

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

安納 爽響¹ 坪内 孝太² 下坂 正倫¹

概要: 本研究では、社会イベントに起因して発生する異常な群衆混雑を、その発生のみならず、いつ終わるのかまでを早期予報する枠組みを提案する。群衆混雑の早期予報は、快適なモビリティやタクシー需要予測など社会的ニーズが強く重要な話題である。しかし、既存手法は主に群衆混雑の発生時点の予報に焦点を当てている。また公示されるイベント開催時刻情報からの類推も、混雑のような非日常的なパターンの元では、参照する過去データが希少であり困難なタスクとなる。そこで本研究では、群衆混雑を発生から終了までの“ライフサイクル”を捉えるための手法である、Cascading-based Hybrid Recurrent Network Incorporating Crowding Life cyclEs (CHRoNICLE) を提案する。提案手法では、群衆混雑の発生過程を生起期・継続期・終了期といった“状態”に分割する。また、各状態における人口変化の傾向を捕捉するため、各状態の時系列性、時間長や、活動人口の増加を明示的に学習し、活動人口の回帰器にカスケードする。人工データ及び実イベントデータを用いた性能評価実験により、提案手法が既存手法と比較して、混雑を発生から終了までをより正確に捉えられることを示す。

1. 序論

群衆混雑は、社会イベント（例えばコンサートやスポーツの試合、お祭りなど）に起因して、ある地域に多くの人々が密集する混雑のことであり、一度発生すると渋滞や交通機関の乱れに繋がる他、群衆雪崩発生リスクもある社会的脅威である。これまで群衆混雑は、人流の時空間依存性に基づいた人流シミュレーション手法 [8], [13], [17] の発展により広く解析が行われてきた。しかし、人流が大幅に変化するイベント下では、シミュレーション可能な人の動きはせいぜい数時間先までである。そのため、イベント参加見送りなどの群衆混雑への対応に十分な期間を前提とした、数日～1週間先といった未来の群衆混雑予報を行うことはできなかった。

本研究では、1週間先の**早期**群衆混雑予報に焦点を当てる。近年、人々の POI^{*1} 訪問予定情報に基づいた将来の混雑の早期予報 [2], [12], [20] が研究されている。しかしこれらの手法は、訪問予定先における人々の滞在については考慮しないため、群衆混雑がいつ始まるのかはわかっても、いつ終わるのかまではわからない。そこで、殆どのイベン

トは開催時刻がスケジュールされているため、イベントの終了時刻を混雑の終了時刻に転用する方法が考えられる。しかし、事前に公示されたイベントの開催時刻通りに混雑が発生するとは限らない。例えば 2019 年ラグビーワールドカップ決勝では、スタジアムの満席が見込まれていたことから、試合開始前には既に多くの観戦者が会場に集まっていた。その上、イベントの余韻に浸る参加者が、即座に帰宅するとは限らない。また、混雑を誘発する社会イベントの希少性ゆえに、開催時刻からずれた時間帯の活動人口の変化を、過去の統計的パターンから類推することも容易ではない。

そこで本研究では、群衆混雑をその始まりから終わりまで、すなわち混雑の“ライフサイクル”を予報するための新しい問題を考える。これは、イベント帰宅時の混雑回避やタクシーの配車需要の予測などにつながる重要な枠組みである。通常、イベントに起因して発生する混雑は、イベントの開始の数時間前から人々が集まり始め人口密度のあるピークに達し、イベント開催中この人口密度が持続した後、イベント終了後に人々が徐々に帰宅していくという典型的な時系列変化を持つ。

我々は、混雑のライフサイクルは、**図 1** に示すような**生起期・継続期・終了期**の3つの状態フェーズに分割できるという仮説を持った。各状態は、活動人口の増加・維持・減少といった傾向が保持されることを想定し、状態の開始

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 Yahoo! JAPAN 研究所 Yahoo Japan Corporation

*1 Point of Interest

時刻、持続時間、状態開始時の人口密度を特性値として持つ。生起期はユーザの訪問予定の増加、継続期はイベントの開催時刻にそれぞれ連動するため、直感的には、混雑下の活動人口を直接予測するよりも、各状態の変化や、その特性値を予測する方が簡単なタスクとなる。そこで、活動人口を直接予測する代わりに、混雑の状態変化を明示的に学習し、予測された状態の特性値に基づいて、活動人口の波形を合成するアプローチを検討する。

そこで本研究では、群衆混雑のライフサイクルにおける時系列変化を明示的に認識する、マルチタスク学習を用いたハイブリッド再帰的ネットワーク: Cascading-based Hybrid Recurrent Network Incorporating Crowding Life cyclEs (CHRoNICLE) を提案する。本手法ではまず、混雑下の活動人口の時系列変化を捉えるための再帰的ニューラルネットワーク (RNN) [3] を構築する。また、混雑時の活動人口の時系列を、非混雑時の活動人口との統計的有意差やイベント開催時刻を考慮しながら、生起期・継続期・終了期といった状態にセグメンテーションする。さらに、各状態における人口変化の傾向を捕捉するため、各状態の時系列性、時間長や、活動人口の増加に関する明示的な学習タスクを導入し、予測結果を活動人口の回帰器にカスケードするよう、RNN を拡張する。こうすることで、データセット内の混雑データが少ない状況下でも、高い精度での混雑ライフサイクル認識が実現できる。

本研究の貢献は以下である。

- 群衆混雑を、その始まりから終わりまで早期に予報するため、混雑の活動人口を生起・継続・終了というフェーズに分割し、そのフェーズの時系列変化を明示的に認識するハイブリッド RNN, CHRoNICLE を提案する。
- 混雑時の人口増加・維持・減少を模倣した人工データ実験により提案手法の有効性を確認する。また実イベントデータを用いた実験により、提案手法がベースラインと比較して1週間前の段階で混雑の発生から終了までをより正確に捉えられることを示す。

関連研究

本研究の新規性と提案手法の動機を明確にするために、(1) 群衆混雑予報 と (2) 楽音の合成及び認識のための技術に関する既存研究について論じる。

群衆混雑予報. 近年の位置情報サービス普及により、都市における群衆の動き、すなわち人流の把握が可能となった。これまで人流は、その時空間依存性をどうモデル化するかに焦点が当てられ、盛んに研究されてきた [8], [9], [13], [17]。これらの手法は、時系列上の自己相関性を仮定したシミュレーションを行うため、数時間先の予報において高性能な予報が可能であるが、1週間先に発生する混雑は、自己相関性が侵される問題設定であるため、予報することが難しい。

この課題を解消するため、近年群衆混雑の早期予報が研

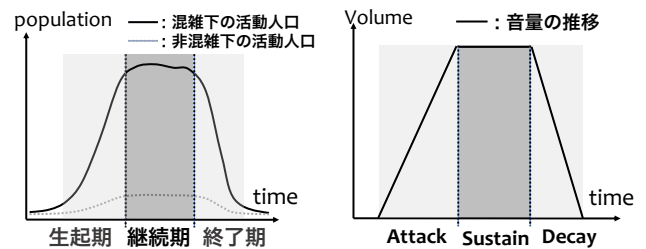


図 1 混雑のライフサイクル。 図 2 ADSR エンベロープ。

究され始めている。将来起こる混雑を類推するためのデータとして、既存研究 [2], [12] ではユーザの将来の移動予定を使用している。Konishi ら [12] は、乗換検索には混雑時の検索量増加が顕著に現れる性質があることを示した。鉄道の乗換案内アプリの検索履歴を用いて、将来どの駅に何人の人々が来訪するかを予見するための手法を提案した。Anno ら [2] らは、Konishi らの手法を乗換検索履歴に基づく訪問予測の不確実性と、異常イベントの希少性による予報の不安定性の課題を解消するための手法を構築した。しかし、これらの手法はユーザのイベント会場への訪問予定のみを使用しているため、訪問後にユーザがどの程度滞在し、離散していくかまでは予報できない。これらの混雑予報の既存研究に対し、本研究では混雑をその始まりから終わりまで予報する問題を考える。

楽音の合成及び認識. 本研究での提案は、従来の楽音合成技術に同期づけられたものである。楽音合成は、電子楽器に利用される楽器音の合成技術であり、最近のポピュラー音楽には欠かせないものである。典型的には、合成器は複数のエンベロープ (音量の時間変化) を組み合わせて音波を合成するが、その代表的なエンベロープの一つに ADSR エンベロープが存在する。ADSR は、音量の増幅期 (Attack), 初期減衰期 (Decay), 継続期 (Sustain), 減衰期 (Release) の頭文字をとったものであり、各状態の時間長やピーク時の音量を変化させることで、さまざまな楽器の音を合成することができる。図 2 には、初期減衰期 (Decay) が存在しない ADSR エンベロープを示す。

近年、この ADSR 法は楽音の合成のみならず、機械学習を用いた楽音認識性能の向上に貢献している。楽音認識技術の代表的な応用先に、オーディオ録音からシンボリックな楽譜を自動で抽出する自動採譜タスクがある [5], [6], [10], [11], [16]。Hawthorne ら [5], [6] はピアノ音の自動採譜において、楽音の立ち上がり (Attack) を明示的に認識する手法を提案し、採譜性能を向上させた。これは、ピアノ音の Attack 状態は極めて短く打楽器的であるため、音波変化を検出しやすいことが背景にある。さらに Kelz ら [10] は、楽音の Attack 状態のみならず、Decay, Sustain, Release と音が存在しない状態を用いることで、採譜性能のさらなる向上を図った。本研究の提案手法は、これらの楽音の合成及び認識技術からヒントを得た。

2. 早期群衆混雑予報の問題設定

2.1 問題設定

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し, 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タ数を $y_\tau^{(l)}$ と定義し, その地域における活動人口と呼ぶ. 本研究では, 混雑下における活動人口の時系列を予測することで, 群衆混雑予報を行う.

また早期予報の実現のため, 先行研究 [1], [2], [12] でも用いられていた, 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 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 $x_{d,t|d'}^{(l)}$ で表す. 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. 到着予定日 d における, 目的地 l の行動予定特徴量 $\mathbf{x}_d^{(l)} \in \mathbb{R}^{p_w T}$ を, l および d における T 個の各到着予定時刻の検索数を並べたベクトルとして表現し, $\mathbf{x}_d^{(l)} = \{x_{d,j|d-i}^{(l)} | i = p_d, p_d + 1, \dots, p_d + p_w, j = 1, \dots, T\}$ と定義する. 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.

2.2 群衆混雑予報のベースライン手法:

RNN による予測モデル

群衆混雑予報のベースライン手法として, 再起的ニューラルネットワーク (RNN) に基づいた予測モデルを用いる. 先行研究 [15] と同様に本研究でも, 活動人口がポアソン分布 Pois に従うと仮定する.

$$\mathcal{P}(y_\tau^{(l)}) = \text{Pois}(y_\tau^{(l)} | \lambda_\tau^{(l)}) = \frac{\lambda_\tau^{(l) y_\tau^{(l)}} \exp(-\lambda_\tau^{(l)})}{y_\tau^{(l)}!}, \quad (1)$$

ここで $\lambda_\tau^{(l)} > 0$ はポアソン分布の平均パラメータである. 本手法では, 曜日や平休日, 対象地域における行動予定情報やイベント開催時刻情報といった, 日付 d に関する入力特徴量群と, 時刻 t に関する特徴量を用いて, 系列 $\lambda_\tau^{(l)}, \dots, \lambda_{\tau+p}^{(l)}$ をモデル化する.

具体的には, 時間ステップ τ に関する入力特徴量群の埋め込みベクトル化を行い, こちらを RNN に渡す. 埋め込みベクトル $\mathbf{h}_\tau^{(0)}$ は, 日付 d に関する特徴量と, 時刻 t に関する特徴量との間の組み合わせを考慮し定式化する.

$$\mathbf{h}_\tau^{(0)} = [h_{\tau,1}, \dots, h_{\tau,H}]^\top \in \mathbb{R}^H, \quad (2)$$

$$h_{\tau,i}^{(0)} = \phi([\mathbf{c}_d^\top, \mathbf{x}_d^{(l)\top}, \mathbf{e}_d^{(l)\top}] \mathbf{W}_i^{(0)} \mathbf{t}), i = 1, \dots, H, \quad (3)$$

ここで, $\mathbf{c}_d \in \mathbb{R}^C$ は日付 d における, 曜日や平休日といった条件を表すコンテキスト特徴量, $\mathbf{t} \in \mathbb{R}^T$ は時間帯 t に関する時間特徴量である. $\mathbf{e}_d^{(l)} \in \mathbb{E}$ は, 日付 d における地域 l でのイベント開催時刻をエンコードした特徴量である. スポーツの試合などのイベント開催下では人々の離散は発生せず, 混雑状態が継続すると考えられることから, イベント開催時刻を入力に使用することは, 混雑状態が継続する時間長の予測に有効であると考えられる. また $\mathbf{W}_i^{(0)} \in \mathbb{R}^{(C+p_w T+E) \times T}$ は行列形式の学習パラメータ, ϕ は非線形変換を施すための活性化関数である. このような定式化を行うことで, 同一の時間帯においても, 曜日の異なりやイベント開催・非開催の違いにより, 活動人口が変化することが表現可能となる [15].

RNN は, 上述の埋め込みベクトル $\mathbf{h}_\tau^{(0)}$ を入力として, 時間ステップ τ から $\tau + p$ までの活動人口系列を出力する One-to-many 形式を採用する.

$$\ln \lambda_\tau^{(l)}, \dots, \ln \lambda_{\tau+p}^{(l)} := \text{RNN}(\mathbf{h}_\tau^{(0)}; \Theta_{\text{RNN}}), \quad (4)$$

ここで, Θ_{RNN} は RNN の学習パラメータである. 本研究では $\text{RNN}(\cdot; \Theta_{\text{RNN}})$ を, 時系列上の短期依存性及び長期依存性の両方の学習に効果的である, 長・短期記憶 (Long Short-Term Memory; LSTM) [7] モジュールを用いて構築する. また, 本手法はポアソン分布の負の対数尤度の最小化を用いて学習する.

$$\mathcal{L}^{(l)}(\Theta_{\text{RNN}}) = - \sum_d \sum_t \sum_{i=0}^p \ln \text{Pois}(y_{\tau+i}^{(l)} | \lambda_{\tau+i}^{(l)}). \quad (5)$$

しかし, RNN を安直に適用したベースライン手法では, 混雑の多段予報を正確に行うことができない. これは, 混雑の開始時刻・終了時刻が正確に捉えられないことが原因である. 本研究では, 混雑の持続長を示す指標として, イベント開催時刻情報をモデルの入力に使用する. しかしそこからわかるのは, 一度混雑がピークに達した後, 継続的に混雑している (と考えられる) 時間帯のみである. 群衆混雑のように, 発生自体が稀でありデータパターンが少ないイベントのモデル化においては, RNN を適用しただけでは, いつから人が来訪し始めるのか, いつ人々の離散が終わり混雑が完全に消滅するのかを, データから十分に学習することができない.

3. 提案手法: CHRoNICLE

この章では, 提案手法: Cascading-based Hybrid Recurrent Network Incorporating Crowding Life cycleEs (CHRoNICLE) を述べる. 図 3 は, 提案手法の概略図を示す. まず 3.1 章にて, 群衆混雑下の活動人口を, 生起期・継続期・終了期・非混雑期という状態に分割する, 混雑状

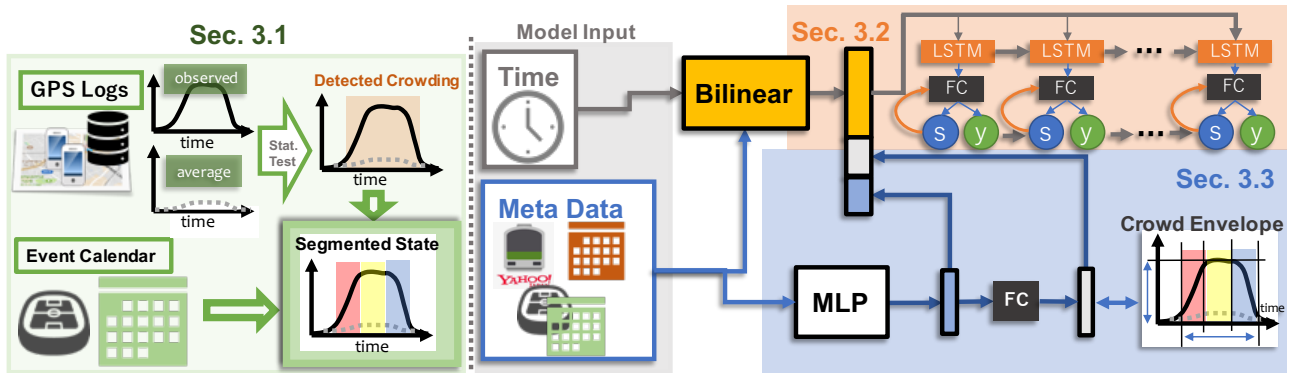


図 3 提案手法 CHRoNICLE の概略図.

態セグメンテーションについて述べる. 次に 3.2 章にて, 各時間ステップの活動人口回帰と混雑状態に関する系列ラベリング問題を同時に解く, 状態認識ハイブリッド RNN を導入する. さらに 3.3 章にて, 各状態の開始時刻, 時間長, 開始時の活動人口といった, 活動人口波形の形状的特性を決定する状態特性値の学習タスクについて述べる.

3.1 混雑状態セグメンテーション

1 章で述べたように, 我々の仮説は混雑が生起期・継続期・終了期という状態に分割可能なライフサイクルを持つというものである. そこで活動人口を, 混雑の生起期・継続期・終了期に加えて, 非混雑期という 4 つの状態ラベルを付与する, 混雑状態セグメンテーションを行う. 我々は, まず活動人口を非混雑期と混雑期に分割し, 混雑期を生起・継続・終了期に分割するアプローチをとる. 活動人口の非混雑期と混雑期は, イベント会場ごとにアクセスの良さやイベント規模が異なるため, 一貫した境界を設けることが難しい. そこで, 対象地域ごとに日常的に観測される活動人口と, イベント開催日に観測される活動人口とを比較し, 活動人口の増分の統計的有意性を元に混雑期を検出する.

日常的な活動人口を, 過去の同じ曜日・時間帯に観測された活動人口平均値 $\bar{y}_\tau^{(l)}$ にて与える. 我々の興味は, 観測された活動人口が $y_\tau^{(l)}$ が, 過去の平均値 $\bar{y}_\tau^{(l)}$ よりも優位に大きいかという点である. そこで以下のような帰無仮説 $H_0 : y_\tau^{(l)} \sim \text{Pois}(\cdot|\bar{y}_\tau^{(l)})$, 対立仮説 $H_1 : y_\tau^{(l)} \sim \text{Pois}(\cdot|\bar{z}_\tau^{(l)})$ ($\bar{z}_\tau^{(l)} > \bar{y}_\tau^{(l)}$) を設定し, ポアソン分布の期待値に基づく尤度比検定 [14], [19] にて検定を行う. 検定統計量である対数尤度比 $\text{LLR}^{(l)}(\tau)$ は以下のよう

$$\text{LLR}^{(l)}(\tau) = \begin{cases} y_\tau^{(l)} \log \frac{y_\tau^{(l)}}{\bar{y}_\tau^{(l)}} + (\bar{y}_\tau^{(l)} - y_\tau^{(l)}) & \text{if } y_\tau^{(l)} \geq \bar{y}_\tau^{(l)}, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

$\text{LLR}^{(l)}(\tau) > 0$ かつ $\text{LLR}^{(l)}(\tau)$ が有意水準 α で統計的に優位である場合, 時間ステップ τ を活動人口の混雑期と定

義する. 通常, 尤度比の統計的優位性の判定には, 高コストなモンテカルロ法の実行が必要であるが, Zhou ら [19] が示したように, ポアソン分布 $\text{Pois}(\cdot|y_\tau^{(l)})$ の累積分布関数 $\Pr(X < y_\tau^{(l)})$ に対し, $1 - \Pr(X < y_\tau^{(l)}) \geq \alpha$ である時, $\text{LLR}^{(l)}(\tau)$ が有意水準 α で統計的に優位といえる. 本研究ではこの事実を利用した.

次に, 混雑期の活動人口を生起期, 継続期, 終了期の 3 状態に分割する. 我々の仮説は, イベント開催下では混雑が継続して発生するというものである. そこで本研究では, イベント開始時刻から終了時刻の一つ前の時間帯までを, 混雑の継続期, その前後を生起期, 終了期とする. 図 4, 5 は, 2019 年 11 月 2 日にラグビー W 杯が行われた, 日産スタジアムにおける混雑状態セグメンテーション結果を表す. このイベントは, イベントの開始時刻が午後 6 時, 終了時刻が午後 8 時であった. それぞれの上図は活動人口の推移を示し, 下図は各時間帯における統計的検定での混雑検出結果 (T は統計的に優位, F は統計的に非優位), ライフサイクル状態のセグメンテーション結果 (A は生起期, S は継続期, R は終了期, N は非混雑期に対応) を示す.

3.2 状態認識ハイブリッド RNN

次に, 2 章で述べた RNN モデルを拡張し, 状態認識ハイブリッド RNN を導入する. 混雑はそのライフサイクル上で, 図 5 のような非混雑期 (N) → 生起期 (A) → 継続期 (S) → 終了期 (R) → 非混雑期 (N) という時系列性を持つ. 1 章で述べたように, この時系列変化はユーザーの行動予定やイベント開催時刻情報に連動して変化する. そこで, モデルがイベント開催時刻などを入力として受け取った際, まず現在の時刻がどの状態に属するのかを学習・予測し, この予測結果に基づいて活動人口を回帰 (合成) する.

対象地域 l における, 各時間ステップ τ に対する混雑状態ラベルを $s_\tau^{(l)} \in \mathcal{S} := \{N, A, S, R\}$ とする. 状態認識は, 状態ラベルに対する系列ラベリング問題として扱え, 本研究ではこの系列ラベリングを RNN で解く. そのため, ベースラインの RNN モデルを以下のように拡張する.

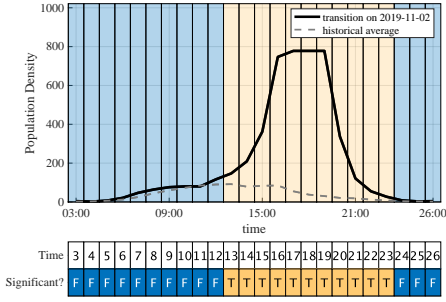


図 4 混雑・非混雑期の検出結果.

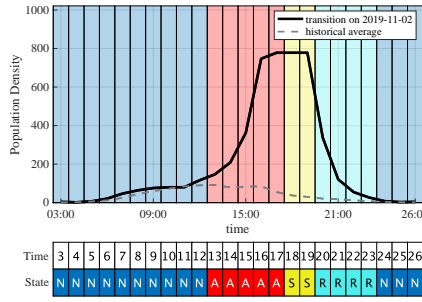


図 5 セグメンテーション結果.

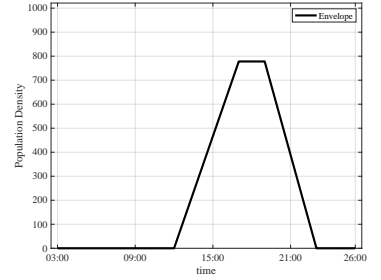


図 6 クラウドエンベロープの例.

$$\mathbf{h}_{\tau}^{\text{RNN}}, \dots, \mathbf{h}_{\tau+p}^{\text{RNN}} := \text{RNN}(\mathbf{h}_{\tau}^{(0)}; \Theta_{\text{RNN}}), \quad (7)$$

$$\hat{s}_{\tau}^{(l)} := \mathbf{f}_{\tau}(\mathbf{h}_{\tau}^{\text{RNN}}; \mathbf{W}_{f,\tau}) = \text{Softmax}(\mathbf{W}_{f,\tau} \mathbf{h}_{\tau}^{\text{RNN}}), \quad (8)$$

$$\ln \lambda_{\tau}^{(l)} := g_{\tau}(\mathbf{h}_{\tau}^{\text{RNN}}, \hat{s}_{\tau}^{(l)}; \mathbf{w}_{g,\tau}) = \mathbf{w}_{g,\tau}^{\top} [\mathbf{h}_{\tau}^{\text{RNN}}; \hat{s}_{\tau}^{(l)}], \quad (9)$$

ここで、 \mathbf{f}_{τ} は状態ラベルの分類モデルであり、 $\mathbf{h}_{\tau} \in \mathbb{R}^H$ 、 $\mathbf{W}_{f,\tau} \in \mathbb{R}^{4 \times H}$ である。また、 g_{τ} は活動人口回帰モデルであり、 $\mathbf{w}_{g,\tau} \in \mathbb{R}^{4+H}$ である。

3.3 人口波形状学習タスク

図 2 に示すように、ADSR エンベロープは各状態の時間長、振幅の最大値が決まれば、エンベロープの形状が決定する。直感的には、直接活動人口を予測するよりも、ADSR エンベロープのような混雑時の活動人口変化の”大まかな形状”を推定した上で、活動人口の波形を合成するほうが簡単なタスクと考えられる。そこで、混雑下の活動人口の形状をクラウドエンベロープと呼び、その時間長といった特性値を学習するタスクを導入する。

まず、3.1 章で述べたセグメンテーション結果より、各状態の特性値を抽出する。 $t_{d,1}^{(l)}, t_{d,2}^{(l)}, t_{d,3}^{(l)}$ を、対象地域 l 、日付 d における各状態の開始時間帯とする。また、 $\Delta t_{d,1}^{(l)}, \Delta t_{d,2}^{(l)}, \Delta t_{d,3}^{(l)}$ を、生起期、継続期、終了期の各状態の持続時間長とし、 $c_1^{(l)}, c_2^{(l)}, c_3^{(l)}$ を各状態開始時の混雑に係る活動人口密度とする。ここで、生起期は混雑の開始時点のため $c_1^{(l)} = 0$ 、継続期と終了期の開始時点での人口密度は等しく、 $c_2^{(l)} = c_3^{(l)} = y_{\tau}^{(l)}, \tau' = dT + t_{d,2}^{(l)}$ とする。図 6 には、2019 年 11 月 2 日のラグビー W 杯における、クラウドエンベロープを示す。

これらの状態特性値は全て非負の離散値をとるため、それぞれ異なるポアソン分布に従うと仮定する。各特性値に共通の多層パーセプトロン (MLP) を使用し、特性値間で特徴空間を共有しながら学習を行う。この時 MLP の入力には、日付 d のみに関連した入力特徴量であるコンテキスト特徴量 \mathbf{c}_d 、イベント開催時刻特徴量 $\mathbf{e}_d^{(l)}$ 、行動予定特徴量 $\mathbf{x}_d^{(l)}$ を使用する。

$$\mathbf{h}_d^{\text{MLP}} = \text{MLP}(\mathbf{c}_d, \mathbf{e}_d^{(l)}, \mathbf{x}_d^{(l)}; \Theta_{\text{MLP}}), \quad (10)$$

$$\ln \pi(\alpha)_d = \mathbf{w}_{\alpha}^{\top} \mathbf{h}_d^{\text{MLP}}, \quad (11)$$

ここで、 $\pi(\alpha)_d$ は、 α が従うポアソン分布のパラメータであり、 $\alpha \in \{t_{d,1}^{(l)}, \Delta t_{d,1}^{(l)}, \Delta t_{d,2}^{(l)}, \Delta t_{d,3}^{(l)}, \Delta t_{d,1}^{(l)} + \Delta t_{d,2}^{(l)} + \Delta t_{d,3}^{(l)}, c_1^{(l)}, c_2^{(l)}, c_3^{(l)}\}$ 、 $\mathbf{h}_d^{\text{MLP}} \in \mathbb{R}^H$ 、 $\mathbf{w}_{\alpha} \in \mathbb{R}^H$ である。

以上を元に、最終的な学習パラメータは Θ_{RNN} 、 Θ_{MLP} 、 $\mathbf{W}_{f,\tau}$ 、 $\mathbf{w}_{g,\tau}$ 、 \mathbf{w}_{α} となる。学習に使用する損失関数は、ポアソン分布の負の対数尤度と、交差エントロピー関数に基づいて以下のように定式化される。

$$\begin{aligned} \mathcal{L}^{(l)}(\Theta_{\text{RNN}}, \Theta_{\text{MLP}}, \mathbf{W}_{f,\tau}, \mathbf{w}_{g,\tau}, \mathbf{w}_{\alpha}) \\ = - \sum_d \left[\sum_t \sum_{i=0}^p \left(\ln \text{Pois}(y_{\tau+i}^{(l)} | \lambda_{\tau+i}^{(l)}) + \sum_{j=0}^{s_{\tau}} s_{\tau} \log \hat{s}_{\tau,j}^{(l)} \right) \right. \\ \left. + \sum_{\alpha_d} \ln \text{Pois}(\alpha_d | \pi(\alpha)_d) \right]. \quad (12) \end{aligned}$$

4. 性能評価実験

本研究の目的である、群衆混雑のライフサイクル早期予測における提案手法の性能を評価するため、人工データおよび実データを用いた実験を行った。実験ではまず人工データを用いた性能評価実験により、(1) データセット内のイベント数による予測性能の変化、(2) イベントの開催時間情報特徴量の設計方法による予測性能の変化という点について、評価を行った。さらに実データを用いた実験により、(3) 実イベントの混雑予測における既存手法からの性能改善について評価を行った。

4.1 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。

- **Historical Average:** 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* [15]:** 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]:** 早期群衆混雑予測の State-of-the-art 手法。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。

- **GCPR*** [20]: 早期群衆混雑予報におけるもう一つの State-of-the-art 手法。駅とイベント会場の地理的な関係を考慮した混雑予報を行なっている。こちらも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, 行動予定特徴量に加え, イベントの開催時刻特徴量を用いる。
- **LSTM**: 2章で述べたベースライン RNN モデル。

データの処理・解析には, Apache Spark [18] 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, Apache Spark 上で動作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。

4.2 性能評価指標

本実験では提案手法を, 混雑の多段予報に即した以下の性能評価指標を用いて評価する。(1) 混雑下における人口密度の予報性能の評価のために, 活動人口予測の平均絶対誤差 (Mean Absolute Error; **MAE**), (2) 混雑の開始時刻の予報誤差を評価するために, 平均絶対開始時刻誤差 (Mean Absolute Starting Time Error; **MASTE**), (3) 混雑の終了時刻の予報誤差を評価するために, 平均絶対終了時刻誤差 (Mean Absolute Ending Time Error; **MAETE**) を用いる。それぞれ, $MAE = \sum_{\tau \in \mathcal{T}_d} |y_\tau^{(l)} - \hat{y}_\tau^{(l)}|$, $MASTE = \sum_{d \in \mathcal{T}_d} |t_{st}^{(l)} - \hat{t}_{st}^{(l)}|$, $MAETE = \sum_{d \in \mathcal{T}_d} |t_{en}^{(l)} - \hat{t}_{en}^{(l)}|$ と定義する。ここで \mathcal{T}_d はイベントが開催された日の時間ステップを表す。また, $t_{st}^{(l)}, t_{en}^{(l)}$ はそれぞれ混雑の開始時刻, 終了時刻を表し, 今回は, 混雑下における活動人口の最大値の 20% に達した時刻を混雑開始時刻, その後 20% を下回った時刻を混雑終了時刻と定義した。

4.3 実験設定

モデルに与えるコンテキスト情報として, 曜日情報, 平日情報, 祝日情報を用いた。曜日情報を $c_d^{(1)} \in \mathbb{R}^7$, 平日情報を $c_d^{(2)} \in \mathbb{R}^2$, 祝日情報を $c_d^{(3)} \in \mathbb{R}^2$ とした時, コンテキスト特徴量を $c_d c_d^{(1)} \otimes c_d^{(2)} \otimes c_d^{(3)} \in \mathbb{R}^{28}$ として与える。時間特徴量は, 先行研究 [15] 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 $t = \{t_j | t_j = \mathcal{N}(j | \tau, \sigma^2), j = 1, \dots, T\}$ と与える。

また, 人工データ実験ではイベント開催時刻特徴量 e_d の設計方法に関しても調査を行う。ここでは, Event Length(EL), Event Time(ET), Event Duration(ED) の, 3種類のイベント開催時刻特徴量について性能評価を行った。日付 d のイベントの開催時刻が t_{on} 時開始, t_{off} 時終了とした場合を考える。EL 特徴量はイベントの開催時間をスカラー値で与える特徴量とし, $e_d^{EL} = t_{off} - t_{on} \in \mathbb{R}$ である。ET 特徴量は, イベントの開始時刻と終了時

刻をそれぞれ 1-of-K 符号化法で与える特徴量とし, $e_d^{ET} = [e_d^{ET(start)\top}, e_d^{ET(end)\top}]^\top \in \mathbb{R}^{2T}$ とする。ここで, $\mathbb{I}[\cdot]$ を, $[\cdot]$ の中の式が満たされた場合に 1, そうでない場合に 0 を返す指示関数とし, $e_d^{ET(start)\top} = \{\mathbb{I}[t = t_{on}] | t = 0, \dots, T\} \in \mathbb{R}^T$, $e_d^{ET(end)\top} = \{\mathbb{I}[t = t_{off}] | t = 0, \dots, T\} \in \mathbb{R}^T$ である。ED 特徴量は, イベントの開催時間中を 1, それ以外を 0 で与える特徴量とし, $e_d^{ED} = \{\mathbb{I}[t \in \{t_{on}, t_{off}\}] | t = 0, \dots, T\} \in \mathbb{R}^T$ と定義する。

深層モデルの学習には早期終了を用いる。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。

4.4 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, $\hat{c}_d \in \mathbb{R}^C$ を一様分布 $\text{Uni}(0, 1)$, $\hat{W} \in \mathbb{R}^{C \times T}$ を一様分布 $\text{Uni}(0, 0.04)$ から生成する乱数により定義し, $\bar{y}_\tau \sim \text{Pois}(\cdot | \lambda_\tau)$, $\ln \lambda_\tau = \hat{c}_d^\top \hat{W} t$ 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 は混雑状態と共に変化する変数とし, 継続期において 500, それ以外では生起期において 0 ~ 500 で線形に増加し, 終了期において 500 ~ 0 で線形に減少する値とした。また $\nu = 0$ の時は $\Delta y_\tau = 0$ とする。またイベントの開始時刻は混雑継続期の開始時間帯, 終了時刻は混雑終了期の開始時間帯とした。モデルの学習時は, 学習データを 180 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 5, 10, 20% と変化させた。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, 133 日分生成した。

人工データにおける実験結果

表 1 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。NaN は, モデルの活動人口の予測値が, 混雑の条件である最大値の 20% に満たなかったため, そもそも混雑を検出できていないことを表す。提案手法である CHRoNICLE が, 比率 5% 下の MAETE を除いて, 最も高い予測性能を達成した。また, 比較手法の LSTM に性能で劣っている比率 5% の MAETE においても, 提案手法は比較的同程度の性能を達成している。このことから学習データ内のイベントデータ数によらず, 提案手法が混雑のライフサイクルを高性能に予報できていることがわかる。

また, イベント開催時刻特徴量毎の予報性能比較では,

表 1 人工データにおける，データセット内のイベント数・イベント特徴量設計の違いによる予報性能の変化.

イベント比率		5%			10%			20%		
model	$e_d^{(l)}$	MAE	MASTE	MAETE	MAE	MASTE	MAETE	MAE	MASTE	MAETE
Historical Average	N/A	138.16	NaN	NaN	145.07	NaN	NaN	138.06	NaN	NaN
BPre [*] [15]	EL	134.50	NaN	NaN	134.83	NaN	NaN	122.80	NaN	NaN
	ET	156.27	NaN	NaN	249.09	NaN	NaN	140.50	NaN	NaN
	ED	153.99	NaN	NaN	243.46	NaN	NaN	134.08	NaN	NaN
LSTM	EL	85.64	3.89	1.58	70.75	2.31	1.53	57.73	0.97	1.04
	ET	61.66	2.19	1.12	68.50	3.00	0.90	54.77	1.53	1.17
	ED	71.53	4.04	1.56	73.41	3.39	1.39	52.56	2.56	1.14
CHRoNICLE	EL	63.55	1.95	1.68	56.52	0.83	1.06	49.48	0.66	0.95
	ET	59.44	1.25	1.15	46.01	1.07	0.95	41.43	0.84	0.74
	ED	52.53	1.11	1.16	49.10	0.94	0.72	48.37	0.69	0.78

ET, ED 特徴量を用いた場合，全てのイベント比率において比較手法と同程度かそれを上回る性能を達成した．特に，イベントデータ数が少ない比率 5% のケースでは，ED 特徴量を用いた場合に最も安定した性能を発揮した．

4.5 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 *2 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．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 *3 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 *4 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，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，実データによる実験結果を示す．実験では，4.5 章における実験結果を元に，イベントデータ数が少ない状況下で高い性能を発揮した ED 特徴量を用いた．プロサッカーの試合における MAETE を除き，いずれの指標においても提案手法が最高性能を達成した．特に提案手法は，MAE の結果が示す活動人口密度の予測性能改善のみならず，混雑の開始・終了時刻の予測誤差を示す MASTE, MAETE 双方において同時に安定した性能を達成している．このことから，提案手法が実イベント下においても，頑健な混雑ライフサイクル予報を実現できることがわかる．

また，図 7 にはラグビー W 杯決勝における混雑予報の可視化結果を示す．上段は活動人口の推移の予測，下段に

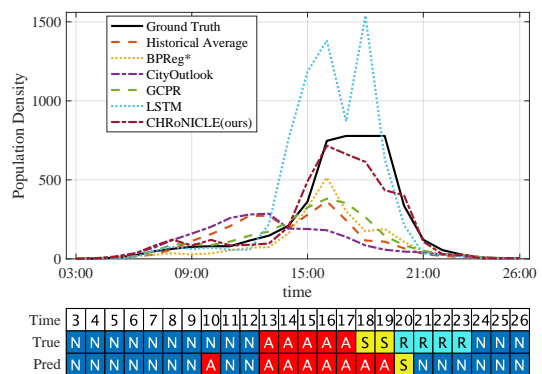


図 7 ラグビー W 杯における予測結果の可視化.

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

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

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

表 2 実データを用いた早期群衆混雑予報の予測誤差.

model	ラグビー W 杯決勝 (2019 年 11 月 2 日)			プロサッカーチームの試合 (2019 年 12 月 7 日)		
	MAE [人]	MASTE [時間]	MAETE [時間]	MAE [人]	MASTE [時間]	MAETE [時間]
Historical Average	134.41	4.0	3.0	175.29	5.0	0.0
BPReg* [15]	109.29	0.0	1.0	135.00	1.0	3.0
CityOutlook [2]	167.20	5.0	4.0	168.66	0.0	1.0
GCPR* [20]	104.16	1.0	2.0	176.37	3.0	2.0
LSTM	144.58	1.0	0.0	152.41	2.0	3.0
CHRoNICLE	47.62	0.0	0.0	94.04	0.0	1.0

は提案手法による混雑状態ラベルの予測結果である。イベントの開催時刻は午後 6 時から 8 時である。上段の予測結果からわかるように、HA や CityOutlook は、予測された混雑の開始時点が、実際の混雑とずれている。また BPReg や GCPR は、混雑の開始時点を検出しているものの、その後の予測人口減少が早く、混雑の終了時刻を捉えられていない。これに対し提案手法では、混雑の開始・継続・終了の捕捉がより正確である。しかし混雑の継続期において、活動人口が減衰を始めており、終了期の時点では真値の半分程度にまで減少している。このことは、下段に示す状態ラベルの予測結果にも表れており、午後 6 時以降の継続期・終了期を誤識別しており、終了期が予測結果から消失している。この状態ラベルの予測精度を向上させることで、さらなる予報性能改善につながると考えられる。

5. 結論

本研究では、混雑のライフサイクル予報という新しい問題を考えた。混雑予報の既存手法は主に群衆混雑の発生時点の予報を行っており、またイベント開催時刻情報からの類推も、混雑のような非日常的なパターンの元では、参照する過去データが希少であり困難なタスクであった。そこで本研究では、混雑下の活動人口を生起・継続・終了というフェーズに分割し、その時系列変化を明示的に捕捉する状態認識型ハイブリッド RNN, CHRoNICLE を提案した。人工データ・実データを用いて提案手法の性能評価実験を行い、実データ上での実験では、既存の混雑予報手法の最も性能が良かった GCPR と比較して、最大で 54.6% の性能改善を実現した。一方、予報の可視化結果は、提案手法のさらなる性能改善の可能性を示唆している。特に、混雑状態系列の生起期・継続期・終了期という一連のサイクルの存在をモデリングに組み込むことで、さらなる性能改善が期待でき、将来課題として考えられる。

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] K. Benidis et al. Deep learning for time series forecasting: Tutorial and literature survey. *ACM Computing Surveys*, 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] C. Hawthorne et al. Onsets and frames: Dual-objective piano transcription. *arXiv preprint arXiv:1710.11153*, 2017.
[6] C. Hawthorne et al. Enabling factorized piano music modeling and generation with the maestro dataset. In *Proc. of ICLR*, 2018.
[7] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 1997.
[8] R. Jiang et al. Deepurbanevent: A system for predicting citywide crowd dynamics at big events. In *Proc. of SIGKDD*, 2019.
[9] R. Jiang et al. Dl-traffic: Survey and benchmark of deep learning models for urban traffic prediction. In *Proc. of CIKM*, 2021.
[10] R. Kelz et al. Deep polyphonic adsr piano note transcription. In *Proc. of ICASSP*, 2019.
[11] J. W. Kim and J. P. Bello. Adversarial learning for improved onsets and frames music transcription. *arXiv preprint arXiv:1906.08512*, 2019.
[12] T. Konishi et al. Cityprophet: City-scale irregularity prediction using transit app logs. In *Proc. of UbiComp*, 2016.
[13] Z. Lin et al. DeepSTN+: Context-aware spatial-temporal neural network for crowd flow prediction in metropolis. In *Proc. of AAAI*, 2019.
[14] D. B. Neill. Expectation-based scan statistics for monitoring spatial time series data. *International Journal of Forecasting*, 2009.
[15] M. Shimosaka et al. Forecasting urban dynamics with mobility logs by bilinear poisson regression. In *Proc. of UbiComp*, 2015.
[16] B. L. Sturm, J. F. Santos, O. Ben-Tal, and I. Korshunova. Music transcription modelling and composition using deep learning. *arXiv preprint arXiv:1604.08723*, 2016.
[17] Z. Wang et al. Event-aware multimodal mobility nowcasting. In *Proc. of AAAI*, 2022.
[18] M. Zaharia, M. Chowdhury, M. J. Franklin, S. Shenker, I. Stoica, et al. Spark: Cluster computing with working sets. *HotCloud*, 2010.
[19] X. Zhou et al. A traffic flow approach to early detection of gathering events. In *Proc. of SIGSPATIAL*, 2016.
[20] 安納爽響 et al. 地域の幾何的關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. *研究報告ユビキタスコンピューティング (UBI)*, 2020.