

# 集団の行動履歴情報に基づく滞留の時空間特性抽出と行動パターン分析\*

## Clustering-based Spatio-Temporal Property Extraction and Behavior Pattern Mining from GPS Trajectories\*

吉田圭吾\*\*・小池亜弥\*\*\*・松崎和賢†・小関悠\*\*\*・白井康之††

By Keigo YOSHIDA\*\*・Aya KOIKE\*\*\*・Kazutaka MATSUZAKI†・Yu KOSEKI\*\*\*・Yasuyuki SHIRAI††

### 1. はじめに

GPSセンサを搭載した携帯端末から位置情報を取得することが容易となり、個人の移動軌跡のログを蓄積することが可能となりつつある。同時に、foursquareのような、MLBSと呼ばれる移動ログデータに基づく位置情報連動サービスの展開が盛んに行われている。膨大な個人の行動履歴情報を収集、解析して、個人の行動パターンや集団としての行動特性を把握することは、今後想定されるさまざまなサービス展開において有用である。

本研究では、GPSセンサを搭載した携帯端末から取得された集団の行動履歴情報を基に、ある場所を時空間的に特徴付け、頻出する行動遷移パターンを抽出する、というタスクを扱う。「場所を時空間的に特徴付ける」、とは具体的に言えば、行動履歴情報から滞留が発生しやすい地点を検出すると同時に、滞留が発生した時間帯や滞留時間などの滞留状況に関する情報と、滞留者の性別などの属性データを用いて、言わば「いつ、どこが、誰にとって、どのような特徴を持つのか」を明らかにすることを意味する。また、行動遷移パターンの抽出では、既に求められた場所の時空間特性に関する情報を基に、「20代の独身男性は休日昼まで自宅周辺で過ごした後に、休日の夜に繁華街へ向かう」といった、あるユーザ層に特徴的な移動のパターンの発見を試みる。

個人の行動履歴を基に、あるユーザにとって重要な意味を持つ地点や、個人の行動履歴パターンを抽出する、という問題は、実は時空間データマイニングの世界ではそれほど目新しい問題ではない<sup>1)-5)</sup>。しかし、従来研究で行われた実験の多くは、数十人から百人程度のモニタに対して大学構内や街中の範囲で観測するもので、集団としての行動特性の把握や頻出行動パターンを抽出する

ために、十分な規模の実験が実施されてはいなかった。本研究では、1300人以上のモニタから行動履歴情報を取得して、首都圏の規模で実験を展開しており、エリアの時空間的な特徴やユーザ属性ごとの行動特性について、数々の興味深い傾向を明らかにした。また、滞留が発生しやすい場所を見つけるためには、従来では空間的な距離のみを考慮したクラスタリングが使われていた。それに対して、我々の方法では性別や年代などユーザ属性の非類似度を加味することで、ユーザ属性ごとの滞留の違いを、より明示的に示すことが可能である。

### 2. 時空間特性抽出

#### (1) 用語の定義

はじめに、本論文で使用する用語の定義を行う。

**[定義1] 位置ログ:** 時刻 $t_i$ における経度 $x_i$ 、緯度 $y_i$ からなるデータ  $p_i = \{x_i, y_i, t_i\}$  を位置ログと呼ぶ。

**[定義2] 行動ログ:**  $p_i$  に時刻 $t_i$ における携帯端末使用時間  $Ph_i$  が紐づいたデータ  $b_i = \{p_i, Ph_i\}$  を行動ログと呼ぶ。

**[定義3] 行動情報:** ユーザ  $u_j (j=1, \dots, N)$  の行動ログの集合  $B^j = \{b_1, \dots, b_{T_j}\}$  を  $u_j$  の行動情報と呼ぶ。ただし、 $N$  は総ユーザ数を、 $T_j$  は  $u_j$  の行動ログの総数を示す。

**[定義4] ユーザ属性:**  $u_j$  の性別  $sex_j$ 、および年齢  $age_j$  からなる集合  $a_j = \{sex_j, age_j\}$  を  $u_j$  のユーザ属性と呼ぶ。

**[定義5] 滞留点:** 部分行動情報  $Sub_{ij} = \{b_i, \dots, b_j\} (i < j)$  が、 $\forall k \in [i, j], d(m_i^k, b_{k+1}) \leq D_{thr} \wedge |t_j - t_i| \geq T_{thr}$  を満たすとする。ただし、 $d(\cdot)$  はWGS84座標系における2地点間の距離を計算する関数を、 $m_i^k$  は部分行動情報  $Sub_{ik}$  の位置ログの重心点を、 $D_{thr}$  は距離に関する閾値を、 $T_{thr}$  は時間に関する閾値を示す。このとき、 $Sub_{ij}$  に対して、重心点経度 $s_x$ 、重心点緯度 $s_y$ 、滞留開始時刻 $t_a = t_i$ 、滞留終了時刻 $t_e = t_j$ 、滞留時間  $\tau = |t_j - t_i|$ 、および携帯端末総使用時間  $\pi = \sum_{i \leq k \leq j} Ph_k$  からなるデータ点  $s = \{s_x, s_y, t_a, t_e, \tau, \pi\}$  を滞留点と呼ぶ。図-1に部分行動情報から滞留点を検出する概念を示す。以下、時空間特性算出の処理手順を述べる。

#### (2) 前処理

##### a) 携帯端末の使用時間の算出

\* キーワーズ: 計画情報, 情報処理

\*\* 非会員、工修、株式会社三菱総合研究所

(東京都千代田区大手町二丁目3-6、

TEL:03-3277-0529、E-mail:keigo@mri.co.jp)

\*\*\* 非会員、工修、株式会社三菱総合研究所

† 非会員、工博、株式会社三菱総合研究所

†† 非会員、工博、株式会社三菱総合研究所 (現・独立行政法人 科学技術振興機構)

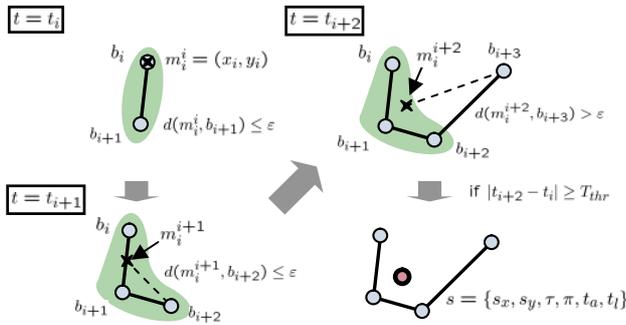


図-1 滞留点検出の概念図（網掛け部に含まれる行動ログが滞留点検出に使用されることを示す。）

携帯端末からは、端末の開、閉を示す情報が時刻と結びついた、開閉ログが取得される。開閉ログには位置情報が含まれないため、事前に最も近い時刻の位置情報と関連付けを行う。位置情報と関連付けられた開閉ログから、開と閉のログが連続している対を検出し、その時刻差の総和を携帯端末使用時間として行動ログに紐つける。

#### b) 誤差成分の除去

位置ログには誤差成分が含まれているため、滞留点を検出する前にカルマンフィルタ、および平滑化を用いて誤差成分の除去処理を施す。使用する状態遷移モデルには、平均移動速度  $v=d(b_i, b_{i+1})/|t_{i+1} - t_i|$  が閾値  $V_{thr}$  よりも小さいときは式(1)の滞留時モデルを、そうでないときは式(2)の移動時モデルを用いる。

$$\begin{bmatrix} X_{t+1} \\ V_{t+1} \end{bmatrix} = \begin{bmatrix} I & O \\ O & I \end{bmatrix} \begin{bmatrix} X_t \\ V_t \end{bmatrix} + W_t^x, \quad (1)$$

$$\begin{bmatrix} X_{t+1} \\ V_{t+1} \end{bmatrix} = \begin{bmatrix} I & \Delta t \cdot I \\ O & I \end{bmatrix} \begin{bmatrix} X_t \\ V_t \end{bmatrix} + W_t^x, \quad (2)$$

ただし、 $X_t = (x_t, y_t), V_t = (\dot{x}_t, \dot{y}_t)$  であり、 $W_t^x$  は時刻  $t$  における状態ノイズに対応する行列を示す。観測モデルは  $Y_t = X_t + W_t^y$  で与えられ、 $W_t^y$  は時刻  $t$  における観測ノイズに対応した行列を示す。

#### (3) 滞留点抽出と滞留時情報の付加

各ユーザの行動情報から滞留点を検出した後に、各滞留点に滞留時の時間帯、日種類、ユーザ属性の情報を付加する。滞留時間帯の判定は、時間閾値  $T_{thr}$  以上の滞留が行われた時間帯、また全ての時間帯での滞留が  $T_{thr}$  未満の場合には、最も長い時間滞留した時間帯とする。日種類は、滞留時間帯が存在する全ての日に関して対応づける。

#### (4) 滞留点のクラスタリング

時間帯、日種類の条件を指定し、条件に合致する滞留点のみを用いてクラスタの抽出を行う。クラスタリング手法には密度ベースの方法であるDBSCAN<sup>6)</sup>を使用する。DBSCANは外れ値に頑健であり、任意の形状のクラスタを抽出することができる特徴を有している。DBSCANでは、 $\epsilon$ とMinPtsという2つのパラメータを指定する必要がある。クラスタの検出は、あるオブジェクトの半径 $\epsilon$ の範囲内に存在する、そのオブジェクトの $\epsilon$ 近傍にMinPts以上のオブジェクトを内包するオブジェクトを、次々と凝集していくことにより実行される。

なお、本研究では、空間的な距離に加えて、性別・年代の属性の差異を考慮して滞留点間の距離を算出するようにした。すなわち、2つの滞留点 $s_i, s_j$ 間の距離が、次式により算出される。

$$D(s_i, s_j) = D_{loc}(s_i, s_j) + \gamma_s D_{sex}(s_i, s_j) + \gamma_a D_{age}(s_i, s_j), \quad (3)$$

ただし、 $D_{loc}$ はWGS84系における空間的な距離を示し、ヒュベニの式により求められる。 $D_{sex}$ は滞留者の性別の違いによる距離を示し、2点の滞留者の性別が異なれば1を、同じであれば0を返す。 $D_{age}$ は滞留者の年代の違いによる距離を示し、2点の滞留者の1年代異なるごとに距離が1ずつ増し、年代が同じ場合は0を返す。

$\gamma_s, \gamma_a$ はそれぞれ性別距離、年代距離の重み係数を示す。

#### (5) 時空間特性値の算出

抽出されたクラスタが示す特性を把握するために、各クラスタの中心点の緯度・経度座標と、クラスタが内包する滞留点数、それらの滞留点が示す平均滞留時間、平均携帯端末使用率、固有滞留者数、平均再滞留回数、滞留者性別比率、および滞留者平均年齢を出力する。なお、平均再滞留回数は固有滞留者ごとの滞留回数の平均を算出し、滞留者性別比率には滞留点数の比率を使用する。

### 3. 行動パターン分析

モニタの属性と行動ログ、および求められた時空間特性情報から、属性別に特徴的な行動パターンを分析する。

#### (1) 滞留点への時空間特性の紐付け

時空間特性抽出で求められたクラスタにおける固有滞留者数の中央値を算出し、中央値以上であればその地点を「人が多い場所」、中央値よりも少ない場合はその地点を「人が少ない場所」と分類した。その後、各ユーザの滞留点集合 $\{s_1, s_2, \dots, s_{|U|}\}$ に属する滞留点がどの地点に対応しているのかを紐づける。ただし、クラスタに所属しない滞留点に対しては、「人がいない場所」という

情報を紐付ける。

(2) 属性を考慮した頻出行動パターン抽出

時空間特性を伴う滞留点集合から、モニタの属性のクラスごとに頻出する行動パターンを分析する。使用するモニタの属性クラスは、表-1に示す項目からなる、実験開始前にアンケートで取得されたモニタの属性、および、行動ログから判定された行動タイプである。頻出行動パターンの抽出には IGVcaesp を使用する。IGVcaesp は高速シーケンシャルパターンマイニングアルゴリズムとして知られている Uno らの LCMseq<sup>7)</sup>を基にしており、イベントの順序関係とモニタの属性を組み合わせることで頻出パターンを検出することが可能である。

表-1 頻出行動パターン抽出に用いるクラス

属性項目	クラス値
性別	男性、女性
年齢	10代、20代、30代、40代、50代、60歳以上
職業	学生、主婦、会社員・自営業、自由業、他
婚姻状況	未婚、既婚
家族構成	単身、夫婦のみ、夫婦と子供、3世帯家族、他
世帯所得	200万円未満、200万円-400万円未満、400万円-600万円未満、600万円-1000万円未満、1000万円-1500万円未満、1500万円以上、不明
平日行動タイプ	平日直行直帰型、平日居住圏中心型、平日勤務地中心型、平日寄り道型
休日行動タイプ	休日自宅型、休日居住圏中心型、休日居住圏近隣型、休日遠方型

4. 実験

(1) 実験の設定

a) 実験期間

実験は2010年1月7日から1月31日の25日間実施した。

b) 行動ログ収集

実験に使用する行動ログは、株式会社エヌ・ティ・ティ・ドコモが法人向けiアプリDXにより実装した、データ収集アプリケーションを各モニタが携帯電話にインストールして、位置ログ、端末開閉ログを取得し、データ収集プラットフォームに送信および蓄積を行うことで収集した。位置ログの取得間隔は、モニタが10分、20分、30分のいずれかから選択可能であるようにした。

c) モニタ

インターネットアンケートを通じて、データ収集アプリケーションの動作検証済みの携帯端末を利用しており、勤務地が東京23区内、かつ居住地が首都圏または首都圏外の都市部であるモニタを募集した。募集モニタのうち、実験期間中に滞留を行ったモニタは1344名であった。表-2にモニタの性別、年代別の内訳を示す。

表-2 モニタの属性別内訳

モニタ	10代	20代	30代	40代	50代	60歳以上	合計
男性	14	121	323	295	93	8	854
女性	8	178	210	77	15	2	490
合計	22	299	533	372	108	10	1344

d) パラメータの設定

滞留点の検出では、距離閾値を $D_{thr}=300[m]$ 、時間閾値を $T_{thr}=30[分]$ 、移動速度の閾値 $V_{thr}=60[m/分]$ とした。また、位置ログと同時に、GPS座標の誤差レベルを示す3クラスの信号が取得され、その信号に対応して49[m]、299[m]、500[m]の3段階の観測ノイズを設定し、誤差成分の除去処理を実行した。

滞留点の検出に関しては、時間帯を{深夜(00:00:00から04:59:59)、朝(05:00:00から08:59:59)、日中(09:00:00から16:59:59)、夜間(17:00:00から23:59:59)、指定なし)の5区分、日種類を{平日、休祝日、指定なし}の3区分と設定した。

抽出されるクラスタ数はパラメータに依存する。例えば、表-3にパラメータを変えた際の、全ての滞留点から抽出されるクラスタ数を示す。 $\epsilon$ に関しては、値が小さ過ぎる場合は近傍の点を十分に探索できないため、また、値が大き過ぎる場合は過剰に点を凝集してしまうため、どちらも抽出されるクラスタ数がやや減少する。表-3からは $\epsilon$ を300mに設定した場合に最も多くのクラスタが抽出されており、一見すると適した値であるように考えられるが、実際には、図-2に示したように、 $\epsilon$ を200m以上に設定すると100mに設定した場合に比べて、郊外で新しいクラスタが抽出される一方で、滞留点が密集している都市部では少数のクラスタに多くの滞留点が凝集してしまう現象が起きてしまう。本実験では、都市部での時空間特性を中心的に扱いたいため、 $\epsilon=100[m]$ を採用した。MinPtsに関しては、その際に最も抽出クラスタ数が多くなる3に設定した。なお、モニタのプライバシー保護のために、クラスタ内の固有滞留者数が3名未満であるクラスタは無視した。

表-3 パラメータと抽出クラスタ数の関係

抽出クラスタ数	MinPts		
	$\epsilon [m]$	3	4
100	400	399	380
200	456	457	450
300	474	459	453
400	429	426	408

(2) 時空間特性抽出の結果

a) 滞留点数と滞留点クラスタの結果

全モニタの行動ログからは、83331点の滞留点が抽出

された。実験期間は25日間であるので、モニタ1名が1日あたり約2.5回の滞留を行ったことを示す。表-4、および表-5に滞留点の性別、年代別の割合を示す。これらの表からは、男性および年代が高いモニタが、滞留回数が多くなる傾向が読み取れる。また、表-6に特徴的な滞留特性を持つクラスタの一例を示す。観光地や郊外の大型商店内で携帯端末が使用されている様子や、大学周辺で滞留者の年齢層が低く、下町では滞留者の年齢層が高い様子、また、男女による滞留しやすい地点の違いなどが表-6から読み取ることができる。

表-4 滞留点の性別割合

性別	滞留数	比率 [%]	モニタ1人平均
男性	54371	65.2	63.7
女性	28960	34.8	59.1
合計	83331	100	62.5

表-5 滞留点の年代別割合

年代	滞留数	比率 [%]	モニタ1人平均
10代	1172	1.4	53.3
20代	17087	20.5	57.1
30代	34010	40.8	63.8
40代	23325	28.0	62.7
50代	7090	8.5	65.6
60歳以上	647	0.8	64.7
合計	83331	100	62.5

#### b) 時空間特性の分析結果

抽出されたクラスタごとの時空間特性から、街の時間帯による特徴の変化を確認した。例えば、図-3は平日の新宿近辺における、昼夜の滞留時間を示しており、日中はオフィス街である西新宿で長時間の滞留を確認できるが、夜間になると歓楽街である歌舞伎町での滞留時間が長くなっている様子が確認できる。

また、図-4に $\gamma_s=0.003$ (性別の違いが座標差10.8秒に相当する)とした場合の日中における抽出クラスタとそこでの男女別滞留時間を、図-5に $\gamma_a=0.003$ (1年代の違いが座標差10.8秒に相当する)とした場合の、平日・夜間における抽出クラスタとそこでの30代と40代の滞留時間を示す。前者からは、男性が霞が関などの官公庁街や虎ノ門などのオフィス街で滞留が長く、女性が青山通り沿いなどでの滞留が長いことが読み取れる。また、後者からは30代が夜間においても大手町や霞が関や溜池山王などオフ

イス街に長く滞留しているのに対して、40代ではそれらのエリアでの滞留自体が少ない。40代には既婚者が多いため、残業を控えて帰宅する傾向があることが、結果から予想できる。

#### (3) 行動パターン分析の結果

表-7に示した結果は、サポート値が0.25以上であり、100人以上の該当者がいて、かつ一度は滞留点クラスタ(人が多い場所、少ない場所)を訪れる、という3つの条件に適合する、属性を考慮した頻出行動パターンである。最も顕著なパターンは、20代に特徴的なもので、未婚で平日夜間に人がいない場所にいる人は、休日日中に人が多い場所へ向かうというものであった。すなわち、20代未婚者における平日夜と休日日中の切り替わりが、最も特徴的であると言える。これは所帯を持つ人間、あるいは30代や40代が休日日中に人が多いところへ向かう可能性が減ることの裏返しとも考えられる。このように、属性と紐付いた頻出行動パターンは、行動の背景を読みとくきっかけになると言える。

## 5. おわりに

本研究では、1300人以上のモニタから行動履歴を収集して、滞留状況から時空間を特徴付けると同時に、ユーザ属性と紐付いた頻出行動パターンを抽出する実験を行った。実験から明らかになった結果は、一般的な事実とみなすことはできないが、ある土地の時空間的な特徴やユーザの行動傾向を把握し、仮説を構築する際に有用な情報になりうると言える。

本研究が残した課題を述べる。まず、時々刻々と生成される行動情報を扱うために、クラスタリングのオンライン化<sup>9)</sup>が必要である。また、人口の疎密性に対処するために、クラスタリングのパラメータは適応的に決定される必要がある<sup>9)</sup>。特定の時間帯や実時間でのクラスタ検出を行うためには、時空間上でのクラスタリング<sup>10)</sup>を実施することが要求される。サービスとしての展開を考えるならば、抽出された時空間クラスタをPOIと結びつけるべきである。また、行動予測では、単純に何パーセントの人がどういう行動をとるといって定量評価だけでなく、なぜそのような行動をとるのかという意味付けを支援する方法を考える必要があるだろう。

## 謝辞

本研究は平成21年度 経済産業省情報大航海プロジェクトの一環として実施された。行動情報の収集には株式会社エヌ・ティ・ティ・ドコモ、実験モニタの募集には

エヌ・ティ・ティレゾナント株式会社にご協力を頂いた。  
 論文中の地図は昭文社から地図使用承認第 51G069 号を  
 頂いている。ここに謝意を表す。

参考文献

- 1) Zhou, C. et al. : Discovering Personal Gazetteers: An Interactive Clustering Approach, Proc. ACM GIS, pp. 266-273, 2004.
- 2) Kang, J. H. et al. : Extracting Places from Traces of Locations, Proc. Wireless Mobile Applications and Services on WLAN Hotspots, pp. 110-118, 2004.
- 3) Zhou, C., Shekhar, S. and Terveen, L. : Discovering Personal Paths from Sparse GPS Traces, Proc. Int'l Workshop on Data Mining, 2008.
- 4) Palma, A. T. et al. : A Clustering-based Approach for Discovering Interesting Places in Trajectories, Proc. SAC '08, pp. 863- 868, 2008.
- 5) Zhang, Y. et al. : Mining Interesting Locations and Travel Sequences from GPS Trajectories, Proc. WWW '09, pp. 791-800, 2009.
- 6) Ester, M. et al. : A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, Proc. Knowledge Discovery and Data Mining, pp. 226-231, 1996.
- 7) Uno, T. et al. : LCM ver.2: Efficient Mining Algorithms for Frequent/Closed/Maximal Itemsets, Proc. IEEE ICDM Workshop FIMI'04, vol.126, 2004.
- 8) Ester, M. et al. : Incremental Clustering for Mining in a Data Warehousing Environment, Proc. the 24th VLDB, pp. 323-333, 1998.
- 9) Ankerst, M. et al. : OPTICS: Ordering Points To Identify the Clustering Structure, Proc. ACM SIGMOD Int'l Conf. Management of Data, ACM Press, pp. 49-60, 1999.
- 10) Birant, D. and Kut, A. : ST-DBSCAN: An algorithm for clustering spatial-temporal data. Data and Knowledge Engineering, vol, 60, No. 1, pp. 208-221, 2007.

表-6 特徴的なクラスタ中心点の一例

滞留者平均年齢が低い地点					男性滞留者率[%]が高い地点				
順位	緯度	経度	年齢	地点	順位	緯度	経度	男性率	地点
1	35.6249	139.3409	17.1	東京工科大学八王子キャンパス	1	35.6998	139.9856	100.0	京成船橋駅前
2	35.6089	140.1201	23.0	千葉都市モノレール菟川公園駅	2	35.6559	139.7968	98.1	東京都江東区豊洲3丁目
3	35.7059	139.7179	23.8	早稲田大学戸山キャンパス	3	35.6996	139.7549	96.3	東京都千代田区神田駿河台
滞留者平均年齢が高い地点					女性滞留者率[%]が高い地点				
順位	緯度	経度	年齢	地点	順位	緯度	経度	女性率	地点
1	35.6443	139.6793	63.2	東京都世田谷区下馬世田谷公園	1	35.6854	139.7293	94.0	四ツ谷駅前
2	35.7134	139.7901	54.8	東京都台東区西浅草2丁目	2	35.6599	139.7555	92.3	東京都港区浜松町1丁目
3	35.7138	139.7938	53.9	東京都台東区浅草2丁目浅草寺	3	35.6256	139.7796	89.3	東京都江東区青海周辺
滞留中の携帯端末使用率[%]が長い地点					滞留者数が多い地点				
順位	緯度	経度	使用率	地点	順位	緯度	経度	滞留者数	地点
1	35.5231	139.5906	6.5	神奈川県横浜市都筑区(家具店)	1	35.6913	139.6999	362	新宿駅西口周辺
2	35.4088	139.5956	5.5	神奈川県横浜市港南区(モール)	2	35.6696	139.7606	322	東京都中央区銀座8丁目
3	35.7119	139.793	5.4	東京都台東区浅草1丁目雷門通り	3	35.7301	139.7138	193	池袋駅東口周辺

表-7 抽出された頻出行動パターン

サポート値	該当人数	モニタの属性	属性を考慮した頻出行動パターン
0.458	136	20代	未婚、平日夜間・人がいない場所 → 休日日中・人が多い場所
0.401	119	20代	未婚、休日日中・人が多い場所 → 休日夜間・人がいない場所
0.387	115	20代	未婚、休日日中・人が多い場所 → 平日夜間・人がいない場所
0.384	114	20代	未婚、休日日中・人が多い場所 → 平日日中・人がいない場所
0.374	111	20代	未婚、平日夜間・人がいない場所 → 休日日中・人が多い場所 → 平日夜間・人がいない場所
0.353	260	未婚	平日夜間・人がいない場所 → 休日夜間・人が多い場所
0.344	133	単身	休日夜間・人が多い場所 → 平日夜間・人がいない場所
0.328	127	単身	平日夜間・人がいない場所 → 平日夜間・人がいない場所 → 休日夜間・人が多い場所
0.323	125	単身	未婚、平日夜間・人がいない場所 → 平日夜間・人がいない場所 → 休日夜間・人が多い場所
0.319	235	未婚	休日日中・人がいない場所 → 休日夜間・人が多い場所
0.307	119	単身	休日日中・人が多い場所 → 休日夜間・人が多い場所
0.307	226	未婚	休日夜間・人が多い場所 → 休日夜間・人がいない場所
0.295	114	単身	会社員、休日夜間・人が多い場所 → 平日夜間・人がいない場所
0.294	101	平日居住圏中心型	会社員、平日朝・人がいない場所 → 休日日中・人が多い場所 → 休日夜間・人がいない場所



図-2 パラメータによるクラスタの違い(MinPts = 3、左:  $\epsilon = 100$  m、右:  $\epsilon = 200$  m)  
 青い円の半径の大きさはクラスタ内に所属する滞留点の個数に比例している。



図-3 新宿駅近辺の滞留時間(左: 平日日中、右: 平日夜間)  
 赤い円の半径の大きさはクラスタ内に所属する滞留点の滞留時間の長さに比例している。

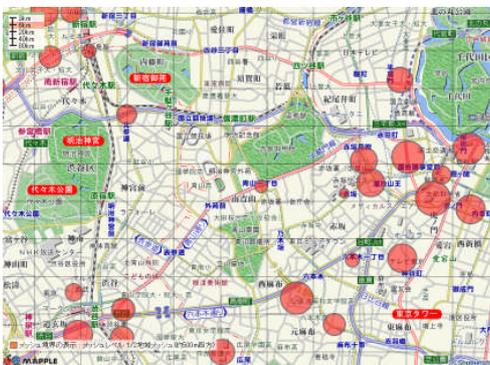


図-4 性別差のなす距離を考慮した時の、男性(左)と女性(右)の日中における滞留時間



図-5 年代差のなす距離を考慮した時の、30代(左)と40代(右)の平日夜間における滞留時間