

ライフログのための位置情報ログデータからの移動モード判定 の検討

青木 政勝[†] 瀬古 俊一[†] 西野 正彬[†] 山田 智広[†] 武藤 伸洋[†] 阿部 匡伸[†]

[†]NTT サイバーソリューション研究所 〒239-0847 神奈川県横須賀市光の丘 1-1

E-mail: [†]{aoki.masakatsu, seko.syunichi, nishino.masaaki, yamada.tomohiro, muto.shinyo,
abe.masanobu}@lab.ntt.co.jp

あらまし 近年、GPS などの位置計測が身近なものになり、個人の位置情報を記録することが容易になった。ユーザが移動した軌跡となる位置情報のログデータはライフログにおいて、最も基本的なデータであり、購買履歴や視聴履歴、写真などの様々なデータと位置情報を組み合わせて蓄積し、ライフログとして活用するサービスが検討されている。一方、ユーザの位置情報に基づく様々な情報の提供が実用化されてきている。現状は現在位置だけであるが、その背後にあるユーザの過去の情報、すなわちライフログと結びつけることで、ユーザの行動特性やライフスタイルを反映した情報の提供が行えるようになる。本研究では基本的なデータである位置情報をライフログとして蓄積することにより、ユーザ状況やプロフィールを推定することを目的としている。ユーザがおかれた状況として、まず位置履歴から移動中と滞在中といった行動モードを切り分ける手法を検討する。

キーワード ライフログ、GPS、位置情報、移動手段

An estimating method for activity modes using location data

Masakatsu AOKI[†] Shunichi SEKO[†] Masaaki NISHINO[†] Tomohiro YAMADA[†] Shinyo MUTO[†]
and Masanobu ABE[†]

[†]NTT Cyber Solutions Laboratories 1-1 Hikarinooka, Yokosuka-shi, Kanagawa, 239-0847 Japan

E-mail: [†]{aoki.masakatsu, seko.syunichi, nishino.masaaki, yamada.tomohiro, muto.shinyo,
abe.masanobu}@lab.ntt.co.jp

Abstract In recent years, it has become familiar a positional measurement such as GPS, and users can easily record there location information. The location data which include moving tracks is the most basic data in life-log. The location data can be combined with various data, purchase histories, attention histories, photographs, and so on. Those data attempt to use for life-log service. While current services only use user's location, more variable services can be provided using user's behavioral trait and lifestyle which can be estimate by user's life-log. Our research aims to estimate user's situation and their profile from accumulated location data. In this paper, we examine the method of extracting activity modes, that is moving or staying, from positional histories.

Keyword LifeLog, GPS, Location information, Transportation mode

1. はじめに

近年、情報機器の高度化により、個人の生活をデジタル化するライフログが注目を集めている。ユーザがパソコンや携帯電話などで閲覧、作成したウェブやメール、各種文書、音楽などの日常操作したファイルやデータなどの情報を蓄積したり、ウェアラブル機材を用いてユーザの日常生活を記録し、後日の振り返りを支援する取り組みがされている[1][2]。ユーザ生活のデ

ータを記録するためには様々なセンサを用いる。各種センサから取得したデータは、そのデータだけではどのような意味があるのかわからない生データであり、ライフログとして利用する場合はこの生データを構造化、言語化して意味があるデータとすることが重要である。

一方、ウェブ上の日記として幅広く普及しているブログの中にライフログが取り入れられつつある。プロ

グは通常ユーザ自らの意図で内容を記入しているが、ユーザの生活をロギングしたデータから自動的にブログを生成する試みがいくつかなされている [3]。さらにライフログサービスとしては、ユーザの日々の生活ログからユーザの行動や好みを予測し、的確な情報をレコメンドするサービスが考えられる。また行動を予測する場合、ユーザがどこにいたかをより正確に把握するためには GPS などのセンサや無線 LAN の電波強度 [4] などを利用して位置情報を取得する必要がある。

位置情報はライフログの基本的な情報であり、位置情報のログからユーザがよく滞在した場所を有意位置として抽出することを主眼としたものが多く取り組まれている [5][6]。しかしながら、滞在と滞在の過程である移動中も重要なライフログデータの一つであり、ユーザに対するレコメンドサービスを提供することを想定した場合、そのユーザが直前までどのような経路で移動したか、どのような移動手段で現在地まで来たのかがわかれば、よりの確かな情報を提供することが可能となる。そこで、本稿では一般的に普及しつつある GPS によって位置情報を取得し、その情報だけを利用して、移動手段の判定の検討を行う。

2. 従来技術

位置情報の履歴から移動手段を判定する代表的な手法として、速度の違いを利用するものと地図データを利用するものの 2 つの手法がある。

速度の違いを利用する手法は、取得する位置情報の各点間の位置差分から移動速度を算出し、それらを比較することで、徒歩か乗り物で移動しているかの判定を行うものである。例えば [7] では、徒歩と車、電車のそれぞれの移動速度の取りうる範囲が異なることに着目し、かつ GPS 計測で生じる誤差について角速度を用いてフィルタリングして、移動手段の判定を行っている。

一方、地図データを利用する手法では、取得した位置情報の軌跡から、軌跡の各座標をベクトル地図上にマッピングし、地図データと照合して移動手段を判定する。例えば、ある一定速度以上の軌跡において、その座標に路線があれば電車、道路があれば車で移動したと判定することができる。しかしながら、地図データを利用して移動手段を判定するためには、端末あるいはネットワーク上のサーバにベクトル地図を保持し、最新状態にしておく必要があり、維持管理の運用にコストがかかる。また、ネットワークを介してサーバ上で判定を行う場合、ユーザの移動軌跡をサーバへ送出手のため、プライバシーの問題も懸念される。

3. アプローチ

本稿ではネットワーク上に移動軌跡を送出することなく、かつローカル端末内で地図データを利用せずに移動手段を判定することを目指す。位置情報以外のセンサ情報、例えば加速度センサを用いて移動手段を判定する手法もいくつか提案されているが [8][9]、本稿では GPS センサから得られる測位データのみを利用して、高精度でかつ周辺環境にロバストに移動手段を判定することを目的とする。

3.1. GPS データの分類モデル作成

GPS による収集データから算出できる移動情報のうち、一定時間区間における平均速度、平均加速度、最高速度、最高加速度に着目し、移動手段に対して各パラメータがどのように影響するのか、時間区間が移動手段の判定精度にどのように影響するのかを分析する。それにより、高精度な移動手段判定を行うために影響度が大きいパラメータおよび適切な分類モデルを決定する。

移動手段として、他手段と明確に区別できるのは、徒歩であり、移動速度は概ね 4Km/h 程度で一定の速度で移動することが想定される。また、電車の場合は駅での停車を除いて一定の移動速度で移動し、発車直後や減速開始時の加速度が一定であると想定される。移動速度は比較的大きく最大 80Km/h 程度となることが想定される。一方、自動車の場合は信号待ちなど道路状況により停止と移動が繰り返されるため、移動中の速度変化が大きく、電車と異なり加速度も一定ではないと想定される。移動速度は電車と比較して低く、変動範囲は 10Km/h ~ 60Km/h となることが想定される。

以上のような各移動手段の特性から、先に示した 4 つパラメータを用いた分類モデルを作成する。ライフログの各データは様々な視点から検索可能にすることが重要であると考え、分類モデルを作成する手法として、判定パラメータを明示できる決定木による手法を採用する。

3.2. 行動パターン特性の利用

移動手段を判定する場合、誤判定は少なからず生じてしまう。例えば、ある一定区間内の移動速度の平均値を比較して、電車と車の移動手段の判定を行う場合、移動手段が電車のとき駅などで停車中あるいはその前後において、移動速度は低速となり、車と誤判定してしまうことがある。このとき移動手段の判定は表 1 に示すような通常ではありえない移動手段の変化となる。

表 1: 移動手段判定例

時刻	移動手段判定
...	...
15:04:01	walk
15:04:31	train
15:05:01	train
15:05:31	car
15:06:01	train
15:06:31	train
....	...

このような誤判定を自動的に補正するために人の行動パターンに着目する。人の行動を「滞在・徒歩」と「移動」の2つに分けると、滞在と移動を交互に繰り返しており、移動と移動の間は滞在あるいは徒歩になる。例えば、電車からバスやタクシーへの乗り換える際、一旦電車を降りて次の移動手段の乗り場まで徒歩で移動し、次の移動手段に乗り込むまでは滞在することになる。つまり、異なる移動手段を連続して変えることはなく、移動手段の前後には必ず滞在あるいは徒歩が行われるということが、人の行動パターンの特性と仮定できる(図1参照)。この仮定を適用し、移動手段の誤判定を補正する方法には確率的にモデル化する手法もあるが、本稿ではまずルールベースによって補正する。移動手段を判定する前処理として、まず移動か滞在(徒歩を含む)かの識別を行い、移動と識別した区間内で判定された移動手段を最頻の手段で統一するというルールを定義する(図2参照)。これにより、移動手段判定の精度向上が期待できる。

4. 予備実験

4.1. データ収集

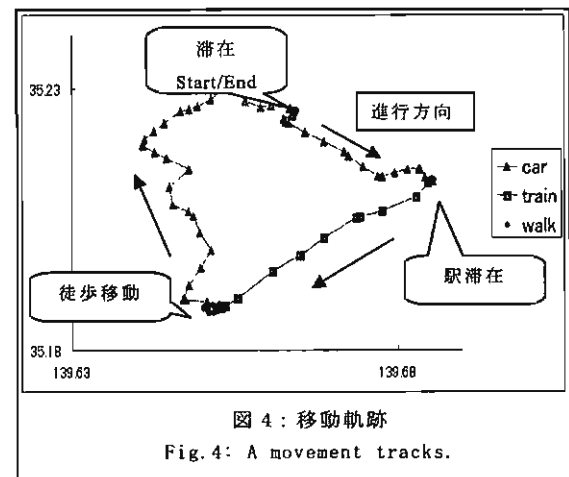
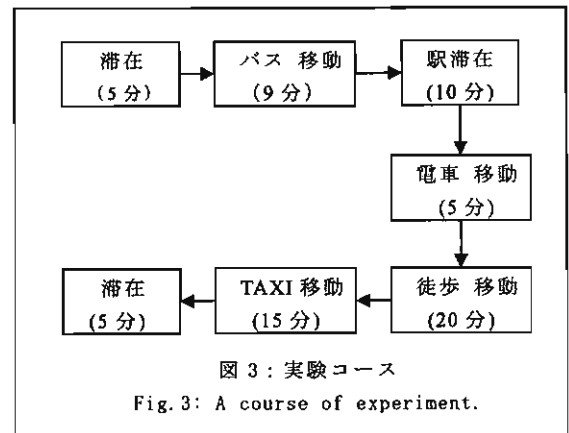
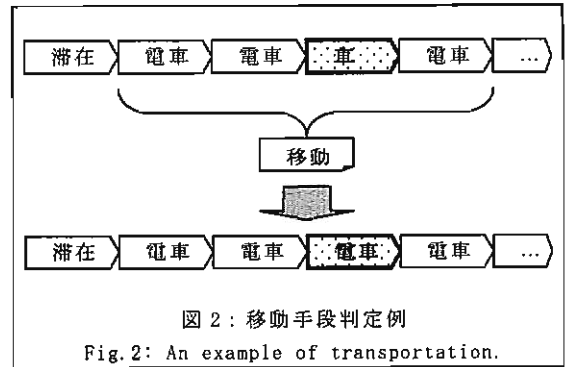
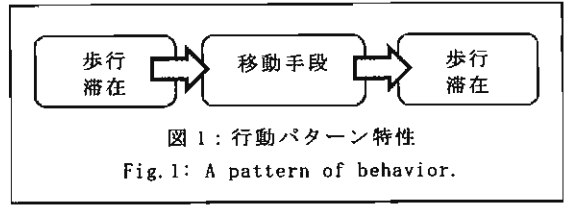
移動手段を判定する決定木の分類モデルを作成するために、位置データ収集を行った。位置を計測するGPSセンサは、市販のハンディ型GPS端末を用いて1秒毎に位置情報(緯度、経度)と時刻を収集した。本手法で判定する移動手段は

- ・滞在および徒歩
- ・自動車(バスやタクシーなど)
- ・電車

の3種類を対象とする。データ収集実験では、各移動手段を効率よく利用できるような設定し、図3に示すような1時間程度の移動行動を行った。図4は実際に収集した位置データを緯度経度平面にプロットした軌跡の例である。

4.2. 決定木モデル

収集した位置データに対し、時間区間を10秒、30秒、60秒、90秒、120秒として、3.1節で述べた4つのパラメータを算出した。複数の時間区間についてパ



ラメータの算出を行うのは、乗物での発車や停車を考慮した場合、時間幅が移動手段の判定に影響を与える

と想定されるためである。その後、データマイニングツール WEKA[10]を利用して、各時間区間での決定木モデルを作成した。一例を図5に示す。WEKAで作成した決定木モデルはそのままでは階層数が多いため、出現頻度が低い事象を例外と判断し、2階層程度に閾値を集約して決定木モデルを修正した。

表2：決定木モデルの閾値

	滞在と移動の識別	自動車と電車の判定
10秒	最高速度 2.92 m/s	最高加速度 1.21 m/s ²
30秒	最高速度 3.55 m/s	最高速度 17.79 m/s
60秒	平均速度 1.68 m/s	最高速度 16.34 m/s
90秒	最高加速度 2.12m/s ²	最高速度 17.79 m/s
120秒	平均加速度 0.012m/s ²	最高速度 17.79 m/s

5. 分析と考察

表2に各時間区間の決定木モデルの閾値を示す。表2から、滞在・徒歩と移動の識別は速度あるいは加速度を用い、自動車と電車の判定はほとんどの場合最高速度によって分類できることがわかる。この決定木モデルを用いて移動手段の判定を行った結果を表3に示す。各時間区間の全体および各移動手段について、実際に移動していた時間と当該移動手段と判定した判定率を算出した。

表3：移動手段判定結果 (単位：%)

	全体	バス	電車	TAXI
10秒	89.11	84.21	42.42	92.63
30秒	89.33	63.16	63.64	84.36
60秒	92.00	90.00	66.67	87.50
90秒	90.00	100.00	75.00	90.00
120秒	86.49	80.00	66.67	75.00

表3において時間区間が10秒の判定結果が相対的に低い。特に電車の場合は誤判定が50%以上となる。これは、判定する時間区間が短い場合、電車の発着時や自動車における信号待ちなど低速移動時を判定することになり、速度値だけでは徒歩と誤判定してしまうことが要因であるといえる。また時間区間を長くすれば、低速移動時が前後の高速移動時のデータ値と平滑化されるため、速度値だけで滞在と移動の判定ができるようになることがわかった。しかしながら、時間区間が長くなりすぎると逆に平滑化しすぎて、各移動手段の特徴の差がなくなり判定率が下がる。図6に各時間区間での各パラメータ値を時系列でプロットしたグ

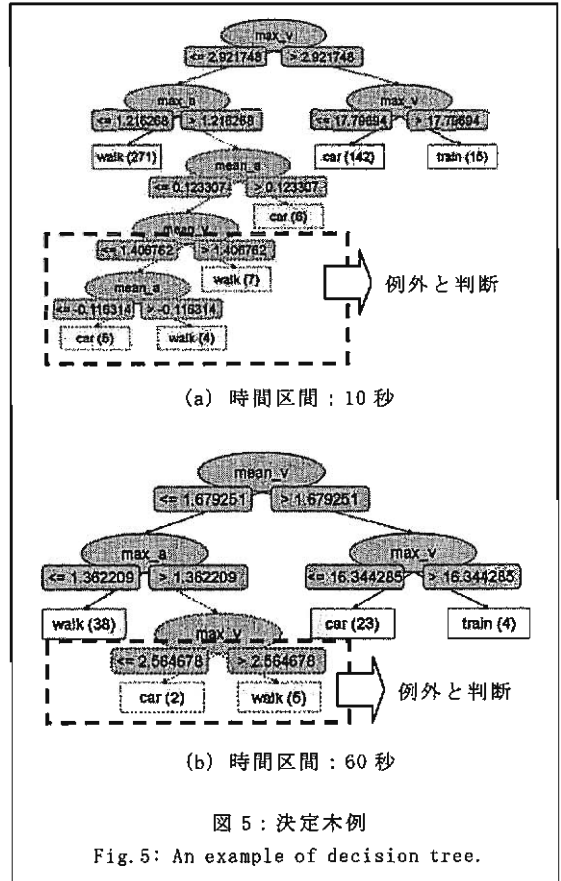
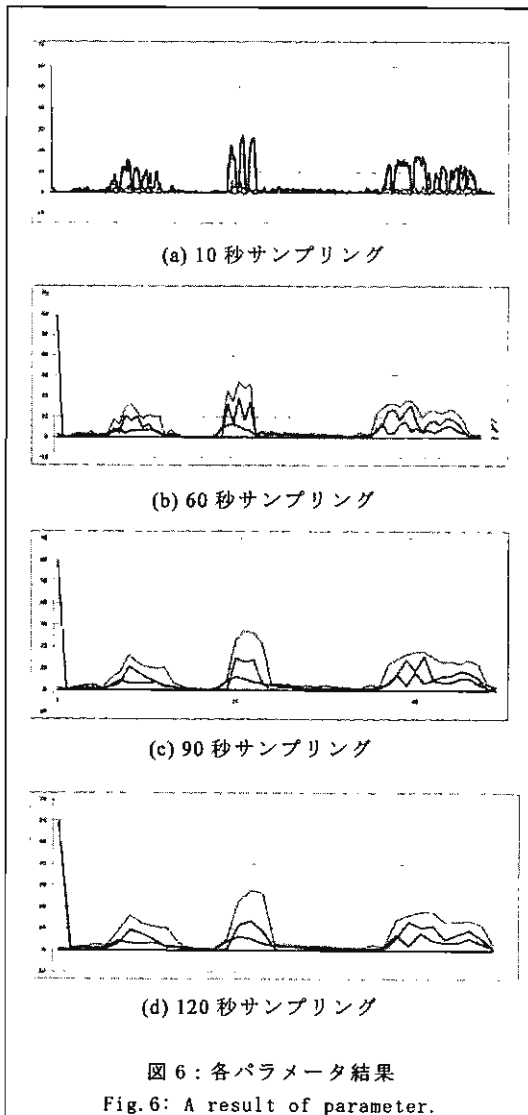


Fig.5: An example of decision tree.

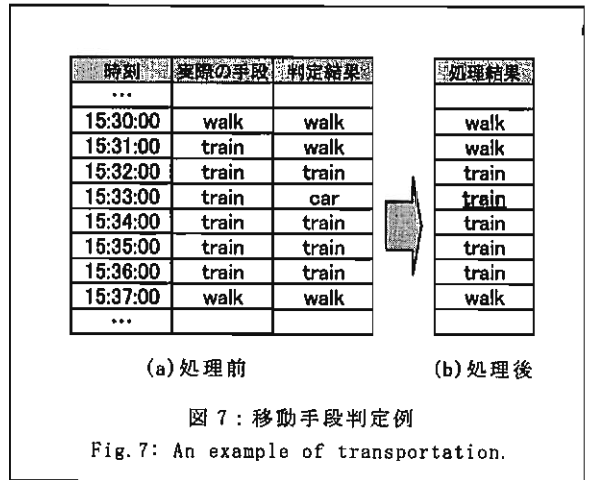
ラフを示す。10秒間隔の場合はパラメータの変動が大きいが、時間区間が長くなるとグラフが平滑化されていることがわかる。今回の結果から、60秒から90秒の時間間隔が各パラメータの特徴を保持しつつ、データの平滑化が行え、誤判定を最も少なくできる間隔であることがわかった。これは、今回のデータ収集コースにおいて、電車やバスが駅やバス停で停車する間隔が1分ないし1分半であったことから、その間隔でサンプリングすることが判定に影響すると予想できる。

加速度については予想とは異なり、自動車と電車の移動手段の判定に大きな影響を与えず、速度だけで判定できることがわかった。しかしながら、速度だけでは十分な判定精度が得られない。例えば判定精度が比較的良好い60秒や90秒の時間間隔であっても、本来電車で移動していたところをバスやタクシーなどの車での移動と誤判定した割合が約30%であった。このような誤判定に対し3.2節で述べた行動パターン特性の仮定を利用した補正処理を行った。時間区間を60秒とした際の判定結果のうち、電車部分は図7(a)になる。「15:33:00」の移動手段の判定結果が「car」と誤判定されている。これに対し、移動部分における移動手段



の最頻手段で統一する処理を実施すると、図 7 (b) になる。この手法では移動手段が切り替わったタイミングの時間帯 (15:31:00) の補正は行えないが、移動中の手段の判定が全て「train」となり、正確な移動手段判定が行えるようになる。

この処理を表 3 の判定結果に適用して、誤判定を補正した結果を図 8 に示す。時間区間が 90 秒の場合、バス、電車およびタクシーの各移動手段の判定が 100% に達するなど、ほとんどの時間区間、移動手段において判定精度が向上した。以上から、行動パターン特性を利用することによって移動手段判定の精度が向上できることが確認できた。



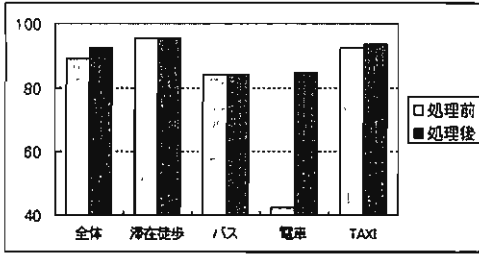
6. まとめ

ネットワーク上に移動軌跡を送出することなく、地図データを利用せずに、GPS のログデータのみから移動手段の判定ができることが確認できた。速度などの移動情報を用いて判定する場合、電車と自動車の判別が困難であるが、行動パターンの特性を利用することで、判定精度を 90% 程度に向上することが確認できた。

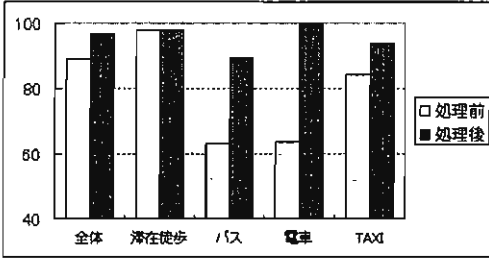
本稿では徒歩や自転車などの低速な移動手段の判定を行わなかったが、今後はそれらを含めた他手段などの判定を可能とするなど判定精度を向上させていく予定である。また移動手段だけではなく、ユーザが高頻度に滞在する場所の抽出や位置情報の日常と非日常の判別など、位置情報からのライフログデータの抽出手法を検討していく予定である。

文 献

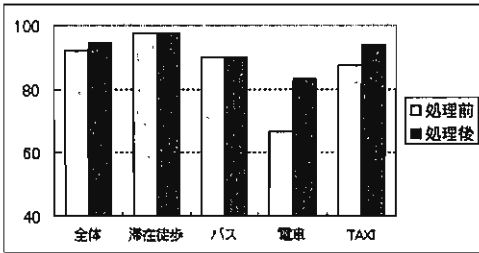
- [1] J. Gemmell, et al., "MyLifeBits: A Personal Database for Everything", *Communications of the ACM*, Vol.49, No.1, pp.88-95, 2006.
- [2] 相澤, "体験の情報処理", 電子情報通信学会技術研究報告, PRMU2003-285, pp.1-9, 2004.
- [3] 沼 他, "学会会議における体験共有のための行動履歴に基づく Weblog システム", 情報処理学会論文誌, Vol.48 No.1, pp.85-97, Jan. 2007.
- [4] 厩本 他, "WHEN-becomes-WHERE: WiFi セルロギングによる継続的位置履歴取得とその応用", *インタラクション 2007*, pp.223-230, 2007.
- [5] D. Ashbrook, et al., "Using GPS to Learn Significant Locations and Predict Movement Across Multiple Users", *Personal and Ubiquitous Computing*, Vol.7, Issue 5, pp.275-286, 2003.
- [6] B. Adams, et al., "Extraction of Social Context and Application to Personal Multimedia Exploration", *Proc. of the 14th annual ACM int. conf. on Multimedia*, pp987-996, 2006.
- [7] 前司 他, "GPS 携帯端末による交通モード自動判定法の開発", *ITS シンポジウム 2005*.



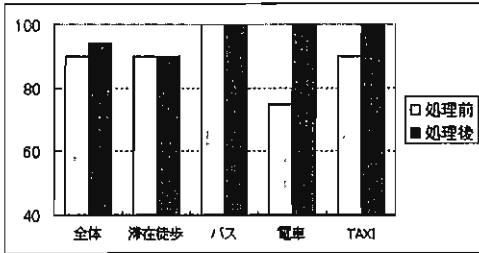
(a) 時間区間：10 秒



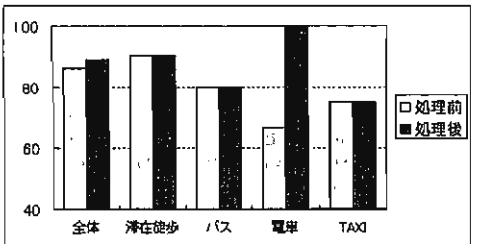
(b) 時間区間：30 秒



(c) 時間区間：60 秒



(d) 時間区間：90 秒



(e) 時間区間：120 秒

図 8：移動手段判定結果

Fig. 8: A result of estimating for activity modes.

- [8] 田名部 他, “移動体端末による交通行動調査の成立の可能性”, 土木計画学シンポジウム 2001.
- [9] 小林 他, “駅迎：携帯電話を用いたユーザ移動状態推定・共有方式”, 電子情報通信学会技術研究報告, MoMuC2008-20, pp.115-120, 2008.
- [10] WEKA, <http://www.cs.waikato.ac.nz/ml/weka/>