



社会的トピックが持つ 株価への影響に関する分析

澁谷 江梨奈

研究背景

どの株銘柄を
購入しよう？



ニュース記事



投資家



社会的トピックが株銘柄に対する
影響の与え方



SNS



データセット

2014年～2015年のデータを対象

毎日新聞社の
1面・2面・経済面記事
35,935件

銘柄

日経シソーラスの
経済用語が含まれる
ツイート279,195件

TOPIX100に含まれる
銘柄の内94銘柄の終値

ニュース記事



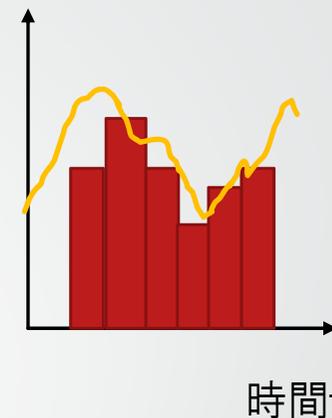
SNS

提案手法

トピックモデル
LDA



抽出されたトピックを
期間毎に集計して時系列化



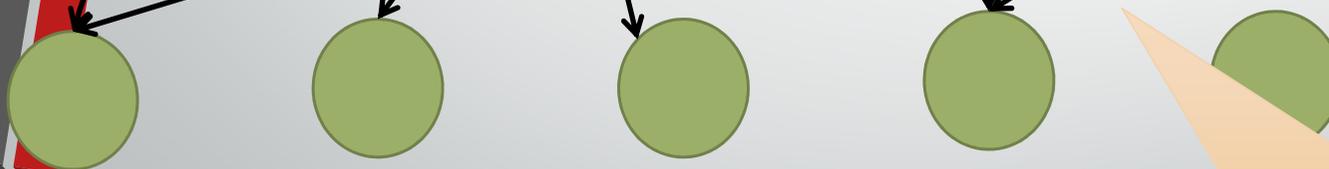
トピック1

トピック2

トピック3

トピック4

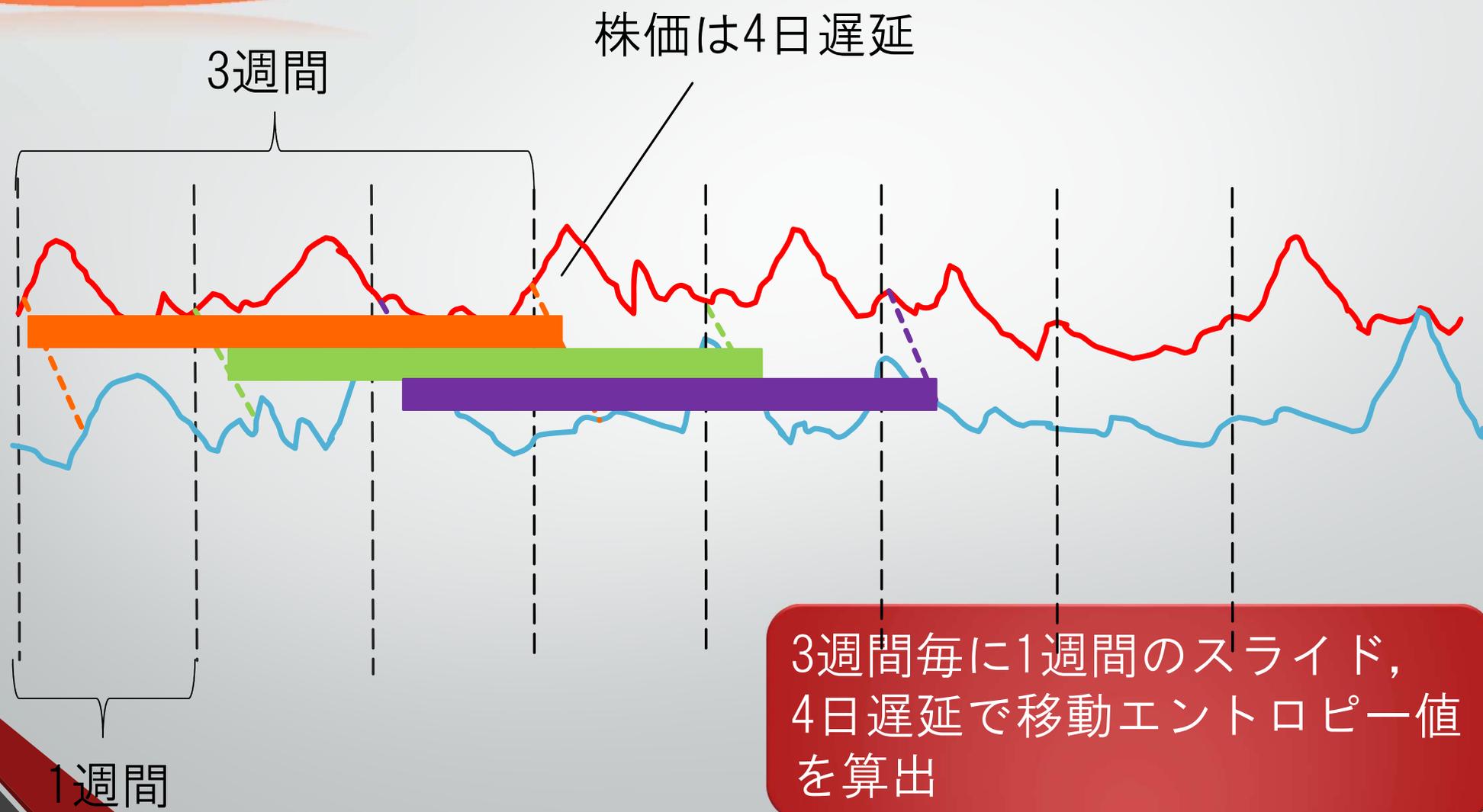
...



銘柄

移動エントロピー
を用いて算出

移動エントロピー



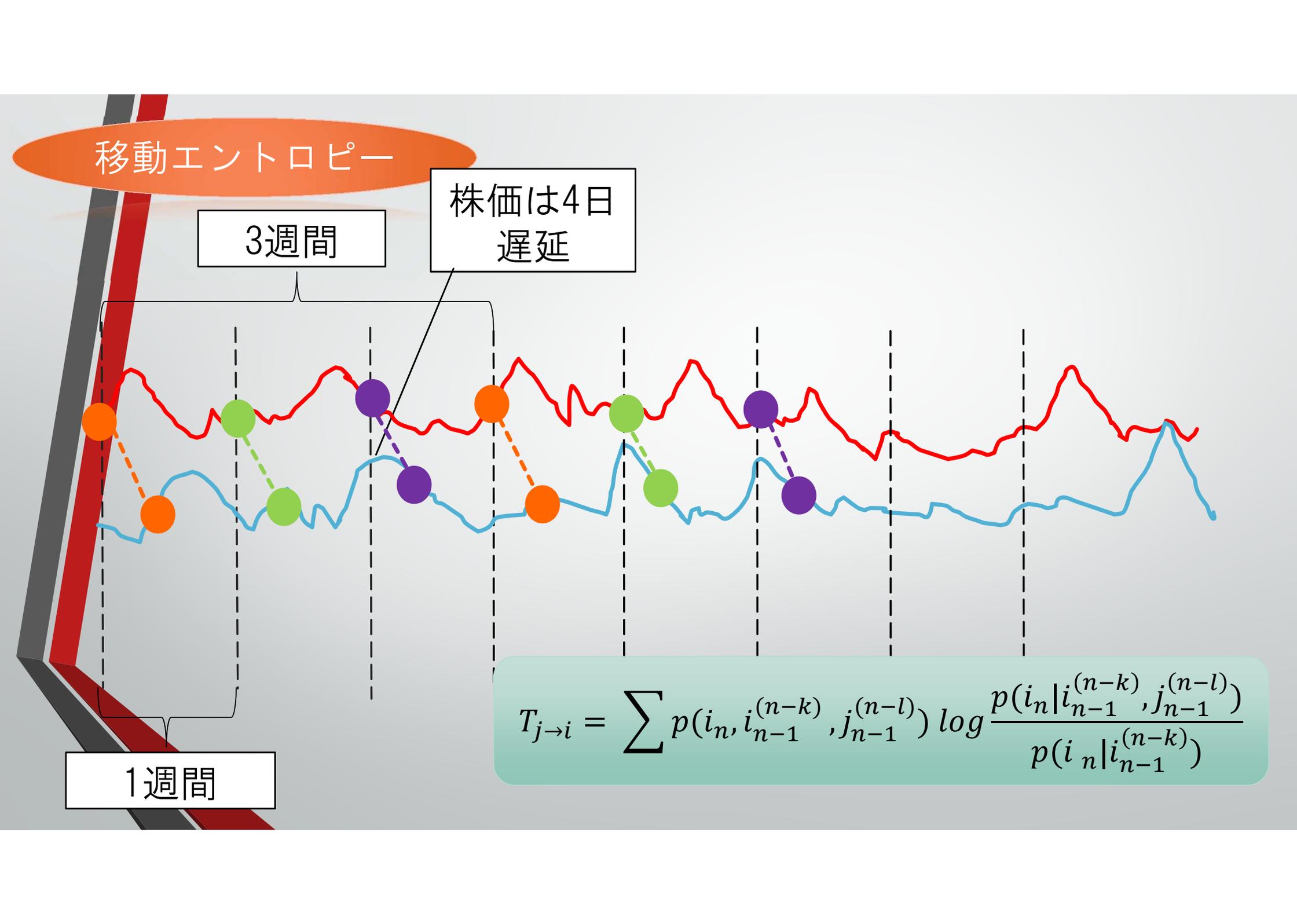
移動エントロピー

3週間

株価は4日
遅延

1週間

$$T_{j \rightarrow i} = \sum p(i_n, i_{n-1}^{(n-k)}, j_{n-1}^{(n-l)}) \log \frac{p(i_n | i_{n-1}^{(n-k)}, j_{n-1}^{(n-l)})}{p(i_n | i_{n-1}^{(n-k)})}$$



ニューストピック ニューストピック数=30(予備実験によるperplexity値より)

トピック0	関税	官僚	経済	t p p	分野	政府	環太平洋	パートナーシップ	農業	コメ
トピック2	容疑	事件	男性	警察	ない	当局	事故	女性	被害	弁護士
トピック16	自治体	災害	国	住民	地震	被害	地域	住宅	県	学校

Twitterトピック Twitterトピック数=50(予備実験によるperplexity値より)

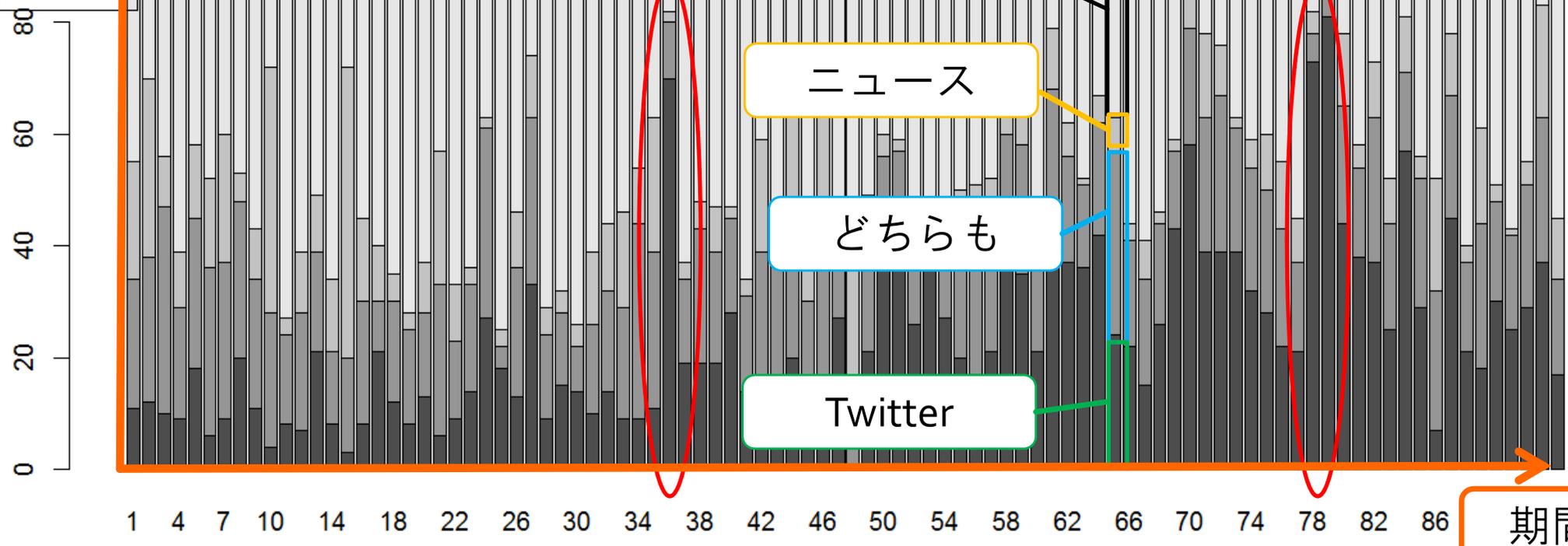
トピック36	店舗	商品	ヤフオク	複数	部分	別	肩	数字	黄金	お客様
トピック3	マネー	電子	ゲーム	サイト	ビジネス	オンライン	手	ネット	ポイント	パック
トピック33	銀行	強盗	バス	いい	大学	ない	郵便	スーパー	事務所	ポイント

集計結果



銘柄数

mid
twee



期間

集計結果

トピック

銘柄

銘柄

銘柄

トピックが銘柄に与えている影響の総量

		1	2	3	4	5
期間 36	Twitter	29.923	26.044	25.733	20.273	18.637
	ニュース	21.859	14.469	11.652	10.781	9.182
期間 77	Twitter	96.361	80.142	75.411	57.934	47.500
	ニュース	62.843	36.024	32.188	28.592	22.739
期間 78	Twitter	122.276	47.405	46.850	43.520	43.114
	ニュース	43.143	39.053	31.233	29.156	27.263

分析1 結果まとめ

集計結果から、

- 特定の時期に関してTwitterのトピックに反応する銘柄が多い。
- 全体で見ると、多くの銘柄はTwitter、ニュースのどちらかだけに常に影響を受けていることはない。

コミュニティ抽出

EdgeBetweenness
コミュニティ

トピック1

トピック2

トピック3

トピック4

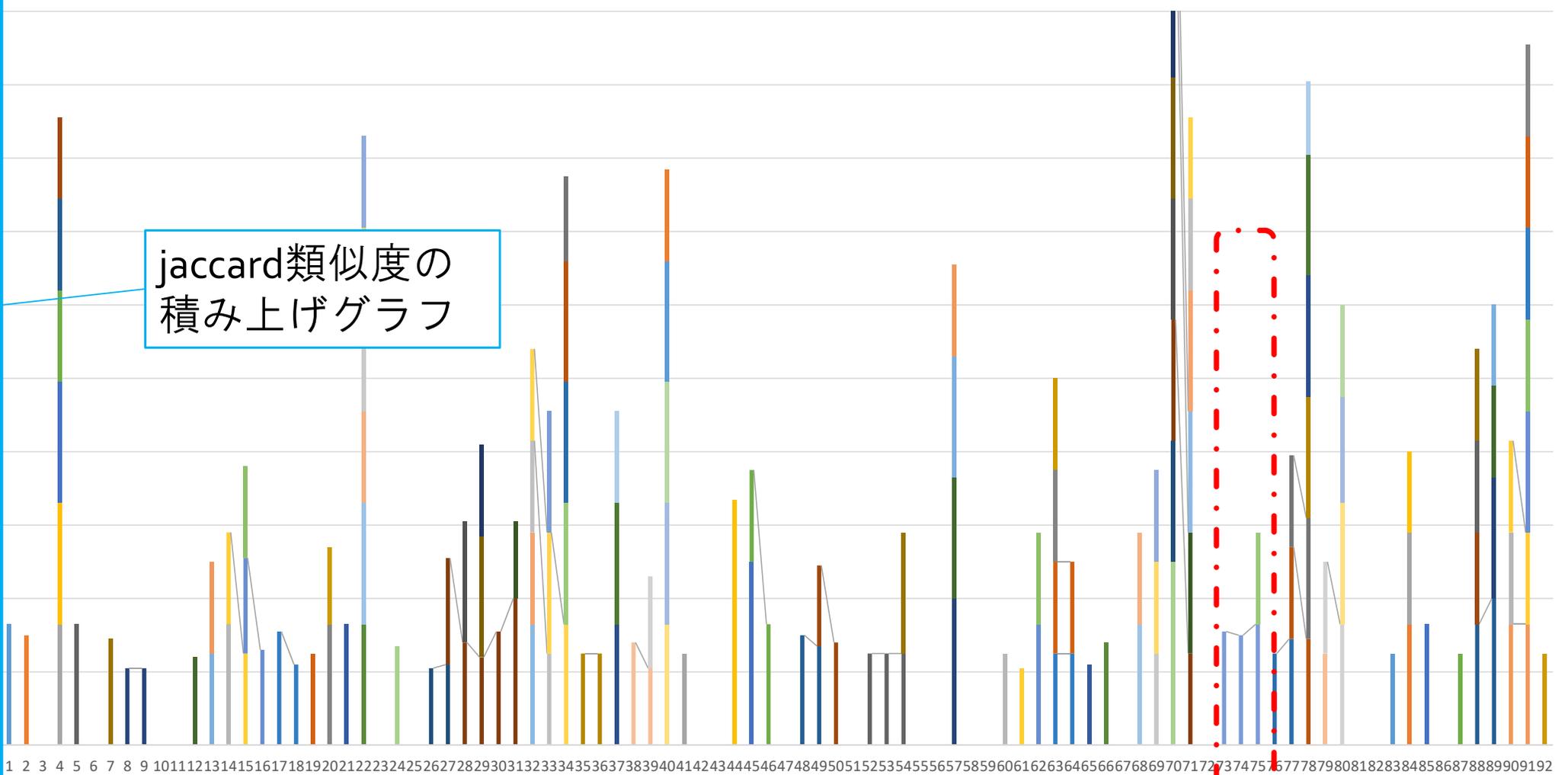


時間推移をjaccard距離により求める

抽出されたコミュニティの例

Twitterコミュニティ抽出の結果

jaccard類似度の
積み上げグラフ



類似度が高かった

コミュニ

Twitter

トピック28
月, 料金, 価格, チ
ケット, 本, 他, 株主,
平成, 無料, 期限

トピック16
自治体, 災害, 国, 住
民, 地震, 被害, 地域,
住宅, 県, 学校

“トピック28”
“kabuka-4063”

“トピック16”
“kabuka-7202”

“トピック28” “トピ
“kabuka-4063”

株価 (4063)
信越化学工業

“トピック16”
“kabuka-7202”

株価 (7202)
いすゞ自動車

まとめ

- 銘柄が一定してトピックに依るのではなく，時期によってTwitter寄りニュース記事寄りになる。
- コミュニティ抽出に関しては，期間毎の連続したコミュニティも多くは見られなかった。

今後の課題

- トピック抽出を動的にする.
- 2部グラフにおいて“銘柄→トピック”の影響も考慮する.