



Department of Risk Engineering  
Faculty of Systems and Information Engineering  
University of Tsukuba

# Twitter におけるデマユーザの 抽出を目的とした特徴空間の構成

2014年度リスク工学グループ演習5班

中川航至 ・ 播岡太郎

土方孝将 ・ 修兵

アドバイザー教員 遠藤靖典





# 発表の構成

2014 Risk Engineering Group Work Practice : Group5

## 1. 背景・目的

- 研究背景～Twitterの特徴
- 関連研究～現在のTwitterにおける研究方法
- 研究目的
- 研究の流れ

## 2. 分析指標の抽出

指標抽出方法

- 形態素解析
- 極性辞書

## 3. データ分析

- 分析手順
- 抽出指標の特徴把握(t検定)
- デマユーザの判別
- 特徴空間の構成

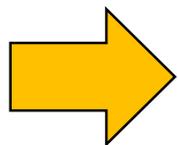
## 4. まとめ・課題

# 研究背景-Twitterの特徴

2014 Risk Engineering Group Work Practice : Group5

- Twitterとは世界中の人々が参加するSNSの一つ
- **アカウント数は2億以上**
- 情報を**リアルタイム**で多くのユーザに伝播or受信することが可能
- しかし不特定多数のユーザがおり、情報の正当性は保証されていない。
  - **デマ** (社会にとって不利益となりうる嘘の情報) が存在する
  - 情報はその信頼性に関わらず、**一気に伝播**してしまう。

実際にデマ情報が多く拡散されている。



**デマツイート**を判断できないか？



# 関連研究

## Twitter上のデマ情報の検出(遠藤ら、2011)

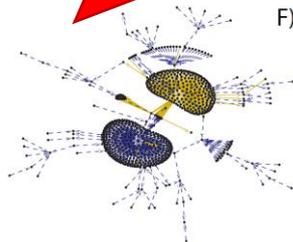
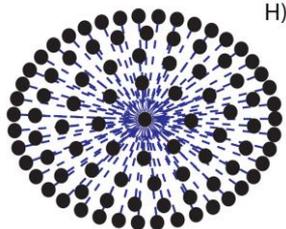
- ▶ ツイート文章
- ▶ 拡散ネットワーク
- デマとニュースの分類指標

ノイズクラスタリングによるデマの検出

	デマ	ニュース	全体
T	48 (89 %)	48 (89 %)	67 (80 %)
F	17 (20 %)	17 (20 %)	17 (20 %)

## マイクロブログ上のデマの検出 (Ratkiewiczら)

Truthyシステムを構成  
→ ツイートデータを監視、



現在  
ツイート情報からの  
デマツイート判定は困難

- ツイートの特徴を評価
- ツイート間のつながり
- ツイート数の多い
- 星のようなネットワークの形
- 2要素間のエッジの重みが大きい

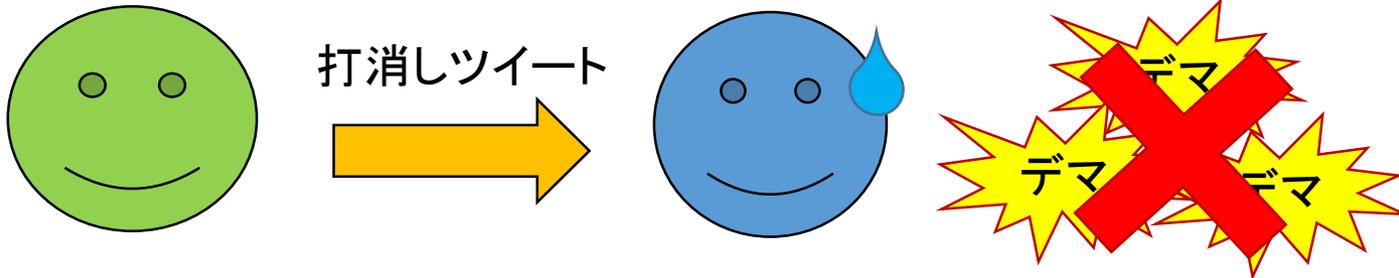


# Twitter機能の仕様変更

2014 Risk Engineering Group Work Practice : Group5

- Twitter

デマに対する耐性が出てきた



一度発信したツイートを完全に削除できるようになった

- Twitter API

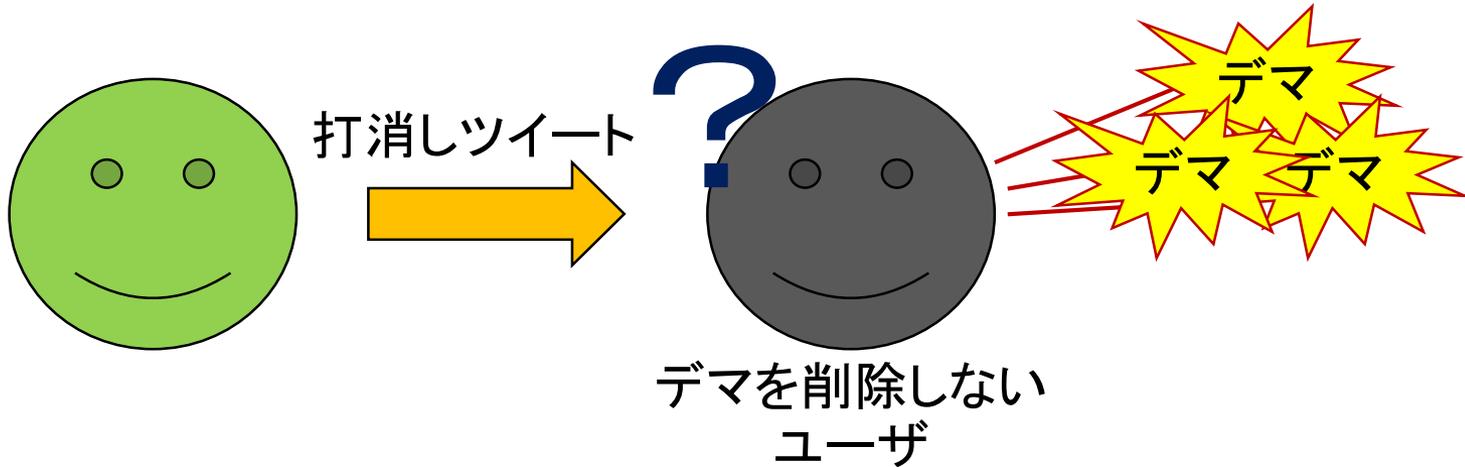
多くの関数が廃止された

制限がかかり、過去、約3000ツイートまでしか抽出できなくなった

→過去のデマツイートがうまく抽出できない



# 調査の目的



## 研究目的

ユーザ情報やそのユーザがツイートした文章(デマユーザの特徴)  
→デマユーザの特徴を分析、デマユーザの判定  
デマ拡散による社会的混乱を回避するための一助となること

- Twitterから得られるデータを元に、デマ情報とそれ以外の違いを探す。
- Twitter APIを用いてデータを取得

デマツイートを探し、通常のツイートとともに情報を取得することで特徴を数値化



抽出した指標をもとに、特徴空間を構成



ユーザの特徴を判断する



# 発表の構成

2014 Risk Engineering Group Work Practice : Group5

## 1. 背景・目的

- 研究背景～Twitterの特徴
- 関連研究～現在のTwitterにおける研究方法
- 研究目的
- 研究の流れ

## 2. 分析指標の抽出

指標抽出方法

- 形態素解析
- 極性辞書

## 3. データ分析

- 分析手順
- 抽出指標の特徴把握(t検定)
- デマユーザの判別
- 特徴空間の構成

## 4. まとめ・課題

# 指標抽出方法

2014 Risk Engineering Group Work Practice : Group5

- TwitterAPIの仕様変更によるツイート情報取得可能項目の減少。  
→ツイートの文章データやフォロワー数など,検出可能データを最大限に用いることで特徴を分析
- ユーザの定義
  - デマユーザ**: デマを流し, その後も削除していない危険ユーザ
  - 健全ユーザ**: デマ情報に対して, 打ち消しツイートをしているユーザ
  - メディアユーザ**: 新聞, テレビ, 雑誌などのユーザ
- 指標抽出
  1. 定義した分類に属する, それぞれ13のユーザを対象として, 過去ツイートの文章データやフォロワー数などを取得
  2. 取得したデータを用いて, 形態素解析ならびに極性辞書評価を行う



# 形態素解析

## 形態素解析～MeCab

文章を単語に分解し品詞を推定する

```

二郎はにんにくマシマシで美味しかった
二郎 名詞,固有名詞,一般,*,*,*,二郎,ニロウ,ニロー
は 助詞,係助詞,*,*,*,は,ハ,ワ
にんにく 名詞,一般,*,*,*,*,にんにく,ニンニク,ニンニク
マシマシ 名詞,一般,*,*,*,*,*
で 助詞,格助詞,一般,*,*,*,で,デ,テ
美味しかっ 形容詞,自立,*,*,形容詞・イ段,連用タ接続,美味しい,オイシカッ,オイシカッ
た 助動詞,*,*,*,特殊・タ,基本形,た,タ,タ
EOS

```

### MeCabの出力

単語、品詞、読み、活用の種類、活用語幹 など

⇒ 活用語幹を利用して辞書とマッチングさせる



# 極性評価表現

## 単語感情極性対応表を用いた極性評価と極性値の算出

◎ 物事に対する評価を表しているような言葉=評価表現

例) 高い 怖い すごい ウザい 美味しい けまらしい

◎ 評価表現は大きく

「ポジティブ」と「ネガティブ」に分けることができる

ポジティブ例) 美味しい 凄い 快適 生きてて良かった!

ネガティブ例) 不味い 退屈 不快 生まれてくるんじゃなかった...

◎ ポジティブ度, ネガティブ度の定義(数値化)が可能!

[単語感情極性対応表]

[http://www.lr.pi.titech.ac.jp/~takamura/pndic\\_ja.html](http://www.lr.pi.titech.ac.jp/~takamura/pndic_ja.html)



# 指標の抽出

➤ デマユーザの特徴を把握したい

→ 健全ユーザ、デマユーザ、メディアユーザ特徴を定量化

(※メディアユーザは健全ユーザに近いのでは?)

対象ユーザ数	39ユーザ(健全13, デマ13, メディア13)	
取得ツイート数	6900	
取得指標	フォロワー数/フォロワー数 単語数 極性/ツイート数 肯定語/ツイート数 否定語/ツイート数 肯定語+否定語/ツイート数 名詞数/ツイート数 助詞数/ツイート数 動詞数/ツイート数	形容詞数/ツイート数 接頭詞数/ツイート数 記号数/ツイート数 肯定語+否定語/全単語数 助詞率/全単語数 動詞率/全単語数 形容詞率/全単語数 接頭詞率/全単語数 記号率/全単語数



# 発表の構成

2014 Risk Engineering Group Work Practice : Group5

## 1. 背景・目的

- 研究背景～Twitterの特徴
- 関連研究～現在のTwitterにおける研究方法
- 研究目的
- 研究の流れ

## 2. 分析指標の算出

- 指標抽出方法
- 形態素解析
- 極性辞書

## 3. データ分析

- 分析手順
- 抽出指標の特徴把握(t検定)
- デマユーザの判別
- 特徴空間の構成

## 4. まとめ・課題

- 選定した指標をもとに健全, デマ, メディアで一元配置分散分析  
→メディア or 健全&デマ の傾向に分かれてしまう.  
プロ or 一般人で傾向が違う(?)

目的: デマユーザの特徴を把握する



- I. メディアは統計分析からは除外, 健全とデマで分析を行う.
- II. 各指標の平均値から特徴を把握.
- III. 把握した指標をもとに, 判別分析を実施.  
影響の大きさと判別モデルを構築

# 分析①-抽出指標の特徴把握

- 指標の特徴を理解するためにt検定を実施.

	t	df	平均値の差
フォロワー数/フォロー数	-0.17	24	-2.242
単語数	-1.95	24	-9.593
極性	1.70	24	1.644
肯定語	-1.18	24	-0.119
<b>1 ツイートあたり</b> 否定語	<b>-2.46</b>	<b>24</b>	<b>-3.403*</b>
肯定語+否定語	<b>-2.43</b>	<b>24</b>	<b>-3.461*</b>
名詞数	<b>-2.95</b>	<b>24</b>	<b>-1572.0**</b>
助詞数	-1.55	24	-373.692
動詞数	-1.32	24	-146.615
形容詞数	-0.18	24	-2.231
接頭詞数	<b>-3.14</b>	<b>24</b>	<b>-28.077**</b>
記号数	<b>-2.42</b>	<b>24</b>	<b>-309.385*</b>
<b>全単語数あたり</b> 肯定語+否定語	-0.89	24	-0.143
助詞率	0.68	24	0.010
動詞率	1.38	19.021	0.011
形容詞率	<b>2.45</b>	<b>24</b>	<b>0.004*</b>
接頭詞率	<b>-2.47</b>	<b>24</b>	<b>-0.001*</b>
記号率	-1.22	24	-0.010

\*:p<0.05, \*\*:p<0.01



# 分析②-ユーザの特徴とその影響

2014 Risk Engineering Group Work Practice : Group5

- t検定にて特徴が抽出された7つの指標を用いて、判別分析を行った。
- 方法:ステップワイズ法により影響の強い変数を理解した後、探索的に最も良い変数を強制投入法により分析。  
(目的変数:デマ→1, 健全→0として分析)

	正準判別関数係数	
	標準化された	標準化されていない
接頭詞数	0.682	0.03
否定語数	0.228	0.07
名詞数	0.177	$0.13 \times 10^{-3}$
(定数)		-2.98

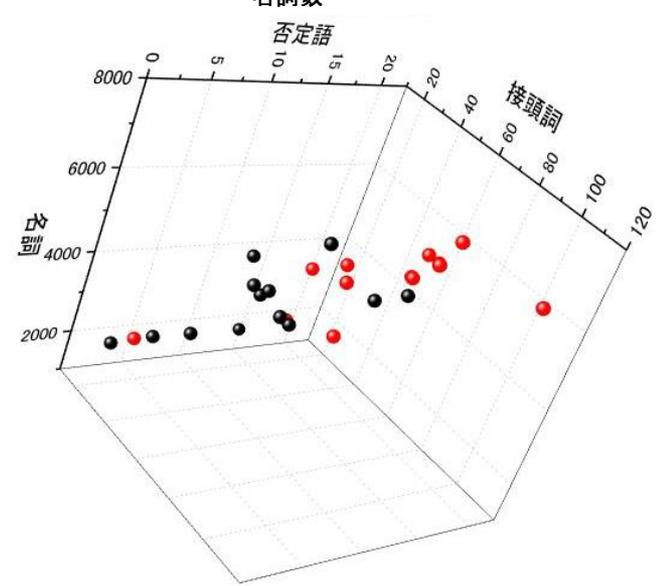
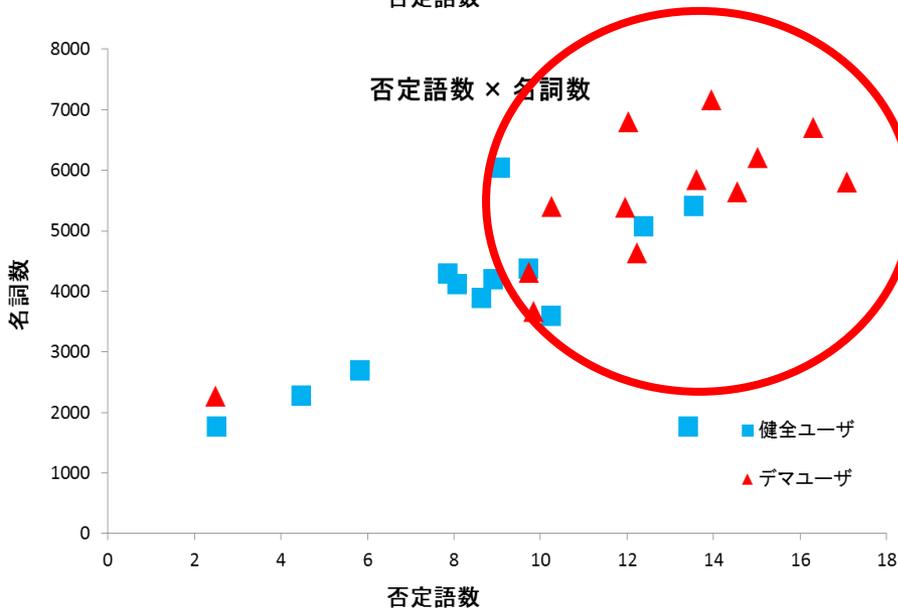
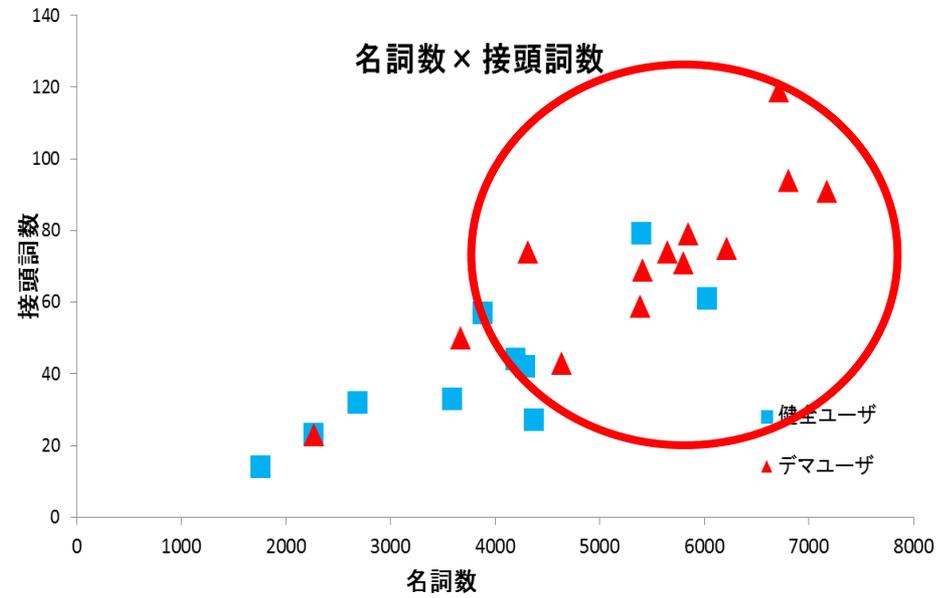
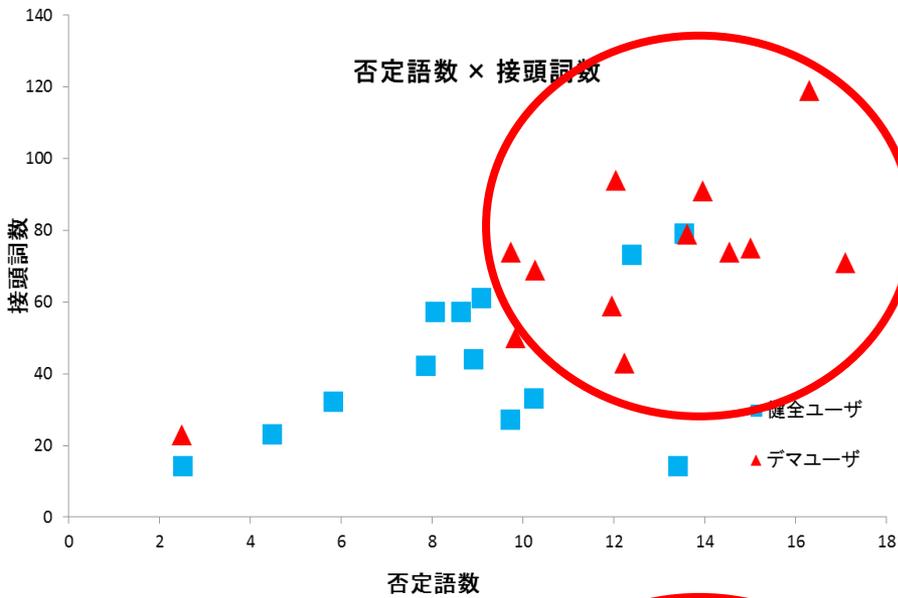
固有値0.43,正準相関0.55, $\lambda=0.70$ , $\chi^2=8.10$ , $p=0.044^*$

最も影響が強いのが接頭詞数であり、  
デマユーザほどツイート文中に用いている



# 特徴空間構造の可視化①

2014 Risk Engineering Group Work Practice : Group5



# 考察①

- デマユーザは、接頭詞、否定語、名詞が多い。  
実際のツイートで確認してみる。
- ① 接頭詞×否定語に着目すると、“不”や“非”，“脱”，“反”
- ② ①を踏まえて接頭詞×名詞に着目すると，“非武装”，“脱原発”，“反日”，“不買”，また，そもそも情報量が多い。
- ③ 否定語×名詞に着目すると，“捏造”，“制裁”，“批判”

考察：“信条”や“信念”を強く発信しているユーザでは？

右か左，どちらかに片寄りのあるユーザでは？

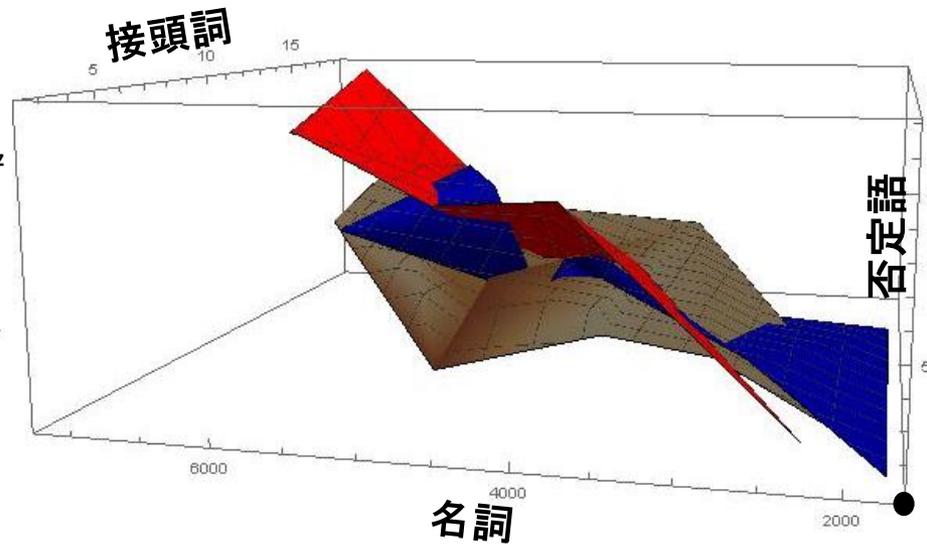
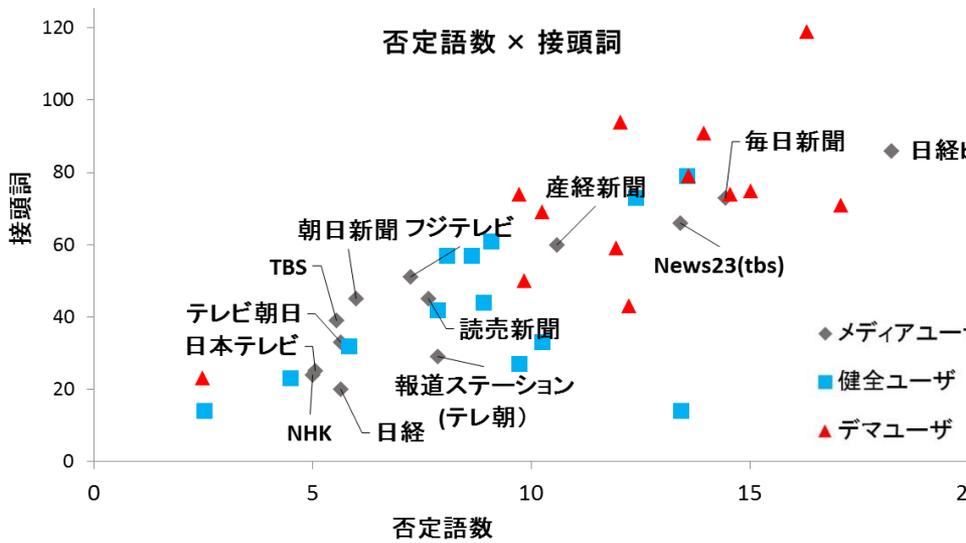
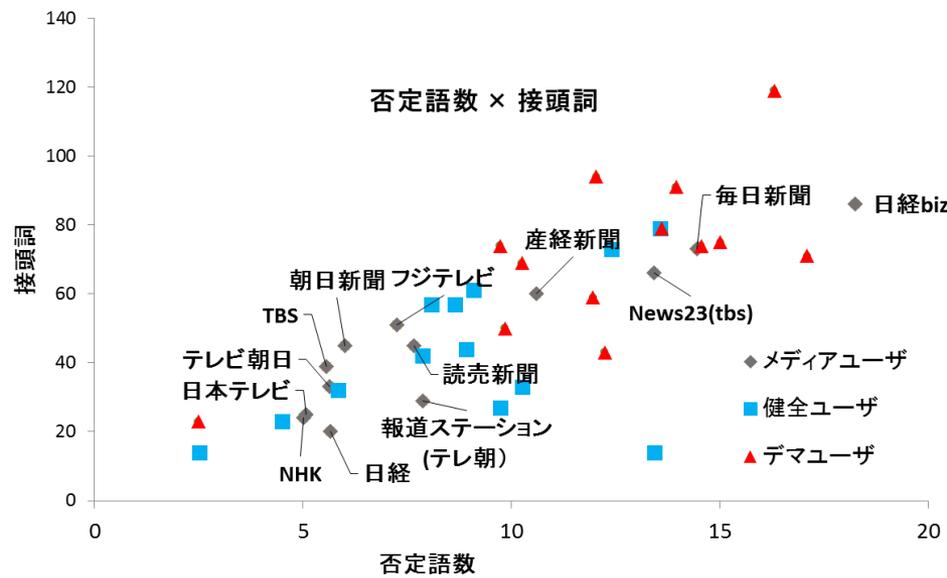
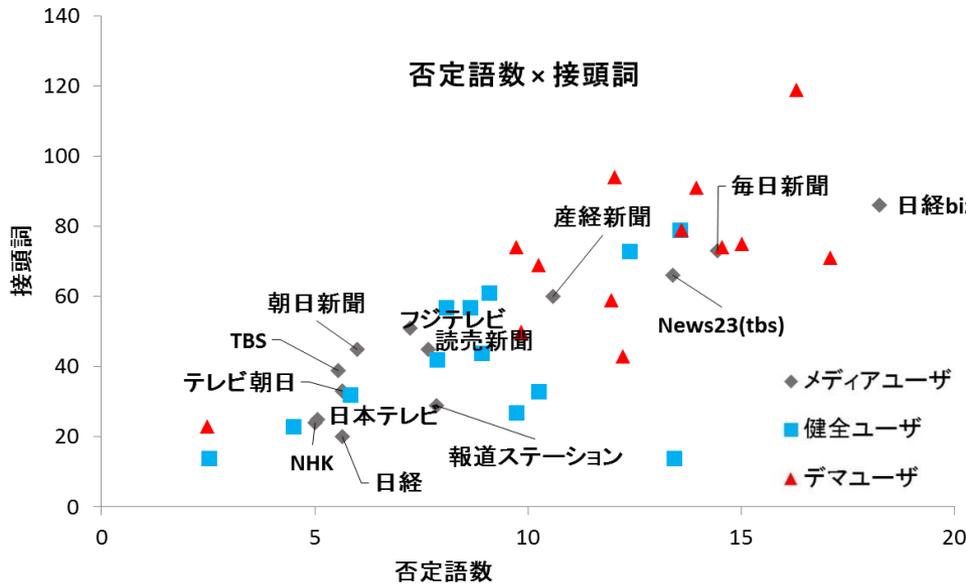
【今回定義したデマユーザ】

デマを流したまま削除しない。たとえデマであっても，そこに“信条”や“信念”があるため削除しない特異な傾向を持っているのではないか。



# 特徴空間構造の可視化②

2014 Risk Engineering Group Work Practice : Group5





## 1. 背景・目的

- 研究背景～Twitterの特徴
- 関連研究～現在のTwitterにおける研究方法
- 研究目的
- 研究の流れ

## 2. 分析指標の算出

- 指標抽出方法
- 形態素解析
- 極性辞書

## 3. データ分析

- 分析手順
- 抽出指標の特徴把握(t検定)
- デマユーザの判別
- 特徴空間の構成

## 4. まとめ・課題

- ツイート情報からユーザの特徴空間の構成を把握することができる可能性が示された.
- ユーザによって, ツイート使用目的, ツイート方法に違いが存在する.
- メディアユーザはそもそも異なる傾向があるが, 指標によっては, デマユーザの特徴と似ているユーザがいる.

デマ情報を削除せず, Twitter上で発信し続けるユーザ,  
今後同様のことを起こす可能性があるユーザの特徴空間を構成



今回の指標では, ある程度デマユーザの抽出が可能に



ただ, 一般化するにはより多くの指標を考える必要がある.

- 取得ツイート数・・・デマユーザ情報が非常に少ない.
- その他の指標を考慮・・・形態素解析, 極性辞書からのデータだけでなく, ユーザのプロフィール情報やフォロワー・フォロワーユーザの情報などからも考慮する必要がある.
- デマユーザ度の判定・・・1か0かの判断であり, “どちらに近い分類であるか”としかできない.
- ツイートとの関連・・・デマユーザ度だけでなく, そのツイート自体の信頼性を考慮しなくては, 発信された情報の信頼性向上には届かない.



# 参考文献

2014 Risk Engineering Group Work Practice : Group5

- [1] “日経BP「Twitter ユーザは新興国で急伸, 4年後にアジア太平洋地域が4割に」”,  
<http://itpro.nikkeibp.co.jp/article/NEWS/20140528/559804/>, 2014.6.2. 確認
- [2] “Twitter上のデマ情報の検出”, 齊藤和孝, 田嶋脩平, 中村裕, 遠藤靖典, リスク工学グループ演習, 2011
- [3] “Detecting and Tracking the Spread of Astroturf Memes in Microblog Streams”, Jacob Ratkiewicz, Michael Conover, Mark Meiss, Bruno Goncalves, Snehal Patil, Alessandro Flammini, Filippo Menczer, (学会名), (ページ数)
- [4] “MeCab”, 京都大学情報学研究科, 日本電信電話株式会社コミュニケーション科学基礎研究所, 共同研究ユニットプロジェクト
- [5] “Togetter”, <http://togetter.com/>



Department of Risk Engineering  
Faculty of Systems and Information Engineering  
University of Tsukuba

ご清聴ありがとうございました。

