インタラクショングラフ

|      | ユーザ数     | 1,066,870   |
|------|----------|-------------|
|      | Mt       | 58,627,341  |
| エッジ数 | RT       | 114,848,093 |
|      | Mt or RT | 153,711,945 |

表 3 1月,2月毎のカスケードの社会的影響力の有無

|         | 1月   | 2月  |
|---------|------|-----|
| 社会的影響力有 | 188  | 106 |
| 社会的影響力無 | 942  | 369 |
| 合計      | 1130 | 475 |

の方向とは逆向きのエッジを追加しインタラクショングラフ  $G$  を得る。今回作成したインタラクショングラフの統計量を表 2 に示す。今回分析の対象としたユーザはこの期間に一度はインタラクション元となっているユーザであり、およそ 100 万ユーザ存在した。

### 3.3 被験者による社会的影響力の有無の注釈付け

3.2 節で得られた情報カスケードに対し (著者を含まない) 3 人の被験者により、「ツイートに書かれた情報を知ったり、その情報を不特定多数に知られたりすることで、直接的あるいは間接的に行動や意思決定に影響を受けるか人がいるか」という観点で社会的影響力の有無を注釈付けした。なお、このアノテーションは実際に RT したユーザが読む/見ると考えられる、元ツイートの本文、画像、元ツイートに含まれる URL のリンク先の情報を基に行ってもらった。これによって得られたラベルにおける三人の被験者間一致度 [8] は 0.69 となり、[13] によれば相当な一致となり、社会的影響力を持つ情報に関して一定の共通認識が得られていることが確認できた。

最終的にラベルの不一致は多数決により解消し、最終的な学習・評価用のデータセットを得た。その結果を表 3 に示す。表 3 から分かる通り、社会的影響力を持つ情報カスケードは 600RT 以上確認されたカスケード全体のおよそ 20%弱、50RT 以上確認されたカスケード全体に対してはおよそ 0.6%と少ないことが確認された。

### 3.4 社会的に影響力のある情報カスケードの分析

前節で抽出した社会的に影響力を持つ情報カスケードに対し、どのような社会的影響力が存在するかを明らかにするため、拡散された情報の内容について分類を行った。

分析の結果、社会的影響力のある情報カスケードは 1) 個人、組織、業界、社会への意見に対する共感・反感と 2) 影響力のある出来事 (事実) の周知、として大別されることが分かった。個人、組織、業界、社会への意見に対する共感・反感は、その意見に対し個人、組織、業界が対応する必要があるため社会的影響力を持ち、社会に対する意見では、そのカスケードが世論を反映していると考えられるため社会的影響力を持つ。一方、影響力のある出来事 (事実) の周知には、事件・事故の速報や知られていない問題の周知が含まれ、それぞれ、世論動向や注意喚起、啓蒙やデマ訂正、問題提起として重要である。詳細な内容については著者らの過去の研究 [20] を参考にされたい。

実際にアノテーションされた社会的影響力の種類を分類した

表 4 個人，組織，業界，社会への意見に対する共感・反感の分布

|    | 明確 | 不明確 |
|----|----|-----|
| 個人 | 23 | 33  |
| 組織 | 44 | 18  |
| 業界 | 32 |     |
| 社会 | 45 |     |

表 5 影響力のある出来事 (事実) の周知の分布

|           |    |
|-----------|----|
| 速報        | 36 |
| 注意喚起・デマ訂正 | 37 |
| 啓蒙        | 26 |

結果，個人，組織，業界，社会への意見に対する共感・反感の個数の合計は 195，影響力のある出来事 (事実) の周知の個数の合計は 99，社会的影響力の無いカスケードの数の合計は 1311 となった．詳細な分類の結果は表 4，表 5 に示す．意見の項目では組織や社会に対する意見が多いことが分かった．特に世代などへの意見や，特定の企業に対する批判が多く見られた．また，社会的影響力を持たないとされた情報カスケードの多くは下に示す例のように日常の面白い出来事やジョークをつぶやくものであり，スパムのような情報カスケードはほとんど見られなかった．

「きょうは 3 月並みの気温」と聞いて驚いたり慌てたりしている皆さん、落ち着いてください。実は... あしたから全国的に 3 月なんです。

(注): <https://twitter.com/Lionbass/statuses/306912821982158848>

被験者による社会的影響力のアノテーションの有無の不一致は啓蒙の項目に多く見られた．啓蒙では，社会的影響力の有無を判断する際に情報の受け手の知識が影響するため，その部分でゆれが見られることが原因であると考えられる．

#### 4. 提案手法

情報カスケードが拡散し，さらに社会的影響力を持つかどうかを同時に分類する手法を提案する．本研究では 1) 直接社会的影響力を持つ情報カスケードを自動抽出する手法と，2) 第一段階としてカスケードサイズが 600 に達するかどうかを予測し，その次に著者らの過去の研究で提案している手法 [20] を用いて分類するという二段階の手順を踏む手法の二種類の手法を提案し，これらと比較する．

それぞれの自動分類においては線形カーネルのサポートベクターマシン (SVM) を用いて情報カスケードの社会的影響力の有無を識別する分類器を学習し，これを用いて社会的影響力を持つ情報カスケードの検知を行う．

分類には，著者等の過去の研究 [20] で有効性が示されたテキスト特徴量，グラフ特徴量，ユーザ特徴量に加え，情報カスケードに対する Mt を手がかりとして用いる．この Mt は情報カスケードに対するユーザの反応であり，情報カスケードがユーザに対してどのような影響を与えるのか，どのようなユー

ザに対して影響を与えるのかを捉えることで効果的な分類ができると考え，Mt に含まれるテキスト，Mt を行ったユーザを特徴量として用いることを提案する．

なお，初期 n 回 RT された情報カスケードに対して用いる事ができる Mt の集合は n 回目の RT よりも (時間的に) 前に元ツイートに対して行われた Mt の集合として得られる．今回対象とする Mt は RT 数やユーザーに登録されたお気に入りの数によって自動で Mt を行うような bot による Mt は社会的影響力の判別には有効でないと考え，@Favstar, @favstar から始まるスクリーンネームのユーザによる Mt は取り除いた．

##### 4.1 テキスト特徴量 (Text)

本研究ではテキスト特徴量としては大きく分けて元ツイートのテキスト特徴量と情報カスケードに付随する Mt のテキスト特徴量を考える．情報カスケードによって伝播する情報の内容を表す特徴量として元ツイートの本文特徴量を，情報カスケードがそれを読んだユーザに対してどのような影響を与えたのかを表す特徴量として Mt のテキスト特徴量を考える．

元ツイートのテキスト特徴量としては，(1) 出現する単語 (BoW) (2) 本文の長さ ( $len_{Text}$ ) (3) 固有表現の有無 ( $has\_ne$ ) (4) URL の有無 ( $has\_url$ ) の 4 種類を利用した．

出現する単語 (BoW) 社会的影響力の有無は伝播する情報内容によって決まるため，特徴的な単語が出現する (例: 事故, 危険)．実験では Bag of words を用い，具体的には，元ツイートの本文から URL，ユーザネームを除き，w, W の連続を一つの文字列とする正規化を行った後 MeCab<sup>(注1)</sup> で mecab-ipadic-NEologd<sup>(注2)</sup> を辞書に用い形態素解析し，自立語の動詞，名詞<sup>(注3)</sup>，形容詞を用いた．

本文の長さ ( $len_{Text}$ ) 意見を述べる社会的影響力を持つ情報カスケードには，本文の長さが長いものが多く存在する．そのため，本文の長さも有効な判断基準と考えられる．実験では，@から始まるユーザネーム，URL を除いた文字数を用い，長さが 70 以上であるかどうかで本文の長さを表す特徴量とした．固有表現の有無 ( $has\_ne$ ) 固有表現が含まれ，その人物や組織に対し意見や批判が述べられている場合，その情報カスケードは社会的影響力を持つ可能性が高いと考えられる．そのため，ツイートに固有表現が含まれるかどうかは有効な判断基準となると考えられる．そこで実験では，CaboCha<sup>(注4)</sup> を用いて固有表現抽出を行い，組織名 (ORGANIZATION)，人名 (PERSON)，地名 (LOCATION)，固有物名 (ARTIFACT) のいずれかの固有表現が含まれているかどうかで特徴量とした．URL の有無 ( $has\_url$ ) 社会的影響力を持たないネタの情報カスケードは画像を含むことが多い．そのため，URL の有無は判断基準として有効だと考えられる．そこで実験では，URL が含まれるかどうかを，1, 0 で表し，特徴量とした．

次に，Mt のテキスト特徴量としては，(1) 出現する単語 ( $BoW_{Mt}$ ) (2) Mt の長さ ( $len_{Mt}$ ) の 2 種類を用いた．テキスト

(注1): <http://taku910.github.io/mecab/>

(注2): <https://github.com/neologd/mecab-ipadic-neologd/>

(注3): ただし接尾辞，数は除く

(注4): <http://taku910.github.io/cabochoa/>

表 6 カスケード構造グラフ

|       | 双方向エッジのみ  | 片方向エッジ         |
|-------|-----------|----------------|
| 全ユーザ  | $G'(R)$   | $\hat{G}(R)$   |
| 直前ユーザ | $G'_2(R)$ | $\hat{G}_2(R)$ |

表 7 グラフに関する特徴量

(a) ルートユーザのグラフ特徴量 (Graph\_root)

|  |
|--|
| $G(V_0), \hat{G}(V_0), \hat{G}_2(V_0)$ の outdegree |
| $G'(V_0), G'_2(V_0)$ の degree                      |

(b) RT したユーザのグラフ特徴量 (Graph\_RT)

|   |
|---|
| $G, \hat{G}, \hat{G}_2$ の outdegree の分布 |
| $G', G'_2$ の degree の分布                 |

(c) 情報伝播過程に関する特徴量 (Graph\_Structure)

|                                       |
|---------------------------------------|
| $G', \hat{G}$ の最大の連結成分の大きさ            |
| $G', G'_2, \hat{G}, \hat{G}_2$ の総エッジ数 |
| $G'$ の深さの平均                           |
| $G'$ の深さの分布                           |
| $G'$ のクラスタリング係数の平均                    |

の正規化の際に、非公式 RT や引用の際に広く用いられる”RT”, ”Rt”, ”rt”, ”QT”, ”Qt”, ”qt”から始まる元ツイートの引用、元ツイートと全く同様の文字列は削除することで、情報カスケードに対する反応のみを抽出した。

出現する単語 (BoW<sub>Mt</sub>) 社会的影響力の有無によってそれを読んだユーザの反応は変化すると考えられ、その反応を表す単語は有用な判断基準と考えられる。実験では元ツイートの本文に出現する単語 (BoW) と同様の方法で特徴量とした。

Mt の長さ (len<sub>Mt</sub>) 意見を述べる情報カスケードに対しては、読んだユーザも自分の意見を投稿することがあるため Mt の長さも長くなることが多いと考えられる。そのため、Mt の長さも有効な判断基準と考えられる。実験では Mt の長さの閾値は、訓練データの分布を参考にし、長さが 0 であるか (引用やユーザ名のみ) の Mt)、長さが 22 以下、長さが 23 以上の 3 種類のどれであるかで特徴量とした。

#### 4.2 グラフ特徴量 (Graph)

カスケードの伝播経路構造を特徴量として捉えることを目的にしてカスケードのグラフ構造を作成する。カスケードのグラフ構造はインタラクショングラフ G の RT したユーザの集合による部分グラフから作成されるが、親密なユーザ間でのやりとりに注目すること、仮想的な伝播経路を定めることを目的とし、エッジの残し方を変えることで 4 種類のグラフ構造を得る。まずエッジの残し方を 1) 以前に RT したユーザで、インタラクショングラフ上でエッジがあるユーザ全てからのエッジを用いる場合 2) インタラクショングラフ上でエッジがあるユーザのうち直前に RT したユーザからのエッジのみを用いる場合の二種類を用いることでインタラクショングラフ G の情報をそのまま残したグラフ  $\hat{G}$  と仮想的に伝播経路を定めたグラフ  $\hat{G}_2$  を作成する。さらにそれぞれに対し双方向のエッジのみを残す

ことで得られるグラフ  $G', G'_2$  を用いることでユーザ間の親密性を捉えたグラフを作成した。以上の 4 種類のグラフをまとめると表 6 のようになる。

今回用いたグラフ特徴量は表 7 に示す。情報カスケードから捉えたグラフ特徴量は元ツイートを投稿したユーザに対する特徴量 (Graph\_root)、RT したユーザのグラフ特徴量 (Graph\_RT)、情報伝播過程に関する特徴量 (Graph\_Structure) に分類できる。前者 2 つはユーザ自身の特徴とインタラクショングラフ上で直接接続しているユーザとの関係を表し、グラフ構造に関する特徴量はカスケード伝播特徴を捉えることを目的としている。ルートユーザのグラフ特徴量 (Graph\_root) 情報カスケードが社会的影響力を持つかどうかは、どのようなユーザが発信元になっているかである程度判断できると考えられる。特にマイクロブログ内でどれだけ影響力を持っているか、どれだけのユーザに直接 RT されたか、などが重要な要素であると考え、実験ではルートユーザのグラフ特徴量として、マイクロブログ内での影響力を表す指数として、G における出次数、情報カスケードを直接どれだけ広めたかを表す指数として  $\hat{G}, \hat{G}_2$  の出次数、 $G', G'_2$  の次数を用いた。

RT したユーザのグラフ特徴量 (Graph\_RT) どのようなユーザが発信元になっているかが、社会的影響力の判断に重要なと同様に、どれだけ影響力を持つユーザが RT しているのかも判断に有効だと考えられる。そのため実験では次数の分布を用いることで特徴量とした。具体的には、この分布は次数の逆累積度数分布を次数の軸において対数軸で 10 個の bin に分け、特徴ベクトルの各次元に対応させるという手法を用いた。情報伝播過程に関する特徴量 (Graph\_Structure) カスケードの成長予測やスパム検知にはグラフ構造に関する特徴量が有効に働くことが知られている [6][7]。そのため、社会的影響力の有無にも有効に働くと考えられる。実験では連結成分、総エッジ数、クラスタリング係数についてはどれだけ密なグラフであるかという指標として、深さはどれだけルートユーザから遠くまで伝播したかという指標として用いた。

#### 4.3 ユーザ特徴量 (User)

ユーザ特徴量もテキスト特徴量と同様に、RT したユーザの特徴量 (user<sub>RT</sub>) と、情報カスケードに付随する Mt のユーザ特徴量 (user<sub>Mt</sub>) を考える。

RT したユーザの特徴量 (user<sub>RT</sub>) マイクロブログではユーザごとに RT するツイートの内容に偏りがあり、社会的影響力のあるカスケードばかりを RT するユーザや、ネタなど社会的影響力のないツイートばかりを RT するユーザが投稿を拡散しているかが分類の手がかりとなる。そこで、実験ではテキストに対する Bag of Words を参考に、カスケードに参加するユーザを、いわば Bag of Users として特徴量にした。つまり、ある情報カスケードのユーザ特徴量を作る際にはその情報カスケードをつぶやいたユーザの対応する次元を 1、それ以外の次元を 0 にすることでユーザの特徴ベクトルとする。このユーザの特徴量の次元は、カスケードサイズ 600 の学習データに含まれるユーザに対応しており、次元数はカスケードサイズ 600 のもので 137,167 次元、50 のもので 33,234 次元となった。

表 8 同時分類を行う際の正例，負例の分布

|    | 正例  | 負例    |       |      |      |      |     |
|----|-----|-------|-------|------|------|------|-----|
|    |     | 50    | 100   | 200  | 300  | 400  | 600 |
| 1月 | 188 | 31291 | 12088 | 4749 | 2758 | 1815 | 939 |
| 2月 | 103 | 16714 | 6014  | 2203 | 1195 | 784  | 364 |

表 9 第一段階，成長予測における正例，負例の分布

|    | 正例   | 負例    |       |      |      |      |
|----|------|-------|-------|------|------|------|
|    |      | 50    | 100   | 200  | 300  | 400  |
| 1月 | 1127 | 31291 | 12088 | 4749 | 2758 | 1815 |
| 2月 | 467  | 16714 | 6014  | 2203 | 1195 | 784  |

情報カスケードに付随する Mt のユーザ特徴量 ( $user_{Mt}$ ) RT するユーザが情報カスケードの社会的影響力の判断に有効なものと同様に，情報カスケードに対し Mt をしているユーザも社会的影響力の判断に有効であると考えられる．そのため Mt を行ったユーザに対しても，情報カスケードのユーザ特徴量と同様にして，Bag of Users として特徴量とした．

## 5. 実験

本節では 2013 年 1 月の情報カスケードによって分類器を学習し，2 月の情報カスケードを自動で分類することで，提案した分類手法でどれだけ正確に社会的影響力を持つ情報カスケードを検知することができるかを評価する実験について述べる．

### 5.1 実験手順

本研究の目的は早期の，つまりカスケードが広がらない段階でのカスケードの社会的影響力の有無の判断であるため，学習・評価の各カスケードは  $n$  回以上 RT が観測されたカスケードとし，先頭から  $n$  回目までの RT より求められる特徴量を用い，分類器の学習・評価を行った ( $n=50, 100, 200, 300, 400, 600$ )． $n$  が小さければ小さいほど候補となる情報カスケード (負例) の数が多くなり，正例と負例のバランスが偏ることから難しい問題となる．表 8 に，今回の学習セット，評価セットの正例負例の分布を示す．

分類器の学習の際，グラフ特徴量は各次元が実数値を取るため，0 から 1 の間の値へ正規化を行い，分類器としては LIBLINEAR<sup>(注5)</sup> を用いて線形カーネルの SVM を学習した．また，分類のラベルに偏りがあるため，学習の際に正例側に重みをかけることで対応した．なお，パラメータチューニングは学習データにおいて 5 分割交差検定を用いて最大の  $F_1$  値を取るパラメータを用いた．

また，二段階で分類を行う場合，第一段階の分類の際は，テキスト特徴量，グラフ特徴量，ユーザ特徴量，全ての特徴量を組み合わせた場合で分類を行い，比較する．その後，最も良い  $F_1$  値を示した分類器を用い，2 月分のテストデータの分類を行い，その結果，著者らの研究室のツイートデータセットにおいて RT 数 600 まで成長すると判別されたカスケードを第二段階の分類の入力として用いる．また，第二段階の分類器は [20] と同様，1 月の訓練データにおいて，実際に 600RT 以上観測さ

表 10 第一段階 (カスケードの成長予測) の  $F_1$  値の変化

|          | カスケードサイズ     |              |              |              |              |
|----------|--------------|--------------|--------------|--------------|--------------|
|          | 50           | 100          | 200          | 300          | 400          |
| ALL      | 0.290        | 0.398        | 0.505        | 0.602        | 0.711        |
| Text     | 0.102        | 0.208        | 0.375        | 0.545        | 0.695        |
| User     | 0.150        | 0.297        | 0.449        | 0.580        | 0.701        |
| Graph    | <b>0.406</b> | <b>0.576</b> | <b>0.732</b> | <b>0.819</b> | <b>0.869</b> |
| Baseline | 0.054        | 0.142        | 0.336        | 0.528        | 0.688        |

表 11 カスケードサイズの変化に伴う  $F_1$  値の変動

|            | カスケードサイズ     |              |              |              |              |              |
|------------|--------------|--------------|--------------|--------------|--------------|--------------|
|            | 50           | 100          | 200          | 300          | 400          | 600          |
| 提案手法 (同時)  | 0.142        | 0.189        | 0.289        | 0.338        | 0.461        | <b>0.762</b> |
| 提案手法 (二段階) | <b>0.268</b> | <b>0.378</b> | <b>0.467</b> | <b>0.545</b> | <b>0.585</b> | <b>0.762</b> |
| Baseline   | 0.012        | 0.033        | 0.086        | 0.147        | 0.208        | 0.361        |
| 被験者 A      | n/a          | n/a          | n/a          | n/a          | n/a          | 0.907        |
| 被験者 B      | n/a          | n/a          | n/a          | n/a          | n/a          | 0.818        |
| 被験者 C      | n/a          | n/a          | n/a          | n/a          | n/a          | 0.900        |

れたカスケードのみを用いることで学習を行う．

### 5.2 実験結果

まず，二段階で分類を行う際の，第一段階の分類結果を表 10 に示す．この結果より，第二段階の分類の際は，グラフ特徴量のみを用いて分類を行った結果成長が予測されたカスケードを入力として用いた．

次に，同時分類，二段階分類を行った際の検知結果を表 11 に示し，比較する．参考のため，各被験者と正解ラベル (多数決) との一致度もともに示す．どれも，全てのカスケードが社会的影響力を持つとした場合 (Baseline) の  $F_1$  値と比較して改善していることが分かる．また，カスケードサイズが小さい (早期である) ほど  $F_1$  値が顕著に小さくなっており，早期の段階で分類するのは難しいことが確認できる．この原因はカスケードサイズの分布が図 1 で示した通り，べき乗分布に従っているためであると考えられる．

また，同時に分類を行う手法と，二段階で分類を行う手法を比較すると，どのカスケードサイズの場合でも，二段階で分類を行う手法の方が分類性能が良くなるということが判明した．この原因としては，第一段階の「カスケードが成長するかどうか」という分類において，テキスト特徴量，ユーザ特徴量が悪影響を及ぼしているからであると推測できる．

### 5.3 考察

#### 5.3.1 分類結果の分析

各カスケードを SVM で分類する際の判断基準である，分離超平面からのマージンを 0 から動かすことで，分類器の適合率と再現率のトレードオフを調査した．なお，二段階分類による手法では，一段階目は開発データで最大の  $F_1$  値を取る閾値を用い，二段階目の分類平面のマージンを動かすことでトレードオフを調査した．図 2，図 3 にそれぞれのカスケードサイズで分類を行った際の適合率と再現率のトレードオフを示す．

まず図 2，図 3 を比較すると，同時分類を行う手法は二段階分類を行う手法に比べて適合率が低くなる傾向が見られる．この原因としては，テキスト特徴，ユーザ特徴によって RT 数 600

(注5): <https://www.csie.ntu.edu.tw/~cjlin/liblinear/>

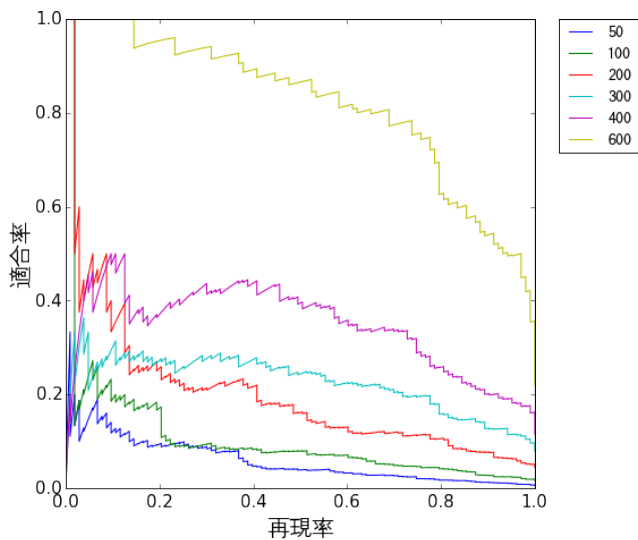


図 2 同時分類による分類結果

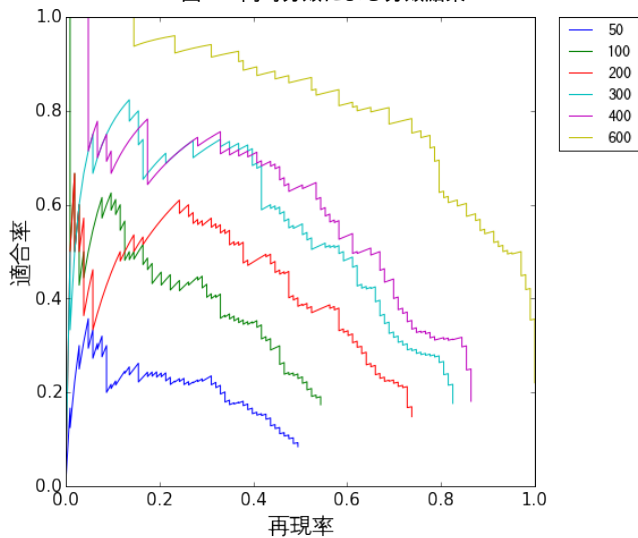


図 3 二段階分類による分類結果

まで拡大しない情報カスケードが多く選ばれることが考えられる。一方で二段階で解いた場合は、図 3 から分かる通り、直接解いた場合に比べ再現率が低くなっていることが分かる。この原因は第一段階で情報カスケードの成長予測を行った段階で正例となる社会的影響力を持つ情報カスケードが落ちてしまっていることである。これを防ぐ方法として、第一段階の成長予測の時点で社会的影響力を持つ情報カスケードに対してはより大きな重みを付けて学習を行うことや、第一段階、第二段階の学習を完全に独立には行わず、第一段階の分類結果を第二段階の分類の特徴量として利用して分類する方法などが考えられる。

### 5.3.2 分類の誤り分析

次に、どのような社会的影響力を持つ情報カスケードを検知できているのか調査を行った。先頭 RT 数 50 の段階で同時分類による結果を表 12, 表 13 に示す。表は 3.4 節の表 4, 表 5 と対応しており、表中のそれぞれの項目の数字は、正しく予測できた情報カスケードの数とテストデータに存在する情報カスケードの数に対応する。

この結果から分かる通り、影響力のある出来事(事実)の周知の項目についてはほとんど予測を行っていない一方で、個人、

表 12 個人, 組織, 業界, 社会への意見に対する共感・反感の予測結果

|    | 明確   | 不明確 |
|----|------|-----|
| 個人 | 2/3  | 2/8 |
| 組織 | 9/19 | 3/4 |
| 業界 | 7/15 |     |
| 社会 | 4/16 |     |

表 13 影響力のある出来事(事実)の周知の予測結果

|           |      |
|-----------|------|
| 速報        | 1/20 |
| 注意喚起・デマ訂正 | 2/7  |
| 啓蒙        | 2/11 |

組織, 業界, 社会への意見に対する共感・反感の項目については 65 個中 27 個と 4 割以上の再現率で予測を行っていることが分かる。このことを踏まえると、炎上等につながる意見や批判を早期検知することはある程度できているが、危険やデマなどを広める情報カスケードはあまり早期検知をすることはできていないということがいえる。また、この原因としては影響力のある出来事(事実)の周知と個人, 組織, 業界, 社会への意見に対する共感・反感を同一視してしまっていることが考えられる。RT するユーザ層やテキスト特徴など、捉えたい特徴が異なると考えると、これらを同一視することなく別々の方法で検知を行うことで、より精度を向上させることができる可能性がある。

## 6. まとめと今後の課題

本研究では社会的影響力を持つ情報カスケードの早期検知という新しいタスクを提案し、これを教師あり学習に基づく分類器を用いて解く手法を提案した。

まず、複数人の被験者によるアノテーションを行い、その一致率を基に社会的影響力を持つ情報カスケードについて問題の妥当性の担保を行った。次に、社会的影響力を持つ情報カスケードを早期に検知を行う手法を提案し、実験により、いずれのカスケードサイズの場合でも  $F_1$  値で全ての情報カスケードが影響力有りだとした場合 (Baseline) と比べ大きな改善が見られた。しかし、実用的にはさらなる精度の改善が求められる。

今後の課題としては、データセットの拡大が大きな課題である。社会的影響力の有無について人手でアノテーションを行っているため、正解データを増やすには大きなコストが伴う。これを解決するためには自動で正解データを増やす手法や、クラウドソーシングなどを検討する必要がある。また、現在では第一段階、第二段階の分類を完全に独立に行ってしまうと、それによって最終的な再現率の低下が起きている。これを改善するために、第一段階の分類結果を第二段階の分類の特徴量として利用することで、現在では第一段階で見落としてしまっている社会的影響力を持つ情報カスケードを判別できるようになる可能性がある。

## 謝 辞

本研究の一部は JSPS 科研費 25280111 の助成を受けたものです。

## 文 献

- [1] Allan, J., Carbonell, J., Doddington, G. et al.: Topic Detection and Tracking Pilot Study Final Report, *Proceedings of the DARPA Broadcast News Transcription and Understanding Workshop*, pp. 194–218 (1998).
- [2] Bakshy, E., Hofman, J. M., Mason, W. A. and Watts, D. J.: Everyone’s an Influencer: Quantifying Influence on Twitter, *Proceedings of ACM International Conference on Web Search and Data Mining*, pp. 65–74 (2011).
- [3] Bikhchandani, S., Welch, I. and Hirshleifer, D. A.: A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades, *Political Economy*, pp. 992–1026 (1992).
- [4] Blei, D. M. and Lafferty, J. D.: Dynamic topic models, *Proceedings of International Conference on Machine Learning*, pp. 113–120 (2006).
- [5] Castillo, C., Mendoza, M. and Poblete, B.: Information credibility on twitter, *Proceedings of International World Wide Web Conference*, pp. 675–684 (2011).
- [6] Chen, P.-C., Lee, H.-M., Tyan, H.-R., Wu, J.-S. and Wei, T.-E.: Detecting spam on Twitter via message-passing based on retweet-relation, *Proceedings of Technologies and Applications of Artificial Intelligence*, pp. 56–65 (2014).
- [7] Cheng, J., Adamic, L., Dow, P. A., Kleinberg, J. M. and Leskovec, J.: Can cascades be predicted?, *Proceedings of International World Wide Web Conference*, pp. 925–936 (2014).
- [8] Fleiss, J. L.: Measuring nominal scale agreement among many raters., *Psychological Bulletin*, No. 5, pp. 378–382 (1971).
- [9] Galuba, W., Aberer, K., Chakraborty, D., Despotovic, Z. and Kellerer, W.: Outtweeting the twitterers-predicting information cascades in microblogs, *Proceedings of Workshop on Online Social Networks* (2010).
- [10] Gao, H., Chen, Y., Lee, K. et al.: Towards Online Spam Filtering in Social Networks, *Proceedings of The Network and Distributed System Security Symposium* (2012).
- [11] Geerajit Rattanaritnont, Masashi Toyoda, M. K.: Characterizing Topic-Specific Hashtag Cascade in Twitter Based on Distributions of User Influence, *Proceedings of Asia-Pacific Web Conference*, pp. 735–742 (2012).
- [12] Hongguang, Z., Nobuhiro, K., Naoki, Y. and Masashi, T.: Proceedings of A Study on Microblog Classification Based on Information Publicness, *DEIM Forum* (2012).
- [13] Landis, R. J. and Koch, G. G.: The measurement of observer agreement for categorical data, *Biometrics*, Vol. 33, No. 1, pp. 159–174 (1977).
- [14] Ren, Z., Peetz, M.-h., Liang, S., Dolen, W. V. and Rijke, M. D.: Hierarchical Multi-Label Classification of Social Text Streams, *Proceedings of Special Interest Group on Information Retrieval*, pp. 213–222 (2014).
- [15] Rill, S., Reinel, D., Scheidt, J. and Zicari, R. V.: PoliTwI: Early detection of emerging political topics on twitter and the impact on concept-level sentiment analysis, *Knowledge-Based Systems*, Vol. 69, No. 1, pp. 24–33 (2014).
- [16] Sriram, B., Fuhry, D., Demir, E., Ferhatosmanoglu, H. and Demirbas, M.: Short text classification in twitter to improve information filtering, *Proceedings of Special Interest Group on Information Retrieval*, pp. 841–842 (2010).
- [17] Tsur, O. and Rappoport, A.: What’s in a Hashtag?: Content Based Prediction of the Spread of Ideas in Microblogging Communities, *Proceedings of ACM International Conference on Web Search and Data Mining*, pp. 643–652 (2012).
- [18] Zhou, X. and Chen, L.: Event detection over twitter social media streams, *The International Journal on Very Large Data Bases*, Vol. 23, No. 3, pp. 381–400 (2013).
- [19] 斎藤 翔太, 富岡 亮太, 山西 健司: ソーシャルネットワークにおける長期間流行する話題の早期検出, 電子情報通信学会技術研究報告. *IBISML*, 情報論的学習理論と機械学習, Vol. 111, No. 480, pp. 77–84 (2012).
- [20] 川本 貴史, 豊田 正史, 吉永 直樹: マイクロブログにおける社会的影響力を持つ情報カスケードの早期検知に向けて, *Proceedings of WebDBForum*, pp. 48–55 (2015).