

強化学習を用いた畳み込みニューラルネットワークによる 巡回セールスマン問題の解法

三木 彰馬[†]榎原 博之[‡]関西大学大学院理工学研究科[†]関西大学システム理工学部[‡]

1 はじめに

組合せ最適化問題は計算機科学における基本的な問題領域の1つであり、輸送や通信、製造などさまざまな分野における応用が期待されている。組合せ最適化問題では高精度かつ高速な解法の開発が重要であるが、人の手による高性能なヒューリスティクスの開発には経験的な専門知識や多大な労力がしばしば必要である。

一方、近年さまざまな分野で深層学習の技術を用いた研究が活発に行われている。さまざまな問題に対してデータの表現方法を自動的に獲得できる深層学習の能力を組合せ最適化問題に適用することで、より性能の高い解法が得られるのではないかと期待されている。先行研究として再帰型ニューラルネットワークにより入力要素を評価・選択する手法 [1] やグラフ構造を考慮してグラフ埋め込みを計算する手法 [2] などが挙げられる。

著者らはこれまでに巡回セールスマン問題 (TSP: Traveling Salesman Problem) に対して深層学習による画像処理技術を利用した解の構築手法として優良エッジ分布と呼ばれる手法を提案している [3]。この手法では TSP が持つ局所空間的な情報に着目し、畳み込みニューラルネットワーク (CNN) を用いて最適経路の画像を予測し、より最適解に近い解の構築に利用する。

この手法では最適解が判明している問題例を教師データとして用いることで優良エッジ分布を学習していたが、TSP では頂点数が増えるほど最適解を短時間で求めることが難しく、十分な数の教師データを用意することができない。そこで本研究では強化学習の仕組みを利用することで教師データを必要としない優良エッジ分布の学習方法を提案し、より大規模な問題例に対して高い性能を持つ解法の開発を目指す。

Reinforcement Learning for Convolutional Neural Network to Solve the Traveling Salesman Problem

[†] Shoma Miki, Graduate School of Science and Engineering, Kansai University

[‡] Hiroyuki Ebara, Faculty of Engineering Science, Kansai University

2 優良エッジ分布

2.1 TSP の画像表現

平面 TSP を表すグラフ $G(V, E)$ が与えられたとき、その頂点と辺および巡回路 z を画像として表現する。解像度 (W, H) の画像空間 S を考えたとき、 S の各画素 $a \in S$ について、任意の頂点集合 V' を描画したときの画素値を $d_a(V')$ 、任意の辺集合 (または巡回路) E' を描画したときの画素値を $t_a(E')$ とする。このとき頂点集合 V' と辺集合 E' が描画される画素の集合をそれぞれ $D(V')$ 、 $T(E')$ とし、そこに含まれる画素に対して 1、それ以外の背景に対して 0 の画素値で描画する。

2.2 優良エッジ分布と優良エッジ値

頂点画像 $d_a(V)$ を入力し、その最適経路 z^* の画像 $t_a(z^*)$ を近似するような全畳み込みニューラルネットワーク (CNN) を考えたとき、その出力 $p_a(V)$ を優良エッジ分布と呼ぶ。各辺 (u, v) が占める画素において優良エッジ分布の値を集約することで、その辺が最適解に含まれる尤度 $v_{u,v}$ を求める。その方法としては出力画素値の濃度を求める手法 (式 (1)) と、出力値を確率とみなして確率積を求める手法 (式 (2)) が考えられる。

$$\bar{v}_{u,v} = \frac{\sum_{a \in S} p_a(V) t_a(\{(u, v)\})}{\sum_{a \in S} t_a(\{(u, v)\})} \quad (1)$$

$$v_{u,v} = \prod_{a \in t_a(\{(u, v)\})} p_a(V) \quad (2)$$

優良エッジ値が高い辺を優先的に選択することで最適経路に近い解が得られることが期待できる。優良エッジ値を用いた解法として、優良エッジ値が高い辺を優先的に選択して巡回路を構築する EV-greedy, 2-opt 近傍探索法において辺のコストの代わりに優良エッジ値の総和を最大化する EV-2opt 法をこれまでに提案している。

3 提案手法

提案手法では強化学習の仕組みを利用し、学習中に観測した解の精度が高ければその巡回路上に存在する画素

値を増加，精度が悪ければその画素値を減少させるように CNN を更新する。

3.1 解の価値関数

訓練用問題例 s に対して，観測によって得られた解 z の精度を式 (3) を用いて評価する． $L(z)$ は巡回路の経路長 (総コスト) を表し，今までに得られた解の経路長の指数移動平均 $b(s)$ と最良経路長 $L^*(s)$ をそれぞれ記録する． ρ は傾斜係数 ($\rho = 2$ を使用)， ϵ は微小項を表す。

$$V(z) = 1 - \frac{L(z) - L^*(s)}{\rho(b(s) - L^*(s)) + \epsilon} \quad (3)$$

この価値関数 $V(z)$ をニューラルネットワークで近似することで，各問題例に対する観測データの精度のばらつきの影響を抑えられることが期待できる。

3.2 画素単位の辺の評価

観測時の解構築アルゴリズムによっては 1 つの選択が解全体の精度を大きく悪化させる場合が多く，その影響を受けて本来選ぶべき辺への評価値を低く見積もってしまう可能性がある．そこで，既知の最良経路長を画素ごとに記録した最良コストマップ $M_a(s)$ を利用して式 (4)，式 (5) のように画素単位の評価を反映させる。

$$U_a(z) = 1 - \frac{M_a(s) - L^*(s)}{\rho(b(s) - L^*(s)) + \epsilon} \quad (4)$$

$$V'_a(z) = (1 - \alpha)V(z) + \alpha U_a(z) \quad (5)$$

3.3 強化学習による CNN の更新

リプレイメモリ内に蓄えられた観測データ z を用いて，CNN の出力 $p_a(s)$ に対して式 (6) のような損失関数を与え勾配降下法を適用する．このとき出力を評価値に近づけるような勾配を与えるための誤差関数 $\Delta(t, y)$ には式 (7) のようなクロスエントロピーを用いる。

$$J(z) = \frac{1}{|T(z)|} \sum_{a \in T(z)} \Delta(V'_a(z), p_a(s)) \quad (6)$$

$$\Delta(t, y) = -t \ln y - (1 - t) \ln(1 - y) \quad (7)$$

4 検証実験

提案した学習方法による効果を調べるため計算機実験を行う．訓練用問題例として頂点数 11–200 で 9,500 個，検証用問題例として頂点数 11–500 で 980 個のランダム座標問題例を生成する．観測した解は優先度付き経路再生に格納され，式 (6) の損失関数に従い CNN $p_a(s)$ を更新するために使用される．優良エッジ値の計算には式 (2) を使用し，解の観測時には ϵ -greedy 法を併用した EV-greedy と 2-opt 探索法を順番に適用する。

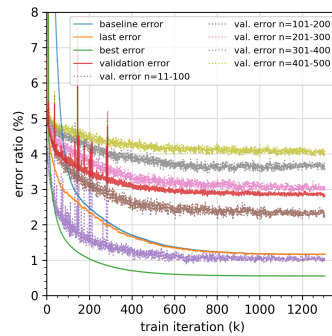


図1 訓練用・検証用問題例に対する誤差率の推移

表1 検証用問題例に対する誤差率

頂点数	誤差率
11–50	0.66%
51–100	1.34%
101–150	2.10%
151–200	2.30%
201–300	3.08%
301–400	3.64%
401–500	4.15%

観測した解および検証用問題例に対する解における最適解に対する誤差率の推移を図1，学習終了時の検証誤差率を表1に示す．学習が進むにつれて観測時の平均誤差率 (baseline error) と検証精度 (破線) が減少していることから，解の精度を向上させる方向へ学習が進んでいることがわかる。

5 まとめ

本研究では，畳み込みニューラルネットワークを用いた TSP の解法において強化学習を利用した新たな学習方法を提案し，実験により教師データを用いずにある程度良好な解の精度を達成できることを確認した．今後は既存の解法との精度の比較や，より大規模な問題例に対する性能について検証が必要である。

謝辞

本研究の一部は，JSPS 科研費 18K11484 と関西大学大学院理工学研究科高度化推進研究費，関西大学先端科学技術推進機構「緊急救命避難支援のための災害情報通信ネットワークに関する研究開発」研究グループの助成を受けている。

参考文献

- [1] Vinyals, O., Fortunato, M. et al.: Pointer Networks, *arXiv preprint arXiv:1506.03134* (2017).
- [2] Khalil, E., Dai, H. et al.: Learning Combinatorial Optimization Algorithms over Graphs, *Advances in Neural Information Processing Systems*, Vol. 30, Curran Associates, Inc. (2017).
- [3] 三木彰馬，榎原博之: 深層学習を用いた巡回セールスマン問題の解法，*情報処理学会論文誌*, Vol. 60, No. 2, pp. 651–659 (2019).