

A11 公共交通ビッグデータを用いた 優良顧客の発見手法に関する研究

公共システム研究室 廣畑健也

1. はじめに

公共交通は通勤や通学、買い物など、人々の日常的な移動には欠かせないサービスである。しかし地方では、少子高齢化や人口減少の影響に伴って公共交通の利用者が従来に比べて少なくなっている。近年では交通 IC カードの普及が進み、そこで得られるビッグデータを活用したマーケティングが可能となっており、利用者数の確保に資する対策が期待されている。そこで本研究では、香川県の高松琴平電気鉄道を対象に、長期の利用履歴に基づいて顧客をクラスタリングすることで優良顧客の特性を明らかにし、その上で、各地域で予測される顧客数と実際の利用者数を比較することで、どの地域で優良顧客の発見が可能かを特定する手法を開発する。

2. 分析方法

本研究ではこれまでの利用履歴を用いて、利用頻度の高い優良顧客とその他の一般的な顧客に分類するためにクラスタ分析を用いる。次いで、その分析結果から得られる各利用者の所属クラスタと年齢から、クラスタ別の年齢分布を求める。その分布から、クラスタ c における年齢 a の確率 $P(a|c)$ を求め、この確率と各クラスタの全体数に占める割合 $P(c)$ から、年齢別の各クラスタ所属確率 $P(c|a)$ をベイズの公式を用いて式(1)より導出する。なお、 K はクラスタ数である。任意の地域 j における年齢 a の人口を $N_j(a)$ で表すと、各地域 $P(c|a)N_j(a)$ によってクラスタ c の利用者数が推計できる。また、各地域においてそもそも利用しない可能性を考慮するために地域 j における利用率 λ_j を予測に乗じる。以上より、地域 j におけるクラスタ c の利用者数の予測 $u_j(c)$ は式(2)より求められる。その上で、各地域で利用者数の予測と実績の比と差の 2 つの観点で比較し、優良顧客が見込める地域を特定する。すなわち、 $v_j(c)$ を地域 j におけるクラスタ c の実績の利用者数とすると、そこで、 $u_j(c)-v_j(c)$ 、ならびに $u_j(c)/v_j(c)$ の 2 つの指標で更なる利用が見込める地域を特定する。

$$P(c|a) = \frac{P(a|c)P(c)}{\sum_{c=1}^K P(a|c')P(c')} \quad (1)$$

$$u_j(c) = \sum_a P(c|a)N_j(a)\lambda_j \quad (2)$$

3. 分析結果

クラスタ分析の結果を図 1 に示す。この図より、利用回数が多いもしくは増加している、クラスタ 1, 3, 5, 8, 9 を優良顧客とする。

高松琴平電気鉄道の沿線市町を対象に、大字で区分される 365 の地域で 15 歳以上を対象に優良顧客の予測と実績を比較する。各クラスタに関して地域ごとの予測の利用者数と利用者数の実績の比及び差の 2 つの観点で比較をした。なお、比が 1 以上で差が 0 以上の地域をさらなる利用の見込みがある地域とし、その中でも比が 2 以上かつ、差が 10 人以上であった地域を特に利用が見込める地域として取り上げる。利用の見込みがあると特定された地域を図 2 に示す。なお、地図上には 3 本の路線ならびに黒色の点で駅を示している。この図から、各クラスタに関して、利用が見込める地域はどこかという情報が一目でわかる結果が得られ、これらの地域を対象に重点的に利用の促進を行なうことが効果的である。

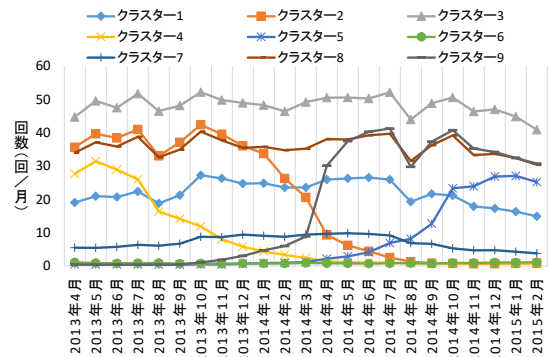


図 1 クラスタ分析の結果

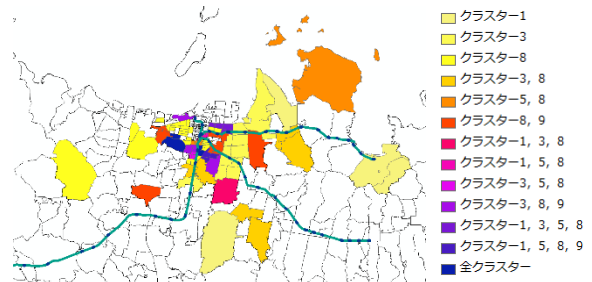


図 2 特に利用が見込める地域