

Twitterからの列車遅延情報収集手法の検討

新井 誠也^{1,a)} 平川 豊^{2,b)} 大関 和夫^{2,c)}

概要: マイクロブログサービスの Twitter が注目を集めている。Twitter が持つ特徴の 1 つにリアルタイム性に優れていることが挙げられるが、近年ではその特徴を生かし、実世界で起きているイベントを Twitter から観測しようとする研究が活発に行われている。本研究では、列車の遅延というイベントに焦点を当て、鉄道利用者から発信されるツイートをもとに各路線の列車の運行状況を判断するシステムを提案する。加えて、リアルタイム性の観点からシステムの有効性を確認する。遅延検出の閾値を、実際の遅延発生時における単位時間当たりのツイート数と各路線・各時間帯の鉄道利用者数を考慮して定める。その結果、鉄道会社の公式運行情報よりも早く列車遅延の第一報を配信可能としたことを示す。

A train delay detection method through the analysis of twitter messages

Abstract: The micro-blogging service called Twitter, has attracted a large amount of attention. Twitter messages, called "tweets," are posted and read in real-time; such tweets are often relevant to a social event. There are numerous research studies that analyze Twitter messages focused on events occurring in the real world. In this study, we propose a train delay detection method through the analysis of Twitter messages, and present the evaluation results. A threshold value is used to detect train delays; this value depends on the number of tweets per unit time and the number of passengers on each railroad, which varies across train lines and time zones. Our results show that train delays can be detected earlier than formal announcements from railroad companies.

1. はじめに

近年、Twitter[1] が注目を集めている。Twitter とは、2006 年にサービスが開始されたマイクロブログサービスの 1 つである。利用者は、身の回りの状況・体験や雑記などを最大 140 字の「ツイート (tweet)」「つぶやき」と呼ばれる短文 (本稿では以下、「ツイート」と表記) にのせて、他の利用者に向けて投稿・発信することができる。リアルタイム性・速報性に優れ、全世界における 1 日あたりの総ツイート数は 4 億件、月間アクティブユーザー数は 2 億人 (2013 年 3 月時点) を超える [2] 大規模なサービスである。

Twitter を利用した研究は様々な分野で数多く行われている。特に近年では Twitter が持つリアルタイム性に着目した研究が盛んに行われている。その中でも、あらゆる利

用者からのツイートを常に収集することで、「実世界では今、何が起きているのか」を観測する試みが多く行われている [3]。これらは、あらかじめターゲットとするイベントを定め、そのイベントの発生をより正確に、より素早く検出を行うことに焦点を当てている。例えば、Twitter を用いて地震の検出を目的とした研究 [4], [5]、交通情報の抽出を目的とした研究 [6], [7]、インフルエンザや花粉症の流行予測を目的とした研究 [8], [9]、犯罪情報の集約を目的とした研究 [10] が挙げられる。一方、近年では、駅に直接行かなくても Web 経由で簡単に各鉄道路線の列車運行情報を入手できる。しかし、入手できる情報は各鉄道会社から公式に提供される運行情報またはその引用であることが多い。それらは遅延が発生してからしばらく経過しないと配信されない場合がある。刻一刻と状況が変化中、リアルタイム性に優れた情報提供が求められる。

本研究では、列車の遅延というイベントに対象を絞り、より早い情報提供を行うことを目的とし、鉄道利用者から投稿されたツイートを情報源に各鉄道路線の運行状況を配信するシステムを提案する。そして、正確かつ素早く列車

¹ 芝浦工業大学大学院理工学研究科
Graduate School of Engineering and Science, Shibaura Institute of Technology

² 芝浦工業大学工学部
College of Engineering, Shibaura Institute of Technology

a) ma13004@shibaura-it.ac.jp

b) hirakawa@shibaura-it.ac.jp

c) ohzeki@shibaura-it.ac.jp

の遅延の検出が行えたか評価を行う。

以降、第2章で関連研究、第3章で列車遅延時に投稿されるツイートの分析、第4章で提案するシステムの詳細、第5章で評価実験について述べ、第6章でまとめとして本研究の有効性と今後の課題を示す。

2. 関連研究

築瀬らは、JR常磐線の運行状況を一般のTwitterの利用者のツイートから予測して、列車が遅延していると推測される場合にはメールで自動通知を行うシステムを提案している(以下、関連研究)[11]。その結果、31件中13件の遅延において、「Yahoo!運行情報」よりも10分~30分程度早く通知することができたと評価している。

関連研究では、常磐線に遅延が発生すると、「常磐線」という単語を含むツイートの投稿件数が増加するという現象に注目した。分析の結果、直前5分間あたりに「常磐線」という単語を含むツイートが6件以上が投稿され、かつそれらに「遅延」「遅れ」「見合わせ」という単語が含まれていると、運行状況に異常がある可能性が高いとしている(ただし、Botからのツイート及びリツイートは判定条件の対象から外している)。よって、その条件を満たしたときに通知を行っている。

しかし、先行研究には以下の問題点がある。

(1) 通知の対象とする路線が1路線のみである

JR常磐線に限定して運行情報の通知を行っている。そのため、他の路線の運行状況については考慮されおらず、汎用性に欠ける。

(2) 早朝や深夜に発生した遅延には対応できない

列車遅延の通知条件を時間帯によって変化させていない。従って、Twitterの利用者数が少なくなる早朝や深夜に列車遅延が起きると、設定した閾値を上回らずに正しく通知を行ってこないケースがある。

(3) 列車遅延を示すツイートの抽出洩れ

ツイートの本文中に「遅延」「遅れ」「見合わせ」という3語のいずれかを含まなければ、列車遅延を示すツイートであると判断してくれない。例えば、「線が人身事故で止まっている」というツイートは、列車遅延を示すツイートではあるが、上記のいずれの語も含んでいないため抽出洩れが起こる。

(4) 閾値の定め方に曖昧性が残る

通知条件の最適な閾値を6件としている。その理由として、閾値を5件以下とすると、通常運転時にも誤って通知条件を満たしてしまう可能性が高くなるためとしている。しかし、そのときに通知ミスが起きてしまう件数・確率などの具体的な数値には言及していない。ま

た、閾値を7件以上と設定したときの不都合点も明らかにしていない。

3. 列車遅延時に投稿されるツイートの分析

3.1 事故発生直後のツイート数の推移

図1は、2012年7月12日にJR京葉線で発生した風速計故障事故に関連して、キーワード「京葉線」を含んだツイート数^{*1}の推移である。図1より以下のことがわかる。

- 事故発生及び運転見合わせ(21時12分)後から急激な伸びを示している。
- JR東日本の公式運行情報の配信時刻^{*2}(21時35分)よりも前に、JR京葉線に関する多くのツイート(21時12分から21時34分の間に計90ツイート)が投稿されている。

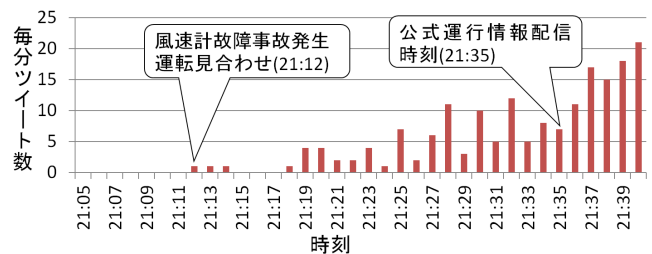


図1 事故発生直後のツイート数の推移

本研究では、遅延発生時刻と公式運行情報配信時刻に大きな差があるという問題点を解決するために、Twitterに流れるこのような情報を活用する。

3.2 時間帯や路線によるツイート数の違い

表1は、いずれも午前11時10分台に発生した人身事故において、発生後30分間に投稿された各路線名を含むツイート数である。利用者数とは、その路線の1日あたりの利用者数を示したもので、文献[12]を参考とした。利用者数の多い路線では、遅延発生時におけるツイート数が多くなることがわかる。両者の相関係数は0.925であった。

表1 人身事故発生後30分間の路線別ツイート数

発生路線	発生日時・場所	ツイート数	利用者数
JR中央線	2012/06/25 11:13	1,015	2,171,822
	武蔵境駅で発生		
東急大井町線	2012/07/06 11:13	232	416,684
西武池袋線	2012/07/11 11:19	167	980,844
	清瀬駅で発生		
JR武蔵野線	2012/07/29 11:14	354	876,365

一方、表2は、いずれもJR京浜東北線で発生した人身事故において、発生後30分間に投稿されたキーワード「京

*1 ただし、リツイート(Retweet)は含まない。以下同様とする。

*2 公式Webサイトに運行情報が掲載されたときの時刻とする。

浜東北線」を含むツイート数である。時間帯利用者数とは、下段の時間帯における京浜東北線の利用者数を示したもので、文献 [13] を参考とした。利用者数が多い時間帯（ラッシュ時）に発生した遅延はツイート数が多くなることがわかる。両者の相関係数は 0.839 であった。

表 2 JR 京浜東北線での人身事故発生後 30 分間のツイート数

発生時間	発生日・場所	ツイート数	時間帯利用者数
7 時 32 分	2012/07/24 洋光台駅で発生	437	1,200,080 (7:30-7:59)
9 時 48 分	2012/07/03 上中里駅で発生	181	905,510 (9:30-9:59)
12 時 27 分	2012/06/15 関内駅で発生	86	626,640 (12:00-12:29)
16 時 01 分	2012/07/16 鶴見駅で発生	133	682,750 (16:00-16:29)
0 時 56 分	2012/07/06 川崎駅で発生	36	152,950 (0:30-0:59)

以上より列車遅延発生時におけるツイート数は発生路線や発生時間帯によって左右されることがわかる。そのため、路線や時間帯によって、遅延検出の閾値を調整する必要がある。

3.3 ツイート中の頻出単語

次に、列車遅延時に投稿されるツイートの内容について分析した。図 2 に示す列車遅延に関連して投稿された「武蔵野線」を含むツイートのうち、出現頻度が多い単語を並べた表を表 3 に示す。ただし、列車の運行状況を表現する単語に限定した。表 3 の状態 A(B) とは、図 2 の状態 A(B) 時に投稿されたツイートを対象としたことを示す。単語の抽出には、形態素解析エンジン MeCab[14] を利用した。

発生路線：JR 武蔵野線
発生場所：西国分寺駅 発生日：2012 年 7 月 29 日
11 時 14 分、人身事故が発生。運転見合わせ。(状態 A)
12 時 42 分、運転再開。それに伴い遅延が発生する。(状態 B)
14 時 30 分、平常通りの運行に戻る。

図 2 対象とした列車遅延 (時刻は公式運行情報に準ずる)

表 3 ツイート中の頻出単語

状態 A(運転見合わせ時)		状態 B(列車遅延時)	
出現単語	出現頻度	出現単語	出現頻度
人身	338	遅れ	102
事故	325	人身	94
運転	295	事故	89
見合わせ	169	運転	77
再開	114	遅延	39
止まっ	109	再開	37

結果、状態 A,B に共通して、「人身」「事故」といった遅延の要因となるキーワードを使ったツイートが多く投稿さ

れていることがわかる。状態別にみると、運転見合わせ時(状態 A)には、「止まっ」「見合わせ」という単語を使ったツイートの投稿が多く、列車遅延時(状態 B)には、「遅れ」「遅延」という単語を使ったツイートの投稿が多い。

4. 提案システム

4.1 システム概要

図 3 に提案するシステムの概要を示す。

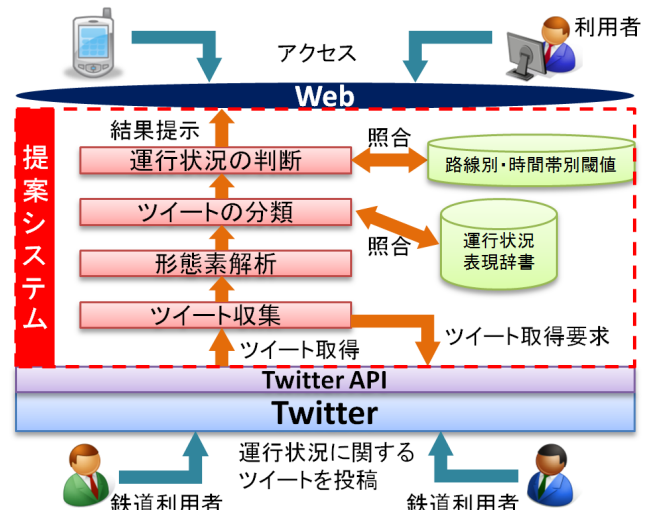


図 3 提案システムの構成図

システムは、Web アプリケーションとして実装する。運行情報の配信対象とする鉄道路線は、首都圏を運行する JR(7 路線：山手線、京浜東北線、総武線、中央線、東海道線、埼京線、武蔵野線)、地下鉄(7 路線：銀座線、丸ノ内線、日比谷線、東西線、千代田線、有楽町線、半蔵門線)、私鉄(7 路線：小田急線、京王線、京急線、田園都市線、東横線、東上線、池袋線)の計 21 路線とする。

はじめに、利用者からのアクセスにより、各路線名を含むツイートを収集する。その後、収集した各ツイートに対して形態素解析をかけ、ツイートを単語レベルまで分割する。その中から列車の運行状況に言及しているツイートのみを抜き出すことで、分類を行う。分類時には運行状況表現辞書(詳細は後述)を用いる。最後に、路線別・時間帯別に定義した閾値に基づいて各路線の運行状況を判断し、結果を提示する。

以下、それぞれの処理の詳細を述べる。

4.2 ツイート収集

Twitter API[15] を用いて、各鉄道路線に関するツイートの監視及び収集を行う。収集対象は、各路線名が含まれているツイートである。ただし、リツイートは除外する。

例えば、JR 山手線の場合は「山手線」、東京メトロ銀座線の場合は「銀座線」、東武東上線の場合は「東上線」というキーワードを含むツイートを収集する。

4.3 形態素解析

4.2 で得られた各ツイートの本文に対して形態素解析を行う。提案システムでは MeCab を使用し、IPA 品詞体系に基づいて解析を行うものとする。

4.4 ツイートの分類

各ツイートを以下の 4 つのカテゴリのいずれかに分類する。以下の「例」は、そのカテゴリに分類できる実際に投稿されたツイートの例、() 内は投稿時刻である。

(1) 見合わせツイート (以下, T_{stop})

列車の運転見合わせに関する情報、列車の運転が止まっている旨が記載されているツイート群。

例:「横須賀線動かないんですけど」(2012/07/23 08:02:09),「小田急線、運転見合わせ、江島線に影響するかな」(2012/07/14 07:08:09)

(2) 遅延ツイート (以下, T_{delay})

列車の遅延に関する情報、列車に遅れが発生している旨が記載されているツイート群。

例:「出勤なう。りんかい線遅れてるー！」(2012/08/08 10:21:19),「京王線遅延してるから吉祥寺まで行って高尾からかーえろ！」(2012/08/08 11:17:19)

(3) 平常運転ツイート (以下, T_{normal})

遅延が解消して平常運転に戻った旨が記載されているツイート群。

例:「さすが常磐線！平常運転に戻るとる ww」(2012/07/05 17:00:23),「宇都宮線、平常ダイヤ回復なう(^_^) #宇都宮線」(2012/07/10 16:51:14)

(4) その他のツイート (以下, T_{other})

上記 3 つに当てはまらない列車の運行情報とは無関係のツイート群。

例:「西武池袋線、車内冷やしすぎ...寒い」(2012/07/05 19:06:56),「横浜線にうるせー非常識なギャルがいっぱいいいた。」(2012/06/17 18:47:22)

なお、列車遅延が起ると、「京浜東北線人身事故かー。」(2012/06/15 12:58:15)「田園都市線、三茶で、人身。」(2012/07/30 19:14:22) のような遅延要因のみで構成されるツイートも多く投稿される。それらは、運行状況に異常があることを示唆しているが、「運転を見合わせている」「遅延している」といった直接的な表現は登場しないため、 T_{stop} であるか T_{delay} であるかを判断できない。従って、この場合に限り、それらのツイートは T_{stop} と T_{delay} の両者のカテゴリに属することとする。

分類作業は、運行状況表現辞書に登録されている各単語が各ツイート中に存在するかを照合しながら行う。存在する場合には T_{stop} または T_{delay} または T_{normal} に分類し、存在しない場合には T_{other} に分類する。

運行状況表現辞書には、3.3 に示したような運行状況が記載されているツイートにおいて出現頻度が高い単語、

[16], [17] をもとに列車が遅延する要因となる単語の計 275 語とその品詞を登録する。データベースは MySQL を使用する。以下に登録する単語の一例を示す。

見合わせに関する語 (T_{stop} へ分類)
見合わせ, 運休, 止まる, 止まっ, 動け, ストップ
遅延に関する語 (T_{delay} へ分類)
遅延, 遅れ, 遅れる, 遅らす, 乱れ, 徐行, 調整
平常運転に関する語 (T_{normal} へ分類)
平常, 通常, 回復, 解消, 戻る, 戻っ
遅延要因に関する語 (T_{stop} と T_{delay} へ分類)
人身, 事故, 異常, 点検, 異音, 障害, 衝突, 大雨, 地震

図 4 運行状況表現辞書の一例

4.5 運行状況の判断

30 秒おきに各路線の運行状況を判断する。提案システムで提示する運行状況は、「平常運転」、「運転見合わせ」、「遅延」の 3 つとする。運行状況を決定するアルゴリズムは以下に従う。 $|T|$ はカテゴリ T に分類されたツイート数、 θ は閾値である。

アルゴリズム

(1) $|T_{stop}| + |T_{delay}| \geq \theta$ であるならば (2) へ。そうでなければ, (4) へ。

(2) $|T_{stop}| > |T_{delay}|$ であるならば、運行状況を「運転見合わせ」と判断する。そうでなければ, (3) へ。

(3) 運行状況を「遅延」と判断する。

(4) 運行状況を「平常運転」と判断する。

閾値 θ は、式 (1) をもって定める。

$$\theta = L \times T \times \alpha \quad (1)$$

α は基本閾値とし、常に固定した値をとる。そして、3.2 より路線や時間帯によって列車遅延発生時のツイート数にばらつきがあることが確認できたため、路線係数 L 及び時間帯係数 T を導入し、基本閾値 α を調整する。調整した結果が最終的な閾値 θ となる。ただし、 $0 < L, T \leq 1.00$ とし、 α は自然数とする。

以下、「運転見合わせ」「遅延」を、 \times 、「平常運転」を \times と表記する。

4.5.1 路線係数 L , 時間帯係数 T の決定

L, T は [12], [13] を参考に路線及び時間帯ごとに値を定義した。ただし、時間帯は 1 時間刻みとする。

L は対象路線の中で利用者が一番多い JR 山手線を 1.00 とし、その他の路線は、(該当路線の利用者数)/(JR 山手線の利用者数) として値を定めた。例えば、JR 京浜東北線は 0.67 であり、小田急線は 0.46 である。 T は首都圏における鉄道利用者数が一番多い時間帯である 8:00~8:59 を 1.00 とし、その他の時間帯は、(該当時間帯の利用者数)/(8:00~8:59 の利用者数) として値を定めた。例えば、7:00~7:59 は 0.88 であり、22:00~22:59 は 0.62 である。

4.5.2 基本閾値 α の決定

α は以下の手順で決定する．まず，本稿で登場する用語を定義する (図 5) ．

- 列車遅延期間...列車の遅延があったとされる期間．公式運行情報に記載されている列車遅延の発生時刻から列車の運行が元に戻って列車遅延に関する公式運行情報の配信が打ち切られた時刻までの間とする．
- 正受入...列車遅延期間において ○ と判断すること．
- 正拒否...列車遅延期間外において × と判断すること．
- 誤受入...列車遅延期間外において ○ と判断すること．
- 誤拒否...列車遅延期間において × と判断すること．

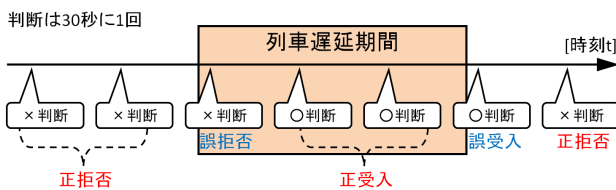


図 5 用語の定義

このとき，誤受入率 F_{accept} と誤拒否率 F_{reject} を式 (2),(3) のように定義する．

$$F_{accept} = \frac{\text{誤受入数}}{\text{正拒否数} + \text{誤受入数}} \quad (2)$$

$$F_{reject} = \frac{\text{誤拒否数}}{\text{正受入数} + \text{誤拒否数}} \quad (3)$$

続いて， α を決めるために実施した実験について述べる．はじめに，2012年11月18日8:00~19日20:00にかけて投稿されたJR山手線に関連するツイート(キーワード「山手線」を含むツイート)計7,417ツイートを収集し，4.4に従って分類した．上記期間に発生した公式運行情報の配信があった列車の遅延は2件である(図6)．

発生した遅延 1	発生日：2012年11月18日
11時07分，JR京浜東北線内の人身事故の影響で運転見合わせ．	
12時15分，運転再開．それに伴い遅延が発生する．	
13時30分，平常通りの運行に戻る．	
発生した遅延 2	発生日：2012年11月18日
20時14分，池袋駅で信号確認のため運転見合わせ．	
20時36分，運転再開．それに伴い遅延が発生する．	
21時40分，平常通りの運行に戻る．	

図 6 対象とした列車遅延 (時刻は公式運行情報に準ずる)

次に，上記を実験材料として30秒ごとに運行状況の判断(または×)を下した．その際， α の候補となる値1~500を変動させながら，各値におけるそれぞれの判断結果及び正受入数・正拒否数・誤受入数・誤拒否数の集計をとった．最終的に $|F_{accept} - F_{reject}|$ が最小となる値を基本閾値 α として採用した．

判断材料としては，まず直前30秒間に投稿されたツイ

トを用いる手法(以下，手法1)の検討を行った(図7)．これは，毎回の判断で用いるツイートはいずれも独立となる手法である．手法1をもとにした F_{accept}, F_{reject} を図9に示す．図9より， $\alpha = 1$ のとき， $F_{accept} = 0.1976, F_{reject} = 0.1441$ であるから，手法1における基本閾値 α は1とした．次に，手法1では列車遅延各件を洩れなくかつ誤りなく検出しているかの評価を行った．評価に使用したのは，公式配信があった列車遅延時に投稿されたツイート群(以下A)と平常運転時に投稿されたツイート群(以下B)の各50時間分のデータである．Bは，最初の誤受入となる判断から30分間は同件とみなすことで誤検出件数を求めた．評価結果を表4に示す．ここで，各用語の定義と適合率 P_d ・再現率 R_d ・F値・正受入率 T_{accept} の計算式は式(4)-(7)とする．F値は適合率と再現率の調和平均を示す．

- 正検出...公式運行情報の配信があった列車遅延において，その列車遅延期間中に正受入となる判断があった．
- 誤検出...平常運転あるいは公式運行情報の配信がない僅かな列車遅延であるのに関わらず誤受入となる判断があった．
- 検出洩れ...公式運行情報の配信があった列車遅延において，その列車遅延期間中に正受入となる判断がなかった．

$$\text{適合率 } P_d = \frac{\text{正検出件数}}{\text{正検出件数} + \text{誤検出件数}} \quad (4)$$

$$\text{再現率 } R_d = \frac{\text{正検出件数}}{\text{正検出件数} + \text{検出洩れ件数}} \quad (5)$$

$$F \text{ 値} = \frac{2 \times P_d \times R_d}{P_d + R_d} \quad (6)$$

$$T_{accept} = 1 - F_{reject} \quad (7)$$

その結果，手法1は公式運行情報の配信があった列車遅延に関してはすべて検出できたものの，適合率が著しく低く，ノイズを誤って拾ってしまうケースが相次いだ(表4)．これは，手法1では毎回の判断をそれぞれ独立な事象として扱っており，実際に列車の遅延が起こっていない場合あるいは短時間でわずかな遅延である場合でも急に θ を上回る，逆に実際に列車の遅延が起こっていても急に θ を下回ることが頻発するためである．

そこで次に，直前 $m(m > 0.5)$ 分間に投稿されたツイートをを用いる手法(以下，手法2)を検討する(図8)．今回は $m = 10$ とする．これは，手法1の問題点を踏まえ，前に起こった事象を一定程度考慮しながら毎回の判断を行う手法である．また，毎回の判断において判断材料とするツイートにも被りが発生するため，急に θ を上回ったり，下回ったりすることを防いだ．手法2をもとにした F_{accept}, F_{reject} を図10に示す．図10より， $\alpha = 24$ のとき， $F_{accept} = 0.0793, F_{reject} = 0.0764$ であるから，手法2における基本閾値 α は24とした．手法1と同様の評価を手法2に対しても行った結果，手法2の適合率とF値が

手法 1 より大きく改善したことがわかった (表 4) . また , T_{accept} も手法 1 より手法 2 の方が高いことから , 列車遅延期間中に正受入の判断を下すことが多い .

従って , 本研究では手法 2 を採用して提案システムを実装することとする .

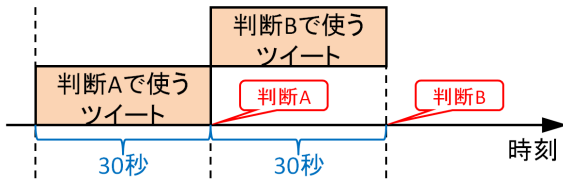


図 7 手法 1 の概要

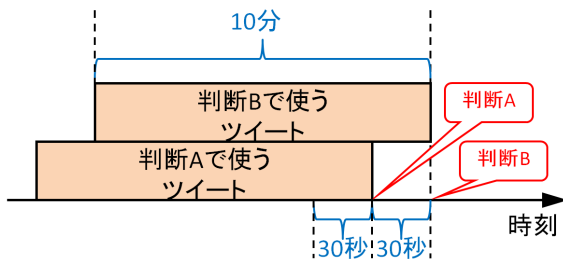


図 8 手法 2 の概要

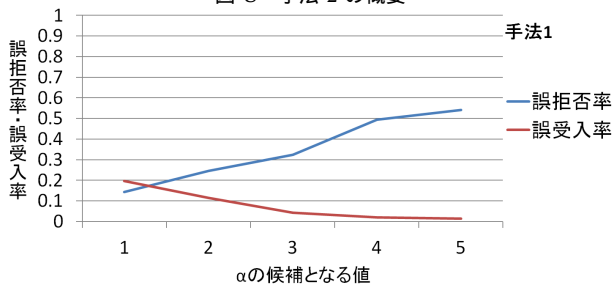


図 9 手法 1 における誤拒否率・誤受入率

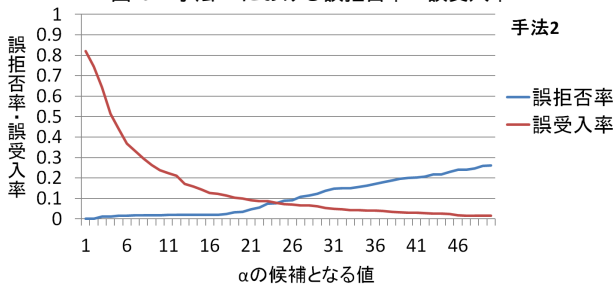


図 10 手法 2 における誤拒否率・誤受入率

表 4 手法 1,2 における適合率・再現率・F 値・ α における正受入率

	適合率	再現率	F 値	α における正受入率
手法 1	0.15	1.00	0.26	0.85
手法 2	0.63	1.00	0.77	0.92

4.6 インターフェース

提案システムのインターフェースは以下とする (図 11) . 画面上的各路線名をクリックすることで , 最新の運行状況が確認できる仕様とした (図 12) . 運行状況は 30 秒おきに更新される .



現在、試験運用中です。(学内ネットワーク内のみで公開しています)

対応路線:

JR 線 JR Line	地下鉄 Subway	私鉄 Private railway
山手線 Yamanote Line	銀座線 Ginza Line	小田急線 Odakyu Line
京浜東北線 Keihin-Tohoku Line	丸ノ内線 Marunouchi Line	京王線 Keio Line
総武線 Sobu Line	日比谷線 Hibiya Line	京急線 Keikyu Line
中央線 Chuo Line	東西線 Tozai Line	東急田園都市線 Tokyu Den-en-toshi Line
東海道線 Tokaido Line	千代田線 Chiyoda Line	東急東横線 Tokyu Tokyoko Line
埼京線 Saikyo Line	有楽町線 Yurakucho Line	東武東上線 Tobu Tojo Line
武蔵野線 Musashino Line	半蔵門線 Hanzomon Line	西武池袋線 Seibu Ikebukuro Line

図 11 提案システムのインターフェース (路線選択)

東急田園都市線
Tokyu Den-en-toshi Line

最新の運行状況:

列車の運転を見合わせているおそれがあります。

図 12 提案システムのインターフェース (運行状況の提示)

5. 評価実験

提案システムに対して以下の評価実験を実施した .

5.1 ツイートの分類に関する性能

2012 年 10 月 23 日 ~ 25 日に発生した公式運行情報の配信があった列車遅延 5 件に関連して , 遅延発生直後に投稿された計 5,574 ツイートを収集し , 4.4 に従い分類した . その際 , 各ツイートを正しく分類できたかを評価する . 評価結果を表 6 に示す . 評価指標は表 6 , 式 (8)-(10) とする .

その結果 , F 値はいずれも 0.79 以上と高い値となった . 特に T_{stop} と T_{delay} の分類における適合率は 0.95 を上回っており , ノイズの誤分類が極めて少ないことがわかる . よって , ツイートの分類に関して一定の性能を持つことが示せた .

しかし , 4.4 ではツイートを字面上での解析しか行っていない . そのため , 皮肉を含むツイートが理解できない , 運行状況表現辞書に登録した単語が違う意味で使われているなど以下のような誤分類があった .

- 「本日も中央線は平常運行の様。(訳 : 順調に電車止まりまくりんぐ)」 (2012/10/24 18:15:23)
→ T_{normal} に誤って分類
- 「中央線新宿・秋葉原の駅間隔の短さは異常。」 (2012/10/24 18:00:07)
→ T_{stop}, T_{delay} に誤って分類
- 「イーモバイル、中央線で繋がらなかったけど通信障害か？」 (2012/10/24 18:53:02)
→ T_{stop}, T_{delay} に誤って分類

表 5 分割表 ($Tweet$ は分類対象のツイート, $A(Tweet)$ は $Tweet$ の手動による分類結果, $B(Tweet)$ は $Tweet$ の 4.4 に従う分類結果, T はカテゴリを示す.)

	$A(Tweet) = T$	$A(Tweet) \neq T$
$B(Tweet) = T$	w	x
$B(Tweet) \neq T$	y	z

$$\text{適合率 } P_c = \frac{w}{w+x} \quad (8)$$

$$\text{再現率 } R_c = \frac{w}{w+y} \quad (9)$$

$$F \text{ 値} = \frac{2 \times P_c \times R_c}{P_c + R_c} \quad (10)$$

表 6 分類性能評価結果

	適合率	再現率	F 値
T_{stop} に分類	0.95	0.79	0.86
T_{delay} に分類	0.98	0.85	0.91
T_{nomal} に分類	0.68	0.95	0.79
T_{other} に分類	0.80	0.96	0.87

5.2 公式運行情報と提案システムの配信時刻の差

提案システムを以下の期間内に稼働させ、期間内に運行情報の配信対象とした全 21 路線で発生した公式運行情報の配信があった列車遅延計 41 件について、公式運行情報と提案システムとの第一報配信時刻の差を検証した。評価結果を表 7 に示す。表中の最大差・平均差は、提案システムの配信時刻から公式運行情報の配信時刻を引いたときの差である。

- 稼働期間：2012 年 12 月 10 日 13:00 ~ 24 日 13:00

表 7 公式運行情報と提案システムの配信時刻差の評価結果

	該当件数 (%)	最大差	平均差
提案システムの方が早い	31 件 (75.6%)	-76 分	-13.2 分
公式運行情報の方が早い	6 件 (14.6%)	+10 分	+4.8 分
同時刻の配信	1 件 (2.4%)	0 分	0 分
未検出	3 件 (7.3%)	-	-
合計	41 件 (100%)	-	-

関連研究 [11] における評価結果は表 8 のとおりである。なお関連研究では、「Yahoo!運行情報」と Bot ツイートの配信時刻を比較対象としている*3。

評価結果より、全体のうち 31 件 (75.6%) で公式運行情報よりも前に、提案システムにおいて列車の遅延を検出することができた。また、関連研究よりも高い割合で、公式運行情報よりも早く列車遅延の第一報を配信することができた。従って、提案システムではリアルタイム性の観点か

*3 「Yahoo!運行情報」及び Bot ツイートも公式運行情報に基づいて運行情報の配信を行っている。

ら一定の有効性を示した。

特に、2012 年 12 月 16 日に JR 山手線で発生した人身事故による遅延と 2012 年 12 月 22 日に東京メトロ千代田線で発生した車両接触事故による遅延は、発生後 1 分で検出できた。更に、2012 年 12 月 13 日に JR 川越線内で発生した異常音確認の影響を受けた JR 埼京線の遅延は、公式運行情報配信時刻よりも 1 時間 16 分前に検出した。

遅延要因別でみると、車両点検・線路確認・踏切確認が原因で発生した遅延や他路線の影響を受けて発生した遅延は、公式運行情報が配信されるまでに時間がかかる傾向があり、提案システムとの配信時刻差が顕著に現れた。一方で、人身事故で発生した遅延は、事故発生直後に公式運行情報が配信される傾向があり、提案システムとの配信時刻差はわずかなものとなった。

発生時刻別でみると、5 時台や 6 時台など早朝に発生した遅延は、閾値 θ を低めに設定してあるものの、投稿されるツイートの絶対数が少ないため、遅延の検出までに時間がかかった場合があった。

表 8 関連研究における配信時刻差の評価結果

	該当件数 (%)
関連研究の方が早い	13 件 (41.9%)
Yahoo!運行情報または Bot ツイートの方が早い	13 件 (41.9%)
未検出	5 件 (16.1%)
合計	31 件 (100%)

6. まとめと今後の課題

本研究では、より早い列車運行情報の提供を行うために、鉄道利用者のツイートを情報源として列車運行状況の判断を行うシステムを提案した。そして、適切な閾値を設定することによって正確かつ素早く列車遅延の検出を行うことを目指した。また、列車遅延発生時におけるツイート数は発生路線や発生時間帯によって左右されることや列車遅延時に投稿されるツイートの頻出単語についての分析を行った。

検討の結果、提案システムでは、30 秒ごとに各路線の運行状況の判断を行い、毎回の判断には直前 10 分間に投稿されたツイートを判断材料とした。その際、遅延を検出するための閾値を最大で 24 として実装した。閾値は路線や時間帯によって変動する。

評価実験の結果、全体の 75.6% で公式運行情報配信時刻よりも前に、提案システムにおいて列車遅延を検出・配信できた。このことから、リアルタイム性の観点において、提案システムの有効性を明らかにでき、関連研究に対する優位性を確認した。

今後は、いつ・どこで・なぜ遅延が発生したかといった詳細な運行情報の表示、教師あり機械学習を用いたツイー

トの分類及び性能評価, m のパラメータの最適化およびノイズの誤検出を更に低減させるための手法の検討, 大量のアクセスに耐え得るシステムの設計が課題となる. また, 同じ名称の路線が全国に複数存在する場合には, 路線名が含まれているツイートを収集しただけでは, 遅延が起きている路線を完全に特定できない, 運行状況が正しく判断されない可能性があるといった課題もある. 例えば, 「中央線」という路線は JR の路線でもあるし, 大阪市営地下鉄の路線でもある. この課題についても解決手法の検討を進めたい.

参考文献

- [1] Twitter,
<https://twitter.com/>
- [2] Twitter Blog: Celebrating #Twitter7,
<http://blog.twitter.com/2013/03/celebrating-twitter7.html>
- [3] 榎剛史, 松尾豊: ソーシャルセンサとしての Twitter - ソーシャルセンサは物理センサを凌駕するか? -, 人工知能学会誌, Vol.27, No.1, pp.67-74 (2012).
- [4] Sakaki, T., Okazaki, M. and Matsuo, Y.: Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors, *Proc. 18th International World Wide Web Conference*, pp.851-860 (2010).
- [5] Sakaki, T., Okazaki, M. and Matsuo, Y.: Tweet Analysis for Real-Time Event Detection and Earthquake Reporting System Development, *IEEE Transactions on Knowledge and Data Engineering*, Vol.25, No.4, pp.919-931 (2013).
- [6] Sakaki, T., Matsuo, Y., Yanagihara, T., Naiwala, P. C. and Nawa, K.: Real-time Event Extraction for Driving Information from Social Sensors, *Proceedings of the 2012 IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems*, pp.221-226 (2012).
- [7] Napong, W., Wasawat, P., Wasan, P. A. and Pimwadee, C.: Social-based Traffic Information Extraction and Classification, *2011 11th International Conference on ITS Telecommunications*, pp.107-112 (2011).
- [8] 荒牧英治, 増川佐知子, 森田瑞樹: Twitter Catches the Flu: 事実性判定を用いたインフルエンザ流行予測, 情報処理学会研究報告, Vol.2011-SLP-86, No.1, pp.1-8 (2011).
- [9] 高橋哲郎, 野田雄也: 実世界のセンサーとしての Twitter の可能性, 電子情報通信学会技術研究報告, Vol.110, No.400, NLC2010-38, pp.43-48 (2011).
- [10] 古川忠延, 阿部修也, 安藤剛寿, 岩倉友哉, 志賀聡子, 高橋哲郎, 井形伸之: Twitter からの犯罪情報抽出の可能性調査, 情報処理学会研究報告, Vol.2011-DD-82, No.3, pp.1-6 (2011).
- [11] 築瀬拓弥, 増田英孝, 山田剛一, 荒牧英治, 中川裕志: Twitter を用いた電車遅延の自動通知, 情報処理学会研究報告, Vol.2013-DD-89, No.1, pp.1-6 (2013).
- [12] 国土交通省: 第 11 回大都市交通センサス 駅別発着・駅間通過人員表 (首都圏) (2010).
- [13] 国土交通省: 第 11 回大都市交通センサス 路線別着時間帯別駅間輸送定員表 (首都圏) (2010).
- [14] MeCab,
<http://mecab.sourceforge.net/>
- [15] Twitter Developers,
<https://dev.twitter.com/>
- [16] 国土交通省: 鉄道トラブルによる影響に関する調査結果の概要 (2009).
- [17] 国土交通省: 鉄軌道輸送の安全にかかわる情報 (2012).