

深層学習ネットワークへのHMM適用による分類パターン解釈 支援

Interpretation Support System for Classification Patterns from Deep Learning Networks using HMM

安藤 雅行^{1*} 河原 吉伸^{2,3} 砂山 渡⁴ 畑中 裕司⁵
Masayuki ANDO¹ Yoshinobu KAWAHARA^{2,3} Wataru SUNAYAMA⁴ Yuji HATANAKA⁵

¹ 滋賀県立大学大学院工学研究科

¹ Graduate School of Engineering, The University of Shiga Prefecture

² 理化学研究所革新知能統合研究センター

² RIKEN Center for Advanced Intelligence Project

³ 九州大学 マス・フォア・インダストリ研究所

³ Institute of Mathematics for Industry, Kyushu University

⁴ 滋賀県立大学工学部

⁴ School of Engineering, The University of Shiga Prefecture

⁵ 大分大学理工学部

⁵ Faculty of Science and Technology, Oita University

Abstract: This paper describes an interpretation support system for classification patterns extracted from deep learning with texts using HMM, and verified its effectiveness. It is well known that classification patterns by deep learning models are often difficult to interpret the reasons derived. In the proposed system, the content of deep learning results is extracted using HMMs, and classification patterns are provided for the system users to interpret the learned features. In verification experiments to confirm the effectiveness of the system, based on the learning result of deep learning classifying sentences, In the experiment, one group used the proposed system. The other group used the system that displays words with high TFIDF values. The results show that the subjects who used the proposed system were able to understand the meanings of the classification patterns of deep learning with texts more deeply than those who used the comparison system.

1 はじめに

インターネットの普及に伴い、また、SNS (Social Networking Service) の出現によって、画像、テキスト、数値データが大規模になり、その処理や情報の抽出に機械学習が使用されるようになってきた。しかし、従来の機械学習は大量のデータから規則などを学習し、分類・予測を行う際、データのどの特徴（画像なら色や形など）に注目するかは人間が指定する必要があった。そこで注目されるようになってきた技術が、深層学習である。深層学習は近年流行りだした機械学習であり、学習を行う層（入力データの規則などを学習する部分）を多層化している。これにより、より人間の

脳の学習に近い段階的な学習ができ、従来の機械学習と比べて学習の精度が高いという利点がある。

一方で、その深層学習による予測・分類基準が人間には不明な点が問題になってきている。特に、医療分野や自動運転では、その分類基準の理解は安全性において重要視されている。仮にテキスト分野においても深層学習の判断基準をより深く理解できれば、医療分野において新人とベテランの書いた電子カルテの違いから、良い電子カルテを書く方法を容易に理解でき、企業においても良い報告書や企画書を書く方法を短時間で習得できるなど、深層学習の新しい活用が期待される。

本研究では、構造が複雑になる代わりに、単語の出現の時系列や順序も考慮した学習が可能な、再帰的深層学習である RNN (Recurrent Neural Network) や LSTM (Long Short-Term Memory) を使用し、テキスト集合

*連絡先：滋賀県立大学大学院工学研究科 先端工学専攻 安藤雅行
〒522-8533 滋賀県彦根市八坂町 2500
E-mail: oh23mandou@ec.usp.ac.jp

の学習によって構築されたネットワークをHMM(Hidden Markov Model)に当てはめ、ネットワークの重みの値から、入力層に時系列順に入力される特徴量(本研究ではテキストを構成する単語)の尤度を算出する。その上で、単語の時系列パターンを尤度順に取り出すことで、再帰的深層学習の学習済みネットワークに蓄積された情報を、分類パターン(単純な単語の順序列)として抽出し、その解釈を支援するシステムを提案する。

以下本論文では、2章で関連研究について述べる。3章でHMMを利用した深層学習による分類パターンの抽出・可視化システムの構成と詳細について述べる。4章で提案システムの評価実験について述べ、5章で本論文を締めくくる。

2 関連研究

インターネットの普及などにより、急速に大規模化しつつあるテキストへの対策として活用され始めているのが、深層学習を用いたテキストマイニングシステムである[1, 2]。深層学習とは、一般に多層から構成されるニューラルネットワークを用いた学習を指し、例えば、深層学習の応用モデルである畳み込みニューラルネットワーク[3]の出現により、画像を用いた場合に限らず多くの場面で高い分類性能を実現できることが報告されている。

その一方で、深層学習には判断根拠のブラックボックス問題が存在している。深層学習は非常に複雑なプロセスによって情報を学習し、高い精度で予測・分類を行える。しかし、そのプロセスの複雑さにより、人間が深層学習の判断基準を説明することが非常に難しい。

深層学習のモデルへの信頼性・公平性の説明や判断基準への理解を重視した研究分野として、XAI(Explainable AI: 説明可能なAI)[4]が注目されてきている。XAIの研究としては、深層学習モデルの動作を理解・信頼するために、何を学習したか説明を行うことの必要性の提唱[5, 6]から始まり、実際に、モデル内のデータや変数間の相関関係から動作の説明を試みたり[7]、反事実的条件文を用いてモデルの動作をユーザに理解させる研究[8]等、モデルの動作自体を説明できないか試みる研究が行われている。また、モデルの動作の解釈だけではなく、モデルの動作の安定性・信頼性に注目し、悪意のあるデータへの対策[9]や、モデルの動作を別の論理回路や決定木に当てはめ、モデルの動作やその安定性を評価する研究[10, 11]も存在する。

そこで、テキストベースのDNN(Deep Neural Network)について、層ごとの学習の流れを単語情報でラベル付けして可視化し、分類基準を人間が理解できる学習ネットワークの解釈支援システムの研究[12]が存在し、一定の成果が確認された。一方で、上記の研究

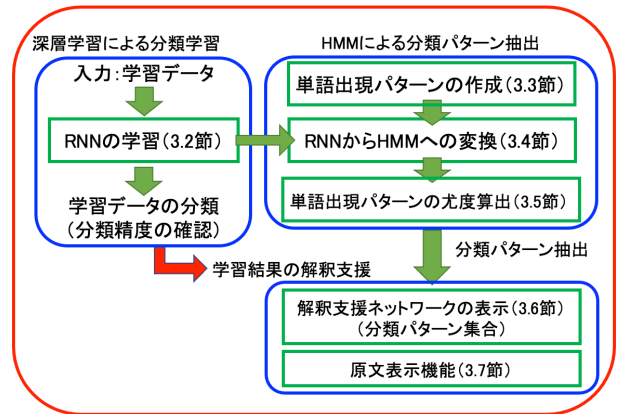


図 1: システムの構成

では深層学習モデルがDNNだったため、単語の時系列情報が失われ、その解釈も一定までしか得られない問題があった。

本研究ではこのような問題意識の下、文章の分類問題を例とし、深層学習モデルとしてRNNを使用することで、単語の時系列情報を含めた分類パターンからの分類基準の解釈支援システムの開発を目指す。

3 HMMを利用した再帰的深層学習ネットワークからの分類パターン抽出・可視化システム

本章では、テキストベースの分類タスクの深層学習ネットワークにおいて、HMMを利用した分類パターンの解釈支援を目的としたシステムの構成とその詳細について述べる。

3.1 提案システムの構成

提案システムでは、まず、図1に示すように、各分類先ごとにラベル付けしたテキスト集合をRNNにて分類し、その分類先を導いた学習ネットワークをHMMに当てはめ、提案システムの分類パターンの抽出処理部によって各出力(分類先)を導く分類パターンの尤度に基づく抽出を行う。最後に、システムの利用者は、システムの可視化処理部によって得られた学習ネットワークの表示を自分が見やすいように調整し、分類パターンを可視化する。また、システムでは分類パターンの意味を理解しやすくするための機能(解釈支援機能)を利用できる。

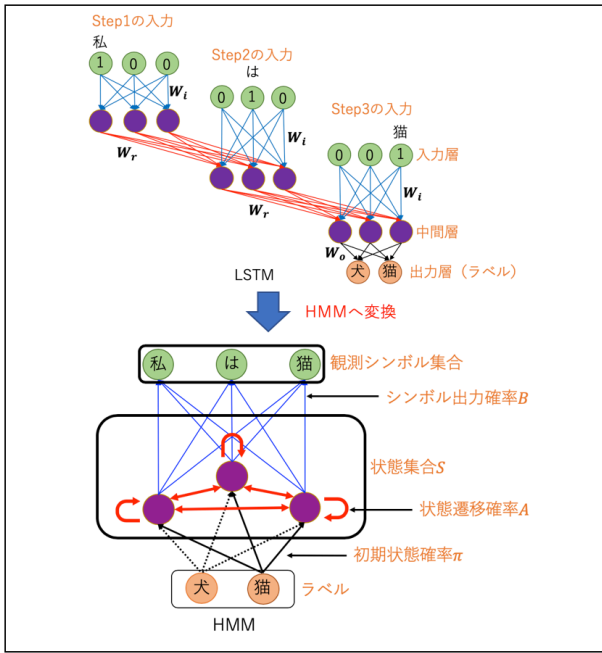


図 2: RNN から HMM への変換

3.2 深層学習による学習ネットワークの形成

3.2.1 テキスト中の単語のベクトル化

深層学習で学習を行う前に、テキストデータはテキスト中の単語を取り出したあと、単語を One hot 法 [13] と呼ばれる手法に従い単語ベクトルの羅列に直す。そして、テキスト中の各単語をその単語ベクトルに置き換え、深層学習への入力データとする。

3.2.2 学習によるネットワークの重み付け

One hot 法によって単語ベクトルの羅列に変換され、分類先ごとにラベル付けされたテキストデータは、RNN でそれぞれの出力層ノード (分類先) を導くネットワークへの重み付けがされていく。入力文章は各単語がベクトル化され、タイムステップごとに単語ベクトルが順番に入力されていく。また、LSTM での分類時は、最後の単語が入力されたタイミングで、出力層から分類結果が出力される。

3.3 HMM を用いた学習ネットワークからの分類パターンの抽出・可視化処理

3.3.1 RNN から HMM への変換

提案システムの分類パターンの抽出処理では、RNN によって得られた学習ネットワークを図 2 のように、ひとつの HMM として処理を行う。

まず、分類パターンの候補として、RNN への入力に使用した全単語の組み合わせを作成する。この時、組み合わせの条件として以下を満たす単語列を候補とする。

- 分類パターン候補の長さ (単語数) は任意で決めた長さで揃えるとする
- 分類パターン候補の単語の順序は実際のテキスト中の単語の出現順序に基づくものとする

次に RNN の入力層ノードを HMM の観測シンボル集合、中間層ノード (LSTM ユニット) を状態集合 $S = \{s\}$ とし、同様に中間層の (再帰的処理による) 時系列間の重みを状態遷移確率 A 、入力層中間層間の重みをシンボル出力確率 B とする。そして、中間層出力層間の重みを初期状態確率 π とするが、この π はその時選択するラベル (分類先) によって変わる。この時、観測シンボルによる観測系列 (前述した分類パターン候補) を $O = o_1, o_2, \dots, o_T$ (T は観測系列の長さ (前述した分類パターン候補の長さ))、状態数 (中間層ノード数) を N (状態番号は i, j) と置くと、状態遷移確率 A は式 (1)、シンボル出力確率 B は式 (2)、初期状態確率 π は式 (3) となる。

$$A = \{a_{ij} | a_{ij} = P(s_{t+1} = j | s_t = i)\} (1 \leq i, j \leq N) \quad (1)$$

$$B = \{b_{ij}(o_t) | b_{ij}(o_t) = P(o_t | s_{t-1} = i, s_t = j)\} \\ (1 \leq i, j \leq N, 1 \leq t \leq T) \quad (2)$$

$$\pi = \{\pi_i | \pi_i = P(s_0 = i)\} (1 \leq i \leq N) \quad (3)$$

この時、ある分類先 x に対して、単語出現パターン O が存在する時、初期状態確率を π_x と表すと、尤度 $P(O | \pi_x, A, B)$ は、式 (4) で算出される。

$$P(O | \pi_x, A, B) = \sum_{all S} P(S | \pi_x, A, B) P(O | S, \pi_x, A, B) \\ = \sum_{all s_0 \dots s_T} \pi_{x s_0} a_{s_0 s_1} b_{s_0 s_1}(o_1) \cdot a_{s_1 s_2} b_{s_1 s_2}(o_2) \cdot \dots \cdot a_{s_{T-1} s_T} b_{s_{T-1} s_T}(o_T) \quad (4)$$

最後に、全ての単語出現パターンについて式 (4) で尤度を算出し、尤度の高い順に、単語出現パターンを分類に寄与する分類パターンとして抽出する。

3.3.2 分類先を導く分類パターンの可視化

提案システムの可視化処理部では、分類先に強く結びつく、尤度の高い分類パターン集合が表示される。例として、5 種類の和菓子の作り方に関するテキスト集

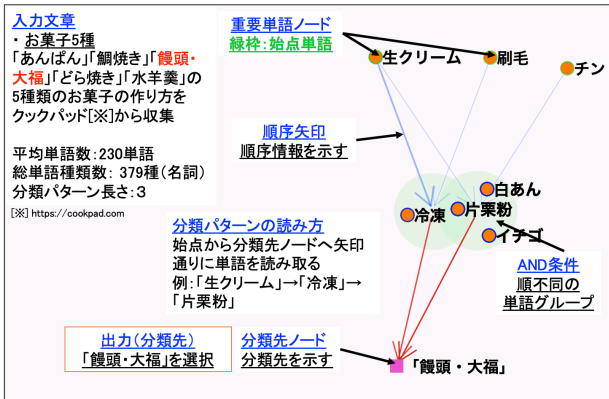


図 3: 提案システムの画面例

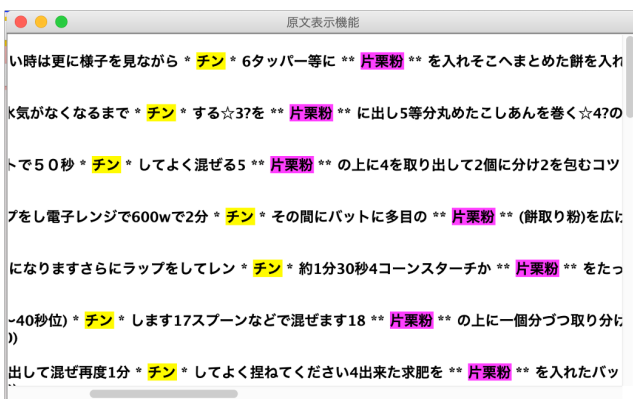


図 4: 原文表示例(「饅頭・大福」についてのテキストに対して単語「チン」と「片栗粉」を選択)

合¹の分類を行った場合の、提案システムのメイン画面を図3に示す。表示分類先は「饅頭・大福」を選択している。図3では、分類パターン中の単語の流れを矢印の向きで表し、分類パターンを構成する単語をノードで表している。また、尤度の大きさを矢印の太さで表している。

3.4 分類パターン解釈支援機能：原文表示機能

システムには、利用者が抽出された分類パターンの解釈行いやすいように、その表示内容の補足を行う機能がある。その中のひとつである原文表示機能について述べる。

分類パターンの解釈に向けて、単語情報だけでは、その単語が実際にどのような文脈で使われていたのかを把握することは難しい。そのため、原文表示機能により、分類パターン中の単語群が、実際に学習に用いたテキスト内でどのように使われているかを表示する。

¹クックパッド (URL:https://cookpad.com) から収集

ユーザは、解釈支援ネットワーク上で単語を選択することで、原文の中でその単語を含む文章が表示され、参照することができる。ただし、見やすさを考慮して、表示されるのは1つの文章につき、選択した単語の前10単語、後10単語までの区間とした。また、単語は最大2種類まで選択でき、その場合は単語間の文章は全て表示される。図4に5種類の和菓子の作り方について分類先「饅頭・大福」のテキストを使用し、単語「チン」と「片栗粉」を順番に選択した時の分類パターンの原文表示例を示す。

4 HMMを適用したテキスト分類パターン解釈支援システムの有効性の検証実験

本章では、深層学習の深い知見を有さない被験者が提案システムの出力する単語の出現パターンをもとに、分類パターンの解釈を行うことができるかを検証した実験について述べる。

4.1 実験手順

実験は、課題1「キャラセリフ」、課題2「家電レビュー」、課題3「ゲームレビュー」の3つの課題(詳細は下記参照)に対して、各課題ごとに指定する「出力ラベル」に分類される文章の分類パターンの解釈を行ってもらった。実験は深層学習についての深い見識がない16名の大学生、大学院生に対して行い、提案システムを用いるグループと、比較システムを用いるグループに8名ずつの2つに分けて行った。

実験では、学習結果を言葉で表して説明する提案システムとの比較として、分類先カテゴリに特有の単語を抽出するTFIDFから、分類先の特徴を解釈するシステムを用意した。TFIDF値を利用する比較システムを採用したのは、TFIDFによる特徴的な単語単体や、単語の組合せから解釈した結果と、提案システムによる単語の時系列関係を参照した解釈を比較することで、分類パターンの解釈における単語の時系列関係を提示することの有効性を明示できると考えたためである。

提案システムを用いるグループでは、提案システムを用いて、分類に寄与する単語(単語単体、組合せ、時系列順序)を見つけてもらった。比較システムを用いるグループでは、比較システムとして、指定する出力ラベルに特有の単語を式(5)のTFIDF値により抽出してリスト形式で提示するシステムを用意し、これらの単語を元に、分類に寄与する単語を見つけてもらった。また比較システムにおいても、提案システムの原文表示機能を利用できるようにした。

ある単語 i の $TFIDF_i =$ 単語 i の文章頻度

$$\times (\log(\frac{\text{出力ラベル数}}{\text{単語 } i \text{ の } DF \text{ 値}}) + 1) \quad (5)$$

実験手順について、以下のステップで両グループの被験者に解釈を行ってもらった。その際、提案システムで表示される分類パターン数は、単語数3個から構成される分類パターンを尤度の高い順に5つとした。また、比較システムでの表示単語数も提案システムに合わせて15個とした。

手順1 解釈対象の出力ラベルを選択する：課題1では「ツンデレ」、課題2と課題3では「役に立つ」に分類されるレビューを対象とした。

手順2 それぞれの選択した出力ラベルに対応した「解釈の目的」を読んで内容を理解する。

手順3 選択した出力について、「解釈支援ネットワーク」を表示させて、出力に寄与すると思われる特徴（単語単体や組合せ、時系列順序等）を10個見つける。

手順4 注目した特徴に対して、原文表示機能を用いながら考案してもらう。

学習データの詳細について述べる。課題「キャラセリフ」においては、Twitterの「ツンデレbot」「デレデレbot」「キャラセリフbot」からそれぞれ「ツンデレ」「デレデレ」「ノーマル」のキャラの特徴を持つセリフを500件ずつ、計1500件を利用し、「ツンデレ」を解釈対象とした。課題「家電レビュー」と課題「ゲームレビュー」においては、amazonの「人気家電製品」の上位50種類と「人気ゲームソフト」の上位100種類より、「役に立つ」（星4以上かつ、「役に立つ人数」が10以上）レビュー、「役に立たない」（星4以上かつ、「役に立つ人数」が0）レビュー、「低評価」（星2以下）レビューをそれぞれ課題「ゲームレビュー」では1036件ずつ、計3108件を利用し、課題「ゲームレビュー」ではそれぞれ1473件ずつ、計4419件を利用した。また、解釈対象は「役に立つ」を選択した。

学習はLSTMによって行い、中間層は1層とした。中間層のノード数は、分類精度が95%を下回らない範囲で、ノード数を減らす操作を行った。学習率は0.1、11ノルム係数、12ノルム係数はともに0.0001、学習回数は50回で学習を行った。

4.2 結果と考察

まず、被験者により記述された解釈の妥当性の内訳（被験者平均）を図5に示す。ただし、解釈の妥当性の内訳は、以下に定義する内容をもとに、著者の1名が分類を行った。

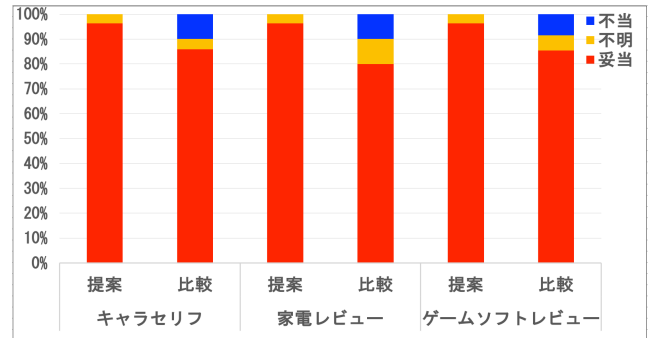


図5: 被験者の解釈の妥当性の内訳 (被験者平均)

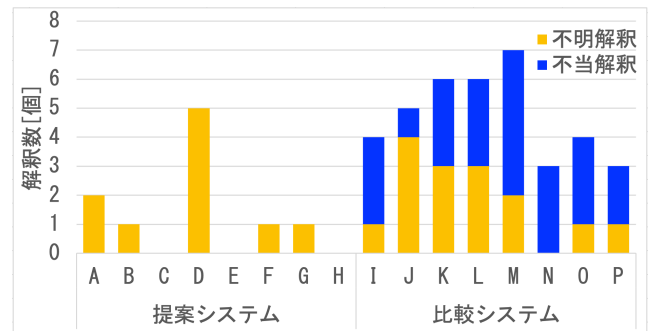


図6: 被験者別の不明, 不当解釈個数

- 妥当な解釈 (妥当)：内容の正しさが原文から確認でき、「解釈の目的」にも合っている。
- 妥当か判断できない解釈 (不明)：内容の意図がはっきりせず、妥当か妥当でないかが判断できない。
- 妥当でない解釈 (不当)：解釈の内容に誤りが確認できたか、または「解釈の目的」に合った内容ではない。

図5の結果から、提案システムでは、97%以上の解釈が妥当な解釈に分類され、その正しさが確認できた。特に、比較システムで10%近く存在している妥当でない解釈については、提案システムの結果ではひとつも見られなかった。また、妥当か判断できない解釈についても、比較システムでは全体の5%から10%程度に含まれていたが、提案システムでは全体の3%以下であった。このことから、提案システムではより意図が明確で妥当な内容の解釈が行われていたと言える。

また、被験者ごとの「妥当でない解釈 (不当解釈)」と「妥当かどうか判断できない解釈 (不明解釈)」の数を図6に示す。図6のAからHは提案システムの被験者8名、IからPは比較システムの被験者8名を表す。

図6より、「不明解釈」を行った被験者の数は、提案システムで5人、比較システムで7人と大きな差はなかった一方で、「不当解釈」を行った被験者は提案シス

テムで0人に対し、比較システムで8人全員となり、1人を除いて複数の「不当解釈」を与えていたことがわかる。そのため、個人差によらず提案システムを用いた方が、より妥当な解釈を与えられたことが確認できる。

以上をまとめると、提案システムでは、比較システムより、より正解率の高い、典型的で妥当な解釈が導き出せることが確認できた。これは、特に複数の単語の時系列関係に注目して解釈が行えることが要因と言える。

5 おわりに

本研究では、複数のテキストデータの分類を単語の時系列関係が学習できるRNNで行い、学習ネットワークの解釈を行うための分類パターンの解釈支援システムを提案した。本研究の特徴として、学習済みの再帰的深層学習のネットワーク構造をHMMに当てはめて処理することで、モデルの構造を変えることなく、容易に、学習された特徴量の時系列情報を抽出できる点が挙げられる。提案システムの有効性を確かめる検証実験では、提案する環境が、深層学習に精通していないユーザでも、時系列情報を含む分類パターンから、容易に原文の内容を広くカバーする妥当な解釈につながられることを確認した。

今後は、BERT (Bidirectional Encoder Representations from Transformers) などのさらに複雑な深層学習ネットワークを対象とした解釈環境を構築することを目指す。

参考文献

- [1] ボレガラ ダヌシカ, “自然言語処理のための深層学習”, 人工知能学会誌, Vol.29, No.2, pp.195-201, 2014
- [2] Ebru Arisoy, Tare N. Sainath, Brian Kingsbury, Bhuvaba Ramabhadran, “Deep Neural Network Language Models”, In Proceedings of the NAACL/HLT Workshop, Will We Ever Really Replace the N-gram Model?, pp.20-28, 2012
- [3] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition”, In Proceedings of the IEEE, 1998
- [4] D. Gunning, ‘Explainable artificial intelligence (xAI)’, Tech. rep., Defense Advanced Research Projects Agency (DARPA), 2017.
- [5] A. Fernandez, F. Herrera, O. Cordon, M. Jose del Jesus, F. Marcelloni, ‘Evolutionary fuzzy systems for explainable artificial intelligence: Why, when, what for, and where to?’, IEEE Computational Intelligence Magazine 14 (1), 69-81, 2019.
- [6] J. Haspiel, N. Du, J. Meyerson, L. P. Robert Jr, D. Tilbury, X. J. Yang, A. K. Pradhan, ‘Explanations and expectations: Trust building in automated vehicles’, Companion of the ACM/IEEE International Conference on Human-Robot Interaction, ACM, pp. 119-120, 2018.
- [7] O. Goudet, D. Kalainathan, P. Caillou, I. Guyon, D. Lopez-Paz, M. Sebag, ‘Learning functional causal models with generative neural networks’, Explainable and Interpretable Models in Computer Vision and Machine Learning, Springer, pp. 39-80, 2018.
- [8] R. M. J. Byrne, ‘Counterfactuals in explainable artificial intelligence (XAI): Evidence from human reasoning’, Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19, pp. 6276-6282, 2019.
- [9] X. Yuan, P. He, Q. Zhu, X. Li, ‘Adversarial examples: Attacks and defenses for deep learning’, IEEE Transactions on Neural Networks and Learning Systems 30 (9), 2805-2824, 2019.
- [10] G. Audemard, F. Koriche, P. Marquis, ‘On Tractable XAI Queries based on Compiled Representations’, KR Proceedings 2020 Special Session on KR and Machine Learning, pp.838-849, 2020.
- [11] Q. Zhang, Y. Yang, H. Ma, Y. N. Wu, ‘Interpreting CNNs via decision trees’, IEEE Conference on Computer Vision and Pattern Recognition, pp. 6261-6270, 2019.
- [12] 安藤雅行, 河原吉伸, 砂山渡, 畑中裕司, ‘テキストベースの深層学習における分類パターンの解釈支援’, 知能と情報 (日本知能情報ファジィ学会誌), Vol.31, No.4, pp.779-787, 2019.
- [13] M. Feurer, A. Klein, K. Eggenberger, J. Springenberg, M. Blum, and F. Hutter. “Efficient and robust automated machine learning”, In Neural Information Processing Systems (NIPS), 2015

CLIP を用いた画像ランキングによる パラメータ最適化に基づいた絵本の挿絵生成

Generated Images for Picture Book

based on Parameter Optimization by Image Ranking Using CLIP

齋藤 優也¹ 黄 潤和²

Yuya Saito¹, Runhe Huang²

¹法政大学大学院 情報科学研究科

¹Graduate School of Computer and Information Sciences, Hosei University

²法政大学 情報科学部

² Faculty of Computer and Information Sciences, Hosei University

Abstract: We propose a system for generating images from the text of a picture book to assist in image formation when reading. In the proposed system, each paragraph of the picture book is summarized, and an image is generated using VQGAN-CLIP of Text-to-Image model based on the summarized text on the top of the original content. In the process of generating an image, we also propose a ranking method for determining parameters (the seed value and learning times) based on the CLIP image score and the loss value of image generation. In the experiments, the relationship between the loss value and the learning time was investigated, and the learning time parameter was determined in consideration of the execution time.

1 はじめに

近年、様々なデータ化が進み、読書をする媒体も紙などの書籍から電子化された電子書籍へと変化している。スマートフォンの普及が進んだ背景もあり、場所や時間を問わずに簡単に読書を行うことが可能となっている。電子書籍を提供する Kindle など様々なサービスでは、読書をサポートするための様々な機能が搭載されている。分からない単語の意味の検索を行う単語検索機能や、メモをしたい箇所を保存しておくハイライト機能、文字を読みやすくするための拡大機能などがサポートされている。これらの機能は紙媒体の時代から行われていることを簡単に行えるようにしただけではあるが、電子書籍ならではの機能として、文章を自動で要約して読書の時間を短縮するような機能も注目を浴びている。

そこで、本研究では文章の内容を解析して、読書のサポートをする機能として、自動で本文から挿絵を生成することによって、読者のイメージ形成をサポートする機能を提案する。本研究によって提案されるシステムでは、各パラグラフに対して挿絵をそれぞれ生成する。各パラグラフの文章の内容を読み

取るために、PEGASUS[1]による文章要約を利用し、パラグラフの内容を 1~2 文にまとめ上げる。要約された文章を基に Text-to-Image と呼ばれる文章からその内容に見合った画像を生成する技術で挿絵を生成する。今回は VQGAN-CLIP[2], [3] と呼ばれるモデルを利用する。挿絵を生成する際は複数枚生成し、ランキングを行うことで最適な挿絵を生成できるようにする。本研究では、特に絵本の文章から画像の生成を行う。絵本は元々文章と画像がペアとなっているため、挿絵として画像を生成することに適している文章であると考えられるからである。最終的に絵本の各パラグラフに対して挿絵を生成し、提案システムによる絵本の作成を行えることを目的とする。

2 関連研究

2.1 文章要約

自然言語処理分野において、文章要約は入力とする文章から簡潔でかつ正確な要約を出力することを目的に研究が行われてきた。一般的に、文章要約に

は、入力の記事内から重要な部分を断片的に得て要約する抽出型要約と、入力の内容に沿った文章を生成することによる抽象的要約の2通りに大きく分けることができる。

抽出型要約には、文章をグラフ構造で表現し、文や単語の関係性を基に要約するグラフベースの TextRank[4]や、文章全体のトピックを算出し、そのトピックにあった文章を抽出するトピックベースの LSA(Latent Semantic Analysis)[5]を利用した手法など様々な手法が取り入れられてきた。

一方で抽象的要約においては、人が要約を作るように、文章の意味を理解したうえで適切な要約を生成する手法である。抽象型要約を実現するためには、長文の内容理解能力、文章の情報整理能力、新たな文章を生成する能力など、自然言語処理分野でも難しい処理が求められる。これらの複雑な処理を実現するために、機械学習が多く取り入れられている。文章データと要約文章を含む高品質な教師付きデータセットが多く公開されていることも機械学習が盛んになっている理由である。特に、RNN や Transformer を応用した Encoder-Decoder モデルが主流となっている。さらに、近年自然言語処理分野で大きな注目を浴びている BERT[6]を応用した PEGASUS によって、高品質な要約を生成することが可能となった。

2.2 Text-to-Image

Text-to-Image タスクとは、任意のテキストから、そのテキストの内容に沿った画像を生成するタスクのことである。従来、このタスクではあらかじめテキストと画像のペアを用意し、入力したテキストに対してそれらの画像を組み合わせる手法がとられてきた。しかし、機械学習分野の発展や豊富なデータセットの取得が可能となったことから、汎用的な画像が生成可能となった。このタスクは、アート生成やコンピュータによるデザインの補助などに応用できると考えられている。

様々な手法が研究されているタスクであるが、近年では特に、深層学習モデルを取り入れた研究が盛んになりつつある。深層学習モデルの一つである敵対性ネットワーク(GAN)を応用し、テキストを入力として与えることで直接画像を生成する手法がある。最初に GAN を応用した GAN-INT-CLS[7]をはじめ、高解像度画像が生成可能となる StackGAN[8]のように GAN を応用した多くの研究が行われている。その他にも機械学習モデルでは、OpenAI から発表された DALL-E[9]と呼ばれるモデルが注目を浴びている。自然言語処理分野で革新的な性能を示した GPT-

3[10]を応用し、膨大なデータセットを学習させることで汎用的な画像を生成することが可能となった。これにより、モデルが未知のデータに対しても画像生成が可能となるゼロショット学習が可能となり、優れた性能を示すことができた。

3 提案手法

本研究では、絵本の文章を入力テキストとして、そのテキスト内容に見合った画像を生成するためのシステムを提案する。提案システムのアーキテクチャは図1となる。

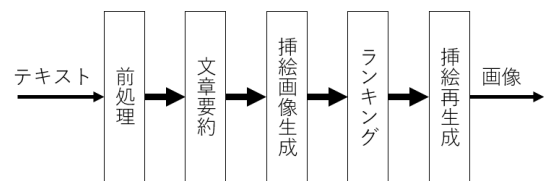


図1：提案システムのアーキテクチャ

入力されたテキストに前処理によって整形したのち、機械学習モデルによって文章要約を行い、要約文章を獲得する。その後、Text-to-Image のモデルによって、要約文章から画像を生成する。また、複数枚画像を生成したのち、ランキング方式で画像生成に必要なパラメータを決定する。最後に、決定したパラメータを用いて出力とする挿絵を再生成する。

なお、今回使用する文章は、英文の絵本を対象としている。前処理や要約、Text-to-Image には英文を利用したデータセットを使用する。

3.1 前処理

対象とする絵本の本文に対して前処理を行う。前処理は自然言語処理分野で一般的な手法で、文章中のノイズを除去し、機械のパフォーマンスを安定させることが目的である。提案システムで行う前処理は以下のとおりである。

- ストップワードの除去

“a”や“the”，“is”などの情報量の少ない単語をテキストから取り除くことで、重要な単語に焦点を当てることが可能となる。

- 不要な記号の除去

文章中に含まれる，“!”や“?”などの記号を除去し，“.”に変換させる。文章内の曖昧性を減らし、パフォーマンスを安定させる。

3.2 文章要約

文章要約では、Text-to-Image のモデルに与える入力の文章を生成することを目的とする。各パラグラフの挿絵を生成するためには、そのパラグラフの内容を最も示す文章を入力とすることが適している。そのための手段として、今回は文章要約による重要文の生成を行う。2.1 で文章要約に関する様々な研究について述べたが、本研究では、文章要約に特化した事前学習モデルである Pretraining with Extracted Gap-sentences for Abstractive Summarization Sequence-to-sequence models (PEGASUS)を用いた。PEGASUSによって、各パラグラフに対して、1~2 文程度の要約文章を得ることが可能となった。

PEGASUS の基本的な構造は、BERT と同様に Encoder-Decoder モデルであるが、最大の特徴は Gap Sentences Generation(GSG)という学習方法である。この GSG という方法は、事前学習に用いる学習方法が適用するタスクに似ていればいるほど、より高速かつ高い性能を発揮できるという方法に基づいている。つまり、文章要約用にモデルを学習するのであれば、学習方法は実際に文章要約に似た方法を利用することが好ましいということである。BERT などの事前学習モデルの学習方法では、Masked Language Model(MLM)と呼ばれる学習方法がある。この方法は、入力文章の一部分をランダムにマスクし、マスクした箇所を学習モデルが予測する方法で学習を行う。GSG では MLM の方法を、より文章要約のタスクに適するように工夫が加えられている。まず、MLM ではランダムにマスクする箇所を決定していたが、文章要約にはランダムに決定する方法は適していない。そこで GSG では、重要な文章をあえてマスクし、その箇所を予測させるように学習を行う。

3.3 画像生成

Text-to-Image の生成モデルは 2.2 で述べたように様々なモデルがあるが、今回は VQGAN と CLIP を組み合わせた VQGAN-CLIP と呼ばれるモデルを使用する。VQGAN-CLIP は様々な画像を生成することが出来ること、オープンソースとして誰でも使用可能となっていることから、AI によるアート生成の分野でも注目を浴びているモデルとなる。

このセクションでは、構成要素となっている VQGAN と CLIP について、そして VQGAN-CLIP としての全体のアーキテクチャについて述べる。

3.3.1 VQGAN

VQGAN は Vector Quantization GAN である。豊富

な表現度を持つ (GAN) によって得た特徴マップに対して、ベクトル量子化 (VQ) のプロセスを加えることで、高解像度の画像を生成したり、合成したりすることが可能となったモデルである。

従来の研究では、より高解像度の画像を生成するためのモデルとして、画像のピクセルの情報を直接 Transformer に与え、Transformer の利点である長距離の依存関係を応用するアプローチが検討された。画像内のコンテキストをより詳しく理解することが出来る手法ではあるが、Transformer の構造上、画像のピクセル数が増えるに応じて、必要となる GPU のスペックが高くなってしまいう問題点があり、高解像度の画像を生成することが難しかった。そこで、VQGAN はベクトル量子化の方法で利用し、画像内のピクセル情報をクラスタリングし、各グループにまとめ上げた。これにより、Transformer に与える情報量が削減し、より高解像度の画像の生成を行うことが可能となった。

本研究では、Variational Autoencoder(VAE)として VQGAN を利用することで、画像を生成するため Image Generator としての役割を担うこととなる。

3.3.2 CLIP

画像とテキストの関係性を得ることのできる事前学習モデルとして Contrastive Language-Image Pretraining(CLIP)が提案された。CLIP はインターネット上から収集した 4 億もの画像とテキストをペアとした事前学習モデルである。テキストには Transformer を用いた Text Encoder と、画像には ResNet-50、Vision Transformer(ViT)を用いて Image Encoder を学習させる。それぞれの Encoder を通して、画像とテキストの Embedding を取得し、それぞれのコサイン類似度が最大となるように Encoder を学習させることで、与えられた画像に対してキャプションを生成する事前学習モデルを実現した。

さらに、一般的な事前学習モデルの精度は、学習したデータセットに依存し、未知のデータに対して精度を上げることが出来なかったが、CLIP は学習したデータ以外でも精度が高くなるような Zero-shot 学習が可能となった。

本研究では、入力テキストの Embedding を得るための Text Encoder と、VQGAN によって生成された画像の Embedding を得るための Image Encoder として使用される。また、CLIP は画像とテキストの相関性をスコアで獲得することが可能であり、後述するランキングによるパラメータ決定の際にも利用する。

3.3.3 VQGAN-CLIP

先ほど説明した、VQGAN と CLIP を組み合わせた

VQGAN-CLIP によって Text-to-Image を行う。
 VQGAN-CLIP のアーキテクチャは、図 2 となる。

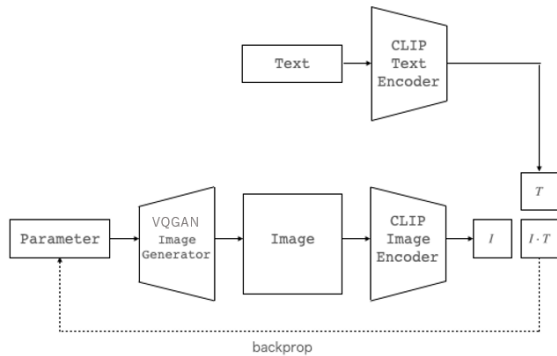


図 2 : VQGAN-CLIP のアーキテクチャ

VQGAN-CLIP は、生成したい内容のテキストと VQGAN が画像生成するために必要なパラメータを入力として与える。パラメータには、学習回数の上限を決める iter 値、生成される画像の内容を決定する seed 値、学習率を示す step size などがある。これらの入力を基に、画像を生成するためのプロセスは次の通りとなる。

①与えられたパラメータから Image Generator である VQGAN が画像を生成する。

②①で生成した画像に対して、CLIP の Image Encoder によって画像の Embedding を得る。同様に、入力として与えられたテキストに対しては、CLIP の Text Encoder によってテキストの Embedding を得る。

③画像 Embedding とテキスト Embedding を比較し、2 つの loss 値を求める。

④③で求めた loss 値が小さくなるように逆誤差伝播法で繰り返し学習を行う。

以上の①~⑤のプロセスを経て、生成する画像が入力されたテキストの内容に近づいていく。ただし、上記で説明したパラメータは自分で決定する必要があり、最適なパラメータを探すことは困難である。そのため、次セクションで述べる方法によって、パラメータを決定する方法を提案する。

3.4 ランキングによるパラメータ決定

3.3.2 で画像生成に必要なパラメータを決定するために複数枚画像を生成し、その中でランキングを行い、ランクが高い画像を生成したときに使用したパラメータを使用する。今回のランキングによって決定するパラメータは、生成される画像の内容を決定する seed 値と、画像を生成する際の学習回数を決める iter 値の 2 つとする。3.3.3 で述べた step size

は学習率を示すパラメータであり、生成される画像の内容に対する影響が少ないため、今回は 0.01 に固定して毎回学習を行う。

Seed 値の決定には、CLIP の画像スコアを利用する。3.3.2 で述べたように CLIP には、テキストと画像を与えることで、画像の内容がテキストにどれくらい即しているかを図るスコアを算出することが出来る。このスコアを生成したすべての画像で比較することによって、最もよい Seed 値を調べることが可能となる。

本研究では、与えられたテキストに対して、ランダムな seed 値を与え、25 枚の画像を生成する。各画像は生成後、CLIP による画像のスコアがそれぞれ算出される。それに加え、各画像を学習し生成した際の loss を基に以下の式(1)でランキングスコアを出す。

$$\text{Rank score} = \frac{\text{CLIPによる画像スコア}}{\text{画像を生成したときのloss}} \quad (1)$$

CLIP による画像スコアは、テキストと画像の相関性を示すスコアで、高い数値であればあるほど相関性があるといえる。また、画像を生成したときの loss は、低い値であればあるほど良い画像が生成できているといえる。つまり、Rank Score は高い値であるほど良い画像が生成されたといえる。

続いての iter 値の決定には、画像が生成されたときの loss 値を利用する。loss 値は先述した通り、低い値であれば、テキストと画像の間で誤差が小さいと評価することが出来る。つまり、loss 値が低い際の iter 値が学習回数として最も適しているといえる。今回は iter 値が 50 ずつ増えるごとに画像を保存し、loss 値を記録する。4 で行う実験によって、各 iter 値での画像の loss を参考にし、最も適した iter 値を決定する。

3.5 再生成

3.4 で決定したパラメータであるランキングが最も高い Seed 値を VQGAN-CLIP のパラメータとし、同じテキストで画像を再度生成する。学習回数 50 回ごとに画像と loss 値を記録し、学習終了後に loss 値が最も低いときの画像をシステムの出力とする。

4 実験

提案システムにおける、パラメータ決定方法の精度を確かめるために、実際に画像を生成した際の iter と loss の関係性を調査する。実験に使う絵本のデータは、無料の公開されている絵本から文章のみを抽出する。抽出した文章は、1 本分の絵本で 21 パラグ

ラフ分を使用し、21枚の画像を生成する。学習回数
 の上限を100回、250回、500回として学習を行う。
 その際に、iter値が50ずつ増えるごとにloss値を計
 測する。

文章要約に用いたPEGASUSは、cnn_dailymailを
 使用する。CNNとDaily Mailによって書かれた30
 万件強のニュース記事を含む英語のデータセットで
 ある。文章要約の機械学習に使われる一般的なデー
 タセットである。

VQGANには、ImageNetのデータセットを学習に
 用いたモデルを使用する。ImageNetは画像認識分野
 における標準的なデータセットであり、訓練用の画
 像データが120万枚用意されており、学習に十分な
 データセットである。

CLIPは、提案元であるOpenAIが公開されている
 事前学習モデルを使用する。3.3.2で述べたようにイ
 ンターネット上から収集した4億にも及ぶ画像とテ
 キストのペアを学習に利用している。

使用するGPUは、NVIDIA GeForce GTX 1080, 8GB
 を使用し、画像生成し、実行時間を計測する。

5 実験結果と考察

学習回数を上限100、250、500回として学習を
 行った際に、学習回数が50ずつ増えるごとにlossを
 計測する。その後、どのiter値が最も低いloss値と
 なったかを、パラグラフごとにまとめる。実行時間
 に関する結果を表1、iter値とloss値の関係を示した
 結果は表2となった。

表1：上限学習回数を行った際の平均実行時間

学習回数	100	250	500
実行時間(s)	48.1166	114.4604	225.50585
100との差		66.34385	177.3893
250との差			111.04545

表1の実行時間から、学習回数が増えると、実行
 時間も増加していることが分かる。これは、画像を
 生成するプロセスを繰り返し行っているためである。

次に、表2のiter値とloss値の関係性についての
 結果を見る。iter値を100とした時は、ほとんどのパ
 ラグラフで100回学習したときのloss値が最も低く
 なった。次に、iter値が250の時も同様に、ほとんどの
 パラグラフで250回学習した時のloss値が最も低
 くなった。続いて、iter値が500とした時は、平均し
 てiter値が350の時が最もloss値が低くなる事が分
 かる。このことから、最も低いloss値の時のiter値
 を探すには最低でも500回程度の学習が必要である

と考えることが出来る。

また、各iter値でのloss値の平均を見ると、上限
 100回から上限250回の時に約0.06下がっているの
 に対して、上限250回から500回の時には0.02ほど
 しか下がっていない。

表2：各パラグラフに対して、学習回数を定めた
 ときのloss値が最も低くなるiter値

パラグラフ	学習回数					
	100		250		500	
	iter	loss	iter	loss	iter	loss
1	100	0.755	250	0.692047	400	0.6839
2	100	0.858606	150	0.777282	400	0.738503
3	100	0.841877	250	0.734142	350	0.710746
4	100	0.783286	250	0.747524	250	0.744691
5	100	0.840254	250	0.78228	250	0.7875
6	50	0.76071	250	0.67835	300	0.662174
7	100	0.776059	250	0.689184	300	0.686342
8	100	0.806283	200	0.801432	250	0.720158
9	100	0.842293	250	0.743442	300	0.740903
10	50	0.75486	250	0.702702	400	0.686614
11	100	0.731423	200	0.689242	300	0.695985
12	100	0.733282	250	0.709531	450	0.702995
13	100	0.775316	250	0.758003	350	0.746047
14	100	0.774019	250	0.708915	500	0.709487
15	100	0.757584	250	0.720492	250	0.710087
16	100	0.824111	250	0.694436	500	0.666173
17	100	0.857855	250	0.758764	450	0.714617
18	100	0.792461	250	0.757109	400	0.721365
19	100	0.789375	250	0.73341	350	0.712016
20	100	0.789269	250	0.737588	300	0.714016
21	100	0.75595	250	0.688265	300	0.676064
平均	95.2381	0.79047	240.4762	0.728769	350	0.710971

表1と表2の結果から、ランキング前の画像を複
 数枚生成する際には、学習回数の上限を250回に
 し、ランキング後の再生成では、500回にすること
 が最も良いと考えた。複数枚の画像を生成するとき
 は、実行時間が増えすぎず、lossも最適なときにほ
 とんど変わらないためである。一方、ランキング後
 は質の良い画像を生成したいため、loss値が最も低
 くなるiter値が見つけれられる500回が最適であると
 考えた。

6 今後の展望

今回提案するシステムの評価として、画像生成の
 精度に関する評価を行う予定である。今回使用する
 モデルでのVQGAN-CLIPは、教師なし学習である
 ためモデルの評価が難しい。一般的にGANなどの
 教師なし学習では、教師あり学習と異なって、正解
 データがないためにモデルをどのように評価するか
 という問題がある。一般的に使われる性能評価の指

標として、Inception Score や FID などの指標があるが、ある程度の実際の画像が必要である。本研究のシステムでは、入力される文章に制限を設けていないため、生成される画像も多種多様となることもあり、特定の画像を用意することが困難であると考えた。そのため、生成された画像の評価には客観的評価として、アンケート方法を用いる予定である。アンケートによって評価する内容は、①文章から画像の整合性があるか(Text→Image)、②画像から文章を想像できるか(Image→Text)の評価を行う。アンケートで評価する内容は以下の通りである。

①要約文章から生成した画像は、その文章の内容にあっているかを検討する。生成された画像内のオブジェクト(人や建物など)、背景や風景、時間などはあっていると思うかについて評価する。

②生成された画像から、どのような文章を生成できるかを検討する。画像内にどのようなオブジェクトがあるように見えるか、風景や背景などの様子はどのように見えるかについて評価を行う。

①と②の評価によって、テキストと画像の双方向からの評価が出来ると考えている。

また、よりシステムの生成する画像の質を向上するために以下のような方法を検討する。

- ・生成画像の統一性

現在の画像は各パラグラフを別々に生成するために、それぞれの画像に統一性がない。生成画像の統一性は、絵本を作成するうえでは必要なものであると考えられる。画像生成モデルには、自分で画像を当てることによってその画像を参考に、入力テキストに合わせた画像を新たに生成することが可能であるため、これによって統一性を保証することが出来ると考えている。

- ・生成画像の画風変換

画像生成モデルは多様なデータを学習させているため、入力テキストに特定のテキストを追加することで、画風を変えることが可能である。例えば、生成する画像を「葛飾北斎」が書いたようにしたい場合は、入力テキストに「by Hokusai Katsushika」と追加することで、あたかも「葛飾北斎」が書いたような画像を生成することができる。

7 まとめ

今回、読者のイメージ形成のサポートとして絵本の文章から画像を生成するシステムを提案した。提案システムでは、生成するための文章を決めるために、文章要約の方法を用いて、事前学習モデルでの

ある PEGASUS を利用して、要約文章の生成を行った。その後、要約文章を入力テキストとして Text-to-Image のモデルである VQGAN-CLIP から、テキストの内容にあった画像を生成することが出来た。また、画像生成の際に必要なパラメータの決定を、複数枚の画像をランク付けし、最も適したパラメータを決定した。今後は、生成された画像の評価をアンケートによって行い、システムの調査を行っていく。

謝辞

本研究を進めるにあたり、ご指導いただいた黄潤和教授に感謝します。また、日頃の議論を通じて多くの知識や示唆をいただいた黄研究室の皆様にも感謝いたします。

参考文献

- [1] Zhang, Jingqing, et al. "Pegasus: Pre-training with extracted gap-sentences for abstractive summarization." International Conference on Machine Learning. PMLR, 2020.
- [2] Esser, Patrick, Robin Rombach, and Bjorn Ommer. "Taming transformers for high-resolution image synthesis." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.
- [3] Radford, Alec, et al. "Learning transferable visual models from natural language supervision." arXiv preprint arXiv:2103.00020 (2021).
- [4] Mihalcea, Rada, and Paul Tarau. "Textrank: Bringing order into text." Proceedings of the 2004 conference on empirical methods in natural language processing. 2004.
- [5] Ozsoy, Makbule Gulcin, Ferda Nur Alpaslan, and Ilyas Cicekli. "Text summarization using latent semantic analysis." Journal of Information Science 37.4 (2011): 405-417.
- [6] Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." arXiv preprint arXiv:1810.04805 (2018).
- [7] Reed, Scott, et al. "Generative adversarial text to image synthesis." International Conference on Machine Learning. PMLR, 2016.
- [8] Zhang, Han, et al. "Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks." Proceedings of the IEEE international conference on computer vision. 2017.
- [9] Ramesh, Aditya, et al. "Zero-shot text-to-image generation." arXiv preprint arXiv:2102.12092 (2021).
- [10] Brown, Tom B., et al. "Language models are few-shot learners." arXiv preprint arXiv:2005.14165 (2020).

逐次推薦状況におけるプロービングアイテム決定戦略に関する 検討

Consideration of Strategy for Selecting Probing Items in Intractive Recommendation

高間 康史^{1*} 徐 銘蔓¹ 柴田 祐樹¹
Yasufumi Takama¹ Mingman Xu¹ Hiroki Shibata¹

¹ 東京都立大学

¹ Tokyo Metropolitan University

Abstract: This paper compares several strategies for determining probing items. Probing items are presented to users especially at their sign-up process to get their feedback, from which user profiles are constructed. Therefore, how to select probing items are important for designing recommender systems. This paper reports the results of comparing several strategies by simulating user behaviors using a dataset.

1 はじめに

本稿では、逐次推薦状況において、ユーザの嗜好を推定するためのプロービングアイテム決定戦略について検討する。情報推薦を利用したサービスでは、アイテムに対するユーザの関心や嗜好に関する情報を収集し、これに基づいて推薦アイテムを決定する。特に、協調フィルタリングを用いたサービスでは、ユーザがアイテムに対して下した行為（閲覧、購入、評価など）に関する情報（インタラクションデータ）に基づいて、推薦アイテムを決定する [12]。研究においては、事前に収集された大規模インタラクションデータを対象とする場合が一般的であるが、実際のサービスを想定すると、サービスを通じてインタラクションデータを逐次収集していく必要がある。

インタラクションデータを収集するアプローチの一つとして、新規ユーザが推薦サービスの利用を始める際に、いくつかのアイテムを提示し、それらに対するフィードバックを得ることでユーザプロファイルを構築する事が考えられ、そのようなプロセスを本稿ではサインアッププロセスと呼ぶ。サインアッププロセスにおいて提示するアイテムは、得られるユーザプロファイルの質に影響を与えることから、提示すべきアイテムの選択は重要と考える。

本稿では、「ユーザのフィードバックを効率よく得られるようなアイテム」をプロービングアイテムと定義し、行為履歴のない新規ユーザにアイテムを逐次的に

推薦していく状況を想定して、プロービングアイテムの決定戦略について検討する。具体的には、ユーザプロファイルを構築するサインアッププロセスと協調フィルタリングを用いる推薦プロセスの2段階に分け、それぞれアイテムの推薦およびユーザによる評価（フィードバック）を繰り返しながらユーザプロファイルを逐次更新していく。実験協力者によるユーザ実験ではなく、データセット（評価値行列）を利用してユーザの行動を模擬することで逐次推薦状況を再現する。これにより、ユーザ実験よりも大規模なユーザ実験を低コスト・短時間で実施可能である。優れた推薦システムは、サインアッププロセスにおけるユーザの労力を最小限に抑えつつ、高い推薦精度を達成できるとの考えに基づき、複数のプロービングアイテム決定戦略の比較検討を行う。

2 関連研究

2.1 協調フィルタリング

情報推薦手法は、アイテムの属性を利用する内容ベースフィルタリングと、ユーザがアイテムに対して下した行為履歴に基づく協調フィルタリングに大別される。協調フィルタリングでは、行為履歴を評価値行列として表現し、これを入力データとしてアイテムに対するユーザの予測評価値を計算する [12]。

協調フィルタリングはメモリベース法とモデルベース法、及びこれらのハイブリッド手法に大別される。メモリベース法はユーザやアイテム同士の類似度を評価

*連絡先： 東京都立大学システムデザイン学部
〒191-0065 東京都日野市旭が丘 6-6
E-mail: ytakama@tmu.ac.jp

値行列から直接計算し、予測評価値を求める。代表的な手法に k-NN (k-Nearest Neighbor) [3] があり、ピアソン相関係数を用いて推薦対象ユーザと類似するユーザ集合を求め、類似ユーザが対象アイテムに対して下した評価値と、推薦対象ユーザとの類似度に基づいて予測評価値を求める。

モデルベース法は評価値行列から構築したユーザやアイテムのモデルを用いて予測評価値を求める。評価値行列から、潜在因子によるユーザベクトル、アイテムベクトルを求める行列分解ベース [8, 13] や深層学習ベースの手法 [4, 5] が提案されている。

2.2 コールドスタート

コールドスタートは協調フィルタリングの課題の一つとして指摘されており、新規ユーザや新規アイテムが増えた際に、行為履歴が蓄積されるまで精度良い推薦が行えなかったり、推薦ができない問題を指す。

これまでに、コールドスタート問題の解決策として様々な手法が提案されており、非個人化推薦と個人化推薦に大別される。非個人化推薦にはランダムな推薦 [6]、人気度やエントロピーに基づいた推薦手法 [10, 11] がある。個人化推薦では、アンケートに基づく推薦 [1, 18]、評価値行列の拡張による推薦 [7, 14] やソーシャル情報を利用する推薦手法 [9, 15] などが挙げられる。

2.3 対話型情報推薦システム

推薦システムの評価は、MAE (Mean Absolute Error) や RMSE (Root Mean Square Error) など予測評価値の誤差や、適合率、再現率など推薦リストに基づく評価が主流である。これらは、蓄積された行為履歴に基づいてユーザの嗜好を予測し、一度だけ推薦の提示を行う状況を想定した評価と言える。

一方、情報推薦システムを運用している状況を想定すると、行為履歴はアイテムの推薦を通じて逐次的に得られるため、ユーザに対しどのようなアイテムを提示するかは、その後の推薦アルゴリズムの挙動に影響を与えることになる。この様な、ユーザとの対話から得られる情報の活用を焦点を当てた、対話型情報推薦システム (conversational recommender systems) に関する研究が行われている。

対話型情報推薦システムにおいて、逐次的に得られるユーザフィードバックを手掛かりに適切な推薦アイテムを決定するプロセスを探索と利用のトレードオフ (exploration-exploitation tradeoff) としてとらえ、多腕バンディット問題 (multi-armed bandit problem) の枠組みを適用する研究も多い。Zhao らは PMF [13] を拡張したインタラクティブな協調フィルタリング手法

を提案している [19]。Thompson Sampling を利用してユーザに推薦するアイテムを決定し、フィードバック (評価) に基づきユーザプロファイルを逐次更新する。Christakopoulou らは Thompson Sampling の他、UCB (Upper Confidence Bond) や Greedy など複数のアイテム提示戦略について比較している [2]。

ユーザとのインタラクション設計に着目した研究として、Sun らは、自然言語による対話を通じた推薦システムを提案している [16]。アイテムが満たす条件・特徴を取得する事を目的としたモデリングのための対話と、推薦のための対話の両方を考慮しており、どちらの対話を行うかを決定する対話戦略を学習している。Wegbar らは、ユーザが満足するまで推薦を繰り返す事を想定した対話型映画推薦システムを提案している [17]。ユーザが満足するまで繰り返す反復的プロセスはユーザによる映画データベースの探索とみなせるため、映画をキャラクタと見立てたストーリーテリングによってより良い探索体験を提供することで、推薦結果に対する満足度向上を目指している。

3 逐次推薦状況の構築

逐次推薦状況を想定して情報推薦システムを評価する場合、実験協力者を募って行うユーザ実験と、データセット (評価値行列) に基づくオフライン実験の 2 つのアプローチが考えられる。前者は、実際のユーザの反応・フィードバックを取得できる利点があるが、大規模な実験を実施したり、多数のシステムやアルゴリズムを比較することが困難である。一方、オフライン実験は評価値行列に含まれない状況、例えば未評価アイテムに対する評価などが不明であるなどの欠点があるが、多数ユーザによる評価や、多数システム・アルゴリズムを同一条件で比較することが可能という利点がある。従って、ユーザ実験を行う前に、検証すべき仮説や評価対象とするシステム・アルゴリズムを絞り込む用途などに、オフライン実験は有効と考える。

以上の考えに基づき、本稿ではデータセットを利用したオフライン実験を採用する。本節では、利用するデータセット、逐次推薦の手順、ユーザフィードバックを取得するためのプロービングアイテム決定戦略について説明する。

3.1 データセット

対象データセットとして Grouplens が提供している MovieLens 100K Dataset¹ を使用する。このデータセットにはユーザ 943 人が映画 1,682 件に対し、評価 1~5 の 5 段階で下した 100,000 件の評価を含んでいる。

¹<https://grouplens.org/datasets/movielens/>

逐次推薦を行う対象ユーザは、その行動を模擬する必要がある事から、評価件数が少ないユーザは対象として不適である。本稿では、評価件数が200以上のユーザを推薦対象ユーザとする。データセット内の全ユーザのうち、本条件を満たすユーザは149人存在した。他のユーザも、推薦対象としては利用しないが近傍ユーザなどとしては利用する。

3.2 逐次推薦手順

推薦対象ユーザは、当該推薦システムを初めて利用する新規ユーザである事を想定する。新規ユーザは推薦アイテムを決定するために必要な行為履歴を持たないため、本稿では以下の2段階に分けて推薦を行う。

- サインアッププロセス：ユーザプロフィール構築を目的とした推薦を行う
- 推薦プロセス：ユーザが好むアイテムを推薦することを目的とした推薦を行う

サインアッププロセスは、推薦プロセスにおいて有効なユーザプロフィールを構築するためのフィードバックを得ることが目的であり、このプロセスで提示するアイテムを本稿ではプロービングアイテムと呼ぶ。推薦プロセスでは、一般的なユーザベース協調フィルタリング(KNNWithMeans²)を利用する。

各推薦対象ユーザ u_t に対して、以下の手順で推薦を逐次実行する。ここで、 V_t は u_t が実際に評価したアイテム、 $R_t \subset R$ (R は評価値行列) は u_t の実際の評価履歴とする。

1. 初期設定：元のデータセットから R_t を除去し、 $V'_t = \{\}, R'_t = \{\}$ とする。
2. サインアッププロセス：以下の推薦を10回繰り返す。
 - (a) V'_t に含まれないアイテムを後述する決定戦略によってソートし、リストとして u_t に提示する。
 - (b) u_t はリストの上位から評価したことがあるアイテムを探して行き、最初に発見したアイテムを $v_i (\in V_t)$ とする。
 - (c) v_i に対する u_t のフィードバック (評価値) $r_{ti} (\in R_t)$ を取得し、 V'_t, R'_t に v_i, r_{ti} をそれぞれ追加する。
3. 推薦プロセス：以下の推薦を10回繰り返す。

- (a) KNNWithMeans を用いて、 R'_t に基づき $v_i (\in V_t - V'_t)$ の予測評価値 \hat{r}_{ti} を求め、最大値をとるアイテム v_m を u_t に提示する。
- (b) v_m に対する u_t のフィードバック $r_{tm} (\in R_t)$ を取得し、 V'_t, R'_t に v_m, r_{tm} をそれぞれ追加する。

サインアッププロセスにおいて、 V_t に含まれるアイテムを u_t が過去に評価したことのあるアイテムとして扱う。提示されたアイテムに対するユーザのフィードバックを取得する必要があるため、推薦プロセスにおいても V_t から推薦アイテムを選択する。

3.3 プロービングアイテム決定戦略

サインアッププロセスにおけるプロービングアイテム決定戦略として、本稿では以下の11種類を採用し、比較する。

1. Random: ランダムにアイテムをソート
2. Popularity: 人気度 (評価件数) f_i が高いアイテムを優先
3. Variance[6]: 評価値の分散が大きいアイテムを優先
4. Mean: 平均評価値が高いアイテムを優先
5. Pure entropy[6]: エントロピーが大きいアイテムを優先

$$H(v_i) = - \sum_j p_{ij} \log(p_{ij}) \quad (1)$$

p_{ij} はアイテム v_i に対する全評価のうち評価値が j である割合を示す。

6. Entropy0[11]: 評価値行列の欠損値を0で置換してから各アイテムのエントロピーを計算し、大きいアイテムを優先
7. Ent*Pop[10]: $H(v_i)$ (Pure entropy) と f_i の積が大きいアイテムを優先
8. Ent*(log Pop)[10]: $H(v_i)$ (Pure entropy) と $\log f_i$ の積が大きいアイテムを優先
9. HELF[11]: エントロピーと評価件数の調和平均が大きいアイテムを優先

$$LF(v_i) = \frac{\log f_i}{\log |R|} \quad (2)$$

$$H'(v_i) = \frac{H(v_i)}{\log 5} \quad (3)$$

$$HELF(v_i) = \frac{2LF(v_i)H'(v_i)}{LF(v_i) + H'(v_i)} \quad (4)$$

²<http://surpriselib.com/>

$|R|$ は全評価件数であり、式 (3) は評価値 (5 段階) の異なり数で $H(v_i)$ を補正している。

10. Item-based (Popularity): 初回は人気度が高いアイテムを優先, 2回目からアイテム近傍ベース協調フィルタリングを用いて予測評価値が高いアイテムを優先
11. Item-based (Random): 初回はランダムにアイテムをソート, 2回目からアイテム近傍ベース協調フィルタリングを用いて予測評価値が高いアイテムを優先

4 評価実験

4.1 評価指標

本稿では、サインアッププロセスにおけるユーザの労力、推薦プロセスにおける推薦性能の二つの観点からプロービングアイテム決定戦略を評価する。前述の通り、サインアッププロセスでは各プロービングアイテム決定戦略に従いアイテムをソートして推薦対象ユーザに提示し、ユーザはリストの上位から自身が視聴したことのある映画を探し、見つけたらそれに対する評価をシステムに伝えるプロセスを繰り返す。このとき、視聴したことのあるアイテムを見つけるまでに要する時間がユーザの労力とみなせる。そこで本実験では、提示されたアイテムのリストにおける、実際に評価したことのあるアイテムの最高順位に基づいてユーザの労力を評価する。この順位が高いほど、サインアッププロセスにおけるユーザの労力は少ないことになる。

推薦プロセスにおける推薦性能の評価には RMSE と MAE を用いる。RMSE と MAE は評価値の予測における精度評価の指標であり、それぞれ以下の式で定義される。

$$RMSE = \sqrt{\frac{1}{|R|} \sum_{r_i \in R} (\hat{r}_i - r_i)^2} \quad (5)$$

$$MAE = \frac{1}{|R|} \sum_{r_i \in R} |\hat{r}_i - r_i| \quad (6)$$

ここで、 r_i , \hat{r}_i はそれぞれ、実際の評価値と予測評価値を表す。トップ k 推薦を想定して適合率、再現率で評価することも考えられるが、本稿では推薦性能の時間的変化を確認する意図もあり、予測誤差による評価を採用した。

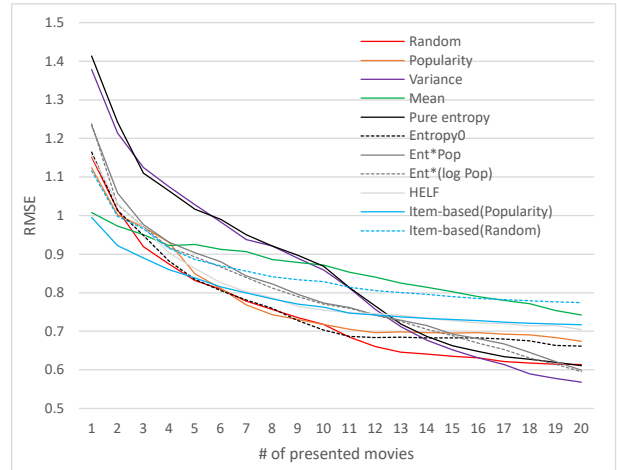


図 1: 各決定戦略による RMSE の結果

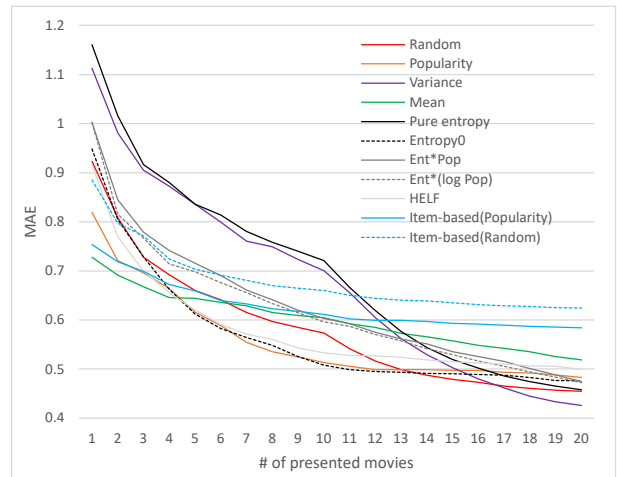


図 2: 各決定戦略による MAE の結果

4.2 実験結果

評価件数が 200 以上のユーザ 149 人に対して、それぞれ 3.2 節に示した手順で全 20 回 (サインアッププロセス、推薦プロセス各 10 回) の推薦を行った。推薦回ごとの全ユーザの平均 RMSE, MAE を求めた結果を図 1, 2 にそれぞれ示す。また、サインアッププロセスにおけるユーザの労力に関する指標として、視聴したことのある映画の最高順位の全ユーザ平均を求めた結果を図 3 に示す。

図に示した結果に基づき、サインアッププロセスにおけるユーザの労力、推薦プロセス開始時および終了時の推薦性能に関して各プロービングアイテム決定戦略を比較した結果を表 1 に示す。

サインアッププロセスでは、人気度を考慮した戦略でユーザ労力が低い傾向にある。人気アイテムはユーザが知っている可能性が高いため、サインアッププロセスにおけるユーザ労力の削減に貢献したと考える。

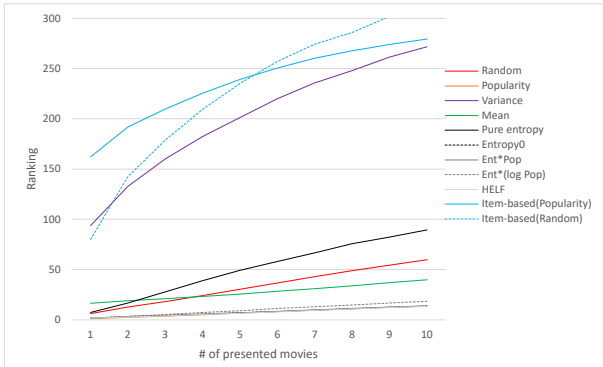


図 3: サインアッププロセスにおけるユーザの労力

表 1: 各決定戦略の比較

決定戦略	ユーザ労力	推薦性能 (開始時)	推薦性能 (終了時)
Random	△	○	○
Popularity	○	○	△
Variance	×	×	○
Mean	△	△	△
Pure entropy	△	×	○
Entropy0	○	○	△
Ent*Pop	○	△	○
Ent*(log Pop)	○	△	○
HELF	○	△	△
Item-based (Popularity)	×	△	×
Item-based (Random)	×	×	×

Entropy0 では未評価が多いアイテムのエントロピーが低下する効果があり、人気度の考慮と同様の効果があったと考える。一方、Variance や Pure entropy など、評価のばらつきのみを考慮した戦略ではユーザ労力が高い傾向があり、人気度との組み合わせなどが必要と考える。

推薦プロセス開始時の推薦性能は、Random, Popularity, Entropy0 で良好な結果が得られている。Random では様々なアイテムを提示することがユーザプロフィールの質向上につながり、推薦プロセス開始時に高い推薦性能が得られていると考える。しかし、サインアッププロセスでユーザが知っているアイテムを上位で提示できない場合があるため、ユーザの労力は Popularity よりも劣る結果となったと考える。

推薦プロセス終了時の推薦性能は、Random, Variance, Pure entropy, Ent*Pop, Ent*(log Pop) で他よりも良好な結果が得られている。Variance と Pure entropy は、多様なアイテムについてのフィードバックを

得ていたことが最終的な性能向上につながったと考える。反対に、Popularity の推薦性能は推薦プロセスを通じてあまり変化しておらず、サインアッププロセスで多様なアイテムについてのフィードバックが得られなかったことが影響した可能性があると考えられる。エントロピーと人気度の両方を考慮した戦略である Ent*Pop, Ent*(log Pop) は、ユーザ労力と推薦プロセス全体を通じた性能の両面で良い結果が得られていると言える。

Item-based (Random, Popularity) はどの評価指標においても他の戦略より劣る結果となった。このことは、新規ユーザに対する推薦が困難という協調フィルタリングの欠点が反映されたものといえ、システム利用開始時の早い段階でユーザに提示するプロービングアイテムの選択が重要であることを示唆していると考えられる。

5 おわりに

本稿では、逐次推薦状況においてユーザプロフィールを構築するために提示するプロービングアイテムの決定戦略について、データセットを用いた比較実験を行った。ランダムな選択や人気度を考慮した選択など、11 種類の決定戦略について実験を行った結果、人気度を考慮した戦略はユーザ労力が抑えられること、エントロピーと人気度の両方を利用した戦略はユーザ労力と推薦性能のバランスが良い結果が得られることを確認した。

本稿の課題として、データセットを利用した実験であるため、ユーザが評価したことのあるアイテムについてしかフィードバックが得られないことが挙げられる。また、ユーザが過去に評価したアイテムは、ユーザの嗜好や知識に影響を与え、結果としてそれ以降のアイテム評価に影響することが考えられるが、今回の実験では考慮することができていない。これらの課題に関しては、実験協力者を募ってのユーザ実験の実施などが必要と考えるが、コストのかかるユーザ実験前に、検証すべき戦略の絞り込みをするためには、本稿で採用したアプローチは有効と考える。

謝辞

本研究の一部は JSPS 科研費 19K22896, 21H03553 の助成を受けたものです。

参考文献

- [1] M. Aharon, O. Anava, N. Avigdor-Elgrabli, D. Drachler-Cohen, S. Golan, O. Somekh. Ex-

- cuseme: Asking users to help in item cold-start recommendations, 9th ACM Conference on Recommender Systems, pp. 83–90, 2015.
- [2] K. Christakopoulou, F. Radlinski, K. Hofmann. Towards Conversational Recommender Systems, KDD'16, pp. 815–824, 2016.
- [3] T. Cover, P. Hart. Nearest Neighbor Pattern Classification, IEEE Transaction on Information Theory, Vol. IT-13, pp. 21–27, 1967.
- [4] T. Ebesu, B. Shen, Y. Fang. Collaborative Memory Network for Recommendation Systems, SIGIR'18, pp. 515–524, 2018.
- [5] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, T.-S. Chua. Neural Collaborative Filtering, WWW2017, pp. 173–182, 2017.
- [6] A. Kohrs, B. Merialdo. Improving Collaborative Filtering for New Users by Smart Object Selection, International Conference on Media Features, 2001.
- [7] X. Lam, T. Vu, T. Le, A. Duong. Addressing Cold-Start Problem in Recommendation Systems, 2nd International Conference on Ubiquitous Information Management and Communication, pp. 208–211, 2008.
- [8] D.D. Lee, H.S. Seung. Algorithms for Nonnegative Matrix Factorization, NIPS'00, pp. 556–562, 2000.
- [9] C. Li, F. Wang, Y. Yang, Z. Li, X. Zhang. Exploring Social Networking Information for Solving Cold Start in Product Recommendation, International Conference on Web Information Systems Engineering, pp. 276–283, 2015.
- [10] A. Rashid, I. Albert, D. Cosley, S. Lam, S. Mcnee, J. Konstan, J. Riedl. Getting to Know You: Learning New User Preferences in Recommender Systems, 7th International Conference on Intelligent User Interfaces, pp. 127–134, 2002.
- [11] A. Rashid, G. Karypis, J. Riedl. Learning Preferences of New Users in Recommender Systems: An Information Theoretic Approach, ACM SIGKDD Explorations Newsletter, Vol. 10, Issue 2, pp. 90–100, 2008.
- [12] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, J. Riedl. GroupLens: An Open Architecture for Collaborative Filtering of Netnews, ACM Conference on Computer Supported Cooperative Work, pp. 175–186, 1994.
- [13] R. Salakhutdinov, A. Mnih. Probabilistic Matrix Factorization, NIPS'07, pp. 1257–1264, 2007.
- [14] M. Saveski, A. Mantrach. Item Cold-start Recommendations: Learning Local Collective Embeddings, 8th ACM Conference on Recommender Systems, pp. 89–96, 2014.
- [15] S. Sedhain, S. Sanner, D. Braziunas, L. Xie, J. Christensen. Social Collaborative Filtering for Cold-Start Recommendations, 8th ACM Conference on Recommender Systems, pp. 345–348, 2014.
- [16] Y. Sun, Y. Zhang. Conversational Recommender System, SIGIR'18, pp. 234–244, 2018.
- [17] K. Wegba, A. Lu, Y. Li, W. Wang. Interactive Storytelling for Movie Recommendation through Latent Semantic Analysis, IUI2018, pp. 521–533, 2018.
- [18] M. Zhang, J. Tang, X. Zhang, X. Xue. Addressing Cold Start in Recommender Systems: A Semi-Supervised Co-Training Algorithm, 37th International ACM Conference on Research; Development in Information Retrieval, pp. 73–82, 2014.
- [19] X. Zhao, W. Zhang, J. Wang. Interactive Collaborative Filtering, CIKM'13, pp. 1411–1420, 2013.

インタラクティブ観光経路推薦手法の構築に対する検討

Prior Investigation Towards Interactive Sightseeing Route Recommendation Method

柴田祐樹^{1*} 高間康史¹
Hiroki Shibata¹ Yasufumi Takama¹

¹ 東京都立大学大学院システムデザイン研究科
¹ Graduate School of System Design, Tokyo Metropolitan University

Abstract: This paper reports a prior investigation towards an interactive sightseeing route recommendation method. While most of the traditional recommendation systems just present a few solutions to its user, this study assumes that users establish their own route from the distribution of degree of recommendation. The degree is formulated as a probability distribution. A fast iterative sampling method called Exchange Monte Carlo is employed. The degree of recommendation is assumed to be updated based on the feedback from users. Through results by simulation, practical use cases are discussed.

1 はじめに

本稿ではインタラクティブ経路推薦手法について検討する。観光経路推薦では多くの不確実性を扱わなければならない。不確実性の原因として、天候の変化、人の行動に起因する外乱、センサの精度不足などがある。この前提の推薦であっても利用者との対話を通して推薦内容を改善していくことが可能であり、その枠組みは Human in the loop と呼ばれ、研究が進められている [1]。観光経路推薦では前述の通り状況が刻一刻と変化していく。くわえて、変化に対応して利用者が待てる時間内で再度推薦を行わなければならない。この問題に関して様々な観点から研究が行われている [6, 7, 3]。

観光経路推薦では Traveling Salesman Problem (TSP) の拡張である Selective TSP (STSP) [8] が問題の定式化としてよく用いられる。TSP は最適化問題に含まれる。最適化問題を定式化として利用した推薦手法では、最適解の近似解のいくらかを候補として推薦が広く行われる [6, 7]。しかしながらこの方法は、利用者に経路を追従する負担を強いることになり、運送計画の決定には適するが、被災地からの避難中や、休暇中の移動など状況の変化が多い用途には適していないと考える。よって本研究では利用者に経路の大まかな方向を与え、利用者自身で実際の経路を決定可能な推薦の提示方法を考える。

本稿では、離散的な経路を提示するのではなく、経路の分布を提示することで、利用者が分布から経路を

描くことができるのではないかと考える。これに対し、不確実性を扱う観光経路推薦の研究の中でも確率分布により経路生成を定式化し、Simulated Annealing (SA) [5] を用いて推薦経路を生成する研究が行われている [3]。しかしながら、[3] は分布の生成のための標本抽出法について論じておらず、確率モデルに適した問題の定式化と、SA による最適化法の提案にとどまる。

SA は温度 0 の極限分布から最適解を抽出するために、モデルの温度と呼ばれる制御変数を複数用いる。同様に多数の温度を用いて確率分布を近似する手法に、Exchange Monte Carlo (EMC) [9] と呼ばれる手法がある。SA が極限分布のみを近似するのに対し、EMC は任意の温度の分布を近似することが可能であるため、確率分布を近似する場合は EMC を用いた方が良い。また、EMC は最適化性能においても優れていることが報告されている。

本稿では、[3] の問題定式化に基づき、EMC を利用して確率分布の近似を行う手法を提案し、数値実験結果を、実際の利用場面と照らし合わせて考察する。評価実験では人工データセットを用い、EMC が SA に比べ優れていることを示す。確率分布を提示することにより、利用者が自ら道順を選択することが可能な推薦手法は著者らの確認する限り見当たらない。また、同時に、インタラクティブに利用者の状態を反映させることが経路最適化手法を目指した研究も見当たらない。以上の課題に本研究は貢献する。

*連絡先：東京都立大学システムデザイン学部情報科学
〒191-0065 東京都日野市旭が丘 6-6
E-mail: hshibata@tmu.ac.jp

1.1 記法

n を自然数としたとき、 $[n] = \{1, 2, \dots, n\}$ であるとする。ある集合 A と添字の集合 N があるとき、 A^N は集合であるとし、その要素 $a \in A^N$ は $\forall i \in N [a_i \in A]$ を満たす系列であるとする。例えば、 $x \in \mathbb{R}^{|n|}$ は $x = (x_1, x_2, \dots, x_n)$ で書かれる実数列を意味する。 $\#A$ は集合 A の要素数を示す。 $\{1, 2, \dots\}$ のように $\{\cdot\}$ を用いて定義されたものは集合であり、 $(1, 2, \dots)$ のように (\cdot) を用いて定義されたものは族であるとする¹。

2 関連研究

2.1 Traveling Salesman Problem

観光経路推薦の問題定式化に TSP がよく使われる。 N を接点の集合、 $E \subset N \times N$ を辺の集合としてグラフ (N, E) を考える。このとき、TSP は各 $n \in N$ を訪れるべき場所、各 $e = (n_1, n_2) \in E$ は n_1, n_2 を結ぶ抽象的な経路であると定義する。普通は n_1, n_2 を実在の経路で結ぶ最短経路がこれに選ばれる。 e には重み $w_e \geq 0$ が割り当てられる。重みはすべての辺に定義される、つまり (N, E) は完全グラフであるとされる。 x が有効な経路であることを

$$P(x) \Leftrightarrow [x \in N^{\{\#N+1\}}, \forall i, j \in [\#N] [x_i \neq x_j], x_1 = x_{\#N+1}]$$

と定義する。この制約 P を用いて、TSP は以下の x^* を求める問題と定義される。

$$\phi(x^*) = \min_{P(x)} \phi(x), \phi(x) = \sum_{i=1}^{\#N} w_{x_i, x_{i+1}} \quad (1)$$

上記中の ϕ は目的関数と呼ばれる。

観光経路の選定では、訪れるべき場所が定まっているよりも、利用可能な時間が定まっている場合の方が多いと言える。利用時間を制約として、より良い経路を選ぶ問題の定式化に Selective TSP (STSP) [8] がある。STSP は与えられた制限コスト内で訪れた地点にある価値の総和を最大化する問題として経路最適化問題を定式化する。

STSP, TSP どちらにおいても、考慮するグラフは実際の交通機関を抽象化したもので、グラフは完全であるとしており、実際の道路の接続関係とは異なる。問題は抽象化され、完全グラフであるため解法を考えやすいという利点がこの定式化にはあるが、観光経路推薦に用いる場合、接点として設定した地点間の経路の情報、つまり移動途中の景観などの情報が失われている、という欠点がある。

2.2 観光経路推薦

観光経路推薦手法は、ユーザ毎の価値観に応じユーザ毎に異なる推薦を行うことができ、またその逆に同じユーザに推薦される経路は、同じ性質を持つものであることが求められる。このために、前節で説明した STSP の価値は、通常ユーザの価値観をモデル化して決定される。つまり、推薦には価値観モデル化と、モデルからの解の推定をそれぞれ行うための 2 つの手法が必要となる。この 2 つをつなぐのが問題の定式化である。

推薦する経路を求める問題を定式化するが、不確実性に対処するかを考える必要がある。経路推薦における不確実性について論じた文献に [6] がある。前述した価値観モデル自体に不確実性が含まれることから、推薦問題の定式化も不確実性を扱えることが望ましい。不確実性を矛盾なくあつかう枠組みとして、Fuzzy 集合論を [6] では用いている。確率モデルによる定式化を用いた研究には Edge Vector (EV) [3] がある。本研究では主に EV を扱うため、こちらについて詳しく紹介する。

前節で述べたとおり、観光経路推薦では STSP が問題の定式化としてよく使われる。STSP は完全グラフを対象とする。これに対し EV では実際の道路地図を用いる。また、STSP が接点に観光地点の価値を割り当てるのに対し、EV ではすべての道に価値があると仮定する。また、価値、移動負荷は両方とも辺上に定義されるとしてしている。図 1 にこれらグラフの比較を示す。図中左側が従来より STSP で使われているグラフである。背面に灰色で元となった道路地図を意図したグラフを示し、興味ある地点を 5 つ仮定し、それぞれ興味の度合いを大きさで表現するとして丸で表している。大きな途中の細かい道はすべて抽象化された経路に置き換えられている。これにより問題の規模はもと道路地図よりずっと小さくなるのがわかる。右側は EV で仮定されるグラフである。すべての道を興味ある地点と仮定して最適化を行う。問題の規模は、左側のモデル化より大きいため、高速な解法が経路生成において必要となる。これに対し、SA と仮想辺と呼ばれる概念の利用が EV で提案されている。

EV の問題定式化をここで説明する。道路地図のグラフを (N, E) とする。 $e \in E$ に対し $w_e^0 \geq 0$ を道に定義された重み、 $b_e^0 \geq [0, 1)$ を道に定義された価値とする。 $n, n' \in N$ に対し、重み w を次の通り定義する。

$$(n, n') \in E \Rightarrow [w_{n,n'} = w_{n,n'}^0, b_{n,n'} = b_{n,n'}^0],$$

$$(n, n') \notin E \Rightarrow [w_{n,n'} = s(n, n'), b_e = -Bs(n, n')].$$

ただし s は以下で定義される。これは実最短経路を与

¹添字が自然数で与えられる場合、族は組や数列という概念に等しい。

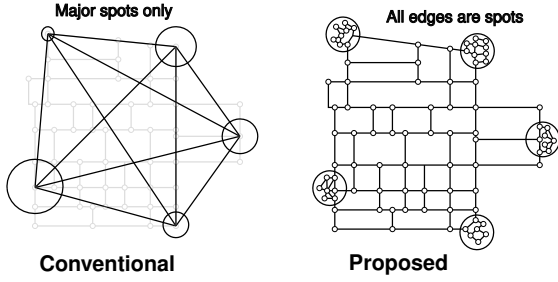


図 1: Conceptual images of the formulations used in Selective Traveling Salesman Problem (left) and proposed method (right).

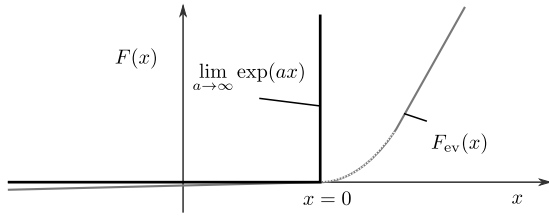


図 2: Comparison of 2 function shapes used in [3].

えるものである。

$$s(n, n') = \min_{x, m} \sum_{i=1}^{m-1} w_{x_i, x_{i+1}}, x \in N^{[m]},$$

$$x_1 = n, x_m = n', \forall i \in [m-1] (x_i, x_{i+1}) \in E$$

B は罰則係数であり、EV では $B > 0$ の値とされる。これは超パラメータである。辺 $e \notin E$ を仮想辺と呼ぶ。対応する b_e は罰則負荷である。以上の重みの元、EV の目的関数は次で定義される。

$$\phi(x, m) = F_r \left(\sum_{i=1}^{m-1} w_{x_i, x_{i+1}} - C \right) - \sum_{i=1}^{m-1} b_{x_i, x_{i+1}} \quad (2)$$

関数 $F_r: \mathbb{R} \rightarrow \mathbb{R}$ は制約関数で、STSP に対応させる場合は、 $F_r(x) = \lim_{a \rightarrow \infty} \exp(ax)$ と定義する。他に、 $F_r(x) = F_{ev}(x)$ の定義も EV では用いられる。図 2 にこれら関数の形状を示す。 F_{ev} の定義の詳細は [3] に記載されている。経路の制約は次で定義される。

$$P(x, m) \Leftrightarrow x \in N^{[m]} \wedge \forall i, j \in [m] [x_i \neq x_j]$$

時間制約に関しては F_r で表現されるため、上記には明示されない。

仮想辺 $e \notin E$ が経路 x に含まれているとき、 w_e の定義より、 x に罰則が追加される。この仮想辺 e を $s(e)$ と等しい E からなる経路に置き換えた場合の経路を x' とすると、 $\phi(x) > \phi(x')$ となるように、 B を定める必要がある。通常 $B > 1$ でこれが満たされる。これにより、仮

想辺を含む状態は取りづらくなり、最終的に最適解からこれは取り除かれる。しかしながら、完全グラフでない本定式化では、途中仮想辺を経由しなければ遷移が行えない場合があり、遷移しづらいことが問題となる。これは後述する meta-Heuristics により解決される。

2.3 Meta-Heuristics

組み合わせ最適化問題の近似解法には、Simulated annealing (SA) [5] や遺伝的アルゴリズム [11], Exchange Monte Carlo (EMC)[9], 蟻コロニー最適化 [4] と呼んだ meta-Heuristics (MHs) が一般に用いられる。MHs は局所最適化法と組み合わせて利用する。いずれも、局所最適化を目的関数を低下させる方向にのみ用いる貪欲法に対し、目的関数の増加を確率的に一時的に認めることで、過適合を防ぐ手法である。TSP 系の問題に対する局所最適化法としては、k-opt 法 [10] が広く用いられる。ここでは MHs のうち、SA と EMC について説明する。

SA, EMC どちらも Markov 連鎖 Monte Carlo (MCMC) 法に基づく手法である。標本を抽出したい対象の確率変数を \mathbf{x} 、その分布を $p(\mathbf{x})$ と表す。MCMC では確率変数の系列 $\mathbf{x}_1, \mathbf{x}_2, \dots$ を用いて \mathbf{x} を近似することを考える。このとき、十分大きな i について、 $\langle \mathbf{x}_i \rangle = \mathbf{x}$ となるような系列を定める遷移確率 $p(\mathbf{x}_{i+1} | \mathbf{x}_i)$ が存在する [2]。実際に X を \mathbf{x} の取りうる状態の空間としたとき、任意の $x, x' \in X$ について、以下の定式化を考えることができる。

$$p(\mathbf{x}_{i+1} = x' | \mathbf{x}_i = x) =$$

$$A(x', x) q(\mathbf{x}_{i+1} = x' | \mathbf{x}_i = x) + C(x) \delta(x, x') \quad (3)$$

ここで、 $(i = j \Rightarrow \delta(i, j) = 1) \wedge (i \neq j \Rightarrow \delta(i, j) = 0)$ は Kronecker のデルタであり、 A は次で定義される。

$$A(x', x) = \min \left(1, \frac{p(\mathbf{x} = x')}{p(\mathbf{x} = x)} \right)$$

また、 C は次で定義される。

$$C(x) = 1 - \sum_{x' \in X} A(x', x) q(\mathbf{x}_{i+1} = x' | \mathbf{x}_i = x)$$

さらに q は次の対称性を満たすとする。

$$q(\mathbf{x}_{i+1} = x' | \mathbf{x}_i = x) = q(\mathbf{x}_{i+1} = x | \mathbf{x}_i = x')$$

以上の式 (3) の系により系列を計算する方法は、Metropolis 法と呼ばれ、MCMC 法を実現するために考案された最初のものである。

目的関数 $\phi: X \rightarrow \mathbb{R}$ が与えられたとき、 \mathbf{x} の確率分布を次の Boltzmann 分布で定める。

$$p(\mathbf{x} = x | T) \sim \exp \left(-\frac{1}{T} \phi(x) \right) \quad (4)$$

この分布を SA と EMC では対象とする。上記分布は T が小さくなるほど $\phi(x)$ の値が小さな x が生成されやすくなる分布である。SA ではこの性質を利用し、 $T \rightarrow 0$ の分布から最適解を抽出することを考える。この極限分布は計算ができないため、実際には MCMC で $p(\mathbf{x} = x|T)$ を近似しながら、 T を低下させ、極限分布を近似する。温度の低下のさせ方としては、系列の添字 i に対し、ある定数を T_0, c 用いて $T \sim T_0 \exp(-ci)$ とする方法がある。他にも提案されているが、 $T_i > T_{i+1}$ であり、 $T_i \leq T_{i+1}$ となる温度列を定義することの有用性を主張している文献は著者らの確認する限り見当たらない。条件が変わった場合に再度最高温度から冷却の過程を実行し直す過程が必要となる。これは、計算済みの近似解を有効利用できないことを意味し、状況の微小な変化への追従性が悪いと言える。つまり、インタラクティブな用途で用いる際の欠点となる。

EMC は複数の温度を ($T_l | l \in L$) として用意する。 L は適当な添字の集合である。対応する複数の分布 $p(\mathbf{x}_l | T_l), l \in L$ を用意する。 $\mathbf{x} = (\mathbf{x}_l | l \in L)$ と置いて、これら $\#L$ の確率変数をまとめて一つの Markov 系列で近似することができる。各 $l \in L$ に対する分布の違いは温度だけに限る必要はないが、本稿では温度のみを対象とする。このような異なる超パラメータを持つ確率モデルのことを複製 (Replica) と呼ぶ。EMC では、複製を含めた全体のモデルに対するパラメータは不変であり、SA と異なる。この性質により、一度得られた解を引き続き改善することが可能であり、さらにそれぞれの複製は独立して計算することが可能なため、並列性に優れる。

SA, EMC ともに q は、用いる局所最適化法と同等な変化を与えるように定式化を行う。

3 提案手法

本稿では、EV の定式化を元に、EMC により経路の確率分布を近似する。完全グラフへ変換してから通常の STSP により最適化を実行すれば、EV の定式化は必要ない。しかしながら、その場合道の価値と負荷を変更するたびに完全グラフへの変換を実行しなければならない。利用者の行動履歴から得られる情報は、本来の道路地図に対応したものであるから、変換をせずに扱えたほうが応用性は高いと考える。また、インタラクティブ性の観点から、追加の計算負荷は避けるべきである。従い、提案手法では完全グラフへの変換をせずとも同等の計算が行える仮想辺による EV の定式化を採用する。

目的関数に式 (2) を使う。問題のグラフは EV と同様に仮想辺を用いて実際の道路地図をそのまま用いるものとする。局所最適化方には 2 変数を対象とした k-Opt 法を用いる。確率分布には式 (4) を用いる。EMC が用

いる複製間で異なる式 (4) の T について以降考える。

温度の組を $T = (T_1, T_2, \dots, T_N)$ とする。複製の分布を $p(\mathbf{x}^l), l \in L$ とする。ただし、添字集合 L は $L = \{l | l = 1, \dots, N\}$ と定義する。このとき、すべての複製をまとめた確率変数を単に \mathbf{x} と表すこととし、その分布を次で定義する。

$$p(\mathbf{x} = x|T) = \prod_{l \in L} p(\mathbf{x}^l = x_l | T_l), \quad (5)$$

この分布を近似するために、Markov 連鎖 $\mathbf{x}_1, \mathbf{x}_2, \dots$ を考える。遷移確率分布に式 (3) を用いる。 q を以降定義する。

ある添字の組 $j, k \in L$ の対を無作為に選び以下の入れ替えを行うように $q(\mathbf{x}_{i+1} = x' | \mathbf{x}_i = x)$ を定めるとする。

$$x' = (x_1, \dots, x_k, \dots, x_j, \dots, x_{\#L}), x = (x_1, \dots, x_j, \dots, x_k, \dots, x_{\#L})$$

上述した無作為な添字対の選択と入れ替えを実現する q は次の通り定義される。これは Gibbs 法に等しい。

$$q(\mathbf{x}_{i+1} = x' | \mathbf{x}_i = x) = \sum_{j,k \in L} \delta(x'_j, x_k) \delta(x'_k, x_j) \prod_{l \neq j,k} \delta(x'_l, x_l) p(\mathbf{j} = j, \mathbf{k} = k).$$

分布 $p(\mathbf{k}, \mathbf{j})$ は添字の選択を行う確率分布である。一様分布とすることもできるが、計算負荷の観点から、近傍の j, k のみ生成される分布とする。式 (3) の A について、式 (5) を代入することで以下を得る。

$$\frac{p(\mathbf{x} = x'|T)}{p(\mathbf{x} = x|T)} = \frac{p(\mathbf{x}_k = x_j | T_k) p(\mathbf{x}_j = x_k | T_j)}{p(\mathbf{x}_k = x_k | T_k) p(\mathbf{x}_j = x_j | T_j)}.$$

目的関数に関して書けば、次のとおりである。丸め誤差の影響を避けるため、実際の計算には以下を用いる。

$$\frac{p(\mathbf{x} = x'|T)}{p(\mathbf{x} = x|T)} = \exp\left(-\frac{1}{T_k} (\phi(x_j) - \phi(x_k)) - \frac{1}{T_j} (\phi(x_k) - \phi(x_j))\right). \quad (6)$$

計算の手続きは次のとおりである。

1. 系列の初期値 x_0 を設定後、以下を反復
2. 局所最適化と同等な q により各系列 $l \in L$ を独立に更新 (通常の MCMC)
3. 式 (6) の系により系列間の状態を交換 (EMC の主な特色)
4. 目的とするモデル l_0 の系列の現在の状態を標本として抽出

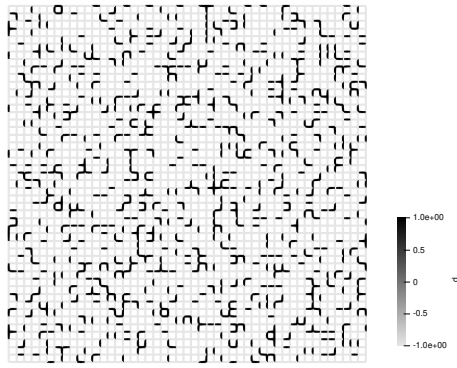


図 3: Synthesized map for experiments. Black and gray denote $b_e = 1, b_e = -1$ respectively.

4 評価実験

4.1 設定

格子状グラフを用いた評価実験を示す。座標点の集合を、 $I = [n] \times [m]$ 、接点の集合を N として、全単射 $h: I \rightarrow N$ により、辺の集合 E を次で定義する。

$$E = \{(h_{i,j}, h_{i+1,j}) | i \in [n-1] \wedge j \in [m]\} \cup \{(h_{i,j}, h_{i,j+1}) | i \in [n] \wedge j \in [m-1]\}$$

負荷は $\forall e \in E [w_e = 1]$ とする。 b_e ($e \in E$) は $b_e \in \{-1, 1\}$ の2値のみを取るとし、そのうち8割は -1 となるように無作為に決定する。以上の定義により生成したグラフに b の値により色をつけた可視化を図3に示す。黒は $b_e = -1$ 、灰色は $b_e = 1$ を示す。市街地では格子状に道が配置されていることが多いため、このグラフを用いた実験は現実的であると考えている。

実験では定常状態に達する過程を観察する。実験の説明において t は計算の進捗度を表すものとし、 $t = 0, t = 1$ でそれぞれ計算開始、計算終了を意味する。 Δt を繰り返しの単位とする。実験では $\Delta t = 2^{-17}$ とする。 Δt 計算が進むごとに、それぞれの系列が 2^6 回、式(3)により更新され、その後、モデルのパラメタの変更が EMC, SA それぞれの場合で行われる。

温度は以下により設定する。

$$T(\tau) = \exp(\ln T_L + (\ln T_H - \ln T_L) \nu) \quad (7)$$

EMC に対して、この式を用いて、 $\forall l \in L, T_l = T(\tau), \tau = (n-1)/(N-1)$ として ($T_l; l \in L$) を定める。 $N = 64$ 、つまり $\#L = 64$ の複製が用意される。分布を近似する標本は $l_0 = 64$ のモデルから抽出した標本のみを用いる。状態の交換は $l' = l + 1, l, l' \in L$ となる組 (l, l') の間で行われる。

SA は $SA_a, a = 1, 2, 4, 8, 16$ で5つの条件について評価を行う。この a について、温度は式(7)により、 $T(at)$ により定められる。つまり、 a の値が大きければ、急速に温度が低下することになる。

以上の設定で EMC, SA_a のいずれから8つの標本を抽出して統計量を算出し、結果を考察する。疑似乱数の計算には Mersenne Twister²、実数の近似には倍精度浮動小数点数を用いる。

4.2 実験結果と考察

目的関数の推移を評価する。ここでは最適化性能のみを計測する。よって、目的関数の値は小さいほうが良い。実験結果を図4に示す。EMC の場合が最も低い目的関数の値 (Energy) を示している。また、いずれの a についても、EMC は SA_a より低い目的関数の値を出している。また、 SA_a は a が大きくなるほど、つまり温度を急激に低下させるほど精度が悪くなっているとわかる。これに比べ、EMC は急速に最適化が進行した後、その後も時間をかけるほど解が改善していることがわかる。全方法の繰り返し数は等しく、EMC は並列化により複製を並列計算できることから、その計算時間は SA と概ね等しい。それにも関わらず、EMC は SA のどの場合よりも早い段階で良い解を得られており、推薦時間にどのような期限が設定された場合でも一般的に計算を開始することが可能であると言える。一方の SA では長い計算時間で精度を確保するか、低い精度でも短時間で結果を出すかを計算開始時に決定しなければならないが、これは一般に困難である。このことから、EMC の利用がインタラクティブな推薦に適している考える。実際の利用場面としては、例えば、不測の事態で経路を再計算しなければならないとき、ユーザがそれを緊急の事情により待てない場合が考えられる。このとき、未知の計算時間で妥当な解の生成を開始でき、加えて時間が確保できた場合は引き続き解の改善を EMC は行える。ただし、このためには問題のグラフを再構成しない EV による定式化が合わせて必要である。

確率分布による可視化について、本稿で検討しているものを図5に示す。(a) は EMC, (b) SA_1 により生成されたものである。正方形を縦横 256 分割した少領域に重なる経路の頻度を可視化している。(a) の方を見れば、複数の経路が重なる色が濃い部分は選択の余地がなくそこを通るべきであるとわかり、濃い経路の途中にある分岐についてはどれを選んでも問題がないことが見てわかりやすい。一方で (b) の結果では経路にまとまりがなく、道筋を見極めることが難しい。(a) がこのような可視化を可能な理由は、用いた EMC が確率

²<http://www.math.sci.hiroshima-u.ac.jp/m-mat/MT/ent.html>

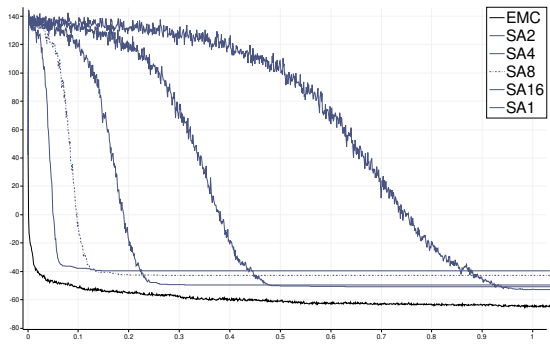


図 4: History of objective function for each method.

分布に従い経路を生成しているため、推薦度が正しくモデル化できているからであると言える。

5 まとめ

本稿では、インタラクティブ観光経路推薦手法の構築のための、経路分布推定手法を提案し、実験結果を考察した。提案手法は Exchange Monte Carlo 法を利用し、広く用いられる Simulated Annealing よりもインタラクティブ性において優れた性質を持つことを実験により示した。また、推薦度の可視化により、ユーザーが経路を自ら設計する場面を検討し、提案手法の想定される場面における応用性を述べた。

今後、インタラクティブ性の評価指標と対応する実験、実データを用意した評価実験を検討する必要がある。また、並列性能の改善、温度の決定法を検討する。

参考文献

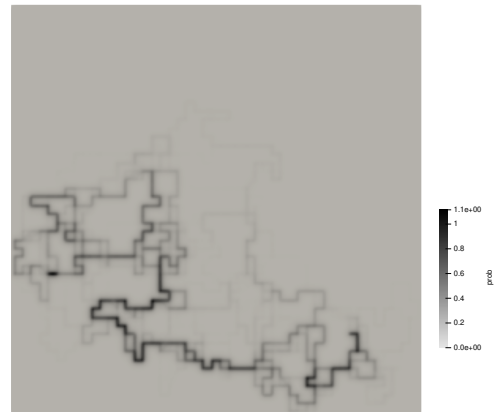
[1] Pinto, R., Mettler, T., Taisch, M.: Managing Supplier Delivery Reliability Risk Under Limited Information: Foundations for a Human-in-the-Loop DDS, *Decision Support System*, Vol. 58, No. 2, pp. 1076-1084 (2013)

[2] Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A.: Equation of State Calculations by Fast Computing Machines, *The Journal of Chemical Physics*, Vol. 21, No. 6, pp. 1087-1092 (1953)

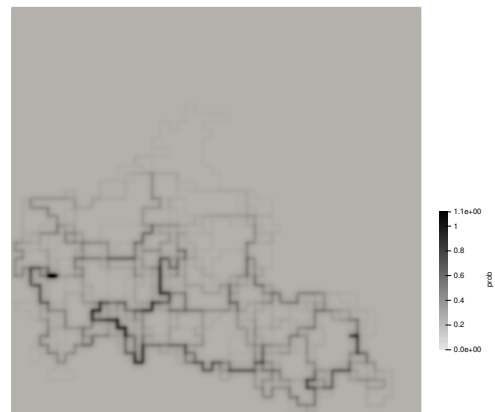
[3] Shibata, H., Takama, Y.: Improvements on Probability Model for Sightseeing Route Recommendation Method Employing Generalized Formulation in terms of Edges, *33th International Conference of Technologies and Applications of Artificial Intelligence*, 6 pages (2019)

[4] Dorigo, M., Maniezzo, V., Colomi, A.: Ant System: Optimization by a Colony of Cooperating Agents, *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 26, No. 1 pp. 29-41 (1996)

[5] Kirkpatrick, S., Gelatt, C., Vecchi, M.: Optimization by Simulated Annealing, *Science*, Vol. 220, No. 4598, pp. 671-680 (1983)



(a) Exchange Monte Carlo: EMC



(b) Simulated Annealing: SA1

図 5: Visualized example of route distribution.

[6] Laporte, G., Martello, S.: Tour Route Planning Problem for Sightseeing with the Multiroute under Several Uncertain Conditions, *2012 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 715-720 (2012)

[7] Ronny, K., Marko, M., Hagen, K.: A City Guide Agent Creating and Adapting Individual Sightseeing Tours Based on Field Trial Results, *International Journal of Computational Intelligence Research*, Vol. 2, No. 2 pp. 191-206 (2006)

[8] Feillet, D., Dejax, P., Gendreau, M.: Traveling Salesman Problem with Profits: An Overview, *Transportation Science*, Vol. 39, No. 2, pp. 188-205 (2005)

[9] Hukushima, K., Nemoto, K.: Exchange Monte Carlo Method and Application to Spin Glass Simulations, *Journal of the Physical Society of Japan*, Vol. 65, No. 6, pp. 1604-1608 (1995)

[10] Sohl-Dickstein, J., Mudigonda, M., DeWeese, M.: General k-Opt Submoves for the LinKernighan TSP Heuristic, *Mathematical Programming Computation*, Vol. 1, No. 2-3, pp. 119-163 (2009)

[11] Goldberg, D.: *Genetic Algorithms in Search, Optimization, and Machine Learning*, 1st edition, Addison-Wesley Professional (1989)

跳躍動作の周波数特性に着目した スポーツ損傷潜在リスクの推定

Estimating the Potential Risk of Sports Injuries Based on the Frequency Characteristics of Jumping Motion

松下 光範^{a,*} 鈴木 雄登^{b,†} 尾崎 優^{a,‡} 堀 寛史^c
Mitsunori Matsushita Yuto Suzuki Suguru Ozaki Hirofumi Hori

a. 関西大学総合情報学部

Faculty of Informatics, Kansai University

b. 関西大学大学院総合情報学研究科

Graduate School of Informatics, Kansai University

c. びわこリハビリテーション専門職大学理学療法学科

Faculty of Physiotherapy, Biwako Professional University of Rehabilitation

Abstract: 本稿では、スポーツ障害の予防支援を目的として、ユーザが抱える怪我のリスク（スポーツ損傷潜在リスク）の推定を試みた。従来、スポーツ損傷潜在リスクは理学療法士などの専門家が三次元動作解析装置や床反力計といった機器を用いて推定するため、継続してリスク管理を行うコストが高い。こうした観点から、本研究ではスマートフォンのセンサを用いて簡易的にスポーツ損傷潜在リスクを評定することを試みる。提案手法では、跳躍動作の周波数特性に着目し、対象者の跳躍動作データにフーリエ変換を用いて周波数成分に分解する。分解することで得た周波数スペクトルを特異値分解を用いて次元削減を行うことで特徴量を抽出し、それを k 近傍法を用いて分類する。本稿では、と提案手法によるスポーツ損傷潜在リスクの推定結果とスポーツの現場で継続的に使用してもらうために試作したスマホアプリケーションについて報告する。

1 はじめに

2020年に蔓延したCOVID-19の影響により、人々の運動やスポーツに関する環境に様々な変化が現れた。例えば、屋内スポーツジムで発生したクラスター感染によるスポーツ施設等の利用制限や、長期間の休校措置による体育授業や部活動への参加機会の減少が、人々の身体活動量の低下を招いている。運動実施の減少が1ヶ月続くと体力の回復には3ヶ月程度かかると言われており^{*1}、それに起因するスポーツ損傷の増加が懸念されている。

スポーツ損傷は、スポーツ外傷とスポーツ障害の二

つに分類される [7]。スポーツ外傷は1回の強い外力によって生じる、明らかな受傷起点がある怪我である。例えば、他人との接触やジャンプの着地失敗による足首の捻挫や靭帯損傷などが該当する。スポーツ障害は、何度も繰り返し強い力が加わることや身体能力の限界を超えた運動を行うことで、徐々に肩や肘、膝、足首などに痛みを感じはじめ、その症状が慢性的に続く状態をいう。スポーツ外傷は多くの場合、大きな痛みを伴うため、発症後すぐに治療へと移行しやすい。一方、スポーツ障害は気付かないうちに症状が進行・悪化することが多く、治療が長期化しがちで明確な原因が判明しない場合も多い。そのため、スポーツ障害を起こさないためには早期発見や予防が重要であり、理学療法士など専門家による継続的な評価や検査を受けることが望ましい [6]。

しかし、これらの評価や検査で使用する機器は価格が

* 連絡先：関西大学総合情報学部

〒569-1095 大阪府高槻市霊仙寺町 2-1-1

E-mail: mat@res.kutc.kansai-u.ac.jp

† 現在、株式会社内田洋行

‡ 現在、パナソニックシステムソリューションズジャパン株式会社

^{*1}<https://www3.nhk.or.jp/news/html/20200827/k10012587111000.html> (2021/11/7 確認)

高額なうえ特定の場所に設置する必要があるため、スポーツの現場に導入するには制約が大きく、継続的かつ容易に評価や検査を受けられる状況にはない [12]. そこで本研究では、スポーツ損傷の予防・早期発見のために、ユーザ自身が日常で簡便に利用できるスポーツ損傷潜在リスク評価ツールの実現を目指す。

2 先行研究

本章ではスマートフォンを活用した身体への負担推定や傷病予測に関する研究について挙げる。

まず、正しい姿勢で運動することがスポーツ損傷のリスク低減に寄与するという観点から、スマートフォンのセンサを活用した姿勢推定について述べる。榎原らは、スマートフォンを用いた歩行動作改善ツール [5] を開発し、スマートフォンに内蔵された加速度センサのみでも、良い姿勢を意識した歩行と無意識な歩行では加速度の波形に差異があることを明らかにした。また、王らは、スマートフォンを用いて、腕立て伏せやスクワットなど、機器をいわずに自重による負荷を利用したトレーニングの支援システムを開発した [4]. このシステムでは、スマートフォンのカメラを用いてユーザのトレーニング時の画像を撮影し、OpenPose^{*1}を利用して姿勢データを取得する。このシステムでは取得した姿勢データからユーザの姿勢が正しいかを確認し、音声フィードバックにより姿勢の修正をユーザに促す。

次に、スポーツ損傷のリスク低減の観点から、活動強度の推定に関する研究について述べる。隅田らは、スマートフォンのセンサを用いて活動時の負担度を推定し、ユーザにとって負担が最小となる歩行ルートを推薦するナビゲーションシステムを提案した [9]. このシステムでは個々の負担度変化を推定するために、機械学習を用いて加速度や歩行速度などの歩行データから心拍数を予測する負担度モデルを作成している。Pernek らは、スポーツ活動強度の推定手法を提案した [1]. 提案手法では、スマートフォンを用いて取得した加速度や心拍のデータに対し、高周波数を除去して特徴抽出を行い、その特徴量を用いて多層パーセプトロン、線形回帰、回帰木、サポートベクトル回帰の4種類の機械学習アルゴリズムを用いてスポーツ活動強度を推定した。これらの研究では、活動時にユーザにかかる負担の推定を試みている。

これらの研究はスポーツ損傷リスクの推定を企図したものではないが、「正しい姿勢で運動ができていますか」

「どのくらいの活動強度で運動しているか」は怪我のしやすさに影響するため、これらの方法や得られた知見を参考にしながらシステム設計を行う。

Wilkerson らは、NCAA フットボールチームの選手45名を対象にスポーツ外傷のリスク推定を行った [2]. この研究では45名の選手を対象に、プレシーズン練習開始1ヶ月前に、以前の筋骨格損傷の持続的影響の自己評価を10項目からなるスポーツ・フィットネス・インデックス (SFI) を用いて調査するとともに、スマートフォンを用いて片足スクワットテスト (UFS) の加速度データを取得した。これらのデータと試合参加の記録 (GP) から、「SFIスコアが低い (Lo SFI)」、「USF値が高い (Hi UFS)」、「試合出場回数が多い (HiGP)」を傷害リスク因子とする傷害予測モデルを作成した。Wilkerson らの研究の目的は我々の研究と合致しており大いに参考になるものの、レベルの高いアスリートを対象にしているため、本研究が想定するユーザには難しい測定項目 (e.g., 片足スクワットテスト) が含まれており、多くの人が日常的に利用するには十分ではない。

3 デザイン指針

本研究で目指すスポーツ損傷潜在リスク推定システムは、専門家の判断を代替するのではなく、スクリーニング手段としての活用を想定している。スクリーニングは、検査や測定を実施する前におおまかに正常か否かを判別し、詳細な検査・測定を実施すべきか否かを判断するものである [3]. スポーツ現場では医師による画像所見の確認および診察が難しいため、初期評価時のスクリーニングによる損傷程度の把握が重要になる [13]. そのため、本研究で目指すシステムは、ユーザ自身が簡易かつ継続的に利用できるようにすることで、自身の身体状況を理解し、予防的に専門家への受診ができるよう支援することを企図する。

こうした目的から、(1) 特別な機器を必要としないこと、(2) 日常的に容易にできる運動動作を対象にすること、(3) 高い運動能力を持つアスリートではなく中学や高校の部活動に参加する生徒などを対象ユーザとすることをシステムのデザイン指針と定めた。こうした観点から、本研究では運動動作を用いた計測のうち、広い場所を必要とせずに気軽に実施でき、下肢の損傷潜在リスクを評価するのに有効である跳躍動作を計測対象とする。跳躍動作を分析することで、下肢の筋力の強さだけでなく、体幹のバランスなどがわかる [11]. 跳躍動作は歩行動作などに比べて運動量が大きいため下肢筋力や体幹などの傾向が現れるという特徴から、理学療法分野

^{*1}<https://github.com/CMU-Perceptual-Computing-Lab/openpose> (2021/11/7 確認)

や医学の分野で検査の際に活用されている。本稿では、跳躍動作の測定手法の一つであるリバウンドジャンプ (Rebound Jump, 以下 RJ と記す) を用いる。RJ は、踏切時間を短くし鉛直方向へできるだけ高く跳躍することで短い時間でのパワー発揮能力を評価する、スポーツパフォーマンス測定手法の一つである [10]。提案システムでは、スマートフォンで獲得可能なセンサ情報に基づいてユーザの RJ の特徴を抽出し、身体状態を推定する。

4 データセット

我々はこれまでに、下肢の筋力を継続的に計測することを目的として、スマートフォンに内蔵のモーションセンサを用いた跳躍動作の計測システムを実装し、跳躍動作の鉛直方向の移動を分析するために加速度センサの値を、姿勢の傾きを分析するためにジャイロセンサから得られるクォータニオンの値を各々取得することで、跳躍動作の特徴取得が可能であることを確認している [8]。本研究では、このシステムを用いて実験参加者の RJ データを収集し、そのデータに対して潜在リスク評価値を付与して教師データとして用いる。

データの収集は、安全性担保の観点から、日常的に運動を行っている 20 代の男子大学生 44 名を対象に行った。実験では、まず倫理事項を口頭及び書面で説明した後同意書を記入してもらい、その後に (1) 質問紙調査、(2) 下肢筋力評価、(3) RJ の順に実施した。質問紙調査では、身長、年齢、体重、利き足、怪我と運動の経歴について尋ねた。次に、下肢筋力評価として、村永が考案した立ち上がり評価 [16] を行ってもらった。立ち上がり評価は、40cm、30cm、20cm、10cm の高さの台を用いて、腰掛け座位から反動を使わずに両脚、片脚のそれぞれで立ち上がることができるかを測るテストである。この下肢筋力評価は、実験の安全性を担保するために実施した。本研究では、この結果を評価指標の一つとして扱うとともに、実験協力者の安全性担保のためのスクリーニング基準としても利用する。そのため、20cm 以下の台座からの左右両方の片脚立ち上がりを遂行できた実験協力者のみ次の手順に進んでもらうこととした。今回の実験では、44 名の実験参加者のうち 40 名が 20cm 以下の片足立ち上がりを両足とも遂行できた。

つぎに、これら 40 名の実験参加者に、両足での RJ、および片足での RJ を課した。このとき、跳躍時のデータを取得するためにスマートフォンをベルトで腰に装着させた。その跳躍の様子は VTR でも撮影した。跳躍時は靴を脱ぎ、安全のために EVA 樹脂製マットないスタイルカーペット上で跳躍してもらった。跳躍時には、以

下の 3 点を注意事項として指示した。

1. できるだけ高く跳ぶこと。
2. 床の反発を利用して間隔をあけずに跳ぶこと。
3. 初期位置からできるだけ移動しないこと。

この跳躍により、加速度、角速度、クォータニオンについて各 3 軸、計 9 種類の時系列データを取得した。一般に腕の振込み動作は跳躍動作を正当に評価するために制約を課することが多いが、本研究で実装を目指すシステムは日常的にユーザが自分自身で測定することを想定しているため、煩雑な動作条件を排除するという観点から腕の振込みについての指示は行わなかった。

収集されたデータに対して本稿第 4 著者がスポーツ損傷潜在リスクの評価を行った。評価に用いたのは

- 質問紙の回答
- 立ち上がりテストの結果
- 跳躍時の正面・側面からの OpenPose ボーン付き動画
- 身体ラインの角度
- OpenPose ボーン付き静止画

の 5 つである。これらのデータを用いて、スポーツ損傷潜在リスクの原因特定に用いられる安定性、再現性、強度、knee-in (以降、KI と記す)、trunk-out (以降、TO と記す) の 5 つの指標、ならびに総合評価についての推定を行う。再現性は、同じ動作の再現ができていないか、安定性は、身体・重心のバランスが保てているか、強度は、対象者が行っている跳躍動作が強いのか、KI は、着地時に膝が内側に曲がっている状態か、TO は、着地時に体感が外側に傾いている状態かを指す指標である。また、総合評価は動作から読み取れる怪我のリスクを表す指標であり、スポーツ損傷潜在リスクと同義である。本研究では、再現性、安定性、強度、総合評価については 0 (できていない) ~5 (とても良い) の 6 段階で評価することとした。また、KI、TO については、傾向が見られる場合は 1、見られない場合は 0 の 2 段階で評価することとした。左右それぞれの動作における評価の分布を表 1 および表 2 に示す。

5 提案手法

連続跳躍動作の周波数特性に着目してスポーツ損傷潜在リスク推定精度を測定するために、本稿ではフーリエ変換を用いる。フーリエ変換では、任意の周期関数 $f(t)$ は特定の周波数を有する波の重ね合わせにより表現される [14]。一次領域における関数 $f(t)$ が周波数領域

表 1: 右足 RJ の評価分布

評価値	再現性	安定性	強度	KI	TO	総合評価
0	0	0	0	31	27	0
1	0	0	1	9	13	0
2	4	6	7	N/A	N/A	7
3	19	14	14	N/A	N/A	14
4	14	18	15	N/A	N/A	15
5	3	2	3	N/A	N/A	4

表 2: 左足 RJ の評価分布

評価値	再現性	安定性	強度	KI	TO	総合評価
0	0	0	0	19	30	0
1	0	0	1	21	10	0
2	5	7	7	N/A	N/A	6
3	16	20	16	N/A	N/A	19
4	17	10	11	N/A	N/A	11
5	2	3	5	N/A	N/A	4

で角振動数 ω を変数とする関数 $F(\omega)$ で表現されたとき、 $F(\omega)$ を $f(t)$ のフーリエ変換といい、次の式で表される。

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt \quad (1)$$

この式のように、取得したモーションデータ $f(t)$ がいくつかの周波数成分に分解される。それによって求められた $F(\omega)$ がスペクトルであり、絶対値がその周波数成分の振幅に対応している。 $F(\omega)$ のスペクトルを並べたものが周波数スペクトルである。スポーツ損傷潜在リスクが低い人ほど跳躍動作に安定性があり、周波数成分が分散しにくいという仮説のもと、本研究では、特徴値としてフーリエ変換後の周波数スペクトルを用いて、スポーツ損傷潜在リスクの推定を行った。

6 スポーツ損傷潜在リスク推定手法

本稿では、(1) モーションデータの整形、(2) 周波数成分への分解、(3) 特異値分解による次元削減、(4) k 近傍法による分類、という流れでスポーツ損傷潜在リスクの推定モデルを作成する。以下で各々の処理の詳細について述べる。

6.1 モーションデータの整形

まず、フーリエ変換を行うために、モーションデータの整形を行った。ある実験協力者の鉛直方向の加速度を図 1 に示す。跳躍動作は自由落下により重心の移動が規則的に生じる動作であるため、跳躍と着地が傾向とし

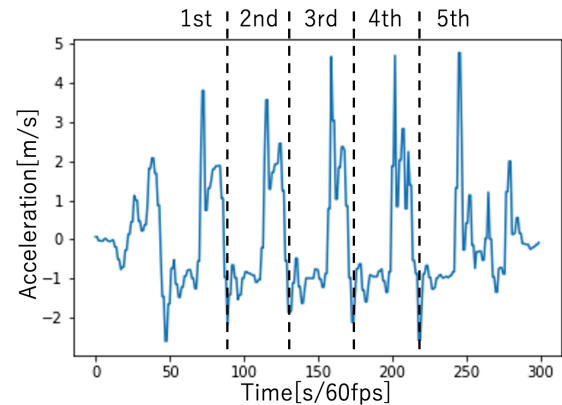


図 1: 跳躍動作の鉛直方向の加速度

て現れる。着地時の極小値を取得することにより、跳躍動作の回数ごとに分割した。動作の開始と動作の終了の影響を排除するために 2~4 回目の跳躍動作を抽出した。同様の分割作業を、鉛直方向の加速度 Y を基準に、他の 8 種類のモーションデータに対しても行った。

6.2 フーリエ変換を用いた周波数分解

データ整形を行ったモーションデータに対し、フーリエ変換を用いて周波数スペクトルに変換した。図 2 にスポーツ損傷潜在リスクが高い実験参加者の周波数スペクトルを示す。

スポーツ損傷潜在リスクが高い跳躍動作ではピーク値が分散していることが確認できる。

次に、スポーツ損傷潜在リスクが低い実験参加者の周波数スペクトルを図 3 に示す。スポーツ損傷潜在リスクが高い実験参加者の波形と比べると、スポーツ損傷潜在リスクが低い評価ではピーク値の違いを確認することができる。スポーツ損傷潜在リスクが低い跳躍動作は周波数成分が分散していないが、スポーツ損傷潜在リスクが高い跳躍動作は周波数成分が分散していることが観察できる。

スポーツ損傷潜在リスクが低い跳躍動作は動作回数ごとに同じような動作を再現できているために周波数成分が分散していないのに対して、スポーツ損傷潜在リスクが高い跳躍動作は動作回数ごとにズレが生じており、結果として周波数成分が分散したのではないかと考える。

6.3 特異値分解を用いた次元削減

フーリエ変換後の周波数スペクトルの次元削減のために特異値分解を用いる。特異値分解は動作識別への適用事例があり、本研究に利用可能であると判断して採用

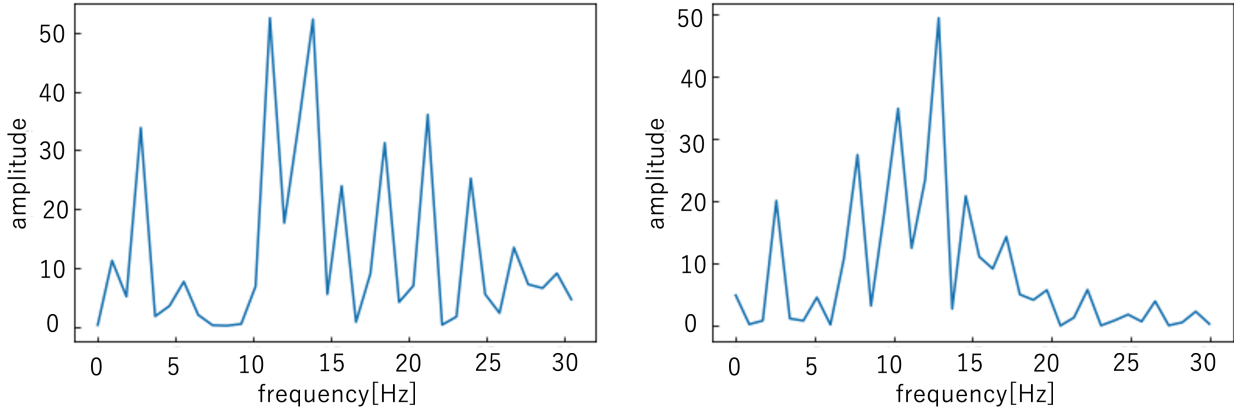


図 2: スポーツ損傷潜在リスクが高い跳躍動作

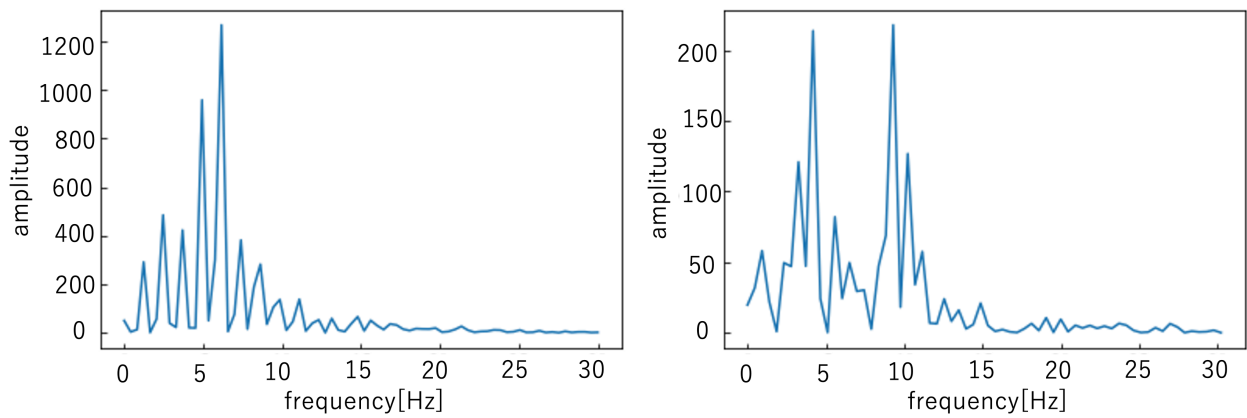


図 3: スポーツ損傷潜在リスクが低い跳躍動作

した [15]. 式 (2) に示すように、特異値分解では時系列データから行列 D を構成し、それを特異値、左特異ベクトル、右特異値ベクトルに分解する。

$$D = U\Sigma V^T \quad (2)$$

$U = (u_1, u_1, \dots, u_m)$ は $m \times m$ のユニタリ行列で、 $V = (v_1, v_1, \dots, v_n)$ は $n \times n$ のユニタリ行列である。また、 Σ は $m \times n$ で、対角要素以外は 0、対角要素は非負となる行列である。 Σ の対角要素は特異値を示し、 $u_i (1 \leq i \leq m)$ が左特異ベクトル、 $v_j (1 \leq j \leq n)$ が右特異ベクトルである。左特異ベクトルが時系列データの特徴を抽出したものであり、特異値行列 Σ が対応する左特異ベクトルの影響の大きさ (重み) を示していることから、特異値の高い左特異ベクトルは対象の時系列データの特徴を良く表現するベクトルとして扱うことができる。

提案手法では、周波数スペクトルから $m \times n$ の行列 D を作成する必要がある。本稿では $m = 9$ とし、9 種類のモーションデータを用いて、9 個の行列を作成した。

これらの結果から第一特異値を抽出し、それらを特徴値とした。

6.4 k 近傍法

特異値分解により得られた 9 種類の特徴値を説明変数に、理学療法士による評価データを目的変数に男子大学生 40 名のデータを学習・テストデータとして k 近傍法を用いて、スポーツ損傷潜在リスクの推定を行った。 k 近傍法は与えられた学習データをベクトル空間上に配置し、リスクの推定対象となるデータ (未知データ) から距離に近い順番に任意の k 個を取得し、その k 個の多数決により、その未知データが属するクラスを推定する。本稿では、leave-one-out 交差検証を用いて推定精度を比較した。

6.5 結果

k 近傍法は、 k の値により推定精度が大きく変動するため、本稿では、 $k=3, 5, 7, 9$ としてスポーツ損傷潜

表 3: 提案手法の左足 RJ

	再現性	安定性	強度	TO	KI	総合評価	全項目平均
k=3	0.225	0.300	<u>0.450</u>	0.525	0.725	0.475	0.450
k=5	0.275	0.325	0.425	0.525	<u>0.775</u>	0.475	0.467
k=7	0.275	0.425	<u>0.450</u>	0.425	0.725	<u>0.550</u>	0.475
k=9	<u>0.375</u>	<u>0.450</u>	0.425	<u>0.575</u>	0.750	<u>0.550</u>	<u>0.521</u>

表 4: 提案手法の右足 RJ

	再現性	安定性	強度	TO	KI	総合評価	全項目平均
k=3	0.350	0.350	0.300	0.725	0.650	<u>0.350</u>	0.454
k=5	<u>0.425</u>	0.400	0.375	0.725	0.650	0.325	0.483
k=7	0.400	0.425	0.425	<u>0.775</u>	0.600	0.250	0.479
k=9	0.375	<u>0.500</u>	<u>0.450</u>	<u>0.775</u>	<u>0.725</u>	0.325	<u>0.525</u>

在リスクの推定を行った。提案手法の推定精度を表 3, 表 4 に示す。精度自体はまだ改善の余地はあるものの、いずれの指標についてもチャンスレベルを上回る結果となっており、提案手法の有用性は一定程度示せたものと考えている。

全項目平均を見ると、いずれの足の RJ でも $k = 9$ が最も高い精度を示した。今回は実験参加者が 40 名であるため、今後継続的に実験参加者を増やすことで精度向上が図れるか検証する必要がある。また、今回の実験結果では、右足 RJ の総合評価の精度が左足 RJ に比べて低い結果になっている。その一因として利き足が影響している可能性が考えられる。今回、事前の質問紙調査で利き足の確認を行わなかったが、今後の調査では実験参加者の利き足についても収集する必要がある。

7 アプリケーション

本稿で示した提案手法に基づき、スマートフォンを用いたアプリケーションとして、スポーツ損傷潜在リスク推定システムを実装した。スマートフォンをプラットフォームとした理由は、可搬性が高く、広く普及しているためである。このシステムの処理手順を図 4 に示す。

本システムは、フロントエンドに Swift (ver.5.0) を、バックエンドに Python (ver.2.7.16) を利用し、iPhone 対応の iOS アプリケーションとして実装した。図 5 にアプリケーション画面を示す。

メニュー画面 (a) では計測、動作説明、記録、設定の 4 種類の項目があり、ボタンを押すことでそれぞれのページへと遷移する。

動作説明の画面では、計測手順の説明と RJ を行う際

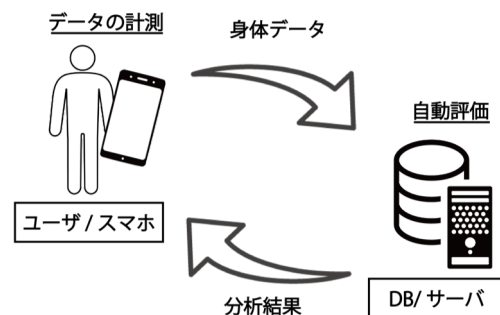


図 4: スマホアプリケーションの処理手順

の注意事項をまとめた動画が再生され、確認することができる。

計測画面では、ユーザ自身に RJ を行ってもらい、スポーツ損傷潜在リスクを確認することができる。計測画面では、動作説明と同様の動画を再確認するためのボタンと、計測を行うための画面へと遷移するためのボタンがあり、計測を行うための画面へ遷移すると、「靴を脱ぎましたか?」と「iPhone をベルトに挿入してください」というメッセージが順に示され、ユーザに計測を行うために必要な行動を促す。計測を開始すると、「右足の RJ を 5 回行ってください」という音声再生されてユーザに計測を行う対象動作を説明する。動作を開始する時には、ユーザ自身に計測のスタートボタンを押してもらい RJ を 5 回行ってもらう、動作の終了後に計測のストップボタンを押してもらうことでモーションデータをサーバに記録する。これらの計測手順を右足 RJ を行った後に、左足 RJ を行ってもらい、スポーツ損傷潜在リスク推定に必要な動作が終了する。これらの収集し

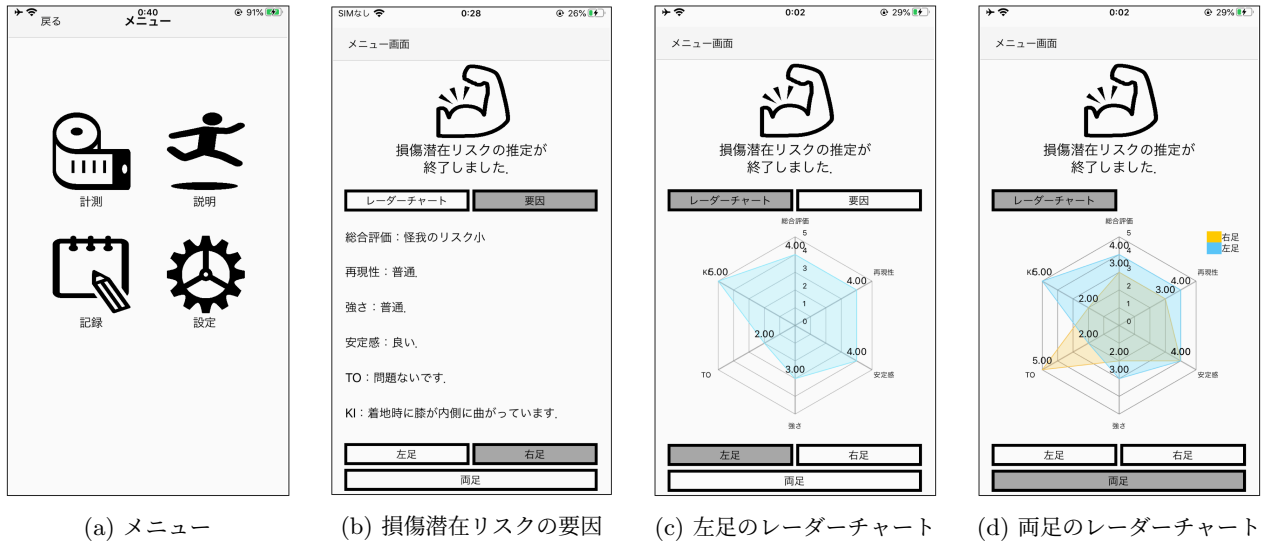


図 5: スマホアプリケーションのスナップショット

たモーションデータから、提案手法で示したスポーツ損傷潜在リスクの推定を用いることで、バックエンドで計算機による自動評価を行い、ユーザのスマートフォンへ提示する。

スポーツ損傷潜在リスクの推定結果を、スポーツ損傷潜在リスクの判断の要因 (b) とそれを可視化したレーダーチャート (c), (d) の 3 つの画面によって提示する。計測結果画面へ遷移した時点では、右足のレーダーチャートが出力される。右足・左足それぞれのボタンを押すことで左右別々に推定結果を確認することができ、両足のボタンを押すことで左右のレーダーチャートを比較して確認することができる。レーダーチャートは、再現性・安定性・強度・総合評価は 5 段階、TO・KI は 2 段階で表示される。

記録画面ではデータを取得した時間ごとにフォルダが分割され、フォルダを選択することで、特定の日時に取得した左右のデータを選択することができる。左右のどちらか一方を選択することで、取得した過去のスポーツ損傷潜在リスク推定結果を確認することができる。推定結果の出力方法は、モーションデータの取得後のスポーツ損傷潜在リスク推定結果の出力画面 (b)~(d) と同様の形で出力される。

8 おわりに

本稿では、スポーツ損傷の予防支援を目的に、スマートフォンを用いた簡易計測による損傷潜在リスクの評価手法の提案を行った。スポーツ損傷のリスクを評価するために理学療法士が用いる機器は、高額かつ特定の場所

に設置する必要があるといった制約があり、日常で継続して検査をすることが容易ではない。本研究では、個人が所有するスマートフォンを活用することで、その問題の解決を試みた。提案手法では、跳躍動作の周波数特性に着目したスポーツ損傷潜在リスクを簡便に評価することを企図し、ユーザの跳躍動作の時系列データ（加速度、角速度、クォータニオンの 3 軸 XYZ における 9 種類の時系列データ）から各々の周波数スペクトルを導出し、特異値分解を用いて第一特異値の抽出を行った。さらに、抽出した第一特異値を特徴値として、 k 近傍法により損傷潜在リスクの推定を行った。leave-one-out 法を用いた評価により、一定程度の有用性が示せたと考えている。

今後、下記の項目の実施により精度の向上を目指す。

1. データセットを増やし、評価ごとの偏りをなくす。今回推定に用いたデータ数は 40 人分で、運動経験のある 20 代男性に限定した上で立ち上がり評価によるスクリーニングも行ったことから、評価ごとの偏りが大きい。また、実験時のデータのみでの判断となり、長期的な推移や質問紙調査の内容の反映も行えていない。これらを評価値に反映することで、精度の改善が期待される。今後、作成したアプリケーションを活用して、より多くのデータの収集を試みる。
2. 複数の理学療法士の評価結果を統合する。現状では第 4 著者の評価のみを用いてスポーツ損傷潜在リスクの推定を行っている。しかし、今後データセットを増加させるにあたり、複数の理学

療法士の評価を学習データとして扱えることが望ましい。今後は、スポーツ理学療法の経験を持つ理学療法士の評価を組み合わせた手法についても検討していく。

3. 各推定項目での説明変数の組み合わせを動的に変更する。

現在は、左右のRJを独立して評価を推定しており、左右の差の影響を考慮できていない。また、TOやKIのように、周期的特徴を持たない要素について、現状の手法を適用することが妥当であるかは検証の必要が残る。今後は、左右のスポーツ損傷潜在リスクや要因の差を複合的に推定できる手法を検討する。

謝辞

本研究の実施に当たり、データ分析手法に関してご示唆を頂いた関西大学総合情報学部林勲教授に感謝する。なお、本研究の一部は2021年度関西大学研修員費によって行われた。

参考文献

- [1] Pernek, I., Stiglic, G. and Kokol, P.: How hard am I training? Using smart phones to estimate sport activity intensity, *2012 32nd International Conference on Distributed Computing Systems Workshops*, pp. 65–68 (2012).
- [2] Wilkerson, G. B., Gupta, A. and Colston, M. A.: Mitigating sports injury risks using internet of things and analytics approaches, *Risk analysis*, Vol. 38, No. 7, pp. 1348–1360 (2018).
- [3] 白田滋: 理学療法評価とは, ビジュアルレクチャー理学療法基礎評価学(白田滋(編)), 医歯薬出版株式会社, chapter 1, pp. 2–10 (2014).
- [4] 王瑞賢, 高橋伸, 志築文太郎: スマートフォンを用いた自重トレーニング時の姿勢支援システム, 情報処理学会研究報告, Vol. 2019-EC-52, No. 5, pp. 1–6 (2019).
- [5] 檜原裕大, 清水裕基, 三好健文, 吉永努, 入江英嗣: スマートフォンを用いた歩行動作改善ツールの開発, 情報処理学会研究報告, Vol. 2011-UBI-32, No. 7, pp. 1–8 (2011).
- [6] 小林寛和: スポーツと理学療法, 理学療法学, Vol. 42, No. 8, pp. 773–774 (2015).
- [7] 小柳磨毅, 橋本雅至, 向井公一, 田中則子, 佐藤睦美, 中江徳彦: スポーツ傷害に対する理学療法の取り組み, 理学療法学, Vol. 32, No. 4, pp. 269–274 (2005).
- [8] 鈴木雄登, 松下光範, 堀寛史: 継続的な評価の支援を目的としたスマートフォンによる跳躍動作の簡易センシング, 情報科学技術フォーラム講演論文集, Vol. 18, No. 4, pp. 457–460 (2019).
- [9] 隅田麻由, 今津眞也, 水本旭洋, 安本慶一: 健康指向歩行ナビゲーションのためのスマートフォンによる身体的負担度の推定, 情報処理学会研究報告, Vol. 2012-MBL-62, No. 27, pp. 1–8 (2012).
- [10] 図子あまね, 荻山靖, 図子浩二: リバウンドジャンプテストを用いた跳躍選手の専門的な下肢筋力・パワーに関する評価, 体力科学, Vol. 66, No. 1, pp. 79–86 (2017).
- [11] 中俣修, 新田收, 占川順光: 健常人における両脚跳躍動作の跳躍高を決定する因子の分析—体幹・下肢の姿勢と運動に着目して—, 理学療法学, Vol. 41, No. 5, pp. 290–300 (2014).
- [12] 新島有信, 水野理, 田中智博: ズボンのポケットに収納したスマートフォンによる歩容解析方法に関する検討, 情報処理学会研究報告, Vol. 2014-UBI-43, No. 13, pp. 1–8 (2014).
- [13] 濱田太郎, 中尾英俊, 稲葉考洋, 高橋孝輔, 来田晃幸, 松本和夫: 高校部活動の支援: 硬式野球部(1), アスリートケア—理学療法士によるスポーツ選手への健康支援—(越智隆弘(編)), 三輪書店, chapter 3-1, pp. 210–216 (2017).
- [14] 古田昌伸, 中山栄浩: 高速フーリエ変換法を用いたセレーションの評価, 軽金属, Vol. 54, No. 2, pp. 51–56 (2004).
- [15] 三嶋賢一, 金田さやか, 中西弘明, 榎木哲夫, 堀口由貴男: 特異値分解を用いた非周期動作における運動特徴の解析, ロボティクス・メカトロニクス講演会講演概要集 2010, 2A1-A26 (2010).
- [16] 村永信吾: 立ち上がり動作を用いた下肢筋力評価とその臨床応用, 昭和医学会雑誌, Vol. 61, No. 3, pp. 362–367 (2001).