

移動体通信機器によって得られる歩行者移動軌跡の誤差モデル

Accuracy of Pedestrian Trajectory Observed by Mobile Communication Instruments

長尾 大^{*}, 井料 隆雅^{**}, 椿 涼太^{***}, 朝倉 康夫^{****}

By Dai NAGAO^{*}, Takamasa IRYO^{**}, Ryota TSUBAKI^{***} and Yasuo ASAKURA^{****}

1. はじめに

近年、ヒトや車の移動を観測する道具としてGPSなどに代表される移動体通信機器が利用されるようになった。従来のアンケート調査などでは、被験者の記憶に頼っていたため、位置や出発時刻などの情報を正確に得ることが難しかった。移動体通信機器を用いることによって被験者の記憶に頼らない、客観的で正確な移動データを入力することが期待されている。

移動体通信機器によって得られるデータは主に移動体の位置情報(座標、緯度・経度 etc)と時間情報である。位置測定はある時間間隔をもって行われるため、時間が離散的なドットデータとなる。一般に、機器の性能、測定する場所の周辺環境によっては位置測定に誤差が発生する。そのため、得られたドットデータの取り扱いに工夫が必要である。

2. 移動体通信機器を用いた交通行動分析

移動体通信機器を用いた研究には都市圏・都市内ネットワークを対象としたものと空間スケールのより細かい広場などを対象としたものがある。また、対象とするエリアのスケールによってデータの取り扱い方法が異なってくる。

(1) 都市圏・都市内ネットワークを対象とした研究

都市圏・都市内ネットワークを対象とする研究には、車や電車を対象とした研究と、歩行者を対象とする研究がある。どちらを対象とした研究でも、移動・滞在判別によるトリップの抽出や、観測データをネットワークに吸着させての移動経路推定などがなされている。

こういった都市圏・都市内ネットワークを対象とした研究では、観測データから、移動体の移動経路、速度、

キーワード: 歩行者, 移動軌跡, 誤差モデル

^{*}学生員, 神戸大学大学院自然科学研究科(神戸市灘区

六甲台町1-1, Tel/Fax 078-803-6208)

^{**}正員, 工博, 神戸大学工学部建設学科

^{***}学生員, 工修, 神戸大学大学院自然科学研究科

^{****}正員, 工博, 神戸大学大学院自然科学研究科

滞在または移動, 旅行時間などの情報を取り出し(あるいは推定し), 交通行動の分析やネットワークのパフォーマンスの評価が行われている。

(2) 広場スケールを対象とした研究

広場などの比較的狭い空間領域を対象とする研究では、移動体は主に歩行者や自転車である。広場スケールを対象とした場合、観測データのもつ情報が失われてしまうため、ネットワークに測定点を吸着させることが適切でない。例えば、移動体がどのリンクにいるかという情報に加え、その移動体が道路の左側、あるいは右側にいるなどの情報も必要になる。つまり、移動軌跡の形状も考慮して分析を行う必要があるといえる。

このようなケースでは、測定時の誤差がより分析に影響すると考えられる。Dieter F. et al¹⁾では、Bayesian filter Techniqueを用いてGPSの観測データから移動体の真の座標を推定する手法を提案している。D. A. Grejner-Brzezinska et al²⁾では、電離層の状態をモデル化することによりGPSの誤差を表現し、観測データの誤差を低減している。しかし、これらの研究では、その情報がどれだけ確からしいものであるかについては、あまり議論されていない。移動体通信機器によって得られる情報はある確率分布に従っていると考えられるため、確率論的にその精度を評価することが可能であると考えられる。

(3) 本研究の位置づけ

本研究では、広場スケールを対象とした移動体の観測データを用いて、推定された行動軌跡がどれだけ真値を表現しているのかの検証を行うことを目的とした。前述のように、移動体通信機器によって得られる情報はある確率分布に従っていると考えられる。そこで、確率統計の考え方を導入し、軌跡総延長を指標として、その信頼区間を設定しその推定の精度を評価する手法を提案する。

3. 精度評価手法の提案

(1) 誤差モデルの作成と適用

本研究では代表的な移動体通信機器として、GPSを利

用した。約 20 分間の定点観測（観測時間間隔は 1 秒）によって得られた時刻、経度・緯度のデータより、機器の特性として以下の 2 つを仮定した。

仮定 1：時刻 $t+1$ における誤差は、時刻 t における誤差の影響を受ける。

仮定 2：連続した時刻における誤差の差は、すべての時刻について独立である。

この仮定より、以下の式で表される GPS の誤差モデルを作成した。

$$\vec{Z}_t = \vec{z}_t + \vec{\varepsilon}_t \quad (3.1)$$

ただし、

$$\vec{Z}_t = (Z_1(t), Z_2(t)) \quad (3.2)$$

$$\vec{z}_t = (z_1(t), z_2(t)) \quad (3.3)$$

$$\vec{\varepsilon}_t = \left(\sum_{s=1}^t \varepsilon_{Z_1,s}, \sum_{s=1}^t \varepsilon_{Z_2,s} \right) \quad (3.4)$$

式 3.2, 式 3.3, 式 3.4 の添え字 1,2 は、平面を 2 次元座標で表現した時の軸の方向を示している。ここで \vec{Z}_t , \vec{z}_t はそれぞれ時刻 t における移動体の測定座標と真の座標を、 $\vec{\varepsilon}_t$ は時刻 t における誤差を表わしている。また、 $\hat{\varepsilon}_{Z_1,t}, \hat{\varepsilon}_{Z_2,t}$ は時刻 t と時刻 $t-1$ の誤差の差（誤差の差分項）を表わしている。本研究では上記の仮定より、時刻 t における誤差は、時刻 t までの誤差の差分項の和で表わすことができると考えた。この誤差の差分項は、測定値より、ほぼ正規分布に従うと仮定することができる。

ここで作成された誤差モデルは、測定データの補正を行うために利用されるのではなく、得られる情報について、その確率分布を再現するために利用される。

(2) 精度評価手法

前述のように、移動体機器によって得られる情報はある確率分布に従うと考えられる。例として、ヒトが移動した軌跡の総延長について考える。軌跡総延長は移動体通信機器によって得られた測定座標より求めることができる。しかし、測定座標には誤差が含まれており、この誤差がある確率分布に従って発生するランダム変数であると考え、得られた軌跡総延長はある出現確率のもとで出現していると考えられる。本研究では以下の手順を用いて測定座標より得られる軌跡総延長の精度を評価する。

<Step.1> 誤差の差分項の再現

誤差の差分項は正規分布に従うと仮定できる。乱数を用いてこの誤差の差分項を発生させる。

<Step.2> 軌跡総延長の計算

全ての軌跡の出現確率が一定であると仮定を置き、Step.1 で発生させた誤差の差分項を式 3.1 の誤差モデルに

適用し、軌跡の総延長を計算する。

<Step.3> 分布の再現

Step.1, Step.2 を繰り返して行うことで軌跡総延長の分布を再現する。

<Step.4> 精度評価

Step.3 で得られる軌跡総延長の確率分布より、信頼区間を設定し、その精度を評価する。

4. 推定された行動軌跡の精度評価

(1) 実証実験

本研究では移動体通信機器によって得られる行動軌跡の精度評価を行うために、広場を移動する歩行者を対象に、実験を行った。その概要は表 4.1 に示すとおりである。

表 4.1 実験概要

項目	内容
実験日時	2005年1月26日 11時から15時まで
実験場所	神戸大学大学院自然科学研究棟3号館前広場
使用器具	GPS2台
被験者	女子大学生2名
取得データ	GPSによる被験者の移動データ ビデオカメラによる広場の歩行者軌跡座標

被験者 2 名それぞれに GPS を携帯してもらい、被験者の移動を対象に位置測定を行った。得られる情報は時刻（1 秒間隔）、経度・緯度である。実験では、それぞれの被験者にあらかじめ設定したいくつかの OD ペアを教え、O から D に自由に移動するように指示を与えた。

同時に、GPS による被験者 2 名の移動データだけでなく、ビデオによる広場の歩行者移動を観測し、画像解析により移動軌跡のデータを作成した。本研究ではビデオによる観測データを真値として取り扱うこととする。

(2) 行動軌跡の精度評価とその結果

測定されたデータのうち、あきらかな異常点を含む 2 ケースを除外した 12 ケースの OD のそれぞれについて、軌跡総延長 D を精度評価の対象とした。異常点の判別は、時速 25km/h を速度の閾値として、この閾値を超える移動を異常点と分類した。時速 25km/h は、50m を 7.2 秒で移動する速度で、ヒトの移動として不可能な速度ではない。しかし、そのような速度で移動するという事は広場の回遊では考えにくいためこの速度を閾値とした。

OD 間の軌跡総延長 D を、以下の式で定義する。

$$d_t = \sqrt{(z_1(t) - z_1(t-1))^2 + (z_2(t) - z_2(t-1))^2} \quad (4.1)$$

$$D = \sum_{t=1}^T d_t \quad (4.2)$$

真の軌跡総延長として、ビデオデータより式4.1, 式4.2を用いて計算した D^* を用いる。図4.1に軌跡総延長 D の定義を図で示し、あわせて真の軌跡総延長 D^* の平均値と最大値・最小値を示す。

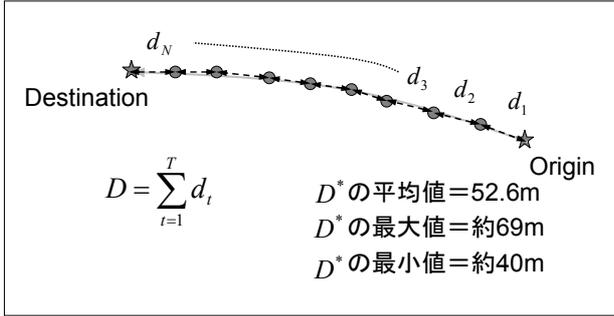


図4.1 OD間の軌跡総延長の定義

提案した精度評価手法より、その繰り返し回数を10000回として D の累積分布を再現した。再現された累積分布の10パーセンタイル値から90パーセンタイル値までを信頼率0.8の信頼区間として設定した。結果を表4.2に示す。

表4.2 信頼区間の設定(その1)

case	真値(m)	測定値(m)	10パーセンタイル値(m)	90パーセンタイル値(m)
1	69.07	72.64	71.91	77.42
2	44.96	46.24	45.17	49.70
3	40.27	42.00	42.79	47.28
4	51.28	54.59	53.76	58.59
5	51.82	53.89	53.06	57.78
6	43.13	41.57	41.00	45.29
7	44.86	50.48	58.11	66.52
8	63.93	68.53	77.54	87.08
9	57.60	59.12	63.58	71.50
10	52.98	62.75	66.26	75.60
11	56.80	48.16	56.59	65.81
12	54.67	53.33	53.81	58.78

真値が信頼区間に入ったケースは灰色で示す3ケースのみであった。信頼区間を外れたケースに着目すると、真値は10パーセンタイル値よりも小さいことがわかる。つまり、推定された分布は真値よりも大きく推定されているといえる。

この理由の1つとして測定された軌跡が直線的ではなく、蛇行しているということが考えられる。測定値と真値について、そのOD間最短距離と軌跡総延長を比較すると、測定値では、その差が平均で約4.3mであった。それに対して真値では、その差は平均で約1.4mであった。これは、実際にヒトが歩くときに、ほぼ直線的な軌跡を描くのに対し、測定された軌跡は真の軌跡のように直線的なものでなく、多少蛇行していることを示している。このことが結果に影響を与えていると考えられる。提案した方法は、分布を再現する際、全ての軌跡形状の出現確率が一定であると仮定している。これは、起点から終点まで、蛇行しながら向かっている軌跡とほぼ直線に近い軌跡が同じ確率で出現するということである。これではヒトの行動を十分に考慮しているとはいえない。そのため、軌跡形状の出現確率にヒトの行動規範を考慮する

ことにより、より良い結果が得られる可能性があるといえる。そのことを考慮して以下のような改良を試みた。

5. ヒトの行動規範を取り入れた誤差モデル

ヒトの行動規範を考慮する手法として、軌跡形状の出現確率にヒトの行動規範を取り入れ、制約条件として誤差モデルに適用する手法を試みる。

(1) ヒトの行動規範の表現

ビデオによって得られた歩行者移動軌跡より、ヒトの行動について分析を行う。分析を行うにあたり、連続する3点の關係に着目し、ヒトの行動を表わす指標として、以下の指標を考えた。

$$(d^* - d_s) / d_s \quad (5.1)$$

ここで d_s, d^* はそれぞれ時刻t-1の点から時刻t+1の点までの最短距離と移動距離を表わしている。(図5.1)

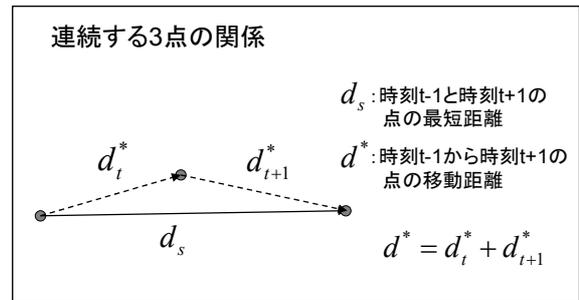


図5.1 連続する3点の關係

この指標は最短経路と移動経路の差と最短経路の比であり、最短距離に対してどれだけ遠回りをしているかを示している。ビデオによって得られた被験者の移動軌跡を対象に全ての連続する1秒間隔の3点を抽出した。抽出した3点より、この指標を計算し、その出現頻度を求めた。(図5.2)

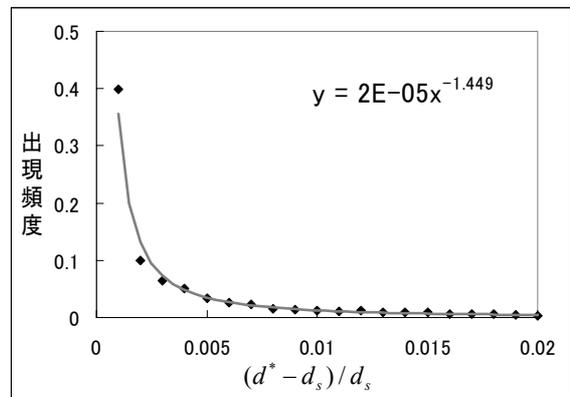


図5.2 指標 $(d^* - d_s) / d_s$ の出現頻度

ただし、図5.2で使われている出現頻度は指標を0.001

ピッチで区切ったときのデータ区間の出現頻度を示している。

図 5.2 より、連続する 1 秒間隔の 3 点のみを考えれば、ヒトはほぼ最短経路を移動しているといえる。この分布図に近似曲線を当てはめると、以下の式が得られた。

$$P(\gamma_{i-1} < (d^* - d_s) / d_s \leq \gamma_i) = 2 \times 10^{-5} \cdot \gamma_i^{-1.449} \quad (5.2)$$

$$\gamma_i = i \times 10^{-3} \quad (i = 1, 2, 3, \dots, I) \quad (5.3)$$

式 5.2 は、i 番目の区間に式 5.1 の指標が入る確率を示している。本研究ではこの式を、時刻 t-1、時刻 t-2 における移動体座標が得られたときの、時刻 t における移動体座標条件付き確率 $P(\bar{z}_t | \bar{z}_{t-1}, \bar{z}_{t-2})$ の代わりとして利用し、ヒトの行動を表現することを試みた。

この確率を誤差モデルに制約条件として適用することにより、蛇行している軌跡よりも直線的な滑らかな軌跡の出現確率を高くすることができると考えられる。

(2) 精度評価とその結果

ヒトの行動規範を考慮していない誤差モデルを用いて情報の精度評価を行った場合と同様に、得られたデータより、異常点を含むケースを除いた 12 ケースの OD 間の軌跡総延長 D を精度評価の対象とした。

制約条件を適用したため、計算が非効率的になった。そのため、 D の確率累積分布を再現するための繰り返し計算の回数を 1000 回とした。信頼区間の設定は同様の方法を用いて行った。結果を表 5.1 に示す。

表 5.1 信頼区間の設定 (その 2)

case	真値(m)	測定値(m)	10パーセンタイル値(m)	90パーセンタイル値(m)
1	69.07	72.64	71.91 71.24	77.42 76.87
2	44.96	46.24	45.17 44.26	49.70 48.90
3	40.27	42.00	42.79 41.17	47.28 45.73
4	51.28	54.59	53.76 52.59	58.59 57.71
5	51.82	53.89	53.06 52.27	57.78 56.89
6	43.13	41.57	41.00 39.93	45.29 44.25
7	44.86	50.48	58.11 53.80	66.52 65.52
8	63.93	68.53	77.54 70.79	87.08 84.64
9	57.60	59.12	63.58 58.33	71.50 68.90
10	52.98	62.75	66.26 60.40	75.60 76.23
11	56.80	48.16	56.59 50.00	65.81 68.78
12	54.67	53.33	53.81 52.35	58.78 57.82

※10パーセンタイル値と90パーセンタイル値の上段は表4.2と同じ値を示している。下段はヒトの行動規範を考慮した誤差モデルを用いた結果を示している。

真値が信頼区間に入ったケースは 4 ケースであり、1 ケース増えたものの、全体の 3 分の 1 に留まっている。人の行動規範を取り入れたことによる変化を見るために、人の行動規範を取り入れる前と取り入れた後の設定された信頼区間を比べた。

12 ケース中 10 ケースに見られるように信頼区間が軌跡総延長の値が小さくなる方向にシフトしている。このことより、人の行動規範を取り入れることにより、取り入れる前の誤差モデルのように軌跡総延長が過大に推定されることがなくなったといえる。しかし、信頼区間に真値が入ったケースが 3 分の 1 ということからもわかるように、改善の余地があるといえる。

提案した手法は、2 秒という限られた時間内での蛇行を少なくするものである。しかし、よりスパンの長い蛇行については低減することができていない。また、本研究では、測定値と真値の起点が一致していると仮定しているが、終点の座標は測定値と真値では一致しない。スパンの長い蛇行と終点座標のズレが結果に影響を与えていると考えられる。

なお、ケース 10、ケース 11 は、人の行動規範を考慮することにより、90 パーセンタイル値が大きくなってしまったケースである。この理由については現時点ではまだわからない。今後研究を進めていく上でその理由を解明したいと考える。

6. まとめ

本研究で提案した手法は、移動体通信機器によって得られるデータより求めた情報が、どれだけ確からしいものかを検証するためのものである。

本研究では、定点観測により機器の誤差をモデル化し、ヒトの行動規範を制約条件として取り入れることで移動軌跡の総延長の分布を再現した。再現された分布より、信頼区間を設定したところ、全体の 3 分の 2 のケースで真値が信頼区間に入らなかった。この理由として、測定値と真値の起点が一致すると仮定した場合の終点座標のズレと 2 秒よりも長いスパンの蛇行が考えられる。これらを考慮した手法については今後検討していくこととする。

今回提案した手法では、乱数を用いた繰り返し計算によって分布を再現している。そのため、計算が非効率的になっている。計算を効率的にする手法として、再現する軌跡総延長の分布を関数で近似することが考えられるが、こちらも今後検討していくこととしたい。

【参考文献】

- 1) Dieter F., Jeffrey H., Henry K., Lin L. and Donald J. P.: Bayesian Techniques for Location Estimation, UbiComp Workshop on Location Estimation, 2003.
- 2) D. A. Grejner-Brzezinska, R. Da, C. Toth: GPS error modeling and OTF ambiguity resolution for high-accuracy GPS/INS integrated system, Journal of Geodesy, 72(11), pp. 628-638, 1998