

# 地域の幾何学的関係を考慮したマルチタスク回帰に基づく 高性能な都市動態予報

安納 爽響<sup>1</sup> 坪内 孝太<sup>2</sup> 下坂 正倫<sup>1</sup>

## 概要：

未来の都市動態の予測，特にイベントが開催される日のいつもとは異なる動態パターン（異常混雑）の予測は，ユーザのモビリティ情報の価値を高めるために非常に重要な課題であり，近年盛んに研究されている．しかし既存手法では，未来の異常混雑のモデル化の難しさから予測可能な地域や時間が制限され，未来の都市動態の定常・異常を含めた高性能な予報は実現されていなかった．この課題の解決のために，本研究では「1週間後」という遠い未来の，「都市部のいかなる場所」における高性能な都市動態予報を実現するためのモデリング手法 **Geometrically Complemented multi-task Poisson Regression (GCPR)** を提案する．GCPR では，遠い未来の予測の実現のために乗換検索履歴を元にしたユーザの駅利用の行動予定をモデル化に組み込む．さらに「都市部のいかなる場所」に予測対象を一般化させるために，駅と予測対象エリアの幾何学的関係を考慮したモデリングを行う．実際の GPS 位置履歴，乗り換え案内アプリの検索履歴を元にした実験により，提案手法が予測地域全体の都市動態予測と，都市部のいかなる場所に存在するイベント会場において発生する異常混雑の早期予報において既存手法よりも優れた予測性能を達成することを示す．

## 1. 序論

近年，GPS 位置をトラック可能な携帯端末の普及により，都市における人々の動き（都市動態）をデータ駆動で理解し，未来の動態の予測を行うことが可能になっている [9], [10], [14]．例えば，あるイベント施設に対して1週間後の夕方に集まる人を予測することを考える．もし1週間後の夕方にイベントがない場合，すなわちいつもと同じようなコンテキストであれば，予測は難しくない．何故ならば，同じコンテキストを持つ過去の日の来訪者数を求め，統計的にいつもの人数を算出すればよいからである [8], [9]．しかし，1週間後の夕方に，同会場にて大きなイベントが開催される場合，予測は大変難しくなる．なぜならば，定常とは異なるパターンである異常混雑の度合いの予知は，普段は観測されないイベント発生を考慮したモデル化となり蓋を開けて見なければ困難である．

未来の都市動態の早期把握，特に異常混雑の未来予測は，ユーザのモビリティ情報の価値を高めるために非常にチャレンジングな課題であり，多くの研究者がこれに挑戦してきた [2], [4], [5], [6], [7]．異常な都市動態の事前予知

に焦点を当てた研究には，現在の動態の観測に基づく手法 [4], [5], [6] と，外的要素を取り入れた予測手法 [2] が存在する．Fan ら [4], Jinag ら [5], [6] らは，ユーザがイベント会場に向かうまでの時間に注目し，現在の動態に基づいて直近「1時間後」の異常動態をシミュレートする手法を提案している．また Anno [2] は，予測対象エリアを駅および駅に隣接する施設に限定し，1週間後の未来予測を可能にした．Konishi らの乗り換え案内の研究 [7] を基に，未来のユーザの行動予定情報として乗換検索履歴を用い，乗換検索のログ解析で推測された異常混雑量に相当する会場のユーザの数を推定し，間遠1週間後の予測を実現している．つまり，Fan ら，Jiang らは「1週間後」という遠い未来を予測することを諦める代わりに「いかなる場所でも」予測できるモデルを実現し，対して Anno らは解析対象ポイントを駅に絞る代わりに「1週間後」という遠い未来の予測を実現している．

これらの既存研究に対し本研究では，高性能な都市動態予報の実現のために，「1週間後」という遠い未来の，「駅」および「都市部のいかなる場所」についての予測という新しい問題を考える．都市部には共通の駅を最寄りとする複数のイベント会場が存在するため，駅周辺から都市部の地域全体に予測対象を一般化することには非常に大きな価値がある．だが一方で，地域全体に一般化された予測には，

<sup>1</sup> 東京工業大学 情報理工学院 情報工学系 Department of Computer Science, School of Engineering, Tokyo Institute of Technology

<sup>2</sup> Yahoo! JAPAN 研究所 Yahoo! Japan Corporation

実現に伴う困難が存在する。駅周辺での予測は乗り換え検索履歴の単純利用で可能であったが、駅から離れた施設の異常混雑では、予測に用いることのできる手がかりを得ることが難しい。また、一つのイベント会場に複数の最寄駅が存在し、イベント会場の違いによって最寄駅の利用パターンも変化することも考慮する必要がある。

このような課題に着目し、我々は駅から離れたイベント会場と、周辺駅利用パタンの「幾何学的関係」を考慮した、パラメータ共有と重み付き最適化に基づくマルチタスク双線型ポアソン回帰モデル、**Geometrically Complemented multi-task Poisson Regression (GCPR)** を提案する。提案手法では、駅から離れたイベント会場における混雑の手がかりを周辺駅におけるユーザの行動予定を用いて補う。さらに、イベント会場と駅の距離やパタンの類似度に基づいた「幾何学的関係」の考慮する。このために、イベント会場の人口動態予測モデルの学習パラメータの一部を、駅の利用パタンの予測モデルのものと共有する。加えて、会場と駅の距離やパタンの類似度に基づいて決定する「重み」を用いた、重み付き最適化を導入する。

本研究の貢献は以下である。

- 高性能な都市動態予報を実現するために、パラメータ共有と重み付き最適化に基づくマルチタスク双線型ポアソン回帰 *GCPR* を提案する。
- 実データを用いた実験により、提案手法が既存手法と比較して、都市のいかなる場所における未来の動態を、1週間前の段階からより高精度に予測可能であることを示す。

## 関連研究

**都市動態予測と異常混雑予測。** 都市動態の予測に焦点を当てた研究には、[1], [9], [16] がある。Zhang ら [16] ら、Akagi [1] らは、都市の人流の予測手法を提案しているが、本研究とは問題設定が異なる。Shimosaka ら [9] らは、外的要素から都市動態の予測を行うための手法を提案している。しかし、異常混雑を捉えるための特徴量を導入していないことにより、日常的に観測される活動人口のみ予測可能である。

将来の異常混雑の予測に焦点を当てた研究には、[2], [4], [5], [6], [7] がある。Fan ら [4], Jiang ら [5], [6] は、人口流の現在の観測に基づいて将来の人の動きを予測することで、異常混雑の予測を可能にしている。しかしこれらの手法は、予測可能な異常が「1時間後」に限られてしまう。Konishi ら、Anno らはユーザの将来の行動情報として乗換検索履歴に着目し、1週間後の異常混雑予測を可能にした。しかし、予測可能な地域は駅周辺に限られてしまう。これらに対し本研究では、「1週間後」の未来の、「都市のいかなる場所」での異常混雑を含めた都市動態の予報という新しい問題を考える。

**地域の関係性を考慮した都市動態のモデリング。** 地域の関係性を考慮した都市動態のモデリングには、[6], [8], [12], [13], [15] が存在する。Xia ら [12], [13] は、地域間で隠れ変数を共有した Hidden Markov model(HMM) による都市動態の予測手法を提案している。また、Shimosaka ら [8] は、細かいメッシュ地域上で隣接メッシュ間の学習パラメータを近づけるような正則化を用いた都市動態のモデリング手法を提案している。しかしこれらの手法は、データ量の不足による過学習の是正や性能の安定化に焦点を当てた研究であり、混雑の予測に焦点を当てた研究ではない。Zhang ら [15] は、地域間の類似度を考慮した都市動態の異常検知手法を提案している。しかしこの手法では、地域ごとの異常度の予測値を集計して類似度を考慮するため、事前に地域間関係を考慮した都市動態のモデリングではない。Zhang ら [16], Jinag ら [6] らは、周辺地域とのデータ共有を Convolutional neural network(CNN) を用いて実現し、地域間関係を考慮した都市動態のモデリング手法を提案している。しかし、これらの手法は、隣接しない地域や駅などの共通パターンは考慮できない。本研究では、メッシュ地域と周辺駅との関係を幾何学的な関係と捉え、距離や動態の類似度に基づいてモデリングを行う手法を検討する。

## 2. 都市動態予報の問題設定と実現に向けた困難

この章では、本研究で扱う都市動態予測の問題設定と、予測可能エリア「駅周辺」から「都市部のいかなる地域」に一般化しようとした際に生じる困難について述べる。

### 2.1 問題設定

#### 2.1.1 変数の定義

本研究では、「駅」および「都市部のいかなる場所」を予測の対象とし、この地域をメッシュで区切って表現する。 $L$  をメッシュの集合とし、各メッシュ  $l \in L$  をその要素と定義する。また、一日の時間帯を  $1, 2, \dots, T$  なる  $T$  個のセグメントに分割し、各時間帯を  $t$  で表す。

**定義 1. 都市動態。**  $d$  日  $t$  時間帯に、位置メッシュ  $l$  において観測される活動人口を  $y_{d,t}^{(l)}$  とし、これを都市動態と呼ぶ。

**定義 2. 乗換検索履歴に基づくスケジュール情報。** 乗換検索履歴は乗換検索クエリ  $q$  の集合からなり、各クエリ  $q$  は到着予定日時  $d, t$ 、検索日  $d'$ 、目的地駅  $s$  を要素として持つタプル  $q = (d, t, d', s)$  として表現される。乗換検索クエリ数  $x_{d,t|d'}^{(s)}$  は、目的地駅を  $s$ 、到着予定を  $d$  日  $t$  時に設定した検索クエリのうち、 $d'$  日に検索されたクエリの数を表す。このとき、一般に  $d'$  は  $d$  と等しいか、それ以前の日付を表す。すなわち、 $x_{d,t|d-i}^{(s)}$  は、到着予定日  $d$  の  $i$  日前に検索されたクエリ数を表す。こうして定義される乗換検索クエリ数は、ユーザの将来の行動予定の集計、すなわちスケ

ジュールのパタンを表す。検索クエリ数を用いて、到着予定日時  $d, t$  における乗換検索履歴に基づくスケジュール情報は、

$$X_{d,t}^{(s)} = \{x_{d,t|d-i}^{(s)} | i = p_d, p_d + 1, \dots, p_d + p_w\}, \quad (1)$$

と表される。このとき  $p_d$  は到着予定より以前の考慮する検索日付の中でもっとも以前の日を表し、 $p_w$  は考慮する検索の日付の範囲を表す。

### 2.1.2 問題の定義

本研究で我々は、これまでの研究で解かれている定常的な都市動態予測に加えて、イベントなどが発生した際における異常混雑についてその際の都市動態を、これまで可能であった「駅」および困難であった「都市部のいかなる場所」について早期予測する「高性能な都市動態予測」を考える。

**定義 3. 高性能な異常混雑予測。** 地域  $L$ 、日時  $d, t$  における都市動態  $y_{d,t}^{(l)}$  を、1週間前から予報する。特に、地域  $L$  において、日時  $d, t$  に異常混雑が起るとき、地域  $L$  に含まれるどのメッシュ  $l \in L$  でどれくらいの規模の異常混雑が起るのかを、都市動態  $y_{d,t}^{(l)}$  の予報を介して予報する。

## 2.2 高性能な異常混雑予測の困難

「1時間後」の予測は人口流の早期観測と将来のシミュレーションを元にした手法 [4], [5], [6] によって実現されつつあるが、予測には異常な人口流の予兆を事前に観測することが必要であり、予測が直近に限られてしまう。また間遠の予測に焦点を当てた際、活動人口のモデル化によく用いられる天気や曜日などのコンテキスト情報のみでは、異常な動態の予測は難しい。この解決のために、解析対象を駅周辺に絞り乗換検索履歴の応用を行うことで、間遠の予測を可能にした手法 [2], [7] が存在するが、これらの手法は乗換検索量の解析結果を駅周辺の動態に紐づける単純なモデル化のため、地域全体における予測に応用することは難しい。そこで本研究では、コンテキストと乗換検索のスケジュール情報を元にした、「都市部のいかなる場所」についての異常混雑予測の手法について検討する。

前提手法として、双線型ポアソン回帰 [9] に基づいた、曜日や天気等のコンテキスト情報と乗換検索履歴に基づくスケジュール情報を用いての都市動態予測モデル: BPreG-with-SI を考える。双線型ポアソン回帰は、コンテキストと時間帯との組み合わせに基づいたモデル化により、予測のための動態の事前観測が不要であるため、1週間後といった間遠の予測が可能となるモデルである。

$\mathbf{c}_d \in \mathbb{R}^C$  を日付  $d$  におけるコンテキスト特徴量ベクトル、 $\mathbf{t} \in \mathbb{R}^T$  を時間帯  $t$  に関する時間特徴量ベクトルとする。Shimosaka ら [9] は双線型ポアソン回帰を用いて、都市動態  $y_{d,t}^{(l)}$  を以下のようにモデル化している。

$$y_{d,t}^{(l)} \sim \text{Pois}(\cdot | \lambda_{\mathbf{c}_d, \mathbf{t}}^{(l)}), \quad (2)$$

$$\ln \lambda_{\mathbf{c}_d, \mathbf{t}}^{(l)} = \mathbf{c}_d^\top \mathbf{W}^{(l)} \mathbf{t}. \quad (3)$$

ここで、 $\mathbf{W}^{(l)} \in \mathbb{R}^{C \times T}$  はパラメータ行列である。また、 $\mathcal{N}(\cdot | \tau, \sigma^2)$  を、平均パラメータ  $\tau$ 、分散パラメータ  $\sigma^2$  により規定される正規分布とし、時間特徴量を  $\mathbf{t} = \{t_s | t_s = \mathcal{N}(s | \tau, \sigma^2), s = 1, \dots, T\}$  と与える。

異常混雑の予測のために、この双線型回帰モデルのスケジュール情報の考慮による単純な拡張として、地域  $L$  に含まれる  $N$  個の鉄道駅  $s_i \in S_L, (i = 1, \dots, N_s)$  を目的地としたスケジュール情報をモデルの入力に取ることを考える。具体的にはスケジュール特徴量ベクトル  $\mathbf{x}_{d,t}^{(S_L)}$  を、乗換検索履歴に基づくスケジュール情報  $X_{d,t}^{(s_i)}$  と定数  $\xi$  を用いて  $\mathbf{x}_{d,t}^{(S_L)} = \{\xi x | x \in X_{d,t}^{(s_i)}, (i = 1, \dots, N_s)\} \in \mathbb{R}^{N_s p_w}$  として与え、以下のようにモデル化する。

$$\ln \lambda_{\mathbf{c}_d, \mathbf{t}}^{(l)} = [\mathbf{c}_d^\top, \mathbf{x}_{d,t}^{(S_L)\top}] \mathbf{Y}^{(l)} \mathbf{t}. \quad (4)$$

ここで、 $\mathbf{Y}^{(l)}$  はパラメータ行列であり、 $\mathbf{Y}^{(l)} \in \mathbb{R}^{(N_s p_w + C) \times T}$  である。しかしこの安直な拡張は、駅利用者の利用パターンや、複数駅とメッシュの距離や利用パターン類似度などの幾何学的関係を考慮できておらず、都市部地域全体の異常混雑に対して頑健さを有するモデル性能の実現は難しい。

## 3. 提案手法: Geometrically Complemented multi-task Poisson Regression

本章では、「都市部のいかなる場所」についての異常混雑予測のために、提案手法: Geometric-Complemented multi-task Poisson Regression (GCPR) について述べる。まず 3.1 章にて提案手法の概略について述べる。次に、3.2, 3.3 章にてポアソン回帰のパラメータ共有および幾何学的関係考慮のための重み付き最適化について述べる。最後に 3.4 章にて、学習パラメータの推論について述べる。

### 3.1 提案手法の概略

図 1 に、提案手法 GCPR の概略図を示す。本研究では、都市部地域全体における予測を実現するために、対象のエリアには周辺の最寄駅利用パターンから補完できる動態と、エリア固有の動態が存在すると仮定する。そこで、対象エリアの都市動態の予測モデル BPreG-with-SI と、スケジュール情報を用いた乗換検索量の予測モデルのパラメータを一部共有し、固有の動態と周辺駅からの割戻による動態を同時にモデリングすることを考える。また、その割戻の程度が、対象エリアと駅との距離、および動態の相関に基づく「幾何学的関係」に依存すると仮定し、関係を考慮した学習のために「重み」を導入した最適化を考える。

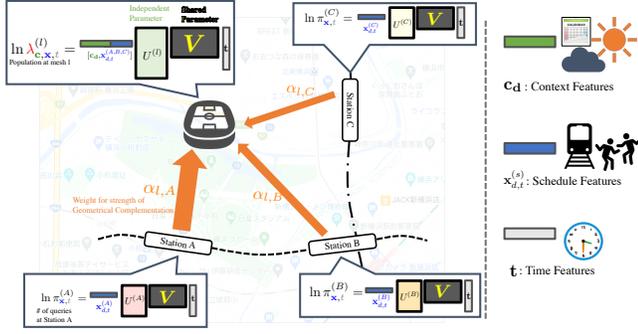


図 1 提案手法 GCPR における施設と周辺駅の幾何学的関係を考慮したモデリングの概略図。

### 3.2 パラメータ共有双線型ポアソン回帰による 駅利用パターンとメッシュ固有動態の考慮

前述の通り，対象とするエリアには周辺の最寄駅利用パターンから補完可能な動態と，エリア固有の動態が存在すると仮定する．周辺の各駅  $s_i$  の乗換検索量のうち，当日に検索された数  $x_{d,t|d}^{(s_i)}$  を，それぞれの駅のスケジュール情報  $X_{d,t}^{(s)}$  から予測するモデルを  $\mathcal{M}^{(s_i)}$  とする．各駅  $s_i$  に対する予測モデル  $\mathcal{M}^{(s_i)}$  を双線型ポアソン回帰に基づき，式 5, 6 のように定式化する．

$$x_{d,t}^{(s_i)} \sim \text{Pois}(\cdot | \pi_{\mathbf{x},t}^{(s_i)}), \quad (5)$$

$$\ln \pi_{\mathbf{x},t}^{(s_i)} = [1, \mathbf{x}_{d,t}^{(s_i)\top}] \mathbf{Y}^{(s_i)} \mathbf{t}, \quad (6)$$

ここで， $\mathbf{Y}^{(s_i)} \in \mathbb{R}^{p_w}$  はパラメータ行列である．また， $\mathbf{x}_{d,t}^{(s_i)}$  は各駅  $s_i$  ごとのスケジュール特徴量ベクトルであり， $\mathbf{x}_{d,t}^{(s_i)} = \{\xi \mid x \in X_{d,t}^{(s_i)}\}$  である． $\mathbf{t}$  は前述の時間特徴量である．

都市部地域全体におけるモデル化において，BPreG-with-SI の学習パラメータ  $\mathbf{Y}^{(l)}$  の一部を，スケジュール情報を用いた各駅  $s_i$  の乗換検索量の予測モデルのパラメータ  $\mathbf{Y}^{(s_i)}$  と共有する．学習パラメータの一部を共有するために，BPreG-with-SI と  $\mathcal{M}^{(s_i)}$  の各モデルのパラメータを行列分解し，式 7, 8 のように記述し直す．

$$\mathbf{Y}^{(l)} = \mathbf{U}^{(l)} \mathbf{V}^\top, \quad (7)$$

$$\mathbf{Y}^{(s_i)} = \mathbf{U}^{(s_i)} \mathbf{V}^\top, \quad (8)$$

ここで， $\mathbf{U}^{(l)} \in \mathbb{R}^{(N_s p_w + C) \times K}$ ， $\mathbf{U}^{(s_i)} \in \mathbb{R}^{(p_w + 1) \times K}$  は対象とするメッシュ  $l$ ，駅  $s_i$  に固有のパラメータ行列であり， $\mathbf{V} \in \mathbb{R}^{T \times K}$  はモデル間で共有するパラメータ行列である．これらのパラメータを用いて，ポアソン回帰モデルのパラメータ  $\lambda_{\mathbf{c},\mathbf{x},t}^{(l)}$ ， $\pi_{\mathbf{x},t}^{(s_i)}$  は，式 9, 10 のように表される．

$$\ln \lambda_{\mathbf{c},\mathbf{x},t}^{(l)} = [\mathbf{c}_d, \mathbf{x}_{d,t}^{(S_L)}]^\top \mathbf{U}^{(l)} \mathbf{V}^\top \mathbf{t}, \quad (9)$$

$$\ln \pi_{\mathbf{x},t}^{(s_i)} = [1, \mathbf{x}_{d,t}^{(s_i)}]^\top \mathbf{U}^{(s_i)} \mathbf{V}^\top \mathbf{t}. \quad (10)$$

### 3.3 幾何学的関係考慮のための重み付き最適化

パラメータ共有によるモデル化により，各メッシュにおいて周辺駅からの割戻による動態要素を考慮することが可能になる．しかし，地域  $L$  内の異なるイベント会場を含んだメッシュを考えた場合，このモデル化では不十分である．何故ならば，単純なパラメータの共有のみによる学習は，周辺駅からの割戻がメッシュの違いに寄らず一定であることを仮定するからである．そこで我々は，イベント会場の違いによって周辺の最寄駅の利用パターンが異なることを考慮するために，幾何学的関係に基づいた重みを考慮した最適化を導入する．

メッシュ  $l \in L$  と，地域  $L$  に含まれる鉄道駅  $s \in S_L$  を考える．幾何学的関係として我々は，1) メッシュと駅の距離，2) 動態の類似度を考える．メッシュ  $l$  と駅  $s$  間の重み  $\alpha_{l,s}$  を，以下のように与える．

$$\alpha_{l,s} = s_{\text{dist}} A_{l,s}^{\text{dist}} + s_{\text{sim}} A_{l,s}^{\text{sim}}, \quad (11)$$

ここで， $A_{l,s}^{\text{dist}}$  はメッシュと駅の距離を考慮して与える重みであり， $A_{l,s}^{\text{dist}} = \exp(-\eta_{\text{dist}} \text{dist}(l, s))$  で与える．このとき， $\eta_{\text{dist}} > 0$  であり， $\text{dist}(\cdot, \cdot)$  は，メッシュ  $l$  と，駅  $s$  を含むメッシュとのマンハッタン距離を表す．また， $A_{l,s}^{\text{sim}}$  は動態の類似度を考慮して与える重みであり， $A_{l,s}^{\text{sim}} = (1 + \text{corr}(l, s))/2$  で与える．ここで  $\text{corr}(l, s)$  は，メッシュ  $l$  で観測される都市動態  $y_{d,t}^l$  と，駅  $s$  を目的地として到着日当日に検索される検索クエリ数  $x_{d,t|d}^{(s)}$  の相関係数である．定義より， $0 < A_{l,s}^{\text{dist}} \leq 1$ ， $0 \leq A_{l,s}^{\text{sim}} \leq 1$  である．また， $s_{\text{dist}} > 0$ ， $s_{\text{sim}} > 0$  は，2 種類の重みのスケールに関する定数である．

### 3.4 パラメータの学習

前章までで定義されるパラメータ共有と幾何学的関係を考慮した重みを用いて，モデルのパラメータの学習を行う．対数尤度に基づいた目的関数  $\mathcal{L}^{(l)}(\mathbf{U}^{(l)}, \mathbf{U}^{(s_1:s_N)}, \mathbf{V})$  は，以下で定義される．

$$\begin{aligned} \ln \mathcal{L}^{(l)}(\mathbf{U}^{(l)}, \mathbf{U}^{(s_1:s_N)}, \mathbf{V}) &= - \sum_d \sum_t \ln \text{Pois}(y_{d,t}^{(l)} | \lambda_{\mathbf{c},\mathbf{x},t}^{(l)}; \mathbf{U}^{(l)}, \mathbf{V}) \\ &\quad - \sum_{i=1}^N \sum_d \sum_t \alpha_{l,s_i} \ln \text{Pois}(x_{d,t|d}^{(l)} | \pi_{\mathbf{x},t}^{(s_i)}; \mathbf{U}^{(s_i)}, \mathbf{V}). \end{aligned} \quad (12)$$

正則化として，L2 正則化を用いる．最終的に，学習パラメータは，以下の最適化問題を解くことにより得られる．

$$\begin{aligned} \arg \min_{\mathbf{U}^{(l)}, \mathbf{U}^{(s_1:s_N)}, \mathbf{V}} & - \ln \mathcal{L}^{(l)}(\mathbf{U}^{(l)}, \mathbf{U}^{(s_1:s_N)}, \mathbf{V}) \\ & + \Omega(\mathbf{U}^{(l)}, \mathbf{U}^{(s_1:s_N)}, \mathbf{V}), \end{aligned} \quad (13)$$

ここで、 $\Omega(\mathbf{U}^{(l)}, \mathbf{U}^{(s_1:s_N)}, \mathbf{V}) = \gamma(\|\mathbf{U}^{(l)}\|_{\text{Fro}}^2 + \|\mathbf{V}\|_{\text{Fro}}^2 + \sum_s \|\mathbf{U}^{(s)}\|_{\text{Fro}}^2)$  であり、 $\gamma > 0$  である。 $\|\cdot\|_{\text{Fro}}^2$  は、行列のフロベニウスノルムである。

ここで導入される最適化は双凸ではないが、固定した  $\mathbf{V}$  の元での  $\mathbf{U}^{(l)}, \mathbf{U}^{(s_1:s_N)}$  の最適化、および、固定した  $\mathbf{U}^{(l)}, \mathbf{U}^{(s_1:s_N)}$  のもとでの  $\mathbf{V}$  の最適化は凸である。従って学習は、非負行列分解 (NMF) の中でよく用いられる、交互最小二乗法 [3] と同様の方法で行うことができる。また、最適化は準ニュートン法 [11] を用いる。

## 4. 性能評価実験

本研究の目的である高性能な都市動態予報の問題における提案手法の性能を評価するために、実際の GPS 位置履歴データ、乗り換え案内アプリの検索履歴を用いて比較実験を行なった。

### 4.1 実験データ

本実験で用いる GPS 位置履歴は、Yahoo! JAPAN から提供されている防災速報アプリ \*1 によって、ユーザの同意のもとで取得された、位置履歴情報である。データはユーザ情報を取り除くことで匿名化し、さらに統計情報化したものであるため、個人を特定できるデータではない。データはタイムスタンプと緯度経度の情報を持ち、日本全国のユーザから取得されている。データ数は東京圏のみで 1 日に約 1500 万件にも及ぶ。データは基本的に移動している状態の端末や基地局が切り替わった端末から取得されるため、人々の動きを表したデータと言える。利用するデータの期間は、COVID-19 による動態の変化を考慮し、2019 年 10 月 1 日から 2020 年 1 月 31 日までの 4 ヶ月間とした。

また、本実験で用いる乗換検索履歴は、同じく Yahoo! JAPAN から提供されている乗り換え案内アプリ \*2 の検索履歴データである。位置履歴と同様、データは完全に匿名化されており、検索時のタイムスタンプとスケジュールされたタイムスタンプ、目的地駅の情報を含む。利用するデータの期間は位置履歴と同様である。図に示すメッシュから構成される地域に含まれている駅は表に示されている通りであり、各駅での検索クエリ数を位置履歴と同様の 60 分間隔でカウントし、訓練データとする。

解析対象とした地域は、複数の鉄道駅周辺に複数のイベント会場が存在する地域を選び、図 2 に示す日産スタジアムの周辺を解析対象とした。この地域には、2019 年にラグビー W 杯が行われた日産スタジアムのほか、アーティストによる音楽ライブが行われる横浜アリーナを含み、また周辺駅として新横浜駅、小机駅、北新横浜駅を含む。この  $3.3 \times 1.2\text{km}$  四方の領域内に落とされた位置履歴情報を、各メッシュ 60 分間隔でカウントし、都市動態の訓練・評

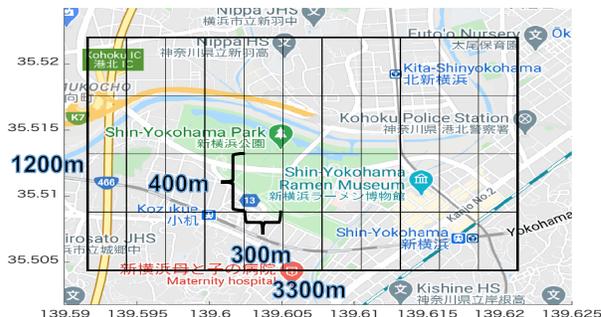


図 2 日産スタジアム周辺の解析対象地域。

価データとする。また、この地域に含まれている各駅での検索クエリ数を位置履歴と同様の 60 分間隔でカウントし、訓練データとする。

### 4.2 モデルの設定

#### 4.2.1 回帰モデルで使用する特徴量

本実験では、回帰モデルには次のような特徴量を与える。時間特徴量  $t$  とスケジュール特徴量  $\mathbf{x}_{d,t}^{(s)}$  は、2.2 章に示すものを使用する。コンテキスト特徴量  $\mathbf{c}_d$  としては、平日か否かを表す 2 次元の 1-of-K 表現特徴量  $\mathbf{c}_{d1} \in \mathbb{R}^2$ 、天気 (晴れ, 曇り, 雨, その他) を表す 4 次元の 1-of-K 表現特徴量  $\mathbf{c}_{d2} \in \mathbb{R}^4$  を用い、 $\mathbf{c}_d = \mathbf{c}_{d1} \otimes \mathbf{c}_{d2}$  とする。

#### 4.2.2 ハイパーパラメータの設定

提案手法に与えるハイパーパラメータには、重み  $\alpha_{l,s}$  を決定するための定数  $s_{\text{dist}}, s_{\text{sim}}, \eta_{\text{dist}}$  であり、それぞれ経験的に  $s_{\text{dist}} = 1.8, s_{\text{sim}} = 1.0, \eta_{\text{dist}} = 0.1$  とした。また、スケジュール情報をベクトル化する際に用いる定数  $\xi$  は  $\xi = 0.05$ 、正則化の強さを制御するパラメータ  $\gamma$  は、 $\gamma = 0.01$  を採用した。

### 4.3 実験設定と比較手法

本実験では、まず実験 1 にて、提案手法の通常時・異常混雑時を問わない活動人口の予測性能を評価するために、実験期間内の全メッシュにおける提案手法の活動人口予測の性能を、既存の活動人口予測手法である双線型ポアソン回帰 (BPRReg) [9]、異常混雑予測の既存手法である Supervised-CityProphet (SCP) [2]、本研究の前提手法としたスケジュール情報を用いた双線型ポアソン回帰 (BPRReg-with-SI),  $\alpha_{s,t} = 1.0$  に固定したパラメータ共有双線型ポアソン回帰 (GCPR( $\alpha = 1$ )) と比較する。実験 2 では、「都市のいかなる場所」で異常混雑発生時における活動人口の予測性能の評価のために、イベントにより異常混雑が起こった日時・その場所における活動人口予測の性能を、SCP, BPRReg-with-SI, GCPR( $\alpha = 1$ ) と比較する。

BPRReg [9], BPRReg-with-SI は、2 章で述べたものを用い

\*1 <http://emg.yahoo.co.jp/>

\*2 <https://transit.yahoo.co.jp/smartphone/app/>

る。GCPR( $\alpha = 1$ ) は、重み  $\alpha_{s,t}$  を全てのメッシュで 1 に固定し、幾何的関係の考慮を用いないパラメータ共有双線型ポアソン回帰である。

実験 2 における評価では、イベントによる異常混雑発生のラベルの真値を定量的に定義するために、Shimosaka ら [9] が提案している異常混雑検知の規範に基づいて、異常混雑のラベルを与える。具体的には、メッシュ  $l$  にて定常的に観測される活動人口を双線型ポアソン回帰に基づく予測値  $\hat{\lambda}_{d,t}^{(l),BP}$  で定義し、実際の活動人口  $y_{d,t}^{(l)}$  に基づいて異常度を以下で定義する。

$$\nu_{d,t}^{(l)} = \frac{y_{d,t}^{(l)} - \hat{\lambda}_{d,t}^{(l),BP}}{\hat{\lambda}_{d,t}^{(l)}}. \quad (14)$$

この異常度に対して、異常混雑かどうかを決める閾値  $\bar{\nu}$  を事前に与え、 $\nu_{d,t}^{(l)} \geq \bar{\nu}$  となる日時  $d, t$  とメッシュ  $l$  を異常混雑の発生した日時・場所とする。本実験では、閾値は  $\bar{\nu} = 10, 15, 20$  に設定する。評価の際は、この規範により検知された異常を 1 週間前から予測可能かということについて評価する。

また、Supervised-CityProphet (SCP) [2] は、都市動態の異常度  $\nu_{d,t}^{(l)}$  を定常的な活動人口  $\hat{y}_{d,t}^{(l)}$  と現在の観測  $y_{d,t}^{(l)}$  に基づいて、 $\nu_{d,t}^{(l)} = (y_{d,t}^{(l)} - \hat{y}_{d,t}^{(l)})/\hat{y}_{d,t}^{(l)}$  で与え、この異常度をモデル化するものである。具体的には、異常度  $\nu_{d,t}^{(l)}$  を、乗換検索履歴を元にして 1 週間前の段階から観測されているスケジュール特徴量から回帰させる。異常度予測のためのスケジュール特徴量  $\zeta_{d,t|t'}^{(S_l)}$  は、メッシュ  $l$  の周辺駅の集合を  $S_l$  とし、乗換検索履歴に基づくスケジュール情報  $X_{d,t}^{(S_l)}$  と、周辺駅の定常的な乗換検索量  $\tilde{x}_{d,t}^{(S_l)}$  を用いて、 $\zeta_{d,t|t'}^{(S_l)} = (x_{d,t|t'}^{(S_l)} - \tilde{x}_{d,t}^{(S_l)})/\tilde{x}_{d,t}^{(S_l)}$  と定義される。この時、異常度は以下の回帰により予測される。

$$\hat{\nu}_{d,t}^{(l)} = \sum_{i=p_d}^{p_d+p_w} \sum_{j=-q}^q w_{i,j} \zeta_{d,t-j|d-i}^{(S_l)} + w_c, \quad (15)$$

この時  $w_{i,j}, w_c \in \mathbb{R}$  は学習パラメータである。本実験では、 $S_l$  には提案手法と同じ駅集合である、新横浜駅、小机駅、北新横浜駅を用いる。予測された異常度  $\hat{\nu}_{d,t}^{(l)}$  と定常的な活動人口  $\hat{y}_{d,t}^{(l)}$  を用いて、最終的な活動人口の予測値  $\hat{y}_{d,t}^{(l),Ano}$  は、 $\hat{y}_{d,t}^{(l),Ano} = (1 + \hat{\nu}_{d,t}^{(l)})\hat{y}_{d,t}^{(l)}$  で与える。

#### 4.4 評価指標

評価指標には、MAE, MAPE を用いる。活動人口の真値を  $y_{d,t}^{(l)}$ 、その予測値を  $\hat{y}_{d,t}^{(l)}$  として、MAE, MAPE は以下の式で定義される。

$$MAE = \frac{\sum_{d=1}^D \sum_{t=1}^T |y_{d,t}^{(l)} - \hat{y}_{d,t}^{(l)}|}{DT}, \quad (16)$$

$$MAPE = \frac{\sum_{d=1}^D \left| \arg \max_{t \in [1, T]} (y_{d,t}^{(l)}) - \arg \max_{t \in [1, T]} (\hat{y}_{d,t}^{(l)}) \right|}{D}. \quad (17)$$

#### 4.5 実験結果

##### 実験 1. 活動人口の予測性能の評価実験

表 1 活動人口の予測性能の評価

model	Evaluation Metric	
	MAE	MAPE
BPReg	8.86 ± 3.30	<b>156.94 ± 10.36</b>
BPReg-with-SI	8.27 ± 2.05	161.88 ± 14.30
SCP	9.95 ± 2.34	158.62 ± 12.10
GCPR( $\alpha = 1$ )	8.19 ± 2.03	164.02 ± 16.93
<b>GCPR(proposed)</b>	<b>8.18 ± 2.03</b>	163.97 ± 15.42

ここでは、提案手法の通常時・異常混雑時を問わない活動人口の予測性能を評価するために、実験期間内の全メッシュにおける提案手法の活動人口予測の性能を、既存の活動人口予測手法との比較により評価する。異常混雑は一般的に発生が稀であるため、この性能評価は主に定常的な活動人口の予測において行われていることに留意されたい。ここでは提案手法が、既存の state-of-the-art な予測手法である双線型ポアソン回帰や、ベースラインである BPReg-with-SI と比較して、MAE, MAPE において同程度の予測性能を担保できていることを確認する。

表 1 に実験結果を示す。まず MAE について、提案手法である GCPR がもっとも良い性能を達成し、GCPR( $\alpha = 1$ ) がそれに追従する予測精度を達成している。また MAPE の予測に関しても、GCPR が既存の予測手法と同程度の予測性能が得られている。

##### 実験 2. 異常混雑発生時における活動人口の予測性能の評価実験

ここでは、「都市のいかなる場所」における異常混雑発生時の活動人口の予測性能の評価のために、異常混雑を定量的に観測できた日時・場所における活動人口予測の性能を、既存の予測手法との比較により評価する。

表 2 に、実験結果を示す。まず異常度に関する閾値  $\bar{\nu}$  が大きくなるにつれて、いずれの予測手法でも MAE が大きくなるのがわかる。これにより、異常混雑の事前の予知というタスクが極めて難しいことがわかる。このような困難の中でも、提案手法である GCPR(Proposed) が、ベースラインである BPReg-with-SI や既存研究における state-of-the-art な手法である SCP、また GCPR( $\alpha = 1$ ) と比較して、高い予測性能を達成している。

#### 4.6 異常混雑予測に関する定性的評価

ここでは、イベントが開催され異常混雑が起こった日時

表 2 異常混雑発生時における活動人口の予測性能の評価

model	Evaluation on MAE		
	$\bar{v} = 10.0$	$\bar{v} = 15.0$	$\bar{v} = 20.0$
BPRReg-with-SI	10.86 ± 9.80	21.54 ± 11.66	23.25 ± 13.38
SCP	19.71 ± 16.84	35.59 ± 17.26	37.23 ± 18.90
G CPR( $\alpha = 1$ )	10.47 ± 9.19	20.66 ± 10.41	21.40 ± 11.15
<b>G CPR(proposed)</b>	<b>10.42 ± 9.24</b>	<b>20.50 ± 10.63</b>	<b>21.38 ± 11.51</b>

におけるメッシュ全体での異常混雑の検知結果と、それを1週間前から予測した結果とを可視化して比較することにより、提案手法が「都市のいかなる地域」にて異常混雑を予測可能であることを定性的に確認する。ここでは、4.3章にも示した Shimosaka [9] らの提案する異常検知規範と異常度の定義を用いて、メッシュ状に異常度を可視化することにより評価を行う。

2019年11月2日、日産スタジアムにてラグビー W 杯 2019 \*3決勝が開催された。試合会場の開場時刻は午後3時、試合開始時刻は午後6時である。当日は約7万人 \*4の観客を動員して試合が行われた。図3には、同日の午後4時における異常混雑予測の比較を示す。図3(中)からわかるように、日産スタジアム周辺で、異常度15を超える異常混雑が観測されていることがわかる。これに対し、図3(右)からわかるように、提案手法は異常混雑の発生を、駅周辺だけでなく、スタジアム周辺においても予測できていることがわかる。

また、2019年11月6日、7日に横浜アリーナにて、アーティスト (ONE OK ROCK) のライブ公演が行われた。このライブ公演は全国ツアーの形で行われ、総動員数はのべ30万人 \*5となった。会場の収容可能人数は1万7千人であり、会場の開場時刻は午後6時、ライブ公演の開演は午後7時である。図4には、同日の午後6時における異常混雑予測の比較を示す。図4(中)からわかるように、会場である横浜アリーナの北側で、異常混雑が観測されていることがわかる。会場において観客は携帯端末の電源を切る必要があるため、一部のメッシュでは異常混雑が観測されていないことに注意されたい。これに対し、図4(右)からわかるように、提案手法は横浜アリーナ周辺での異常混雑の発生をライブの1週間前の段階から捉えられていることがわかる。

## 5. 議論

本章では、実験により得られた定量的・定性的評価を元に、本研究の目的である高精度な都市動態予測に関して、提案手法 GCPR について議論する。実験1により得られた結果から、提案手法が期間内の全メッシュでの、主に定

常的な活動人口の予測タスクにおいて、既存手法に引け劣らない安定した性能を達成できていることがわかる。また、実験2により得られた結果から、「都市のいかなる場所」における異常混雑の1週間前からの予測タスクにおいても、提案手法が既存手法と比較して高い予測精度を提供していることが確認できる。さらに、日産スタジアムと横浜アリーナでの、異なる2つの会場でのイベントにおける可視化結果より、提案手法がイベント会場の違いを考慮したモデリングが可能であることが示唆される。したがって、本研究の目的である高精度な都市動態予測に関して、提案手法 GCPR の学習パラメータ共有と幾何的關係を考慮した重み付き最適化が効果的であることが伺える。

## 6. 結論

本研究では、都市部のいかなる場所における定常および異常な都市動態に焦点を当て、この動態の1週間前からの早期予測を可能にする高性能な都市動態予測のために、パラメータ共有と重み付き最適化に基づくマルチタスク双線型ポアソン回帰モデル GCPR を提案した。一般に、未来の都市動態、特に異常混雑の早期の予知は困難なタスクであり、既存手法では予測可能な異常を時間的 (e.g. 1時間後の予測) あるいは空間的 (e.g. 駅周辺のみ) に制限して予測を可能にしていた。

本研究では、この制限を取り払い、1週間後の都市部のいかなる場所における予測のために、イベント会場と駅の幾何的關係を考慮するための新しいモデリング手法を提案した。実際の GPS 位置履歴と乗り換え案内アプリの検索履歴を用いて性能比較実験を行い、異常混雑時における予測において既存の予測手法である SCP と比較して MAE で 15.85 の改善が見られた。また期間・地域全体における評価でも、提案手法が高精度かつ安定した予測を提供可能であることを示した。

将来課題として、より広範囲な都市圏地域での予測のための、提案手法の拡張が考えられる。より広範囲な地域における予測は、応用状大変重要である。しかし、都市地域は鉄道網が入り組んでおり、考慮すべき幾何的關係がより複雑になると考えられる。したがって、この課題を解決する手法が求められる。

\*3 <https://www.rugbyworldcup.com/>

\*4 <https://www.rugbyworldcup.com/news/538422>

\*5 <https://news.yahoo.co.jp/articles/adda1f03002f906d51612193dbd85e1>

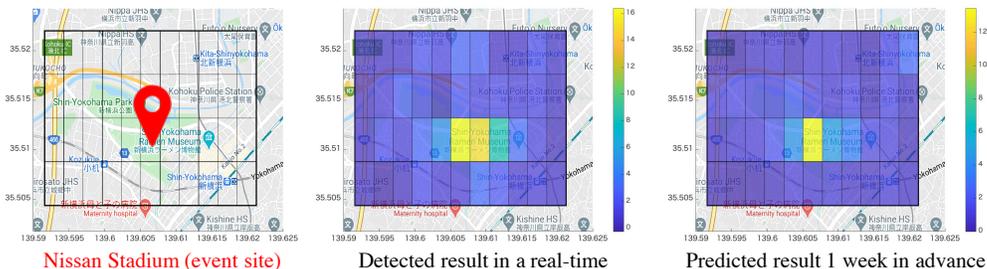


図 3 2019 年 11 月 2 日の日産スタジアムにおけるラグビー W 杯決勝時の、午後 4 時の混雑具合の予測結果の可視化。(左) イベント会場の場所。(中) 当日観測された異常度。(右) 1 週間前に予測した異常度。

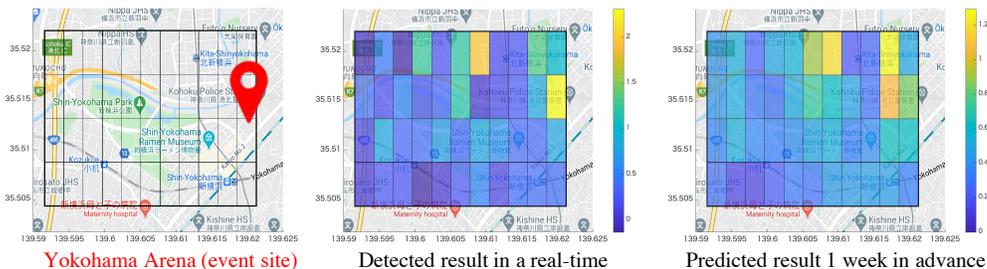


図 4 2019 年 11 月 7 日の横浜アリーナにおけるアーティストのライブ公演時の、午後 6 時における混雑具合の予測結果の可視化。(左) イベント会場の場所。(中) 当日観測された異常度。(右) 1 週間前に予測した異常度。

参考文献

- [1] Y. Akagi, T. Nishimura, T. Kurashima, and H. Toda. A fast and accurate method for estimating people flow from spatiotemporal population data. In *IJCAI*, 2018.
- [2] S. Anno, K. Tsubouchi, and M. Shimosaka. Supervised-cityprophet: Towards accurate anomalous crowd prediction. In *Proceedings of the 28th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2020.
- [3] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons. Algorithms and applications for approximate nonnegative matrix factorization. *Computational statistics & data analysis*, 2007.
- [4] 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*, 2015.
- [5] R. Jiang, X. Song, Z. Fan, T. Xia, Q. Chen, S. Miyazawa, and R. Shibasaki. Deepurbanmomentum: An online deep-learning system for short-term urban mobility prediction. In *AAAI*, 2018.
- [6] R. Jiang, X. Song, D. Huang, X. Song, T. Xia, Z. Cai, Z. Wang, K.-S. Kim, and R. Shibasaki. Deepurbanevent: A system for predicting citywide crowd dynamics at big events. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019.
- [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*, 2016.
- [8] M. Shimosaka, Y. Hayakawa, and K. Tsubouchi. Spatiality preservable factored poisson regression for large-scale fine-grained gps-based population analysis. In *AAAI*, 2019.
- [9] M. Shimosaka, K. Maeda, T. Tsukiji, and K. Tsubouchi. Forecasting urban dynamics with mobility logs by bilinear poisson regression. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2015.
- [10] L. Sun, D.-H. Lee, A. Erath, and X. Huang. Using smart card data to extract passenger’s spatio-temporal density and train’s trajectory of mrt system. In *Proceedings of the ACM SIGKDD international workshop on urban computing*, 2012.
- [11] S. Wright and J. Nocedal. Numerical optimization. *Springer Science*, 1999.
- [12] T. Xia and Y. Li. Revealing urban dynamics by learning online and offline behaviours together. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2019.
- [13] T. Xia, Y. Yu, F. Xu, F. Sun, D. Guo, D. Jin, and Y. Li. Understanding urban dynamics via state-sharing hidden markov model. In *The World Wide Web Conference, WWW ’19*, 2019.
- [14] J. Yuan, Y. Zheng, and X. Xie. Discovering regions of different functions in a city using human mobility and pois. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2012.
- [15] H. Zhang, Y. Zheng, and Y. Yu. Detecting urban anomalies using multiple spatio-temporal data sources. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2018.
- [16] J. Zhang, Y. Zheng, and D. Qi. Deep spatio-temporal residual networks for citywide crowd flows prediction. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 2017.