2017.3.15
【修士論文 研究報告】

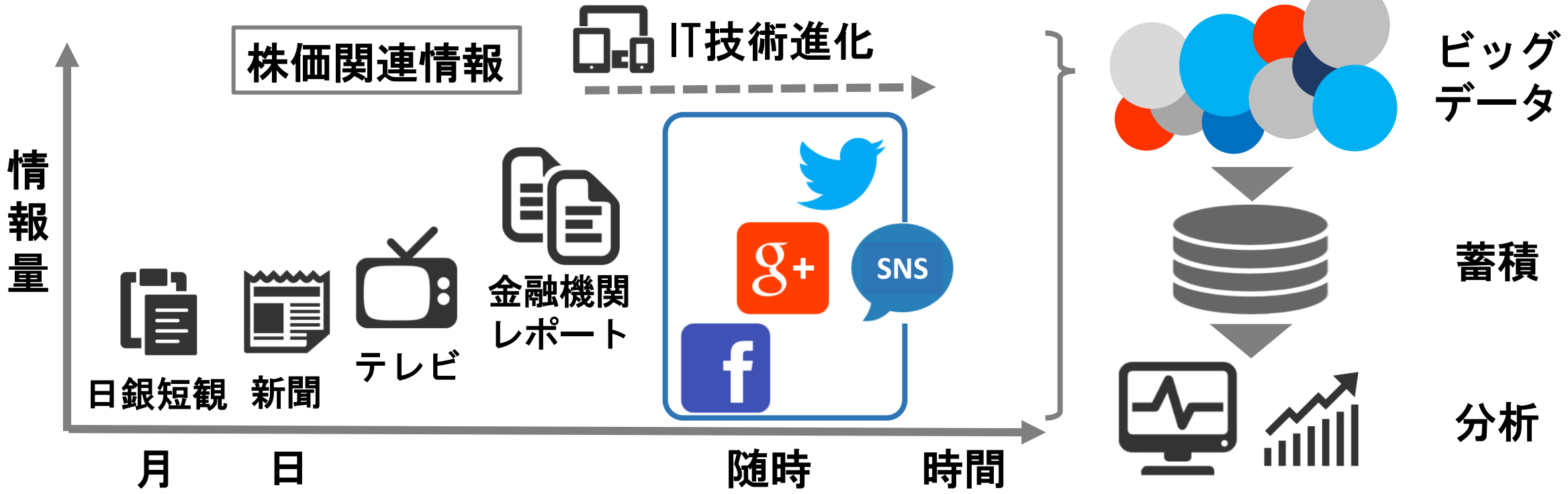
「 SNS情報と高頻度取引データを用いた株価リターンの実証分析」

～Twitterと高頻度取引との関係～

一橋大学大学院国際企業戦略研究科
金融戦略・経営財務コース
磯部 好孝

はじめに

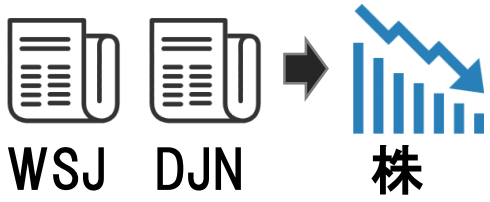
- 今日の株式市場における取引において、金融機関や個人投資家は**様々な情報を元に判断**し、**投資活動**を行っている。
- ＩＴ技術の発達に伴い、発信元が組織から個人へと拡大し**情報量**も**スピード**も**飛躍的に増加**。
- 金融業界においてもＩＴ化にともない莫大なデータが蓄積され、**ビッグデータ**を使い**株式市場の分析**が行われ始めている。



先行研究

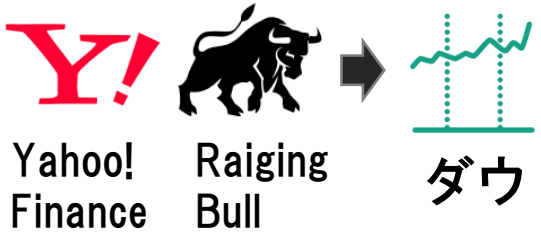
米国市場

・Tetlock（2008）



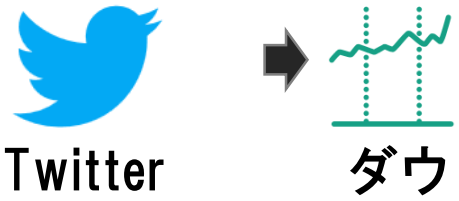
ネガティブな単語の出現頻度が個別企業の株価とリターン低下を予測

・Antweiler and Frank（2004）



投稿数が増加すると株価が変動する確率に影響を与え、確率の上昇を予測

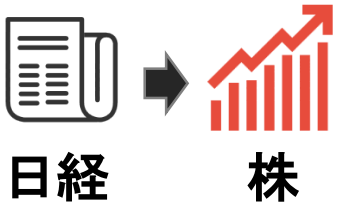
・Bollen et al.（2011）



語句を6つの感情に分類。各出現頻度を計測することで3日後の指数の動向を予測

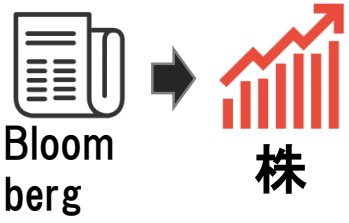
日本市場

・沖本・平澤（2014）



ニュース指標（ネガ・ポジ）を利用
翌営業日の株式リターンや
出来高に対して有意な関係を観測

・岡田・羽室（2011）



センチメント指標（ネガ・ポジ）を利用
株は心配しているときに上昇し、
楽観的になったときに下落

・丸山他（2008）

掲示板指標（数・内容）を利用
強気投稿数と弱気投稿数の差が
株式のリターンやボラティリティ、
出来高に影響

仮説

背景

・先行研究より各メディアと株式市場との間には何らかの関係があることを示唆

・高頻度取引においても、従来の取引同様にメディアと株式市場との間に何らかの関係があるのではないかと考えた

着目

・マーケットのベンチマークとして個別株式のリスクリターンの関係を表す資本資産評価モデル（CAPM）のα※1, β※2とSNSとの関係に着目

※1･･･超過期待投資収益率　※２･･･企業銘柄と株式市場の連動性を示すリスク指標

仮説

仮説１：「SNSの内容がポジティブの場合αはプラスになり、ネガティブの場合αはマイナスになる」

仮説２：「SNSの情報量が多いと、情報が行き渡りβは小さくなり SNSの情報量が少ないと、情報が行き渡らずβは大きくなる」

分析対象

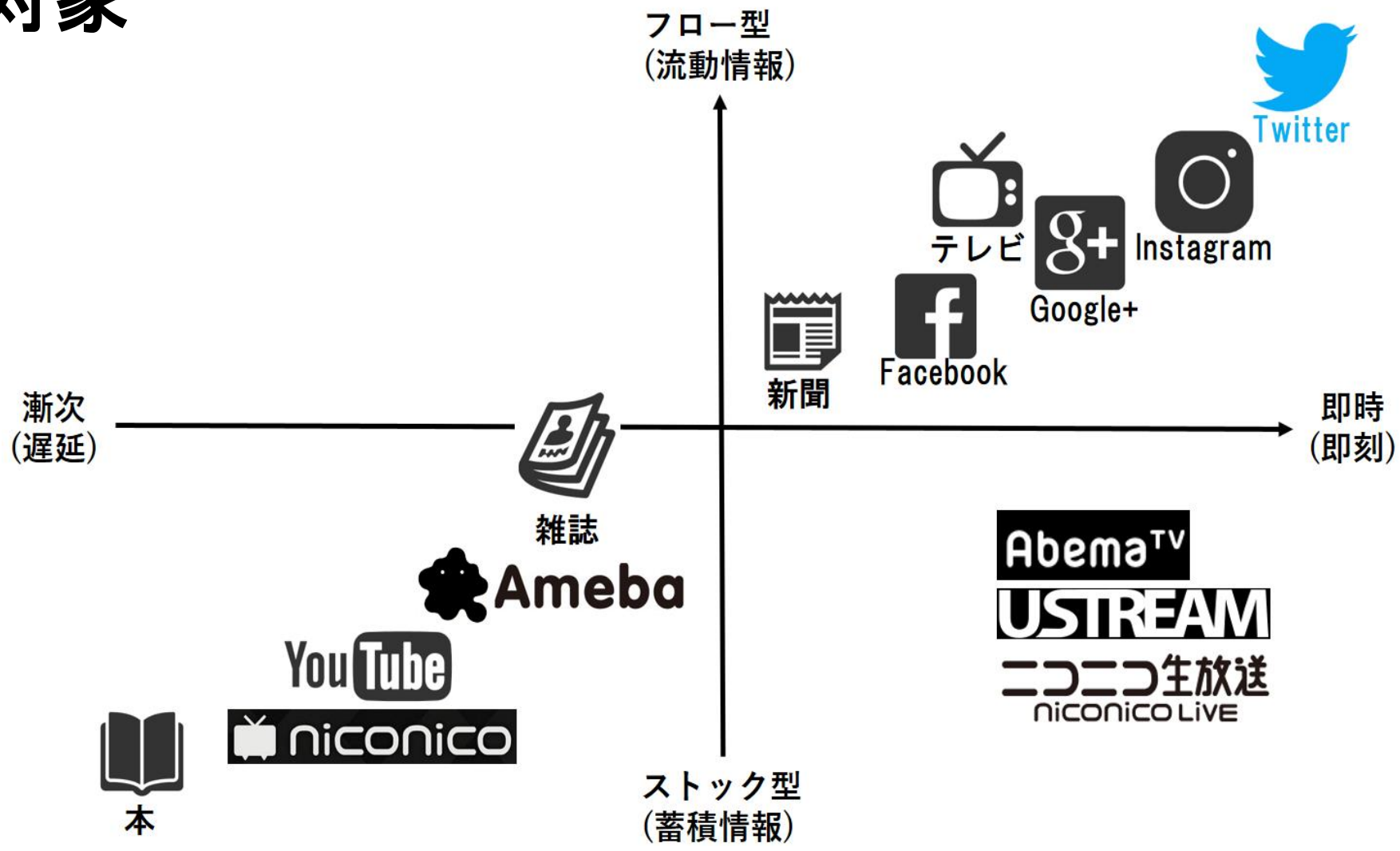


図1. ストック型およびフロー型メディアの関係図

出所：オフィス『仁』「インターネットメディアと既存メディア」
(http://idegene.com/mktg/comparison-internet-media-and-real-media)の図の一部を筆者が加筆修正

分析対象

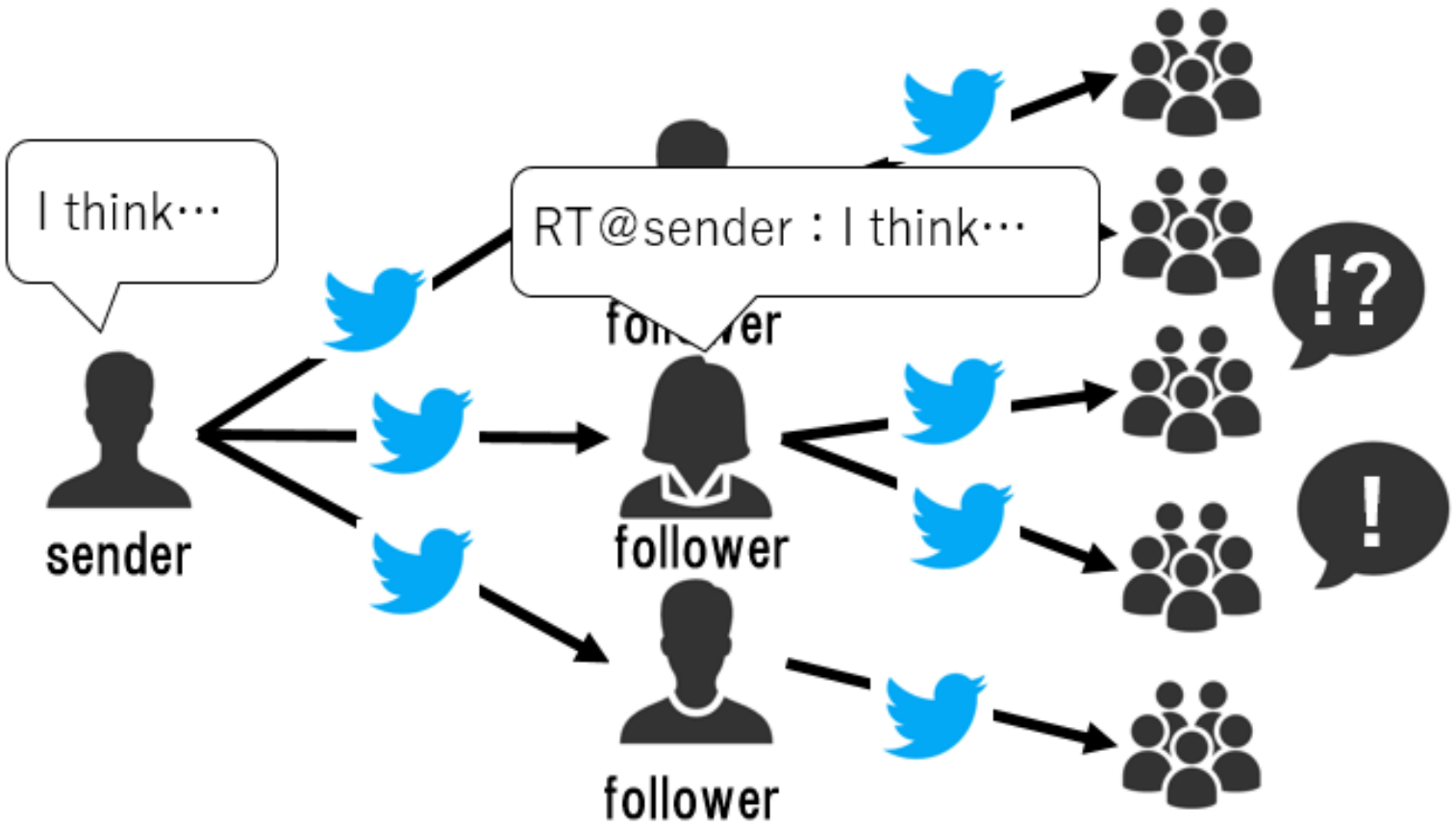


図2. Twitterの人的ネットワークと拡散のしくみ

出所：静岡県立大学 経営情報学部 斉藤和巳研究室「Twitter・Blog上での情報拡散の比較」
(http://pre4306.u-shizuoka-ken.ac.jp/research/)の図の一部を筆者が加筆修正

Twitterは同じ関心を持った人達が、**瞬時**に情報を**配信**、**共有**、**反応**することで**拡散**され、**最新の情報**をスピーディに入手することが可能

分析対象

表1. 各メディアの特徴と評価

No	メディア名	更新頻度	利用者数	信頼性	データ形式	収集環境
1	日銀短観	×	△	◎	◎ (テキスト)	―
2	新聞	△	◎	◎	◎ (テキスト)	―
3	テレビ	○	◎	◎	× (動画)	―
4	ラジオ	○	○	◎	× (音声)	―
5	ネット動画	◎	○	△(※)	× (動画)	―
6	ネットラジオ	◎	△	△(※)	× (音声)	―
7	電子掲示板	○	△	△	◎ (テキスト)	―
8	Webニュース	○	◎	◎	◎ (テキスト)	―
9	ブログ	△	○	△(※)	◎ (テキスト)	○ (API)
10	SNSサイト	○	◎	○(※)	◎ (テキスト)	○ (API)
11	Twitter	◎	◎	△(※)	◎ (テキスト)	◎ (API)

（注1）各メディアに対し項目ごとに評価を実施。評価が高い順に◎、○、△、×で表し掲載。
（注2）※公式チャンネル・アカウントを除く。

出所：松尾太加志(1999)『コミュニケーションの心理学』の表の一部を筆者が加筆修正

分析対象

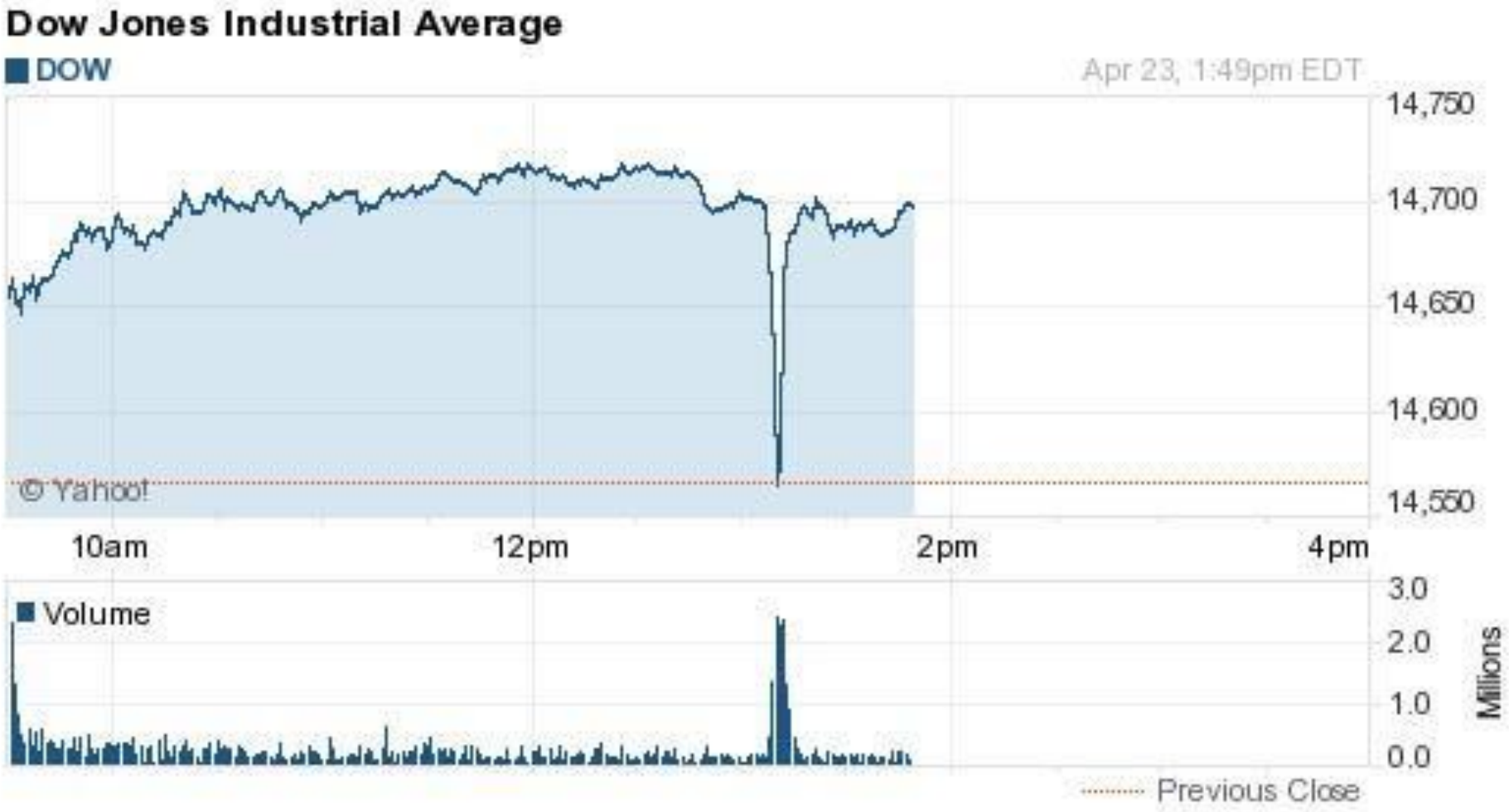
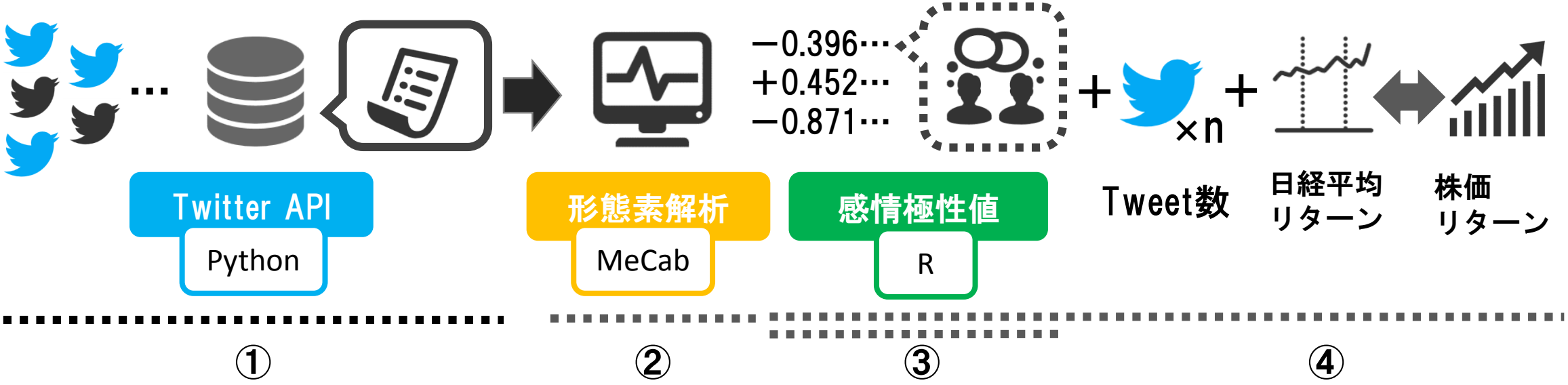


図3. 偽Tweetによるダウ・ジョーンズ平均株価への影響

出所：Yahoo!Finance (http://finance.yahoo.com/)

分析方法(手順)



分析手順

① 対象銘柄に関するTwitterのTweet情報(Tweet数とTweetの内容)を収集

② テキストマイニングを実施し、形態素解析を実施

③ 感情(ﾈｶﾞ･ﾎﾟｼﾞ)辞書に照らし合わせ投稿者の感情極性値を算出

④ 感情極性値・Tweet数・日経平均ﾘﾀｰﾝに対し、株価ﾘﾀｰﾝとの関係を推定

分析方法(感情極性値の算出)

対象銘柄関連のキーワード

Mon Aug 01 10:59:49 +0000 2016 @user1:昨日から株式市場は大きく値下がり模様。
Mon Aug 01 10:59:30 +0000 2016 @user2:建設大手はコスト削減効果などもあり営業・
Mon Aug 01 10:58:55 +0000 2016RT @user3:日経平均が大きく値下がる可能性がある。

抽出・収集

建設大手はコスト削減効果などもあり営業・最終利益の予想を上方修正の見通しで・・・

Twitter API | Python

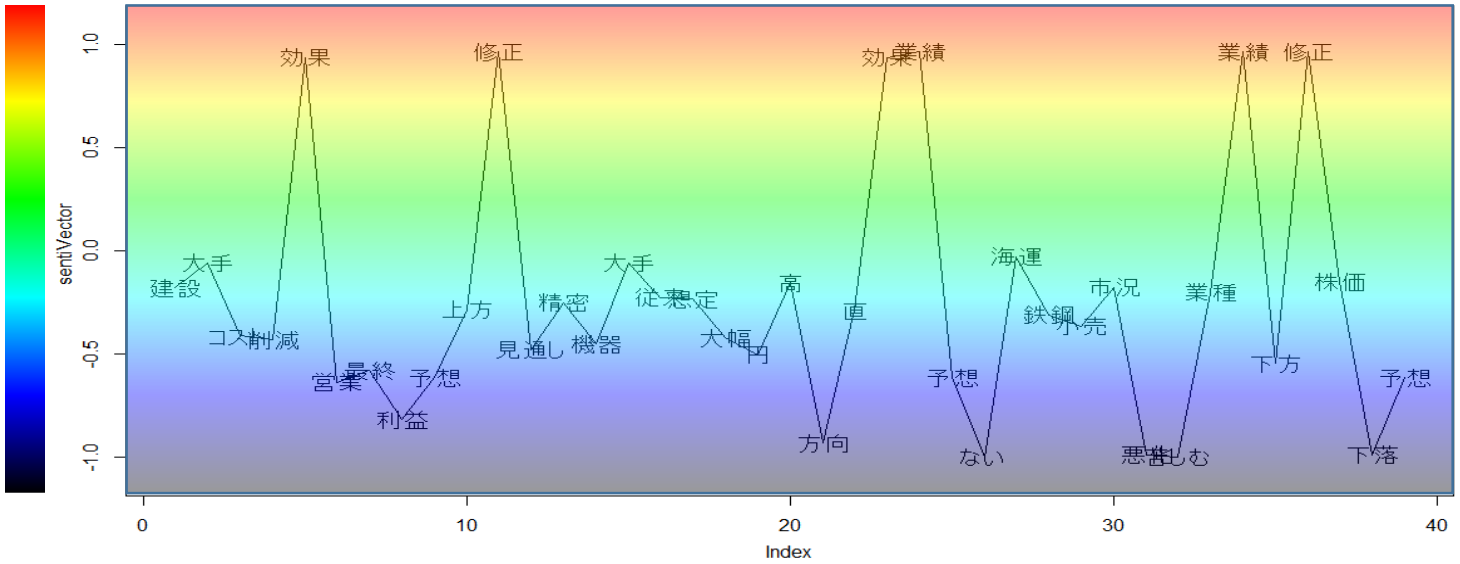
分解

建設(名詞) / 大手(名詞) / は(接続詞) / コスト(名詞) / 削減(名詞) / 効果(名詞) / ・・・

形態素解析 | MeCab

照合・算出・判定

建設(名詞):−0.18 / 大手(名詞): −0.01 / コスト(名詞) : −0.41 / 削減(名詞) : −0.43



ポジティブ
1.0

-1.0
ネガティブ

単語感情極性対応表 | R

+1.0
…
−1.0

結果: −0.391
判定: ネガティブ

分析データ

1. 高頻度取引データ:(秒単位)

●日経平均
●企業銘柄(6社※1)

期間:2016年8月1日～31日

※1･･･日立製作所、東芝、三菱電機、
パナソニック、シャープ、ソニー

2. Tweetデータ:(秒単位)

●対象銘柄に関するTweet情報

期間:2016年8月1日～31日

表2. 基本統計量　【例】日立製作所(6501)

Variable	Obs	Mean	Std. Dev.	Min	Max
price	47542	477.8007	13.22106	441.2	498.4
return	47542	1.31E-06	0.000344	-0.01144	0.007148
volatility	47542	0.000203	0.000287	0	0.009553
nikkei225	47542	16883.93	212.9642	16190	17240
nikkei225_return	47542	4.97E-08	0.00018	-0.00817	0.006135
score	47542	-0.42432	0.202394	-0.99996	0.999995
tweet	47542	0.647533	1.251748	0	44

price･･･････････････企業銘柄の株価
return･･･････････････企業銘柄の株価リターン
volatility･･･････････企業銘柄の株価変動度合
nikkei225･･･････････日経平均(上場インデックスファンド225)
nikkei225_return･･･日経平均のリターン
score･･･････････････Tweetの感情極性値
tweet･･･････････････Tweetの数

分析データ

【例】日立製作所(6501)

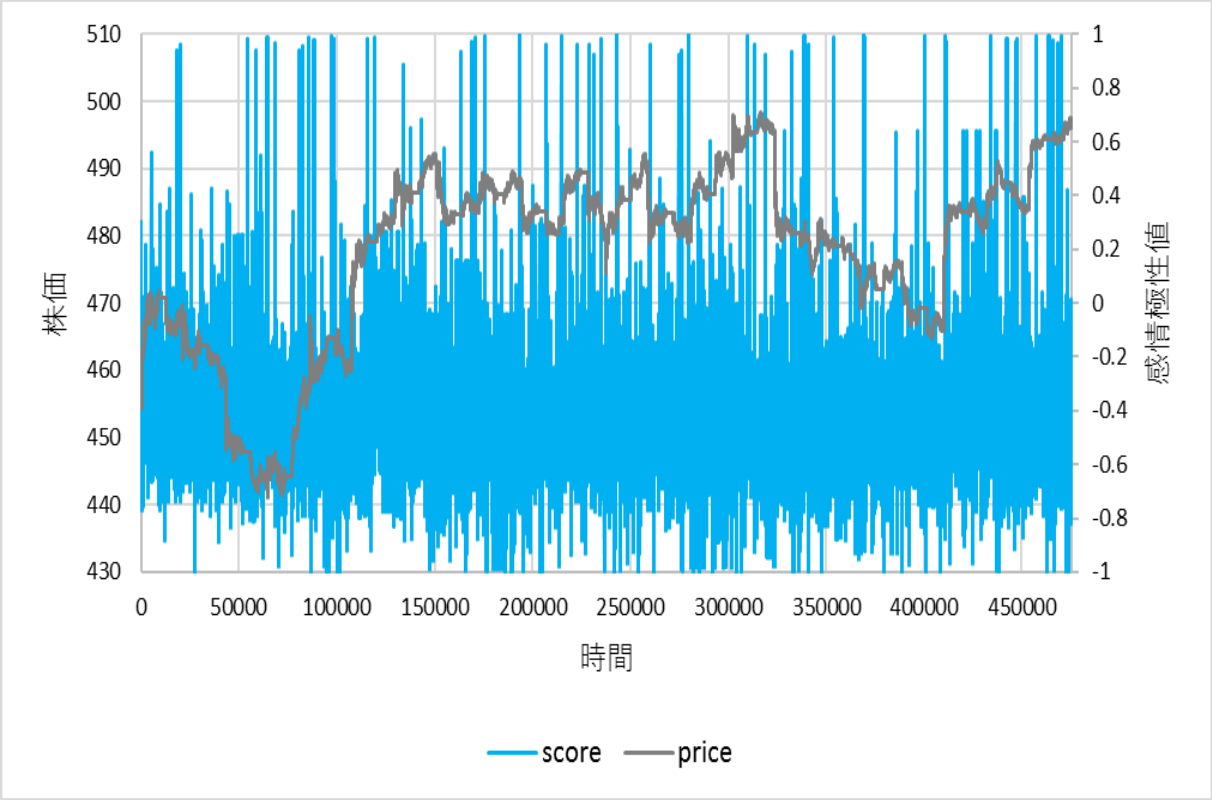


図4. 株価と感情極性値

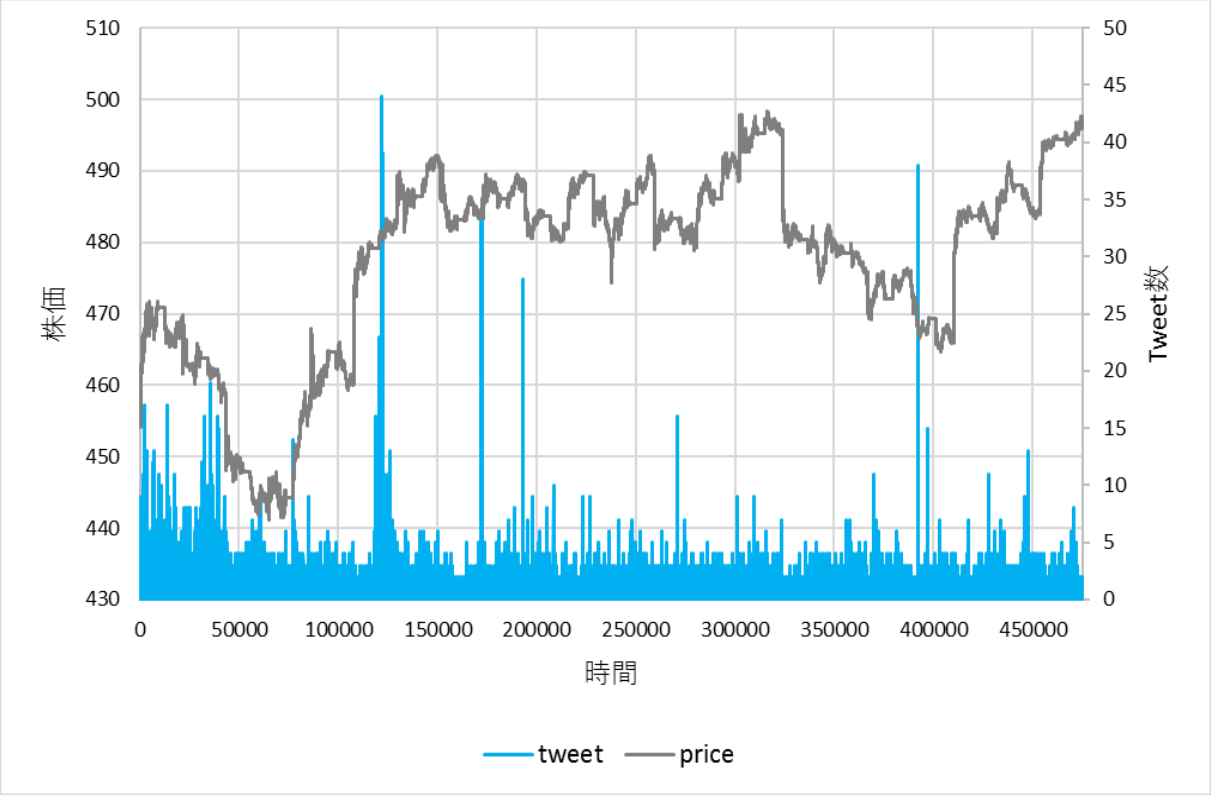


図5. 株価とTweet数

分析データ

【例】日立製作所(6501)

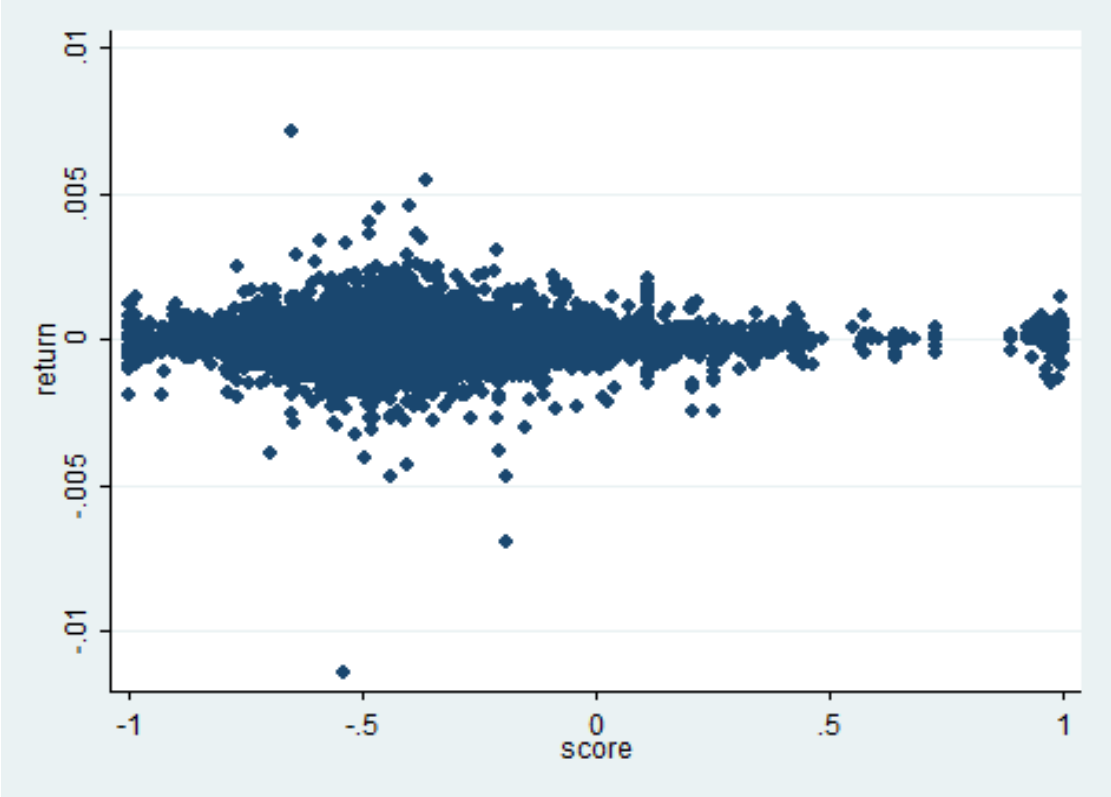


図6. 株価リターンと感情極性値

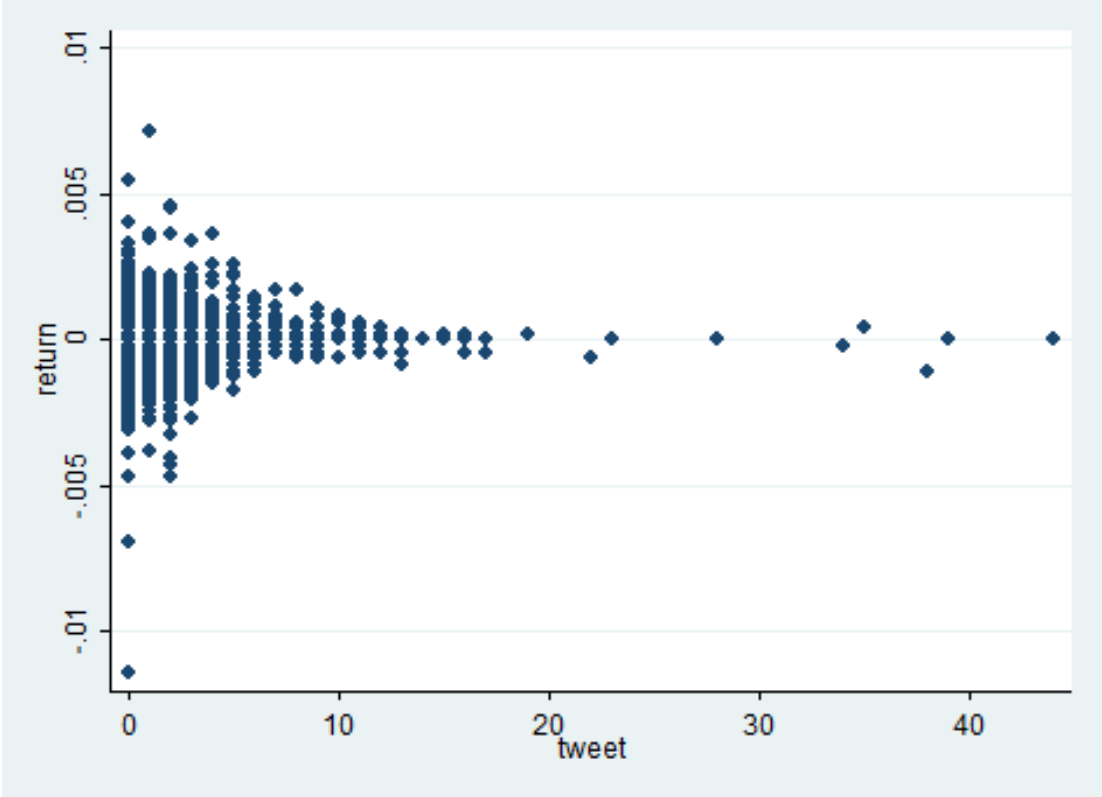


図7. 株価リターンとTweet数

分析モデル

$$R_{(i)(t)} = \frac{(P_{(i)(t)} - P_{(i)(t-1)})}{P_{(i)(t-1)}} \tag{1}$$

$$R_{(i)(t)} = \alpha_{(i)} + \beta_1 Score_{(i)(t)} + R(INDEX)_{(t)}\{\beta_2 + \beta_3 Tweet_{(i)(t)}\} + \varepsilon_{(i)} \tag{2}$$

銘柄・・・*i*　時間・・・ *t -1* ～*t* (10秒間隔)

- $R_{(i)(t)}$ ・・・・・・・・株価リターン
- $R(INDX)_{(t)}$ ・・・日経平均（上場インデックスファンド225）のリターン
- $Score(i)(t)$・・・・Tweetの感情分析スコア（感情極性値）
- $Tweet(i)(t)$・・・Tweetの数

分析結果

表3. 分析モデル(２)の推定結果

	日立製作所	東芝	三菱電機	パナソニック	シャープ	ソニー	6社
score	-0.01	-0.03	-0.01	-0.01	0	-0.02*	-0.01**
nikkei225_return	0.27***	0.31***	0.56***	0.42***	0.01	0.30***	0.31***
nikkei225_return×tweet	0.16***	-0.11**	-0.11***	-0.26***	0	-0.05***	-0.05***
Obs	17563	9924	17612	15076	23520	26868	110563

* p<0.1, ** p<0.05, *** p<0.01

*,**,***はそれぞれ係数が10%、5%、1%の有意水準において有意なものを示す。

① αについては、感情極性値との関係性を示す
有意な関係は観測できず
⇒ 仮説1支持されず

② βについては、Tweet数が関係しており、
Tweet数が多くなるとβの減少が観測された
⇒ 仮説2支持

分析結果

追加分析①：感情極性値とα

対象：ポジティブのみ　⇒　日立製作所 1社のみマイナスに**有意な関係**
対象：ネガティブのみ　⇒　有意な関係は**観測できず**

追加分析②：感情極性値と株価リターンの残差のボラティリティ

$$R_{(i)(t)} = \alpha_{(i)} + R(INDEX)_{(t)}\{\beta_4 + \beta_5 Tweet_{(i)(t)}\} + \varepsilon_{(i)} \quad (3)$$

$$\hat{\varepsilon}_{(i)(t)} = R_{(i)(t)} - \hat{R} \quad (4)$$

$$std(\hat{\varepsilon}_{(i)(t\sim t+\tau)}) = \sqrt{\frac{1}{n}\sum(\hat{\varepsilon}_{(i)(t)} - \mathrm{average}(\hat{\varepsilon}_{(i)(t\sim t+\tau)}))^2} \quad (5)$$

$$std(\hat{\varepsilon}_{(i)(t\sim t+\tau)}) = \partial_0 + \partial_1 Score_{(i)(t)} \quad (6)$$

日立製作所　⇒　マイナスに**有意な関係**　　三菱電機　⇒　プラスに**有意な関係**

考察

●Tweet数が増加すればするほど情報が共有されることになり、
結果としてリスクが減少すると考えられる

●一部の企業では、感情極性値とαおよび
株価リターンの残差のボラティリティとの間に有意な関係を確認

Twitterのセンチメント指数が取引判断の1つの指標になる可能性

検討課題

① 対象銘柄の拡大（大型株以外の中型株、小型株 ／ 業種の拡大）

② 観測期間の延長に伴う市場の『凪』と『時化』の比較

③ 他の感情辞書の再検討

参考文献

岡田克彦・羽室行信（2011）
「相場の感情とその変動 －自然言語処理で測定するマーケットセンチメントとボラティリティー」
『証券アナリストジャーナル』49(8)、37-48頁.

沖本竜義・平澤英司（2014）
「ニュース指標による株式市場の予測可能性」『証券アナリストジャーナル』52(4)、67-75頁.

五島圭一・高橋大志（2016）
「ニュースと株価に関する実証分析 －ディープラーニングによるニュース記事の評判分析－」
『証券アナリストジャーナル』54(3)、76-86頁.

上瀧弘晃・高橋悟・高橋大志（2009）
「クレジット市場におけるヘッドラインニュースの効果」『日本ファイナンス学会第17回大会予稿集』、113-122頁.

高村大也, 乾孝司, 奥村学（2006）
「スピンモデルによる単語の感情極性抽出」『情報処理学会論文誌ジャーナル』、Vol.47 No.02 pp. 627--637, 2006.

中山興, 藤井崇史（2013）
「株式市場における高速・高頻度取引の影響」『日銀レビュー』2013-J-2

丸山健・梅原英一・諏訪博彦・太田敏澄（2008）
「インターネット株式掲示板の投稿内容と株式市場の関係」『証券アナリストジャーナル』46、(11・12)、110-127頁.

参考文献

Antweiler, W., and M. Z. Frank (2004)
"Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards, "
Journal of Finance 59(3), pp.1259-1294.

Bollen, J., H. Mao and X. Zeng (2011),
“Twitter mood predicts the stock market,” *Journal ofComputational Science*, 2 (1), pp. 1 8.

Dougal, C., J. Engelberg, D. Garcia, and C. A. Parsons (2012)
"Journalists and The Stock Market, " *Review of Financial Studies* 25(3), pp.639-679.

Tetlock, P. C. (2007)
"Giving Content to Investor Sentiment: The Role of Media in the Stock Market,"
Journal of Finance 62(3), pp.1139-1168.

Tetlock, P. C., M. Saar-Tsechansky and S. Macskassy (2008)
"More Than Words: Quantifying Language to Measure Firms’ Fundamentals, "
Journal of Finance 63(3), pp.1437-1467.