

敵対的生成モデルに基づく 活動人口の 波形描画を用いた 混雑寿命予報

安納 爽響¹ 坪内 孝太² 下坂 正倫¹

概要: イベント下の群衆混雑は来訪者の動きに応じて人数の増減パターンが見られ、そのような混雑の発生から消滅までの一連の予報は、来訪者・イベント主催者双方にとって重要なアプリケーションである。既存研究では、イベント種ごとに増減パターンの共通点があることを用いて、増減の状態遷移をイベント情報から学習・予測し、これを基に訪問者の人口密度を推定している。しかし、予測された状態列からは具体的な人数変化量が分からず、そのために発生する局所的な推定の誤りが、時系列上の大きな予測誤差に繋がることが課題であった。本研究ではこの課題の対処のため、状態遷移の代わりに人口密度推移の形状をモデル化する。形状には、状態遷移の持つ増減傾向に加えて、概算的な人数の変化量や推移傾向が反映されるため、混雑の一部始終の傾向補足が可能となり、寿命予測に適していると言える。そこで、形状的特性を精緻にモデル化するため、敵対的生成ネットワーク (Generative Adversarial Nets; GAN) を用いて形状を学習・描画するフレームワークを提案する。実データと人工データで、既存の混雑予報手法と比較して、形状を直接描き出す提案手法が、より小さな誤差で混雑を予測可能であることを示す。

1. 序論

本研究では、イベントなどに際してある地域が人々で混み合う現象である“群衆混雑”について、その発生から消滅までの一連を予測する問題を「混雑寿命予報」と称し、この問題に取り組む。群衆混雑は、イベント来訪者の快適かつ安全な移動の妨げになりやすく、実際に 2022 年には韓国梨泰院において、混雑により群衆雪崩が発生し多数の死傷者を出した^{*1}。そのため群衆混雑の予報は実用上重要な課題であり、人流モデリング領域において、これまで様々な方法論が議論されてきている。

人流の予測手法は、1) 人流の自己相関に着目した手法 [5], [11], [13] と、2) 外的要因に基づいた手法 [2], [21], [26] に大別される。一つ目の手法は、人流の自己相関性に基づいたモデル化を行うため、説明変数を現在の人流、目的変数を未来の人流として予測器を構成する。そのため、説明変数に現れる現在人流の変化に基づいて、将来発生する混雑の予測が可能になると考えられるが、その度合いはせいぜい数時間先に限定される。これは、現時点での人流には、数時間後の混雑の予兆は現れても、それ以降の、例えば 2

日後に発生する混雑の予兆は現れないためである。

これに対し、二つ目の手法では、天候や時間帯、ユーザの来訪予定といった外的な要因を説明変数として、人流とその混雑を早期予報するための枠組みが提案されている。特に、安納ら [26] は、イベントの主催者が事前に公示する情報を元に、混雑寿命の 1 週間前段階での予測問題に取り組んだ。安納らは、スポーツの試合や展示会といったイベントの種類ごとに、来訪者数推移の似通ったパターンを見出した。その共通点を記述するため、楽音合成分野で根幹的技術である ADSR エンベロープ^{*2}を応用し、人口密度の「増加」「維持」「減少」といった状態遷移を外的要因から予測し、その予測結果を用いて来訪者数を復元する枠組みを提案した。

しかし、安納ら [26] の枠組みでは、状態遷移中の人数の変化量までは分からない (図 1 上段)。そのため、人数変化の予測誤差が時間方向に伝わり、混雑の継続時間の補足が不安定になりやすい。また、状態遷移の予測誤差が来訪者数の推定に伝播しやすい。例えば、本来「増加」状態である時間帯を「減少」と誤分類すると、増加傾向にあるはずの来訪者数を本来より少なく見積もってしまう。

¹ 東京工業大学 情報理工学系 情報工学系 Department of Computer Science, School of Computing, Tokyo Institute of Technology

² LINE ヤフー株式会社 LY Corporation

^{*1} https://en.wikipedia.org/wiki/Seoul_Halloween_crowd_crush

^{*2} アコースティック楽器の音に現れる形状的特徴を統一的に制御するパラメータ。ADSR は音の立ち上がり (Attack), 一次減衰 (Decay), 維持 (Sustain), 余韻 (Release) の頭文字であり、各フェーズの時間長を変化させることで、任意のアコースティック楽器音の合成が可能。

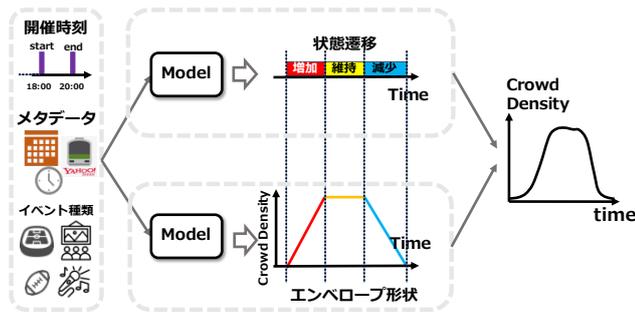


図 1 安納ら [26] が用いた状態表現によるアプローチ (上段) と、提案手法の形状表現によるアプローチ (下段) の違い。

そこで我々は、混雑から抽出した ADSR エンベロープの形状に着目してみる (図 1 下段)。形状表現は、状態表現が反映していた各時間帯における増減傾向に加えて、訪問者数の概算値や混雑全体の時間長を反映している。そのため、この形状的特性を精緻に学習・予測することができれば、混雑全体の傾向捕捉が容易になる。残された課題は、この形状的特性の精緻な予測をどのように実現するか、である。

本研究では、状態表現の代わりに、エンベロープの形状表現を学習する枠組みを提案する。形状的特性の精緻な学習・予測を実現するため、我々は、文章・画像などのコンテンツを現実レベルで合成可能な生成 AI を活用する。具体的には、生成 AI の一種である敵対的生成ネットワーク (Generative Adversarial Nets; GAN) [6] を活用し、混雑から抽出した形状を、イベント時間長や乗換検索に基づく訪問予定者数といった条件に基づいて、GAN の変種である Pix2Pix [10] を用いてモデル化する。

本研究の貢献は以下である。

- 状態表現による混雑寿命予測の課題の解消のため、形状表現の学習・予測による混雑寿命予測を提案する。
- 形状を精緻にモデル化するため、イベント情報を考慮して活動人口のエンベロープ形状を抽出し、敵対的生成ネットワーク (GAN) を用いて形状を学習・描画する枠組みを提案する。
- 実データと人工データで、既存の混雑予測手法と比較して、形状を直接描き出す提案手法が、より小さな誤差で混雑を予測可能であることを示す。

2. 関連研究

人流モデリング。 GPS 位置履歴を搭載したモバイルデバイスの普及により、人流のモデリングが近年盛んに行われている。未来の人流予測に焦点を当てた代表的な手法として、人流の自己相関を利用したモデリング [12], [13] と、外的な要因を説明変数としたモデリング [8], [19], [21], [22] が存在する。前者は、混雑を含め精緻な予測が可能であるが、人流が自己相関を持つ数時間先の予測に限定される。後者は、人流の自己相関の代わりに外的な要因を説明変数

とすることで、より間遠の未来も予測可能となる。本研究では、イベント情報といった外的な要因を説明変数として、混雑の早期の予測を可能とするモデリングを行う。

イベント下における人流モデリング。 イベント下の人流に焦点を当てた研究は、そのイベントの発生が事前にわかるもの (スポーツの試合や花火大会など、主催者が開催を事前公示するもの) [1], [2], [17] と、事前にわからないもの (災害など) [11], [13], [23], [24] とに分類され、本研究は前者に分類される。未来に起こるイベント下の混雑のモデル化は従来、乗換検索アプリの検索履歴をユーザの訪問予定として用いる予測手法が存在した [1], [2], [17]。一方、乗換検索から取得した訪問予定は、イベント下におけるユーザの滞在時間までは反映しておらず、混雑の開始はわかっても終了までは予測できなかった。

また、発生が事前に知らされたイベントから、人々が帰宅をするタイミングをモデル化した研究も存在する [14]。この手法では、人々のイベント下の滞在を生存、帰宅を死亡と捉え、生存解析を適用している。前述の乗換検索ベースの手法と異なり、こちらの手法を用いることで混雑の終了を捉えることが可能と考えられる。しかしながら、生存解析ベースの本手法は、混雑発生後の人流観測を前提としており、イベント当日の混雑開始後にしかモデルを適用することができない。

一方、混雑寿命予測に取り組んだ先駆的研究として、安納ら [26] のものがある。彼らは ADSR エンベロープを活用した状態遷移に基づく手法が存在する。状態遷移を活用することで様々なイベントの混雑動態が予測可能となり、本研究ではこの手法をベースラインとする。一方、1 章でも述べたように状態遷移からは訪問者数の変化量まではわからず、予測が不安定になりやすい。

GAN とその応用。 GAN [6] は、変分オートエンコーダ [15] や拡散モデル [9] などと同様に、潜在変数モデルと呼ばれるクラスの一つである。その特徴は、データの生成を担う Generator と、生成されたデータと実データを見分ける Discriminator を競い合わせながら行う敵対的学習の規範である。GAN には広範な応用先が存在し、こちらのサーベイ [7] が詳しい。

GAN を人流分野に適用した研究としては COVID-GAN [3] があり、COVID-19 流行下における感染者数や政府施策などの外的要因が、人流に与える影響を GAN を用いてモデル化している。一方、本手法の Generator は人流の空間分布に対して最適化され、時間方向の依存性については考慮しない。そのため、外因にイベントの開催情報を与えても、イベント前後の人流の時間変化を捉えることはできない。これに対し我々の提案手法の利点は、イベント種類ごとにパターン化した人数変化をモデル化するため、種々のイベントに対して、混雑の発生、継続、終了を精緻に予測可能にすることである。

3. 問題設定と変数定義

l を混雑予報を行う対象地域, d を日付を表す変数とし, 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. $\tau = dT + t$ を, データの開始時点から数えた時間帯数を表す時間ステップ変数として導入する. 日付 d , および時間帯 t において, 地域 l において観測される携帯端末位置履歴のカウント数を群衆密度と定義し, $y_{l,\tau}$ で表す.

早期の混雑寿命予報の実現のため, 乗換検索案内サービスの検索履歴より得られる人々の行動予定を用いる. 人々の行動予定は, ある目的地 l に対して, 到着予定を時間ステップ τ 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 $x_{l,\tau|d'}$ で表す. イベントの1週間前の時点で訪問者数を予測する場合, $d' = d - 7$ となる検索クエリ数を用いる. 以降は表記上の簡単のため, $x_{l,\tau|d-7}$ を単に $x_{l,\tau}$ と表す.

また, スポーツの試合, 展示会といったイベント種類を表す変数を s_d とする. さらにイベントの長さを表す指標として, 主催者によって事前に公示されるイベントの開催時刻情報を用いる. $e_{l,\tau} \in \{0, 1\}$ を, 地域 l において時間ステップ τ の間のイベント開催の有無を表す変数とし, 開催中であれば 1, そうでなければ 0 を取る変数とする.

最後に, 時間ステップ τ における時刻 t の特徴量を $t \in \mathbb{R}^T$, 日付 d における平休日や曜日といったコンテキストによる特徴量を $c_d \in \mathbb{R}^C$ とする.

本研究では, 乗換検索数 $x_{l,\tau}$, イベントの種類 s_d , イベントの開催時刻情報 $e_{l,\tau}$, 時刻特徴量 t , コンテキスト特徴量 c_d に基づいて, 群衆密度 $y_{l,\tau}$ を予測する問題に取り組む.

4. 提案手法: Generative Envelope Depiction

本研究では, ADSR エンベロープを応用した既存手法 [26] の課題の解決に取り組む. その課題とは, 群衆密度の「増加」「減少」といった状態表現を用いるため (図 1 上部), 訪問者数増減の有無までは分かっても, 具体的な人数変化までは分からないというものである.

この課題の対処のため, 形状表現を用いた群衆密度変化の学習規範である, Generative Envelope Depiction (GED) を提案する (図 1 下部). 提案手法では, まず混雑時の群衆密度の推移から, エンベロープ形状の抽出を行う (4.1 章). 次に, 抽出された形状の精緻な描画を実現するために, GAN を用いた形状の学習と描画を行う (4.2 章). 最後に, GAN により描画された形状から, 実際に群衆密度動態を復元するための差分予測器を導入する (4.3 章).

4.1 混雑動態からのエンベロープ形状の抽出

本コンポーネントは, イベント時の群衆密度推移からエンベロープ形状の抽出を自動で行い, GAN に基づく形状描画モデルを学習するための教師データの作成を担うものに位置付けられる.

安納ら [26] は, イベント種類ごとに, 群衆密度の増減パターンに共通点があることを見出している. それは,

- **ASR タイプ: スポーツの試合・コンサート:** 典型的にスタジアムなどで行われるこのイベントは, イベント開催中は観客がその場所に留まりやすく, 群衆密度は横ばいに推移する. 一方, イベント開始前と終了後に人の集まりが観測され, 群衆密度は「増加」「維持」「減少」のパターンで推移する (図 4(a), 図 4(b)).
- **ASAR タイプ: 展示会:** コミックマーケット*3や東京モーターショー*4などが本イベントの典型例であり, (1) イベントの開始前に待機列が発生し, 群衆密度増加が観測される, (2) イベント開始後に再び群衆密度が増加する, (3) イベント終了前にも関わらず, 群衆密度が減少する, という特徴がある. その結果, 全体として群衆密度は「増加 (一段階目)」「維持」「増加 (二段階目)」「減少」と推移する (図 4(c), 図 4(d)).

既存手法では, これらのイベント種ごとの特徴, 乗換検索数, 及びイベントの開催時刻情報を用いて, 各時間ステップ τ の群衆密度に対して「増加 (Attack; A)」「維持 (Sustain; S)」「減少 (Release; R)」「非混雑 (Not crowded; N)」の 4 ラベルにセグメンテーションする [26]. 一方, この状態ラベルのみでは, A または R ラベルの際, 各時間ステップ τ において活動人口がどれだけ変化するのが分からない.

我々はこのセグメンテーション結果を拡張し, エンベロープ形状の抽出を行う. 各時間ステップにおけるエンベロープ形状の値を $\psi_{l,\tau}$ とする. 既存手法 [26] によって抽

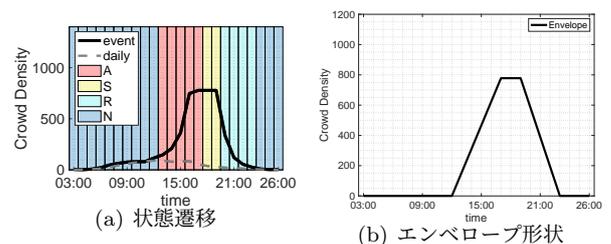


図 2 ラグビー W 杯の状態遷移・形状の抽出結果

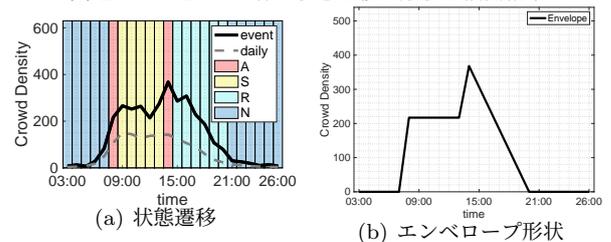


図 3 東京モーターショーの状態遷移・形状の抽出結果

*3 <https://www.comiket.co.jp/>

*4 <https://www.tokyo-motorshow.com/>

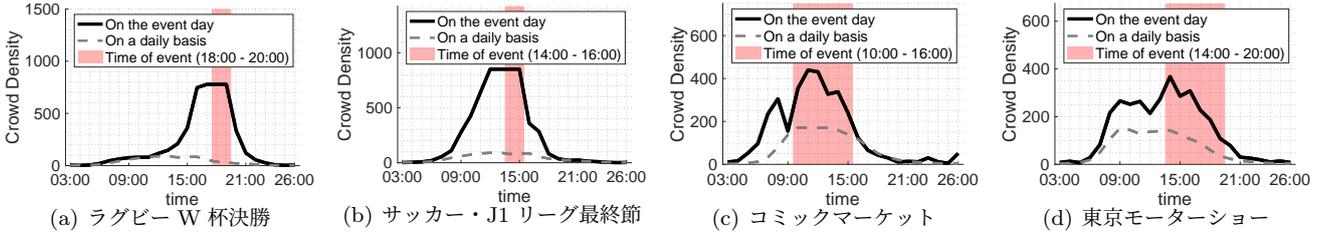


図 4 各イベントの開催時刻と開催日におけるイベント会場周辺の群衆密度の推移。

出された A~N の状態ラベルに対し、我々は以下の仮定の元で形状を規定する。

ASR タイプ: スポーツの試合・コンサート

- A 状態: 直後にくる S 状態開始時点の群衆密度を $y_{i,\tau}^{\text{sust.}}$ とし, 0 から $y_{i,\tau}^{\text{sust.}}$ に線形に増加させる. すなわち, $\psi_{l,\tau} = (\tau / (\tau_S - \tau_A)) y_{i,\tau}^{\text{sust.}}$ である. ここで, τ_A, τ_S はそれぞれ A, S 状態開始時の時間ステップとする.
- S 状態: $y_{i,\tau}^{\text{sust.}}$ のまま横ばいに変化させる. すなわち $\psi_{l,\tau} = y_{i,\tau}^{\text{sust.}}$ である.
- R 状態: $y_{i,\tau}^{\text{sust.}}$ から 0 に向かって線形に減少させる. すなわち $\psi_{l,\tau} = ((\tau - \tau_N) / (\tau_R - \tau_N)) y_{i,\tau}^{\text{sust.}}$, ここで τ_R, τ_N はそれぞれ R, N 状態開始時の時間ステップとする.

ASAR タイプ: 展示会

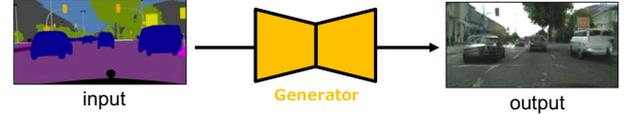
- A 状態 1 回目: 直後にくる S 状態開始時点の群衆密度を $y_{i,\tau}^{\text{sust.}}$ とし, 0 から $y_{i,\tau}^{\text{sust.}}$ に線形に増加させる. すなわち, $\psi_{l,\tau} = ((\tau - \tau_A) / (\tau_S - \tau_A)) y_{i,\tau}^{\text{sust.}}$ である.
- S 状態: $y_{i,\tau}^{\text{sust.}}$ のまま横ばいに変化させる. すなわち $\psi_{l,\tau} = y_{i,\tau}^{\text{sust.}}$ である.
- A 状態 2 回目: $y_{i,\tau}^{\text{sust.}}$ から, 群衆密度のピーク値 $y_{i,\tau}^{\text{peak}}$ に向かって, 線形に増加させる. すなわち $\psi_{l,\tau} = y_{i,\tau}^{\text{sust.}} + ((\tau - \tau_A) / (\tau_{\text{peak}} - \tau_A)) y_{i,\tau}^{\text{peak}}$ である.
- R 状態: $y_{i,\tau}^{\text{peak}}$ から 0 に向かって線形に減少させる. すなわち $\psi_{l,\tau} = ((\tau - \tau_N) / (\tau_R - \tau_N)) y_{i,\tau}^{\text{peak}}$ である.

またいずれのタイプでも, N 状態においては $\psi_{l,\tau} = 0$ とする. 図 2 には, タイプ 1 に属するイベントであるラグビー W 杯, 及び図 3 には, タイプ 2 に属するイベントである東京モーターショーにおける, 既存手法 [26] により抽出された状態遷移, 及び提案手法により抽出されたエンベロープ形状を示す. 次節では, 抽出された形状を描画の対象とした, GAN モデルの学習について述べる.

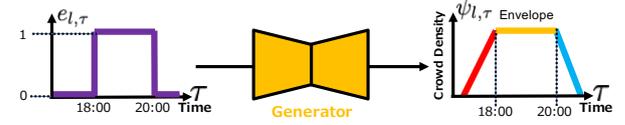
4.2 敵対的学習による形状描画

前節で述べた抽出された形状は, 状態遷移表現にはない, 各状態における群衆密度の変化量が反映されている. 我々は, この形状表現の精緻なモデル化のために, GAN の一種で, 画像変換タスクにおいて成功を収めている Pix2Pix [10] に着目する.

Pix2Pix は図 5(a) のように, あるドメインの画像 (例:



(a) Pix2Pix [10] を用いた画像変換



(b) 提案手法による形状描画

図 5 Pix2Pix を用いた形状描画と, Pix2Pix を適用した提案手法の概要.

「道路」や「車両」などのセマンティクスを, ピクセルごとにラベル付した画像から, 別ドメインの画像 (例: 実際に道路を移した風景画像) に変換する. モデルに入力された画像のピクセル値 (ラベル情報を保持) は, 畳み込みニューラルネットワークを介して周辺のピクセル値との相関が考慮されながら, 風景写真上の新たなピクセル値へ変換される. 我々の問題は, 一つの時間ステップ τ をピクセルと見ることで, イベント開催情報 $e_{l,\tau}$ や乗換検索数 $x_{l,\tau}$ を, 前後の時間ステップとの相関を加味しながら, エンベロープ形状 $\psi_{l,\tau}$ に変換するタスクと捉えることができる (図 5(b)).

具体的には, イベント開催時刻情報の時系列 $e_{l,\tau} = \{e_{l,\tau}\}_{\tau=\tau_A}^{\tau_S}$ ($p > 0$) の列, 乗換検索数時系列 $x_{l,\tau} = \{x_{l,\tau}\}_{\tau=\tau_A}^{\tau_S}$ から, エンベロープ形状の時系列 $\psi_{l,\tau} = \{\psi_{l,\tau}\}_{\tau=\tau_A}^{\tau_S}$ への変換問題を, Pix2Pix を用いて解く.

モデルは, $[e_{l,\tau}, x_{l,\tau}] \in \mathbb{R}^{p \times 2}$ から $\psi_{l,\tau} \in \mathbb{R}^p$ への変換を行う Generator $G: \mathbb{R}^{p \times 2} \rightarrow \mathbb{R}^p$ と, G により生成された形状の予測値 $\hat{\psi}_{l,\tau} = G(e_{l,\tau}, x_{l,\tau})$ と真値 $\psi_{l,\tau}$ との判別を行う Discriminator D により構成される. 学習のための損失関数は, 敵対的学習損失 [6] $\mathcal{L}_{\text{GAN}}(G, D)$ と L1 再構成損失 $\mathcal{L}_{\text{L1}}(G)$ から構成され, それぞれ以下のように表される.

$$\mathcal{L}_{\text{GAN}}(G, D) = \mathbb{E}[\log D(\psi_{l,\tau})] + \mathbb{E}[\log(1 - D(G(e_{l,\tau}, x_{l,\tau})))], \quad (1)$$

$$\mathcal{L}_{\text{L1}}(G) = \mathbb{E}[\|\psi_{l,\tau} - G(e_{l,\tau}, x_{l,\tau})\|_1]. \quad (2)$$

最適な Generator G^* は, 以下の最適化問題を解くことにより得られる.

$$G^* = \arg \min_G \max_D (\mathcal{L}_{\text{GAN}}(G, D) + \lambda \mathcal{L}_{\text{L1}}(G)), \quad (3)$$

ここで λ は、L1 損失の学習への寄与度を調整するハイパーパラメータである。 G 及び D には、それぞれ U-Net [20] と PatchGAN [10] を用いる。ただし、Pix2Pix を時間軸のみを持つ 1 次元データに適用するため、我々は 2 次元畳み込み層の代わりに、1 次元畳み込み層 [16] を用いる。また我々は、イベントの種類 s_d を G に与える代わりに、 s_d ごとにデータを分けて Pix2Pix モデルの学習を行い、イベントの種類ごとの形状パターンの学習を行った。

4.3 描画された形状からの訪問者数復元

G^* により描画されたエンベロープ形状から、実際の訪問者数を復元するための差分予測器を定式化する。本コンポーネントは、差分 $\delta_{l,\tau} := y_{l,\tau} - \psi_{l,\tau}$ をモデル化するために使用し、時刻特徴量 \mathbf{t} とコンテキスト特徴量 \mathbf{c}_d を定式化する。

$$\hat{\delta}_{l,\tau} = \mathbf{c}_d^\top \mathbf{W} \mathbf{t}, \quad (4)$$

ここで $\hat{\delta}_{l,\tau}$ は差分の予測値を表す。また $\mathbf{W} \in \mathbb{R}^{C \times T}$ は学習パラメータであり、最小二乗誤差の最小化により学習を行う。最終的な群衆密度の予測値 $\hat{y}_{l,\tau}$ は、 $\hat{y}_{l,\tau} = \hat{\psi}_{l,\tau} + \hat{\delta}_{l,\tau}$ により求められる。ここで、 $\hat{\psi}_{l,\tau}$ は G^* による形状予測値である。

5. 性能評価実験

混雑寿命予測における、提案手法の性能を評価するため、人口データ及び実イベントデータを用いた性能評価実験を行った。

5.1 比較手法

本実験では、提案手法を以下のベースラインと比較した。
Event-aware Historical Average (EHA): 学習データにおけるイベントが発生した日における活動人口推移の平均値。
BPReg [21]: 外的要因を用いた活動人口予測手法である双線形ポアソン回帰。外的要因として、先行研究でも用いられたコンテキスト特徴量、時刻特徴量に加え、行動予定特徴量、イベント開催時刻特徴量を用いる。
CityOutlook [2]: 自己回帰モデルを用いて、乗換検索ベースの行動予定情報から活動人口への割戻を行う手法である。
CHRoNICLE [26]: 状態遷移を用いて混雑寿命予測を行う手法であり、混雑寿命予測問題の State-of-the-art 手法と位置付けられる。

データの処理・解析には、Apach Spark [25] の MapReduce フレームワークを用いた。また、LSTM および提案手法における深層学習モデルの実装には、Apach Spark 上で

動作をする分散型深層学習フレームワークである、BigDL 1.0 [4] を用いた。

5.2 実験設定と性能評価指標

モデルに与えるコンテキスト情報として、曜日情報、平休日情報、祝日情報を用いた。曜日情報を $\mathbf{c}_d^{(1)} \in \mathbb{R}^7$ 、平休日情報を $\mathbf{c}_d^{(2)} \in \mathbb{R}^2$ 、祝日情報を $\mathbf{c}_d^{(3)} \in \mathbb{R}^2$ とした時、コンテキスト特徴量を $\mathbf{c}_d \mathbf{c}_d^{(1)} \otimes \mathbf{c}_d^{(2)} \otimes \mathbf{c}_d^{(3)} \in \mathbb{R}^{28}$ として与える。時間特徴量は、先行研究 [21] に従い、 $\mathcal{N}(\cdot | \tau, \sigma^2)$ を、平均パラメータ τ 、分散パラメータ σ^2 により規定される正規分布とし、時間特徴量を $\mathbf{t} = \{t_j | t_j = \mathcal{N}(j | \tau, \sigma^2), j = 1, \dots, T\}$ と与える。

また、予測モデルの学習には、早期終了を用いた。この時、テストデータ以外のデータを 9 対 1 に分割し、前者を学習データ、後者を評価データとし、評価データに対する損失値が 10 エポック間改善されなくなるまで学習データに対しモデルのパラメータ更新を行う。その後、学習データと評価データを双方用いて、10 エポックの固定回学習を行い、モデルの収束とする。

性能評価指標には、各状態ごとの群衆密度の絶対予測誤差 (Mean Absolute Error; MAE)、混雑開始時刻の予測誤差 (Mean Absolute Starting Time Error; MASTE)、混雑終了時刻の予測誤差 (Mean Absolute Ending Time Error; MAETE) を用いる。ここで、混雑の開始・終了時刻は、既存研究 [26] と同様に、尤度比検定 [18] による検出法を用いる。

5.3 人工データを用いた性能評価実験

実験データの生成方法。人工データにおける活動人口 y_τ は、日常的に観測される活動人口 \bar{y}_τ と、イベント下における活動人口の増加分 Δy_τ の和により定義する。ここで、 $\hat{\mathbf{c}}_d \in \mathbb{R}^C$ を一様分布 $\text{Uni}(0, 1)$ 、 $\hat{\mathbf{W}} \in \mathbb{R}^{C \times T}$ を一様分布 $\text{Uni}(0, 0.04)$ から生成する乱数により定義し、 $\bar{y}_\tau \sim \text{Pois}(\cdot | \lambda_\tau)$ 、 $\ln \lambda_\tau = \hat{\mathbf{c}}_d^\top \hat{\mathbf{W}} \mathbf{t}$ として、日常の活動人口を生成する。イベント下における活動人口の増加分 Δy_τ は、正規分布 $\mathcal{N}(\cdot | \mu_\tau, \mu_\tau/10)$ から生成した乱数とする。ここで μ_τ は混雑状態と共に変化する変数である。イベントの開始時刻 t_{start} は午前 8 時から午後 2 時の間、イベントの継続時間は 5~10 時間の間から、それぞれ 1 時間単位でランダムに選択する。

ASR タイプにおいて、 μ_τ は次のように変化させる。 μ_τ は S 状態において $\mu_{\text{peak}} \in [250, 500]$ に設定し、A 状態の時間長を、 $\mu_{\text{peak}} < 300$ なら 2 時間、 $300 \leq \mu_{\text{peak}} < 400$ なら 3 時間、それ以外なら 4 時間に設定した。N 状態の間は $\mu_\tau = 0$ とし、A 状態の間は 0 から μ_{peak} まで線形に増加させ、R 状態の間は μ_{peak} から 0 まで線形に減少させた。ASAR タイプでは、2 回目の A 状態の開始時刻を t_{start} とし、その長さを $\mu_{\text{peak}} < 300$ なら 1 時間、 $300 \leq \mu_{\text{peak}} < 400$

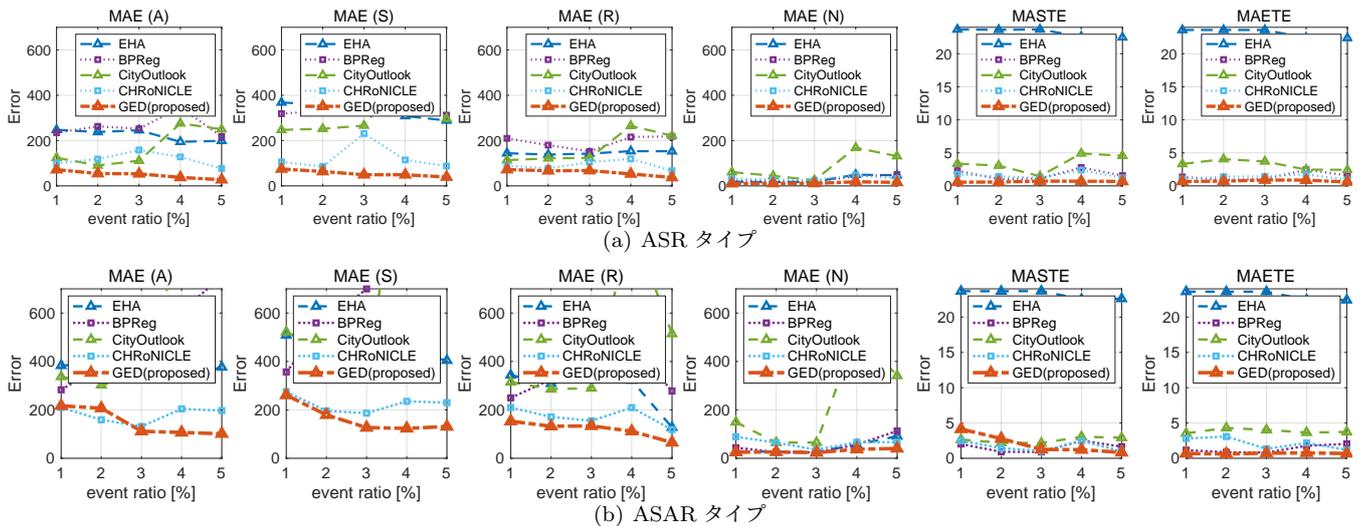


図 6 人工データにおける、データセット内のイベント数率の違いによる予報性能の変化。

なら 2 時間、それ以外なら 3 時間に設定した。2 回目の A 状態の前に、時間長が同じ 1 回目の A 状態と、時間長を {1, 2, 3} からランダムに選択した S 状態を設定した。また、S 状態時において、 $\mu_r = \mu_{peak}/2$ とした。

モデルの学習時は、学習データを 120 日分生成し、その中でデータに含まれるイベント日の比率を 1 ~ 5% の間で変化させた。またテストデータはイベント開催を模倣した日のみのデータとし、120 日分生成した。

人工データにおける実験結果。 図 6 に、学習データセット内のイベント比率を変化させた場合の人工データによる実験結果を示す。ASR タイプにおいては、提案手法は学習データ内のイベント数の比率に寄らず、常に最高性能を達成している (図 6(a))。また、ASAR タイプに対しては、S, R, R 状態の MAE, 及び MAETE において、学習データ内のイベント数の比率に寄らず最高性能を達成した。一方、MAE, MASTE では、イベント比率が 1~2 で比較手法に性能で劣り、3% で比較手法と同等、4% 以上で比較手法を超える性能を達成できた。この結果は、学習データ内のイベント数が極端に少ない場合、予報性能が悪化する示唆を与えるものである。

5.4 実イベントデータを用いた性能評価実験

実験データ。 本実験では、実世界におけるイベント開催下での混雑のライフサイクル早期予報性能を評価するため、実際の GPS 位置履歴データ、乗り換え案内アプリ、イベントの開催時刻情報データを使用した。今回解析の対象と

したのは、神奈川県横浜市に所在する日産スタジアム周辺の群衆混雑動態である。

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^{*5}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。データは事前にユーザ情報を取り除くことで匿名化し、さらに解析対象地域におけるログ数を 60 分間隔でカウントし、統計情報化したものである。そのため、最終的なデータは、日付、時刻、ログのカウント数からなるデータであり、個人を特定できるものではない。解析対象地域は、表 1 に示すイベント会場周辺の、500 × 500 m 四方の正方形エリアとした。また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^{*6}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。日産スタジアム周辺に存在する新横浜駅、小机駅における検索履歴のみを抽出し、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、行動予定情報として利用した。そのため、こちらも統計情報化されたものであり、個人を特定できる情報は含まれない。

またイベントの開催時刻情報として、日産スタジアムのイベントカレンダー^{*7}をスクレイピングすることで収集した、イベントの開催時刻情報を利用した。データは日付、

表 1 実データにおける解析対象イベント。

イベント	会場	開催日時	イベント種
ラグビー W 杯決勝	日産スタジアム	11/2	ASR
ラグビー W 杯準決勝	日産スタジアム	10/27	ASR
J1 リーグ最終節	日産スタジアム	12/27	ASR
J1 リーグ開幕節	日産スタジアム	2/23	ASR
コミックマーケット	東京ビッグサイト	12/28-31	ASAR
東京モーターショー	東京ビッグサイト	10/25-11/4	ASAR

*5 <http://emg.yahoo.co.jp/>

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

*7 <https://www.nissan-stadium.jp/calendar/>

表 2 実イベントデータにおける予報性能の評価.

Model	ASR タイプ (real)						ASAR タイプ (real)					
	MAE				MASTE	MAETE	MAE				MASTE	MAETE
	A	S	R	(N)			A	S	R	(N)		
EHA	251.9	535.3	117.1	41.2	4.0	2.2	173.4	94.8	113.3	16.1	3.0	3.0
BPre ^g * [21]	157.0	456.2	120.8	22.5	2.0	1.2	119.1	242.6	85.6	12.6	0.0 (*)	0.5 (*)
CityOutlook [2]	308.1	611.9	88.4	30.7	5.3	4.8	183.6	199.3	132.1	24.6	1.0	4.0
CHRoNICLE [26]	173.8	294.0	61.2	33.3	3.4	0.8	117.2	70.7	94.8	8.8	1.0	2.7
GED	137.9	128.9	115.7	12.0	0.8	1.2	101.9	124.3	125.5	12.0	1.3	1.6

イベント名, 開始・終了時刻からなり, 今回の実験では各日付における開催時刻, 終了時刻を特徴量として使用した.

実イベントデータにおける実験結果. 表 2 に, 実イベントデータによる実験結果を示す. 各イベント種類ごとに性能評価値の平均を取り掲載している. 混雑の予報開始・終了時刻は尤度比検定を用いて行うため, 群衆密度の予測値が下振れを起こした場合に検出されない. 混雑の検出率に関しては表 3 に記載した. 検出率が 100% に満たないケースは, 少なくとも 1 つ以上のイベントで混雑検出に失敗したことを意味する. このようなイベント事例は MASTE, MAETE の計算時, 除外した.

まず ASR タイプに対しては, 表 2 より R 状態の MAE, MAETE を除いて提案手法が最高性能を達成し, 表 3 より, 混雑発生を漏れなく検出できている. 特に, S 状態の群衆密度の予報性能では, 既存手法の CHRoNICLE を 56.2% 上回る予報性能を達成した. 一方, 混雑の終焉期における予報である, R 状態における MAE 及び MASTE では, CHRoNICLE に劣る性能となった. こちらは, 5.5 章の可視化の章で再度考察する.

ASAR タイプに対しては, 表 2 より A 状態における MAE を除いて, 既存手法に予報性能で劣る形となった. 一方, MASTE, MAETE にて最高性能性能を達成した BPre^g* (表内 (*)) は, 表 3 より混雑の検出率が 100% に満たない. BPre^g* は混雑の見落としを起しているこの事実から, BPre^g* は予報が不安定であると言える. BPre^g* や, 同様に混雑検出率が 100% を下回る EHA を除けば, 提案手法は既存手法の CHRoNICLE に対し, A 状態の予報性能, 及び MAETE で優位な性能となり, S, R 状態の MAE と MASTE で劣る性能となり, 同等程度の性能を達成できていると言える. こちらも, 5.5 章の可視化の章で再度考察する.

表 3 実イベントデータにおける混雑の検出率.

Model	混雑検出率 [%]	
	ASR	ASAR
EHA	100.0	66.7
BPre ^g * [21]	100.0	66.7
CityOutlook [2]	80.0	100.0
CHRoNICLE [26]	100.0	100.0
GED	100.0	100.0

5.5 混雑寿命予報の可視化

図 7 には, 各イベント下における混雑寿命予測の可視化例を示す. 上段 (a)~(c) が ASR タイプのイベントであり, 下段 (d)~(f) が ASAR タイプのイベント事例である.

図 7(a), 図 7(b) からわかるように, 提案手法 GED は既存手法と比較して, 群衆密度の真値により近い値をトレースするような予測が可能になったことがわかる. 特に, 形状を学習する提案手法の優位性が現れた点として, 図 7(a) では, 状態遷移を用いる CHRoNICLE の予測値が下振れしてしまっていることに対して, 提案手法ではこの下振れが抑えられている. 一方, これら 2 つのケースでは, 提案手法の R 状態の減少が真値と比較して急速であり, このことが前節でも述べた提案手法の混雑終焉期の予報性能劣化に繋がっている. 図 7(c) は, COVID-19 の流行が首都圏で顕著となってきた 2020 年 2 月 23 日のイベント事例であり, 観客数は (a),(b) のケースと比較しても少ない. 学習データには, 流行下におけるイベント事例は含まれず, 提案手法に限らずこのケースの予報には失敗している.

図 7 (d)~(f) の ASAR ケースでは, 提案手法はいずれのイベントにおいても, ASAR イベントの特性である二段階の群衆密度増加や, イベント終了前の人数減少などを捉えることはできている. 一方, 形状をとらえた上での人数スケールの予測が不安定であることから, それを反映可能な特徴量の追加が重要と示唆される. 今後は, イベントの種類や形状の情報に加えて, そのイベントに対する参加者の熱意や評価などをモデルの入力に組み込むことで, こういった人数スケールや, さらには他のイベント種類 (花火大会やお祭り) などの混雑も予報可能になると考えられる.

6. 結論

本研究では, イベントに際して発生する混雑を, その始まりから終わりまで予報する問題である, 混雑寿命予報に取り組んだ. 混雑寿命予報に取り組んだ既存の研究は, 群衆密度の増減パターンがイベントの種類ごとに共通することを見出し, 増減を状態遷移としてモデル化する手法を提案していた. しかし, 状態遷移からは増減の傾向はわかっても実際の訪問者数の変化量まではわからず, 状態遷移を用いる既存手法では実際の訪問者数の予測が不安定になる恐れがあった. そこで本研究では, 状態遷移の代わりに

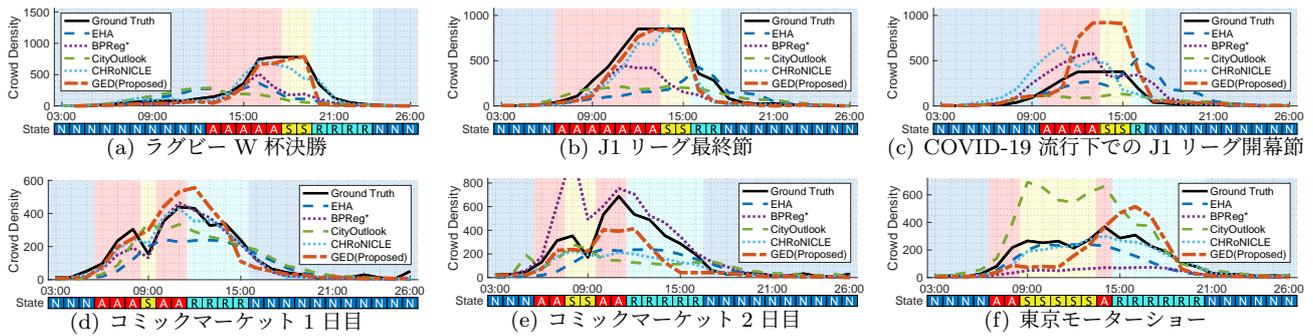


図 7 予報の可視化を伴う混雑寿命予測ケーススタディ.

訪問者数推移の大まかな傾向を形状としてモデル化する Generative Envelope Depiction(GED) を提案した。GED では、形状の精緻なモデル化のために、データの生成と画像の描画において成功を収めている、Pix2Pix を用いた。人工データ及び実イベントデータを用いた評価実験を行い、(1) 人工データにおいて、2つのイベント種類に対して既存手法を上回る性能を達成し、(2) 実イベントデータを用いた実験において、ASR タイプのイベントに対して MAE で最大 56.2%の性能改善が見られ、(3)ASAR タイプのイベントに対しても形状の予測の優位性を示す実験結果が得られた。一方、形状の捕捉が可能であったケースでも、訪問者数スケールの予測が不安定であるケースも存在した。この課題に対しては、来訪者数のスケールを反映可能な追加の特徴量、例えばイベントに対するユーザの反応や評価などを取り込むことが考えられる。

参考文献

- [1] S. Anno et al. Supervised-cityprophet: Towards accurate anomalous crowd prediction. In *Proc. of SIGSPATIAL*, 2020.
- [2] S. Anno et al. Cityoutlook: Early crowd dynamics forecast towards irregular events detection with synthetically unbiased regression. In *Proc. of SIGSPATIAL*, 2021.
- [3] H. Bao et al. Covid-gan+: Estimating human mobility responses to covid-19 through spatio-temporal generative adversarial networks with enhanced features. *ACM Trans. Intell. Syst. Technol.*, 13(2), 2022.
- [4] J. J. Dai et al. Bigdl: A distributed deep learning framework for big data. In *Proc. of the ACM Symposium on Cloud Computing*, 2019.
- [5] Z. Fan et al. Citymomentum: An online approach for crowd behavior prediction at a citywide level. In *Proc. of UbiComp*, 2015.
- [6] I. Goodfellow et al. Generative adversarial nets. *Proc. of NIPS*, 2014.
- [7] J. Gui et al. A review on generative adversarial networks: Algorithms, theory, and applications. *IEEE TKDE*, 35(4):3313–3332, 2021.
- [8] Y. Hayakawa et al. Simultaneous multiple poi population pattern analysis system with hdp mixture regression. In *Proc. of PAKDD*, 2021.
- [9] J. Ho et al. Denoising diffusion probabilistic models. *Proc. in NeurIPS*, 2020.
- [10] P. Isola et al. Image-to-image translation with condi-

- tional adversarial networks. In *Proc. of CVPR*, 2017.
- [11] R. Jiang et al. Deepurbanevent: A system for predicting citywide crowd dynamics at big events. In *Proc. of SIGKDD*, 2019.
- [12] R. Jiang et al. Dl-traffic: Survey and benchmark of deep learning models for urban traffic prediction. In *Proc. of CIKM*, 2021.
- [13] R. Jiang et al. Learning social meta-knowledge for now-casting human mobility in disaster. In *Proc. of TheWebConf*, 2023.
- [14] A. V. Khezerlou et al. Dilsa+: Predicting urban dispersal events through deep survival analysis with enhanced urban features. *ACM Trans. Intell. Syst. Technol.*, 12(4):1–25, 2021.
- [15] D. P. Kingma and M. Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.
- [16] S. Kiranyaz et al. 1d convolutional neural networks and applications: A survey. *Mechanical systems and signal processing*, 2021.
- [17] T. Konishi et al. Cityprophet: City-scale irregularity prediction using transit app logs. In *Proc. of UbiComp*, 2016.
- [18] D. B. Neill. Expectation-based scan statistics for monitoring spatial time series data. *International Journal of Forecasting*, 2009.
- [19] M. Okawa et al. Online traffic flow prediction using convolved bilinear poisson regression. In *Proc. of MDM*, 2017.
- [20] O. Ronneberger et al. U-net: Convolutional networks for biomedical image segmentation. In *Proc. of MICCAI*, 2015.
- [21] M. Shimosaka et al. Forecasting urban dynamics with mobility logs by bilinear poisson regression. In *Proc. of UbiComp*, 2015.
- [22] M. Shimosaka et al. Spatiality preservable factored poisson regression for large-scale fine-grained gps-based population analysis. In *Proc. of AAAI*, 2019.
- [23] A. Sudo et al. Particle filter for real-time human mobility prediction following unprecedented disaster. In *Proc. of SIGSPATIAL*, 2016.
- [24] T. Yabe et al. A framework for evacuation hotspot detection after large scale disasters using location data from smartphones: Case study of kumamoto earthquake. In *Proc. of SIGSPACIAL*, 2016.
- [25] M. Zaharia et al. Spark: Cluster computing with working sets. *HotCloud*, 2010.
- [26] 安納爽響 et al. 混雑の生起・継続・終了を考慮した状態認識型 RNN に基づく 早期群衆混雑予報. *研究報告ユビキタスコンピューティングシステム (UBI)*, 2023(16):1–8, 2023.