

# Wi-Fi位置情報データを用いた 3次元ネットワークにおける移動経路の推計

後藤 祥孝<sup>1</sup>・羽藤 英二<sup>2</sup>

<sup>1</sup>東京大学工学部 社会基盤学科 (〒 113-8656 東京都文京区本郷 7-3-1)  
E-mail: goto@bin.t.u-tokyo.ac.jp

<sup>2</sup>東京大学工学部 社会基盤学科 (〒 113-8656 東京都文京区本郷 7-3-1)  
E-mail: hato@bint.t.u-tokyo.ac.jp

本研究では、経路選択行動の観測を3次元ネットワークに拡張し、機械学習の手法を取り入れたマップマッチング手法を提案する。屋内においても十分な観測精度を保つことの可能なWiFiによる位置情報観測データを用いる形でのランダムフォレストによるノード、リンクへのマッチングを行った。さらにリンクへのマッチングに対して経路選択モデルと経路選択枝を導入した経路推計手法の提案を行った上で、数値実験によりWiFiアクセスポイント数の変化による推定精度の変化を示した。

**Key Words :** Route choice model, WiFi data, Map matching, Machine learning

## 1. はじめに

駅一まちのような複雑化した動線の立体的な空間構成と、人々の多様な回遊行動を分析する上で、3次元での位置特定技術は必要不可欠であるといえよう。既往のマップマッチング技術はGPSの位置情報を前提としたもので、3次元の移動軌跡の特定において十分に機能するとは言い難い。GPSの位置特定精度は屋内において著しく低く、高さ方向の精度を確保することは難しい。こうした問題に対して精度の異なる屋内位置観測データを組み合わせた解析手法のニーズが高い。特にWiFiやBluetoothによる観測では数十cm数m単位での位置特定が可能であるものの、取得間隔距離が長く移動速度によって欠損のある場合に移動経路をどのように特定するかという問題を扱う研究は多くない。Danalet et al.<sup>1)</sup>では観測方程式とシステム方程式を組み合わせた位置特定の研究が行われているものの、アクティビティ場所の推定にとどまっており、アクティビティ間の移動経路の推計までは行っていない。本研究ではこうした既往研究に着目し、経路選択行動の観測を3次元の立体ネットワークに拡張するとともに、機械学習の手法を取り入れた位置特定手法の提案を試みる。

## 2. 既往研究の整理

### (1) マップマッチング手法

GPSやWiFiといった誤差を持つ位置情報が得られたとき、取得された位置情報を用いてある時点で旅行

者がどの経路を利用しているのか、また経路上のどの位置(x,y,z)にいるのかを推定する手法がマップマッチングである。また経路は、リンクとノードから構成されるネットワーク上の軌跡を通過したリンク-ノード配列によって表現する。マップマッチングは位置情報データとネットワークデータの扱い方によって以下の3種類に大別される。

幾何解析マップマッチングは、点を表す観測値とネットワークデータのリンク、ノードとの幾何的情報のみを利用して行う方法である。Bernstein et al.(1998)<sup>2)</sup>によるpoint to point手法、White et al.(2000)<sup>3)</sup>によるpoint to curve手法、curve to curve手法に大別される。

point to point手法は、観測データとノードとの距離が最短となるように観測データをノードにマッチングするのに対して、一方、point to curve手法は観測データとリンクの距離が最短となるように観測データをリンクにマッチングする。

こうしたマッチング手法はいずれも観測データとの距離のみを利用しているため、計算量が抑えられる一方で、ネットワーク上のリンク接続条件を考慮しないため、実際には接続していないリンクやノードに対してマッチングされる可能性がある。またネットワークデータの密度によって精度が左右される。といったことが課題といえよう。

位相幾何解析マップマッチングは、点の観測値とネットワークの幾何的情報に加え、リンクとリンクの接続性や連続性を考慮してマッチングを行う手法といえる。連続する観測データとリンクとの交差角度や相対角度

距離を用いて位置情報が乗るリンクを決定したのちに、速度や進行方向角度を用いてリンク上での位置を決定する。

確率的マップマッチングは観測値が測位誤差を持つことを前提に、その測位精度の分布をもとに居たであろう位置とその尤もらしさに着目してマッチングを行う手法を総称して確率的マップマッチングと呼ぶ。

一般状態空間モデルは実状態の時系列での遷移を表現するシステム方程式とある時点における実状態と観測データとの関係を表す観測方程式の2つの式を連立させたモデルである。本研究において利用する一般状態空間モデルを用いた手法もここに含まれる。観測者は実状態を知り得ず、観測値のみが入手できるという前提の上で、観測方程式によって観測データと実状態の関係を表現をすることによって、観測データを実状態の時間遷移の式に導入することを可能にしており、交通分析で用いられてきた共分散構造モデルも基本的に同様のモデルと言えよう。Danalet et al.(2014)<sup>1)</sup>はWiFi観測データを用いて歩行者の時系列でアクティビティの連続推定に一般状態空間モデルを適用している。観測方程式を観測値ベクトルと活動場所の関係の尤度関数として記述し、システム方程式は次状態のアクティビティに対するアクセシビリティを表現するものとなっている。

## (2) 経路選択モデル

ネットワーク上の個人の行動を記述するために様々なモデルが研究されてきた。経路選択モデルはネットワーク上での個人の移動経路の選択を扱う。交通機関選択のような一般的な離散選択モデルとの違いとして選択肢集合の複雑さがあげられよう。格子状ネットワークの場合でも簡単に組み合わせ爆発を起こしてしまうという問題を抱えている。こうした課題への対処として、選択肢を明示的に列挙する方法と選択肢列挙を必要としない選択確率の導出という2つのアプローチで経路選択問題が扱われてきた。ここではそれぞれ、選択肢列挙型経路選択モデル、選択肢非列挙型経路選択モデルと分類する。

### a) 選択肢列挙型経路選択モデル

選択肢列挙型の特徴は、ネットワークデータ上で出発地と到着地が定まっている状態で、出発地と到着地を結ぶような経路を列挙し、選択確率を求めるという点である。選択肢間に誤差相関がないとする場合は選択確率は以下の式(1)によって表される。

$$P_i = \frac{\exp(V_i)}{\sum_{j=1}^J \exp(V_j)} \quad (1)$$

経路選択肢は取り得る選択肢数が膨大になることから、

全てを列挙することは非現実的である。従って限定的に経路を列挙する手法が検討されてきた。最短経路からk番目までを列挙し選択肢とする手法や、距離以外のパラメータをネットワークに設定し、各パラメータでの最短距離を取り出すことで選択肢を生成するラベリング法が挙げられよう。

### b) 選択肢非列挙型経路選択モデル

選択肢非列挙型では、各ノードにおいて次状態のノードの選択が確率的に行われて遷移するという、逐次的な経路の選択が行われる。

佐佐木(1965)<sup>4)</sup>はノード間の遷移を確率過程と捉えて遷移確率行列を与えることで、吸収マルコフ連鎖として記述している。この場合、目的地を明示しないが、遷移の繰り返しの中で周回構造も考えることができるという特徴がある。

また、逐次のノード選択を離散選択モデルとして捉えることによる動的離散選択モデルによるアプローチも行われている。こちらはノード遷移の際に目的地までの期待効用を考慮することで、目的地からの再帰的な構造で確率を求めることにより経路列挙を不要としている。ネットワーク内に正の効用を持つ周回構造が含まれる場合には期待効用の計算が不能となる問題を抱えていたが、大山・羽藤(2016)<sup>5)</sup>ではネットワークに対して時間構造化を行うことで正の効用を持つ周回構造のあるネットワークに対しても、時間構造化ネットワーク上では周回構造を除去し計算可能にしている。 $a$ から $a'$ に遷移する際の経路選択確率は直接的効用 $v(a'|a)$ と期待効用の評価関数 $V(a')$ と期待効用に対する時間割引率 $\beta$ を用いて式(2)で表される。

$$P(a'|a) = \frac{\exp\{\frac{1}{\mu}(v(a'|a) + \beta V^d(a))\}}{\sum_{a' \in A(k)} \exp\{\frac{1}{\mu}(v(a'|a) + \beta V^d(a'))\}} \quad (2)$$

### (3) 機械学習

機械学習は画像データからの物体認識や言語処理の分野で多く用いられており、観測データからの位置特定の研究分野において用いられることは多くない。Shafique et al.(2014)<sup>9)</sup>ではプローブパーソンデータを用いて移動状態での交通手段推定において機械学習を適用し、交通期間の自動識別で高い精度を得ている。

ランダムフォレスト<sup>10)11)</sup>は複数の決定木を組み合わせることで森のようにすることによって分類・回帰を行う機械学習アルゴリズムである。学習データからランダムサンプリングしてサブサンプルを生成して各決定木に当てはめること、学習データの説明変数が複数存在するときランダムに選択された変数を用いて決定木でのノードにおける分割を行うことの2つのランダム性を有している。以下に、フローを示す。

$d$ 次元の入力データ  $\mathbf{v} = (x_1, x_2, \dots, x_n)$ , 学習データ集合を  $S = \{\mathbf{v}\}$ , 生成する決定木の本数を  $T$  とする.  $t$  番目 ( $1 < t < T$ ) の決定木の根ノードにランダムサンプリングされた学習データ集合  $S_0^t \in S$  が与えられたとき, ノードにおける最適な分割関数を求めて子ノードにデータを分割し, 子ノードでも分割関数を決定することを繰り返して決定木を生成する.

ここで分割関数とは, 決定木のノードにおいて学習データを分ける関数であり  $h(\mathbf{v}, \theta)$  という形で書ける.  $\theta$  は分割を決定するパラメータであり, 入力  $\mathbf{v}$  からランダムに特徴量を抽出するフィルタ, 分割基準の幾何特徴と閾値を決定するパラメータとして働く.

ノード  $i$  の学習に用いるデータ  $S_i$  とし, 分割された左右の子ノードの学習データ  $S_i^L, S_i^R$  とする. ノードにおける分割関数は  $\theta$  の最適パラメータ  $\theta^*$  によって決定されるため式 (3) から求める.

$$\theta^* = \arg \max I(S_i, S_i^L, S_i^R, \theta) \quad (3)$$

目的関数  $I$  はエントロピーと情報利得の式によって定義される.

集合  $S$  内のデータにクラスラベル  $c \in C$  が対応しているときエントロピーは式 (4) のように表される.  $p(c)$  は集合  $S$  内のデータのクラス  $c$  に分類される割合である.

$$H(S) = - \sum_{c \in C} p(c) \log(p(c)) \quad (4)$$

$S$  内が単一のクラスの時,  $H(S) = 0$ ,  $S$  内のデータが全クラスで等確率の時  $H(S)$  は最大となる. また, ノードの分割によって得られる情報利得は式 (5) のように表される.

$$I = H(S) - \sum_{i \in L, R} \frac{|S^i|}{|S|} H(S^i) \quad (5)$$

式 (5) の第 1 項は分割前エントロピー, 第 2 項は分割後のエントロピーを意味している.  $I$  が大きいほど分割により得られる利得が大きくなるため,  $I$  が最大となるようにパラメータ  $\theta$  を決定する.

以上のような分割の繰り返しによりノードに割り当てられたデータが一定値以下になったところで木の生成が終了しクラスラベル  $c \in C$  に対して, 確率  $p_t(c|\mathbf{v})$  が決定される. 全ての木の生成が終了すると相加平均  $p(c|\mathbf{v}) = \frac{1}{T} \sum p_t(c|\mathbf{v})$  によってランダムフォレストによるクラス識別を行う.

### 3. WiFi 位置情報データに対するマップマッチング手法の導入

#### (1) 位置情報の形式

WiFi の受信端末はネットワーク上に設置された WiFi アクセスポイントから発信される電波によって通信を

可能としており, そうした WiFi 受信の記録を本研究では位置特定に利用することを仮定する. 従って, 観測データとして与えられるのは端末が受信したアクセスポイント (複数も許容する) の名前とそこからの電波強度である. 但し, アクセスポイント側で記録される端末 ID と時系列データを用いても同様の解析ができることに注意したい. 時系列のデータを用いる場合, 時刻, アクセスポイント名, 電波強度の列を解析に用いる. 電波強度は距離を変数として減衰する関数に従うとすれば, 電波強度のデータは距離に変換することができる. 従前の位置情報の取得やマップマッチングに用いられてきた GPS によるデータは 3 次元座標上での絶対的な位置を表現するのに対し, WiFi によるデータの場合アクセスポイントとの距離という相対的な位置関係の情報しか得ることができない. 従って, マップマッチングを行う上では相対的な位置情報を絶対的な位置情報へと変換する必要がある.

2 つのアプローチが考えられる. 一つはアクセスポイントの座標系上での位置情報を既知として与えることによって端末の位置を割り出す手法であり, スマートフォンにおける WiFi を用いた位置特定の方法として用いられている. この場合, 対象地域内をカバーするために十分なだけのアクセスポイントの位置情報データベースが必要となる. 二つ目のアプローチとしては機械学習の教師あり学習を用いた手法である. この場合, 座標系上での位置が既知である WiFi データを教師データとして利用することで未知データの位置を推定する手法である. 教師データの多寡が推定精度を左右するものの, ネットワーク上の離散的なノード, リンクといった情報にマッチングする際には, 分類木を用いることで座標を介さずにデータを扱うことができるという利点がある.

以下では機械学習を用いたアプローチに着目して, マップマッチング手法の導入を行う.

#### (2) 機械学習を用いたアプローチ

本研究において利用するのは, 機械学習のアルゴリズムの中でも教師あり学習の分類器である. 分類の場合, 特徴量を表す入力ベクトルと出力に対応する分類を学習することによって, 分類未知の入力データの分類を返す. WiFi データでのマップマッチングであれば特徴量は端末が受信したアクセスポイント名とその電波強度ということになる. 従ってアクセスポイント数を  $N$  個とすれば, 入力データベクトル  $\mathbf{v} = (x_1, x_2, \dots, x_N)$  とかけて  $n \in N$  について  $x_n =$  (アクセスポイント  $n$  の電波強度) となる. また, 分類に用いるラベルとしてはノードあるいはリンクということになる. 従って, 時系列での WiFi データが与えられた場合の出力

は、各時点における滞在ノードまたは滞在リンクである。マップマッチングの目的である経路データを得るためには本研究では以下の 2 つの方法を提案する。

a) 幾何的接続

機械学習において得られる出力は各観測値に対するノードあるいはリンクである。従って、幾何的マップマッチング手法である、point to point や point to curve の手法と同様に、ノードあるいはリンクを順に結ぶことによってネットワーク上の経路とすることができる。

b) 経路選択モデルの導入

機械学習による滞在リンクの推定結果と経路選択確率を掛け合わせることで経路尤度を各経路に対して算出し、経路尤度が最大となるものを採択することで経路の推計を行う。図 1 に計算フローを示す。まず経路選択肢の決定によって、推計する経路の選択肢を決定する。列挙された各選択肢に対して、経路効用を設定することにより、ロジット型離散選択モデルを適用することで選択確率を計算することが可能である。同時に機械学習による各時点での位置推計結果と経路選択肢を合わせることで各データに対して経路ごとの観測尤度を計算する。本研究では、推計された時系列の滞在リンクが各経路の構成リンクに含まれているか否かを調べることで、リンクの一致率を計算し観測尤度とする。さらに、経路尤度 = (経路選択確率) × (経路観測尤度) とすることで経路尤度を計算する。最終的に計算された各経路尤度が最大となるような移動経路とする。

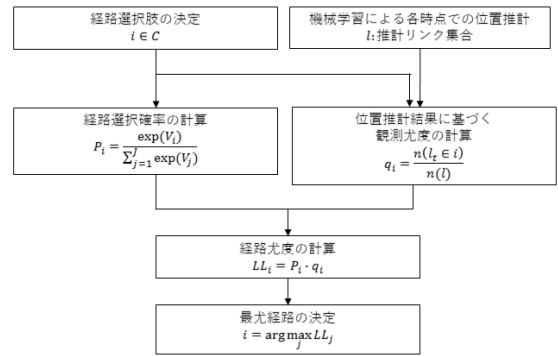


図-1 計算フロー

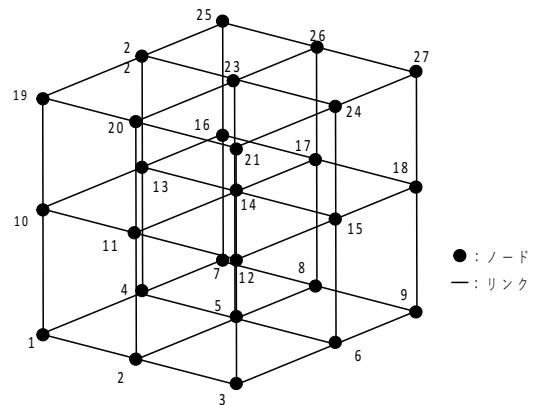


図-2 数値実験におけるネットワーク図

4. 数値実験

本章では前章で導入したマップマッチング手法の検証として数値実験を行う。既知の実状態を仮定した上で、それらの実状態に対しての認知や観測による誤差を加味した観測データを発生させることにより、推定手法の比較検討を可能にしている。ここではネットワークデータの作成、シミュレーションデータの発生、各手法の適用の順に進めていく。

(1) ネットワークデータ

本研究では図 2 のような 10 × 10 × 10 の立方体を 8 個積み重ねた立方格子状のネットワークを考える。各頂点、各辺をそれぞれノード、リンクとするためノードは全 27 個、リンクは全 54 本となる。

(2) シミュレーションデータの発生

本研究で想定するのは WiFi データが得られた際の位置推定であるため、数値実験でも同様に WiFi のアクセスポイントを設定する必要がある。アクセスポイント

トには名前と 3 次元座標軸での位置座標を与える。

本来の WiFi データは、時刻、アクセスポイント名、電波強度からなるため、数値計算でも時刻、アクセスポイント名、電波強度を模したパラメータを発生させる。時刻はデータ取得間隔と考えると、 $t=1,2,\dots,T$  として整数で順に与える。電波強度を模したパラメータについては、ノードとアクセスポイントの距離の 2 乗に反比例する形の関数を設定する。これは式 (6) のフリスの伝達公式から距離の 2 乗に反比例することがわかるため、式 (7) のように簡略化して与える。

$$P_R = \left( \frac{\lambda}{4\pi D} \right)^2 G_t G_r P_T \tag{6}$$

$P_R, P_T$ : 電力  
 $G_T, G_R$ : アンテナ絶対利得,  $D$ : 通信距離

$$x_{pn} \propto \left( \frac{1}{d_{pn}} \right)^2 \tag{7}$$

$x_{pn}$ :  $p$  番目データのパラメータ  $n$   
 $d_{pn}$ :  $p$  番目データの実状態の位置とアクセスポイント  $n$  との距離

ケース	教師データ数	カバー率	アクセスポイント位置
1-1	100	4/27	1,9,21,25
1-2	1000	4/27	1,9,21,25
1-3	10000	4/27	1,9,21,25
1-4	100	5/27	1,9,14,21,25
1-5	1000	5/27	1,9,14,21,25
1-6	10000	5/27	1,9,14,21,25
1-7	100	6/27	5,11,13,15,17,23
1-8	1000	6/27	5,11,13,15,17,23
1-9	100	7/27	5,11,13,14,15,17,23
1-10	100	8/27	1,3,7,9,19,21,25,27

表-1 ノードにマッチングする際のアクセスポイントの位置と教師データ数

誤差はノードを選択した際の  $x, y, z$  の位置座標のそれぞれに対して平均 0.0 標準偏差 1.0 の正規分布で与えられる白色誤差を足すことによって与えている。また、アクセスポイントからの距離が遠くなることや現実には壁が存在することなどにより受信不可能となる条件を考慮して、アクセスポイントからの距離が 20 以上となった場合には電波強度  $x = 0$  とした。

推定に用いるデータは、時刻 0 から T までの各時刻における位置座標を設定したうえで、各時刻でのアクセスポイント名と電波強度からなる WiFi 模擬データを発生させることにより得る。

また、機械学習を利用する場合は学習用の教師データと推定に用いるデータとは別に必要となる。ここで推定するのは各時刻において位置するノードまたはリンクであり、特徴量としては各アクセスポイントからの電波強度を与える。従って教師用データは時刻の情報は不要であり、実状態(ノード番号あるいはリンク番号)と各アクセスポイントからの電波強度からなるデータを作成すればよい。

### (3) 機械学習による数値実験

ここでは機械学習を用いた WiFi データのマッチングの実験と同時にアクセスポイント数、教師データ数を変化させたときの精度の変化を数値実験により検証する。

#### a) ノードに対するマッチング

シミュレーションデータはネットワークデータ上でランダムにノードを選択し、ノードの座標に白色誤差を加えることで位置座標を決定し、その点における各アクセスポイントとの距離を利用して模擬 WiFi データを出力する、という操作を繰り返し行うことによって作成する。

表 1 において各ケースにおけるアクセスポイントの位置とノードのカバー率、教師データの数を示す。アクセスポイントの位置は図 2 中の数字と対応している。

アクセスポイントの位置は立方体格子上で対称となるように配置している。例えば、4 個の場合は正四面体

ケース	教師データ数	カバー率	アクセスポイント位置
2-1	10000	4/27	1,9,21,25
2-2	10000	5/27	1,9,14,21,25
2-3	10000	6/27	5,11,13,15,17,23
2-4	10000	7/27	5,11,13,14,15,17,23
2-5	10000	8/27	1,3,7,9,19,21,25,27
2-6	10000	14/27	1,3,5,7,9,11,13,15,17,19,21,23,25,27
2-7	10000	18/27	2,4,5,6,8,10,11,12,13,15,16,17,18,20,22,23,24,26

表-2 リンクにマッチングする際のアクセスポイントの位置と教師データ数

表-3 ノードにマッチングしたときの推計結果

教師データ数	アクセスポイント数				
	4	5	6	7	8
100	48%	52%	1%	0%	6%
1000	40%	51%	1%	-	-
10000	35%	41%	-	-	-

表-4 リンクにマッチングしたときの推計結果

ケース	誤判別率
2-1	64.2%
2-2	62.7%
2-3	25.2%
2-4	24.1%
2-5	25.1%
2-6	20.5%
2-7	15.9%

の頂点、5 個の場合は正四面体の頂点と重心に置く。表 3 において各ケースでの機械学習による推定結果を示す。分類不能となっているものは、どのアクセスポイントも受信不可能、すなわち全アクセスポイントに対する電波強度が 0 となったものを表す。

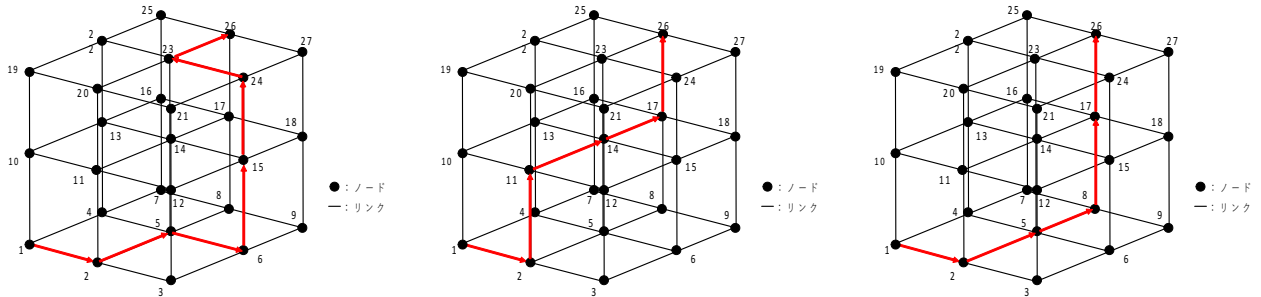
結果よりアクセスポイント数が 6 より大きくなることで急激に判別率が向上していることが分かる。一方でアクセスポイントが最も多いケース 8 において誤判別率が增大している。これはケース 10 のようなアクセスポイントを頂点においた場合、ケース 7, 9 のように面の中心に置いた場合と比較して、ネットワークが作る立方体と各アクセスポイントのカバー範囲が共有する体積が小さくなることに起因するのではないかと考えられる。結果として各データにおける意味のあるパラメータ数が減ることで全体の精度を下げる要因になるだろう。

#### b) リンクに対するマッチング

シミュレーションデータの作成方法について以下に述べる。ネットワークデータ上でランダムにリンクを選択した上で、一様分布の乱数を発生させることで

表-5 経路選択枝とその効用

経路	通過ノード	効用 (確定部分)
1	1 → 2 → 5 → 6 → 15 → 24 → 23 → 26	-46
2	1 → 2 → 11 → 14 → 17 → 26	-45.5
3	1 → 2 → 5 → 8 → 17 → 26	-47.5



[1] 経路 1

[2] 経路 2

[3] 経路 3

図-3 経路選択枝

リンク上での位置座標を決定し 1 つのデータの観測地点とする。観測地点における各アクセスポイントとの距離を利用して模擬 WiFi データを出力する、という操作を繰り返し行うことによってデータを作成する。

表 2 に各ケースにおけるアクセスポイント数と教師データ数を示す。結果を表 4 に示す。ノードに対するマッチングに比較して全体として精度が下がっている一方で、アクセスポイントが 6 つ以上になったケース 2-3 以降で精度が大きく向上していることはノードの場合と同様である。アクセスポイントを多くすることによって、判別精度に向上は見られるものの、入力データの特徴量は増加することによる計算時間の増大が起こるというトレードオフが存在している。

c) 経路選択モデルの導入

ここではシミュレーションデータの発生と経路推計において経路選択モデルを導入する。ここで計算された経路選択確率を経路推計においても利用する。経路選択確率の計算では、ネットワークデータの各リンクに移動コストとリンク魅力度を設定することで以下の式 (8) で経路  $i$  による効用を定義する。

$$V_i = \beta_1 \times (\text{経路 } i \text{ の移動コスト合計}) + \beta_2 \times (\text{経路 } i \text{ のリンク魅力度の合計}) + \varepsilon \quad (8)$$

また、ここでは  $\beta_1 = -1.5, \beta_2 = 1$  とする。経路選択確率はロジット型離散選択モデルを用いて式 (9) で計算される。

$$P_i = \frac{\exp(V_i)}{\sum_{j=1}^J \exp(V_j)} \quad (9)$$

計算された経路選択確率に基づき発生させた乱数によって経路を決定したうえで、経路上での移動の位置データ生成を繰り返すことによって、位置データ集合を発生させる。本実験においては 3 枝の経路選択モデルを考える。各経路とその効用の確定部分を表 5, 各経路のネットワーク上の位置を図 3 に示す。

以上で発生させた推定データに対してリンクに対するマッチングで用いたものと同様の教師データを用いることにより、リンク推計と経路推計を行う。ここでもアクセスポイント数を変更した実験を行うが、アクセスポイント数と位置は表 2 と同様に対応している。また、教師データ数は全て 10000 データとしている。表に結果を示す。

全てのケースにおいて、リンク誤判別率に対して経路一致率の値には精度の向上が見られ、経路という概念を導入することによって滞在リンク推定の誤差を吸収できていると考えられる。全て 90% を超える経路一致率であることから経路推計としては十分に働いていることが分かる。また、アクセスポイント数の大小が経路一致率に与える影響が小さいことも分かる。しかし、個別のデータを見ていったときに、経路が不一致となっているのはすべて経路 3 の場合であるという結果が得られた。経路選択モデルを導入したことにより、選択確率が極端に小さい選択枝が含まれている場合には単純に観測尤度を掛けるだけでは、経路推計の際には不十分であると言えよう。また、機械学習による観

表-6 経路推計結果

ケース	カバー率	リンク誤判別率	経路一致率	観測尤度のみによる一致率
1	4/27	60.3%	93%	100%
2	5/27	59.6%	94%	100%
3	6/27	24.4%	92%	100%
4	7/27	25.5%	90%	100%
5	8/27	26.6%	94%	100%

測尤度のみを利用した場合と比べて一致率が下がる結果となった。本実験においてカバー率の設定が十分に高かったことと経路選択肢の類似度が低かったことが原因として考えられよう。アクセスポイント数がさらに少ない場合、重複するリンクを多く含む経路選択肢が存在する場合に経路選択確率を導入する利点があると考えられる。

## 5. まとめ

本研究の成果として、1)WiFi データを用いて 3 次元における位置推定を行う手法を提案した上で、2)経路選択行動モデルと組み合わせることによって経路推定が十分な精度で可能であることを計算例により示した点にある。実データを扱う際には、ネットワーク上には WiFi のアクセスポイントが多数存在しておりそのすべての 3 次元位置座標を調べることが困難な場合であっても、一定のカバー率で記録端末側に残る情報のみを用いて経路特定が可能であると考えられる。一方で全ての取得データを利用することは特徴量の増加による計算量の増大をもたらすため、取得データ全体でのアクセスポイントの総数を抑えることも可能になると考える。

マップマッチングを行う場合には時系列上で移動をするようなデータとなるため、時系列での前後のデータで同じノード、リンクに属していることやせいぜい一つ隣のリンクに移る程度といったことが想定される。本研究においては機械学習における特徴量として各時点における観測データのみを扱ったが、機械学習においての精度向上に向けては、前後の時点での観測データや前時点での推定位置データを利用していくことが考えられよう。

## 参考文献

- 1) Danalet, A., Farooq, B., Bierlaire, M. (2014). A Bayesian approach to detect pedestrian destination-sequences from WiFi signatures. *Transportation Research Part C: Emerging Technologies*, 44, 146-170.
- 2) Bernstein, D., Kornhauser, A. (1998). An introduction to map matching for personal navigation assistants.
- 3) White, C. E., Bernstein, D., Kornhauser, A. L. (2000). Some map matching algorithms for personal

- navigation assistants. *Transportation Research Part C: Emerging Technologies*, 8(1), 91-108.
- 4) 佐佐木綱, (1965). 吸収マルコフ連鎖による交通流配分理論. *土木学会論文集*, 121, 28-32
- 5) 大山雄己, 羽藤英二 (2016). 移動軌跡情報に基づく時間構造化ネットワーク上の交通配分. *土木計画学研究・講演集 53(CD-ROM)*
- 6) Fosgerau, M., Frejinger, E., Karlstrom, A. (2013). A link based network route choice model with unrestricted choice set. *Transportation Research Part B: Methodological*, 56, 70-80.
- 7) Bierlaire, M., Chen, J., Newman, J. (2013). A probabilistic map matching method for smartphone GPS data. *Transportation Research Part C: Emerging Technologies*, 26, 78-98.
- 8) Chen, J., Bierlaire, M. (2015). Probabilistic multimodal map matching with rich smartphone data. *Journal of Intelligent Transportation Systems*, 19(2), 134-148.
- 9) Shafique, M. A., Hato, E., Yaginuma, H. (2014). Using Probe Person Data for Travel Mode Detection. *Int. J. Comput. Inf. Syst. Control Eng. World Acad. Sci. Eng. Technol*, 94, 1501-1505.
- 10) Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- 11) 波部齊. (2012). ランダムフォレスト. *情報処理学会研究報告*, 2012.