

グラフマイニングとテキストマイニング技術の統合によるソーシャルセンサからの実社会イベント予測手法の開発

研究代表者

津川 翔

筑波大学 システム情報系 助教

1 はじめに

電気通信技術の発展により、多くの人が日常的に様々なネットワークサービスを利用している。本研究では、ネットワークサービスのユーザを、実社会を観測するセンサであるソーシャルセンサとして捉える。ソーシャルセンサから得られる様々な情報、すなわちユーザの検索履歴や Web 閲覧の履歴、ソーシャルメディアにおける投稿などの情報は、実社会におけるトレンドを反映していると考えられている。そのため、ソーシャルセンサから得られる情報を用いて、病気の流行や経済状況の変動などの実社会におけるイベントを予測する技術の実現に期待が高まっている [1-6]。

ソーシャルセンサを用いて実社会イベントを予測する試み自体は既に国内外の学会において報告されている。これまでに、Web 検索の履歴を用いた景気動向の予測 [2] やインフルエンザの流行の予測 [6]、選挙結果の予測 [7] などが試みられている。このような技術によって、例えば、病気の流行を予測することができれば、国や自治体がワクチンの備蓄や伝染病の対策などを行う上で有用である。ただし、このような研究分野は未だ発展途上の段階であるため、現在も、予測精度の向上や予測の適用範囲拡大を目指して、研究開発が進められている。

本研究は、従来研究のトレンドを踏襲し、ソーシャルセンサから実社会イベントを予測する技術を確立することを目指している。これまで、2015 年度の電気通信普及財団の支援を受け、Google Trends という Web 検索ユーザの行動を反映したデータと、Wikipedia のページの閲覧数のデータを用いて、自動車の販売台数を予測するモデルを構築している。また、実社会イベントの予測において有用な投稿を行うユーザを特定するための、ユーザの影響力推定手法の基礎的な評価を実施している。本年度は、これまでの研究をさらに発展させ、研究期間内において、主に以下の 2 つの研究課題に取り組んだ。

- A) ソーシャルメディアデータのシグナル検出技術の研究
- B) ソーシャルメディアデータからの流行予測技術の研究

研究課題 (A) においては、2015 年度の研究を発展させ、ソーシャルメディアユーザの中でも、実社会のトレンドなどに影響を与え得るような影響力の強いユーザを抽出するための技術の開発を行った。研究課題 (B) では、ソーシャルメディア上の流行を予測するための研究を行った。テキストマイニングの手法を用いて抽出したソーシャルメディアにおける投稿のセンチメントと、グラフマイニングの手法を用いて抽出したソーシャルメディアユーザのコミュニティの情報が、ソーシャルメディアにおける流行の予測に有用であることを示した。

以降、2 章では研究課題 (A) について、3 章では、研究課題 (B) について報告する。4 章において全体のまとめと今後の課題について述べる。

2 ソーシャルメディアデータのシグナル検出技術の研究

2-1 研究の背景と目的

ソーシャルメディアでは多くのユーザが様々なトピックの投稿を行っている。これらの投稿は、実社会のトレンドを反映していると考えられており、ソーシャルメディアのデータから景気動向や選挙結果の予測などが試みられている。ただし、ソーシャルメディアのデータには予測にとって有用でないノイズも含まれている。このようなノイズを除去し、予測にとって有用なシグナルを抽出することが、予測精度の向上に有用であると考えられる。

本研究課題では、ソーシャルメディアユーザの中でも、実社会のトレンドに影響を与え得るような影響力の強いユーザを抽出するための技術を開発することを目指す。影響力の強いユーザの投稿は実社会イベント予測において有用なシグナルとなるのではないかと考えている。

今年度は、昨年度から引き続き、ソーシャルメディアユーザ間の関係を表現したソーシャルネットワーク

から、影響力の強いユーザを推定する手法を検討した。ただしソーシャルネットワークは非常に大規模であり、その構造を正確に取得することは現実的ではない。そこで、ソーシャルネットワークの部分的な構造をサンプリングし、一部の構造から影響力の強いユーザを推定する手法を検討した。具体的には、Sample Edge Count と呼ばれるネットワークサンプリング手法を用いて得られた部分グラフから、次数中心性、PageRank などの指標を求めることで、影響力の強いユーザを抽出した。これによって抽出されたユーザは、実際に影響力の強いユーザのおおよそ 70% 程度の影響力を有しており、十分に影響力の強いユーザを発見できることを確認した。

さらに、影響力の強いユーザの推定精度を高めるため、ソーシャルネットワークの構造を補完するリンク予測手法の研究を行った。影響力の強いユーザの推定には、ソーシャルネットワークの構造を用いるが、そのソーシャルネットワークの構造を正確に得ることは難しい。そこで、リンク予測によって、ソーシャルネットワークの構造を補完する手法を提案した。従来のリンク予測では用いられてこなかったユーザ間の情報拡散の履歴を用いることで、リンク予測の精度が向上することを示した。

2-2 サンプリングされたソーシャルネットワークからの影響力の強いユーザの推定

サンプリングによって得られた不完全なソーシャルネットワークから推定したユーザの影響力と、各ユーザの実際の影響力の関係を調査する。情報の欠落がないソーシャルネットワークからノードをサンプリングすることにより、不完全なソーシャルネットワークを得る。得られた不完全なソーシャルネットワークから、次数中心性・近接中心性・媒介中心性・PageRank・k-core 指標を用いて、各ユーザの影響力を推定する。各ユーザの実際の影響力として情報拡散数を用いる。特に影響力の強いユーザをどの程度推定できるかに着目して評価を行う。推定したユーザの影響力のランキング上位 1% 以内に位置するユーザの情報拡散数を、実際の情報拡散数に基づくランキングの上位 1% 以内に位置するユーザの情報拡散数で正規化した値 $I_{1\%}$ を求める。

本課題では、我々の収集した Twitter (follow) ならびに公開データセットである、Twitter (mention)、Facebook、API Journals の 4 種類のデータセットを用いる。代表的なネットワークのサンプリング手法である BFS (Breadth First Search: 幅優先探索)、DFS (Depth First Search: 深さ優先探索)、SEC (Sample Edge Count) [8]、およびランダムサンプリングを用いて、ネットワークからノードをサンプリングする。なお、BFS、DFS、SEC ではランダムサンプリングと比較して、次数の高いノードがサンプリングされることが知られている [8]。影響力の強いノードは次数が高い場合が多いため、これらのサンプリング手法を用いることで、サンプリングするノードの割合が小さい場合でもサンプリングされた部分ネットワークから影響力の強いユーザを推定できると予想される。本実験では、各サンプリング手法を用いた場合の、サンプリングするノードの割合がユーザの影響力を推定する指標にどのような影響を与えるかを明らかにする。

図 1 に SEC 法を用いた場合のサンプリングの割合と、推定した影響力の強いユーザの正規化した影響力の関係を示す。この結果より、SEC 法でサンプリングした場合には、小さいサンプルサイズで影響力の強いユーザを特定できていることが分かる。特に Facebook や Twitter-mention などのソーシャルメディアのデータセットでは、10% 程度ノードをサンプリングするだけで、ネットワーク全体を用いた場合と同程度の影響力の強いノードを特定できている。このことは、大規模なソーシャルメディアのネットワークから影響力の強いユーザをある程度の精度で推定可能であることを示唆している。

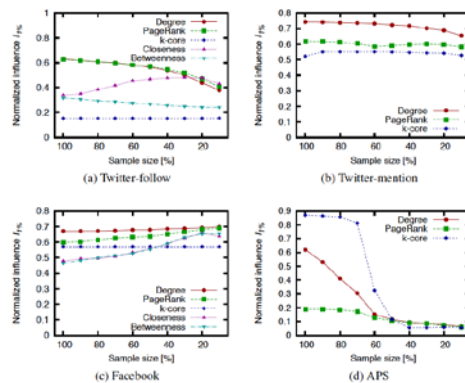


図 1:各データセットにおいて SEC 法によりサンプリングした時の推定したユーザの正規化した影響力

2-3 情報拡散履歴を用いたリンク予測手法

前節では、不完全なソーシャルネットワークの構造のみからでも影響力の強いユーザをある程度の精度で推定できることを示した。ただし、ネットワークの不完全性を補完することができれば、さらに高精度で影響力の強いユーザを特定できる可能性がある。そこで、ソーシャルメディアにおけるユーザ間のソーシャルネットワークにおいて欠損している可能性の高いリンクを予測するリンク予測手法を検討する。

これまでにソーシャルネットワークにおけるリンク予測手法は数多く提案されているが、それらの手法の多くは、ソーシャルネットワークの構造のみを用いてリンク予測を行う。それに対して本研究では、ユーザ間の情報拡散の履歴を用いたリンク予測手法を提案した。その結果、Twitter においては、リツイートに基づくリンク予測が、ソーシャルネットワークの構造のみに基づくリンク予測よりも有効であることを示した。なお、提案手法ならびに結果の詳細については、文献[9]を参照されたい。

2-4 まとめ

本研究課題では、ソーシャルメディアユーザの中でも、実社会のトレンドに影響を与え得るような影響力の強いユーザを抽出するための技術を開発することを目指して研究を行った。SEC 法などのネットワークサンプリングの手法を用いることで、大規模なソーシャルメディアユーザ間のネットワークから影響力の強いユーザをある程度の精度で推定できることを示した。さらに、影響力の強いユーザの推定精度向上を目指して、ソーシャルネットワークの構造を補完するためのリンク予測手法を提案した。提案手法は、ユーザ間の情報拡散の履歴を用いてソーシャルネットワークにおけるリンクの存在を予測する。評価の結果、情報拡散に基づくリンク予測は、ネットワークのトポロジ構造のみに基づく従来のリンク予測よりも予測精度が高いことを示した。

3 ソーシャルメディアデータからの流行予測技術の研究

3-1 研究の背景と目的

Twitter や Facebook などのソーシャルメディアに投稿された情報の一部は、ユーザのソーシャルネットワークを通じて大規模に拡散され、実社会における世論やトレンドの形成に影響を与えうる[10]。ソーシャルメディアのユーザは、Twitter における「リツイート」や、Facebook における「シェア」と呼ばれる機能を用いて、他のユーザに対して情報を拡散することができる。拡散された情報を見たユーザは、さらに自身の友人やフォロワーに対して、その情報を拡散することができる。その結果、多くのユーザの興味を惹いた情報は、ソーシャルメディア上で大規模に拡散される。ソーシャルメディア上の流行を早期に検知することができれば、実社会における流行やトレンドの変化を予測するのにも有用であると考えられる。

本研究では、ソーシャルメディアにおける情報拡散の履歴を分析することにより、ソーシャルメディア上の流行の予測に有用な特徴量を明らかにする。Twitter を対象とし、リツイートの履歴と、ユーザ間のフォロー関係を表現したソーシャルネットワークを分析する。分析の結果、ソーシャルメディア上の投稿の有する感情と、ソーシャルネットワークにおけるコミュニティ構造がソーシャルメディア上の流行を予測する上で有用な特徴量であることを示す。さらに本研究では、ソーシャルメディア上の情報拡散のモデルを構築し、ソーシャルメディア上の流行をどの程度予測できるか検証する。

3-2 投稿の感情と拡散力の関係の分析

代表的なソーシャルメディアである Twitter における投稿(ツイート)を収集し、各ツイートの有する感情と、そのツイートの拡散力、すなわち拡散規模ならびに拡散速度との関係を調査する。ここでは各ツイートの有する感情として、Stieglitz らの研究と同様に、正の感情、負の感情、および中立の感情の3種類を対象とする[11]。

各ツイートが有する感情の判定には、日本語における感情極性が正の単語ならびに感情極性が負の単語のリストである単語感情極性対応表を用いた判定方法と、複数の判定者によって人手で判定する方法の二通りの方法を用いる。まず、ツイートに含まれる単語のうち単語感情極性対応表に掲載されている単語の割合を求めることによって、各ツイートが有する感情を判定する。ただし、このような判定には誤りが含まれると考えられるため、収集したツイートの一部については、さらに人手でツイートの有する感情を判定する。

ツイートが有する感情によって、各ツイートを 3 つのカテゴリ(正の感情を有するツイート、負の感情を有するツイート、中立の感情を有するツイート)に分類し、それぞれのカテゴリごとに、ツイートがリツイートされた回数(リツイート数)、およびツイートが N 回リツイートされるまでの時間 (N リツイート時間) の平均を求める。これにより、ツイートの有する感情と拡散規模および拡散速度との関係を分析する。

収集した全てのツイート(DA) ならびに、収集したツイートから(人手による判定のために) 抽出したツイート(DS) を対象に、ツイートの有する感情とその拡散の規模および拡散との速度の関係を明らかにする。

まず、ツイートの有する感情に基づく 3 つのカテゴリ (正の感情を有するツイート、負の感情を有するツイート、中立の感情を有するツイート) ごとの平均リツイート数を図 2 に示す。図 2 左はデータセット DS から、図 2 右はデータセット DA から求めたツイートの平均リツイート数である。

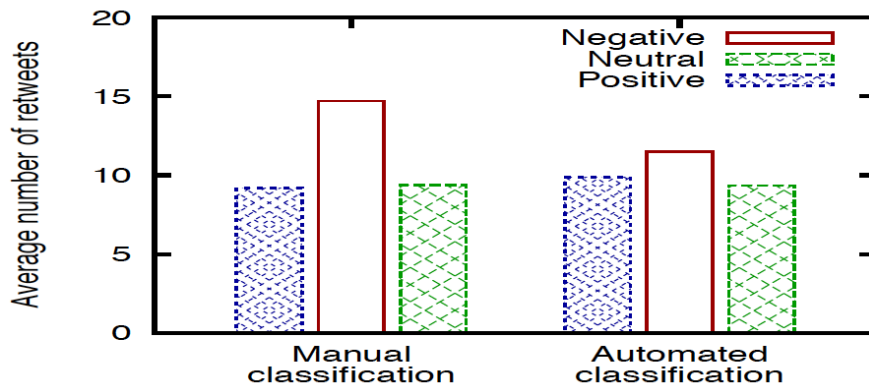
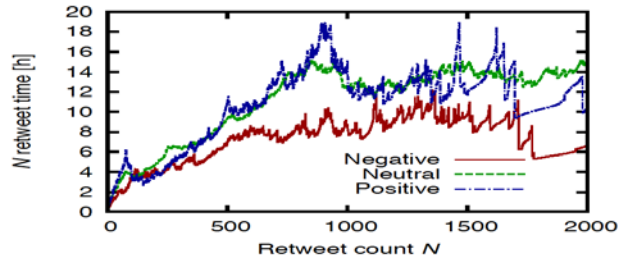


図 2: ツイートの有する感情ごとの平均リツイート数 [12]

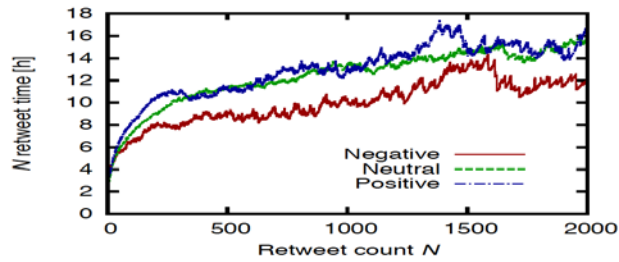
図 2 より、負の感情を有するツイートの平均リツイート数は、正の感情および中立の感情を有するツイートの約 1.3 ~ 1.6 倍であることが分かる。なお、データセット DA において Steel-Dwass 法によって多重比較を行った結果、ツイートの有する感情に基づくどのカテゴリ間にも、リツイート数に有意な差が認められた ($p < 0.05$)。このことから、負の感情を有する投稿は中立の感情を有する投稿の約 1.3 ~ 1.6 倍拡散されること、正の感情を有する投稿は中立の感情を有する投稿よりもやや拡散されにくいことが示された。なお、DS から求めた平均リツイート数と DA から求めた平均リツイート数の違いは、人手による判定と単語感情極性対応表による判定の違いに起因するものと考えられる。

次に、各カテゴリの平均 N リツイート時間を図 3 に示す。平均 N リツイート時間とは、N 回以上リツイートされたツイートの N リツイート時間の平均である。これらの図において、特に N が大きい場合に、平均 N リツイート時間の値が大きく上下しているのは、平均 N リツイート時間を求めるためのサンプル(N 回以上リツイートされたツイート) の数が、N が大きくなるにつれて少なくなるためである。

図 3 より、負の感情を有するツイートの平均 N リツイート時間は、正の感情および中立の感情を有するツイートの平均 N リツイート時間に比べて短いことが分かる。特に、おおよそ N が 100 以上の場合に注目すると、負の感情を有するツイートの平均 N リツイート時間は、正の感情および中立の感情を有するツイートの N リツイート時間よりも、20% 程度短いことが分かる。N が 100 未満の場合には、各カテゴリで N リツイート時間にそれほど大きな違いはないが、負の感情を有するツイートの N リツイート時間は、他の感情を有するツイートの N リツイート時間と比べて 5% 程度短いことも分かる。なお、100 回以上リツイートされたツイートは収集したツイートの約 1% のみである。つまり、100 回以上リツイートされたツイートは、拡散規模の大きなツイートであると考えられる。したがってこの結果から、投稿の拡散がある程度以上の規模に到達した場合には、負の感情を有する投稿は中立の感情を有する投稿と比べて、約 20% 短い時間で拡散する、すなわち約 1.2 倍の拡散速度で拡散することが分かる。



(a) Manual classification



(b) Automated classification

図3: ツイートの有する感情ごとの平均 N リツイート時間 [12]

一方で、これらの結果から、正の感情を有するツイートの平均 N リツイート時間は中立の感情を有するツイートの平均 N リツイート時間よりやや長いことも分かる。このことは、正の感情を有する投稿の拡散速度が中立の感情を有する投稿と比べてやや遅いことを示唆している。

以上の分析結果より、Twitter における投稿の内容から推定した投稿の有する感情（センチメント）は、投稿の拡散規模ならびに拡散速度を予測するのに有用な特徴量であることが示唆された。

3-3 ソーシャルネットワークにおけるコミュニティ構造と情報拡散の関係の分析

本研究課題では、ソーシャルメディアにおける情報拡散に影響を与える要因として、ユーザ間の関係を表現したソーシャルネットワークにおける「コミュニティ構造」（図 4）に着目する。多くのソーシャルネットワークは、互いに密に接続されたノードの「コミュニティ」と、それらのコミュニティ間を接続する少数のリンクで構成されるというコミュニティ構造を有している [13]。ソーシャルネットワークにおけるコミュニティは、興味の近い人同士のグループや、現実社会における何らかの集団を反映しているため、ソーシャルネットワークにおけるコミュニティと、人と人との交流の間には強い関係が存在することが報告されている [14]。文献 [14] は、電話、電子メール、SNS (Social Networking Services) におけるメッセージの 3 種類の交流の履歴を分析し、数百人から数万人程度の大規模なソーシャルネットワークにおいて、ユーザ間の交流の 8 ~ 9 割程度がコミュニティ内での交流であり、強いコミュニティ内交流局所性が存在することを示している。直感的には、ソーシャルメディアにおいても、ユーザは自身と同一のコミュニティに属するユーザの投稿を拡散させやすいという、コミュニティ内交流局所性が存在すると考えられる。ただし、ソーシャルメディアにおける情報拡散と、コミュニティ構造の関係を、実データに基づき分析した研究は限られている。

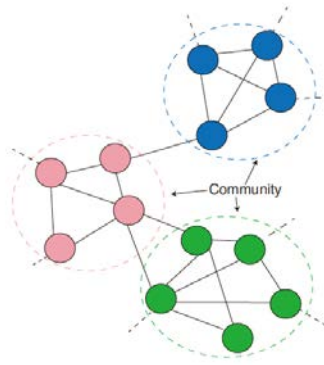


図 4: コミュニティ構造のイメージ [15]

本課題では特に、代表的なソーシャルメディアとして Twitter を分析の対象とし、コミュニティ構造が情報拡散に与える影響を明らかにする。

まず、分析の対象とするユーザを決定する。ここでは、ある程度活発にリツイートを行うユーザを分析対象とすることを意図して、次の手順により対象とするユーザを決定した。Twitter API (Application Programming Interface) を用いて、2013 年 12 月 11 日～ 17 日の 1 週間の日本語のリツイートを収集した。これには、Twitter REST API v1.1 の検索 API で $q=RT, lang=ja$ というクエリを用いた。その結果、14,220,864 件のツイート (オリジナルツイート) と、それらの 52,129,804 件のリツイートを得た。上記の期間においてリツイート数が 10 回から 100 回であるツイートを抽出し、それらのツイートを 10 件以上リツイートしたユーザを、分析の対象とした。これにより、356,453 ユーザが分析の対象となった。

次に分析対象の 356,453 ユーザの 2014 年 1 月上旬時点におけるフォロー関係を表現したソーシャルネットワークを取得する。Twitter API を用いて、2014 年 1 月 1 日～ 11 日の期間に、これらのユーザがフォローしているユーザのリストを得た。356,453 ユーザをノード、これらのユーザ間のフォロー関係をリンクとしたソーシャルネットワークを構築した。

次に、得られたソーシャルネットワークに対して、代表的なコミュニティ抽出アルゴリズムである Louvain 法 [16] を適用することにより、ユーザのコミュニティを得る。Louvain 法は大規模なネットワークに適したコミュニティ抽出アルゴリズムとして広く用いられている。コミュニティの抽出にあたっては、有向グラフであるソーシャルネットワークのリンクの向きを無視して無向グラフとした。

最後に、2014 年 1 月 1 日～ 31 日の期間における、対象ユーザ 356,453 人の投稿したリツイートを、Twitter API を用いて取得した。これにより、対象ユーザの投稿した 14,001,533 件のツイート (オリジナルツイート) と、それらの 63,462,387 件のリツイートを得た。これらのリツイートの履歴と抽出したコミュニティの情報を用いて分析を行う。

ツイートがリツイートされた時、その後のリツイート数の期待値を、リツイートしたユーザがツイートを投稿したユーザと同一コミュニティに属する場合、異なるコミュニティに属する場合で比較する。図 5 に、ツイートがちょうど N 回リツイートされた時点において、そのツイートがその

後リツイートされる回数の平均を示す。ここでは、 N 回目のリツイートを、ツイートの投稿者と同一のコミュニティに属するユーザが行った場合、異なるコミュニティに属するユーザが行なった場合で比較している。

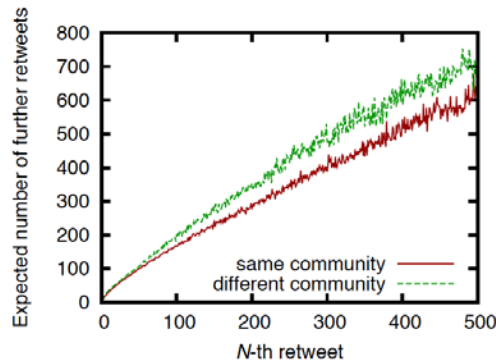


図 5: ツイートが N 回リツイートされた時点におけるその後のリツイート数の期待値 [17]

図 5 より、コミュニティをまたがってツイートが拡散された場合の方が、コミュニティ内で拡散された場合と比べて、その後のリツイート数の期待値が高いことが分かる。例えば、1 回目のリツイートが行われた時点におけるその後のリツイート数の期待値は、ツイートの投稿者とリツイートしたユーザの属するコミュニティが異なる場合で約 5.3、同一の場合で約 3.6 である。このようにツイート拡散の比較的初期の段階においても、異なるコミュニティに属するユーザによって拡散された場合と、同一のコミュニティに属するユーザによって拡散された場合で、その後のリツイート数の期待値が 1.5 倍程度異なることが分かる。なお、t 検定の結果、これらには有意な差が存在することを確認している ($p < 0.01$)。このことは、ソーシャルネットワークのコミュニティ構造が、投稿の拡散規模を予測する上で有用な特徴量であることを示唆している。

3-4 投稿の拡散規模予測

最後に、教師あり機械学習のアプローチによって、ツイートの拡散規模を予測するモデルを構築する。機械学習のアルゴリズムとして、Random Forest を用いる。

ここではユーザのコミュニティやユーザの影響力などのソーシャルネットワークにおける特徴量と、ソーシャルメディアにおける投稿から抽出した投稿の有する感情などを用いて、ツイートの拡散規模を予測するモデルを構築した。ツイートに対する 10 回目のリツイートが投稿された時点で、そのツイートの将来の拡散規模を予測する実験を行った。実験の結果、大規模に拡散するツイート（100 回以上リツイートされるツイート）を適合率 0.7 再現率 0.2 程度の精度で予測できることを確認した。

3-5 まとめ

本研究課題では、ソーシャルメディアにおける流行の予測に有用な特徴量について分析した。投稿に含まれる感情と、ユーザ間の関係を表現したソーシャルネットワークにおけるコミュニティの情報がソーシャルメディアにおける流行の予測に有用な情報であることを示した。さらに、これらの特徴を用いて、大規模に拡散する投稿を予測する実験を行い、その精度を検証した。

4 まとめと今後の課題

研究期間内において、(A) ソーシャルメディアデータのシグナル検出技術の研究、(B) ソーシャルメディアデータからの流行予測技術の研究の研究に取り組んだ。研究課題 (A) では、部分的なソーシャルネットワークの構造のみから影響力の強いユーザを推定する手法を検討した。さらに、影響力の強いユーザの推定精度を向上させるため、ソーシャルネットワークの構造を補完するためのリンク予測手法を提案した。研究課題 (B) では、テキストマイニング技術およびグラフマイニングの技術を用いて、ソーシャルメディアの流行予測に有用な特徴量を分析した。その結果、投稿の有する感情と、ソーシャルネットワークにおけるユーザのコミュニティがソーシャルメディアにおける流行の予測に有用な特徴であることを示した。さらに、ソーシ

ャルメディア上で大規模に拡散される投稿を予測する実験を行い、予測精度を検証した。

今後は、ソーシャルメディア上の流行と実社会における流行の関係を調査し、実社会における流行を予測する技術の開発を目指す。

【参考文献】

- [1] S. Asur, B. Huberman et al., “Predicting the future with social media,” in Proc. of WI-IAT, 2010, pp. 492–499.
- [2] H. Choi and H. Varian, “Predicting the present with Google trends,” *Economic Record*, vol. 88, no. s1, pp. 2–9, 2012.
- [3] E. Aramaki, S. Maskawa, and M. Morita, “Twitter catches the flu: detecting influenza epidemics using Twitter,” in Proc. of EMNLP, 2011, pp. 1568–1576.
- [4] W. Fan and M. D. Gordon, “The power of social media analytics,” *Communications of the ACM*, vol. 57, no. 6, pp. 74–81, 2014.
- [5] D. Gayo-Avello, “A meta-analysis of state-of-the-art electoral prediction from Twitter data,” *Social Science Computer Review*, vol. 31, no. 6, pp. 649–679, 2013.
- [6] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant, “Detecting influenza epidemics using search engine query data,” *Nature*, vol. 457, no. 7232, pp. 1012–1014, 2009.
- [7] L. Granka, “Using online search traffic to predict us presidential elections,” *PS: Political Science & Politics*, vol. 46, no. 02, pp. 271–279, 2013.
- [8] A. S. Maiya and T. Y. Berger-Wolf, “Benefits of bias : Towards better characterization of network sampling,” in Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD’11), Sep. 2011, pp. 105–113.
- [9] S. Tsugawa, and K. Kito, "Retweets as a predictor of relationships among users on social media," *PLoS ONE* 12(1): e0170279 19 pages, Jan. 2017.
- [10] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts, “Everyone’s an influencer: Quantifying influence on Twitter,” in Proceedings of the 4th ACM International Conference on Web Search and Data Mining (WSDM’11), Feb. 2011, pp. 65–74.
- [11] S. Stieglitz and L. Dang-Xuan, “Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior,” *Journal of Management Information Systems*, vol. 29, no. 4, pp. 217–247, 2013.
- [12] S. Tsugawa, and H. Ohsaki, "On the relation between message sentiment and its virality on social media," *Social Network Analysis and Mining*, Vol.7, No.1, pp.19:1–19:14, Dec. 2017.
- [13] M. E. J. Newman and M. Girvan, “Finding and evaluating community structure in networks,” *Physical Review E*, vol. 69, no. 2, p. 026113, Feb. 2004.
- [14] S. Tsugawa, and H. Ohsaki, "Community structure and interaction locality in social networks," *Journal of Information Processing*, Vol.23, No.4, pp.402-410, Jul. 2015.
- [15] 津川 翔, "ソーシャルメディアにおける情報拡散と重複コミュニティ構造の分析," 電子情報通信学会技術研究報告 (CQ2016-99), pp.53-58, Jan. 2017.
- [16] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast unfolding of communities in large networks,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, Oct. 2008.
- [17] 津川 翔, "Twitter における情報拡散とコミュニティ構造の関係に関する一考察," 電子情報通信学会技術研究報告 (CQ2016-54), pp.33-38, Aug. 2016.

〈発表資料〉

題名	掲載誌・学会名等	発表年月
On the relation between message sentiment and its virality on social media	Social Network Analysis and Mining	2017年12月(掲載決定)
Retweets as a predictor of relationships among users on social media	PLoS ONE	2017年1月
Estimating influence of social media users from sampled social networks	Proceedings of the Workshop on Social Influence	2016年8月
ソーシャルメディアにおけるユーザのコミュニティを用いた投稿の拡散規模予測に関する一検討	電子情報通信学会総合大会講演論文集	2017年3月
ソーシャルメディアにおける情報拡散と重複コミュニティ構造の関心の分析	電子情報通信学会技術研究報告	2017年1月
Twitterにおける情報拡散とコミュニティ構造の関心に関する一考察	電子情報通信学会技術研究報告	2016年8月