

2012 年度 修 士 論 文

携帯電話の移動履歴を用いた生活パターンの分析

Analysis of Life Patterns using Long-Term Trajectory of
Mobile Phones

柴崎 真理子

Shibasaki, Mariko

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

目次

第1章 序論.....	3
1.1. 背景.....	3
1.2. 本研究の目的.....	4
1.3. 既往研究.....	4
1.4. 概念の定義.....	5
1.4.1. 活動量の定義.....	5
1.4.2. 位置情報測位点に対する属性ラベルの定義.....	6
1.4.3. 就業日と非就業日の定義.....	6
1.4.4. 繰返し性の定義.....	6
1.5. 使用する用語の定義.....	7
1.6. 本論文の構成.....	7
第2章 携帯電話の移動履歴データの特徴と比較データ.....	9
2.1. データの収集方法.....	9
2.2. 移動履歴収集対象者の属性におけるバイアス.....	12
2.3. 評価に用いる統計データ.....	19
第3章 移動履歴データの前処理.....	21
3.1. ステイポイントの抽出方法.....	21
3.2. ホームポイント・ワークポイントラベルの付与.....	21
3.3. 内挿データの作成.....	22
第4章 位置情報測定頻度を用いた一年間を通した生活パターン分類.....	24
4.1. 各時間帯の特徴.....	26
4.1.1. 各時間帯における相関関係.....	26
4.1.2. 各時間帯に対する主成分分析.....	29
4.2. クラスター分析による生活パターン分類.....	33
4.2.1. 平日の生活パターン分類.....	34
4.2.2. 休日の生活パターン分類.....	41
4.3. クラスター分析を用いた生活パターンのまとめ.....	46
4.4. 社会生活基本調査による生活パターンの整合性の評価.....	52
4.4.1. 社会生活基本調査による生活パターンの可視化.....	52
4.4.2. 位置情報測定頻度の生活パターンの整合性.....	54
4.5. 生活パターンごとの自宅・勤務地推定における問題点.....	55
4.5.1. 在宅時間の分布に着目した分析.....	56
4.5.2. 勤務地への滞在時間に着目した分析.....	59
4.6. 生活パターンに関する考察.....	62

第5章 活動圏に着目した年間生活パターン分類の特徴分析	64
5.1. 活動圏による生活パターン分析	64
5.1.1. 活動圏からみた平日の生活パターン	65
5.1.2. 活動圏からみた休日の生活パターン	69
5.2. 活動圏の空間的広さに関する考察	72
第6章 繰返し性に着目した個人の日々の生活パターン分析	74
6.1. 個人の生活における日パターン繰返し性の分析	74
6.1.1. 活動量を用いた生活の日パターンの抽出に関する予備分析	74
6.1.2. 活動量を用いた生活の日パターンの抽出	78
6.2. 滞在地点における個人の日々の生活パターンの繰返し性	82
6.3. 年間生活パターンごとの繰返し性の比較	86
6.3.1. 平日の生活パターンごとの活動量を用いた繰返し性の評価	86
6.3.2. 平日の生活パターンごとの滞在地点を用いた繰返し性の評価	87
6.4. 国民生活時間調査による生活パターンの整合性の評価	89
6.4.1. 国民生活時間調査による就業日と非就業日の可視化	89
6.4.2. 平日の生活パターンごとの滞在地点を用いた繰返し性の評価	90
6.4. 生活パターンの繰返し性に関するまとめ	91
第7章 結論	92
謝辞	94
発表文献	95
参考文献	96

第1章 序論

1.1. 背景

元来，“生活”に対する関心および需要は高い。人は様々な環境に巧みに適応して“生活”し、その“生活”はまさしく居住環境、帰属する社会環境から個人的な嗜好までを顕在化する。多くの組織や人々が、その“生活”を社会の管理もしくは改善、学究的な目的から調査し、その実態を把握する努力が行われてきた。

現在においても，“生活”への需要は、より様々な分野、組織や人から喚起されている。しかしながら、他者・もの・金・情報といった社会環境と相互行為から成り立つ複雑な人の“生活”を如何にして把握し、評価するのか。一つに“生活”の計量化があげられる。個人の移動の側面に着目したパーソントリップ調査、OD表、時間帯ごとの行動タイプに対する行動者率で一日の“生活”を把握しようとする社会生活基本調査や国民生活時間調査。個人のお金の動きから“生活”を計量化する家計消費状況調査など多くの調査が行われてきている。しかし、そのような大規模で信頼性の高い社会調査は実施するコストが高い。ましてや、個人に関して、1年間以上の長期間にわたり継続的に情報を蓄積するのは非常に手間のかかることである。

近年、GPS機能付き携帯電話端末などのモバイル端末が普及し人々の生活に必要なものとなっている。紙媒体のアンケート調査と違い、モバイル端末は個人の利用履歴や行動に関するデータをリアルタイムに送信し、かつ長期間にわたって蓄積することができ、大規模解析が比較的低コストで可能である。

この携帯電話端末から収集される移動履歴などの情報は、まさしく個人の日々の行動を表し、その利用は様々な分野で行われている。

まず、個人向けサービスとしては、日々の行動パターンから健康状態を推定し、ユーザの心情に密着する推薦システムサービスの『ココロ支援プロジェクト』¹。ネット上のアバターに現実空間でのユーザの行動を反映することで、コミュニケーションの円滑かつ作り出すツール『眺めてつながるツール』¹が開発されている。また、位置情報からクーポンを配布するマーケティング分野など民間での利用がある²。

さらに、都市計画³や交通計画立案のための基礎データなど行政機関が利用する可能性や、モバイル空間統計⁴に代表されるようにリアルタイムにある特定の地域における人口を推定することで災害対応支援などに対する社会的なアプリケーションの可能性はある。

このように、モバイル端末から個人の動きに関する情報を収集し解析することに対して、個人向けサービスの構築⁵および組織や社会の動きの予測や社会的なアプリケーションへの応用できるのではないかという期待が高まっている⁶。

携帯端末による大量の移動履歴データと、個人の属性と関連付けられた生活調査による基礎データから、対象者の生活の傾向やライフイベントによる生活パターンの変化などが

検出できる可能性がある。

しかし、今後の携帯電話端末によるデータからの個人の行動を解析するには、まず人の行動、つまり生活のパターンの仮説生成やモデル構築が必要である。人の“生活”のモデル化については、行動の実行回数と費やした平均時間や行動する人数の割合、行動に配分された時間量、時刻と行動との関連性を用いたモデル構築および分析が行われてきた⁷。

ただ、これらの調査やモデルは、人のリアルタイムな行動を記録する携帯電話端末によるデータほど、多様な人間活動を網羅している保証はない。

つまり、携帯電話から得られる膨大な位置情報から、どの程度人の生活パターンを把握できるのか、そもそもどのくらいの多様なパターンが存在するのか、詳細な解析やモデル構築の前に俯瞰に関する研究は少なく、研究することは非常に意義がある。

1.2. 本研究の目的

移動履歴から対象者の行動文脈解析および生活パターン解析は数多く行われている。しかしながら、実際の個々の携帯電話利用者が送った様々な“生活”には、どのような種類・パターンがあるのか、それらの“生活”が移動履歴および携帯電話で収集されるデータにどのように表れるか明らかではない。

社会調査によるデータ以外を用いた人の行動、ないしに生活のモデル化に関する研究は数多くある。ランダムウォークモデルの一つである **Levy-flight** モデルに基づいて人の行動の空間的広さに関するモデル化⁸や人間の行動パターンの予測可能性を示唆した研究⁹がある。しかし、実際の生活、日中働き夜間帰宅、夜間労働し日中在宅、勤務日がばらつく生活パターンなど個人の性別、家族構成、職業、勤務形態によって異なる生活パターンをどのように移動履歴などモバイル端末によって収集されるデータからモデル化するのかはまだ明らかではない。

そこで、本研究では、一年間という長期間にわたる携帯端末の移動履歴を用いて、対象者の生活パターンがどの程度分かるのか、そしてどのような生活パターンがあるのか探索的に分析し明らかにすることを目的とする。

1.3. 既往研究

位置情報と時刻が付与された点の集まりである移動履歴データから、人の行動を読み取るために研究は数多くされてきた。点の集まりである移動履歴データに人の行動に関するデータを付与し“**Semantic trajectory**”を作ることで、よりリッチなデータにして管理・分析するシステムに関する研究¹⁰¹¹¹²がある。しかし、そのようなシステムを用いたボトムアップ的な研究以外にも、大量に蓄積された携帯電話端末の移動履歴から直接人の行動や生活に関する情報を抽出しようという試みがある。ある地点での行動をその地点の属性（レ

ストラン、マーケットなど)で推定し、日々の行動を記述するシステムが提案されている¹³。基地局における移動履歴を用いて移動パターンおよびその空間的広さが居住地ごとに異なることを明らかにした研究¹⁴があるが、それは個人の移動パターンに着目していない。

個人の行動や生活を予測・推定するためには、自宅や勤務地など生活に深く関わる場所の特定が重要である。このような場所に基づいた生活・行動パターン解析の研究には、実際のデータを分析・学習し、構築したモデルに沿って重要な場所の抽出¹⁵¹⁶を行ったものがある。あるいは重要な場所に対する訪問回数に基づいたモデル¹⁷や重要な場所に関する位置情報と行動のタイプが記述された正解データから、より実際の人の行動に基づいた確率モデルを構築した研究¹⁸などがある。そして、抽出した重要な場所から生活パターンを解析する研究には、重要な場所への移動履歴から移動の頻出パターンを生活パターンとして考える研究¹⁹や滞在する時間帯別分布から生活パターンモデルを構築した研究²⁰がある。

このように場所に基づいて生活を記述し生活パターンモデルを構築する試みの一方で、多様な生活を移動履歴から予測できるのか、その限界について検討する研究がある。

基地局による移動履歴から、滞在地点の訪問回数の情報量を算出し、人の行動が予測できるか検討した研究では、行動パターンは 93%の確率で規則性があることを報告している²¹。

また、実際の携帯電話端末の移動履歴を用いて、ユーザの生活パターンの予測可能性を評価するエントロピーが提案されている²²。しかしながら、これらの研究は人の行動の繰返し性、予測可能性にのみ着目しており、多様な人の生活がそもそもどのようなパターンから構成されているのか、分類できるのかといった分析は行っていない。

さらに、人の移動パターンや生活パターンには居住地などの環境によって、特徴がありかつ生活の類似度に影響があることが報告されている²³。Nathan Eagle et al(2009)²³は、同じ大学に所属する学生を対象に、滞在場所のクラスタリングを用いて 100 人の学生同士のつながりを評価した。その結果、生活パターンを 6 パターンに類型化できると報告している。この研究は、対象者はすべて学生であるという属性のバイアスがあり、対象者数も 100 人と少ない。似通った属性を持つ人同士は生活パターンに何らかの類似性があり、類型化の可能性を指摘している。したがって、より大規模な人数で、属性のバリエーションが多い集団では、生活パターンがどのような多様性を持っているのか、そしてどの程度類型化できるか研究することは意義がある。

1.4. 概念の定義

本章では、本論文で用いられる重要な用語の概念について定義する。

1.4.1. 活動量の定義

本研究では、移動履歴の GPS 測位を行うシステムから位置情報の測定頻度は、ユーザの活動状況を表す指標であると考えている。位置情報の測定頻度は、そのユーザの活

動の有無だけではなく、測定頻度の増加と減少およびその持続時間から、ユーザの携帯電話を伴ったうえでの活動内容の変化を表している。位置情報測定の頻度が少ない状態から、急激に増加した場合、それは活動をしていない状態から移動する状態になったことを示唆している。位置情報測定頻度が少ない状態が続いている場合は、移動（この場合はトリップである）ではなく、何らかの滞在もしくは携帯電話をさわる等といった加速度が生じる活動を行っていると考えられる。

そこで、1時間当たりの位置情報測定頻度を活動量と定義する。活動量は、その1時間におけるユーザが移動を含む活動を行っているかどうかを表していると考えられる。位置情報の測定は最短で5分間隔で行われるため、活動量は最も多くて一時間で12回を上限としているが、まれに13回となるユーザがいる。

ただし、この活動量の数値および増減が、移動や活動の中身、つまり徒歩や鉄道による移動方法、およびテレビを見ているもしくはご飯を食べているという詳細な活動内容および激しい運動をしているといった活動の程度を示していない。

1.4.2. 位置情報測位点に対する属性ラベルの定義

位置情報測位点（以降、ポイントとする）に対して3つの属性ラベルを付与している。ステイポイント、ホームポイントおよびワークポイントの3つである。

ステイポイントとは、一定範囲内に一定時間以上滞在していると考えられるポイントに付与されるラベルである。ステイポイントとは、パーソントリップ調査にて用いられている用語「トリップ」間に存在するポイントという概念を有する。「トリップ」とは、出勤や買い物と行ったある目的に基づいて、一方向の移動を表す用語である。つまり、ステイポイントは一方向の「目的地に対する」移動の間に、あるエリアに滞在している行動を表すラベルである。ある地点に止まっていることを必ずしも意味しない。

ホームポイントとは、自宅とされる地点から一定範囲内にいるポイントである。ホームポイントは在宅を意味する。

ワークポイントとは、勤務地とされる地点から一定範囲内にいるポイントである。ワークポイントは勤務地への滞在を意味する。

1.4.3. 就業日と非就業日の定義

本研究では、4つの日付に関する概念を用いている。平日はカレンダー上の月曜日から金曜日をさす。休日はカレンダー上の土曜日・日曜日および祝祭日のことである。必ずしも平日は就業や学校がある日、休日は就業や学校がない日を意味しない。就業や学校がある日を就業日、就業や学校がない日を非就業日として定義する。

1.4.4. 繰返し性の定義

個人の一日一日に着目し、個人がどのような「一日」を繰り返しているのかを示す。つま

り、個人の一日一日の過ごし方が何通りあるのかを意味する。

1.5. 使用する用語の定義

以下に本研究で使用する用語の定義を以下の一覧表に記す。

表 1 - 1：用語の一覧

用語	意味
活動量	位置情報測定頻度のことを示す。ユーザの移動を含む身体運動などの活動を行っていることを表す。
ポイント	位置情報測位点のことを示す。
ステイポイント	一定時間以上に渡って、一定範囲内に滞在している位置情報測位点を示す。
ホームポイント	自宅に滞在していると考えられるポイントを示す。
ワークポイント	勤務地もしくは学校に滞在していると考えられる位置情報測位点を示す。
活動圏	ユーザが訪問および通過したエリアを示す。 本研究では、単位は標準地域メッシュである。
平日	月曜日から金曜日を示す。
休日	土曜日と日曜日および祝祭日を示す。
就業日	仕事および学校がある日を示す。
非就業日	仕事および学校がない日を示す。
繰返し性	個人の一日一日の日パターンが何通りあり、どれほど繰り返されているか示す。

1.6. 本論文の構成

本論文の構成について説明する。

第 1 章では、本研究の背景と目的を述べる。そして本論文の用語の解説を行っている。

第 2 章では、使用した移動履歴の収集対象者の属性にあるバイアスについて述べる。

第 3 章では、分析を行うために、元の移動履歴データに付与したラベルや前処理について述べる。

第 4 章では、位置情報測定頻度を用いて年間の生活パターンの類型化を行った

第 5 章では、一日の活動範囲の観点から、生活パターンに属するユーザの移動の特徴を分析した。

第 6 章では、個人の一日一日に着目し、個人がどのような「一日」を繰り返しているのか

を明らかにすることを試みる。

第7章では、分析から得られた知見と今後の課題について述べる。

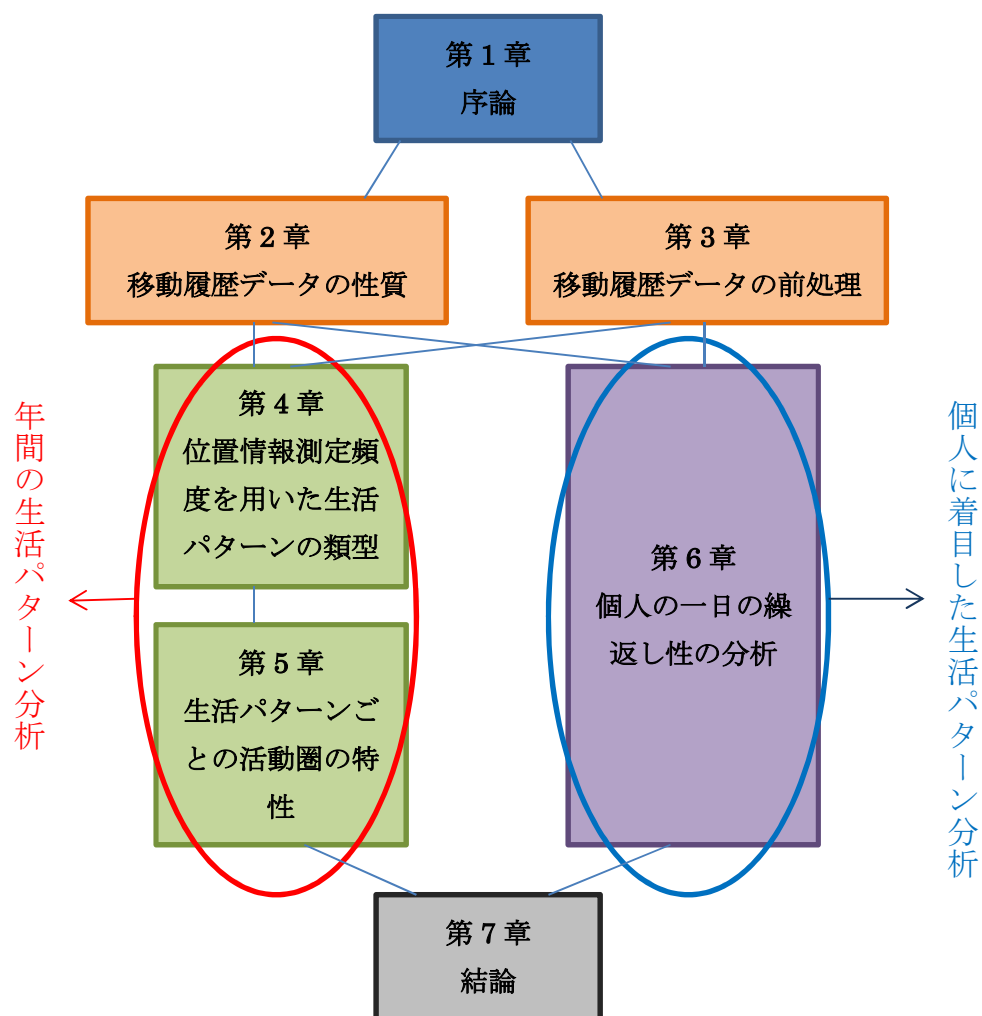


図 1-1：論文の構成

第 2 章 携帯電話の移動履歴データの特徴と比較データ

第 2 章では，本研究で使用する携帯電話の移動履歴データの特徴と，分析結果である生活パターンの整合性を比較・検討する比較データについて述べる．

2.1. データの収集方法

本研究で使用する移動履歴データは，（株）ゼンリンデータコムからご提供いただいた混雑統計®（URL：<http://www.zenrin-datacom.net/business/campaign/s01.html>）を使用した．混雑統計®は携帯端末に付属している GPS レシーバーから取得されている．移動履歴は，地図ナビゲーションアプリを携帯にダウンロードし，かつ位置情報の収集に同意をしていたいただいた一部のユーザのものである．この移動履歴を収集した位置情報測定システムについて以下の特徴がある．

表 2 - 1 : 位置情報測定システムの特徴

特徴：

1. 位置情報は最短 5 分間隔で取得される．
2. 携帯に付属している加速度センサーが加速度を検知すると、位置情報測定システムが自動的に現在位置情報の計測を始める．
3. ナビゲーションアプリが起動していなくても、携帯の電源が入っていれば、位置情報測定システムは位置情報の計測を継続的に行っている．

上記の位置情報測定システムの特徴から，位置情報が取得されているということは，携帯を身に着けたうえで，ユーザが自動車及び電車などによる移動を含むなんらかの加速度を発生させる“活動”を行っていると考えられる．現在位置情報がないときは，①ユーザは活動を行っていたが，携帯を身に着けていなかった．②ユーザは活動していたが，GPS 取得ができる環境にいなかった．③携帯の電源が切られているという 3 つのユーザの状態が考えられる．これを念頭に起き，位置情報から分析された結果をユーザの状態推定と結びつけなくてはならない．なお，本研究で使用した期間の移動履歴は，ガラパゴス携帯に付属している GPS レシーバーから取得されているのが大部分である．

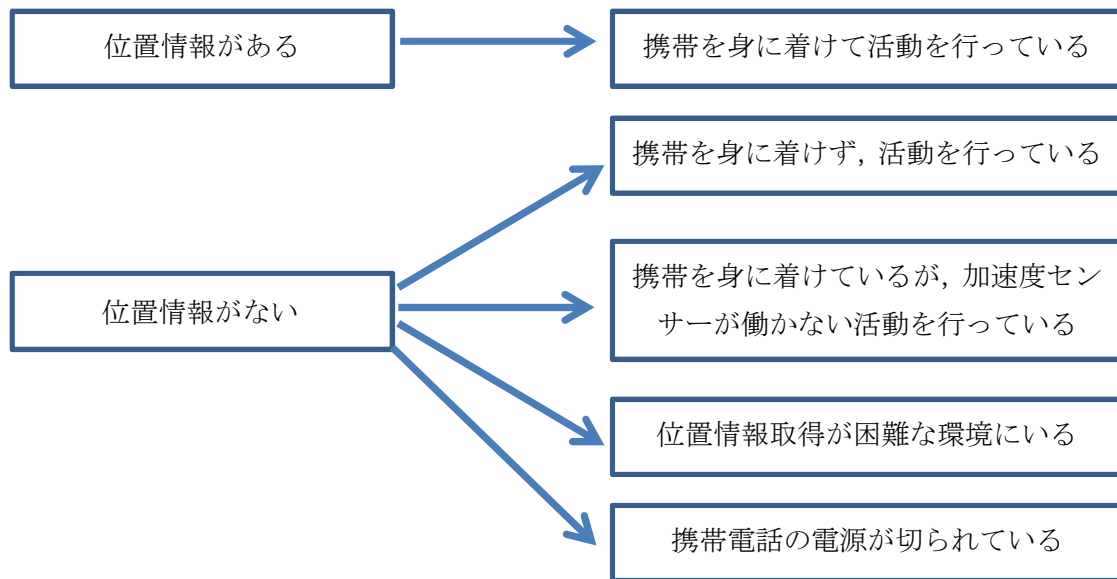


図 2-1：位置情報測定とユーザの活動の関係

混雑統計⑩には、日本全国にわたる約 100 万人のユーザの移動履歴が含まれている。各都道府県のユーザ数はそれぞれの都道府県人口の平均 1.19%にあたる人数と同じである。しかしながら、ユーザは、ナビゲーションアプリの利用者であることから、その抽出にはバイアスがかかっている。カードライバーを対象としたナビゲーションアプリをダウンロードおよび使用するのは自動車運転免許を持つ資格がある人間だと考えられるため、若年層および自動車運転をしない人間などはあまり含まれていないと考えられる。また、男女比率の偏りなどが考えられる。このバイアスについては、次節にてナビゲーションアプリ使用に関するアンケート調査結果に基づいて検討する。

混雑統計⑩には位置情報の他、日時、高度 (Altitude)、位置情報の誤差 (Estimation Error)、都道府県コード (prefecture code) と市町村コード (city code) が含まれる。位置情報の誤差は、①100m 以内、②100m 以上 200m 以内、③200m 以上 300m 以内の三つがある。

本研究では、混雑統計⑩から各都道府県の人口と、各都道府県内に存在すると考えられる移動履歴収集対象となったユーザの人数とその都道府県人口に対する比率を考慮し、都道府県ごとの最低抽出人数を 100 人として、各都道府県からランダムにユーザを抽出し、合計 5020 人のユーザの移動履歴を用いた。県人口が特に多い東京都、神奈川県、大阪府、福岡県は算出された抽出人口を四捨五入し、それぞれ 350 人、120 人、140 人、110 人抽出している。かつ移動履歴収集の期間は 2010 年 8 月 1 日から 2011 年 7 月 31 日である。移動履歴データに含まれる位置情報の誤差は図 2-2 に記した。

移動履歴に含まれる位置情報の誤差

■ 100m以下 ■ 100m以上200m以下 ■ 200m以上300m以下

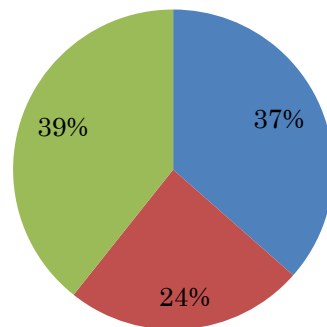


図 2-2 : 移動履歴に含まれる位置情報の誤差

次に位置情報の測定頻度から、どの程度活動が網羅されているか検討した。1年間にわたる移動履歴から生活パターンという日々の活動を抽出する際に、1年間のうちどれほどの日数において位置情報が測定されているか検討した。『ケータイ白書』（2011 年）²⁴においては、アプリケーションの使用回数は少なかったが、図 2-3 から 5020 人中 1627 人のユーザが 365 日間携帯を身につけて活動していることが分かる。5020 人の 53%にあたるユーザが 1 年間のうち 8 割以上の日数において携帯を身に着けて活動している。

したがって、おおよそ 1 年間の生活を網羅することは可能だと考えられる。

ユーザの位置情報測位日数の割合ごとの人数

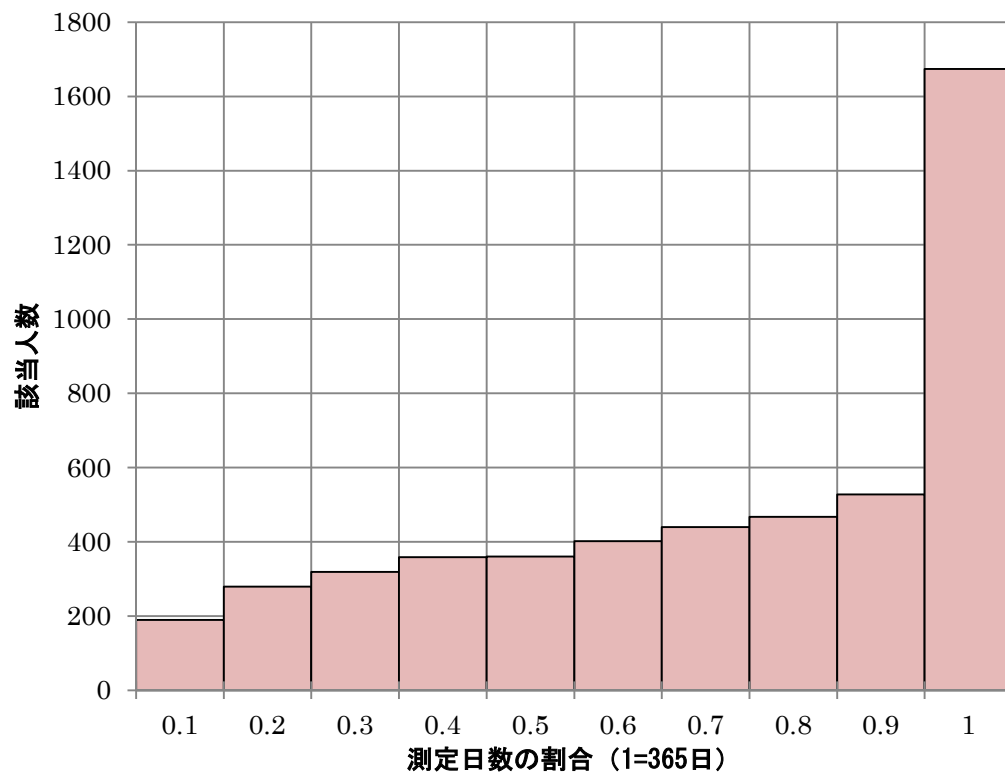


図 2-3：ユーザの位置情報測位日数の割合ごとの人数

2.2. 移動履歴収集対象者の属性におけるバイアス

本研究で使用した移動履歴データは、地図ナビゲーションアプリ（以降、地図ナビアプリとする）の利用者で、位置情報収集の許諾を得たなかの一部の利用者の移動履歴である。この移動履歴データには、データ収集対象者の性別および年齢、居住地、職業形態、勤務形態といった個人属性に関するデータは、その対象者の移動履歴データと関連付けられていない。また、その移動履歴データの全体の男女比率および職業構成比率や年齢構成比率に関しても公表されていない。

この移動履歴データは、地図ナビアプリ利用者の移動履歴データであるため、その移動履歴収集対象者（以降、ユーザとする）の属性構成にはバイアスが存在する。

そこで、どんな属性を持つユーザの移動履歴が収集されているか考察するため、アンケート調査を行った。このアンケート調査は日経の調査チームが行い、web アンケートという形式で収集されている。アンケートには、性別・年齢・住所および1年間に使用したアプリケーションを質問しており、1年間に地図ナビゲーションアプリを使用した回答者を地図ナビアプリ利用者として考えている。

このアンケート調査では、5万人の回答者を得ることができた。この5万人のうち、1年以内に、地図ナビゲーションアプリを使用していると回答したのは1356人であり、アンケート回答者の2.7%にあたる。しかし、このアンケート調査はweb アンケート形式であるため、その回答者自体にバイアスがある。混雑統計®の各都道府県における抽出率（＝ユーザ数／各都道府県人口）は平均1.19%であり、2.7%がより多い。このアンケートで得られた1356人の属性分布（性別、年代構成、職業構成）が、混雑統計®の属性分布に対する考察を与えるものとして、集計した。

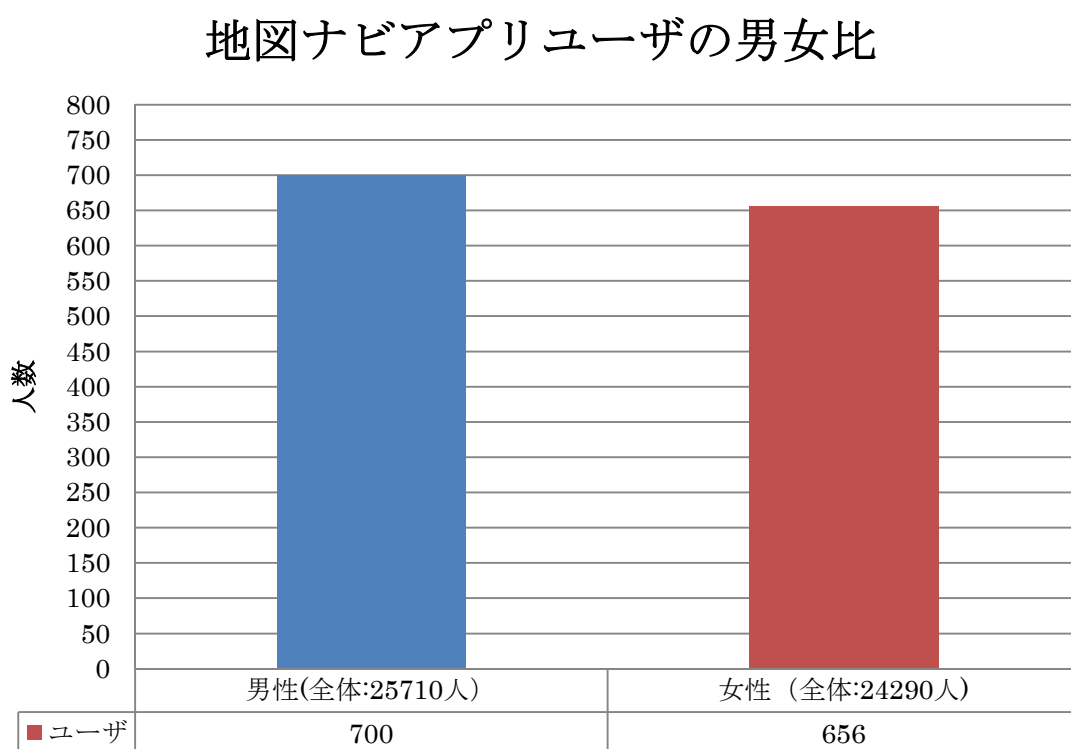


図 2 - 4 : 地図ナビアプリユーザの男女比

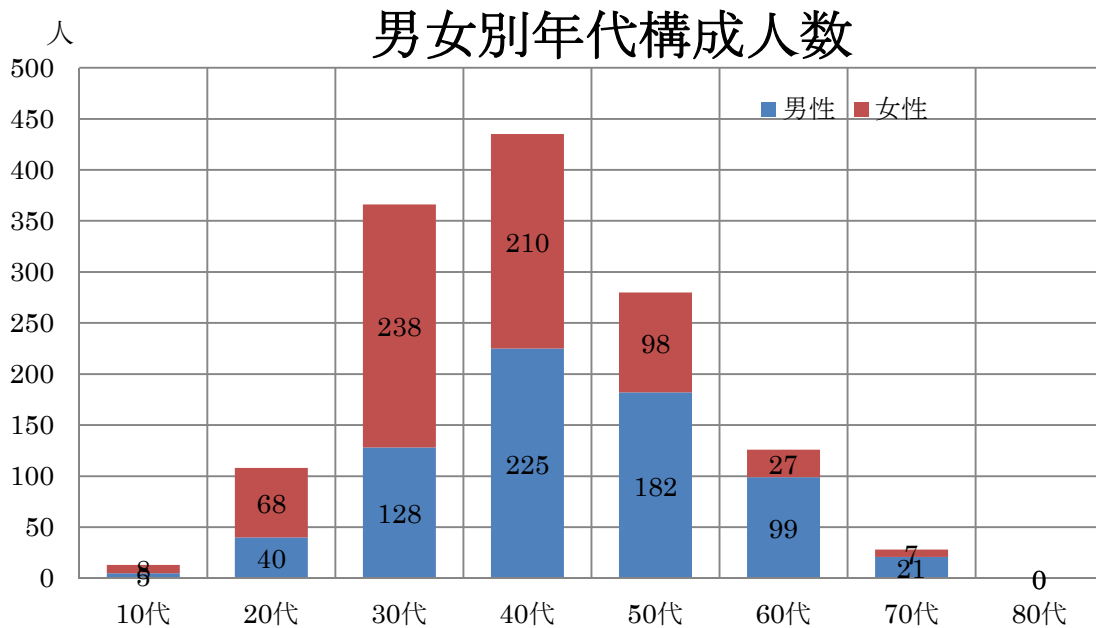


図 2 - 5 : 男女別年代構成人数

男女構成比および男女別年代構成比をグラフ化した。

地図ナビは、ナビゲーションアプリであるため、地図ナビアプリユーザの男女比は、男性が圧倒的に多いと思われたが、男女比がおおよそ 1 : 1 となっており、女性も多く含まれていることが分かる。

また、地図ナビアプリユーザの男女構成人数は、30代 40代が主に多く、高齢者層は含まれていないと考えられたが、実際には 40代が最も多く、50代、60代、70代の高齢者も少なくない。男女別の年代構成比では、高齢になるにしたがって男性の方が圧倒的に多くなる。反対に、20代の男女比は女性の方が多くなっている。

次に、男女別職業構成に着目してみると、男性では正社員が圧倒的に多いが、パートや商工自営・自由業および専業主婦、パート・アルバイトも少なくない。女性では、専業主婦・主夫がもっとも多いがパート・アルバイト、正社員など多様な雇用形態に属する女性が含まれていることが分かる。

男女別職業構成人数

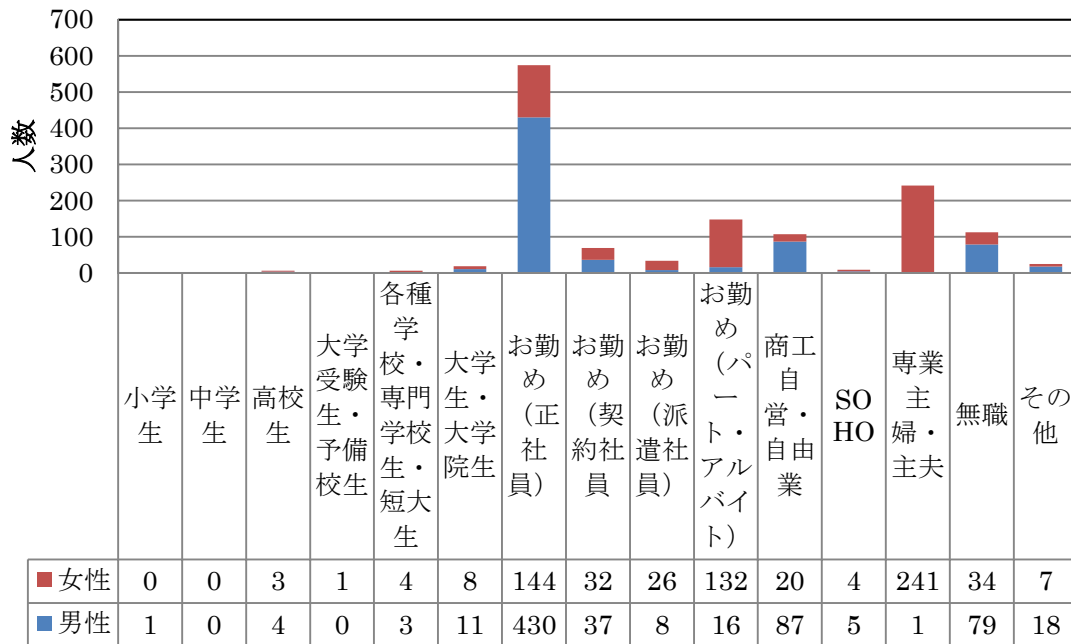


図 2-6：男女別職業構成人数

したがって、本研究で用いられている移動履歴データも、性別、年齢、職業によって極端な偏りがなく、ナビゲーションアプリユーザとバイアスがあるが特定の属性を持つ集団が欠落しているということではなく、様々な属性を持つユーザの移動履歴が含まれていると考えられる。

しかしながら、このデータが web アンケートによって収集されているため、回答者の集団自体も日本国民の属性構成から見て偏っていると考えられる。正確な地図ナビアプリユーザの男女構成比、年代構成比が推定することはできない。そこで構成比率に基づいた補正を行うことで、特定の属性を持つユーザのみが地図ナビアプリユーザを構成しているわけではないことを示す。

まず、日経アンケートによる男女ごとの年代構成人数比率と平成 22 年度国勢調査における年代構成人数比率に基づいて、補正指数を算出した。補正係数の定義式は以下に記す。なお、この補正は、日経アンケートの男女別年代別構成比率が平成 22 年度の国勢調査の男女別年代構成比率と等しくなるように調節している。この構成比率に基づいた補正を行ったうえで集計することを、母集団補正集計という。

Web アンケートという対象者に対するバイアスによる性別の構成比率を補正することで、なるだけ、実際の人口に近い男女別年代構成比率において、地図ナビアプリユーザが性別、年代別、職業別にどれほどの割合で含まれているのか検討するために行っているの、こ

の補正集計結果が実際の地図ナビアプリユーザの人数を表しているものではない。

補正指数定義式：

$$\text{補正指数} = \frac{\text{基準となる母集団における構成比率}}{\text{調査結果の集団における構成比率}}$$

これらの各年代の補正指数に基づいて、日経アンケートにおける地図ナビアプリユーザの属性に含まれるバイアス補正のために、日経アンケートの母集団に平成 22 年度国勢調査に基づいた補正指数を算出し、補正を行う。各年代における補正指数は以下のようになる。

表 2-2：各年代における補正指数

	男性	女性
10代	14.48	8.73
20代	2.15	0.94
30代	0.84	0.45
40代	0.44	0.44
50代	0.58	0.88
60代	1.31	3.03
70歳以上	4.41	20.63

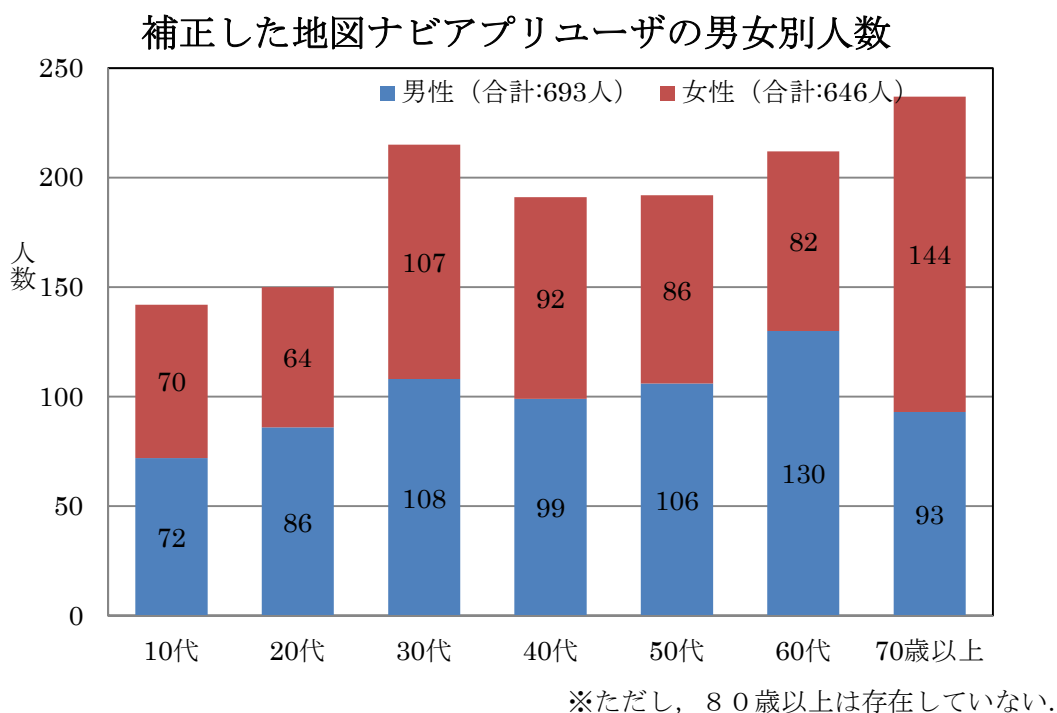


図 2-7：補正した地図ナビアプリユーザの男女別人数

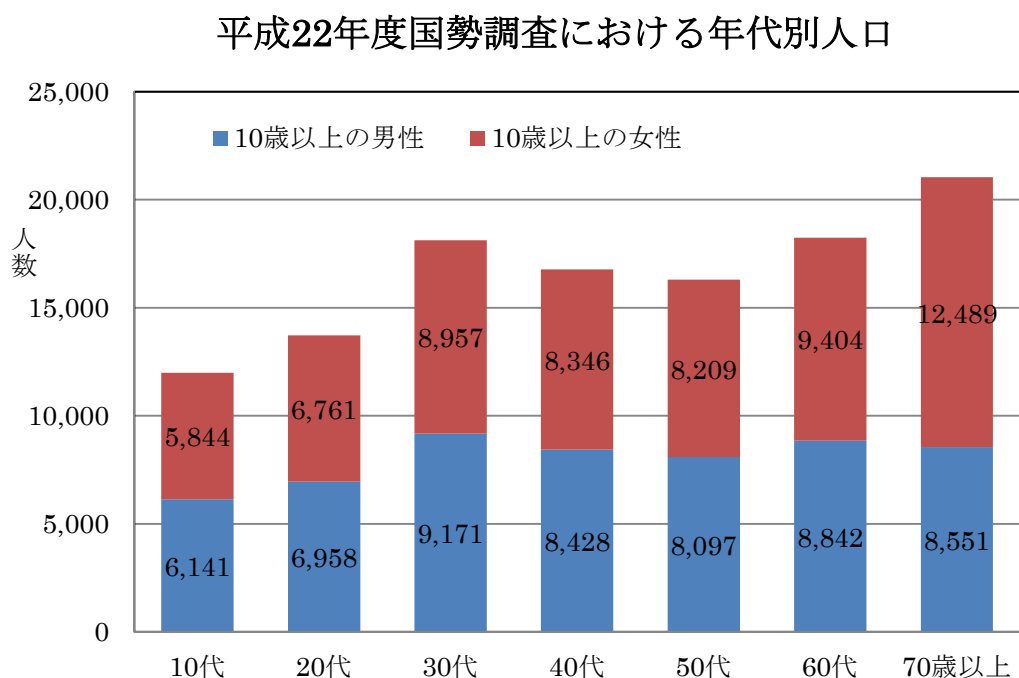


図 2-8：平成22年度国勢調査における年代別人口

(総務省統計局ページ 「平成22年国勢調査」 人口等基本集計結果の統計表より作成

URL：<http://www.e-stat.go.jp/SG1/estat/List.do?bid=000001034991&cycode=0>)

補正した地図ナビアプリユーザの男性の年代別職業構成

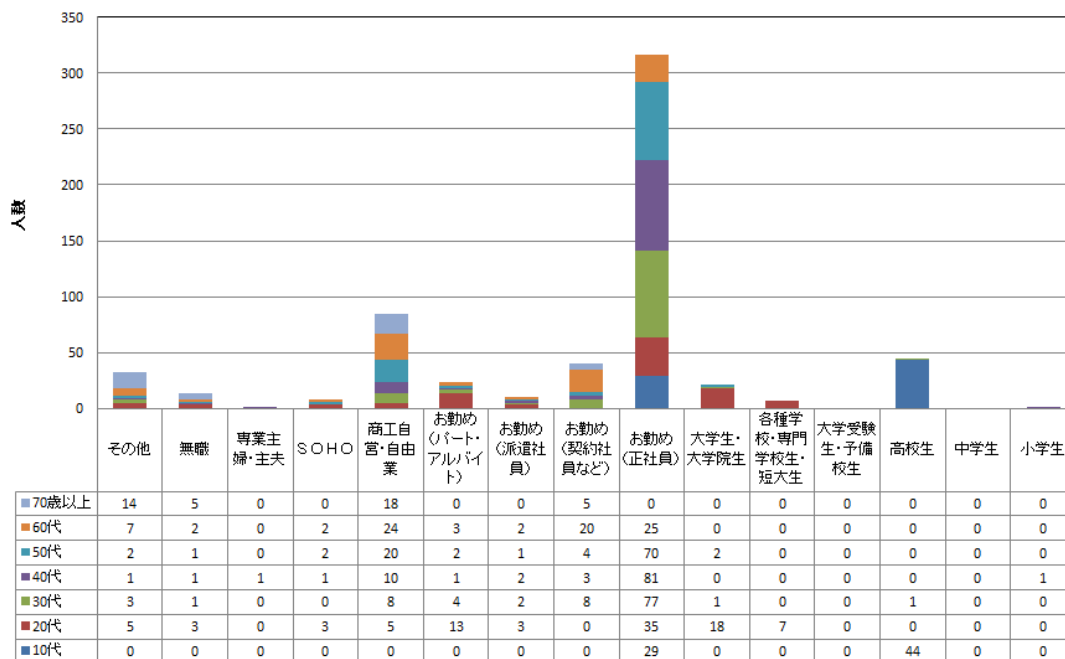


図 2-9：補正した地図ナビアプリユーザの男性の年代別職業構成

補正した地図ナビアプリユーザの女性の年代別職業構成

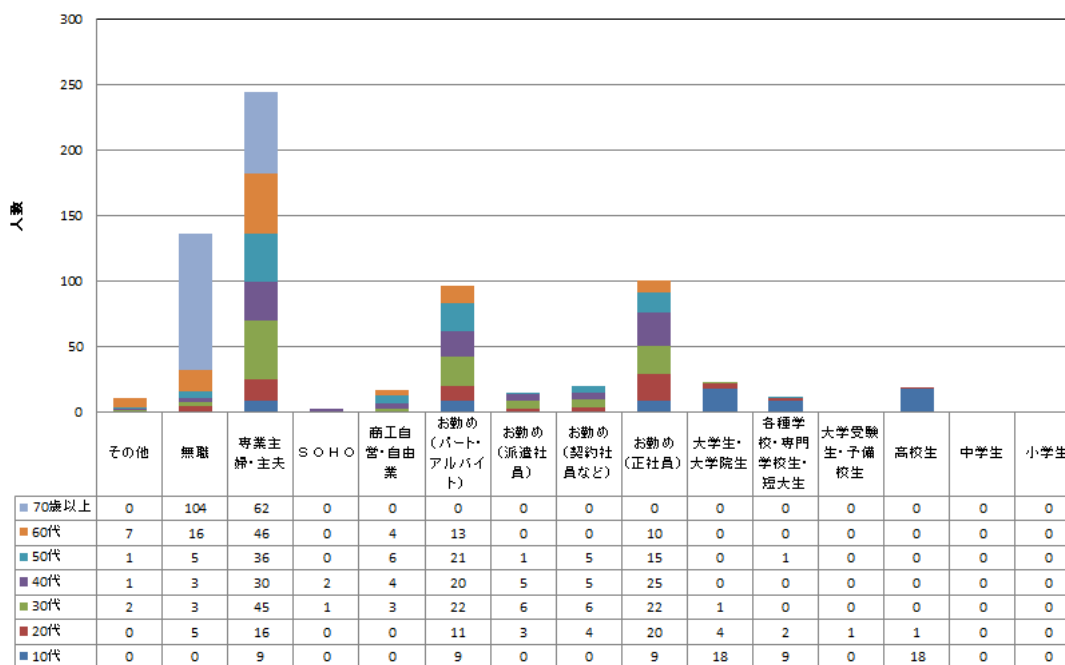


図 2-10：補正した地図ナビアプリユーザの女性の年代別職業構成

図 2-7 と図 2-8 から、補正した地図ナビアプリユーザの集計結果と国勢調査による男女の構成比率を比較した。国勢調査において、70 歳以上の年代を除く他の年代は、男性と女性の構成人数比率は、1:1 となっている。しかしながら、地図ナビアプリユーザの男女比率は、60 代から高齢になるにつれて男女比率は 1:1 から崩れ、その人数差は開いていく。

また、男女別に年代ごとの職業構成も補正した。図 2-9 と図 2-10 より、年齢及び性差が職業構成に影響を与えていることが分かる。専業主婦・主夫は女性が圧倒的に多く、70 歳以上の女性が無職の半分以上を占めている。男性では正社員が一番多い職業であるが、女性では専業主婦が最も多い。パート・アルバイトは男性よりも女性が多い。生活パターンの分類の際に、この性別および年代別の偏りを考慮する。

次に、生活行動および生活時間に関する統計調査について述べる。これらのデータは本研究で使用している移動履歴データと違い、性別・年齢・居住地などに関する属性と生活に関する情報をともに集めているため、分析結果の検証として用いられると考えられる。

2.3. 評価に用いる統計データ

本研究では、移動履歴データから推定された生活パターンの結果を検証および比較するために、二つの統計調査を用いた。NHK 放送文化研究所によって行われる NHK 生活時間調査と日本政府によって 5 年ごとに行われている社会生活基本調査を用いた。

NHK 生活時間調査は、人々の生活を 1 日の生活を時間面からとらえ、生活実態を把握し、放送サービスに役立てることを目的としている。この生活時間調査は、住民基本台帳から層化無作為 2 段抽出法にて抽出された 10 歳以上の国民を対象とし、調査票を用いて、連続する 2 日間の 0 時から 24 時間の時刻別(15 分刻み)の生活に関する行動を記入してもらい、生活行動と在宅状況をまとめている。本研究では、2010 年にて実施された生活時間調査の報告書である『2010 年国民生活時間調査報告書』²⁵を用いた。この調査において、一曜日でも有効な回答があった人を有効調査相手数（率）4905 人（68.1%）としている。

社会生活基本調査は、1 日の生活時間の配分と過去 1 年間における主な活動状況調査したもので、対象となった国民の性別、年齢、居住地、職業などの属性ごとに集計され、各属性の生活の傾向を明らかにしている。平成 23 年度にて実施されたものを用いた。『平成 23 年度社会生活基本調査』²⁶は 10 月 20 日に行われており、生活行動別平均時間、時間帯別の生活行動の状況や趣味や娯楽活動などに関する調査である。生活時間の配分についての調査は 10 月 15 日から 23 日までの 9 日間のうちから調査区ごとに指定された連続する 2 日間としている。

対象者は、層化 2 段抽出法によって抽出されている。層化 2 段抽出法とは、母集団をいくつかの部分母集団に分割し（層化）、各部分母集団から標本を抽出する層化抽出法を二段階行っている抽出法である。まず、第一次抽出単位は平成 17 年度国勢調査区(全国で約 6,900

調査区)を用いて、47 都道府県ごとに人口に基づく確率比例系統抽出により抽出する。第
二次抽出単位は等確率系統抽出により各調査区から 12 世帯を抽出している。世帯ごとに行
っている。ただし、岩手県と宮城県および福島県の一部は震災の影響で調査区から外れて
いる。このように選定された約 8 万 3 千世帯の 10 歳以上の世帯員約 20 万人を対象として
いる。

第3章 移動履歴データの前処理

本章では，生活パターンが推定するために，移動履歴データに対して行ったデータ処理について解説する．移動履歴における各位置情報測位点（以降，ポイントという）に対して，属性を付与した．また，位置情報の測定は不定期に行われるため，ポイント間に内挿を行った．

3.1. ステイポイントの抽出方法

ステイポイントとは，一定範囲内に滞留していると考えられるポイント（＝位置情報測位点）に付与される属性ラベルである．

ステイポイントは，次のポイントが 150m 以内にあり，かつその位置情報測定間隔が 20 分以上のポイントである．ステイポイント算出式を以下に記す．このステイポイント算出式は Apichon et al.(2010)²⁷から引用したものである．また，このステイポイント算出式は Montoliu et al.(2012)²⁸においても用いられている．Apichon et al.(2010)²⁷ で用いられているステイポイントを本研究でも使用している．

ステイポイント算出式

$$D_{threh} = p_{start} - p_{end} \text{ and} \\ T_{thresh} < \text{TimeDiff}(p_{start}, p_{end})$$

D_{threh} ：ステイポイントと考えられる範囲をカバーできる最大距離（ $D_{threh} = 150\text{m}$ ）

T_{thresh} ：滞留ポイントと考えられる範囲に滞留する最小の時間（ $T_{thresh} = 20$ 分間）

Apichon et al.(2010)²⁷では，算出したステイポイント群を DBSCAN アルゴリズムによるクラスタリングを用いて，ステイポイント群をまとめ，そのステイポイント群の重心点を算出している．DBSCAN アルゴリズムとは，データの密度を重みづけに用いて，データ密度も考慮したクラスタリング手法である．位置情報測位点に対するクラスタリング手法としては，ノイズがあるデータにおいても k-means 法よりも精度のよいクラスタリングをすることが知られている²⁹．Apichon et al.(2010)²⁷にて用いられているクラスタの重心点をステイポイント群の位置として本研究では使用している．

3.2. ホームポイント・ワークポイントラベルの付与

本研究では，Apichon. et al(2010)²⁷にて推定された自宅および勤務地を利用している．

ホームポイントは自宅から 500m 以内にあるポイントに付与される属性ラベルである。ワークポイントは勤務地もしくは学校から 500m 以内にあるポイントに付与される属性ラベルである。

ユーザの自宅および勤務地は朝に出勤して日中は就業し夜は帰宅して深夜を自宅で過ごすという一日の生活パターンという仮定に沿ってホームポイントラベルとワークポイントラベルは算出されている。Apichon et al.(2010)²⁷では、前日の 22 時から 6 時と 8 時から 18 時という二つの夜間時間帯と日中時間帯に分けて、この二つ時間帯における滞在時間を用いて夜間時間帯にいるステイポイントを自宅の候補地と仮定し、日中時間帯にいるステイポイントを勤務地の候補地と仮定して、自宅と勤務地を推定している。

自宅と勤務地の推定には、滞在時間のほかに、各ステイポイントのクラスタ内に含まれるステイポイントの数、夜間に存在しているステイポイント数の割合、滞在している日数、および日中に存在しているステイポイント数の割合、最大滞在時間と平均滞在時間と最少滞在時間の順番に変数の重要度を高く設定し、トレーニングデータを活動日記付きの移動履歴から作成し、自宅と勤務地を算出するアルゴリズムを機械学習によって学習させている。

このような日中は就業し深夜は自宅にいるという生活パターンのモデルは他の既往研究でも自宅と勤務先の推定に使われている¹⁵¹⁷²⁰。

このような生活パターンの前提に基づいた自宅と勤務先推定手法によって、使用している 5020 人分の移動履歴データのうち、3660 人のみが自宅と勤務先を算出されている。

3.3. 内挿データの作成

移動履歴データには、ユーザによって位置情報の蓄積の度合いが幅広い。例えば、ほぼ 1 年間において位置情報測定が行われ、1 時間当たり平均 4 回から 5 回の位置情報があるユーザがいれば、数ヶ月間で 1 時間当たり平均 1 回に満たないユーザがいる。また、位置情報がないということは、必ずしもユーザの活動がないとは断定できない（前章を参照）。携帯を身に着けた活動ではないため、位置情報測定がなくても、ユーザが何らかの活動を行っている可能性がある。そこで、移動履歴の補完を行った。

内挿アルゴリズムは以下のとおりである。

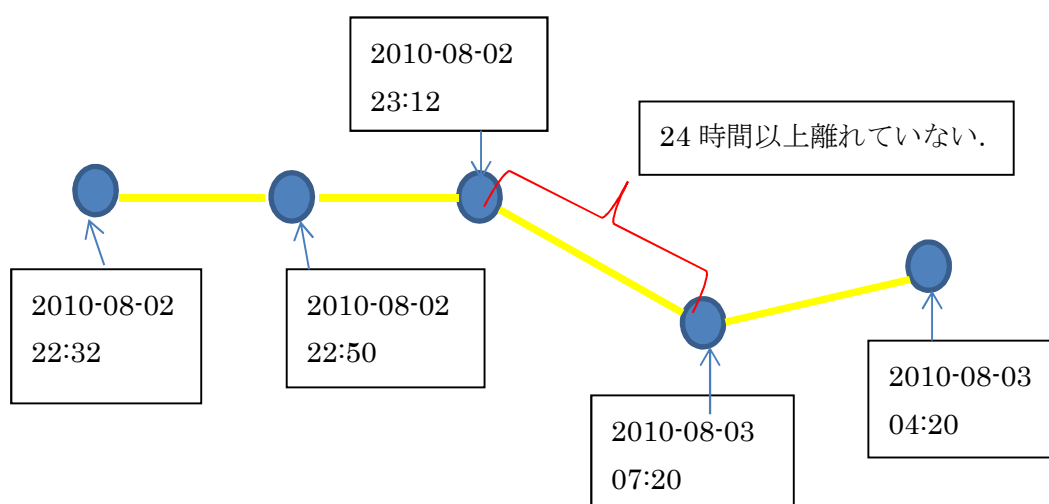
1 分ごとに、ポイント（緯度経度）を内挿する。内挿は座標系に変換せず、時間差分から算出した。内挿の重要性の一つには、連続した日数にまたがる移動履歴をシームレスにつなげることがある。それは、前章（移動履歴データについて）で述べたように人が起きて活動を行っていたとしても、携帯を身に着けていなければ、それは位置情報測定として、収集されない。よって、帰宅し家に滞在している間に、携帯に触れなければ、その人は位置情報測定が行われないため、移動履歴データ上では移動等を含む活動を行っていないことになる。そのため、日付をまたいで、移動履歴をシームレスにつなげることで、位置情

報測定がない夜間時などにおいて、位置情報を仮想的に付与することで、その時間帯の活動を拾い上げることにした。

しかしながら、位置情報測定は何日間も行われないことがある。その間にユーザは別の移動や活動を行っている可能性があるため、1日以上の日数が空けて位置情報測定が行われていない場合は、緯度経度の内挿を行っていない。

(1) 内挿するケース

位置情報が24時間以上離れていない場合



(2) 内挿しないケース

24 時間以上離れている場合

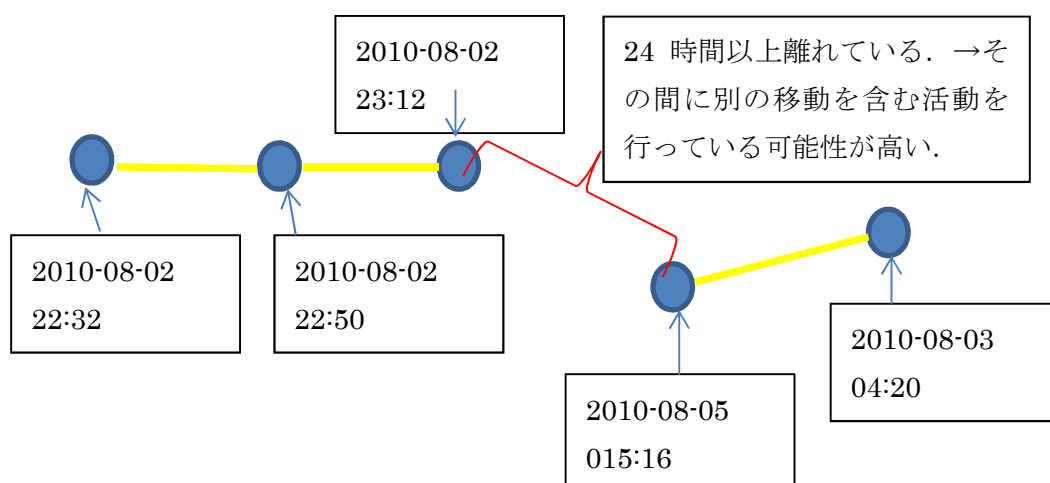


図 3-1 : 内挿の事例

第 4 章 位置情報測定頻度を用いた一年間を通した生活パターン分類

位置情報が得られているということは、ユーザが携帯電話を身に着けたうえで、移動を含む身体運動などの活動を行っていることを示している。つまり、一定時間当たりの位置情報測定頻度はユーザが携帯電話を身に着けたうえでの身体的な活動を行っていると考えられる。

本章では、位置情報測定頻度を活動量と定義し、ユーザの活動がどの時間帯に分布しているのかについて分析する。すなわち 1 年間にわたる移動履歴を用いて、各 1 時間あたりの位置情報測定頻度（＝活動量）を平日（月曜日から金曜日）と休日（土曜日から日曜日および祝祭日）ごとの長期的な傾向を分析・可視化した。

例えば、下図にあるユーザ A は平日の 6 時から 9 時、17 時から夜間にかけて活動量が見られるが、休日は午後の時間帯から継続的に活動量がある。この図から想像できるユーザ A の生活パターンは、平日は 6 時から活動し、7 時に出勤し（移動）、勤務地についてから、自分の携帯電話の電源を切っているか携帯電話をロッカーやカバンなどにいれているか、携帯電話をもって会社内をあまり動き回らない仕事についている。そして、午後 17 時以降に帰宅（移動）するため、携帯電話の位置情報測定システムが位置情報を測位しはじめ、帰宅し就寝するなど携帯電話に加速度が生じない活動を行うと位置情報がなくなる。休日は日中から何らかの活動を行っており、夕方の帰宅時間帯と考えられる位置情報の測定の増加が平日ほど規則正しくないため、帰宅する時間は、変動するといった生活パターンである。

一週間の各曜日の位置情報測定頻度

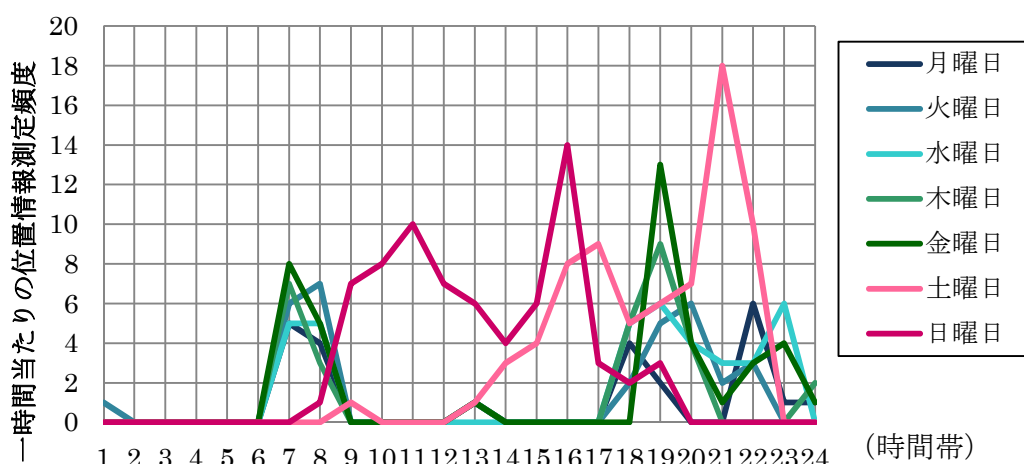


図 4-1：一週間の各曜日の位置情報測定頻度

平日の位置情報測定頻度

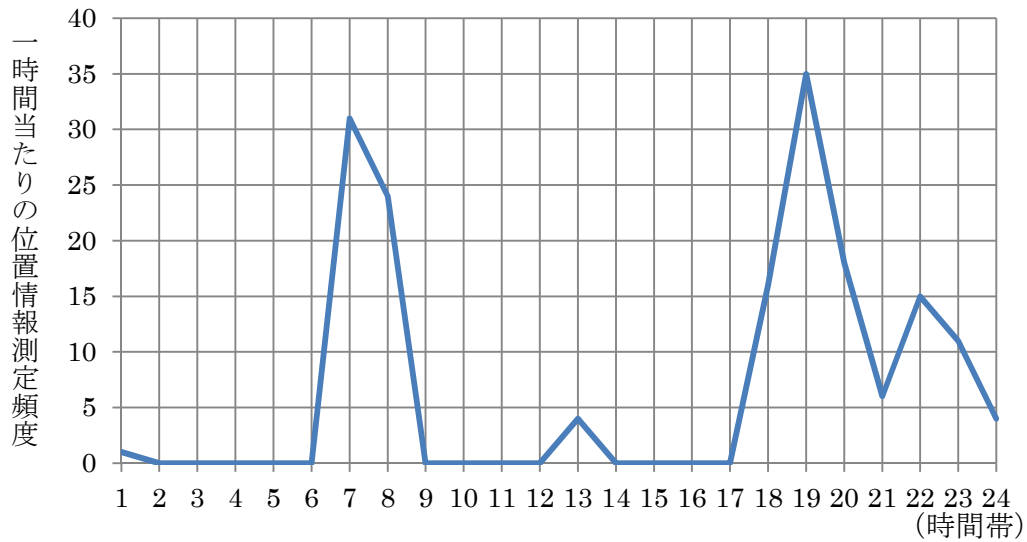


図 4-2：平日の位置情報測定頻度

休日の位置情報測定頻度

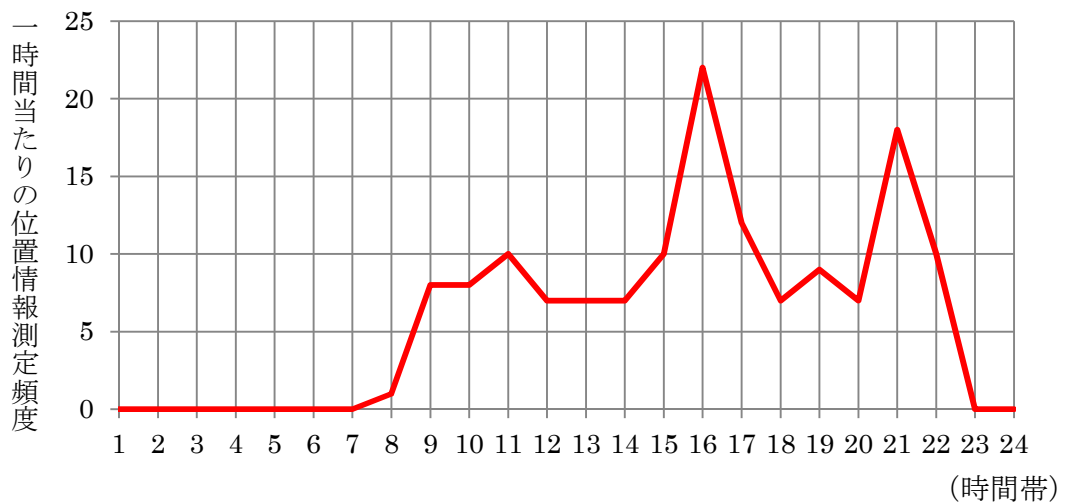


図 4-3：休日の位置情報測定頻度

この考え方をを使い、活動時間帯からみた生活パターンがどのような分類が行われるか検討した。

4.1. 各時間帯の特徴

まず、各ユーザの位置情報測定回数を平日と休日ごとに集計し、各時間帯の平均値を算出した。k は測定日数、i は時間帯を表す。

(平日)

$$i\text{時台の平均活動量} = \frac{1}{k} \sum_{n=1}^k (n_i)$$

(休日)

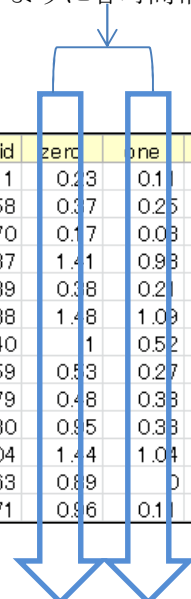
$$i\text{時台の平均活動量} = \frac{1}{k} \sum_{n=1}^k (n_i)$$

以上のように算出された各ユーザの平日・休日ごとの各時間帯（0 時台、1 時台など）の平均位置情報測定頻度を平均活動量とする。

各時間帯同士の相関係数の算出および主成分分析を行い、一日の時間の関連性や属性について検討した。

4.1.1. 各時間帯における相関関係

このように各時間帯同士の相関係数を算出する



user_id	zero	one	two	three	four	five	six	seven	eight	nine	ten	eleven	twelve
1	0.23	0.11	0.06	0.11	0.1	0.07	0.33	3.35	1.23	0.35	0.29	0.5	0.88
58	0.37	0.25	0.28	0.24	0.28	0.18	2.8	4.05	0.2	0.29	0.23	0.17	1.79
70	0.17	0.03	0.06	0.09	0.03	0.03	0.07	0.03	0.15	0.3	1.75	2.31	1.36
87	1.41	0.93	0.71	0.42	0.31	0.57	1.63	2.7	3.67	4.65	4.68	4.77	3.71
89	0.38	0.21	0.14	0.16	0.15	0.2	1.32	2.69	1.02	0.78	0.77	0.58	0.53
138	1.48	1.03	0.98	0.44	0.31	0.3	0.16	0.21	0.45	0.29	0.37	0.47	0.46
140	1	0.52	0.24	0.14	0.13	0.18	0.18	0.18	0.03	0.06	1.35	1.8	0.38
159	0.53	0.27	0.15	0.05	0.06	0.04	0.05	0.28	1.13	1.59	1.31	1.26	1.73
179	0.48	0.33	0.43	0.25	0.31	0.16	0.1	0.12	0.49	1.45	2.12	2.17	1.78
180	0.55	0.33	0.21	0.14	0.07	0.09	0.18	0.77	1.52	1.96	1.59	1.11	2.07
204	1.44	1.04	0.84	0.51	0.34	0.14	0.11	0.17	0.25	0.32	0.59	0.93	1.15
263	0.89	0	0.06	0	0	0	0	0.61	1.5	0.89	0.33	0.11	0.44
271	0.56	0.11	0	0.04	0	0	0.96	0.79	0.79	0.61	0.86	0.46	0.57

図 4-4：各時間帯の相関係数のとりかた

各ユーザにおいて、0 時台、1 時台というように各時間帯における平均活動量の値を用いて、各時間帯のピアソン相関係数を以上の図のように算出した。相関係数が強いほど、その時間帯における携帯電話を身に着けたうえでの活動（移動を含む）の連続性があると考えられる。

そこで相関係数の散布図行列を、正の相関係数が強ければ青く、負の相関係数が強ければ赤くなるように可視化し、活動の連続性が維持されている時間帯の可視化を行った。

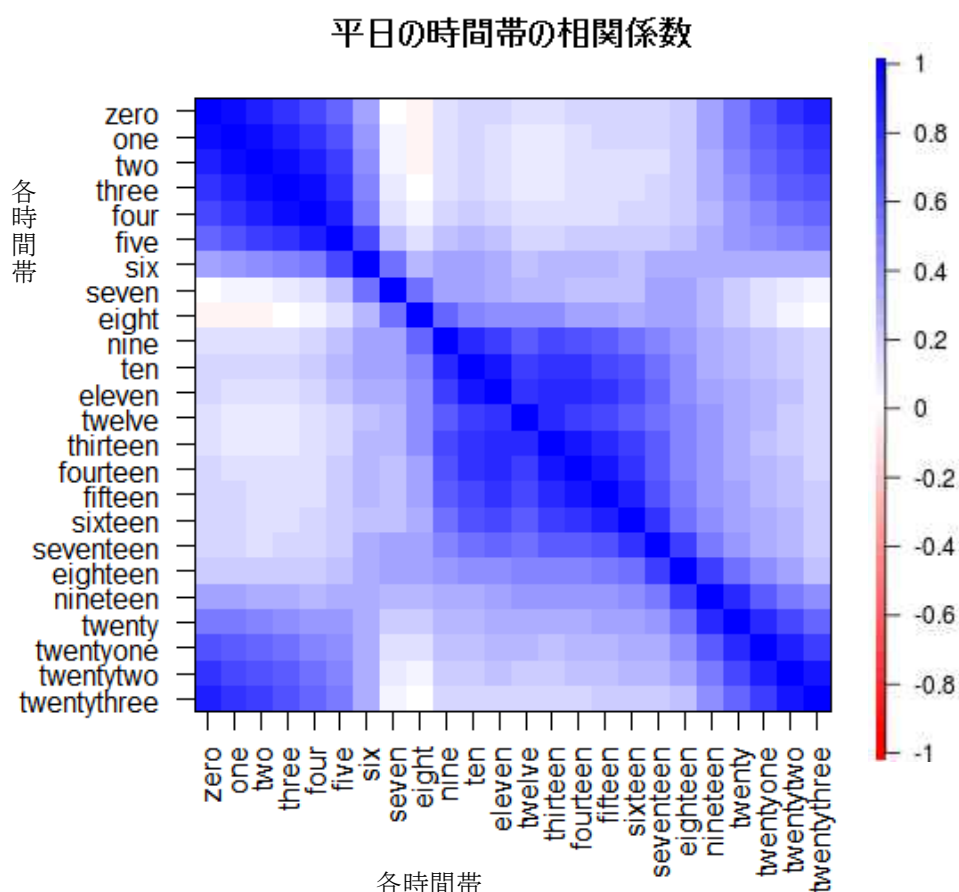


図 4-5：平日の相関係数の散布図行列の可視化

図 4-5 から、0 時から 5 時と 20 時もしくは 21 時から 23 時が夜間の時間帯に分けられる。これらの時間帯は就寝、ないしは休養と考えられ、同じ活動が継続している。

0 時間帯との相関係数から、日中は、7 時から 8 時と 9 時から 17 時までという二つの時間帯ブロックに分けられる。『平成 23 年度社会生活基本調査』(2012)²⁶の報道資料「生活時間に関する結果要約」によると、7 時から 8 時は多くの人（特に正規雇用者にとって）にとって通勤時間帯にあたっている。そのため、活動の継続時間は 1 時間から 2 時間程度であり、それまでの夜間の就寝・休養、9 時以降の就業・修学などとは異なる活動パターンとな

っている。これは、7時から8時の活動が他の時間帯の活動と相関が低いことから推察できる。したがって、位置情報測定頻度から、移動と移動以外の活動の区別がつくことが示唆される。

その後9時から17時くらいまでのブロックは就業・修学などの活動時間にあたっているためか比較的同じような活動パターンが連続している。12時から13時の時間帯には、昼休みによる活動パターンの変化がいくらか見られる。

18時から19時もしくは20時の時間帯は、帰宅時間にあたっていると考えられるが、午前7時から8時帯の通勤・通学時間に比べ、活動の変化は緩やかである。これは帰宅時刻が人により分散していると想像できる。一方、朝の通勤・通学時間帯は、活動の変化はより鮮明であり、これは通勤・通学の時間帯が集中していることを示している。これは、『平成23年度社会生活基本調査』(2012)²⁶報道資料にて報告されている帰宅時間の分散性を立証していると考えられる。

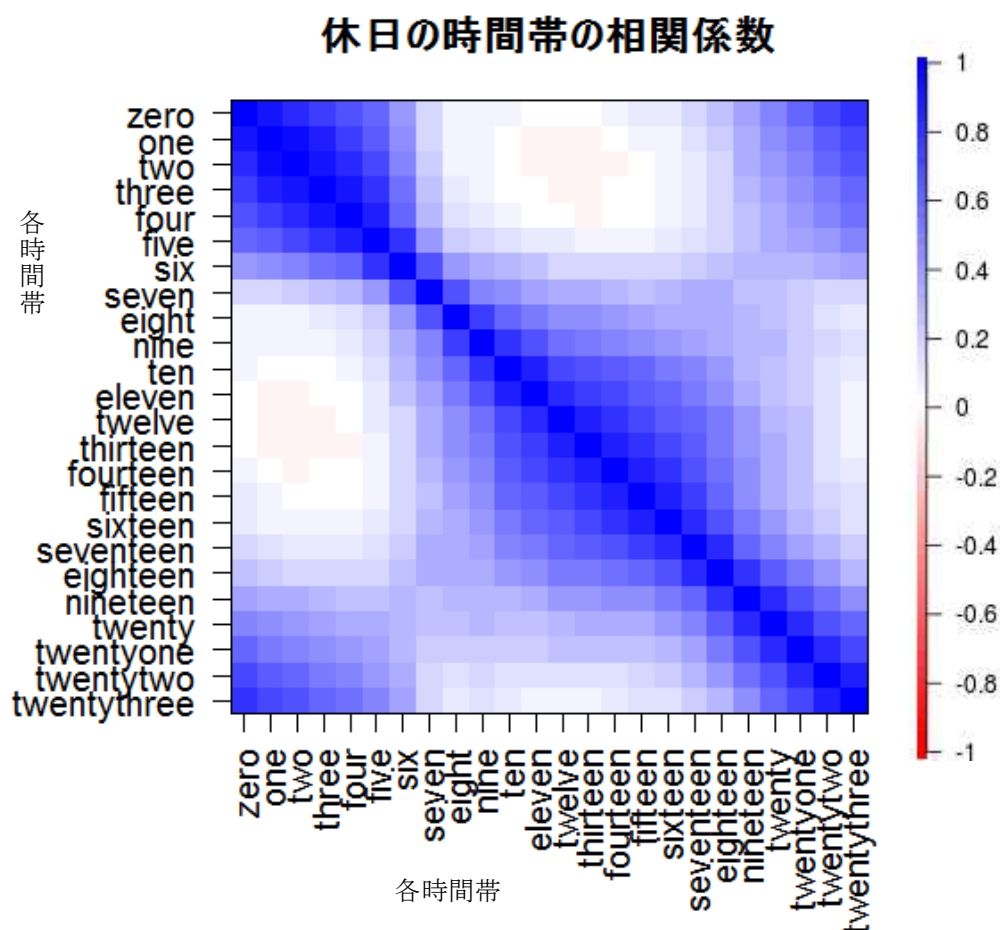


図 4-6 : 休日の相関係数の散布図行列の可視化

また図 4-6 から、休日は平日と同様に、0 時から 5 時と 20 時もしくは 21 時から 23 時の夜間の時間帯ブロックと、7 時から 18 時の日中の時間帯ブロックの二つの時間帯ブロックに分けられる。

日中の時間帯ブロックにおける相関係数をみると、休日の相関係数は平日に比べて、緩やかに変化しており、明確な活動の区切れ目が存在しない。したがって、9 時から 18 時までの日中の活動の開始時刻、終了時刻が分散していることが分かる。

4.1.2. 各時間帯に対する主成分分析

これまでの分析から、朝の通勤・通学の時間帯や、夕方から夜にかけての帰宅時間帯が活動の「区切れ目」になり、人の行動パターンを特徴付ける可能性があることがわかった。そこで、次に、平日と休日ごとに主成分分析を行い、平日・休日の活動パターンを特徴付ける時間帯はどこに存在するのかをより詳細に分析する。時間帯はこれまでと同様、1 時間ごとに 24 個という時間帯（変数として扱っている）を設定した。合成された主成分の固有値と累積寄与率を図 4-7 から図 4-11 に記す。

平日における各主成分の固有値

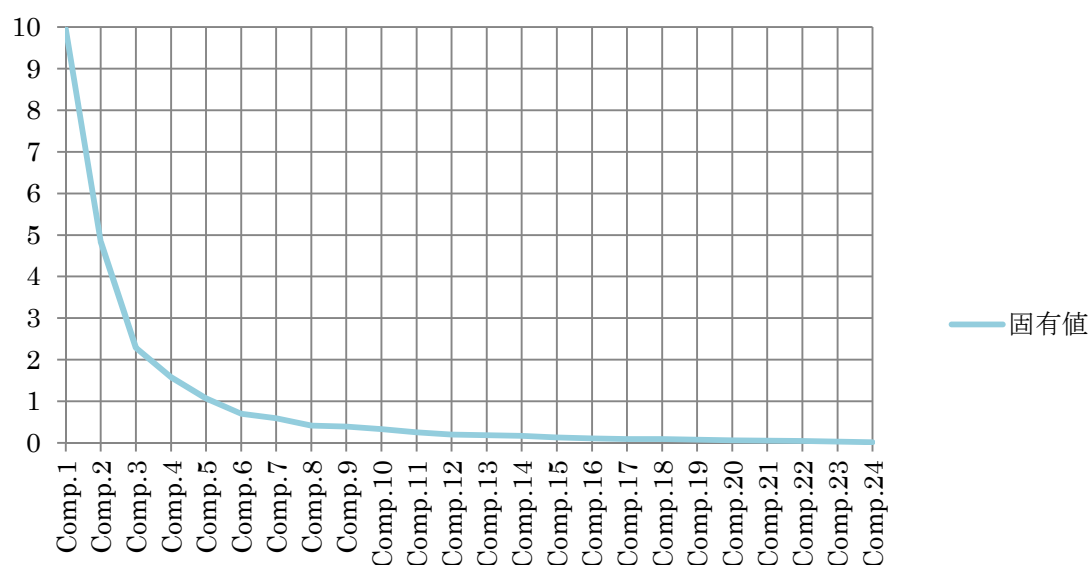


図 4-7：平日における各主成分の固有値

休日における各主成分の固有値

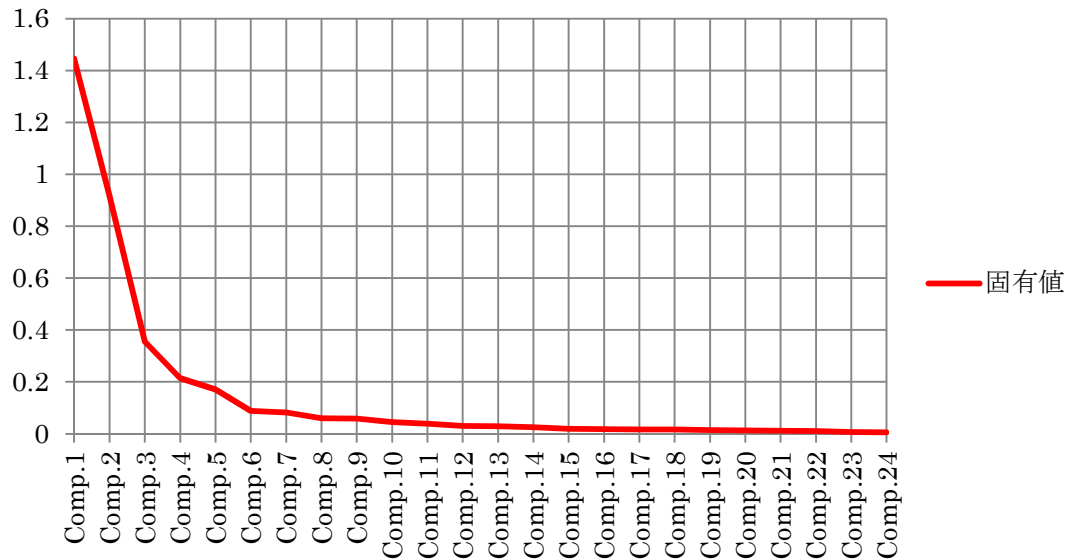


図 4-8：休日における各主成分の固有値

平日における各主成分の寄与率と累積寄与率

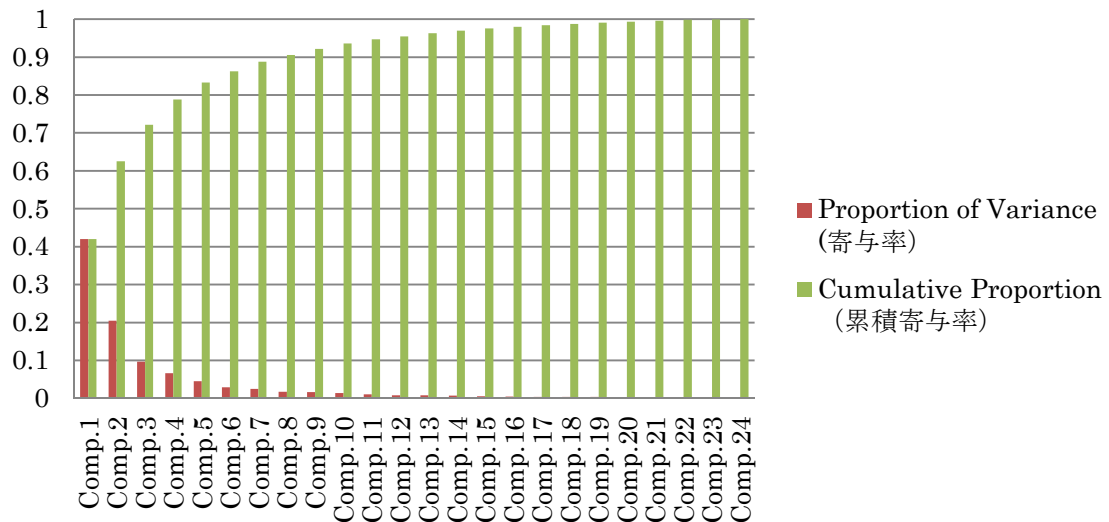


図 4-9：平日における各主成分の寄与率と累積寄与率

休日における各主成分の寄与率と累積寄与率

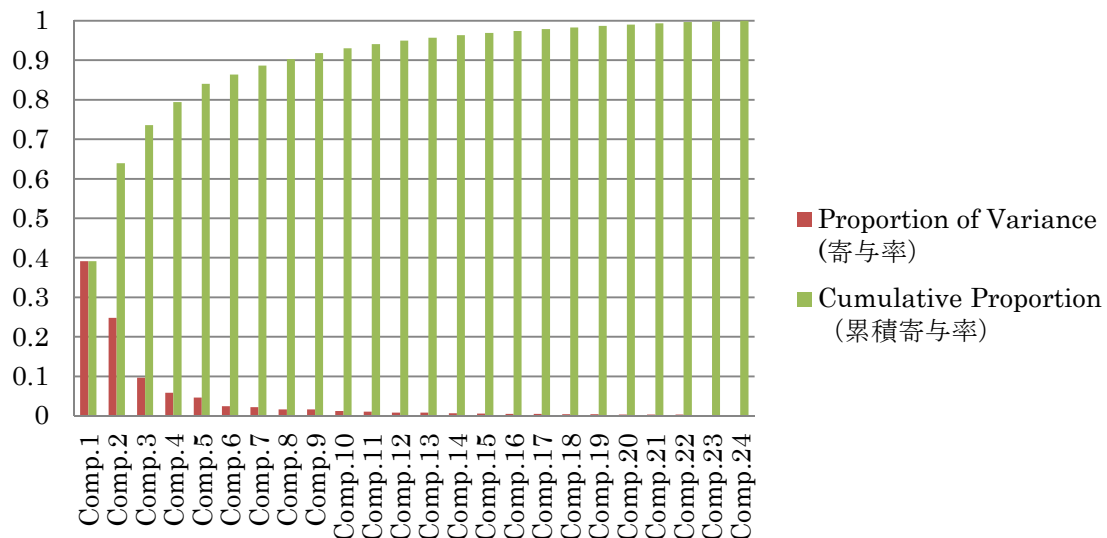


図 4-10：休日における各主成分の寄与率と累積寄与率

平日および休日の主成分分析において、ともに各主成分の固有値が 1 以上であるのは第 5 主成分までであり、累積寄与率が 0.8 を超えるのは第 1 主成分から第 5 主成分までであることから、第 1 主成分から第 5 主成分を採用する。第 1 主成分から第 5 主成分における変数の寄与率を図 4-11、図 4-12 にまとめた。

平日の各主成分における変数の寄与率

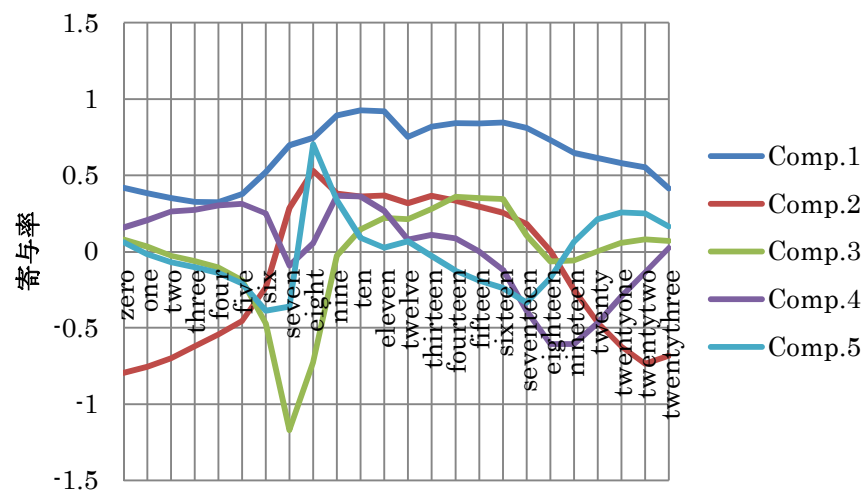


図 4-11：平日の各主成分における変数の寄与率

休日の各主成分における変数の寄与率

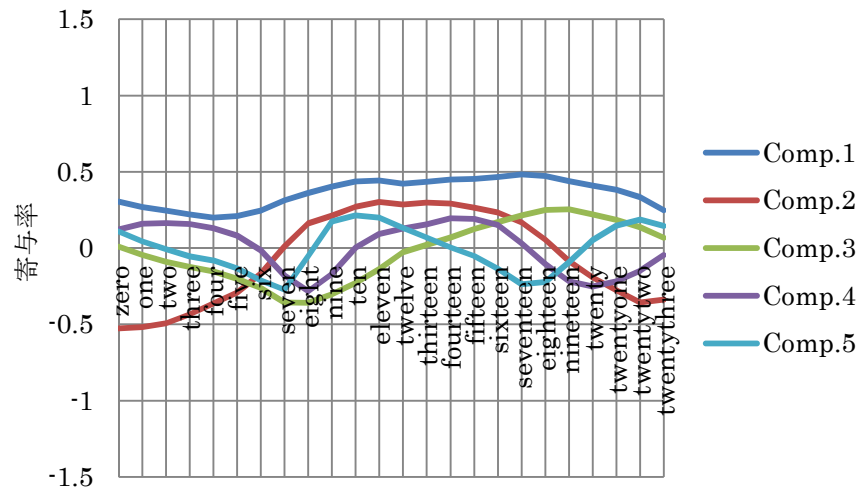


図 4-12：休日の各主成分における変数の寄与率

平日の第一主成分において 0 時から 6 時が最も寄与率が低く、9 時から 18 時台までが寄与率が比較的高い。反対に休日の第一主成分は平日ほど寄与率に差が見られない。この第一主成分は日中の時間帯と夜間の時間帯を分ける軸であることが分かる。

平日および休日の第二主成分は 0 時から 5 時と 19 時から 23 時が負の寄与率であり、7 時から 17 時は正の寄与率である。したがって第二主成分は日中の時間帯と夜間の時間帯を明確に分けるものだと考えられる。

平日の第三主成分は、7 時台を中心に突出した負の寄与率となっており、10 時から 17 時までが正の寄与率となっている。したがって 5 時から 8 時までの活動と 10 時から 17 時までの活動を区別する軸であると考えられる。休日の第三主成分は 8 時を中心に午前中が負の寄与率を示し、14 時から 23 時までが正の寄与率を示し、19 時台がもっとも寄与率が高い。したがって休日の第三主成分は午前中の活動と夕方の活動を区別する軸であることが分かる。

平日の第四主成分は、7 時から 8 時の寄与率が低く、0 時から 5 時と 9 時から 11 時が高い正の寄与率をとる。反対に 16 時から 22 時は負の寄与率をとり、18 時と 19 時は最も低い寄与率となる。この平日の第四主成分は 7 時と 8 時の時間帯の活動以外の 0 時から 5 時と 9 時から 11 時、つまり深夜および早朝と午前中の時間帯と 16 時から 22 時の夜間の時間帯を区別する軸であると考えられる。休日の第四主成分においては、0 時から 5 時までと 11 時から 16 時までの時間帯が正の寄与率をとるが、7 時から 9 時と 18 時から 23 時は負の寄与率をとる。

平日の第五主成分は、8 時台が最も高い寄与率であり、6 時と 7 時、16 時と 17 時が負の

寄与率となっている。平日の第五主成分は 8 時台の活動と 6 時から 7 時および 16 時から 17 時の活動の区別をする軸であることが分かる。休日の第五主成分は 9 時から 13 時および 20 時から 23 時は正の寄与率になっており、9 時から 11 時および 22 時が高い正の寄与率である。反対に、3 時から 7 時および 15 時から 19 時は負の寄与率になっており、6 時から 7 時および 17 時が最も低い寄与率となっている。休日の第五主成分は平日の第五主成分と違い、8 時台の活動とそれ以外の時間帯の活動を分けるものではなく、6 時から 7 時の朝の活動と 9 時から 13 時までの活動を分けるものである。

これらの主成分分析による有効な主成分の各時間の寄与率から、平日と休日の各時間における活動の性質およびそれらの分類が考察できる。

4.2. クラスタ分析による生活パターン分類

次に k-means クラスタリングを用いて、5020 人（うち 1 人は活動量が 0 だったため省いた）の一年間の活動量の分布が平日と休日ごとに何パターンに分かれるか検討した。活動量は携帯電話を持ったうえで、指定するクラスタ数はギャップ統計量から算出した。

まず、5019 人からランダムに 1000 人抽出し、それらの平日と休日におけるギャップ統計量を 100 回推定し、最適クラスタ数を求めたところ、平日では最適クラスタリング数は 11 個、休日では最適クラスタリング数は 9 個となった。

1000 人の k-means 法によって、平日と休日ごとクラスタリングを行った。k-means 法は、データセットからあらかじめ指定されたクラスタ数 (i) だけのクラスタの代表点 (Ci) を無作為に選び、各データに対し Ci との距離を測り、各データをもっとも近くにある Ci のクラスタに分類する。そして、そのクラスタ内部で中心点を取り再度クラスタに分割しなおすというアルゴリズムで行われている³⁰。そのため、この代表点の選び方によって、クラスタリングの結果は毎回違う。そこで、k-means によるクラスタリングを 100 回繰り返し、評価関数によって最適なクラスタリング結果を採択する。k-means 法は各クラスタ内の分散が最小となることが最適である。評価関数の定義式はウェブサイト WIKIBOOKS の「Data Mining Algorithms In R」³¹から引用した。

評価関数定義式

$$\text{Within-cluster sum of squares} = \sum_j^k \sum_{i=1}^n \|x_i^j - c_j\|^2$$

(k : クラスタ数, n : データ数, c : クラスタの中心点)

平日のクラスタリングの評価関数の値は 9187.988 である．休日 は、1391.032 であった．平日および休日の各クラスの測定日数の平均と分散を算出した結果，測定日数が低いユーザのみのクラスなどは形成されず，どのクラスにも，測定日数が多いユーザと少ないユーザが含まれており，測定日数の差異はクラスタリング結果に影響を与えていないと考えられる．

4.2.1. 平日の生活パターン分類

以下に，クラススタリングに分けられた各グループの時間帯ごとの平均活動量の推移とクラスの人数を記す．

平日：クラス 1（55人）

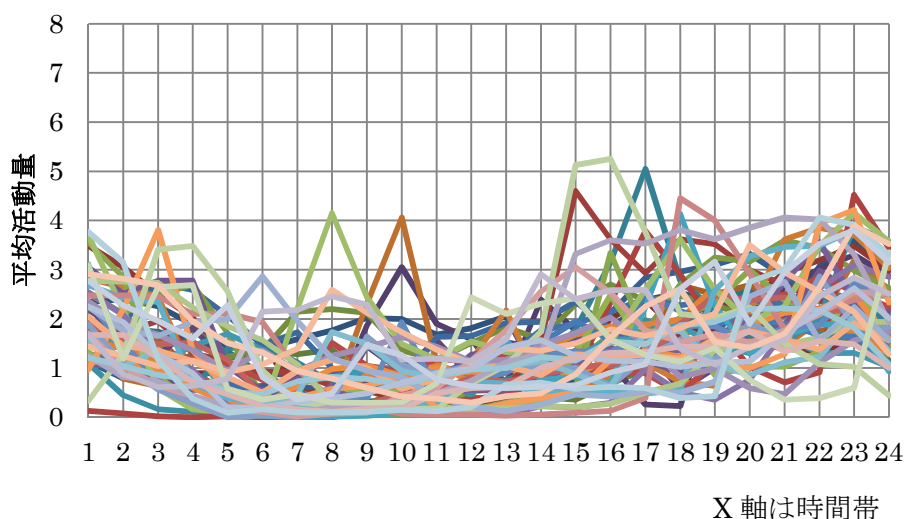


図 4-13：平日のクラス 1 の平均活動量の時間帯別分布

クラス 1 では日中の活動量より夜間の方がやや多いユーザが分類されている．1000 人中 55 人がこのクラスに分類される．このクラスのユーザは平日の午前の時間帯より午後の時間帯および夜間において，携帯電話を持ったうえで移動を含む活動を行っていることが分かる．

平日：クラス2（86人）

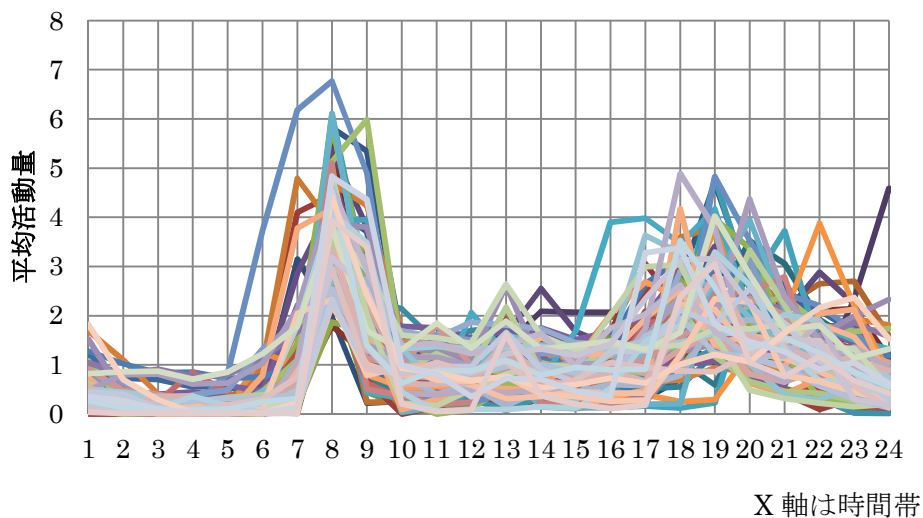


図 4-14：平日のクラス2の平均活動量の時間帯別分布

クラス2では8時を中心に急激に活動量が増加しており、定期的にこの時間帯に携帯電話を持ったうえで、移動を含む活動を行っているため、この位置情報の測定は通勤のため携帯電話を持って移動したため行われていると考えられる。その後、活動量が落ちており、携帯電話を伴った移動および活動をあまり行っていないことが分かるため、オフィスワーク等を行っていることが分かる。16時台から夜間にかけて、再度活動量が急激に増加し、0時から朝にかけて活動量は極端に低くなり保たれている。夕方から夜間にかけての活動量の増加は帰宅の移動のため位置情報の測定頻度が上がったためと考えられる。

このため、クラス2には夜間は活動を停止し、8時に出勤および通学し、その後勤務先・通学先では、携帯電話を持ったうえでの活動をあまり行っていないと推定され、夕方から夜間の間に帰宅するという平日は比較的規則正しい生活を行っているユーザがこのクラスに分類されることが考えられる。このクラスには1000人中86人が分類された。

平日：クラス3（1人）

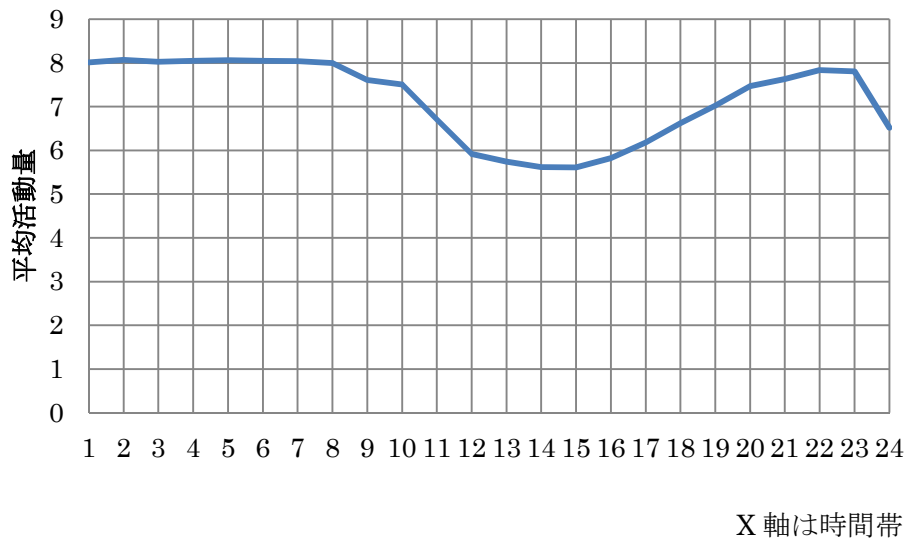


図 4-15：平日のクラス3の平均活動量の時間帯別分布

クラス3は1000人中1人しかいないが、このユーザは日中も夜間も携帯電話を持ったうえで活動を行っていることが分かる。

平日：クラス4（98人）

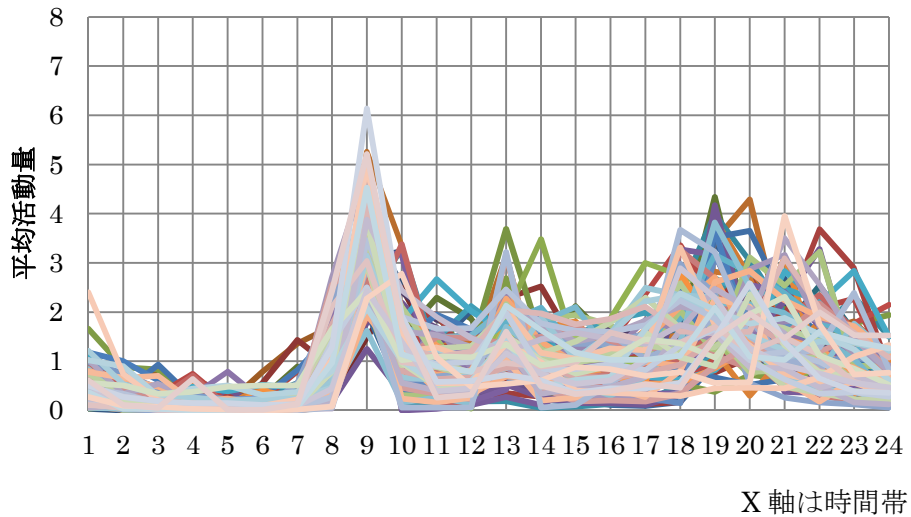


図 4-16：平日のクラス4の平均活動量の時間帯別分布

クラス4は0時から7までの活動量は低いままだが、9時を中心に活動量が急激に上がるため、この時間帯に携帯電話をもって移動を含む活動を行っていると考えられる。その後、活動量は下がるが、13時ごろ再び活動量が増加する。この増加も移動を含む活動を行った

ためであるので、昼休みのため携帯電話を持って移動するなど何らかの活動を行ったためと考えられる。その後再び活動量が下がり、夕方から夜間にかけての時間帯に、活動量が急激に増加するため、移動を行っていると考えられる。このクラスには定期的に朝出勤し、その後移動先では携帯電話を伴った大幅な移動を含む活動を行わないが昼頃に再び移動を含む活動を行い、夕方から夜間にかけて帰宅し、深夜就寝するという定期的な生活を平日に送っているユーザが分類されている。

平日：クラス5（63人）

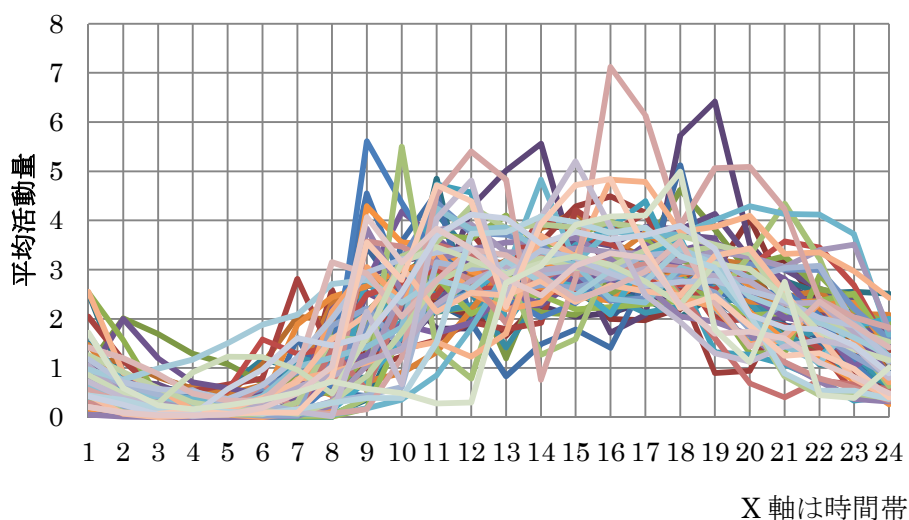


図 4-17：平日のクラス5の平均活動量の時間帯別分布

クラス5は、日中に活動しているがクラス2やクラス4と違って定期的に通勤する時間帯が同じユーザが集まっておらず、朝の時間帯の急激な活動量の増加（移動のためと考えられる）後、活動量が低くなっておらず、増加し続けているユーザが集まっている。

平日：クラス6（294人）

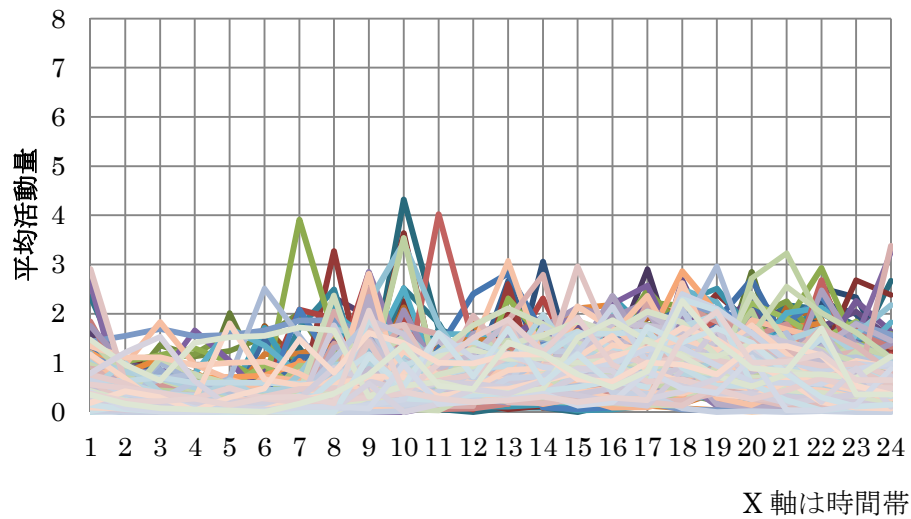


図 4-18：平日のクラス6の平均活動量の時間帯別分布

クラス6は日中と夜間の時間帯における活動量の明確な差が存在しない。しかし分類されたクラスの中では、最も多くのユーザがこのクラス6に分類されており、1000人中295人がこのクラスである。

平日：クラス7（43人）

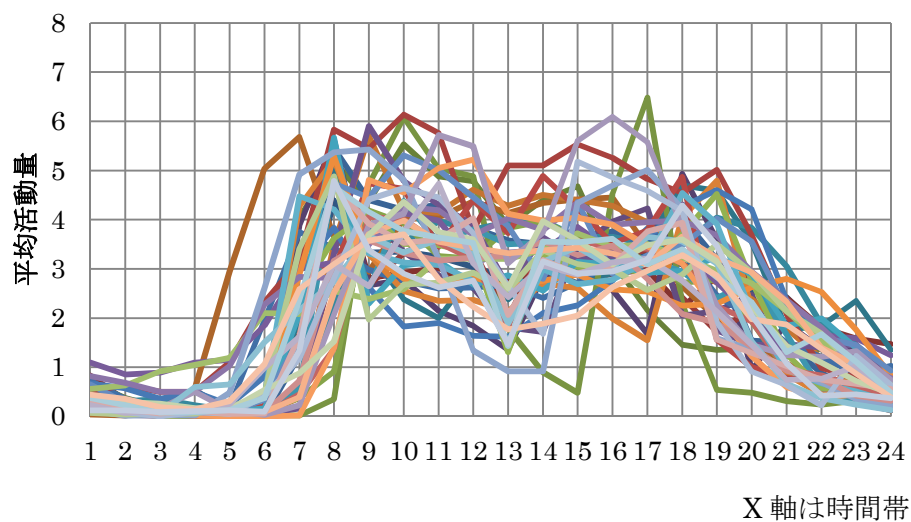


図 4-19：平日のクラス7の平均活動量の時間帯別分布

クラス7は朝の時間帯から活動量が急激に上がり、活動量は多少減少するも、活動量を比較的保つ。13時台にいったん活動量が減少し、再び同じ水準まで増加する。夕方から夜

間にかけて減少し深夜では、活動量がほとんどないため、就寝していると考えられる。

平日：クラス8（47人）

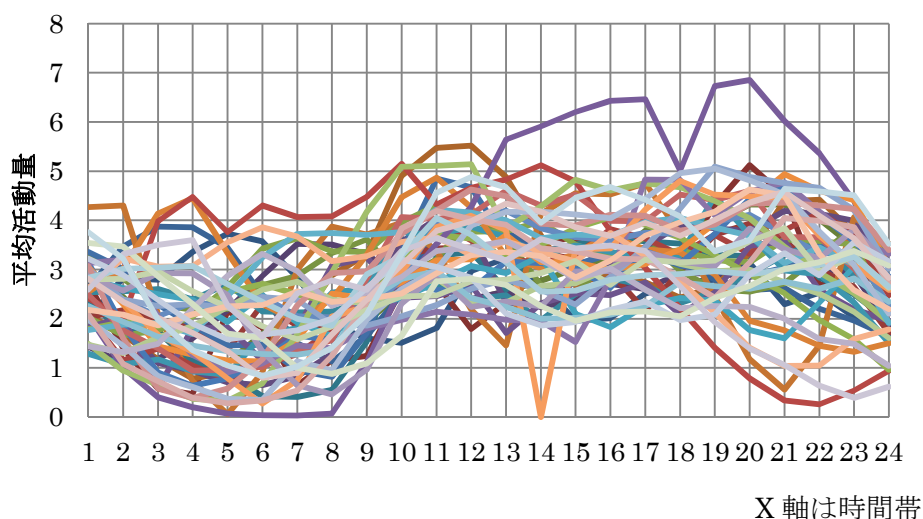


図 4-20：平日のクラス8の平均活動量の時間帯別分布

クラス8には、0時から9時の間に活動量の減少がピークを迎え、クラス2やクラス4などと比べると緩やかに増加にしている。増加して以降は18時ごろまでは活動量は2以上を保っているが、その後減少するユーザが少数いるが、大部分のユーザは2以上を保ったまま0時まで活動量が見られる。

平日：クラス9（103人）

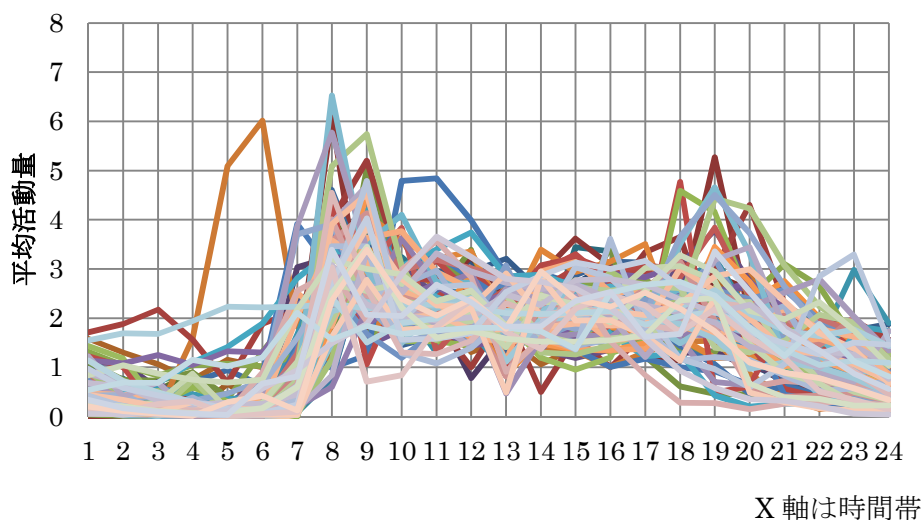


図 4-21：平日のクラス9の平均活動量の時間帯別分布

クラス9にはクラス2やクラス4と同様に活動量の急激な増加が6時から10時まで見られる。その後活動量は1以上を保たれている。18時から20時にかけて再び急激に増加するユーザもいるが、大半のユーザは緩やかに活動量を減少させていき0時から朝までは活動量が低下している。

平日：クラス10（190人）

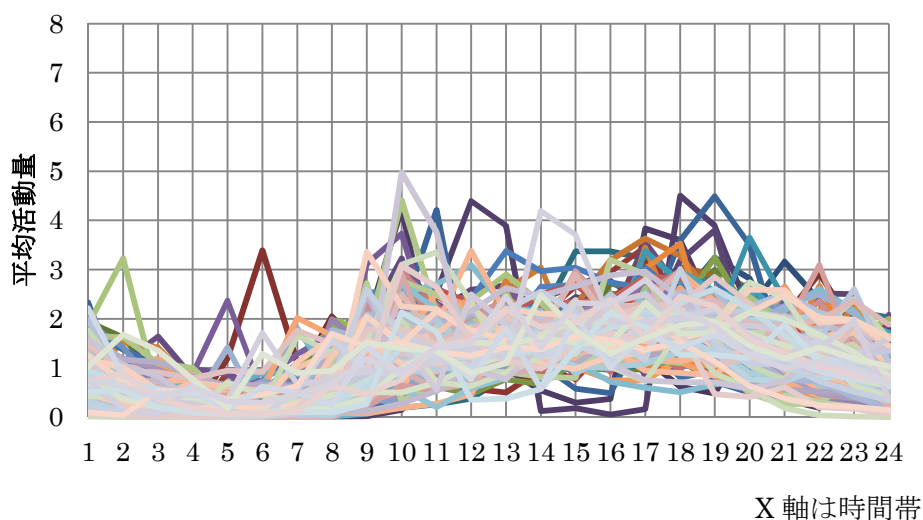


図 4-22：平日のクラス10の平均活動量の時間帯別分布

クラス10は、日中に活動量が高く保たれており、日中活動が行われているが、その増加や減少は非常に緩やかであり、活動量も全体を通して低い。

平日：クラス11（20人）

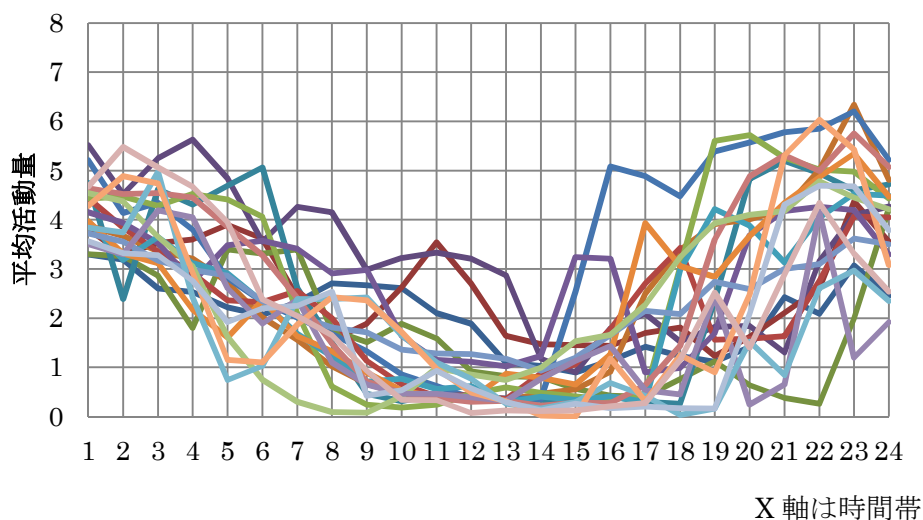


図 4-23：平日のクラス11の平均活動量の時間帯別分布

クラス1は日中よりも夜間に活動量が多くあり、夜間に携帯電話を身に着けた移動を含む活動が行われていることが分かる。

4.2.2. 休日の生活パターン分類

次に、休日の平均活動量の推移のクラスタリング結果をまとめる。

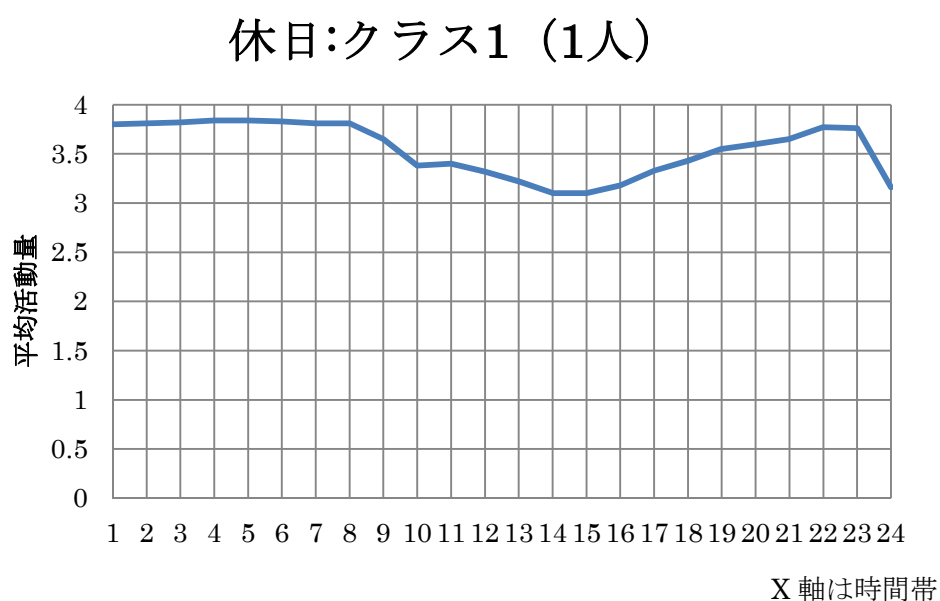


図 4-24 : 休日のクラス1の平均活動量の時間帯別分布

クラス1は、他のクラスと比べて平均活動量が多い。これは平日のクラス3と同一人物であり、同じように夜間の方が日中よりも平均活動量が多い。

休日:クラス2 (117人)

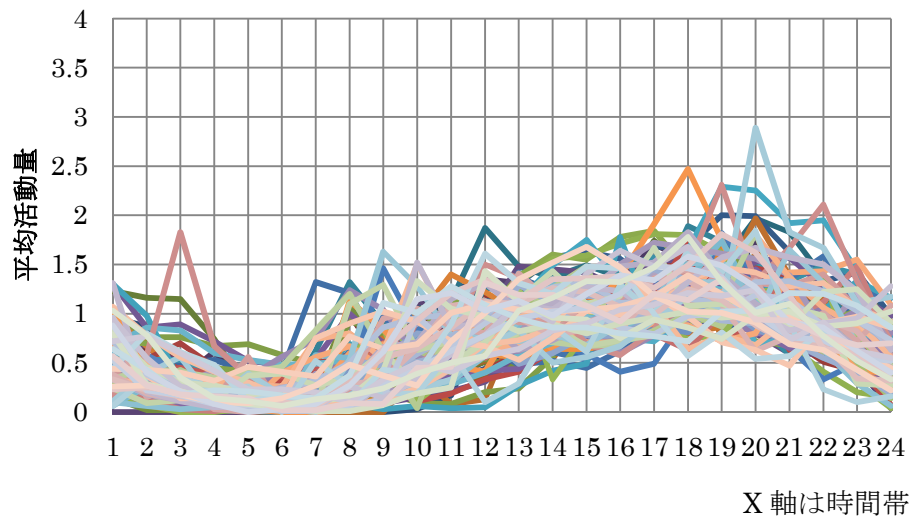


図 4-25 : 休日のクラス2の平均活動量の時間帯別分布

クラス2は午前中から緩やかに平均活動量が増加し続け18時ごろをピークに、翌日の5時ごろに向かって徐々に減少していく。

休日:クラス3 (173人)

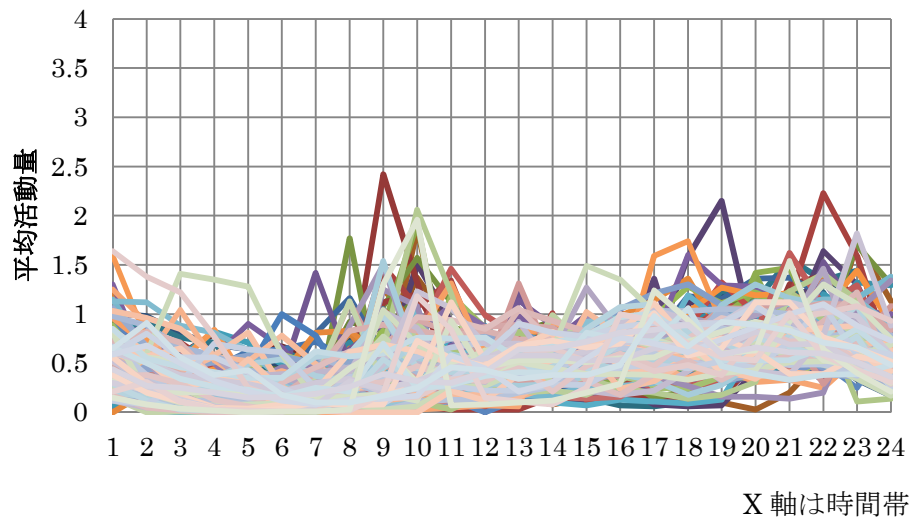


図 4-26 : 休日のクラス3の平均活動量の時間帯別分布

クラス3は、クラス2同様、午前中から夕方に向かって平均活動量が徐々に増加しているユーザもいる。クラス全体の傾向としては平均活動量1.5以下である、他クラスと比べて少ない。

休日:クラス4 (42人)

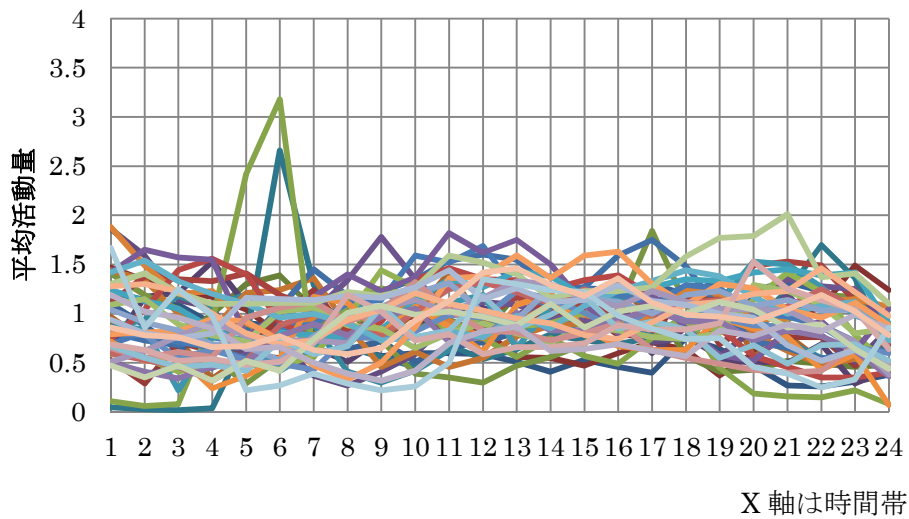


図 4-27 : 休日のクラス4の平均活動量の時間帯別分布

クラス4には、どの時間においても、およそ0.3以上1.5以下に保たれており、定期的な平均活動量分布の変動が見られない。

休日:クラス5 (116人)

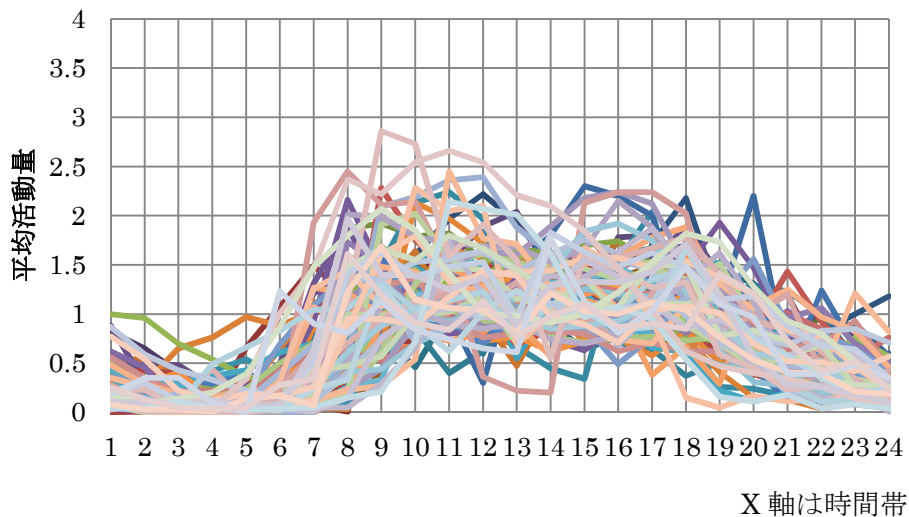


図 4-28 : 休日のクラス5の平均活動量の時間帯別分布

クラス5には、クラス3と同様に日中に活動量が増えているが、クラス3より日中の活動量が多い。

休日:クラス6 (242人)

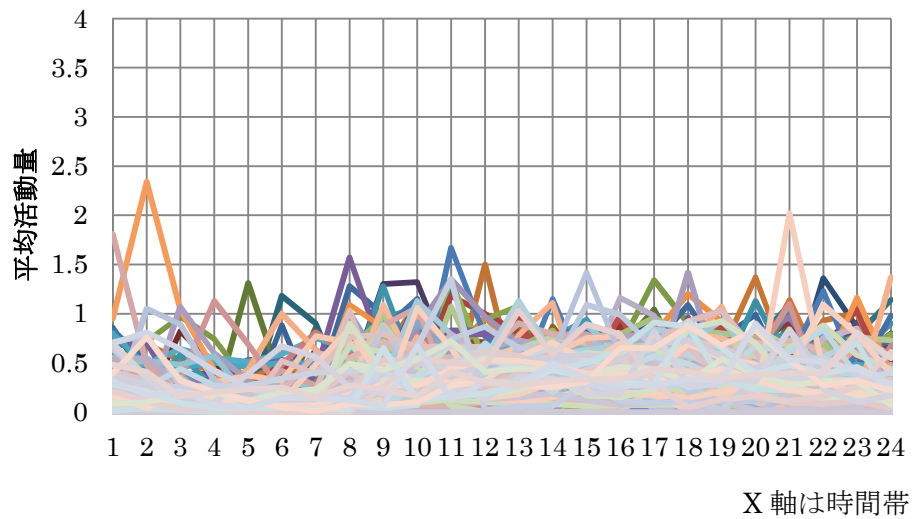


図 4-29：休日のクラス6の平均活動量の時間帯別分布

クラス6では、日中と夜間の平均活動量の分布に明確な差がなく、平均活動量の大半が1以下に抑えられている

休日:クラス7 (26人)

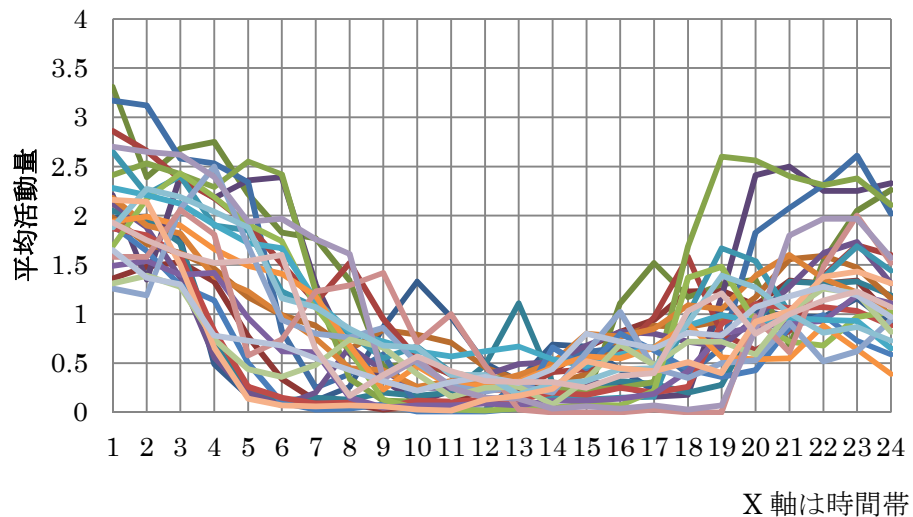


図 4-30：休日のクラス7の平均活動量の時間帯別分布

クラス7は、日中の時間帯よりも夜間の時間帯の方が平均活動量が多い。

休日:クラス 8 (37人)

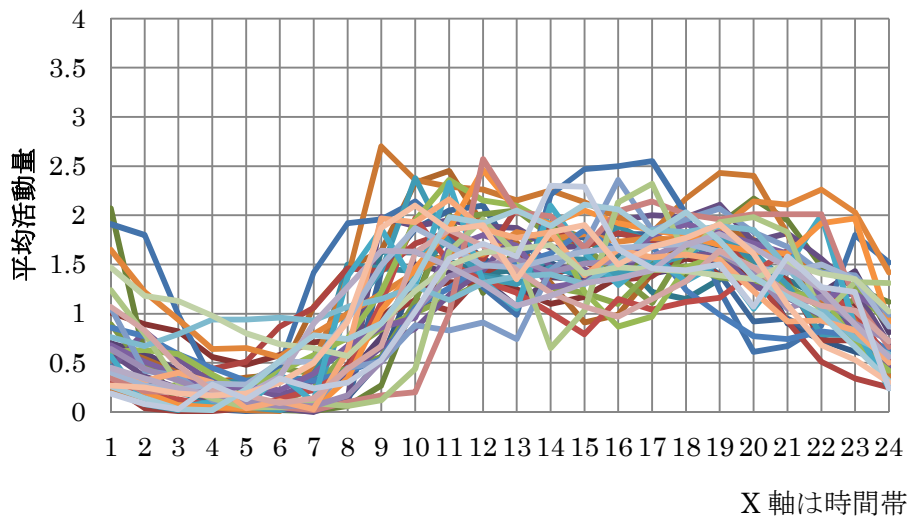


図 4-31 : 休日のクラス 8 の平均活動量の時間帯別分布

クラス 8 は、日中の方が夜間よりも平均活動量が多い。したがって、朝から定期的に日中活動を行っているユーザが分類されていると考えられる。

休日:クラス 9 (246人)

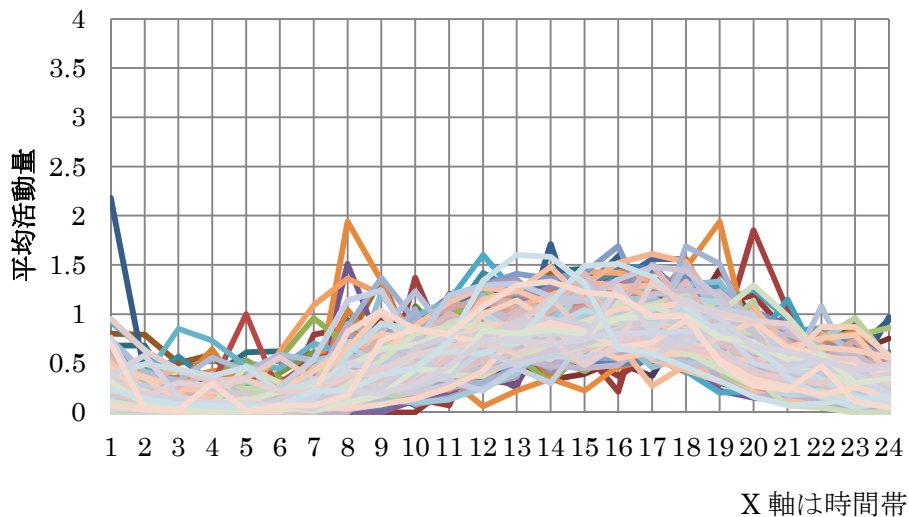


図 4-32 : 休日のクラス 9 の平均活動量の時間帯別分布

クラス 9 は、クラス 8 と同様に日中に活動を行っているユーザが分類されている。しかし活動を行っていないと考えられる深夜の平均活動量と活動を行っていると考えられる日中の平均活動量の差が小さいため、クラス 8 ほど定期的に、日中に携帯電話を伴った活動

を行っているわけではないと考えられる。

4.3. クラスター分析を用いた生活パターンのまとめ

以上の平日と休日のクラスタリング結果から、日中と夜間の時間帯のどちらに活動量の増加のピークがあるか、どの時間帯に急激な増加があるか、平均活動量が低い時間帯と高い時間帯における平均活動量の差から、クラスタリング結果から生活パターンが分類する。平日と休日の分類されたクラスから大まかに以下の生活パターンに分類してみた。

まず、平日は5つのグループに分けられる。一つ目は日中よりも夕方から夜間および深夜の方が、平均活動量が多く、夕方から翌日の早朝にかけて活動のピークがある『昼夜逆転型』の生活パターンのグループである。二つ目は、8時および9時を中心に急激な増加があるため、定常的にその時間帯に通勤もしくは通学をしていると考えられ、その後活動量が落ちる『固定勤務型』の生活パターンのグループである。3つ目は、朝に通勤および通学があるがその後も同程度の活動量を維持する『日中外回り型』の生活パターンである。4つ目は、夜間と日中の活動量にあまり差が出ないユーザと早朝に活動量が落ちるユーザが混合しており、定時的な活動量の分布が見られない『低レベルの活動型』の生活パターンである。5つ目は、日中と夜間の活動量の差が見られず、一日にわたり活動量が少ない『非活動型』の生活パターンである。

休日の生活パターンを5つに分ける。1つ目は日中よりも夜間の方が活動量の増加がある『夜型』の生活パターンである。2つ目は、日中においては平均活動量が常にピークのままを保持しているが夕方から夜間にかけて減少する傾向のある『日中活動型』である。3つ目は、日中の方が夜間よりも活動量がやや多く、休日にも日中活動を行うが、わずかな増加のピークが夕方にある『夕方中心日中活動型』である。4つ目は、日中と夜間において平均活動量が少なく、変動の幅も小さい『低レベルの活動型』である。5つ目は、日中と夜間において平均活動量があるが、その変動が見られない『非定時活動型』である。

平日のクラスタリング結果から大別された5つの生活パターングループである。

1) 夜間に活動する『昼夜逆転型』

該当するクラスは、クラス1、クラス3とクラス11が該当する。

2) 8時に出勤（移動）する『固定勤務型』

該当するクラスは、クラス2、クラス4とクラス9が該当する。

3) 8時に出勤したのち移動を含む活動が持続する『日中外回り型』

該当するクラスは、クラス5とクラス7である。

4) 1日の平均活動量が低い『低レベルの活動型』

該当するクラスは、クラス6とクラス10である。

5) 日中と夜間に移動および活動分布の明確な差異は見られないが、一定以上の活動量が1日に渡ってある『非定時活動型』

該当するクラスはクラス 8 である.

以下は, 休日のクラスタリング結果から大別された 5 つの生活パターングループである.

1) 夜間に活動する『昼夜逆転型』

該当するクラスは, クラス 1 とクラス 7 である.

2) 日中に活動する『日中活動型』

該当するクラスは, クラス 5 とクラス 8 である.

3) 夕方に活動のピークが来る『夕方中心活動型』

該当するクラスは, クラス 2 とクラス 9 である.

4) 日中にやや活動量が増加するが, 1 日を通して平均活動量が低い『低レベルの活動型』

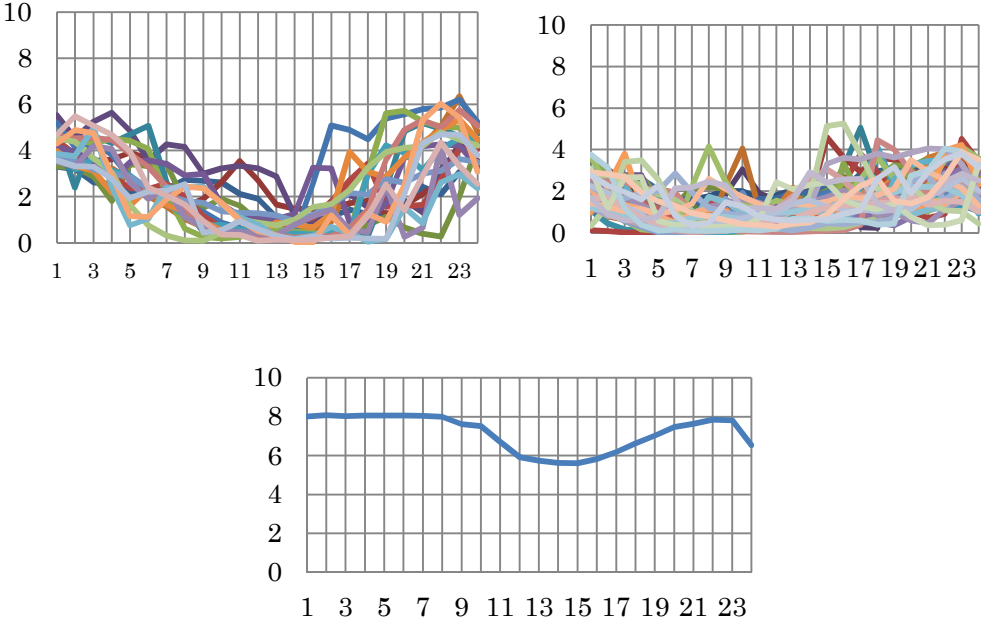
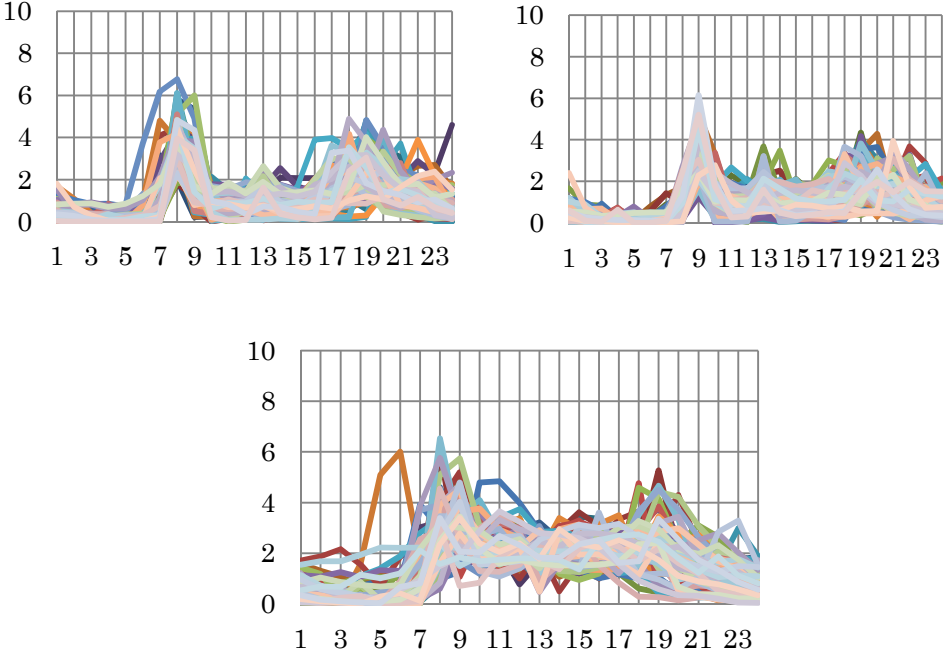
該当するクラスは, クラス 3 とクラス 6 である.

5) 日中と夜間に移動および活動分布の明確な差異は見られないが, 一定以上の活動量が 1 日に渡ってある『非定時活動型』

該当するクラスは, クラス 4 である.

以下の表 4-1 と表 4-2 にてまとめる.

表 4-1：平日のクラスタリング結果の生活パターン分け

<p>昼夜逆転型</p> <p>76人</p>	<p>クラス 1 およびクラス 3 およびクラス 11 がこれにあたる。1000 人中 7.6% のユーザが属する。主に夜間に活動を行っている。</p> 
<p>固定勤務型</p> <p>287人</p>	<p>クラス 2 およびクラス 4 およびクラス 9 が該当する。1000 人中 28.7% のユーザが属する。7時から8時にかけて定期的に出勤し、帰宅時間は各ユーザによって異なる。出勤後は活動が抑えられ、12時から13時に、移動を含む活動が見られるユーザがいる。有業者・通学者が多く分類されていると考えられる。</p> 

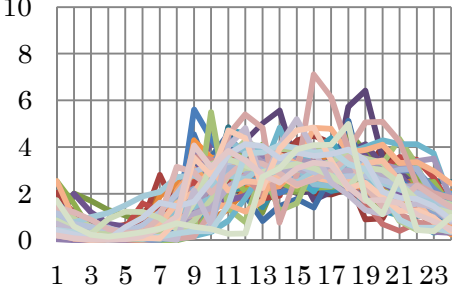
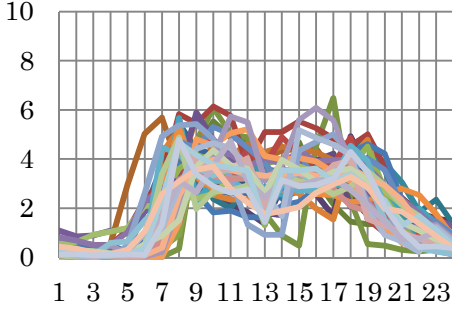
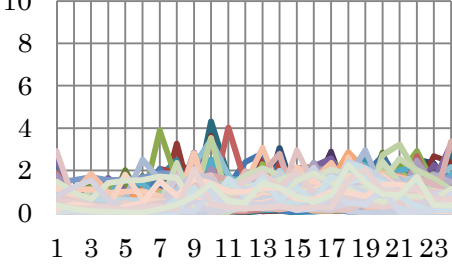
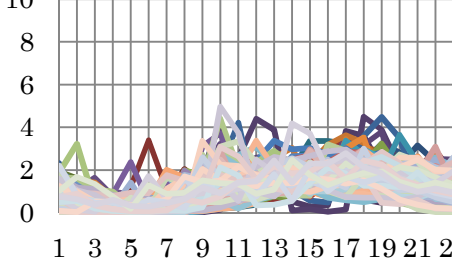
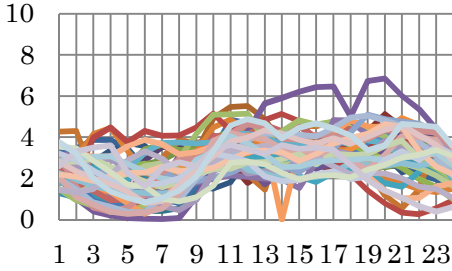
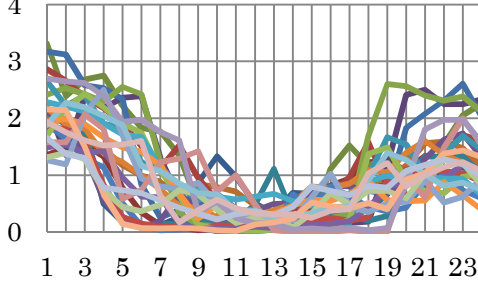
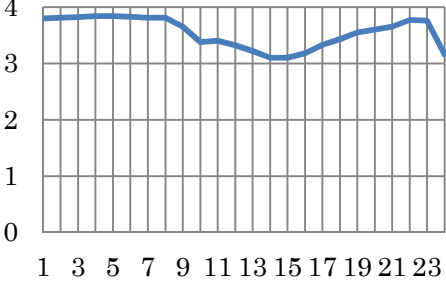
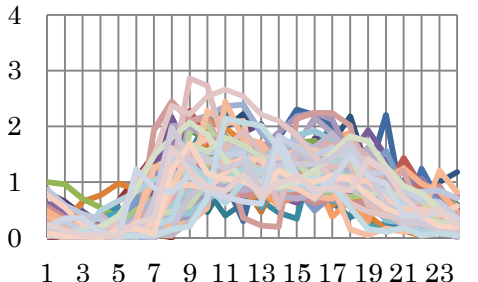
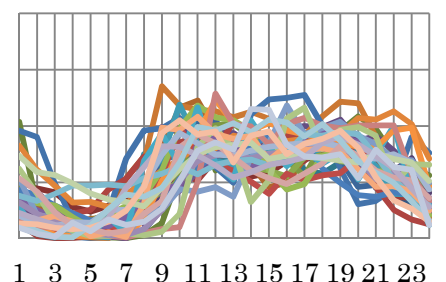
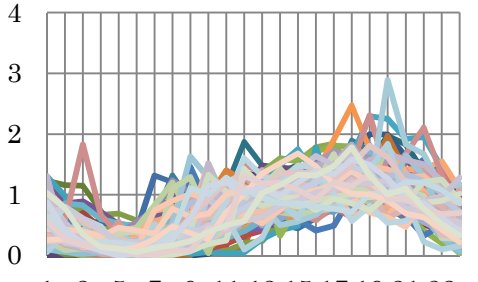
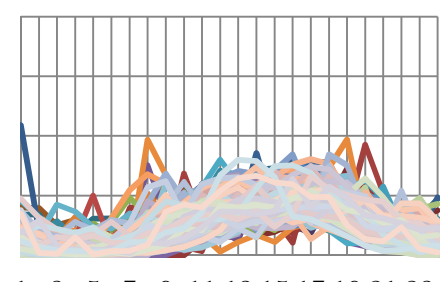
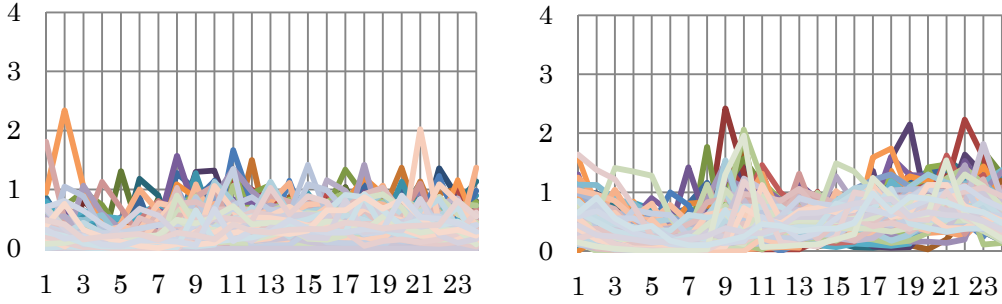
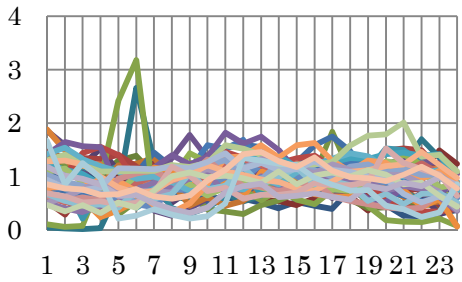
日 中 外 回 り 型 106 人	<p>クラス 5 およびクラス 7 が該当する。1000 人中 10.6%のユーザが属する。7 時から 8 時の出勤後も、移動を含む、何らかの活動を続けていると考えられるユーザが分類されている。有業者が多く分類されていると考えられる。</p> <div>   </div>
低 レ ベ ル の 活 動 型 484 人	<p>クラス 6 およびクラス 10 が該当する。1000 人中 48.4%のユーザが属する・移動距離が少ないユーザが含まれている。つまり、勤務地と自宅が近い人および主婦など自宅を中心に活動をする人が含まれていると考えられる。</p> <div>   </div>
非 定 時 活 動 タ イ プ 47 人	<p>クラス 8 が該当する。1000 人中 4.7%のユーザが属する。常に移動する人および携帯電話を持ったうえで何らかの活動が多い人が分類されていると考えられる。</p> <div>  </div>

表 4-2：休日のクラスタリング結果の生活パターン分け

<div>昼夜逆転型</div> <div>27人</div>	<p>クラス 1 およびクラス 7 が該当する。1000 人中 2.7% のユーザが属する。</p> <div>   </div>
<div>日中活動型</div> <div>153人</div>	<p>クラス 5 およびクラス 8 が該当する。1000 人中 15.3% のユーザが属する。</p> <p>平日の日中に活動するグループと違い、出勤後の活動量の急激な減少がない。</p> <div>   </div>
<div>夕方中心日中活動型</div> <div>363人</div>	<p>クラス 2 およびクラス 9 が該当する。1000 人中 36.3% のユーザが属する。</p> <p>夕方に活動がもっとも多くなる。</p> <div>   </div>

低 レ ベ ル の 活 動 型	<p>クラス 3 およびクラス 6 が該当する。1000 人中 41.5% のユーザが属する。日中を通して、活動量が少なく、移動などが少ないと考えられる。</p> 
415 人	
非 定 時 活 動 タ イ プ	<p>クラス 4 が該当する。1000 人中 4.2% のユーザが属する。移動や携帯電話を伴った何らかの活動を常に行っている。</p> 
42 人	

また、平日の生活パターンごとに休日の生活パターンがどのように変わるのか集計をした。

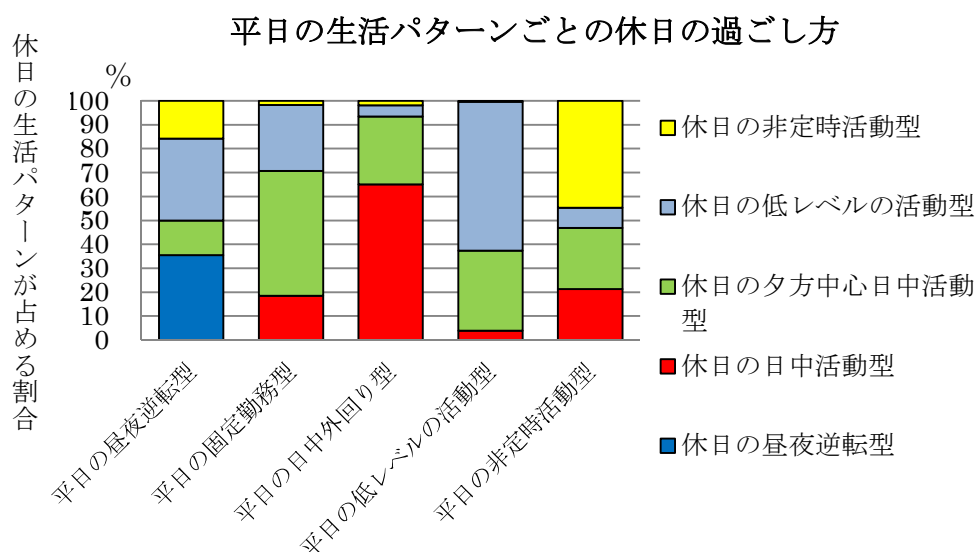


図 4-33：平日の生活パターンごとの休日の過ごし方

図 4-33 から、平日の昼夜逆転型は休日では低レベルの活動型になるユーザと休日にも昼夜逆転型となるユーザが多くいることが分かる。平日は昼夜逆転型でも、休日はあまり移動をしない低レベルの活動型や夕方中心日中活動型になるユーザは、平日は夜間に働き、休日は非就業日となる生活パターンを持っていると考えられる。しかし、平日の昼夜逆転型で休日にも昼夜逆転型や常に移動および活動している非定時型活動型のユーザは、平日と休日で就業日と非就業日が分かれていないかユーザであると考えられる。

平日において低レベルの活動型であるユーザは、休日にも低レベルの活動型であることが多い。したがって、低レベルの活動型には平日と休日の活動パターンが似通っているユーザが多く含まれている可能性がある。

また、平日の非定時活動型は、休日にも非定時活動型となっていることが多い。非定時活動型は、平日と休日の活動パターンが同じであるか、不定期に移動や活動を繰り返すため、平均活動量の時間帯別の分布では、明確なパターンが見いだせなかった可能性がある。

4.4. 社会生活基本調査による生活パターンの整合性の評価

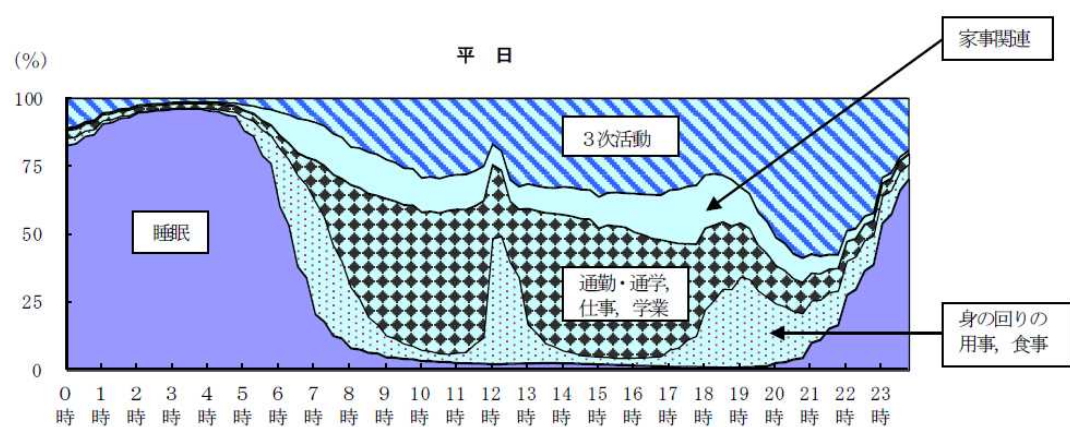
移動履歴データから抽出した生活パターンが、『平成 23 年度社会基本調査報告書の生活時間に関する結果』(2012)²⁶ で得られた生活パターンと比較することにより、その分類の整合性を評価する。

4.4.1. 社会生活基本調査による生活パターンの可視化

4.3 節にて平日・休日ごとに大別された生活パターンが、実際に行われた生活に関する全国調査と比較し、どの程度、調査結果を反映しているかその整合性を検証した。

社会生活基本調査から読み取れる生活パターンから、移動履歴データから算出した活動量の時系列な増加・減少から読み取った活動パターンが妥当であるか検証した。そのため、本研究で使用している移動履歴データ収集対象者と社会基本調査対象者の性別・年代構成は違うが、社会基本調査対象者に対して補正係数を用いて集計結果を補正していない。

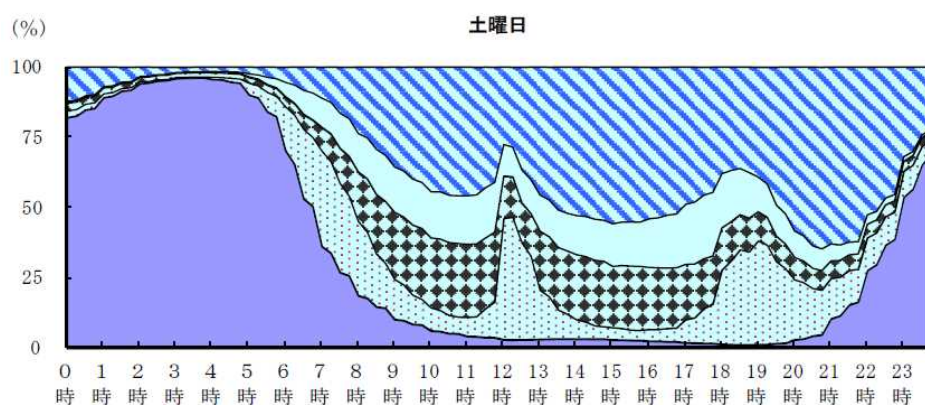
比較に用いた統計データは、平成 23 年度の社会生活基本調査である。図 4-34 は、『平成 23 年度社会基本調査報告書の生活時間に関する結果』(2012)²⁶ の報道資料「生活時間に関する結果要約」の p.6 における図 1-4「曜日、行動の種類、時間帯別行動者率」を引用したものである。行動者率(%)とは、その行動をとっている行動者のその属性別人口に対する割合を示している。なお、グラフは男女ともに集計した結果である。



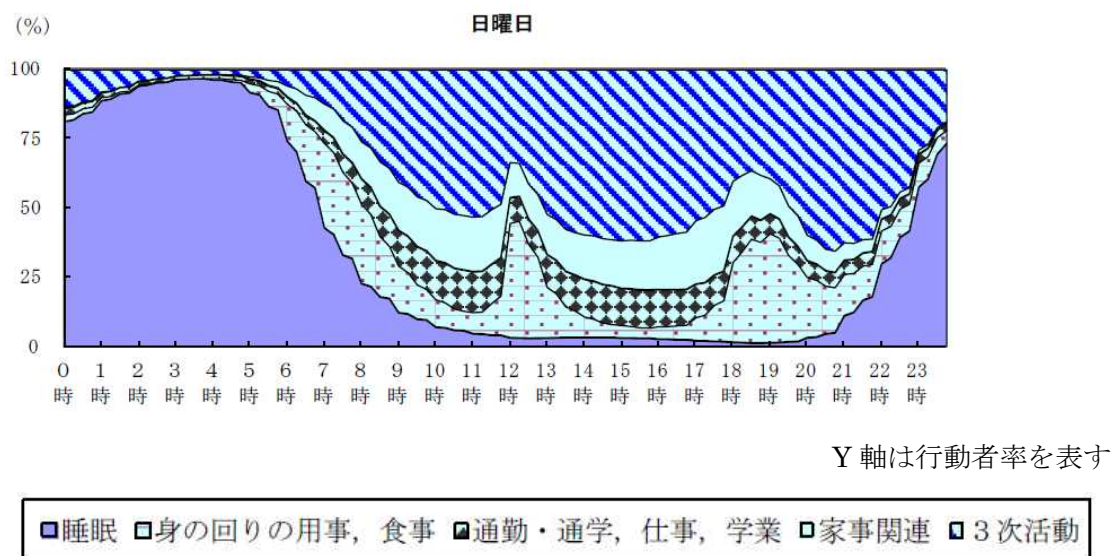
Y 軸は行動者率を表す



※3次活動とは当人が自由に使える時間における活動を指す



Y 軸は行動者率を表す



※3次活動とは当人が自由に使える時間における活動を指す

図 4-34：曜日ごとの時間帯別にみる行動者率の一日の推移

図 4-34 から、平日は、7 時および 8 時から移動は、通勤・通学を指し、その後の日中の活動は主に仕事・学業である。12 時から 13 時にかけて身の回りの用事、つまり食事などが行われている。19 時から帰宅が始まり、就寝となる。平日の出勤時刻および帰宅時刻については、雇用形態によって異なる傾向があることが同報道資料(2012)³⁰にて報告されている。正規の職員・従業員および契約社員、嘱託、労働者派遣事業所の派遣社員は 7 時台に出勤する傾向がある。パートは 8 時台に出勤する傾向があり、アルバイトは、8 時台が他の時間帯と比べて出勤する人数が多いが、おおむね出勤時刻が分散している。帰宅時刻は、どの雇用形態においても 17 時から 19 時にかけて集中しているが、パートでは 14 時前に帰宅する人数も同程度いる。また、帰宅時間は分散していることが報告されている。

休日は、7 時および 8 時における通勤・通学を行う人が平日に比べて減る。3 次活動の行動者が増えている。15 時から 16 時にかけて、3 次活動を行う人が増え、18 時から 20 時にかけて身の回りの用事および家事関連の行動を行う人が増えるため、3 次活動を行う人が急減する。この午後の活動内容の変化は、平日の同じ時間帯では活動内容の変化は見られない。したがって、休日の 15 時以降の活動内容の変化は、移動履歴データの位置情報測定頻度にも何らかの影響があると考えられる。

4.4.2. 位置情報測定頻度の生活パターンの整合性

4.1.1 節から、有業者の 7 時から 8 時における出勤時刻および帰宅時間の分散性が、一時間当たりの平均位置情報測定頻度(＝活動量)から読み取れることが分かった。また、4.4.1 節の『平成 23 年度社会生活基本調査報道資料』(2012)²⁶から引用された図 4-34 から、

4.2 節の平日の生活パターンにおける 7 時から 8 時と夕刻から夜間の活動量の増加が出勤と帰宅のためであることがわかる。このように、活動量の増加が時間帯によっては移動を意味し、活動量が一定に維持されていれば、その時間帯は長距離の移動もしくは携帯電話に加速度が生じる何らかの活動を行っていることを意味することが分かった。

したがって、平日の活動量の増加のピークとその時間帯から、クラスタリング結果を 5 つの生活パターンに大別し、『固定勤務型』と『日中外回り型』という出勤・帰宅時間に対応した生活パターンと夜間活動する『昼夜逆転型』に分けることは妥当性がある。また、上記の 3 つのグループに比べて、時間帯別の活動量の分布が異なる『低レベルの活動型』と『非定時活動型』に分けることができる。

同様に、休日においても平日と同様に活動量増加のピークとその時間帯から、夜間に活動するグループと日中に活動をするグループに分けた。さらに、日中の活動のピークから『日中活動型』と『夕方中心日中活動型』に分けることは、図 4-35 から見られる活動の行動者率の推移から妥当性があることが分かる。

4.5. 生活パターンごとの自宅・勤務地推定における問題点

自宅および勤務地・学校などに人間の生活に深く関わる場所は、個人の生活パターンや生活行動の推定に重要である。このように携帯電話から収集される情報を用いて生活に重要な場所を推定するのに、以下のように、時間帯に基づいた特徴が用いられてきた。

Zhu et al.(2012)³²では、携帯電話の Wi-Fi の強弱、平日か休日か、ある一定のエリアに滞在している時間帯、一日におけるそのエリアへの訪問回数、加速度、およびアプリケーションの使用履歴、通話ログなどの特徴量のうち、自宅と勤務地、滞在した他人の家を分類するには、どの特徴量が一番重要となるか解析している。ユーザがある一定のエリアに滞在するときに Wi-Fi を使用するため、滞在しているか移動しているか状態の推定に有効であると考えている。特徴の重要度の解析結果では、8 時から 10 時での時間帯、Wi-Fi の値、20 時から 22 時が 3 つのロケーション（自宅・勤務地・友達の家）の分類に有効な特徴であった。

本研究で使用している Apichon. et al(2010)²⁷における自宅と勤務地の場所推定においても、活動日記付きの移動履歴から明らかになった生活パターン、すなわち夜間時間帯には自宅、日中時間帯には勤務地という生活パターンを元に自宅・勤務地推定アルゴリズムを作成している。なお、活動日記とはユーザの滞在場所と滞在時間、滞在目的の記録である。

しかしながら、4.2 節にて、自宅・勤務地を推定するには、日中に仕事および学校に行き、深夜は自宅で過ごすという生活パターンを持つユーザのほかに、夜間に移動行い、日中に自宅で過ごしている夜間就業者の生活パターンも全体の 8%程度は存在する。しかし、夜間就業者は日中に活動（就業を含む）するユーザに比べて人数が少ない。そのため、機械学習などを用いて推定アルゴリズムを得るために使用する正解データ（移動履歴と活動内容

が記録されているデータ）に含まれている夜間就業者が少ないため，日中に活動するユーザの生活パターンに基づいた自宅・勤務地推定が行われていると考えられる．

活動日記付き移動履歴データから明らかになった日中に仕事および学校に行き，夜間は自宅で過ごすという生活パターンに基づいた推定アルゴリズムで算出された自宅と勤務地が実際の携帯電話の移動履歴に適用すると，どのような結果になるか，4.3 節の生活パターンごとに検討した．この検討には平日の一日の自宅と勤務地の各時間帯の滞在時間の割合を用いる．滞在時間は，自宅および勤務地の半径 500m 以内に滞在している時間である．なお 1000 人中 726 人が自宅および勤務地が推定された．

4.5.1. 在宅時間の分布に着目した分析

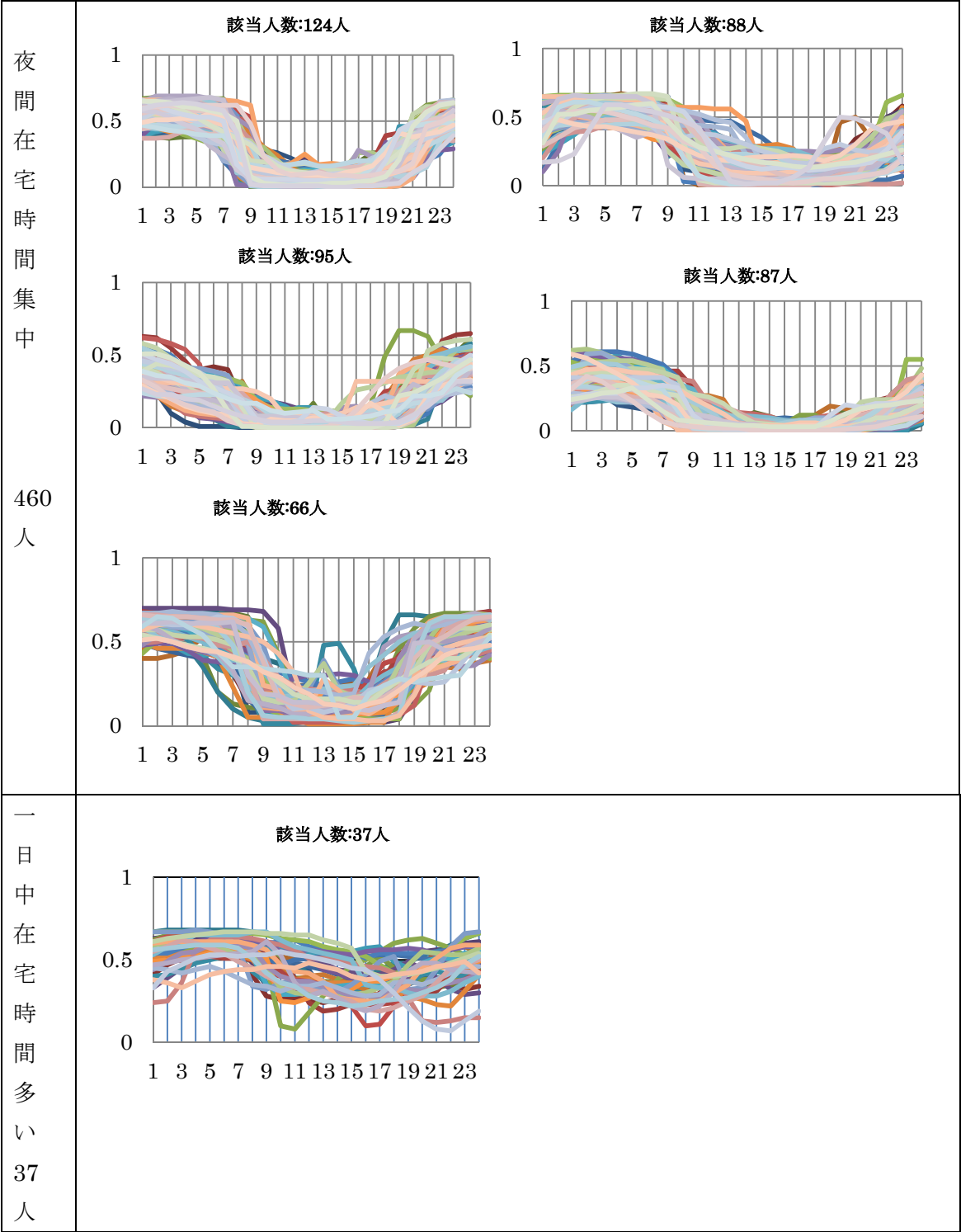
まず，夜間就業者が分類されている『昼夜逆転型』のユーザと日中に活動を行っている生活パターンのユーザが，算出された自宅と勤務地における一日の滞在時間がどのように分布しているか明らかにした．

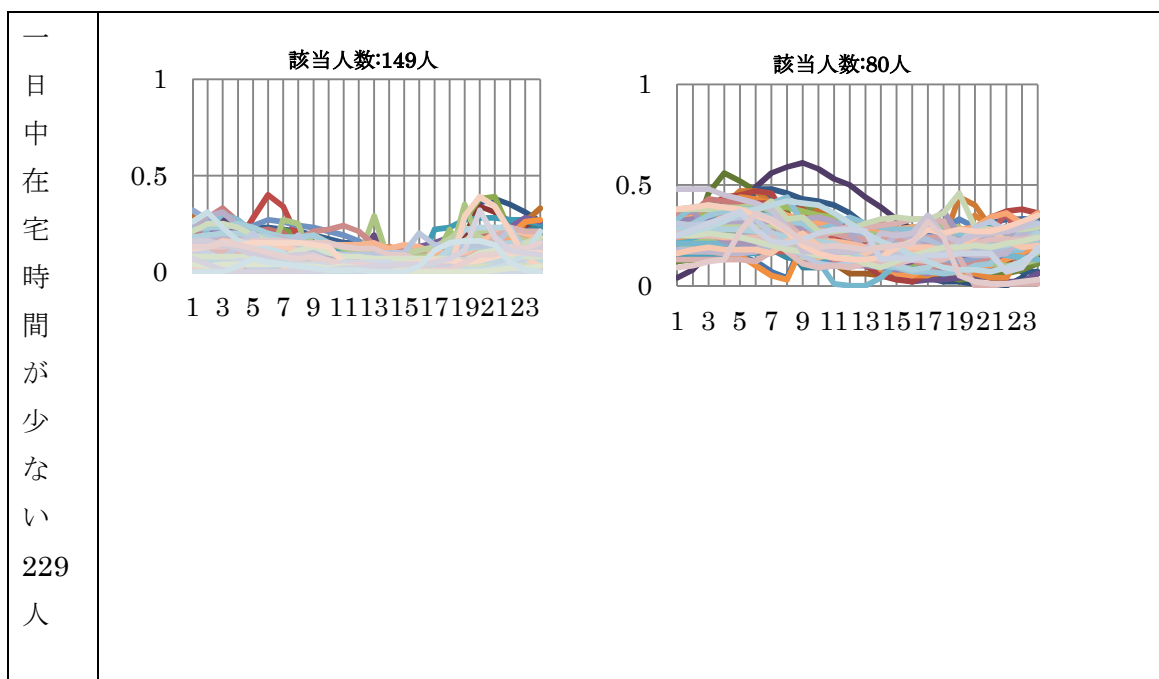
4.2 節の位置情報測定頻度を用いた生活パターン分類と同様に，一日の在宅時間分布をギャップ統計量にて最適クラスタ数を求めてから，k-means 法によってクラスタリングを行った結果，8 つのクラスに分類できた．それらのクラスは 3 つのクラスに大別することができる．

夜間に在宅時間が集中するグループと一日に渡って常に一時間当たり 18 分から 42 分の在宅時間を維持し，夜間と日中の在宅時間の差が少ないグループがある．さらに，一日に渡って一時間当たりの在宅時間 18 分以下のクラスと 30 分以下のクラスの二つを合わせた．上記にクラスと比べると在宅時間が少ないグループである．

以下の表 4-3 に在宅時間のグループを記す．

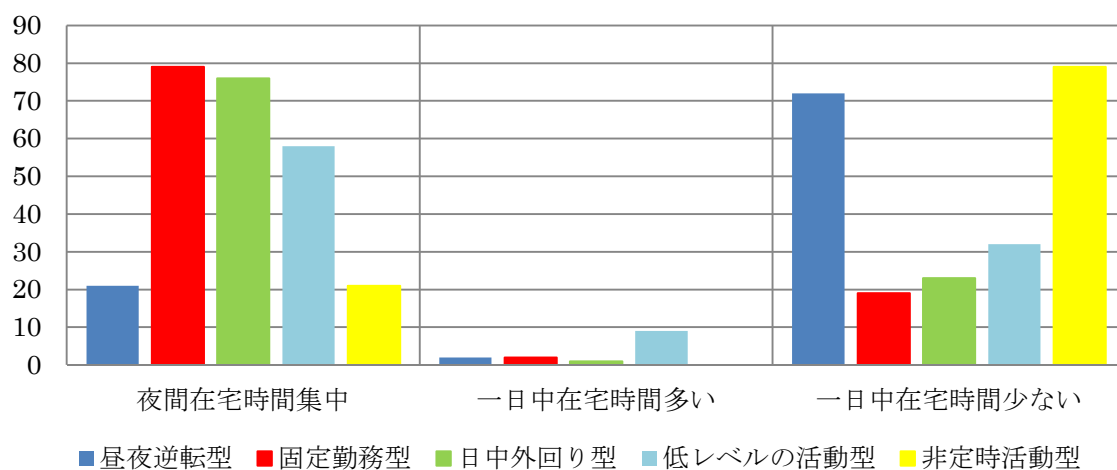
表 4-3 : 在宅時間のパターン類型化





4.2 節で大別した 5 パターンの平日の生活パターンに属するユーザが、上記の表の在宅時間のどの分布パターンに分類されているかグラフで可視化した。

生活パターンごとの在宅時間の分布



Y 軸は該当ユーザの人数を表す。

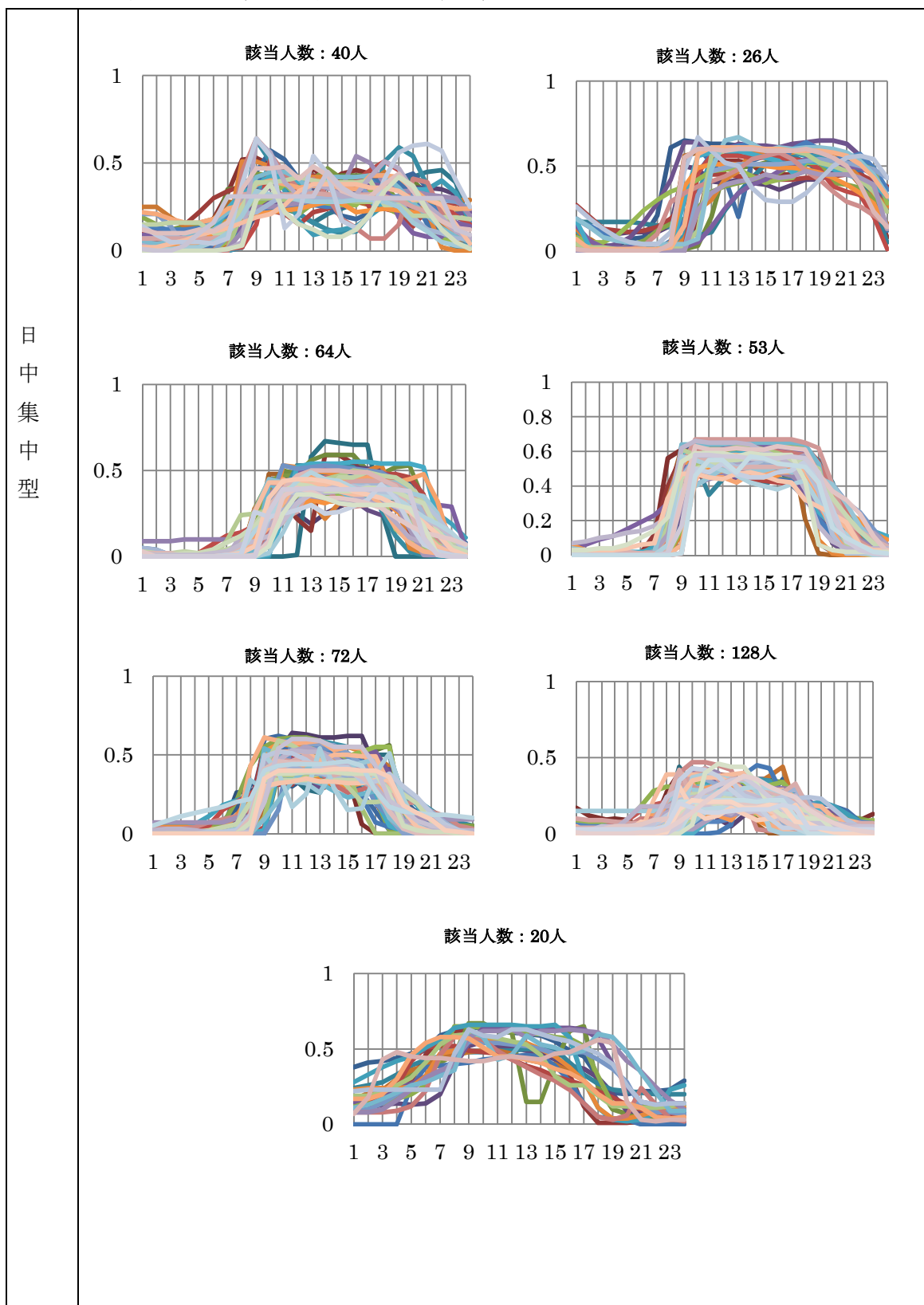
図 4-35：生活パターンごとの在宅時間の分布

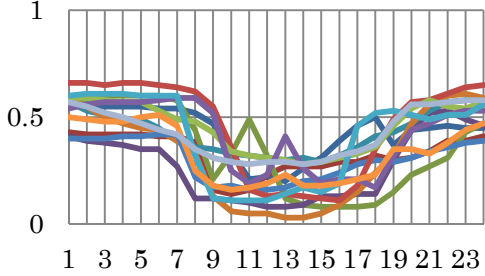
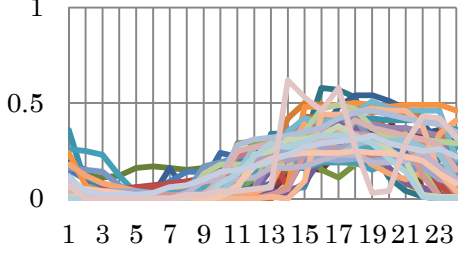
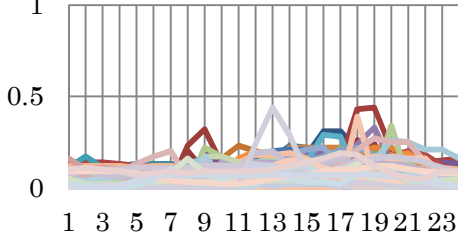
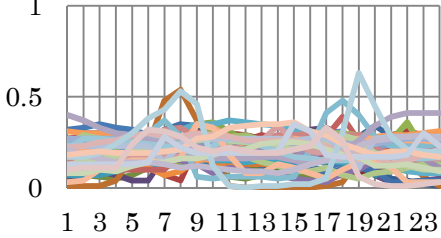
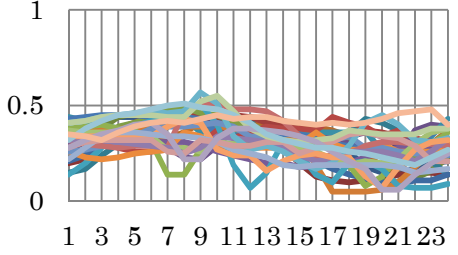
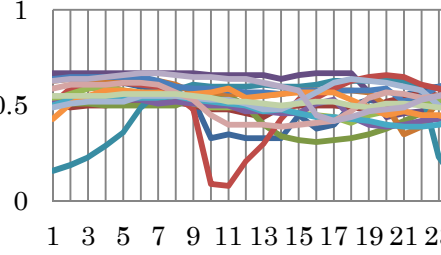
平日の生活パターンが『昼夜逆転型』のユーザの72%が1日の在宅時間の分布が『一日中在宅時間が少ない』グループに属している。日中に活動の中心がある『固定勤務型』と『日中外回り型』と『低レベルの活動型』の3つは夜間の時間帯に在宅時間が集中している。夜間の時間帯に活動し、日中に活動をしていない『昼夜逆転型』のユーザは、自宅には日中に滞在すると考えられるが、ほとんどの『昼夜逆転型』のユーザが日中と夜間の在宅時間の分布に差が出ていない。分析の対象となった『昼夜逆転型』のユーザに関して、本当の自宅がどこかに関する検証データが存在しないため、信頼性の高い評価をすることは困難であるが、『昼夜逆転型』のユーザの自宅の推定の精度が悪いと推察される。

4.5.2. 勤務地への滞在時間に着目した分析

次に、勤務地に滞在時間の分布を平日の生活パターンごとに分析した。平日の勤務地の滞在時間分布は在宅時間分布と同様に、クラスタリングを行い、13クラスに分類することができた。それらのクラスは、日中に滞在時間が集中するグループと夜間に集中するグループ、夕方に集中するグループ、一日を通して滞在時間が少ないグループ、一日を通して滞在時間が長いグループの5グループに分けることができる。以下の表にまとめる。

表 4-4：勤務地への滞在時間のパターン類型化



夜間集中型	<p>該当人数：13人</p> 
夕方集中型	<p>該当人数：50人</p> 
一日中滞在時間が少ない	<div> <div> <p>該当人数：158人</p>  </div> <div> <p>該当人数：96人</p>  </div> </div>
一日中滞在時間が多い	<div> <div> <p>該当人数：30人</p>  </div> <div> <p>該当人数：16人</p>  </div> </div>

次に、4.2 節で大別した 5 パターンの平日の生活パターンに属するユーザが、上記の表の勤務地滞在時間のどの分布パターンに分類されているかグラフで可視化した。

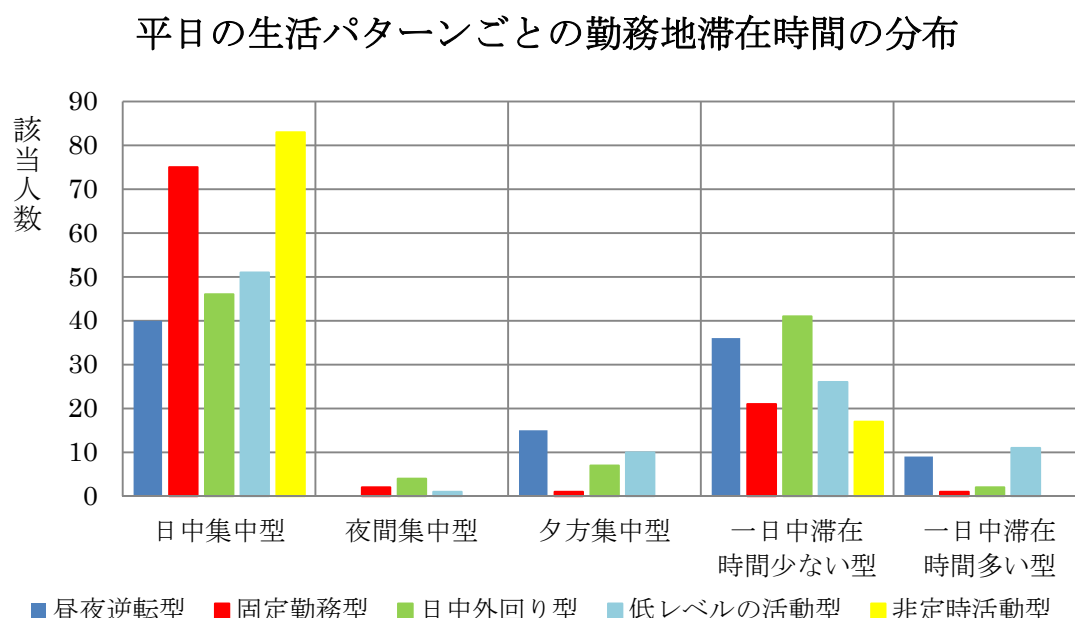


図 4-36：平日の生活パターンごとの勤務地滞在時間の分布

勤務地は日中によく滞在するステイポイントの中から推定される。そのため、日中に勤務地滞在時間が集中する。しかし、『昼夜逆転型』のユーザも本来、日中にはあまり移動を含む活動を行っていないにもかかわらず、40%のユーザが日中に勤務していることになっている。また、『日中外回り型』という日中に活動量がある一定上保たれており、移動を含む活動の持続性が高いユーザは、『1 日中勤務地滞在時間が少ない』グループに 41%も属している。このことから『日中外回り型』のユーザは日中に移動を主に繰り返しているため、勤務地を 1 か所と推定すると勤務地推定の精度はよくないのではないかと推察される。

したがって、移動履歴から自宅・勤務地などを生活に密接にかかわるロケーションを推定する場合は、まず、ユーザが夜間に主に活動するユーザなのか、日中活動するユーザなのか、日中に移動を繰り返すユーザなのかといった、生活パターンを考慮してから行わなければならない。

4.6. 生活パターンに関する考察

第 4 章では、移動履歴収集システムの特性から、位置情報測定頻度の相対的な増減がユーザの移動を含む活動の有無を表していると考えた。そこで位置情報測定頻度を活動量と定義した。

4.1 節では、各時間帯の活動量が、その時間帯における時系列なユーザの活動を表していることが『平成 23 年度社会生活基本帳』（2012）²⁶ から明らかにすることができた。

活動量を用いた平日・休日ごとの生活パターンの分類に、既存の活動日記付きの移動履歴などを正解データとして学習された「日中は勤務し、オフィスにいて夜間在宅する」というルールに基づいた自宅・勤務地推定アルゴリズムを適用し、実際の平日の在宅時間や勤務地滞在時間が 1 日でどのように分布しているか検討した。

結果、『昼夜逆転型』では在宅時間が 1 日中平均的に低くなっており、時間帯に対して明確な差が出ていない。その原因の一つには、『昼夜逆転型』は自宅の場所の推定精度が悪いことがあると考えられる。

同時に、勤務先滞在時間帯別分布の分析では、『昼夜逆転型』は日中に移動を含む活動をあまり行っていないのにもかかわらず、日中にいる滞在場所が勤務先と推定されていることを示した。また、『日中外回り型』は、日中に移動を行っていることが多いために、勤務先を 1 か所と仮定すると推定アルゴリズムの精度が落ちることを示した。移動履歴から自宅・勤務地などを生活に密接にかかわるロケーションを推定には、まず、ユーザの活動の分布が夜間なのか日中なのかといった生活パターンの分類情報が必要である。

第 5 章 活動圏に着目した年間生活パターン分類の特徴分析

第 5 章では，第 4 章にて平均活動量の分布を用いて分類された生活パターンを，ユーザが訪問した活動圏の広さおよび滞留したロケーションの散らばりからみた特徴を述べる．ここで言う活動圏とは一定期間にユーザが訪問あるいは通過した空間範囲を意味している．

5.1. 活動圏による生活パターン分析

本節では，ユーザが訪問した、あるいは通過した空間範囲を活動圏と呼び、その広さを標準地域メッシュで表現している．活動圏の算出には，第 3 章の 5 節で作成した，一分ごとに位置情報測位点を内挿した移動履歴データを用いている．内挿した移動履歴データを用いることにより，位置情報測位点がなくともユーザが通過したであろう標準地域メッシュを拾い上げることができ，実際の活動圏に近い活動圏を算出することができる．

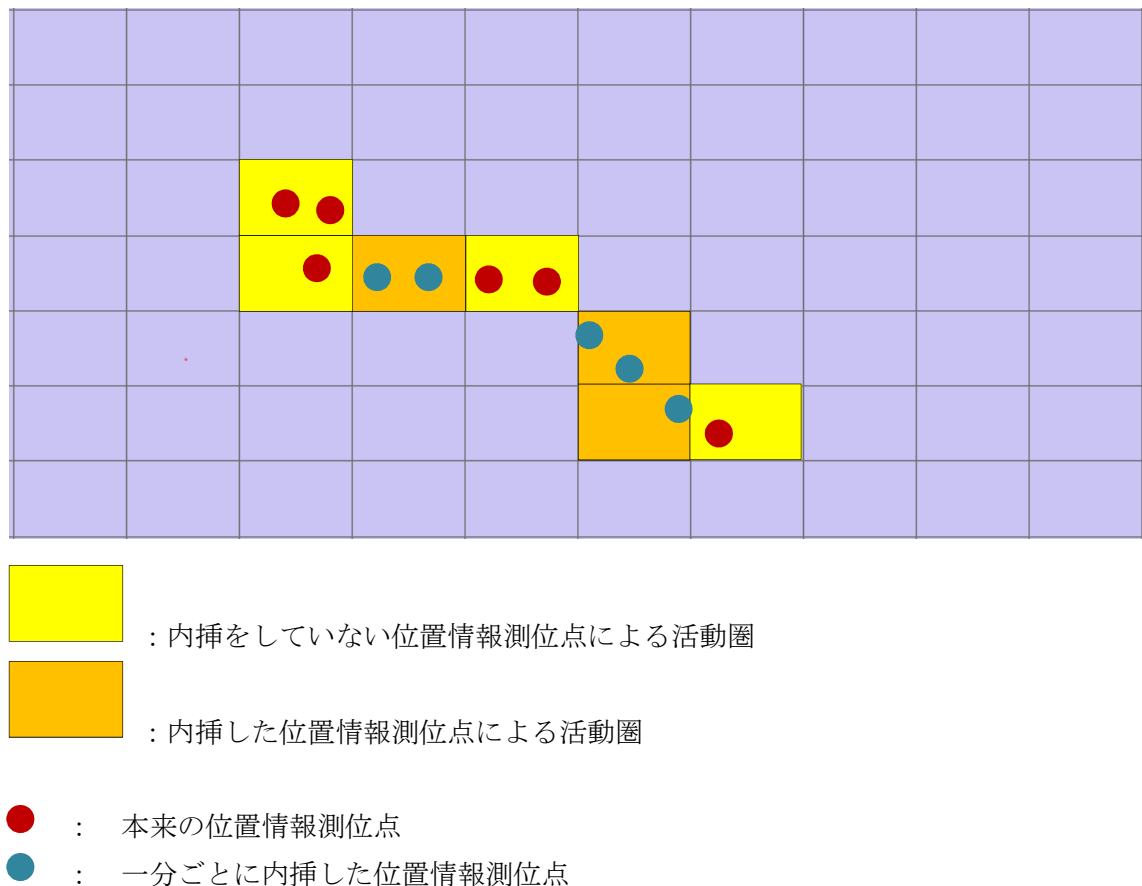


図 5-1：内挿した移動履歴データからの活動圏算出の事例

活動圏は、平日と休日ごとに算出される。第4章で分類した個人の生活パターンごとに、1年間の活動圏の広さを集計し、日数で除して平均化した。内挿した位置情報測位点から算出しているため、この活動圏の広さは、単に面積を示しているだけではなく、活動圏が空間的に広い範囲に及んでいるのかということを示す。これを「1日の平均活動圏」と呼ぶこととする。そして、生活パターンが活動圏から見るとどのような特徴があるのか、さらに同じパターンの中にも、どのようなバリエーションがあるのか分析した。

なお、元の移動履歴データに1分ごとに内挿を行っているため、各ユーザは1日に1440万点の位置情報測位点を持っている。したがって、活動圏が広いユーザは、各メッシュの訪問・通過回数は必然的に少なく（薄く）なり、活動圏が狭いユーザはより少ないメッシュを何度も訪問・通過・滞留していると解釈できる。また、生活パターン間の差の有意差を検定する手法としては、分散分析を用いている。

5.1.1. 活動圏からみた平日の生活パターン

最初に、平日の生活パターンについて分析する。まず、平日の平均活動量を用いてバートレットの検定法を行い、各クラスタ群の母分散が等しいかどうか、つまり各生活パターン群の等分散性を検討する。

ただし、このバートレットの検定法は各クラスタ群内のデータ数が2個以上あることが前提となっている。平日のクラス3はデータ数が1つであるので、生活パターンのクラス3を除いて、バートレットの検定法を行っている。バートレットの検定法の結果より、p値が $2.2e-16$ となり、帰無仮説である「各群の母分散が等しい」が棄却される。各生活パターン群の母分散に差があることが分かる。そこで、等分散を仮定しない一元配置の分散分析を行った。

(バートレットの検定法による結果)

Bartlett's K-squared = 993.0835, df = 9, p-value < $2.2e-16$

(等分散を仮定しない一元配置の分散分析の結果)

F = 81.3848, num df = 9.000, denom df = 200.619, p-value < $2.2e-16$

図 5-2：平日の生活パターンの分析結果

p値が有意水準5%以下となったため、生活パターンごとに一日の平均活動圏の広さに有意差があることが分かる。

次に、第 4 章で大別した平日の生活パターンごとの一日の平均活動圏の散らばりを可視化した。

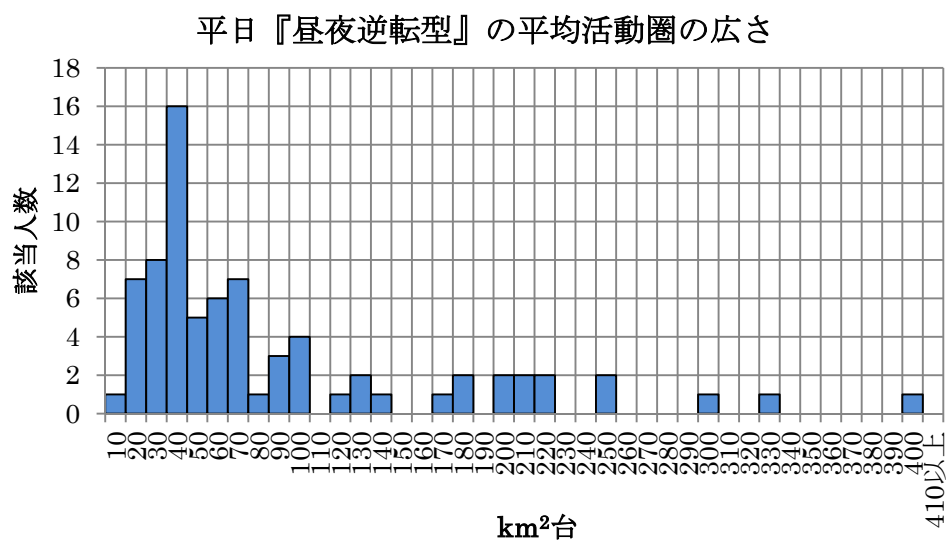


図 5-3：平日『昼夜逆転型』の平均活動圏の広さ

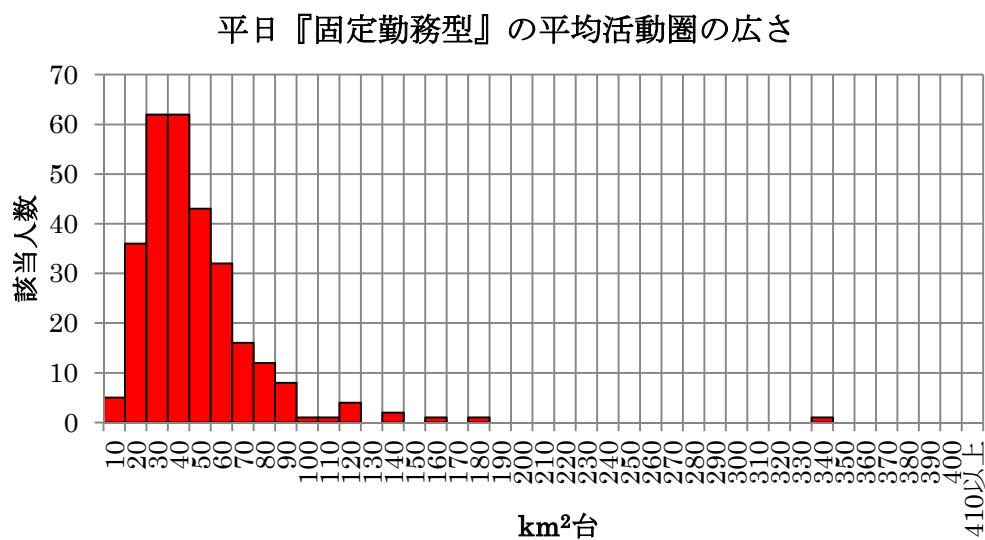


図 5-4：平日『固定勤務型』の平均活動圏の広さ

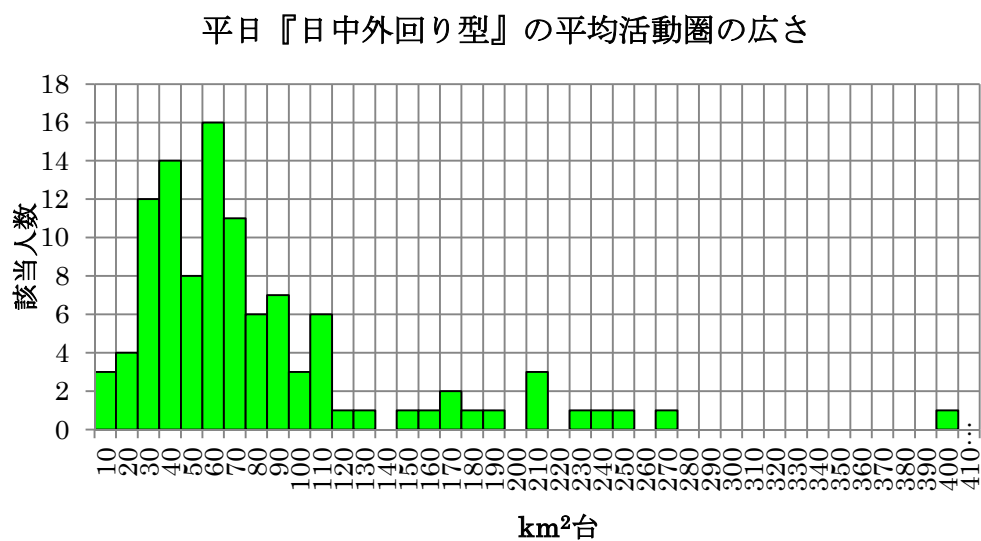


図 5-5：平日『日中外回り型』の平均活動圏の広さ

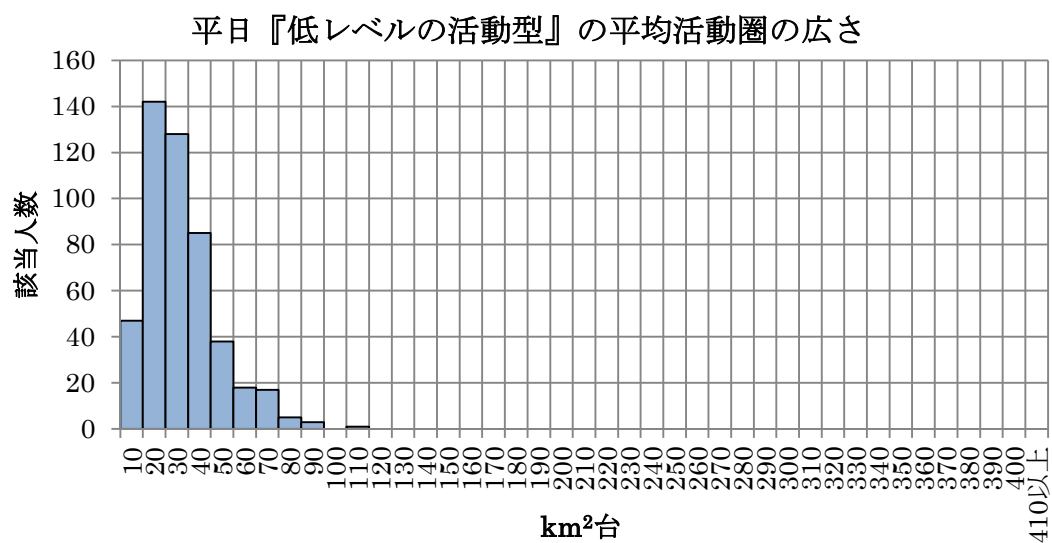


図 5-6：平日『低レベルの活動型』の平均活動圏の広さ

平日『非定時活動型』の平均活動圏の広さ

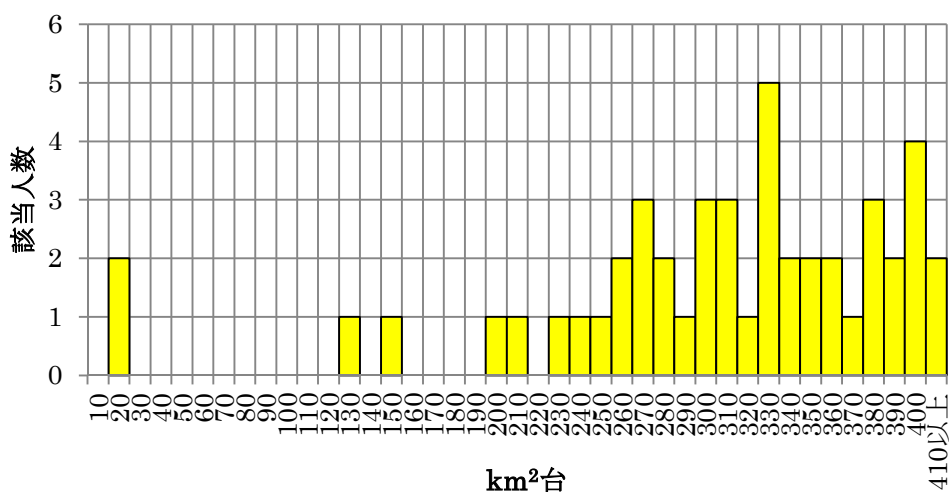


図 5-7：平日『非定時活動型』の平均活動圏の広さ

表 5-1：平日の生活パターンごとの平均活動圏の中央値

	昼夜逆転型	固定勤務型	日中外回り型	低レベルの活動型	非定時活動型
中央値	51.92	37.42	56.30	24.26	311.22

『昼夜逆転型』は 40km² 台を中心に分布しているが、100km² 以上にも約 30%のユーザが分布している。『昼夜逆転型』に属するユーザは平均活動圏の空間的広さが分散していることが分かる。『昼夜逆転型』には夜間の移動距離が短いユーザと長いユーザが混在している。

『固定活動型』は 30km² 台と 40km² 台を中心に分布しており、100km² 以上の活動圏をもつユーザは 4%ほどであり、ほとんどいない。したがって、『固定活動型』には、8 時を中心とした通勤時間帯に移動できる空間的広さをもつユーザが分類されていることが分かる。

『日中外回り型』は、30km² 台と 60km² 台の 2 つを中心に分布しているが、100km² 台以上にも 24%のユーザが分布しており、分布の裾が広いことが分かる。『日中外回り型』には、『固定活動型』のユーザより広い活動圏を持つユーザが分類されている。『日中外回り型』は『固定活動型』と同様に 8 時ごろから出勤（移動）するが、通勤した後も勤務地などにか所留に留まることなく日中も移動を行っていることから、『固定活動型』より広い活動圏を持つと考えられる。

『低レベルの活動型』は 20km² および 30km² を中心に分布しており、一日の平均活動圏は他の生活パターンに比べて極めて狭いことが分かる。したがって、『低レベルの活動型』は活動量も他の生活パターンと比べて低いことから、平日の移動する総距離が短いユーザが分類されていることが分かる。

『非定時活動型』は 100 km² 台以上の平均活動圏を持つユーザが大半を占めており、他の生活パターンに属するユーザに比べて、非常に広い活動圏を持っていると考えられる。1 日中、平均活動量が他の生活パターンより常に多く保たれているので、一日の行動に移動が多く含まれている生活が反映した結果となっている。

5.1.2. 活動圏からみた休日の生活パターン

次に、休日の生活パターンごとに活動圏の空間的広さがどのように違っているのか分析した。まず、同様にバートレットの検定法にて各生活パターン群の等分散性を検討した結果、p 値が 2.2e-16 となり、帰無仮説が棄却される。各パターン群の母分散に差があることが分かる。

したがって、各パターン群の等分散性を仮定しない一元配置の分散分析を行った。結果として、p 値が 2.2e-16 となり、各生活パターンの活動圏の空間的広さに有意差があることが分かる。

(バートレット法の結果)
Bartlett's K-squared = 976.8917, df = 7, p-value < 2.2e-16
(一元配置の分散分析結果)
F = 76.0134, num df = 7.000, denom df = 181.652, p-value < 2.2e-16

図 5-8：休日の生活パターンの分析結果

以下に、休日の生活パターンごとの空間的広さを可視化した。

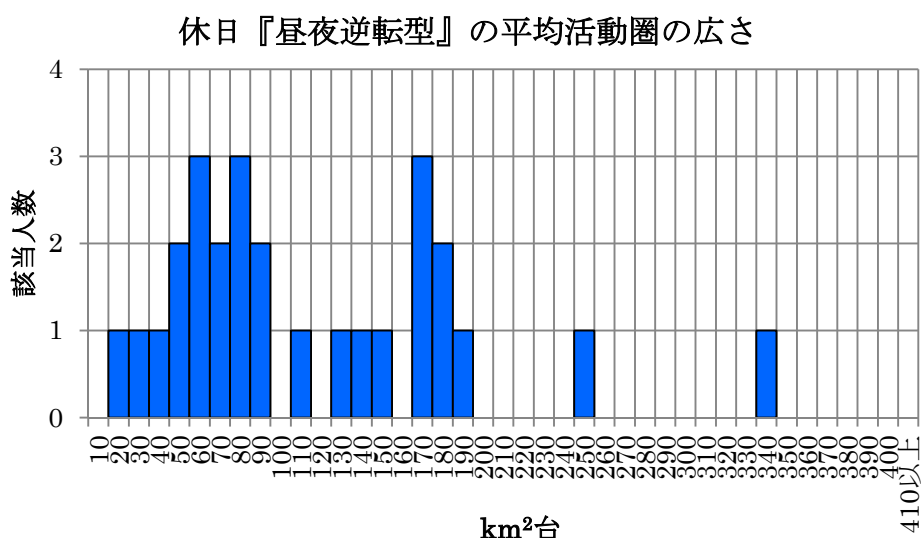


図 5-9：休日『昼夜逆転型』の平均活動圏の広さ

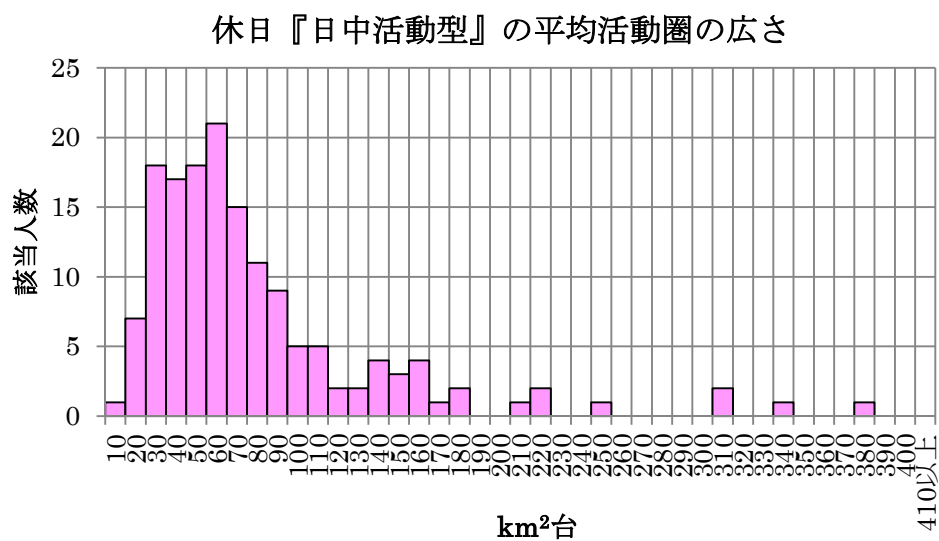


図 5-10：休日『日中活動型』の平均活動圏の広さ

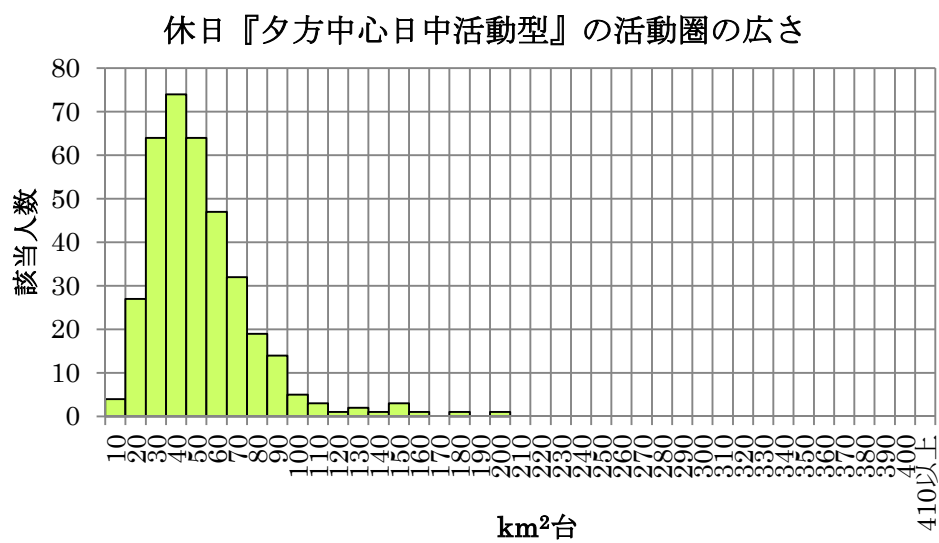


図 5-11：休日『夕方中心日中活動型』の活動圏の広さ

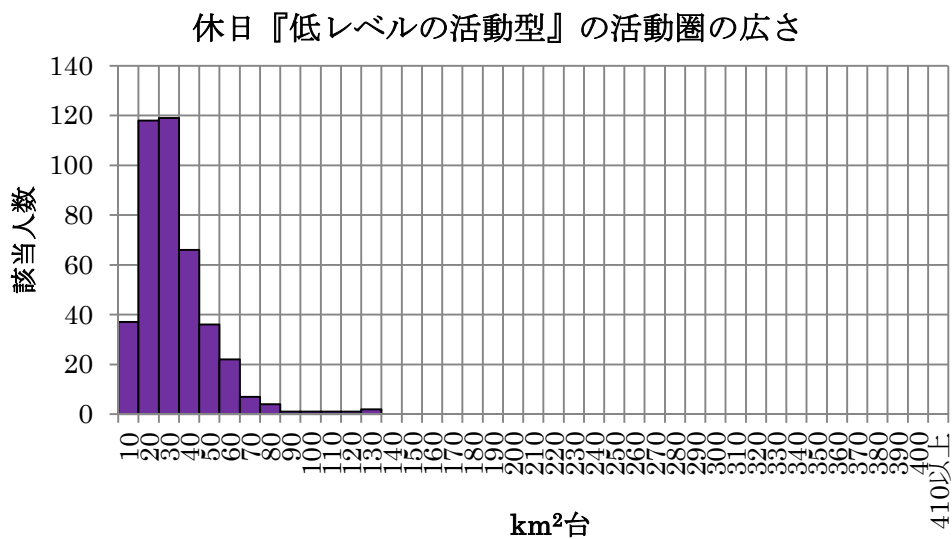


図 5-12：休日『低レベルの活動型』の活動圏の広さ

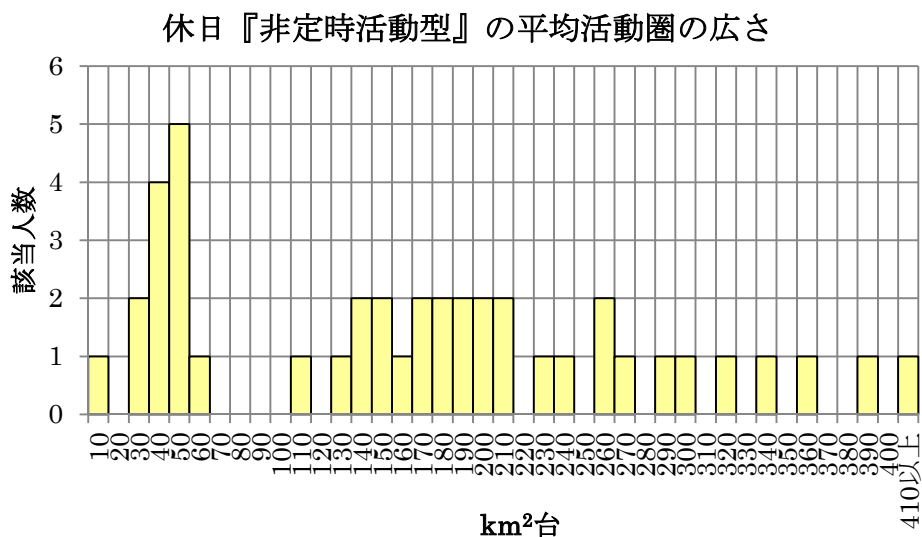


図 5-13：休日『非定時活動型』の平均活動圏の広さ

表 5-2：休日の生活パターンごとの平均活動圏の中央値

	昼夜逆転型	日中活動型	夕方中心日中活動型	低レベルの活動型	非定時活動型
中央値	80.9	57.69	42.24	23.81	167.21

『昼夜逆転型』は、60km²から80km²と170 km²の二つの分布の中心があり、100 km²以下の平均活動圏の空間的広さを持つユーザは『昼夜逆転型』の56%いる。100 km²以上

の空間的広さを持つユーザは『昼夜逆転型』の44%いる。したがって、一日の活動圏の空間的広さが広いユーザが『昼夜逆転型』には分類されていることが分かる。また、平日の『昼夜逆転型』の生活パターンであるユーザの空間的広さと比較しても、休日の『昼夜逆転型』の方が、平均的に空間的広さが広く、広範囲のデータ区間にもユーザが比較的多く存在している。

『日中活動型』は30km²から60km²にユーザが集中している。午後から徐々に活動を始め、夕方に活動量の増加のピークを迎え、その後活動があまり見られなくなる『夕方中心日中活動型』の活動圏の空間的広さを比較すると、『日中活動型』の方がより空間的範囲が広いデータ区間にまでユーザが存在している。午前中から移動を含む活動を行っているユーザが『日中活動型』に分類されていると考えられる。『日中活動型』の活動圏の空間的広さが広いユーザは、朝から離れた場所に出かけ、夜や夕方に帰宅するという休日の過ごし方をしていると考えられる。反対に、午後から徐々に活動し、夕方を過ぎると移動を含む活動が減少する『夕方中心日中活動型』には、40km²を中心に分布しており、分布の広さは狭い。したがって、各時間帯における平均活動量及び活動量が持続する時間から、『夕方中心日中活動型』には、あまり休日に遠出をしないユーザが分類されていると考えられる。

同様に『低レベルの活動型』も休日の活動圏の範囲が狭いことから、遠出をあまりしない傾向があるユーザが該当していると考えられる。

『非定時活動型』は、活動圏の空間的広さが狭い範囲から広い範囲まで幅広くユーザが分散していることが分かる。平日の『非定時活動型』は200km²以上にユーザの大半が集中していたが、休日はユーザが分散している。

5.2. 活動圏の空間的広さに関する考察

第4章において、位置情報の測定頻度がユーザの移動を含んだ活動の有無と継続性を表すことから、位置情報測定頻度を活動量と定義した。その一日の時間帯ごとの分布をクラスタリングすることで、何時、何をしているのかという生活パターンを類型化した。

さらに、訪問もしくは通過した標準地域メッシュを活動圏と定義した。第4章の類型化された生活パターンのクラスに属するユーザが、一日の活動圏にどれほどの広さを持っているのか算出した。

どういった生活パターンをもっているユーザがどのような活動圏を持っているのか、もしくは生活パターンごとにユーザの活動圏の空間的な分布に差が生じ、移動範囲などに特徴が表れるか分析した。

結果として、平日の『固定勤務型』と『低レベルの活動型』は活動圏の広さの分布は単峰となり、分布の裾は非常に狭い。活動圏の広さが違うユーザが混合しておらず、ある一定の広さを持つ活動圏内だけで過ごしているユーザが分類されている。

また、平日の『昼夜逆転型』と休日の『昼夜逆転型』も狭い活動圏を持つユーザと広い

活動圏を持つユーザが混合している。しかしながら、平日の『非定時活動型』と休日の『非定時活動型』を比べると、平日の『非定時活動型』は 200km^2 以上にユーザの大半が集中し分布が偏っているが、休日はユーザが分散している。4.3 節にて平日の非定時活動型は、休日においても非定時活動型に分類されることが多いが、活動圏の広さからは、平日と休日では移動する総距離が異なるユーザがいることが分かる。

以上の第4章および第5章の分析では、カレンダー上の平日と休日で分けて、生活パターンを出している。カレンダー上の平日が就業日、休日が非就業日に当たるとすれば、平日ごと休日ごとに一日の過ごし方が大きく異なる。そのため、土曜日・日曜日・祝祭日が就業日となっているユーザやフレックスタイム性の勤務体系のユーザでは、カレンダー上の平日・休日の枠組みでは生活パターンを正確に理解することができない。平日・休日の生活パターンの比較や分析をするときこれを念頭に置かなくてはならない。

第 6 章 繰返し性に着目した個人の日々の生活パターン分析

第 5 章までは一年間を通じての平均的な活動パターンや活動圏の広さに着目して、各個人がどのように類型化されるのか、分類されるのかを分析した（年間生活パターン分析）。

第 6 章では、個人の一日一日に着目し、個人がどのような「一日」を繰り返しているのかを明らかにすることを試みる。つまり、第 6 章では、各個人に着目し、彼あるいは彼女の一日一日がどのような生活パターンに分類され、それが「いつもの月曜日」「いつもの火曜日」「いつもの水曜日」と 7 パターンの繰返しで 1 年が過ぎているのか否か、もし 7 パターンではないとすると何パターンになるのかをまず分析する。次に、いつも繰り返して立ち寄る場所、テリトリーはどのくらいの広さを有するのかを分析する。これらにより、人の生活の「繰返し性」を時間軸、空間軸の両面から明らかにできる。

こうした繰返し性の推定はサービスの提供などにも役に立つ。例えば、生活の日パターンが個人の仕事・学校がある日（以降、就業日とする。）と仕事・学校がない日（以降、非就業日とする）に関連づけできれば、個人の一日一日の行動文脈がより深く理解でき、休日出勤をする人に、買い物情報を送ったりする過ちを犯さずにすむ。また、誰かがある場所に訪れているとき、その人がその場所のことをよく知っているにも関わらず、初めて訪問したかのような基本的な情報を送ってしまうといった誤りを避けられる。さらにその日が仕事の日であるか否かが分かれば、その訪問目的が「出張」か「旅行」かが区別がつき、出張者向けの案内情報、あるいは観光旅行者向けの案内情報を送り分けることができる。

6.1. 個人の生活における日パターン繰返し性の分析

一日一日の活動量時間帯別分布をクラスタリングし、何パターンの生活パターンの分類できるか検討した。さらに、そのユーザの活動量時間帯の分布に関して、平日および休日に特有なパターンが見つかれば、その日が就業日なのか非就業日なのか判定ができる。

また、分類されるクラス数が少なければ、そのユーザは、活動時間帯の分布の観点から日パターン数が少なく、同一の日パターンを繰り返していると考えられる。同時に、平日と休日に対応する日パターンが明確に分離できれば、そのユーザは平日と休日の生活パターン、つまり過ごし方が明確に異なるといえる。

6.1.1. 活動量を用いた生活の日パターンの抽出に関する予備分析

予備分析として、一年間を対象に、あるユーザ（ユーザ A とする）の一日ごとの活動量の時間帯別分布を第 4 章と同様に k-means 法によってクラスタリングした。クラスタ数はギャップ統計量を 100 回推定して決定している。

ユーザ A については、各クラスを曜日ごとに集計した結果を以下に記す。

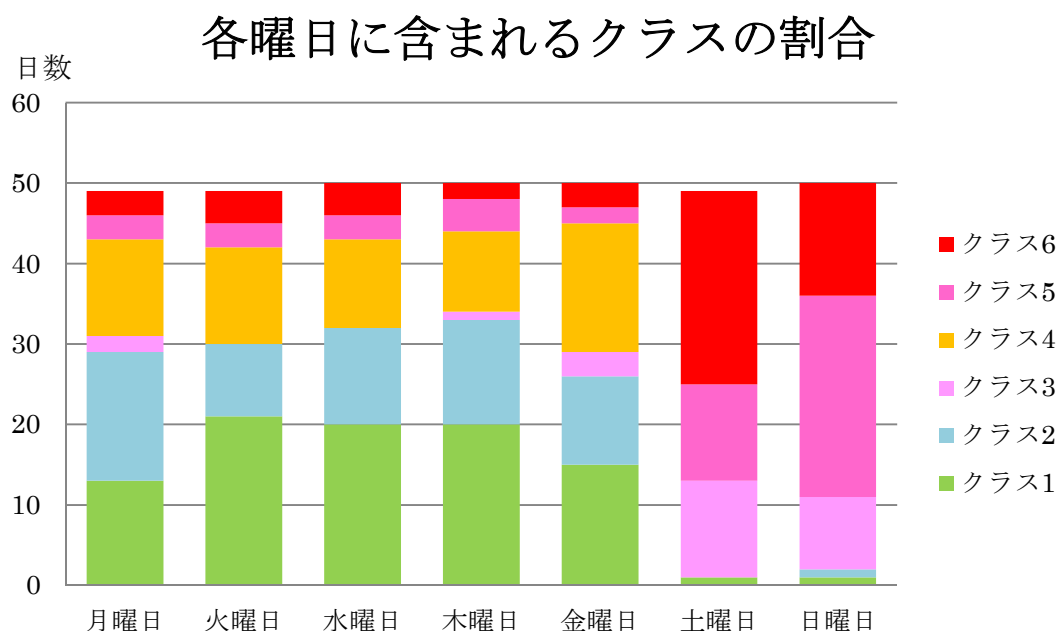


図 6-1：各曜日に含まれるクラスの割合

図 6-1 からクラス 05 とクラス 06 に日曜日が集中していることが分かる。これらの一日がどのクラスに分類されているかというクラスターリング結果を、カレンダーを用いて可視化した。土曜日と日曜日および祝祭日が多く含まれるクラス（クラス 05 とクラス 06 が該当する）と平日が多く含まれるクラス（クラス 01 とクラス 04 が該当する）と平日と休日が両方含まれるクラス（クラス 03 など）が読み取れる。土曜日はクラス 06 に多く分類されていることが分かる。

表 6-1：ユーザ A の 1 日を対象にしたクラスタリング結果

月曜日	火曜日	水曜日	木曜日	金曜日	土曜日	日曜日
2010-08-16	2010-08-17	2010-08-18	2010-08-19	2010-08-20	2010-08-21	2010-08-22
なし	なし	5	1	4	3	3
2010-08-23	2010-08-24	2010-08-25	2010-08-26	2010-08-27	2010-08-28	2010-08-29
4	1	1	1	1	6	6
2010-08-30	2010-08-31	2010-09-01	2010-09-02	2010-09-03	2010-09-04	2010-09-05
2	5	1	1	4	3	5
2010-09-06	2010-09-07	2010-09-08	2010-09-09	2010-09-10	2010-09-11	2010-09-12
4	1	1	5	5	5	5
2010-09-13	2010-09-14	2010-09-15	2010-09-16	2010-09-17	2010-09-18	2010-09-19
1	1	1	1	4	6	6
2010-09-20	2010-09-21	2010-09-22	2010-09-23	2010-09-24	2010-09-25	2010-09-26
6	2	4	5	4	5	5
2010-09-27	2010-09-28	2010-09-29	2010-09-30	2010-10-01	2010-10-02	2010-10-03
2	1	1	1	5	5	5
2010-10-04	2010-10-05	2010-10-06	2010-10-07	2010-10-08	2010-10-09	2010-10-10
4	4	4	1	4	3	3
2010-10-11	2010-10-12	2010-10-13	2010-10-14	2010-10-15	2010-10-16	2010-10-17
3	2	2	2	2	6	6
2010-10-18	2010-10-19	2010-10-20	2010-10-21	2010-10-22	2010-10-23	2010-10-24
1	2	1	2	1	6	6
2010-10-25	2010-10-26	2010-10-27	2010-10-28	2010-10-29	2010-10-30	2010-10-31
2	1	4	1	4	3	3
2010-11-01	2010-11-02	2010-11-03	2010-11-04	2010-11-05	2010-11-06	2010-11-07
3	6	6	4	2	3	3
2010-11-08	2010-11-09	2010-11-10	2010-11-11	2010-11-12	2010-11-13	2010-11-14
2	4	2	4	1	3	5
2010-11-15	2010-11-16	2010-11-17	2010-11-18	2010-11-19	2010-11-20	2010-11-21
4	4	1	2	4	3	6
2010-11-22	2010-11-23	2010-11-24	2010-11-25	2010-11-26	2010-11-27	2010-11-28
2	6	2	2	2	3	5
2010-11-29	2010-11-30	2010-12-01	2010-12-02	2010-12-03	2010-12-04	2010-12-05
1	1	6	4	1	6	6
2010-12-06	2010-12-07	2010-12-08	2010-12-09	2010-12-10	2010-12-11	2010-12-12
2	1	4	2	1	3	3
2010-12-13	2010-12-14	2010-12-15	2010-12-16	2010-12-17	2010-12-18	2010-12-19
4	1	4	1	1	6	5
2010-12-20	2010-12-21	2010-12-22	2010-12-23	2010-12-24	2010-12-25	2010-12-26
2	1	1	3	2	5	1
2010-12-27	2010-12-28	2010-12-29	2010-12-30	2010-12-31	2011-01-01	2011-01-02
1	1	5	5	3	3	3
2011-01-03	2011-01-04	2011-01-05	2011-01-06	2011-01-07	2011-01-08	2011-01-09
5	1	1	1	2	6	5
2011-01-10	2011-01-11	2011-01-12	2011-01-13	2011-01-14	2011-01-15	2011-01-16
5	2	2	2	2	3	3
2011-01-17	2011-01-18	2011-01-19	2011-01-20	2011-01-21	2011-01-22	2011-01-23
2	1	1	4	4	6	5
2011-01-24	2011-01-25	2011-01-26	2011-01-27	2011-01-28	2011-01-29	2011-01-30
1	1	1	4	1	6	5
2011-01-31	2011-02-01	2011-02-02	2011-02-03	2011-02-04	2011-02-05	2011-02-06
4	2	2	2	6	6	2
2011-02-07	2011-02-08	2011-02-09	2011-02-10	2011-02-11	2011-02-12	2011-02-13
2	5	4	4	3	3	3
2011-02-14	2011-02-15	2011-02-16	2011-02-17	2011-02-18	2011-02-19	2011-02-20
1	2	2	2	1	6	5
2011-02-21	2011-02-22	2011-02-23	2011-02-24	2011-02-25	2011-02-26	2011-02-27
1	2	1	4	1	6	5

月曜日	火曜日	水曜日	木曜日	金曜日	土曜日	日曜日
2011-02-28	2011-03-01	2011-03-02	2011-03-03	2011-03-04	2011-03-05	2011-03-06
1	1	2	4	1	1	6
2011-03-07	2011-03-08	2011-03-09	2011-03-10	2011-03-11	2011-03-12	2011-03-13
2	1	2	1	6	6	5
2011-03-14	2011-03-15	2011-03-16	2011-03-17	2011-03-18	2011-03-19	2011-03-20
1	4	1	1	4	6	6
2011-03-21	2011-03-22	2011-03-23	2011-03-24	2011-03-25	2011-03-26	2011-03-27
6	5	5	1	2	6	5
2011-03-28	2011-03-29	2011-03-30	2011-03-31	2011-04-01	2011-04-02	2011-04-03
4	4	1	1	1	6	6
2011-04-04	2011-04-05	2011-04-06	2011-04-07	2011-04-08	2011-04-09	2011-04-10
4	4	4	2	4	6	6
2011-04-11	2011-04-12	2011-04-13	2011-04-14	2011-04-15	2011-04-16	2011-04-17
4	2	2	2	2	5	5
2011-04-18	2011-04-19	2011-04-20	2011-04-21	2011-04-22	2011-04-23	2011-04-24
2	4	4	4	2	5	5
2011-04-25	2011-04-26	2011-04-27	2011-04-28	2011-04-29	2011-04-30	2011-05-01
1	2	4	2	3	6	6
2011-05-02	2011-05-03	2011-05-04	2011-05-05	2011-05-06	2011-05-07	2011-05-08
6	6	6	6	2	5	5
2011-05-09	2011-05-10	2011-05-11	2011-05-12	2011-05-13	2011-05-14	2011-05-15
1	4	2	1	1	6	6
2011-05-16	2011-05-17	2011-05-18	2011-05-19	2011-05-20	2011-05-21	2011-05-22
1	1	1	1	4	6	5
2011-05-23	2011-05-24	2011-05-25	2011-05-26	2011-05-27	2011-05-28	2011-05-29
2	4	1	1	1	5	5
2011-05-30	2011-05-31	2011-06-01	2011-06-02	2011-06-03	2011-06-04	2011-06-05
2	1	1	4	1	6	5
2011-06-06	2011-06-07	2011-06-08	2011-06-09	2011-06-10	2011-06-11	2011-06-12
4	6	2	1	2	5	5
2011-06-13	2011-06-14	2011-06-15	2011-06-16	2011-06-17	2011-06-18	2011-06-19
2	1	4	1	4	6	6
2011-06-20	2011-06-21	2011-06-22	2011-06-23	2011-06-24	2011-06-25	2011-06-26
1	1	1	2	4	5	5
2011-06-27	2011-06-28	2011-06-29	2011-06-30	2011-07-01	2011-07-02	2011-07-03
2	1	6	6	6	6	3
2011-07-04	2011-07-05	2011-07-06	2011-07-07	2011-07-08	2011-07-09	2011-07-10
4	4	2	2	4	5	5
2011-07-11	2011-07-12	2011-07-13	2011-07-14	2011-07-15	2011-07-16	2011-07-17
4	1	1	1	4	5	5
2011-07-18	2011-07-19	2011-07-20	2011-07-21	2011-07-22	2011-07-23	2011-07-24
5	4	1	5	1	なし	6
2011-07-25	2011-07-26	2011-07-27	2011-07-28	2011-07-29	2011-07-30	2011-07-31
2	4	4	1	4	6	5

一日一日の活動量の時間帯別分布をクラスタリングした結果、土曜日・日曜日及びゴールデンウィークなどの祝祭日はクラス 5 とクラス 6 が集中していることが分かる。この結果から、一日の活動量の時間帯別分布のクラスタリングから、平日に特徴的な時間帯ごとの活動量分布と休日に特徴的な時間帯ごとの活動量分布に分けられるユーザがいることが分かる。

6.1.2. 活動量を用いた生活の日パターンの抽出

そこで、2.1 節で述べた全国都道府県ごとに 100 人を無作為抽出（ただし東京都、大阪府などはその各都道府県のユーザ数に応じて 100 人以上を無作為抽出した）した 5020 人のユーザを対象として、ユーザの一日の活動量の時間帯別分布パターンがいくつかのクラスタ数に分かれるかギャップ統計量を用いて明らかにした。計算資源の関係から 20 回のギャップ統計量の推定を行い、最頻度のクラスタ数を最適クラスタ数として算出した。この最適クラスは日パターンとして考えられる。以下の図 6-2 は 5020 人のユーザにおいて日パターン数の人数の分布である。

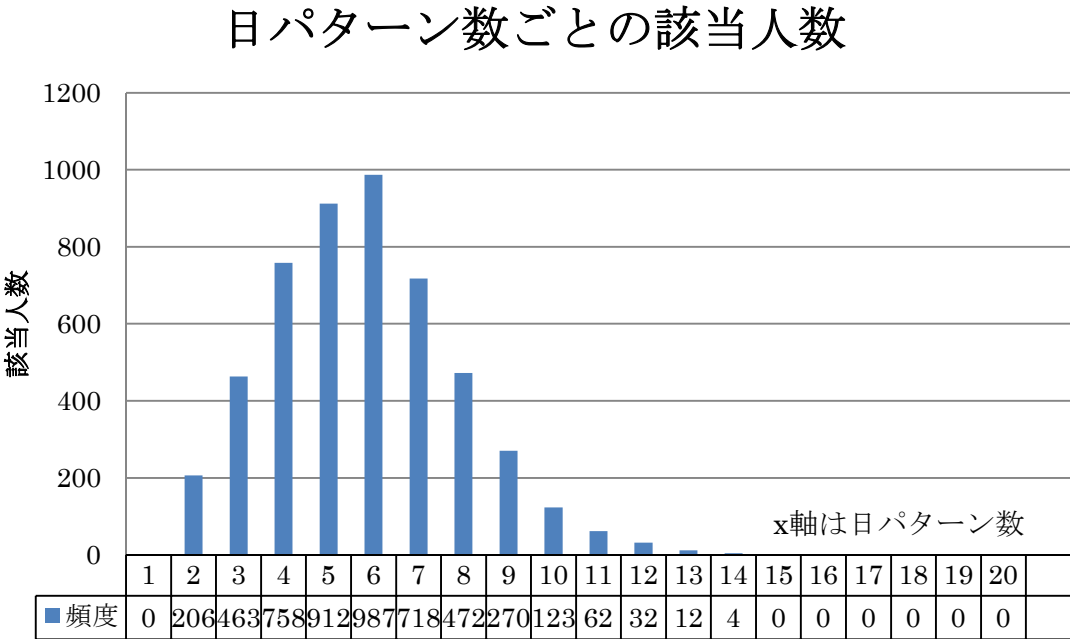


図 6-2：日パターン数ごとの該当人数

図 6-2 より、活動量の観点から見た生活の日パターンは 6 パターンに分かれるユーザが 1000 人近くおり、次に 5 パターン、4 パターンとなっている。活動量による日パターンの数は 10 以上になることはそれほど多くないことが分かる。つまり、活動の時間帯別分布からみた人の一日の過ごし方は、7、6、5、4 パターンで説明できることが分かる。

次に、図 6-2 に記された各個人の最適クラスタ数に基づいて、一日ごとの活動量の時間

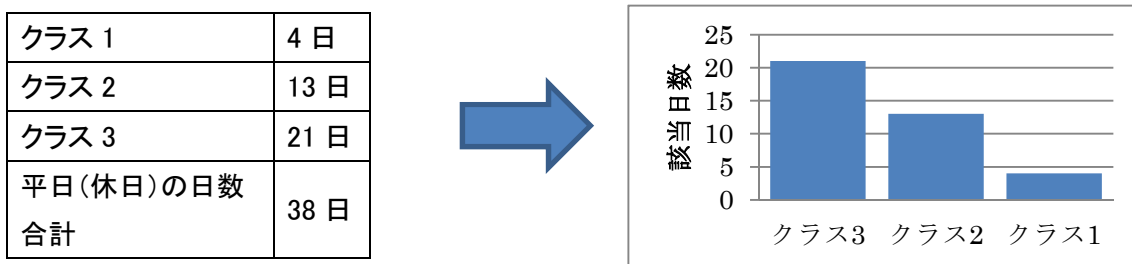
帯別分布を k-means 法によってクラスタリングを行った。

平日と休日が何個のクラスによって表すことができるか、可視化した。平日と休日をそれぞれが何個のクラスによって表すことができるか分かれば、平日の生活パターンが複雑であるが、休日は過ごし方が単調であるユーザが分かる。もしくは平日の生活パターンが少ないユーザであれば、平日においては同じ生活パターンを送っていると分かる。

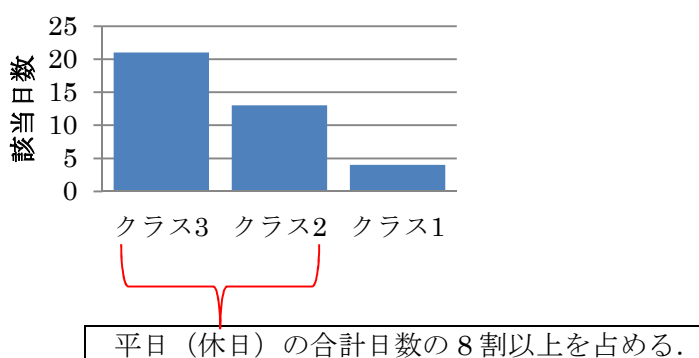
まず、平日の 8 割以上の日数を何個のクラスによって表すことができるか、個人ごとに算出した。図 6-3 に算出方法を記す。

手順

- 1) クラスタリングによって、各クラスに振り分けられた平日（休日）が合計日数を算出し、合計日数が多い順にクラスを整理する。



- 2) 該当日数が多いクラスから累積該当日数を算出する。累積該当日数が平日（休日）の合計日数の 8 割以上になったとき、それらのクラスの個数が、平日（休日）の主な生活パターン数とする。



- 3) 例では、クラス 3 とクラス 2 が平日の主な生活パターンと考えられる。したがって、平日は主に 2 パターンの生活パターンがあると考ええる。

図 6-3：平日の生活パターン数の算出の手順

以上のように、各個人の平日・休日が何パターンの生活パターンで8割以上説明できるか、下の図6-4のように可視化した。

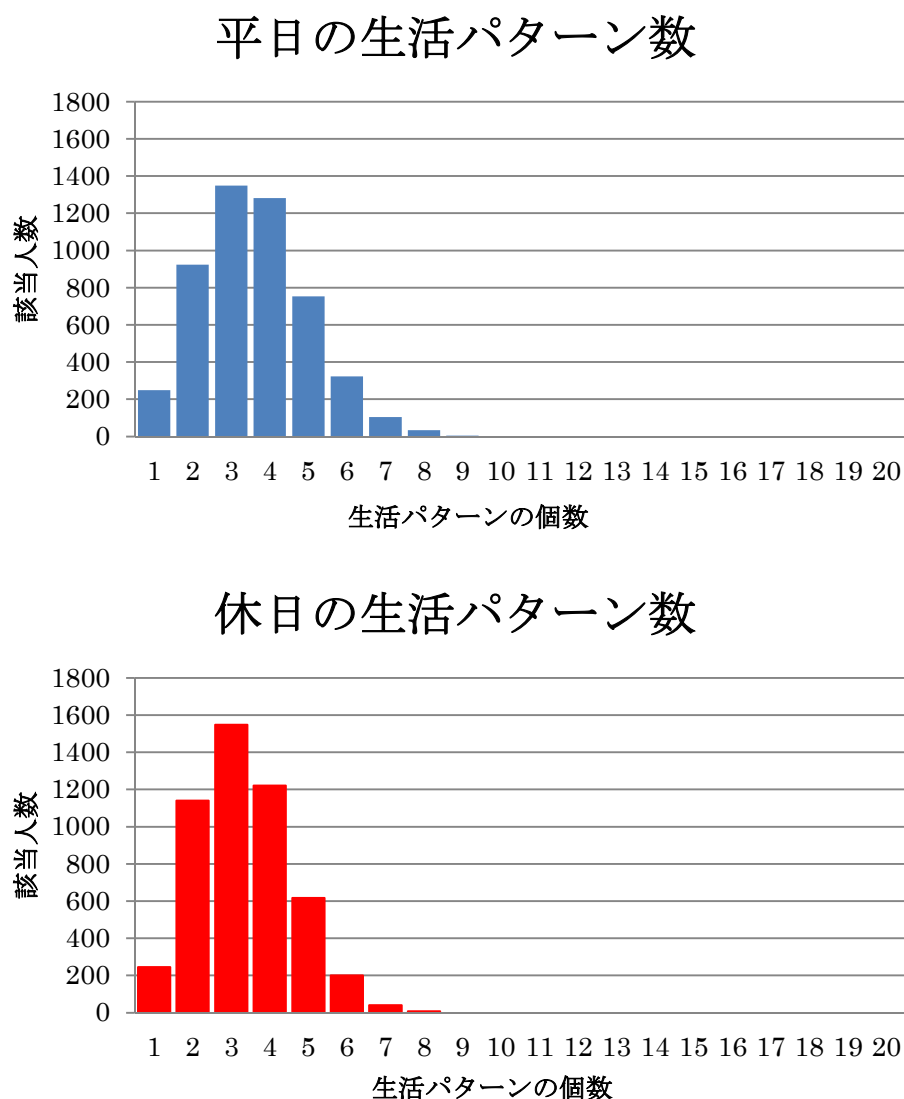


図 6-4：平日と休日別の生活パターン数

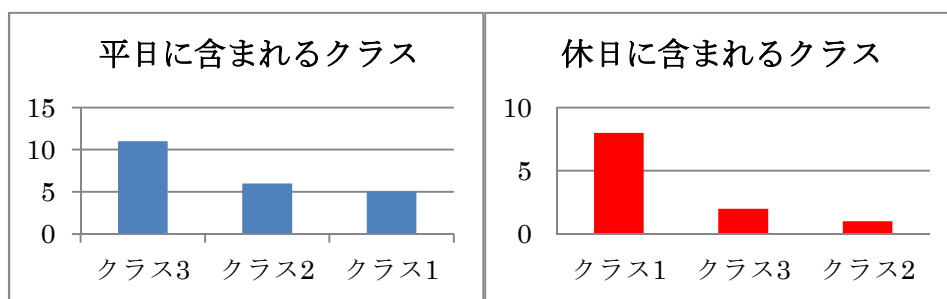
平日は、3、4パターンが最も多い生活パターン数となっている。休日では3パターンが最も多くなっている。平日と休日はともに3、4パターンの過ごし方があることが分かる。

さらに、各個人において、平日と休日がどれほど明確に区別できるか検討した。各個人の一日ごとの活動量の分布のクラスタリングを行い、平日と休日を含むクラスがどれほど一致しているか表すことで、平日と休日の区別がついているか示した。図6-5にその手順を記す。

手順)

1) クラスタリングによって、各クラスに振り分けられた平日（休日）が合計日数を算出し、合計日数が多い順にクラスを整理する。

	平日	休日
クラス 1	5 日	8 日
クラス 2	6 日	1 日
クラス 3	11 日	2 日



2) 平日と休日ごとに、該当日数が多いクラスから累積該当日数を算出する。累積該当日数が平日（休日）の合計日数の 6 割以上になったとき、そのクラスが平日および休日を代表するクラスと考える。

例においては、平日は「クラス 3」と「クラス 2」である。休日は「クラス 1」である。

3) 平日と休日のそれぞれの代表クラスで一致しているクラス数 (i) を、平日の代表クラス数 (m) と休日の代表クラス数 (n) の和で割り、一致度を算出する。

$$\text{一致度}(\%) = \frac{i}{m+n} \times 100$$

(一致度の下限値は 0%であり、上限値は 100%である。)

図 6-5 : 平日と休日を占めるクラスの一致の程度を算出する手順

以上の算出方法によると、一致度が 0%であればあるほど、平日と休日のクラスはかぶらないので、平日と休日の過ごし方は明確に区別がつくと考えられる。また、一致度が 1 であれば、平日と休日のクラスが一致しているため、カレンダー上の平日と休日では、活動時間帯の分布からでは、そのユーザの就業日と非就業日における過ごし方の違いがでないことが分かる。

まず、5020 人において一致度を算出した結果を図 6-6 に示す。

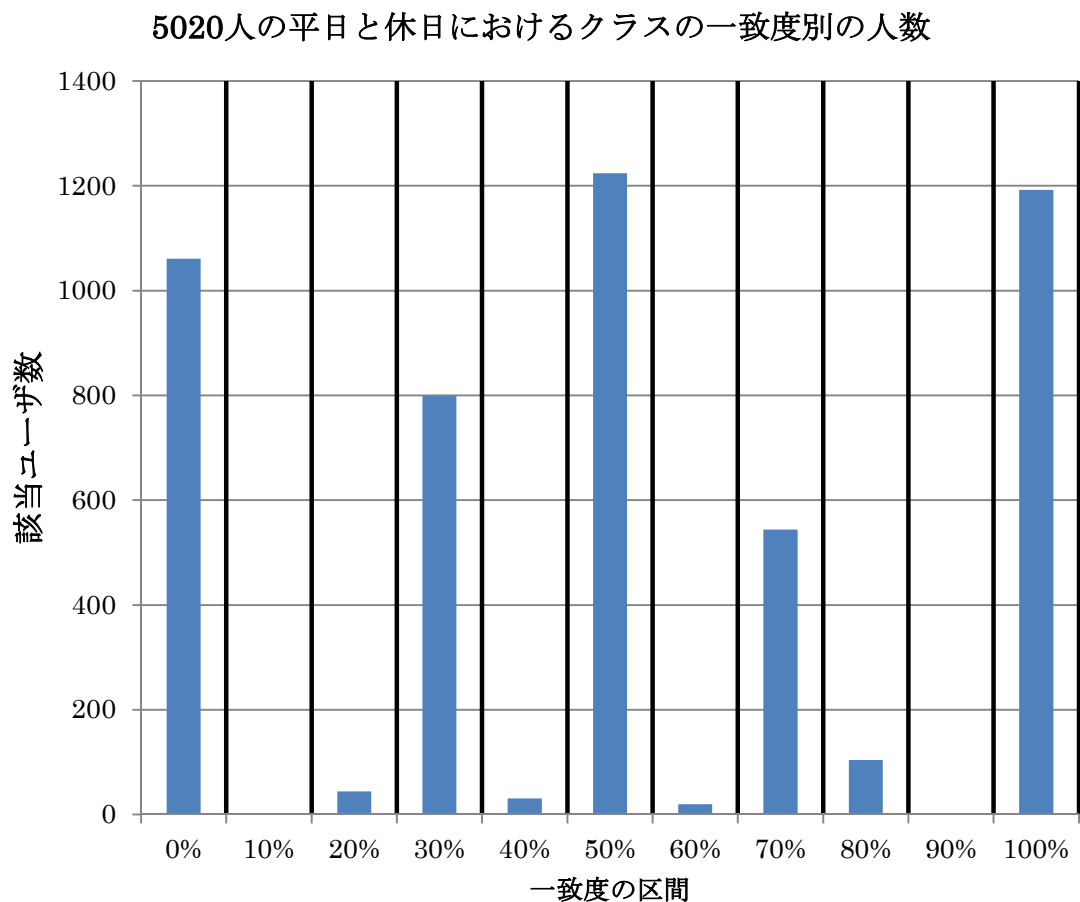


図 6-6 : 5020 人の平日と休日におけるクラス的一致度別の人数

5020 人中，平日のクラスと休日のクラスが一致しなかったのは，1061 人（全体の 21% にあたる）である．平日のクラスと休日のクラスがすべて一致したのは，1192 人（全体の 24% にあたる）であった．活動量のクラスタリングを用いて，一日一日の活動量の時間帯別分布の類似度を評価した結果とでは，平日と休日で活動の時間帯別分布が全く異なるユーザは 5020 人のうち 21% のみである．反対に 5020 人のうち 24% のユーザが平日と休日で活動の時間帯別分布が一致する．このように個々の人々の生活パターン（昼夜逆転型など）を考慮せずに，平日と休日の分離状況を検討すると，完全分離から完全一致まで幅広く分布してしまうことがわかる．

6.2. 滞在地点における個人の日々の生活パターンの繰返し性

6.1 節では，活動量の分布からみて一日一日の生活パターンがどの程度定常的なのかを観察した．6.2 節では，3.3 節にて述べた Apichon et al.(2010)²⁷ が算出したステイポイントをを用いて，どこに滞在するのかに着目して，個人の生活パターンの繰返し性を分析した．

ステイポイントとは、一定範囲内に滞留していると考えられる位置情報測位点に付与される属性ラベルであり、ある目的のために一方向性の移動を意味するトリップ間にある位置情報測位点を意味する。そのため、ステイポイント群はユーザが滞留していることを意味すると考えられる。

このステイポイント群の分散が多いほど、そのユーザは様々な場所に滞在していると考えられる。例えば、平日勤務で朝 8 時に出勤し、オフィスでデスクワークをし、帰宅するという生活パターンのユーザは自宅とオフィス（勤務地）での滞在時間が長いと仮定できる。また、輸送業者のユーザは、様々な物流地点へ移動を繰り返すため、各ステイポイント群の滞在時間は分散し、かつステイポイント群の数も多くなると考えることができる。

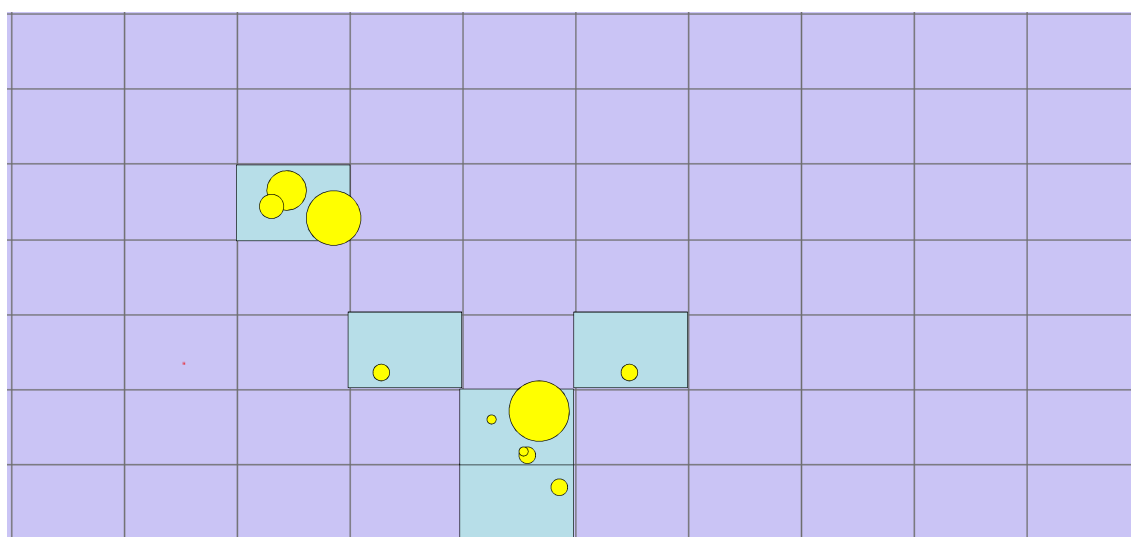
そこで、1 年間わたるステイポイント群を標準地域メッシュ（一辺 1km）ごとに、それらの滞在時間を集計した。なお、ステイポイント群の位置は 3.1 節にて述べたように Apichon et al.(2010)²⁷ が算出しているステイポイントクラスタの重心点をそのステイポイントクラスタの位置として、本研究では使用している。

次に標準地域メッシュの合計滞在時間の降順から滞在時間の累積滞在時間を出し、累積滞在時間が層滞在時間の 8 割を超えた時点で、累積滞在時間に加わった標準地域メッシュを、1 年間において訪問する頻度が高いと考える。

以下にその手順を記す。

手順)

1. 標準地域メッシュごとに、存在するステイポイント群の滞在時間を集計する。



● : ステイポイント群を示し、大きさは滞在時間に比例する。

2. 各メッシュに存在するステイポイント群の合計滞在時間を集計し、その降順に標準地域メッシュを整理する。
3. 降順に累積滞在時間を算出していき、総滞在時間の 8 割を超えた段階で、累積滞在時間にある標準地域メッシュを、1 年間で滞在することが多い地域として考える。

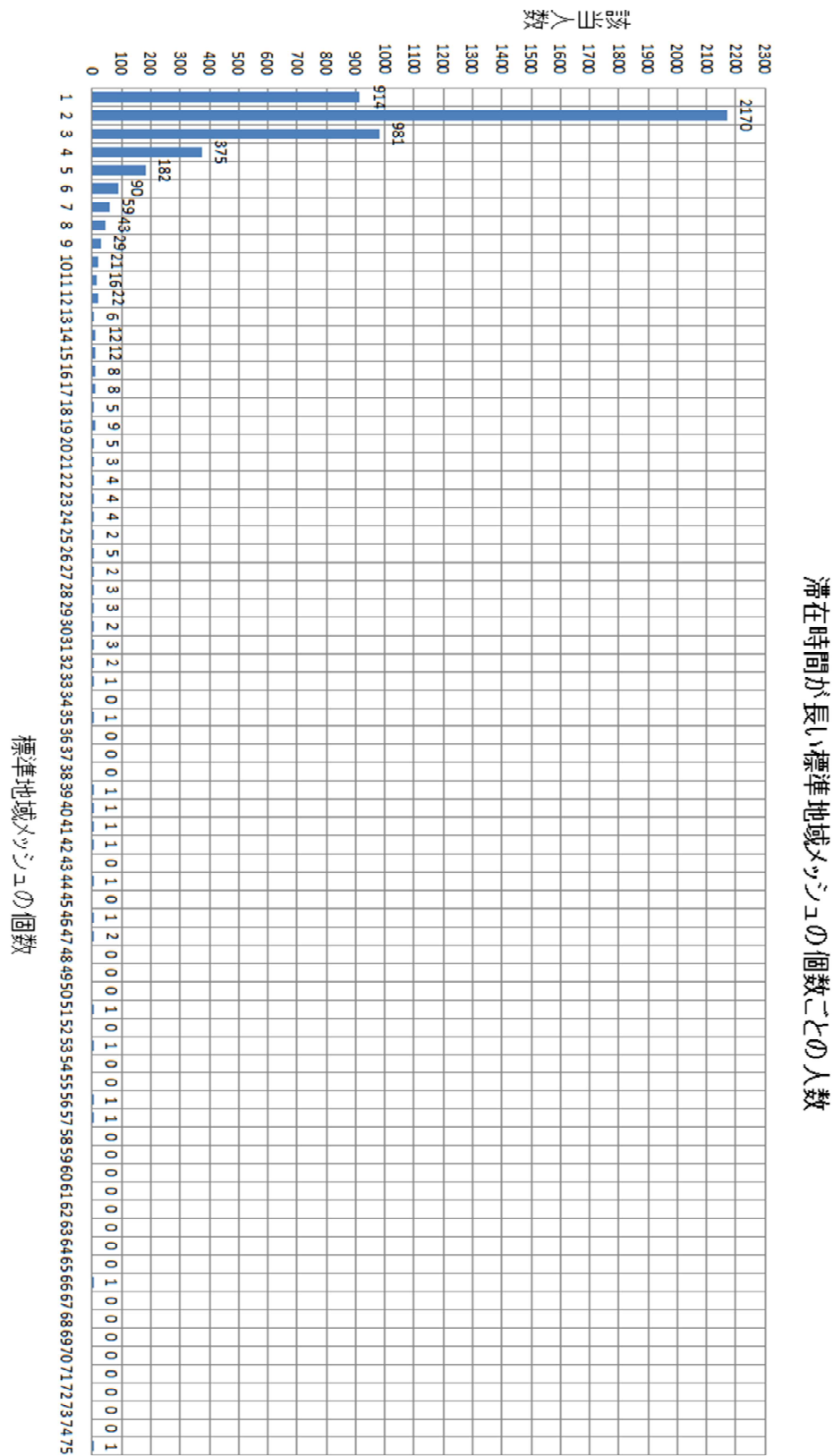
図 6-7：滞在時間の長い標準地域メッシュの算出の手順

図 6-7 のように 1 年間で滞在することの多い標準地域メッシュを算出する。この滞在することの多い標準地域メッシュの数で、滞在地点の分散性を論じる。以下に 5020 人のユーザの滞在することが多い標準地域メッシュの個数ごとの該当人数を図 6-8 にて記す。

滞在時間が総滞在時間の 8 割を超える標準地域メッシュの数が 2 個であるユーザだけで 5020 人の 43% を占める。滞在時間が多い標準地域メッシュの数が 1 個のユーザと 2 個のユーザ、3 個のユーザは全体の 5020 人の 8 割以上を占めている。滞在時間が総滞在時間の 8 割を超える標準地域メッシュの数が 1 個であるユーザは、自宅近辺のみを活動しているユーザ、もしくは滞在時間が短いステイポイント群が分散している、つまり様々なところへ移動しているが自宅で過ごすというユーザの 2 種類が考えられる。滞在時間が総滞在時間の 8 割を超える標準地域メッシュの数が 3 個のユーザに関しても、自宅および勤務地と他の勤務地を持つユーザか自宅および勤務地と他のよく滞在・訪問する場所があるユーザと考えられる。

しかしながら、標準地域メッシュの数が多いユーザは、「在宅時間と勤務地の滞在時間が長い」といった場所の属性による滞在時間の仮定に当てはまらない。滞在地点と滞在時間が分散していることから、移動履歴データ上では、常に移動して短い滞在を繰り返しているユーザとなる。

図 6-8：滞在時間が長い標準地域メッシュの個数における該当人数



6.3. 年間生活パターンごとの繰返し性の比較

6.1 節では各一日一日の活動量の時間帯別分布を k-means 法クラスタリングによって、一年間の日々が何個のクラスで説明できるか分析し、活動量から見た生活パターンの繰返し性を明らかにした。そして 6.2 節では、滞在地点を用いて個人の日々の生活パターンの繰返し性を評価した。

そこで、6.3 節では、第 4 章におけるクラスター分析より明らかになった年間生活パターンごとに、活動量および滞在地点からみた生活パターンの繰返し性を分析する。

6.3.1. 平日の生活パターンごとの活動量を用いた繰返し性の評価

生活パターンごとに、各ユーザの一日の活動量の時間帯別分布に基づいて、一年間の日々が何パターンに分けることができるか図 6-9 に記す

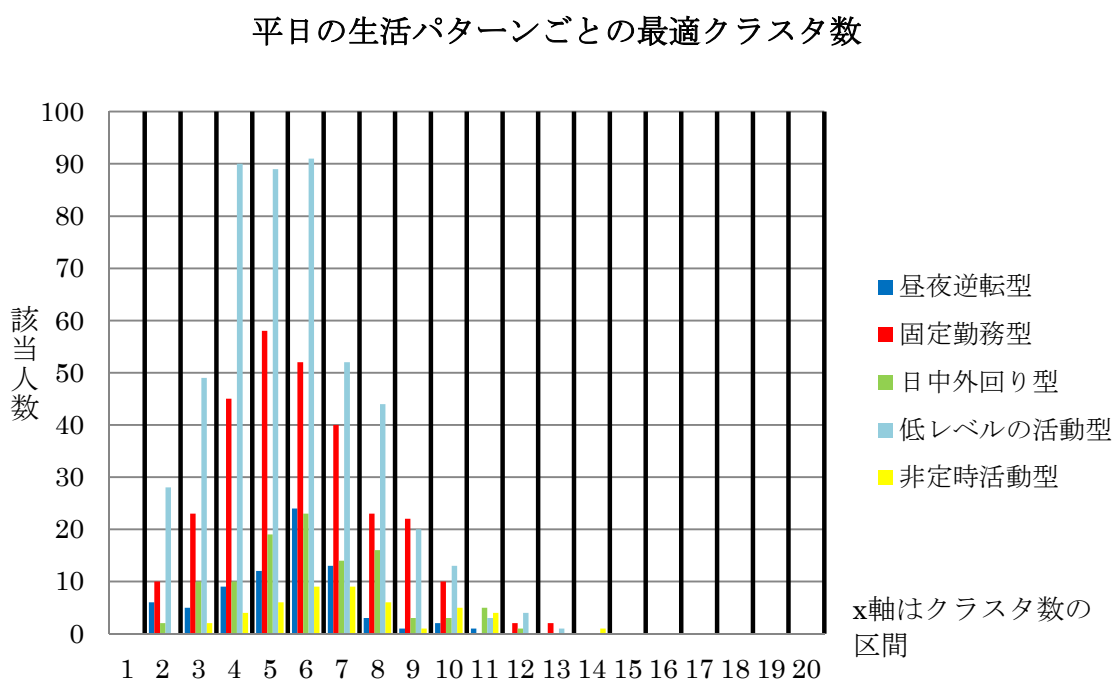


図 6-9：平日の生活パターンごとの最適クラスタ数

低レベルの活動型では、4 から 6 パターンが最も該当人数が多くなっている。低レベルの活動型以外の生活パターングループは、日々の生活パターンが 5 もしくは 6 パターンとなっており、6.1.1 節における 5020 人の結果と等しい。

次に、平日と休日の過ごし方が明確に区別できるか分析した。図 6-10 に生活パターンごとの平日と休日の日パターンが一致するかどうかの指標である一致度に対する該当人数を示した。

平日の生活パターンごとの平日と休日を占めるクラスの一緻度

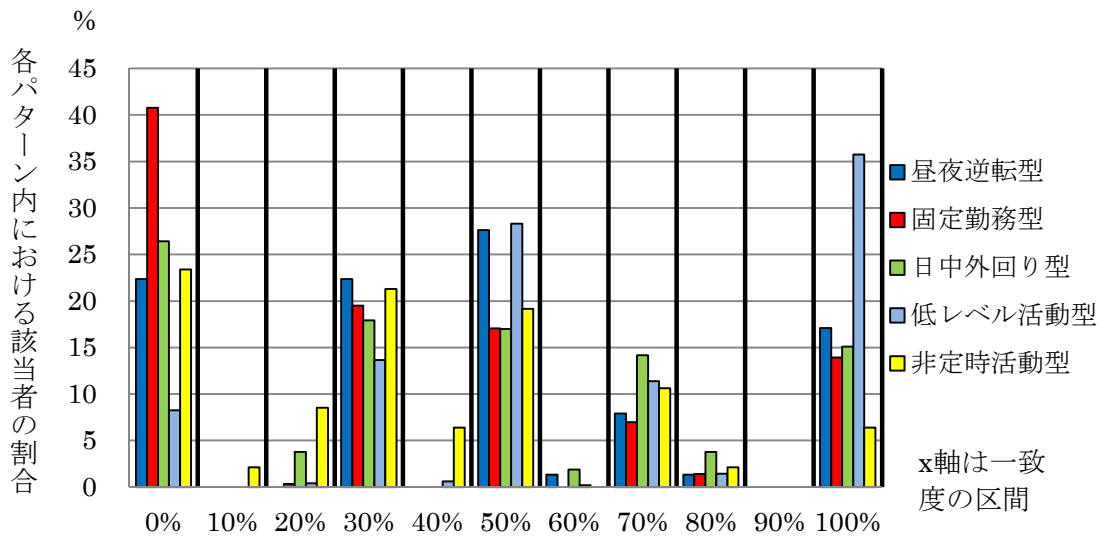


図 6-10：平日の生活パターンごとの平日と休日を占めるクラスの一緻度

図 6-10 から、平日と『固定勤務型』に属するユーザのうち 40%ほどのユーザは平日と休日の過ごし方が全く異なることが分かる。反対に『低レベル活動型』のユーザは活動の時間帯別分布だけでは、平日と休日の区別がつきにくいことが分かる。

このように、就業者および学生など日中に規則的に通勤・通学する場所があるユーザである可能性が高い『固定勤務型』と『日中外回り型』は、平日と休日の活動パターンが比較的是っきり区別される。

また、一日を通して移動および活動が行われている非定時活動型も、平日および休日における活動パターンの一緻度が 30%以下のユーザが非定時活動型グループの約 55%である。4.3 節にて、平日の非定時活動型のユーザの 45%が休日にも移動および活動が一日を通して行われている非定時活動型に分類されていることが明らかになっているが、このグラフから非定時活動型は一日一日の活動量のパターンに着目すれば、平日と休日の活動パターンが異なっている傾向がある。

なお、ここでの休日はカレンダー上の休日であり、実際に各ユーザが休みを取っている日を休日としているわけではない。そのため、非定時活動型ではカレンダー上の平日と休日が十分分離できていない可能性もある。

6.3.2. 平日の生活パターンごとの滞在地点を用いた繰返し性の評価

次に、6.1.2 節と同様に滞在地点の分散性から、平日の生活パターンごとに、日々の生活パターンがどれほど異なっているのか評価した。

図 6-11：滞在時間が長い標準地域メッシュの個数に対する平日の生活パターンごとの該当者数

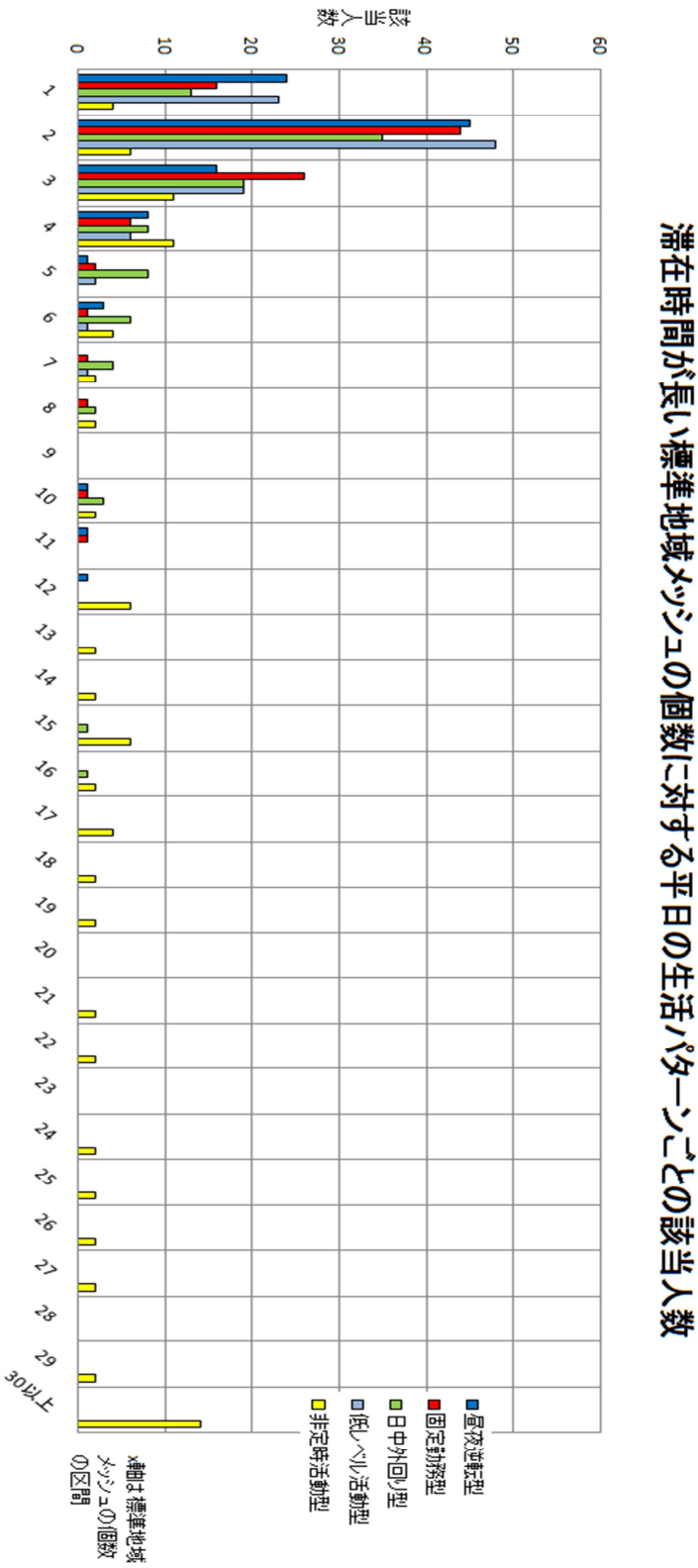


図 6-11 から、低レベルの活動型および固定勤務型、昼夜逆転型、日中外回り型では滞在時間が多い標準地域メッシュ数が 2 個のユーザが各グループ内で最も多いことが分かる。非定時活動型は滞在時間が短い標準地域メッシュ数が多いことから、一年間において移動が多いユーザが分類される傾向があることが分かる。

6.4. 国民生活時間調査による生活パターンの整合性の評価

第 6 章では、個人の一日一日の活動量の時間帯別分布を用いて、一日一日の過ごし方のバリエーションおよびその繰返し性について検討している。

6.1.節では、平日と休日に日パターンを振り分け、平日における過ごし方と休日における過ごし方がどの程度一致しているのか 5020 人と平日の生活パターンごとに検討している。そこで、本節では、NHK 放送文化研究所が作成した『2010 年国民生活時間調査報告書』（2010）²⁵ における職業別に非就業日の取得制度に関する統計から、平日の生活パターンごとに分析された活動量のクラスター分析による平日と休日の過ごし方の違いについて整合性を評価する。

6.4.1. 国民生活時間調査による就業日と非就業日の可視化

図 6-12 は NHK 文化研究所による『2010 年国民生活時間調査報告書』（2010）²⁵ の図表 46「休みの曜日（職業別）」を引用したものである。

(%)	月曜日	火曜日	水曜日	木曜日	金曜日	土曜日	日曜日	決まっていない
有職者	6	5	6	5	4	48	65	31
男有職者	5	4	5	3	3	49	68	28
女有職者	7	6	8	6	5	46	60	35
勤め人	6	5	6	5	4	55	71	26
勤め人以外の有職者	6	4	5	4	2	23	46	47
農林漁業者	1	2	1	1	0	4	12	84
自営業者	8	5	6	4	1	18	50	38
販売職・サービス職	13	11	12	8	8	32	43	42
技能職・作業職	2	3	4	3	2	48	74	26
事務職・技術職	3	2	5	4	2	76	86	16
経営者・管理職	3	2	2	3	1	77	86	9
専門職・自由業・その他	4	5	7	6	5	39	57	43

※有職者は農林業業者、自営業者、販売業者・サービス職、技能職・作業職、事務職・技術職、経営者・管理職、専門職・自由業・その他を指す。

※勤め人は販売職・サービス職、技能職・作業職、事務職・技術職、経営者・管理職を指す。

図 6-12：職業別の休みである曜日

図 6-12 から、職業によって非就業日（図 6-12 では休みの曜日）が異なる傾向があることが分かる。

例えば、農林漁業関係者は非就業日が決まっていない人がその集団の 8 割以上に上り、土曜日・日曜日が非就業日である人は他の集団と比べて極めて少ない。農林漁業関係者は非就業日を不定期にとると考えられる。

自営業者は非就業日が決まっていない人がその集団の 4 割弱ではあるが、土曜日と日曜日が非就業日である人が 18%、50%と、販売職・サービス職を除く勤め人の集団に比べてやや低い、同様に、販売職・サービス職も土曜日と日曜日が非就業日である人が 32%、43%と他の勤め人の集団と比べると低くなっている。専門職・自由業・その他も同様である。

販売職・サービス職を除く勤め人の中で、土曜日が非就業日であるのはその集団の平均 67%であり、日曜日が非就業日であるのは平均 82%である。したがって、販売職・サービス職を除く勤め人は土曜日・日曜日は非就業日だと考えられる。

6.4.2. 平日の生活パターンごとの滞在地点を用いた繰返し性の評価

一日一日の活動量の時間帯別分布を用いて、クラスタリングを行い、平日と休日の過ごし方（生活の日パターン）の一致度を算出した。その一致度が平日の生活パターンごとによって、どのように変化するか 6.2.1 節で分析している。以下の図 6-13 に、平日と休日の過ごし方の一致度ごとに平日の年間生活パターンがどれほどの割合を占めているか可視化した。

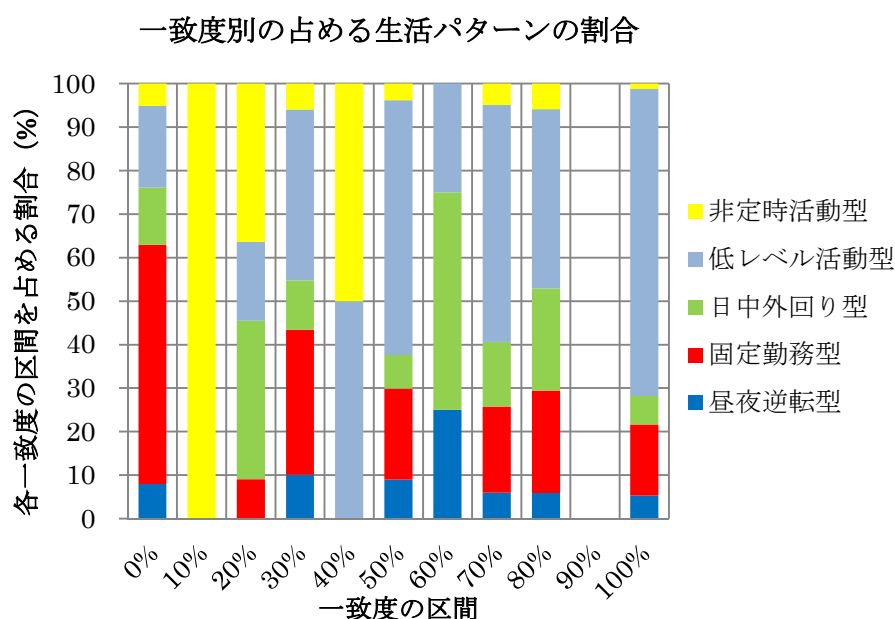


図 6-13：一致度別の占める生活パターンの割合

平日と休日的一致度が低い 0%から 30%の区間において、勤め人が多く属していると考えられる固定勤務型と日中外回り型が 57%を占めていることが分かる。『2010 年国民生活時間調査報告書』(2010) ²⁵によると、勤め人（販売職・サービス職を除く）で土曜日・日曜日が非就業日であるのは、その集団の 82%にあたる。したがって、固定勤務型と日中外回り型が、平日が就業日で休日が非就業日という過ごし方のグループの中で 57%を占めることは妥当である。したがって、活動量に着目した平日と休日の過ごし方の違いについて整合性を評価されたと考える。

6.4. 生活パターンの繰返し性に関するまとめ

第 6 章では、個人に着目して、一日一日が個人にとってどのような一日が繰り返されるのか、活動量の時間帯別分布を用いて分析した。何種類の日、平日と休日ではどの程度過ごし方が異なっているかで、活動量から見た生活パターンごとにその繰返し性を評価した。さらに、滞在時間の多い標準地域メッシュは、日常的な行動圏と考えられるが、その標準地域メッシュ数から生活パターンの繰返し性を評価した。

その結果、一日一日の過ごし方のバリエーションはかなり限られる。たとえば 6 パターンが最も多く、10 パターン以上のユーザは少数である。また、平日の年間生活パターン類型ごとにそこに属する人々の生活の日パターンを見ると、平日の生活パターンと休日の過ごし方の一致度が異なる。一般に固定勤務型の年間生活パターンに分類される人は、休日と平日の生活パターンが異なる傾向が他のタイプに比べ大きい。

平日と休日の過ごし方の違いは、ユーザの職業および雇用形態に依存する。平日と休日の過ごし方によって、ユーザのこれらの属性がある程度推定できる可能性がある。

なお、生活における滞在場所および移動距離などに関する調査は資料が少なかったために、ユーザの職業などの属性と結びつけることが難しい。

第7章 結論

個人の移動履歴に関するデータが携帯電話のログ情報の一部として大量に利用可能になってきた。これらは人の行動を数年間にわたり追跡することができる情報であるのと同時に、数百万人もの人々をカバーしていることから人々の行動を明らかにする。

また行動文脈に応じたタイムリーな情報サービスを、携帯電話を通じて提供するために非常に貴重な情報である。多くの研究が近年開始されているが、研究は大きく二つに分けられる。一つは移動履歴情報を予測可能性、繰返し性という観点から分析し、「人の行動は多様に見えるが、予測可能である」といった結論を導こうとするものであり^{9,21,22}、もう一つは個々の移動履歴情報から活動の内容や移動の方法を推定しようとする研究である

^{15,16,17,18,19,20}。しかしながら、前者は時間軸方向に見て各人の行動がそれぞれ比較的安定し、予測可能性が高いことを結論づけてはいるが、人々の行動の類型がどのくらいの数存在するのか、どのような生活のパターンが存在するのかについては分析していない。一方、個別の移動履歴を分析する研究では、夜、長時間滞在している場所は居住地であるといったステレオタイプの「生活パターン」を想定して分析をしている例もある。もしどのくらいの生活パターンが存在しているか事前にわかっていたら、そのパターンごとに分析の方法や推定のアルゴリズムなどを適切に選択することができる。本研究は、多量・多様な移動履歴データを分析し、どのような生活パターンが存在するのか、どのような特徴を持っているのかを明らかにすることで、生活モデルの構築の試みであると言える。

具体的には、GPS 付き携帯電話から得られた全国 5020 人の移動履歴情報をもとに、1 年間にわたる平日と休日ごとの平均活動量の時間帯別分布、空間的な分布を用いて、日本全体を対象に、年間の平均生活パターンを類型化した。その結果、平日の行動、休日の行動のそれぞれについて 5 パターン程度（平日は 11 クラスター、休日は 9 クラスター）で網羅できることが示された。その類型パターンごとに活動圏の空間的広さなどからどのように特色づけられるのかを明らかにした。

それに加えて、各個人の日々の生活パターンがどのくらいの類型数で記述されるのか、滞在場所にはどのくらいの繰返し性が存在するのかを明らかにした。日々の生活パターン類型は、その個人がどの年間平均生活パターン類型に属しているかにより異なるものの、概ね 4 から 6 類型であることがわかった。言い換えれば、4 から 6 の日パターンが 1 年 365 日を構成していると言ってよい。なお、休日と平日との間に大きなパターン差があるかどうかについては、必ずしも明確な結論得られなかったが、固定勤務地に通うタイプの人々は休日と平日の生活パターンの間に比較的是っきりした差異が見られた。滞在場所についてはほぼ 95% の人が 1 から 3 メッシュ（1km 四方）の空間で滞在時間（移動に要している時間を除く）の 8 割以上を過ごしている。これらの人々は滞在時間が長い場所の数が少なく、同じところ行ったり来たりしていると考えられる。

平日と休日の生活パターンの違い、滞在場所の繰返し性の観点から、各類型がどのよう

に特色づけられるのかを明らかにした。

その結果、生活パターンによっては、構成員の勤務形態などが似通っているグループがある。そして、生活パターンごとに一日の過ごし方、休日の過ごし方、活動圏の広さに相対的な差があることが分かる。

謝辞

本論文を執筆するに当たり、多くの方々からお力をいただき、誠に世話になりました。研究に対する助言はもちろん、論文執筆の基本的な作法から研究の取り組み方など様々な面でご指導していただいた有川正俊教授、そして日常的に研究の方法やリサーチクエスションの立て方などアドバイスをしていただき、また多くの研究分野の違う方々にお会いするご機会をいただいた木實新一准教授、プログラミングの勉強やサーバの構築など技術的な面から研究のまとめ方において幅広くご指導していただいた東京大学空間科学研究センター藤田秀之助教、研究の方向性についてご相談にのっていただいた柴崎亮介研究室研究員金杉洋さん、学会発表に関する事柄から研究のご相談にのっていただいた有川研究室博士課程鍛冶秀紀さん、鶴岡謙一さん、Lu Min さん、そして高い意識をもち、共に励みしあいながら研究生活を送った同期の笹尾知世さん、吉村大希さん、そして空間情報学系研究室をはじめ社会文化環境学専攻のみなさまに、感謝の言葉を述べさせていただきたいと思います。

また、この研究に大変貴重なデータをご提供していただいたゼンリンデータコム株式会社の足達龍太郎さま、森永久行さま、稲毛隆行さまに深く御礼申し上げます。

そして、最後に学生生活を支えてくださった家族への感謝を述べて、本論文の謝辞とさせていただきます。

発表文献

柴崎 真理子，藤田秀之，木實新一，有川正俊. 「長期に渡る携帯ナビ移動履歴を用いたユーザの生活スタイル推定手法の検討」. 地理情報システム学会第 21 回研究大会講演論文集，2012 年.

柴崎真理子，藤田秀之，木實新一，有川正俊. 「1 年間にわたる移動履歴を用いた生活パターン分類」 電子情報通信学会ヒューマンプローブ研究会公開シンポジウム「ヒューマンプローブの新たな展開」 ポスター発表，2012 年

参考文献

- ¹ 宮崎雄一郎, 山田直治, 住谷哲夫, 磯田佳徳. ユーザの行動に合わせたサービス実現のための行動推定技術の開発. NTT DOCOMO テクニカルジャーナル, Vol. 17, No3, pp. 55 - 61, 2009.
- ² 日本経済新聞. 『ビッグカメラ, 位置情報を活用したクーポン配信実験を開始』, http://www.nikkei.com/article/DGXNASFK1401X_U2A111C1000000/, 参照日時 2012.12.09
- ³ Jonathan Reades, Francesco Calabrese. Cellular Census: Explorations in Urban Data Collection. Pervasive Computing, IEEE, Vol. 6, No. 3, pp.30-38, 2007.
- ⁴ NTT ドコモホームページ「モバイル空間統計に関する情報」
http://www.nttdocomo.co.jp/corporate/disclosure/mobile_spatial_statistics/ 参照日時 2012.01.21
- ⁵ Jason J.Jung, Hojin Lee, and Kwang Sun Choi. Contextualized recommendation based on reality mining from mobile subscribers. Cybernetics and Systems: An International Journal, Vol. 40, No. 2, pp.160-175, 2009.
- ⁶ Nathan Eagle, Alex Pentland. Reality mining: Sensing complex social systems. Personal and Ubiquitous Computing, Vol. 10, No.4, pp.255-266, 2006.
- ⁷ 伊藤滋, 原芳男, 岡村駿, 腰塚武志, 佐枝三郎, 矢野真和. 土木工学体系 21 都市環境論. 彰国社. 1982.
- ⁸ D. Brockman, L.Hufnagel, and T. Geisel. The scaling laws of human travel. Nature Vol. 439, pp.462-465, 2006.
- ⁹ Marta C. Gonzalez, Cesar A. Hidalgo, and Albert-Laszlo Barabasi. Understanding individual human mobility patterns. Nature Vol. 453, pp. 779-782, 2008.
- ¹⁰ Stefano Spaccapietra, Christine Parent, Maria Luisa Damiani, Jose Antonio de Macedo, Fabio Porto, and Christelle Vangenot. A conceptual view on trajectories. Data&Engineering, Vol. 65, Issue.1, pp.128-146. 2008.
- ¹¹ Baris Guc, Michael May, Yucel Saygin and Christine Korner. Sematic Annotation of GPS Trajectory. 11th AGILE International Conference on Geographic Information Science. 2008.
- ¹² Miriam Baglioni, Jose Antonio Fernandes de Macedo, Chiara Renso, Roberto Trasarti, and Monica Azchowicz. Towards Semantic Interpretation of Movement Behavior. Advances in GIScience Lecture Notes in Geoinformation and Cartography, pp.271-288, 2009.

- ¹³ Santi Phithakkitunkoon, Teerayut Horanont, Giusy Di Lorenzo, Ryosuke Shibasaki, and Carlo Ratti. Activity-Aware Map: Identifying human daily activity pattern using mobile phone data. Human Behavior Understanding Lecture note in Computer Science Vol. 6219, pp. 14-25, 2010.
- ¹⁴ Richard Becker, Ramon Caceres, Karrie Hanson, Sibren Isaacman, Ji Meng Loh, Margaret Martonosi, James Rowland, Simon Urbanek, Alexander Varshavsky, and Chris Volinsky. Human Mobility Characterization from Cellular Network Data. Communications of the ACM CACM, Vol. 56, Issue. 1, pp.74-82, 2013.
- ¹⁵ Changqing Zhong, Nupur Bhatnagar, Shashi Shekhar, and Loren Terveen. Mining Personally Important Place from GPS Tracks. Data Engineering Workshop, 2007 IEEE 23rd International Conference, pp.517-526, 2007.
- ¹⁶ Andrey Tietbohl Palma, Vania Bogorny, Bart Kuijpers, and Luis Otavio Alvares. A clustering-based approach for discovering interesting places in trajectory. SAC '08 Proceedings of the 2008 ACM symposium on Applied computing, pp.863-868, 2008.
- ¹⁷ Daniel Ashbrook, Thad Starner. Using GPS to learn significant locations and predict movement across multiple users. Personal and Ubiquitous Computing Vol. 7, Issue. 5, pp.275-286, 2003.
- ¹⁸ Lin Liap, Dieter Fox, and Henry Kautz. Extracting Places and Activity from GPS Traces Using Hierarchical Conditional Random Fields. The International Journal of Robotics Research, vol. 26, No. 1 pp.119-134, 2007.
- ¹⁹ Yang Ye, Yu Zheng, Yukun Chen, Jianhun Feng, and Xing Xie. Mining Individual Life Pattern Based on Location History. Mobile Data Management: Systems, Services and Middleware, MDM '09 Tenth International Conference, pp.1-10, 2009.
- ²⁰ Yohan Chon, Yungeun Kim, Hyojeong Shin, and Hojung Cha. Topic Modeling-based Semantic Annotation of Place using Personal Behavior and Environmental Features. Mobile Data Challenge 2012 Workshop, 2012.
- ²¹ Chanoming Song, Zehui Qu, Nicholas Blumm, and Albert-Laszlo Barabasi. Limits of Predictability in Human Mobility. Science, Vol. 327, No. 5968 pp. 1018-1021, 2010.
- ²² James McInerney, Sebastian Stein, Alex Rogers, and Nicholas R. Jennings. Exploring Periods of Low Predictability in Daily Life Mobility. Mobile Data Challenge Workshop, 2012.
- ²³ Nathan Eagle, Alex Sandy Pentland. Eigenbehaviors: identifying structure in routine. Behavioral Ecology and Sociobiology, Vol. 63, Issue. 7, pp. 1057-1066, 2009.
- ²⁴ 株式会社インプレス R&D. ケータイ白書 2011. 2011.
- ²⁵ NHK 放送文化研究所. 2010 年国民生活時間調査報告書.
URL : <http://www.nhk.or.jp/bunken/yoron/lifetime/index.html> 参照日時 : 2012.12.05

²⁶ 総務省『平成 23 年度社会生活基本調査 調査の結果 結果の概要』の「調査 A に関する結果 生活時間に関する結果」の報道資料 URL :

<http://www.stat.go.jp/data/shakai/2011/gaiyou.htm#a03> 参照日時:2012.12.09

²⁷ Apichon Witayangkurn, Teerayut Horanont, Yoshihide Sekimoto, Ryosuke Shibasaki., Large Scale Mobility Analysis: Extracting Significant Place using Hadoop/Hive and Spatial Processing. Conference'10, ACM, 2010 (投稿中)

²⁸ Raul Montoliu, Jan Blom, Daniel Gatica-Perez. Discovering places of interest in everyday life from smartphone data. Multimedia Tools and Applications, Vol. 62, Issue. 1, pp. 179-207, 2012.

²⁹ Martin Ester, Hans-Peter Kriegel, Jorg Sander, and Xiaowei Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining, 1996.

³⁰ 神寫敏弘. データマイニング分野のクラスタリング手法(1)ークラスタリングを使ってみよう!ー. 人工知能学会誌 18 巻 1 号 pp.59-65, 2003.

³¹ WIKIBOOKS : Data Mining Algorithms In R URL:

http://en.wikibooks.org/wiki/Data_Mining_Algorithms_In_R/Clustering/K-Means#Algorithm 参照日時 : 2012.11.28

³² Yin Zhu, Erheng Zhong, Zhongqi Lu. and Qiang Yang, 2012. *Feature Engineering for Place Category Classification*. The Mobile Data Challenge 2012 Workshop, 2012.

付図および付表

表 1 - 1 : 用語の一覧.....	7
図 1-1 : 論文の構成.....	8
表 2 - 1 : 位置情報測定システムの特徴.....	9
図 2-1 : 位置情報測定とユーザの活動の関係	10
図 2-2 : 移動履歴に含まれる位置情報の誤差.....	11
図 2-3 : ユーザの位置情報測位日数の割合ごとの人数	12
図 2 - 4 : 地図ナビアプリユーザの男女比.....	13
図 2 - 5 : 男女別年代構成比.....	14
図 2-6 : 男女別職業構成比.....	15
表 2-2 : 各年代における補正指数.....	16
図 2-7 : 補正した地図ナビアプリユーザの男女別人数.....	17
図 2-8 : 平成 2 2 年度国勢調査における年代別人口.....	17
図 2-9 : 補正した男性の年代別職業構成.....	18
図 2-10 : 補正した女性の年代別職業構成.....	18
図 3-1 : 内挿の事例.....	23
図 4-1 : 一週間の各曜日の位置情報測定頻度.....	24
図 4-2 : 平日の位置情報測定頻度.....	25
図 4-3 : 休日の位置情報測定頻度.....	25
26	
図 4-4 : 各時間帯の相関係数のとりかた	26
図 4-5 : 平日の相関係数の散布図行列の可視化.....	27
図 4-6 : 休日の相関係数の散布図行列の可視化.....	28
図 4-7 : 平日における各主成分の固有値.....	29
図 4-8 : 休日における各主成分の固有値.....	30
図 4-9 : 平日における各主成分の寄与率と累積寄与率.....	30
図 4-10 : 休日における各主成分の寄与率と累積寄与率.....	31
図 4-11 : 平日の各主成分における変数の寄与率.....	31
図 4-12 : 休日の各主成分における変数の寄与率.....	32
図 4-13 : 平日のクラス 1 の平均活動量の時間帯別分布	34
図 4-14 : 平日のクラス 2 の平均活動量の時間帯別分布.....	35
図 4-15 : 平日のクラス 3 の平均活動量の時間帯別分布.....	36
図 4-16 : 平日のクラス 4 の平均活動量の時間帯別分布.....	36
図 4-17 : 平日のクラス 5 の平均活動量の時間帯別分布.....	37

図 4-18：平日のクラス 6 の平均活動量の時間帯別分布.....	38
図 4-19：平日のクラス 7 の平均活動量の時間帯別分布.....	38
図 4-20：平日のクラス 8 の平均活動量の時間帯別分布.....	39
図 4-21：平日のクラス 9 の平均活動量の時間帯別分布.....	39
図 4-22：平日のクラス 10 の平均活動量の時間帯別分布.....	40
図 4-23：平日のクラス 11 の平均活動量の時間帯別分布.....	40
図 4-24：休日のクラス 1 の平均活動量の時間帯別分布.....	41
図 4-25：休日のクラス 2 の平均活動量の時間帯別分布.....	42
図 4-26：休日のクラス 3 の平均活動量の時間帯別分布.....	42
図 4-27：休日のクラス 4 の平均活動量の時間帯別分布.....	43
図 4-28：休日のクラス 5 の平均活動量の時間帯別分布.....	43
図 4-29：休日のクラス 6 の平均活動量の時間帯別分布.....	44
図 4-30：休日のクラス 7 の平均活動量の時間帯別分布.....	44
図 4-31：休日のクラス 8 の平均活動量の時間帯別分布.....	45
図 4-32：休日のクラス 9 の平均活動量の時間帯別分布.....	45
表 4-1：平日のクラスタリング結果の生活パターン分け.....	48
表 4-2：休日のクラスタリング結果の生活パターン分け.....	50
図 4-33：平日の生活パターンごとの休日の過ごし方.....	52
図 4-34：曜日ごとの時間帯別にみる行動者率の一日の推移.....	54
表 4-3：在宅時間のパターン類型化.....	57
図 4-35：生活パターンごとの在宅時間の分布.....	58
表 4-4：勤務地への滞在時間のパターン類型化.....	60
図 4-36 平日の生活パターンごとの勤務地滞在時間の分布.....	62
図 5-1：内挿した移動履歴データからの活動圏算出の事例.....	64
図 5-2：平日の生活パターンの分析結果.....	65
図 5-3：平日『昼夜逆転型』の平均活動圏の広さ.....	66
図 5-4：平日『固定勤務型』の平均活動圏の広さ.....	66
図 5-5：平日『日中外回り型』の平均活動圏の広さ.....	67
図 5-6：平日『低レベルの活動型』の平均活動圏の広さ.....	67
図 5-7：平日『非定時活動型』の平均活動圏の広さ.....	68
表 5-1：平日の生活パターンごとの平均活動圏の中央値.....	68
図 5-8：休日の生活パターンの分析結果.....	69
図 5-9：休日『昼夜逆転型』の平均活動圏の広さ.....	69
図 5-10：休日『日中活動型』の平均活動圏の広さ.....	70
図 5-11：休日『夕方中心日中活動型』の活動圏の広さ.....	70

図 5-12：休日『低レベルの活動型』の活動圏の広さ.....	71
図 5-13：休日『非定時活動型』の平均活動圏の広さ.....	71
表 5-2：休日の生活パターンごとの平均活動圏の中央値	71
図 6-1：各曜日に含まれるクラスの割合	75
表 6-1：ユーザ A の 1 日を対象にしたクラスタリング結果.....	76
図 6-2：日パターン数ごとの該当人数	78
図 6-3：平日の生活パターン数の算出の手順.....	79
図 6-4：平日と休日別の生活パターン数.....	80
図 6-5：平日と休日を占めるクラスの一致の程度を算出する手順.....	81
図 6-6：5020 人の平日と休日におけるクラスの一致度別の人数.....	82
図 6-7：滞在時間の長い標準地域メッシュの算出の手順	84
図 6-8：滞在時間が長い標準地域メッシュの個数における該当人数	85
図 6-9：平日の生活パターンごとの最適クラス数.....	86
図 6-10：平日の生活パターンごとの平日と休日を占めるクラスの一致度.....	87
図 6-11：滞在時間が長い標準地域メッシュの個数に対する平日の生活パターンごと の該当者数	88
図 6-12：職業別の休みである曜日	89
図 6-13：一致度別の占める生活パターンの割合	90