

マイクロブログおよび携帯電話人口統計を用いた 大規模イベント会場における人口変化の長期予測

塚田涼太郎[†] 詹 浩森[†] 石渡祥之佑^{††} 豊田 正史^{††}

[†] 東京大学大学院情報理工学系研究科 〒113-8656 東京都文京区本郷 7-3-1

^{††} 東京大学生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1

E-mail: [†], ^{††}{tsukada, hos, ishiwatari, toyoda}@tkl.iis.u-tokyo.ac.jp

あらまし 都市における混雑予測は、モビリティの改善や群集事故の防止等の観点から重要である。しかしながら、東京ドームのように、不定期に野球の試合やコンサート等の多様なイベントが開催されるイベント会場については、来場者数が日々変化するため混雑予測が難しい。本稿では未来の日付とイベント会場名の両方に言及しているマイクロブログ上の投稿と、イベント会場の過去の混雑度からなる異種データを手がかりとして、イベント会場の未来の混雑度およびイベントの有無を長期にわたって予測する手法を提案する。実験には研究室で収集している Twitter データと、携帯電話の GPS ログデータに個人の非特定化処理を加えた形で提供される混雑統計データを用い、東京・神奈川の主要 9 イベント会場の 24 時間先までの混雑度を予測するタスクについて、いずれの会場においても提案手法がベースライン手法を上回る予測精度を達成することを確認した。

キーワード 混雑予測, マイクロブログ

1 はじめに

野球の試合やコンサートなどの大規模イベントの開催時には、参加者が会場付近に集中することで混雑が突発的に発生する。そのような混雑は、イベント参加者自身だけでなく、周囲を巻き込んで様々な悪影響を及ぼす。例えば、通常なら空いているはずの時間帯の電車が、途中駅で乗り込んできた大量のイベント帰りの客のため満員になると、混雑した車内で乗客は身体的・精神的苦痛を強いられることになるだけでなく、乗降時間の増大に伴う遅延により経済的損失まで被る事態となる。また、イベントによる混雑が起きることを知らずに観光地を訪れた観光客にとっては、周遊性が低下することになり、旅行の満足度が下がってしまう。さらに、2014 年の年末に上海で開催された年越しカウントダウンイベントにおいて、集まりすぎた群衆が制御不能になり、雑踏の中で約 35 名もの死者を出す大惨事となる [1] など、混雑は時として人命に関わる深刻な事故を引き起こすこともある。

このような混雑に伴う諸問題への一つの解決策として、混雑を予測することが考えられている。もし混雑の時間・場所・程度が事前に予測できれば、鉄道の乗客は混雑に巻き込まれないように時間帯をずらしたり他の路線へ迂回したりして乗車できるほか、遅延の程度を考慮して早めに出発するなどの対策をとれる。観光客は混雑する時間帯と場所を避けるようにして旅行計画を立てることで、時間の有効活用ができる。乗換案内アプリケーション内の「異常混雑予報」機能のように既に実用化されているものもあり、実際に鉄道の乗客が混雑を避けるために利用している [2]。こうした背景のもと、GPS を搭載した端末から収集されたデータを利用して、都市全体の 1~2 時間先

までの混雑を予測する研究が行われている [3, 4]。しかしながら、これらの研究で提案されている手法は、1 日~数日先までの比較的長期の混雑予測には適用されていない。予測期間が長期になるほど、混雑の原因となる大規模イベントなどの外的要因 [5] による非定常的な人の流れの影響が無視できなくなるため、これらの手法を混雑の長期予測に適用することは困難である。実用上、混雑が予測できてからその対策を実施するまでの時間の確保という観点から、混雑の長期予測の実現は重要な課題である。

他方、Twitter に代表されるマイクロブログ上の投稿を活用し、実世界の様々な現象を把握する試みがなされており、地震の検出 [6] などの分野で有効性が示されている。投稿の中には未来のイベントに言及しているものも含まれており、それらの投稿から、いつ・どこで・どのようなイベントが開催予定であるかを知ることができる。例えば「誰か、明日の東京ドーム、巨人対横浜に一緒に行きませんか」という内容の投稿が 2018 年 7 月 2 日に作成されていた場合、2018 年 7 月 3 日に東京ドームにおいて野球の試合が予定されていることが読み取れる。

本稿では、イベントの影響を強く受けるイベント会場周辺の混雑の長期予測が難しいことを確認し（第 3 節）、その問題を解決するために、上述のように未来のイベントに言及しているマイクロブログ上の投稿と、イベント会場の過去の混雑度を組み合わせて予測器を学習させる手法を提案する（第 4 節）。さらに、実世界データを用いた実験で提案手法の有効性を検証した（第 5 節）。

2 関連研究

2.1 位置情報を含むデータを用いた混雑予測

GPS を搭載した携帯電話や自動車から収集された位置情報データを活用して、混雑予測を行う研究がされている。

Fan ら [3] は、従来の混雑予測に関する研究が、大規模イベントの開催に伴う例外的な混雑を外れ値として無視してきた点を指摘し、大規模イベント開催期間における都市全体の人々の動きのモデル化に取り組んだ。Fan らの手法は、大規模イベントの開催時には会場に早く到着する人と遅れて到着する人がいると仮定し、直近の人々の移動経路が含まれる位置情報データからオンラインでマルコフ連鎖のモデルを作成するもので、イベントの影響下においても従来手法より高い精度での混雑予測が可能であることを示した。

将来の人々の移動行動を把握するためのより直接的なアプローチとして、乗換案内や地図のアプリケーションに入力される検索クエリのログデータを混雑予測に活用する取り組みも行われている。Konishi ら [7] は、乗換案内アプリケーションの検索クエリログを活用し、花火大会などのイベントに伴う都市全体規模の混雑を予測した。Liao ら [4] は、地図アプリケーションにおけるイベント会場を目的地とした検索クエリの件数が、その会場でのイベント開始時刻の数時間前から増加するという現象に着目して、イベント会場周辺の人口変化の予測を行った。

2.2 マイクロブログ上の投稿からのイベント情報の抽出

マイクロブログ上の投稿を実世界で起きているイベントを観測するための手段として利用する研究が行われている。

Sakaki ら [6] は、地震に言及している Twitter 上の投稿から、地震の発生をリアルタイムで検知して速報するシステムを提案した。Yamada ら [8] は、旅行のプランニングで利用するために、Twitter 上の投稿からローカルなイベント情報を抽出する研究を行った。Yamada らの手法は、イベント会場名を含む投稿からイベントの開催期間を取得し、会場名の表記揺れを考慮して集約するもので、従来手法より正確なイベント情報抽出ができることを示した。Jatowt ら [9] は、Twitter 上の投稿において時間に言及している表現に着目することで、現在起きているイベントだけではなく、未来や過去のイベントを可視化して分析するフレームワークを提案した。

2.3 マイクロブログ上の投稿を用いた混雑度の推定と予測

マイクロブログ上の投稿を利用し、実世界の混雑度を推定・予測する研究が行われている。

Onishi ら [10] は、現実空間の混雑度と、仮想空間であるマイクロブログ上の投稿数が相互作用していると仮定して分析を行った。具体的には、秋葉原の商業施設においてカメラの映像から取得された混雑度と、「秋葉原」という単語が含まれている Twitter 上の投稿数が互いに影響を与えているとするモデルを作成し、モデルが学習したパラメータの大小から、混雑度が投稿数に与える影響と、投稿数が混雑度に与える影響の説明を試みた。ただし、Onishi らのモデルは現在の混雑度を現在の投稿

数および過去の投稿数・混雑度を用いて推定するものであるため、予測には適用されていない。

He ら [11] は、Twitter 上の投稿に交通状況への言及があることに着目し、投稿に含まれる単語の情報をを用いて自動車の交通量を比較的長期にわたって予測する手法を提案した。

He らの手法は、投稿に付与されている位置情報を用い、交通量の予測対象となる区画内で発信された投稿のみをフィルターして利用している。しかしながら、2019 年に Twitter は投稿に正確な位置情報を付与する機能を廃止¹したため、この手法は現在適用不可能である。本論文では、イベント会場の名前に言及している投稿を収集し利用することで、この問題を解決することを試みた。

3 予備実験

本節では、不定期に野球の試合やコンサート等の多様なイベントが開催される大規模イベント会場における人口変化の長期予測が、それ以外の場所と比較して難しいことを確認する。

3.1 データセット

実験には混雑統計データ²を使用した。混雑統計データには、約 250m 四方の「メッシュ」と呼ばれる矩形領域ごとに、そのメッシュ内に滞在している人数の推定値を 1 時間単位で集計した値（混雑度）が含まれる。このうち 2015 年 9 月から 2018 年 11 月までの 3 年 3 ヶ月分のデータを用いて予備実験を行い、最初の 3 年分（2015 年 9 月～2018 年 8 月）を訓練データ、残りの 3 ヶ月分（2018 年 9 月～2018 年 11 月）をテストデータとした。対象とするメッシュは、東京都内の 2018 年 8 月における月間累計混雑度において上位 1,500 位までのメッシュとした。これらのメッシュには野球場や展示場などの大規模イベント会場のほか、主要駅や幹線道路が含まれている。

3.2 予測手法

予測には時刻 t の予測値 $\hat{X}_t$ を過去 1 週間の実測値 $\{X_{t-24 \times 7}, X_{t-24 \times 7+1}, \dots, X_{t-1}\}$ で線形回帰する自己回帰 (AR) モデルを各メッシュごとに用いた。

3.3 評価指標

異なるメッシュ間での予測結果を比較可能にするために、予測結果の評価には式 (1) で定義される加重絶対パーセント誤差 (WAPE) を用いた。

$$WAPE = \frac{1}{N} \sum_t \left| \frac{\hat{X}_t - X_t}{\bar{X}} \right| \quad (1)$$

ここで、 $X_t, \hat{X}_t$ はそれぞれ時刻 t における実測値、予測値であり、 N はデータの総数、 $\bar{X}$ は実測値の平均である。

1: <https://twitter.com/TwitterSupport/status/1141039841993355264>

2: 「混雑統計@」データは、NTT ドコモが提供するアプリケーション (※) の利用者より、許諾を得た上で送信される携帯電話の位置情報を、NTT ドコモが総体的かつ統計的に加工を行ったデータ。位置情報は最短 5 分毎に測位される GPS データ (緯度経度情報) であり、個人を特定する情報は含まれない。※ドコモ地図ナビサービス (地図アプリ・ご当地ガイド) 等の一部のアプリ。

3.4 予測結果と考察

予測結果はヒートマップで可視化した(図1)。このヒートマップでは、WAPEの値が大きい、すなわち予測が難しいメッシュが赤に近い色で、WAPEの値が小さい、すなわち予測が簡単なメッシュが青に近い色で表示されている。

WAPEの値が小さいメッシュは主に東京駅や新宿駅などの主要駅周辺に分布している。一方、WAPEの値が特に大きいメッシュには東京ドームや神宮球場といった大規模イベント会場が含まれている。このように、場所ごとに予測誤差の差異が生じる理由を明らかにするため、新宿駅、東京ドームおよび神宮球場の3地点に着目し、それぞれの混雑度の予測値および実測値を比較する。上記3地点それぞれを含むメッシュにおける、2018年9月16日からの1週間の混雑度の予測値および実測値の時系列データを図2に示す。図2から、新宿駅では主に通勤・通学の時間帯にピークを迎える同じパターンの混雑が毎日繰り返り起きているため、ARモデルもそのパターンに追従して正確な予測ができていることが分かる。東京ドームや神宮球場では、大規模イベントに起因する非定常的な人口変化が断続的に発生しているため、予測に失敗していることが分かる。実際に、東京ドームでは16日、17日、19日、22日にプロ野球の巨人戦が開催されており、急激な人口変化をもたらして予測を妨げる原因になっていると考えられる。

4 提案手法

第3節で確認したように、都市においては特にイベント会場の混雑予測が難しいという問題がある。すなわち、過去の混雑度情報のみを用いて、未来に開催されるスポーツの試合やコンサートといった非定常的な大規模イベントに起因する人口変化を推定することは困難である。この問題を解決するために、イベントに言及しているマイクロブログ上の投稿およびイベント会場の過去の混雑度を組み合わせた異種データを用いて、イベント会場の未来の混雑度およびイベントの有無を長期にわたって予測する手法を提案する。

4.1 イベントに言及しているマイクロブログ上の投稿

大規模イベント会場で将来発生する人口変化を予測するための手がかりとして、未来の日付とイベント会場名の両方に言及しているマイクロブログ上の投稿を用いる。イベント会場 v のある日 d の混雑度を予測するために用いる投稿は、以下の条件をすべて満たす投稿として定式化される。

- d に言及する日付の表現を含む
- イベント会場名 v を含む
- d の前日以前に投稿されている

予測には、この条件を満たす投稿から、重複を取り除いたものをすべて連結した文書のBag-of-Wordsを特徴として用いる。

4.2 イベント会場の過去の混雑度との融合

Bag-of-Wordsのベクトルと、 d の1日前の日から n 日前の日までにおける当該会場の1時間ごとの混雑度を表現する、 $24n$ 次元のベクトルを連結したベクトルを後述の回帰の説明変数と

した。ここで、 n は過去の混雑度を何日前まで遡って用いるかを表すハイパーパラメータである。

4.3 各時間についての回帰による予測

予測は上述のベクトルを説明変数、混雑度を目的変数として、24時間の各時間について別々に回帰することで行う。今回は、回帰にGradient Boosting Regressor [12]を用いる。Gradient Boosting Regressorとは、アンサンブル学習の一種である勾配ブースティングに基づく回帰モデルである。Gradient Boosting Regressorでは、弱学習器として決定木を用い、前の弱学習器の予測誤差を打ち消すような出力値を学習する次の弱学習器を直列に接続して学習が進行する。決定木を用いているため、説明変数の分布を仮定しないことが特長である。

5 評価実験

第4節で提案した手法の有効性を実世界の混雑データを用いて評価した。

5.1 実験設定

実験には第3節と同じ混雑統計データのうち、2014年12月から2018年11月までの4年分のデータを用い、最初の3年分(2014年12月~2017年11月)を訓練データ、残りの1年分(2017年12月~2018年11月)をテストデータとした。予測の対象とする大規模イベント会場は、東京都と神奈川県における収容人数(キャパシティ)が上位に入る9会場を選んだ。なお、上位に入っている営業休止期間のある会場(国立競技場、代々木競技場)および投稿数が極端に少ない会場(茅ヶ崎公園野球場)については、対象外とした。

マイクロブログ上の投稿データは、研究室で収集しているTwitterデータに対して、各会場ごとに会場名による検索を実行することで作成した。会場名による検索では、正式名称と通称や旧称を紐付けの名寄せ辞書を作成して表記揺れに対応した。検索して得られた投稿は、MeCab [13]で分かち書き処理を行い、頻度上位5,000語までのBag-of-Wordsを生成した。

Gradient Boosting Regressorのパラメータとして、ブースティングのステージ数を100、個々の決定木の深さを5、損失関数をHuber損失とした。

ベースラインとしては、第3節の予備実験と同じくARモデルを用いた。

5.2 評価指標

評価は2つの指標を用いて行った。初めに、イベントの有無が予測可能であることが実用上最も必要と思われることから、イベントの有無を分類する2値分類問題として解いた際の性能評価を行った(第5.3節)。続いて、より詳細に1日の人口変化が分かるかどうかという観点で、式(2)で定義される平均絶対誤差(MAE)を用いた評価を行った(第5.4節以降)。

$$MAE = \frac{1}{N} \sum_t |\hat{X}_t - X_t| \quad (2)$$

ここで、 X_t , $\hat{X}_t$ はそれぞれ時刻 t における実測値、予測値であ

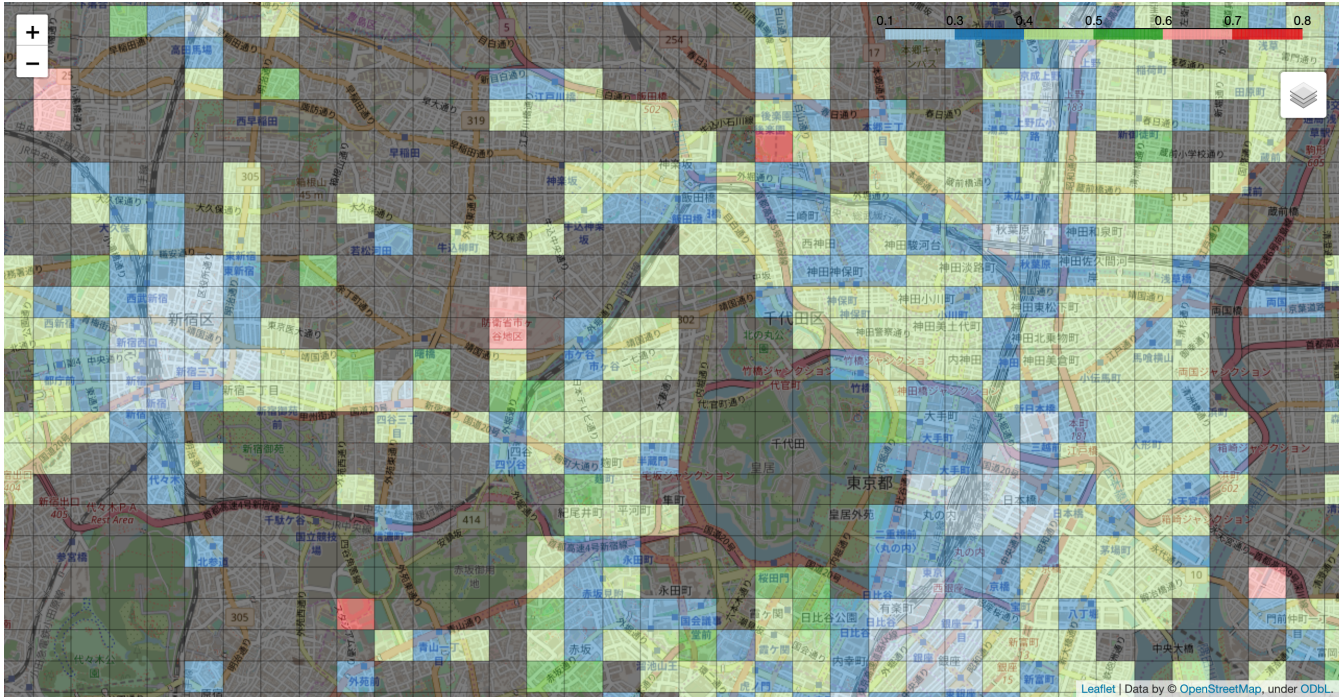
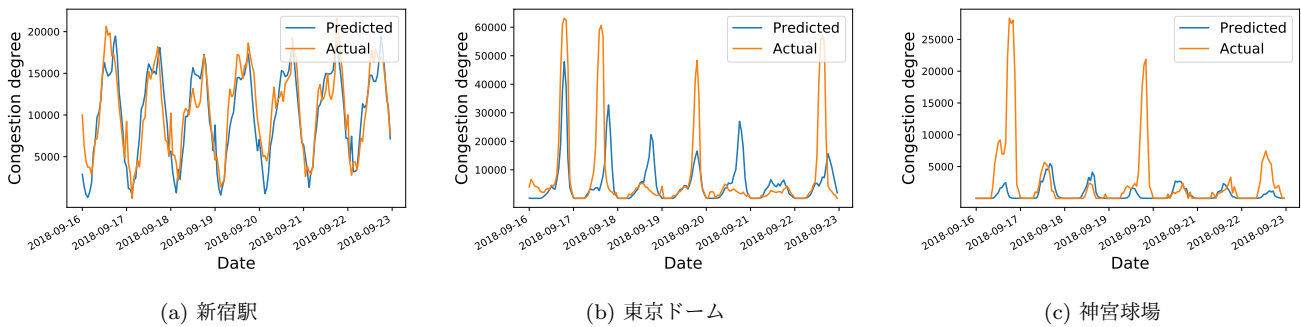


図 1: 都内の混雑度上位 1,500 地点に対する混雑度の予測誤差。
「混雑統計®」©ZENRIN DataCom CO., LTD.



(a) 新宿駅 (b) 東京ドーム (c) 神宮球場

図 2: 都内の主要地点における 2018 年 9 月 16 日からの 1 週間の混雑度の予測値および実測値。
「混雑統計®」©ZENRIN DataCom CO., LTD.

り, N はデータの総数である。

5.3 イベントの有無の予測

各イベント会場ごとに、訓練データにおける 1 日の混雑度の最大値の中央値をイベントが開催されていない平常時の人口のピーク値とみなし、テストデータにおいて 1 日の混雑度の最大値がそのピーク値よりも 5,000 人以上増加している日を正例（イベントがある日）、そうでない日を負例（イベントがない日）と定義し、2 値分類問題として解いた際の性能を評価した。ここで、過去の混雑度は 7 日前まで遡って用いた ($n=7$)。イベントの有無の予測結果を表 1 に示す。ベースラインの AR モデルはほとんどの会場でイベントがある日の予測に失敗しているのに対して、提案手法は比較的高精度な予測ができることが示された。AR モデルは、いずれの会場においても Recall の値がほとんど 0 になっており、イベントの有無判定の閾値を超える混雑度を予測値として出力できていない。このことは、イベントがない日が多数を占める不均衡なデータの中で、前日のイ

ベント会場の混雑度だけを入力として与えられる AR モデルでは、当該会場における非定常的な大規模イベントに起因する人口変化を捉えられないためだと考えられる。提案手法は、ある程度の割合でイベントがある日を判別できていることから、マイクロブログ上の投稿に含まれている情報を活用できていると思われる。

会場の種別ごとに Precision, Recall を見ると、東京ドーム、明治神宮野球場、横浜スタジアムなどの野球場では比較的高い数値となっているのに対して、横浜アリーナや日本武道館といった主にコンサートが開催される会場においては比較的低い数値となっていることが分かる。原因として、未来の野球の試合に関する投稿はチーム数が限定されているため、いつも同じような特徴を持っており、学習が容易であるのに対して、コンサートの出演者は様々であり学習が困難であることが考えられる。

表 1: イベントの有無の予測結果.

会場名	AR				提案手法			
	Precision	Recall	Accuracy	F1	Precision	Recall	Accuracy	F1
日産スタジアム	0.00	0.00	0.97	0.00	0.71	0.50	0.98	0.59
東京ドーム	0.69	0.09	0.66	0.15	0.76	0.78	0.83	0.77
味の素スタジアム	0.00	0.00	0.92	0.00	0.88	0.48	0.95	0.63
明治神宮野球場	1.00	0.01	0.78	0.02	0.75	0.77	0.89	0.76
横浜スタジアム	0.00	0.00	0.79	0.00	0.92	0.83	0.95	0.88
秩父宮ラグビー場	0.00	0.00	0.94	0.00	0.64	0.39	0.96	0.48
パシフィコ横浜	0.33	0.03	0.90	0.05	0.72	0.64	0.94	0.68
横浜アリーナ	0.00	0.00	0.90	0.00	0.64	0.24	0.91	0.35
日本武道館	0.00	0.00	0.88	0.00	0.50	0.17	0.88	0.25

「混雑統計®」©ZENRIN DataCom CO., LTD.

5.4 人口変化の予測

人口変化の予測結果を表 2, 表 3 に示す. 表 2 は 1 日前までの混雑度を用いる設定 ($n = 1$) であり, 表 3 は 7 日前までの混雑度を用いる設定 ($n = 7$) である. ベースライン手法は, イベント会場を含む単一メッシュの混雑度を利用する場合と, イベント会場を含むメッシュの混雑度だけでなく, その周辺の 8 近傍³のメッシュの混雑度も説明変数として入力に加えた場合の 2 手法を用意した. 表において, 単一メッシュを考慮して予測を行うモデルはベースラインとなる「AR」, 提案手法である「中心のみ」であり, 近傍メッシュを考慮して予測を行うモデルはベースラインとなる「VAR」である. VAR モデルは, AR モデルを複数の説明変数を用いるように拡張したモデルである. 実験の結果, いずれの会場においても, 提案手法はベースラインの AR, VAR モデルより小さい予測誤差を達成した.

表 2, 表 3 の比較から, 会場によって差異はあるものの, 7 日前までの混雑度を使用するモデル ($n = 7$) は 1 日前までの混雑度を使用するモデル ($n = 1$) よりも予測誤差が概ね小さい傾向にあることが分かった. 使用する過去の混雑度の期間を広げると予測が改善される理由を明らかにするために, 実際の予測事例を見て分析を行った. 図 3 は, 明治神宮野球場における 2018 年 5 月 27 日から 2018 年 6 月 3 日までの混雑度の予測値および実測値をプロットしたグラフである. 予測値は, 1 日前までの混雑度を使用するモデルによるものと, 7 日前までの混雑度を使用するモデルによるものの 2 種類で比較を行った. なお, いずれのモデルも, 中心のメッシュの混雑度のみを使用するモデルである. 図 3 の 6 月 3 日では, 実際に野球の試合が行われており, 実測値では最大 15,000 人近くの混雑度となっているが, 1 日前までの混雑度を使用するモデルは, イベントが開催されていない日と同程度の予測値しか出力できていない. 一方, 7 日前までの混雑度を使用するモデルは, 実測値より小さいものの, ある程度近い予測値を出力できている. この違いは, 野球の試合の連戦における 1 週間単位での周期性が重要な手がかりになっていることが一因であると考えられる.

以降の実験は $n = 7$ として行った.

続いて, 第 5.3 節において「イベントがある日」として定義

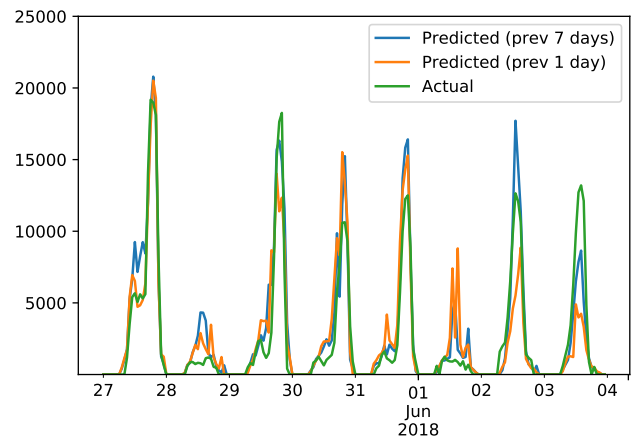


図 3: 明治神宮野球場における 2018 年 5 月 27 日から 2018 年 6 月 3 日までの混雑度の予測値および実測値

「混雑統計®」©ZENRIN DataCom CO., LTD.

した日のみに絞って同様に評価した結果を表 4 に示す. 日産スタジアムのように, 表 3 で特に MAE が小さい会場は, 予測誤差を生じる原因となるイベントの絶対数が少ないと考えられる. 表 4 では, そのような会場においても, ベースラインと提案手法との差が大きくなっていることから, イベントが少ない会場においても, 提案手法は比較的正確な予測ができることが分かった.

また, 「イベントがある日」以外の日を「イベントがない日」として定義して, その日のみに絞って同様に評価した結果を表 5 に示す. 会場によっては, 提案手法の予測誤差がベースラインよりも大きくなっている場合があるが, 収容人数に対して最大でも約 6%程度の誤差 (パシフィコ横浜) に抑えられている.

5.5 予測期間と予測精度との関係

第 5.4 節より, 提案手法は先行研究 [3, 4] で行われている数時間先までの対象とした混雑予測と比較して, より長期である 24 時間先までの混雑予測が可能であることが分かった. 提案手法がさらに長期にわたる混雑予測にも適用できるかを検証するために, 予測期間を変化させて実験を行った. 言い換えると, 予測時点で利用可能な情報を使って, 1 日後, 2 日後, 3 日後の人口変化の予測を行う実験である. 予測期間と予測精度との

3: 対象メッシュを中心とする 3×3 の範囲に含まれるメッシュ.

表 2: 人口変化の予測結果 (1 日前までの混雑度を使用).

会場名	収容人数	ベースライン		提案手法
		AR	VAR	中心のみ
日産スタジアム	72,327	300.8	276.4	255.3
東京ドーム	55,000	4,504.9	4,335.6	2,400.0
味の素スタジアム	50,000	435.0	434.8	336.7
明治神宮野球場	35,133	1,317.7	1,332.6	1,015.9
横浜スタジアム	30,000	1,428.9	1,327.7	1,025.4
秩父宮ラグビー場	24,871	856.3	865.0	631.3
パシフィコ横浜	18,000	1,735.4	1,593.6	1,237.8
横浜アリーナ	17,000	1,362.0	1,331.5	1,086.4
日本武道館	14,471	1,127.2	1,059.6	950.0

「混雑統計®」©ZENRIN DataCom CO., LTD.

表 3: 人口変化の予測結果 (7 日前までの混雑度を使用).

会場名	収容人数	ベースライン		提案手法
		AR	VAR	中心のみ
日産スタジアム	72,327	272.4	271.8	244.4
東京ドーム	55,000	4,291.6	4,331.5	2,409.5
味の素スタジアム	50,000	429.5	428.4	339.8
明治神宮野球場	35,133	1,302.5	1,375.5	987.2
横浜スタジアム	30,000	1,338.6	1,389.6	985.0
秩父宮ラグビー場	24,871	759.6	898.0	552.7
パシフィコ横浜	18,000	1,488.9	1,571.2	1,158.4
横浜アリーナ	17,000	1,216.5	1,446.3	1,091.4
日本武道館	14,471	1,053.6	1,045.7	927.2

「混雑統計®」©ZENRIN DataCom CO., LTD.

表 4: 人口変化の予測結果 (イベントがある日のみ).

会場名	収容人数	ベースライン		提案手法
		AR	VAR	中心のみ
日産スタジアム	72,327	2,439.7	2,439.4	1,152.1
東京ドーム	55,000	6,790.2	6,823.0	3,457.8
味の素スタジアム	50,000	2,697.6	2,696.1	1,777.5
明治神宮野球場	35,133	3,423.5	3,383.3	1,867.5
横浜スタジアム	30,000	3,372.9	3,386.3	2,075.2
秩父宮ラグビー場	24,871	1,935.9	2,098.7	1,488.9
パシフィコ横浜	18,000	4,230.2	4,147.1	1,963.1
横浜アリーナ	17,000	2,928.3	2,927.0	2,306.8
日本武道館	14,471	2,641.4	2,654.8	1,906.8

「混雑統計®」©ZENRIN DataCom CO., LTD.

関係を表 6 に示す。「1 日前まで」の実験は表 3 の実験と同じ設定, すなわち, 1 日前までに入手可能な投稿と 1 日前のイベント会場の混雑度の時系列を用いる設定であり, 「2 日前まで」, 「3 日前まで」の実験はそれぞれ 2 日前, 3 日前までに入手可能な投稿および 2 日前, 3 日前のイベント会場の混雑度の時系列を用いる設定である.

実験の結果, いずれの会場においても予測期間が長くなるほど予測誤差が大きくなる傾向があることを確認できた. 特に, 「1 日前まで」と「2 日前まで」の実験設定での予測誤差の差が大きかった. このことは, 「明日」について言及している投稿が重要な手がかりになっている可能性や, イベント設営などの準

表 5: 人口変化の予測結果 (イベントがない日のみ).

会場名	収容人数	ベースライン		提案手法
		AR	VAR	中心のみ
日産スタジアム	72,327	211.4	210.7	218.8
東京ドーム	55,000	2,015.3	2,061.9	1,454.5
味の素スタジアム	50,000	189.0	187.9	187.3
明治神宮野球場	35,133	468.6	586.2	641.2
横浜スタジアム	30,000	767.6	829.1	679.0
秩父宮ラグビー場	24,871	695.0	832.1	501.3
パシフィコ横浜	18,000	1,053.7	1,162.4	1,030.7
横浜アリーナ	17,000	987.9	1,248.6	929.1
日本武道館	14,471	696.6	683.9	707.0

「混雑統計®」©ZENRIN DataCom CO., LTD.

備作業の影響が前日の混雑度の時系列に手がかりとして出てきている可能性を示唆している.

表 6: 予測期間と予測精度との関係.

会場名	収容人数	1 日前まで	2 日前まで	3 日前まで
日産スタジアム	72,327	244.4	248.1	248.5
東京ドーム	55,000	2,409.5	2,738.2	2,759.2
味の素スタジアム	50,000	339.8	347.3	356.1
明治神宮野球場	35,133	987.2	1,061.1	1,088.3
横浜スタジアム	30,000	985.0	1,152.1	1,189.6
秩父宮ラグビー場	24,871	552.7	566.7	567.5
パシフィコ横浜	18,000	1,158.4	1,271.8	1,329.5
横浜アリーナ	17,000	1,091.4	1,102.7	1,117.9
日本武道館	14,471	927.2	954.2	953.6

「混雑統計®」©ZENRIN DataCom CO., LTD.

5.6 投稿数と予測精度との関係

投稿数の多さが予測精度に寄与するかどうかを調べるために, 訓練データとテストデータ中の両方の投稿を一定の割合でランダムにサンプリングする条件で評価を行った. 投稿数と予測精度との関係を表 7 に示す. 使用する投稿の比率を 1.0 とした実験は表 3 の実験と同じ設定であり, 比率を 0.5, 0.1 とした実験はそれぞれ訓練データとテストデータ中の両方の投稿を 50%, 10% だけランダムにサンプリングして用いる設定である. また, 比率を 0.0 とした実験は, 前日のイベント会場の 1 日の混雑度の時系列の 24 次元ベクトルだけを Gradient Boosting Regressor の説明変数として用いる設定である.

実験の結果, 使用する投稿数が少なくなるほど予測誤差が大きくなる傾向が明らかになった. 投稿を全く使用しない場合は特に予測誤差が大きくなることから, マイクロブログ上の投稿にはイベント会場の人口変化の予測精度向上に寄与する何らかの情報が含まれていることが分かった.

6 おわりに

本論文では, 都市において大規模イベント会場周辺の人口変化の長期予測が特に難しいことを確認し, その問題に対処するために, マイクロブログ上の投稿およびイベント会場の過去の

表 7: 投稿数と予測精度との関係。

会場名	投稿数	使用する投稿の比率			
		1.0	0.5	0.1	0.0
日産スタジアム	12,360	244.4	259.2	244.4	272.9
東京ドーム	143,267	2,409.5	2,585.7	2,465.4	4,065.3
味の素スタジアム	31,763	339.8	355.4	336.5	463.7
明治神宮野球場	14,337	987.2	1,041.3	1,018.2	1,193.9
横浜スタジアム	34,465	985.0	1,102.3	1,034.3	1,298.1
秩父宮ラグビー場	4,136	552.7	556.1	545.6	596.4
パシフィコ横浜	38,742	1,158.4	1,226.6	1,156.9	1,401.8
横浜アリーナ	100,228	1,091.4	1,137.5	1,097.7	1,177.9
日本武道館	69,321	927.2	939.8	937.5	1,067.1

「混雑統計®」©ZENRIN DataCom CO., LTD.

混雑度を組み合わせて会場の未来の混雑度を長期にわたって予測する手法を提案した。提案手法の有効性は、実世界データを用いた 24 時間先までの予測実験によって検証した。実験を通して、提案手法は、いずれの会場においても既存手法である AR モデルと比較して高精度で予測できることを明らかにした。

今後は、投稿数が多い会場のデータを活用して投稿数が少ないイベントや会場の予測を改善する手法の提案や、イベントの種類や出演者の来場者数に対する影響力の可視化、さらには来場者数の公式発表がないような小規模イベントや花火大会のような正確な来場者数を把握することが困難な超大規模イベントへの本手法の適用等を行いたい。

謝 辞

本研究を進めるにあたり共同研究を通じて支援、コメントを頂いた国立研究開発法人情報通信研究機構の是津様、梅本様に感謝いたします。また、本研究の一部は、情報通信研究機構の高度通信・放送研究開発委託研究（課題 201：異分野データ連携によるスマートモビリティ基盤の研究開発）の成果です。

文 献

- [1] CNN, “‘I failed to protect you’ – Details emerge of victims in deadly Shanghai stampede.” <https://edition.cnn.com/2015/01/02/world/asia/china-shanghai-new-years-stampede/>, 2015.
- [2] 山. 達雄, 坪. 孝太, 丸. 三喜也, 山. 優樹, and 岡. 宏一朗, “乗換案内ログと SNS の融合による未来に発生する混雑原因の特定,” 人工知能学会全国大会論文集, vol. JSAI2018, pp. 4Pin141–4Pin141, 2018.
- [3] Z. Fan, X. Song, R. Shibasaki, and R. Adachi, “CityMomentum: An online approach for crowd behavior prediction at a citywide level,” in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 559–569, ACM, 2015.
- [4] B. Liao, J. Zhang, C. Wu, D. McIlwraith, T. Chen, S. Yang, Y. Guo, and F. Wu, “Deep sequence learning with auxiliary information for traffic prediction,” in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 537–546, ACM, 2018.
- [5] J. Zhang, Y. Zheng, and D. Qi, “Deep spatio-temporal residual networks for citywide crowd flows prediction,” in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- [6] T. Sakaki, M. Okazaki, and Y. Matsuo, “Earthquake shakes Twitter users: real-time event detection by social sensors,” in *Proceedings of the 19th International Conference on World Wide Web*, pp. 851–860, ACM, 2010.
- [7] T. Konishi, M. Maruyama, K. Tsubouchi, and M. Shimosaka, “CityProphet: City-scale irregularity prediction using transit app logs,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 752–757, ACM, 2016.
- [8] W. Yamada, D. Torii, H. Kikuchi, H. Inamura, K. Ochiai, and K. Ohta, “Extracting local event information from micro-blogs for trip planning,” in *2015 Eighth International Conference on Mobile Computing and Ubiquitous Networking (ICMU)*, pp. 7–12, IEEE, 2015.
- [9] A. Jatowt, É. Antoine, Y. Kawai, and T. Akiyama, “Mapping temporal horizons: Analysis of collective future and past related attention in Twitter,” in *Proceedings of the 24th International Conference on World Wide Web*, pp. 484–494, ACM, 2015.
- [10] M. Onishi and S. Nakashima, “Mutual interaction model between the number of people in real space and the number of tweets in virtual space,” in *2016 23rd International Conference on Pattern Recognition (ICPR)*, 2016.
- [11] J. He, W. Shen, P. Divakaruni, L. Wynter, and R. Lawrence, “Improving traffic prediction with tweet semantics,” in *Twenty-Third International Joint Conference on Artificial Intelligence*, 2013.
- [12] J. H. Friedman, “Greedy function approximation: a gradient boosting machine,” *Annals of statistics*, pp. 1189–1232, 2001.
- [13] T. Kudo, K. Yamamoto, and Y. Matsumoto, “Applying conditional random fields to Japanese morphological analysis,” in *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 230–237, Association for Computational Linguistics, 2004.