

岡山市における観光経路の推薦ウェブサービスの提案

鹿毛 宗一郎[†] 小田 哲也[†]

[†] 岡山理科大学工学部 〒700-0005 岡山県岡山市北区理大町 1-1

E-mail: [†] t15j030ks@ous.jp, oda.tetsuya.fit@gmail.com

あらまし 近年、岡山市内の観光客数は年々増加傾向にある。しかしながら、岡山市は、岡山県内の他の市町村と比較して観光地が少ない。そのため、観光客がどこを観光すればいいのかわからない可能性がある。そこで、著者らは岡山市における観光経路の推薦ウェブサービスを開発している。また、岡山市の観光地や、観光に関する店舗をまとめたデータベースが少ないことから、岡山市の観光と、観光に関する店舗の情報をまとめた **Linked Open Data (LOD)** を作成した。そして、この **LOD** を、観光経路の推薦のために利用し、加えて、岡山市における観光経路案内の一手法として、観光経路の推薦ウェブサービスを提案する。本論文では、岡山市における観光地 **LOD** の作成と、**Twitter** における岡山市の観光関連ツイートの共起語抽出、観光経路の最短巡回回路部分を実装する。また、岡山市の観光関連ツイートにおける共起語を抽出するとともに、岡山市における観光地 **LOD** と **OpenStreetMap (OSM)** を利用し、その地理情報に基づいて、焼きなまし法で最短巡回回路を導出する。

キーワード 観光経路, **Linked Open Data**, 焼きなまし法, **Twitter**, 共起語

1. はじめに

近年、日本には海外からの旅行、いわゆるインバウンドをはじめ、旅行者は年々増加傾向にある。岡山県においても、海外からの旅行者は増加傾向であり[1]、国内の旅行者も増加傾向である[2]。しかし、旅行者が増加傾向にある一方で、岡山市は岡山県内の他の市町村と比較して観光地が少ない[3]。そのため、岡山市に訪問する観光客がどこを観光すればいいのかわからないという状況が予見される。そこで著者らは、岡山市の観光支援のために、岡山市における観光経路の推薦ウェブサービスを提案している[4]。本ウェブサービスの開発のために、岡山市の観光地と観光に関係のある店舗の情報を **Linked Open Data (LOD)**[5,6]として作成している。この **LOD** の情報を、**Twitter**[7]における岡山市の観光に関するツイートの共起頻度[8]と組み合わせ、観光経路を推薦する。また、ユーザが、ある時間内で観光できるように、**OpenStreetMap (OSM)** から取得した地理情報に基づき、焼きなまし法で最短巡回回路となる観光経路を導出する。これらを統合し、ウェブシステムとして拡張することで、岡山市における観光経路の推薦ウェブサービスを開発する。本稿では、提案システムの構成、岡山市における観光地 **LOD**、岡山市の観光関連ツイートの共起語抽出、焼きなまし法を用いた観光経路の導出手法について述べる。

2. 岡山市における観光地 LOD

LOD とは、**LOD** とは、ウェブ上に存在するリソースを記述するための、統一されたフレームワークである **Resource Description Framework (RDF)**[9]を用いて作成される、グラフ構造を有するオープンデータである。

福井県鯖江市は、地域活性化のために、統計データを **LOD** 形式で公開し活用している[10]。一方、岡山市は **LOD** として使用できるファイルが少ない。そのため、この **RDF** をテキストで記述できるファイル形式である **Turtle** ファイルとして、岡山市における観光地 **LOD** を作成した。また、本研究では、観光地と観光に関係する店舗の情報を **Turtle** として **LOD** 化し、観光経路の推薦のための地理情報と、観光地と観光に関係する店舗情報をウェブサービスに用いる。**Turtle** として作成した、岡山市における観光地と観光に関係する **LOD** の記述例を下記する。

```
@prefix bp: <http://www.city.okayama.jp/(以下省略)> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema> .
@prefix cc: <http://creativecommons.org/ns#> .
@prefix geo: <http://www.w3.org/2003/01/(以下省略)> .
<1> rdfs:label "test"@ja ;
bp:郵便番号 "000-000" ;
bp:住所 "岡山県岡山市北区 0-0" ;
bp:電話番号 "086-XXX-XXX" ;
geo:lat "XX.XXXX"^^xsd:double ;
geo:long "XX.XXXX"^^xsd:double .
```

@prefix は、ある変数を指定した URL 先で定義されている **RDF** を使用するための記法である。<1>はデータの番号を表し、データ内に含まれる **bp** と **geo** は、それぞれ店舗情報と地理情報を表している。また、作成した岡山市の観光地と観光に関係する店舗の **LOD** を、**PostgreSQL** を用いてデータベースを構築した。

3. 岡山市における観光関連ツイートの共起語抽出手法

岡山市では旅行者へ向けた様々なイベントが開かれている。しかしながら、観光者は観光地に訪れていても、観光者が求めている観光に関する有益な情報を入手しているとは限らない。一方、Twitterへツイートしているユーザーはハッシュタグをつけてツイートしていることが多くみられる[11]。ハッシュタグは、様々な人に興味を持ってもらうために用いており、岡山市の観光に関するツイートから、任意の文書や文に、ある文字列とある文字列が同時に出現する語である、共起語を導出する。そこで、Twitterから岡山市に関するツイート情報を利用して、観光者に観光地を推薦する手法を提案する。本提案手法では、Twitterを用いて、岡山市の観光に関する有益な情報をユーザーのツイートから共起語を抽出する。そして、抽出した共起語において、日本語評価極性辞書[12]を用いてポジティブ/ネガティブを判定する感情分析[13]を行った後に、岡山市の観光地 LOD と合致させる。日本語評価極性辞書は、-1 から 1 までの値を各単語に付与した辞書である。評価極性タグは次の 4 つに分類されている。ポジティブ/ネガティブを主観的な「～である・になる（評価・感情）」、客観的な「～である・になる（状態）」である。今回、ツイートの文章を名詞、形容詞、動詞、副詞に品詞分解をして感情分析を行っている。ポジティブ/ネガティブの判定を行う際に、語句の値がニュートラル、すなわち 0 より低い値を持つ語句の場合であれば旅行者への推薦すべき観光地に適さないと捉え、そのような語句の入っているツイート文の選択をしない。語句の値が 0 以上の値を持つ語句の場合であれば、旅行者への推薦すべき観光地に最適に近いような値であるため、推薦の候補に選択する。これは、岡山に初めて観光しに訪れた旅行者に、可能な限り旅行者へ良い観光をしてもらうための手法である。

4. 焼きなまし法を用いた観光経路の導出手法

本研究では、ユーザーが、ある時間内で観光できるよう、決められた時間内で観光できる巡回路を導出しなければならない。そこで、閾値に達するまで計算を繰り返して、最適解もしくは準最適解を得ようとするメタヒューリスティックアルゴリズムである、焼きなまし法を用いる。本研究における焼きなまし法は、ある時間内に観光地を巡る最短巡回路を導出することを目的としており、OSMの地理情報を用いた Travelling Salesman Problem (TSP) の解決を行う[14,15]。焼きなまし法のパラメータを表1のように設定した。焼きなまし法では、「温度」というパラメータを設定する。「温度」とは仮想的な値であり、計算が進行すれば下がっ

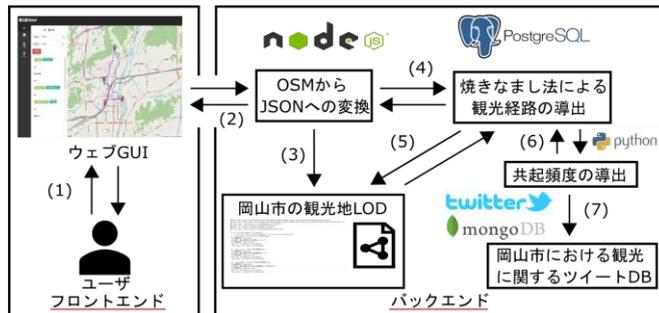


図1 提案システムの構成

ていく変数「冷却因子」を設定する。本研究の場合、「温度」を用いて言い換えると、温度が高いときは、経路が長くなったときであり、温度が低くなるにつれて経路が最短に近づいていることを意味する。試行回数とは、焼きなまし法における処理を開始してから、停止するまでに繰り返し行われる回数である。初期温度を 100 度と設定し、試行回数が 500 回に達したときまたは最終温度が 0.1 度になるときに処理は停止する。停止するまで処理は繰り返し行われる。焼きなまし法を用いた観光経路の導出は 100 回とする。TSP を解決するため、入力値を岡山市における観光地 LOD に記載された緯度経度とし、OSM の道路情報を基に、最短巡回路を導出するように実装した。

5. 提案システム

提案する、岡山市における観光経路の推薦ウェブサービスの構成を図1に示す。以下、提案システムの機能について、図1を参照しながら述べていく。はじめに、ユーザーが行きたい場所をウェブ GUI 上で選択する(図1(1))。このとき、OSMのデータをJSONに変換し(図1(2))、LODと複合して選択できるようにする(図1(3))。次に、選択した場所を基として、岡山市における観光地LODをOSMと複合して、最短巡回路となる観光経路を焼きなまし法で導出する(図1(4, 5))。また、ユーザーが選択した場所を基として、Twitterにおける岡山市の観光に関するツイートの共起頻度とマッチングすることで、観光地を自動生成し、ユーザーに推薦する(図1(6, 7))。ここで得られた観光地も含めて、観光地の最短経路を導出する(図1(4, 6))。そして、得られた観光経路をJSONでユーザーへと送信し、ウェブ GUI で可視化する(図1(1))。また、ユーザーが新たに行きたい場所を決めたとき、ウェブ GUI 上の入力フォームからカテゴリ名(美術館、ラーメン等)を入力するとデータベース内に構築済みのツイートより、Twitterの利用者によって評価されている観光地、または飲食店を導出する(図1(1, 2, 4, 5, 6, 7))。Twitterのツイート情報から得られた共起語が観光地と観光に係る店舗の情報がLOD化されてないものであれば、共起

表 3 最短巡回路の導出に用いる地理データ

地点	地名	緯度	経度	最短巡回路におけるソート順	元地点	地名
1	岡山駅	34.666121	133.917734	1	1	岡山駅
2	家電量販店	34.665908	133.919637	2	2	家電量販店
3	岡山県総合グラウンド	34.679223	133.915391	3	10	飲食店
4	岡山大学	34.687387	133.922223	4	8	ボーリング場
5	岡山理科大学	34.696164	133.927288	5	3	岡山県総合グラウンド
6	法界院駅	34.684793	133.927532	6	4	岡山大学
7	スーパーマーケット	34.679451	133.928678	7	5	岡山理科大学
8	ボーリング場	34.674849	133.919633	8	6	法界院駅
9	岡山済生会病院	34.67246	133.919757	9	7	スーパーマーケット
10	飲食店	34.669005	133.92161	10	9	岡山済生会病院

焼きなまし法を用いて、巡回路を導出したシミュレーション結果を図 3 に示す。シミュレーションの 100 回中、全て図 3 と同じ結果が得られたことで、実験的に準最適、もしくは最適に、距離の短い巡回路を導出できることがわかった。また、導出した観光経路を OSM に適用し、推薦ウェブサービス上の GUI で可視化した結果を図 4 に示す。図 4 より、巡回路を焼きなまし法によって導出できたことを視覚的に確認した。

7. まとめ

近年、岡山県は国内外の観光者が増加傾向にあり、その影響で観光案内の需要が高くなっている。そこで本稿では、岡山県において他の市町村と比較すると観光地が少ない岡山市を対象に、観光経路の推薦ウェブサービスを提案した。この観光経路の推薦のために、岡山市における観光地と観光に関する店舗の地理情報等をまとめた LOD を作成した。また、Twitter を用いて岡山市の観光に関するツイートの共起頻度の導出に加え、共起語から感情分析を行い、KH Coder を用いて共起ネットワークを可視化することで、飲食店の共起を確認することができた。加えて、岡山市における観光に関する LOD の地理情報を入力値として最短巡回路を導出する焼きなまし法を実装し、シミュレーションを行った。シミュレーション結果より、実装した焼きなまし法が、実験的に準最適、もしくは最適に、距離の短い巡回路を導出できることがわかった。そして、焼きなまし法によって導出した巡回路を、ウェブ GUI により可視化した。

今後は、提案したシステムを統合し、ウェブサービスとして実装する。また、現段階で Twitter から得たツイート数が 3,962 ツイートと少なく、共起語の出現が偏っていたため、引き続き岡山市に関するツイートの収集を行っていき、「くそ」+「おいしい」等の、スラングの選別を行う。加えて、観光地と観光に関する情報をまとめた LOD の種類を増やす。

表 4 最短巡回路の導出に用いるシミュレーションパラメータ

機能	値
試行回数	500
温度の初期値	100
温度の最終値	0.1
冷却因子	0.9
実験回数	100

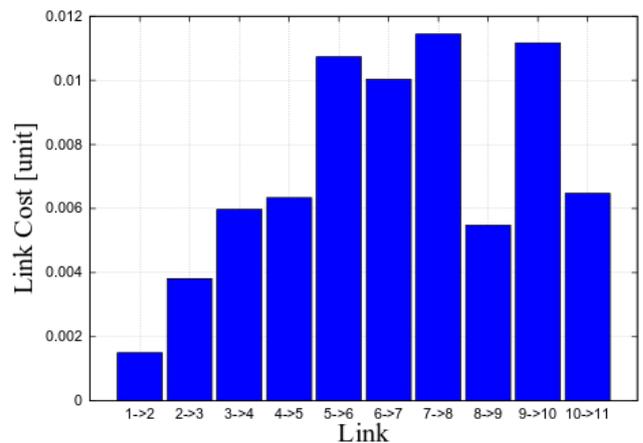


図 3 焼きなまし法を用いた最短巡回路の導出結果

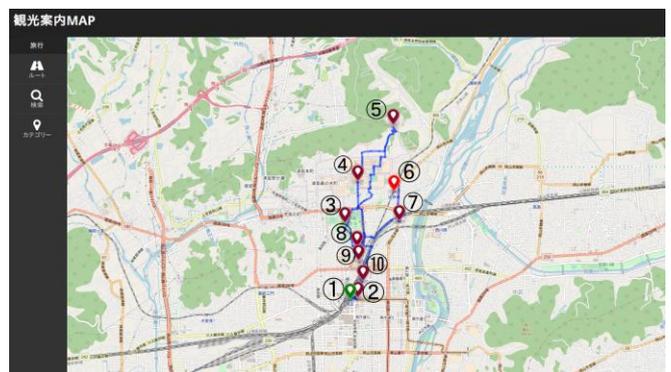


図 4 導出した観光経路のウェブ GUI による可視化

参 考 文 献

- [1] 日本政策投資銀行岡山事務所, “岡山のインバウンド観光の現状と今後の方向性 -DBJ・JTBF アジア・欧米豪 訪日外国人 旅行者の意向調査 (平成 29 年版)”, 拠点レポート (中国), pp. 1-25, 2018.
- [2] 佐藤 彩生, “都道府県別にみた日本人国内観光客数のトレンド -観光客数の少ない県でも 2011 年以降好調に推移-”, 金融市場, 28 巻, 12 号, pp. 42-45, 2017.
- [3] 岡山県産業労働部観光課, “観光客・その流れと傾向 -平成 29 年岡山県観光客動態調査報告書-” pp.8-11, 2018.
- [4] 鹿毛 宗一郎, 小田 哲也, “岡山市における観光経路の推薦ウェブサービスの開発”, 観光情報学会第 18 回研究発表会講演論文集, pp.5-7, 2018.
- [5] 小舘 亮之, 佐野 未来, ボルト ティモシー, 榎俊孝, 若原 俊彦, 田中 康裕, “オープンデータを活用した地域の観光関連情報発信サービスにおける課題”, 電子情報通信学会技術研究報告, Vol. 116, No. 488, pp. 123-128, 2017.
- [6] 若原 俊彦, 榎 俊孝, 高橋 和生, 河原 直樹, 中井 健登, 平井 麻未, 平城 裕大, 古賀 大騎, 小舘 亮之, 曾根原 登, “LOD を用いた観光 CMS プロトタイプの構成とネットワーク特性”, 電子情報通信学会技術研究報告, Vol. 117, No. 388, pp. 117-122, 2018.
- [7] “Twitter”, <https://twitter.com/>, Available Online: 10 January 2019.
- [8] 新谷 研, 角田 達彦, 大石 巧, 長尾 眞, “単語の共起頻度と出現位置による新聞の関連記事の検索手法”, 情報処理学会論文誌, Vol. 38, No. 4, pp. 855-862, 1997.
- [9] D. Brickley and R. Guha, “RDF Schema 1.1”, <https://www.w3.org/TR/rdf-schema/>, Available Online: 10 January 2019.
- [10] 榎 俊孝, 高橋 和生, 若原 俊彦, 山口 明宏, 曾根原 登, “地域の Linked Open Data の現状課題とその解決法に関する考察”, 信学技報, Vol. 116, No. 138, pp. 51-56, 2016.
- [11] 総務省情報通信政策研究所, “平成 29 年情報通信メディアの利用時間と情報行動に関する調査報告書”, pp. 1-96, 2018.
- [12] 小林 のぞみ, 乾 健太郎, 松本 裕治, 立石 健二, 福島 俊一, “意見抽出のための評価表現の収集. 自然言語処理, Vol. 12, No. 3, pp. 203-222, 2005.
- [13] 川島 崇秀, 佐藤 哲司, 神門 典子, “Twitter からの消費者ニーズの抽出手法に関する提案”, 第 8 回データ工学と情報マネジメントに関するフォーラム, B5-1, 2016.
- [14] Ł. Strąk, W. Wiczołek, A. Nowakowski, “Simulated Annealing for Finding TSP Lower Bound”, Proc. of ICCCI-2017, Vol. 10449, pp 45-55, 2017.
- [15] 山崎 紘揮, 宗久 知男, 宗久 保子, “巡回セールスマン問題への確率的状態選択法: 集団的焼きなまし法”, 人工知能と知識処理, Vol. 106, No. 617, pp. 33-35, 2007.
- [16] “Nominatim”, <http://nominatim.org/>, Available Online: 10 January 2019.
- [17] T. Kudo, K. Yamamoto, Y. Matsumoto, “Applying Conditional Random Fields to Japanese Morphological Analysis”, Proc. of EMNLP-2004, pp.230-237, 2004.
- [18] P. Jaccard, “The Distribution of The Flora in The Alpine Zone”, New Phytologist, Vol. 11, pp. 37-50, 1912.