

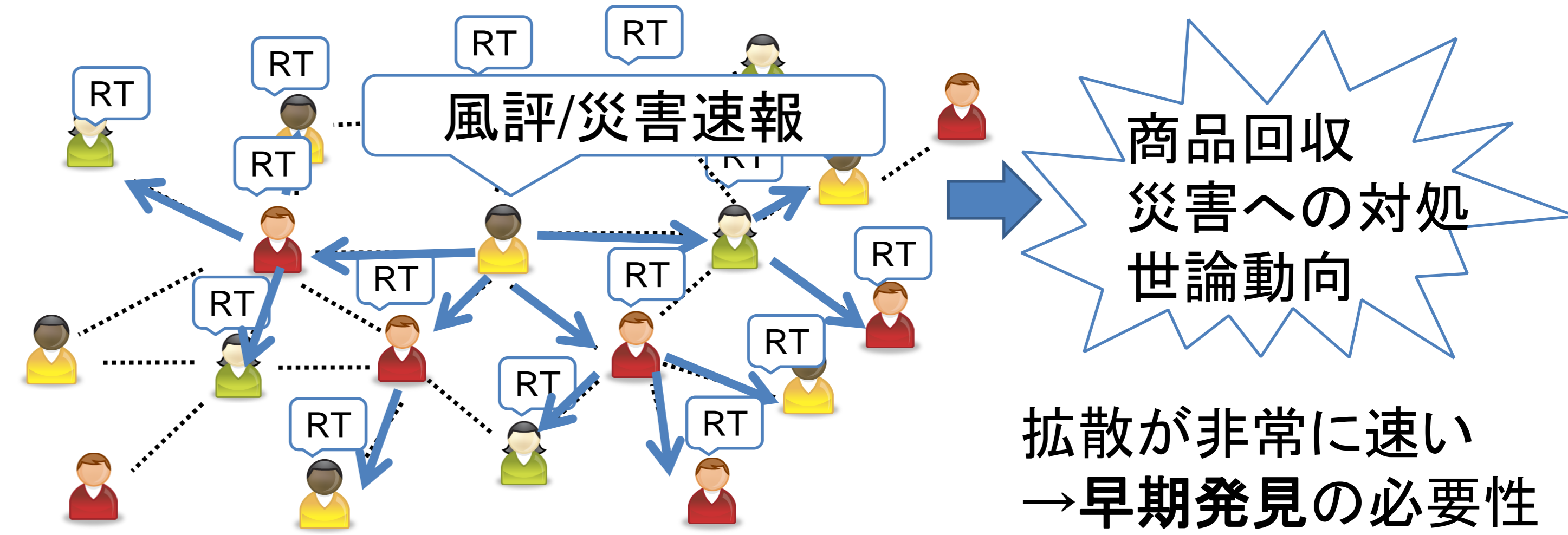
# マイクロブログにおける情報カスケードの社会的影響力に基づく分類手法に関する検討

川本貴史, 豊田正史(東大), 吉永直樹(東大,NICT)({kawamoto,toyoda,ynaga}@tkl.iis.u-tokyo.ac.jp)

## はじめに

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。

Task : **社会的影響力を持つ情報カスケードの(早期)発見**



## 社会的影響力に基づく情報カスケードのアノテーション

### 社会的影響力のあるカスケード

個人/組織/業界/社会に対する意見に対する反応

- その意見に対し**個人**, **組織**, **業界**が対応する必要

<人名>が授賞式に行くのにご本人の飛行機はエコノミー、ホテルは一泊12000円だったのに、必要もないのについてった英語もしゃべれない**渉外部10人のオヤジ**らの飛行機はビジネス、ホテルは一泊5万円だったんだと! ('D') 言いふらしてやってくれと言われたので、言いふらすわ!

<URL> 2月3日の記事からご覧ください。全国チェーン、<社名>の対応です。私は許すことは出来ませんし、今後、関わることは一生ないでしょう。こんな会社が拡大することを望みませんし、全国のライダーにこの事実が届くことを切に願う。

- 世論**を反映している

「ゲームでなら人殺しができる!」という人と、「ゲームの中でも人殺しはちょっと……」という人では、明らかに後者の方が「ゲームと現実の区別がついていない」のだが、**世間的**には何故か前者こそが「ゲームと現実の区別がつかなくなって殺人を犯す!」と囃し立てられる。

影響力のある出来事(事実)の周知

- 事件・事故の速報**や知られていない**問題の指摘**

【**北朝鮮核実験情報**】本日11時19分頃、気象庁が北朝鮮を震源とする地震波を観測、自然地震ではない可能性があります。北朝鮮による核実験の可能性もあるので、官邸対策室を設置しました。今後も随時情報をお知らせします。

- 世論動向**や**注意喚起**・**デマ訂正**, **啓蒙**

【重要】昨日辺りから流行ってる「**日頃の行いおみくじ**」、どうやら勝手にフォローを行ったリツイートをする権限も解放してしまう**スパムらしい**。使ってしまった人は承認アプリリストから速やかに削除した方がいい。削除ページは公式サイトのココ→<URL>

### アノテーション基準

アノテータの用いる手がかり

- ツイートの本文
- ツイートに含まれるURLのリンク先のメディア

個人/組織/業界/社会に対する意見に対する反応

→ **意見の対象の粒度, 明確性**により分類

**影響力のある出来事(事実)の周知**

→ 周知している出来事の**事件性**, その出来事がどのような人, 集団に対して**重要なのか**をもとに分類

### アノテート結果

意見に対する共感・反感

粒度	個人	組織	業界	社会
明確	33	48	36	46
不明確	32	16	36	46

月ごとの比較

	1月	2月
共感・反感	147	69
事実の周知	64	43
影響力無	935	363

影響力のある事実の周知

種類	注意喚起		
	速報	デマ訂正	啓蒙
個数	32	42	33

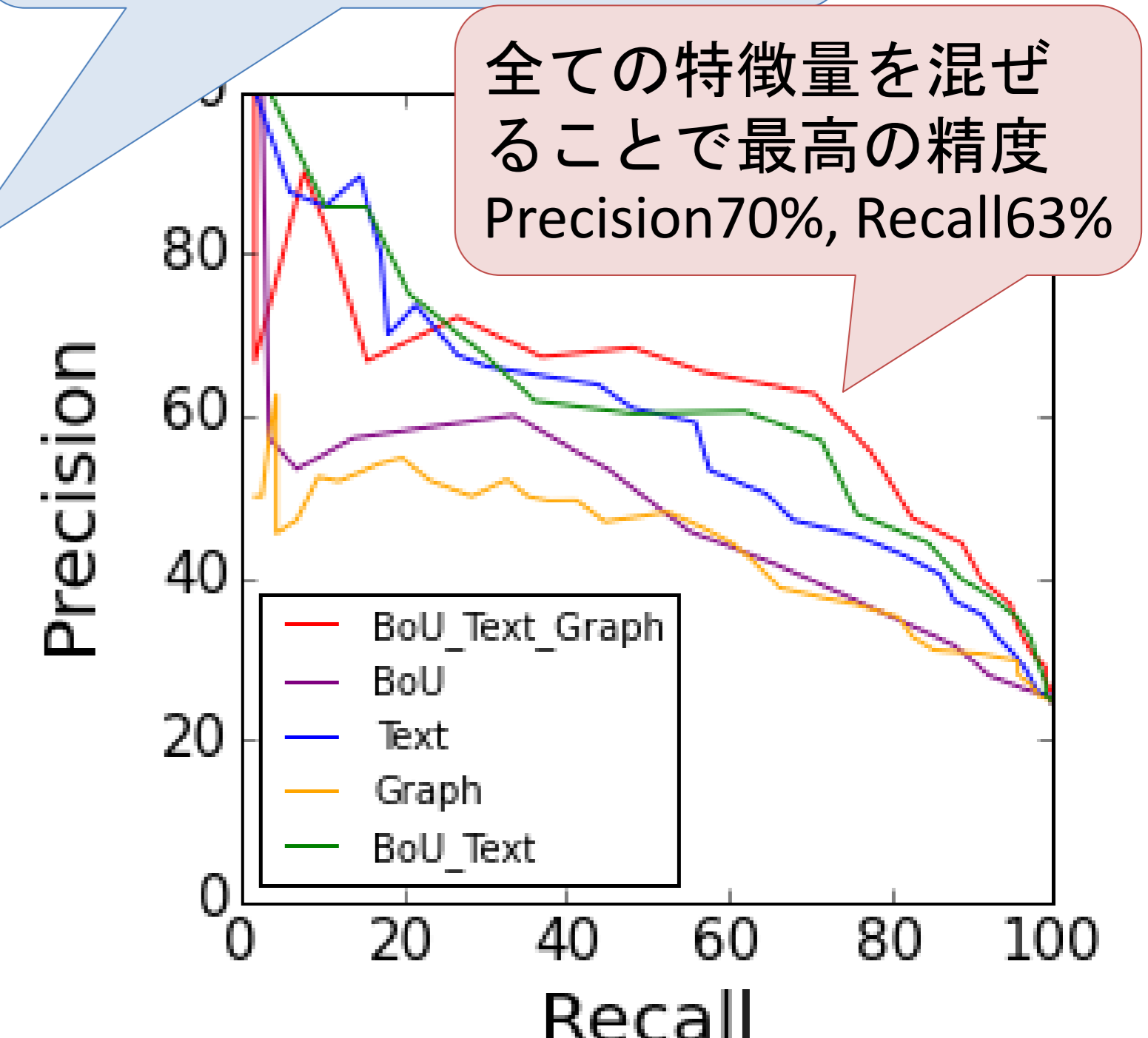
## 自動分類の実験

初期のNリツイートに限定してカスケードを分類

- Training Data:** 2013/01, 1146 Cascades
- Test Data:** 2013/02, 475 Cascades
- Classifier:** RBF kernel SVM
- Feature:**
  - Text:** Bag of Words + 固有表現の有無 + URLの有無
  - Graph:** カスケードのグラフ特徴量 (degreeの分布, 深さ, クラスタリング係数, etc)
  - Bag of Users:** カスケードに参加するユーザ集合の特徴. ユーザごとのRTの内容の偏りを捉える.

特徴量	カスケードサイズ				
	50	100	200	400	600
BoU	0.51	0.50	0.63	0.66	0.71
Text	0.58	0.58	0.58	0.58	0.58
Graph	0.51	0.51	0.55	0.54	0.55
BoU + Text	0.63	0.66	0.70	0.70	0.72
All	0.66	0.66	0.70	0.69	0.72

カスケードサイズ大: BoU効果大(単独で良精度)



カスケードサイズ=50の詳細

カスケードサイズ小: 多様なFeatureが必要