

# 個別化された道路交通情報のプッシュ型配信に向けた定常的な自動車利用パターンの抽出

賈 舒陽<sup>1</sup>・山本 俊行<sup>2</sup>・森川 高行<sup>3</sup>

<sup>1</sup>非会員 名古屋大学 大学院環境学研究科 (〒464-8603 名古屋市千種区不老町)

E-mail: jia.shuyang@d.mbox.nagoya-u.ac.jp

<sup>2</sup>フェロー 名古屋大学教授 エコトピア科学研究所 (同上)

E-mail: yamamoto@civil.nagoya-u.ac.jp

<sup>3</sup>正会員 名古屋大学教授 未来社会創造機構 人とモビリティ社会の研究開発センター (同上)

E-mail: morikawa@nagoya-u.jp

現在のカーナビなどの道路交通情報提供システムは、運転者自ら目的地を手動で入力する必要がある。また、VICSなどの道路交通情報を提供する際に情報の個別化は十分に行われていない。本研究では、プローブ情報収集技術により収集した、個々の運転者の自動車利用データを用いて、定常的な自動車の利用パターンを抽出するためのアルゴリズムを構築した。また、これらの定常性を持つ代表トリップパターンを対象として、将来のトリップを予測し、かつ予測の精度を検証した。さらに、データ学習期間の変化や予測前のユーザー位置確認などによる予測精度への影響を考察した。

**Key Words** : mobility pattern, GPS observation, traffic information, trip prediction

## 1. はじめに

現在、高度道路交通システム (ITS) 技術の発展によって、道路網上での交通状況をリアルタイムで把握する仕組みが実現しつつある。自動車運転者に対してその情報を提供することで、運転者自身はリアルタイムの道路交通状況を把握することができる。このような情報通信技術を利用した自動車運転者へ道路交通情報を提供する方法の一つとしてVICS (Vehicle Information and Communication System) が挙げられる。しかしながら、カーナビを通じた情報提供では、運転席に座ってからでしか情報を得ることができない。また、トリップ経路に関係した情報を取得するには、カーナビにトリップの目的地を運転者が入力する必要がある。道路交通情報のより有効な活用方法を考えた時、運転者が自ら入力する必要がなく、運転者のトリップに合わせた情報が提供されることや、運転席に座ってトリップを開始する以前に必要な情報が提供されることが望ましい。

一方、人々の運転行動や交通行動を把握する方法として、プローブカー調査やプローブパーソン調査が数多く実施されており、人々の行動の長期的な観測が可能となっている。さらに、プローブパーソンデータに関する研究では、被験者による追加的な入力を必要とせず、

GPS データや加速度データだけから滞在・移動判別や交通手段、トリップ目的の特定を行うアルゴリズムの開発が進められている。

本研究では、自動車ユーザーがトリップを開始するよりも前に、そのユーザーの利用トリップに対応した情報を提供するシステムの開発を目的として、プローブカーデータを用いて運転者の詳細な位置情報や運転情報を長期間取得し分析を行うことで、運転者の自動車利用パターンを把握し、運転者の定常的なトリップパターンを抽出、トリップ開始前に当該トリップの発生を予測するアルゴリズムの構築を行う。

## 2. 既往研究および本研究の目的と位置付け

GPS による運転履歴データを用いて自動車利用パターンの定常性を分析して将来トリップパターンを予測した研究として、Froehlich and Krumm<sup>1)</sup>は、自動車を利用する約 250 人の自動車走行 GPS データを用いて、経度と緯度から自動車の地理的な位置に注目し、自動車利用経路を予測するアルゴリズムを構築した。そして、40 日間以上の連続観測においては、自動車ユーザーのおよそ 60% のトリップが重複していることを明らかにした。また、一部のユーザーに対し、トリップ開始直後の約 2 マイル

(約3.2キロメートル)の経路が分かれば、そのユーザーの利用経路はほぼ100%の確率で予測可能であることを示した。Simmons et al.<sup>2)</sup>は隠れマルコフモデル (Hidden Markov Model) を用いて、およそ1ヶ月間の観測データから自動車ユーザーの利用経路を予測し、98%以上の正解率を得た。これらの研究は、分析データの収集期間が短く、運転者の自動車利用パターンを分析する際に、利用経路に着目している。

本研究では、プローブカー情報収集技術により収集した個々の運転者の自動車利用パターンから将来トリップを予測し、その精度を検証することを目的とする。これまでの研究で使われたデータの観測期間が短いのに対して、本研究では8ヶ月にわたる長期的な観測に基づく自動車走行GPSデータを用いた。また、運転者の自動車利用パターンを分析する際に、観測期間を変化させて予測精度への影響を考察していることが本研究の特徴として挙げられる。さらに、自動車ユーザーの利用経路ではなく、トリップの起点と終点の情報から自動車利用パターンを分析することも本研究の特徴の一つである。

### 3. 月毎の代表トリップパターン抽出

#### (1) 使用データの概要

本研究では、表-1に示すように、豊田市内在住あるいは在勤の268人の運転者を対象として平成23年3月から平成23年12月までの10ヶ月間に実施された車載器を活用したエコドライブに関する実験で得られたデータを使用した。元データからすべてのユーザーの全トリップの起点(O点)と終点(D点)のデータをそれぞれ抜き出し、月毎に起点と終点を分別しcsvファイルを生成した。

#### (2) 月ごと代表トリップパターンの抽出

抜き出した全トリップの起点と終点のデータを用いて、月毎に各ユーザーの定常性を持つ代表トリップパターンを抽出する。

##### a) 代表トリップパターン

本研究では、代表トリップパターンを抽出する際に、距離と時間の2つの基準から判断する。つまり、起点間距離と終点間距離は一定の範囲内に、さらにトリップ開始時刻が一定の範囲内であれば、同じトリップパターンとして認識する。月毎の各ユーザーの定常性を持つトリップパターン抽出手順を図-1に示す。

##### b) 距離基準での起点データ選定

3(1)で生成した起点データの集合を用いて、捜査半径radius1を半径とした円に囲まれる起点ポイントがもっとも多くなる円を選定する。ユーザーごとに、O点グループの各点*i*に対し、その他の全ての点*t*との距離を経度と

表-1 使用データの概要

対象ユーザー	豊田市内在住、 在勤の268人自動車ユーザー
データ種類	自動車走行GPSデータ
記録機器	社団法人交通工学研究会より無償で 貸与された専用車載器
記録間隔	1秒ごと
観測期間	2011年3月—2011年12月
使用データ期間	2011年4月—2011年11月
使用データ項目	ユーザーコード、測位データID、 測位日時、トリップID、緯度、経度

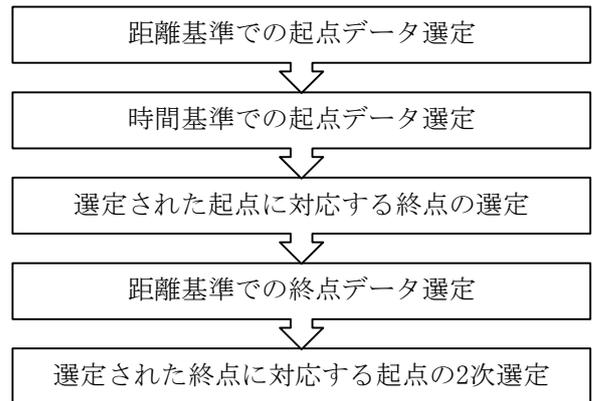


図-1 月ごとの代表トリップパターン抽出手順

緯度から計算し、 $O(i, t)$ とする。また、 $O(i, t) < radius1$ となる点の数を*m*とする。 $m \geq minpoint$ の点*i*に対し、最小平均分散法を用いて次の式の値を計算する。なお、*minpoint*は最少抽出点数である。

$$\frac{\sum(O(i, t)^2)}{m} \quad (1)$$

上式の値が一番小さい点*i*を円心とし、radius1を半径とした円の内にはO点がもっとも集中していると判断する。この円*i*に囲まれるすべてのO点データを抜き出し、3(2-b)に移る。なお、本研究では、感度分析の結果、捜査半径radius1を450mに設定した。

##### c) 時間基準での起点データ選定

距離基準で抽出されたO点グループは、次に時間基準で選定する。各O点の測位日時の年月日を取り除き、単位を秒に換算する。一日分の時間を0sから86400sとした場合、幅3600s、間隔60sのタイムウインドウが合計1380個得られる。これらのタイムウインドウを用いて、各タイムウインドウ内のO点の数を*n*とする。 $n \geq minpoint$ のタイムウインドウに対し、各O点の時刻を*ti*とし、タイムウインドウの中心点の時刻を*T*とし、最小平均分散法を用いて次の式の値を計算する。

$$\frac{\sum((t_i - T)^2)}{n} \quad (2)$$

上式の値が一番小さいタイムウインドウの中に含まれるO点グループを一つのトリップパターンとして出力する。この作業は $n \geq \text{minpoint}$ のタイムウインドウがなくなるまで繰り返す。

時間基準での選定作業が終了したら、3-(2)-b)に戻り距離基準で2回目の選定作業を行う。次はまた3-(2)-c)に移し時間基準で選定する。このように選定作業は $m \geq \text{minpoint}$ のO点なくなるまで繰り返す。

#### d) 選定された起点に対応する終点の選定

3-(2)-c)で選定されたO点グループに対応するD点のデータを3-(1)で生成した終点ファイルから抜き出す。

#### e) 距離基準での終点データ選定

3-(2)-d)で抜き出したD点のデータを用いて、3-(2)-b)のO点と同じ作業を距離基準でD点の選定を行う。D点に対する捜査半径はO点の捜査半径 $\text{radius1}$ より狭く、 $\text{radius2} = 220\text{m}$ に設定した。

終点データに対する選定作業は距離基準のみで行い、時間基準での選定作業は行わない。これは道路交通の事情で目的地への到着時間がばらばらになり、定常性が見つからなくなってしまう危険性があるためである。

#### f) 選定された終点に対応する起点の2次選定

3-(2)-e)で選定されたD点グループに対応するO点のデータを3-(2)-c)で生成した起点ファイルから抜き出す。

### 4. 代表トリップパターンの月変動解析

3章では、ユーザー毎に各月の定常性を持つ複数の代表トリップパターンの抽出手法について述べた。本章ではユーザー毎に、連続した数ヶ月間で定常性を持つトリップパターンを選定する。

#### (1) 代表O、D点位置の計算

ユーザー毎に、3章で抽出できた各トリップパターンのO点とD点の地理的中心点をそれぞれ計算する。計算方法については3-(2)-b)とほぼ同様である。あるユーザーのあるトリップパターンには $m$ 個のトリップがある。このトリップパターンに含まれる各O点 $i$ に対し、その他の全ての点 $t$ との距離を計算し、 $O(i, t)$ とする。最小平均分散法を用いて次の式の値を計算する。

$$\frac{\sum(O(i, t)^2)}{m} \quad (3)$$

上式の値が一番小さい $i$ をこのトリップパターンの代表的なO点とみなし、この点 $i$ の経度と緯度をこのトリップパターンのO点の代表的な位置にする。D点グループ

も同様である。

#### (2) 代表O点出発時間の計算

ユーザー毎に、3章にて抽出した各トリップパターンのO点の代表的な出発時間を計算する。ここで、各トリップパターンに含まれるO点の平均出発時間をこのトリップパターンの代表的な出発時間とした。

#### (3) 定常性を持つトリップパターンの抽出

4(1)と4(2)で計算した各トリップパターンの代表的なO、D点の位置と代表的なO点の出発時間を用いて、ユーザー毎に連続した数ヶ月間に定常性を持つトリップパターンを選定する。判断の基準として、代表O点間の距離は500m以内、代表D点間の距離は500m以内、O点の代表出発時間の差が30分以内とし、この3つの条件を全て満足すれば、当該ユーザーの分析対象期間における定常性を持つトリップパターンとして抽出する。

### 5. トリップの予測と精度の解析

#### (1) トリップの予測

4章までのアルゴリズムを用いて選定した、ユーザー毎の定常性を持つトリップパターンに含まれるトリップをさらに曜日毎に分別し、各曜日に発生したトリップ数が最少トリップ回数 $\text{minnum}$ より多ければ、当該ユーザーの予測対象トリップとする。これによって得られた曜日別のトリップパターンを4(1)と4(2)と同様な方法を利用し、代表的なO、D点の位置、また代表的なO点の出発時間を計算する。

ユーザー毎の予測トリップを用いて、翌月の実際に発生したトリップと曜日別に比較し、予測トリップの代表O点と実際に発生したトリップのO点間距離が500m以内、予測トリップの代表D点と実際に発生したトリップのD点間距離が500m以内、予測トリップのO点の代表出発時間と実際に発生したトリップの開始時間との差が30分以内、といったこの3つの条件を満足すれば、トリップ予測に正解したと見なす。

また本研究では、上述した最少トリップ回数 $\text{minnum}$ の設定値を最初に3回/月と設定し、後ほど2回/月に変化させて予測精度への影響を考察した。

#### (2) 予測精度の解析

ある月において、ユーザー $i$ の予測トリップ総数を $A_i$ とし、その月に予測が当たったトリップ数を $B_i$ とすると、その月のユーザー $i$ の予測精度は次の式で表せる。

$$\frac{B_i}{A_i} \quad (4)$$

また、その月の全ユーザーの予測精度は次の式で表せる。

$$\frac{\sum(B_i)}{\sum(A_i)} \quad (5)$$

本研究では、4月から10月までで抽出された定常性を持つ

トリップパターンを予測トリップとし、11月の実際に発生したトリップと比較する。

#### a) データ学習期間の変化による精度への影響

本研究では、2011年4月から2011年11月まで合計8ヶ月間のデータを使用するため、データ学習期間を6種類に分ける。表-2にはデータ学習期間とその概要を示す。

また、最少トリップ回数minnumの値は、3回/月と2回/月の設定により変化するため、ここで、最少トリップ回数minnumの概要を表-3に示す。

図-2と図-3に最少抽出点数minpoint = 4, 最少トリップ回数minnumは3回/月の条件で得られた結果を示す。

図-2では、横軸がデータ学習期間、縦軸が予測精度を示している。図-3では、横軸がデータ学習期間、縦軸が定常トリップ抽出ユーザーの割合を示している。定常トリップ抽出ユーザーの割合というのは、実際に予測トリップが抽出可能であったユーザー数を11月の参加ユーザー総数で割ったものである。

図-2から、データ学習期間が長くなることに伴い、予測の精度が向上していることがわかる。これに対して、図-3から定常トリップを抽出できたユーザーは、データ学習期間が長くなるとともに、減少していることがわかる。予測精度を高めることが重要であるが、実用性から

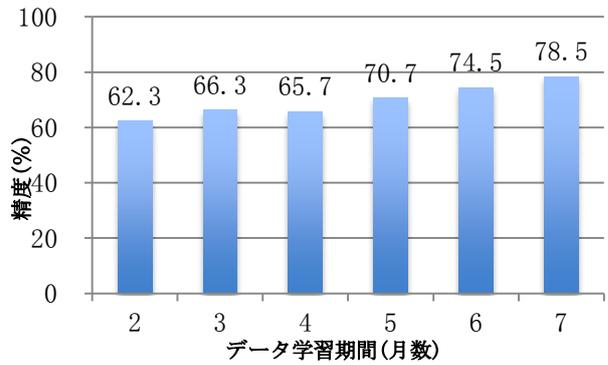


図-2 データ学習期間の変化による精度の変化

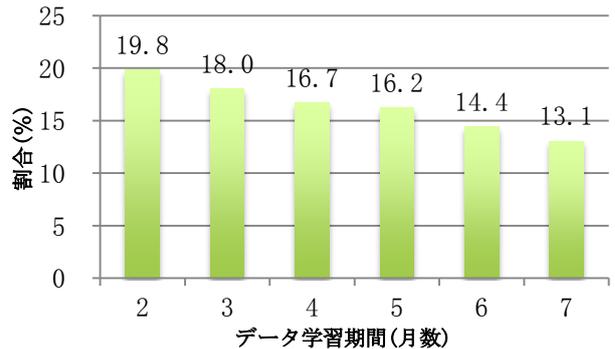


図-3 データ学習期間の変化による定常トリップ抽出ユーザー数の変化

みると、定常トリップを抽出できたユーザーが少なくなると、広範囲に適用することができなくなってしまう。そのために、予測精度と定常トリップを抽出できるユーザー数の両方も配慮しながら、データ学習期間を設定する必要があると考える。

表-2 データ学習期間の概要

表示方法	データ学習期間(月数)	概要
2	9月-10月	9月から10月までの定常性を持つトリップパターンを予測トリップとし、11月の実際に発生したトリップと比較
3	8月-10月	8月から10月までの定常性を持つトリップパターンを予測トリップとし、11月の実際に発生したトリップと比較
4	7月-10月	7月から10月までの定常性を持つトリップパターンを予測トリップとし、11月の実際に発生したトリップと比較
5	6月-10月	6月から10月までの定常性を持つトリップパターンを予測トリップとし、11月の実際に発生したトリップと比較
6	5月-10月	5月から10月までの定常性を持つトリップパターンを予測トリップとし、11月の実際に発生したトリップと比較
7	4月-10月	4月から10月までの定常性を持つトリップパターンを予測トリップとし、11月の実際に発生したトリップと比較

表-3 最少トリップ回数 minnum の概要

データ学習期間(月数)設定	2	3	4	5	6	7
3回/月	6	9	12	15	18	21
2回/月	4	6	8	10	12	14

## b) 予測前の位置確認による精度の向上

トリップを予測する際に、実際に発生したトリップの起点が予測トリップの起点と遠く離れるために予測が外れてしまうといった傾向にあることが分かった。したがって、予測精度を向上させるために、トリップの予測を開始する前にユーザーの位置を確認する手順を追加した。

具体的には、毎回トリップの発生を予測する前に、そのユーザーが実際に実施した直前のトリップの終点位置を確認するといった方法である。直前トリップの終点位置が予測トリップの起点と600m以上離れている場合には、定常トリップ発生を予測を行わない。予測する前の位置確認による精度の変化を以下の図に示す。なお、最少抽出点数minpointは4、最少トリップ回数minnumは3回/月に設定した。

図-4から、データ学習期間に関わらず、予測する前にユーザーの位置を確認することにより、予測の精度が向上したことがわかる。さらに、図-5には定常トリップ抽出ユーザーの割合を示している。位置確認を行わない場合と比較すると、定常トリップを抽出できたユーザーは少し減少したが、これは位置確認を行う際に上述した条件を満たさない場合に予測が行わなかったため、実際に運転者へ道路情報を送信しないと判断する。したがって、予測する直前のユーザーの位置確認は、予測精度の向上に有効と言える。なお、以降の分析はすべて予測する前の位置確認を考慮した結果を示している。

## c) 最少トリップ回数minnumの変化による精度への影響

以上までに最少トリップ回数minnumを3回/月に設定し

た場合について述べた。ここでは、最少トリップ回数minnumを2回/月に変化させ、予測精度への影響を考察する。その結果を以下の図-6、図-7に示す。なお、最少抽出点数minpointは4に設定した。

図-6より、最少トリップ回数minnumの減少に伴い、予測の精度が低減していることがわかる。加えて、観測期間が長くなるとともに、その差が大きくなることがわかる。これに対して、図-7から予測できたユーザー数は、最少トリップ回数minnumの減少に伴い、増加していることがわかる。これは制約が緩和されるため、定常トリップを抽出できたユーザーが増加したことが考えられる。

## d) 最少抽出点数minpointの変化による精度への影響

本研究では、最少抽出点数minpointを3、4、5へ変化させ、予測精度の推移を検証する。その結果を図-8、図-9に示す。なお、最少トリップ回数minnumは3回/月に設定した。

図-8より、データ学習期間に関わらず、最少抽出点数minpoint=3と4の場合の精度がほぼ一緒であることがわかる。しかしながら、最少抽出点数minpointは5に設定すると、精度が低下してしまうことがわかる。さらに、データ学習期間を長くするにつれ、最少抽出点数minpointが大きいほど、定常トリップを抽出できたユーザーは減少したことを、図-9からわかる。以上のことを踏まえ、最少抽出点数minpointは比較的小さい値に設定することで、より多くのユーザーを予測可能になり、予測精度も向上することが期待できる。

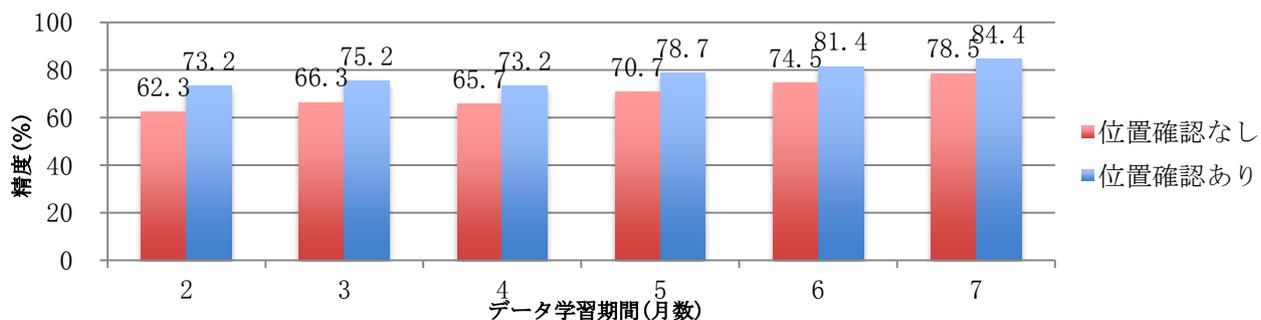


図-4 予測前の位置確認による精度の変化

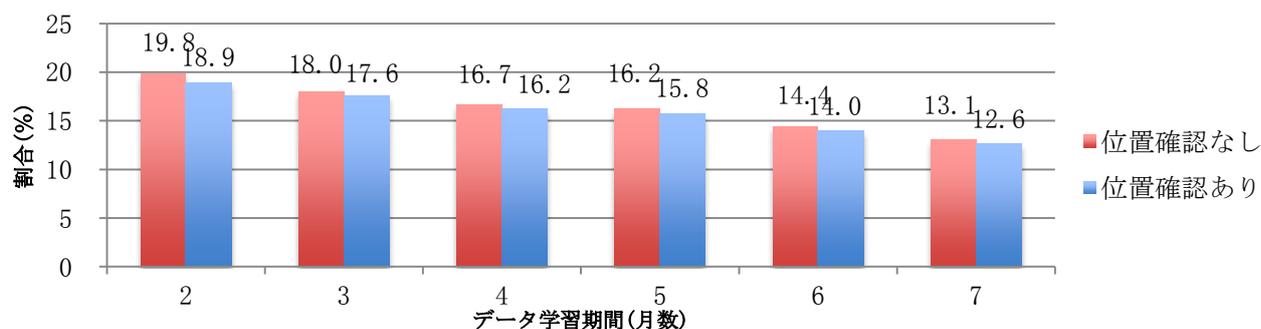


図-5 予測前の位置確認による定常トリップ抽出ユーザー数の変化

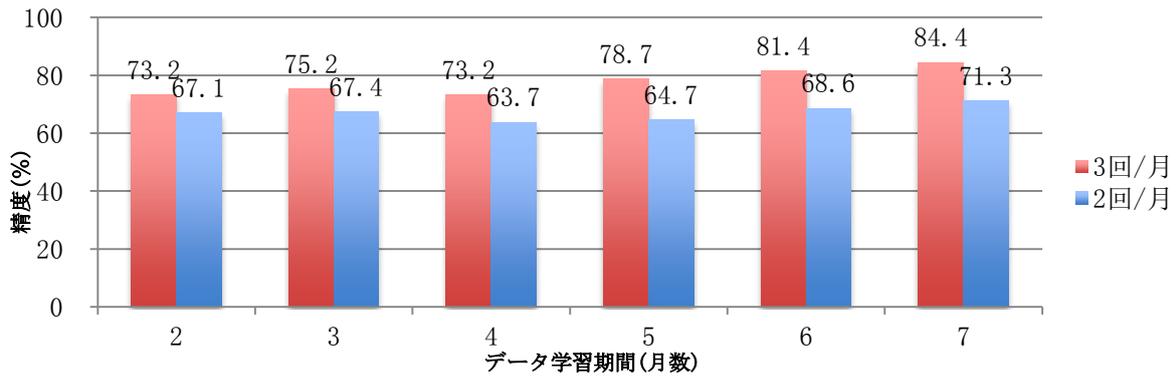


図-6 最少トリップ回数 minnum の変化による精度の変化

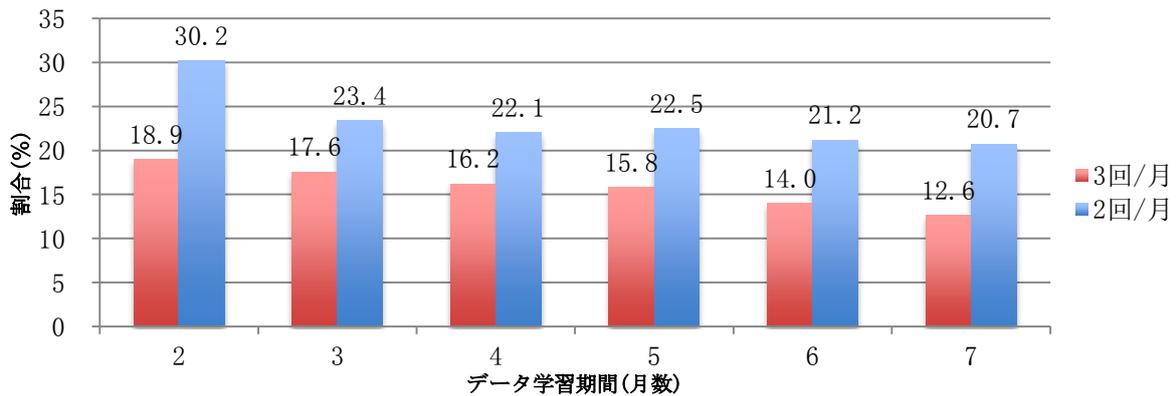


図-7 最少トリップ回数 minnum の変化による定常トリップ抽出ユーザー数の変化

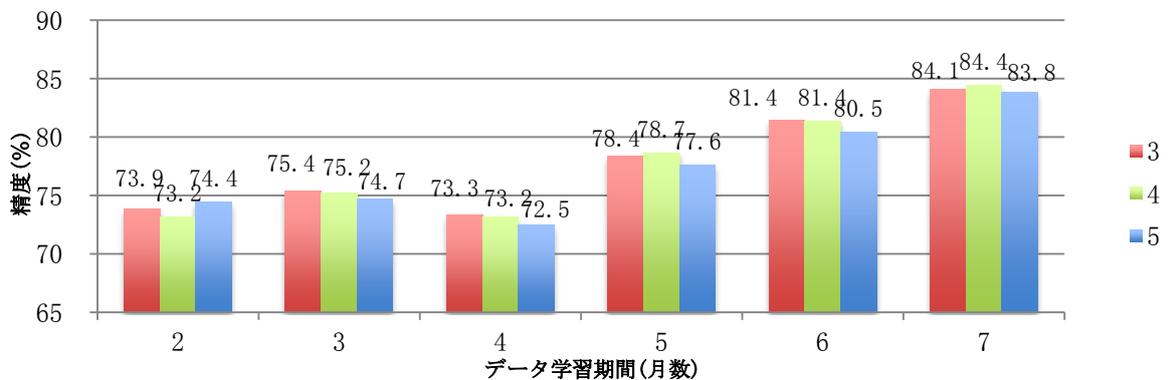


図-8 最少抽出点数 minpoint の変化による精度の変化

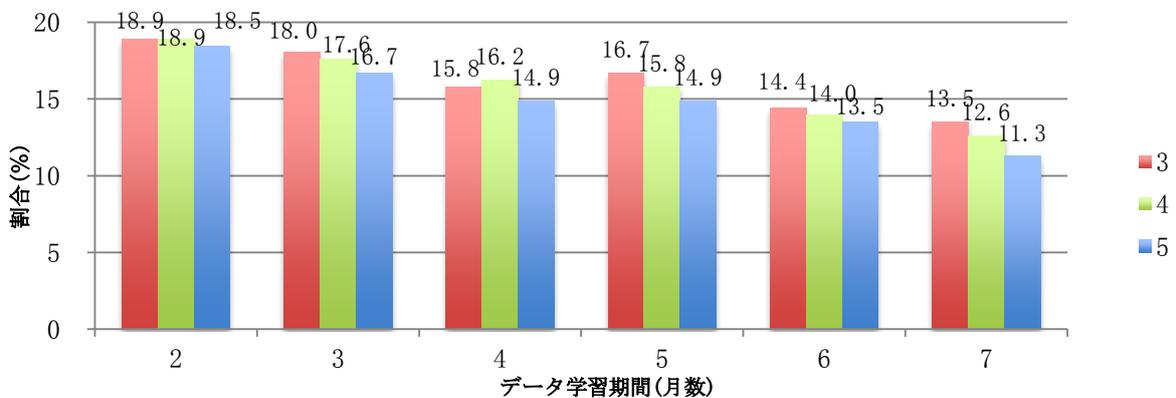


図-9 最少抽出点数 minpoint の変化による定常トリップ抽出ユーザー数の変化

## 6. 終わりに

本研究では、プローブ情報収集技術により収集した、個々の運転者の自動車利用データを用いて、ユーザー毎に自動車の利用パターンを推定するためのアルゴリズムを構築した。また、これらの定常性を持つトリップパターンを対象として、ユーザー毎に将来のトリップ発生を予測した。さらに、予測したトリップを実際に発生したトリップと比較することにより、予測の精度を検証することができた。

本研究で使われたデータは、8ヶ月間にわたる長期間に観測した上で収集されたものである。これを踏まえた上で、トリップの予測精度を検証するとき、データ学習の期間を変化させて精度への影響を考察できた。研究の結果により、データ学習期間を長くすることに伴い、予測の精度が向上していることがわかる。これに対して、定常トリップを抽出できたユーザー数は、データ学習期間が長くなるとともに、減少していることがわかる。

また、予測精度を向上させるために、本研究ではトリ

ップの予測を開始する前にユーザーの位置を確認する方法を考えた。この方法により、予測の精度がおよそ5% - 10%向上したことがわかる。さらに、本研究では、ユーザー毎の定常性を持つトリップパターンを抽出する際に、最少抽出点数`minpoint`と最少トリップ回数`minnum`、この二つのパラメーターを用いた。トリップの予測精度を検証するとき、`minpoint`と`minnum`を変化させ、精度への影響を考察できた。

### 参考文献

- 1) Froehlich, J., F. and Krumm, J. : Route Prediction from Trip Observations, *Society of Automotive Engineers SAE 2008 World Congress*, 2008.
- 2) Simmons, R., Browning, B., Zhang, Y. and Sadekar, V. : Learning to Predict Driver Route and Destination Intent, *Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE*, pp.127-132, 2006.

(2014. 8. 1 受付)

## EXTRACTION OF RECURRENT CAR TRIP PATTERNS FOR PERSONALIZED PUSH-TYPE ROAD TRAFFIC INFORMATION PROVISION

Shuyang JIA, Toshiyuki YAMAMOTO, Takayuki MORIKAWA

Current personal road traffic information providing systems, such as car navigation, require drivers manually input the destination. On the other hand, road traffic information such as VICS is not sufficiently personalized. In this research, an algorithm was developed for analyzing individual car users' mobility patterns, based on GPS observations of the vehicles' past trips. By extracting recurrent car trip patterns, this study was able to predict users' future trips and verify the accuracy as well. In addition, this study also shows the effects of learning periods and checking user's location on the accuracy.