

マイクロブログを用いた鉄道の運行トラブル状況抽出に関する一検討

土屋 圭[†] 豊田 正史^{††} 喜連川 優^{††}

[†] 東京大学大学院情報理工学系研究科

^{††} 東京大学生産技術研究所

E-mail: †{tsuchiya,toyoda,kitsure}@tkl.iis.u-tokyo.ac.jp

あらまし 鉄道の運行トラブルが発生した際、マイクロブログには鉄道の運行状況に関する情報が投稿されている。マイクロブログの投稿には運行トラブルについて、公式情報よりも詳細な情報が含まれることがある。鉄道の運行状況について詳しく知ることは意思決定を行う上で非常に重要である。本論文では Twitter の投稿を解析することによって、首都圏の鉄道運行トラブル状況に関する詳細な情報を抽出する手法について提案し、その評価を行う。提案手法は、東京メトロ 9 路線における運行トラブルの発生状況、復旧状況、混雑状況を対象に情報の抽出を行う。運行トラブルの発生状況については、運行トラブルを全線運転見合わせ、一部区間運転見合わせ、その他の異常の 3 段階に分けて抽出を行った。実験結果から、全線運転見合わせ、一部区間運転見合わせを抽出する場合において提案手法が有効であることが確認できた。また、復旧状況および混雑状況の抽出によって、公式情報だけではできないような意思決定を支援できることが確認できた。

キーワード データマイニング, ウェブマイニング, マイクロブログ, イベント抽出

Extracting Details of Train Troubles from Microblogs

Kei TSUCHIYA[†], Masashi TOYODA^{††}, and Masaru KITSUREGAWA^{††}

[†] Graduate School of Information Science and Technology, The University of Tokyo

^{††} Institute of Industrial Science, The University of Tokyo

E-mail: †{tsuchiya,toyoda,kitsure}@tkl.iis.u-tokyo.ac.jp

Abstract When train troubles occur, users of microblogs often post information about these troubles. Microblogs often include more fine-grained information of train troubles than official information. Such information could make out decision better. In this paper, we propose a method for extracting details of train troubles by analyzing Twitter. We classify train troubles into suspensions, partial suspensions and other troubles, then detect whether these troubles occur or not. Experimental results show that proposed method work well in case of suspensions and partial suspensions. We also find that by analyzing twitter, we can know information which is not reported officially such as congestion and recovery of train troubles.

Key words Data mining, Web mining, Microblog, Event detection

1. はじめに

近年、スマートフォンやソーシャルネットワークサービスの普及によってリアルタイム性のある、粒度の細かい情報がウェブにアップロードされるようになった。スマートフォンの普及率は世界中で増加の一途を辿っており、特に日本においては 2016 年までに 70% を超えると推定されている。世界中で広く利用されているソーシャルネットワークの 1 つである Twitter では、2012 年 6 月に 1 日の投稿件数が 4 億を超えた。また、Twitter のアクティブユーザの 60% がモバイル端末から利用しているということも分かっている。人々は Twitter のような

リアルタイムに情報を共有するウェブサービスを、スマートフォンによっていつでも、どこでも利用することができる。その結果、実世界で発生した出来事は即座にウェブ上のコンテンツに反映される。

我々は実世界の出来事の情報テレビや新聞、インターネット上のニュースサイトなどを通じて知ることができる。これらのメディアは、出来事が発生してから情報が伝えられるまでに時間がかかり、得られる情報も限られている。一方で、ソーシャルメディアには実世界の出来事は即座に反映され、他のメディアが報じている内容よりも詳細な情報が含まれている場合がある。例えば、地震や台風などの自然災害、交通渋滞、鉄道

の運行トラブルなどが発生すると、ソーシャルメディアにはそれらの状況がリアルタイムに反映される。実世界の出来事について、リアルタイムに詳細に把握することは意思決定を下す際に非常に重要である。そこで、ソーシャルメディアから実世界の出来事の情報抽出するという機運が高まっている。しかし、ソーシャルメディアは膨大な情報を含んでおり、情報のフォーマットも不揃いである。したがって、ソーシャルメディアから実世界の出来事について有益な情報を抽出することは困難である。そのため、近年ではソーシャルメディアを解析して実世界のイベントを抽出したり、抽出した情報をもとに将来の出来事を予測する研究が盛んに行われている。

本論文では、公式情報よりも詳細な情報がリアルタイムにソーシャルメディアに反映されている出来事として鉄道の運行トラブルに焦点を当てる。ソーシャルメディアの中でもリアルタイム性の強い Twitter の投稿を解析することによって、首都圏の鉄道運行トラブル状況に関する詳細な情報を抽出する手法について提案し、その評価を行う。提案手法は、東京メトロ 9 路線における運行トラブルの発生状況、復旧状況、混雑状況を対象に情報の抽出を行う。運行トラブルの発生状況については、運行トラブルを全線運転見合わせ、一部区間運転見合わせ、その他の異常の 3 段階に分けて抽出を行う。

本論文の構成は以下の通りである。2 章ではソーシャルメディアから実世界の出来事の抽出を行った関連研究について述べる。3 章では、Twitter を解析して鉄道の運行トラブルを抽出する手法について提案し、4 章で提案手法の評価実験について述べる。5 章では評価実験の考察を行い、最後に 6 章でまとめと今後の課題を述べる。

2. 関連研究

ソーシャルメディアを解析して実世界の出来事を抽出する研究は、対象としている出来事によって、人の行動や状態、社会や経済の動き、自然現象の 3 つに分けることができる。

人の行動パターンや状態などを予測する研究として、Eytan ら [1] によるインターネットユーザの行動を予測する研究や、Adam ら [2] による個々人の健康状態を予測する研究がある。Eytan らは、過去のユーザの検索履歴を解析して、インターネットユーザの今後の行動を予測した。Adam らは Twitter の投稿内容、投稿位置、ソーシャルリンクを解析して、Twitter のユーザの健康状態を最大 7 日先まで予測した。

次に、社会や経済の動きを抽出する研究として、Eric ら [3] によるソーシャルメディアと株価変動の相関を調査した研究や、Kira ら [4] による今後のニュースを予測する研究、Panagiotis ら [5] による選挙結果を予測する研究がある。Eric らはブログ上の不安表現と株価指数の相関を見つけ、株価変動の予測を行った。Kira らは過去 150 年分のニュース記事から因果関係を抽出し、将来の出来事を予測した。Panagiotis らは Twitter を解析して、選挙結果の予測を行った。

最後に、自然現象を抽出する研究として、Sakaki ら [6] の地震と台風を対象にした研究や、Hapipeng ら [7] の積雪や植生を対象にした研究がある。Sakaki らは Twitter の投稿内容と投

稿位置を用いて、地震と台風それぞれの発生時間および発生位置を推定した。Haipeng らは写真共有サイトである Flickr に投稿された写真およびそのタグを解析し、積雪の有無や植生を抽出する手法を提案した。

関連研究の多くは実世界の出来事が発生したかどうかや予測について焦点を当てており、実世界の出来事の詳細な情報を抽出している研究はあまりなされていない。例えば地震の場合は、地震の発生および発生位置の特定については焦点が当てられていたが、その震度や被害の状況については言及されていない。そのため、対象の出来事が発生したかどうかだけでなく、その出来事の詳細な情報を抽出することが課題として挙げられる。

3. 鉄道の運行トラブル状況の抽出

鉄道の運行トラブルが発生した際、どのようなトラブルが発生しているかを知ることができれば、我々はより良い意思決定を行うことができる。本論文では Twitter を解析することによって、東京メトロ 9 路線の運行トラブルをそれぞれの路線について全線運転合わせ、一部区間運転見合わせ、運転見合わせ以外の異常（直通運転中止や遅延）という 3 段階の粒度で抽出する。

3.1 データセット

本論文では、Twitter API を用いて東京メトロの一部または全路線で何らかの異常が生じた 2011 年 9 月 21 日^(注1)、2012 年 4 月 3 日^(注2)、2012 年 5 月 10 日^(注3)、2012 年 5 月 21 日^(注4)、2012 年 9 月 30 日^(注5) に投稿された、東京メトロのいずれかの路線名を含むツイートを用いた。また、収集した全てのツイートに対して、i) 検索時に使用した路線名を“Queryword”という文字に置換、ii) 投稿に含まれる URL を“URL”という文字に置換、iii) 投稿に含まれる @ から始まるユーザ名を“MENSION”という文字に置換、という処理を行った。それぞれの日の路線ごとのツイート数を表 1 に示す。

3.2 ツイートの分類

路線名を含むツイートには、“今、千代田線が全線で運転見合わせている”というように、現在の運行トラブルについて述べているものが含まれている。しかし、“これから千代田線に乗る”、“今夜は千代田線遅れるだろう”のように、現在の運行トラブルとは関係ないツイートもあるため、ツイートが現在の運行トラブルについて述べているかどうかを分類する必要がある。そこで、SVM を用いてツイートの分類を行った。SVM のライブラリには LIBSVM^(注6)、SVM のカーネルには線形カーネル、特徴量にはツイートに出現する品詞を用いた。また、日

(注1)：台風 15 号の影響で一部路線で運転見合わせ、全路線で遅延などが発生した。

(注2)：強風の影響で、銀座線を除く全線で遅延などが生じた。

(注3)：送電トラブルの影響で東西線で一部区間運転見合わせ、遅延などが生じた。

(注4)：刺傷事件の影響で副都心線で全線運転見合わせ、遅延などが生じた。

(注5)：台風 17 号の影響で銀座線、半蔵門線、丸ノ内線を除く 6 路線で遅延などが生じた。

(注6)：<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

表 1 路線名を含むツイート数

Table 1 The number of tweets including line name.

	千代田線	副都心線	銀座線	半蔵門線	日比谷線	丸ノ内線	南北線	東西線	有楽町線
2011年9月21日	3,492	3,205	3,385	2,459	4,771	2,805	2,810	8,763	4,031
2012年4月3日	1,165	333	538	667	549	434	8,524	446	-
2012年5月10日	-	-	-	-	-	-	-	2,684	-
2012年5月21日	-	12,504	-	-	-	-	-	-	-
2012年9月30日	602	209	-	248	-	-	123	1,537	245

本語文章を品詞に分解する形態素解析器には MeCab^(注7)を用いた。ツイートの分類実験の結果は 4.1 で述べる。

3.3 運行トラブル状況の抽出

運行トラブルの推定は、時間を m 分間隔に区切り、各時間帯で運行トラブルが発生しているかどうかを判定することによって行う。まず、推定対象の時間帯に投稿されたツイートの分類を行う。次に、分類器によってポジティブ（運行トラブルについて述べている）と判定された m 分ごとのツイートの数 T_p を求める。 T_p が閾値よりも大きい場合、その時間帯に運行トラブルが発生したと推定する。

4. 評価実験

4.1 ツイートの分類実験

ツイートの分類のために、5000 件のツイートに対して人手でラベル付けを行った。これらのツイートは 2011 年 9 月 21 日、2012 年 4 月 3 日、2012 年 9 月 30 日の 3 日間に投稿された、東京メトロのいずれかの路線名を含むツイートの中からランダムにサンプリングしたものである。ラベル付けは運行トラブル状況、混雑状況、復旧状況それぞれについて独立に行った。運行トラブル状況についてのラベルは、 T_0 : 運行トラブル状況について言及なし、 T_1 : 全線運転見合わせ、 T_2 : 一部区間運転見合わせ、 T_3 : 運転見合わせ以外の異常の 4 種類、混雑状況のラベルは C_0 : 混雑状況について言及なし、 C_1 : 混んでいる、 C_2 : 空いているの 3 種類、復旧状況のラベルは R_0 : 復旧状況について言及なし、 R_1 : 復旧状況について言及ありの 2 種類である。ラベル付けの結果を表 2 に示す。

表 2 ラベル付けの結果

Table 2 The result of labeling.

T_0	T_1	T_2	T_3	C_0	C_1	C_2	R_0	R_1
3,010	868	461	661	4,488	328	182	4,745	255

次に、運行トラブルが発生した時間帯が既知の場合に、その時間帯に投稿されたツイート全てを正例とした学習データを作成し、これを用いた半自動分類もあわせて試みる。半自動の結果の有効性が確認できれば、過去の運行状況だけをを用いたツイートの分類ができ、人手でラベル付けを行う手間が大幅に削減される。手動の手法では 2 値分類の 10-分割交差検定を行った。表 3 に手動の手法で T_1 をターゲットラベルにした際に学習した SVM の特徴量とその重みの例を示す。また、ベースラ

インとして、全てのツイートをターゲットラベルとする手法を用いた。

表 4 にツイートの分類実験の結果を示す。F 値においては全ての場合で手動が最も高く、運行トラブル状況においては、半自動が手動よりも適合率が高い場合があった。特に $T_1 + T_2$ 、 $T_1 + T_2 + T_3$ をターゲットラベルとした際は、8 割以上の適合率が得られた。また、半自動とベースラインの F 値を比較すると半自動の方が高いが、手動の F 値よりは低く、今回の実験においては半自動の手法は実用的ではないことが分かった。

表 3 ターゲットラベルが T_1 のときの SVM の特徴量の重みTable 3 Example of positively and negatively weighted features of SVM targeting T_1 .

Positive Features			Negative Features		
品詞	特徴量	重み	品詞	特徴量	重み
動詞	まっ	2.0251	名詞	一部	-1.5174
動詞	止まっ	2.0000	副詞	よく	-1.2009
名詞	見合わせ	1.7811	助詞	たり	-1.1436
動詞	とまっ	1.6911	副詞	ほとんど	-1.0931
動詞	見合せ	1.6841	助動詞	なかつ	-1.0883
名詞	ストップ	1.5479	副詞	そろそろ	-1.0655
副詞	なんで	1.4419	名詞	東陽	-1.0220
動詞	とまり	1.4231	名詞	以外	-1.0018
名詞	運休	1.2678	感動詞	ひゃー	-1.0000
副詞	なぜ	1.2435	動詞	限ら	-1.0000
名詞	経過	1.2155	名詞	(^ - ^ ~ ^ - ^)	-1.0000
動詞	見合わせ	1.2117	名詞	ガラガラ	-1.0000
名詞	全滅	1.1663	名詞	成功	-1.0000
動詞	動か	1.1407	名詞	文章	-1.0000
名詞	不通	1.0967	名詞	両方	-1.0000

4.2 運行トラブル状況の推定

運行トラブル状況の推定は 3.1 で述べた 5 日間について行った。運行トラブルは T_1 : 全線運転見合わせ、 T_2 : 一部区間運転見合わせ、 T_3 : 運転見合わせ以外の異常、 $T_1 + T_2$: 全線または一部区間運転見合わせ、 $T_1 + T_2 + T_3$: 何らかの異常の 5 種類とし、それぞれを 4.1 で作成した分類器を用いて 5 分間隔で推定を行った。推定結果は閾値を変化させたときの、最も良い F 値によって評価する。ここで、F 値は式 (1) によって求める。

$$F = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (1)$$

ただし、推定区間数を $I_{estimation}$ 、正解区間数を I_{answer} 、推定と正解が一致した区間数を $I_{consistency}$ としたとき、

(注7) : <http://mecab.sourceforge.net/>

表 4 ツイートの分類結果

Table 4 Performance of classification.

Target Label	Method	Precision	Recall	F-value
T_1	手動	0.6966	0.6694	0.6827
	半自動	0.4505	0.6616	0.5360
	ベースライン	0.1736	1.000	0.2958
T_2	手動	0.6067	0.5857	0.5960
	半自動	0.7267	0.3215	0.4458
	ベースライン	0.0922	1.0000	0.1688
T_3	手動	0.5943	0.5386	0.5651
	半自動	0.4416	0.1819	0.2568
	ベースライン	0.1322	1.0000	0.2335
$T_1 + T_2$	手動	0.7165	0.7171	0.7168
	半自動	0.8623	0.5403	0.6643
	ベースライン	0.2658	1.0000	0.4200
$T_1 + T_2 + T_3$	手動	0.7711	0.7568	0.7639
	半自動	0.9548	0.4392	0.6016
	ベースライン	0.3980	1.0000	0.5694
C_1	手動	0.6799	0.6280	0.6529
	ベースライン	0.0656	1.0000	0.1231
C_2	手動	0.7911	0.6793	0.7310
	ベースライン	0.0368	1.0000	0.0710
$C_1 + C_2$	手動	0.7596	0.6973	0.7271
	ベースライン	0.1024	1.0000	0.1858
R_1	手動	0.7155	0.6706	0.6923
	ベースライン	0.0510	1.0000	0.0971

$Precision = \frac{I_{consistency}}{I_{estimation}}$, $Recall = \frac{I_{estimation}}{I_{answer}}$ である。ベースラインには、分類器を用いず、推定時間帯に投稿された路線名を含むツイートの数に対して提案手法と同様の処理を行う手法を用いた。

実験結果を表 5 に示す。 T_1 , T_2 および $T_1 + T_2$ の推定においては、2012 年 5 月 21 日の副都心線の T_1 を除いた全ての場合において、提案手法の有効性を確認することができた。一方で、 T_3 や $T_1 + T_2 + T_3$ の推定では、提案手法が有効に機能している場合も見られたが、全体としてベースラインの方が高い F 値が得られていた。

5. 考 察

5.1 ツイートの誤分類

誤分類されたツイートには大きく 3 つの特徴が見られた。1 つ目は、“ついに東西線が止まった、か?” のような、投稿の内容が運行状況についての疑問文のツイートや想像・予測のツイートである。今回の実験では投稿に含まれる品詞のみを特徴量としているため、疑問文の判定や、予測であるかどうかの判定ができず、誤分類が生じたと考えられる。2 つ目は、“東西線の地上区間は終了しました。” のような、運転状況について婉曲的な表現を用いているツイートである。学習データに存在しない表現を含むツイートを正しく分類することは難しい。学習データを増やして、数多くの表現に対応することが解決策として考えられる。3 つ目は、“中央線・有楽町線・半蔵門線・千代田線まだ動いてません。丸ノ内線・銀座線・大江戸線は動いて

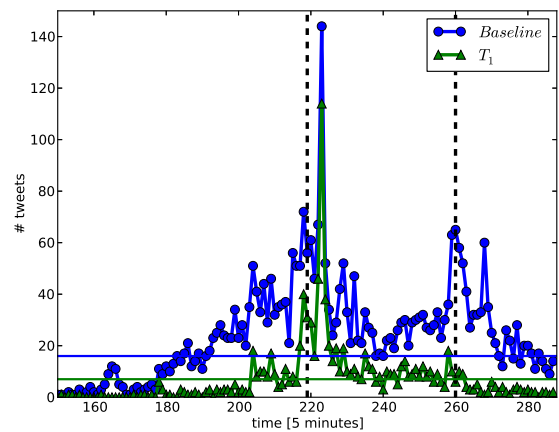


図 1 2011 年 9 月 21 日 千代田線の全線運転見合わせ推定時のツイート数のヒストグラム

Fig.1 The histogram of chiyo-da line detecting suspension at 2011/09/21.

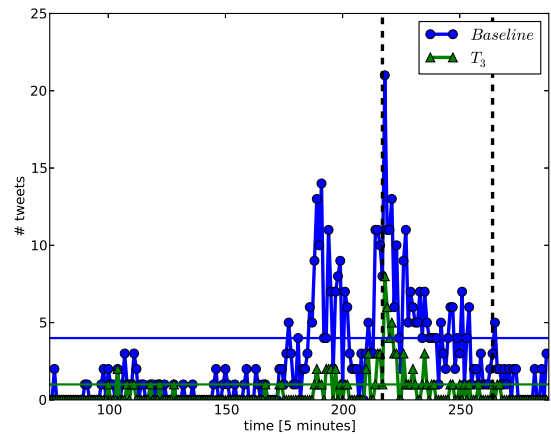


図 2 2012 年 4 月 3 日 半蔵門線の運転見合わせ以外のトラブル推定時のツイート数のヒストグラム

Fig.2 The histogram of hanzomon line detecting troubles except suspension at 2012/4/3.

るみたい” のように、1 つのツイートの中に複数の路線の運行状況が述べられているツイートである。今回は文脈を考慮する特徴量を用いておらず、どの路線についての情報なのかを明確に区別することができないため、誤分類が生じたと考えられる。

5.2 運行トラブル状況推定

図 1 に、提案手法の有効性が確認できた 2011 年 9 月 21 日の千代田線の全線運転見合わせを推定した際のヒストグラムを示す。提案手法の閾値は 7、ベースラインの閾値は 16 であり、点線で囲まれた領域が実際に全線で運転を見合わせた区間である。提案手法では、運転見合わせを行っていない $time = 200$ 近傍や、 $time = 260$ 以降において、ツイート数が閾値を超えていないのに対し、ベースラインでは閾値を超えていることが確認できる。このことから、分類器が有効に機能したと言える。

次に、提案手法の有効性が確認できなかった 2012 年 4 月 3 日の半蔵門線の運転見合わせ以外の異常を推定した際のヒスト

表 5 運行トラブルの推定結果

Table 5 The result of detecting train troubles.

Target Trouble	Method	千代田線	副都心線	銀座線	半蔵門線	日比谷線	丸ノ内線	南北線	東西線	有楽町線
2011年9月21日										
T_1	Proposed	0.75	0.50	0.86	-	0.89	-	0.77	0.57	0.81
	Baseline	0.6290	0.40	0.73	-	0.73	-	0.62	0.39	0.56
T_2	Proposed	-	0.84	0.75	-	0.92	-	0.85	0.83	0.75
	Baseline	-	0.79	0.53	-	0.82	-	0.70	0.68	0.67
T_3	Proposed	0.59	0.40	0.32	0.91	0.43	0.60	0.41	0.54	0.50
	Baseline	0.57	0.42	0.49	0.95	0.50	0.75	0.62	0.60	0.57
$T_1 + T_2$	Proposed	0.73	0.86	0.70	-	0.92	-	0.89	0.91	0.97
	Baseline	0.63	0.83	0.60	-	0.92	-	0.79	0.72	0.92
$T_1 + T_2 + T_3$	Proposed	0.81	0.74	0.77	0.93	0.71	0.65	0.85	0.89	0.71
	Baseline	0.80	0.78	0.78	0.95	0.74	0.75	0.90	0.95	0.76
2012年4月3日										
T_1	Proposed	-	-	-	-	-	-	-	1.00	-
	Baseline	-	-	-	-	-	-	-	1.00	-
T_2	Proposed	-	-	-	-	-	-	-	0.89	-
	Baseline	-	-	-	-	-	-	-	0.72	-
T_3	Proposed	0.42	0.46	-	0.49	0.51	0.17	0.37	0.52	0.40
	Baseline	0.81	0.27	-	0.64	0.53	0.37	0.53	0.58	0.47
$T_1 + T_2$	Proposed	-	-	-	-	-	-	-	0.95	-
	Baseline	-	-	-	-	-	-	-	0.75	-
$T_1 + T_2 + T_3$	Proposed	0.77	0.39	-	0.50	0.50	0.19	0.39	0.92	0.42
	Baseline	0.81	0.27	-	0.63	0.53	0.37	0.53	0.95	0.47
2012年5月10日										
T_3	Proposed	-	-	-	-	-	-	-	0.84	-
	Baseline	-	-	-	-	-	-	-	0.82	-
$T_1 + T_2 + T_3$	Proposed	-	-	-	-	-	-	-	0.83	-
	Baseline	-	-	-	-	-	-	-	0.82	-
2012年5月21日										
T_1	Proposed	-	0.60	-	-	-	-	-	-	-
	Baseline	-	0.67	-	-	-	-	-	-	-
T_2	Proposed	-	0.81	-	-	-	-	-	-	-
	Baseline	-	0.70	-	-	-	-	-	-	-
T_3	Proposed	-	0.69	-	-	-	-	-	-	-
	Baseline	-	0.76	-	-	-	-	-	-	-
$T_1 + T_2$	Proposed	-	0.92	-	-	-	-	-	-	-
	Baseline	-	0.89	-	-	-	-	-	-	-
$T_1 + T_2 + T_3$	Proposed	-	1.00	-	-	-	-	-	-	-
	Baseline	-	0.99	-	-	-	-	-	-	-
2012年9月30日										
T_2	Proposed	-	-	-	-	-	-	-	0.77	-
	Baseline	-	-	-	-	-	-	-	0.63	-
T_3	Proposed	0.75	0.24	-	0.48	-	-	0.44	0.50	0.47
	Baseline	0.77	0.32	-	0.51	-	-	0.52	0.52	0.54
$T_1 + T_2$	Proposed	-	-	-	-	-	-	-	0.74	-
	Baseline	-	-	-	-	-	-	-	0.63	-
$T_1 + T_2 + T_3$	Proposed	0.75	0.25	-	0.47	-	-	0.53	0.77	0.41
	Baseline	0.77	0.32	-	0.51	-	-	0.52	0.80	0.54

グラムを図2に示す。この場合，“半蔵門線”というワードを含むツイート数自体が少なく，提案手法の閾値が1になった。そのため， $time = 100$ 近傍でツイート数が閾値を超え，F 値が

ベースラインを下回ったと考えられる。

5.3 混雑状況および復旧情報の抽出

4.1 で作成した，混雑状況および復旧状況の分類器によって

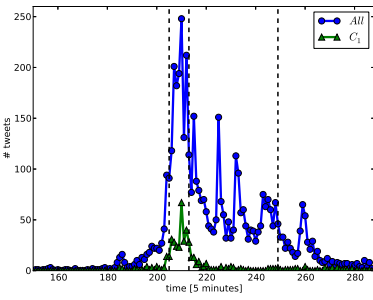


図3 2011年9月21日日比谷線 C_1
Fig.3 2011/9/21 Hibiya line C_1

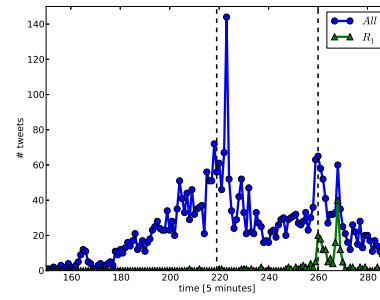


図4 2011年9月21日千代田線 R_1
Fig.4 2011/9/21 Chiyoda line R_1

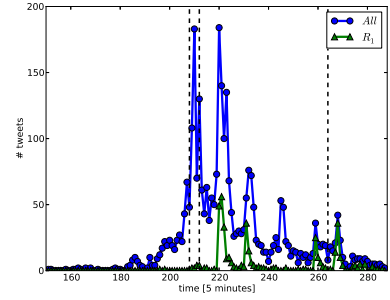


図5 2011年9月21日副都心線 R_1
Fig.5 2011/9/21 Fukutoshin line R_1

ポジティブと判定されたツイートのヒストグラムの例を図3, 4, 5に示す。図3では、 $time = 210$ 近傍の破線で囲まれた区間が全線で運転を見合わせ時間帯である。この時間帯において、分類器によって C_1 と分類されたツイートが増えている。実際のツイートを確認したところ、“中目黒駅のホームが混雑している影響で午後5時から中目黒～北千住の全線で運転見合わせ(略)”というツイートのリツイートによるものであり、正しくツイートを分類できていることが分った。また、 $time = 215$ 以降、運転を再開した後も混んでいることを示すツイートが見られた。この情報によって、電車は動き始めたが混雑していることが分かるため、少し時間が経ってから駅に向かうなどの意思決定ができる。次に図4では、運転が再開された $time = 260$ において R_1 と分類されたツイートが増えていることが確認できる。実際のツイートは、千代田線の運転再開について述べているものがほとんどであった。 $time = 270$ 近傍でもポジティブツイートのピークが見られる。これはNHK報道局生活情報部の公式 Twitter アカウントによる“【東京メトロ、全路線で再開】東京メトロは、運転を見合わせていた東西線と千代田線、それに副都心線の一部で運転を始め、すべての路線で運転を再開しました。(22:17)”というツイートが大量にリツイートされたことによるものであった。このことから、実際に運転を再開してから、報道局によって運転再開が伝えられるまで40分以上のタイムラグがあることが分かった。また、全線で運転を見合わせていた点線で囲まれた時間帯においては、 R_1 と判定されたツイートはほとんど存在していない。このことから、復旧状況を分類する分類器が運転見合わせを推定する場合において有効に働くことが予想される。最後に、図5では、 $time = 220, 230, 260, 270$ において、 R_1 と分類されたツイート数が増えていることが分かる。実際には、 $time = 220, 230$ 近傍において、副都心線が渋谷と池袋の間で運転を再開したという内容のツイートが大多数だった。 $time = 260$ におけるポジティブツイートの増加は、数多くリツイートされているツイートを誤分類したことによるものだった。 $time = 270$ におけるピークは、副都心線の全線運転再開を伝えるツイートによるものであった。この例では、全線運転見合わせから一部区間運転再開、その後全線で運転再開という流れを Twitter から抽出することができた。

6. まとめと今後の課題

本論文では、鉄道の運行トラブル状況を全線運転見合わせ、一部区間運転見合わせ、その他の異常という3段階の粒度で推定する手法について提案した。また、鉄道の混雑状況や運行トラブルの復旧状況の抽出についても言及した。評価実験では、全線運転見合わせおよび一部区間運転見合わせを推定する場合において、提案手法の有効性を確認できた。

今後の課題として、路線ごとに分類器を学習することが挙げられる。本論文では、東京メトロ9路線の名前を含むツイート全てからランダムに5000件をサンプリングし、ラベル付けを行ったため、路線ごとに存在する特徴が捉えられていないと考えられる。例えば、直通運転中止を伝えるツイートであれば、そのツイートの中に直通運転先の路線名が含まれると考えられる。直通運転先の路線は東京メトロ各路線ごとに異なる。したがって、路線ごとに分類器を作成することで、運転見合わせ以外の運行トラブルについて提案手法の性能向上を見込むことができる。また、運行トラブルの推定方法の検討も今後の課題として挙げられる。5.3で述べたように、復旧状況の分類器も用いるなど、様々な方法を検討したい。

文 献

- [1] Eytan Adar, Daniel S. Weld, Brian N. Bershad and Steven D. Gribble, Why We Search: Visualizing and Predicting User Behavior, WWW, 2007
- [2] Adam Sadilek, Henry Kautz and Vincent Silenzio, Predicting Disease Transmission from Geo-Tagged Micro-Blog Data, AAAI, 2012
- [3] Eric Gilbert and Karrie Karahalios, Widespread Worry and the Stock Market, AAAI, 2010
- [4] Kira Radinsky, Sagie Davidovich and Shaul Markovitch, Learning Causality for News Events Prediction, WWW, 2012
- [5] Panagiotis T. Metaxas and Eni Mustafaraj, How (Not) to Predict Elections, IEEE third international conference on social computing, 2011
- [6] Takeshi Sakaki, Makoto Okazaki and Yutaka Matsuo, Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors, WWW, 2010
- [7] Haipeng Zhang, Mohammed Korayem and David J. Crandall, Mining Photo-sharing Websites to Study Ecological Phenomena, WWW, 2012