

辺ベクトルによる定式化を用いた柔軟な経路推薦手法の提案

Proposal of Flexible Route Recommendation Method Employing Edge Vector-Based Formulation

柴田祐樹 *1

Hiroki Shibata

高間康史 *1

Yasufumi Takama

*1 首都大学東京

Tokyo Metropolitan University

Many studies for sightseeing route recommendation use the formulation of Selective Traveling Salesman Problem or its extension. However, this formulation uses the costs of the shortest paths between spots as the weights of edges in a graph, which means that the variety of routes between spots is ignored. Users might prefer different routes than the shortest path, such as the route along the river, forest, or sea if they have enough time. This paper proposes a method to recommend routes by considering the balance between staying time and traveling time. Extending the Edge Vector-based formulation, the proposed method utilizes inner routes in a spot to represent the staying time as the costs of those inner routes. An experiment using an artificial data set shows that the proposed method can optimize the staying time and traveling time concurrently.

1. はじめに

観光旅行をする際、ガイドブックや旅行会社が提供する観光ツアーを利用するのではなく、SNSなどのWebから入手した情報を用いて旅行者自身で観光計画を立案することも増えてきている。観光情報サイトなどでは、エリアなどの指定条件を満たす観光スポットをランキング形式で提示するサービスを提供しているが、知名度の高いスポットや宣伝に力を入れているスポットが上位に表示される傾向がある。ユーザの検索行動として、ランキング上位の件数のみしか確認しないことが一般的であるため、結果として個人の趣味嗜好に合った観光地を訪れる機会の損失を招き、一部の観光地へ旅行者が集中するなど、旅行者、観光地側双方に不利益を生じる。こういった問題を防ぐためにも、ユーザの好みに応じたスポットを選び、限られた時間で効率よく巡回する経路を提案する観光経路推薦手法は重要であり、研究が進められている [松田 04][Lim 17]。

与えられた節点全てを巡回する最短経路を求める問題は Traveling Salesman Problem (TSP) [Gutin 06] として古くから研究されている。NP-Hard [Yannakakis 91] に属する TSP を解くためには近似解法が用いられ、代表的なものには 2-Optimization (2-Opt) 法 [Helsgaun 09]、Simulated Annealing (SA) [Kirkpatrick 83] がある。さらに、スポットに価値を付加し、与えられた総負荷制限の中で価値の総和が最大となる経路を選ぶ問題として、Selective Traveling Salesman Problem (STSP) [Laporte 90] が定式化されている。また、STSP を拡張し、訪れた時間に依存して価値が変化する問題を文献 [松田 04] では最適観光経路問題として定式化しており、文献 [Lim 17] では同様の定式化において、スポットの価値を Web から取得したデータを用いて推定する方法を提案している。

これらの既存研究はいずれも、節点にスポットを割り当て、節点間の辺の重みに、スポット間を結ぶ最短経路負荷 (コスト) を割り当てる、STSP の定式化を利用しているが、この定式化では特定の交通機関や道路しか用いられないことに等しい。これに対し文献 [柴田 19a] ではスポットのみでなく、途中の経路における街並みや景観等もユーザにとって重要な因子であると

し、これらを考慮可能な観光経路推薦手法 (以降 Edge Vector: EV と呼ぶ) を提案している。スポットおよび経路に関する価値や負荷すべてを辺に対する重みとして定式化することで、スポットと経路の情報を統一的に扱い、探索の過程において両者を同時に調整しながら最適経路の発見を可能としている。

本稿では EV を更に拡張し、スポット内に配置した冗長な経路を割り当てられた価値に応じて選択することによりスポットへの滞在時間をモデルする手法を提案する。提案手法により、滞在時間、途中経路を同時に決定可能となるため、柔軟な経路推薦手法が実現できる。人工データセットを用いた評価実験により、提案手法が有効に動作することを示す。

2. 辺ベクトルを用いた観光経路推薦問題の定式化と SA による近似解法

本節では EV の定式化と、同文献で用いられる近似解法について述べる。また、各用語の命名が曖昧であると考えられるものについては、対応関係を示しながら、本稿で新たに定義したものを用いて説明する。

$\mathcal{N}$ を節点の集合、 $\mathcal{E}_p$ を辺の全体集合とする。 $\mathcal{E}_p$ は問題に与えられた辺 (定義辺) と後述する自己回帰辺、仮想辺 *1 を含む。辺 $e \in \mathcal{E}_p$ に対し、 $n_{st}(e)$ を e の始点、 $n_{ed}(e)$ を e の終点としたとき、これらは $\bigcup_{e \in \mathcal{E}_p} (n_{st}(e), n_{ed}(e)) = \mathcal{N}^2$ の関係を満たすものとする。

$|\mathcal{N}|$ 個の辺からなる経路を含む辺ベクトルを $\mathbf{e} := (e_i | i = 1, 2, \dots, |\mathcal{N}|)$ と定義する。 $n_{st}(e_1)$ は経路の始点、 $n_{ed}(e_{|\mathcal{N}|})$ は経路の終点であり、最適化中変化しない。辺 $e \in \mathcal{E}_p$ のうち、 $n_{st}(e) = n_{ed}(e)$ であるものを自己回帰辺と呼ぶ。辺ベクトルは自己回帰辺を含むことで、要素数を固定としながら様々な経路を表現することができる。図 1 では、定義辺を黒線で表し、輪を描いているものが自己回帰辺、青で示すものは辺ベクトルに含まれる定義辺である。この図は、節点 1, 2, 7, 8 をたどる経路の例であり、対応する辺ベクトルは $((n_{st}(e_i), n_{ed}(e_i)) | i = 1, 2, \dots, 8) = ((1, 2), (2, 7), (3, 3), (4, 4), (5, 5), (6, 6), (7, 8))$ となる。

また、この手法では、仮想辺を用いることで、2 辺の操作の

連絡先: 高間康史, 首都大学東京, 東京都日野市旭が丘 6-6,
email: ytakama@tmu.ac.jp

*1 文献 [柴田 19a] では実在辺, 自己ループ辺, 非存在辺と呼ばれる



図 1: グラフの例.

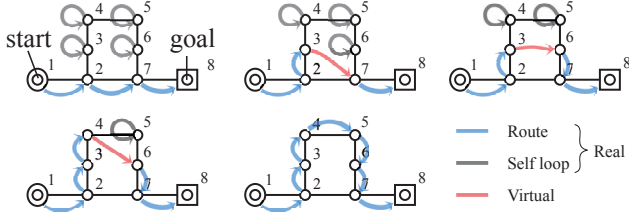


図 2: 仮想辺を用いた経路拡張の例.

みにより経路の変更をしながら探索を行う。図 2 に示す経路変更の例では、赤線により仮想辺を示している。定義辺に乗らない仮想辺を経由することで、本来は 4 つの自己回帰辺を定義辺に変換する必要がある変化を、2 辺の置換を 4 回繰り返すことで表現している。本稿では、定義辺と自己回帰辺を合わせて実辺と呼び、その全体集合を $\mathcal{E}$ ($\subset \mathcal{E}_p$) とする。仮想辺の集合は $\mathcal{E}$ の補集合となる。

経路は Boltzmann 分布 [Metropolis 53] から確率的に生成されるとモデル化し、最適解を SA により近似的に求める。 $\mathbf{e}$ の確率変数を $\mathbf{e}_A$ 、その分布 $p(\mathbf{e}_A)$ を式 (1) により定義する。 $A = \{1, 2, \dots, |N|\}$ は辺ベクトルの各要素に対する添字の集合である。本稿において、同じ添字集合を持つ確率変数は独立同分布に従うものとする。

$$p(\mathbf{e}_A) = \frac{1}{\sum_{\mathbf{y}_A} \exp(-\frac{1}{kT} \phi(\mathbf{y}_A))} \exp\left(-\frac{1}{kT} \phi(\mathbf{e}_A)\right) \quad (1)$$

T は温度である。また、 k は Boltzmann 定数であり、問題ごとに適切な値を設定する。 $\phi(\mathbf{e}_A)$ は目的関数であり、この値が小さくなる程、 $\mathbf{e}_A$ の出現率が高くなり、 T が小さくなるほどこの傾向が強くなる。SA ではこの性質を用いて最適解の近似解を生成するため、目的関数の値が小さな状態が望ましい経路を表すように定式化を行う。 ϕ を式 (2) により定義する。文献 [柴田 19a] の定式化と多少異なるが、本質的には同じであり、3. 節における提案手法の説明に同一の式を利用可能としている。 $\alpha_s = 0$ としたものが、EV の定式化に対応する *2。

$$\phi(\mathbf{e}_A) = F_{\text{res}} \left(\sum_{i \in A} f_c(\mathbf{e}_i) + \alpha_s \sum_{i \in A} f_s(\mathbf{e}_i) - C_c \right) - \sum_{i \in A} f_d(\mathbf{e}_i) + \alpha_a \sum_{i \in A} f_s(\mathbf{e}_i) \quad (2)$$

$f_c(e)$ は e の負荷、 $f_d(e)$ は e の価値、 $f_s(e)$ は仮想辺に対する負荷であり、式 (3) により定義される。

$$f_s(e) = f_c(e) - \sum_{s \in S(e) \cap \mathcal{E}_c} f_c(s) \quad (3)$$

各値は辺の向きによらず、 $e \notin \mathcal{E} \Rightarrow f_c(e) = 0, f_d(e) = 0$ および $\forall e, f_d(e) < f_c(e)$ を満たすものとする。後者の条件は次

*2 文献 [柴田 19a] ではさらに f_s に対し微小な定数項を加算している等の違いがあるが、それらの影響がほぼ無いことを事前実験に確認しているため、詳細については省略する

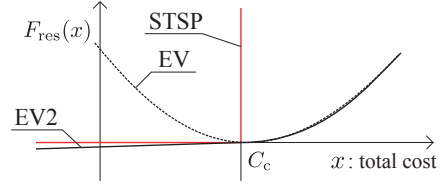


図 3: 制約関数の概形. STSP: STSP, EV: 文献 [柴田 19a], EV2: 提案手法.

の段落で説明する STSP の制約を無制限に超えた経路が生成されることを防ぐものである。 $S(e)$ ($\subset \mathcal{E}$) は辺 e の終始点を結ぶ最短経路を構成する辺の集合である。また、 $S(e) = \emptyset \Rightarrow f_s(e) = \infty$ とする。EV の定式化において、 α_a は仮想辺を消滅させるための罰則係数であり、1 以上で、仮想辺が最終的に問題なく消滅する程度に 1 に近い値を用いる。

$F_{\text{res}}: \mathbb{R} \rightarrow \mathbb{R}$ は経路負荷に対する STSP の制約を表現する関数であり、 C_c は制約値である。観光経路推薦においてこの制約値は、経路長や旅行時間に対応する。STSP, EV, 提案手法 (EV2) で用いる制約関数の概形を図 3 に示す。STSP による定式化は C_c を超える経路を認めないため、目的関数は C_c を堺に無限大の値を取る。このような目的関数を用いた最適化では、経路長が C_c 付近となる状態の近傍探索における効率が低下し、また、興味のあるスポットがあれば、多少の制約違反が許容される観光旅行において妥当なモデルとは言えない。これに対し、EV ではある程度の誤差を認め、最適化効率を向上させたいので、より現実的なモデルとしている。EV2 の持つ性質については 3. 節において説明する。EV, EV2 に共通して利用する定式化を式 (4) に示す。

$$F_{\text{res}}(x) = \begin{cases} \alpha_g x + c_1 & , x < x_m \\ c_2 x^2 + c_3 x & , x_m \leq x \leq \sigma^2 \\ x + c_4 & , \sigma^2 < x \end{cases} \quad (4)$$

各設定値を、 $F'_{\text{res}}(0) = \max(\alpha_g, 0)$, $c_1 = c_2 x_m^2 + c_3 x_m - \alpha_g x_m$, $c_2 = (1 - c_3)/(2\sigma^2)$, $c_3 = \max(\alpha_g, 0)$, $x_m = (\alpha_g - c_3)/(2c_2)$, $c_4 = c_2(\sigma^2)^2 + c_3\sigma^2 - \sigma^2$ と定めたとき、 $\alpha_g = -1$ において、この関数は EV のものに対応する。 σ^2 により違反の許容度合いを調節することが可能である。

SA 法において、Boltzmann 分布を近似するために Gibbs Sampling (GS) 法 [Geman 84] を用いる。経路変更を行うための最小要素は二つの辺であるため、GS 法において利用される部分空間として辺ベクトル中の 2 辺を選ぶ。この 2 辺に対する添字を i, j ($\in A$) としたとき、これら以外の添字 $A \setminus \{i, j\}$ で定義される変数により条件付けられた分布は式 (5) と書ける。

$$p(\mathbf{e}_{\{i,j\}} | \mathbf{e}_{A \setminus \{i,j\}}) = \frac{\exp(-\frac{1}{kT} \phi(\mathbf{e}_A))}{\sum_{\mathbf{y}_{\{i,j\}}} \exp(-\frac{1}{kT} \phi(\mathbf{y}_{\{i,j\}}, \mathbf{e}_{A \setminus \{i,j\}}))} \quad (5)$$

添字 A の部分集合 M を添え字に持つ確率変数は対応する部分空間を構成し、式 (6) の性質が満たされるものとする。

$$p(\mathbf{e}_{A \setminus M}) = \sum_{\mathbf{y}_M} p(\mathbf{y}_M, \mathbf{e}_{A \setminus M}), M \subset A \quad (6)$$

ここで、 $\sum_{\mathbf{y}_M}$ は $\mathbf{y}_M$ の状態空間についての総和を示す [Bishop 06]。この状態空間には 2-Opt 法に対応する、経路としての制約を満たすものを与える。

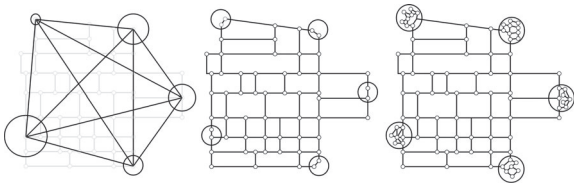


図 4: 道路ネットワークに対するモデル化. (a): STSP ベースの手法, (b): EV, (c): 提案手法のモデル化.

3. 提案手法

本稿では, EV の定式化を拡張し, スポットにおけるユーザの滞在時間と, 途中経路を同時に最適化可能な推薦手法を提案する. EV では, スポットに一つの辺を割り当て, その移動時間として滞在時間を表現しており, スポットにおける滞在時間は固定的である. しかしながらこの定式化において, 「興味のあるスポットは時間の許す限り長く滞在したい」 などのような, 実際によく行われる柔軟な調整を考慮することができない. そこで本稿ではスポット内に冗長な経路を含むグラフを配置し, 辺に与える価値に応じて生成される経路を決定する手法を提案する.

図 4(a) に, STSP ベースの従来手法, (b) に EV, (c) に提案手法による定義辺配置の様子を示す. (a) のモデル化は, スポットに対応する節点への重みの一つとして固定的な滞在時間を表現する. (b) においては, スポットに対し辺を配置し, その負荷 (f_c) として滞在時間を表現している. 辺か節点かの違いはあるが, スポットに対する滞在時間が固定値となる点で (a) と同様である.

これらに対し, 経路全体を考慮して各スポットの滞在時間を柔軟に調節可能とするため, 本稿では (c) で示されるような, 冗長な経路をスポット内に配置するモデル化を提案する. スポットの滞在時間を, 「スポット内に存在する辺のうち, 経路として選択されたものが持つ負荷の総和」とみなすことで, 長い滞在時間はスポット内の冗長な経路をめぐり行動としてモデル化する. このとき, 人気のスポットや関心のあるスポット内にある辺の価値 (f_d) を高くすることで, スポット内の辺を選択されやすくなることが可能である. 滞在時間を最適化中に決定することができれば, 行きたいスポットをすべて巡っても若干の余裕がある場合に, 好きなスポットの滞在時間を長くしたり, そのほかのスポットを短時間寄り道したり, あるいは土産物店やレストランで時間を潰すといったような柔軟な経路を生成可能となる.

ここで, 提案手法で用いる目的関数 ϕ について説明する. 定式化には 2. 節でも説明したとおり式 (2) を用い, 式 (4) で定義する制約関数 F_{res} の形状には図 3 中の EV2 に該当する, $\alpha_g > 0 \wedge \alpha_g \approx 0$ と設定したものをを用いる [柴田 19b]. EV は C_c に極小値を持つ形状となっており, 訪れるべきスポットがない場合でも経路長が C_c に近づくように最適化が行われるため, 無駄な経路が生成されることから望ましい定式化ではない. これに対し EV2 は経路長に対し常に傾きが正となるため, 無駄な経路を削減するものとなっている.

α_g には, C_c 以下の経路長において, 価値がそれほど大きくないスポットを無視してしまうことがない程度に小さな値を用いる. この値は, $\min_e f_d(e)/f_d(e)$ に対し, 定義辺の与え方にも依存するため, 事前実験により決定することができる.

式 (2) 中のパラメータについては, 文献 [柴田 19b] に従い, $\alpha_s = 1, \alpha_a \approx 0 \wedge \alpha_a > 0$ と設定する. これにより EV に比べ

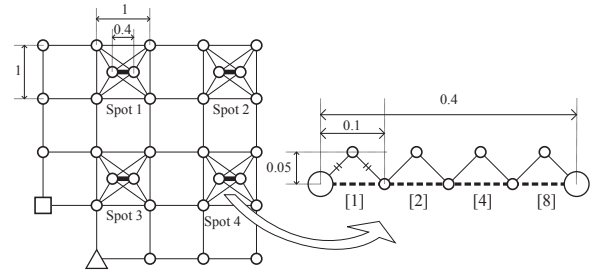


図 5: 4 つのスポットを持つ人工データセット. 太線で描かれた辺は右側のグラフに対応する. 黒破線下の [] 内の数字は各辺の f_c の値を示す. 四角は始点, 三角は終点をそれぞれ表す.

探索効率が向上し, 冗長経路の導入による問題の大規模化に対応することができる. α_a には事前実験を行い, 定義辺上を十分に探索可能である程度の微小な値を設定する.

4. 評価実験

図 5 に示す人工データセットを用い, 経路生成を行った結果を示す. このデータセットはスポット内に経路が配置され, スポット間の経路探索も含む, 提案手法が対象とするモデルに該当する. 太線で描かれた辺はスポットを示し, 図の右側にあるグラフに対応する. [] 内の数字は直上にある辺の f_c の値である. このグラフには 4 つのスポット, Spot 1, Spot 2, Spot 3, Spot 4 があり, 破線で描かれる辺 $e \in \mathcal{E}$ における f_d の値は以下のように設定する.

$$\text{Spot1: } f_d(e) = 0.8f_c(e) \quad \text{Spot2: } f_d(e) = 0.6f_c(e)$$

$$\text{Spot3: } f_d(e) = 0.4f_c(e) \quad \text{Spot4: } f_d(e) = 0.2f_c(e)$$

また, これら以外のすべての辺 e において, $f_d(e) = 0$ とし, $f_c(e)$ は図中に示す幾何学情報に基づいた距離に等しいものとする. 四角は始点, 三角は終点をそれぞれ示す. この人工データセットに対して, 制約値 (予定旅行時間) $C_c = 0.00, 0.25, 0.50, \dots, 110$ についてそれぞれ 32 回の試行を行い, 各スポットの滞在時間について平均値を求め比較する. パラメータの設定には $\alpha_s = 1, \alpha_g = 0.01, \sigma^2 = 0.1, k = 2, \alpha_a = 0.005$ を用い, 温度変化は $T(t) = 10^{-3t}$ とした. t が 10^{-5} 進むごとにサンプリングを行い, 文献 [柴田 19a] で提案されている近似的棄却サンプリングの有効範囲には $K = 16$ を用いた.

実験結果を図 6 に示す. Move はスポット間の移動時間 [] を示す. Spot 3 が終始点に近い場所にあるため, 制約値 (予定旅行時間) C_c が小さい場合にこのスポットが優先されているが, 大きくなるにつれ, より f_d の値 (価値) の高いスポットが優先的に選択されている. また, 大きくなった C_c のあまりが生じた場合にも Spot 3 を経由する, 寄り道のような経路を生成していると言える. スポット間の移動時間は $C_c = 10, 30$ 付近など, スポットが経路に追加されるときにのみ増加しており, このことから提案手法は不必要に経路を拡張しないこともわかる. これらの結果から, 提案手法では, スポット間の経路探索を行いつつ, 冗長な経路による滞在時間表現が可能であることがわかる.

5. おわりに

本稿では, 辺ベクトルによる定式化を用いた観光経路推薦手法を拡張し, スポット間の途中経路とスポットにおける滞在時

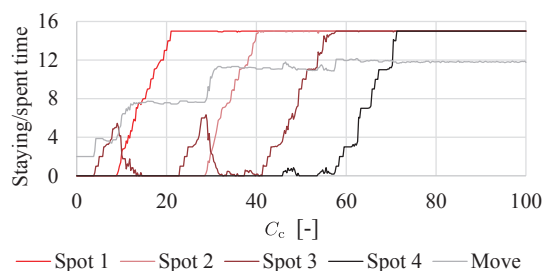


図 6: C_c の各値に対し, 人工データセットにおいて生成された経路に含まれる各スポットの滞在時間とスポット間の移動時間の総和。

間を同時に調整可能であるような, 柔軟な経路推薦手法を提案した。今後の展望として, 各スポットに対するユーザの飽きを考慮することで, 1 日程度に渡る長期間の旅行計画にも適用可能な手法に拡張することが考えられる。また, 道路ネットワークを用いた評価実験を行い, ユーザの求める旅行計画を実際に生成可能であることを検証する予定である。

謝辞

本研究は JSPS 科研費 16K12535, 15H02780 の助成および首都大学東京傾斜的研究費 (全学分) 学長裁量枠戦略的研究プロジェクト戦略的研究支援枠「ソーシャルビッグデータの分析・応用のための学術基盤の研究」によるものです。

参考文献

- [Gutin 06] Gutin, G. and Punnen, A.: The traveling salesman problem and its variations, *Springer Science & Business Media*, 2006.
- [Yannakakis 91] Yannakakis, M.: Expressing combinatorial optimization problems by linear programs, *Journal of Computer and System Sciences*, Vol. 43, No. 3, 1991.
- [柴田 19a] 柴田祐樹, 高間康史: 辺ベクトルを用いた観光経路推薦問題の定式化と焼きなまし法による解法の提案, 知能と情報, Vol. 31, No. 1, 2019.
- [Laporte 90] Laporte, G., Martello, S.: The selective travelling salesman problem, *Discrete applied mathematics*, Vol. 26, No. 2-3, pp.193-207, 1990.
- [松田 04] 松田善臣, 名嘉村盛和, 姜 東植, 宮城 隼夫: 最適観光経路問題とその解法, 電気学会論文誌 C, Vol. 124, No. 7, pp. 1507-1514, 2004.
- [Lim 17] Lim, K., Chan, J. and Karunasekera, S.: Personalized itinerary recommendation with queuing time awareness, *Proc. 40th SIGIR*, pp. 325-334, 2017.
- [Helsgaun 09] Helsgaun, K.: General k-opt submoves for the LinKernighan TSP heuristic, *Mathematical Programming Computation*, Vol. 1, No. 2-3, pp. 119-163, 2009.
- [Kirkpatrick 83] Kirkpatrick, S., Gelatt, C., Jr. and Vecchi, M.: Optimization by simulated annealing, *Science*, Vol. 220, No. 598, pp. 671-680, 1983.

[Metropolis 53] Metropolis, N., Rosenbluth, A., Rosenbluth, M. and Teller, A.: Equation of state calculations by fast computing machines, *The Journal of Chemical Physics*, Vol. 21, No. 6, pp. 1087-1092, 1953.

[Geman 84] Geman, S. and Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No. 6, pp. 721-741, 1984.

[Bishop 06] C. M. Bishop: Pattern recognition and machine learning, Chapter 11, *Springer*, 8th printing, 2006.

[柴田 19b] 柴田祐樹, 高間康史: 辺ベクトルを用いた経路推薦問題の解法における目的関数の改善に関する提案, 第 21 回インタラクティブ情報アクセスと可視化マイニング研究会, 2019.