

携帯電話からのGPSデータを用いた歩行者の経路選択肢集合の推定*

Pedestrians' Route Choice Set Generation using GPS Data from Mobile Phone*

田中伸治**・藤原直生***・桑原雅夫****

By Shinji TANAKA**・Naoki FUJIWARA***・Masao KUWAHARA****

1. はじめに

(1) 研究背景と目的

歩行は、日々の生活の一部として誰もが利用する交通手段であるため、より快適な歩行環境の整備や計画は多くの効用をもたらすといえる。その快適な歩行環境を実現するためには歩行者の行動、つまりどの経路を選択するかという経路選択行動と選択した経路の中でどのように歩行するかというミクロな歩行行動の両者を正確に把握することが求められる。

これら二つの行動分析は前者の行動をモデル化することで各経路あるいはリンクの歩行者交通量を推定し、その交通量下での歩行者流動を後者の分析に基づいたシミュレーションモデルによって再現するというように密接につながっている。ゆえに経路選択行動のより正確な把握と歩行者交通量の予測により、近年活発に開発されている歩行者流シミュレーションモデルの有効活用が促され、より適切な歩行環境整備が可能になると考えられる。このように経路選択行動は歩行環境整備のために重要であり、これまでにロジットモデルに代表される離散選択モデルの構築をはじめ、数多くの研究が行われてきた。しかし、それら選択モデルの推定結果に大きな影響を及ぼすと考えられる選択肢集合の生成に関しては、詳しい分析が行われていない。そこで本研究では、この歩行者の経路選択肢集合を推定することを目的とする。

また、歩行者の利用経路を得るために、後述する既往研究ではアンケートや歩行者の追跡が行われてきた。しかし、近年では携帯電話端末にGPSが内蔵され、周辺地図の確認やナビゲーションといったサービスが提供されているため、サービスの提供者にはユーザーの位置データが蓄積されており、これを用いることで大量のサンプルを簡便に得ることができると考えられる。そこで本

研究ではこのGPSからのデータを用いることで、多数の歩行者の実流動データに基づいて経路選択肢集合を推定する手法を提案する。

(2) 既往研究の整理

歩行者の経路選択モデルの構築にあたっては、ほとんどの研究でロジットモデルが採用され、モデルにおける各経路の確定効用項の定式化、つまり経路特性を表す説明変数や歩行者の属性変数の定式化と導入に主眼を置いている。

中村ら(1984)は歩行行動の活性化にかかわる要件を「安全性の確保」、「歩行者空間の快適性の確保」、「交通手段としての効率化」の3つに分類しており、岡田(2006)はそれに基づく既存研究の整理を行っている。まず、「安全性の確保」の観点からは吉永ら(2001)が歩道、電柱の有無や自動車交通量などの道路条件といった制御可能な要因を説明変数としてロジットモデルを構築し、その有意性を立証している。次に「歩行空間の快適性確保」とは、肉体的、心理的に負担を感じさせない歩行空間条件にかかわる項目である。外井(1981)は歩行者の快適性を表現する要因として歩道の有無、その幅員、道路の傾斜、歩行者専用の当否などを列挙し、それらが歩行距離に与える効果を算出している。最後に「交通手段としての効率化」とはより効率的に歩行を機能させることにかかわるものであり、経路長だけでなく目的地方向の選択や、直進性などが含まれる。前述した吉永ら(2001)のモデルにおいてももちろん経路長は説明変数として扱われている。一方で、毛利・塚口(1979)は同時に歩行者が目的地方向を好む性質、歩行方向をできるだけ維持しようとする性質をそれぞれ「目的地指向性」、「方向保持性」、それらをまとめて「空間的安定」と呼び、経路の距離差が小さい場合におけるその特性の存在を示している。また塚口ら(2002)によって、空間的安定を説明変数としたロジットモデルによるモデル化が様々な地域を対象として行われた。なお、このモデルは各交差点における選択行動に着目したものであり、そのノードに接続しているリンクが選択肢となり、それぞれのリンクを含む経路が経路として認識されているか、つまり選択肢集合に含まれるか否かはそれほど考慮されていない。また、前述の吉永

*キーワード: 歩行者, 経路選択, GPS

**正会員, 博士(工), 東京大学生産技術研究所,
(東京都目黒区駒場 4-6-1, TEL:03-5452-6419,
E-mail: stanaka@iis.u-tokyo.ac.jp)

***非会員, 修士(工), 大成建設(株)

****正会員, Ph.D, 東北大学情報科学研究科

らのモデルでは各経路を選択肢としているが、選択肢集合の生成過程に関する考察は行われていない。

このように、これまで歩行者の経路選択に影響を与える要因の特定や、それに基づく選択モデルの構築に関する研究は多数行われてきている。しかし前述したように、その経路選択肢集合の生成段階における研究は行われていない。

2. データ概要

(1) データの種類

本研究では携帯電話を通して現在地付近の地図の確認やナビゲーションシステムの利用に際して得られる位置情報、取得時刻などのデータを用いる。データの範囲は東京23区を含む矩形エリア、期間は2008年1月1日から2009年1月1日まで（ただし12月24日から12月31日の期間分はない）である。このデータは表1のように分類できるが、ナビゲーション起動時ではユーザーはおおむね示された経路通りに進むと考えられるため、本研究では非起動時データの活用する。

表1 データ概要

	ナビゲーション起動時	非起動時		
		足あと機能	自動現在位置更新	地図確認
ログ取得間隔	1秒	1秒	1秒	1分
データ内容	位置情報			
	データ取得時刻			
	個人識別番号			
	携帯電話機種情報			
	利用サービス情報			
	ナビゲーション識別番号			
ルートマッチング	あり	なし		

(2) 対象エリア

データが集中して取得されている渋谷駅周辺（東経139.694843度～139.705897度、北緯35.654971度～35.663983度）に対象を絞って分析を行った。この範囲内に含まれるデータ数は37258トリップである。ただしこれらは対象エリア内にトリップの一部が含まれるトリップ数を表しており、非起動時のデータにはトリップを識別する情報が含まれないためデータ取得間隔が3分以内のものを同一トリップとして集計している。自動現在位置更新、足あと機能利用時のデータではサンプル数が少なく、またログ取得時間データにエラーを含むものが多く見受けられたため、地図確認時のデータのみを分析対象とし、以下ではこれを非起動時データと呼ぶ。

図1に非起動時データの各トリップにおける取得ログ数とトリップ数の関係を示す。ナビゲーション非起動時データでは取得ログ数が少ないトリップ数の割合が非常に高くなっているが、それは取得間隔が長いことに加えて、周辺地図の確認という行為が一時的なものであり、確認後すぐにアプリケーションを終了してしまう行動をとるユーザーが多いことに起因しているのではないかと

推測される。

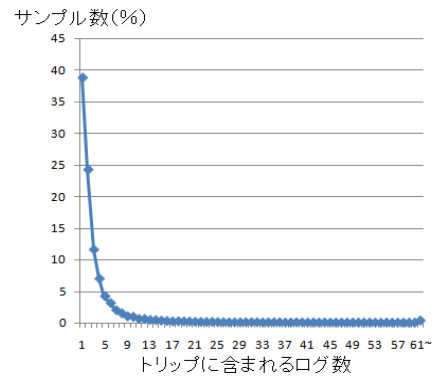


図1 含まれるログ数でみたトリップの分布

図2はログ間の直線距離をログ取得間隔で除したものの分布を示しており、0.1m/s未満のものが多いことが見て取れる。実際のログ間の歩行速度は直線で結ぶことができないためこの値よりも大きくなる可能性が高いと考えられるが、それを考慮しても移動速度が小さいほうに分布が偏っていると言える。これは信号待ちなどにより歩行中に立ち止まっている場合に加えて、地図の確認のために立ち止まっている状態や、アプリケーションを起動したまま、店舗に立ち寄っている状態などの存在が原因として考えられる。一方で移動速度が2.0m/sを超えるようなものも存在も見受けられる。このような移動は一般的な歩行速度から考えると現実的ではない。原因としてはGPSの測位誤差によりログ間距離が実際の値よりも大きくなっている点が存在していることが考えられる。この問題を解決するため、マップマッチングによる補正を行う必要がある。

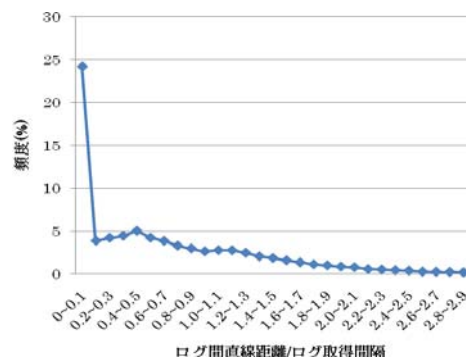


図2 ログ間移動速度の分布

3. マップマッチング

(1) マッチング方法

GPSの普及に伴い、近年、様々なマップマッチング方法が提案されてきた。しかし、これらは基本的に自動車のナビゲーション時を対象としており、そのログ取得間隔は1秒程度であるが、本研究で用いたGPSデータはログ取得間隔が約1分と比較的長いという特性がある。

また既存の研究ではリアルタイムのマップマッチングが求められているものが主であるのに対して、本研究ではトリップに含まれるすべてのログが得られた後にマッチングを行うという違いも存在する。

そこで本研究では比較的簡便で、上記の特性を考慮したマッチングを行うことが可能であるトータルウェイトスコア (TWS) を使ったアルゴリズムを採用する。これは Greenfeld(2002), Velaga, N.R. et al.(2009) などにおいて用いられている。

そのアルゴリズムは大きくログの近傍に存在するリンクを候補リンクとする段階と、それらのひとつひとつのリンクに関して TWS を計算し、それが最も大きく (あるいは小さく) なる候補リンクに垂直的にマッチングを行う段階の二つに分けられる。また、TWS は各候補リンクがマッチング先として正しいか否かを判断する指標の線形和の形であらわされ、その指標の最たる例はログとマッチング先となるリンクとの距離である。

上で述べた2段階に加えて今回はマッチング点の修正という段階を加えている。これはログ取得間隔が長いことに起因して必要となるものであるが詳しくは後ほど述べる。以降では各段階について詳しく述べる。

a) 候補ポイントの検索

まず、マッチング先の候補となるリンク (以降、候補リンクと呼ぶ) を検索する範囲を設定し、ログからの距離がその範囲内に収まるリンクを候補リンクとし、点としてのマッチング先の候補 (以降、候補ポイントと呼ぶ) をそのリンク上のログから最も近い点におく。またそのリンクとログとの距離はリンク上の候補ポイントとログとの距離、つまり最短距離とする。ただし、前後のログの候補ポイントからの移動において必要な歩行速度が2.5m/sを超える場合はその点を候補ポイントとみなさない。これはそのようなログ間の移動は現実的な歩行では起こる可能性が低いと考えられるからである。このように前後の候補ポイントからの歩行速度を考慮しているため、i番目のログ (以降、ログ*i*と呼ぶ) の候補ポイント検索時にはログ*i+1*の候補ポイントも同時に検索する必要がある。同じようにログ*i*の候補ポイントはログ*i-1*の候補ポイント検索時に既に検索されているため、ログ*i*の候補ポイント検索時にはログ*i+1*の候補ポイントを検索し、上記の必要歩行速度の条件を満たさないものを除外する作業を行うだけで十分である。

b) マッチング先の決定

以下にログ*i*のリンク*l*上にある TWS を示す。ただし、ログ間の歩行速度を考慮しているためこの値はログ*i-1*の位置に依存しており、下の式ではログ*i-1*の候補ポイントがリンク*m*上にある場合を示している。式中の*S*がログ間歩行速度を示しているのだが、これはパラメータ*a*により場合分けしている。これは歩行速度

が*a*以下であればTWSに影響を与えない、つまり、歩行速度が現実的な値である2つの候補ポイントがある場合は、歩行速度の影響を無視したマッチングが行われることを示している。

$$TWS_{i,l} = \begin{cases} D_{il} + bS & (i \neq 0) \\ D_{il} & (i = 0) \end{cases}$$

$$S = \begin{cases} s(C_{i-1,m}, C_{i,l}) (s(C_{i-1,m}, C_{i,l}) > a) \\ a (s(C_{i-1,m}, C_{i,l}) \leq a) \end{cases}$$

また上記のように本研究では事後的なマップマッチングが求められているため、上記のトータルウェイトスコアを候補リンクごとに計算した上でトリップに含まれるログすべてに関して足し合わせ、その和が最小となる候補リンクの組み合わせをマッチング結果とする。つまりトリップに含まれるログの集合を*I*とすると $\sum_{i \in I} TWS_{i,l_i}$

を最小とする*l_i*を全ての*i*に関して求める

c) マッチング点の修正

上記の方法で求めたマッチング結果では不自然な経路をたどる可能性がある。これはログ取得間隔が比較的長いこと、垂直的なマッチングを行っている、つまり候補ポイントを候補リンク上のログから最も近い点においていること、計算時間短縮のために候補リンク数の最大値を設けていることに起因していると考えられる。そこで前後のログからの移動において用いられるリンクを比較し、それらに重複がないようにマッチング点の修正を行った。

(2) パラメータの決定と精度

アルゴリズム内で用いられているパラメータは以下の3つである。

- ①候補リンクを検索する際の検索半径
- ②TWSの式中、歩行速度にかかる係数(*b*)
- ③TWSの式中、歩行速度決定に用いる速度閾値(*a*)

これらを決定するためにGPS携帯で位置情報を得つつ実際に歩行するという実験を行った (サンプル数720)。その結果、①に関しては実験で観測された点と実際に歩行したリンクの距離の半数近くの点は10m以内に収まっている一方で、50mを超えるようなログも存在したことから、すべてのログにおいて実際に利用されたリンクが候補リンクのひとつとなるためには非常に大きな検索範囲が必要となると言え、検索距離が大きくなれば正しいリンクがマッチング候補に含まれる可能性が高くなることは容易に想像できる。その一方で、検索距離が大きくなるにつれ、検索範囲内に含まれるリンク数も大きくなり、最適な候補リンクの組み合わせを探す際の組み合わせ数が増大するため、マッチングに要する処理時間が大幅に増大する。反対にネットワークの密度が場所に

よって異なることから検索範囲が小さすぎると、その範囲内にリンクが存在しない状態が生じる。ゆえに今回は検索範囲の初期値を20mとし、前後のログからの移動に必要な歩行速度の条件を満たす候補リンクが検索範囲内に含まれるまで検索範囲を5m刻みで大きくし、かつ候補ポイント数に上限を設ける方法をとった。

残り2つのパラメータに関しては様々な組み合わせを試してみても精度が高いものを採用した。(b=79, a=1.05の時であり、78.922%のログが実際に歩行したリンク上にマッチングされた。)

4. マッチング結果の特性

(1) マッチングの精度による問題

前章で述べたマップマッチングの精度の問題から実際にユーザーが歩行した経路とマップマッチングにより推定された経路に違いが生じている可能性がある。そこで経路が利用されたか否かに着目するのではなく、推定された経路中に含まれるリンクに着目する。つまりそれぞれのリンクが利用されたか否かに着目する方法で選択肢集合の生成を試みる。よって以降では経路選択肢集合の生成を経路選択肢集合に含まれるリンク集合の生成と言い換える。それに加え、一定割合以上のトリップが通過するリンクを求めるリンクとすることでマッチングのエラーによる影響を考慮することにした。以降では5%以上のトリップが利用したリンクを求めることを試みた。

(2) ODの定義

マップマッチングの結果として得られた歩行経路は通過したノードの集合として表現することができる。しかし、起点ノード(トリップに含まれる最初のノード)と終点ノード(トリップに含まれる最後のノード)が一致しているトリップ数が10以上のノードの組み合わせは存在しておらず、これは選択肢集合を生成するには不十分なサンプル数であると考えられる。そこで通過するノードすべてにおいてそのノードまでの経路とは独立に経路選択行動が行われるという仮定を置き、ひとつのトリップを複数のトリップに分解することにした。これにより十分なトリップ数を有するODが生まれ、経路選択肢集合の生成が可能になる。今回はOD間のトリップ数が50を越え、かつ最短経路長が200m以上のODを対象とした。このようなODは対象エリア内に118個存在する。

5. リンク集合の生成

(1) リンクの評価指標

ここでは以下に示す二つの指標をもとにリンク集合の生成を行った。

①迂回率

経路を対象として経路選択肢集合を求める場合、各経路の経路長を説明変数とすることが最も一般的である。しかし、上記のように利用リンク集合を求めるため、経路長に代わる説明変数を次のように定義する。

あるOD間の移動に着目した場合、すべてのリンクに関してそのリンクを通過した場合の最短経路を引くことが可能である。その経路長をOD間の最短経路長で除したものをそのリンクの迂回率と定義し、「指標A」と呼ぶ。

②インテグレーション値を考慮した値

ここでは経路長以外のリンクの利用されやすさを表す指標としてインテグレーション値を導入する。これはHillier et al.(1984)によって提案されたものであり、空間をつなぐ意味での重要性を示す値である。詳しくはHillier et al.(1984)を参照されたい。ただし、経路選択にはその経路長が大きく影響を与えており、このインテグレーション値のみで利用リンクを推定することは困難である。そこで以下で説明する平均インテグレーション値をその経路の迂回率で除した値を用いる。この値は最短経路と比較して単位割合あたり遠回りをした場合に享受できるインテグレーション値と解釈することができ、インテグレーション値と空間の利用人数には正の相関があるのでこの値が大きいくほど経路として選択されやすいと考えられる。各リンクの迂回率計算時と同様にリンクを通過する経路のうち平均インテグレーション値/迂回率が最大となる経路を検索し、検索された最大値をそのリンクの平均インテグレーション値/迂回率と定義し、「指標B」と呼ぶことにする。

(2) 利用リンク集合の特性把握

以下の手順で利用されたリンク集合の特性を把握した。

- ①すべてのリンクを指標に基づきソートする。
- ②求めるリンク集合 S が完全に含まれる順位までを、その指標から求められるリンク集合 S' とする。
- ③ S' 中に含まれる誤差や、 S が完全に含まれる順位における指標の値の分布に関する考察を行う。ただし、ここでの誤差とは S' に含まれるリンクのうち全く利用されていないリンクの総リンク長を S の総リンク長で除したものを指している。

Aによるソートのケースで S が完全に含まれる順位における指標の値とは S' に含まれるリンクの迂回率の最大値であり、これを a_{\max} と呼ぶことにする。

また、Bによるソートのケースで S が完全に含まれる順位における指標の値の分布をみる場合、ソート結果、最上位となるBの値との比をとることにした。これは、迂回率が最短経路との比で計算されているのに対して、Bに含まれるインテグレーション値は相対的な値ではなく、OD周辺に存在するリンクのインテグレーション値

に依存していることを考慮するためである。つまり、上記のように比をとることで、ODの相違が分布に与える影響をなくしているといえる。ここでその比の値を b_r と呼ぶこととし、 S が完全に含まれる順位における b_r は S' 中で最小の値を示すため、 $b_{r, \min}$ と呼ぶことにする。

(3) 指標Aによるソート

まず、誤差は非常に小さいことから利用されたリンク集合がある程度、迂回率で説明できると推測できる。

また、最大迂回率も2を超えているものがいくつか見受けられるが、1.6付近に集中している。(図4) また S' の総リンク長を最短経路長で除したものの分布を見てみると図5のようになっており、最大迂回率が1.6付近に集中している一方でそれはばらついており、経路数の多いODペアと少ないODペアの存在が見取れる。つまり、OD周辺のネットワーク密度に関係なく、一定迂回率以下の経路が利用されており、経路集合の生成を考えた場合、経路の数には関係なく、歩行者が許容できる迂回率以下の経路が選択肢となっていると言える。

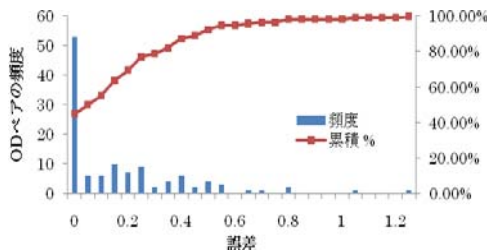


図3 Aソートによる完全推定時の誤差分布

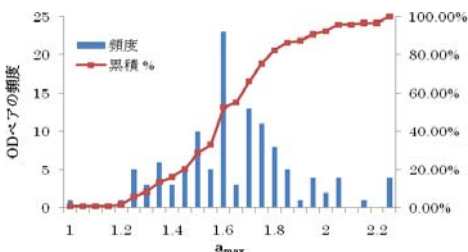


図4 Aソートによる完全推定時の a_{\max} の分布

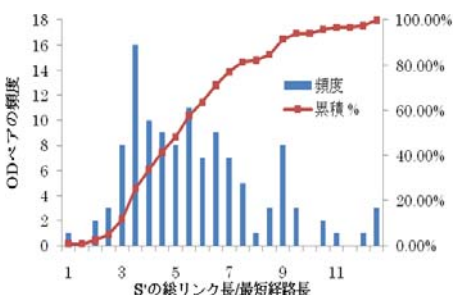


図5 Aソートによる完全推定時の総リンク長/最短経路長分布

(4) 指標Bによるソート

迂回率の場合と同様に求めるリンク集合を S とすると誤差は非常に小さくなっていることから、95%の人

が利用するリンクはこの平均インテグレーション値/迂回率で説明できると言える。また $b_{r, \min}$ の分布を見ると、0.6付近に集中している。(図7) 一方で S' に含まれる総リンク長をOD間の最短経路長で除したもののばらつきは大きい迂回率の場合と同様に歩行者には許容できる b_r の値が存在していると考えられる。

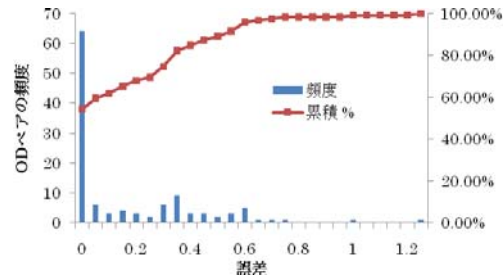


図6 Bソートによる完全推定時の誤差分布

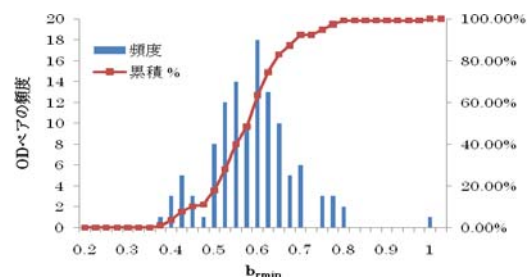


図7 Bソートによる完全推定時の $b_{r, \min}$ の分布

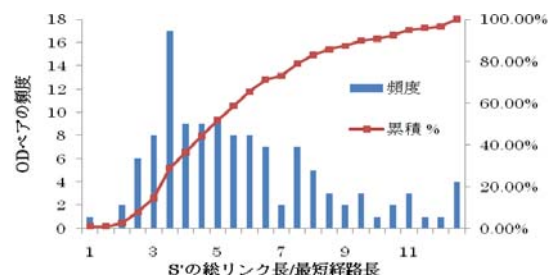


図8 Bソートによる完全推定時の総リンク長/最短経路長の分布

(5) リンク集合の推定

a_{\max} , $b_{r, \min}$ それぞれの平均値をその閾値とし、Aを用いる場合、迂回率がその値以下、Bを用いる場合、 b_r がその値以上のリンク集合を推定リンク集合 S_e とする方法で経路選択肢集合に含まれるリンク集合の推定を試みた。 S のうち S_e に含まれるリンクの割合を的中率と呼び、図9に、 S_e に含まれる全く利用されていないリンクの $S_e \cap S$ に対する割合を誤差とし、図10に示した。ただし、ともにリンク長ベースで計算している。Bを用いたほうが誤差が小さくなっていることが見て取れる。

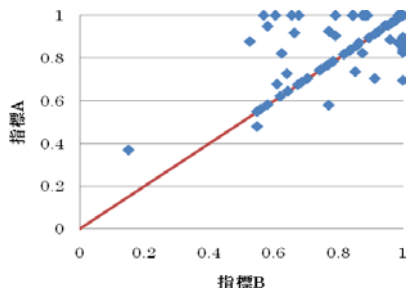


図 9 推定方法の違いによる利用リンク的中率の比較

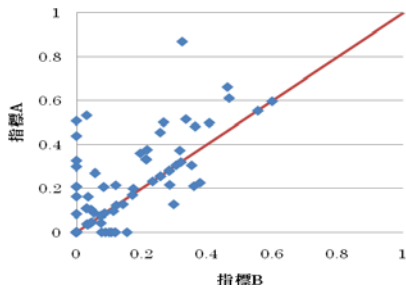


図 10 推定方法の違いによる誤差の比較

6. 終わりに

(1) まとめ

まず、今回提案したマップマッチング方法で、ログ取得間隔が比較的長いデータであってもある程度の精度で利用したリンクの推定が可能であることがわかった。

また、マッチング結果を用いて利用リンク集合の推定を試みた結果、インテグレーション値を用いることで経路長の場合と比較して、精度の向上が見られた。そこで本研究ではリンク集合の推定方法として以下の手順を提案する。

- ① 周辺エリアのインテグレーション値をアクシアル分析から求める。
- ② 各リンクに対して平均インテグレーション値／迂回率が最大となる経路を引き、その値を、そのリンクの平均インテグレーション値／迂回率とする。
- ③ 全てのリンクのうち平均インテグレーション値／迂回率が最大のものを検索し、それを平均インテグレーション値／迂回率の基準値とする。
- ④ 全てのリンクの平均インテグレーション値／迂回率を上記の基準値で除したものが閾値以上のリンクを経路選択肢集合に含まれ

るリンクとする。閾値としては今回の研究から 0.577 が得られた。

(2) 今後の課題

本研究では経路選択集合を直接推定するのではなく、その集合を構成するリンク集合を推定した。経路は連続したリンクの組み合わせで表現されるため、推定されたリンク集合から実際には経路選択肢集合に含まれない経路が生成される可能性が残っている。この点について、リンクの接続関係などから候補となる経路を抽出する手法等を今後検討する必要がある。

参考文献

- 1) 岡田絵里子, 2006, 集計ロジットモデルに基づく鈍者流動モデルによる歩行回遊性向上に関するシミュレーション評価, 東京大学修士論文。
- 2) 外井哲志, 1981, 街路における歩行経路に関する研究, 交通工学, Vol.16, No.4, pp.11-18.
- 3) 塚口博司, 松田浩一郎, 2002, 歩行者の経路選択行動分析, 土木学会論文集, No.709/IV-56, 117-126.
- 4) 中村和夫, 小林實, 1984, 交通環境における歩行行動, 国際交通安全学会誌, vol.10, No5, pp.300-313.
- 5) 毛利正光, 塚口博司, 1979, 歩行者の経路選択特性について, 土木学会関西支部年次学術講演会講演梗概集, IV28-1, IV28-2.
- 6) 吉永誠, 武田欣弘, 松永千晶, 厲国権, 角知憲, 2001, 路上条件を考慮した歩行者流動モデルに関する研究, 土木学研究・論文集, Vol.18, No.3, pp.463-469.
- 7) Greenfeld, J.S., 2002. Matching GPS observations to locations on a digital map. In proceedings of the 81st Annual Meeting of the Transportation Research Board, January, Washington D.C.
- 8) Velaga, N.R., et al. 2009. Developing an enhanced weight-based topological map-matching algorithm for intelligent transportation systems. Transportation Research Part C.
- 9) Bill Hillier, Julienne Hanson, 1984, The Social Logic of Space