

各駅の相互作用に着目した駅併設複合施設の 利用状況の評価手法に関する考察

佐々木 竜也*¹ 池谷 風馬*² 長谷川 悠*² 大口 敬*²

東京大学大学院 工学系研究科*¹

東京大学生産技術研究所*²

近年、駅やその周辺に商業機能を持った複合施設が設置されており、複合施設に必要とされる機能に注目が集まっている。本研究では鉄道利用者の属性や付近の駅に存在する別の複合施設の影響に着目した複合施設の評価手法を提案する。交通系 IC カード利用履歴から得られる個人属性や、複合施設の調査から得られる各駅の複合施設の店舗数といった情報を統合し、非負値テンソル因子分解を適用することでパターン抽出を行った。検証の結果、シナリオに沿うように意図的に内容を偏らせたダミーデータから抽出されたパターンは、シナリオに沿うもの、沿わないもの、更に細かく分岐したものを表現できており、本手法が複合施設の評価に有用であることが示唆された。

Discussion on evaluation methods for the use of station-anchored complexes, focusing on the interaction of each station

Tatsuya SASAKI*¹ Fuma IKEYA*² Yu HASEGAWA*² Takashi OGUCHI*²

Graduate School of Engineering, The University of Tokyo*¹

Institute of Industrial Science, The University of Tokyo*²

In recent years, complex facilities having commercial functions have been installed in and around stations, and the functions required for these complex facilities are drawing attention. This study proposes an evaluation method for complexes that focuses on the attributes of railway users and the influence of other complexes existing at nearby stations. Patterns were extracted by applying Non-negative Tensor Factorization to integrating data obtained from smart card data and survey of complex facility. Validation using dummy data is intentionally biased to follow the two scenarios that show that the proposed approach is able to represent each of the characteristics, demonstrating its effectiveness.

Keyword: station facilities, smart card data, pattern extraction

1. はじめに

近年、駅に商業施設を複合することや駅周辺にこれらを設置することがある（以後、このような施設を複合施設と呼ぶ）。複合施設に入居するテナントの種類についても多様化しており、様々な機能が求め

られている。また、利用者の数が膨大な駅における大規模な複合施設だけでなく、様々な複合施設が設置されている。そして、各駅に設置されている複合施設同士の相互の影響もあると考えられるとともに、人々がどの駅でどの複合施設を利用するのか、とい

った行動についても複雑化していると考えられる。

ここで、少子高齢化や環境への配慮から「コンパクト・プラス・ネットワーク」という都市のあり方が近年注目されており、その考え方に基づく都市の核となるのが鉄道の駅である。したがって、駅に商業をはじめとする様々な機能をもたせた複合施設についても、重要な役割を持っていると言える。都市のあり方を検討していくにあたり、複合施設の実態を適切に評価することで、どのような駅にどのような複合施設があるべきなのかを考える必要がある。

さて、複合施設の評価手法に関する研究はいくつか存在する。大抵は小田急電鉄 成城学園前駅の駅ビルを事業者へのヒアリングや質問票を用いて評価した。その結果、地元商店街との競合回避および世田谷区という広い商業圏域での差別化のため、生活関連施設や子育て層に向けた施設といった「地域拠点機能」が充実されていることを示した。横田ら²⁾は近畿圏のある鉄道事業者の 295 駅に対して現地調査等を行い、駅構内の施設立地が駅近傍の施設発生を抑制していることや乗降人員、乗換タイプ、用途地域、駅形態が施設分布と関係が強いことを示した。越後ら³⁾は都市空間データと PT 調査データを用いて、鉄道駅を中心とするエリアの活動地選択モデルを構築した。その結果、詳細な建物用地・延床面積が私事の活動目的の推定に有用であることなどを示した。

一方で、これら既往研究では利用者個人の属性や、分析対象の駅・複合施設とは別の駅に設置されている複合施設との関係については十分に考慮されていない。そこで本研究では交通系 IC カード利用履歴を用いることで複合施設のある駅に訪れた鉄道の利用者の属性に着目するとともに、周辺の駅に存在する複合施設について調査・集計した結果を勘案した、複合施設の評価手法を提案する。また、評価方法は非負値テンソル因子分解を用いたパターン抽出を採用する。非負値テンソル因子分解とは、商業施設における購買行動データのように、時刻や場所、属性などの膨大な情報量から特徴的なパターンや特徴量同士の関係を明らかにすることが可能な手法である。

2. IC カードデータを用いた複合施設の評価手法

交通系 IC カード利用履歴を用いた複合施設の評価手法の大まかな流れを図-1 に示す。まず、対象とする路線の各駅において存在する複合施設について調査を行う。その後、交通系 IC カード利用履歴と複合施設の調査結果を統合したデータを作成する。そ

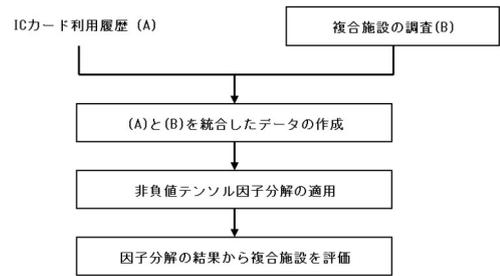


図-1 複合施設の評価の手順

して、このデータに対して非負値テンソル因子分解を適用することによってパターン抽出を行う。これにより、鉄道利用者の属性や近隣の駅に存在する複合施設との関係性を勘案した複合施設の評価が可能となる。そして最後にその結果を考察する。

1-1 IC カードデータ

本研究では次に示すような情報を持つ交通系 IC カード利用履歴を想定する（以後、この形式の交通系 IC カード利用履歴を IC カードデータと呼ぶ）。

- (1) IC カードの識別 ID
- (2) 入場駅の情報
- (3) 入場駅を運営する事業者の情報
- (4) 入場の日時
- (5) 退場駅の情報
- (6) 退場駅を運営する事業者の情報
- (7) 退場の日時
- (8) IC カード保持者の年齢層
- (9) IC カード保持者の性別
- (10) IC カード保持者が学生かどうかの判定

ここで、IC カードデータは一般に公開される例は少ないため、本研究では IC カードデータを再現した自作のダミーデータを用いる。-

1-2 複合施設の調査

複合施設の評価手法の提案にあたり、複合施設そのものに関する次の 5 個の特徴量について調査するとともに集計を行う。

- (1) 最寄駅
- (2) 最寄駅からの所要時間
- (3) 店舗数
- (4) 店舗のジャンル数
- (5) 複合施設店舗数が最大の近隣駅の合計店舗数

まず、複合施設の立地に関して(1)最寄駅、(2)最寄駅からの所要時間 について、Google Map の経路検索を用いて算出する。そして複合施設の基本情報の(3)店舗数、(4)店舗のジャンル数 について、各施設管理者がインターネット上に公開する施設のプロ

アマップから各店舗をカウントする。ここで、店舗のジャンル数とは商業施設のテナントとなりうる小売業およびサービス業である。具体的には表-1に示すように、日本標準産業分類⁴⁾を基に店舗を12のジャンルに分類したものである。また、対象の駅の利用に対し、付近の駅の複合施設からの影響を検証するため、(5)複合施設店舗数が最大の近隣駅の合計店舗数を考える。これは分析対象駅から数えて5駅以内の駅(以後、このような駅を近隣駅と呼ぶ)のうち(3)店舗数が最も多いものとする。そして、複合施設が駅併設か否かという点に着目し、最寄駅からの所要時間が徒歩10分以内の複合施設を対象とした。

1-3 非負値テンソル因子分解

行動データのような情報量が大きくなりやすいデータの分析においては、ベクトルや行列の形式での表現には限界があり、高次元の表現が可能なテンソルを用いることがある⁵⁾。このような高次元のデータの特徴を保持したままパターン抽出が可能な手法として近年利用されているのが非負値テンソル因子分解である。本手法は購買行動⁶⁾や観光行動⁷⁾、交通

行動⁸⁾の分析をはじめとする多くの分野で用いられている。本研究では鉄道利用と複合施設を統合した次数の大きいデータを分析することから、非負値テンソル因子分解を用いたパターン抽出による複合施設の評価手法の提案を試みる。

図-2に示す非負値テンソル因子分解の概念図のような、 $I \times J \times K$ の大きさを持つ3次テンソル $\mathbf{X} = [\hat{x}_{ijk}] \in \mathbb{R}_+^{I \times J \times K}$ を考える。3次テンソル \mathbf{X} をランク数を R とした3つの行列 $\mathbf{U} = [u_{ir}] \in \mathbb{R}_+^{I \times R}$ 、 $\mathbf{V} = [v_{jr}] \in \mathbb{R}_+^{J \times R}$ 、 $\mathbf{W} = [w_{kr}] \in \mathbb{R}_+^{K \times R}$ の内積として分解する。 $\mathbf{U}, \mathbf{V}, \mathbf{W}$ は各要素の因子行列であり、因子行列 $\mathbf{U}, \mathbf{V}, \mathbf{W}$ を用いて次に示す式(1)のように

$$\hat{\mathbf{X}} = [\hat{x}_{ijk}] = \mathbf{U} \otimes \mathbf{V} \otimes \mathbf{W} \quad (1)$$

として得られるテンソル $\hat{\mathbf{X}}$ に対して、

$$D = \sum_i \sum_j \sum_k d(x_{ijk}, \hat{x}_{ijk}) \quad (2)$$

を最小化するように、 $\mathbf{U}, \mathbf{V}, \mathbf{W}$ を求める。ここで、 $d(x_{ijk}, \hat{x}_{ijk})$ は距離関数を表しており、一般的にこの距離関数としては、信号処理などの数値の取りうる値が大きい場合には一般化KLダイバージェンスが、5段階評価点などの値のばらつきが小さいデータにはユークリッド距離が用いられる。よって本研究では一般化KLダイバージェンスを用いる。非負値テンソル因子分解はこれらの因子行列 $\mathbf{U}, \mathbf{V}, \mathbf{W}$ の各要素の非負性を保ったまま、 D を最小にする $\mathbf{U}, \mathbf{V}, \mathbf{W}$ を求める手法である。

ここでは3次テンソルを考えたが、本研究では表-2に示す特徴量を持つ10次テンソルを用い、計算量削減のため、分解する因子の数は3つとした。

表-1 店舗のジャンルの分類

カテゴリ	日本標準産業分類 産業分類コード一覧(小分類)	産業分類内容
① 百貨店	各種商品小売業	百貨店(全分類)
② 衣料品店	繊維・衣服・身の回り品小売業	(全分類)
③ 雑貨屋	各種商品小売業 その他の小売業 その他の小売業	均一価格店 書籍・文具小売業 他に分類されない小売業
④ ブランド品店	その他の小売業	写真機・時計・眼鏡小売業
⑤ スポーツ用品店	その他の小売業	スポーツ用品
⑥ 食料品店	各種商品小売業 各種商品小売業 飲食料品小売業	スーパーマーケット コンビニエンスストア(全分類)
⑦ 医薬品・化粧品店	各種商品小売業 その他の小売業	ドラッグストア 医薬品・化粧品小売業
⑧ 医療機関	医療業 保健衛生 社会保険・社会福祉・介護事業	(全分類) (全分類) (全分類)
⑨ 美容・フィットネス店	洗濯・理容・美容・浴場業 娯楽業 スポーツ施設提供業	(全分類) (全分類)
⑩ 娯楽業	娯楽業(スポーツ施設提供業を除く)	(全分類)
⑪ 飲食店	飲食店 持ち帰り・配達飲食サービス業	(全分類) (全分類)
⑫ その他サービス業	無店舗小売業 銀行業 保険業 不動産取引業 その他の生活関連サービス業 学校教育 その他の教育、学習支援業	(全分類) (全分類) (全分類) (全分類) (全分類) (全分類) (全分類)

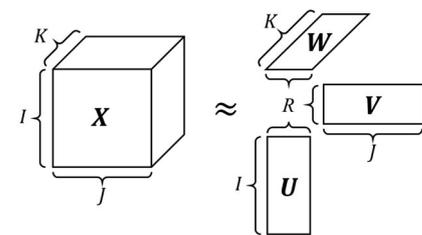


図-2 非負値テンソル因子分解の概念図

表-2 特徴量の詳細

項目名	内容	分類
(1) 祝日ダミー	ICカードデータにおける利用日時が、祝日であるか祝日でないかを表現したダミー変数	①0. ②1
(2) 時間帯	ICカードデータにおける分析対象駅への到着時間を、次に示す6つの分類に基づいて表現したもの	①16:00から8:59. ②9:00から11:59. ③12:00から14:59. ④15:00から17:59. ⑤18:00から20:59. ⑥21:00から23:59
(3) 滞在時間	ICカードデータにおける分析対象駅への到着時間から次の利用の出発時間までの時間を、次に示す4つの分類に基づいて表現したもの	①10分未満. ②10分以上30分未満. ③30分以上60分未満. ④60分以上
(4) 年齢層	ICカードデータにおける利用者の年齢層を、次に示す5つの分類に基づいて表現したもの	①20歳未満. ②20歳以上40歳未満. ③40歳以上60歳未満. ④60歳以上80歳未満. ⑤80歳以上
(5) 性別	ICカードデータにおける利用者の性別を、次に示す2つの分類に基づいて表現したもの	①男性. ②女性
(6) 学生ダミー	ICカードデータにおける利用者が、学生かであるか学生でないかを表現したダミー変数	①0. ②1
(7) 出発駅から分析対象駅までの所要時間	ICカードデータにおける出発駅から到着駅(分析対象駅)までの所要時間を、次に示す4つの分類に基づいて表現したもの	①10分未満. ②10分以上30分未満. ③30分以上60分未満. ④60分以上
(8) 全複合施設の店舗数の合計	先述の複合施設の調査における、分析対象駅のすべての複合施設の店舗数の合計を、次に示す5つの分類に基づいて表現したもの	①50未満. ②50以上100未満. ③100以上400未満. ④400以上1000未満. ⑤1000以上
(9) 全複合施設が網羅するジャンル数	先述の複合施設の調査における、分析対象駅の複合施設が網羅するジャンルの数を、次に示す3つの分類に基づいて表現したもの	①4以上7未満. ②7以上9未満. ③9以上
(10) 複合施設店舗数が最大の近隣駅の合計店舗数	先述の複合施設の調査における、分析対象駅の近隣の駅のうち、すべての複合施設の店舗数の合計が最も多い近隣の駅のすべての複合施設の店舗数の合計を、次に示す5つの分類に基づいて表現したもの	①10分未満. ②10分以上30分未満. ③30分以上60分未満. ④60分以上

3. ダミーデータを用いた検証

3-1 ダミーデータの作成

手法の検証として、合計2つのシナリオに基づいたダミーデータに対する本手法の適用結果が、各シナリオに対応しているかを考察する。検証の対象は図-3に示す阪急電鉄の路線を対象とする。また、デ

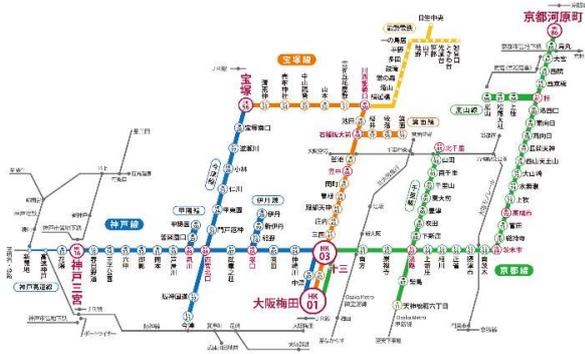


図-3 阪急電鉄の路線図

表-3 対象地域の複合施設の調査結果

駅	店舗数	ジャンル数	近隣駅店舗数
大阪梅田*	2715	12	123
園田	31	7	2715
塚口	123	9	2715
西宮北口*	586	12	156
夙川	92	9	586
神戸三宮*	370	12	0
三国	31	4	2715
庄内	24	7	2715
曾根駅	15	7	261
豊中	40	6	261
蛍池	81	9	261
池田	68	9	261
川西能勢口*	261	12	68
売布神社	52	7	261
宝塚*	186	11	188
淡路	3	3	2715
茨木市	52	7	58
高槻市	58	8	52
洛西口	24	4	73
桂	22	7	432
西院	12	5	432
烏丸	73	9	432
河原町*	432	10	73
稲野	156	9	586
伊丹	115	8	586
今津	16	6	586
甲東園	14	3	586
逆瀬川	188	9	586
宝塚南口	17	8	186
千里山	6	5	68
南千里	68	8	67
山田	30	8	68
北千里	67	8	68

備考：*は複合施設が百貨店を含むことを表す。また、百貨店の店舗数については、それ以外の複合施設の店舗数と店舗面積を回帰分析した次式と店舗面積から推定する。

$$\text{店舗数} = 5.263 \times 10^{-3} \times \text{店舗面積} [\text{m}^2] + 10$$

表-4 ダミーデータのフォーマット

祝休日	時間帯	滞在時間	年齢層	性別	学生	出発駅から	店舗数	ジャンル数	近隣駅店舗数
0	6:00~8:59	10分未満	20歳以上40歳未満	女性	1	30分以上60分未満	50未満	4以上7未満	50以上100未満
1	18:00~20:59	10分以上30分未満	80歳以上	男性	0	10分未満	50以上100未満	7以上9未満	100以上400未満
0	12:00~14:59	60分以上	40歳以上60歳未満	女性	0	10分以上30分未満	400以上1000未満	9以上	100以上400未満

ータの制約上、本研究では第13回大都市交通センサス⁹⁾を用いてICカードデータを再現する。そして、大都市交通センサスには利用者の属性や鉄道を利用した時間についての記載はないため、厚労省の人口動態統計¹⁰⁾や国交省の鉄道利用に関する統計¹¹⁾を参考に利用者属性や各時間帯の利用を確率変数として記述し、その確率に基づき大都市交通センサスに情報を付加した。更に、このデータに表-3に示す対象地域の複合施設の調査結果を統合し、結果的に得られた表-4に示す形式のデータに対して非負値テンソル因子分解を適用する。

3-2 シナリオ1への適用

1つ目のシナリオとして、40歳以上の女性の利用が多い状況を想定し、40歳以上60歳未満の女性、60歳以上80歳未満の女性、80歳以上の女性の利用割合がそれぞれ21.67%、それ以外の属性の利用割合がそれぞれ5%となる確率変数に基づいたデータセットを作成し、非負値テンソル因子分解を適用した。

その結果、図-4に示すように、得られた3つの因子ともに、年齢層、性別の特徴量は主に40歳以上の女性を表現している。ここで他の特徴量を確認することでパターン別の40歳以上の女性の利用について考察することができる。具体的には図-4より、因子1, 2, 3はそれぞれ店舗数が50未満、50以上400未満、100以上の複合施設を持つ駅への利用を表現している。また、店舗数が50未満、50以上400未満といった中・小規模の複合施設を持つ駅への利用は女性の割合が多いといった傾向や、その近隣駅にはそれよりも大きな規模の複合施設があることが捕捉できる。ここで因子1および因子2に見られる傾向は40歳以上の女性に見られる特徴であるが、因子3に見られる傾向については若年層をはじめとする学生や男性の特徴として捕捉が可能である。

3-3 シナリオ2への適用

2つ目のシナリオとして、近隣駅に大型の複合施設があることを想定し、複合施設店舗数が最大の近隣駅の合計店舗数について、1000以上である割合が80%、それ以外の割合がそれぞれ5%となるような確率変数に基づいてデータセットを作成し、非負値テンソル因子分解を適用した。

その結果、図-5に示すように、得られた2つの因子ともに、複合施設店舗数が最大の近隣駅の合計店舗数が1000以上である駅での利用を表現している。ここで他の特徴量を確認することでパターン別の近隣駅に大規模複合施設がある場合の利用について考

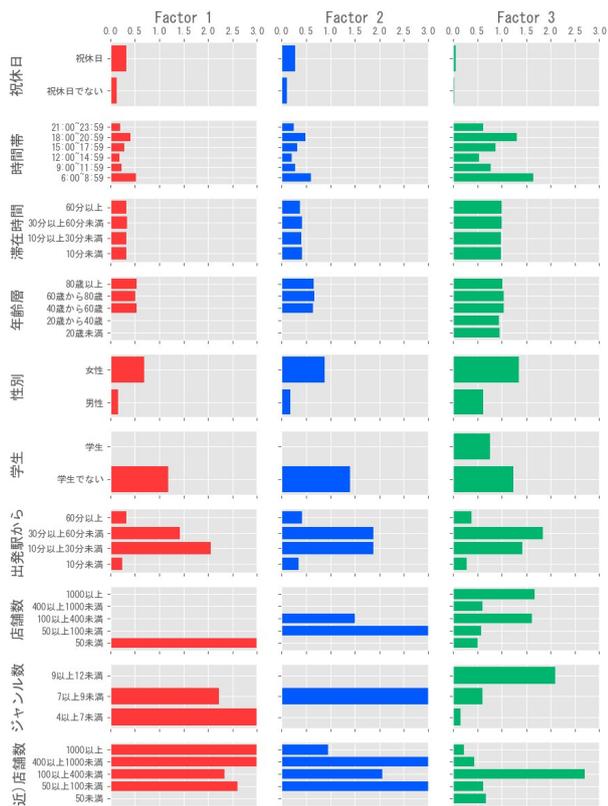


図-4 シナリオ 1 への適用結果

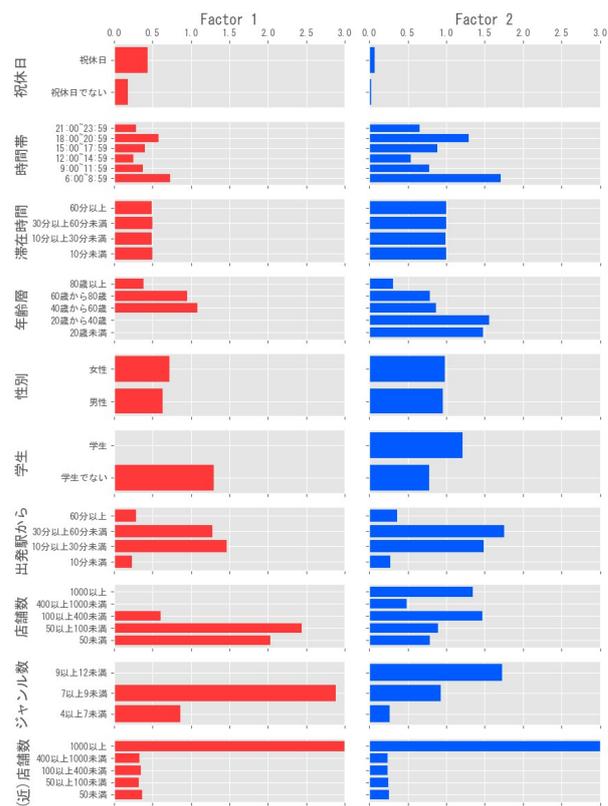


図-5 シナリオ 2 への適用結果

察することができる. 具体的には図-5 より, 因子 1, は利用者の年齢が 40 歳以上かつ店舗のジャンル数が 9 未満の複合施設を持つ駅への利用, 因子 2 は主に学生を含む 0 歳から 80 歳未満の利用かつ, 店舗のジャンル数が 7 以上の複合施設を持つ駅への利用が表現されている.

3-4 手法の適用結果

シナリオ 1,2 への非負値テンソル因子分解の適用から, 分析対象のデータが持つ利用者属性や複合施設に関する意図的な偏りが, 結果として得られるパターンに反映されている. このことから IC カードデータ, 複合施設の調査結果を用いた, 非負値テンソル因子分解による複合施設の評価は効果的であることが示唆された.

4. 結論

本研究では, 交通系 IC カード利用履歴を用いることで複合施設のある駅に訪れた鉄道の利用者の属性に着目するとともに, 周辺の駅に存在する複合施設を考慮に入れた複合施設の評価手法の提案を行った. IC カードデータおよび複合施設の調査を統合したデータに対して, 非負値テンソル因子分解を適用することでパターン抽出をし, その結果から分析対象

の駅の利用に関して, 利用者属性や近隣駅に存在する複合施設の影響について考察した. 具体的には評価の検証において, 2 つのシナリオ別に作成した偏りのあるデータに対して手法を適用することにより, シナリオに基づいた評価ができていたかを考察した. ここで, データの制約上, IC カードデータについては大都市交通センサスなどを用いて再現した. 結果的にシナリオを用いた検証のすべてにおいて, 因子分解の結果が意図的に偏らせたデータの特徴を補足していることや, その偏りの中において更に詳細に分岐したパターン, 偏りの対象でなかったデータの傾向について本手法は表現ができることが示された. 具体的には, シナリオ 1 の 40 歳以上の女性の利用を意図的に多くしたデータの因子分解の結果, 彼女らの利用が複合施設の店舗数の大小別にパターン抽出され, これを更に細かく分解して見ると複合施設の店舗数が 50 未満といった場合には, 近隣駅最大の複合施設店舗数が 50 以上, 中でも特に 400 以上である傾向が強いことを補足した. シナリオ 2 の近隣駅に大規模な複合施設があることを想定したデータの因子分解においても, 比較的高い年齢層の利用者は規模の小さい複合施設のある駅を利用し, 比較的低い年齢層の利用者は規模の大きい複合施設のある駅の

利用を行うといった傾向を補足した。以上より、ICカードデータおよび複合施設の調査結果を統合したデータに非負値テンソル因子分解を適用することで、既存の複合施設の評価手法では十分に考慮されていなかった利用者の属性や近隣の駅に存在する複合施設の影響を、複合施設の評価に加味できることが示唆された。

今後の課題としては、実際のICカードデータを用いたケーススタディや実証分析を行うことで、実データへの適用の可否を検証する必要がある。また、複合施設の利用は通勤通学などの定期的な鉄道利用と比較して、そうでない非定期的な鉄道利用を通して行われると推測できるため、ICカードデータから非定期的な利用を抽出したデータに手法を適用する必要性についても示唆される。そして、今回は阪急電鉄の路線のみを対象としたが、現実的には付近の別の路線の影響や、別の路線の駅に存在する複合施設の影響もあると考えられるため、その点についても評価に加味する必要がある。更には、ICカードデータだけでは複合施設を利用したかどうかを断定することはできないため、決済履歴をはじめとする複合施設での利用者の行動がわかるようなデータを用いることも有用であると考えられる。

REFERENCES

- 1) 大家亮子：「地域拠点機能」を発揮する地域貢献型・提案型の駅ビルづくりに関する研究 成城学園前駅の事例を対象に、都市計画論文集，45 巻，pp. 1-8，2010.
- 2) 横田隆司，柏原土郎，吉村英祐，阪田弘一：鉄道駅構内における地域施設の複合化実態と鉄道駅への施設成立可能性の評価モデルに関する研究，日本建築学会計画系論文集，65 巻，538 号，pp. 117-123，2000.
- 3) 越智健吾，石井良治，末木祐多，羽藤英二，筒井祐治：都市計画基礎調査・PT 調査データ活用による鉄道駅を中心とした私事目的別の活動地選択モデルの開発 コンパクト・プラス・ネットワーク政策評価のための実務的なモデル構築とケーススタディによる検証，都市計画論文集，56 巻，2 号，pp. 360-367，2021.
- 4) 総務省，日本標準産業分類（令和 5 年 7 月告示），2023. (https://www.soumu.go.jp/toukei_toukatsu/index/seido/sangyo/R05index.htm)
- 5) 久保基，田中宏季，中村哲：非負値テンソル因子分解を用いた観光行動データからの情報抽出，マルチメディア，分散，協調とモバイルシンポジウム 2019 論文集，pp. 1259-1263，2019.
- 6) 松林達史，幸島匡宏，林亜紀，澤田宏：非負値テンソル因子分解を用いた購買行動におけるブランド選択分析，人工知能学会論文誌，30 巻，6 号，pp. 713-720，2015.
- 7) 熊谷雄介，今井良太，松林達史，佐藤吉秀，堀岡力：非負値複合テンソル因子分解を用いた訪日外国人観光客の回遊行動分析，研究報告バイオ情報学（BIO），3 巻，pp. 1-5，2015.
- 8) 細江美欧，桑野将司，谷本圭志：非負値テンソル因子分解を用いた交通系 IC カードデータからの移動パターンの抽出に関する研究，都市計画論文集，53 巻，3 号，pp. 1320-1326，2018.
- 9) 国土交通省，第 13 回大都市交通センサス，2023. (<https://www.e-stat.go.jp/stat-search/files?page=1&toukei=00600020&tstat=000001103355>)
- 10) 厚生労働省，令和 3 年人口動態統計，2021. (<https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/kakutei21/>)
- 11) 国土交通省，第 VII 章 利用時刻別にみた鉄道利用者数の推計. (<https://www.mlit.go.jp/common/001001534.pdf>)