

# 大規模ユーザの時空間滞在・経路パターンに基づく 早期目的地予測

今井 遼<sup>†1</sup> 坪内 孝太<sup>1</sup> 小西 達也<sup>†2</sup> 下坂 正倫<sup>†1</sup>

**概要:** ユーザの行動予測により、ユーザに適したサービスの提供を行うことが可能になり、行動予測の中でも、ユーザの未来の位置を予測する目的地予測は、様々なアプリケーションへの応用可能性があり、実際に利用されている。事故や遅延の際の迂回路提示のような GIS サービスへの応用を考えると、出発から到着までになるべく早期に目的地を正しく予測する必要がある。一方で、出発から早い段階での予測は利用できる経路情報が少なく精度が低くなる。すなわち、早期性と精度はトレードオフの関係にある。本研究では、我々が最近提案した移動手段によらない早期目的地予測手法を概説し、これまでの研究報告で行っていなかった、既存の経路・滞在情報それぞれのみに基づく予測と、早期目的地予測手法の予測の様子の違いに関する定性的な評価を行うことを目的とする。この定性的実験により、既存手法である経路・滞在情報それぞれのみに基づく予測手法と比較し、早期に目的地が予測できることを確認し、その理由を考察した。

**キーワード:** 目的地予測, 位置情報, 滞在パターン, 経路パターン, 早期予測

## 1. 緒論

スマートデバイスの普及によりユーザの状況分析が容易になってきており、ユーザの状況、すなわちコンテキストに応じたサービスの提供を行う研究やアプリケーションが登場している。その中でも、位置情報は多くのスマートデバイスのセンサで取得可能であり、多くのサービスや研究の対象となっている。Senら [1] は、現在位置や天気、心拍数など、現在のコンテキストに応じた音楽の推薦に関する研究を発表している。Woerndlら [2] は協調フィルタリングに現在のコンテキストを組み込んだ推薦システムを構築し、その実験において POI (Point Of Interest) の推薦システムによる評価を行った。実アプリケーションへの応用例として、myThings<sup>\*1</sup>や IFTTT<sup>\*2</sup>は、GPS を利用し、設定しておいたエリアへの出入りをトリガーとし、プッシュ通知やメール送信などのアクションを行うサービスを提供している。

さらに、現在の状況に加え、ユーザの行動予測を組み込

むことで、提供するサービスを事前に準備し、ユーザの入力の手間を省いたり、ユーザ自身が気付いていない嗜好やルーチンに基づいたサービスを提供したりすることが可能になる。Parateら [3] は、モバイルアプリケーション利用の際の待ち時間に着目し、次にユーザが利用するアプリケーションと、その利用タイミングを予測するアルゴリズムを設計し、ユーザの利用するアプリを事前に準備する研究を発表した。Okoshiら [4] は、モバイルフォンが時と場合を選ばずに通知を発することがユーザの負担となっていることを問題視し、ユーザが通知を受けるべきタイミングを予測する研究を発表した。

ユーザの行動予測の中でも、ユーザの未来の位置情報、すなわち目的地を予測することは重要な課題の一つであり、多くの研究者の関心を引きつけている。目的地を予測することで、事故や電車遅延の際、ユーザの入力を待たずに迂回路を提示することや、目的地に応じた検索ワードや広告の提示、目的地への到着時間に応じた丁度良い長さの記事や動画の推薦などが可能になる。実際に、先読み案内サービス<sup>\*3</sup>は、これまでの走行履歴から、エンジン始動時の曜日・時間帯に応じてユーザの目的地を予測し、「いつもの道」であれば目的地の設定なしに渋滞情報や天候などを通知するサービスを提供している。また、Google Now<sup>\*4</sup>で

<sup>1</sup> ヤフー株式会社 Yahoo! JAPAN 研究所  
Yahoo Japan Corporation, Yahoo! JAPAN Research, Chiyoda, Tokyo, Japan

<sup>†1</sup> 東京工業大学  
Tokyo Institute of Technology, Meguro, Tokyo, Japan

<sup>†2</sup> 東京大学  
University of Tokyo, Bunkyo, Tokyo, Japan

<sup>\*1</sup> <https://mythings.yahoo.co.jp/>

<sup>\*2</sup> <https://ifttt.com/>

<sup>\*3</sup> <http://tconnect.jp/detail/3389602/>

<sup>\*4</sup> <https://www.google.com/intl/ja/landing/now/>

は、ユーザの過去の位置履歴、チェックアウト、カレンダーで繰り返されている予定などから、ユーザに適した情報を提供する。

目的地予測問題において、目的地予測の精度と予測の早期性はトレードオフの関係にある。すなわち、移動の完了度合いが小さいほど予測精度は低くなり、大きいほど高くなるということである。これは、目的地予測において利用できる情報として、直前の滞在地や出発時点、出発日時などの滞在情報と、どのような経路を辿ってきたかの経路情報が利用できるが、移動初期の場合、利用できる経路情報が少なくためである。このトレードオフを解決することは、実アプリケーションへの応用という観点で重要である。目的地予測を利用したアプリケーション例として、先程述べた通り、事故や電車遅延の際の迂回路提示が考えられる。この時、ユーザが事故や遅延が発生した路線に到達してから迂回路を提示しても意味がなく、移動初期からの高精度な目的地の絞り込みが必要となる。また、目が不自由なユーザへの目的地への誘導・経路補正というアプリケーション例も考えられる。このアプリケーションでは早期の予測が出来なければ、ユーザに長い迂回路を歩かせるという負担を強いるおそれがある。

本研究では、目的地予測を簡単な式で定式化するため、これらの情報を、移動経路から得られる経路情報と曜日・時間帯といった出発時点で得られる滞在情報の2種類に分類する。このように利用できる情報を分類すると、既存の目的地手法も経路情報に基づく予測手法 [5], [6], [7] と滞在情報に基づく予測手法 [8], [9], [10], [11], [12] の2種類に分類することが出来る。これらの手法には、前者には移動初期には考慮できる情報が少なく予測精度が低い、後者には移動が進んでも利用できる情報が増えないため精度が上がらないという問題がある。小西ら [13] は目的地予測のトレードオフに着目し、移動初期から利用できる滞在情報と、移動が進むに連れて重要となる経路情報の両方を考慮した予測手法を構築した。しかし、提案された予測手法は、ユーザの移動方法を電車移動と仮定し予測を行うため、その他の移動手段である車や徒歩による移動に対しては予測が困難であるという問題がある。また、筆者ら [14] は小西らのモデルを拡張し、移動手段によらない予測手法を構築したが、一方でその評価は、定量的な評価にとどまっておらず、なぜ性能が向上するのかに関する定性的な評価が行われていない点が問題であった。

本研究の目的は、筆者らが最近提案した早期目的地予測手法 [14] を概説し、既存の報告でなされていなかった、定性的な評価結果を報告し、この手法の性能向上がもたらされる理由を考察することである。

本論文では、2節で、目的地予測の問題設定を行い、常に参照可能な情報である滞在情報と、徐々に利用できる情報が増えていく移動情報のそれぞれ単体を考慮した場合

の問題点を述べ、さらに、早期予測手法である小西らの手法 [13] を紹介し、その問題点を述べる。また、3節では、筆者らの最近の提案手法である、経路・滞在情報の両方を考慮した、移動手段によらない予測手法を概説する。実際に、東京都を中心とした、およそ 4000 km<sup>2</sup> の範囲を日常の活動場所とする、1600 人を超えるユーザの2ヶ月の位置情報データに対して行った、移動段階ごとの定量的な精度比較実験および定性的な予測の様子 of 可視化により、滞在パターンと移動パターンの両方を考慮することで、既存手法と比較してより早期目的地予測に適した予測が可能となることを確認した (4節)。

#### 関連研究

目的地予測の研究には、滞在情報に基づく予測に関する研究と、経路情報に基づく予測に関する研究に加え、滞在・経路情報に関わらずコンテキストを取り入れることができる研究や、滞在・経路情報の両方を考慮した研究が存在する。

**経路情報に基づく研究：**経路情報に基づく研究では、現在の経路から将来の経路を予測し、移動が進むにつれて目的地を絞り込むことができる。経路情報に基づく研究として、運転手の目的地の履歴と、目的地の種類や辿ってきた経路の効率、運転時間などに着目し、尤度最大の目的地を予測する手法 [5] や、ある地点からある地点への遷移確率を元に目的地を予測する手法 [6], [7] が発表されている。これらの研究には、出発直後は考慮できる経路情報が少ないため、精度が低いという共通の問題がある。

**滞在情報に基づく研究：**滞在情報に基づく予測として、決定木を用いた多クラス識別 [8] により、運転手の目的地を予測する手法や、ユーザの滞在情報を特徴量とした多クラス識別に加え、曜日・時間帯ごとの過去のパターンを考慮した予測手法 [9] などがある。特徴量として用いられる滞在情報としては、直前の滞在場所や現在の時間帯、週末か否か、直前の場所に滞在した長さ [10], [11] などが利用されている。また、ユーザの滞在位置の遷移パターンをモデリングすることに着目し、隠れマルコフ手法を用いて滞在情報をモデリングする手法 [12] も考案されている。ユーザ間の共通パターンを利用し少ない学習データで正確な予測をすることに着目した研究も存在し、2項分布を用いてユーザが滞在しているか、外出しているかを予測する研究 [15] や、位置情報への利用のため拡張した階層ディリクレ過程である LocHDP [16] などが発表されている。これらの研究には、移動に伴う目的地予測の補正ができないという欠点が存在する。

様々なコンテキストを取り入れることができる研究として、逆強化学習の特徴量を移動パターンモデリングのために設計することで目的地予測を行う研究が発表されている [17], [18]。これらの研究の問題点は、多くのコンテキストを取り入れることができるが、モデルが複雑になりがち

であったり、一部の移動手段に適したモデル構築が容易でも、汎用的なモデルの構築が難しくなる点である。滞在・経路情報をともに考慮した早期目的地予測に適した手法も発表されており [13]、ここでは、逆強化学習の特徴量を電車移動に適した設計にすることで、経路情報に基づく予測として利用している。しかし、徒歩や車などの予測には適さないという問題があり、汎用的な手法となっていない。また、筆者らはこの問題を解決するため、移動手段によらない早期目的地予測手法を構築している [14]。

## 2. 目的地予測の定式化と既存の予測手法の問題点

筆者らが提案した最新の予測手法の紹介の前に、目的地予測の問題設定と、既存の経路・滞在情報に基づく予測手法、および小西ら [13] の早期目的地予測手法の問題点を述べる。

### 2.1 問題設定

本論文では、目的地予測を、ある目的地から次の目的地までにたどり着くまでに、次の目的地を当てる問題として設定する。まず、目的地の定義を行う。データセットに含まれる位置情報に対し他の位置情報との距離、周囲の点の個数のような類似度に基づきラベリングを行う。この際に得られたラベルの個数を  $L$  個とする。その後、滞在時間の長さに基づき、移動・滞在を示すラベリングを行う。滞在を示すラベル付けがなされた位置データのうち、直前の位置データが移動とラベル付けされているものを目的地と定義する。得られた目的地の集合を  $\mathcal{Y}$  とする。その後、連続する移動の系列を抽出し、それぞれを  $T$  ステップに分割する。この移動の 1 ステップ目から  $t$  ステップ目までを  $t$  ステップ目における経路情報とし、 $l_{1:t}$  と表す。また、曜日や時間帯、出発地などの滞在情報を  $\mathbf{x}$  とおく。これらの情報を用い、目的地予測は次のように定式化される。

$$\hat{y} = \arg \max_y p(y|\mathbf{x}, l_{1:t}), \quad (1)$$

早期目的地予測とは、より小さいステップ  $t \ll T$  において高い精度で目的地を予測することである。

### 2.2 既存の予測手法の定式化

上記の問題設定において、経路情報に基づく予測、滞在情報に基づく予測がどのように定式化されるかを述べ、その問題点を示す。

経路情報に基づく予測では、ステップ  $t \in \{1, 2, \dots, T\}$  において目的地が  $y \in \mathcal{Y}$  である確率が、以下のように定義される。

$$p(y|l_{1:t}). \quad (2)$$

(2) が示すように、経路情報に基づく予測手法は、滞在情

報  $\mathbf{x}$  を考慮できていない。

一方で、滞在情報に基づく予測では、ステップ  $t \in \{1, 2, \dots, T\}$  において目的地が  $y \in \{1, 2, \dots, Y\}$  である確率が、以下のように定義される。

$$p(y|\mathbf{x}). \quad (3)$$

(3) が示すように、滞在情報に基づく予測手法は、経路情報  $l_{1:t}$  を考慮できていない。

また、小西ら [13] は、経路・滞在情報をともに考慮できる手法として、同時確率と条件付き確率の関係を用い、以下のような予測手法を構築した。

$$\begin{aligned} p(y|\mathbf{x}, l_{1:t}) &= \frac{p(y, l_{1:t}|\mathbf{x})}{p(l_{1:t})} \\ &\propto p(y, l_{1:t}|\mathbf{x}) \\ &= p(y|\mathbf{x})p(l_{1:t}|y, \mathbf{x}) \\ &= p(y|\mathbf{x})p(l_{1:t}|y). \end{aligned} \quad (4)$$

この予測手法では、目的地の確率を目的地に対する経路の尤度  $p(l_{1:t}|y)$  と、滞在情報に基づく目的地の確率  $p(y|\mathbf{x})$  に分解することができ、経路・滞在情報に基づく手法それぞれを構築することでより早期目的地予測に適した手法の構築が可能となる。

小西らは、経路情報に基づく手法として、電車移動の際の 3 種の理想経路を、最も安い経路、時間が短い経路、乗り換え回数が少ない経路とし、これらの理想経路と辿ってきた経路との距離を特徴量とした逆強化学習を用いた。また、滞在情報に基づく手法として曜日・時間帯・出発地点を特徴量とした多クラスロジスティック回帰を用いた。この手法は、電車移動の際の理想的な経路に実際の近いほど良い経路と判定でき、目的地を絞り込むことができるが、徒歩や車移動に関する特徴量を含まないため、これらの移動手段に対しては正しい予測ができないという問題がある。また、他の移動手段に拡張する際、移動手段ごとに特徴量を設計する必要があり、モデルが複雑になってしまうという問題がある。

筆者ら [14] は、この問題を解決するため、経路情報に基づく予測手法として、移動手段によらない手法の導入を行っている。

## 3. 移動手段によらない早期目的地予測手法

本節では、筆者らが提案した、移動手段によらない早期目的地予測手法を概説する。この手法は、小西らの手法 [13] の問題点である、電車移動以外に対して正しく予測ができないという問題と、様々な移動手段を考慮しようとするとモデルが複雑になりすぎるといった問題の解決を目的としている。この目的のため、(4) における経路情報に基づく手法として、移動手段によらず、辿ってきた経路の頻度のみに基づく手法である SubSynE [7] を導入する。

このモデルでは、移動の際にマルコフ性を仮定し目的地を予測する。M<sup>Total</sup> を L × L の、ある地点からある地点への遷移確率を表す行列とする。ここで、状態とは、離散化された位置である。M<sup>Total</sup> は状態 i から j への遷移確率 p<sub>i→j</sub> を用い、(5) のように表される。

$$M^{Total} = \begin{pmatrix} p_{1 \rightarrow 1} & \cdots & p_{1 \rightarrow j} & \cdots & p_{1 \rightarrow L} \\ \vdots & \ddots & & & \vdots \\ p_{i \rightarrow 1} & & p_{i \rightarrow j} & & p_{i \rightarrow L} \\ \vdots & & & \ddots & \vdots \\ p_{L \rightarrow 1} & \cdots & p_{L \rightarrow j} & \cdots & p_{L \rightarrow L} \end{pmatrix}. \quad (5)$$

この遷移確率行列は、ある状態から隣接する状態への遷移確率行列である M の積により計算される。

また、本手法においては、SubSynE アルゴリズムにより各目的地に対する経路の尤度を計算することが目的となる。目的地 y に対する経路 l<sub>1:t</sub> の尤度は式 (6) のようになる。ここで、n<sub>l</sub> は l における状態番号である。また、 $\prod_{i=1}^{t-1} M_{n_i n_{i+1}}$  は l<sub>1:t</sub> のみに依存するため、すべての目的地 y に共通することに注意する。

$$\begin{aligned} p(l_{1:t}|y) &= \frac{M_{n_t n_y}^{Total}}{M_{n_1 n_y}^{Total}} \prod_{i=1}^{t-1} M_{n_i n_{i+1}} \\ &= \frac{p_{n_t \rightarrow n_y}}{p_{n_1 \rightarrow n_y}} \prod_{i=1}^{t-1} M_{n_i n_{i+1}} \\ &\propto \frac{p_{n_t \rightarrow n_y}}{p_{n_1 \rightarrow n_y}}. \end{aligned} \quad (6)$$

また、滞在情報に基づく予測手法として、小西らの手法と同様、曜日・時間帯・出発地点を特徴量として多クラスロジスティック回帰を用いる。

## 4. 性能評価実験

本論文では、目的地予測の性能を確認するため [14] で行った定量的実験を概説し、筆者らの最近の提案手法である、移動手段によらない早期目的地予測手法の予測の様子や問題点を確認するための定性的実験の結果を報告する。

### 4.1 実験目的

定量的実験では、移動段階ごとの精度比較実験により、移動手段によらない早期目的地予測手法が、経路・滞在情報に基づく予測の両手法を組み合わせることにより、経路・滞在情報に基づく手法と比較して、早期目的地予測に適した手法となっており、かつ、移動手段によらない予測が可能であることを確認することを目的としている。

定性的実験では、目的地予測の様子をプロットした図を確認することにより、移動手段によらない早期目的地予測手法による予測がどのように行われ、経路・滞在情報に基づく予測との差を視覚的に確認する。

### 4.2 データセット\*5

実験に用いる位置情報は、Yahoo!防災速報\*6により得られたデータのうち、東京周辺在住者 100 人の、2016 年 5 月 1 日から 2016 年 6 月 30 日までに取得されたものである。取得されたデータには、ユーザ ID、緯度、経度、取得時間 (JST)、速度、GPS の精度が含まれ、このうち、ユーザ ID、緯度、経度、取得時間を利用する。データが取得されているユーザは日本全国に数十万~数百万人規模で存在しているが、データのハンドリングのため、東京周辺在住者を抽出した。この際、東京周辺を、北緯 35.5 度から 35.9 度、東経 139.2 度から 139.95 度の範囲とした。その後、データが毎日取得できているユーザ約 1690 名を抽出した。

### 4.3 定量的実験：精度比較実験

実験の目的である、筆者らの最近の提案手法が既存手法と比較して早期目的地予測に適した手法となっていることを示すため、移動終了割合ごとの精度比較実験を行う。ここで、移動終了割合は、総ステップを T、現在のステップを t としたとき、t/T で表され、これを、fraction と呼び以下のように離散化する。

$$\text{fraction} = \begin{cases} 0 \\ 0.2 \\ 0.4 \\ 0.6 \\ 0.8 \end{cases}. \quad (7)$$

#### 4.3.1 評価指標

評価指標として、精度を用いる。本実験では、精度を、予測された目的地が、真の目的地と一致していた割合とする。縦軸に精度、横軸に移動終了割合を取る。移動終了割合が小さく、精度が高いほど、より早期に正確な予測ができており、早期目的地予測に適した手法と言える。

#### 4.3.2 比較手法

移動手段によらない早期目的地予測手法の利点を確認するため、以下の 4 種類の手法を比較手法として用いる。

- Baseline: 曜日や時間帯などに関係なく、常に過去最も頻度が高かった目的地を予測目的地とする手法。目的地予測のベースラインとして用いる。この手法と他の手法との比較により、経路・滞在情報を考慮することの効果を確認する。
- SubSynE: 経路情報に基づく予測手法として用いる。SubSynE による目的地予測は、次の式による。

$$p(y = d|l_{1:t}) = \frac{P(l_{1:t}|y = d)P(y = d)}{\sum_{d'=1}^K P(l_{1:t}|y = d')P(y = d')} \quad (8)$$

\*5 氏名や住所等、個人を特定する情報とは切り離れた上で、データを分析を行っている。

\*6 <https://itunes.apple.com/jp/app/yahoo!-fang-zai-su-bao-zhenya/id481914139>

- 多クラスロジスティック回帰: 滞在情報に基づく予測手法として用いる。
- ナイーブな組み合わせ手法: 小西らが提案した組み合わせ手法に対し, 単純に経路・滞在情報に基づく予測それぞれを計算し予測確率が高かったほうの結果を用いる手法が考えられる。この組み合わせ手法に対し, 同時確率と条件付き確率の関係を用いた組み合わせ手法が優れた予測を行うことを示す。

#### 4.3.3 実験結果

実験結果は図 1 である。ここでは, データをランダムに 5 分割し 5 交差検定を行った。縦軸は 5 分割交差検定の各結果の平均であり, エラーバーはこの母平均の 95%信頼区間を表す。

まず, Baseline と他の予測手法の比較により, 滞在情報・移動情報を考慮することで, 目的地予測の精度が向上することが確認できる。SubSynE に焦点を当てると, 移動情報を考慮することにより, 精度が移動に従い上昇していくことがわかる。一方で, 多クラスロジスティック回帰に焦点を当てると, 初期から精度が高いが, 移動が進んでも精度が変化しないことがわかる。また, ナイーブな組み合わせ手法と比較することにより, 移動手段によらない予測手法が移動情報と滞在情報の両方を同時に加味できていることで, 高い精度をもつことが確認された。

ここで, SubSynE を内包する予測手法の予測精度の移動に対する上昇が小さい点に関しては, 移動距離が短い移動が結果に影響を与えているためであると考えられる。これは, 移動距離が小さい場合, 移動情報が十分でなく, SubSynE の精度が低くなることが予想されるためである。そこで,  $T$  が 5 以上である移動のみを抽出し, 予測結果のプロットを行った。図 2 は  $T$  が 5 以上である移動のみの予測結果である。SubSynE を内包する手法に着目すると, 図 1 の結果と比較して, 移動に対する精度の上昇が急であることがわかる。

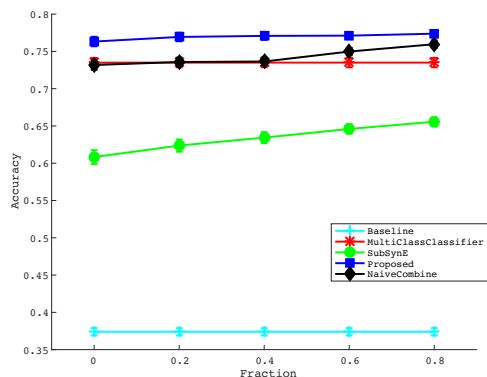


図 1: 精度比較実験結果

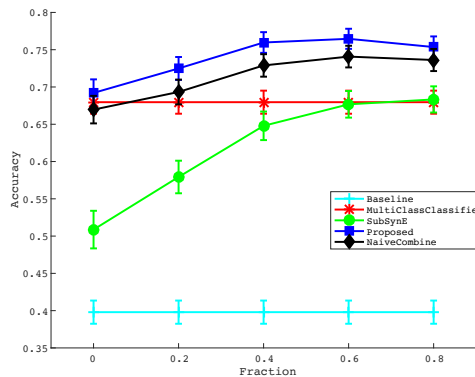


図 2: 精度比較実験結果 ( $T \geq 5$ )

#### 4.4 定性的実験: 移動に対する予測の様子の確認

実験目的である, 移動手段によらない早期目的地予測手法による予測がどのように行われ, 経路・滞在情報に基づく予測手法との差を視覚的に確認するため, 移動手段によらない早期目的地予測手法と, 経路・滞在情報に基づく予測手法の移動につれての予測の様子の変化をプロットした。ここで, 黒線が移動経路, 黒点が現在地, 赤点が真の目的地, 青点が真の目的地以外の目的地候補である。赤点, 青点のサイズの大きさが予測確率の大きさを表しており, 予測確率が 5%以上の点のみプロットしている。また, 匿名性の観点から, 背景の地図や緯度経度情報, 縮尺情報等は除外している。

図 3 は経路・滞在情報が正しく考慮されている場合の移動の様子を表している。SubSynE による予測に着目すると, 移動につれて目的地の候補が絞り込まれているが, 一方で, 移動初期は目的地の絞り込みがあまりできていないことがわかる。一方で, 多クラスロジスティック回帰による予測に着目すると, 移動初期から目的地が 2 点に絞り込まれているが, 移動が進んでも予測を修正できておらず, 目的地に到着するまで予測が間違っただまとなっている。筆者らの最近の提案手法に着目すると, 移動初期から目的地の候補が 2 点に絞り込まれており, かつ, 予測結果が移動が進むに連れ修正され, 移動の終盤には真の目的地を予測できていることがわかる。

また, 経路・滞在情報単体では予測が絞り込めない場合でも, 組み合わせにより予測結果を絞り込むことができる。図 4 では, 多クラスロジスティック回帰による予測では十分に目的地を絞り込むことができておらず, SubSynE による予測でも, 移動が進むに従って真の目的地の予測確率が高くなっているが, 他の目的地候補も同様に確率が高くなっているものがあり, やはり絞り込みが十分にできていない。一方で, 筆者らの最近の提案手法による予測では, SubSynE, 多クラスロジスティック回帰による予測両方で予測確率が高い候補を抽出していることから, 目的地の絞り込みができていくことがわかる。

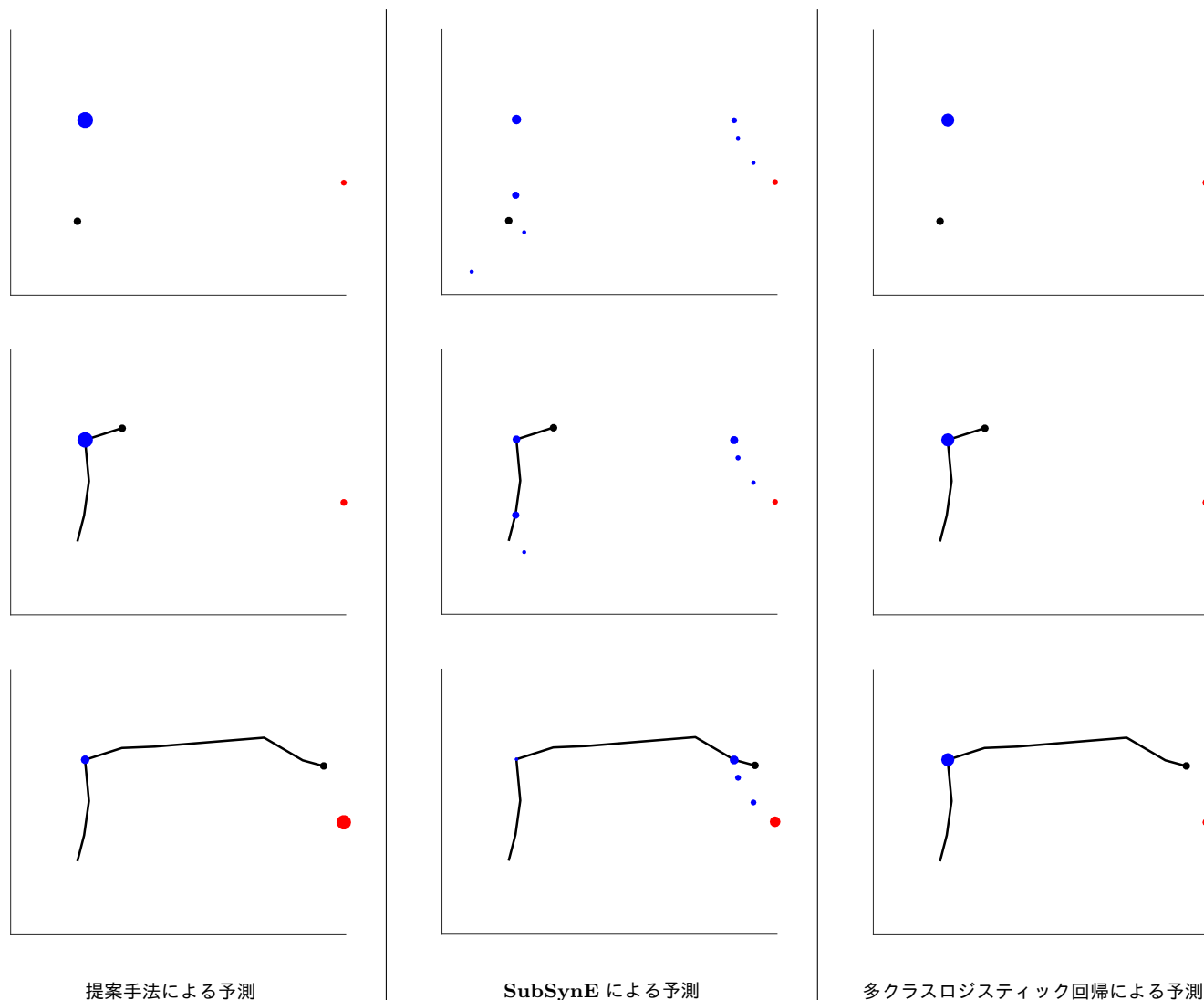


図 3: 経路・滞在情報が正しく予測に利用されているときの移動手段によらない早期目的地予測手法 (左), SubSynE (中央), 多クラスロジスティック回帰 (右) による予測の様子

一方で、経路・滞在情報の両方を考慮しているため、片方の予測がもう片方の予測を阻害するが発生することもある。図 5 は経路情報による予測が滞在情報による予測を阻害している例である。多クラスロジスティック回帰による予測は正しい予測を行っているが、SubSynE による予測では、移動につれて真の目的地に予測確率が低くなっていき、移動の終盤では予測確率が 5% 未満となってしまっている。これは、SubSynE が移動頻度に基づく予測であるため、予測を行った位置から真の目的地までの過去の経路の組み合わせが少ない場合、位置が真の目的地に近づいていたとしても、真の目的地を予測できないためである。筆者らの最近の提案手法に着目すると、経路情報による予測が滞在情報による予測を阻害し、正しく目的地を予測できなくなっていることがわかる。しかし、滞在情報による予測により、経路情報による予測が補正されているとも考えられ、実際に、移動手段によらない早期目的地予測手法による真の目的地の予測結果は上位 2 位以内に含まれている。

## 5. 結論

本論文では、筆者らが提案した、既存の早期目的地予測手法の、電車移動にしか利用できない、モデルの複雑さという 2 つの問題点を解決する、移動手段によらない早期目的地予測手法を概説し、既存の研究報告でなされていなかった、定性的な評価を行った。

今後の課題として、頻度が低い目的地に対する精度向上が挙げられる。このような低い頻度の目的地には、遊園地やレストラン、観光地での行き先など、予測することでユーザにとって有益な情報を提供できる可能性がある目的地が含まれていると予想され、重要な課題と言える。このような目的地を予測する手段として、ユーザ情報のような追加データの利用が考えられる。

謝辞 4 節における大規模ユーザの位置情報を用いた実験は、ヤフー株式会社との共同研究にて実施したものである。また、2 節および 3 節のモデルの設計は本研究は JSPS

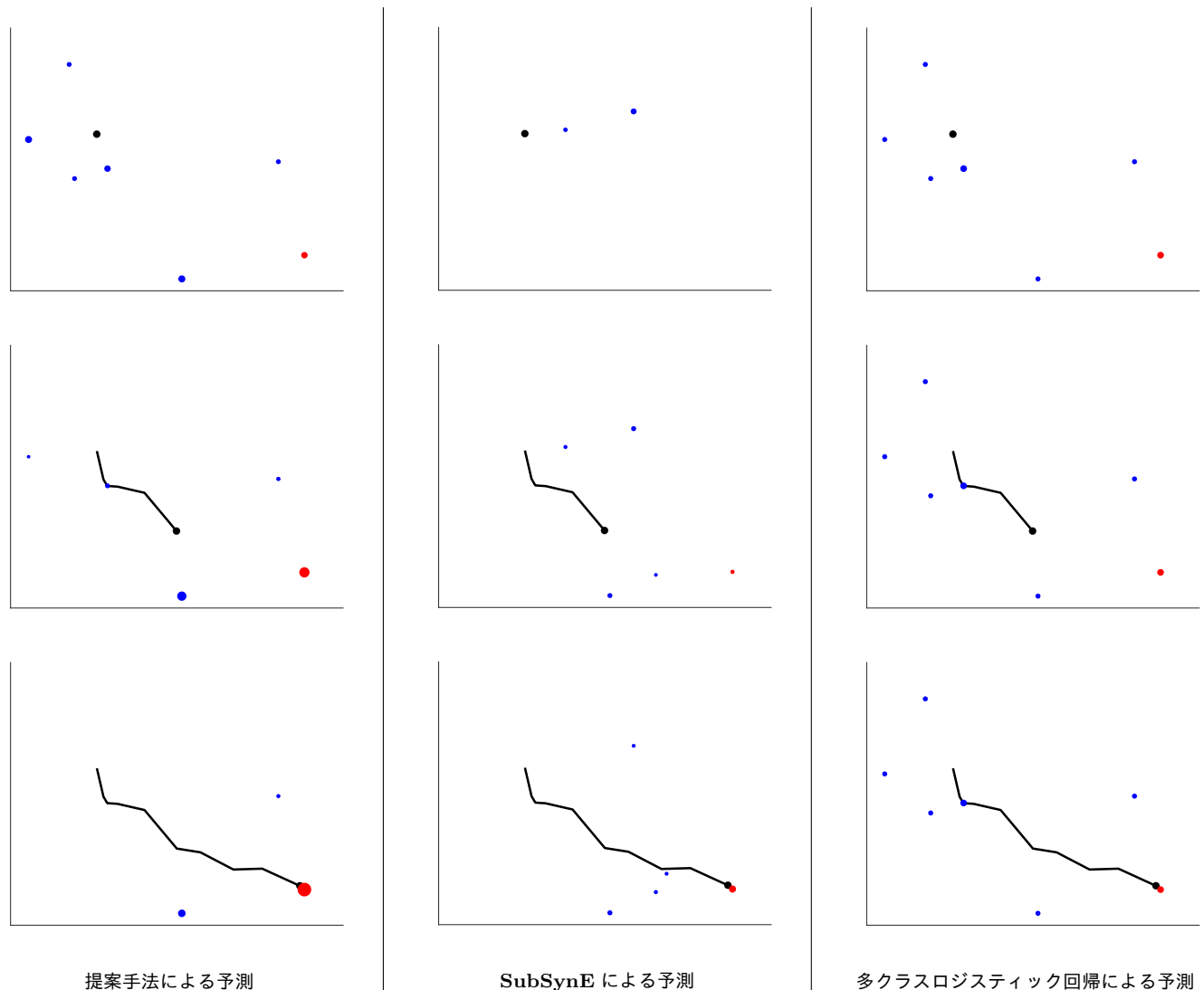


図 4: 経路・滞在情報単体では絞り込めなかった目的地が、組み合わせにより絞り込まれる様子。移動手段によらない早期目的地予測手法 (左), SubSynE (中央), 多クラスロジスティック回帰 (右) による予測

KAKENHI JP25700026, および JST, CREST の支援を受けたものである。

#### 参考文献

- [1] Sen, A. and Larson, M.: From Sensors to Songs: A learning-free novel music recommendation system using contextual sensor data, *Proc. of RecSys 2015 Workshop LocalRec*.
- [2] Woerndl, W., Schueller, C. and Wojtech, R.: A Hybrid Recommender System for Context-aware Recommendation of Mobile Applications, *Proc. of ICDE 2007*, pp. 871–878.
- [3] Parate, A., Bohmer, M., Chu, D., Ganesan, D. and Marlin, B. M.: Practical Prediction and Prefetch for Faster Access to Applications on Mobile Phones, *Proc. of UbiComp 2013*, pp. 275–284.
- [4] Okoshi, T., Ramos, J., Nozaki, H., Nakazawa, J., Dey, A. K. and Tokuda, H.: Reducing users' perceived mental effort due to interruptive notifications in multi-device mobile environments, *Proc. of UbiComp 2015*, pp. 475–486.
- [5] Krumm, J. and Horvitz, E.: Predestination: Inferring Destinations from Partial Trajectories, *Proc. of UbiComp 2006*, pp. 243–260.
- [6] Xue, A. Y., Zhang, R., Zheng, Y., Xie, X., Huang, J. and Xu, Z.: Destination Prediction by Sub-Trajectory Synthesis and Privacy Protection Against Such Prediction, *Proc of ICDE 2013*, pp. 254–265.
- [7] Xue, A. Y., Qi, J., Xie, X., Zhang, R., Huang, J. and Li, Y.: Solving the data sparsity problem in destination prediction, *J. of VLDB 2015*, pp. 219–243.
- [8] Manasseh, C. and Sengupta, R.: Predicting Driver Destination using Machine Learning Techniques, *Proc. of ITSC 2013*, pp. 142–147.
- [9] Nadembega, A., Taleb, T. and Hafid, A.: A Destination Prediction Model based on Historical Data, Contextual Knowledge and Spatial Conceptual Maps, *Proc. of ICC 2012*, pp. 1416–1420.
- [10] Gao, H., Tang, J. and Liu, H.: Mobile Location Prediction in Spatio-Temporal Context, *Proc. of Pervasive Computing 2012 Workshop Nokia Mobile Data Challenge*.
- [11] Zhu, Y., Sun, Y. and Wang, Y.: Nokia Mobile Data Challenge: Predicting Semantic Place and Next Place



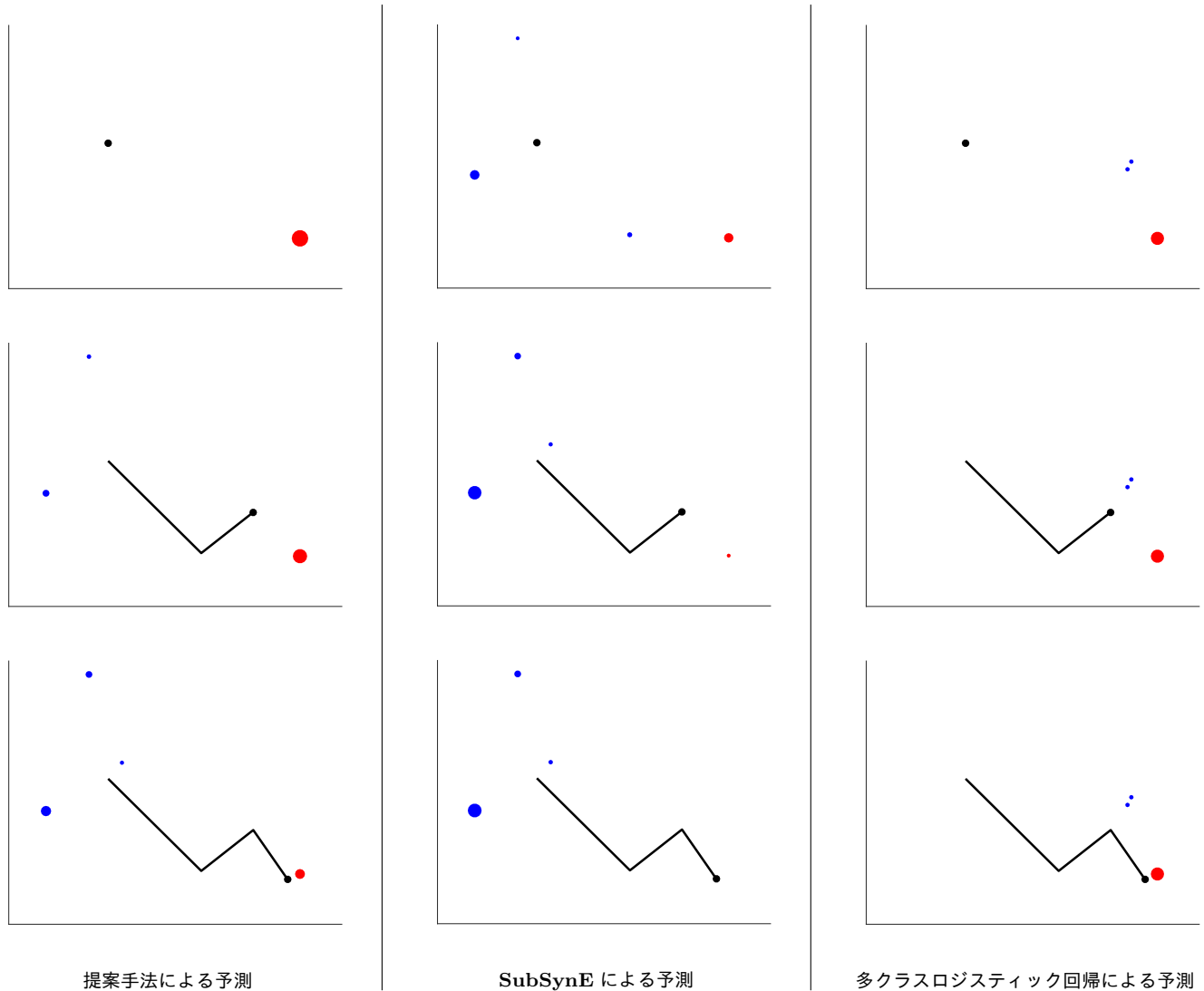


図 5: 経路情報による予測が滞在情報による予測を阻害しているときの移動手段によらない早期目的地予測手法 (左), SubSynE (中央), 多クラスロジスティック回帰 (右) による予測の様子

via Mobile Data, *Proc. of Pervasive Computing 2012 Workshop Nokia Mobile Data Challenge*.

[12] Huang, W., Li, M. and Hu, W.: Hierarchical destination prediction based on GPS history, *Proc. of FSKD 2013*, pp. 972–977.

[13] 小西達也, 下坂正倫: 滞在時間帯と経路情報を用いた混合最大エントロピー逆強化学習に基づく早期目的地予測, 情報処理学会研究報告 (2016).

[14] Imai, R., Tsubouchi, K., Konishi, T. and Shimosaka, M.: Early Destination Prediction with Spatio-temporal User Behavior Patterns, *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. Vol. 1, Issue 4, 2017*, pp. 142:1–142:19.

[15] Tominaga, S., Shimosaka, M., Fukui, R. and Sato, T.: A Unified Framework for Modeling and Predicting Going-Out Behavior, *Proc. of Pervasive Computing 2012*, pp. 73–79.

[16] McInerney, J., Zheng, J., Rogers, A. and Jennings, N. R.: Modelling Heterogeneous Location Habits in Human Populations for Location Prediction Under Data Sparsity, *Proc. of UbiComp 2013*, pp. 469–478.

[17] D.Ziebart, B., L.Maas, A., Bagnell, J. and Dey, A. K.: Maximum Entropy Inverse Reinforcement Learning, *Proc. of AAAI 2008*, pp. 1433–1438.

[18] D.Ziebart, B., L.Maas, A., Dey, A. K. and Bagnell, J.: Navigate Like a Cabbie: Probabilistic Reasoning from Observed Context-Aware Behavior, *Proc. of UbiComp 2008*, pp. 322–331.