

道路毎のタクシー乗客需要予測の道路粒度の分析

甲斐 大雅† 神山 剛† 石黒 慎† 山田 尚志†

長崎大学 情報データ科学部†

1. はじめに

近年のタクシー業界では全国的にドライバー不足が進行している[1]. 特に、乗客を獲得するノウハウや経験があるドライバーは高齢化などにより減少傾向にあり、一方で、ドライバー育成の時間やコストから、新人ドライバーを増やすことが困難な状況にある。

この問題を解決すべく、ドライバーの乗客獲得を支援するタクシー乗客需要予測サービスが提供されている.[2]. このサービスでは、タクシー運行エリア全体を500メートル四方ごとに区切り、領域ごとに獲得が見込まれる乗車数を予測し、可視化している[2].

しかし、この空間粒度での可視化は、ドライバーが回遊もしくは待機する対象の道路を具体的に示すものではない. ドライバーの経験値に依存しない顧客獲得を実現するためには、さらに細かな粒度での可視化が必要と考える。

本研究では、より詳細な粒度での可視化を実現すべく、道路レベルでの需要予測を行うAIの実現を目的とする。

道路毎の需要予測AIの作成には、過去の乗車実績データから、データの集計単位となる道路毎に十分な乗車数やその傾向が確認できる必要がある. しかし、集計単位となる道路が短すぎると、傾向を確認できるだけのデータが不足する可能性がある。

以上より、本稿では、実際のタクシー運行実績データを用いた道路レベルでの乗車実績の傾向分析を行うことで、学習・予測が可能な道路粒度を決定することを課題とする。

2. 道路粒度の検討

前述したように、道路毎の需要予測を実現するためには、乗車数の集計単位となる道路の粒度が重要である. 仮に道路を細かく区切り過ぎた場合、その道路に紐づく乗車数が少なくなり学習が困難となる. 一方、道路が長すぎる場合、学習には十分なデータが集まるが、予測の対象空間が広がってしまう. よって、学習可能かつ、できるだけ細かな粒度を考える必要がある。

Spatial Analysis of Taxi Demand for Road-level Demand Forecast

Kai Taiga†, Kamiyama Takeshi†, Ishiguro Shin†, Yamada Hisashi†

† Nagasaki University

表1 1週間の平均乗車数の階級別の道路数

1週間の平均乗車数	道路数	備考
0以上100未満	27253	低需要
100以上500未満	138	中需要
500以上	38	高需要

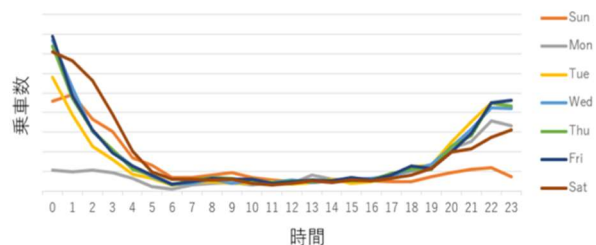


図1 高需要道路の時間帯別の平均乗車数の推移

今回は、名古屋市周辺における2023年4月～6月のタクシー運行実績データと、Open Street Mapの道路データを用い、道路毎の乗車数の傾向を分析する. なお、乗車数の集計単位は前述の道路データに収録されているレコード単位とする。

表1は、道路データのレコード単位で、その該当場所での1週間の平均乗車数を集計し、階級別に道路数を示したものである. 表に示す通り、ほとんどの道路の平均乗車数が0以上100未満と、極めて低需要な道路であることがわかる. これらの道路は、データ量の観点から学習が困難であることに加え、高需要な場所での運行を誘導するためのサービスでは対象外としてよい道路であると考えられる。

また、100以上500未満の低需要、中需要な道路は、名古屋市全体に分布していた. 500以上の高需要な道路は、名古屋市都心部の駅周辺、繁華街に分布していた。

図1は、ある高需要な道路の各時間帯における平均乗車数の推移である. なお、乗車数の具体的な値は営業情報にあたり公開できないため記載を省略する. この道路においては、夕方から夜にかけて乗車数が増加し、0～1時台がピークで、そこから翌朝まで減少している. また、曜日ごとに見ると、他の曜日と比べて日曜の夜から月曜の早朝は乗車数が少ない。

この道路は夜の飲食店が多い繁華街にあるため、店舗の利用者の行動あわせた傾向であると考えられる。

以上より、高需要道路に関しては、曜日、時間帯ごとに乗車需要の傾向が明確であることから、それぞれの道路単体で需要予測ができる可能性が高い。

3. 中需要道路の集約

中必要な道路の場合、データ量の問題から個別の道路単体では学習が難しい可能性がある。その場合、複数の道路をグループ化し、グループ毎に乗車実績を集約することで、道路グループ毎の需要予測を行うアプローチが考えられる。

道路の集約方法は、単に近くの道路同士でグループ分けするのではなく、乗車需要の傾向が類似した道路同士をグループ分けすることで、精度の良いモデルが学習できると期待される。

そこで、道路同士のグループ化が可能か確認すべく道路に対し、各曜日の時間帯別の平均乗車数を特徴量とした時系列クラスタリングを実施した。

図 2 は、時系列クラスタリングの結果、ある地域について、クラスタ毎に色分けして道路を示したものである。このように同一のクラスタに所属する道路同士が、地理的に同じエリアにまとまる形で分布していることがわかる。

また、図 3 に示す通り、それぞれのエリアは実際に繁華街やオフィス街という風に街区の特性に対応しており、実世界の事象と照らし併せてクラスタの特徴を定性的に説明できている。

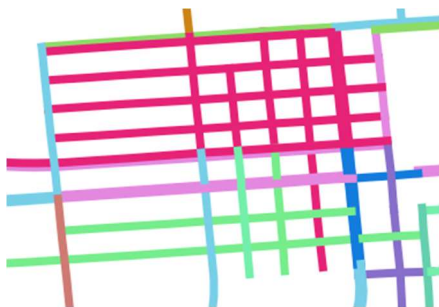


図 2 時系列クラスタリングによりグループ化された道路群



図 3 街区の特性

以上より、中必要な道路については、各道路の曜日、時間帯ごとの乗車需要の傾向から、道路をグループ化することが可能であったため、道路グループ毎に乗車実績を集約して、学習に十分なデータ量が確保されれば、グループ毎の需要予測を実現できる可能性がある。

4. おわりに

本稿では、タクシー乗車需要の学習・予測が可能な道路粒度を決定するため、実際のタクシー運行実績データを用いた道路レベルでの乗車実績の傾向分析を行った。

道路毎に 1 週間の平均乗車数を確認すると、ほとんどの道路が学習困難である低必要な道路であり、これらは本研究の想定ではスコープ外であることから検討対象として除外した。高必要な道路では、学習・予測に十分な乗車数と傾向が確認された。データ量の問題から個別の道路単体では学習が困難と思われる中必要な道路は、乗車実績に基づく時系列クラスタリングの分析の結果、傾向が類似した道路同士を実際の街区特性にあわせてグループ化できたことから、道路グループ毎に集約することで学習・予測に十分なデータ量を確保できるようになると考える。

今後は、更なる傾向分析を進めていき道路粒度を決定していくとともに、決定した道路粒度に基づいたモデルを作成し、需要予測 AI の構築を進めていく。

参考文献

[1]東京タクシー 2023:東京ハイヤー・タクシー協会. 入手先([https://www. taxi-tokyo. or. jp/assets/pdf/datalibrary/hakusyo2023all. pdf](https://www.taxi-tokyo.or.jp/assets/pdf/datalibrary/hakusyo2023all.pdf))(参照 2023-01-04)

[2]石黒 慎, 菊池 悠, 深澤 佑介:リアルタイム人口を用いた Stacked denoising Autoencoders によるタクシー将来需要予測. 情報処理学会研究報告, Vol. 2017-CDS-20, No. 21, pp.1-6