

スマートフォン位置情報から読み解く観光地における人の流動 ～ビッグデータはどこまで人の行動を追えるのか～

Understanding Tourists' Activity through Analyzing Smartphone Data

廣瀬 健¹ 井村祥太郎¹ 絹田裕一² 矢部 努³

By Takeshi HIROSE, Shotaro IMURA, Yuichi KINUTA and Tsutomu YABE

1 はじめに

総務省によれば、2017年の個人のモバイル端末の保有率は84.0%、その7割に上る60.9%がスマートフォンを保有しており、急速に普及が進んでいる¹⁾。そのような中、それらから得られる位置情報を活用したマクロ・ミクロの両面から移動実態を把握する技術が注目されている。スマートフォンから取得可能なデータ（以下、スマートフォン位置情報と表記）からは、任意の地域への来訪実態（いつ、どこにいたか）や移動の軌跡、任意の地点への集中状況などが把握可能である。そのため、行動が多岐に渡ることで、今まで特に把握しづらいとされていた都市圏内・都市内の流動が把握可能となることが期待されている²⁾。

本稿では、スマートフォン位置情報の特性を整理した後、神奈川県鎌倉市の鎌倉地区をフィールドとした、来訪者の行動特性に関して実施した分析結果について述べる。検証の結果を踏まえ、データ活用にあたる今後の展望について取りまとめる。

2 スマートフォン位置情報とは

スマートフォン位置情報には、代表的なものとして、携帯電話の基地局情報（以下、基地局データと表記）、GPS位置情報（以下、GPSデータと表記）、Wi-Fiデータの3種類がある。データごとに計測箇所が異なることで、測位精度の異なり、特性の違いが生じる。以下に一般的な特性について述べる³⁾。なお、これらのデータには一部フィーチャーフォン（所謂「ガラケー」）から取得された情報も含まれていることを留意されたい。

基地局データは、携帯電話が定期的に基地局と交信している履歴をもとに、人の移動を把握する。携帯電話の電源をオンにしていれば、データ取得の対象となるため、ビッグデータの中でも特に大量のサンプルが

取得できる。交信が記録された基地局の位置や時刻から、滞在エリアや移動を把握することができる。

GPSデータは、スマートフォン等のGPS機能で取得される位置情報をもとに、通信事業者やアプリ提供事業者が提供する移動データである。収集頻度はサービス提供事業者によって異なり、アプリ起動時のみ収集するもの（現在地測位）と定期的に移動履歴を収集するもの（定期測位）がある。GPSは見通しの良い屋外の位置情報を測位する機能であるため、地下や建物内では位置情報が取得できない場合がある。地下街、アーケード街、ペDESTリアンデッキの下部などの分析には留意が必要である。

Wi-Fiデータは、スマートフォン等がWi-Fiアクセスポイントに接続した履歴をもとに、人の移動を把握する。Wi-Fi機能をオンにしていれば、データ取得の対象となり、アクセスポイントが密に設置されていれば、滞在場所や移動経路を推定することが可能である。GPSデータと比較すると、地下や建物内における移動も把握できるという特徴があるが、あくまでアクセスポイントの設置位置をもとに推定したものである。

本検討で対象とした鎌倉地区は、アクセス経路が限られる地域で、高い建物や地下空間などがあまり存在しない。よって、基地局データ及び、GPSデータの活用に関する検証を行うこととした。本検討においては、基地局データはソフトバンク株式会社、GPSデータは株式会社Agoopのデータを活用した。

3 分析内容と使用データの概要

(1) 分析内容

2. で述べた各データの特性を踏まえ、基地局データを単独で活用した分析、GPSデータを単独で活用した分析、二者を組み合わせた分析を実施した。活用したデータの概要を表-1に示す。なお、個人情報保護の観点から端末IDをはじめとした個人情報の紐づけによ

¹交通・社会経済部門 研究員 ²交通・社会経済部門 グループマネジャー

³交通・社会経済部門 担当部門長兼グループマネジャー 博士(工学)

るデータの連携はできないことを留意されたい。

基地局データ単独での活用では、総量を把握可能なメリットを勘案し、以下の2項目の分析を実施した。

- ・日別来訪者数
- ・来訪者の出発地（平休別）

GPSデータ単独での活用では、各端末の詳細な位置情報が把握可能なことのメリットを勘案し、以下の4項目の分析を実施した。

- ・来訪者の利用交通手段（方面別）
- ・来訪者の来訪時間帯と滞在時間分布
- ・来訪者の利用交通手段別訪問施設
- ・鎌倉地区内の利用交通手段別周遊パターン

基地局データとGPSデータを組み合わせた分析として、以下の2項目を実施した。

- ・ヒートマップの作成
- ・ヒートマップを活用した周遊行動

(2) 基地局データの概要

本検討では、鎌倉中心部の2018年11月分のメッシュ別・日別・時間帯別の滞留人口データを用いた（鎌倉中心部の判定エリアは図－1参照）。

データに付与される個人属性は、年齢（10代～80代の8分類）・性別（男女）・鎌倉市内居住/非居住・推定居住地（都道府県に集約）である。基地局データは、付与する個人属性の数が増えると秘匿処理が施されるが、今回用いたデータの秘匿率は2.4%であった。

基地局データは500mメッシュ別に集計されたデータであり、分析にあたって特殊な処理は実行しない。

(3) GPSデータの概要とデータ処理

GPSデータは、1日のうちに鎌倉市内で1度でも観測されているIDのその日一日の全観測データ（以下、観測データをログと表記）を用いた。データ取得期間は、基地局データ同様2018年11月の1か月分である。

分析にあたって、まず鎌倉地区への来訪者を抽出するため、オリジナルのデータから、居住者と思われるログを除外した。さらに、データの活用にあたって以下の観点から2種類のデータを作成した。GPSデータを単独で活用した行動特性の分析にあたっては、利用交通手段や滞在地などを分類するための処理を実施した。一方で、ヒートマップの作成にあたっては、記録

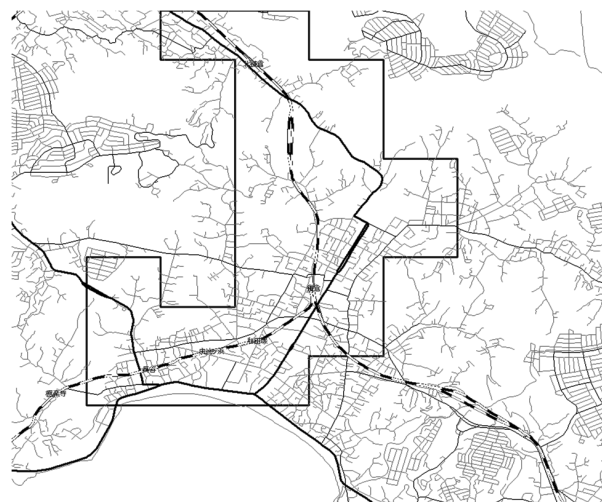
表－1 本検討で用いるデータの概要※1

	基地局	GPS
データ取得の対象者	携帯端末利用者	特定アプリの利用者
計測箇所単位	基地局ごと	緯度経度
提供データの空間解像度	任意の広さのメッシュ単位	緯度経度情報
計測時間間隔	基地局との通信時	数秒～数十分
提供データの時間解像度	1時間単位	分単位
個人情報	契約者情報 (性・年齢) 推定居住地※2 (市区町村)	推定居住地※2 (市区町村)
拡大方法	性・年齢・居住地で拡大	拡大なし

※1 データ特性は提供者によって異なる

※2 データ提供者の判定方法で推定された居住地

(出典：筆者作成)



図－1 鎌倉中心部の判定エリア

(出典：筆者作成)

されているログの重みを平準化する処理を施した。

a) データ処理

まず、GPSデータ単独による行動特性の分析にあたり、データ数による抽出に加えて、利用交通手段・方面の判別、鎌倉地区内での立寄り施設の判定を行った。

データ数による抽出では、精度確保の観点から、鎌倉中心部エリア内において一定以上のログが記録されているIDを抽出した（判定エリアは図－1参照）。

利用交通手段・方面の判別は、鎌倉中心部への流入

前/流出後のログ情報を用いて行った。具体的には、エリア内への流入前/流出後のログに着目し、GPSデータを国土数値情報の鉄道リンクやデジタル道路地図の道路リンク群とマップマッチングすることで、対応づいたリンクが最も多い交通手段・方面をそのIDの利用交通手段・方面と判定した。利用交通手段と方面は、全9パターン設定しており、鉄道利用として、A.横須賀線(横浜)方面、B.横須賀線(逗子)方面、C.江ノ電方面の3パターン、自動車利用として、①北鎌倉方面、②朝比奈方面、③藤沢方面、④江ノ島方面、⑤逗子(R134)方面、⑥逗子(県道)方面の6パターンとした。道路に関しては、代表路線の周辺道路を含めてリンク群を設定している(図-2)。

鎌倉地区内での立寄り施設もGPSデータの位置情報による判定を実施した。具体的には、鎌倉地区内の観光施設の領域を示すポリゴンを作成(以下、施設ポリゴンと表記)し、その上に緯度経度情報を保持したGPSデータをプロットする。GPSデータの各ログが施設ポリゴンの領域に内包されるかを判定し、ログが連続で一定時間以上存在していたと判断される場合、そのログのIDは当該施設に「立寄った」と判定した(図-3)。なお、今回の検証では、鎌倉駅周辺、小町通り、若宮大路、鶴岡八幡宮、長谷駅周辺、長谷寺、高德院(鎌倉大仏)、北鎌倉駅周辺、明月院、円覚寺、建長寺、東慶寺の計12施設の施設ポリゴンを作成し、それぞれの施設の立寄りの判定を行った(図-4)。

b) 基地局データと組み合わせた分析に向けた処理

GPSデータは、携帯端末によってログの取得頻度が異なり、空間的・時間的な重みが同等ではない。例えば、ログの間隔が10分間だとして「2点間に距離がある」場合と「ほぼ同じ位置で観測される」のでは、前者は10分間毎分異なる位置に存在していたと思われるのに対し、後者は同じ場所に10分間継続して存在していたと捉えることができる。そのため、ヒートマップを作成するにあたって、後者のログの周辺のほうがその空間に存在している人・時間の値が大きいことを表現するため、ログの補完を行うことでログの重みを同等にする処理を施した。

具体的には、実際のログの観測時間と位置からこの2点間を直線的に動いたと仮定し、2点の間にログを追加することで補完とした。

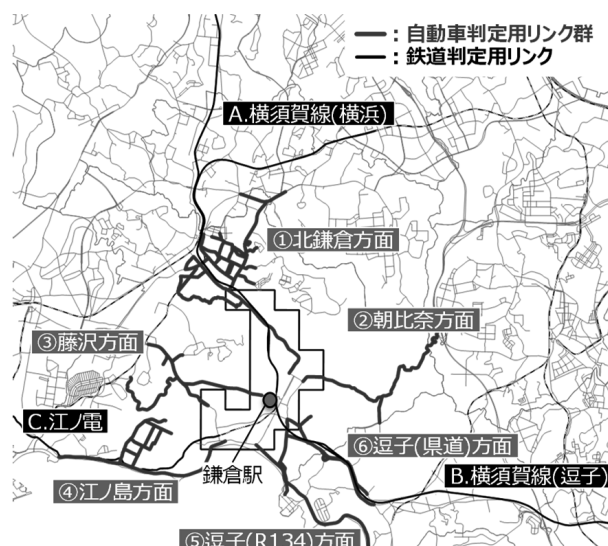


図-2 利用交通手段・方面判定用のリンク群

(出典：筆者作成)

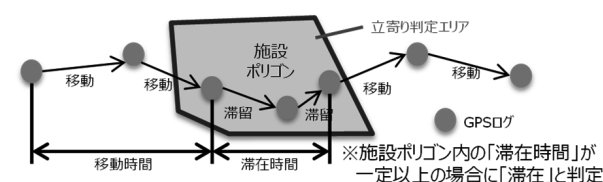


図-3 滞在判定の仕方

(出典：筆者作成)

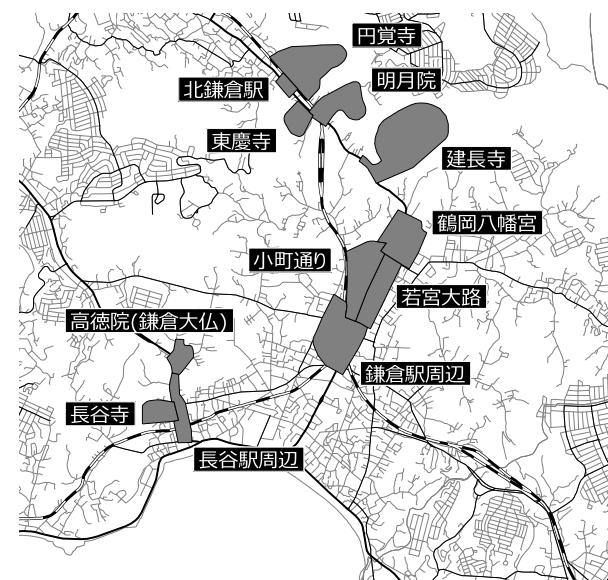


図-4 滞在判定を行った施設と各施設のポリゴン

(出典：筆者作成)

4 来訪者の行動特性の把握

本章では、基地局データと処理済みのGPSデータを用いて実施した分析の結果、来訪者の行動について把握できた事柄を取りまとめる。

(1) 基地局データによる分析

a) 鎌倉地区への日別来訪者数分析

図-5は、基地局データを集計して得られた、2018年11月の鎌倉地区への来訪者数の日別の推移である。2018年11月は、11月23日から25日が3連休だったことから、他の土日と比較しても、この3日間の来訪者数が突出して多くなっていることが確認できる。

平日の棒グラフに着目すると、11月の前半から後半にかけて徐々に来訪者数が増加しており、紅葉シーズンが深まるにつれ、来訪者数が多くなっていることが表現されている。さらに、11月30日(金)の来訪者数は、休日である11月4日(日)の来訪者数を上回り、3連休を除く他の休日と匹敵する来訪者数を記録しており、紅葉シーズンは平日も混雑していることが想像される結果となっている。

b) 鎌倉地区への来訪者の居住地分析

基地局データに付与される推定居住地を真の居住地と仮定し、鎌倉地区への来訪者の居住地の構成比を集計したものが図-6である。平日は神奈川県内の居住者が5割強と大きな割合を占めているが、休日は神奈川県外からの来訪者割合が大きくなっていることが分かる。平休共に東京都内と神奈川県内からの来訪者の合計が7割程度を占めており、休日は都内からの来訪者が大きく増加することが確認できる。その他の地域の構成割合は、平休で概ね同程度の数値を示しており大きく変動しないものの、2018年11月の来訪者数は、平日が平均約31,000人であるのに対し、休日は平均約65,000人となっており、休日の来訪者数は平日の約2倍程度となっている。これはつまり、神奈川県・東京都を除いた他の地域からも休日は平日の2倍程度の来訪者がいることとなる。

(2) GPSデータによる分析

a) 利用交通手段(方面別分布)の分析

GPSデータにおいて各IDについて、利用交通手段・流入/流出方向の判定を行い、鎌倉地区への来訪者の利用交通手段の分析を実施した。

はじめに、方面別の利用交通手段の集計結果について示す。なお、この分析では交通手段間の比較が可能なように、方面の数を統一するため、集計は図-7に示す3方向に分類して行った。

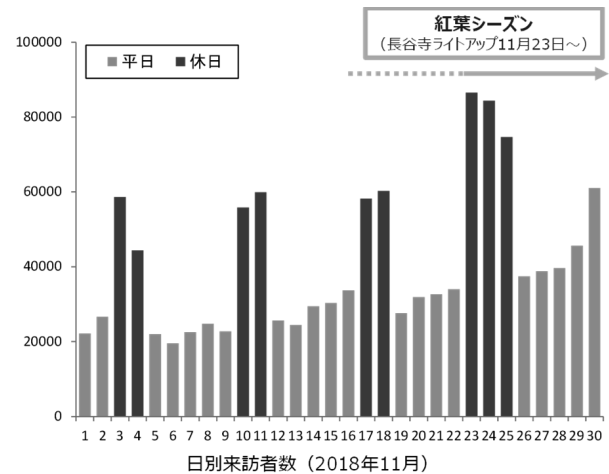


図-5 日別来訪者数の推移

(出典：基地局データを基に作成)

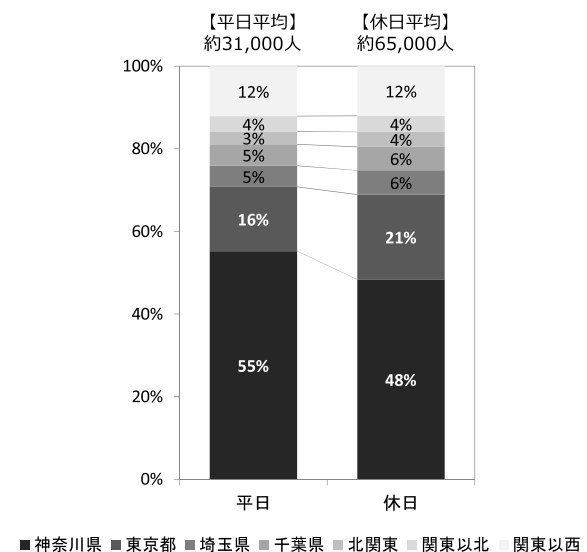


図-6 平休別来訪者の居住地の構成

(出典：基地局データを基に作成)

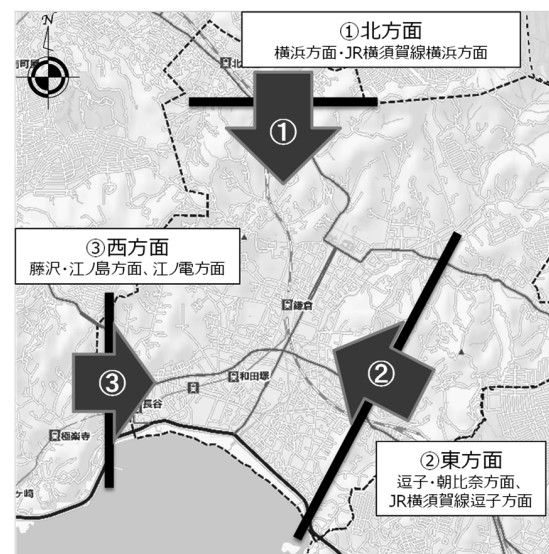


図-7 GPSデータ分析時の方面の分類

(出典：筆者作成)

全方面合計での来訪者の利用交通手段は、約5割が鉄道利用となっており、3割が自動車利用である。不明は、2割弱となっている（図－8）。

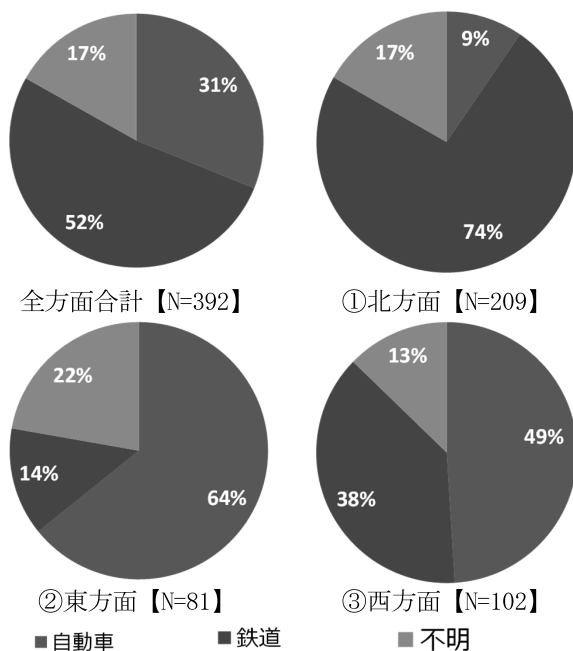
各方面の来訪者の構成比は、①北方面（横浜方面、JR横須賀線横浜方面）が53.3%、②東方面（逗子・朝比奈方面、JR横須賀線逗子方面）が20.7%、③西方面（藤沢・江ノ島方面、江ノ電方面）が26.0%となっており、圧倒的に北方面からの来訪者が多い。

方面別に利用交通手段の内訳を見ると、北方面では鉄道利用が7割強と大きな割合を占める一方、東方面では自動車利用が6割、西方面でも自動車利用が5割程度と自動車利用が大きな割合を占めている。

次に、利用交通手段別に各断面別流入/流出経路の構成比について示す（図－9、図－10）。

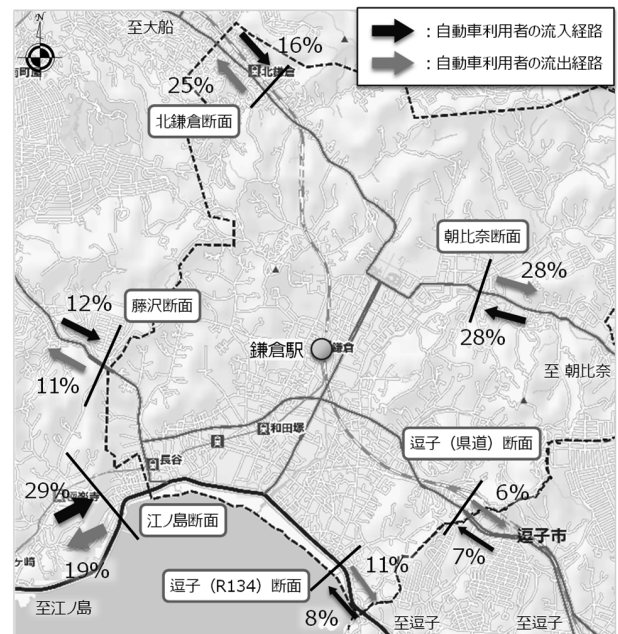
自動車による来訪者は、各方面に分散しており、特に大船に至る北鎌倉断面や横浜横須賀道路の朝比奈ICに至る朝比奈断面といった鶴岡八幡宮の北側の経路を多くの来訪者が利用していることが確認できる。これらの方面の利用割合に江ノ島断面の利用を加えることでおよそ6割となり、国道134号を含む逗子方面の道路を利用する来訪者は比較的少数である。

鉄道による来訪者は、8割弱がJR横須賀線（北鎌倉断面）、残りの2割程度は、江ノ電を利用している。JR横須賀線（逗子断面）を利用する来訪は少ない。



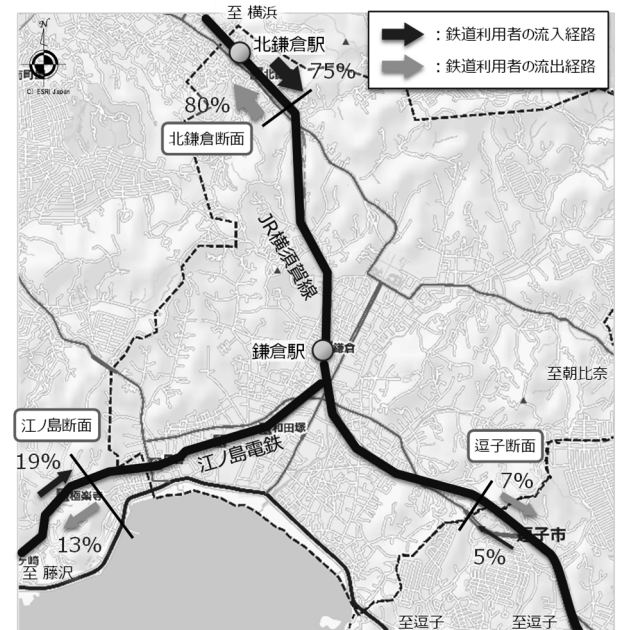
図－8 来訪方面別の利用交通手段の内訳

（出典：GPSデータを基に作成）



図－9 自動車利用者の流入/流出方面内訳

（出典：GPSデータを基に作成）



図－10 鉄道利用者の流入/流出方面内訳

（出典：GPSデータを基に作成）

b) 利用交通手段別の滞在時間の分析

次に、鎌倉地区への来訪者の利用交通手段別の滞在時間の分析を行った。

図－11は、利用交通手段別の来訪時間帯分布と帰宅時間帯分布である。それぞれ、鎌倉地区内で最初に観測された時刻を来訪時間、最後に観測された時刻を帰宅時間と仮定して分析している。

鉄道による来訪者は午前中に、自動車による来訪者

は午前9時台や11時台に集中しており、その後は分散傾向にある。帰宅については、いずれも14時頃から増え始め、17時台には多くが帰宅する傾向にある。

利用交通手段別にみると、自動車利用者は13時台に帰宅する来訪者が突出している一方で、鉄道利用者は19時台になっても一定数帰宅者が存在する傾向から、自動車利用者は比較的早い時間帯に帰宅し、鉄道利用者は比較的夜間まで滞在していることが示された。

図-12は、利用交通手段別の鎌倉地区内での滞在時間分布を示したものであり、色の濃い右側の割合が大きいほど滞在時間が長いことを表している。この図から、鉄道利用者が自動車利用者と比較して滞在時間が長い傾向にあることが確認できる。特に鉄道利用者においては、7時間以上滞在している来訪者が3割強と

もっとも大きな割合を占めている。

各利用交通手段による来訪者の平均滞在時間は、自動車来訪者が3.9時間、鉄道来訪者が5.0時間となっており、こちらでも鉄道利用者と自動車利用者の間で滞在時間に差がみられた。

c) 利用交通手段別訪問施設の分析

GPSデータを用いて判定を行った立寄り施設の情報から鎌倉地区内での来訪者の利用交通手段別の訪問施設の分析を行った。先述の通り、立寄り施設の判定は、鎌倉駅周辺、小町通り、若宮大路、鶴岡八幡宮、長谷駅周辺、長谷寺、高德院（鎌倉大仏）、北鎌倉駅周辺、明月院、円覚寺、建長寺、東慶寺の計12施設で実施している。各利用交通手段による全来訪者のうち各施設周辺に立寄っている来訪者の割合を集計し、図化したものが図-13、図-14である。

鉄道による来訪者、自動車による来訪者共に来訪箇所が多いのは鎌倉駅周辺（鉄道：54%、自動車：38%）・小町通り（鉄道：41%、自動車：31%）・若宮大路（鉄道：23%、自動車：35%）・鶴岡八幡宮（鉄道：19%、自動車：29%）となっている。鉄道による来訪者は小町通りに立ち寄っていると判定されている割合が高い一方で自動車による来訪者は、若宮大路に立ち寄っていると判定されている割合が高い。

しかし、各施設に対応づくサンプルが総じて少なく、利用交通手段間で差異が見られなかった。加えて、鉄道利用者の多くは、鎌倉駅を利用しているものと推察されるが、立寄り割合は54%と低い印象を受ける。今回の分析では、施設の立寄りの判定条件を「図-4の施設ポリゴンに15分以上滞在した場合」としているため、15分間未満の滞在では立寄りとは判定されないためである。短時間の滞在を把握するためには、判定条件やデータ取得方法の改良が課題である。

(3) 基地局データとGPSデータを組合せた分析

来訪者の総量を把握可能だが解像度の粗い基地局データと詳細な位置情報が把握可能だがサンプル数が少ないGPSデータを組み合わせることで、より実態に即したヒートマップの作成を試みた。

以下の手順で、データを組み合わせている。

手順①：ログデータ補完を行ったGPSデータを、「100mメッシュ・15分単位」に割付

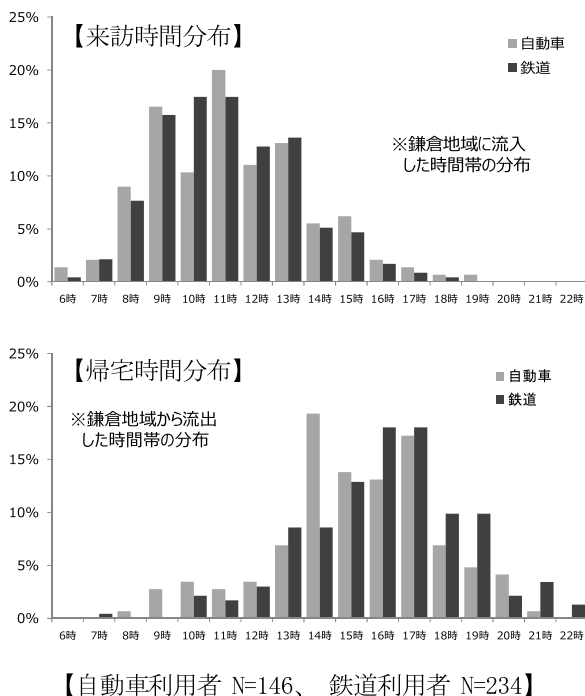


図-11 利用交通手段別来訪・帰宅時間帯分布
(出典：GPSデータを基に作成)

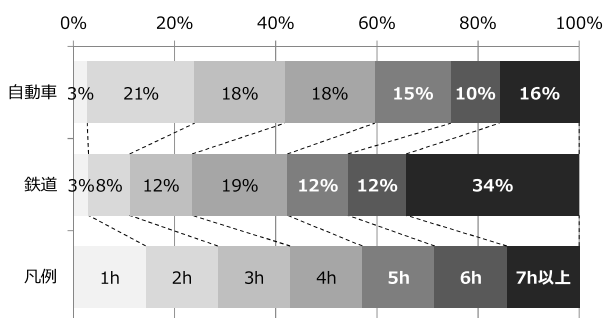


図-12 利用交通手段別滞在時間分布
(出典：GPSデータを基に作成)

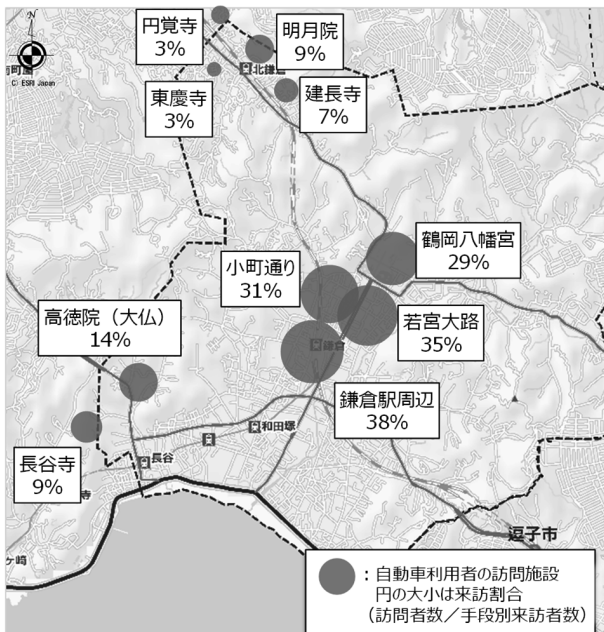


図-13 自動車による来訪者の訪問施設の割合 [N=153]
(出典: GPS データを基に作成)

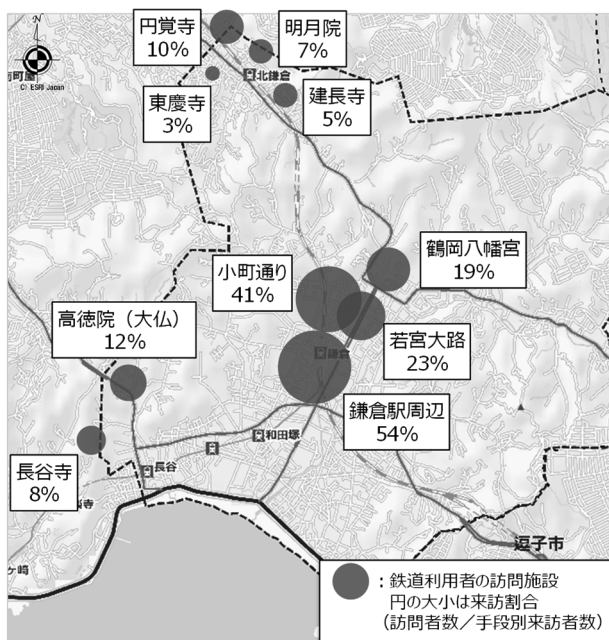


図-14 鉄道による来訪者の訪問施設の割合 [N=239]
(出典: GPS データを基に作成)

手順②: 手順①で割付したログから、「500mメッシュ・1時間単位」のログ数 (母数) を集計
 手順③: 手順②の補完後ログ数を母数に、「100m・15分単位」の滞留人口按分比率を作成
 手順④: 手順③で作成した按分比率と基地局データの集計地である「500mメッシュ・1時間単位」の滞留人口に掛け合わせ、「100m・15分単位」の滞留人口 (ヒートマップ) として詳細化

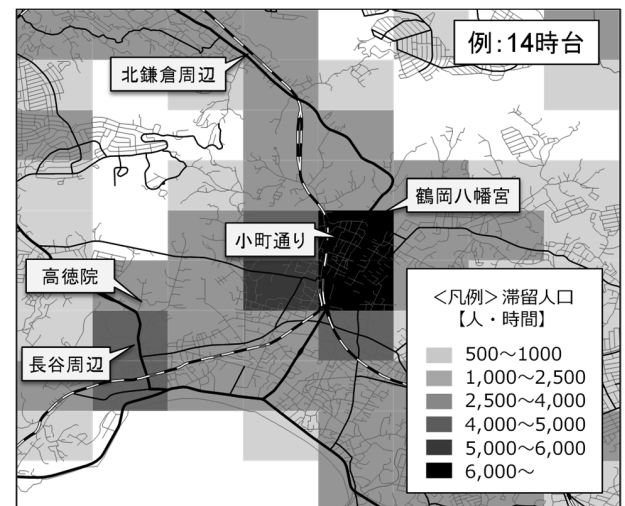
a) ヒートマップの作成結果

上記の手順で作成したヒートマップの前後比較を、滞留人口が最も多かった14時台を例として、図-15に示す。基地局データ単独では、小町通りと鶴岡八幡宮への来訪者の集中度合いが分からなかったのに対し、データを組み合わせることにより、小町通りと鶴岡八幡宮周辺の詳細化ができています。

さらに、詳細化により、鎌倉駅周辺や南西の長谷駅や高徳院 (鎌倉大仏) 周辺も、この時間帯に来訪者が集中していることが表現できており、より実態に即したヒートマップになっていると考えられる。

b) ヒートマップを活用した周遊行動の分析

作成したヒートマップをもとに、北鎌倉周辺、長谷寺周辺、鶴岡八幡宮周辺、小町通りのメッシュの滞留



↓ 時間・空間的に詳細化

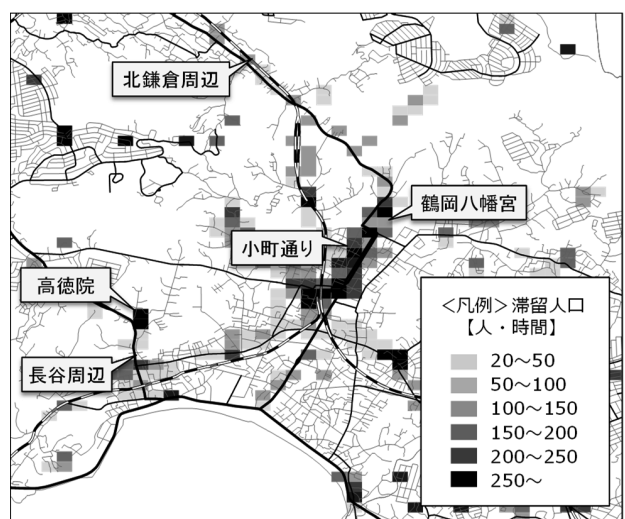
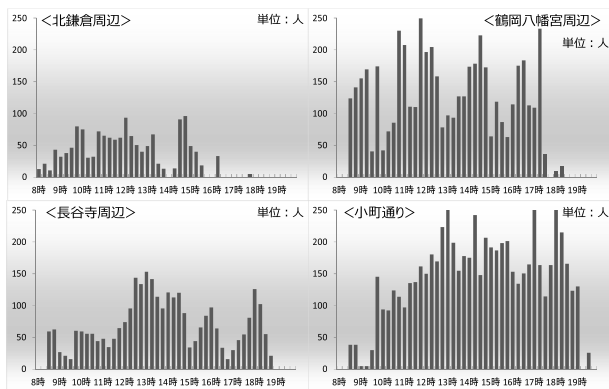
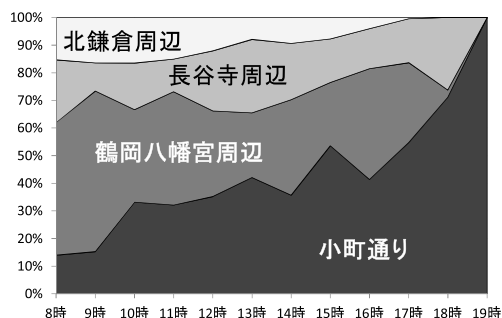


図-15 基地局データとGPSデータの統合による詳細化
(出典: 基地局データ・GPSデータを基に作成)



図－16 主要観光地周辺の来訪者数の推移

(出典：基地局データ・GPSデータを基に作成)



図－17 観光エリア毎の滞留人口の時間構成比

(出典：基地局データ・GPSデータを基に作成)

人口を時間帯別に集計した。なお、施設の広がりや、集計対象メッシュ数が異なることに留意されたい。

各施設周辺の時間帯による来訪者数の推移を見ると、特に小町通りにおいて日中は常に15分間に150人程度が存在していることが分かる。北鎌倉周辺は、午前中に来訪者が集中し、長谷寺周辺は午後に来訪者が増加していることが分かる(図－16)。

この集計値から時間帯別の構成比を算出した(図－17)。北鎌倉・鶴岡八幡宮周辺は午前中、長谷寺周辺は正午過ぎ、小町通りは午後以降に来訪者が相対的に集中する傾向がよりはっきり示されている。

5 検証結果のまとめと今後の展望

(1) 検証結果のまとめ

本検討で実施した、基地局データ、GPSデータ及びそれらを組み合わせたデータによる分析を通して、以下の知見が得られた。

個々のデータの活用について、基地局データからは、鎌倉地区への来訪者数の日変動や来訪者の居住地の構成比を把握できた。GPSデータからは、来訪者の

利用交通手段や流入/流出方面を把握できた。一方で、来訪施設については十分に把握ができず、周遊行動の分析にあたっては課題が残る結果となった。

基地局データとGPSデータを組み合わせることで、それぞれのデータの特徴を生かし、より実態に即した滞留人口の推移を表現することができ、さらにそれを集計することで個別施設周辺の来訪者の集中度合、来訪者の周遊行動の傾向も把握することができた。

(2) 今後の展望

検証結果より、現時点のデータでも人が集中する時期や時間帯の混雑予報や、コンサートなどのイベント時の流動把握への活用は充分に考えられる。

今後、GPSデータに性別やおおまかな年齢階層(若年層・生産年齢・高齢者など)の情報が付与されれば、属性別の基地局データと組み合わせることで、都市内における属性別の周遊行動が把握可能となり、ターゲットを絞った効果的な交通計画・観光施策の検討が可能になると考えられる。

また、仮にGPSデータのサンプル数の増加や、時間解像度の詳細化が叶えば、より正確で詳細な都市内の周遊や利用交通手段の分析が可能となり、来訪目的の分析への応用も考えられる。

謝辞：本稿は、国土交通省が公募した「交通マネジメントに活用する新たな技術」の検証の成果の一部をとりまとめたものである。当該検討にあたり、ソフトバンク株式会社、株式会社Agoopにご協力を頂いた。ここに記して関係各位に謝意を表する。

参考文献

- 1) 総務省：平成30年度情報通信白書，第5章 第2節，pp.235-236，2018
- 2) 森尾淳，石井良治，中野敦，萩原剛，田中啓介：都市交通分野におけるビッグデータの活用とその展開～ビッグデータの特性と活用可能性～，IBS研究活動報告2018，pp.29-36，2018
- 3) 国土交通省都市局：総合都市交通体系調査におけるビッグデータ活用の手引き，2018