

## 平均移動速度比に基づく位置情報からの滞留期間推定方式 A Method to Predict Duration of Stay Using Average Speed Ratio from Location Information

梅津 圭介<sup>†</sup> 多賀戸 裕樹<sup>†</sup>  
Keisuke Umezu Hiroki Tagato

### 1. はじめに

近年、位置情報の分析技術に注目が集まっている。背景には、スマートフォンやカーナビゲーションシステムなどの GPS を備えた機器、3G や LTE などの通信環境の普及により、位置情報の収集、蓄積が容易になったことがある。

[1][2][3]では、位置情報からユーザが滞留した期間である滞留期間を推定する技術が提案されている。滞留期間を推定できると、滞留期間における位置情報をもとにユーザの滞留したエリアが特定できる。ここで、特定したエリアをレストラン等のスポットと紐付けることで、各スポットへの滞留頻度によるユーザの嗜好の推定も可能となる。

しかしながら、[1][2][3]に示す従来の滞留期間推定方式は、滞留期間を正確に推定できない。なぜなら、滞留期間前後のユーザの移動手段に対する推定性能のロバスト性が低いためである。滞留期間を推定するためには事前に複数のパラメータを設定する必要があるが、従来方式では、どのような移動手段においても滞留期間の推定性能が高くなるパラメータは設定できない。例えば、徒歩移動時における滞留期間の推定性能が高くなるパラメータを設定すると、電車移動時の推定性能は低くなる。

そこで、本稿では滞留期間を正確に推定することを目的とし、滞留期間前後の移動手段に対して推定性能がロバストな滞留期間推定技術を提案する。

### 2. 滞留期間推定方式

#### 2.1 従来方式

[1][2][3]に示す従来方式は、事前に設定するパラメータにより定義される条件を満たす期間を、滞留期間と推定する。[1]に記載の方式は、時間  $t$ 、半径  $r$ 、の 2 つのパラメータを設定し、位置情報が一定時間  $t$  以上、半径  $r$  の円内に収まっている期間を滞留期間と推定する。[2]では、時間  $t$ 、半径  $r$ 、速度  $v$ 、の 3 つのパラメータを設定し、位置情報が一定時間  $t$  以上、半径  $r$  の円内に収まっている期間のうち、さらに位置情報間の距離と測位時刻から算出される位置情報間の差分速度がすべて  $v$  以下である期間を滞留期間と推定する。[3]では、時間  $t$ 、速度  $v1$ 、 $v2$ 、の 3 つのパラメータにより定義される 2 つの条件を満たす期間を滞留期間と推定する。

前述のとおり、各従来方式は滞留期間前後の移動手段に対して推定性能がロバストでないため、滞留期間を正確に推定できない。例えば、[2]の方式において  $t=20sec$ 、 $r=10m$  とパラメータを設定するとき、電車移動時における滞留期間は高い確率で正しく推定できる。しかし、徒歩移動時には移動中であっても半径 10m のエリア内に 20 秒以上留まることがよくあるため、実際には移動中であつた期

間を誤って推定することが多い。一方、パラメータに対し  $t=100sec$ 、 $r=10m$  が設定される場合、徒歩移動時に移動中である期間を誤って推定することは少ないが、電車移動時に実際には滞留した期間が推定できないことが多い。理由は、電車移動時には駅ホームにおける数十秒程度の滞留が頻繁に起こるためである。

#### 2.2 提案方式

本稿では、滞留期間を正確に推定するため、滞留期間前後の移動手段に対して推定性能がロバストな滞留期間推定技術を提案する。本方式では、滞留期間を推定するため、時間  $t$ 、半径  $r$ 、移動速度比  $\alpha$ 、周辺時間  $\beta$ 、の 4 つのパラメータを設定する。提案方式のアルゴリズムを以下に示す。

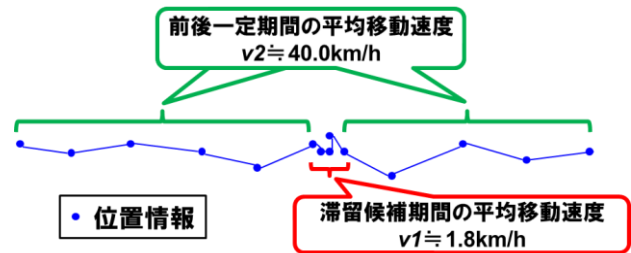


図 1 平均移動速度

ステップ 1: 各パラメータを設定する。

- ①時間  $t$
- ②半径  $r$
- ③移動速度比  $\alpha$
- ④周辺時間  $\beta$

ステップ 2: 位置情報を読み込む。

ステップ 3: 滞留期間が  $t$  以上かつ半径  $r$  の円内に収まっている期間を候補期間としてリストアップする。

ステップ 4: 各候補期間における平均移動速度  $v1$  と、滞留候補期間の前後の一定期間における平均移動速度  $v2$  を算出する。なお、一定期間の長さは周辺時間  $\beta$  とする。

ステップ 5: 各候補期間における移動速度比 ( $v2/v1$ ) を算出し、移動速度比が  $\alpha$  より小さい候補期間をリストより削除する。

ステップ 6: リストに含まれる候補期間を時系列順にソートし、期間の重複する場合にはより移動速度比の小さい候補期間をリストより削除する。

ステップ 7: リストに含まれる候補期間を、滞留期間として出力する。

移動速度比  $\alpha$  をパラメータとすることにより、その移動距離や速度によらず一定の速度で移動している期間は滞留期間と推定されない。そのため、提案方式では  $t$  に対して電車移動時における滞留期間の推定性能が高くなるような小さい値を設定した場合であっても、歩行移動時に移動中である期間を誤って推定することは少ない。

<sup>†</sup> NEC 情報・ナレッジ研究所, NEC Knowledge Discovery Research Laboratories

### 3. 評価実験

#### 3.1 評価用データの収集

評価実験のため、被験者3人、延べ8日分の位置情報と、位置情報を収集する間に実際に滞留していた期間である正解滞留期間と、移動手段を収集した。われわれは、10秒間隔で位置情報を測位、記録する Android アプリを実装し、自宅を出発してから戻るまでの間の位置情報を収集した。正解滞留期間は5秒以上その場に立ち止まった期間と定義し、計338個の正解滞留期間をアンケート形式で収集した。移動手段は、車、電車、バス、自転車、徒歩のなかから選択してもらう形式で収集した。

#### 3.2 評価方法

本実験では、[1][2][3]に示す従来方式と提案方式の4つの方式の滞留期間推定性能を評価する。推定性能は、適合率、再現率、F値、推定時間誤差 $E$ の4つの指標により評価する。本実験では、滞留期間の推定結果である推定滞留期間と正解滞留期間とを比較したときに1秒以上重なっている期間がある場合を推定成功と定義する。推定時間誤差 $E$ は、正解滞留期間と推定滞留期間の一致度合いを表す。推定時間誤差 $E$ は正解滞留データの滞留時間 $ct$ 、推定滞留データの滞留時間 $et$ 、正解滞留データと推定滞留データの重なる時間の長さ $ot$ を用いて、以下の数式で算出する。

$$E = \frac{ct + et - (2 \times ot)}{ct + et - ot}$$

また、従来方式と提案方式における滞留期間前後の移動手段に対する推定性能のロバスト性を検証する。本稿では、[2]に記載の方式と提案方式の2つの方式において、パラメータを変化させたときの各移動手段の推定性能を検証する。[2]の方式を選択する理由は、F値と推定時間誤差 $E$ の逆数を乗じた結果からなる滞留期間の推定性能が、[1][2][3]の中で最も高いためである。

#### 3.3 評価結果

各方式の推定性能の評価結果を表1に示す。なお、表1に示す値は、各方式においてパラメータを変化させながら試行を繰り返したとき、F値が最も大きくなった試行における結果である。結果、提案方式のF値は従来方式と比較して0.52~0.20ポイント程度良いことが分かった。また、提案方式の適合率は従来方式と比べて0.63~0.22ポイント程度改善されている。[3]の方式は、従来方式のなかで最もF値は大きいですが、提案方式と比べて推定時間誤差の値が0.39ポイント程度低い。一方、[2]の方式は提案方式との推定時間誤差の差が0.04ポイント程度と小さいが、F値は0.31ポイント程度低い値となった。

表1 推定性能 (推定方式別)

	適合率	再現率	F値	推定時間誤差
提案方式	0.785	0.778	0.782	0.139
従来方式[1]	0.156	0.843	0.264	0.316
従来方式[2]	0.353	0.716	0.473	0.171
従来方式[3]	0.565	0.607	0.585	0.533

次に、滞留期間前後の移動手段に対する推定性能のロバスト性を検証する。図2に、[2]の方式においてパラメータを変化させたときのF値を移動手段別に示す。結果、徒歩

移動時のF値は $t$ を増やしていくと大きくなるが、他の移動手段においては $t=20$ を境に $t$ を増やすほどF値は小さくなる。[2]の方式はパラメータを変化させたときの各移動手段のF値の変動が大きく、移動手段ごとに最適なパラメータも異なるため、[2]の方式の滞留期間前後の移動手段に対する推定性能のロバスト性は低いと言える。次に、提案方式においてパラメータを変化させたときの移動手段別のF値を図3に示す。提案方式ではパラメータを変化させたときのF値の変動はどの移動手段であっても小さい。そのため、提案方式は滞留期間前後の移動手段に対する推定性能のロバスト性は高いと言える。

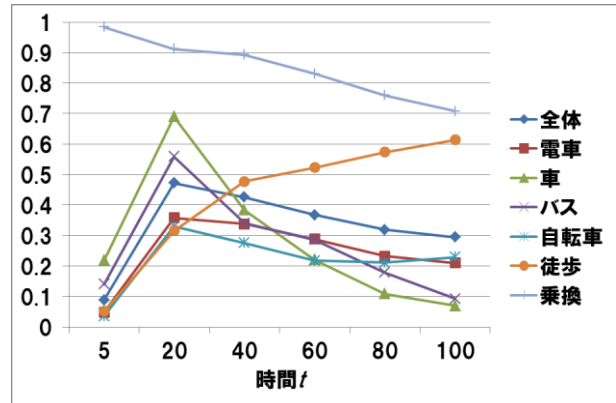


図2 F値 (移動手段別, 従来方式[2])

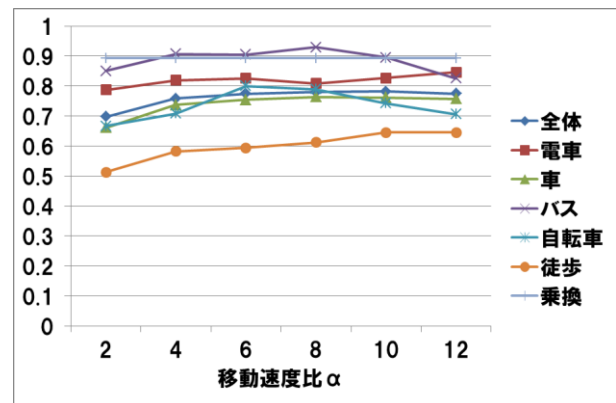


図3 F値 (移動手段別, 提案方式)

### 4. まとめ

本稿では、滞留期間を正確に推定することを目的とし、滞留期間前後の移動手段に対して推定性能がロバストな滞留期間推定技術を提案した。また、従来方式と提案方式の推定性能を評価した。結果、従来方式と比べて提案方式の推定性能は良く、ロバスト性が高いことが明らかになった。

#### 参考文献

- [1] Zheng Yu, Zhang Lizhu, Xie Xing, Ma Wei-Ying Ma, "Mining interesting locations and travel sequences from GPS trajectories", Proceedings of the 18th international conference on World wide web, pp791-800(2009).
- [2] 堀口良太, 長岡亨, 畑成年, "GPS携帯電話による大規模パーソナルプローブ調査のためのトリップ情報抽出手法に関する研究", 土木計画学研究・講演集, Vol.33(2006).
- [3] 長尾光悦, 川村秀憲, 山本雅人, 大内東, "GPSログからの周遊型観光行動情報の抽出", 情報処理学会研究報告. ICS, 知能と複雑系(78), pp23-28(2005)