

平滑化によるPTデータの時空間内挿の高精度化
渡邊淳人・中村敏和・関本義秀・薄井智貴・柴崎亮介

High-accuracy spatiotemporal interpolation of PT data by smoothing
Atsuto WATANABE, Toshikazu NAKAMURA, Yoshihide
SEKIMOTO, Tomotaka USUI and Ryosuke SHIBASAKI

Abstract: Reconstruction of people flow using person trip surveys needs spatio-temporal interpolation between origin and destination because surveys such as person trip surveys only have information on origin and destination. The time notation is often biased in surveys such as person trip surveys into round number because these surveys are usually conducted by questionnaire. Due to this tendency reconstruction of people flow using person trip surveys is different from actual situation of transportation. For these reasons, smoothing biased departure and arrival time acquired from person trip survey need to be done in order to reconstruct people flow that is not biased as less as possible. This study proposes method of smoothing departure and arrival time by using kernel density estimation.

Keywords: カーネル密度推定(kernel density estimation)、平滑化(smoothing)、PTデータ(PT data)

1. はじめに

近年、防災、環境、マーケティング、交通計画等の観点から、都市空間における日々の人々の流れを把握することが重要になっている。そこで、東京大学空間情報科学研究センターは、研究サービス「人の流れプロジェクト¹⁾」を立ち上げ、パーソントリップ調査（以下、PT調査）等の時空間データから人の流れの再現に取り組んでいる。PT調査は、都市内の交通状態を把握するための調査であり、通常アンケート

との個人情報と各トリップの出発時刻、出発ゾーンID、到着時刻、到着ゾーンID、利用した交通モードなどであり、出発・到着の間にに関する情報は含んでいないため、このPT調査を利用して、人の流れを再現するためには、出発と到着の間の時空間内挿が必要である。また、PT調査はアンケートにより実施されるため集計される時間表記は0分や30分に偏りがちである。そのため、再現される人の流れは現実の人の流れとは離れたものになってしまう。

本研究は、現実に近い高精度な人の流れを再現することを目的に、出発・到着時刻の平滑化、データの加工をし、時空間内挿を行うものである。

連絡先：

千葉県柏市柏の葉 5-1-5 総合研究棟 4 階 435 号室
東京大学大学院新領域創成科学研究科

社会文化環境学専攻柴崎研究室

Tel : 04-7136-4290 (内線 64307)

Email:atsuto@csis.u-tokyo.ac.jp

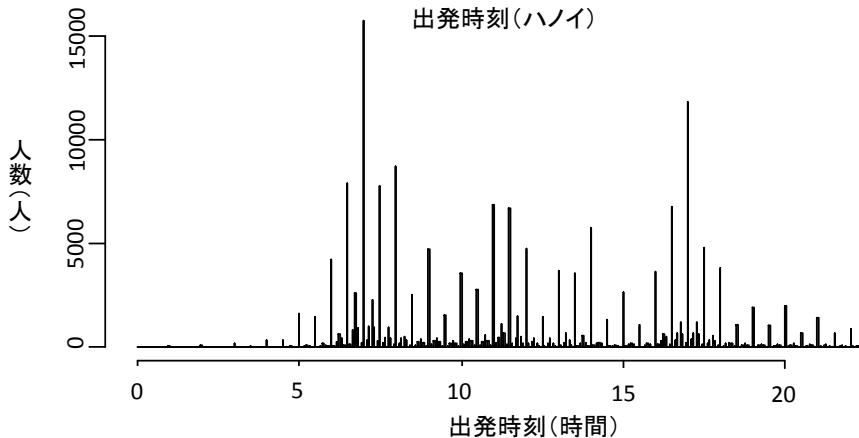


図1 出発時刻のヒストグラム(ハノイ)

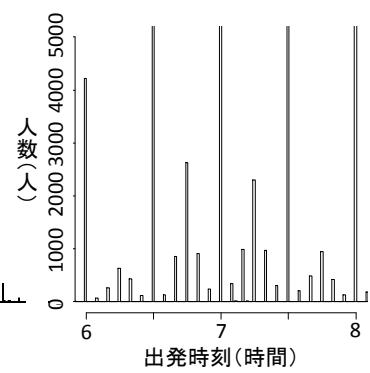


図2 6~8時のヒストグラム

2 本研究のアプローチ

2.1 既存の手法

PTの時空間内挿については、関本ら²⁾が行ってきたものとほぼ同様に OpenStreetMap³⁾、 pgRoutingProject⁴⁾などを用いて行った。しかし、先述したように時間表記の偏りがあり、人の流れの再現の精度が低いことが問題点としてあげられる。図1はJICAがハノイで行ったPT調査の結果の出発時刻ごとの人数を示したものであり、図2はそのうちの6~8時を示している。図1、図2からPTデータの出発時刻が●●時00分や30分などのきりの良い時刻に出発時刻が集中していることがわかる。これは、調査方法がアンケートであることに起因していると考えられる。高精度な時空間内挿には、PTデータがより現実の状況に近い必要があり、00分や30分に集中している出発時刻を補正する必要がある。

そこで、本研究では一部の時刻に集中している出発・到着時刻を平滑化し、高精度な時空間内挿を実現する手法を提案する。

2.2 本研究の提案手法

次に、本研究で提案する平滑化の手法を説明する。具体的にはカーネル密度推定^{5),6)}を用いる。カーネル密度推定法は、中島ら⁷⁾や中谷ら⁸⁾が行っているように、時空間データの平滑化

に有効で、以下の(1)式によって求められる。

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (1)$$

x_i ：同じ分布からの n 個の観測値

$(i = 1, 2, \dots, n)$

n ：観測値数 h ：バンド幅

$K(\cdot)$ ：カーネル関数

カーネル関数とバンド幅の算出方法は

Epanechnikov、Gaussian、Rectangular（カーネル関数）、Plug-in法、Silverman（バンド幅の算出方法）など複数あることが知られている。また、カーネル関数、バンド幅の選定が密度推定の結果を大きく左右することが知られている⁹⁾。本研究では、平滑化の結果がより滑らかになるカーネル関数は Gaussian、バンド幅は、Silverman の経験則として紹介されている方法を採用した。

3 本研究で用いた手法

3.1 出発時刻の平滑化

次に、PTデータの出発時刻を平滑化する具体的な手法について説明する。

$$K(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad (2)$$

$$h = 0.9 \min\left[\sigma, \frac{\text{四分位範囲}}{1.34}\right] n^{-0.2} \quad (3)$$

本研究で採用する Gaussian 関数(2)と、バンド幅を算出する(3)を Java アプリケーションで

実装し、出発時刻の平滑化をゾーンごとに行つた。また、カーネル密度推定では、合計が 1 となる 1 分おきの出発時刻のトリップ発生頻度が算出されるので、これを出発時刻のトリップ発生人数に補正した。

3.2 トリップデータの加工

前節では、各個人のトリップデータの出発時刻を平滑化した。本節では、PT データの出発時刻を平滑化された出発時刻に書き換える手法を説明する。

まず、平滑化前・後での 1 分おきの出発時刻 (1440 分) ごとのトリップ発生人数を算出する。この 1 分おきのトリップ発生人数をもとに、出発時刻を書き換える人数を算出する。ただし、3.1 で行った平滑化によって算出される 1 分おきのトリップ発生人数は少数なので、これを整数に変換する必要がある。この際、具体的な変換方法は、小数点以下の数の大きさを基準に行った。まず、少数の人数の整数部分を残し、小数点以下を取り除いた。そして、取り除いた小数点以下の合計を、小数点以下の数の大きさ順に振り分けた。

そして、上で求めた人数だけ出発時刻を書き換える。書き換える個人の選定は個人 ID の小

さい順に行った。

4. 実験

本研究で提案する手法を用い、JICA の PT データを平滑化する実験を行った。実験を行った各都市の概要を表 1 に記す。

まず、3.1 で提案した手法でハノイ、ホーチミン、ジャカルタ、トリポリの PT データの平滑化を行った。そして、3.2 で提案した手法で、平滑化された出発時刻を反映した PT データを作成し、時空間内挿を行った。時空間内挿の具体的な手法は関本ら²⁾を参照されたい。得られた結果を都市ごとに図 3 から図 12 に示す。図 3 から図 6 は、平滑化前後の 1 分おきの出発時刻のトリップ発生人数、平滑化パラメータ h を示す。図 7 から図 12 は平滑化前後の時空間内挿の結果を示している。ただし、平滑化の効果が顕著に表れている時間における時間を任意に選定した。図 3 から図 6 から出発時刻は有効に平滑化されていることがわかる。また、図 7 から図 12 のように、平滑化前には、見られなかったトリップ発生が平滑後に見られることから、平滑化が有効であることがわかる。

表 1 各都市の PT データの概要

| | ハノイ | ホーチミン | ジャカルタ | トリポリ |
|--------------|--|-------------------------------------|--|---|
| 居住人口 | 約 318 万人 | 約 716 万人 | 約 2000 万人 | 約 33 万人 |
| 調査人口 | 約 7.9 万人(約 2.5%) | 約 17 万人(約 2.3%) | 約 42 万人(約 2.0%) | 約 3000 人(約 1.1%) |
| ゾーン数 | 約 220 | 約 260 | 約 350 | 約 50 |
| トリップ/人 | 2.3 | 1.7 | 2.6 | 2.1 |
| 主要交通機関分担率(%) | 自転車(25.3%), オートバイ(63.2%), 車・タクシー(3.6%), バス(6.7%) | オートバイ(58.8%), 徒歩(20.9%), 自転車(15.4%) | バス(49.1%), バイクタクシー(31.2%), 三輪自転車(4.5%) | 自転車・オートバイ(33.4%), 車(27.9%), タクシー(17.8%), バス・バス(14.2%) |

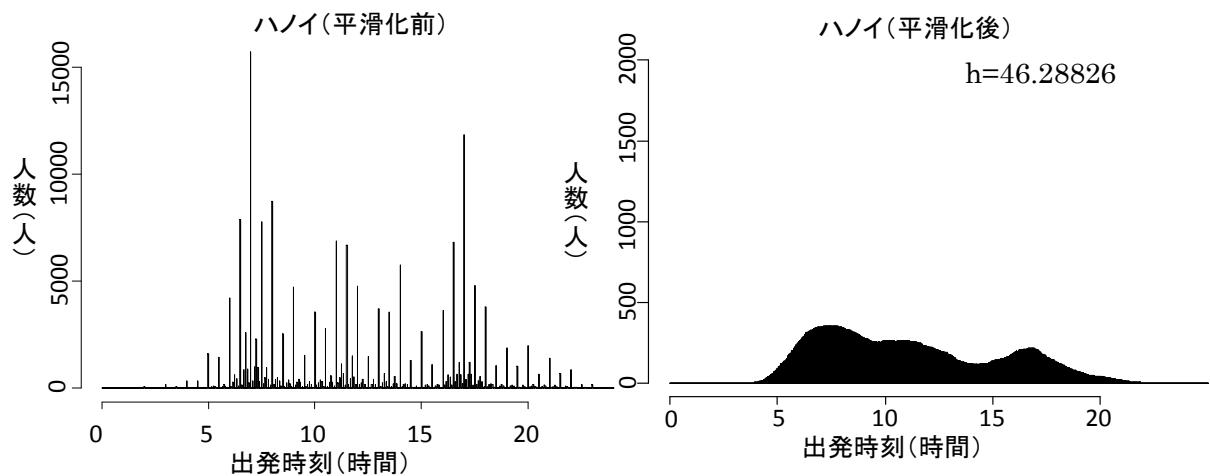


図3 平滑化前後のトリップ発生人数の比較（ハノイ）

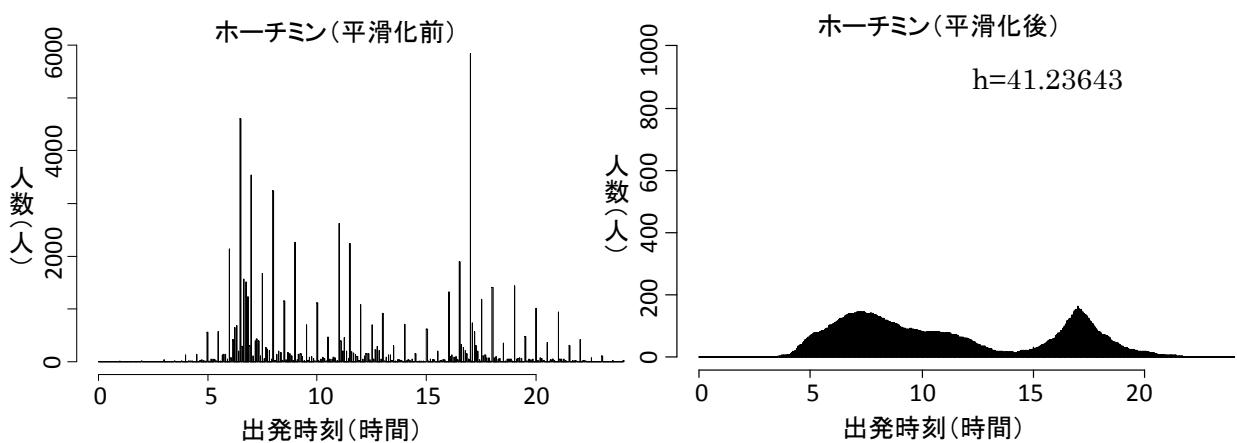


図4 平滑化前後のトリップ発生人数の比較（ホーチミン）

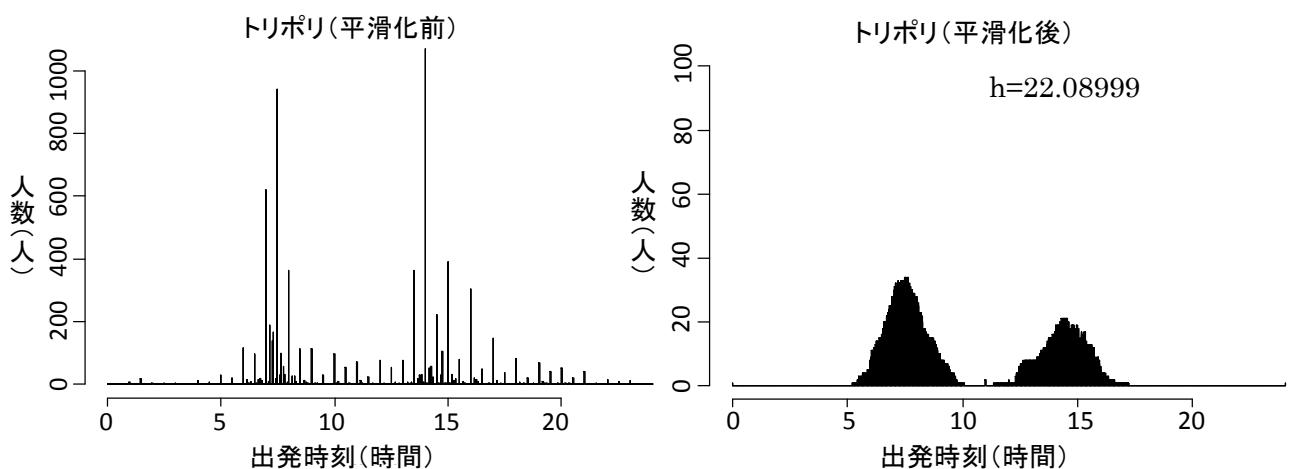


図5 平滑化前後のトリップ発生人数の比較（トリポリ）

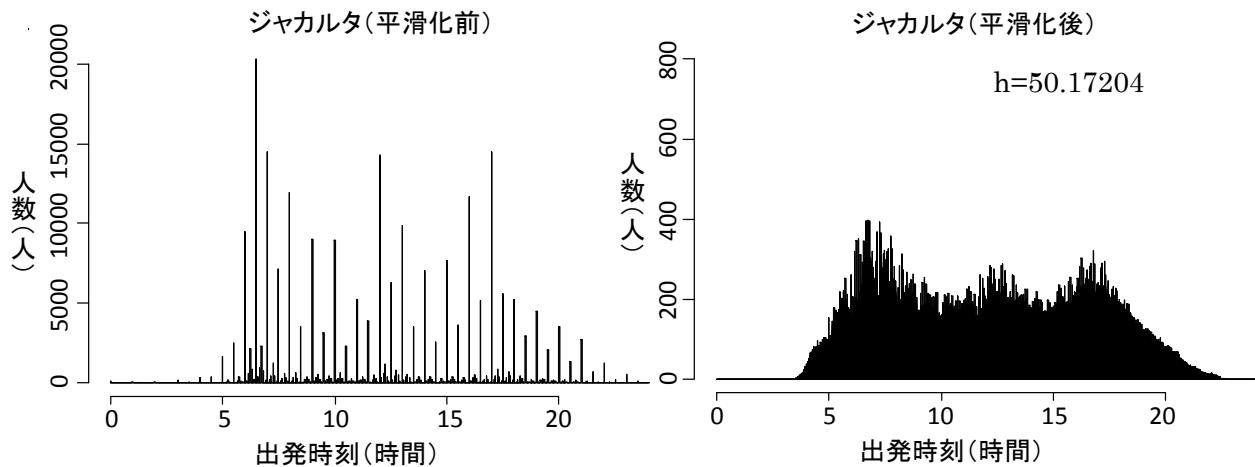


図 6 平滑化前後のトリップ発生人数の比較（ジャカルタ）



図 7 平滑化前後の時空間内挿の結果の比較（ハノイ）

（赤：歩行、黄：自転車、紫：商用車、水色：自動二輪、黄緑：自家用車）



図 8 平滑化前後の時空間内挿の結果の比較（ホーチミン）

（赤：歩行、黄：自転車、紫：商用車、水色：自動二輪、黄緑：自家用車、青：鉄道）



図9 平滑化前後での時空間内挿の結果の比較（トリポリ）

（赤：歩行、黄：自転車・自動二輪、水色：自家用車、黄緑：商用車、紫：トラック、青：バス）

5. おわりに

本研究では、より現実の状況に近い人の流れの再現を目指し、既存の時空間内装の問題点を解消するために、PTデータの時間情報の平滑化による高精度な時空間内挿を行い、カーネル密度推定を用いた平滑化の有効性と高精度な時空間内挿の実現を確認した。今後は、PTデータの内、本研究で使用していない情報を利用した、より高精度な人の流れの再現の研究を進めていきたい。なお、データの不足により時空間内挿を行うことができなかったジャカルタのPTデータの時空間内挿も完了させたい。また、今回対象にした4都市以外の他都市での適用実験も進めていきたい。

謝辞

本研究では、朝日航洋株式会社の大伴様、杉森様に大変お世話になりました。この場を借りて、お礼をさせて頂きたいと思います。

参考文献

- 1) 東京大学空間情報科学研究センター：人の流れプロジェクト、
<http://pflow.csis.u-tokyo.ac.jp/>（アクセス：2010年8月27日）
- 2) 関本義秀・中村敏和・薄井智貴・金杉洋：海外における人々の時空間位置の詳細化－ハノイのPT調査を事例に、交通工学研究発表会論文集、2010（掲載決定）
- 3) OpenStreetMap:<http://www.openstreetmap.org/>（アクセス：2010年8月27日）
- 4) pgRoutingProject：
<http://pgrouting.postgis.org/>（アクセス：2010年8月27日）
- 5) Bowman, A. and Azzelini, A. (1997) Applied Smoothing Techniques for Data Analysis. Oxford: Oxford University Press.
- 6) Simonoff, J.S. (1996) Smoothing Methods in Statistics. New York: Springer-Verlag
- 7) 中島亮一・貞弘幸雄（2007）点分布の空間的平滑化が地図利用者の点の位置推定に与える影響の評価「GIS－理論と応用」 Vol. 15, No. 2, pp. 23-32
- 8) 中谷友樹・矢野桂司（2008）犯罪発生の時空間3次元地図－ひったくり犯罪の時空間集積の可視化－「地学雑誌」117(2) 506-521
- 9) 谷崎久志(2005)密度関数のカーネル推定量におけるバンド幅の選択について：モンテカルロ実験による小標本特性『国民経済雑誌』第191巻、第1号、pp. 59-70