

ウォーキング経路推薦のための口コミ分析に基づくエリア評価手法

山内 克之[†] 石坪 史帆[†] 桐生 拓海^{††} Panote Siriaraya^{†††} 栗 達^{†††}

河合 由起子^{††††} 中島 伸介^{††††}

[†] 京都産業大学 コンピュータ理工学部 〒603-8555 京都府京都市北区上賀茂本山

^{††} 京都産業大学大学院 先端情報学研究科 〒603-8555 京都府京都市北区上賀茂本山

^{†††} 京都工芸繊維大学 情報工学・人間科学系 〒606-8585 京都府京都市左京区松ヶ崎橋上町

^{††††} 京都産業大学 情報理工学部 〒603-8555 京都府京都市北区上賀茂本山

E-mail: [†]{g1745248,g1744096}@cc.kyoto-su.ac.jp, ^{††}i1986043@cc.kyoto-su.ac.jp, ^{†††}spanote@kit.ac.jp,
^{††††}{lida,kawai,nakajima}@cc.kyoto-su.ac.jp,

あらまし 近年、ダイエットや体力維持を目的としたウォーキング人口が増加しているものの、継続することに難しさを感じるユーザも多く、エクササイズとしてのウォーキングを支援することの意義は大きい。そこで我々は、特徴の異なるユーザにとって前向きに行えるウォーキング支援システムの開発に取り組んでいる。本稿では主にグルメスポットの口コミ分析によるポジティブエリア評価に基づくウォーキング経路推薦方式を提案し、評価実験ではポジティブエリア評価が妥当かどうかの被験者実験を行う。

キーワード ウォーキング支援, 口コミ分析, ポジティブエリア, ルート推薦

1. はじめに

近年の健康ブームによるダイエットや体力維持を目的としたウォーキング人口が増加している [1]。また、東京オリンピックに向けたスポーツ庁による運動を促進する取り組み [2] が行われており、エクササイズとしてのウォーキングを支援するシステムを開発する意義は大きいと考えている。しかし、楽しさを見出せず、ウォーキングを継続して行うことは非常に困難と感じるユーザも少なくはない [3]。また、新型コロナウイルスの影響で、多くのユーザが家で過ごす時間が増加した。しかし、家でのエクササイズも浸透していない為、運動不足と感じるユーザが増えつつある [4]。そこでウォーキングは安全で誰でも比較的行いやすい運動であると考えられるが、既存の歩行者ナビゲーションシステムは、ユーザが指定した目的地へ最短で移動するルートを推薦するものであり、ウォーキングそのものを楽しくさせる事を目的としたものは少ない。

そこで我々は、口コミや Twitter のジオタグ付きツイート内容、交通情報、事故データといったデータを用いて、ポジティブやネガティブといったユーザの心理を考慮したウォーキングルート推薦手法を提案することで、決められた範囲の中で、ユーザが安全で且つ、楽しいと感じるウォーキングを支援することを目指す。

本稿では主にグルメスポットの口コミを分析し、エリアの評価を行なった上で、ユーザにとって継続的にできるウォーキングを提案する。グルメスポットの口コミを使用することでユーザの特徴を考慮でき、さらにウォーキングに対するモチベーションにつながるのではないかと考えた。そして、エリア評価の妥当性を測る為、被験者にアンケートを答えてもらい評価実験を行う。

以下、2章では関連研究との比較を述べ、3章ではスポットエリアの評価手法及び経路探索アルゴリズムについての説明を行い、4章ではそのスポットエリア評価の妥当性を測るための評価実験、そして5章でまとめを述べる。

2. 関連研究

ウォーキングを行う上で、ウォーキングを継続または開始する「動機付け」が重要である。なぜならウォーキングは運動であり、多少なりとも負荷を伴う。そして一般的に辛い、つまらない、といった印象を抱き、動機付けや継続が難しいからである。

前田らの研究では、勾配データや心拍数、ジオタグ付きツイートをを用いて、ユーザにとって過度な負荷をかけ過ぎないウォーキングルート推薦の提案を行っている [5]。しかし、ジオタグ付きツイートの内容は考慮せず、ツイートの量だけでルート推薦を行っている。その為、「暗い」、「汚い」といったネガティブなツイートであったり、ネガティブでもポジティブでもない内容のツイートであっても、ツイートのデータ数に含まれてしまう恐れが考えられる。

武藤らは、ウォーキングコースの見どころポイントや、ウォーキングコースの共有機能を基に、スマートフォンを活用したウォーキング支援サービスを提案している [6]。見どころポイントや危険情報を共有機能により、情報を取得するが、ユーザの身体的負荷情報は個人により様々であるため共有機能では考慮されていない。しかし、見どころポイントといったポジティブエリアは考慮されているが、危険情報といったネガティブエリアについては考慮されていない。また、交通情報や事故データといったユーザが感じ取り難いデータについても考慮されていない。

Daniele らは、二枚の写真を基にどちらが好ましいのかを判定し、どの様な道を好むのかを機械学習させる。それにより、従来のナビの様に最短経路を推薦するだけでなく、綺麗に感じるルートを紹介している [7]。しかし、交通情報や事故データ、勾配については考慮されていないので、十分に安全性を考慮できていない。

川俣らは、四つの景観要素 (田園系, 山林系, 水辺系, 都市系) を元に、景観ベクトルに基づいたクラスタリングを行い、クラスタ間での大まかなルート探索を行う。そのルート探索結果より選ばれたクラスタに含まれるノード、リンクからさらに詳細なルート探索を行い、各景観のルート推薦を行う [8]。しかし、景観要素のみを考慮している為、交通情報、事故データといった安全性については考慮されていない。その為、どんなに景観が良くても、安全でない意味がない。また、この研究はドライブにおけるルート推薦手法であるのに対し、本研究はウォーキングにおけるルート推薦手法である。

Kim らはリアルタイムのジオタグ付きツイートから推測される感情より、極端にネガティブな感情を持つエリアを避け、最短距離よりわずかに増加した安全で楽しいルートを見つけることを提案している [9]。それに加え、シカゴ市ポータルの犯罪履歴データ内の犯罪率と、ネガティブなツイートデータが多く収められた地域間で有意な相関関係があることから、ソーシャルメディアの感情の意を利用し、犯罪スポットを迂回するルート推薦を検証している。しかし、交通情報については触れられていない。

Johnson らは美しさを重視したシーニックルート、安全性を重視したセーフティルート、そして経路の簡単さを重視したシンプルルートを作成し、従来の経路推薦手法との比較を行った [10]。しかしこの研究は、3つの観点をそれぞれ重視した経路の作成を行っており、本研究では、これら3つのルートを提示するのではなく、ウォーキングユーザに適した経路の推薦を行う。

3. 継続的にできるウォーキング経路の推薦方式

本稿ではユーザにとって継続的にできるウォーキング経路の推薦方式について述べる。3.1 節では、グルメスポットの口コミを分析し、その結果に応じたスポットの評価を算出した。それを元に3.2 節では経路探索アルゴリズムの作成手法を述べる。

3.1 スポット総合評価の算出法

ここではスポット評価の算出式を示す。本研究では口コミサイトじゃらん [11] のデータを使用した。スポット S_i の最終評価を $Score(S_i)$ とすると式 (1) で表すことができる。

$$Score(S_i) = MGenre(W_k) \times (RAve(S_i) + AScore(S_i))/2 \quad (1)$$

スポット S_i の最終評価を $Score(S_i)$ は、他のユーザによるスポット S_i の平均評価値 $RAve(S_i)$ と、口コミ分析に基づくスポット S_i の平均評価値 $AScore(S_i)$ 、これらの平均をスポッ

ト S_i のグルメジャンルがウォーキングユーザ W_k の好きなグルメジャンルとの合致率 $MGenre(W_k)$ で考慮したものである。 $AScore(S_i)$ は式 (2) で表すことができる。

$$AScore(S_i) = \frac{1}{Rn(S_i)} \sum_{U_j=1}^{Rn(S_i)} (mid + (UScore(U_j) - mid) \times UTrust(U_j)) \quad (2)$$

$$(1 \leq UScore(U_j) \leq 5, 0 \leq UTrust(U_j) \leq 1)$$

全ての口コミから類推できるスポットの評価 $AScore(S_i)$ の基本的な式構成はスポットページ内の任意のユーザ U_j が投稿した口コミページから類推できるスポットの評価 $UScore(U_j)$ の平均値であるが、口コミユーザ U_j の信頼度 $UTrust(U_j)$ を考慮したものである。 $UTrust(U_j)$ は 0 から 1 の範囲で値を取り、値が 1 に近づくほど信頼度が高い設定をした。信頼度が高いほど $UScore(U_j)$ の値が保たれ、低いほどスポット評価範囲の中央値 mid に収束する構成である。以下 3.1.1 節では口コミから類推できるスポットの評価 $UScore(U_j)$ の算出法を、3.1.2 節では口コミユーザの信頼度 $UTrust(U_j)$ の算出法を述べる。

3.1.1 口コミから類推できるスポットの評価の算出法

口コミユーザ U_j が投稿した口コミから類推できるスポットの評価 $UScore(U_j)$ は式 (3) で表すことができる。

$$UScore(U_j) = (RAnalysis(U_j) + GScore(U_j))/2 \quad (3)$$

$RAnalysis(U_j)$ は口コミ内容の感情分析値に 5 を掛けた値であり、 $GScore(U_j)$ は口コミユーザ U_j によるスポットの評価であり、この2つの項目の平均が $UScore(U_j)$ となる。口コミの感情分析値は Microsoft Azure の Text Analytics [12] という感情分析を行うツールで算出した、この値の範囲は 0 から 1 で、口コミの文章内容がネガティブであれば 0、ポジティブであれば 1 に近い値を取る。また五段階評価に統一するために 5 を掛けている。口コミユーザ U_j によるスポットの評価 $GScore(U_j)$ は基本的に口コミ内にある五段階評価されたご当地感、味、価格、サービス、雰囲気の平均値とした。しかし口コミによってはそれらが明記されていないものもある。その場合、同じく五段階評価されたそのスポットの全体評価値を使用する。

3.1.2 口コミユーザの信頼度の算出法

口コミユーザ U_j の信頼度 $UTrust(U_j)$ は式 (4) で表すことができる。

$$UTrust(U_j) = G \times GExi(U_j) + C \times CExi(U_j) + T \times TExi(U_j) + N \times NExi(U_j) + L \times LExi(U_j) + WP \times WPdep(U_j) \quad (4)$$

$$(G + C + T + N + L = 1)$$

口コミユーザ U_j の信頼度 $UTrust(U_j)$ は投稿した口コミより様々な観点から信頼度を測り、値が 1 に近いほど信頼度が高い

ユーザをした。その観点では、3.1.1 節で触れたご当地感などの評価の存在有無 $GExi(U_j)$ 、混雑具合の存在有無 $CExi(U_j)$ 、滞在時間の存在有無 $TExi(U_j)$ 、訪れた人数の存在有無 $NExi(U_j)$ 、口コミユーザのページリンクの存在有無 $LExi(U_j)$ 、スポットに行った時期と口コミを投稿した日の離れ具合 $WPdep(U_j)$ である。 $GExi(U_j)$ 、 $CExi(U_j)$ 、 $TExi(U_j)$ 、 $NExi(U_j)$ 、 $LExi(U_j)$ は存在する場合は値を 1 に、存在しない場合は値を 0 に定義し、 $WPdep(U_j)$ は初期値 0 で 1 年離れるごとに値を 1 減らすよう定義した。なお、 $WPdep(U_j)$ の値によって $UTrust(U_j)$ になることもあるが、その場合は $UTrust(U_j)=0$ とする。

3.2 経路推薦への応用

前項で算出した値を使用し、ウォーキング経路探索アルゴリズムを作成方法について述べる。スポットの総合評価 $Score(S_i)$ によってそのスポット付近の道路の仮想的なコストを変更する。 $Score(S_i)$ が高い場合、仮想的なコストを下げ、 $Score(S_i)$ が低い場合、仮想的なコストを上げるにより、評価の高いスポットへ優先的にルート推薦できるようになると考える。

4. 評価実験

この項では、3.1 節で算出したスポットの総合評価の算出結果及び妥当性を考察する。妥当性を測るため被験者 18 名を対象に、異なる特徴を持つウォーキングユーザ 3 名 (A さん、B さん、C さん) が、継続してウォーキングできる為に京都市内に実在する 10 件のスポットのうちどこを訪れれば良いかを五段階評価をした上で、上位 3 位のスポットを選択してもらったアンケートを実施した。ウォーキングユーザ 3 名の特徴は表 1、京都市内に実在する 10 件のスポットの概要は表 2 の通りに設定した。また、本実験では $MGenre(W_j)$ はウォーキングユーザの好きなグルメジャンルと一致しているスポットの場合、 $MGenre(W_j)=1.2$ 、一致しない場合は値を変えないために $MGenre(W_j)=1.0$ とする。

表 1 ウォーキングユーザの特徴

ユーザ名	年齢	性別	特徴	
			ウォーキング理由	好きなグルメジャンル
A	40	男	ダイエット	居酒屋
B	20	女	健康維持	スイーツ
C	60	男	長生き	和食

表 2 選択した 10 個のスポットの概要

スポット	ジャンル	評価分布の平均	口コミ数
1	その他, その他軽食・グルメ	1.00	1
2	居酒屋	3.00	1
3	居酒屋	3.87	15
4	居酒屋	4.67	6
5	カフェ・スイーツ, カフェ	2.34	3
6	カフェ・スイーツ, カフェ	3.75	4
7	カフェ・スイーツ, カフェ	4.67	6
8	和食, うどん・そば	3.00	2
9	和食, しゃぶしゃぶ・すき焼き	3.75	4
10	和食, しゃぶしゃぶ・すき焼き	4.60	5

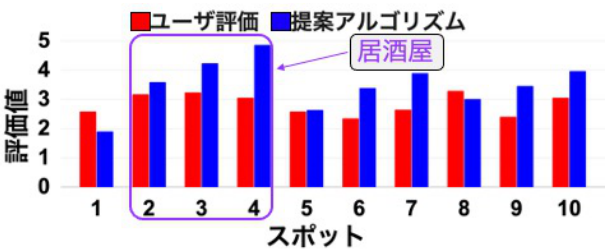


図 1 A さん (居酒屋好き) を想定したユーザ評価と提案手法の比較

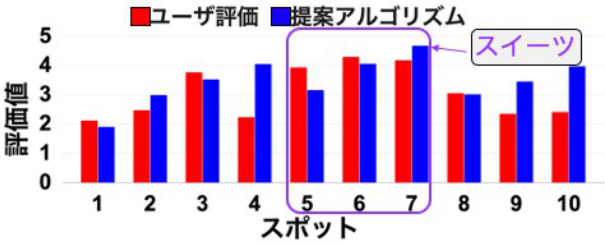


図 2 B さん (スイーツ好き) を想定したユーザ評価と提案手法の比較

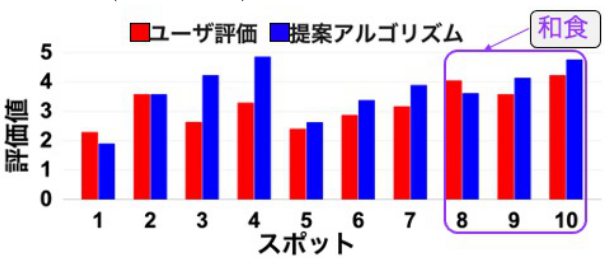


図 3 C さん (和食好き) を想定したユーザ評価と提案手法の比較

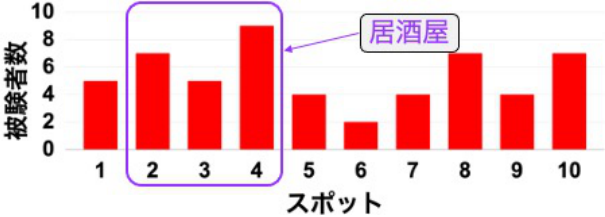


図 4 A さんにとってそのスポットが上位 3 件以内と選択した被験者数

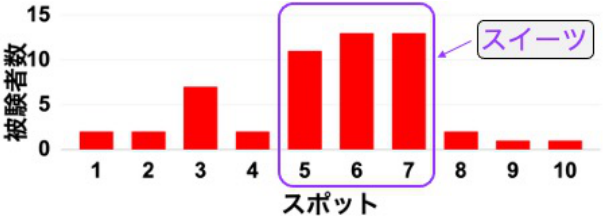


図 5 B さんにとってそのスポットが上位 3 件以内と選択した被験者数

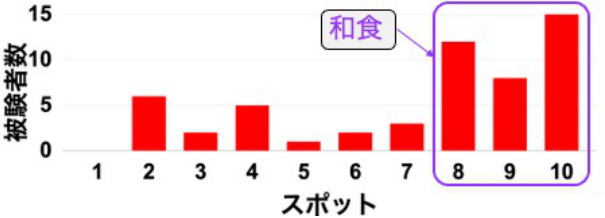


図 6 C さんにとってそのスポットが上位 3 件以内と選択した被験者数

図 1, 図 2, 図 3 はそれぞれの被験者を想定したユーザ評価と提案手法を比較した結果で、図 4, 図 5, 図 6 はそれぞれの被験者にとってそのスポットが上位 3 件以内と選択した被験者数の結果となっている。それぞれのグラフの中で紫で囲まれた

スポットがあるが、これは想定したウォーキングユーザの好きなグルメジャンルと一致したスポットである。

図1, 図2, 図3の提案アルゴリズムでの結果では紫で囲まれたスポット, 及びスポット4のような元々評価の高かったスポットの算出結果が高くなっている。よってある程度適切な結果を得たと言える

ユーザ評価の結果を見る上でまず, 図2及び図3を見てみると, 紫で囲まれたスポットの評価が比較的高くなり, 図5及び図6においても紫で囲まれたスポットに選択が集中していることがわかる。しかし, 図1では紫で囲まれたスポットは提案アルゴリズムより低い結果を得ている。また図4でも選んだスポットがばらけていることがわかる。これはAさんのウォーキング理由のダイエットが関係しているのではないかと考える。我々が意図していた継続的にウォーキングを行えるように, 好きなグルメジャンルと一致したスポットに選択した被験者と, あくまでダイエットという観点から好きなグルメスポットである居酒屋を避けた被験者がいたのではないかと考える。

5. ま と め

本項では, ユーザが継続的にウォーキングを行うための経路推薦アルゴリズムの提案を行った。このシステムの実現のために, まず口コミ分析によるエリア評価手法の開発を行い, それと同時にグルメスポットの評価の算出法の妥当性を測るため, 被験者による評価実験を行なった。実験の考察としては, まだ被験者数が少ないが, ウォーキングユーザの特徴に応じた評価の算出ができていることを確認できた, 今後は, エリア評価手法に基づくウォーキング経路推薦アルゴリズムの作成, 実際の道路ネットワークを使った評価実験, 提案アルゴリズムを用い, ジオタグ付きツイートなどのSNSデータで応用したエリア評価手法の拡張, さらにこのアルゴリズムに交通情報や事故データを用いて安全面も考慮できるような経路推薦アルゴリズムの作成を行い, 将来的には実用的な経路推薦の構築を行う。

謝 辞

本研究の一部は, 科研費基盤研究(B)(課題番号: 17H01822, 19H04118, 20H04293) および Society 5.0 実現化研究拠点支援事業, 京都産業大学先端科学技術研究所(ヒューマン・マシン・データ共生科学研究センター)の研究活動による。ここに記して謝意を表す。

文 献

- [1] 笹川スポーツ財団
”<https://www.ssf.or.jp/report/slodata/tabid/1404/Default.aspx>”
- [2] スポーツ栄養 web
”<https://sndj-web.jp/news/000369.php>”
- [3] お客様生活文化研究所”<https://www.asahigroup-holdings.com/company/research/hapiken/maian/bn/200611/00159.html>”
- [4] 水野映子:”コロナ禍”としての運動不足”<http://group.dai-ichi-life.co.jp/dlri/pdf/ldi/2020/wt2005b.pdf>”, LIFE DESIGN REPORT 2020.5
- [5] 前田幸道, 桐生拓海, Panote Siriaraya, 河合百合子, 中島伸介:”運動負荷および心理負荷を考慮したスマートウォーキングナビの提案”, DEIM Formu 2018, H3-5, 2018.
- [6] 武藤武, 佐々木喜一郎, 安田考美:”スマートフォンを活用したウォーキング支援サービスの検討”, 情報処理学会 第76回全国大会, 6V-5, 2014.
- [7] Daniele Quercia, Rossano Schifanella, Luca Maria Aiello:”The Shortest Path to Happiness: Recommending Beautiful, Quiet, and Happy Routes in the City”, HT’14 Proceedings of the 25th ACM conference on Hypertext and social media, Pages 116-125, 2014.
- [8] 川俣光司, 奥健太:”景観クラスタリングに基づく景観アウェアルート推薦システム”, DEIM Formu 2018, D1-1, 2018.
- [9] Jaewoo Kim, Meeyoung Cha, Thomas Sandholm:”SocRoutes: Safe Routes Based on Tweet Sentiments”, Proceedings of the 23rd International Conference on World Wide Web, Pages 179-182, 2014.
- [10] I. JOHNSON, J. HENDERSON, C. PERRY, J. SCHÖNING, B. HECHT:”Beautiful...but at What Cost? An Examination of Externalities in Geographic Vehicle Routing”, Proceedings of the ACM on Interactive, Mobile, Wearable, and Ubiquitous Technologies, 2017.
- [11] じゃらん.net, <https://www.jalan.net/>
- [12] Microsoft Azure Text Analytics, <https://azure.microsoft.com/ja-jp/services/cognitive-services/text-analytics/>

- [1] 笹川スポーツ財団
”<https://www.ssf.or.jp/report/slodata/tabid/1404/Default.aspx>”
- [2] スポーツ栄養 web
”<https://sndj-web.jp/news/000369.php>”
- [3] お客様生活文化研究所”<https://www.asahigroup-holdings.com/company/research/hapiken/maian/bn/200611/00159.html>”
- [4] 水野映子:”コロナ禍”としての運動不足”<http://group.dai-ichi-life.co.jp/dlri/pdf/ldi/2020/wt2005b.pdf>”, LIFE DESIGN REPORT 2020.5
- [5] 前田幸道, 桐生拓海, Panote Siriaraya, 河合百合子, 中島伸