

# パーソナルデータの統合による個々人の快適な交通行動の検討

種村京介, 松原剛, 金杉洋, 柴崎亮介

## A Study on the Quality of Transportation by Using Combined Personal-Data

**Kyosuke TANEMURA, Go MATSUBARA, Hiroshi KANASUGI,  
and Ryosuke SHIBASAKI**

**Abstract:** Due to recent development of smart devices, personal data are easily recorded in many aspects. However, not all of the personal data are used effectively on individual basis. Therefore, the purpose of this research is to improve individual living standards, by using combined and analyzed personal data acquired from several devices and services. In this research, we considered on suggesting the higher quality transportation from personal traffic behavior tendency derived from personal data such as GPS log.

**Keywords:** 移動(transportation), ライフマネジメント(life management), パーソナルデータ(personal data)

### 1. 背景と目的

近年のスマートフォン、タブレット端末などのスマートデバイスの普及により、個人に関するデータは様々な切り口から観測され、蓄積することが可能となった。事実、世界経済フォーラム(World Economic Forum)が2011年1月に公表した報告「Personal Data: The Emergence of New Asset Class; パーソナルデータ: 新たな資産カテゴリーの出現」では、2020年のデジタルデータの量は、2009年の44倍になるであろうと予測した。また、パーソナルデータはデジタル世界における新しい通貨になるとし、その効果や経済価値に言及している。わが国ではこの現状を踏まえ、平成28年にIT総合戦略本部の世界最先端IT国家創造宣言にて、パーソナルデータの利活用による新サービスの創出から、個人の多様なニーズへの適応や社会全体での利便性の向上を推進する必要

があると明記されている。

これらの背景から、パーソナルデータを個人レベルで分析し、個々人のライフマネジメントにおいて活用することの重要性が伺える。そこで本研究は、都市圏の日々の生活に欠かせない“鉄道移動”に着目し、パーソナルデータの活用を検討した。平成27年度全国都市交通特性調査結果によると、三大都市圏および地方都市圏での鉄道利用率は平成17年から平成27年にかけて増加しており、鉄道移動の質を向上することは人々のQOL向上においては必須であると言える。また、近年MaaS(Mobility as a Service)などの登場により、人々の移動はよりサービスとしての側面を持つようになってきた。そのためこれまで鉄道利用時に経路選択する条件として移動時間やコストの観点のみで判断されていたが、移動の質として、鉄道利用時の快適性を把握することは、より快適な鉄道移動を提供する上で、今後必要となる。

そこで本研究では、位置情報をはじめとした多様な個人データから、個人にとって最適で快適な

移動の傾向を分析し、そこから鉄道利用時の快適性を推定・検討する。

時間帯名	通勤・通学ラッシュ					オフピーク					帰宅ラッシュ										
乗車時間帯(時)	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
表記	-2	-1	0	+	+								-1	0	+	+					+3
				1	2										1	2					

図1 時間帯の定義

## 2. 実験およびデータ操作・分析

### 2.1 パーソナルデータ

本研究では、スマートフォンのGPS機能から得られる移動軌跡データ、ヘルスケアデバイスであるFitbit Charge2によって測定された心拍数と歩数、そしてGoogle Calendarにて登録された各日の予定をパーソナルデータとして用いた。なお、登録された日々の予定から、鉄道利用の理由を「通勤・通学」、「仕事・営業」、「行楽」、「帰宅」、「その他」に分類した。

### 2.2 鉄道利用時の快適性の記録

データ提供者は、実験期間内で鉄道に乗車した際に、「車内の混雑度」、「車内の快適度」、「着席の有無」、「駅までの移動手段（徒歩、自転車、自動車）」を記録している。車内の混雑度は、100%以下、100%、125%、150%、175%、200%の六段階から主観的判断で申告し、その時間帯の混雑度とした。また、車内の快適度は、1（非常に不快）から5（非常に快適）の5段階評価で回答している。

### 2.3 その他データ

これらのデータの他に、鉄道利用時に考えるストレス要因をデータ化した。本研究では、その他データとして、乗車時の「時間帯」、「気温」、「直近10日間の平均気温との気温差」、「天候」、「飲酒の有無」、「遅延の有無」、データ取得時までの「累計乗車時間」、「累計乗り換え回数」を用いた。「時間帯」とは、5時～10時を通勤・通学ラッシュ、11時～16時をオフピーク、17時から25時を帰宅ラッシュに3つの時間帯に分類したものである。さらに、通勤・通学ラッシュの時間帯では、8時を中心に8時を「0」、7時を「-1」、9時を「+1」のようにし、帰宅ラッシュの時間帯では、同様に18時を中心に18時を「0」、17時を「-1」という形で表記した（図1）。

### 2.4 データの統合

全てのデータを測定されたタイムスタンプによって統合した。Fitbit Charge2はバイタルデータを1分間隔で記録しており、これをデータ統合する際の基準とした。一方で、Google Mapのロケーション履歴による記録は、計測間隔は不定期であることから、前後の位置情報から線形補間を用いて毎分ごとの位置情報を推定した。その上で全ての移動データを含んだ上記データから、図2のプロセスによって、鉄道利用区間のみを抽出した。

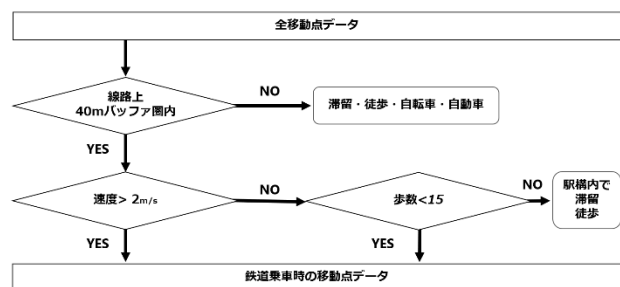


図2 鉄道移動データ抽出フロー

### 2.5 予測モデル

上述の蓄積された統合データを用いて、多変量解析の一つである重回帰分析から、任意の鉄道乗車時の心拍数を数学モデルで予測した。重回帰分析は、被説明変数（心拍数）を複数の説明変数（心拍数、快適度以外のパーソナルデータおよびその他データ）から予測する数学モデルで、以下の式によって求められる。

$$\text{被説明変数} = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i + \alpha$$

（ただし、 $X_1, X_2 \dots X_i$ は説明変数、 $\beta_1, \beta_2 \dots \beta_i$ は各説明変数の係数、 $\alpha$ は定数）

## 3. 結果・考察

### 3.1 鉄道利用区間の抽出

本論文では健常男性成人1名（学生）の2018年8月から2019年6月までの間のデータを用いて

分析を行った。抽出された鉄道移動データは期間中の総鉄道利用時間 12, 598 分のうち、約 65.3% にあたる 7, 621 分の点データであった。抽出精度が低下した要因として、「GPS 精度」と「位置情報データの不規則な計測間隔」によるものだと考えられる。また、Yan et al. (2013)による方法では位置と移動速度によって鉄道による移動モードが推定されていたが、この方法では乗車しているにも関わらず駅に近づく際の減速は歩行として捉えられてしまう可能性があった。しかし本フローでは歩数を活用し、鉄道移動データの抽出と同時に、駅構内の移動（乗降車、乗換等）も抽出することができた（図 3）。

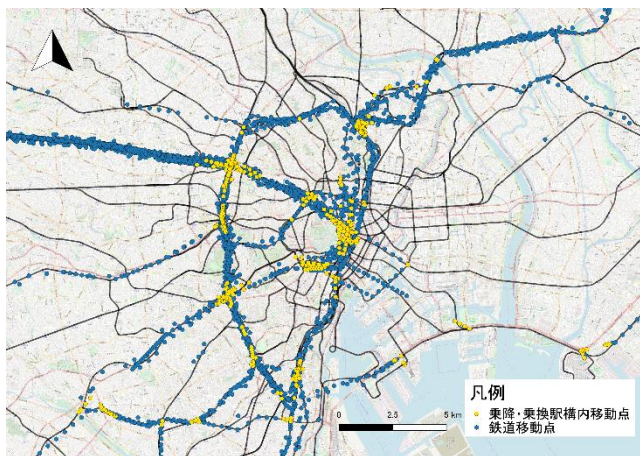


図 3 抽出された鉄道および駅構内移動点

### 3. 2 快適度と心拍数の関係

車内での快適度と心拍数には相関係数 0.653 のやや強い相関が見られた。また右の図 4a からは、快適性が向上するにつれて心拍数が低くなる傾向であることが言える。図 4b の日常生活での心拍数の変化と比較してみると、快適度 5（非常に快適）での心拍数の範囲は安静時と同じである。一方、快適度 2, 3 は身体的負荷のかかる歩行時と同等の心拍数の範囲であることから、快適度 5（非常に不快）における心拍数は非常に高いことが言える。

図 4a から、本研究では快適度 1~5 におおよそ対応する心拍数の範囲として[快適度 5:~75], [快

適度 4: 75~85], [快適度 3: 85~95], [快適度 2: 95~105], [快適度 1: 105~]として議論する。

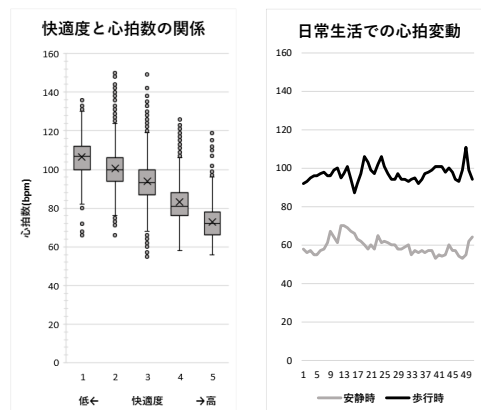


図 4a

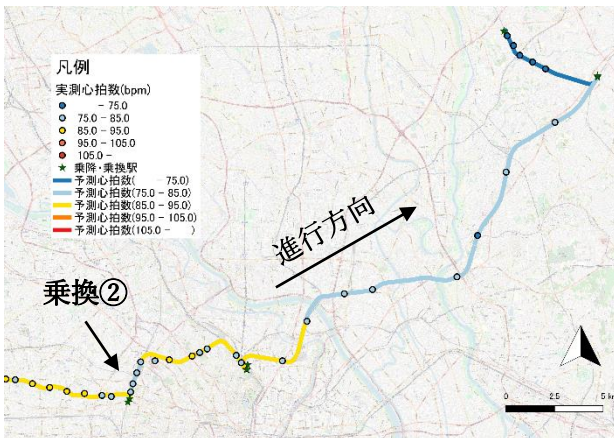
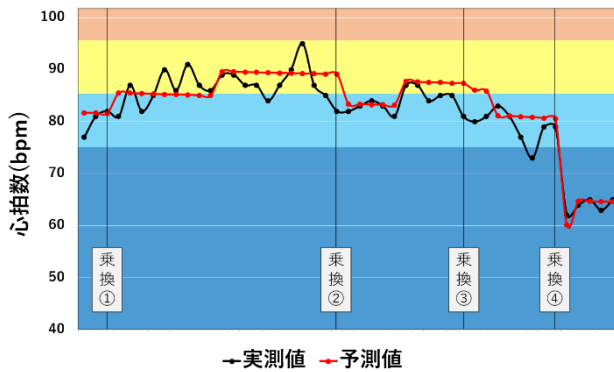
図 4b

### 3. 3 予測モデルの検証

心拍数を被説明変数とした重回帰分析を行うにあたり、上述した「通勤・通学ラッシュ」、「オフピーク」、「帰宅ラッシュ」の 3 つの時間帯ごとにそれぞれ分析した。各時間帯の重回帰モデルの予測能は、調整済み決定係数 ( $R^2$ ) によって表すとそれぞれ 0.63, 0.49, 0.72 となり、帰宅ラッシュにおける重回帰モデルが最も回帰式の当てはまりの度合いが高い結果となった。

これらの重回帰モデルを検証すべく、同一人物の 2019 年 7 月から 2019 年 8 月の新規データをモデルにあてはめ、心拍数を予測した。その結果、心拍数の実測値と予測値の平均絶対誤差 (MAE) は 5.89 (bpm)、平均絶対誤差率 (MAPE) は 7.58 (%) となった。図 5a, 5b は 7 月下旬（時間帯：オフピーク、予定：通学）のある一日の鉄道利用時の実測・予測心拍数を表したものである。一日を通して MAE は 2.51 (bpm)、MAPE は 3.06 (%) と精度は非常に高い。一方で実測値は突発的に上下しており、予測値から乖離している点もある。この日の最大絶対誤差は 7.85 (bpm) であり、これは本研究で定義した快適度の段階を 1 段階上下しかねない誤差である。その要因として車内での外的要因だと考えられ、今後それらを特定し、より精度の高いモ

デルに改良する必要があるだろう。



(上)図 5a 心拍数の実測・予測値  
(下)図 5b 心拍数変化の地理的特徴

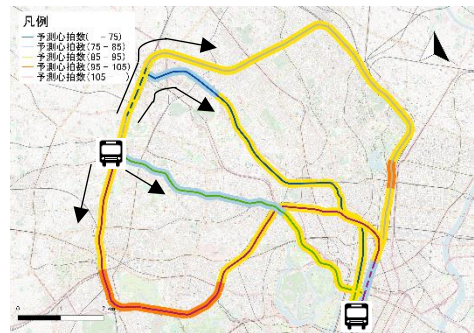
### 3. 4 予測モデルの活用

この予測モデルを用いることで、働き方改革において時差ビズが推奨されている昨今、複数の路線と複数の時間帯から快適な経路案内に有用だろう。図 6 は高田馬場駅から東京駅までの 4 経路 (全て起立していると仮定) における 3 つの乗車時間帯の予測心拍数を可視化したものである。図からは、時間帯や利用路線によって心拍数の取りうる範囲が異なることが確認された。

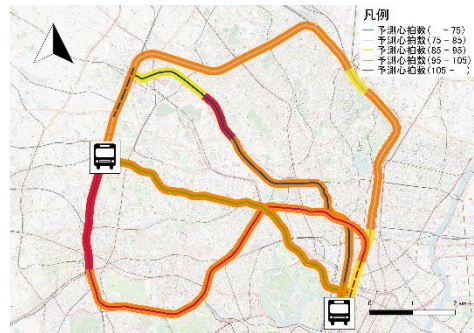
### 4. おわりに

本研究では、GPS、歩数、予定などのパーソナルデータと混雑率や温度といった定量データを組み合わせて、鉄道移動における快適度と比較的強い相関のある心拍数を予測する数学モデルの検討を行った。その結果、ある程度の精度を以て個人の心拍数の変動は予測可能となった。このこと

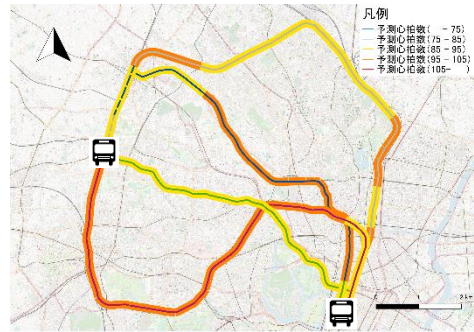
から、パーソナルデータを活用することで鉄道利用時の快適度を把握することができ、個人レベルでより快適な鉄道移動を提案することに繋がると考える。



[7 時台]



[8 時台]



[9 時台]

図 6 高田馬場・東京駅間の各経路の予測心拍数

### 参考文献

IT 総合戦略本部 (2014) 「世界最先端 IT 国家創造宣言」 報告書, <<https://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20160520/siryoul.pdf>>  
Yan, Z., Chakraborty, D., Parent, C., Spaccapietra, S., Aberer, K. (2013). Semantic trajectories: Mobility data computation and annotation. ACM Transactions on Intelligent Systems and Technology (TIST), 4(3), 49.