

RBF ネットワークを用いたコミュニティサイクル利用状況の短期予測

Short-term prediction of community cycle usage by the RBF network

室蘭工業大学工学部建築社会基盤系学科 学生員 ○岡田和洋 (Kazuhiro Okada)
 室蘭工業大学大学院工学研究科公共システム工学科専攻 学生員 丸山翔大 (Shota Maruyama)
 室蘭工業大学大学院工学研究科 正 員 有村幹治 (Mikiharu Arimura)

1. はじめに

コミュニティサイクルは、都心部において共用自転車を提供し、スムーズな移動環境と低炭素なまちづくりの一端を担っている。例えば、札幌都心部で展開中の「ポロクル」では、ポートと呼ばれる拠点で貸し出し返却を行っており、利用データ数も年間約 14 万回を数える成長を見せ、駐輪マナーの向上や都心部来訪の機会の増加といった社会的効果も確認されてきている。

一方、コミュニティサイクル事業の継続により明らかになった課題として、ピーク時の特定ポートに偏る需給ギャップ対応、コミュニティサイクルの利用を通じて顕在化した新たな公益性をユーザーと共有する仕掛けづくり、収支バランスの調整等、改善の余地が挙げられる。

本研究は、これらの課題のうち、ピーク時に特定ポートに偏る需給ギャップへの対応に関して、利用履歴データをもとに RBF ネットワークを用いて将来予測を行い、その有用性を検討し運営の効率化の足掛かりを得ることを目的とする。

2. 既存研究の整理

丸山ら¹⁾は、ポロクルのポート間 OD をモジュラリティ指標を用いて抽出した。また、利用パターンを可視化し、現状の把握を行った。松田ら²⁾は、時間帯により変化する需給ギャップに関する現状分析を行い、ポート増設による改善効果を検証し、運営の効率化に向けた方向性を示した。数値計算による将来予測の研究として斉藤ら³⁾は、確率的在庫管理問題をシェアリングネットワークに拡張し、料金によって内生的に変化する需要関数を設定することで異なる需要パターンに対して安定的に将来収益を増加させられる可能性を示した。

本研究の特徴は、自転車の再配置の効率化を見据え、RBF ネットワークを用いたポート毎の利用状況の短期予測可能性の検討を行う点にある。

3. RBF ネットワークの概要

RBF ネットワークは有限個の入出力データを補完する方法として提案されたニューラルネットワークである。入力層、中間層、出力層の 3 層から構成され、中間層にあたる複数の RBF (放射基底関数) に重みをもたせ足し合わせることで、任意の非線形関数の近似が可能である。

多層パーセプトロン等によく挙げられるローカルミニマムの問題がなく、最小二乗法によって任意の非線形関数の最良近似法を導くことができるという特徴がある。なお、入出力層に対応する数値 n は $0 < n < 1$ である必要がある。概念図を以下に示す(図-1)。

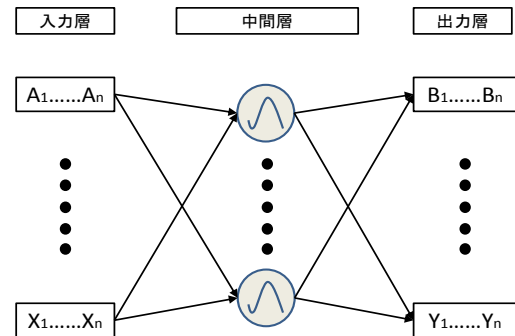
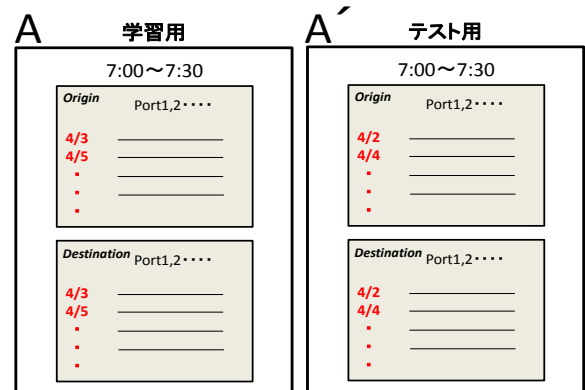


図-1 RBF ネットワークの概念

4. 利用データの整理

短期予測を行うにあたり、自転車の還流の傾向が時間帯により変化することを考慮する必要がある。松田らは、時間帯別の満車と空車の発生状況を可視化し、ラッシュアワーである平日の 7 時~9 時、18 時~21 時に需給ギャップがピークに達することを明らかにしている。このことから、本研究では札幌都心部で展開中のコミュニティサイクル「ポロクル」の 2012 年度利用履歴を用いた。この全利用データのうち、朝の利用ピークである平日の 7 時~9 時を抽出し分析に用いる。

データ整理の手順は以下の通りである。30 分ごとの利用データを $A \sim D$ 、 $A' \sim D'$ の 8 つのボックスを作成し整理する。例としてボックス AA' を図-2 に示す。

図-2 整理したデータ ボックス AA'

7 時から 7 時 30 分までに発生が起こったトリップデータを取り出し、発生量をポートごとにまとめる。これを 30 分区分切りで繰り返し、集重量も同じ手順で整理する。学習用データとテスト用データに大きな条件の差が生まれないよう、行(ここでは日付)を交互に抜き出して二分し、学習用データボックスを $A \sim D$ 、テスト用データボックスを $A' \sim D'$ とした。

最後に、RBF ネットワークのパラメータとして適切な数値 n ($0 < n < 1$) に変換するため正規化を行う。各利用データを x 、全利用データの平均値、標準偏差をそれぞれ y 、 z として以下の式を用いて変換する。また、 $0 < n < 1$ に落とし込むため、最大値を 0.001、最小値を 0.999 に置き換えた。

$$n = \frac{a - MINa}{MAXa - MINa} \quad \begin{matrix} MAXn = 0.001 \\ MINn = 0.999 \end{matrix} \quad \dots(1)$$

$$a = \frac{x - y}{z} \quad \dots(2)$$

4.1 RBF ネットワークを使った学習

学習用パラメータとして作成した4つのボックスを用いる。入力層に ABC を、出力層に BCD を図-3のように並べる。RBF ネットワークは行列の関係を保持したままパラメータの変化を学習するため、特定の時間から30分後の利用状況の予測が可能になる。

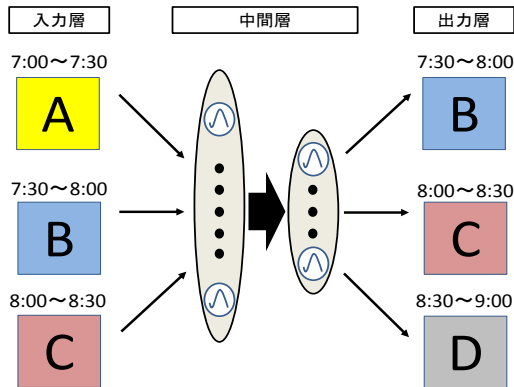


図-3 RBFネットワークによる学習のイメージ

この際、中間層のユニット数を任意の値で決定する。本研究では中間層を2層、ユニット数の比を10:6に固定し、各ユニット数を(5:3)(10:6)(20:12)(40:24)(60:36)(120:72)の6通りで学習させた。

4.2 精度の検討

テスト用パラメータのうち、7時から8時30分までを同様に入力層に並べ、学習に用いた中間層パターンによって6通りの予測結果を得た。結果の値は正規化されたパラメータが算出されるため、正規化と逆の手順を行い、実際の利用データと比較することで、得られた予測値の精度を検討する。

まず、使用した6通りの中間層パターンを、誤差の絶対値の平均(以下、平均誤差)と、AIC(赤池情報量基準)について比較する(図-4)。ユニット数の多い組み合わせほど平均誤差が小さくなる傾向が見られたが、(120:72)では誤差が大きくなった。

次に、平均誤差の最も小さい(60:36)について精度を検討する。ポートに焦点を当て、各ポートの平均誤差を縦軸、通勤ピーク時にそのポートが使用される回数の平均を横軸にとり、分布を調べた(図-5)。図-5より、利用頻度が高くなるにつれて誤差が大きくなることがわか

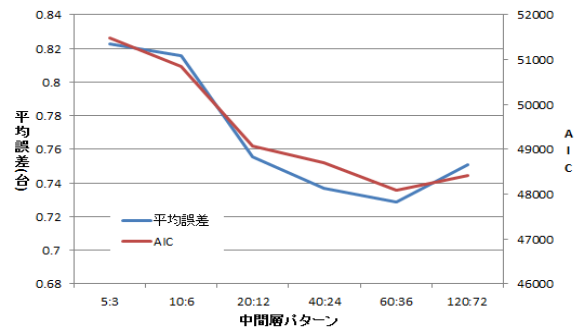


図-4 中間層パターンと平均誤差、AICの比較

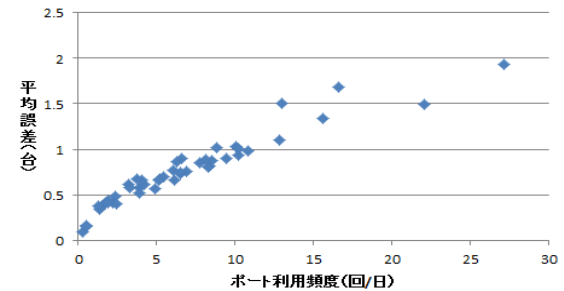


図-5 ポートの利用頻度と平均誤差の分布

る。しかしながら、利用頻度が低い場合は概ね正確な予測が可能であり、自転車の流入出の大まかな流れは予測できていると考えられる。

5. まとめ

本研究では、2012年度ポロクル利用履歴をもとに、RBF ネットワークのにより利用状況の短期予測を行い、その精度を検討した。

利用頻度の上昇に応じて平均誤差が大きくなる問題については、学習に使用した正規化前のデータ（ここでは利用回数）の大半を0や1のような、最小値とそれに近い数値が占めていることで学習の際の重みづけが極端になったことが原因と考えられる。この問題の解決方法として、RBF ネットワークに利用状況を学習させる際に日平均利用頻度の高低による差別化、学習データの小規模化をはかること、発生量と集中量の差をとりパラメータにばらつきをもたせることで、極端な重みづけを回避するといった方法が考えられる。

以上を踏まえ、今後の課題として、利用頻度の高いポートに焦点を置いた局所的な予測、各ポートのラック数を考慮した、より詳細で実用性の高い予測方法の考案等があげられる。

参考文献

- 丸山翔大、松田真宜、長谷川裕修、有村幹治：札幌都心部におけるサイクルシェアリングの自転車還流における基礎的研究、土木計画学研究、講演集、Vol.48,CDROM,2013
- 松田真宜、平川貴志、有村幹治：利用履歴データに基づくサイクルシェアリングの戦略的な運営手法の構築土木計画学研究、講演集、Vol.47,CDROM,2013
- 斎藤いつみ、羽藤英二：シェアリングシステムの短期オペレーションにおける確率的在庫管理手法の導入、土木計画学研究、講演集、Vol.48,CDROM,2013