

街路構造の類似性に着目した鉄道駅周辺エリアの分類

大塚昇・山本佳世子

Classification of Areas around Railway Stations Focusing on the Similarity of Street Structures

Noboru Otsuka and Kayoko Yamamoto

Abstract: In general, it is difficult to imagine a landscape that has never been visited. Therefore, it is also difficult to decide which city to live in when moving. In order to solve such problems, this research finally aims to develop a system that calculates the similarity between cities and recommends similar cities from known ones. For this purpose, in the paper, we first verified the algorithm for classifying railway stations. The area within 800m from each station on the Chuo and Keio lines of Tokyo was treated as one area, and the data of the street structure was extracted from the open road map. Clustering analysis was conducted based on the feature amount obtained from the street network of each area, and an attempt was made to classify the two routes.

Keywords: 街路構造 (Street Structure), オープンストリートマップ (Open Street Map), グラフネットワーク (Graph Network), 鉄道駅エリア (Railway Station Area), 類似性 (Similarity)

1. はじめに

1.1. 研究の背景

進学や就職を機に引越し、新しい生活を始めるという人々は毎年多くみられる。しかし、そうした人々がオンラインで引越し先を検索しようとすると、すぐに戸惑いを覚えることになるだろう。それは多くの不動産検索サービスでは最初に住みたいエリアまたは駅を選択する必要があることに起因する。一般的に、土地勘のない場所の街並みを想像することは困難である。そのため住む場所

を選ぶためには、多くの街の特徴を事前に調査する努力を要する。しかし、実際にその土地を訪れずに街の特徴を把握することは難しい。引越し先の候補となる街を全て訪れて比較・検討することには、時間とコストの両面において現実的な方法ではない。その結果、多くの人々は家賃や間取りといった物件自体のスペックのみに基づいて引越し先を選択してしまっている。それは「住む街を選択する」という過程が抜けており、実際に生活を始めたあとでライフスタイルとの齟齬を生む可能性がある。

* 大塚 昇 学生会員 電気通信大学大学院情報理工学研究科 (University of Electro-Communications · Graduate School of Informatics and Engineering)

〒135-0033 東京都江東区深川1-6-16深川住宅B-803 E-mail: noboru.otsuka@go.wt-jdpsr.jp

** 山本 佳世子 正会員 電気通信大学大学院情報理工学研究科 (University of Electro-Communications · Graduate School of Informatics and Engineering)

本研究では、最終的には、既知の街から類似した街を推薦するシステムを構築することで、上記の問題を解決することを目指す。そのために、本稿では、まずは、各街の都市的類似度を算出するアルゴリズムを構築する。

1.2. 研究の位置付け

街並みのイメージを捉えようとする研究は主に建築学や都市計画の分野で行われてきた。例えば西尾ら（1995）は、LOGMAPと呼ばれるイメージ分析手法を用いて京王線沿線の街に対して人々が想起するイメージに対して計量的に分析している。このように、街の景観やイメージといった主観的な観点から街の特徴や類似性を把握しようとする研究は多く試みられてきた。

しかし、本研究では街と街との類似度を測るために街路構造を用いる。ここで街路構造に着目している理由は大きく2点挙げられる。(1)街路構造をグラフネットワークとして扱うことで、街同士の特徴を定量的に比較・検討することができる。(2)ある街に長期間住んだ場合、駅や商業地など街の主要機能に至るまでの導線が、街に対する心象に大きく寄与していると考えられるためである。街路構造という客観的な指標を用いて街の特徴を把握することはこれまで研究されてきた主観的な都市イメージに新たな知見をもたらすことができると考える。

本稿では、類似度算出アルゴリズムを構築するための下準備として、街路構造を街の特徴を把握するための客観的な指標とし、鉄道駅周辺エリアの分類を行った結果について報告する。

2. 実験方法と対象エリア

2.1. 仮説

本研究では、街と街の類似性を測定するために次のような仮説をたてた。それは、”中央線や京王線といった特定の路線沿いの駅前空間は各鉄道

会社の開発方針にのっとって開発されており、同一沿線上の駅前空間は似通っている”という仮説である。この仮説に基づき、街路構造のグラフネットワークから得られた統計量を用いて、各沿線に存在する駅ごとに、駅前街路空間の分析を行う。

2.2. 実験方法

まず、決定木モデルによる学習を通じて、ある駅が中央線沿線であるのか、京王線沿線であるのかを分類するモデルを構築した。決定木モデルは教師あり学習と呼ばれる機械学習モデルの1つであり、説明変数に使用したデータを用いて条件付き分岐を繰り返すことにより分類を行う。決定木モデルが学習した分類ルールを可視化し、分類に有効な説明変数を確認した。

次に、決定木モデルによって有効性が認められた説明変数を用いて、K-Means法によるクラスタリングを試みた。K-Means法は、あるデータセットの中から類似した特徴を持つデータを自動的に分類するアルゴリズムであり、教師なし学習と呼ばれる機械学習手法の1つである。このクラスタリングの結果から、決定木モデルで分類に用いられた変数が二路線間の駅前街路空間の特徴を説明できているかを検証した。

2.3. 対象エリア

本稿では中央線と京王線の二つの沿線を分析対象とした。

中央線は1889年に開通した。当初は新宿、中野、武蔵境、国分寺、立川、八王子の6駅がのみで利用者の数も少なかった。1930年に全線が電化すると利用価値が高まり、1935年ごろにはだいたい現在の姿が完成した（井出、1967）。京王線は1913年に笹塚駅-調布駅間が開通した。近年の京王線沿線の開発では、1960年代後半の住宅需要を背景とした多摩ニュータウン計画などがその発展に大きく影響している（福島、1975）。本稿ではこの異

なる発展を遂げてきた二つの沿線を対象に各駅の街路ネットワークの特徴量から二路線の分類を試みた。

表1 対象とした中央線および京王線の駅一覧

中央線 沿線の駅	東京 神田 御茶ノ水 四ツ谷 新宿 中野 高円寺 阿佐ヶ谷 荻窪 西荻窪 吉祥寺 三 鷹 武蔵境 東小金井 武蔵小金井 国分寺 西国分寺 国立 立川 日野 豊田 八王子 西八王子 高尾
京王線 沿線の駅	新宿 笹塚 代田橋 明大前 下高井戸 桜上 水 上北沢 八幡山 芦花公園 千歳烏山 仙 川 つつじヶ丘 柴崎 国領 布田 調布 西 調布 飛田給 武蔵野台 多磨霊園 東府中 府中 分倍河原 中河原 聖蹟桜ヶ丘 百草 園 高幡不動 南平 平山城址公園 長沼 北 野 京王八王子

また、各駅周辺の街路ネットワークを抽出する範囲を次のように定めた。課題の節で述べたように、本研究で作成する推薦システムのユースケースを「引越し先の推薦」と定める。そのため、類似度を測る街の範囲を「駅から徒歩10分以内」と設定する。不動産業界の慣習では一般的に「徒歩1分=80m」として計算することが多い。そのため本稿では、駅から800mを分析の対象とする。

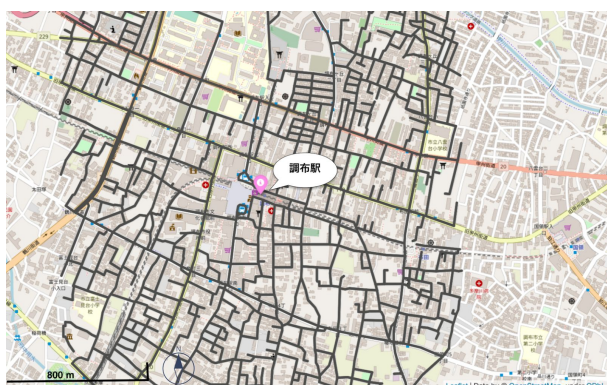


図1 調布駅から800m圏内の街路ネットワーク

2.4. データソース

街路のデータを取得するために OpenStreetMap のデータを利用した。OpenStreetMap はインターネット上で無料で公開されている地図データであり、API やライブラリを使用することで誰でも簡単に地図データを取得することができる。さらに街路データをグラフネットワークとして扱うために、OSMnx という Python 言語のライブラリを利用した。OSMnx を使うと、街路ネットワーク内のノード数やエッジ数、平均距離などの基本的な統計量を簡単に計算し取得することができる。例えば、OSMnx の `basic_stats` 関数を用いると、下記のような統計量を取得することができる。ここでは誌面の都合上、本稿で言及する変数の紹介のみに留める。

表2 OSMnx から得られる統計量の抜粋

n	グラフ内のノード(交差点)数
m	グラフのエッジ(道)の数
circuitry_avg	各エッジのノード間の大圏距離の合計で割ったエッジ長の合計
streets_per_node_avg	各ノードから平均していくつの通りが繋がっているか

本稿では OSMnx から得られるグラフの基本統計量を説明変数として利用した。

また、OSMnx を用いて街路ネットワークデータを取得する際に各駅の緯度経度情報が必要になる。そのため、各駅の緯度経度情報を利用するために駅データ.jp (<https://www.ekidata.jp/>) というサイトから駅データの CSV ファイルをダウンロードした。



図2 中央線沿線および京王線沿線の駅の位置関係

利用したデータの一部を表3に抜粋した。また、本稿での実験の対象とした中央線および京王線の駅を地図上にプロットすると図2のようになる。

以降の実験では、取得した駅データの中から中央線と京王線に該当する行だけを取り出し、OSMnxによる街路ネットワークに付加して作成したデータを用いて検証を行った。

表3 駅データ。jpから取得した駅データの抜粋

station_name	line_cd	address	lat	lon
東京	11312	東京都千代田区丸の内一丁目	35.681391	139.766103
神田	11312	東京都千代田区鍛冶町二丁目13-1	35.691173	139.770641
御茶ノ水	11312	千代田区神田駿河台2丁目	35.699605	139.764955

3. 実験結果

3.1. 決定木モデルによる有効な説明変数の抽出

2つ沿線の特徴をよりよく説明する変数を調べるために、決定木モデルによる学習を行った。説明変数にはOSMnxから取得できる基本統計量を用

いた。対象とした駅が中央線に属するのか京王線に属するのかを目的変数として学習を行った。モデルの学習後、学習された分類ルールを可視化すると図3のような結果が得られた。

図3を見ると、X[14]とX[4]という変数が分類の初期に用いられていることがわかる。X[14]はOSMnxによる統計量の一つであるcircuitry_avgという変数であり、X[4]はstreets_per_node_avgという変数である。中央線か京王線かを分類するためにこの二つの変数が有効だと判断し、次に説明するK-Means法によるクラスタリングで用いることにする。

3.2. K-Means法によるクラスタリング

K-Means法によるクラスタリングを試みた。その結果、次のような結果が得られた。

表4 K-Means法によって予測された駅と実際の駅のクロス集計表

	実際の中央線沿線駅の数	実際の京王線沿線駅の数
中央線沿線と予測された駅の数	12	3
京王線沿線と予測された駅の数	12	29

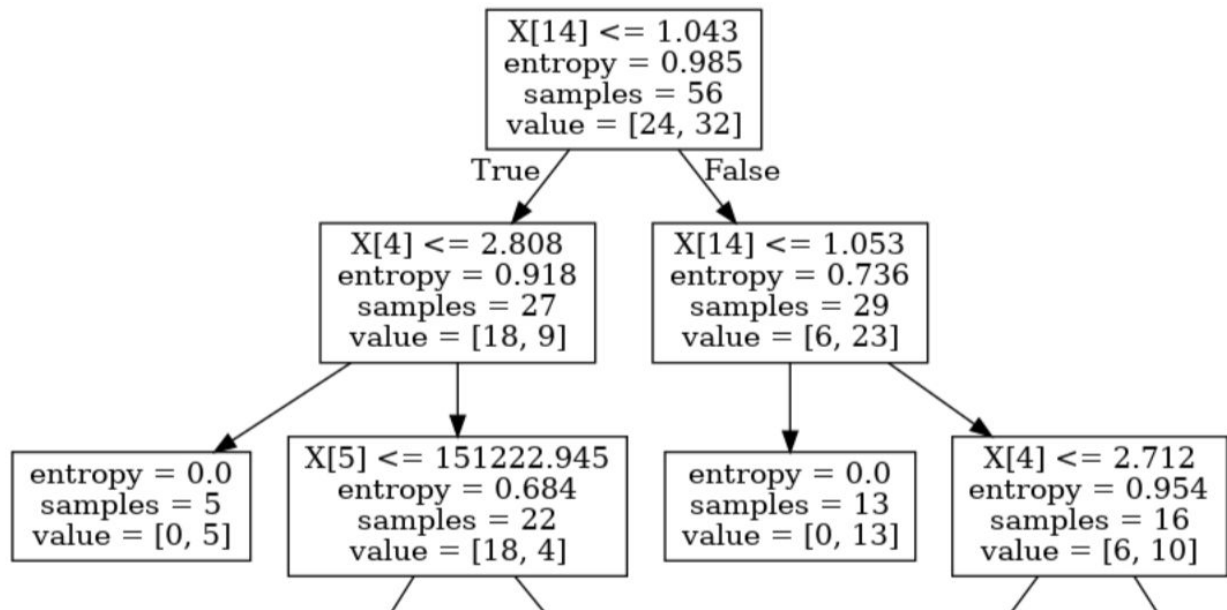


図3 決定木モデルによって抽出された分類ルール

京王線沿線と予測された駅の正解率に着目すると、 $12 / (12 + 29) \times 100 = \text{約}29\%$ とあまりうまく分類できていない。一方で中央線沿線と予測された駅の正解率に着目すると、 $12 / (12 + 3) \times 100 = \text{約}80\%$ と高いことがわかる。さらに、分類された駅の内訳を見てみると興味深い結果が得られた。表5はK-Means法によって分類された各駅の一覧である。

表5 K-Means法によって予測された中央線沿線および京王線沿線の駅の一覧

中央線沿線と予測された駅	東京 神田 御茶ノ水 新宿 高円寺 西荻窪 吉祥寺 三鷹 立川 豊田 八王子 西八王子 新宿 府中 京王八王子
京王線沿線と予測された駅	四ツ谷 中野 阿佐ヶ谷 荻窪 武蔵境 東小金井 武蔵小金井 国分寺 西国分寺 国立 日野 高尾 笹塚 代田橋 明大前 下高井戸 桜上水 上北沢 八幡山 芦花公園 千歳烏山 仙川 つつじヶ丘 柴崎 国領 布田 調布 西調布 飛田給 武蔵野台 多磨霊園 東府中 分倍河原 中河原 聖蹟桜ヶ丘 百草園 高幡不動 南平 平山城址公園 長沼 北野

注) 太字は中央線の特別快速停車駅

中央線沿線であると判断された駅には、特別快速停車駅が多く分類されている。circuitry_avgおよびstreets_per_node_avgの二つの変数は中央線の特別快速停車駅に見られる特徴を説明している可能性があるといえる。さらに、この分類された中に府中駅と京王八王子駅が分類されている。この結果は府中駅および京王八王子駅が、中央線の特別快速停車駅と特徴が類似している可能性があるといえる。

4. まとめと今後の研究課題

中央線および京王線の各駅から街路構造データを取得し、circuitry_avgとstreets_per_node_avgの二つの変数を利用してクラスタリングを行った。その結果、中央線と分類された駅に着目すると、約80%程度の正解率で中央線特別快速が停車する駅とその他の駅を分類することができた。しかし、その中には京王八王子駅と府中駅のように、中央線の特別快速停車駅として誤って分類されている駅も存在する。これは、誤って分類された駅が中央線の特別快速停車駅と何らかの似た特徴を持っている可能性を示唆している。

また、本稿で行った実験にはいくつかの改善点がある。本稿で取り扱った以外の路線、例えば小田急線などを分類対象として追加することが考えられる。これらは、今後の研究課題としたい。

参考文献

- 西井和夫・土井勉・木内徹・三浦啓江（1995）
LOGMAPモデルの鉄道沿線イメージ構造分析
への適用 京王線路線間比較「都市計画論文
集」， **30**， 247-252.
- 井出策夫（1967）中央線沿線における都市化。
「新地理」， **15**（**3**）， 1-20.
- 福島達夫（1975）多摩ニュータウンと多摩市：首
都圏における巨大住宅都市の形成過程(1)。
「経済地理学年報」， **21**（**3**）， 22-36.