

ソーシャルセンサからの鉄道運行情報検出システムの開発*

長野 伸一[†] 上野 晃嗣[†] 長 健太[†]

Development of a System Detecting Train Status Information from Social Sensors*

Shinichi NAGANO[†], Koji UENO[†], and Kenta CHO[†]

あらまし 近年、エネルギーの効率的利用という観点から注目を浴びているスマートコミュニティを構成するモジュールの一つに、センサから取得した情報を利用してリアルタイムに交通情報を把握する次世代交通システムがある。本論文では、ソーシャルセンサとしての Twitter から交通情報、特に鉄道の運転見合わせや遅延などの運行情報をどの程度正確、且つ、迅速に検出できるのか検証した。ツイートがもつ情報の不確実性についてはヒューリスティックなルールによるテキスト処理、路線による情報量の違いについては統計処理におけるパラメータ値最適化により、検出結果の F 値が 0.85、検出までに掛かる時間が 3 分台となることを確認した。また、鉄道事業者が発表する公式情報よりも早く、また、鉄道事業者が発表しない小規模な事象も検出できる場合があることを確認した。

キーワード ソーシャルセンサ, Twitter, 鉄道運行情報, 次世代交通システム

1. ま え が き

東日本大震災以降、スマートコミュニティが注目を浴びている。スマートコミュニティとは、様々な ICT ソリューションをシステムとして有機的に結びつけることで、エネルギーの効率的利用を目指す社会インフラを指す [1], [2]。スマートコミュニティの構成要素の一つに次世代交通システムがある。次世代交通システムは、交通インフラの環境に設置したカメラやセンサからの情報、あるいはインフラ利用者の GPS 情報を収集して、リアルタイムに交通情報を把握する。その交通情報を利用者に還元することで、人々の移動経路や移動手段を最適化し、結果的に全体としてのエネルギー使用量を削減するというものである [3]~[5]。

次世代交通システムでは、GPS やカメラなどの物理センサに加えて、ソーシャルセンサの利用も検討されている。ソーシャルセンサとは、ソーシャルメディア上でユーザが投稿した発言をセンサの一種とみなし、発言内容を解析して交通インフラに関する情報を検出

する手法である [6]~[8]。代表的なソーシャルメディアの一つに Twitter がある。Twitter は、ユーザがツイートと呼ばれる最大 140 文字のテキストを投稿するサービスで、日本国内での利用者数が 2011 年 10 月時点で約 1400 万人以上にのぼる [9]。スマートフォンの普及と相まって、ユーザが外出中に投稿する利用形態が増えており、身の回りで見聞したり経験したりした出来事も数多く投稿されている。そのため、ユーザの目線で観察された情報が、物理センサで観察された情報を補完するものとして期待されている。すなわち、物理センサでは取得できない情報を取得したり、物理センサが設置されていない場所でも情報を取得できる可能性をもつ。一方で、Twitter 上では誰もが自由に投稿できるため、発言内容に含まれる情報の正確性は担保されないという欠点もある。

本論文では、ソーシャルセンサとしての Twitter から交通情報、特に鉄道運行情報をどの程度正確、且つ、迅速に検出できるのかその可能性を探る。自動車運転中、運転者はスマートフォンや携帯電話を使えないため、道路交通情報についてのツイートは少ないと考え、本論文では鉄道運行情報に着目する。鉄道運行情報には一般には運転再開や運転見合わせ区間などの情報も含まれるが、ソーシャルセンサとしての可能性を探る

[†] 株式会社東芝研究開発センター, 川崎市
Corporate R&D Center, Toshiba Corporation, 1, Komukai
Toshiba-cho, Saiwai-ku, Kawasaki-shi, 212-8582 Japan

* 本論文はシステム開発論文である。

という観点から、まずは運転見合わせと遅延に関する情報のみを扱う。

鉄道事業者は、鉄道ダイヤに影響する事故などの事象が発生すると、各事業者の Web サイト上^(注1)で、運転見合わせ、運転再開、遅延、通常ダイヤ復帰に関する情報を順次発表する。実際には、発生事象による利用者への影響の大きさや、事象の把握状況に応じて、運行情報の開示をコントロールしている。そのため、鉄道利用者へ情報が伝わるのが遅れたり、事象発生時に当該路線にいない利用者へ情報が伝わらないことが起きている。我々が確認したところ、事象発生後1時間以上経過してから第一報が発表された場合もあった。また、鉄道会社によっては30分以上の遅延の発生、あるいは遅延の見込みがある場合にのみ発表している。一方、鉄道事業者側の責任で30分程度の遅延時間が発生した場合、それを許容できる利用者はわずか11.6%しかない[16]。

利用者の視点で考えると、事象発生による影響の大きにかかわらず、鉄道の運行情報に対するアクセシビリティを高めることが望ましいと考えられる。鉄道利用者が運行情報をいち早く知ることができれば、迂回経路を利用する、別の移動手段を利用する、次の予定を変更するなど、その状況に応じた適切な行動を選択できる。更に、地震、台風、大雪などの自然災害で鉄道ダイヤが大幅に乱れた場合は、特に貴重な情報となる。

2. 関連研究

関連研究として、まずソーシャルセンサとしての Twitter から特定の事象を検出する研究について述べる。また、Twitter が出現する以前から、ニュースやブログなどの時系列文書を対象とした分析が行われている。本節では、特定の期間に高い頻度で出現するキーワードを検出する研究について述べる。

2.1 ソーシャルセンサからの事象検出

Sakaki ら [10] は、Twitter から地震の発生を検出する手法を提案した。Support Vector Machine (SVM) による分類器を利用し、地震の発生に関するツイートを抽出することにより、高い確率で地震の発生を検出できることを示した。Aramaki ら [11] は、Twitter からのインフルエンザ流行推定の可能性を示した。具体的には、SVM による分類器を利用して抽出した、イン

フルエンザに掛かった人物が存在することを表現するツイートの数と、インフルエンザの症例数との間に相関関係があることを示した。高橋ら [12] は、Twitter からの花粉飛散量推定の可能性を示した。具体的には、SVM による分類器を利用して抽出した、花粉症に掛かったことを表現するツイートの数と、花粉の飛散量との間に相関関係があることを示した。これらの研究より、特定の分野においてはソーシャルセンサとしての Twitter の有用性が確認されている。

2.2 時系列文書からのトピックワード検出

Swan ら [13] は、時系列文書としてのニュース記事集合から、トピックワードを抽出する Time Mines というシステムを開発した。全文書に対する文書頻度の割合が他の期間と比較して有意に高いキーワードをカイ2乗検定により抽出し、抽出したキーワードをその期間のトピックワードとみなした。Kleinberg [14] は、特定のキーワードを含む文書の出現間隔が短くなっている部分をバースト区間とし、そのキーワードをその区間のトピックワードとして抽出する手法を提案した。状態遷移コストを導入することにより、小さなバースト区間が連続して出現する場合にでも、一つのバースト区間として扱える。これにより、文書頻度を利用するよりも、高い精度でトピックワードを抽出できることを示した。藤木ら [15] は、Kleinberg の手法を拡張し、単位時間当たりの文書頻度の変化を考慮する手法を提案した。これにより、文書頻度が時間によって変化するブログ記事集合や電子掲示板への書き込み記事集合からも正確にバースト区間を検出でき、高い精度でトピックワードを抽出できることを示した。

3. 予備実験と課題

3.1 予備実験

交通情報、特に鉄道分野におけるソーシャルセンサとしての Twitter の有用性を確認するため、簡単な調査を行った。具体的には、鉄道の路線ごとの輸送人員とツイート数の相関関係を調べた。輸送人員は、関東交通広告協議会^(注2)が公表している2010年度の統計データを用いた。一方、ツイート数は、特定期間(2011/07/28~2011/10/11)に収集したツイートから算出した。「山手線」「東横線」などの呼称・愛称キーワード(以下、「路線名キーワード」と呼ぶ)を検索クエリとしてツイートを収集した。1日当たりの路線

(注1) : http://traininfo.jreast.co.jp/train_info/kanto.aspx

(注2) : <http://www.train-media.net/report/index.html>

ごとの輸送人員とツイート数の散布図を図1(a)に示す。輸送人員が公表されており、且つ、我々がツイートを収集した首都圏の68路線を対象とした。結果、輸送人員とツイート数の相関係数は0.83となり、正の相関があると分かった。一方、災害や事故が発生した場合の例として、爆弾低気圧と呼ばれる急発達した低気圧が日本各地を襲った、2012年4月3日に着目する。首都圏での強風のピーク時間が帰宅ラッシュ時間と重なり、大規模な鉄道の制限や運休が発生した。この日のツイート数を図1(b)に示す。相関係数は0.13となり、相関性は見られないと分かった。また、1路線あたりの平均ツイート数を比較すると、(a)では429、(b)では9330あり、1万ツイート/日を超えた路線は、(a)は0、(b)は21あった。特に、影響の大きかった相鉄本線、京成本線では、それぞれ約700倍、約250倍も多く、平常時ではツイート数が非常に少ない東武鬼怒川線、東武小泉線においても、それぞれ約10倍もあった。これらの調査から、鉄道分野において、実世界で発生したイベントに応じてツイート内容が変化し、Twitterは実世界の状況がある程度反映したソーシャルセンサとしての可能性があるとと言える。

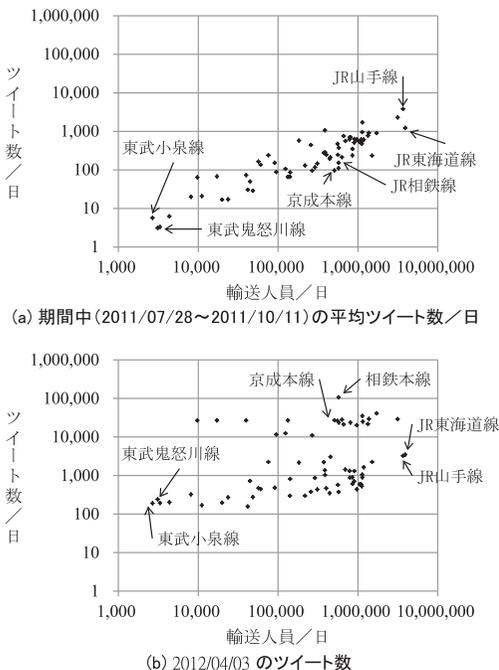


図1 各路線の輸送人員とツイート数の散布図
Fig.1 Scatter diagram of the # of passengers and the # of tweets.

3.2 鉄道運行情報検出の課題

Twitterから鉄道運行情報を検出するにあたり、大きく二つの課題が存在する。一つ目の課題は、情報の不確実性である。例えば、「山手線」と「止まった」を含むツイートを検索すると、以下のようなツイートが収集される。

1. 山手線が止まったので帰れない。
2. 埼京線が止まったので、山手線に乗り換えた。
3. 最近、よく止まるな。RT @taro あれ、山手線が止まった。
4. @hanako 山手線が止まったので、カフェで時間潰してた。

ツイート1からはそのときに山手線が止まっているらしいと分かるが、その他のツイートからはそのようなことは分からない。ツイート2では止まっているのは埼京線であり、山手線は通常運行している可能性が高い。ツイート3、4では山手線が止まっていたのは過去の話題であり、そのときは通常運行に戻っているかもしれない。Twitterから鉄道の運行情報を検出する場合、両者を区別して扱う必要がある。本研究では、運行情報に関するツイートを次のように定義する。

運行情報陽性ツイート：路線のリアルタイムな運行情報に言及しているツイート

運行情報陰性ツイート：注目路線に言及しているが、運行情報が現在のものではない、または別の路線のものであるツイート

ツイート1は運行情報陽性ツイート、ツイート2、3、4は運行情報陰性ツイートである。

二つ目の課題は、路線によってツイート数が異なる点にある。例えば、首都圏の主要路線で長時間電車が止まった場合、多くの人がその事をTwitterで呟くことが予想される。一方、利用者が少ない路線で数十分

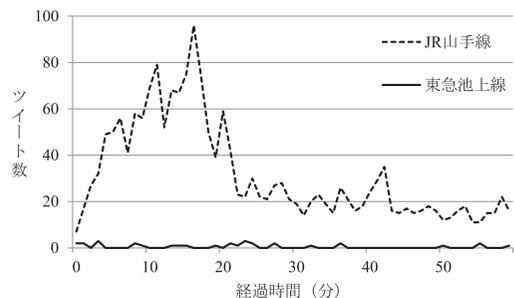


図2 事故発生直後のツイート数分布
Fig.2 Distributions of the #s of tweets after the accidents.

程度電車が遅れた場合、その事を Twitter で呟く人はあまり多くはない。JR 山手線で人身事故が起きて公式運行情報が発表された際の「山手線」という単語が含まれるツイート数、及び、東急池上線で車両故障が起きて公式運行情報が発表された際の「池上線」という単語が含まれるツイート数の事象発生直後から 60 分後までの 1 分ごとの分布を図 2 に示す。ツイート数とその分布は、JR 山手線と東急池上線で大きく異なると分かる。Twitter から鉄道運行情報を検出する場合、このように路線ごとのツイート数の違いを考慮することが重要となる。

4. 鉄道運行情報検出手法

鉄道運行情報検出は、路線ごとに独立して行う。まず、特定の路線に関する最新のツイートを収集する。次に、収集したツイートから、その路線のリアルタイムな運行情報に言及している運行情報陽性ツイートだけを抽出する。そして、抽出した運行情報陽性ツイートに基づいて運行情報の有無を推定する。以下、ツイート収集、運行情報陽性ツイート抽出、及び、運行情報推定について詳細に述べる。

4.1 ツイート収集

Twitter API を利用し、特定の路線に関する最新のツイートを収集する。具体的には、一つの路線名に対する路線名キーワードを 1 回の検索クエリとし、3 分おきに最大 100 件のツイートを取得する。路線名キーワードを表 1 に示す。JR 横須賀線のように複数の路線名キーワードがある場合、OR 検索でツイートを収集する。なお、ツイート数が少ない路線には本手法は適用できない。3.1 で述べたように、輸送人数とツイート数との正の相関性に着目して、1 日あたりの平均輸送人数が 10 万人以上の路線に限定した。

4.2 運行情報陽性ツイート抽出

4.1 で収集したツイートから、運行情報に関連したキーワード（以下、「運行情報キーワード」と呼ぶ）のいずれかを含むものを運行情報陽性ツイート候補として抽出する。平常時のツイートと比較して運行情報発生時のツイートのみ出現回数が多くなるという観点から、次の 14 件のキーワードを運行情報キーワードとして選定した。

運行情報キーワード：止まって、とまって、止まった、とまった、止まる、止まり、遅れる、遅れて、遅れた、ストップ、見合わせ、事故、停止、遅延。

ただし、この処理だけでは、3.2 で説明したように

表 1 路線名キーワード

Table 1 Keywords of train lines.

公式路線名	路線名キーワード
JR 山手線	山手線
JR 中央線	中央線
JR 総武線	総武線
JR 横須賀線	横須賀線, スカ線
JR 高崎線	高崎線
JR 川越線	川越線
京王井の頭線	井の頭線
東京メトロ副都心線	副都心線
東武東上線	東上線
東急池上線	池上線

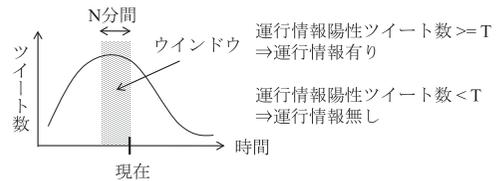


図 3 運行情報推定

Fig. 3 Estimation of train status information.

運行情報陰性ツイートも含まれてしまう。そこで、以下の三つのルール全てを満たすものだけを運行情報陽性ツイートとして取り出す。

ルール 1: 路線名キーワードの後に運行情報キーワードが出現する。

ルール 2: ユーザ発言部分（引用でない部分）に路線名が出現する。

ルール 3: 返信ツイート（@から始まるツイート）ではない。

ルール 1 は、「埼京線が止まったから、山手線に乗り換える。」のように他の路線の運行情報に言及しているツイートを除去するためである。JR 埼京線では運行情報陽性ツイートとなり、JR 山手線では運行情報陰性ツイートとなる。ルール 2 とルール 3 は、情報の信頼性が低いツイートを除去するためである。ユーザ発言部分に路線名が出現しない場合、ユーザはその路線に乗っていない可能性が高い。また、返信ツイートはコミュニケーションに時間のズレが生じるため、情報の鮮度が低下する傾向がある。

4.3 運行情報推定

4.2 で抽出した個々の運行情報陽性ツイートを見ても、運行情報が含まれていると言い切ることはできない。なぜなら、ツイートには常に正しい情報が書かれているとは限らないからである。ユーザが意図せずに、誤った情報を書いてしまう場合もある。そこで、運行

情報陽性ツイートの頻度に基づいて、運行情報の有無を推定する. 具体的には、直近 N 分間 (ウィンドウ) での運行情報陽性ツイートの頻度がしきい値 T を超えると運行情報有り, 超えなければ運行情報無しと判断する (図 3).

5. 評価実験

Twitter からどの程度正確, 且つ, 迅速に鉄道運行情報を検出できるのか確認するため, 鉄道運行情報検出の評価実験を実施した. 以下, 評価用に作成したデータセット, 評価指標, 評価結果について述べる.

5.1 データセット

首都圏 10 路線を対象に, 4.1 で述べた方法で 1 日分のツイートを収集し, 1 分単位で運行情報の有無を手動でラベル付けた. 鉄道事業者の公式運行情報が発表された 1 日を対象とした. 各路線の運行情報発生時間, 運行情報発生件数, ツイート数を表 2 に示す. 運行情報が発生してから消滅するまでを運行情報発生区間とし, 一つの運行情報発生区間を 1 件の運行情報とみなした.

5.2 評価指標

運行情報陽性ツイート抽出と運行情報推定の評価指標について述べる. 運行情報陽性ツイート抽出の評価指標には, 再現率, 適合率, F 値を利用する. それぞれ, 以下の式で計算される.

$$(\text{再現率}) = TP / (TP + FN)$$

$$(\text{適合率}) = TP / (TP + FP)$$

$$(\text{F 値}) = 2 \times (\text{再現率}) \times (\text{適合率}) / ((\text{再現率}) + (\text{適合率}))$$

TP は運行情報陽性ツイートを運行情報陽性ツイートとしてシステムが正しく抽出した件数, FN は運行情報陽性ツイートを運行情報陰性ツイートとしてシステムが誤って抽出した件数, FP は運行情報陰性ツイートを運行情報陽性ツイートとしてシステムが誤って抽出した件数である. 本論文では, 抽出漏れの少なさ, 抽出誤りの少なさをバランスした指標として F 値を重視する.

次に, 運行情報推定の評価指標は, 正確性と検出時間という二つの観点から見る. 正確性は運行情報をどの程度正確に検出できたのかを示す. 4. で述べた運行情報検出による 1 分ごとの結果が, 1 分ごとに手動でラベル付けた結果とどの程度一致するかで決まり, 一致している時間が長ければ正確性は高くなり, 短ければ正確性は低くなる. ここで, 一致の程度は, 運行

表 2 運行情報発生時間, 運行情報発生件数, ツイート数
Table 2 The elapsed time while train status information had occurred, the # of the occurrences, and the # of tweets.

路線	運行情報発生時間 (分)	運行情報発生件数	ツイート数
JR 山手線	188	1	2,394
JR 中央線	518	3	3,198
JR 総武線	218	2	504
JR 横須賀線	82	1	306
JR 高崎線	157	1	266
JR 川越線	296	1	233
京王井の頭線	77	1	338
東京メトロ副都心線	15	1	16
東武東上線	234	2	406
東急池上線	118	1	57

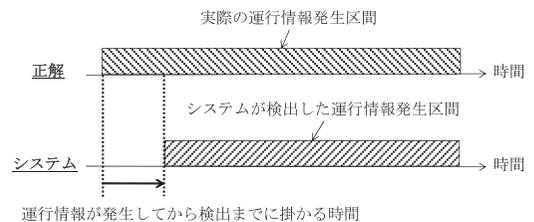


図 4 運行情報の検出時間

Fig. 4 Detection time of train status information.

表 3 評価結果 (運行情報陽性ツイート抽出)
Table 3 Evaluation results (Positive tweet extraction).

	再現率(%)	適合率(%)	F 値(%)
比較手法	71.6	81.8	76.4
本手法(ルールなし)	82.4	69.0	75.1
本手法(ルール 1)	82.4	71.8	76.7
本手法(ルール 2)	77.8	77.8	77.8
本手法(ルール 3)	81.8	73.8	77.6
本手法(ルール 1,2,3)	77.3	85.5	81.2

情報陽性ツイート抽出の評価指標と同様, 再現率, 適合率, F 値で表現される. TP は運行情報の発生をシステムが正しく検出できた時間, FN は運行情報の発生をシステムが正しく検出できなかった時間, FP は運行情報の発生をシステムが誤って検出した時間である. 一方, 検出時間は運行情報をどの程度迅速に検出できたのかを示す. 運行情報が実際に発生してから検出までに掛かる時間で決まり, この時間が短いほどよい (図 4).

5.3 評価結果 (運行情報陽性ツイート抽出)

運行情報陽性ツイート抽出の評価結果を表 3 に示す. JR 山手線についてのツイートから, ランダムに抽出した 1,000 件のツイートを評価に用いた. それぞれのツイートに対し, 運行情報陽性ツイート・運行情

表 4 評価結果 (運行情報推定)
Table 4 Evaluation results (Expectation of train status information).

	ウィンドウの長さ(N)	閾値(T)	検出時間(分)	再現率(%)	適合率(%)	F 値(%)
比較手法 1	-	-	*	32.3	99.4	48.8
比較手法 2	10	0.01	0.0	98.0	49.1	65.4
		0.05	0.0	97.2	53.2	68.8
		0.10	0.1	93.3	56.4	70.4
		0.15	0.2	88.2	57.1	69.3
	20	0.01	0.0	100.0	38.9	56.1
		0.05	0.0	99.9	44.5	61.5
		0.10	0.9	96.8	49.7	65.7
		0.15	1.0	91.3	50.9	65.4
	30	0.01	0.0	100.0	33.3	50.0
		0.05	0.0	100.0	39.1	56.2
		0.10	0.9	97.5	46.5	63.0
		0.15	1.0	92.7	47.6	62.9
本手法	10	1	0.0	98.0	49.1	65.4
		5	5.6	68.4	95.3	79.7
		10	*	44.5	99.9	61.5
	20	1	0.0	100.0	38.8	55.9
		5	5.2	87.6	80.3	83.8
		10	*	68.6	95.8	80.0
	30	1	0.0	100.0	33.2	49.9
		5	4.2	95.3	72.2	82.2
		10	*	80.6	88.6	84.4

*: 運行情報を全く検出できなかった区間が存在するため、検出時間の平均値を算出できなかった。

報陰性ツイートを手動でラベル付けした。比較手法には、ツイートの長さや路線名キーワード、運行情報キーワード、引用記号 (RT, QT)、ユーザ名のそれぞれの出現位置を素性とする SVM による分類を設定し、5-fold cross validation により結果を評価した。本手法では、どのルールも適用しなかった場合の評価結果、ルール 1、ルール 2、ルール 3 のみをそれぞれ適用した場合の評価結果、三つのルールを適用した場合の評価結果を算出した。表 3 より、SVM を利用した比較手法は、本手法で提案した三つのルールを適用した場合と比較して再現率、適合率ともに低下した。また、本手法は三つのルールを適用することで、再現率はやや低下するが、適合率が大幅に向上し、F 値は約 6 ポイント程度向上した。再現率の低下は主にルール 2 の副作用に起因するが、ルール 2 はそれ以上に適合率の向上に寄与している。ルール 1 とルール 3 は副作用がほとんどなく、それぞれ適合率の向上に寄与している。これより、運行情報陽性ツイート抽出において、本手法で提案した三つのルールが有効であると言える。

5.4 評価結果 (運行情報推定・パラメータ共通)

4.2 で述べた三つのルールを適用して抽出した運行情報陽性ツイートに基づく、運行情報推定の評価結果を表 4 に示す。関連研究を参考に、二つの比較手法を設定した。一つ目は、運行情報陽性ツイートの出現間

隔からバースト区間を検出し、バースト区間を運行情報発生時間と推定する手法である。具体的には、運行情報陽性ツイート集合の状態列を、通常状態とバースト状態の 2 状態からなるオートマトンを用いて表し、全ての運行情報陽性ツイートとその状態に関して必要なコストを計算した後、最小コストとなるような状態列を選択した。コスト計算には、関連研究で紹介した Kleinberg [14] が提案した計算式をそのまま利用した。二つ目は、全ツイート数に対する運行情報陽性ツイート数の割合がしきい値よりも高い時間を運行情報発生時間と推定する手法である。具体的には、本手法と同様、直近 N 分での運行情報陽性ツイート数の割合がしきい値 T を超えると運行情報有り、超えなければ運行情報無しと判断した。比較手法 2 ではウィンドウの長さ N は 10 分、20 分、30 分、しきい値 T は 0.01, 0.05, 0.10, 0.15 の合計 12 通りの組み合わせ、本手法ではウィンドウの長さ N は 10 分、20 分、30 分、しきい値 T は 1, 5, 10 の合計 9 通りの組み合わせで実験を行い、それぞれ平均値を算出した。

表 4 より、バースト区間を運行情報発生時間とする比較手法 1 は、適合率は 99% 以上と非常に高くなったが、再現率は 30% 台と非常に低くなった。これは、運行情報発生直後のツイート数が急激に増える時間をバースト区間として検出できる一方、運行情報発生区

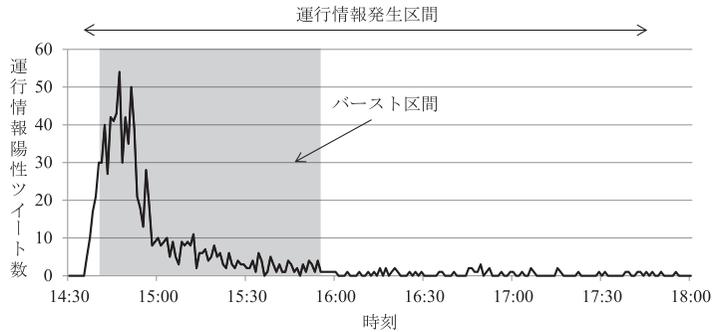


図 5 比較手法 1 における運行情報の検出漏れ例 (JR 山手線, 2011/07/31)

Fig. 5 Example of false negatives by comparative method 1 (JR Yamanote Line, 2011/07/31).

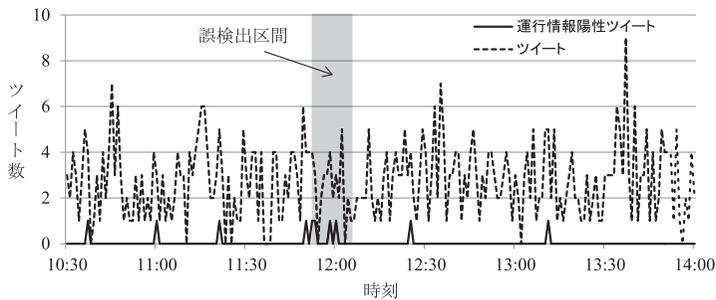


図 6 比較手法 2 における運行情報の誤検出例 (JR 山手線, 2011/07/31)

Fig. 6 Example of false positives by comparative method 2 (JR Yamanote Line, 2011/07/31).

間終盤のツイート数が少なくなる区間をバースト区間として検出できなかったためである (図 5)。更に、中央線の 2 件目の運行情報発生区間を全く検出できなかったため、検出時間の平均値を算出することができなかった。運行情報陽性ツイートの割合から運行情報発生を検出する比較手法 2 は、どのパラメータの組み合わせでも再現率が 90%前後と高くなり、且つ、運行情報の検出に掛かる時間は 1 分以内とリアルタイム性も高くなった。ところが、適合率は 50%前後と低くなった。これは、運行情報陽性ツイートが誤って投稿された場合に、それが 2, 3 件でもあれば運行情報が発生していると誤検出されたためである (図 6)。運行情報陽性ツイートの頻度から運行情報を検出する本手法は、ウィンドウの長さ N が 20 分、しきい値 T が 5 の場合に、再現率、適合率ともに 80%を超えた。リアルタイム性についても、比較手法 2 には及ばないが、検出時間は 5 分程度と短くなった。これより、本手法の有効性を確認できた。なお、しきい値 T が 10 の場合は、どのウィンドウの長さでも東京メトロ副都心線

の運行情報発生区間を全く検出されず、検出時間の平均値を算出できなかった。

5.5 評価結果 (運行情報推定・パラメータ最適化)

表 4 はどの路線でもパラメータを共通にした場合の結果であるが、路線ごとにツイート数は異なるためパラメータの最適値は路線ごとに異なることが予想される。そこで、本手法でウィンドウの長さ N が 20 分、しきい値 T が 5 の場合を基準に、平均ツイート数に応じてパラメータ値を変更した。予備実験からウィンドウの長さよりもしきい値の方が影響は大きいと分かったので、しきい値のみを路線ごとに変更した (ウィンドウの長さは 20 分に固定)。これにより、ツイート数が少ない路線ではしきい値を下げて検出漏れを減らし、ツイート数が多い路線ではしきい値を上げて誤検出を減らし、結果的に再現率と適合率を向上できると考えた。一日当たりの平均ツイート数が 1,000 以上の路線と 1,000 以下の路線に分け、それぞれでしきい値を 1 から 10 の間で変更して最適値を求めた。路線ごとの平均ツイート数、及び、路線ごと、全体平均の評価結

表 5 平均ツイート数でしきい値を変えた場合の評価結果
Table 5 Evaluation results shifting the parameter value (T) based on the average number of tweets.

	平均ツイート数/日	ウィンドウの長さ(N)	閾値(T)	検出時間(分)	再現率(%)	適合率(%)	F 値(%)
JR 山手線	3,761	20	5	0.0	96.8	84.7	90.3
			8	1.0	82.4	96.3	88.8
JR 中央線	2,262		5	1.3	98.1	81.4	89.0
			8	3.3	92.9	94.5	93.7
JR 総武線	1,212		5	4.5	95.9	79.8	87.1
			8	5.0	93.6	89.9	91.7
JR 横須賀線	567		5	2.0	97.6	63.0	76.6
			3	1.0	98.8	61.8	76.1
JR 高崎線	283		5	5.0	90.4	64.8	75.5
			3	4.0	97.5	49.8	65.9
JR 川越線	85		5	21.0	63.9	95.0	76.4
			3	6.0	88.9	88.0	88.4
京王井の頭線	473		5	2.0	97.4	72.8	83.3
			3	2.0	97.4	71.4	82.4
東京メトロ副都心線	438		5	13.0	13.3	10.5	11.8
			3	6.0	60.0	33.3	42.9
東武東上線	618		5	4.5	86.3	88.6	87.4
			3	3.0	97.4	79.4	87.5
東急池上線	128		5	8.0	66.1	98.7	79.2
			3	3.0	88.1	68.9	77.3
全体平均	-	5	5.2	87.6	80.3	83.8	
		3.8	3.5	92.1	79.5	85.3	

果を表 5 に示す。1,000 以上の路線ではしきい値を 8, 1,000 以下の路線ではしきい値を 3 に設定した場合に全体平均の F 値が最も高くなった。適合率は若干低下したが、再現率は 5 ポイント近く向上し、更に、リアルタイム性を 3 分台にまで短縮できた。これより、運行情報推定において、路線ごとの情報量を考慮することに意味があったと言える。再現率が向上したのは、しきい値を下げた路線での検出漏れを大幅に削減でき、しきい値を上げた路線での副作用としての検出漏れの増加が大きくなかったからである。一方、適合率が向上しなかったのは、しきい値を上げた路線での誤検出を大きく減らすことができた一方、しきい値を下げた路線での副作用としての誤検出の増加が同程度に大きかったからである。例えば、東急池上線では、しきい値を下げることで再現率が約 22 ポイント向上しているが、適合率が約 30 ポイント低下した。

6. 公式運行情報との比較

公式運行情報と本手法による運行情報検出結果がどの程度一致するのか調査した。本手法ではウィンドウの長さを 20 分に固定し、平均ツイート数によって路線ごとにしきい値を変更した。

まず、JR 中央線で公式運行情報が発表されたとき

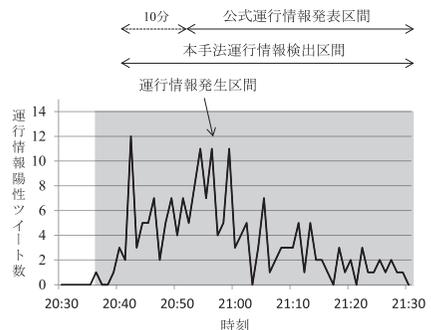


図 7 公式運行情報発表区間、及び、本手法運行情報検出区間 (JR 中央線, 2011/08/03)

Fig.7 Period official train status information has published and period the information detected by our method has occurred (JR Chuo Line, 2011/08/03).

の、公式運行情報発生区間、及び、本手法による運行情報検出区間を図 7 に示す。運行情報陽性ツイート数の分布とツイート内容から人手で判断した運行情報発生区間も併記した (図 7 のグラフ中の網かけ部分)。本手法では、公式運行情報よりも 10 分早く検出できていることが分かる。次に、JR 総武線で公式運行情報が発表されず運行情報が発生していたと推測されるとき、本手法による運行情報検出区間を図 8 に示す。

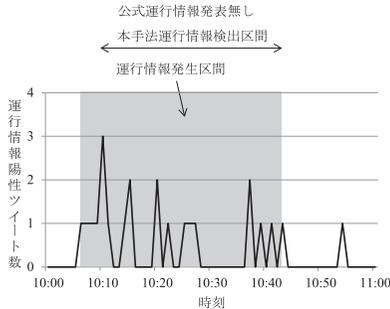


図 8 本手法運行情報検出区間 (JR 総武線, 2011/07/31)
 Fig. 8 Period train status information detected by our method has occurred (JR Sobu Line, 2011/07/31).

表 6 運行情報陽性ツイート (JR 総武線, 2011/07/31)
 Table 6 Positive tweet (JR Sobu Line, 2011/07/31).

ツイート内容	投稿時刻
なんだろう？窓の外みたら、総武線乗り電車が変な所でかれこれ 10 分以上は止まっている…	10:08:44
電車遅れてる…総武線よく遅れるなあ、	10:09:54
総武線、本八幡で転落事故 10 分以上の遅れだっ。	10:10:21
週末の総武線、またトラブルで遅れてる…もうやだ。	10:11:03
総武線各駅 本八幡駅で線路に転落した人の救護にあたった為、ほんのり遅れている。	10:15:58

図 7 と同様に、人手で判断した運行情報発生区間も併記し、運行情報陽性ツイートを表 6 に掲載する。運行情報陽性ツイートの内容から、本八幡駅の転落事故の影響で最大で 10 分以上の遅延があったと推測される。本手法では、5 分程度のタイムラグが見られるが、このような小規模な事象も検出できている。これより、本手法を適用することで、公式運行情報が発表される場合はそれよりも早く、また、鉄道事業者が発表しない小規模な事象も検出できる場合があることを確認した。これらの情報は、鉄道利用者がいち早く運行情報の有無を確認できるという点で有用である。

7. Android アプリの開発

本論文で提案した鉄道運行情報検出機能を組み込んだ Android アプリ「路線実況」を開発し、2012 年 3 月 29 日から同年 5 月 29 日まで株式会社駅探と共同で、一般公開する実証実験を実施した。

7.1 アプリ機能の概要

本アプリは、Twitter 上の大量の書き込みの中から、鉄道の運行情報に関連するものだけを抽出し、時系列 (タイムライン) に表示する。首都圏の 158 の鉄道路線を対象として、利用者が路線一覧から鉄道路線を手動で選ぶほか、加速度センサーの情報による乗車判定、



ダイヤの乱れが発生している路線を検出

その運行状況に遭遇した人のつぶやきだけを抽出

図 9 Android アプリ「路線実況」の機能概要
 Fig. 9 Functional overview of the application “Train status live”.

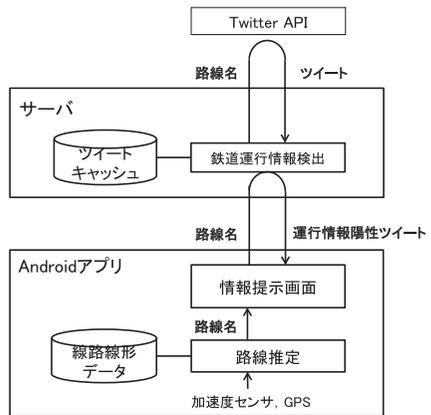


図 10 システム構成
 Fig. 10 System overview.

及び GPS から得た位置情報を用いて推定された乗車中・近隣路線からも選択できる。

アプリ画面を図 9 に示す。路線一覧では、運行情報が検出された路線名の右側に「！」マークが表示される。路線名を一つ選択すると、その路線に関する最新の運行情報陽性ツイートを閲覧できる。

7.2 システム上の工夫点

本アプリは、鉄道運行情報検出機能を組み込んだサーバと、クライアントとしての Android アプリから構成される。システム構成を図 10 に示す。サーバは、本アプリで対象とするそれぞれの路線名を検索キーワードとして、Twitter API から定期的にツイートを取得し、運行情報陽性ツイートを抽出し、DB にキャッシュしておく。一方、クライアントは、路線推定機能により推定した路線名をサーバへ送り、DB から当該路線に関する運行情報陽性ツイートを取得して画面に表示する。

路線推定機能は、GPS 及び加速度センサを利用し

てユーザの移動状況（静止，乗車，歩行，走行の4状態）を判定し，乗車中若しくは近隣の鉄道路線を推定する[17]．路線の推定には，線路線形データを利用している．線路線形データは，線路に関する地理空間データで，線路を複数の線形データの集まりで表現したものである．国土交通省が提供の「国土数値情報（平成20年度作成・鉄道（線）データ）」^(注3)をもとに開発した．乗換駅周辺や，並行路線など，複数路線が存在して特定が難しい地域の線形データを残し，その他の地域の線形データを間引くことで，アプリ本体と線形データとを合わせた使用メモリ量を，19MBから3MBまで軽量化した^(注4)．

7.3 実証実験結果

実証実験期間中に計1万件以上のダウンロードがあった．運行情報検出機能の有用性を評価する目的で，アプリのメニュー内にアンケートのページを設置し，アプリの一般利用者に対して二つの質問を提示し，それぞれ四つの選択肢の中から一つを回答してもらう主観評価を実施した．実証実験期間中，利用者は任意のタイミングで回答が可能で，有効回答者数は17人であった．

アンケート回答の集計結果を表7に示す．運行情報の有用性に関する質問1では，全員が「役に立った」と回答した．また，アプリの運行情報と実際の運行情

報の一致度合いに関する質問2では，1人を除く16人「一致」と回答した．自由記述には，「現地の詳しい状況を早く入手できる」「悪天候時に役立つ」など，利便性の理由について回答があった．また，「乗り換えて次に利用する路線の情報も知りたい」「運行状況の程度変化が分かると，待つべきか代替手段を利用するべきかの判断ができる」といった機能改善に対する要望も見られた．ダウンロード総数に対して回答数が少なく，十分に客観的な評価とは言い難いが，鉄道運行情報検出機能に対して一定の有用性を確認できたと言える．

8. 考 察

ソーシャルセンサとしてのTwitterからの鉄道運行情報検出において，検出結果のF値が0.85，また，検出までに掛かる時間が3分台という高い正確性とリアルタイム性を実現できた．また，鉄道事業者が発表する公式運行情報よりも早く，鉄道事業者が発表しない小規模な事象も検出できる場合があることを確認した．更に，提案した鉄道運行情報検出機能を組み込んだAndroidアプリを開発し，利用者にとって有用な情報を提供できたことを確認した．これより，本手法で提案した鉄道運行情報検出が一定の精度を達成し，ソーシャルセンサの潜在的可能性を確認したと言える．

これらの結果をもとに，ソーシャルセンサとしての可能性について三つの観点から考察する．

精度向上の可能性

鉄道運行情報検出の更なる精度向上には，運行情報陽性ツイート抽出，及び運行情報推定それぞれの洗練化が不可欠である．前者については，例えば，運行情報キーワードの網羅性向上が挙げられる．運行情報に遭遇した際の言い回しは本論文で設定した運行情報キーワード以外にも，「死亡」「終わった」など数多く存在する．ただし，これらは鉄道運行情報が発生した以外の場面でも多く使われるため，ツイート中に出現する他のキーワードも考慮する必要がある．

一方，後者については，例えば，事故の規模による情報量の違いを考慮することが挙げられる．本論文では路線による情報量の違いのみ着目したが，本来，事故の規模によっても情報量は異なる．数十分程度の遅れの場合と，終日運転を見合わせた場合など，事故の影響に応じた解析処理が必要となる．

運行情報検出の汎用性

本論文で対象としなかった鉄道運行情報の検出の可能性について考察する．運転再開については，「運転

表7 アンケート回答
Table 7 Questionnaire answers.

質問1. アプリの運行情報は役に立ちましたか？

選択肢	回答数
かなり役に立った	11
まあまあ役に立った	6
あまり役に立たなかった	0
役に立たなかった	0

質問2. アプリの運行情報と実際の運行情報はどの程度一致していましたか？

選択肢	回答数
かなり一致	8
まあまあ一致	8
あまり一致しない	1
一致しない	0

(注3) : <http://nlftp.mlit.go.jp/ksj/jpgis/datalist/KsjTmplt-N02-v1.1.html>

(注4) : 一般に，Android 端末の利用者にとって許容できる容量と言われている．

再開「動き出した」などを利用者が投稿するため、これらのキーワードの出現を手掛かりにすれば、本手法と同じ仕組みで検出できる。一方、運転見合わせ区間(駅名など)の詳細情報は、表現のばらつきが見られるため、駅名の出現位置を見るなどの一工夫が必要となる。また、Android 端末に搭載の GPS から取得した位置情報など実世界情報と併用すれば、区間検出の精度を向上できるであろう。このように、鉄道に関してはソーシャルセンサからの情報検出のポテンシャルは非常に高いといえる。

他の交通インフラへの適用可能性

鉄道以外の交通インフラの例として、道路の渋滞情報の検出について考察する。Twitter 以外のソーシャルメディアを含めて、道路の情報に関するユーザ投稿は非常に少ない。そのため、現時点ではソーシャルセンサとしてのポテンシャルは高いとはいえない。運転手による音声情報を活用したり、同乗者に聞いてもらうサービスの開発、自動車に搭載した各種物理センサの情報の活用、道路に設置したカメラなどの物理センサから取得できる情報の利用が不可欠であろう。すなわち、ソーシャルセンサと物理センサの情報を相互に補完する技術の開発が望まれる。

9. む す び

鉄道の運転見合わせ、遅延を対象に、ソーシャルセンサとしての Twitter からの交通情報検出の潜在的可能性を確認した。今後は、本論文で対象としなかった鉄道運行情報の検出を進めるとともに、鉄道以外の交通情報検出の可能性を検討していく。

なお、その他本論文に掲載の商品、機能等の名称は、それぞれ各社が商標として使用している場合がある。

文 献

- [1] X. Li, R. Lu, X. Liang, X. Shen, J. Chen, and X. Lin, "Smart Community: An Internet of things application," *IEEE Communications Magazine*, vol.49, no.11, pp.68–75, 2011.
- [2] 篠原哲哉, 松井洋子, "震災復興におけるコミュニティ作りへの取組み," *東芝レビュー*, vol.66, no.8, 2011.
- [3] S.H. An, B.H. Lee, and D.R. Shin, "A survey of intelligent transportation systems," *Proc. 3rd International Conference on Computational Intelligence, Communication Systems and Networks*, pp.332–337, 2011.
- [4] 本田健司, "次世代交通情報システムを担うプローブ技術," *IT ソリューションフロンティア*, vol.25, no.3, pp.14–15, 2008.
- [5] 野村総合研究所, *ビッグデータ革命*, pp.77–125, アスキー・メディアワークス, 2012.
- [6] A. Sheth, "Citizen sensing, social signals, and enriching human experience," *IEEE Internet Computing Magazine*, vol.13, no.4, pp.80–85, 2009.
- [7] M. Nagarajan, K. Gomadam, A.P. Sheth, A. Ranabahu, R. Mutharaju, and A. Jadhav, "Spatio-temporal-thematic analysis of citizen sensor data: Challenges and experiences," *Proc. 10th International Conference on Web Information Systems Engineering*, pp.539–553, 2009.
- [8] 榑 剛史, 松尾 豊, "ソーシャルセンサとしての Twitter," *人工知能学会誌*, vol.27, no.1, pp.67–74, 2012.
- [9] ネットレイティングス株式会社, 2011 年 10 月の日本の主要 SNS サイトの動向, http://www.netratings.co.jp/news_release/2011/11/sns-report-Oct-2011.html (2012/06/30 アクセス)
- [10] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes Twitter users: Real-time event detection by social sensors," *Proc. of the 19th International Conference on World Wide Web*, pp.851–860, 2010.
- [11] E. Aramaki, S. Maskawa, and M. Morita, "Twitter catches the flu: detecting influenza epidemics using Twitter," *Proc. 2011 International Conference on Empirical Methods in Natural Language Processing*, pp.1568–1576, 2011.
- [12] 高橋哲朗, 野田雄也, "実世界のセンサーとしての Twitter の可能性," *信学技報*, NLC2010-38, 2011.
- [13] R. Swan and D. Jensen, "TimeMines: Constructing timelines with statistical models of word usage," *Proc. 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.73–80, 2000.
- [14] J. Kleinberg, "Bursty and hierarchical structure in streams," *Proc. 8th ACM SIGKDD International Conference and Knowledge Discovery and Data Mining*, pp.91–101, 2002.
- [15] 藤木稔明, 南野朋之, 鈴木泰裕, 奥村 学, "document stream における burst の発見," *情報処理学会研究報告(自然言語処理研究会)*, 2004-NL-160, pp.85–92, 2004.
- [16] 株式会社 NTT データ経営研究所, *社会インフラにおける停止許容時間についての調査*, <http://www.keikeiken.co.jp/aboutus/newsrelease/090907/index.html> (2012/06/30 アクセス)
- [17] 岡本雄三, 長 健太, 池谷直紀, 瀬戸口久雄, 服部正典, "ユーザ状況推定技術を応用した乗換案内アプリケーションの実装と評価," *FIT2010*, 2010.

(平成 25 年 1 月 11 日受付, 5 月 15 日再受付)



長野 伸一 (正員)

1999年大阪大学院基礎工学研究科博士後期課程了。同年、(株)東芝入社。現在、同社研究開発センター知識メディアラボラトリー主任研究員、博士(工学)。2004-2006年国立情報学研究所特任講師。2010-2012年法政大学兼任講師。LOD チャレンジ Japan

2011, 2013 実行副委員長。主に Linked Data, セマンティック Web の研究・開発に従事。情報処理学会, 人工知能学会, 各会員。



上野 晃嗣

2000年東京工業大学大学院総合理工学研究科知能システム科学専攻博士前期課程了。同年(株)東芝入社。現在、同社研究開発センター知識メディアラボラトリー所属。主に音声認識技術応用アプリケーションの研究開発に従事。



長 健太 (正員)

1997年早稲田大学大学院理工学研究科修士課程了。同年(株)東芝入社。エージェント技術, コンテキストウェア技術の研究開発に従事。博士(工学)。情報処理学会, 電子情報通信学会, 日本ソフトウェア科学会各会員。